



ANALYSE GÉOSPATIALE, SPECTRALE ET MULTIVARIÉE DU NDVI DANS LE BASSIN VERSANT DE BETSIBOKA

¹ RASOANAINA Jacques, ² RAKOTOSON Andriatiana Tolontsoa, ³ TOVONIRINA Mamiharizo Jackie, ⁴ RAZAFIARISERA RALAY TIANA, ⁵ RASOLOMANANA Eddy Harilala

Ecole Doctorale en INGENIERIE ET GEOSCIENCES-UNIVERSITE D'ANTANANARIVO

RESUME : Cette étude analyse la variabilité spatiale, spectrale et multivariée du NDVI dans le bassin versant de Betsiboka à partir de données satellitaires et d'outils géostatistiques. L'interpolation spatiale met en évidence des valeurs de NDVI comprises entre 0,375 et 0,825, révélant une forte hétérogénéité de la couverture végétale. Les valeurs élevées, supérieures à 0,60, caractérisent principalement les secteurs à végétation dense, tandis que les valeurs inférieures à 0,45 signalent des zones potentiellement dégradées et plus sensibles à l'érosion hydrique. L'analyse spectrale montre une relation linéaire très forte entre les réflectances des bandes rouge M05 et proche infrarouge M07, avec un coefficient de détermination de $R^2 = 0,943$. La comparaison entre le NDVI fourni et le NDVI recalculé indique également une cohérence satisfaisante, avec $R^2 = 0,778$, malgré certains écarts liés aux effets atmosphériques, aux pixels mixtes et aux différences de traitement radiométrique. L'analyse en composantes principales explique 66,4 % de la variabilité totale, dont 42,9 % pour l'axe 1 et 23,5 % pour l'axe 2. Les profils standardisés, variant de -3 à +3, distinguent clairement les sites à végétation dense des zones dominées par de faibles NDVI.

Mots-clés : NDVI ; télédétection ; analyse géospatiale ; analyse spectrale ; bassin versant de Betsiboka.

ABSTRACT: This study examines the spatial, spectral, and multivariate variability of the Normalized Difference Vegetation Index (NDVI) in the Betsiboka River basin using satellite data and geostatistical methods. Spatial interpolation reveals NDVI values ranging from 0.375 to 0.825, indicating marked heterogeneity in vegetation cover. High values above 0.60 mainly characterize densely vegetated areas, whereas values below 0.45 identify potentially degraded zones that may be more vulnerable to water erosion. Spectral analysis shows a very strong linear relationship between reflectance in the red M05 band and the near-infrared M07 band, with a coefficient of determination of $R^2 = 0.943$. The comparison between the supplied NDVI and the NDVI recalculated from M05 and M07 also indicates satisfactory consistency, with $R^2 = 0.778$, despite discrepancies associated with atmospheric effects, mixed pixels, and radiometric processing differences. Principal component analysis explains 66.4% of the total variability, including 42.9% for the first component and 23.5% for the second. Standardized profiles ranging from -3 to +3 clearly distinguish densely vegetated sites from areas dominated by low NDVI values.

Keywords: NDVI; remote sensing; geospatial analysis; spectral analysis; Betsiboka River basin.

Digital Object Identifier (DOI): <https://doi.org/10.5281/zenodo.21739872>

1 Introduction

La couverture végétale joue un rôle essentiel dans le fonctionnement hydrologique et écologique des bassins versants. Elle intervient dans l'interception des précipitations, l'évapotranspiration, l'infiltration, le ruissellement et la protection des sols contre l'érosion. Dans les régions tropicales soumises à une forte variabilité climatique et à des changements rapides d'occupation du sol, son suivi constitue donc un enjeu important pour la gestion durable des ressources naturelles. La télédétection satellitaire permet d'observer la végétation de manière régulière, homogène et sur de vastes superficies. Parmi les indicateurs disponibles, l'indice de végétation par différence normalisée, ou NDVI, est largement utilisé pour caractériser l'activité photosynthétique et la densité relative du couvert végétal. Il repose sur le contraste entre l'absorption du rayonnement rouge par la chlorophylle et la forte réflectance de la végétation dans le proche infrarouge. Son interprétation reste toutefois influencée par les propriétés du sol, les conditions atmosphériques, la structure du couvert et les traitements radiométriques appliqués aux images satellitaires (Huete et al., 2002) [1]. Le NDVI constitue également un indicateur pertinent pour étudier les variations de productivité végétale, les perturbations du couvert et certaines formes de dégradation des terres. Cependant, sa fiabilité dépend directement de la qualité des observations. Les nuages, les ombres, les brumes, les pixels mixtes ou les erreurs de calibration peuvent introduire des biais dans les valeurs obtenues. La comparaison entre un NDVI fourni par un produit satellitaire et un NDVI recalculé à partir des bandes rouges et proche infrarouge représente ainsi une étape importante de validation méthodologique (Pettorelli et al., 2005) [2]. À Madagascar, la déforestation, les feux, l'extension agricole et la fragmentation des formations naturelles modifient profondément les paysages et les services écosystémiques. Ces transformations renforcent la nécessité d'un suivi spatial régulier de la végétation afin de mieux comprendre les trajectoires de changement et d'appuyer les politiques de conservation et d'aménagement du territoire (Harper et al., 2007) [3]. Le bassin versant de Betsiboka est particulièrement concerné par ces enjeux. Il est marqué par une forte érosion des sols et par d'importants transferts de sédiments depuis les hautes terres jusqu'à l'estuaire de Betsiboka et à la baie de Bombetoka. Ces processus résultent de l'interaction entre les précipitations, le relief, les caractéristiques des sols, les pratiques d'occupation des terres et l'état de la couverture végétale. Dans ce contexte, l'analyse du NDVI peut contribuer à identifier les secteurs faiblement protégés par la végétation et potentiellement exposés au ruissellement et à l'érosion hydrique. L'intérêt des données satellitaires multitemporelles pour le suivi environnemental de ce système hydrographique a déjà été démontré (Raharimahefa et Kusky, 2010) [4]. Toutefois, l'utilisation du seul NDVI moyen peut masquer la diversité des situations écologiques et radiométriques présentes dans le bassin. Des valeurs moyennes similaires peuvent correspondre à des proportions différentes de végétation dense, de couverture clairsemée ou de sols nus. De plus, la disponibilité et la qualité des pixels exploitables peuvent varier selon les secteurs. Il est donc nécessaire d'associer l'analyse spatiale du NDVI, l'étude de la relation entre les bandes spectrales, la vérification de la cohérence des indices et l'examen conjoint de plusieurs indicateurs. L'analyse en composantes principales constitue, à cet égard, une méthode adaptée pour résumer des variables corrélées et identifier les principaux gradients de différenciation entre les unités spatiales. Combinée à la standardisation des données, elle facilite la mise en évidence des similitudes, des oppositions et des profils caractéristiques de végétation et de qualité des observations satellitaires (Abdi et Williams, 2010) [5]. Cette étude vise ainsi à analyser les dimensions géospatiale, spectrale et multivariée du NDVI dans le bassin versant de Betsiboka. Elle cherche à caractériser sa distribution spatiale, à examiner la relation entre les réflectances des bandes M05 et M07, à évaluer la cohérence entre le NDVI fourni et le NDVI recalculé, puis à identifier les principales structures de variation des indicateurs de végétation et de qualité des données. Cette approche doit contribuer à renforcer la fiabilité du suivi satellitaire de la végétation et à améliorer la compréhension des relations entre couverture végétale, fonctionnement hydroécologique et vulnérabilité des sols à l'érosion.

