

Analyse Multidimensionnelle Des Déterminants De La Production Forestiere Au Sein De La Société Fanalamanga

ANDRIAMANIRAKA Lantonirina Harry Liva^{1,1}, FARAALINORO Rakotomangarivelo^{1,2}, MIRANA Vivia Claudette^{1,3}

^{1,2,3}, Université d'Antananarivo

Auteur Correspondant : ANDRIAMANIRAKA Lantonirina Harry Liva



Résumé : Cette étude propose une analyse multidimensionnelle des déterminants de la production forestière au sein de la société FANALAMANGA, spécialisée dans la gestion des plantations de pins à Madagascar. L'objectif est d'identifier les facteurs structuraux, environnementaux et dendrométriques influençant la productivité des parcelles. Une approche intégrée combinant la régression linéaire (OLS), l'analyse en composantes principales (ACP), le clustering et l'analyse des corrélations a été mobilisée. Les résultats révèlent une très forte relation entre la superficie et la production ($r = 0,96$), confirmée par un modèle OLS performant ($R^2 = 0,966$; R^2 ajusté = 0,955). La production est également fortement corrélée à l'exploitation ($r = 0,99$), traduisant une relation quasi proportionnelle. L'ACP montre que les deux premiers axes expliquent 66,6 % de la variance totale (39,1 % pour PC1 et 27,5 % pour PC2), mettant en évidence un gradient de production lié à la superficie et un gradient structurel associé à la densité. Le clustering identifie trois groupes de parcelles, allant de faibles productions inférieures à 3500 m³ à des niveaux élevés dépassant 6500 m³. Par ailleurs, le risque de feu présente une corrélation négative modérée avec la production ($r = -0,56$). Ces résultats contribuent à une meilleure compréhension des mécanismes de production et à l'optimisation de la gestion durable des plantations.

Mots-clés : Production forestière ; ACP ; Régression OLS ; Clustering ; FANALAMANGA

Abstract: This study proposes a multidimensional analysis of the determinants of forest production within the FANALAMANGA company, which specializes in the management of pine plantations in Madagascar. The objective is to identify the structural, environmental, and dendrometric factors influencing plot productivity. An integrated approach combining linear regression (OLS), principal component analysis (PCA), clustering, and correlation analysis was employed. The results reveal a very strong relationship between area and production ($r = 0.96$), confirmed by a high-performing OLS model ($R^2 = 0.966$; adjusted $R^2 = 0.955$). Production is also strongly correlated with harvesting ($r = 0.99$), indicating a near-proportional relationship. PCA shows that the first two components explain 66.6% of the total variance (39.1% for PC1 and 27.5% for PC2), highlighting a production gradient linked to area and a structural gradient associated with density. Clustering identifies three groups of plots, ranging from low production levels below 3,500 m³ to high levels exceeding 6,500 m³. Furthermore, fire risk shows a moderate negative correlation with production ($r = -0.56$). These results contribute to a better understanding of production mechanisms and to the optimization of sustainable plantation management.

Keywords: Forest production; PCA; OLS regression; Clustering; FANALAMANGA.

INTRODUCTION

La gestion durable des ressources forestières constitue aujourd'hui un enjeu majeur face aux pressions croissantes exercées par les activités humaines et les changements climatiques. À Madagascar, les plantations forestières jouent un rôle essentiel dans la satisfaction des besoins en bois, la réduction de la pression sur les forêts naturelles et la contribution au développement économique local. Parmi les acteurs clés du secteur, la société FANALAMANGA se distingue par l'importance de ses plantations de pins, qui

couvrent plusieurs dizaines de milliers d'hectares et constituent une base stratégique pour la production de bois industriel. Les plantations de pins, notamment les espèces *Pinus kesiya*, *Pinus elliottii* et *Pinus patula*, sont largement utilisées en raison de leur croissance rapide et de leur adaptabilité aux conditions écologiques locales [1]. Toutefois, la productivité de ces plantations dépend de multiples facteurs, incluant la superficie des parcelles, la densité du peuplement, l'âge des arbres ainsi que les contraintes environnementales telles que les risques d'incendie [2]. L'identification et la compréhension de ces facteurs sont essentielles pour optimiser les rendements et assurer une gestion forestière durable. Dans ce contexte, les méthodes statistiques multivariées apparaissent comme des outils particulièrement adaptés pour analyser la complexité des interactions entre variables forestières. La régression linéaire permet d'évaluer les relations quantitatives entre la production et ses déterminants, tandis que l'analyse en composantes principales (ACP) facilite la réduction de la dimension des données et la mise en évidence des structures sous-jacentes [3]. Par ailleurs, les techniques de classification, telles que le clustering, permettent d'identifier des groupes homogènes de parcelles en fonction de leurs caractéristiques productives et structurales [4]. L'intégration de ces approches offre une vision globale et cohérente des systèmes forestiers, en combinant analyse descriptive, modélisation et interprétation des données [5]. De plus, la prise en compte des dynamiques de croissance, notamment à travers des modèles biologiques, permet de mieux comprendre l'évolution temporelle des peuplements forestiers et d'anticiper leur potentiel de production [6]. Ainsi, l'objectif de cette étude est de réaliser une analyse multidimensionnelle des déterminants de la production forestière au sein de la société FANALAMANGA, en mobilisant une approche intégrée combinant régression, ACP, analyse de corrélation et classification. Cette démarche vise à identifier les principaux facteurs influençant la productivité des plantations de pins et à fournir des éléments d'aide à la décision pour une gestion durable et optimisée des ressources forestières [7].

