



Suivi des algues vertes par drone multispectral

De la réflectance des pixels à la compréhension du littoral

QGIS • SCP • OTB • GRASS • Python • Deep Learning

Juin 2026

Atilio FRANCOIS

sigterritoires.fr

Introduction

Les littoraux figurent parmi les environnements les plus dynamiques et les plus complexes de notre planète. Soumis à l'influence des marées, des conditions météorologiques, des apports continentaux et des activités humaines, ils évoluent en permanence. Comprendre ces transformations constitue un enjeu majeur pour la gestion des espaces côtiers et la préservation des écosystèmes.

Parmi les phénomènes qui préoccupent aujourd'hui les gestionnaires du littoral, les proliférations d'algues vertes occupent une place particulière. Leur apparition résulte de mécanismes écologiques complexes impliquant la qualité de l'eau, les apports en nutriments, les conditions hydrodynamiques et les caractéristiques du milieu. Leur suivi nécessite des observations régulières, précises et reproductibles.

Pendant longtemps, ces observations reposaient principalement sur des relevés de terrain. Bien que indispensables, ces méthodes présentent certaines limites : elles sont souvent coûteuses en temps, difficiles à répéter fréquemment et parfois insuffisantes pour appréhender l'ensemble d'un site.

L'émergence des drones et des capteurs multispectraux a profondément transformé cette situation. Il est désormais possible d'acquérir des images à très haute résolution, de cartographier de vastes surfaces en quelques minutes et d'observer avec une précision inédite les structures présentes sur l'estran.

Cependant, produire des images ne suffit pas.

Entre l'acquisition d'une photographie aérienne et la production d'une information environnementale utile, de nombreuses étapes sont nécessaires : préparation des données, analyse spectrale, calcul d'indices, classification, validation des résultats et interprétation écologique.

Ce livre propose d'explorer l'ensemble de cette chaîne de traitement à travers un cas d'étude concret : le suivi des algues vertes à partir d'images acquises par drone multispectral.

L'objectif n'est pas seulement de présenter des outils ou des méthodes. Il s'agit également de comprendre les principes scientifiques qui les sous-tendent, leurs possibilités, mais aussi leurs limites. Derrière chaque carte produite se cachent des choix méthodologiques, des hypothèses et des incertitudes qu'il est essentiel de connaître pour interpréter correctement les résultats.

Au fil des chapitres, le lecteur découvrira comment préparer des données multispectrales dans QGIS, calculer des indices de végétation, construire des classifications supervisées, utiliser des méthodes avancées de machine learning et explorer les possibilités offertes par l'intelligence artificielle et le deep learning. Il apprendra également à évaluer la qualité des résultats obtenus, à développer un regard critique sur les cartes produites et à automatiser certaines étapes grâce à Python.

Plus qu'un simple manuel technique, cet ouvrage se veut une introduction à la démarche scientifique en télédétection environnementale. Les outils évoluent rapidement, les algorithmes se perfectionnent et de nouvelles méthodes apparaissent chaque année. Pourtant,

les principes fondamentaux restent les mêmes : observer, mesurer, comparer, valider et interpréter.

Car la finalité de la télédétection n'est pas la production de cartes.

Sa véritable ambition est de transformer des images en connaissances et de fournir aux scientifiques, aux gestionnaires et aux décideurs des informations fiables pour mieux comprendre et préserver les milieux naturels.

C'est ce voyage, depuis le pixel jusqu'à la compréhension du territoire, que ce livre vous invite à entreprendre.

Comment est organisé cet ouvrage ?

Cet ouvrage suit le déroulement complet d'un projet de télédétection appliqué au suivi des algues vertes sur l'estran. Chaque partie correspond à une étape importante de la chaîne d'analyse, depuis l'acquisition des données jusqu'à l'interprétation des résultats et leur automatisation.

Première partie — Les bases scientifiques et les données drone

La première partie introduit le contexte environnemental et les principes fondamentaux de la télédétection multispectrale. Le lecteur découvre les spécificités des milieux littoraux, les caractéristiques du drone DJI Mavic 3 Multispectral ainsi que les notions essentielles de signature spectrale, de réflectance et de calibration radiométrique.

Une attention particulière est portée aux contraintes propres à l'estran : influence des marées, humidité du substrat, présence d'eau résiduelle, ombres et conditions d'éclairage. Ces éléments conditionnent directement la qualité des données et leur interprétation.

Cette première partie fournit les bases scientifiques indispensables à la compréhension des traitements présentés dans la suite de l'ouvrage.

Deuxième partie — Pré-traitements géomatiques sous QGIS

Avant toute analyse, les données doivent être préparées et contrôlées.

Cette deuxième partie présente les opérations de pré-traitement réalisées dans QGIS : vérification des bandes multispectrales, inspection des histogrammes, analyse des valeurs de réflectance, création de masques, calcul d'indices spectraux et exploration des signatures spectrales.

Le lecteur apprend également à identifier les principales sources d'erreur susceptibles d'affecter la qualité des résultats. Cette phase de préparation est souvent moins spectaculaire

que la classification elle-même, mais elle constitue l'une des étapes les plus importantes de toute étude de télédétection.

Troisième partie — SCP : le cœur de la chaîne de classification

Cette partie est consacrée au plugin Semi-Automatic Classification Plugin (SCP), véritable centre de la chaîne de classification supervisée dans QGIS.

Le lecteur découvre comment préparer un Band Set, calculer des indices spectraux, créer des régions d'intérêt (ROI), analyser les signatures spectrales et entraîner différents modèles de classification.

Une attention particulière est portée à la qualité des échantillons d'apprentissage, à la validation des résultats et à l'interprétation des matrices de confusion.

Cette partie permet de comprendre comment transformer des données multispectrales en cartes thématiques exploitables.

Quatrième partie — Comparaison des méthodes de classification

Une même image peut être analysée de nombreuses façons.

Cette partie compare plusieurs approches : classifications statistiques traditionnelles, segmentation orientée objet, machine learning et deep learning.

Les outils présentés incluent notamment Orfeo ToolBox, GRASS GIS, Random Forest et le plugin Deepness. Le lecteur découvre les avantages, les limites et les domaines d'application de chaque méthode.

L'objectif n'est pas de rechercher un algorithme universel, mais de comprendre comment choisir l'approche la plus adaptée à une problématique donnée.

Cinquième partie — Intelligence artificielle et amélioration des modèles

Les développements récents du deep learning ouvrent de nouvelles perspectives pour l'analyse des images drone.

Cette partie introduit les principes du transfert d'apprentissage et du fine-tuning à travers l'amélioration progressive d'un modèle de segmentation des algues. Elle montre comment adapter un modèle existant à de nouveaux sites d'étude et comment intégrer ces modèles dans des workflows géospatiaux opérationnels.

Sixième partie — Interpréter, valider et comprendre

Produire une carte n'est qu'une étape. Encore faut-il savoir l'interpréter.

Cette partie développe une réflexion sur l'évaluation des résultats, l'analyse des erreurs, les matrices de confusion et la discussion scientifique des classifications obtenues.

Le lecteur est progressivement amené à considérer les cartes non comme des vérités absolues, mais comme des représentations du territoire qu'il convient de comprendre, de valider et de questionner.

Cette démarche critique constitue l'un des fondements de la télédétection environnementale.

Septième partie — Automatiser les traitements avec Python

La dernière partie aborde la question de l'automatisation.

À mesure que les campagnes d'acquisition se multiplient, certaines opérations deviennent répétitives et gagnent à être automatisées. À travers l'exemple de la validation des classifications, le lecteur découvre comment Python peut être utilisé pour préparer les points de contrôle, calculer automatiquement les matrices de confusion et produire des rapports reproductibles.

Cette approche illustre le passage progressif d'une analyse ponctuelle à un véritable système de suivi environnemental.

Du pixel à la connaissance du territoire

Au-delà des outils et des méthodes présentés, ce livre poursuit un objectif plus large : montrer comment des images acquises par drone peuvent être transformées en informations utiles pour comprendre les dynamiques environnementales du littoral.

De la réflectance d'un pixel jusqu'à l'interprétation écologique des résultats, chaque étape contribue à construire une connaissance plus fine du territoire observé.

Première partie- Données drone et contexte intertidal

Chapitre 1 — Les drones au service du suivi des algues vertes

Observer le littoral autrement

Le suivi des algues vertes sur les zones littorales nécessite des outils capables d'observer rapidement des milieux complexes, souvent difficiles d'accès et fortement influencés par les marées. Depuis quelques années, les drones occupent une place de plus en plus importante dans les études environnementales côtières.

Leur principal intérêt réside dans leur capacité à produire des images extrêmement détaillées du terrain. Grâce à eux, il devient possible d'analyser avec précision la répartition des tapis algaux, leur extension ou encore leur évolution au fil du temps.

Contrairement aux méthodes d'observation traditionnelles, les drones permettent une acquisition rapide, flexible et ciblée des données, directement au moment le plus pertinent pour l'étude.

Une résolution spatiale très élevée

L'un des principaux avantages des drones est leur très haute résolution spatiale.

Avec un drone multispectral comme le DJI Mavic 3 Multispectral, les images acquises atteignent souvent une résolution comprise entre 3 et 5 centimètres par pixel. Cette finesse d'observation change considérablement la qualité des analyses environnementales.

Une telle précision permet de détecter des patches d'algues de petite taille, mais aussi d'observer la structure détaillée des tapis algaux. Les images produites révèlent avec précision les différentes surfaces présentes sur l'estran, qu'il s'agisse des algues, du sable, des roches ou encore des flaques d'eau résiduelle.

À titre de comparaison, les satellites d'observation classiques proposent généralement des résolutions comprises entre 10 et 30 mètres par pixel. Les drones permettent donc une observation beaucoup plus fine des phénomènes locaux.

Une grande flexibilité d'acquisition

Les milieux intertidaux évoluent rapidement sous l'effet des marées, de l'ensoleillement ou encore des conditions météorologiques. Dans ce contexte, la flexibilité des drones constitue un avantage majeur.

Les acquisitions peuvent être réalisées à des dates très précises et à des horaires adaptés aux besoins de l'étude. Il devient ainsi possible de programmer des relevés en fonction des marées, notamment à marée basse lorsque les zones d'échouage sont visibles.

Cette souplesse permet également de suivre un événement ponctuel ou de répéter des acquisitions à intervalles rapprochés afin d'observer l'évolution des surfaces algales dans le temps. Le drone devient ainsi un outil particulièrement adapté aux suivis environnementaux nécessitant des observations fréquentes et réactives.

Une observation locale et ciblée

Les drones sont également très utiles pour l'étude de secteurs localisés. Ils permettent d'observer des zones spécifiques, des sites sensibles ou encore des espaces difficilement accessibles à pied.

Cette capacité d'observation ciblée en fait un outil particulièrement pertinent pour les études locales, les expérimentations scientifiques et les suivis environnementaux détaillés. Les drones sont aussi largement utilisés pour valider des méthodes d'analyse ou produire des données de référence destinées à la recherche.

Le drone agit alors comme un véritable outil d'investigation de terrain capable de fournir des informations très précises sur des espaces restreints.

Chapitre 2 — Drone et satellite : des approches complémentaires

Deux échelles d'observation différentes

Les données acquises par drone ne remplacent pas les images satellitaires. Les deux approches sont au contraire complémentaires, car elles correspondent à des échelles d'analyse différentes.

Les satellites permettent une vision globale et régulière des territoires, tandis que les drones offrent une observation locale extrêmement détaillée. Ensemble, ils constituent des outils particulièrement efficaces pour le suivi environnemental des zones côtières.

Sentinel-2 : l'observation à grande échelle

Le programme européen Sentinel-2 fournit des images multispectrales particulièrement utiles pour le suivi environnemental.

Les satellites Sentinel-2 proposent une résolution spatiale d'environ 10 mètres tout en assurant une couverture très large des territoires. Leur fréquence d'acquisition, proche de cinq jours, permet de suivre régulièrement l'évolution des milieux naturels.

Ces caractéristiques rendent les données Sentinel-2 particulièrement adaptées à la détection de phénomènes à grande échelle, au suivi régional des proliférations algales ainsi qu'à l'analyse des dynamiques saisonnières.

Les images satellitaires constituent donc un excellent outil de surveillance globale des zones littorales.

Le drone : l'observation à très haute résolution

À l'inverse des satellites, les drones permettent une analyse extrêmement fine des surfaces observées.

Grâce à leur résolution centimétrique, ils facilitent l'identification détaillée des tapis algaux et l'observation de structures fines parfois invisibles sur les images satellitaires. Les drones permettent également de produire des cartographies très précises des zones étudiées.

Ils jouent ainsi un rôle essentiel dans la caractérisation des surfaces algales, mais aussi dans la validation des observations issues des satellites.

Une complémentarité essentielle

Dans de nombreux projets scientifiques, une approche combinée est aujourd'hui privilégiée.

Le principe consiste généralement à utiliser les satellites pour repérer les zones d'intérêt à grande échelle, puis à mobiliser les drones pour analyser ces secteurs avec davantage de précision.

Les données drone servent également de référence pour améliorer l'interprétation des images satellitaires et affiner les méthodes de classification.

Cette complémentarité entre observation globale et observation locale constitue aujourd'hui l'un des fondements des approches modernes de suivi environnemental.

Chapitre 3 — La chaîne de traitement des données drone

De l'acquisition à l'analyse

Le suivi des algues vertes à partir d'images drone repose sur une chaîne de traitement composée de plusieurs étapes successives. Chaque étape joue un rôle essentiel dans la qualité des résultats finaux.

L'acquisition des images

La première étape correspond à la réalisation du vol drone et à l'acquisition des images.

Plusieurs paramètres influencent directement la qualité des données produites. L'altitude de vol détermine notamment la résolution des images, tandis que le recouvrement entre les clichés conditionne la qualité des traitements photogrammétriques. Les conditions d'éclairage ainsi que l'état de la marée jouent également un rôle important dans la qualité des observations.

Les drones multispectraux enregistrent plusieurs bandes spectrales permettant d'analyser la végétation et les surfaces biologiques. Ces bandes constituent la base des traitements ultérieurs.

La photogrammétrie et la production d'orthomosaïques

Les images individuelles acquises pendant le vol doivent ensuite être assemblées afin de produire une représentation cohérente du site étudié.

Cette étape de photogrammétrie permet de générer une orthomosaïque géoréférencée et, dans certains cas, un modèle numérique de surface.

Plusieurs logiciels spécialisés sont couramment utilisés pour réaliser ces traitements, notamment DJI Terra, Pix4D ou encore Agisoft Metashape.

L'orthomosaïque constitue alors une image continue du terrain, prête à être analysée dans un système d'information géographique.

L'analyse dans un SIG

Les données produites sont ensuite intégrées dans un système d'information géographique comme QGIS.

Plusieurs traitements peuvent alors être réalisés, notamment l'extraction des différentes bandes spectrales, le calcul d'indices de végétation ou encore l'analyse des signatures spectrales.

Parmi les indices fréquemment utilisés figurent notamment le NDVI et le NDRE, qui facilitent l'identification des surfaces végétalisées ou algales.

Ces traitements permettent d'exploiter pleinement les informations contenues dans les images multispectrales.

La classification des surfaces

Une étape de classification permet ensuite de distinguer les différentes catégories de surfaces présentes sur l'estran.

L'objectif est de séparer les zones couvertes par les algues des surfaces minérales, des zones d'eau ou des autres éléments présents dans l'image.

Cette classification peut être réalisée à l'aide de méthodes supervisées ou non supervisées selon les objectifs de l'étude et les données disponibles.

Les résultats obtenus prennent généralement la forme de cartes thématiques facilitant l'analyse spatiale des phénomènes observés.

La production d'indicateurs

Une fois les cartes produites, différents indicateurs peuvent être calculés afin de quantifier les phénomènes étudiés.

Il devient alors possible d'estimer la surface totale couverte par les algues, d'analyser leur distribution spatiale ou encore de suivre leur évolution temporelle.

Ces résultats sont particulièrement utiles pour le suivi environnemental, l'étude des dynamiques saisonnières et l'aide à la décision dans les politiques de gestion des zones littorales.

Les drones deviennent ainsi de véritables outils scientifiques capables de produire des informations précises, quantitatives et répétables sur l'état des milieux côtiers.

Chapitre 4 — Comprendre les capteurs du drone multispectral

Deux systèmes de caméras complémentaires

Le DJI Mavic 3 Multispectral repose sur deux systèmes de caméras distincts, chacun ayant un rôle spécifique dans l'acquisition et l'analyse des données.

Le premier système correspond à une caméra RGB classique de 20 mégapixels. Le second est constitué de plusieurs capteurs multispectraux dédiés à l'analyse spectrale de la végétation.

Cette combinaison permet au drone de produire à la fois des images visuelles détaillées et des données scientifiques exploitables pour l'analyse environnementale.

La caméra RGB : une vision proche de l'œil humain

La caméra RGB enregistre les longueurs d'onde visibles par l'œil humain. Elle capture ainsi trois bandes spectrales principales correspondant au bleu, au vert et au rouge.

La bande bleue se situe autour de 450 nanomètres, la bande verte autour de 560 nanomètres et la bande rouge autour de 650 nanomètres.

Cette caméra joue un rôle essentiel dans les traitements photogrammétriques. Elle est principalement utilisée pour produire des orthomosaïques visuelles de très haute qualité ainsi que des modèles photoréalistes du terrain.

Les images RGB facilitent également l'interprétation visuelle des sites étudiés. Elles permettent notamment de repérer rapidement les zones d'échouage, les structures du littoral ou les variations visibles de surface.

Les caméras multispectrales : l'analyse scientifique des surfaces

En complément de la caméra RGB, le drone possède quatre capteurs multispectraux spécialement conçus pour l'analyse de la végétation et des surfaces biologiques.

Ces capteurs enregistrent différentes bandes spectrales stratégiques : le vert autour de 560 nanomètres, le rouge autour de 650 nanomètres, le red edge autour de 730 nanomètres et le proche infrarouge autour de 860 nanomètres.

Contrairement aux images classiques, ces bandes permettent d'analyser la manière dont les surfaces réfléchissent ou absorbent le rayonnement électromagnétique.

Ces informations sont particulièrement utiles pour :

- l'analyse de la végétation ;
- le calcul d'indices spectraux ;
- la détection de biomasse ;
- l'agriculture de précision ;
- le suivi environnemental.

Les bandes rouge, red edge et proche infrarouge jouent notamment un rôle central dans l'étude de l'activité chlorophyllienne.

Pourquoi le multispectral ne comporte-t-il pas de bande bleue ?

Une question revient souvent lors de l'étude des capteurs multispectraux : pourquoi la bande bleue n'est-elle pas présente parmi les capteurs dédiés à l'analyse spectrale ?

La raison est principalement liée aux propriétés physiques de cette longueur d'onde.

La bande bleue est fortement affectée par la diffusion atmosphérique, ce qui peut réduire la stabilité des mesures. Elle pénètre également davantage dans l'eau, ce qui complique l'interprétation des surfaces en milieu humide ou intertidal.

Enfin, cette bande est généralement moins discriminante pour l'analyse de la chlorophylle que les bandes rouge, red edge ou proche infrarouge.

Pour cette raison, les principaux indices de végétation reposent surtout sur les relations entre la bande rouge et le proche infrarouge, ou entre la bande red edge et le proche infrarouge.

Les indices les plus utilisés dans ce contexte sont notamment :

Le NDVI :

$$NDVI = \frac{NIR - Rouge}{NIR + Rouge}$$

et le NDRE :

$$NDRE = \frac{NIR - RedEdge}{NIR + RedEdge}$$

Ces indices permettent d'évaluer l'activité végétale et de mettre en évidence certaines propriétés biologiques des surfaces observées.

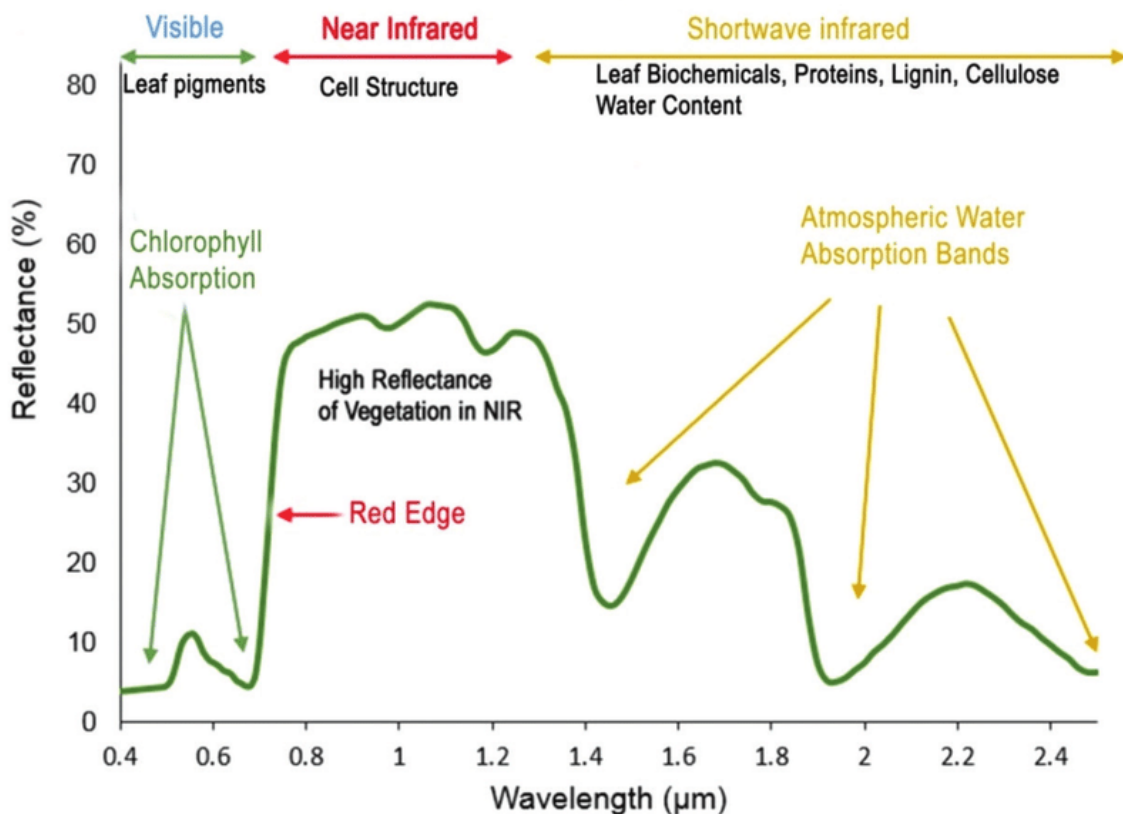
Signature spectrale des algues vertes

Les algues vertes présentent une signature spectrale caractéristique qui permet leur détection à partir des données multispectrales.

Comme les végétaux terrestres, elles absorbent fortement le rayonnement dans le rouge en raison de la présence de chlorophylle. À l'inverse, elles réfléchissent davantage le rayonnement dans le proche infrarouge.

Une transition marquée apparaît également dans la zone du red edge, correspondant au passage entre les fortes absorptions du rouge et les fortes réflectances du proche infrarouge.

Cette signature spectrale constitue la base de nombreux traitements de classification et d'identification des surfaces algales.



The spectral reflectance curve of vegetation. The major absorption and reflectance features are indicated. Roman, Anamaria & Ursu, Tudor. (2016). Multispectral satellite imagery and airborne laser scanning techniques for the detection of archaeological vegetation marks.

Les particularités du milieu intertidal

L'interprétation des signatures spectrales en milieu littoral reste toutefois plus complexe qu'en milieu terrestre classique.

Dans les zones intertidales, plusieurs facteurs peuvent perturber les mesures spectrales. L'humidité du substrat influence fortement la réflectance des surfaces. La présence d'eau résiduelle peut également modifier le signal observé, notamment dans les bandes proches du proche infrarouge.

Les biofilms présents sur les sédiments constituent un autre facteur de complexité, car ils peuvent produire des signatures proches de celles des algues.

Ces éléments expliquent pourquoi l'analyse des algues vertes sur l'estran nécessite souvent des traitements adaptés et une interprétation prudente des indices spectraux.

Chapitre 5 — Les contraintes du milieu intertidal

Un environnement complexe à observer

Le milieu intertidal constitue un environnement particulièrement complexe pour l'acquisition et l'analyse d'images multispectrales. Située entre les niveaux de marée haute et de marée basse, cette zone littorale évolue constamment sous l'effet des variations marines, des conditions météorologiques et de la nature du substrat.

Contrairement aux surfaces continentales relativement homogènes, l'estran présente une très forte variabilité spatiale. Cette complexité influence directement la qualité des mesures spectrales et complique parfois l'interprétation des données acquises par drone.

Pour produire des analyses fiables, il est donc indispensable de comprendre les principales contraintes propres à ce type d'environnement.

Une forte variabilité spatiale

L'une des premières difficultés du milieu intertidal réside dans son hétérogénéité spatiale.

Les surfaces observées présentent souvent une micro-topographie complexe composée de petites dépressions, de rides sédimentaires, de blocs rocheux ou encore de zones de transition entre différents types de substrats.

À très haute résolution, ces variations deviennent particulièrement visibles sur les images drone. Deux pixels très proches peuvent ainsi correspondre à des surfaces complètement différentes.

Cette hétérogénéité influence fortement les signatures spectrales et peut compliquer les traitements de classification automatique.

La présence d'eau résiduelle

Même à marée basse, l'eau reste omniprésente sur l'estran.

Des flaques résiduelles persistent souvent dans les dépressions du terrain, tandis que certaines surfaces demeurent partiellement saturées en eau. La turbidité peut également modifier la réponse spectrale des zones humides.

Or, l'eau absorbe fortement le rayonnement dans certaines longueurs d'onde, notamment dans le proche infrarouge. Cette propriété perturbe directement les indices de végétation et peut masquer la signature réelle des algues.

Dans certains cas, une même zone peut présenter des réponses spectrales très différentes selon la quantité d'eau présente au moment de l'acquisition.