2 Matériels et méthodes

2.1 Matériels

Les données géospatiales utilisées dans cette étude proviennent principalement du Foiben-Taosarintanin'i Madagasikara, organisme national chargé de la production et de la gestion de l'information géographique à Madagascar. Elles comprennent les limites administratives, le réseau hydrographique, les localités, les axes de communication ainsi que les fonds topographiques nécessaires à la délimitation et à la caractérisation du bassin

versant de Betsiboka. Avant leur exploitation, les couches vectorielles ont fait l'objet d'un contrôle de cohérence géométrique et topologique, puis ont été harmonisées dans un système commun de coordonnées géographiques et projetées. Elles ont ensuite été découpées selon l'emprise du bassin versant et mobilisées pour la localisation des unités spatiales, l'extraction des centroïdes et la représentation cartographique des indicateurs environnementaux [6]. Les données de télédétection reposent sur des observations multispectrales acquises par le capteur VIIRS. Elles comprennent notamment les réflectances enregistrées dans la bande rouge M05 et dans la bande du proche infrarouge M07, ainsi que les valeurs de NDVI fournies par le produit satellitaire. Les couches relatives à la qualité des observations et à la fiabilité des pixels ont également été prises en compte. Le NDVI a été recalculé à partir du contraste normalisé entre les réflectances du proche infrarouge et du rouge, sans recourir directement aux valeurs fournies par le produit. Cette procédure a permis d'évaluer la cohérence entre les deux séries d'indices. Les pixels affectés par les nuages, les ombres, les valeurs de remplissage ou les anomalies radiométriques ont été identifiés à partir des indicateurs de qualité, puis exclus des traitements statistiques et cartographiques [7]. L'ensemble des traitements numériques et géospatiaux a été réalisé dans un environnement Python 3. La bibliothèque Rasterio a été utilisée pour lire, écrire, découper, masquer et reprojeter les données raster, ainsi que pour extraire les valeurs spectrales correspondant aux différentes unités spatiales étudiées [8]. La gestion des données vectorielles a été assurée avec GeoPandas, conjointement avec Shapely et PyProj. Ces bibliothèques ont permis d'effectuer les transformations de systèmes de coordonnées, les jointures spatiales, les calculs géométriques, la détermination des centroïdes et l'agrégation des statistiques zonales [9]. Les bibliothèques NumPy et Pandas ont servi à organiser, nettoyer et traiter les matrices ainsi que les tableaux de données. SciPy a été mobilisée pour l'interpolation spatiale et certains calculs statistiques, tandis que Matplotlib et Seaborn ont permis de produire les cartes, les diagrammes de densité, les graphiques de régression et les cartes thermiques. La standardisation des variables et l'analyse en composantes principales ont été effectuées avec Scikit-learn afin d'identifier les principaux gradients de variation et les similitudes entre les unités géographiques étudiées [10].

2.2 Méthodes

La démarche méthodologique a reposé sur l'intégration de traitements radiométriques, géostatistiques et multivariés appliqués aux données satellitaires couvrant le bassin versant de Betsiboka. Après l'harmonisation des systèmes de coordonnées, les images ont été découpées selon l'emprise de la zone d'étude. Les observations affectées par les nuages, les ombres, les valeurs de remplissage ou des anomalies radiométriques ont été exclues à partir des indicateurs de qualité. Les réflectances des bandes M05 et M07 ont ensuite été extraites pour chaque pixel considéré comme exploitable. Le NDVI a été recalculé à partir du contraste normalisé entre la réflectance du proche infrarouge M07 et celle du rouge M05. Cette méthode repose sur le comportement spectral de la végétation chlorophyllienne, qui absorbe fortement le rayonnement rouge tout en réfléchissant davantage le proche infrarouge. Les valeurs obtenues ont été contrôlées afin d'écarter les observations situées en dehors du domaine théorique de l'indice ou associées à des données spectrales invalides. Les indicateurs de NDVI moyen, de faible NDVI et de NDVI dense ont ensuite été calculés par agrégation spatiale des pixels valides à l'intérieur des unités géographiques étudiées (Tucker, 1979) [11]. La distribution continue du NDVI moyen a été estimée à partir des valeurs associées aux points géoréférencés au moyen d'une interpolation spatiale bidimensionnelle pondérée par la distance. Dans ce modèle, chaque valeur estimée dépend des observations voisines, auxquelles un poids décroissant est attribué à mesure que leur distance au point d'estimation augmente. Les observations les plus proches exercent ainsi une influence plus importante que les points éloignés. Une grille régulière a été générée sur l'ensemble du bassin, puis limitée à son contour géographique afin d'éviter la représentation de valeurs interpolées en dehors de la zone d'étude. Cette méthode permet de restituer les principaux gradients spatiaux du NDVI, mais les estimations restent plus incertaines dans les zones périphériques ou éloignées des points disponibles (Shepard, 1968) [12]. La relation entre les réflectances M05 et M07 a été étudiée au moyen d'un modèle de régression linéaire simple estimé par la méthode des moindres carrés ordinaires. La réflectance du proche infrarouge a été considérée comme variable dépendante, tandis que la réflectance du rouge constituait la variable explicative. La pente et l'ordonnée à l'origine ont permis de décrire le sens et l'intensité de la variation conjointe des deux bandes. Le coefficient de corrélation de Pearson a été utilisé pour mesurer le degré d'association linéaire, tandis que le coefficient de détermination a servi à quantifier la proportion de variabilité expliquée par le modèle. Une représentation par mailles hexagonales a été privilégiée afin de visualiser la densité des observations sans superposition excessive des pixels. La cohérence entre le NDVI fourni par le produit satellitaire et le NDVI

recalculé a également été évaluée par une régression linéaire. La droite ajustée a été comparée à la ligne théorique d'égalité, correspondant à une concordance parfaite entre les deux séries. Une pente proche de l'unité et une ordonnée à l'origine proche de zéro indiquent une forte cohérence, tandis qu'un écart entre les deux droites traduit un biais systématique. Le coefficient de détermination a été utilisé pour apprécier la concordance globale, alors que la dispersion des observations autour de la ligne d'égalité a permis d'identifier d'éventuelles différences liées aux pixels mixtes, aux corrections atmosphériques ou aux traitements radiométriques (Montgomery, Peck et Vining, 2012) [13]. L'analyse multivariée a été réalisée à partir d'une matrice regroupant le NDVI moyen, la proportion de faibles NDVI, la proportion de NDVI denses, la qualité des observations et la disponibilité des pixels exploitables. Comme ces indicateurs possédaient des amplitudes et des dispersions différentes, ils ont été préalablement centrés et réduits. Chaque valeur a ainsi été exprimée relativement à la moyenne et à l'écart-type de la variable correspondante. Cette standardisation a empêché les variables présentant les plus fortes variances de dominer artificiellement l'analyse. Une analyse en composantes principales a ensuite été appliquée à la matrice standardisée. Cette méthode transforme les variables initiales corrélées en composantes orthogonales successives, construites de manière à expliquer une proportion décroissante de la variance totale. Les premières composantes ont été retenues pour représenter les principaux gradients de différenciation entre les unités géographiques. Les coordonnées factorielles ont permis de positionner les sites, tandis que les charges factorielles ont servi à interpréter la contribution et l'orientation des indicateurs. La proximité entre deux sites traduit une similarité de leurs profils, alors que l'opposition entre deux vecteurs indique une relation négative entre les variables correspondantes (Jolliffe, 2002) [14]. Les profils standardisés ont enfin été représentés sous la forme d'une carte thermique. La similarité entre les unités géographiques a été déterminée à partir des distances euclidiennes calculées sur les variables centrées et réduites. Un regroupement hiérarchique ascendant selon la méthode de Ward a été employé pour ordonner les sites. À chaque étape, cette procédure fusionne les groupes dont la réunion entraîne la plus faible augmentation de la variance interne. L'organisation des lignes de la carte thermique met ainsi en évidence les sites caractérisés par des profils comparables de végétation et de qualité des observations. Les valeurs positives correspondent à des niveaux supérieurs à la moyenne générale, les valeurs négatives à des niveaux inférieurs, et les valeurs proches de zéro à des situations intermédiaires (Ward, 1963) [15].