MATÉRIELS ET MÉTHODES

La présente étude s'appuie sur un ensemble de données issues de sources primaires et secondaires afin d'analyser les déterminants de la production forestière au sein de la société FANALAMANGA. Les données primaires ont été collectées lors de descentes sur terrain effectuées dans les principales zones d'exploitation, notamment à Moramanga, Ambatondrazaka et Amparafaravola. Ces observations ont permis de relever des variables dendrométriques essentielles telles que la superficie des parcelles, l'âge moyen des peuplements, la densité de tiges ainsi que les volumes de production et d'exploitation. Parallèlement, des enquêtes semi-structurées ont été menées auprès des responsables techniques et des agents forestiers afin de recueillir des informations qualitatives relatives aux pratiques sylvicoles et aux contraintes environnementales, en particulier les risques d'incendie. Ces données ont été complétées par des informations secondaires issues de rapports techniques et de documents de gestion forestière, permettant ainsi de constituer une base de données cohérente et exploitable. Le traitement et l'analyse des données ont été réalisés à l'aide du langage de programmation Python, en utilisant la distribution Anaconda qui fournit un environnement intégré pour le calcul scientifique et l'analyse de données [8]. Les scripts ont été développés dans l'interface Jupyter Notebook, facilitant la reproductibilité des analyses et l'intégration des résultats [9]. La manipulation et la structuration des données ont été assurées par la bibliothèque pandas [10], tandis que les opérations numériques ont été effectuées à l'aide de NumPy [11]. Les visualisations graphiques ont été réalisées avec Matplotlib et Seaborn, permettant de représenter de manière claire les relations entre variables [12]. Les analyses statistiques, notamment la régression linéaire, ont été conduites avec la bibliothèque Statsmodels [13], et les méthodes d'analyse multivariée telles que l'analyse en composantes principales (ACP) et le clustering ont été implémentées à l'aide de Scikit-learn [14]. Les données collectées ont fait l'objet d'un processus de nettoyage et de validation visant à éliminer les valeurs aberrantes et à corriger les incohérences éventuelles. Une phase de transformation a ensuite été réalisée, incluant la normalisation des variables quantitatives et l'encodage des variables catégorielles, notamment la variable « zone », sous forme de variables indicatrices afin de permettre leur intégration dans les modèles statistiques. L'analyse des données s'est appuyée sur une combinaison de méthodes complémentaires. La régression linéaire ordinaire (OLS) a été utilisée pour modéliser la relation entre la production forestière et ses principaux déterminants, en évaluant l'effet de variables telles que la superficie, l'âge et la densité des peuplements. L'analyse en composantes principales a permis de réduire la dimension du jeu de données et de mettre en évidence les structures de corrélation entre variables, facilitant ainsi l'interprétation des relations multivariées. Par ailleurs, une classification non supervisée par la méthode des k-moyennes a été appliquée afin de regrouper les parcelles en ensembles homogènes selon leurs caractéristiques productives et structurales. Enfin, un modèle de croissance logistique a été mobilisé pour décrire la dynamique

temporelle des plantations, en représentant l'évolution du volume de bois en fonction de l'âge des peuplements, conformément aux approches classiques de modélisation de la croissance forestière [15].

RESULTATS ET DISCUSSIONS

1- Analyse des corrélations entre variables dendrométriques, environnementales et productives des plantations de pins

La figure suivante présente une matrice de corrélation sous forme de heatmap, illustrant les relations linéaires entre les principales variables caractérisant les plantations de pins. Cette représentation permet d'identifier l'intensité et le sens des corrélations entre les facteurs structuraux, écologiques et productifs, facilitant ainsi la compréhension des interactions qui influencent la production forestière.