Les effets du substrat humide

L'humidité du substrat constitue un autre facteur majeur de complexité.

Les sédiments humides présentent souvent des signatures spectrales proches de celles des surfaces algales, en particulier lorsque les algues sont fines ou dispersées.

Cette similarité peut entraîner des confusions lors des traitements de classification. Certaines zones humides peuvent ainsi être interprétées à tort comme des surfaces couvertes par des algues.

Ce phénomène est particulièrement marqué dans les environnements vaseux ou sableux où la teneur en eau varie fortement sur de courtes distances.

Les contraintes liées à l'éclairage

Les conditions lumineuses influencent également fortement la qualité des données multispectrales.

Les ombres portées par les reliefs, les rochers ou les variations de terrain peuvent modifier localement la réflectance observée. Les reflets spéculaires provoqués par l'eau ou les surfaces humides constituent également une source importante de perturbation.

Ces reflets produisent parfois des zones extrêmement lumineuses qui dégradent la qualité des mesures spectrales et compliquent l'analyse automatique des images.

L'angle du soleil, la présence de nuages ou encore l'heure de l'acquisition jouent donc un rôle important dans la qualité des données obtenues.

Une interprétation qui nécessite de la prudence

L'ensemble de ces contraintes montre que l'analyse multispectrale en milieu intertidal ne peut pas être abordée de la même manière qu'en milieu terrestre classique.

La variabilité des substrats, l'humidité, la présence d'eau et les effets lumineux introduisent de nombreuses perturbations dans les signatures spectrales observées.

Pour cette raison, l'interprétation des données nécessite une approche prudente associant souvent plusieurs méthodes complémentaires : observations de terrain, analyses visuelles, indices spectraux et classifications supervisées.

La compréhension des particularités du milieu intertidal constitue ainsi une étape essentielle pour produire des cartographies fiables des surfaces algales.

Chapitre 6 — L'importance des conditions d'acquisition

Pourquoi le moment de l'acquisition est-il crucial ?

En télédétection par drone, la qualité scientifique des résultats dépend directement des conditions d'acquisition des images. Cette réalité est encore plus marquée en milieu intertidal, où les surfaces observées évoluent constamment sous l'effet des marées, de l'humidité et des conditions météorologiques.

Une acquisition réalisée dans de mauvaises conditions peut produire des données difficiles à interpréter, voire inutilisables pour certaines analyses. Les signatures spectrales enregistrées par les capteurs sont en effet fortement influencées par l'état du milieu au moment du vol.

Trois facteurs principaux doivent donc être pris en compte lors de la planification des acquisitions :

- le niveau de la marée ;
- les conditions d'éclairage ;
- les conditions météorologiques, notamment le vent.

La préparation du vol constitue ainsi une étape essentielle du protocole scientifique.

L'importance des fenêtres de marée

Observer l'estran à marée basse

Le premier facteur déterminant pour l'acquisition d'images de l'estran est le niveau de la marée.

Pour observer correctement les surfaces intertidales, les acquisitions doivent être réalisées à marée basse. À marée haute, une grande partie des surfaces algales reste immergée. Or, l'eau absorbe fortement le rayonnement dans le proche infrarouge, ce qui modifie considérablement les signatures spectrales enregistrées par les capteurs.

Dans ces conditions, les indices de végétation deviennent beaucoup moins fiables et certaines zones deviennent pratiquement impossibles à analyser.

Une surface recouverte d'eau peut ainsi masquer totalement la présence des algues sous-jacentes.

Les avantages des marées de vives-eaux

Les marées de vives-eaux correspondent aux marées les plus basses du cycle lunaire. Elles offrent des conditions particulièrement favorables pour l'acquisition de données sur l'estran.

Lors de ces périodes, l'exposition des surfaces intertidales est maximale. Les zones couvertes par les algues deviennent plus accessibles à l'observation et la quantité d'eau résiduelle est souvent réduite.

Cette situation permet d'observer la plus grande surface possible et d'améliorer la qualité des analyses spectrales.

Dans de nombreux projets scientifiques, les acquisitions sont volontairement programmées pendant ces fenêtres de marée afin d'optimiser les conditions d'observation.

Une surface observable très variable

La surface réellement visible dépend directement du niveau de la marée.

Même une différence de quelques dizaines de centimètres peut modifier fortement l'étendue des surfaces exposées. Certaines zones de l'estran ne deviennent visibles que lors des marées les plus basses, tandis que d'autres restent immergées plus longtemps.

Cette variabilité influence directement la comparabilité des acquisitions dans le temps.

Pour cette raison, il est essentiel de documenter précisément les conditions de marée associées à chaque vol drone.

La question de la reproductibilité temporelle

Dans les études de suivi environnemental, l'objectif est souvent de comparer des acquisitions réalisées à différentes dates.

Pour garantir la cohérence des analyses, il est important de répéter les vols dans des conditions aussi similaires que possible. Les acquisitions doivent idéalement être réalisées :

- à une période de marée comparable ;
- durant la même saison ;
- sous des conditions météorologiques proches.

Cette standardisation permet de limiter les variations liées aux conditions d'acquisition et de mieux isoler les véritables évolutions écologiques observées sur le terrain.

Les conditions d'éclairage

Un facteur majeur en télédétection

L'éclairage influence directement la quantité de rayonnement réfléchi par les surfaces et enregistrée par les capteurs du drone.

Toute variation lumineuse peut modifier les valeurs spectrales observées et compliquer l'interprétation des images.

La qualité des conditions d'éclairage joue donc un rôle essentiel dans la fiabilité des analyses multispectrales.

Les avantages de la lumière diffuse

Les conditions idéales correspondent souvent à un ciel légèrement couvert produisant une lumière diffuse.

Dans cette configuration, l'éclairage est plus homogène et les ombres sont moins marquées. Les variations d'illumination entre les différentes parties de l'image sont réduites, ce qui améliore la comparabilité des pixels.

Les analyses spectrales deviennent alors plus stables et plus fiables.

Les limites d'un fort ensoleillement

À l'inverse, un soleil très intense peut produire des contrastes importants dans l'image.

Les reliefs, les rochers ou les irrégularités du terrain génèrent alors des ombres marquées qui modifient localement la réflectance observée.

Ces variations lumineuses peuvent perturber les traitements automatiques et produire des valeurs spectrales atypiques.

Les effets des ombres

Les ombres apparaissent fréquemment sur l'estran, notamment derrière les rochers ou dans les micro-dépressions du terrain.

Dans ces zones, la quantité de lumière reçue par la surface est fortement réduite. Les capteurs enregistrent alors des signatures spectrales différentes de celles observées sur les surfaces éclairées.

Ces anomalies peuvent compliquer la classification des images et introduire des erreurs dans les indices de végétation.

Les reflets spéculaires

Les reflets spéculaires, souvent appelés *glint*, constituent une autre difficulté importante en milieu littoral.

Ils apparaissent lorsque la lumière du soleil se réfléchit directement sur la surface de l'eau ou sur des surfaces très humides.

Ces phénomènes sont particulièrement fréquents sur l'estran en présence de flaques résiduelles ou de sédiments saturés en eau.

Les reflets spéculaires produisent parfois des pixels extrêmement lumineux dont les valeurs radiométriques deviennent aberrantes. Ils peuvent ainsi perturber fortement l'analyse des images multispectrales.

L'influence du vent

Un facteur souvent sous-estimé

Le vent influence directement l'état de surface de l'eau présente sur l'estran.

Lorsque le vent devient important, les surfaces d'eau s'agitent et les reflets lumineux deviennent plus irréguliers.

Cette agitation introduit du bruit radiométrique dans les images et réduit la stabilité des mesures spectrales.

Les conditions les plus favorables

Les meilleures acquisitions sont généralement réalisées dans des conditions de vent faible.

Une surface d'eau relativement calme limite les perturbations lumineuses et améliore la qualité des données enregistrées par les capteurs multispectraux.

La stabilité atmosphérique contribue également à améliorer la précision des traitements photogrammétriques.

Les conséquences d'une mauvaise acquisition

Des impacts directs sur les traitements

Une acquisition réalisée dans de mauvaises conditions peut entraîner de nombreuses difficultés lors du traitement des données.

Les images peuvent présenter :

- des valeurs spectrales incohérentes ;
- des artefacts dans les indices de végétation ;
- des erreurs de classification ;
- des anomalies radiométriques.

Dans certains cas, ces problèmes rendent les résultats difficilement exploitables scientifiquement.

Un principe fondamental en télédétection

En télédétection, la qualité des résultats dépend avant tout de la qualité des données acquises sur le terrain.

Les traitements numériques permettent de corriger certaines imperfections, mais ils ne peuvent pas compenser totalement une acquisition réalisée dans de mauvaises conditions.

C'est pourquoi la planification du vol constitue une étape fondamentale du protocole d'étude.

Conclusion de la séquence

Les acquisitions drone en milieu littoral doivent être préparées avec une grande rigueur.

Le niveau de la marée, les conditions d'éclairage et les conditions météorologiques influencent directement la qualité des données multispectrales.

Une planification adaptée permet de limiter les perturbations liées à l'environnement et de garantir la qualité scientifique des analyses réalisées par la suite.

Dans les études environnementales, une bonne acquisition constitue toujours la première étape vers des résultats fiables.

Chapitre 7 — La calibration radiométrique : rendre les images scientifiquement exploitables

Une image “belle” n’est pas forcément une bonne donnée scientifique

Lorsqu’on observe une orthomosaïque produite à partir d’un drone, la première réaction consiste souvent à juger sa qualité visuelle. Une image nette, contrastée et agréable à l’écran peut donner l’impression d’être parfaitement exploitable.

Pourtant, en télédétection scientifique, l’aspect visuel ne suffit pas.

Une image peut sembler excellente tout en étant inutilisable pour des analyses quantitatives, notamment lorsqu’il s’agit de comparer plusieurs dates, plusieurs saisons ou plusieurs sites. La qualité scientifique d’une image dépend avant tout de la fiabilité radiométrique des valeurs enregistrées par le capteur.

C’est précisément le rôle de la calibration radiométrique.

DN et réflectance : deux notions différentes

Les DN : des valeurs brutes dépendantes des conditions d’acquisition

Les images produites par un capteur multispectral contiennent d’abord des valeurs numériques appelées DN, pour *Digital Number*.

Ces valeurs correspondent à l’intensité lumineuse enregistrée par le capteur pour chaque pixel. Elles ne représentent pas directement les propriétés physiques réelles des surfaces observées.

Les DN dépendent de nombreux facteurs extérieurs, notamment :

- la luminosité ambiante ;
- l’exposition du capteur ;

- l'heure du vol ;
- les conditions météorologiques ;
- l'angle du soleil.

Ainsi, une même surface observée à deux moments différents peut présenter des valeurs DN très différentes sans que ses propriétés réelles aient changé.

Les DN constituent donc des valeurs relatives et non directement comparables dans le temps.

La réflectance : une mesure physique comparable

La réflectance correspond à la proportion de lumière réfléchie par une surface pour une longueur d'onde donnée.

Contrairement aux DN, elle représente une grandeur physique beaucoup plus stable et comparable. Elle est généralement exprimée sous forme de valeur comprise entre 0 et 1, ou parfois en pourcentage.

Les différentes surfaces présentent des comportements spectraux distincts. L'eau possède par exemple une très faible réflectance dans le proche infrarouge, tandis que les algues présentent souvent une réflectance élevée dans cette bande spectrale. Le sable sec occupe généralement une position intermédiaire.

La réflectance permet ainsi de comparer des acquisitions réalisées à différentes dates ou sur différents sites. Elle constitue également la base du calcul fiable des indices de végétation.

Pourquoi les DN ne suffisent pas

Deux images acquises sur le même site peuvent présenter des valeurs DN très différentes simplement en raison des conditions d'acquisition.

Une variation de l'angle solaire, la présence d'un voile nuageux, une différence d'humidité ou encore une modification automatique de l'exposition du capteur peuvent modifier fortement les valeurs enregistrées.

Par exemple, une zone couverte d'algues peut présenter une valeur moyenne NIR de 120 lors d'un premier vol et de 170 lors d'un second vol.

Cette différence ne signifie pas nécessairement que l'état des algues a évolué. Elle peut simplement résulter d'un changement des conditions lumineuses au moment de l'acquisition.

Sans correction radiométrique, il devient donc très difficile d'interpréter correctement les variations observées dans les images.

Le rôle du panneau de calibration

Convertir les DN en réflectance

Pour transformer les valeurs DN en réflectance exploitable scientifiquement, on utilise un panneau de calibration radiométrique.

Ce panneau possède des propriétés connues et stables. Sa réflectance a été mesurée précisément pour les différentes bandes spectrales utilisées par le capteur.

Le protocole est relativement simple. Avant le vol, le panneau est photographié par le drone dans les mêmes conditions lumineuses que celles du site étudié. Une seconde photographie est généralement réalisée après le vol afin de prendre en compte les éventuelles variations d'éclairage survenues pendant l'acquisition.

Ces images servent ensuite de référence pour corriger les valeurs enregistrées par le capteur multispectral.

Pourquoi le panneau est-il si important ?

Le panneau de calibration joue un rôle essentiel dans la qualité scientifique des données.

Il permet de corriger les variations liées :

- aux changements de luminosité ;
- aux différences d'exposition ;
- aux variations météorologiques ;
- aux écarts entre plusieurs acquisitions.

Sans cette étape, il devient pratiquement impossible de déterminer si une différence observée entre deux images correspond réellement à une évolution écologique ou simplement à une variation des conditions d'acquisition.

La calibration radiométrique constitue donc une condition indispensable pour réaliser des comparaisons temporelles fiables.

Le capteur de lumière incident

Le DJI Mavic 3 Multispectral possède également un capteur de lumière embarqué destiné à améliorer la qualité radiométrique des acquisitions.

Ce capteur mesure l'irradiance, c'est-à-dire la quantité de lumière reçue pendant le vol. Il enregistre également les variations d'éclairement susceptibles d'apparaître au cours de l'acquisition.

Ces informations permettent d'affiner les corrections radiométriques appliquées aux images.

Cependant, même si ce système améliore considérablement la qualité des données, il ne remplace pas totalement l'utilisation d'un panneau de calibration radiométrique au sol.

Les problèmes fréquemment rencontrés

L'absence de calibration

L'un des problèmes les plus fréquents est l'absence de calibration radiométrique.

Lorsque aucun panneau n'est utilisé, les valeurs obtenues restent relatives. Les comparaisons temporelles deviennent alors très fragiles et les interprétations scientifiques beaucoup moins fiables.

Les variations d'éclairage pendant le vol

Les conditions lumineuses peuvent évoluer rapidement pendant une acquisition.

Le passage de nuages, la diminution progressive de la hauteur du soleil ou encore les variations d'angle solaire modifient directement les valeurs enregistrées par les capteurs.

Même au cours d'un même vol, certaines parties de l'orthomosaïque peuvent ainsi présenter des différences radiométriques importantes.

Les ombres et les nuages

Les ombres réduisent localement la quantité de lumière reçue par les surfaces observées. Les nuages peuvent quant à eux provoquer une diminution générale de l'éclairement ou produire des variations lumineuses localisées.

Certains indices comme le NDVI résistent relativement bien aux variations homogènes d'éclairage :

$$NDVI = \frac{NIR - Rouge}{NIR + Rouge}$$

Cependant, d'autres traitements radiométriques ou classifications plus fines peuvent être fortement perturbés par ces variations lumineuses.

Les conséquences pour les comparaisons temporelles

L'objectif principal d'un suivi environnemental consiste souvent à comparer plusieurs acquisitions réalisées à différentes dates, saisons ou années.

Sans calibration radiométrique rigoureuse, il devient impossible de savoir si une différence observée correspond :

- à une véritable évolution écologique ;
- ou simplement à une variation des conditions d'acquisition.

Cette ambiguïté limite fortement la valeur scientifique des analyses.

Les études temporelles nécessitent donc des protocoles d'acquisition standardisés et une calibration radiométrique systématique.

A retenir

Une image visuellement impressionnante n'est pas nécessairement une image scientifiquement exploitable.

La télédétection ne consiste pas uniquement à produire de belles orthomosaïques, mais surtout à obtenir des mesures fiables, comparables et reproductibles dans le temps.

Sans calibration radiométrique, les comparaisons temporelles restent extrêmement fragiles.

Conclusion

La calibration radiométrique constitue une étape essentielle de toute démarche scientifique utilisant des données multispectrales.

Elle permet de produire des cartes fiables, de comparer plusieurs dates d'acquisition et de suivre l'évolution des surfaces algales dans le temps.

Bien que parfois négligée dans les projets débutants, cette étape conditionne directement la qualité scientifique des analyses réalisées par la suite.

Chapitre 8 — Les limites de l'interprétation en milieu intertidal

Comprendre les limites des observations multispectrales

L'utilisation de drones multispectraux offre aujourd'hui des possibilités remarquables pour le suivi des algues vertes sur l'estran. La finesse spatiale des images et la richesse des informations spectrales permettent d'observer des phénomènes difficilement visibles avec d'autres outils.

Cependant, malgré les performances des capteurs et la qualité des orthomosaïques produites, l'interprétation des données reste complexe en milieu intertidal.

Le littoral constitue un environnement particulièrement dynamique où de nombreux facteurs physiques, biologiques et radiométriques modifient les signatures spectrales enregistrées par les capteurs.

Il est donc essentiel de développer un regard critique sur les cartes produites et de comprendre que les images multispectrales ne représentent jamais directement les objets biologiques observés sur le terrain.

Les limites physiques

La présence permanente d'eau résiduelle

Même à marée basse, l'estran n'est jamais totalement sec.

De nombreuses surfaces restent partiellement recouvertes par des flaques d'eau, des zones saturées ou de très minces films d'eau invisibles à distance mais suffisants pour modifier fortement les mesures spectrales.

Cette présence d'eau constitue une difficulté majeure en télédétection multispectrale, car l'eau absorbe fortement le rayonnement dans le proche infrarouge.

Ainsi, une zone couverte d'algues mais recouverte de quelques millimètres d'eau peut produire une signature spectrale proche de celle de l'eau libre ou d'un substrat humide.

Ce phénomène peut conduire à une sous-détection des surfaces algales et introduire une forte variabilité dans les indices de végétation calculés à partir des images.

Les effets de la turbidité

Lorsque l'eau est présente sur l'estran, sa turbidité influence également le signal enregistré par le capteur.

Une eau chargée en particules ou en sédiments diffuse davantage la lumière et modifie la réflectance apparente des surfaces observées.

Dans ces conditions, le signal spectral mesuré ne provient plus uniquement du fond ou des algues présentes sous l'eau. Il résulte d'un mélange complexe entre :

- la colonne d'eau ;
- les particules en suspension ;
- le substrat ;
- les éventuelles algues présentes.

L'interprétation des données devient alors beaucoup plus délicate.

L'influence de la profondeur d'eau

Même une très faible profondeur d'eau peut modifier fortement la réflectance observée.

Lorsque la lumière traverse l'eau, une partie du rayonnement est absorbée tandis qu'une autre partie est diffusée. Plus la profondeur augmente, plus le signal provenant du fond s'atténue.

Dans certaines situations, les algues présentes sous une faible lame d'eau deviennent pratiquement invisibles pour le capteur multispectral.

Cette limitation constitue l'un des principaux défis du suivi des algues en milieu intertidal.

Les limites biologiques

La présence de biofilms

Les surfaces intertidales sont souvent recouvertes de biofilms composés de micro-organismes tels que des micro-algues, des bactéries ou des diatomées.

Ces biofilms modifient les propriétés optiques du substrat et peuvent produire des signatures spectrales proches de celles de certaines algues vertes.

Lors des traitements automatiques, ils risquent donc d'être interprétés comme des tapis algaux alors qu'ils correspondent à des phénomènes biologiques différents.

Cette confusion est particulièrement fréquente sur les vasières et les substrats fins humides.

Les algues partiellement immergées

Les algues présentes sur l'estran ne sont pas toujours complètement exposées à l'air libre.

Certaines restent partiellement immergées, recouvertes par un film d'eau ou situées dans de petites dépressions où l'humidité persiste plus longtemps.

Dans ces situations, le signal spectral enregistré par le capteur correspond à un mélange complexe entre :

- la signature des algues ;
- celle de l'eau ;
- celle du substrat.

Cette combinaison réduit fortement la capacité des indices spectraux à détecter correctement les surfaces algales.

Les limites radiométriques

Les effets du substrat humide

L'humidité du substrat constitue une autre source importante de confusion.

Un sable humide ou un sédiment saturé en eau peut présenter une signature spectrale relativement proche de celle des algues, notamment dans certaines bandes du proche infrarouge.

L'humidité modifie en effet les propriétés radiométriques du sol en augmentant l'absorption du rayonnement et en réduisant les contrastes spectraux.

Certaines surfaces humides peuvent alors produire des valeurs d'indices de végétation comparables à celles de zones réellement couvertes par des algues.

Ces situations génèrent des faux positifs dans les cartes de classification.

Les perturbations liées aux ombres

Les ombres constituent également une difficulté importante dans l'analyse des images drone.

Elles peuvent être produites par des rochers, des variations de micro-topographie ou simplement par l'angle du soleil au moment de l'acquisition.

Dans les zones ombragées, la quantité de lumière reçue par les surfaces est fortement réduite. Les valeurs spectrales enregistrées deviennent alors plus faibles et parfois atypiques.

Ces anomalies peuvent perturber les indices de végétation ainsi que les algorithmes de classification automatique.

Les conséquences pour l'interprétation

Ce que représentent réellement les cartes multispectrales

Les cartes produites à partir d'images multispectrales ne représentent pas directement des objets biologiques observés sur le terrain.

Elles représentent avant tout des signatures spectrales.

Autrement dit, les capteurs enregistrent des comportements optiques liés à la manière dont les surfaces réfléchissent la lumière dans différentes longueurs d'onde.

La présence d'algues est ensuite déduite à partir :

- des caractéristiques spectrales observées ;
- des indices de végétation calculés ;
- du contexte écologique du site étudié.

Cette distinction est fondamentale pour interpréter correctement les résultats.

A retenir

On ne cartographie jamais directement les algues elles-mêmes.

On cartographie une signature spectrale dont l'interprétation nécessite des connaissances écologiques, des observations de terrain et une compréhension des limites physiques du milieu intertidal.

Les implications scientifiques

La production de cartes fiables nécessite donc une approche rigoureuse et critique.

Il est indispensable de connaître le contexte écologique du site étudié, de confronter les résultats aux observations de terrain et d'interpréter les classifications avec prudence.

La télédétection doit être considérée comme un outil d'aide à l'observation et à l'analyse environnementale, mais non comme une mesure directe et absolue du phénomène étudié.

Cette approche critique constitue une étape essentielle dans toute démarche scientifique utilisant des données multispectrales en milieu littoral.

Deuxième partie — Pré-traitements géomatiques sous QGIS

Chapitre 9 — Pourquoi les pré-traitements sont essentiels

La partie invisible des projets de télédétection

Lorsqu'on découvre la télédétection, l'attention se porte souvent sur les cartes finales, les indices de végétation ou les classifications automatiques. Pourtant, dans la réalité des projets scientifiques, la plus grande partie du travail ne concerne pas directement l'analyse elle-même.

Dans la majorité des projets de télédétection, une grande partie du temps est consacrée à la préparation des données. Les étapes de vérification, de nettoyage et d'organisation représentent souvent l'essentiel du travail technique avant même le début des analyses.

Cette phase est parfois sous-estimée par les débutants, alors qu'elle conditionne directement la qualité scientifique des résultats obtenus.

Pourquoi cette étape est-elle si importante ?

Les méthodes d'analyse utilisées en télédétection — calcul d'indices, classifications ou modèles statistiques — supposent que les données respectent plusieurs conditions fondamentales.

Les rasters doivent être cohérents géométriquement, cohérents radiométriquement et réellement représentatifs du phénomène étudié.

Si ces conditions ne sont pas respectées, les traitements peuvent produire des résultats scientifiquement incorrects, même lorsque les calculs sont techniquement réalisés sans erreur.

Autrement dit, un indice calculé correctement sur des données incorrectes produira malgré tout une interprétation fausse.

Les principales sources d'erreurs

Les problèmes d'alignement entre bandes multispectrales

Dans un capteur multispectral, chaque bande est généralement enregistrée par un capteur distinct. Malgré les traitements photogrammétriques réalisés après le vol, de légers décalages peuvent subsister entre les bandes.

Ces décalages semblent parfois invisibles à l'œil nu, mais ils deviennent problématiques lors du calcul des indices spectraux.

Les indices comme le NDVI supposent en effet que chaque pixel correspond exactement à la même surface dans toutes les bandes utilisées.

Le NDVI repose par exemple sur la relation suivante :

$$NDVI = \frac{NIR - Rouge}{NIR + Rouge}$$

Si les bandes ne sont pas parfaitement alignées, les pixels comparés ne représentent plus les mêmes objets au sol. Cela peut provoquer :

- des valeurs aberrantes ;
- du bruit spatial ;
- des contours instables ;
- des artefacts dans les cartes produites.

Même un décalage de quelques pixels peut dégrader fortement la qualité des résultats.

Les problèmes radiométriques

Une autre erreur fréquente consiste à travailler avec des données radiométriquement incorrectes.

Certaines images peuvent par exemple être compressées en 8 bits avec des valeurs comprises entre 0 et 255, alors que les traitements scientifiques nécessitent souvent des données en réflectance et en précision Float32.

L'absence de calibration radiométrique ou des incohérences entre bandes peuvent également produire des résultats erronés.

Dans ces situations, les indices calculés ne correspondent plus à la réalité physique observée sur le terrain et les comparaisons temporelles deviennent très fragiles.

Pour cette raison, il est indispensable de vérifier :

- le type de données ;
- la plage de valeurs ;

- les histogrammes des bandes ;
- la cohérence radiométrique générale.

Les surfaces perturbatrices

Certaines surfaces présentes sur l'estran peuvent fortement perturber les analyses multispectrales.

L'eau résiduelle, les ombres, les substrats humides ou encore les surfaces saturées produisent parfois des signatures spectrales ambiguës.