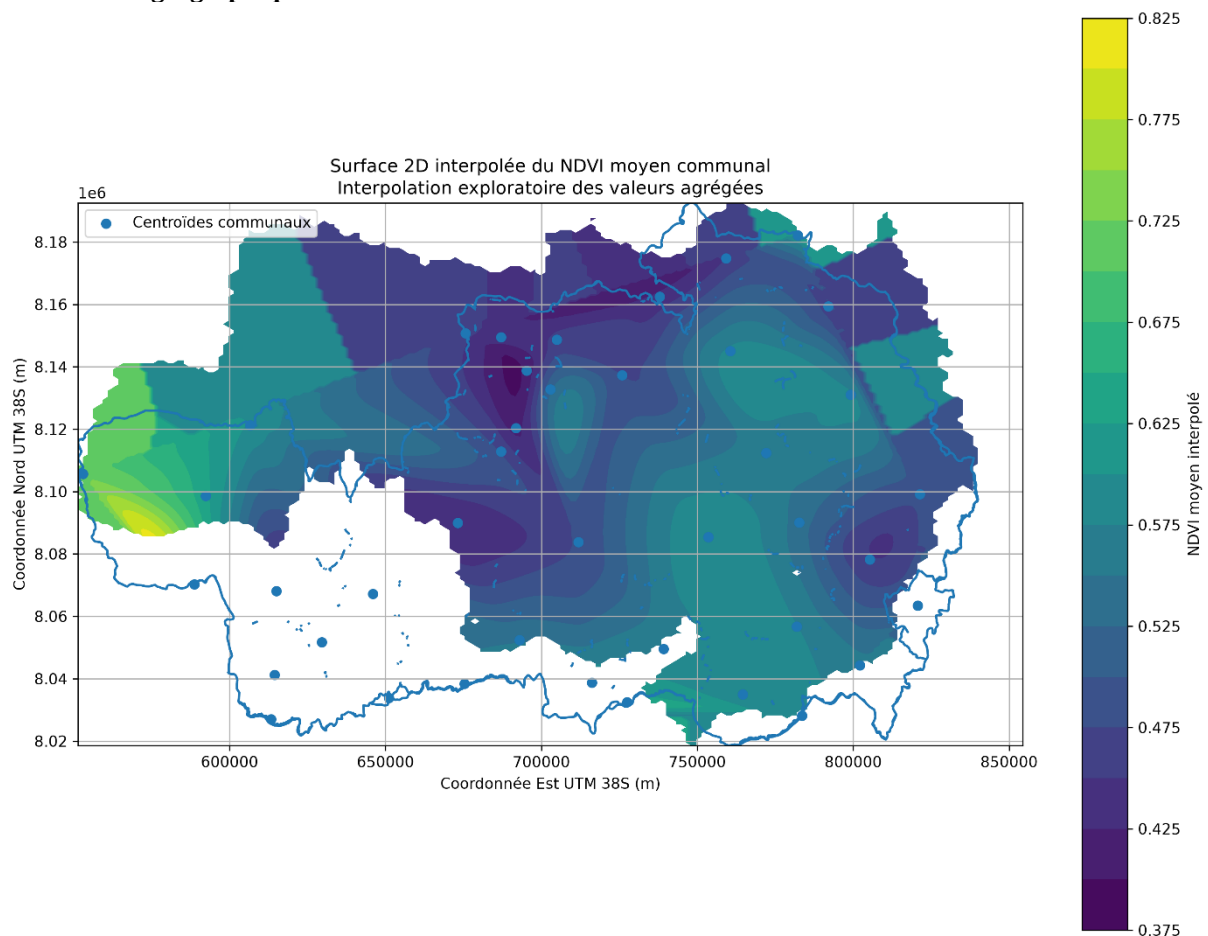
3 Résultats et discussions

3.1 Résultats

3.1.1 Distribution spatiale interpolée du NDVI moyen dans le bassin versant de Betsiboka selon les coordonnées géographiques

La figure 1 présente la distribution spatiale interpolée du NDVI moyen dans le bassin versant de Betsiboka. La zone représentée s'étend approximativement entre 45,47° et 48,28° de longitude Est et entre 16,35° et 17,91° de latitude Sud. Les valeurs mesurées au niveau des différents points géoréférencés ont été utilisées pour générer une surface continue permettant de mettre en évidence les principaux gradients spatiaux de la couverture végétale. Le NDVI interpolé varie globalement de 0,375 à 0,825, ce qui traduit une importante hétérogénéité de la densité et de l'activité chlorophyllienne de la végétation.

Figure 1: Distribution spatiale interpolée du NDVI moyen dans le bassin versant de Betsiboka selon les coordonnées géographiques



Source : Auteur 2026

Interprétation

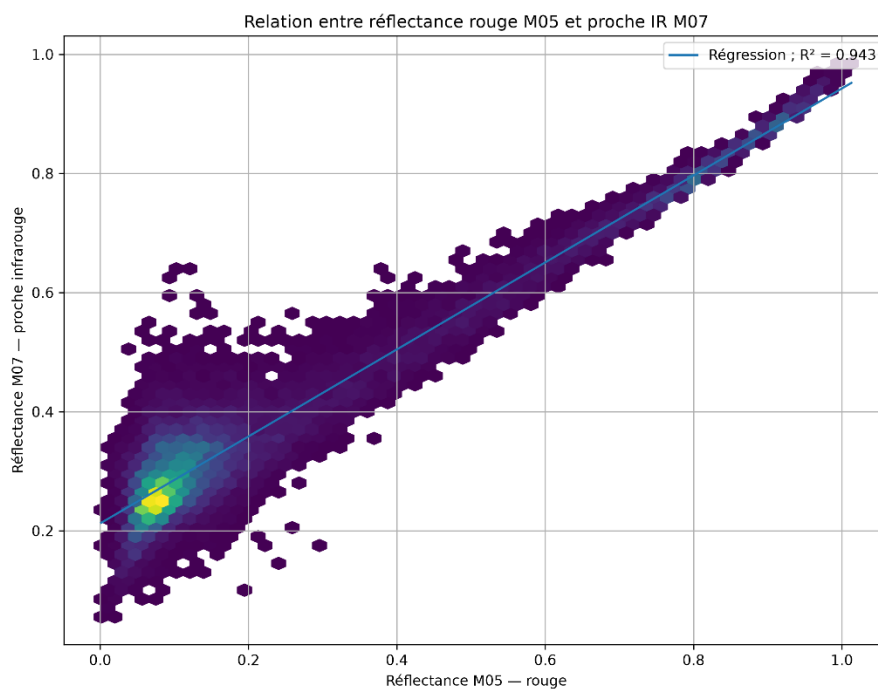
Les valeurs les plus élevées, comprises approximativement entre 0,675 et 0,825, sont principalement observées dans la partie occidentale et sud-occidentale du bassin, autour de 45,75° de longitude Est et 17,23° de latitude Sud. Ces valeurs indiquent une couverture végétale relativement dense et active, pouvant correspondre à des formations forestières, des savanes bien développées, des zones humides ou des espaces agricoles en période de forte croissance végétative. Quelques secteurs méridionaux, notamment autour de 47,40° Est et 17,80° Sud, présentent également des valeurs relativement élevées, généralement comprises entre 0,625 et 0,725. À l'inverse, les valeurs les plus faibles, situées entre environ 0,375 et 0,475, apparaissent principalement dans la partie centrale et nord-centrale du bassin. Un noyau de faible NDVI est notamment observé autour de 46,88° de longitude Est et 16,84° de latitude Sud. Ces faibles valeurs peuvent traduire une végétation clairsemée, la présence de sols nus, de zones dégradées, de surfaces récemment brûlées ou de parcelles agricoles faiblement végétalisées. Un autre secteur de faible NDVI est visible dans la partie sud-est, autour de 47,92° Est et 17,34° Sud, où les valeurs descendent localement jusqu'à environ 0,40. La majeure partie du bassin présente néanmoins des valeurs intermédiaires comprises entre 0,475 et 0,625. Ces niveaux de NDVI peuvent être associés à des mosaïques de savanes, de cultures, de jachères et de formations végétales partiellement dégradées. Dans la partie orientale, autour de 47,63° de longitude Est et 16,90° de latitude Sud, les valeurs se situent généralement entre 0,525 et 0,625, indiquant une couverture végétale modérée. Sur le plan hydroécologique, les secteurs caractérisés par un NDVI supérieur à 0,60 disposent généralement d'une couverture végétale capable de réduire l'impact direct des précipitations, de ralentir le ruissellement et de favoriser l'infiltration de l'eau dans le sol. À l'opposé, les zones présentant un NDVI inférieur à 0,45 peuvent être considérées comme plus sensibles au ruissellement superficiel, au détachement des particules et à l'érosion hydrique, particulièrement lorsqu'elles sont associées à de fortes pentes ou à des pluies fortement érosives. Ces résultats peuvent ainsi contribuer à l'estimation du facteur de couverture végétale C dans le modèle RUSLE et à l'analyse des relations entre végétation, ruissellement, pertes en sols et fonctionnement hydrologique du bassin versant. Il convient toutefois de préciser que les axes de la figure sont initialement exprimés en coordonnées projetées UTM zone 38S. Les coordonnées en longitude et latitude présentées dans cette

interprétation constituent donc des conversions approximatives destinées à faciliter la localisation géographique des principaux secteurs observés.