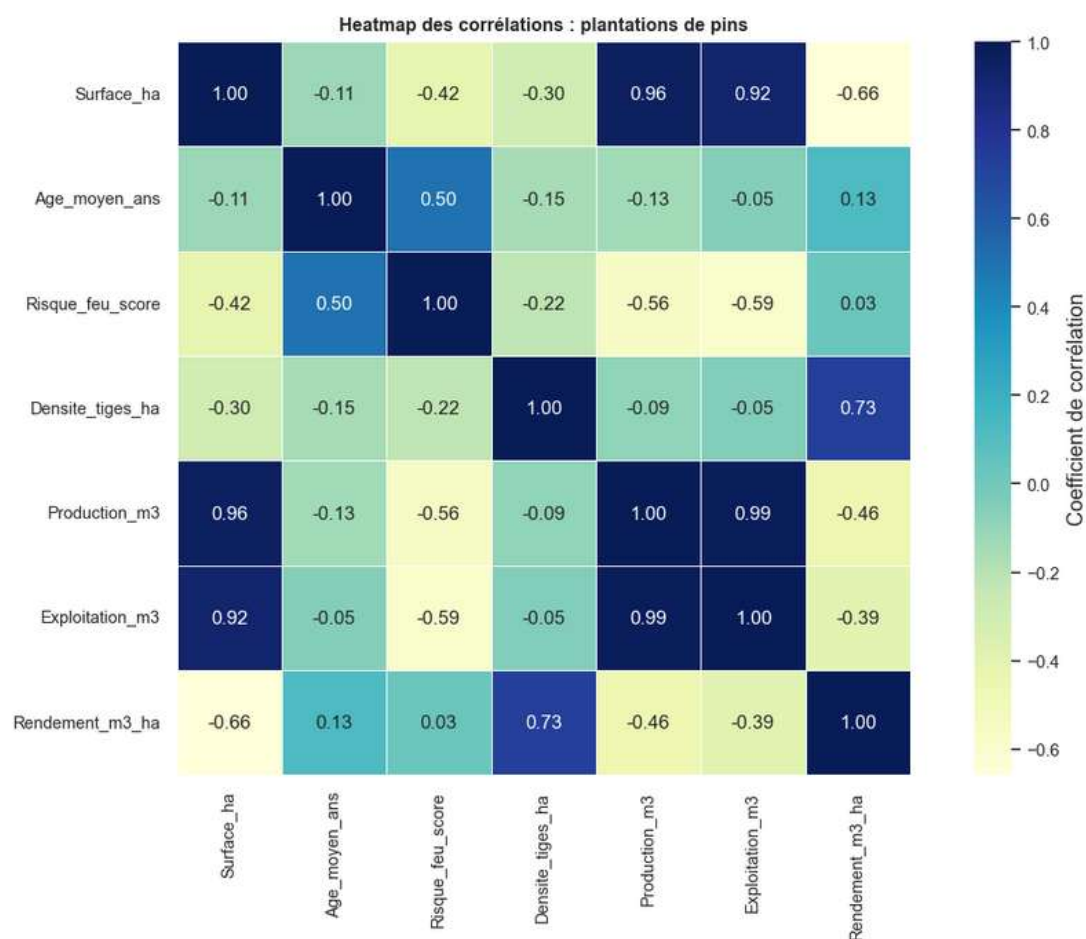


Figure 1: Analyse des corrélations entre variables dendrométriques, environnementales et productives des plantations de pins

Source : Auteur

Interprétation

L'analyse de la matrice de corrélation met en évidence une très forte corrélation positive entre la production et l'exploitation, avec un coefficient de 0,99, indiquant une relation quasi proportionnelle entre ces deux variables. Cette forte

dépendance suggère que le volume exploité est directement déterminé par le niveau de production, ce qui est cohérent avec un système d'exploitation intensif et bien structuré. La superficie présente également une corrélation très élevée avec la production (0,96) et avec l'exploitation (0,92), confirmant son rôle central dans la détermination du volume total de bois produit. Cette relation traduit un effet d'échelle important, où l'augmentation de la surface forestière se traduit directement par une augmentation de la production globale. En revanche, le rendement en m^3 par hectare présente une corrélation négative relativement forte avec la superficie (-0,66), indiquant que les grandes parcelles tendent à avoir une efficacité productive moindre à l'unité de surface. Cette relation peut s'expliquer par des contraintes de gestion, une hétérogénéité accrue des peuplements ou une densité plus faible dans les grandes surfaces. La densité de tiges montre une corrélation positive significative avec le rendement (0,73), ce qui indique que les peuplements plus denses tendent à produire davantage de volume par hectare. Cette relation souligne l'importance de la structure interne du peuplement dans l'optimisation de la productivité surfacique. Le risque de feu présente des corrélations négatives modérées avec la production (-0,56) et l'exploitation (-0,59), ce qui suggère que les zones les plus exposées aux incendies ont des performances productives plus faibles. Cette influence négative met en évidence le rôle des facteurs environnementaux dans la limitation de la production forestière. L'âge moyen des peuplements montre des corrélations relativement faibles avec la production (-0,13) et l'exploitation (-0,05), indiquant que, dans ce jeu de données, l'âge ne constitue pas un facteur déterminant de la production globale. Toutefois, sa corrélation positive avec le risque de feu (0,50) suggère que les peuplements plus âgés pourraient être plus exposés aux risques d'incendie, possiblement en raison de l'accumulation de biomasse inflammable.