Ces zones peuvent générer :

- des valeurs de NDVI aberrantes ;
- des faux positifs ;
- du bruit dans les classifications.

Si elles ne sont pas identifiées et masquées avant l'analyse, elles diminuent considérablement la qualité des résultats.

Chapitre 10 — Les produits issus de la photogrammétrie

Des données multiples pour des usages différents

Après le traitement photogrammétrique des images drone avec des logiciels comme Pix4D, Agisoft Metashape ou DJI Terra, plusieurs types de produits géospatiaux sont générés.

Chaque produit correspond à une information spécifique et répond à des objectifs différents dans la chaîne de traitement.

Les principaux produits utilisés en télédétection sont :

- les orthomosaïques RGB ;
- les orthomosaïques multispectrales ;
- les bandes spectrales individuelles ;
- les modèles numériques de surface.

L'orthomosaïque RGB

L'orthomosaïque RGB est une image géoréférencée composée des trois bandes visibles : rouge, vert et bleu.

Elle est produite par assemblage des images acquises par le drone après correction des déformations géométriques, des effets de perspective et du relief.

Chaque pixel correspond ainsi à une position précise au sol.

Ces orthomosaïques sont principalement utilisées pour :

- l'interprétation visuelle ;
- la cartographie ;
- le contrôle qualité ;
- la numérisation d'objets.

Elles permettent notamment de repérer les limites des zones d'algues, les flaques d'eau ou les structures visibles du substrat.

Cependant, elles restent limitées pour les analyses de végétation car elles ne contiennent pas de bande proche infrarouge.

L'orthomosaïque multispectrale

Dans le cas d'un drone comme le DJI Mavic 3 Multispectral, la photogrammétrie peut produire une orthomosaïque multispectrale contenant plusieurs bandes spécialisées.

Ces bandes correspondent aux capteurs du drone :

- Green ;
- Red ;
- Red Edge ;
- NIR.

Elles sont principalement utilisées pour le calcul des indices de végétation tels que :

Le NDVI :

$$NDVI = \frac{NIR - Rouge}{NIR + Rouge}$$

Ou le NDRE :

$$NDRE = \frac{NIR - RedEdge}{NIR + RedEdge}$$

L'orthomosaïque multispectrale peut être fournie sous forme d'un raster multibande ou sous forme de fichiers séparés pour chaque bande.

Les bandes spectrales individuelles

Dans de nombreux logiciels photogrammétriques, les bandes multispectrales sont exportées séparément.

On obtient alors des fichiers comme :

- reflectance_green.tif ;
- reflectance_red.tif ;
- reflectance_red_edge.tif ;
- reflectance_nir.tif.

Chaque raster contient une seule bande spectrale exprimée en réflectance.

Ces données sont généralement stockées en Float32 afin de conserver la précision radiométrique nécessaire aux analyses scientifiques.

Cette organisation présente plusieurs avantages. Elle permet de manipuler les bandes indépendamment, de calculer facilement des indices dans QGIS et d'effectuer des analyses spectrales avancées.

Le modèle numérique de surface

La photogrammétrie produit également un modèle numérique de surface, souvent appelé DSM (*Digital Surface Model*).

Le DSM représente l'altitude des surfaces visibles par le drone : sol, rochers, végétation ou structures diverses.

Chaque pixel contient une valeur d'altitude.

Ces données peuvent être utilisées pour analyser :

- la micro-topographie ;
- les reliefs ;
- les chenaux ;
- les zones d'accumulation ;
- la rugosité du terrain.

Dans les milieux littoraux, ces informations peuvent aider à mieux comprendre les dynamiques hydrologiques et les processus d'accumulation des algues.

Chapitre 11 — Organiser correctement un projet QGIS

Une organisation indispensable

Dans un projet de télédétection, l'organisation des données constitue une étape essentielle.

Une structure claire facilite la compréhension du projet, améliore la reproductibilité des traitements et réduit les risques d'erreurs.

Une organisation typique peut par exemple comporter plusieurs dossiers spécialisés :

- orthomosaic ;
- multispectral ;
- DSM ;
- indices ;
- scripts.

Le dossier des orthomosaïques

Le dossier *orthomosaic* contient généralement les orthomosaïques RGB produites par le logiciel photogrammétrique.

Ces images servent principalement à la visualisation, à l'interprétation visuelle et au contrôle qualité des données.

Elles permettent de vérifier rapidement la cohérence générale du projet.

Le dossier multispectral

Le dossier *multispectral* contient les bandes utilisées pour les analyses spectrales.

Il est important de conserver ces données originales sans modification afin de préserver leur intégrité scientifique.

Ces rasters servent notamment au calcul des indices et aux classifications.

Le dossier DSM

Le dossier *DSM* regroupe les modèles numériques de surface et les produits dérivés liés au relief.

Ces données permettent d'analyser la topographie du terrain et certains paramètres géomorphologiques.

Le dossier des indices

Le dossier *indices* contient les rasters dérivés produits lors des traitements.

On y retrouve par exemple :

- les cartes de NDVI ;
- les cartes de NDRE ;
- les masques d'eau ;
- les produits de classification.

Ces données peuvent être recrées à partir des bandes originales et ne remplacent jamais les données sources.

Le dossier des scripts

Dans les projets avancés, il est souvent utile d'ajouter un dossier *scripts* contenant :

- les modèles QGIS ;
- les scripts Python ;
- les commandes GDAL ;
- les procédures automatisées.

Cette organisation améliore fortement la traçabilité et la reproductibilité des traitements scientifiques.

Chapitre 12 — Importer et vérifier les rasters dans QGIS

Charger les données dans le projet

La première étape du pré-traitement consiste à importer les rasters dans QGIS.

Cette opération s'effectue via le menu :

Layer → Add Raster Layer

Les différentes bandes multispectrales peuvent alors être chargées dans le projet.

Vérifier la cohérence géométrique des données

Vérifier le système de coordonnées

Une fois les rasters chargés, il est indispensable de vérifier leur système de coordonnées.

Dans QGIS :

Clic droit sur la couche → Properties → Information

Les données drone utilisent fréquemment des systèmes UTM comme :

EPSG:32630

Toutes les couches utilisées ensemble doivent partager un système de coordonnées cohérent. Sinon, les rasters risquent d'être mal positionnés et les analyses spatiales deviennent incorrectes.

Vérifier la résolution spatiale

La résolution correspond à la taille réelle d'un pixel au sol, également appelée GSD (*Ground Sampling Distance*).

Une résolution de :

0.0137 m

signifie qu'un pixel représente environ 1,37 cm au sol.

Toutes les bandes multispectrales doivent posséder exactement la même résolution.

Si ce n'est pas le cas, les indices spectraux deviennent incohérents car les pixels ne correspondent plus aux mêmes surfaces.

Vérifier l'étendue spatiale

Les bandes utilisées ensemble doivent également partager la même emprise spatiale.

Dans le cas contraire, certaines zones peuvent manquer dans certaines bandes et produire des erreurs lors des calculs raster.

Certaines versions de QGIS affichent parfois des informations incomplètes concernant l'étendue spatiale. Dans ces situations, les outils GDAL constituent une solution plus fiable.

Les outils suivants peuvent être utilisés :

- Informations raster ;
- gdalinfo.

La vérification visuelle

Enfin, une vérification visuelle reste indispensable.

Une méthode simple consiste à afficher alternativement différentes bandes et à vérifier si les contours des objets coïncident parfaitement.

Les rochers, les flaques d'eau ou les limites entre sable et algues doivent apparaître exactement au même endroit dans toutes les bandes.

Si ce n'est pas le cas, un problème d'alignement est probablement présent dans les données.

Chapitre 13 — Vérification des valeurs radiométriques

Pourquoi vérifier les valeurs des rasters ?

Avant d'utiliser des données multispectrales dans QGIS, il est indispensable de vérifier la nature des valeurs contenues dans les rasters.

Les bandes spectrales doivent représenter la réflectance des surfaces observées, c'est-à-dire la proportion de lumière réfléchiée pour chaque longueur d'onde.

Cette étape paraît parfois technique, mais elle conditionne directement la validité scientifique des analyses réalisées par la suite.

Des données incorrectement interprétées peuvent conduire à des indices biaisés, à des classifications erronées ou à des comparaisons temporelles impossibles.

Les plages de valeurs attendues

Deux formats couramment rencontrés

Selon le logiciel de photogrammétrie utilisé, la réflectance peut être stockée sous différentes formes numériques.

Dans certains cas, elle est exprimée directement sous forme de fraction comprise entre 0 et 1.

Une valeur de :

0.25

signifie alors que la surface réfléchit 25 % de la lumière incidente.

Dans d'autres cas, les valeurs sont simplement multipliées par 10000 afin de faciliter leur stockage sous forme entière.

Ainsi :

2500

correspond également à une réflectance de 25 %.

Ces deux formats sont scientifiquement valides, à condition d'identifier correctement l'échelle utilisée avant toute analyse.

Vérifier les valeurs dans QGIS

Les propriétés du raster

Dans QGIS, les premières vérifications peuvent être réalisées directement dans les propriétés de la couche raster.

Le menu suivant permet d'accéder aux informations radiométriques :

Clic droit sur la couche → Properties → Information

Les paramètres importants à observer sont notamment :

- les valeurs minimales ;
- les valeurs maximales ;
- la moyenne des pixels.

Ces informations donnent une première indication sur la nature des données.

L'inspection directe des pixels

Une autre méthode très utile consiste à examiner directement les valeurs des pixels sur différentes surfaces.

L'outil *Value Tool* permet par exemple de cliquer sur :

- des zones d'eau ;
- des surfaces sableuses ;
- des tapis d'algues.

Cette approche aide à mieux comprendre le comportement spectral des différentes surfaces présentes sur l'estran.

Le problème des données 8 bits

Des images adaptées à la visualisation, mais pas à l'analyse scientifique

Lorsque les valeurs observées sont comprises entre 0 et 255, cela signifie généralement que les images ont été compressées en 8 bits.

Ce type de données est fréquemment utilisé pour :

- la visualisation ;
- la cartographie ;
- les orthomosaïques RGB destinées à l'affichage.

Cependant, ces images ne sont pas adaptées aux analyses spectrales précises.

La compression en 8 bits réduit fortement la précision radiométrique et limite la capacité à détecter de faibles variations spectrales entre les surfaces.

Les conséquences de la compression

Lorsque les données sont compressées en 8 bits, plusieurs problèmes apparaissent.

Les différences fines entre surfaces deviennent moins visibles, les signatures spectrales perdent en précision et les indices de végétation peuvent être biaisés.

Dans les projets scientifiques, on privilégie donc généralement des formats comme :

- Float32 ;
- UInt16.

Ces formats conservent une précision radiométrique beaucoup plus élevée et permettent des analyses quantitatives fiables.

Chapitre 14 — Comprendre les signatures spectrales des surfaces littorales

Les bandes du drone multispectral

Le DJI Mavic 3 Multispectral enregistre quatre bandes principales :

- Green autour de 560 nm ;
- Red autour de 650 nm ;
- Red Edge autour de 730 nm ;
- NIR autour de 860 nm.

Chaque type de surface réfléchit différemment ces longueurs d'onde, ce qui permet leur identification par télédétection.

Les signatures spectrales typiques

L'eau

L'eau absorbe fortement le rayonnement dans le proche infrarouge.

Cette propriété entraîne des valeurs NIR très faibles, voire quasiment nulles.

C'est la raison pour laquelle les surfaces d'eau produisent souvent des valeurs de NDVI négatives :

$$NDVI = \frac{NIR - Rouge}{NIR + Rouge}$$

Le sable

Le sable présente généralement une réflectance relativement homogène entre les bandes visibles et le proche infrarouge.

Sa signature spectrale est souvent modérée et relativement plate.

Cependant, l'humidité du sable modifie fortement cette signature et peut compliquer l'interprétation des images.

Les algues vertes

Les algues vertes contiennent de la chlorophylle, ce qui provoque une forte absorption dans le rouge et une réflectance élevée dans le proche infrarouge.

Cette différence spectrale constitue la base de nombreux indices de végétation utilisés en télédétection.

La transition rapide entre les faibles valeurs du rouge et les fortes valeurs du NIR correspond à la zone appelée *Red Edge*.

Une astuce de visualisation multispectrale

Les compositions colorées dans QGIS

Dans QGIS, il est possible de créer des compositions colorées multispectrales afin de mieux visualiser certaines surfaces.

Une combinaison fréquemment utilisée consiste à afficher :

- le NIR dans le canal rouge ;
- la bande rouge dans le canal vert ;
- la bande verte dans le canal bleu.

Dans cette représentation, la végétation et les algues apparaissent souvent en rouge vif.

Cette méthode facilite l'identification visuelle des zones biologiquement actives.

Chapitre 15 — Utilisation du modèle numérique de surface (DSM)

Une information complémentaire aux données spectrales

En plus des bandes multispectrales, les traitements photogrammétriques produisent également un modèle numérique de surface, ou DSM.

Ce raster représente l'altitude des surfaces visibles par le drone :

- le sol ;
- les rochers ;
- les structures du substrat ;
- la végétation.

Chaque pixel contient donc une information topographique.

L'intérêt du DSM pour l'analyse de l'estran

Comprendre les micro-structures du terrain

Dans les milieux littoraux, le DSM permet de détecter des structures parfois difficiles à percevoir sur les orthomosaïques classiques.

Il peut révéler :

- des chenaux de drainage ;
- des rides sableuses ;
- de petites dépressions ;
- des zones d'accumulation d'algues.

Ces structures influencent fortement :

- la circulation de l'eau ;
 - l'humidité du substrat ;
 - la répartition des algues sur l'estran.
-

Visualiser le relief dans QGIS

Le calcul d'ombrage du relief

Pour améliorer la lecture du relief, il est possible de calculer un ombrage du terrain, appelé *hillshade*.

Dans QGIS, cet outil est accessible via :

Raster → Analyse du terrain → Ombrage

Le principe consiste à simuler un éclairage rasant afin de faire ressortir les variations de relief.

Cette visualisation met en évidence :

- les micro-reliefs ;
- les structures du substrat ;
- les zones de dépression ;
- les chenaux.

Une approche multi-source de l'interprétation

Combiner données spectrales et topographiques

Dans certaines situations, les accumulations d'algues sont directement liées à la topographie locale.

Une dépression du terrain ou un chenal de drainage peut favoriser l'accumulation d'eau et modifier la distribution des surfaces algales.

La combinaison des informations :

- spectrales ;
- topographiques ;

permet alors une interprétation beaucoup plus complète de l'environnement littoral.

Cette approche multi-source constitue aujourd'hui une pratique essentielle dans les analyses environnementales avancées.

Chapitre 16 — Comprendre les histogrammes des rasters

Pourquoi analyser les histogrammes ?

Lorsqu'on travaille avec des images multispectrales dans QGIS, l'analyse visuelle de l'image ne suffit pas pour évaluer la qualité des données.

Deux images peuvent sembler très similaires à l'écran tout en présentant des comportements radiométriques très différents.

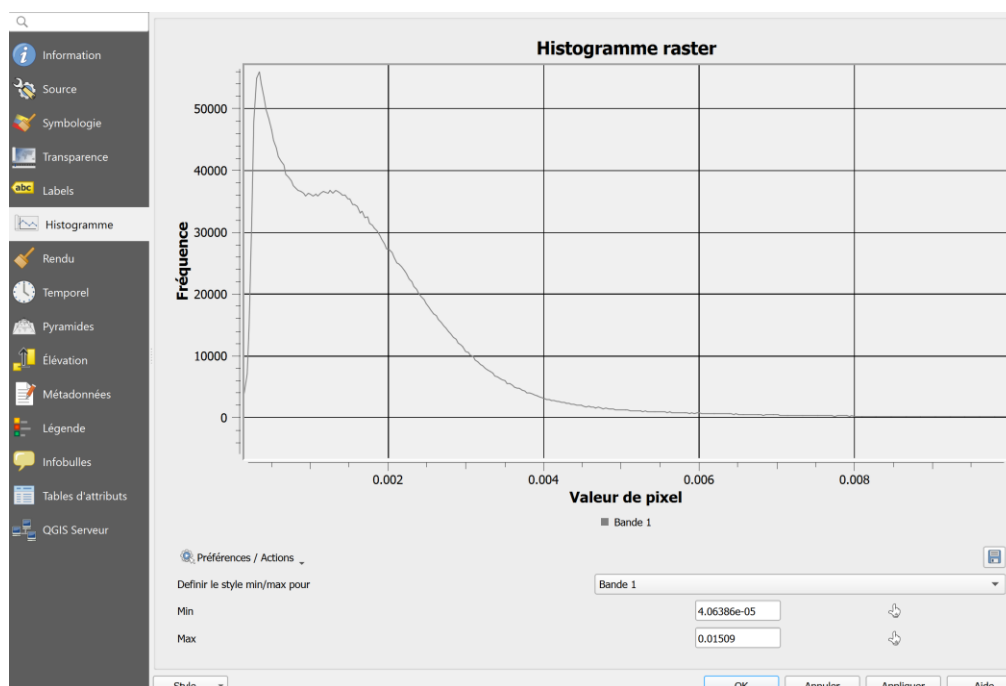
C'est pourquoi l'analyse des histogrammes constitue une étape essentielle des pré-traitements géomatiques.

Un histogramme représente la distribution statistique des valeurs des pixels contenus dans un raster. Chaque barre correspond à une plage de valeurs ainsi qu'au nombre de pixels possédant ces valeurs.

Dans QGIS, l'histogramme d'une bande peut être affiché via :

Layer Properties → Histogram

Cet outil fournit une vue globale du comportement radiométrique de l'image et permet de détecter plusieurs types de problèmes susceptibles d'affecter les analyses.



Histogramme raster de QGIS

Comprendre la forme d'un histogramme

Une représentation de la distribution des pixels

Dans une image multispectrale, chaque pixel possède une valeur numérique correspondant à la réflectance mesurée par le capteur.

L'histogramme représente la manière dont ces valeurs sont réparties dans l'ensemble de l'image.

Une distribution étalée indique généralement une bonne diversité radiométrique, tandis qu'une distribution très resserrée peut révéler un manque de contraste ou un problème de calibration.

L'histogramme constitue ainsi une sorte de “signature statistique” du raster.

Détecter les phénomènes de saturation

Quand le capteur atteint ses limites

La saturation apparaît lorsqu'un grand nombre de pixels atteignent les valeurs maximales ou minimales enregistrables par le capteur.

Dans un histogramme, ce phénomène se traduit souvent par un pic très marqué aux extrémités de la distribution.

Ces saturations peuvent être provoquées par :

- des surfaces extrêmement réfléchissantes ;
- un problème d'exposition du capteur ;
- des conditions lumineuses excessives ;
- des zones très sombres dépourvues d'information utile.

Lorsqu'une zone est saturée, une partie de l'information radiométrique disparaît définitivement. Le capteur n'est alors plus capable de distinguer les variations réelles entre surfaces.

Ce problème peut fortement perturber les calculs d'indices spectraux et les classifications automatiques.

Identifier les valeurs aberrantes

Détecter les anomalies radiométriques

Les valeurs aberrantes correspondent à des pixels dont les valeurs diffèrent fortement du reste de l'image.

Dans l'histogramme, elles apparaissent souvent sous forme :

- de petites barres isolées ;
- de valeurs très éloignées du cœur principal de la distribution.

Ces anomalies peuvent avoir plusieurs origines :

- bruit électronique du capteur ;
- erreurs de traitement ;
- mauvaise gestion des pixels NoData ;
- artefacts photogrammétriques.

Même peu nombreuses, ces valeurs peuvent perturber certains traitements statistiques ou modifier les plages d'affichage des rasters.

L'analyse de l'histogramme permet donc d'identifier rapidement ce type de problème avant de commencer les analyses.

Détecter des problèmes de calibration radiométrique

Une image visuellement correcte peut être scientifiquement incohérente

L'histogramme constitue également un outil précieux pour évaluer la cohérence radiométrique des données.

Certaines anomalies apparaissent lorsque :

- les valeurs sont globalement très faibles ;
- la distribution est excessivement compressée ;
- les contrastes radiométriques sont anormalement réduits.

Ces situations peuvent révéler :

- une mauvaise calibration radiométrique ;
- une conversion incorrecte en réflectance ;
- un problème d'exposition lors de l'acquisition.

Il est important de comprendre qu'une image peut sembler correcte à l'écran tout en présentant une distribution radiométrique incohérente pour une analyse scientifique.

Les outils de visualisation ajustent souvent automatiquement les contrastes, ce qui peut masquer certains problèmes présents dans les données brutes.

Interpréter les histogrammes des images littorales

Une distribution liée à la diversité des surfaces

Dans les milieux littoraux, les histogrammes reflètent souvent la grande diversité des surfaces présentes sur l'estran.

Les différentes catégories de surfaces possèdent des comportements radiométriques distincts :

- l'eau présente généralement des valeurs très faibles ;
- le sable produit des valeurs intermédiaires ;
- les algues présentent souvent des valeurs plus élevées dans certaines bandes comme le proche infrarouge.

Cette diversité peut produire plusieurs zones de concentration dans l'histogramme, correspondant aux différents types de surfaces présents dans l'image.

L'interprétation de ces distributions aide à mieux comprendre le comportement spectral global du site étudié.

Les histogrammes comme outil de diagnostic

Une étape essentielle avant toute analyse

L'analyse des histogrammes constitue un outil de diagnostic simple mais extrêmement puissant dans les projets de télédétection.

Elle permet :

- d'évaluer rapidement la qualité radiométrique des données ;
- de détecter des anomalies ;
- de vérifier la cohérence des bandes ;
- d'identifier des problèmes de saturation ou de calibration.

Cette étape doit être considérée comme une vérification systématique avant tout calcul d'indice ou toute classification.

En télédétection scientifique, comprendre les données avant de lancer les traitements est souvent aussi important que les traitements eux-mêmes.

Chapitre 17 — Inspection des pixels et compréhension des signatures spectrales

Observer directement les valeurs des pixels

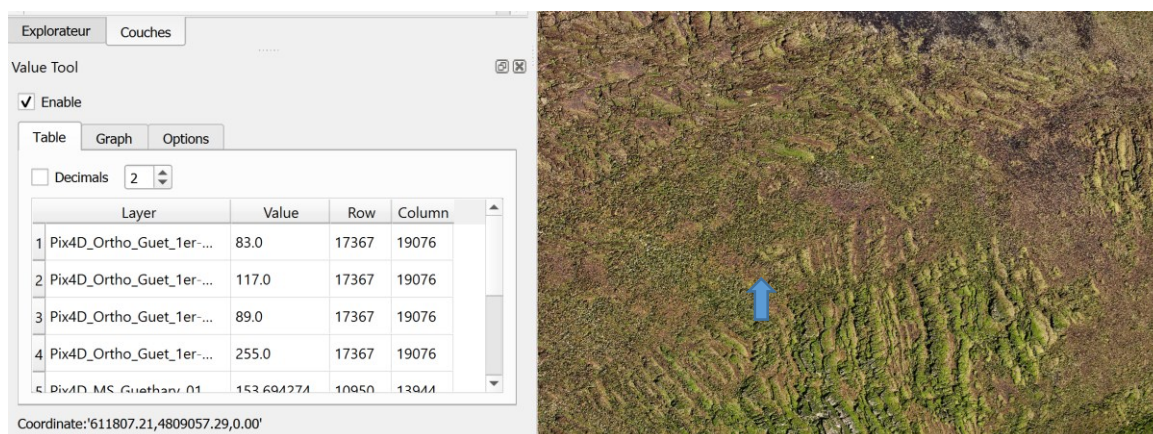
Une étape importante des pré-traitements consiste à inspecter directement les valeurs des pixels dans les différentes bandes spectrales.

Cette opération permet de mieux comprendre le comportement spectral des surfaces présentes dans l'image et de vérifier rapidement la cohérence des données.

Dans QGIS, cette analyse peut être réalisée avec l'outil :

Value Tool

Cet outil affiche la valeur du pixel situé sous le curseur pour chaque bande raster chargée dans le projet.



Utilisation de l'outil Value Tool

Explorer les surfaces de l'image

Une fois l'outil activé, il suffit de déplacer le curseur sur l'image et de cliquer sur différentes surfaces afin d'observer les valeurs spectrales correspondantes.

La fenêtre de l'outil indique alors :

- la couche raster concernée ;
- la valeur du pixel ;
- la bande observée.

Cette approche constitue une méthode simple mais extrêmement utile pour interpréter les données multispectrales.

Comparer différentes surfaces littorales

L'eau

Lorsque l'on inspecte une zone d'eau, les valeurs observées sont généralement très faibles dans toutes les bandes.

La réflectance devient particulièrement faible dans le proche infrarouge, car l'eau absorbe fortement cette longueur d'onde.

Dans certaines situations, les valeurs NIR deviennent presque nulles.

Cette propriété explique pourquoi les surfaces d'eau produisent souvent des valeurs de NDVI négatives .

Le sable

Le sable présente généralement des valeurs modérées dans toutes les bandes spectrales.

Contrairement à la végétation ou aux algues, il ne montre pas de différence très marquée entre le rouge et le proche infrarouge.

Sa signature spectrale apparaît donc relativement homogène et assez plate.

Cependant, l'humidité du substrat peut modifier fortement ce comportement radiométrique.

Les algues vertes

Les algues possèdent un comportement spectral proche de celui de la végétation terrestre en raison de leur teneur en chlorophylle.

On observe généralement :

- une absorption importante dans la bande rouge ;
- une forte réflectance dans le proche infrarouge.

Cette relation se traduit souvent par :

$NIR > Red$

Cette différence constitue la base des indices de végétation utilisés en télédétection.

Pourquoi cette inspection est-elle importante ?

Comprendre les données avant les traitements

L'inspection directe des pixels permet :

- de vérifier la cohérence des valeurs ;
- de comprendre les signatures spectrales ;
- d'identifier les différentes surfaces présentes ;
- de détecter rapidement certaines anomalies.