3.1.2 Relation spectrale entre la réflectance dans le rouge M05 et le proche infrarouge M07 dans le bassin versant de Betsiboka

La figure 2 présente la relation entre la réflectance mesurée dans la bande rouge M05 et celle enregistrée dans la bande du proche infrarouge M07. La représentation en mailles hexagonales permet de visualiser simultanément la distribution des valeurs et leur densité d'occurrence. Les teintes jaunes et vertes correspondent aux concentrations les plus élevées de pixels, tandis que les teintes violettes indiquent des combinaisons spectrales moins fréquentes. La droite de régression met en évidence une relation linéaire positive très forte entre les deux bandes spectrales.

Figure 2: Relation spectrale entre la réflectance dans le rouge M05 et le proche infrarouge M07 dans le bassin versant de Betsiboka



Source : Auteur 2026

Interprétation

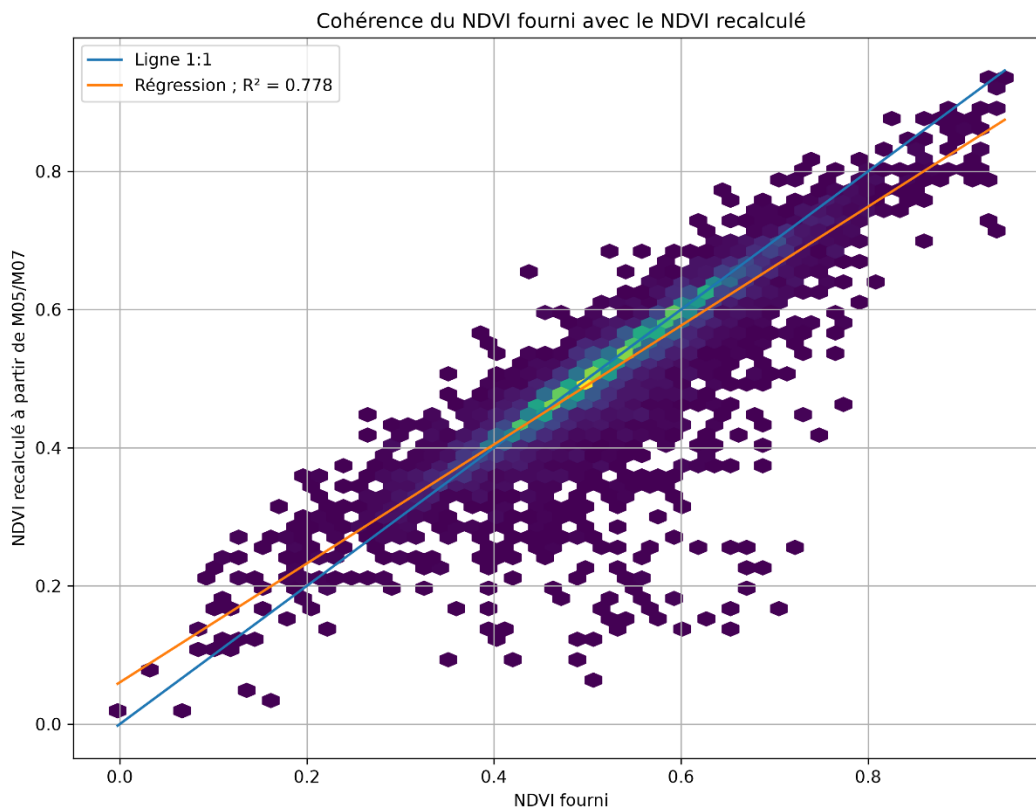
Le coefficient de détermination atteint une valeur de $R^2 = 0,943$, ce qui signifie qu'environ 94,3 % de la variabilité de la réflectance dans le proche infrarouge M07 est statistiquement expliquée par les variations de la réflectance dans le rouge M05 dans le cadre du modèle linéaire ajusté. Le coefficient de corrélation correspondant est approximativement égal à $r = 0,971$, indiquant une association positive particulièrement élevée entre les deux variables. Cette relation traduit une augmentation presque systématique de la réflectance dans le proche infrarouge lorsque la réflectance dans le rouge augmente. Elle ne doit toutefois pas être interprétée comme une relation causale directe, mais comme une covariation spectrale des surfaces observées. La concentration maximale des observations apparaît approximativement pour des réflectances rouges comprises entre 0,05 et 0,15 et des réflectances dans le proche infrarouge situées entre 0,20 et 0,35. Le noyau de densité le plus important se situe autour de 0,08 à 0,10 dans le rouge et de 0,24 à 0,28 dans le proche infrarouge. Cette configuration, caractérisée par une faible réflectance dans le rouge et une réflectance sensiblement plus élevée dans le proche infrarouge, peut correspondre à des surfaces végétalisées. La chlorophylle absorbe en effet une partie importante du rayonnement rouge, tandis que la structure interne des feuilles favorise une réflexion plus forte dans le proche infrarouge. Pour une réflectance rouge comprise entre 0,20 et 0,60, la réflectance dans le proche infrarouge varie principalement entre 0,30 et 0,70. Cette zone intermédiaire peut traduire la présence de mosaïques paysagères associant végétation peu dense, cultures, savanes, sols partiellement nus et surfaces dégradées. Lorsque les réflectances rouges dépassent 0,70, les valeurs du proche infrarouge deviennent généralement supérieures à 0,70 et peuvent atteindre environ 0,95 à 1,00. Ces fortes réflectances simultanées peuvent être associées à des surfaces très lumineuses, à des sols nus secs, à certaines formations géologiques ou à des pixels affectés par des nuages, des brumes ou des

effets de saturation radiométrique. Une vérification à partir des masques de nuages et des cartes d'occupation du sol reste donc nécessaire. La droite de régression passe visuellement d'une réflectance M07 proche de 0,21 lorsque M05 est voisine de zéro à une valeur d'environ 0,95 lorsque M05 atteint 1,00. La pente graphique peut ainsi être estimée à environ 0,74, ce qui indique que l'augmentation de la réflectance dans le proche infrarouge est légèrement moins rapide que celle de la réflectance dans le rouge sur l'ensemble de la gamme étudiée. La présence d'une ordonnée à l'origine positive reflète le fait que les surfaces présentant une très faible réflectance rouge conservent généralement une réponse mesurable dans le proche infrarouge. Quelques observations s'écartent cependant de la tendance principale. Pour des valeurs rouges proches de 0,10, certaines réflectances dans le proche infrarouge atteignent environ 0,60 à 0,64, ce qui peut correspondre à une végétation particulièrement active ou à des pixels spectralement atypiques. À l'inverse, certaines valeurs du proche infrarouge inférieures à 0,20 sont observées pour des réflectances rouges comprises entre 0,20 et 0,35. Ces valeurs peuvent être liées à des surfaces en eau, à des zones fortement ombragées, à des mélanges de pixels ou à des anomalies résiduelles provenant du prétraitement des images. Sur le plan hydroécologique, cette relation spectrale constitue une base importante pour l'analyse de la végétation et le calcul des indices tels que le NDVI, défini par le rapport entre la différence et la somme des réflectances du proche infrarouge et du rouge. Par exemple, un pixel présentant une réflectance rouge de 0,10 et une réflectance proche infrarouge de 0,28 possède un NDVI d'environ 0,47, caractéristique d'une couverture végétale modérée. En revanche, un pixel dont les réflectances rouges et proche infrarouge sont respectivement égales à 0,60 et 0,65 présente un NDVI proche de 0,04, indiquant une faible différenciation spectrale entre les deux bandes et probablement une couverture végétale limitée. La forte corrélation observée doit donc être complétée par l'analyse de la différence relative entre les deux bandes afin de distinguer efficacement les surfaces végétalisées, les sols nus et les autres catégories d'occupation du sol.