2- Relation entre la superficie des parcelles et la production forestière des plantations de pins

La figure ci-dessous illustre la relation entre la superficie des parcelles forestières et leur niveau de production en bois, à travers un nuage de points accompagné d'une droite de régression linéaire. Cette représentation permet d'évaluer l'existence d'une relation fonctionnelle entre ces deux variables clés de la production sylvicole, tout en tenant compte de la variabilité liée aux différentes zones géographiques étudiées.

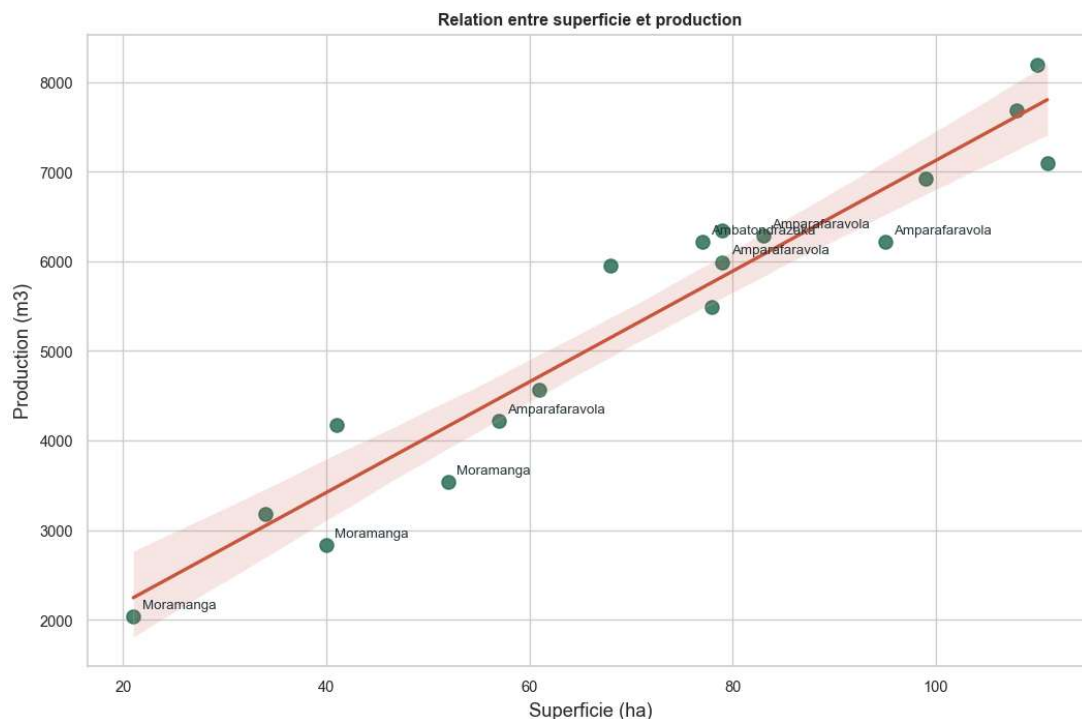


Figure 2: Relation entre la superficie des parcelles et la production forestière des plantations de pins

Source : Auteur

Interprétation

L'analyse graphique met en évidence une relation linéaire positive marquée entre la superficie des parcelles et la production en bois, indiquant que l'augmentation de la surface exploitée s'accompagne d'une hausse significative du volume produit. La pente ascendante de la droite de régression confirme l'existence d'un lien direct entre ces deux variables, traduisant un effet d'échelle caractéristique des systèmes de production forestière. Les parcelles de faible superficie, situées autour de 20 à 40 hectares, présentent des niveaux de production relativement faibles, généralement compris entre 2000 et 3500 m³. À l'inverse, les parcelles de grande superficie, supérieures à 80 hectares, affichent des volumes de production nettement plus élevés, atteignant des valeurs comprises entre 6000 et plus de 8000 m³. Cette progression quasi proportionnelle suggère que la superficie constitue un facteur déterminant de la production totale. Cependant, la dispersion des points autour de la droite de régression, bien que modérée, indique l'existence d'une variabilité non négligeable entre les parcelles de taille similaire. Par exemple, pour des superficies proches de 80 hectares, la production varie approximativement entre 5500 et 6500 m³, ce qui traduit l'influence d'autres facteurs tels que la densité du peuplement, l'âge des arbres, les conditions écologiques ou encore les pratiques sylvicoles. L'observation des annotations révèle également des différences entre zones géographiques. Les parcelles localisées à Amparafaravola tendent à se situer dans les niveaux de production les plus élevés pour des superficies comparables, suggérant des conditions favorables ou une meilleure gestion. À l'inverse, certaines parcelles de Moramanga présentent des productions légèrement inférieures à celles attendues selon la tendance générale, ce qui pourrait refléter des contraintes locales ou des différences de structure du peuplement.