Cette étape aide également à repérer :

- des erreurs de calibration ;
- des valeurs incohérentes ;
- des problèmes radiométriques.

Avant de lancer des classifications ou des calculs d'indices, il est toujours utile de comprendre concrètement ce que contiennent réellement les rasters.

Les signatures spectrales typiques des algues

Le comportement des différentes bandes

Les algues vertes présentent généralement :

- une réflectance moyenne dans le vert ;
- une absorption dans le rouge ;
- une transition marquée dans le Red Edge ;
- une forte réflectance dans le proche infrarouge.

Cette signature résulte directement de l'activité chlorophyllienne.

Le DJI Mavic 3 Multispectral enregistre précisément les bandes nécessaires pour observer ce comportement spectral.

Chapitre 18 — L'importance de l'alignement des bandes

Une condition essentielle des analyses multispectrales

Les analyses multispectrales reposent sur une hypothèse fondamentale :

un même pixel doit représenter exactement la même surface dans toutes les bandes spectrales.

Autrement dit, lorsqu'un indice combine plusieurs bandes, les valeurs utilisées doivent correspondre au même point au sol.

Cette condition est indispensable pour obtenir des résultats cohérents.

Pourquoi cette hypothèse est-elle si importante ?

Les indices reposent sur plusieurs bandes

Les indices de végétation combinent plusieurs bandes spectrales.

Le NDVI repose par exemple sur la relation suivante :

$$NDVI = \frac{NIR - Rouge}{NIR + Rouge}$$

Dans ce calcul, les valeurs NIR et Rouge doivent provenir exactement du même pixel au sol.

Si les bandes sont décalées, le calcul mélange des informations provenant de surfaces différentes.

Les particularités des capteurs multispectraux

Plusieurs capteurs pour plusieurs bandes

Dans les systèmes multispectraux, les différentes bandes ne sont généralement pas enregistrées par un seul capteur.

Chaque bande provient souvent :

- d'une caméra différente ;
- ou d'un capteur indépendant.

Dans le cas du DJI Mavic 3 Multispectral, plusieurs capteurs spécialisés enregistrent simultanément les différentes longueurs d'onde.

Même si ces systèmes sont calibrés ensemble, de légers décalages peuvent apparaître entre les bandes.

Les origines possibles des décalages

Des erreurs parfois très faibles mais importantes

Plusieurs phénomènes peuvent provoquer un décalage entre bandes :

- la position légèrement différente des capteurs ;
- les erreurs d'alignement photogrammétrique ;
- les différences de perspective ;
- les interpolations liées au reprojetage.

Même un décalage très faible peut perturber les analyses multispectrales.

Les conséquences d'un mauvais alignement

Des indices incorrects et du bruit spatial

Lorsque les bandes ne sont pas parfaitement alignées :

- les contours des objets deviennent instables ;
- les indices présentent du bruit spatial ;
- certaines zones produisent des valeurs aberrantes.

Par exemple, si un pixel correspond à du sable dans la bande rouge mais à une algue dans la bande NIR, le calcul du NDVI produira une valeur artificiellement élevée.

Ce type d'erreur peut entraîner des problèmes dans :

- la détection des algues ;
 - la classification des surfaces ;
 - l'interprétation des cartes.
-

Une sensibilité encore plus forte avec les drones

L'effet de la très haute résolution

Les drones produisent des images à très haute résolution spatiale.

Une résolution de :

5 cm par pixel

signifie qu'un décalage d'un seul pixel représente déjà plusieurs centimètres au sol.

Ces petits décalages deviennent alors visibles dans les indices calculés et peuvent produire des artefacts importants.

Les analyses haute résolution sont donc particulièrement sensibles à la qualité de l'alignement géométrique.

Vérifier l'alignement dans QGIS

Une vérification visuelle indispensable

Dans QGIS, l'alignement peut être vérifié visuellement en affichant différentes bandes et en comparant les contours des objets.

Les limites suivantes sont particulièrement utiles :

- les frontières sable/algues ;
- les bords de rochers ;
- les contours des flaques d'eau.

Si ces éléments coïncident parfaitement dans toutes les bandes, l'alignement est probablement correct.

Chapitre 19 — Alignement raster avec GDAL

Harmoniser les rasters avant les analyses

Pour garantir la cohérence géométrique des bandes multispectrales, il est souvent nécessaire de réaligner les rasters sur une même grille de pixels.

Dans QGIS, cette opération peut être réalisée avec l'outil :

GDAL → Align Raster

Cet outil permet d'harmoniser :

- l'emprise spatiale ;
- la résolution ;
- la grille des pixels.

Pourquoi cet alignement est-il nécessaire ?

Des différences parfois invisibles mais problématiques

Même lorsque les rasters semblent parfaitement superposés à l'écran, ils peuvent présenter :

- des résolutions légèrement différentes ;
- des origines de grille différentes ;
- de faibles décalages géométriques.

Ces différences suffisent à perturber les calculs raster.

L'alignement permet de garantir que chaque pixel correspond exactement à la même position au sol dans toutes les bandes.

Fonctionnement de l'outil Align Raster

Utiliser un raster de référence

L'outil fonctionne en utilisant un raster de référence qui définit :

- la résolution ;
- la taille des pixels ;
- l'emprise spatiale ;

- la grille géométrique.

Les autres rasters sont ensuite recalculés pour correspondre exactement à ces paramètres.

Dans QGIS, l'outil est accessible via :

Processing → Toolbox → GDAL → Align Raster

Harmoniser la résolution spatiale

Une étape essentielle avant les indices

L'outil permet également d'harmoniser les tailles de pixels.

Par exemple, deux rasters peuvent présenter des résolutions légèrement différentes :

- 0.05 m ;
- 0.049 m.

Même si cette différence semble minime, elle peut perturber les calculs multispectraux.

L'alignement raster permet alors d'obtenir une résolution commune cohérente pour toutes les bandes.

Cette étape est indispensable avant :

- le calcul d'indices ;
- les classifications ;
- les analyses multispectrales avancées.

Résultat final

Des rasters parfaitement comparables

Après l'alignement :

- les pixels correspondent exactement aux mêmes positions ;
- les rasters deviennent directement comparables ;
- les calculs raster gagnent en cohérence géométrique.

Cette étape constitue l'un des fondements des pré-traitements multispectraux sous QGIS.

Chapitre 20 — Création des indices spectraux

Transformer les bandes spectrales en indicateurs interprétables

Les images multispectrales contiennent une grande quantité d'informations radiométriques. Cependant, l'interprétation directe des différentes bandes reste souvent complexe.

Pour faciliter l'analyse, la télédétection utilise des indices spectraux, c'est-à-dire des combinaisons mathématiques entre plusieurs bandes spectrales.

Ces indices permettent de mettre en évidence certaines propriétés des surfaces observées, notamment :

- la présence de végétation ;
- l'humidité ;
- la chlorophylle ;
- la biomasse ;
- la présence d'eau.

Ils constituent aujourd'hui l'un des outils fondamentaux de l'analyse multispectrale.

Le NDVI : l'indice de végétation le plus utilisé

Principe du NDVI

Le NDVI (*Normalized Difference Vegetation Index*) est probablement l'indice spectral le plus connu en télédétection.

Sa formule repose sur la différence entre le proche infrarouge et le rouge :

$$NDVI = \frac{NIR - Rouge}{NIR + Rouge}$$

Rouge correspond à la bande rouge.

Le principe physique du NDVI

Le NDVI exploite le comportement spectral caractéristique de la végétation et des algues.

Les surfaces riches en chlorophylle absorbent fortement le rayonnement dans la bande rouge, tandis qu'elles réfléchissent fortement le proche infrarouge.

Cette différence crée un contraste spectral important entre les deux bandes.

Plus ce contraste est marqué, plus la valeur du NDVI devient élevée.

Interprétation des valeurs

Le NDVI varie généralement entre :

-1 et +1

Les surfaces d'eau produisent souvent des valeurs négatives en raison de leur très faible réflectance dans le NIR.

Les surfaces sableuses donnent fréquemment des valeurs proches de zéro.

Les algues et les végétations produisent généralement des valeurs positives, tandis que les végétations très denses atteignent les valeurs les plus élevées.

Dans les milieux littoraux, cet indice permet souvent de distinguer :

- l'eau ;
- le sable ;
- les zones couvertes par des algues.

Calcul du NDVI dans QGIS

Dans QGIS, le calcul peut être réalisé avec la calculatrice raster.

Le menu utilisé est :

Raster → Raster Calculator

L'expression utilisée est :

```
("reflectance_nir" - "reflectance_red") /  
("reflectance_nir" + "reflectance_red")
```

Le résultat est un nouveau raster contenant la valeur de NDVI pour chaque pixel.

Les limites du NDVI

Le NDVI constitue un indice robuste et simple à interpréter, mais il présente certaines limites en milieu littoral.

Certaines surfaces humides ou certains biofilms peuvent produire des valeurs proches de celles des algues.

Le substrat humide peut également générer des faux positifs.

L'interprétation du NDVI doit donc toujours être réalisée avec prudence et replacée dans son contexte environnemental.

Le NDRE : exploiter la bande Red Edge

Un indice plus sensible à la chlorophylle

Le NDRE (*Normalized Difference Red Edge*) utilise la bande Red Edge à la place de la bande rouge.

Sa formule est :

$$NDRE = \frac{NIR - RedEdge}{NIR + RedEdge}$$

La bande Red Edge correspond à la zone de transition entre :

- l'absorption dans le rouge ;
- la forte réflectance du proche infrarouge.

Cette région spectrale est particulièrement sensible à la présence de chlorophylle.

Différences entre NDVI et NDRE

Le NDVI est principalement utilisé pour détecter la végétation de manière générale.

Le NDRE est souvent plus performant pour :

- les végétations denses ;
- les variations subtiles de biomasse ;
- l'analyse de la chlorophylle ;
- le suivi du stress végétal.

Dans les milieux littoraux, le NDRE peut parfois mieux distinguer :

- les tapis d'algues ;
- les substrats minéraux ;
- certaines zones partiellement immergées.

Il peut également être moins sensible à certains effets de saturation observés avec le NDVI.

Calcul du NDRE dans QGIS

Dans la calculatrice raster :

```
("reflectance_nir" - "reflectance_red_edge") /  
("reflectance_nir" + "reflectance_red_edge")
```

Le raster produit peut être enregistré sous le nom :

ndre.tif

Comparer visuellement le NDVI et le NDRE permet souvent d'obtenir une interprétation plus robuste des surfaces observées.

Le GNDVI : utiliser la bande verte

Une variante sensible à la chlorophylle

Le GNDVI (*Green NDVI*) repose sur la bande verte au lieu de la bande rouge.

Sa formule est :

$$GNDVI = \frac{NIR - Green}{NIR + Green}$$

Cet indice reste basé sur le contraste entre le proche infrarouge et le visible, mais il exploite ici la sensibilité de la bande verte aux pigments photosynthétiques.

Intérêt du GNDVI

Le GNDVI est souvent plus sensible :

- à la concentration en chlorophylle ;
- à certaines variations physiologiques ;
- à la densité de végétation.

Dans certaines situations où le NDVI tend à saturer, le GNDVI peut offrir une meilleure discrimination des différences de biomasse.

En milieu littoral, il peut parfois mieux détecter certains tapis d'algues denses.

Cependant, il reste sensible :

- à l'humidité ;
- à la turbidité ;
- à la présence d'eau.

Calcul du GNDVI

Dans QGIS :

```
("reflectance_nir" - "reflectance_green") /  
("reflectance_nir" + "reflectance_green")
```

Le raster obtenu représente les variations spatiales du GNDVI.

Le NDWI : détecter l'eau

Un indice dédié aux surfaces humides

Le NDWI (*Normalized Difference Water Index*) est conçu pour mettre en évidence la présence d'eau.

Il exploite le contraste spectral entre :

- la bande verte ;
- le proche infrarouge.

Sa formule est :

$$NDWI = \frac{Green - NIR}{Green + NIR}$$

Le comportement spectral de l'eau

L'eau réfléchit modérément la lumière dans le visible mais absorbe fortement le proche infrarouge.

Ainsi, dans les zones d'eau :

Green > NIR

Le NDWI devient alors positif.

À l'inverse, les surfaces terrestres ou végétalisées produisent souvent des valeurs faibles ou négatives.

Utilité du NDWI sur l'estran

Dans les milieux intertidaux, le NDWI permet de détecter :

- les flaques d'eau ;
- les chenaux ;
- les zones saturées ;
- les surfaces partiellement immergées.

Ces surfaces perturbent souvent les analyses des algues.

Le NDWI est donc fréquemment utilisé pour produire un masque eau avant les classifications.

Calcul du NDWI

Dans la calculatrice raster :

```
("reflectance_green" - "reflectance_nir") /  
("reflectance_green" + "reflectance_nir")
```

Le raster peut être enregistré sous le nom :

ndwi.tif

Le Simple Ratio (SR)

Un indice historique de végétation

Le Simple Ratio constitue l'un des premiers indices utilisés en télédétection.

Sa formule est extrêmement simple :

$$SR = \frac{NIR}{Red}$$

Cet indice exploite directement le contraste entre :

- la forte réflectance du NIR ;
 - l'absorption dans le rouge.
-

Caractéristiques du Simple Ratio

Contrairement au NDVI, le Simple Ratio n'est pas normalisé.

Ses valeurs ne sont donc pas limitées entre -1 et +1.

Il peut être très sensible dans les zones à forte biomasse, mais également plus sensible au bruit radiométrique.

Le NDVI peut d'ailleurs être considéré comme une version normalisée du Simple Ratio.

Le Chlorophyll Index (CI)

Estimer la chlorophylle

Le Chlorophyll Index est conçu pour estimer la concentration en chlorophylle.

Une formule couramment utilisée est :

$$CI = \frac{NIR}{RedEdge} - 1$$

Cet indice exploite la forte sensibilité de la bande Red Edge à la chlorophylle.

Intérêt du CI pour les algues

Comme les algues contiennent elles aussi de la chlorophylle, cet indice peut être utilisé pour :

- détecter les tapis algaux ;
- estimer leur densité ;
- analyser certaines variations physiologiques.

Dans certaines situations, le CI peut être plus sensible que le NDVI pour détecter des variations de biomasse.

Comparer plusieurs indices

Aucun indice n'est parfait

Chaque indice spectral possède ses avantages et ses limites.

Le **NDVI** est robuste mais sensible au substrat humide.

Le **NDRE** est souvent plus performant pour analyser la chlorophylle.

Le **GNDVI** peut mieux détecter certaines végétations denses.

Le **NDWI** permet de détecter l'eau mais n'est pas spécifique aux algues.

Comparer plusieurs indices permet généralement d'obtenir une interprétation plus robuste et plus fiable des surfaces observées.

Une interprétation toujours critique

Les indices ne sont pas des mesures directes

Les indices spectraux constituent des indicateurs indirects construits à partir du comportement radiométrique des surfaces.

Ils ne représentent jamais une mesure directe des algues elles-mêmes.

Leur interprétation doit toujours prendre en compte :

- le contexte écologique ;
- les conditions d'acquisition ;
- l'humidité du substrat ;
- les limites des capteurs.

Cette approche critique reste essentielle dans toute démarche scientifique de télédétection environnementale.

Chapitre 21 — Création de masques raster

Pourquoi utiliser des masques ?

Dans les analyses multispectrales, toutes les surfaces présentes dans une image ne sont pas forcément utiles pour l'étude.

Certaines zones peuvent même perturber fortement les traitements et introduire du bruit dans les résultats.

C'est notamment le cas :

- des surfaces en eau ;
- des zones saturées ;
- des ombres ;
- des substrats très humides.

Ces surfaces produisent souvent des signatures spectrales atypiques susceptibles de perturber :

- les indices de végétation ;

- les classifications automatiques ;
- les analyses statistiques.

Pour limiter ces problèmes, les pré-traitements utilisent des masques raster.

Le principe d'un masque raster

Isoler les zones utiles

Un masque est généralement un raster binaire permettant de distinguer :

- les pixels à conserver ;
- les pixels à exclure.

Dans sa forme la plus simple :

- la valeur 1 représente les zones conservées ;
- la valeur 0 représente les zones exclues.

Le masque agit donc comme un filtre appliqué aux données raster avant l'analyse.

Cette étape permet de nettoyer les données et d'améliorer la robustesse des traitements.

Le masque d'eau

Pourquoi l'eau pose problème

L'eau constitue l'une des principales sources de perturbation en milieu intertidal.

Elle absorbe fortement le rayonnement dans le proche infrarouge, ce qui modifie fortement les valeurs des indices de végétation.

Les zones d'eau peuvent ainsi produire :

- des valeurs de NDVI très faibles ;
- des signatures ambiguës ;
- des erreurs de classification.

Pour cette raison, il est souvent utile de créer un masque spécifique permettant d'exclure les surfaces immergées avant les analyses.

Détection de l'eau avec le NDWI

Utiliser le contraste Green / NIR

Le NDWI (*Normalized Difference Water Index*) constitue un outil efficace pour détecter les surfaces en eau.

Sa formule est :

$$NDWI = \frac{Green - NIR}{Green + NIR}$$

L'eau réfléchit relativement peu le proche infrarouge, ce qui produit généralement des valeurs positives de NDWI.

Un seuil simple peut alors être appliqué :

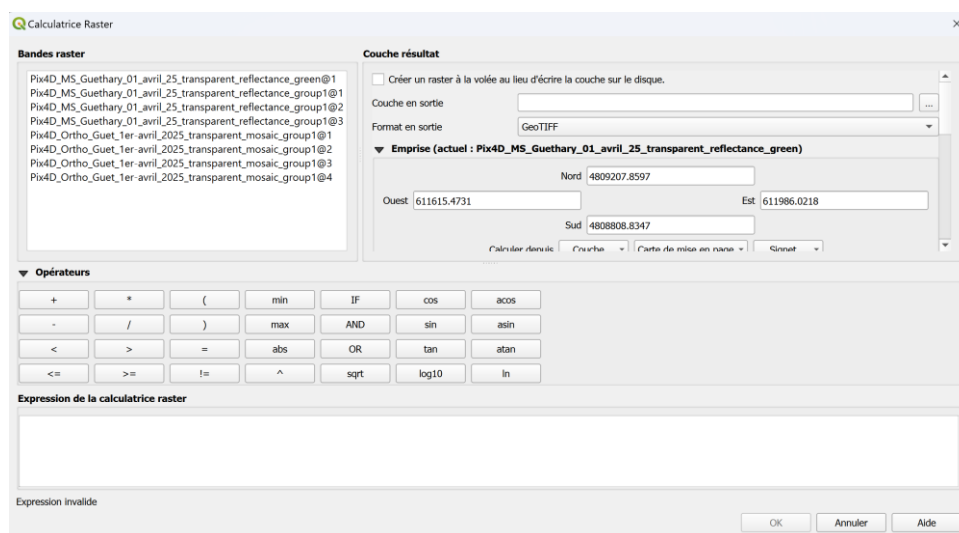
$$NDWI > 0$$

Les pixels correspondant à l'eau sont ainsi isolés.

Création du masque dans QGIS

Utiliser la Raster Calculator

Dans QGIS, le masque peut être créé avec la calculatrice raster.



Calculatrice raster de QGIS

L'expression suivante permet par exemple de détecter l'eau :

```
("ndwi" > 0)
```

Le raster produit contient :

- la valeur 1 pour l'eau ;
- la valeur 0 pour les autres surfaces.
-

Application du masque

Exclure les pixels perturbateurs

Le masque peut ensuite être utilisé pour exclure certaines zones lors du calcul des indices.

Par exemple :

$NDVI * (mask = 0)$

Cette opération permet de conserver uniquement les pixels considérés comme exploitables.

Les masques constituent ainsi une étape essentielle du nettoyage des données raster.

Les principaux types de masques

Des masques adaptés aux différentes perturbations

Plusieurs types de masques sont couramment utilisés dans les projets de télédétection littorale.

Les plus fréquents sont :

- les masques d'eau ;
- les masques d'ombre ;
- les masques de saturation ;
- les masques de zone d'étude.

Chaque masque répond à un objectif spécifique :

- supprimer les zones immergées ;
- éliminer les ombres ;
- exclure les pixels extrêmes ;
- limiter l'analyse à certaines zones.

Un bon pré-traitement inclut presque toujours plusieurs étapes de masquage.

Chapitre 22 — Le masque des ombres

Pourquoi les ombres perturbent les analyses

Les zones d'ombre constituent une autre difficulté importante en télédétection.

Dans les zones ombragées, les surfaces reçoivent moins de lumière. Les valeurs enregistrées deviennent faibles dans toutes les bandes spectrales.

Cette diminution d'illumination peut produire :

- des valeurs de NDVI faibles ou négatives ;
- des confusions avec l'eau ;
- des erreurs dans les classifications automatiques.

Les ombres ne représentent pourtant pas une surface écologique différente. Elles correspondent simplement à une condition d'éclairage particulière.

Détection simple des ombres

Utiliser un seuil sur le NIR

Dans une zone d'ombre, les valeurs restent faibles dans toutes les bandes.

Une méthode simple consiste donc à utiliser un seuil sur la bande NIR.

Par exemple :

```
"reflectance_nir" < 0.02
```

Le raster obtenu contient :

- la valeur 1 pour les ombres ;
- la valeur 0 pour les autres surfaces.

Une approche plus robuste

Utiliser plusieurs bandes

Une méthode plus robuste consiste à utiliser la moyenne ou la somme de plusieurs bandes spectrales.

Par exemple :

$(\text{Green} + \text{Red} + \text{NIR}) / 3 < \text{seuil}$

Cette approche permet de mieux distinguer :

- les ombres ;
- l'eau ;
- les surfaces naturellement sombres.

Le masque d'ombre peut ensuite être appliqué aux indices spectraux afin d'exclure les zones problématiques.

Les limites du masque d'ombre

Des confusions possibles avec l'eau

Les ombres peuvent parfois être confondues avec :

- des surfaces en eau ;
- des substrats très humides ;
- certaines surfaces sombres.

Pour cette raison, il est souvent nécessaire de combiner plusieurs approches de détection.

Un masque basé uniquement sur un seuil radiométrique reste souvent insuffisant en milieu littoral.

Chapitre 23 — Distinguer l'ombre et l'eau

Deux surfaces sombres mais des signatures différentes

L'eau et les ombres présentent un problème similaire : elles produisent toutes deux des valeurs faibles dans plusieurs bandes spectrales.

Cependant, leur comportement spectral reste différent.

L'eau conserve généralement un certain signal dans la bande verte, alors que les ombres présentent des valeurs faibles dans toutes les bandes.

Cette différence peut être exploitée pour améliorer la qualité des masques.

Une approche combinée recommandée

Combiner NDWI et NIR

Une méthode robuste consiste à utiliser plusieurs critères simultanément.

La première étape consiste à calculer le NDWI :

$$NDWI = \frac{Green - NIR}{Green + NIR}$$

Ensuite, deux masques distincts peuvent être créés.

Le masque d'eau repose sur :

$$NDWI > 0$$

Le masque d'ombre peut être défini par :

$$("ndwi" < 0) \text{ AND } ("reflectance_nir" < 0.02)$$

Cette combinaison permet de distinguer :

- l'eau ;
- les ombres ;
- les surfaces végétalisées.

Pourquoi cette méthode fonctionne-t-elle ?

Différence entre structure spectrale et absence de lumière

L'eau possède une signature spectrale propre liée à ses propriétés physiques.

Même si le NIR est fortement absorbé, la bande verte conserve souvent un certain niveau de signal.

L'ombre, au contraire, correspond principalement à une diminution générale de l'éclairage. Toutes les bandes deviennent faibles simultanément.

Autrement dit :

- l'eau conserve une structure spectrale identifiable ;
- l'ombre "écrase" l'ensemble du signal spectral.

Cette différence explique l'intérêt d'une approche multi-indices.

Combiner plusieurs masques

Conserver uniquement les zones exploitables

Dans la pratique, plusieurs masques sont souvent combinés.

Par exemple :

```
(mask_eau == 0) AND (mask_ombre == 0)
```

Cette approche permet de conserver uniquement les zones considérées comme fiables pour l'analyse.

Les pixels perturbateurs sont exclus avant les calculs d'indices ou les classifications.

Une étape essentielle des pré-traitements

Nettoyer les données avant les analyses

Sur l'estran, la confusion entre eau, ombre et substrat humide constitue l'une des erreurs les plus fréquentes.

L'utilisation d'un seuil unique est généralement insuffisante pour traiter correctement ces situations complexes.

Les approches combinant plusieurs indices et plusieurs masques offrent des résultats beaucoup plus robustes.

Le masquage doit donc être considéré comme une étape fondamentale des pré-traitements géomatiques en télédétection littorale.

Chapitre 24 — Inspection des signatures spectrales avec SCP

Comprendre le concept de signature spectrale

Toute l'analyse multispectrale repose sur une idée fondamentale : chaque surface possède un comportement spectral particulier.

Une signature spectrale correspond à la variation de la réflectance d'une surface selon les différentes bandes spectrales observées par le capteur.

Autrement dit, chaque matériau réfléchit la lumière de manière différente selon les longueurs d'onde.

L'eau, le sable, les algues ou les rochers possèdent ainsi des signatures spectrales distinctes qui permettent leur différenciation dans les images multispectrales.

Comprendre ces signatures constitue une étape essentielle avant toute classification ou interprétation des données.

Le plugin SCP dans QGIS

Un outil dédié à l'analyse multispectrale

Dans QGIS, l'analyse des signatures spectrales peut être réalisée avec le plugin :

Semi-Automatic Classification Plugin

Ce plugin, souvent appelé SCP, constitue l'un des outils les plus utilisés pour l'analyse d'images satellitaires et multispectrales dans QGIS.

Il permet notamment :

- de créer des zones d'intérêt ;
- d'extraire des signatures spectrales ;
- de visualiser les courbes spectrales ;
- de préparer des classifications supervisées.

SCP joue ainsi un rôle central dans les workflows de télédétection sous QGIS.