3.1.3 Cohérence entre le NDVI fourni et le NDVI recalculé à partir des bandes spectrales M05 et M07

La figure 3 présente la relation entre les valeurs de NDVI fournies dans le produit satellitaire et celles recalculées à partir des réflectances de la bande rouge M05 et de la bande du proche infrarouge M07. La ligne diagonale 1:1 représente la situation théorique dans laquelle les deux estimations du NDVI seraient parfaitement identiques. La droite de régression traduit, quant à elle, la tendance statistique observée entre les deux séries de valeurs. La représentation sous forme de mailles hexagonales permet également d'identifier les plages de valeurs les plus fréquemment rencontrées, les couleurs les plus claires correspondant aux plus fortes densités d'observations.

Figure 3: Cohérence entre le NDVI fourni et le NDVI recalculé à partir des bandes spectrales M05 et M07



Source : Auteur 2026

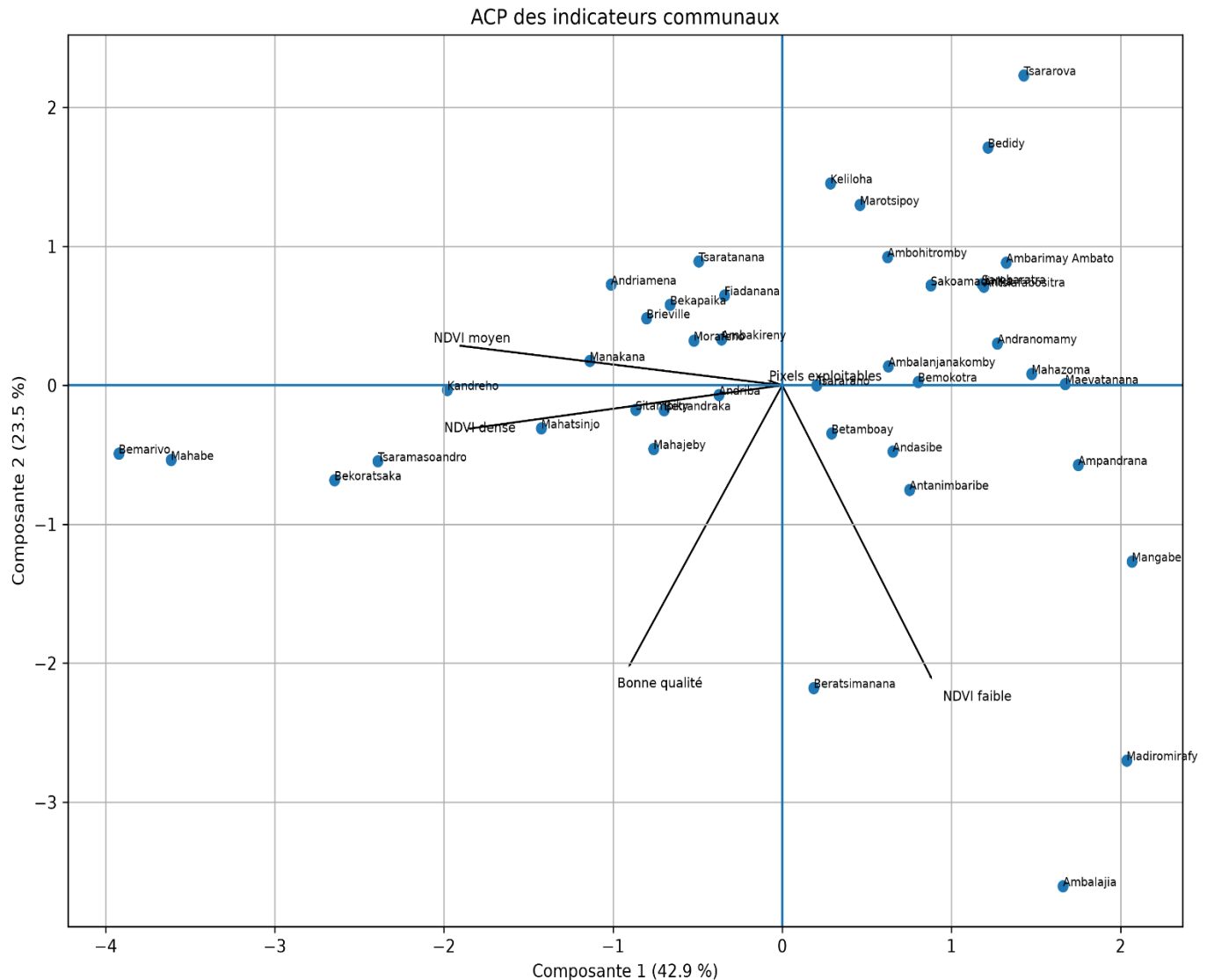
Interprétation

Le coefficient de détermination de la régression est égal à $R^2 = 0,778$. Cette valeur indique qu'environ 77,8 % de la variabilité du NDVI recalculé est expliquée par les variations du NDVI fourni. Le coefficient de corrélation correspondant est approximativement égal à $r = 0,882$, ce qui traduit une relation positive forte entre les deux variables. Ainsi, lorsque le NDVI fourni augmente, le NDVI recalculé tend également à augmenter. Cette concordance confirme que les bandes M05 et M07 reproduisent globalement la dynamique spatiale de l'activité végétale représentée par le produit NDVI initial. La plus forte concentration des observations se situe approximativement entre 0,45 et 0,60 pour le NDVI fourni et entre 0,45 et 0,58 pour le NDVI recalculé. Le noyau de densité principal apparaît autour de 0,50 pour les deux variables. Ces valeurs intermédiaires sont généralement associées à une couverture végétale modérée à relativement dense, pouvant correspondre à des savanes, des formations arbustives, des cultures en phase de développement ou des mosaïques de végétation et de sols partiellement découverts. La droite de régression peut être estimée graphiquement par une relation approximative de la forme $\text{NDVI recalculé} = 0,055 + 0,86 \times \text{NDVI fourni}$. La pente inférieure à 1 indique que l'amplitude des valeurs recalculées est légèrement réduite par rapport à celle du NDVI fourni. Pour un NDVI fourni égal à 0,20, le modèle prévoit par exemple un NDVI recalculé proche de 0,23. Pour une valeur fournie de 0,50, le NDVI recalculé estimé est d'environ 0,49, tandis que pour un NDVI fourni de 0,80, la valeur recalculée attendue est voisine de 0,74. La comparaison entre la droite de régression et la ligne 1:1 met en évidence un biais systématique variable selon le niveau du NDVI. Pour les faibles valeurs, approximativement inférieures à 0,40, la droite de régression se situe au-dessus de la ligne 1:1. Le NDVI recalculé tend donc à être légèrement supérieur au NDVI fourni. À l'inverse, pour les valeurs supérieures à environ 0,40, la droite de régression passe progressivement sous la ligne d'égalité, ce qui traduit une sous-estimation relative des NDVI moyens et élevés. Cette sous-estimation devient plus visible pour les valeurs supérieures à 0,70. À titre d'exemple, un NDVI fourni de 0,90 correspond à une valeur recalculée moyenne proche de 0,83, soit une différence d'environ 0,07. La dispersion autour de la droite de régression est relativement limitée dans la partie centrale du nuage, mais elle augmente pour certaines plages de valeurs. Plusieurs observations caractérisées par un NDVI fourni compris entre 0,40 et 0,70 présentent des NDVI recalculés nettement plus faibles, parfois situés entre 0,10 et 0,30. Ces écarts peuvent résulter de la présence de pixels mixtes, de surfaces en eau, de zones ombragées, de nuages résiduels, d'effets atmosphériques ou de différences dans les procédures de correction radiométrique. Ils peuvent également être liés à des décalages spatiaux ou temporels entre les données utilisées pour produire le NDVI fourni et celles employées pour effectuer le nouveau calcul. Les valeurs élevées, comprises entre 0,75 et 0,95, restent globalement proches de la ligne 1:1, bien qu'une légère sous-estimation soit observée dans le NDVI recalculé. Ces valeurs correspondent généralement à une végétation dense et photosynthétiquement active. Toutefois, quelques observations s'écartent sensiblement de la tendance générale, notamment lorsque le NDVI fourni est supérieur à 0,85 alors que le NDVI recalculé descend localement autour de 0,70 à 0,80. Ces différences peuvent être liées à une saturation partielle du NDVI dans les formations végétales denses ou à des différences de résolution, de calibration et de composition temporelle des produits satellitaires.