3- Analyse en composantes principales des caractéristiques dendrométriques et productives des plantations de pins : cercle des corrélations et structuration des individus

La figure ci-dessous présente les résultats de l'analyse en composantes principales (ACP) appliquée aux variables dendrométriques, environnementales et productives des plantations de pins. Elle combine, d'une part, le cercle des corrélations permettant d'identifier les relations entre variables et axes factoriels, et, d'autre part, le graphe des individus illustrant la distribution des parcelles selon ces mêmes axes. Les deux premières composantes principales expliquent respectivement 48,3 % et 21,0 % de la variance totale, soit un cumul de 69,3 %, ce qui garantit une représentation satisfaisante de la structure des données dans le plan factoriel.

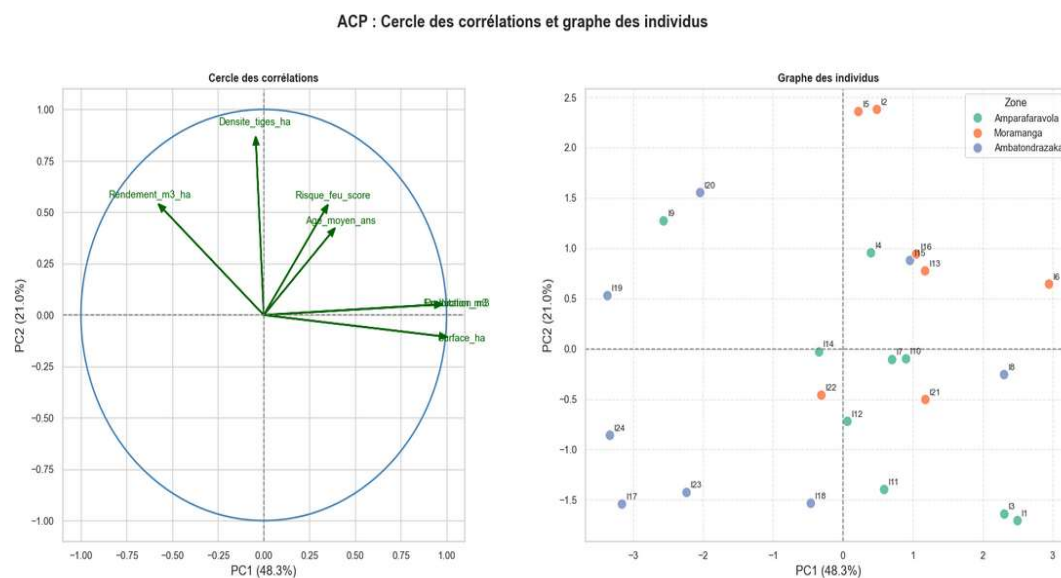


Figure 3: Analyse en composantes principales des caractéristiques dendrométriques et productives des plantations de pins : cercle des corrélations et structuration des individus

Source : Auteur

Interprétation

L'analyse du cercle des corrélations montre que la première composante principale (PC1), qui explique 48,3 % de la variance, est fortement associée aux variables de production et de structure du peuplement. La production en m³ et la superficie présentent des corrélations positives élevées avec cet axe, indiquant que PC1 traduit principalement un gradient d'intensité productive et d'extension des parcelles. Cette relation suggère que les parcelles les plus étendues tendent à générer des volumes de production plus importants, ce qui est cohérent avec les principes de production forestière extensive. La seconde composante principale (PC2), expliquant 21,0 % de la variance, est davantage liée aux variables écologiques et structurales telles que la densité de tiges et le rendement en m³ par hectare. La forte projection de la densité sur cet axe indique que PC2 reflète un gradient de structure interne du peuplement, opposant des peuplements denses à forte productivité surfacique à des peuplements plus ouverts. Le rendement, corrélé positivement à PC2, confirme que les peuplements plus denses tendent à présenter une meilleure efficacité productive par unité de surface. Les variables d'âge moyen et de risque de feu présentent des corrélations intermédiaires avec les deux axes, traduisant leur rôle secondaire mais non négligeable dans la structuration des données. Leur position suggère qu'elles influencent simultanément la production globale et la dynamique interne du peuplement, sans constituer des facteurs dominants. Le graphe des individus met en évidence une structuration spatiale des parcelles en fonction de leur localisation géographique. Les individus situés à droite du plan factoriel, caractérisés par des valeurs positives de PC1, correspondent aux parcelles à forte production et grande superficie, tandis que ceux situés à gauche présentent des niveaux plus faibles de production. Par ailleurs, la dispersion verticale selon PC2 traduit des différences de densité et de rendement entre les parcelles. On observe également une certaine différenciation entre les zones géographiques, bien que celle-ci ne soit pas strictement séparée. Les parcelles appartenant à Moramanga tendent à se regrouper dans les zones de production plus élevée, alors que celles d'Ambatondrazaka présentent une plus grande dispersion, traduisant une hétérogénéité des conditions de production. Les parcelles d'Amparafaravola apparaissent réparties de manière intermédiaire, suggérant une diversité de structures et de performances.