Création des zones d'intérêt (ROI)

Définir des surfaces représentatives

La première étape consiste à créer des ROI (*Regions of Interest*), c'est-à-dire des zones représentatives des différentes classes présentes dans l'image.

Ces classes peuvent par exemple correspondre à :

- de l'eau ;
- du sable ;
- des tapis d'algues ;
- des surfaces rocheuses.

Dans SCP, l'utilisateur dessine directement des polygones sur l'image afin de sélectionner des surfaces homogènes.

Chaque ROI est ensuite associée à une classe spécifique.

L'importance du choix des ROI

La qualité des signatures spectrales dépend directement de la qualité des ROI sélectionnées.

Les zones choisies doivent être :

- homogènes ;
- représentatives ;
- dépourvues de mélange entre plusieurs surfaces.

Une ROI contenant à la fois du sable humide et des algues produira une signature spectrale ambiguë qui compliquera les analyses ultérieures.

Le choix des ROI constitue donc une étape critique dans tout processus de classification supervisée.

Extraction des signatures spectrales

Transformer les pixels en courbes spectrales

Une fois les ROI créées, SCP permet d'extraire automatiquement les valeurs spectrales correspondantes.

Pour chaque classe, le plugin calcule généralement :

- une valeur moyenne par bande ;
- une variance ;
- une signature spectrale complète.

Ces informations décrivent le comportement radiométrique moyen des surfaces sélectionnées.

L'analyse ne porte donc plus directement sur l'image visible, mais sur les propriétés spectrales des différentes surfaces.

Visualisation des courbes spectrales

Lire les signatures comme des profils radiométriques

SCP permet d'afficher les signatures sous forme de courbes spectrales.

Dans ces graphiques :

- l'axe horizontal représente les bandes spectrales ;
- l'axe vertical représente la réflectance.

Chaque courbe correspond à une classe définie par les ROI.

Cette représentation facilite grandement l'interprétation des comportements spectraux.

Exemples de signatures spectrales

L'eau

L'eau présente généralement :

- des valeurs faibles dans toutes les bandes ;
- une réflectance particulièrement faible dans le proche infrarouge.

Cette absorption du NIR constitue l'une des caractéristiques spectrales les plus importantes des surfaces en eau.

Le sable

Le sable présente le plus souvent :

- des valeurs modérées ;
- une signature relativement homogène ;
- une courbe assez plate.

Contrairement aux surfaces végétalisées, il ne montre pas de forte augmentation de réflectance dans le proche infrarouge.

Les algues vertes

Les algues possèdent une signature spectrale plus proche de celle de la végétation.

On observe généralement :

- une faible réflectance dans le rouge ;
- une forte réflectance dans le NIR ;
- une transition marquée dans la zone du Red Edge.

Cette transition spectrale constitue la base de nombreux indices de végétation.

Le contraste entre le rouge et le proche infrarouge peut notamment être observé avec :

$$NDVI = \frac{NIR - Rouge}{NIR + Rouge}$$

Interpréter les courbes spectrales

Comparer les classes

L'analyse comparative des courbes permet de vérifier si les différentes classes sont réellement distinctes.

Cette étape aide notamment à :

- identifier les bandes les plus discriminantes ;
- détecter les recouvrements entre classes ;
- évaluer la difficulté future de la classification.

Si deux classes présentent des signatures très proches, les algorithmes de classification auront du mal à les distinguer correctement.

Pourquoi cette étape est essentielle

Préparer correctement une classification

L'analyse des signatures spectrales constitue une étape fondamentale avant toute classification supervisée.

Elle permet :

- de valider les classes sélectionnées ;
- d'améliorer la qualité de l'apprentissage ;
- de limiter les erreurs d'interprétation ;
- de comprendre le comportement spectral réel des surfaces.

Sans cette étape, la classification risque d'être construite sur des classes mal définies ou mal séparées.

Une idée fondamentale en télédétection

On ne classe pas directement des objets visibles

En télédétection, les algorithmes ne “voient” pas directement des objets comme un observateur humain.

Ils analysent uniquement :

des signatures spectrales

Autrement dit, les classifications reposent sur des différences radiométriques entre pixels.

Comprendre ces signatures est donc indispensable pour interpréter correctement les résultats produits par les traitements multispectraux.

Chapitre 25 — Les limites de l'interprétation spectrale

Des signatures parfois ambiguës

Même lorsque les signatures spectrales semblent bien séparées, certaines surfaces peuvent rester difficiles à distinguer.

Dans les milieux intertidaux, plusieurs éléments peuvent produire des signatures proches de celles des algues :

- les biofilms ;
- les substrats humides ;
- les algues partiellement immergées.

Ces surfaces peuvent être confondues lors des classifications automatiques.

Une distinction essentielle

Cartographier une signature, pas directement les algues

En télédétection, on ne cartographie jamais directement les objets biologiques eux-mêmes.

On cartographie :

une signature spectrale interprétée

Cette distinction est fondamentale.

Les cartes produites représentent des comportements radiométriques que l'on interprète ensuite à la lumière :

- du contexte écologique ;
- des observations de terrain ;

- des connaissances du milieu.

La télédétection constitue ainsi un outil puissant d'aide à l'observation, mais son interprétation nécessite toujours une approche critique et scientifique.

Chapitre 26 — Les 10 erreurs fréquentes en pré-traitement sous QGIS

Pourquoi les erreurs de pré-traitement sont critiques

En télédétection multispectrale, les erreurs les plus importantes ne proviennent pas toujours des algorithmes de classification ou des calculs d'indices.

Dans de nombreux cas, les problèmes apparaissent bien plus tôt, lors des étapes de pré-traitement des données.

Une erreur discrète dans l'organisation des rasters, dans l'alignement des bandes ou dans l'interprétation des valeurs peut produire des résultats scientifiquement incorrects tout en donnant l'impression d'un traitement réussi.

Comprendre les erreurs les plus fréquentes permet donc d'éviter de nombreux pièges dans les analyses sous QGIS.

Des bandes multispectrales mal alignées

Une erreur fréquente mais souvent invisible

L'une des erreurs les plus courantes consiste à utiliser des bandes multispectrales qui ne sont pas parfaitement alignées.

Dans ce cas, les pixels des différentes bandes ne correspondent plus exactement à la même surface au sol.

Ce problème peut sembler mineur à l'écran, mais il devient très problématique lors du calcul des indices spectraux.

Par exemple, le calcul du NDVI suppose que les bandes NIR et Rouge représentent exactement les mêmes pixels.

Lorsque ce n'est pas le cas, les indices produisent :

- des artefacts ;
- du bruit spatial ;
- des erreurs de classification.

L'alignement raster constitue donc une étape fondamentale des pré-traitements.

Utiliser des données 8 bits

Des images “jolies” mais scientifiquement limitées

Certaines orthomosaïques sont exportées sous forme d'images 8 bits contenant des valeurs comprises entre :

0 et 255

Ces données sont adaptées à la visualisation mais beaucoup moins aux analyses scientifiques.

La compression en 8 bits réduit fortement la précision radiométrique et limite la capacité à détecter de faibles variations spectrales.

Pour les analyses multispectrales, il est préférable d'utiliser des formats comme :

- Float32 ;
- UInt16.

Vérifier le type de données raster doit devenir un réflexe systématique.

Mélanger plusieurs projections

Un problème classique dans les SIG

Une autre erreur fréquente consiste à utiliser des couches possédant des systèmes de coordonnées différents.

Par exemple :

- certaines couches en WGS84 ;
- d'autres en UTM.

Ces incohérences peuvent produire :

- des décalages spatiaux ;
- des erreurs de superposition ;
- des calculs incorrects.

Avant toute analyse, il est indispensable de vérifier que toutes les couches utilisent un système de coordonnées cohérent.

Ignorer la taille des pixels

La résolution influence directement les calculs

Les rasters utilisés ensemble doivent posséder la même résolution spatiale.

Si les tailles de pixels diffèrent, les calculs raster deviennent incohérents car les valeurs ne correspondent plus aux mêmes surfaces.

Même une différence très faible peut perturber :

- les indices spectraux ;
- les classifications ;
- les statistiques spatiales.

Il faut donc toujours vérifier :

- la taille des pixels ;
- l'alignement des grilles raster.

Calculer des indices sur des zones d'eau

Une source importante de bruit

L'eau absorbe fortement le proche infrarouge.

Lorsque des indices comme le NDVI sont calculés sur des surfaces immergées, les résultats deviennent souvent aberrants.

Ces pixels peuvent produire :

- des valeurs extrêmes ;
- du bruit spatial ;
- des erreurs de classification.

Pour limiter ce problème, il est fortement recommandé de créer un masque d'eau avant les analyses.

Le NDWI est souvent utilisé pour cette étape.

Ne pas inspecter les valeurs des pixels

Une erreur fréquente chez les débutants

Avant de calculer des indices ou de lancer une classification, il est indispensable d'examiner les valeurs contenues dans les rasters.

Cette étape permet notamment de vérifier :

- les valeurs de réflectance ;
- la cohérence radiométrique ;
- la distribution statistique des pixels.

Dans QGIS, plusieurs outils facilitent cette inspection :

- Value Tool ;
- Histogram.

Comprendre les données avant de les traiter constitue une règle fondamentale en télédétection.

Ignorer les zones saturées

Quand le capteur atteint ses limites

Certaines surfaces peuvent être surexposées ou saturées.

Dans ces zones, le capteur atteint sa valeur maximale et perd une partie de l'information radiométrique.

Les pixels saturés doivent souvent être :

- identifiés ;
- masqués ;
- exclus des analyses.

L'analyse des histogrammes permet généralement de détecter facilement ces problèmes.

Interpréter directement un indice spectral

Un indice n'est pas un objet réel

Une erreur fréquente consiste à considérer un indice comme une représentation directe d'un phénomène biologique.

Par exemple, le NDVI ne représente pas directement des algues.

Il représente une signature spectrale calculée à partir de deux bandes

Cette nuance est essentielle.

L'interprétation des indices doit toujours prendre en compte :

- le contexte écologique ;
- les conditions d'acquisition ;
- les limites du milieu étudié.

Oublier les conditions d'acquisition

Des variations parfois indépendantes du phénomène étudié

Les valeurs spectrales peuvent varier fortement en raison :

- de la marée ;
- de l'humidité ;
- de l'illumination ;
- des conditions météorologiques.

Une variation observée entre deux dates ne correspond pas nécessairement à une évolution écologique réelle.

Sans protocole d'acquisition rigoureux, les comparaisons temporelles deviennent très fragiles.

Ne pas vérifier les signatures spectrales

Une étape indispensable avant la classification

Avant toute classification supervisée, il est indispensable de vérifier les signatures spectrales des différentes classes.

Cette étape permet :

- de vérifier que les classes sont distinctes ;
- d'identifier les bandes les plus discriminantes ;
- de détecter les recouvrements entre surfaces.

Dans Semi-Automatic Classification Plugin, cette analyse peut être réalisée avec l'outil :

Spectral Signature Plot

Si les signatures des classes sont trop proches, la classification sera probablement difficile et peu fiable.

Conclusion

La qualité des résultats en télédétection dépend directement de la qualité des pré-traitements réalisés dans le SIG.

Les erreurs de préparation des données sont souvent discrètes mais leurs conséquences peuvent être importantes sur les indices, les classifications et les interprétations scientifiques.

Développer une démarche rigoureuse de vérification constitue donc une compétence essentielle dans tout projet de télédétection environnementale sous QGIS.

Troisième partie — Semi-Automatic Classification Plugin : le cœur de la chaîne de classification

Ce que nous allons apprendre dans cette partie

Dans les chapitres suivants, nous allons apprendre à utiliser SCP pour construire une chaîne complète de classification multispectrale.

Les principales étapes abordées seront :

- la préparation des bandes multispectrales dans SCP ;
- le calcul des indices spectraux utiles ;
- la création de ROI de qualité ;
- l'analyse des signatures spectrales ;
- le lancement d'une classification supervisée ;
- l'interprétation et la validation des résultats.

Cette troisième partie constituera le cœur opérationnel de la chaîne de traitement des données drone sous QGIS.

Chapitre 27 — Pourquoi utiliser SCP ?

Du raster multispectral à la carte thématique

Les images multispectrales acquises par drone contiennent une grande quantité d'informations spectrales. Chaque pixel possède des valeurs de réflectance dans plusieurs bandes, mais ces données brutes restent difficiles à interpréter directement.

L'objectif de la télédétection n'est pas uniquement d'afficher des images ou de calculer quelques indices spectraux. L'enjeu principal consiste à transformer ces données en informations thématiques exploitables.

C'est précisément le rôle de la classification.

Dans QGIS, l'un des outils les plus utilisés pour cette tâche est le :

Semi-Automatic Classification Plugin

souvent appelé simplement SCP.

Ce plugin est spécialement conçu pour l'analyse d'images satellitaires et multispectrales issues de drones.

Il permet de transformer des bandes raster en véritables cartes thématiques représentant les différentes surfaces observées.

Les limites des outils raster classiques

Ce qu'il est possible de faire sans SCP

Les outils raster standards de QGIS permettent déjà de réaliser plusieurs opérations utiles.

Il est notamment possible :

- d'afficher les différentes bandes ;
- de calculer des indices spectraux ;
- de comparer visuellement plusieurs images ;
- d'explorer les valeurs des pixels.

Ces outils constituent une base importante pour l'analyse multispectrale.

Les difficultés rencontrées sans outil spécialisé

Cependant, dès que l'on souhaite réaliser une véritable classification supervisée, les limites des outils classiques apparaissent rapidement.

Sans environnement spécialisé, il devient difficile :

- de gérer des échantillons d'apprentissage ;
- d'organiser les signatures spectrales ;
- de comparer les classes ;
- de reproduire précisément une méthode de travail ;
- d'automatiser les classifications.

La gestion complète d'un workflow de télédétection devient alors complexe et peu reproductible.

Ce que SCP apporte

Une plateforme complète de classification

Le principal intérêt de SCP réside dans le fait qu'il centralise l'ensemble du workflow de télédétection dans une seule interface intégrée à QGIS.

Le plugin permet notamment de gérer :

- les *Band Sets* ;
- les indices spectraux ;
- les ROI (*Regions of Interest*) ;
- les signatures spectrales ;
- les classifications supervisées ;
- la validation des résultats.

Cette centralisation facilite considérablement le travail d'analyse multispectrale.

Les Band Sets

Organiser les bandes multispectrales

Dans SCP, les bandes multispectrales sont regroupées dans des ensembles appelés *Band Sets*.

Ces ensembles permettent de définir quelles bandes seront utilisées dans les traitements et dans les classifications.

Par exemple, un *Band Set* peut contenir :

- Green ;
- Red ;
- Red Edge ;
- NIR.

Cette organisation simplifie le calcul des indices et l'utilisation des algorithmes de classification.

Les indices spectraux dans SCP

Calculer rapidement les indicateurs

SCP intègre également des outils permettant de calculer automatiquement plusieurs indices spectraux courants.

Parmi les plus utilisés figurent :

le NDVI :

$$NDVI = \frac{NIR - Rouge}{NIR + Rouge}$$

le NDRE :

$$NDRE = \frac{NIR - RedEdge}{NIR + RedEdge}$$

ou encore le NDWI :

$$NDWI = \frac{Green - NIR}{Green + NIR}$$

Ces indices peuvent ensuite être utilisés directement dans les classifications.

Les ROI et les signatures spectrales

Construire les classes d'apprentissage

SCP permet également de créer facilement des zones d'intérêt servant d'échantillons d'apprentissage.

Ces ROI représentent des surfaces connues telles que :

- l'eau ;
- le sable ;
- les rochers ;
- les tapis d'algues ;
- les substrats humides.

À partir de ces échantillons, SCP extrait les signatures spectrales des différentes classes.

L'utilisateur peut ensuite visualiser les courbes spectrales et comparer les comportements radiométriques des surfaces observées.

Cette étape constitue la base de toute classification supervisée.

L'intérêt de SCP pour le suivi des algues vertes

Produire des cartes objectives

Dans les projets de suivi des algues vertes, SCP permet de distinguer automatiquement plusieurs types de surfaces présents sur l'estran.

Le plugin peut notamment aider à différencier :

- les surfaces en eau ;
- les zones sableuses ;
- les rochers ;
- les substrats humides ;
- les tapis d'algues.

L'objectif n'est plus simplement d'observer les images, mais de produire :

une carte thématique objective et reproductible des surfaces observées

Cette approche permet de quantifier les phénomènes étudiés et d'améliorer la comparabilité des résultats dans le temps.

Pourquoi cette approche est importante scientifiquement

Une méthode reproductible

L'un des principaux avantages d'une classification réalisée avec SCP est sa reproductibilité.

Contrairement à une simple interprétation visuelle, la méthode devient :

- documentée ;
- répétable ;
- comparable entre plusieurs dates ;
- plus robuste scientifiquement.

Les mêmes paramètres peuvent être réutilisés sur plusieurs acquisitions afin de suivre l'évolution des surfaces algales dans le temps.

Cette standardisation constitue un élément fondamental des études environnementales quantitatives.

A retenir

L'objectif de la télédétection n'est pas uniquement de "voir" les algues sur une image.

L'objectif est de :

les identifier de manière rigoureuse, reproductible et quantifiable

Cette distinction marque le passage entre une simple observation visuelle et une véritable démarche scientifique de classification.

Chapitre 28 — Préparer les bandes multispectrales dans Semi-Automatic Classification Plugin

Le rôle central des *Band Sets*

Dans le plugin SCP, toutes les opérations de télédétection reposent sur une organisation particulière des données appelée *Band Set*. Cette notion est fondamentale, car elle structure l'ensemble du workflow de classification.

Un *Band Set* correspond à un regroupement cohérent des différentes bandes spectrales issues d'un même capteur. Dans le cas d'un drone multispectral, cela signifie que les différentes bandes enregistrées lors du vol sont réunies dans un ensemble unique afin que SCP puisse les interpréter correctement.

Cette étape peut sembler purement technique, mais elle conditionne en réalité tout le reste de l'analyse. Sans *Band Set* correctement défini, SCP ne peut ni calculer les indices spectraux de manière fiable, ni analyser les signatures spectrales, ni lancer une classification supervisée cohérente.

Le *Band Set* agit donc comme une structure de référence permettant au logiciel de comprendre quelles bandes doivent être utilisées et dans quel ordre elles doivent être interprétées.

Une condition indispensable : la cohérence géométrique

Le fonctionnement du *Band Set* repose sur un principe fondamental de la télédétection multispectrale : chaque pixel doit représenter exactement la même position au sol dans toutes les bandes.

Autrement dit, lorsqu'un pixel est analysé dans la bande verte, la bande rouge, la bande Red Edge ou la bande proche infrarouge, il doit toujours correspondre à la même surface réelle sur l'estran.

Cette contrainte explique pourquoi les pré-traitements géomatiques réalisés précédemment sont si importants. Avant même de créer un *Band Set*, il est indispensable de vérifier soigneusement l'alignement des rasters, leur résolution spatiale, leur emprise géographique ainsi que leur système de coordonnées.

Si une seule bande présente un léger décalage, les indices spectraux deviendront incohérents et les classifications pourront produire des résultats erronés.

Les bandes du DJI Mavic 3 Multispectral

Dans notre projet, le drone multispectral fournit quatre bandes principales spécialement conçues pour l'analyse environnementale.

Ces bandes correspondent au vert, au rouge, au Red Edge et au proche infrarouge. Elles constituent le cœur de l'analyse multispectrale et permettent de calculer les principaux indices utilisés pour détecter les algues et analyser leur biomasse.

Les bandes RGB classiques peuvent également être utilisées pour certaines visualisations ou pour l'interprétation visuelle des orthomosaïques, mais ce sont principalement les quatre bandes multispectrales qui seront utilisées dans SCP.

Création du *Band Set* dans SCP

La première étape pratique consiste à ouvrir le plugin SCP dans QGIS. Le plugin apparaît généralement sous la forme d'un panneau latéral intégré à l'interface du logiciel.

Dans l'onglet consacré aux *Band Sets*, les différentes bandes raster doivent être ajoutées une par une. Cette étape demande de la rigueur, car l'ordre des bandes est extrêmement important.

Dans notre workflow, l'ordre correct est le suivant : Green, Red, Red Edge puis NIR.

Cet ordre détermine la manière dont SCP attribue automatiquement les identifiants internes des bandes, généralement notés b1, b2, b3 et b4.

Ainsi :

- b1 correspondra à Green ;
- b2 correspondra à Red ;
- b3 correspondra à Red Edge ;
- b4 correspondra à NIR.

Une erreur dans cet ordre peut avoir des conséquences importantes sur l'ensemble des calculs ultérieurs. Un NDVI calculé avec des bandes inversées produira par exemple des valeurs totalement incohérentes.

Les erreurs fréquentes lors de la création du *Band Set*

Plusieurs erreurs apparaissent régulièrement lors de cette étape.

Il arrive fréquemment que certaines bandes soient légèrement décalées, que des rasters provenant de traitements différents soient mélangés ou encore qu'une bande soit oubliée dans le *Band Set*.

Une autre erreur très fréquente consiste à ne pas vérifier la résolution spatiale ou le système de coordonnées des rasters avant leur intégration dans SCP.

Ces problèmes peuvent sembler mineurs, mais ils suffisent à rendre les classifications instables ou scientifiquement incorrectes.

Le résultat attendu

Une fois correctement configuré, le *Band Set* devient la base de toute l'analyse multispectrale dans SCP.

Il permet de calculer les indices spectraux, d'extraire les signatures spectrales, de créer des ROI cohérentes et de lancer les classifications supervisées.

Le *Band Set* constitue donc le véritable socle technique de la chaîne de traitement sous SCP.

Chapitre 29 — Calcul des indices spectraux dans SCP

Transformer les bandes en indicateurs exploitables

Une fois le *Band Set* correctement défini, SCP permet de calculer rapidement différents indices spectraux.

Ces indices jouent un rôle essentiel dans l'analyse des algues vertes, car ils permettent de transformer les valeurs radiométriques des bandes en indicateurs plus facilement interprétables.

L'objectif n'est plus simplement d'observer des bandes individuelles, mais de mettre en évidence certains comportements biologiques liés à la chlorophylle, à la biomasse ou à l'humidité des surfaces.

Dans ce module, nous utiliserons principalement le NDVI, le NDRE et le GNDVI.

Où calculer les indices dans SCP ?

Dans SCP, les calculs sont généralement réalisés dans l'outil appelé *Band calc.* Selon la version du plugin, cet outil peut être accessible directement dans le panneau principal ou dans l'onglet du *Band Set*.

Le principe est simple : l'utilisateur écrit une formule mathématique utilisant les identifiants des bandes du *Band Set*.

Ces identifiants correspondent directement à l'ordre défini lors de la préparation des bandes.

Cette étape montre une nouvelle fois pourquoi l'organisation correcte du *Band Set* est si importante.

Le calcul du NDVI

Le NDVI est l'un des indices de végétation les plus utilisés en télédétection.

Il repose sur la différence entre la bande proche infrarouge et la bande rouge :

$$NDVI = \frac{NIR - Rouge}{NIR + Rouge}$$

Dans SCP, cette formule devient :

$$(b4 - b2) / (b4 + b2)$$

Cet indice exploite le comportement spectral caractéristique des surfaces riches en chlorophylle. Les algues absorbent fortement la lumière rouge tout en réfléchissant davantage le proche infrarouge.

Plus ce contraste est important, plus la valeur du NDVI devient élevée.

Le calcul du NDRE

Le NDRE utilise la bande Red Edge à la place de la bande rouge :

$$NDRE = \frac{NIR - RedEdge}{NIR + RedEdge}$$

Dans SCP, l'expression correspondante est :

$$(b4 - b3) / (b4 + b3)$$

Le NDRE est souvent plus sensible aux variations fines de biomasse et à la chlorophylle que le NDVI classique.

Dans les environnements littoraux, il peut parfois mieux distinguer certaines zones algales ou réduire certains effets de saturation.

Le calcul du GNDVI

Le GNDVI exploite quant à lui la bande verte :

$$GNDVI = \frac{NIR - Green}{NIR + Green}$$

Dans SCP :

$$(b4 - b1) / (b4 + b1)$$

Cet indice peut être utile pour analyser certaines variations physiologiques ou certaines différences de densité de biomasse.

Les étapes pratiques dans SCP

Le workflow de calcul reste relativement simple. Une fois le *Band Set* sélectionné, il suffit d'ouvrir *Band calc*, de saisir la formule souhaitée puis de lancer le traitement.

Le résultat produit est un nouveau raster contenant les valeurs de l'indice calculé pour chaque pixel de l'image.

Ces rasters peuvent ensuite être visualisés, analysés ou utilisés dans les classifications supervisées.

Interpréter les indices avec prudence

Les indices spectraux produisent généralement des valeurs comprises entre -1 et +1.

Les valeurs négatives correspondent souvent à l'eau ou aux ombres, tandis que les valeurs positives indiquent généralement des surfaces riches en chlorophylle.

Cependant, il est essentiel de rappeler qu'un indice spectral ne constitue pas une carte directe des algues.

Il s'agit uniquement d'un indicateur construit à partir du comportement radiométrique des surfaces observées.

Son interprétation doit toujours prendre en compte le contexte écologique, les conditions d'acquisition, l'humidité du substrat ainsi que les nombreuses perturbations propres au milieu intertidal.

Les indices spectraux sont donc des outils extrêmement puissants, mais ils nécessitent toujours une interprétation critique et rigoureuse.

Chapitre 30 — Création d'indices personnalisés pour le suivi des algues vertes

Les limites des indices spectraux classiques

Les indices spectraux classiques comme le NDVI, le NDRE ou le GNDVI sont aujourd'hui largement utilisés en télédétection. Ils permettent de mettre en évidence l'activité chlorophyllienne, la biomasse ou encore certaines caractéristiques physiologiques de la végétation.

Ces indices sont particulièrement efficaces dans des environnements relativement homogènes, comme les cultures agricoles ou les surfaces végétalisées terrestres. Pourtant, lorsqu'on travaille en milieu intertidal, les conditions deviennent beaucoup plus complexes.