3.1.4 Analyse en composantes principales des indicateurs de végétation et de qualité des données dans le bassin versant de Betsiboka

La figure 4 présente la projection des sites étudiés et des principaux indicateurs de télédétection sur les deux premiers axes d'une analyse en composantes principales. La composante 1 explique 42,9 % de la variabilité totale, tandis que la composante 2 en représente 23,5 %. Le plan factoriel restitue ainsi 66,4 % de l'information contenue dans les données initiales, ce qui permet d'identifier les principales structures spatiales et les associations entre le NDVI moyen, le NDVI dense, le NDVI faible, la qualité des observations et la disponibilité des pixels exploitables. Les 33,6 % restants sont portés par les composantes suivantes et ne sont donc pas représentés sur cette figure.

Figure 4: Analyse en composantes principales des indicateurs de végétation et de qualité des données dans le bassin versant de Betsiboka

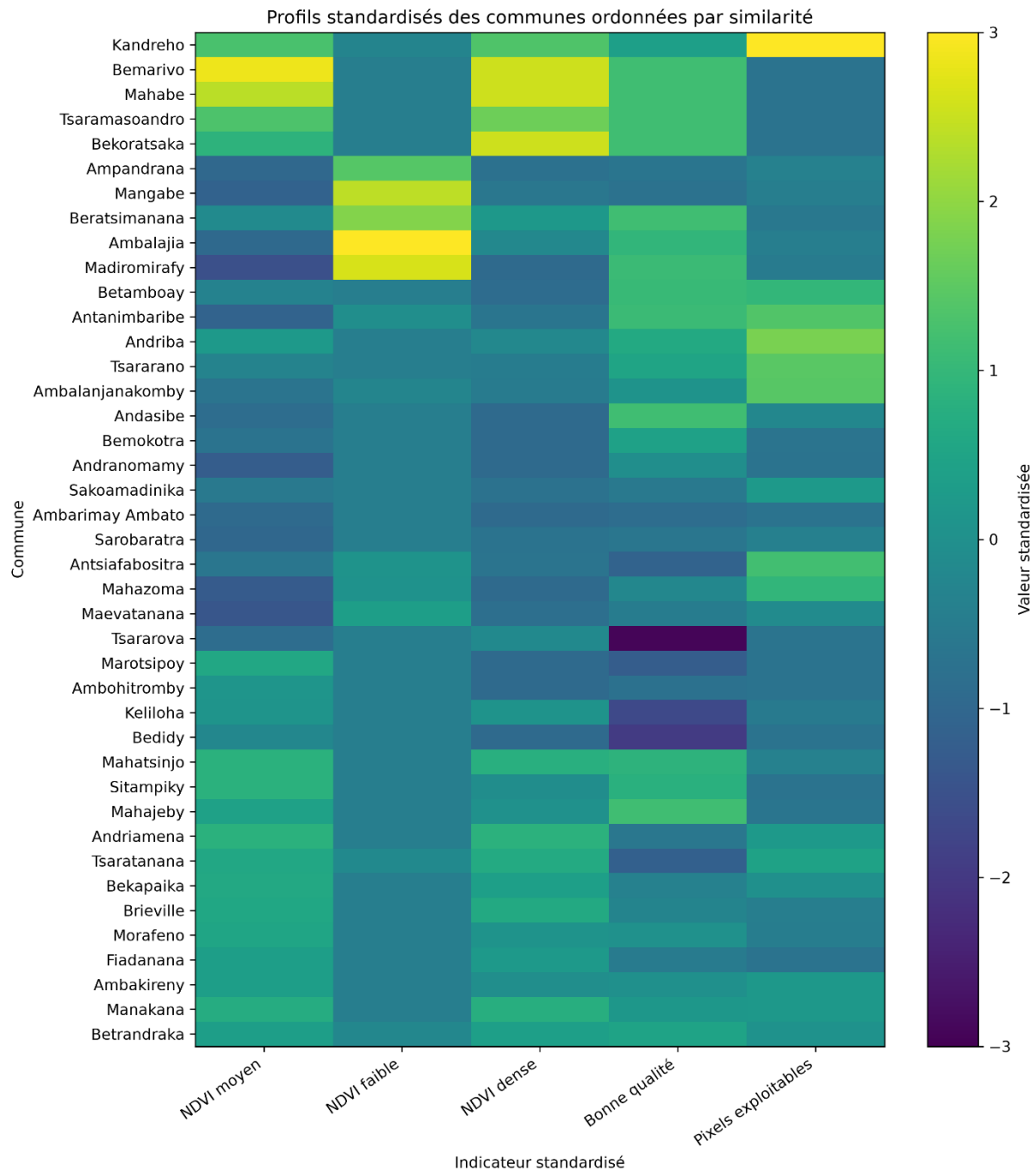


localités situées dans la partie gauche du graphique, telles que Bemarivo, Mahabe, Bekoratsaka, Tsaramasoandro et Kandrehio, présentent des scores négatifs élevés sur la composante 1, compris approximativement entre $-4,0$ et $-2,0$. Leur position dans la direction des vecteurs du NDVI moyen et du NDVI dense suggère qu'elles sont relativement associées à une couverture végétale plus développée. Bemarivo et Mahabe constituent les localités les plus éloignées sur le côté négatif du premier axe, avec des scores proches respectivement de $-3,9$ et $-3,6$. Cette position indique un profil spectral nettement différent de celui des localités situées dans la partie droite du graphique. Les localités comme Andriamena, Tsaratanana, Bekapika, Brieville, Fiadanana, Morafeno et Ambakireny occupent la partie supérieure gauche ou centrale du plan. Leurs scores sur la composante 2 sont généralement compris entre $0,3$ et $0,9$. Elles sont davantage associées au NDVI moyen qu'au NDVI faible. Leur position intermédiaire suggère toutefois des conditions moins contrastées que celles observées pour Bemarivo ou Mahabe. Elles peuvent correspondre à des mosaïques paysagères associant formations végétales, savanes, cultures et surfaces temporairement découvertes. À l'opposé, plusieurs localités de la partie inférieure droite sont positionnées dans la direction du vecteur du NDVI faible. Mangabe présente des scores proches de $2,1$ sur la composante 1 et de $-1,3$ sur la composante 2, tandis que Madiromirafy atteint environ $2,0$ sur la première composante et $-2,7$ sur la seconde. Ambalajia constitue le cas le plus extrême sur la composante 2, avec un score voisin de $-3,6$. Beratsimanana se distingue également par un score d'environ $-2,2$ sur le deuxième axe. Ces localités sont donc relativement associées à une plus forte représentation des faibles valeurs de NDVI, pouvant traduire une végétation clairsemée, des sols nus, des zones agricoles, des espaces brûlés ou des surfaces soumises à une dégradation plus importante. La partie supérieure droite regroupe notamment Tsararova, Bedidy, Keliloha et Marotsipoy. Tsararova présente les scores positifs les plus élevés sur la composante 2, avec une valeur proche de $2,2$, alors que Bedidy atteint environ $1,7$. Keliloha et Marotsipoy présentent respectivement des scores proches de $1,4$ et $1,3$. Leur éloignement par rapport à l'origine indique des profils particuliers, mais leur position n'est directement alignée avec aucun des vecteurs principaux. Leur différenciation pourrait donc résulter d'une combinaison spécifique des indicateurs ou de variables davantage représentées sur les composantes non affichées. Les localités proches de l'origine, telles qu'Ambakireny, Ambodimanga, Sitampiky, Andranofasika, Bemokotra ou certaines localités voisines de Maevatanana, possèdent des scores relativement faibles sur les deux composantes. Elles présentent donc des profils proches de la moyenne générale des données et ne sont fortement associées à aucun indicateur particulier. La proximité de deux localités sur le graphique traduit une similarité de leurs caractéristiques statistiques, tandis qu'une distance importante indique des profils plus différents. Sur le plan hydroécologique, l'opposition entre les indicateurs de végétation dense et de NDVI faible peut être utilisée pour distinguer les secteurs relativement protégés par la couverture végétale des zones potentiellement sensibles au ruissellement et à l'érosion hydrique. Les localités projetées vers les NDVI moyens et denses pourraient présenter une meilleure interception des précipitations, une infiltration plus importante et une protection accrue des sols. À l'inverse, les localités associées au NDVI faible, notamment Ambalajia, Madiromirafy, Mangabe et Beratsimanana, peuvent être considérées comme des secteurs prioritaires pour les investigations de terrain, en particulier lorsque leurs faibles niveaux de végétation coïncident avec de fortes pentes, des sols érodibles ou une pluviométrie fortement érosive.