4- Approche intégrée de modélisation statistique et d'analyse multidimensionnelle des plantations de pins : régression OLS, ACP, classification et dynamique de croissance

La figure suivante présente une analyse intégrée des plantations de pins à travers plusieurs approches complémentaires, combinant un modèle de régression linéaire avec effet de zone, une analyse en composantes principales (ACP), une classification non supervisée par clustering et un modèle de croissance biologique. Cette approche multidimensionnelle permet d'explorer simultanément les relations entre variables, la structuration des parcelles forestières et la dynamique de production en fonction du temps.

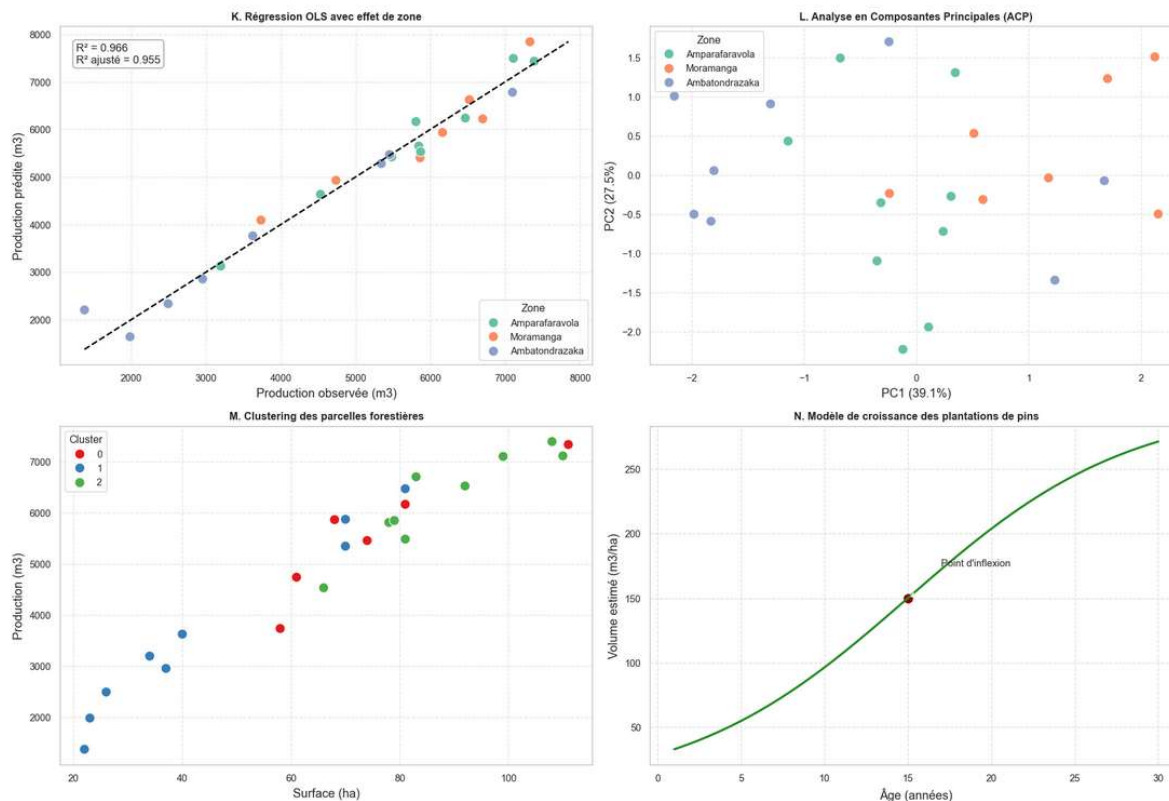


Figure 4: Approche intégrée de modélisation statistique et d'analyse multidimensionnelle des plantations de pins : régression OLS, ACP, classification et dynamique de croissance