L'estran constitue un environnement extrêmement hétérogène où les surfaces observées changent rapidement sur de très courtes distances. Une même image peut contenir simultanément des zones de sable sec, des substrats saturés en eau, des flaques résiduelles, des biofilms, des rochers et des tapis d'algues partiellement immergés.

Dans ce contexte, les indices classiques deviennent parfois insuffisants pour distinguer correctement les surfaces.

Pourquoi les algues vertes sont difficiles à détecter

Les algues vertes présentent des caractéristiques particulières qui compliquent leur détection spectrale.

Contrairement à une végétation terrestre dense et bien structurée, les tapis algaux sont souvent très humides, relativement fins et parfois partiellement transparents. Dans de nombreuses situations, le substrat reste visible sous les algues, ce qui modifie fortement le signal enregistré par le capteur.

Le problème devient encore plus complexe lorsque les algues sont recouvertes par un film d'eau ou mélangées à des substrats humides.

Le signal spectral observé par le drone ne correspond alors plus uniquement aux algues elles-mêmes, mais à un mélange complexe associant eau, sable, matière organique et surfaces biologiques.

Dans certaines situations, un substrat humide peut produire des valeurs de NDVI proches de celles d'une zone réellement couverte par des algues.

Ainsi, les indices standards deviennent parfois trop généraux ou insuffisamment discriminants pour analyser correctement les surfaces intertidales.

L'idée des indices personnalisés

Pour répondre à ces limites, il devient souvent nécessaire de construire des indices mieux adaptés au contexte local.

Le principe d'un indice personnalisé consiste à exploiter les particularités spectrales du site étudié afin d'améliorer la séparation entre les différentes surfaces présentes sur l'image.

Contrairement aux indices standards, qui utilisent une formule fixe applicable à de nombreux contextes, un indice personnalisé peut être adapté spécifiquement aux caractéristiques du milieu étudié.

Cette approche permet de tenir compte :

- de l'humidité du substrat ;
- de la présence d'eau ;
- des conditions d'éclairage ;
- des signatures particulières des algues locales ;
- des caractéristiques du capteur utilisé.

L'objectif n'est plus seulement de détecter de la végétation au sens général, mais de mieux isoler certaines surfaces spécifiques comme les tapis d'algues présents sur l'estran.

Combiner plusieurs informations spectrales

Un indice personnalisé peut combiner plusieurs types d'informations.

Il peut s'appuyer sur différentes bandes spectrales, mais également intégrer plusieurs indices calculés séparément.

Par exemple, il peut être utile de combiner le comportement du NDVI avec celui du NDWI afin de limiter les confusions entre algues et surfaces immergées.

Dans d'autres cas, on peut également ajouter des seuils sur certaines bandes afin d'exclure les pixels correspondant aux ombres ou aux zones saturées.

L'intérêt de cette approche est de construire progressivement un indicateur plus spécifique au phénomène réellement étudié.

L'importance des conditions logiques

Les indices personnalisés ne reposent pas uniquement sur des formules mathématiques classiques.

Ils peuvent également intégrer des conditions logiques permettant de sélectionner certains types de pixels et d'en exclure d'autres.

Par exemple, un pixel peut être considéré comme une zone potentiellement algale si plusieurs conditions sont réunies simultanément :

- le NDVI est positif ;
- le NDWI est négatif ;
- la valeur du proche infrarouge reste suffisamment élevée.

Cette logique permet de mieux distinguer certaines surfaces qui présentent des comportements radiométriques proches mais des propriétés physiques différentes.

L'eau, par exemple, absorbe fortement le proche infrarouge mais conserve généralement un certain signal dans le vert. Les ombres, au contraire, produisent une diminution globale du signal dans toutes les bandes.

En combinant plusieurs critères, il devient donc possible d'améliorer la robustesse de la détection.

Une démarche expérimentale et scientifique

Contrairement aux indices standards, les indices personnalisés ne sont pas universels.

Un indice efficace sur un site particulier peut devenir beaucoup moins performant dans un autre contexte environnemental.

Les résultats dépendent fortement :

- du type de substrat ;
- de l'humidité ;
- des conditions d'acquisition ;
- de la saison ;
- du capteur utilisé ;
- des caractéristiques écologiques locales.

La création d'indices personnalisés relève donc d'une véritable démarche scientifique et expérimentale.

Elle nécessite :

- l'analyse des signatures spectrales ;
- des observations de terrain ;
- des comparaisons entre classes ;
- des tests successifs ;
- une validation rigoureuse des résultats.

Dans la pratique, les chercheurs construisent souvent plusieurs versions d'un indice avant d'identifier la combinaison la plus pertinente pour leur site d'étude.

Le rôle de QGIS et de SCP

Les outils comme QGIS et Semi-Automatic Classification Plugin facilitent énormément cette approche expérimentale.

Grâce à la Raster Calculator, au module *Band calc* ou encore aux outils de masquage, il devient possible de tester rapidement différentes combinaisons de bandes, de seuils et de conditions logiques.

Cette flexibilité permet de construire des workflows spécifiquement adaptés au suivi des algues vertes en milieu littoral.

A retenir

Les indices spectraux standards constituent une base très utile pour l'analyse multispectrale, mais ils ne suffisent pas toujours dans les environnements complexes comme l'estran.

Dans le suivi des algues vertes, il est souvent nécessaire d'adapter les indices au contexte local afin d'améliorer la séparation entre les différentes surfaces observées.

La télédétection environnementale ne consiste donc pas simplement à appliquer automatiquement des formules existantes.

Elle implique également une démarche d'interprétation, d'expérimentation et d'adaptation aux réalités physiques et écologiques du terrain.

Chapitre 31 — Les ROI et la stratégie d'échantillonnage : le cœur de la classification supervisée

Pourquoi les ROI sont si importantes

Dans Semi-Automatic Classification Plugin, les ROI (*Regions Of Interest*) constituent la base même de la classification supervisée.

Les algorithmes de classification ne “devinent” pas automatiquement ce qu'est une zone d'algues, une surface sableuse ou une zone d'eau. Ils apprennent à reconnaître ces surfaces à partir des exemples fournis par l'utilisateur.

Ces exemples sont précisément les ROI.

Autrement dit, les ROI définissent les signatures spectrales qui serviront de référence pendant toute la classification.

Cette idée est fondamentale : dans une classification supervisée, la qualité des résultats dépend avant tout de la qualité des échantillons d'apprentissage.

Une mauvaise sélection des ROI produira presque toujours une mauvaise classification, même avec un algorithme performant.

À l'inverse, des ROI bien construites permettent souvent d'obtenir des résultats robustes et scientifiquement cohérents.

Le principe des ROI

Une ROI correspond à une zone représentative d'une classe particulière présente dans l'image.

Dans le cadre du suivi des algues vertes, les principales classes peuvent correspondre à :

- des tapis d'algues ;
- du sable ;
- des surfaces rocheuses ;
- de l'eau ;
- des zones humides.

Chaque ROI sert à décrire le comportement spectral moyen d'une surface donnée.

L'algorithme de classification apprend ensuite à reconnaître dans toute l'image les pixels présentant des signatures similaires.

Les ROI jouent donc un rôle beaucoup plus important que l'algorithme lui-même.

L'algorithme applique simplement les règles construites à partir des exemples fournis.

La notion d'homogénéité

La première règle essentielle dans la création des ROI concerne l'homogénéité spectrale.

Chaque ROI doit correspondre à une surface aussi homogène que possible.

Si une ROI contient simultanément des algues, du sable humide et quelques pixels d'eau résiduelle, la signature spectrale obtenue deviendra ambiguë.

L'algorithme apprendra alors une signature "moyenne" qui ne représentera correctement aucune des surfaces réelles.

C'est l'une des erreurs les plus fréquentes chez les débutants.

Les bordures entre classes doivent être évitées autant que possible, car elles contiennent souvent des pixels mixtes particulièrement problématiques en télédétection.

L'importance de la répartition spatiale

Une autre erreur fréquente consiste à créer toutes les ROI dans une seule petite zone de l'image.

Or, une même classe peut présenter des signatures légèrement différentes selon les conditions locales.

Par exemple, des algues situées dans une zone humide ne produiront pas exactement le même signal spectral que des algues plus sèches situées ailleurs sur l'estran.

Pour cette raison, les ROI doivent être réparties sur l'ensemble de la zone d'étude.

Cette stratégie permet de mieux capturer la variabilité réelle des surfaces présentes sur le terrain.

Capter la variabilité des classes

Une bonne classification ne repose pas uniquement sur des classes "moyennes".

Elle doit également prendre en compte la variabilité naturelle des surfaces observées.

Dans le cas des algues vertes, cette variabilité peut être importante.

Certaines zones présentent des tapis très denses, tandis que d'autres correspondent à des accumulations fines ou partiellement immergées.

L'humidité du substrat peut également modifier fortement les signatures spectrales.

Pour cette raison, il est souvent préférable de créer plusieurs ROI correspondant à différentes situations :

- algues denses ;
- algues fines ;
- zones très humides ;
- surfaces partiellement recouvertes.

Cette approche améliore fortement la robustesse des classifications supervisées.

La taille des ROI

La taille des ROI doit également être choisie avec soin.

Des ROI trop petites risquent d'être fortement influencées par le bruit radiométrique ou les anomalies locales.

À l'inverse, des ROI trop grandes augmentent le risque d'inclure plusieurs types de surfaces différentes.

Dans la pratique, il est souvent préférable d'utiliser plusieurs petites ROI homogènes plutôt qu'une seule grande zone couvrant des surfaces variées.

Cette stratégie permet de mieux contrôler la qualité des signatures spectrales.

Les erreurs fréquentes

Plusieurs erreurs apparaissent régulièrement lors de la création des ROI.

Certaines classifications reposent sur un nombre insuffisant d'échantillons, ce qui limite fortement la capacité de généralisation de l'algorithme.

D'autres problèmes proviennent de ROI mal placées, contenant des mélanges de surfaces ou positionnées dans des zones bruitées.

Les zones d'ombre, les bordures ou les surfaces saturées doivent être évitées autant que possible.

Ces erreurs produisent souvent des signatures spectrales incohérentes et dégradent fortement les résultats de classification.

Les particularités du site de Guéthary

Dans le cas d'un site littoral comme Guéthary, plusieurs difficultés spécifiques doivent être prises en compte.

Les tapis d'algues sont souvent partiellement mélangés au substrat sableux. Certaines zones restent humides même à marée basse, tandis que des ombres résiduelles peuvent modifier localement les signatures spectrales.

Le signal enregistré par le capteur correspond alors à un mélange complexe entre :

- algues ;
- sable ;
- eau ;
- humidité du substrat.

La création des ROI doit donc être réalisée avec beaucoup de prudence.

Procédure détaillée pour créer une ROI “algues”

Préparer correctement l’affichage

Avant même de commencer à dessiner les ROI, il est important de choisir un affichage facilitant l’identification des surfaces algales.

Les indices comme le NDVI, le NDRE ou le GNDVI constituent souvent une bonne base de travail.

Une composition colorée multispectrale peut également être utilisée.

Dans une composition de type :

- NIR dans le rouge ;
- Red dans le vert ;
- Green dans le bleu ;

les algues apparaissent souvent sous des teintes rouges très visibles.

Cette représentation facilite fortement le repérage des zones biologiquement actives.

Identifier des zones fiables

Une fois l’affichage préparé, il faut rechercher des zones clairement identifiables.

Les meilleures ROI correspondent généralement à des tapis d’algues bien visibles, homogènes et éloignés des zones de transition.

Les bordures entre sable et algues doivent être évitées, tout comme les zones contenant de l’eau résiduelle ou des ombres.

Les pixels mixtes constituent l’un des principaux problèmes des classifications supervisées.

Dessiner les ROI dans SCP

Dans SCP, l’outil de création de ROI permet de dessiner directement des polygones sur l’image.

L’utilisation de petits polygones homogènes est généralement préférable à celle de points isolés.

Une ROI contenant entre une dizaine et quelques dizaines de pixels constitue souvent un bon compromis entre stabilité statistique et homogénéité.

Il est recommandé de privilégier plusieurs petites ROI réparties sur l’image plutôt qu’une seule grande zone.

Multiplier les ROI pour une même classe

Une seule ROI ne suffit généralement pas pour représenter correctement une classe complexe comme les algues vertes.

Il est préférable de créer plusieurs ROI correspondant à différentes situations observées sur le terrain.

Certaines ROI peuvent représenter des tapis très denses, d'autres des zones plus humides ou des accumulations plus fines.

Cette stratégie permet de mieux capturer la variabilité réelle des signatures spectrales.

Vérifier les signatures spectrales

Une fois les ROI créées, SCP permet d'afficher les courbes spectrales correspondantes.

Cette étape est essentielle.

Les différentes ROI d'une même classe doivent présenter des signatures relativement cohérentes.

Il faut également vérifier que les classes restent suffisamment séparées les unes des autres.

Si les courbes se recouvrent fortement, la classification deviendra beaucoup plus difficile.

Contrôler les ROI avant la classification

Avant de lancer la classification, plusieurs vérifications sont indispensables.

Il faut s'assurer que les ROI ne contiennent pas :

- d'eau ;
- de sable ;
- d'ombres ;
- de pixels mixtes.

Une ROI mal définie peut perturber l'ensemble du modèle de classification.

Cette étape de contrôle constitue donc un élément essentiel du workflow scientifique.

Une stratégie avancée : les sous-types de classes

Dans les projets plus avancés, il peut être utile de subdiviser certaines classes.

Par exemple, les algues peuvent être séparées en plusieurs sous-types :

- algues denses ;
- algues fines ;
- algues humides.

Cette approche permet à SCP de mieux représenter la variabilité réelle des surfaces et améliore souvent la qualité des résultats.

Sauvegarder les ROI

Enfin, les ROI doivent être enregistrées afin de garantir la reproductibilité du travail.

La sauvegarde permet :

- de réutiliser les mêmes échantillons ;
- de comparer plusieurs dates ;
- de documenter précisément la méthode utilisée.

Cette étape est particulièrement importante dans les projets scientifiques nécessitant des analyses reproductibles dans le temps.

A retenir

Dans une classification supervisée, les ROI constituent le véritable cœur du modèle.

L'algorithme ne fait qu'appliquer les règles construites à partir des échantillons fournis.

La qualité scientifique des résultats dépend donc avant tout :

- de la qualité des ROI ;
- de leur représentativité ;
- de leur homogénéité ;
- et de la compréhension du contexte écologique du site étudié.

Chapitre 32 — La matrice de confusion : évaluer la qualité d’une classification

Pourquoi valider une classification ?

Une classification supervisée peut produire une carte visuellement très convaincante tout en contenant de nombreuses erreurs.

En télédétection, il ne suffit donc jamais de regarder une carte pour conclure qu’un modèle fonctionne correctement.

Une classification doit être validée de manière objective.

C’est précisément le rôle de la matrice de confusion.

Cet outil permet de comparer :

ce que prédit le modèle avec la réalité observée sur le terrain

La matrice de confusion constitue ainsi l’un des principaux outils d’évaluation des classifications supervisées.

Le principe général

Le fonctionnement est relativement simple.

Pour chaque point de validation, deux informations sont nécessaires :

- la classe réelle observée ;
- la classe prédite par la classification.

Par exemple, un point situé sur une zone d’algues peut être correctement classé comme “algues”, tandis qu’un autre point situé sur du sable peut être incorrectement classé comme “algues”.

La matrice de confusion permet précisément de comptabiliser ces correspondances et ces erreurs.

Comprendre la structure de la matrice

La matrice est organisée sous forme de tableau croisant :

- les classes réelles ;
- les classes prédites.

Chaque cellule indique le nombre de points correspondant à une combinaison donnée.

La diagonale principale représente les bonnes classifications.

Autrement dit, lorsque la classe prédite correspond à la classe réelle, le point est comptabilisé sur la diagonale.

Les cellules situées hors de cette diagonale correspondent aux erreurs de classification.

Par exemple, si plusieurs zones de sable sont classées comme des algues, ces erreurs apparaîtront dans la cellule correspondant à :

réel = sable

prédit = algues

Cette organisation permet d'identifier rapidement quelles classes sont correctement reconnues et quelles classes sont fréquemment confondues.

Interpréter les erreurs

L'intérêt de la matrice de confusion ne réside pas uniquement dans le calcul d'un pourcentage global de réussite.

Elle permet surtout de comprendre la nature des erreurs produites par la classification.

Dans les environnements littoraux, certaines confusions sont particulièrement fréquentes.

Les substrats humides peuvent être confondus avec des algues, tandis que certaines ombres peuvent être interprétées comme des zones d'eau.

La matrice permet d'identifier précisément ces problèmes et d'améliorer progressivement le modèle de classification.

La précision globale

L'indicateur le plus simple dérivé de la matrice est la précision globale.

Elle correspond au rapport entre :

- le nombre de bonnes classifications ;
- le nombre total de points validés.

Autrement dit, on additionne les valeurs de la diagonale puis on divise le résultat par le nombre total de points.

Cet indicateur fournit une estimation générale de la qualité de la classification.

Cependant, il doit être interprété avec prudence.

Une bonne précision globale ne garantit pas forcément que toutes les classes sont correctement détectées.

Certaines classes rares peuvent rester très mal classifiées malgré une précision globale élevée.

Les limites de la validation

La qualité de la validation dépend directement de la qualité des points utilisés.

Si le nombre de points est trop faible, les résultats deviennent statistiquement fragiles.

La validation dépend également fortement de la qualité des observations terrain.

Un point mal interprété produira une erreur dans la matrice, même si la classification était correcte.

Il est donc essentiel de construire soigneusement les jeux de validation.

Une règle fondamentale : ne jamais utiliser les ROI

L'une des erreurs les plus fréquentes consiste à utiliser les ROI d'apprentissage comme points de validation.

Cette pratique est incorrecte.

Les ROI ont déjà servi à entraîner le modèle. Les utiliser pour l'évaluation introduit un biais très important.

Le modèle “connaît déjà” ces pixels.

La validation doit toujours être indépendante de l’apprentissage.

Cette séparation constitue un principe fondamental de toute démarche scientifique de classification supervisée.

Création des points de validation dans QGIS

Une méthode simple et manuelle

La méthode la plus simple consiste à créer manuellement une couche de points dans QGIS.

Une nouvelle couche ponctuelle peut être créée au format GeoPackage ou Shapefile. Cette couche contient généralement un champ texte permettant d’indiquer la classe réelle observée sur le terrain.

Une fois la couche créée, le mode édition est activé afin d’ajouter les différents points de validation directement sur l’image.

L’utilisateur clique alors sur différentes surfaces de l’orthomosaïque :

- zones d’algues ;
- sable ;
- eau ;
- ombres ;
- rochers.

Chaque point reçoit ensuite une classe réelle basée sur :

- l’interprétation visuelle ;
- les connaissances terrain ;
- les observations de terrain disponibles.

Cette approche est simple et très pédagogique.

Une approche plus scientifique : les points aléatoires

Dans les études plus rigoureuses, il est souvent préférable d’utiliser des points générés aléatoirement.

Dans QGIS, l’outil :

Créer des points aléatoires

permet de générer automatiquement un ensemble de points répartis sur l'image.

L'utilisateur peut définir :

- le nombre de points ;
- l'étendue spatiale ;
- une distance minimale entre les points.

Une fois les points créés, chaque point doit être interprété manuellement afin d'attribuer sa classe réelle.

Cette méthode réduit certains biais liés à la sélection manuelle des points.

Récupérer la classe prédite

Une fois les points de validation créés, il faut récupérer la classe produite par la classification raster.

Dans QGIS, cela peut être réalisé avec des outils comme :

- Sample Raster Values ;
- Point Sampling Tool.

Ces outils ajoutent automatiquement à chaque point la valeur du raster de classification.

On obtient alors, pour chaque point :

- une classe réelle ;
- une classe prédite.

C'est cette comparaison qui permettra de construire la matrice de confusion.

Construire la matrice

Une fois les données préparées, il devient possible de comparer :

réalité terrain vs prédiction

Chaque correspondance est reportée dans la matrice afin de comptabiliser :

- les bonnes classifications ;
- les erreurs ;
- les confusions entre classes.

Cette étape permet de transformer une impression visuelle subjective en véritable évaluation quantitative du modèle.

Les règles importantes pour une bonne validation

Une validation correcte repose sur plusieurs principes essentiels.

Les points de validation doivent être indépendants des ROI utilisées pour l'apprentissage.

Ils doivent également être répartis sur l'ensemble de l'image afin de couvrir toutes les classes présentes.

Les bordures entre surfaces doivent être évitées autant que possible, car elles contiennent souvent des pixels mixtes difficiles à interpréter.

Les zones ambiguës ou incertaines doivent également être exclues.

L'objectif est de construire un jeu de validation aussi fiable que possible.

Une idée fondamentale

Une classification ne doit jamais être évaluée uniquement à partir de son apparence visuelle.

Une carte peut sembler correcte tout en contenant de nombreuses erreurs invisibles.

La matrice de confusion permet précisément de mesurer objectivement la qualité du modèle et de comprendre les limites de la classification produite.

En télédétection scientifique, une classification doit être :

validée, quantifiée et interprétée de manière critique

Utilisation d'un fichier Excel de validation

Automatiser les calculs

Dans les projets pédagogiques ou opérationnels, il est fréquent d'utiliser un fichier tableur pour simplifier le calcul des matrices de confusion.

Le fichier contient généralement une première feuille permettant de saisir :

- l'identifiant du point ;

- la classe réelle ;
- la classe prédite.

Une seconde feuille calcule automatiquement :

- la matrice de confusion ;
- les totaux ;
- la précision globale.

Cette approche facilite grandement l'analyse des résultats et permet de se concentrer davantage sur l'interprétation scientifique des erreurs de classification.

Chapitre 33 — Peut-on distinguer les algues vertes des algues rouges ?

Une question fréquente en télédétection littorale

Lorsqu'on commence à travailler sur les algues avec des images multispectrales drone, une question revient presque systématiquement :

est-il possible de distinguer automatiquement les algues vertes des algues rouges ?

Intuitivement, cela semble logique. Les algues présentent des couleurs différentes à l'œil humain et possèdent des pigments biologiques distincts. On pourrait donc penser qu'un capteur multispectral est capable de les séparer facilement.

Pourtant, la réalité est beaucoup plus complexe.

Avec un capteur comme celui du DJI Mavic 3 Multispectral, la distinction entre algues vertes et algues rouges reste souvent difficile, parfois même impossible dans certaines situations.

Cela ne signifie pas que les algorithmes sont mauvais.

Le problème provient surtout de la quantité d'information spectrale réellement disponible.

Pourquoi la séparation est difficile

Les algues vertes et les algues rouges possèdent effectivement des pigments différents.

Les algues vertes sont dominées par la chlorophylle, tandis que les algues rouges contiennent d'autres pigments photosynthétiques comme les phycoérythrine.

Ces différences biologiques modifient leur comportement spectral.

Cependant, avec les seules bandes :

- Green ;
- Red ;
- Red Edge ;
- NIR ;

les signatures restent souvent relativement proches.

Le problème principal vient du fait que certaines longueurs d'onde importantes ne sont pas mesurées par le capteur.

En particulier, le drone ne dispose pas :

- d'une véritable bande bleue multispectrale ;
- de bandes SWIR (*Short Wave Infrared*) ;
- ni d'informations hyperspectrales fines.

Or, ce sont souvent ces longueurs d'onde supplémentaires qui permettent de mieux différencier certains pigments biologiques.

Une limite liée au capteur

Cette situation illustre une idée fondamentale en télédétection :

un algorithme ne peut pas inventer une information que le capteur ne mesure pas

Si deux classes possèdent des signatures très proches dans les bandes disponibles, aucun algorithme ne pourra les séparer parfaitement.

Le problème devient alors un problème d'information spectrale et non un problème de puissance de calcul.

Cette distinction est extrêmement importante.

De nombreux débutants pensent qu'il suffit d'utiliser un algorithme "plus intelligent" pour résoudre le problème.

En réalité, la qualité de la séparation dépend d'abord de la qualité et de la richesse des données acquises.

Tester des indices plus spécifiques

Explorer les différences spectrales subtiles

Même si la séparation reste difficile, certaines approches permettent parfois d'améliorer légèrement la discrimination entre différents types d'algues.

L'une des premières stratégies consiste à tester plusieurs indices spectraux.

Le NDRE, par exemple, peut parfois être plus sensible à certaines variations liées à la chlorophylle :

$$NDRE = \frac{NIR - RedEdge}{NIR + RedEdge}$$

La bande Red Edge se situe précisément dans une zone de transition spectrale très sensible aux variations biologiques.

Dans certains cas, cet indice permet de mieux révéler certaines différences subtiles entre tapis algaux.

Utiliser le GNDVI

Le GNDVI peut également être intéressant dans les milieux humides :

$$GNDVI = \frac{NIR - Green}{NIR + Green}$$

Comme cet indice exploite la bande verte, il peut parfois mieux réagir à certaines différences de pigmentation ou d'humidité.

Cependant, les résultats restent très dépendants :

- des conditions d'acquisition ;
- de l'humidité ;
- de l'état des algues ;
- du contexte local.

Tester des ratios simples

Dans certains projets, des ratios spectraux très simples peuvent parfois révéler des différences locales intéressantes.

Par exemple :

$$\frac{RedEdge}{Red}$$

ou encore des différences simples comme :

Green - Red

et :

RedEdge - Red

Ces approches peuvent parfois amplifier de faibles variations spectrales invisibles dans les indices classiques.

Les indices personnalisés

Une approche souvent plus intéressante

Dans les environnements complexes comme l'estran, les indices personnalisés deviennent souvent plus pertinents que les indices standards.

L'idée consiste à combiner plusieurs bandes ou plusieurs indices afin d'amplifier certaines différences spectrales faibles.

Par exemple, un indice combinant :

$$\left(\frac{NIR}{RedEdge} \right) \times (Green - Red)$$

peut parfois renforcer certaines variations locales entre différents types de surfaces biologiques.