3.1.5 Profils standardisés des indicateurs de végétation et de qualité des observations

La figure 5 présente une classification visuelle des sites étudiés à partir de cinq indicateurs standardisés : le NDVI moyen, la proportion de faibles valeurs de NDVI, la proportion de NDVI élevés ou denses, la qualité des observations satellitaires et la disponibilité des pixels exploitables. Les localités sont ordonnées selon la similarité de leurs profils, de sorte que les sites présentant des caractéristiques proches apparaissent les uns à côté des autres. L'échelle de couleur varie de -3 à $+3$. Une valeur standardisée égale à 0 correspond à la moyenne de l'ensemble des sites, une valeur de $+1$ indique un niveau supérieur d'un écart-type à cette moyenne et une valeur de -1 traduit un niveau inférieur d'un écart-type. Les teintes jaunes et vert clair représentent ainsi les valeurs relativement élevées, tandis que les teintes bleues et violettes correspondent aux valeurs faibles.

Figure 5: Profils standardisés des indicateurs de végétation et de qualité des observations



Source : Auteur 2026

Interprétation

Les cinq premières localités, Kandreho, Bemarivo, Mahabe, Tsaramasoandro et Bekoratsaka, forment un groupe caractérisé par des NDVI moyens et des NDVI denses supérieurs à la moyenne. Bemarivo et Mahabe présentent notamment des NDVI moyens standardisés proches de +2,5 à +3,0 et des valeurs de NDVI dense comprises approximativement entre +2,2 et +2,7. Bekoratsaka se distingue également par un NDVI dense voisin de +2,5. Parallèlement, leur indicateur de NDVI faible reste généralement inférieur à la moyenne, avec des valeurs standardisées proches de -0,5. Ces résultats traduisent une couverture végétale relativement développée et une proportion limitée de surfaces faiblement végétalisées. Kandreho se distingue particulièrement par une disponibilité très élevée des pixels exploitables, dont la valeur standardisée atteint environ +3,0, ce qui indique des conditions satellitaires exceptionnellement favorables par rapport à l'ensemble des sites. Bemarivo, Mahabe, Tsaramasoandro et Bekoratsaka présentent cependant des valeurs de pixels exploitables légèrement inférieures à

la moyenne, généralement comprises entre $-0,5$ et $-0,8$. Cette configuration montre qu'un niveau élevé de végétation ne correspond pas nécessairement à une plus grande disponibilité des observations satellitaires. Les caractéristiques écologiques et la qualité des données doivent donc être interprétées séparément. Une faible disponibilité de pixels peut notamment résulter de la couverture nuageuse, des ombres, de la présence de données invalides ou des procédures de masquage appliquées aux images. Un deuxième ensemble, comprenant Ampandrana, Mangabe, Beratsimanana, Ambalajia et Madiromirafy, présente un profil opposé. Ces localités sont caractérisées par des NDVI moyens inférieurs à la moyenne, avec des valeurs standardisées comprises approximativement entre $-0,5$ et $-1,5$, ainsi que par une forte représentation des faibles NDVI. Mangabe présente un indicateur de NDVI faible proche de $+2,4$, Ambalajia atteint environ $+3,0$ et Madiromirafy se situe autour de $+2,6$. Dans le même temps, les NDVI denses sont généralement faibles, avec des valeurs comprises entre environ $-0,2$ et $-1,0$. Ce profil traduit une couverture végétale plus clairsemée et une proportion relativement importante de surfaces peu végétalisées, cultivées, brûlées ou partiellement dénudées. Ambalajia et Madiromirafy apparaissent comme les localités les plus fortement associées aux faibles valeurs de NDVI. Ambalajia combine un NDVI moyen proche de $-0,8$, un NDVI faible d'environ $+3,0$ et un NDVI dense voisin de $-0,2$. Madiromirafy présente un NDVI moyen encore plus faible, approximativement égal à $-1,5$, accompagné d'un NDVI faible supérieur à $+2,5$. Ces configurations concordent avec leur position dans la direction du vecteur « NDVI faible » observée dans l'analyse en composantes principales. Elles peuvent indiquer une plus forte exposition des sols au ruissellement et à l'érosion hydrique, particulièrement si ces localités sont également caractérisées par des pentes élevées et des précipitations érosives. Les localités comprises entre Betamboay et Maevatanana présentent des profils plus intermédiaires. Leurs NDVI moyens sont généralement légèrement inférieurs à la moyenne, avec des valeurs comprises entre environ $-1,5$ et 0 , tandis que les NDVI faibles restent proches de la moyenne ou légèrement négatifs. Betamboay, Antanimbaribe, Andasibe, Bemokotra et Andranomamy présentent des valeurs positives de l'indicateur de bonne qualité, généralement comprises entre $+0,2$ et $+1,2$. Ces résultats indiquent que les images utilisées pour ces sites contiennent une proportion relativement satisfaisante d'observations valides, même lorsque les indicateurs de végétation ne sont pas particulièrement élevés. Andriba et Tsararano se distinguent par une disponibilité élevée des pixels exploitables, avec des valeurs standardisées respectivement proches de $+1,9$ et $+1,4$. Ambalanjanakomby et Antsiafabositra présentent également des valeurs supérieures à $+1$ pour cet indicateur. Ces sites disposent donc d'un volume d'information satellitaire supérieur à la moyenne, ce qui peut renforcer la fiabilité des statistiques de NDVI qui y sont calculées. En revanche, une valeur élevée de pixels exploitables ne signifie pas nécessairement que la végétation est dense ; elle indique principalement que les observations sont suffisamment nombreuses et conformes aux critères de qualité définis. Tsararova constitue un cas particulier, avec une valeur de bonne qualité proche de $-3,0$, soit environ trois écarts-types en dessous de la moyenne. Cette valeur indique une proportion particulièrement faible d'observations répondant aux critères de qualité. Les résultats relatifs au NDVI de cette localité doivent donc être interprétés avec prudence. Keliloha et Bedidy présentent également des valeurs faibles pour cet indicateur, comprises approximativement entre $-1,5$ et $-2,0$. Ces écarts peuvent être liés à une couverture nuageuse persistante, à des effets atmosphériques, à des ombres topographiques ou à un nombre insuffisant d'images utilisées pendant la période analysée. Les localités situées dans la partie inférieure de la figure, notamment Mahatsinjo, Sitampiky, Mahajeby, Andriamena, Tsaratanana, Bekapaika, Brieville, Morafeno, Fiadanana, Ambakireny, Manakana et Betrandraka, présentent généralement des NDVI moyens légèrement supérieurs à la moyenne. Les valeurs standardisées sont principalement comprises entre $+0,4$ et $+1,2$. Leurs indicateurs de NDVI faible sont simultanément négatifs, généralement autour de $-0,3$ à $-0,6$, ce qui confirme une proportion relativement limitée de surfaces faiblement végétalisées. Andriamena et Tsaratanana se distinguent par des NDVI denses compris approximativement entre $+0,8$ et $+1,2$. Manakana présente également un NDVI moyen et un NDVI dense proches de $+1,0$. Ces localités peuvent donc être associées à une végétation relativement active, mais moins exceptionnelle que celle observée à Bemarivo, Mahabe ou Bekoratsaka. Les profils plus proches de zéro, comme ceux de Morafeno, Fiadanana, Ambakireny et Betrandraka, indiquent des conditions proches de la moyenne générale du bassin versant. L'organisation des lignes met ainsi en évidence plusieurs types de profils hydroécologiques. Un premier type associe des NDVI moyens et denses élevés à une faible proportion de NDVI bas, traduisant une couverture végétale relativement protectrice. Un deuxième type est dominé par les faibles NDVI et correspond potentiellement à des surfaces plus ouvertes ou dégradées. Un troisième type rassemble des localités intermédiaires, dont les indicateurs de végétation sont proches de la moyenne mais dont la qualité ou la disponibilité des données peut être élevée. Enfin, certains sites se distinguent principalement par des anomalies dans la qualité des observations satellitaires. Sur le plan de l'érosion hydrique, les localités présentant des NDVI moyens et denses supérieurs à $+1$ peuvent bénéficier d'une meilleure protection du sol par l'interception des précipitations, la réduction de la vitesse du ruissellement et la stabilisation des particules par les racines. À l'inverse, les sites dont l'indicateur de NDVI faible dépasse $+2$, notamment Ambalajia, Madiromirafy, Mangabe et Beratsimanana, peuvent présenter une sensibilité plus importante au détachement et au transport des particules.