Source : Auteur

Interprétation

Le modèle de régression OLS avec effet de zone met en évidence une très forte capacité explicative, avec un coefficient de détermination R^2 de 0,966 et un R^2 ajusté de 0,955, indiquant que plus de 96 % de la variabilité de la production observée est expliquée par les variables incluses dans le modèle. La forte proximité des points par rapport à la droite d'ajustement confirme la robustesse du modèle et la pertinence des variables explicatives, notamment la superficie, l'âge moyen et la densité. La distribution des points selon les zones géographiques montre également que l'effet spatial contribue à expliquer les différences de production, bien que celles-ci restent globalement cohérentes au sein du modèle. L'analyse en composantes principales montre que les deux premiers axes expliquent respectivement 39,1 % et 27,5 % de la variance, soit un total de 66,6 %, ce qui permet une représentation satisfaisante de la structure des données. La dispersion des individus dans le plan factoriel met en évidence une structuration partielle selon les zones, sans séparation stricte, traduisant une certaine homogénéité des caractéristiques des plantations tout en conservant des différences locales. Les individus situés sur les extrémités des axes correspondent aux parcelles présentant des caractéristiques extrêmes en termes de production ou de structure. Le clustering des parcelles forestières permet d'identifier trois groupes distincts. Le premier groupe, caractérisé par des surfaces faibles comprises approximativement entre 20 et 40 hectares et des productions inférieures à 3500 m³, correspond à des parcelles à faible intensité productive. Le second groupe regroupe des parcelles intermédiaires, avec des surfaces comprises entre 50 et 80 hectares et des productions variant entre 4000 et 6000 m³. Le troisième groupe se distingue par des surfaces élevées, supérieures à 80 hectares, associées à des productions supérieures à 6500 m³. Cette classification met en évidence un gradient clair de productivité lié à la superficie et à la structure des peuplements. Le modèle de

croissance logistique met en évidence une dynamique typique des plantations forestières, caractérisée par une phase initiale de croissance lente, suivie d'une phase d'accélération puis d'un ralentissement à l'approche de la capacité maximale. Le point d'inflexion, situé autour de 15 ans pour un volume estimé d'environ 150 m³/ha, correspond à la phase de croissance maximale, où l'accroissement du volume est le plus rapide. Au-delà de ce stade, la croissance ralentit progressivement pour atteindre une asymptote proche de 270 à 280 m³/ha vers 30 ans.

DISCUSSION

Les résultats obtenus dans cette étude mettent en évidence le rôle prépondérant des facteurs structuraux dans la détermination de la production forestière au sein de la société FANALAMANGA. La forte relation observée entre la superficie des parcelles et la production, confirmée à la fois par l'analyse de corrélation ($r = 0,96$) et par le modèle de régression ($R^2 = 0,966$), souligne l'importance de l'effet d'échelle dans les systèmes de plantations forestières. Ce résultat est cohérent avec les principes fondamentaux de la production sylvicole, selon lesquels l'augmentation de la surface exploitée conduit mécaniquement à une augmentation du volume total produit, bien que cela n'implique pas nécessairement une amélioration de la productivité à l'unité de surface. L'influence de la densité de tiges sur le rendement, mise en évidence par une corrélation positive significative ($r = 0,73$), confirme le rôle essentiel de la structure du peuplement dans l'optimisation de la production forestière. Des peuplements plus denses permettent une meilleure utilisation de l'espace et des ressources disponibles, ce qui se traduit par une augmentation du rendement en m³/ha. Toutefois, cette relation doit être interprétée avec prudence, car une densité excessive peut également entraîner une compétition accrue entre les arbres et, à terme, une diminution de la croissance individuelle. Le risque de feu apparaît comme un facteur limitant non négligeable, avec une corrélation négative modérée avec la production ($r = -0,56$). Ce résultat met en évidence la vulnérabilité des plantations de pins face aux incendies, en particulier dans des formations monospécifiques caractérisées par une forte inflammabilité. Cette contrainte environnementale souligne la nécessité de mettre en place des stratégies de gestion adaptées, notamment en matière de prévention et de lutte contre les feux de forêt. L'analyse en composantes principales apporte un éclairage complémentaire en révélant une structuration des variables autour de deux axes principaux expliquant 66,6 % de la variance totale. Le premier axe, associé à la production et à la superficie, traduit un gradient d'intensité productive, tandis que le second axe, lié à la densité et au rendement, reflète la structure interne des peuplements. Cette organisation confirme que la production forestière résulte de l'interaction entre des facteurs extensifs, liés à la surface, et des facteurs intensifs, liés à la structure du peuplement. Le clustering des parcelles permet de distinguer trois groupes homogènes correspondant à différents niveaux de productivité. Les parcelles à faible production, inférieure à 3500 m³, se caractérisent généralement par des surfaces réduites et des conditions moins favorables, tandis que les parcelles à forte production, dépassant 6500 m³, sont associées à des surfaces plus importantes et à une structure de peuplement plus optimale. Cette classification met en évidence l'existence de profils types de parcelles, qui pourraient être utilisés pour orienter les stratégies de gestion et d'aménagement. Par ailleurs, le modèle de croissance logistique met en évidence une dynamique de production typique des plantations forestières, avec un point d'inflexion situé autour de 15 ans et un volume estimé d'environ 150 m³/ha. Ce résultat indique que la phase de croissance maximale se situe à un stade intermédiaire du cycle de développement, ce qui constitue un élément clé pour la planification des opérations sylvicoles, notamment les éclaircies et les coupes. Dans l'ensemble, ces résultats confirment que la production forestière au sein de FANALAMANGA est principalement déterminée par des facteurs structurels, tout en étant influencée par des contraintes environnementales et des dynamiques biologiques. L'approche multidimensionnelle adoptée dans cette étude permet ainsi de mieux comprendre la complexité des systèmes de plantation et de fournir des éléments pertinents pour l'amélioration des pratiques de gestion durable. Toutefois, certaines limites doivent être soulignées, notamment le caractère illustratif de certaines données utilisées, qui pourrait affecter la généralisation des résultats. Des études futures basées sur des données empiriques plus étendues permettraient d'affiner les modèles et de renforcer la robustesse des conclusions.