L'objectif n'est pas de créer une "formule magique", mais d'explorer différentes combinaisons capables de mieux révéler certaines signatures.

Cette démarche reste expérimentale et dépend fortement du site étudié.

L'approche multi-indices

Comparer plusieurs représentations simultanément

Une autre stratégie très efficace consiste à analyser plusieurs indices simultanément.

Plutôt que de chercher un unique indice capable de tout résoudre, il peut être plus intéressant d'observer ensemble :

- le NDVI ;
- le NDRE ;

- le GNDVI.

Certaines zones peuvent apparaître différemment selon l'indice utilisé.

Cette approche multi-indices permet souvent de mieux comprendre les comportements spectraux des différentes surfaces.

Ajouter davantage d'informations à la classification

Les indices comme nouvelles variables

Dans une classification supervisée, les indices peuvent également être ajoutés directement dans le *Band Set* de SCP.

Autrement dit, la classification ne repose plus uniquement sur les bandes originales, mais aussi sur des indices calculés.

On obtient alors une classification multidimensionnelle utilisant :

- les bandes spectrales ;
- les indices ;
- éventuellement d'autres variables dérivées.

Cette approche augmente parfois la capacité de séparation entre classes proches.

Le rôle des algorithmes

Certains modèles gèrent mieux les signatures proches

Même si les signatures restent similaires, certains algorithmes de classification sont plus performants que d'autres pour gérer des séparations complexes.

Les approches comme :

- Random Forest ;
- SVM (*Support Vector Machine*) ;

parviennent parfois à mieux exploiter de faibles différences spectrales.

Cependant, il est essentiel de comprendre leurs limites.

Ces algorithmes n'inventent pas de nouvelle information.

Ils exploitent simplement plus efficacement l'information déjà présente dans les données.

Si les signatures sont réellement trop proches, même les meilleurs algorithmes atteindront leurs limites.

L'importance de l'information spatiale

Une approche souvent sous-estimée

La distinction entre algues ne repose pas uniquement sur la couleur ou la signature spectrale.

Certaines différences apparaissent également dans :

- la texture ;
- la rugosité ;
- l'organisation spatiale ;
- la structure des tapis algaux.

Ces informations peuvent parfois être exploitées avec :

- des analyses de texture raster ;
- des modèles de surface (*DSM*) ;
- des indicateurs de rugosité ;
- des analyses spatiales avancées.

Dans certains cas, cette approche spatiale apporte plus d'information que les bandes spectrales elles-mêmes.

La véritable solution scientifique

Quand les bandes multispectrales ne suffisent plus

Si l'objectif scientifique est de distinguer précisément plusieurs types d'algues, il devient souvent nécessaire d'utiliser des capteurs plus riches spectralement.

Les approches les plus performantes utilisent généralement :

- des bandes bleues spécialisées ;
- des bandes SWIR ;
- des capteurs hyperspectraux.

Ces systèmes mesurent beaucoup plus finement les variations spectrales liées aux pigments biologiques.

Ils permettent donc une discrimination beaucoup plus précise entre espèces ou groupes d'algues.

A retenir

Lorsqu'on ne parvient pas à distinguer correctement deux classes, le problème ne vient pas forcément de l'algorithme.

Très souvent, il s'agit simplement d'un manque d'information spectrale disponible dans les données.

En télédétection, on ne peut pas distinguer ce que le capteur ne mesure pas.

Cette idée constitue l'un des principes fondamentaux de toute analyse multispectrale scientifique.

Chapitre 34 — Peut-on utiliser la bande bleue du RGB ?

Une question logique... mais plus complexe qu'elle n'en a l'air

Lorsque l'on travaille avec le DJI Mavic 3 Multispectral, une remarque revient souvent : le drone possède déjà une caméra RGB classique contenant une bande bleue. Il semble donc naturel de vouloir utiliser cette information supplémentaire pour améliorer la discrimination des algues.

Cette idée est pertinente, car la bande bleue peut effectivement contenir des informations intéressantes sur certains pigments biologiques ou sur certaines différences de surface.

Cependant, l'utilisation de cette bande pose un problème important :

la bande bleue issue du RGB n'est pas une véritable bande multispectrale calibrée en réflectance

Cette distinction change profondément la manière dont les données peuvent être utilisées.

Ce que contient réellement l'orthomosaïque RGB

L'orthomosaïque RGB produite par des logiciels comme Pix4D contient bien trois canaux correspondant :

- au rouge ;
- au vert ;
- au bleu.

La bande bleue existe donc réellement dans les données produites par la photogrammétrie.

Techniquement, il est tout à fait possible de l'extraire et de l'utiliser dans QGIS.

Extraire la bande bleue dans QGIS

Dans QGIS, la récupération de cette bande reste relativement simple.

Après avoir chargé l'orthomosaique RGB, l'outil :

Raster → Extraction → Extraction de bande

permet d'extraire individuellement chaque canal.

Dans la plupart des cas :

- la bande 1 correspond au rouge ;
- la bande 2 correspond au vert ;
- la bande 3 correspond au bleu.

Il reste néanmoins indispensable de vérifier cet ordre, car certains exports peuvent utiliser des organisations différentes.

Le raster extrait peut ensuite être enregistré sous un nom comme :

blue_rgb.tif

Vérifier la cohérence des valeurs

Une fois la bande extraite, il est important d'inspecter les valeurs avec des outils comme *Value Tool*.

Cette étape permet de vérifier que les valeurs évoluent de manière cohérente selon les surfaces observées.

Les zones claires doivent produire des valeurs plus élevées, tandis que les surfaces sombres ou ombragées doivent présenter des valeurs plus faibles.

Cette vérification est essentielle avant toute tentative d'intégration dans un workflow multispectral.

Le problème fondamental : le RGB n'est pas calibré

Une différence majeure avec les bandes multispectrales

La principale difficulté vient du fait que les bandes RGB issues de l'orthomosaique ne correspondent pas à des mesures physiques calibrées en réflectance.

Les bandes multispectrales du drone sont conçues pour produire des données scientifiquement exploitables et comparables entre acquisitions.

Elles reposent sur :

- une calibration radiométrique ;
- des bandes spectrales étroites ;
- des mesures physiques plus stables.

À l'inverse, l'image RGB est principalement conçue pour la visualisation.

Les valeurs dépendent fortement :

- de l'exposition ;
- des conditions d'éclairage ;
- des traitements automatiques de l'appareil ;
- de la balance des blancs ;
- des corrections internes de l'image.

Autrement dit, le RGB produit une image "visuellement agréable", mais pas nécessairement une mesure radiométrique rigoureuse.

Pourquoi cela pose problème

Lorsque l'on combine des bandes multispectrales calibrées avec une bande RGB non calibrée, on mélange des données de nature très différente.

Cette hétérogénéité peut introduire plusieurs problèmes.

La classification peut devenir biaisée si l'information RGB domine artificiellement certaines signatures spectrales.

Des incohérences radiométriques peuvent également apparaître entre les bandes.

Dans certains cas, l'ajout du bleu améliore visuellement la séparation des classes sans réellement améliorer la validité scientifique du modèle.

Cette situation peut être trompeuse.

Peut-on malgré tout utiliser cette bande ?

Oui... mais avec prudence

Malgré ces limites, la bande bleue peut tout de même être utile dans certaines situations.

Elle peut apporter une information spectrale supplémentaire absente des bandes multispectrales classiques du drone.

Dans certains projets exploratoires, cette bande permet parfois d'améliorer légèrement la séparation entre certaines surfaces algales.

Elle peut également être intéressante pour tester des indices personnalisés ou certaines classifications expérimentales.

Cependant, il est important de comprendre que cette approche reste essentiellement empirique.

Les cas où cette approche peut être utile

L'utilisation de la bande bleue peut être pertinente lorsqu'on cherche principalement :

- à améliorer visuellement certaines séparations ;
- à tester des hypothèses exploratoires ;
- à rechercher des différences spectrales faibles ;
- à enrichir temporairement un workflow expérimental.

Dans certains cas, elle peut effectivement apporter une information complémentaire utile.

Les situations où il faut rester prudent

En revanche, cette approche devient beaucoup plus problématique lorsqu'on cherche à produire des analyses quantitatives rigoureuses.

L'utilisation du bleu RGB doit être évitée pour :

- les estimations précises de biomasse ;
- les comparaisons temporelles quantitatives ;
- les modèles physiques de réflectance ;
- les analyses nécessitant une forte rigueur radiométrique.

Dans ces situations, les données non calibrées peuvent produire des interprétations scientifiquement fragiles.

Les bonnes pratiques

Normaliser les valeurs

Lorsqu'on souhaite malgré tout utiliser la bande bleue RGB, il est préférable de normaliser les valeurs afin de limiter certaines différences d'échelle.

Une normalisation simple peut par exemple être réalisée avec :

$$\frac{Blue - min}{max - min}$$

Cette opération ne transforme pas la bande en véritable réflectance, mais elle peut réduire certains effets liés aux différences de dynamique radiométrique.

Comparer les résultats

Il est également conseillé de tester les classifications :

- avec la bande bleue ;
- puis sans la bande bleue.

Comparer les résultats permet d'évaluer si l'ajout du bleu apporte réellement une amélioration ou simplement du bruit supplémentaire.

Cette démarche comparative est essentielle dans les workflows expérimentaux.

Garder une approche critique

L'utilisation de la bande bleue RGB doit toujours être interprétée avec prudence.

Le bleu apporte effectivement une information supplémentaire, mais cette information ne possède pas le même niveau de rigueur radiométrique que les bandes multispectrales calibrées.

Il s'agit donc davantage :

d'une amélioration empirique que d'une mesure scientifique stricte

Cette nuance est importante pour éviter de surinterpréter les résultats obtenus.

A retenir

Oui, il est possible d'utiliser la bande bleue issue de l'orthomosaïque RGB dans un workflow multispectral sous QGIS.

Cependant, cette bande ne correspond pas à une véritable réflectance calibrée.

Elle peut enrichir certaines analyses exploratoires, mais elle doit être utilisée avec prudence et toujours interprétée de manière critique.

En télédétection scientifique, la qualité des données est souvent aussi importante que la puissance des algorithmes utilisés.

Chapitre 35 — Les trois grandes familles d'algorithmes de classification

Comprendre la logique des classifications supervisées

Dans Semi-Automatic Classification Plugin, plusieurs algorithmes de classification supervisée sont disponibles. Lorsqu'on débute, il peut être tentant de chercher immédiatement "le meilleur algorithme".

Pourtant, cette approche est souvent trompeuse.

En télédétection, il n'existe pas d'algorithme universel capable de résoudre parfaitement tous les problèmes. Chaque méthode repose sur une logique différente et répond à des situations particulières.

Comprendre cette logique est beaucoup plus important que mémoriser une liste d'algorithmes.

De manière générale, les méthodes de classification utilisées dans SCP peuvent être regroupées en trois grandes familles :

- les méthodes classiques statistiques ;
- les méthodes spectrales ;
- les approches de machine learning.

Chacune de ces familles possède ses avantages, ses limites et ses domaines d'application.

Les méthodes classiques : comparer les pixels aux signatures moyennes

Une approche statistique historique

Les méthodes classiques constituent les approches historiques de la télédétection supervisée.

Elles reposent sur une idée relativement simple : comparer les valeurs spectrales d'un pixel aux signatures moyennes des différentes classes définies par les ROI.

Autrement dit, l'algorithme cherche à déterminer à quelle classe un pixel ressemble le plus.

Parmi les méthodes les plus connues figurent :

- le *Maximum Likelihood* ;
- le *Minimum Distance* ;
- le *Nearest Neighbor*.

Ces méthodes ont été largement utilisées pendant plusieurs décennies avant l'arrivée du machine learning moderne.

Le principe du Maximum Likelihood

Le *Maximum Likelihood* — ou maximum de vraisemblance — constitue probablement l'algorithme statistique le plus célèbre en télédétection.

Cette méthode ne se contente pas de comparer des moyennes. Elle modélise également la distribution statistique des signatures spectrales de chaque classe.

L'algorithme estime alors la probabilité qu'un pixel appartienne à une classe donnée.

Autrement dit, il cherche :

la classe la plus probable pour chaque pixel

Cette approche produit souvent des résultats robustes lorsque les classes sont bien définies et que les signatures spectrales restent relativement stables.

Le principe du Minimum Distance

Le *Minimum Distance* fonctionne de manière encore plus simple.

Chaque classe possède une signature spectrale moyenne, et l'algorithme calcule simplement la distance entre le pixel analysé et ces signatures moyennes.

Le pixel est ensuite affecté à la classe la plus proche.

Cette méthode est rapide, facile à comprendre et peu coûteuse en calcul.

Cependant, elle reste relativement sensible au bruit et aux variations internes des classes.

Les avantages des méthodes classiques

Les méthodes statistiques possèdent plusieurs qualités importantes.

Elles sont généralement :

- simples à mettre en œuvre ;
- robustes ;
- rapides ;
- faciles à interpréter.

Leur fonctionnement reste relativement transparent, ce qui facilite l'analyse scientifique des résultats.

Dans de nombreux projets environnementaux, elles produisent encore aujourd'hui des résultats tout à fait satisfaisants.

Leurs limites

Ces méthodes reposent toutefois sur certaines hypothèses simplificatrices.

Elles supposent souvent que les signatures spectrales suivent des distributions relativement simples et homogènes.

Dans des environnements complexes comme l'estran, cette hypothèse devient parfois problématique.

Les classes littorales présentent souvent :

- une forte variabilité ;
- des mélanges de surfaces ;
- des signatures partiellement recouvertes.

Les méthodes classiques deviennent alors sensibles :

- à la qualité des ROI ;
 - au bruit radiométrique ;
 - aux variations d'humidité ;
 - aux pixels mixtes.
-

Les méthodes spectrales : analyser la forme des signatures

Une approche plus “physique”

Les méthodes spectrales reposent sur une logique différente.

Au lieu de comparer directement les valeurs absolues des pixels, elles analysent la forme des signatures spectrales.

L’algorithme le plus connu de cette famille est :

Spectral Angle Mapping (SAM)

Cette méthode est particulièrement intéressante en télédétection multispectrale et hyperspectrale.

Le principe du SAM

Le SAM considère chaque signature spectrale comme un vecteur dans un espace multidimensionnel.

L’algorithme mesure alors l’angle entre :

- la signature du pixel ;
- la signature de référence.

Plus l’angle est faible, plus les signatures sont considérées comme similaires.

Cette approche présente une propriété importante :

elle est relativement robuste aux variations globales d’illumination

Autrement dit, deux signatures présentant la même forme mais des intensités légèrement différentes resteront proches dans l’espace spectral.

Les avantages du SAM

Le SAM peut être particulièrement utile dans des conditions d’éclairage variables.

En milieu littoral, cette propriété devient intéressante car les conditions lumineuses changent rapidement :

- ombres ;
- humidité ;
- reflets ;
- variations d’illumination.

L'analyse de la forme des signatures permet parfois de mieux résister à ces perturbations.

Les limites du SAM

Le principal défaut du SAM vient du fait qu'il privilégie la forme des signatures plutôt que leur intensité absolue.

Deux surfaces présentant des signatures de forme similaire mais des niveaux de réflectance différents peuvent alors être considérées comme proches.

Dans certaines situations, cette propriété réduit la capacité de discrimination entre certaines classes.

Le SAM devient donc moins performant lorsque les différences importantes résident principalement dans l'intensité du signal.

Le machine learning : apprendre des relations complexes

Une approche moderne basée sur l'apprentissage

Les approches de machine learning constituent aujourd'hui les méthodes les plus performantes dans de nombreux projets de télédétection.

Contrairement aux méthodes classiques, ces algorithmes ne reposent pas sur des hypothèses statistiques simples.

Ils apprennent directement des relations complexes entre les variables spectrales.

Dans SCP, plusieurs méthodes de machine learning peuvent être utilisées, notamment :

- Random Forest ;
- SVM (*Support Vector Machine*) ;
- Multi-Layer Perceptron.

Le Random Forest

Le Random Forest est aujourd'hui l'un des algorithmes les plus populaires en télédétection environnementale.

Il repose sur un ensemble d'arbres de décision combinés entre eux.

Chaque arbre apprend une partie des relations entre les variables, puis l'ensemble du modèle vote pour produire la classification finale.

Cette approche est particulièrement robuste face :

- au bruit ;
- aux variables nombreuses ;
- aux signatures complexes ;
- aux interactions non linéaires.

Dans les milieux intertidaux, le Random Forest produit souvent d'excellents résultats.

Le SVM

Le *Support Vector Machine* cherche à séparer les classes dans un espace multidimensionnel en construisant des frontières optimales entre elles.

Cette méthode peut être extrêmement performante lorsque les classes présentent des séparations complexes.

Cependant, elle demande souvent davantage de réglages et peut devenir sensible aux paramètres utilisés.

Les avantages du machine learning

Les approches de machine learning sont capables de gérer :

- des signatures complexes ;
- des distributions non linéaires ;
- des données bruitées ;
- des interactions multiples entre variables.

Elles exploitent souvent beaucoup mieux les données que les méthodes classiques.

Dans les projets complexes, elles produisent généralement les meilleures performances.

Les limites du machine learning

Ces algorithmes présentent toutefois plusieurs limites importantes.

Leur fonctionnement devient souvent difficile à interpréter scientifiquement.

Contrairement aux méthodes classiques, il est parfois compliqué de comprendre précisément pourquoi un pixel est classé dans une catégorie donnée.

Ces modèles peuvent également devenir sensibles :

- à la qualité des ROI ;
- au déséquilibre des classes ;
- au bruit dans les données ;
- au sur-apprentissage (*overfitting*).

Le sur-apprentissage apparaît lorsque le modèle apprend trop précisément les données d'entraînement et perd sa capacité de généralisation.

Comparer les trois familles

Trois philosophies différentes

Les trois familles d'algorithmes reposent finalement sur des logiques très différentes.

Les méthodes classiques privilégient la simplicité statistique et l'interprétabilité.

Les méthodes spectrales cherchent à exploiter directement la structure physique des signatures.

Le machine learning privilégie la capacité d'apprentissage et la performance.

Chaque approche répond donc à des problématiques différentes.

Une idée essentielle

Il n'existe pas de "meilleur" algorithme universel.

Le choix dépend toujours :

- du type de données ;
- du capteur ;
- de la qualité des ROI ;
- de la complexité du site ;
- des objectifs scientifiques.

Dans certains cas, une méthode simple comme le Maximum Likelihood peut produire d'excellents résultats.

Dans d'autres situations, un Random Forest sera beaucoup plus performant.

L'essentiel n'est donc pas de choisir l'algorithme "le plus moderne", mais celui qui correspond le mieux au problème étudié.

A retenir

En télédétection, la qualité des données et des ROI influence souvent davantage les résultats que le choix de l'algorithme lui-même.

Un excellent algorithme ne compensera jamais :

- des signatures spectrales incohérentes ;
- des ROI mal construites ;
- ou des données de mauvaise qualité.

La classification supervisée reste avant tout une démarche scientifique fondée sur :

- la compréhension du milieu ;
- la qualité des échantillons ;
- et l'interprétation critique des résultats.

Chapitre 36 — Introduction à Random Forest en télédétection

Pourquoi Random Forest est devenu si populaire

Depuis plusieurs années, l'algorithme Random Forest occupe une place centrale dans les projets de télédétection environnementale.

On le retrouve aujourd'hui dans des domaines extrêmement variés :

- cartographie de l'occupation du sol ;
- agriculture de précision ;
- suivi forestier ;
- analyse urbaine ;
- classification hyperspectrale ;
- télédétection drone ;
- suivi des algues vertes.

Cette popularité s'explique par une combinaison rare de qualités : Random Forest est à la fois performant, relativement simple à utiliser et robuste face aux données complexes.

Dans les milieux littoraux, où les signatures spectrales sont souvent bruitées et très variables, cet algorithme donne fréquemment de meilleurs résultats que les méthodes statistiques classiques.

Qu'est-ce que Random Forest ?

Une “forêt” d’arbres de décision

Le nom Random Forest signifie littéralement :

forêt aléatoire

L’algorithme repose sur un ensemble de nombreux arbres de décision travaillant ensemble.

Un arbre de décision fonctionne comme une suite de questions logiques.

Par exemple, un arbre peut décider :

- si le NDVI est supérieur à une certaine valeur ;
- si le NIR est élevé ;
- si la texture est homogène ;
- ou si certaines bandes présentent des signatures particulières.

Chaque arbre construit progressivement ses propres règles pour séparer les classes.

Le principe du vote

Dans Random Forest, il n’existe pas un seul arbre, mais souvent des dizaines ou des centaines.

Chaque arbre produit sa propre prédiction pour un pixel donné.

Ensuite, tous les arbres “votent”.

La classe obtenant le plus grand nombre de votes devient la classe finale attribuée au pixel.

Cette approche collective constitue l’une des grandes forces du modèle.

Un arbre individuel peut se tromper, mais l’ensemble de la forêt produit généralement une décision beaucoup plus robuste.

Pourquoi “Random” ?

Le rôle de l’aléatoire

Le terme “random” provient du fait que chaque arbre est construit à partir :

- d’un sous-ensemble aléatoire des données ;
- et d’un sous-ensemble aléatoire des variables.

Autrement dit, tous les arbres ne voient pas exactement les mêmes informations.

Cette diversité réduit fortement le risque qu'un arbre apprenne des erreurs spécifiques présentes dans les données d'entraînement.

L'algorithme devient ainsi :

- plus robuste ;
- moins sensible au bruit ;
- plus stable face aux variations des données.

Pourquoi Random Forest fonctionne bien en télédétection

Gérer des données complexes

Les données de télédétection sont souvent difficiles à analyser.

Les surfaces naturelles présentent :

- des signatures variables ;
- des effets d'humidité ;
- du bruit radiométrique ;
- des mélanges de pixels ;
- des relations non linéaires entre variables.

Les méthodes classiques supposent souvent des comportements statistiques relativement simples.

Random Forest, au contraire, est capable d'apprendre des relations beaucoup plus complexes.

C'est l'une des raisons pour lesquelles il fonctionne particulièrement bien avec :

- les indices spectraux ;
- les données drone ;
- les environnements littoraux ;
- les analyses multisources.

Les variables utilisées par Random Forest

Bien plus que les bandes spectrales

Random Forest peut utiliser un très grand nombre de variables simultanément.

Dans un projet de télédétection, ces variables peuvent inclure :

- les bandes spectrales ;
- le NDVI ;
- le NDRE ;
- le GNDVI ;
- des textures raster ;
- des modèles de hauteur ;
- des informations topographiques.

L'algorithme peut donc apprendre des relations complexes entre plusieurs dimensions des données.

En analyse orientée objet

Dans les approches orientées objet, Random Forest devient encore plus puissant.

Au lieu de classifier uniquement des pixels individuels, l'algorithme peut utiliser des propriétés calculées sur des segments d'image.

Il peut alors exploiter :

- la moyenne du NDVI d'un objet ;
- la texture ;
- la forme ;
- le voisinage ;
- la rugosité ;
- la compacité des segments.

Cette approche est particulièrement intéressante dans les projets drone à très haute résolution.

Les paramètres importants

Le nombre d'arbres

Le principal paramètre de Random Forest correspond au nombre d'arbres utilisés dans la forêt.

Plus le nombre d'arbres augmente, plus le modèle devient généralement stable.

Cependant, augmenter indéfiniment le nombre d'arbres n'améliore pas toujours les résultats.

Au-delà d'un certain seuil, les gains deviennent faibles tandis que les temps de calcul augmentent fortement.

Dans de nombreux projets de télédétection, quelques centaines d'arbres suffisent largement.

Des valeurs comprises entre :

100 et 500 arbres

donnent souvent de très bons résultats.

La profondeur des arbres

La profondeur maximale des arbres contrôle la complexité du modèle.

Des arbres très profonds apprennent des règles extrêmement détaillées.

Cela peut améliorer les performances sur les données d'entraînement, mais augmente également le risque de sur-apprentissage.

Le nombre de variables utilisées

Chaque arbre n'utilise qu'une partie des variables disponibles.

Ce mécanisme introduit davantage de diversité dans la forêt et améliore généralement la robustesse globale du modèle.

Les limites de Random Forest

La qualité des ROI reste fondamentale

Même si Random Forest est très performant, il possède plusieurs limites importantes.

La première est fondamentale :

la qualité des données d'apprentissage reste plus importante que l'algorithme lui-même

Si les ROI sont mauvaises, bruitées ou mal représentatives, le modèle apprendra des signatures incorrectes.

Dans ce cas, même un algorithme très puissant produira une mauvaise classification.

Le déséquilibre des classes

Random Forest peut également devenir sensible au déséquilibre des classes.

Si une classe domine fortement les données d'apprentissage, le modèle aura tendance à favoriser cette classe.

Par exemple, si la majorité des pixels correspondent à de l'eau, le modèle peut produire une excellente précision globale tout en détectant très mal les algues.

Ce problème est extrêmement fréquent en télédétection.

Le sur-apprentissage

Le sur-apprentissage — ou *overfitting* — correspond à une autre limite importante.

Le modèle apprend alors trop précisément les données d'entraînement et perd sa capacité à généraliser sur de nouvelles données.

Autrement dit, il devient excellent sur les ROI d'apprentissage mais moins performant sur le reste de l'image.

Ce phénomène apparaît souvent lorsque :

- les ROI sont trop peu nombreuses ;
- les arbres deviennent trop complexes ;
- les données sont trop bruitées.

Évaluer un modèle Random Forest

Nous verrons cette évaluation en détail dans le chapitre 41.

Pourquoi la précision globale ne suffit pas

Lorsqu'on évalue un modèle, il est tentant de regarder uniquement la précision globale.

Pourtant, cet indicateur peut être très trompeur.

Imaginons un site contenant :

- 950 pixels d'eau ;
- 50 pixels d'algues.

Un modèle qui classe tous les pixels comme "eau" obtiendra :

95 % de précision globale

et pourtant :

aucune algue ne sera détectée

Cette situation montre pourquoi une seule métrique ne suffit jamais.