3.2 Discussions

Les analyses mettent en évidence une forte hétérogénéité spatiale de la couverture végétale dans le bassin versant de Betsiboka. Les valeurs de NDVI comprises entre 0,375 et 0,825 traduisent la coexistence de secteurs densément végétalisés et de zones présentant une couverture plus faible. Les faibles valeurs peuvent correspondre à des sols nus, des espaces agricoles, des surfaces brûlées ou des zones dégradées, potentiellement plus exposées au ruissellement et à l'érosion hydrique. À l'inverse, les secteurs où le NDVI dépasse 0,60 disposent généralement d'une couverture végétale susceptible de mieux protéger les sols. La forte relation entre les réflectances des bandes M05 et M07, caractérisée par un coefficient de détermination de 0,943, confirme la cohérence générale du signal spectral. Toutefois, cette corrélation élevée ne suffit pas, à elle seule, à différencier les types de couverture du sol, car certaines surfaces peuvent présenter des réflectances importantes dans les deux bandes. L'utilisation du NDVI demeure donc nécessaire pour mesurer leur contraste relatif. La comparaison entre le NDVI fourni et le NDVI recalculé montre une cohérence satisfaisante, avec un coefficient de détermination de 0,778. Les écarts observés peuvent être liés aux corrections atmosphériques, aux facteurs d'échelle, aux pixels mixtes, aux nuages résiduels ou aux différences de traitement radiométrique. La tendance à sous-estimer certaines valeurs élevées indique que les deux produits ne sont pas strictement interchangeables sans validation préalable. L'analyse en composantes principales, qui restitue 66,4 % de la variabilité totale sur les deux premiers axes, met en évidence une opposition nette entre les sites dominés par une végétation dense et ceux caractérisés par de faibles NDVI. Les profils standardisés confirment cette différenciation et permettent également d'identifier les zones où la qualité ou la disponibilité des observations est insuffisante.

4 Conclusion

Cette étude a permis de caractériser la variabilité géospatiale, spectrale et multivariée du NDVI dans le bassin versant de Betsiboka. Les analyses ont révélé une importante hétérogénéité de la couverture végétale, avec des secteurs à végétation relativement dense et d'autres présentant des niveaux plus faibles, potentiellement associés à des surfaces dégradées, cultivées ou partiellement dénudées. La relation étroite observée entre les bandes M05 et M07 confirme la cohérence générale des données spectrales utilisées. La comparaison entre le NDVI fourni et le NDVI recalculé montre également une concordance satisfaisante, malgré certains écarts pouvant résulter des conditions atmosphériques, des pixels mixtes ou des traitements radiométriques. L'analyse en composantes principales et les profils standardisés ont, quant à eux, permis d'identifier les principaux gradients de différenciation entre les sites et de distinguer les zones dominées par une végétation dense de celles caractérisées par de faibles valeurs de NDVI. L'intégration de ces approches constitue ainsi une méthode pertinente pour le suivi de la végétation et l'évaluation de la qualité des observations satellitaires dans le bassin versant de Betsiboka. Ces résultats peuvent contribuer à l'identification des secteurs sensibles à la dégradation des terres et à l'érosion hydrique. Leur interprétation gagnerait toutefois à être complétée par des données de terrain, des cartes d'occupation du sol et des variables topographiques afin de renforcer la portée écologique et opérationnelle de l'analyse.

REFERENCES

- [1] Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X. et Ferreira, L.G. (2002) 'Overview of the radiometric and biophysical performance of the MODIS vegetation indices', *Remote Sensing of Environment*, 83(1–2), pp. 195–213. doi: 10.1016/S0034-4257(02)00096-2.
- [2] Pettorelli, N., Vik, J.O., Mysterud, A., Gaillard, J.-M., Tucker, C.J. et Stenseth, N.C. (2005) 'Using the satellite-derived NDVI to assess ecological responses to environmental change', *Trends in Ecology & Evolution*, 20(9), pp. 503–510. doi: 10.1016/j.tree.2005.05.011.
- [3] Harper, G.J., Steininger, M.K., Tucker, C.J., Juhn, D. et Hawkins, F. (2007) 'Fifty years of deforestation and forest fragmentation in Madagascar', *Environmental Conservation*, 34(4), pp. 325–333. doi: 10.1017/S0376892907004262.
- [4] Raharimahefa, T. et Kusky, T.M. (2010) 'Environmental monitoring of Bombetoka Bay and the Betsiboka Estuary, Madagascar, using multi-temporal satellite data', *Journal of Earth Science*, 21, pp. 210–226. doi: 10.1007/s12583-010-0019-y.
- [5] Abdi, H. et Williams, L.J. (2010) 'Principal component analysis', *WIREs Computational Statistics*, 2(4), pp. 433–459. doi: 10.1002/wics.101.

- [6] Foiben-Taosarintanin'i Madagasikara (FTM) (2023) Madagascar Report by FTM. 19th Southern African and Islands Hydrographic Commission Meeting. Monaco: International Hydrographic Organization.
- [7] Didan, K., Barreto Munoz, A., Compton, T.J. et Pinzon, J.E. (2018) Suomi National Polar-orbiting Partnership Visible Infrared Imaging Radiometer Suite Vegetation Index Product Suite: User Guide and Abridged Algorithm Theoretical Basis Document. Version 2.1.2. Tucson: University of Arizona and NASA.
- [8] Gillies, S. et al. (2013–) Rasterio: Geospatial Raster I/O for Python Programmers. Logiciel informatique. Mapbox.
- [9] Jordahl, K. et al. (2020) geopandas/geopandas: version 0.8.1. Zenodo. doi: 10.5281/zenodo.3946761.
- [10] Pedregosa, F. et al. (2011) 'Scikit-learn: Machine Learning in Python', Journal of Machine Learning Research, 12, pp. 2825–2830.
- [11] Tucker, C.J. (1979) 'Red and photographic infrared linear combinations for monitoring vegetation', Remote Sensing of Environment, 8(2), pp. 127–150. doi: 10.1016/0034-4257(79)90013-0.
- [12] Shepard, D. (1968) 'A two-dimensional interpolation function for irregularly-spaced data', Proceedings of the 23rd ACM National Conference, pp. 517–524. doi: 10.1145/800186.810616.
- [13] Montgomery, D.C., Peck, E.A. et Vining, G.G. (2012) Introduction to Linear Regression Analysis. 5e éd. Hoboken, NJ: John Wiley & Sons.
- [14] Jolliffe, I.T. (2002) Principal Component Analysis. 2e éd. New York: Springer-Verlag. doi: 10.1007/b98835.
- [15] Ward, J.H. Jr. (1963) 'Hierarchical grouping to optimize an objective function', Journal of the American Statistical Association, 58(301), pp. 236–244. doi: 10.1080/01621459.1963.10500845.