CONCLUSION

Cette étude avait pour objectif d'analyser de manière multidimensionnelle les déterminants de la production forestière au sein de la société FANALAMANGA, en mobilisant une approche intégrée combinant des méthodes statistiques et analytiques. Les résultats obtenus ont permis de mettre en évidence le rôle central des facteurs structuraux, notamment la superficie des parcelles et la densité de tiges, dans la détermination de la production forestière. L'analyse de régression a montré une forte capacité explicative

du modèle, avec un coefficient de détermination élevé ($R^2 = 0,966$), confirmant la pertinence des variables retenues pour expliquer la variabilité de la production. Les analyses de corrélation ont également révélé des relations significatives, en particulier entre la superficie et la production ($r = 0,96$), ainsi qu'entre la densité et le rendement ($r = 0,73$), tandis que le risque de feu apparaît comme un facteur limitant avec une influence négative modérée ($r = -0,56$). L'analyse en composantes principales a permis de réduire la complexité du jeu de données en identifiant deux axes principaux expliquant 66,6 % de la variance totale, traduisant respectivement un gradient de production lié à la superficie et un gradient structurel associé à la densité. Le clustering a mis en évidence trois groupes distincts de parcelles correspondant à différents niveaux de productivité, soulignant l'existence de profils homogènes au sein des plantations. Par ailleurs, le modèle de croissance logistique a permis de caractériser la dynamique temporelle des peuplements, avec un point d'inflexion situé autour de 15 ans, correspondant à la phase de croissance maximale. Ces résultats apportent une contribution importante à la compréhension des mécanismes de production des plantations de pins et offrent des perspectives concrètes pour l'amélioration des pratiques de gestion forestière. Ils suggèrent notamment la nécessité d'optimiser la structure des peuplements, de maîtriser les risques environnementaux et d'adapter les stratégies d'exploitation en fonction des stades de développement des plantations. Cependant, certaines limites liées à la nature des données utilisées invitent à la prudence dans l'interprétation des résultats. Des travaux futurs basés sur des données empiriques plus étendues et intégrant d'autres variables, telles que les facteurs climatiques et édaphiques, permettraient de renforcer la robustesse des analyses et d'affiner les recommandations pour une gestion durable et efficiente des ressources forestières au sein de FANALAMANGA.

REFERENCE

- [1] Evans, J. & Turnbull, J.W. (2004). *Plantation Forestry in the Tropics*. Oxford University Press.
- [2] FAO (2015). *Global Forest Resources Assessment 2015*. Food and Agriculture Organization of the United Nations.
- [3] Jolliffe, I.T. (2002). *Principal Component Analysis*. Springer.
- [4] Kaufman, L. & Rousseeuw, P.J. (2005). *Finding Groups in Data: An Introduction to Cluster Analysis*. Wiley.
- [5] Hair, J.F., Black, W.C., Babin, B.J. & Anderson, R.E. (2010). *Multivariate Data Analysis*. Pearson.
- [6] Vanclay, J.K. (1994). *Modelling Forest Growth and Yield*. CAB International.
- [7] Pretzsch, H. (2009). *Forest Dynamics, Growth and Yield*. Springer.
- [8] Anaconda Inc. (2020). *Anaconda Software Distribution*. DOI: <https://doi.org/10.5281/zenodo.3632056>
- [9] Kluyver, T. et al. (2016). *Jupyter Notebooks – a publishing format for reproducible computational workflows*. DOI: <https://doi.org/10.3233/978-1-61499-649-1-87>
- [10] McKinney, W. (2010). *Data Structures for Statistical Computing in Python*. DOI: <https://doi.org/10.25080/Majora-92bf1922-00a>
- [11] Harris, C.R. et al. (2020). *Array programming with NumPy*. DOI: <https://doi.org/10.1038/s41586-020-2649-2>
- [12] Waskom, M.L. (2021). *Seaborn: statistical data visualization*. DOI: <https://doi.org/10.21105/joss.03021>
- [13] Seabold, S. & Perktold, J. (2010). *Statsmodels: Econometric and statistical modeling with Python*. DOI: <https://doi.org/10.25080/Majora-92bf1922-011>
- [14] Pedregosa, F. et al. (2011). *Scikit-learn: Machine Learning in Python*. DOI: <https://doi.org/10.5555/1953048.2078195>
- [15] Richards, F.J. (1959). *A flexible growth function for empirical use*. DOI: <https://doi.org/10.1093/jxb/10.2.290>