La précision par classe

Une question essentielle

La précision par classe répond à la question suivante :

quand le modèle prédit une classe, a-t-il souvent raison ?

Par exemple, une précision de :

0.90 pour les algues

signifie que 90 % des objets classés comme “algues” sont réellement des algues.

Cet indicateur mesure donc la fiabilité des prédictions produites.

Le rappel (*Recall*)

Le modèle retrouve-t-il réellement les objets ?

Le rappel mesure la capacité du modèle à retrouver tous les objets appartenant réellement à une classe.

Par exemple, un rappel de :

0.40 pour les algues

signifie que seulement 40 % des véritables zones algales ont été détectées.

Le modèle oublie donc une grande partie des surfaces concernées.

En télédétection littorale, ce problème est très fréquent pour :

- les algues ;
- les ombres ;
- les substrats humides ;
- les classes rares.

Le F-score

Combiner précision et rappel

Le F-score combine :

- la précision ;
- le rappel.

Il permet d’obtenir une vision plus équilibrée des performances du modèle.

Un bon F-score indique généralement que le modèle :

- détecte correctement la classe ;
- tout en produisant peu de faux positifs.

L'indice Kappa

Une métrique très importante en télédétection

L'indice Kappa constitue un indicateur particulièrement utile en classification environnementale.

Contrairement à la précision globale, il tient compte du hasard.

Il mesure donc la qualité réelle du modèle.

Un modèle peut afficher :

99 % de précision globale

tout en possédant :

un Kappa faible

Cette situation indique généralement que le modèle favorise fortement une classe dominante.

Un exemple typique en télédétection

Quand l'eau domine les résultats

Dans les milieux littoraux, les classes dominantes comme l'eau obtiennent souvent d'excellentes performances.

À l'inverse, les classes plus complexes comme les algues produisent :

- des précisions plus faibles ;
- des rappels plus bas ;
- davantage de confusions.

Cette situation est normale.

Les signatures des algues sont souvent plus proches :

- des substrats humides ;
- des biofilms ;
- des surfaces ombragées.

L'interprétation des métriques doit donc toujours être réalisée classe par classe.

Workflow typique avec Random Forest

Une démarche complète

Dans un projet de télédétection, l'utilisation de Random Forest suit généralement plusieurs étapes :

- préparation des données ;
- création des ROI ;
- entraînement du modèle ;
- classification ;
- validation des résultats.

Chaque étape influence directement la qualité finale du modèle.

A retenir

Random Forest est un algorithme extrêmement puissant, mais il ne remplace jamais une bonne méthodologie scientifique.

La qualité des résultats dépend avant tout :

- de la qualité des données ;
- de la qualité des ROI ;
- de la validation ;
- et de l'interprétation critique des résultats.

En télédétection, un modèle performant ne se résume pas à un pourcentage élevé de précision globale.

Il doit être :

- interprété ;
- validé ;
- analysé classe par classe ;
- et replacé dans son contexte écologique.

Quatrième partie — Comparaison de méthodes de classification

Chapitre 37 — Mise en œuvre de classifications alternatives

Aller au-delà des classifications pixel par pixel

Dans les chapitres précédents, nous avons principalement travaillé avec des approches de classification supervisée reposant sur l'analyse directe des pixels. Ces méthodes constituent une excellente introduction à la télédétection multispectrale, car elles permettent de comprendre le rôle des signatures spectrales, l'importance des ROI et la logique générale des algorithmes de classification.

Cependant, les environnements littoraux présentent une complexité particulière. Les images acquises avec le DJI Mavic 3 Multispectral possèdent une résolution extrêmement fine, souvent de quelques centimètres seulement. À cette échelle, chaque détail du terrain devient visible : micro-topographie, humidité du substrat, structures des tapis algaux, petites flaques résiduelles ou encore variations de texture.

Cette richesse d'information ouvre la voie à des approches plus avancées que les simples classifications pixel par pixel.

Pourquoi comparer plusieurs méthodes ?

En télédétection, il n'existe pas de méthode universellement meilleure que toutes les autres. Chaque approche possède ses propres avantages, ses limites et ses hypothèses de fonctionnement.

Certaines méthodes fonctionnent très bien sur des images satellitaires de résolution moyenne, mais deviennent moins efficaces sur des images drone très détaillées. À l'inverse, certaines approches spécialement conçues pour les images haute résolution deviennent particulièrement intéressantes dans les milieux littoraux.

L'objectif de cette partie est donc de comparer différentes stratégies de classification afin de mieux comprendre leur comportement et leur capacité à représenter les surfaces observées sur l'estran.

Nous allons explorer des approches statistiques classiques, des méthodes orientées objet, des techniques de machine learning ainsi que certaines approches hybrides combinant plusieurs logiques d'analyse.

Les outils utilisés

Les traitements seront réalisés principalement avec QGIS, qui servira d'environnement principal de travail et de visualisation.

Nous utiliserons également Orfeo ToolBox, souvent abrégé OTB, spécialisé dans les traitements avancés d'images de télédétection.

Enfin, certaines analyses seront réalisées avec GRASS GIS, particulièrement puissant pour les traitements spatiaux et morphologiques.

Ces outils complémentaires permettront d'explorer différentes philosophies de classification.

Chapitre 38 — Présentation de Orfeo ToolBox

Un outil conçu pour la télédétection avancée

Orfeo ToolBox est une bibliothèque open source spécialisée dans le traitement des images de télédétection. Développé initialement dans le cadre du programme spatial français ORFEO, cet outil est aujourd'hui largement utilisé dans les projets scientifiques liés aux satellites, aux drones et aux analyses multispectrales.

Contrairement à certains logiciels SIG plus généralistes, OTB a été conçu spécifiquement pour manipuler des données de télédétection complexes. Il propose ainsi des outils avancés de segmentation, de classification, de machine learning et d'analyse d'images haute résolution.

Cette spécialisation le rend particulièrement adapté aux données issues des drones multispectraux.

Pourquoi OTB est intéressant pour les images drone

Les images drone très haute résolution contiennent énormément de détails spatiaux. Cette richesse d'information constitue à la fois une force et une difficulté.

Avec une résolution de quelques centimètres, chaque variation locale devient visible : humidité du sable, petites ombres, textures fines ou micro-structures des tapis algaux.

Les méthodes pixel par pixel classiques deviennent alors parfois instables ou très bruitées.

OTB propose justement des approches mieux adaptées à ce type de données, notamment grâce aux méthodes orientées objet et aux algorithmes avancés de segmentation.

Intégration dans QGIS

L'un des grands avantages d'OTB est son intégration dans la boîte à outils de traitements de QGIS.

Les algorithmes peuvent être exécutés directement depuis l'interface de QGIS sans utiliser la ligne de commande. Cette intégration facilite énormément les workflows pédagogiques et permet de tester rapidement différentes approches de classification.

Chapitre 39 — Introduction à la segmentation orientée objet

Pourquoi segmenter une image ?

Les classifications pixel par pixel considèrent chaque pixel indépendamment des pixels voisins. Cette logique fonctionne relativement bien sur des images satellitaires de résolution moyenne.

Mais sur des images drone très haute résolution, cette approche atteint rapidement ses limites.

Deux pixels appartenant à une même nappe d'algues peuvent présenter de petites différences liées à l'humidité, à l'éclairage ou à la texture locale. Une classification pixel par pixel peut alors produire une image très bruitée, souvent décrite comme un effet "sel et poivre".

La segmentation orientée objet cherche précisément à réduire ce problème.

Le principe de la segmentation

La segmentation adopte une logique différente de celle des classifications classiques.

Au lieu de classer directement chaque pixel, elle commence par regrouper les pixels voisins présentant des caractéristiques similaires. L'image est ainsi transformée en un ensemble d'objets homogènes.

Ces objets peuvent correspondre à des nappes d'algues, des surfaces sableuses, des zones d'eau ou des secteurs de vase humide.

La classification est ensuite réalisée non plus sur les pixels individuels, mais sur ces objets segmentés.

L'algorithme LargeScaleMeanShift

Dans Orfeo ToolBox, l'un des algorithmes de segmentation les plus utilisés pour les images haute résolution est LargeScaleMeanShift.

Cet algorithme repose simultanément sur la proximité spatiale des pixels et sur leur similarité spectrale.

Autrement dit, les pixels proches les uns des autres et présentant des signatures similaires sont regroupés dans un même objet.

Cette approche permet de créer des zones cohérentes correspondant beaucoup mieux aux structures réellement présentes sur le terrain.

Une méthode particulièrement adaptée au suivi des algues

La segmentation orientée objet devient particulièrement intéressante dans les environnements littoraux.

Les tapis d'algues ne se distinguent pas uniquement par leur signature spectrale. Leur texture, leur continuité spatiale et leur organisation sur l'estran apportent également des informations importantes.

En regroupant les pixels en objets cohérents, la segmentation permet d'intégrer cette dimension spatiale dans l'analyse.

Cette approche produit souvent des résultats plus stables et plus lisibles que les classifications pixel par pixel classiques.

Une nouvelle manière de voir l'image

La segmentation change profondément la logique d'analyse.

Dans une approche classique, chaque pixel est traité individuellement. Avec une approche orientée objet, ce sont désormais les objets segmentés qui deviennent les unités d'analyse.

Chaque objet peut alors posséder des propriétés comme :

- une moyenne spectrale ;
- une texture ;
- une forme ;
- une surface ;
- ou encore une relation avec les objets voisins.

Cette approche se rapproche davantage de la manière dont un humain interprète naturellement une image.

A retenir

Les classifications pixel par pixel restent très utiles, mais elles ne sont pas toujours les plus adaptées aux images drone très haute résolution.

Les approches orientées objet permettent d'intégrer des informations supplémentaires liées à la structure spatiale, à la texture et à l'organisation des surfaces.

Dans les milieux littoraux complexes, cette information supplémentaire peut améliorer fortement la qualité et la stabilité des classifications.

Chapitre 40 — Classification Random Forest avec Orfeo ToolBox

Une méthode devenue incontournable en télédétection

Parmi les méthodes modernes de classification supervisée, le Random Forest occupe aujourd'hui une place centrale dans les projets de télédétection environnementale.

Cet algorithme est largement utilisé pour la cartographie :

- de l'occupation du sol ;
- des cultures agricoles ;

- des zones forestières ;
- des milieux urbains ;
- des habitats littoraux ;
- et des surfaces algales.

Sa popularité s'explique par une combinaison particulièrement intéressante de robustesse, de simplicité d'utilisation et de bonnes performances sur des données complexes.

Dans les environnements littoraux, où les signatures spectrales sont souvent perturbées par l'humidité, les ombres ou les mélanges de surfaces, le Random Forest produit fréquemment des résultats plus stables que les méthodes statistiques classiques.

Le principe général du Random Forest

Le Random Forest appartient à la famille des méthodes de machine learning supervisé.

Son fonctionnement repose sur un ensemble de nombreux arbres de décision travaillant collectivement.

Chaque arbre construit progressivement une série de règles permettant de séparer les différentes classes présentes dans l'image.

Par exemple, un arbre peut apprendre qu'un pixel présentant :

- un NDVI élevé ;
- un NIR important ;
- et une texture homogène ;

correspond probablement à une zone d'algues.

Un autre arbre peut privilégier d'autres combinaisons de variables ou d'autres seuils.

Chaque arbre produit ensuite sa propre prédiction.

Le résultat final correspond au vote majoritaire de l'ensemble des arbres.

Cette logique collective constitue l'une des grandes forces du Random Forest.

Même si certains arbres individuels produisent des erreurs, la combinaison de nombreux arbres permet généralement d'obtenir un modèle beaucoup plus robuste.

Pourquoi cette méthode fonctionne bien en télédétection

Les données de télédétection sont rarement simples.

Les signatures spectrales peuvent être influencées par :

- l'humidité ;
- la présence d'eau ;
- les ombres ;
- la texture ;
- les variations d'éclairage ;
- ou encore les mélanges de surfaces.

Les méthodes classiques reposent souvent sur des hypothèses statistiques relativement simples qui deviennent parfois insuffisantes dans des environnements complexes comme l'estran.

Le Random Forest, au contraire, est capable d'apprendre des relations beaucoup plus complexes entre les variables.

Il gère très bien :

- les données bruitées ;
- les signatures non linéaires ;
- les variables nombreuses ;
- les interactions complexes entre bandes spectrales et indices.

C'est l'une des raisons pour lesquelles il est devenu extrêmement populaire dans les projets utilisant des images multispectrales drone.

Étape 1 — L'apprentissage supervisé

Fournir des exemples au modèle

Comme toutes les méthodes supervisées, le Random Forest doit d'abord apprendre à reconnaître les classes présentes dans l'image.

Cette phase d'apprentissage repose sur les ROI créées précédemment.

L'utilisateur fournit des zones représentatives correspondant aux différentes surfaces à identifier.

Dans le cadre du suivi des algues vertes, ces classes peuvent par exemple correspondre :

- aux tapis d'algues ;
- au sable sec ;
- au sable humide ;

- à la vase ;
- à l'eau ;
- ou à la végétation littorale.

Chaque ROI fournit au modèle des exemples de signatures spectrales associées à une classe particulière.

Le Random Forest apprend ensuite les relations entre ces signatures et les classes correspondantes.

L'importance des données d'apprentissage

La qualité des ROI reste essentielle.

Même un algorithme très performant ne peut pas produire de bons résultats avec des données d'apprentissage mauvaises ou incohérentes.

Si les ROI contiennent :

- des mélanges de surfaces ;
- des pixels bruités ;
- des ombres ;
- ou des signatures ambiguës ;

le modèle apprendra des règles incorrectes.

Cette idée est fondamentale :

la qualité des données d'apprentissage influence souvent davantage les résultats que le choix de l'algorithme lui-même

Étape 2 — Construction du modèle

Une forêt d'arbres de décision

Une fois les données d'apprentissage fournies, le Random Forest construit automatiquement un ensemble d'arbres de décision.

Chaque arbre est créé à partir :

- d'un sous-ensemble aléatoire des données ;
- et d'un sous-ensemble aléatoire des variables disponibles.

Cette diversité joue un rôle très important.

Tous les arbres ne voient pas exactement les mêmes données ni les mêmes combinaisons de variables.

Cela réduit fortement le risque que le modèle apprenne des erreurs spécifiques présentes dans certaines ROI.

La forêt devient ainsi :

- plus robuste ;
- moins sensible au bruit ;
- plus stable face aux variations des données.

Les variables utilisées par le modèle

Le Random Forest peut utiliser simultanément un grand nombre de variables.

Dans un projet multispectral, ces variables peuvent inclure :

- les bandes Green, Red, Red Edge et NIR ;
- des indices comme le NDVI ou le NDRE ;
- des textures raster ;
- des modèles de surface ;
- des informations spatiales supplémentaires.

Cette capacité à combiner plusieurs dimensions des données constitue l'une des grandes forces du modèle.

Étape 3 — Production de la carte classifiée

Appliquer le modèle à toute l'image

Une fois le modèle entraîné, le Random Forest peut être appliqué à l'ensemble de l'image.

Chaque pixel — ou chaque objet segmenté dans une approche orientée objet — est analysé par l'ensemble des arbres de la forêt.

Chaque arbre propose une classe, puis le système de vote détermine la classe finale attribuée.

Le résultat produit est une carte classifiée dans laquelle chaque pixel reçoit une catégorie :

- algues ;
- sable ;
- eau ;
- vase ;
- végétation ;
- ou toute autre classe définie pendant l'apprentissage.

Les probabilités de confiance

Dans certains cas, le modèle peut également produire des cartes de probabilité.

Ces cartes indiquent le niveau de confiance associé à chaque classification.

Une zone peut par exemple être classée comme “algues” avec :

- une forte certitude ;
- ou au contraire une probabilité plus faible.

Ces informations sont très utiles pour identifier :

- les zones ambiguës ;
 - les signatures incertaines ;
 - les secteurs nécessitant une validation supplémentaire.
-

Une méthode puissante... mais pas magique

Les limites du Random Forest

Même si le Random Forest est très performant, il possède certaines limites importantes.

Le modèle reste fortement dépendant :

- de la qualité des ROI ;
- de la représentativité des classes ;
- de la qualité des données d’entrée.

Il peut également devenir sensible :

- aux classes déséquilibrées ;
- au sur-apprentissage ;
- aux signatures spectrales trop proches.

Par exemple, si les signatures des algues humides et du sable humide sont très similaires, même un modèle performant peut produire des confusions importantes.

A retenir

Le Random Forest est aujourd’hui l’un des algorithmes les plus efficaces pour la classification d’images multispectrales.

Sa robustesse et sa capacité à gérer des données complexes expliquent sa popularité en télédétection environnementale.

Cependant, cet algorithme ne remplace jamais :

- une bonne préparation des données ;
- des ROI de qualité ;
- une validation rigoureuse ;
- et une interprétation critique des résultats.

En télédétection, la puissance du modèle dépend toujours de la qualité des informations utilisées pour l'entraîner.

Chapitre 41 — Évaluer un modèle Random Forest : comment interpréter les résultats ?

Pourquoi l'évaluation est indispensable

Lorsqu'un modèle de classification produit une carte visuellement convaincante, il peut être tentant de conclure immédiatement que les résultats sont bons.

Pourtant, en télédétection, une classification peut sembler très précise tout en contenant :

- de nombreuses erreurs ;
- des classes mal détectées ;
- des confusions importantes ;
- ou encore un fort sur-apprentissage.

C'est particulièrement vrai dans les environnements littoraux où certaines classes présentent des signatures spectrales très proches.

Les tapis d'algues, les substrats humides, les biofilms ou les zones d'ombre peuvent facilement être confondus.

Une simple observation visuelle de la carte ne suffit donc jamais.

Il devient indispensable d'évaluer le modèle avec plusieurs indicateurs complémentaires.

La précision globale (*Overall Accuracy*)

L'indicateur le plus connu

La précision globale correspond au pourcentage total de bonnes classifications produites par le modèle.

Elle représente la proportion de pixels correctement classés par rapport à l'ensemble des pixels validés.

Cet indicateur est souvent le premier affiché après l'entraînement d'un modèle Random Forest.

De manière générale :

- une précision supérieure à 95 % est souvent considérée comme excellente ;
- entre 85 % et 95 %, les résultats sont généralement très bons ;
- entre 70 % et 85 %, la classification reste correcte ;
- en dessous de 70 %, les performances deviennent faibles.

Cependant, cet indicateur doit être interprété avec beaucoup de prudence.

Pourquoi la précision globale peut être trompeuse

La précision globale peut masquer des problèmes importants lorsqu'une classe domine fortement les données.

Imaginons par exemple une image contenant :

- 950 pixels d'eau ;
- 50 pixels d'algues.

Si le modèle classe absolument tous les pixels comme "eau", il obtiendra :

95 % de précision globale

et pourtant :

aucune algue ne sera détectée

Le modèle semblera donc excellent alors qu'il échoue complètement sur la classe réellement importante.

Cette situation est extrêmement fréquente en télédétection environnementale.

La précision par classe (*Precision*)

Une question fondamentale

La précision par classe répond à une question très importante :

quand le modèle prédit une classe, a-t-il réellement raison ?

Autrement dit, lorsqu'un pixel est classé comme "algues", quelle est la probabilité qu'il corresponde réellement à une zone d'algues ?

Cet indicateur mesure donc la fiabilité des prédictions produites par le modèle.

Interpréter la précision

Une précision proche de :

- 0.95 indique généralement un excellent résultat ;
- 0.80 reste considéré comme bon ;
- 0.60 devient moyen ;
- en dessous de 0.50, les résultats deviennent faibles.

Par exemple, une précision de :

0.90 pour la classe "algues"

signifie que :

90 % des pixels classés comme algues correspondent réellement à des algues

Le modèle produit donc peu de faux positifs pour cette classe.

Le rappel (*Recall*)

Une autre question essentielle

Le rappel mesure une dimension différente du problème.

Il répond à la question suivante :

le modèle retrouve-t-il correctement tous les objets appartenant réellement à cette classe ?

Autrement dit, parmi toutes les véritables zones d'algues présentes sur le terrain, combien ont réellement été détectées par le modèle ?

Comprendre le rappel

Un rappel élevé indique que très peu d'objets réels sont oubliés.

Par exemple :

- un rappel supérieur à 0.90 est excellent ;
- entre 0.70 et 0.90, les résultats restent bons ;
- entre 0.50 et 0.70, la détection devient moyenne ;
- en dessous de 0.50, le modèle oublie une grande partie des objets.

Si :

Recall algues = 0.40

cela signifie que :

le modèle détecte seulement 40 % des véritables zones algales

Autrement dit, une grande partie des algues présentes sur le terrain n'est pas identifiée.

La différence entre précision et rappel

Deux indicateurs complémentaires

La précision et le rappel mesurent deux choses différentes.

Un modèle peut être très prudent et ne classifier comme "algues" que quelques zones très évidentes.

Dans ce cas :

- la précision peut être excellente ;
- mais le rappel très faible.

À l'inverse, un modèle très permissif peut détecter presque toutes les algues mais produire énormément de faux positifs.

Il est donc indispensable d'analyser simultanément ces deux indicateurs.

Le F-score

Combiner précision et rappel

Le F-score combine la précision et le rappel dans une seule métrique.

Il fournit une vision plus équilibrée de la qualité du modèle.

Un bon F-score indique généralement que le modèle :

- détecte correctement les objets ;
- tout en limitant les faux positifs.

En pratique :

- un F-score supérieur à 0.90 est excellent ;
- entre 0.75 et 0.90, les résultats sont bons ;
- entre 0.50 et 0.75, les performances deviennent moyennes ;
- en dessous de 0.50, le modèle reste peu fiable.

Le F-score est particulièrement utile lorsque les classes sont déséquilibrées.

L'indice Kappa

Une métrique très importante en télédétection

L'indice Kappa occupe une place particulière dans les projets de télédétection.

Contrairement à la précision globale, il tient compte du hasard.

Autrement dit, il mesure la qualité réelle du modèle en estimant ce qui dépasse une simple classification aléatoire.

Cette propriété rend le Kappa particulièrement intéressant lorsque certaines classes dominent fortement les données.

Interpréter le Kappa

De manière générale :

- un Kappa supérieur à 0.80 correspond à un excellent modèle ;
- entre 0.60 et 0.80, les résultats sont bons ;
- entre 0.40 et 0.60, la qualité devient moyenne ;
- entre 0.20 et 0.40, les performances restent faibles ;
- en dessous de 0.20, le modèle devient mauvais.

Un exemple typique de piège statistique

Un modèle peut parfois afficher :

- 99 % de précision globale ;
- mais un Kappa de seulement 0.30.

Cette situation indique généralement que le modèle favorise fortement une classe dominante.

Autrement dit, le modèle semble très performant, mais il produit en réalité une classification peu équilibrée et peu informative.

Ce phénomène est extrêmement fréquent lorsque les classes rares sont mal représentées.

Exemple simplifié en télédétection littorale

Quand les algues deviennent difficiles à détecter

Imaginons une classification contenant trois classes :

- eau ;
- sable ;
- algues.

Le modèle peut obtenir d'excellents résultats pour l'eau avec :

- une précision de 0.98 ;
- un rappel de 0.99.

En revanche, les algues peuvent présenter :

- une précision de 0.55 ;
- un rappel de 0.40.

Cela signifie que le modèle reconnaît très bien l'eau, mais détecte mal les surfaces algales.

Cette situation est très fréquente dans les environnements littoraux.

Pourquoi les classes rares sont difficiles

Certaines classes restent naturellement plus difficiles à détecter.

C'est notamment le cas :

- des algues ;
- des ombres ;
- des sols humides ;
- des biofilms.

Plusieurs facteurs expliquent cette difficulté.

Ces classes disposent souvent :

- de peu d'exemples d'apprentissage ;
- de signatures spectrales proches ;
- d'un signal fortement perturbé par le bruit radiométrique.

Même un algorithme performant comme Random Forest peut alors produire des confusions importantes.

Le rôle central de la matrice de confusion

Comprendre les erreurs du modèle

Toutes ces métriques proviennent finalement de la matrice de confusion.

La matrice permet de comprendre précisément :

- quelles classes sont correctement détectées ;
- quelles classes sont confondues ;
- quelles erreurs dominent la classification.

Elle reste l'outil central d'interprétation des performances d'un modèle supervisé.

A retenir

Une forte précision globale ne garantit jamais qu'un modèle est réellement bon.

Un modèle doit toujours être évalué :

- avec plusieurs métriques complémentaires ;
- classe par classe ;
- et à l'aide de la matrice de confusion.

En télédétection environnementale, comprendre les erreurs du modèle est souvent plus important que connaître simplement son pourcentage global de réussite.

Chapitre 42 — Classification avec GRASS GIS

Un environnement historique de la télédétection

Parmi les grands outils open source utilisés en géomatique, GRASS GIS occupe une place particulière.

Développé depuis plusieurs décennies, GRASS GIS possède une longue histoire dans le domaine de l'analyse spatiale et de la télédétection scientifique. Bien avant l'arrivée des bibliothèques modernes de machine learning, il proposait déjà des outils avancés pour :

- la classification supervisée ;
- la segmentation ;
- l'analyse raster ;
- les traitements morphologiques ;
- l'analyse multispectrale.

Aujourd'hui encore, plusieurs de ses algorithmes restent très puissants et largement utilisés dans les workflows scientifiques.

La classification supervisée avec `i.maxlik`

Une approche statistique historique

L'un des outils classiques de GRASS GIS est :

`i.maxlik`

Cette commande réalise une classification supervisée de type *Maximum Likelihood*, ou maximum de vraisemblance.

Cette méthode constitue l'une des grandes approches historiques de la télédétection multispectrale.

Son principe repose sur une idée relativement simple :

chaque classe possède une signature spectrale statistique

Le modèle cherche ensuite à déterminer la probabilité qu'un pixel appartienne à chacune des classes connues.

Le rôle des signatures spectrales

Dans cette approche, chaque classe est décrite statistiquement à partir des ROI d'apprentissage.

Pour chaque classe, le modèle calcule :

- une moyenne spectrale ;
- une variance ;
- et une distribution statistique des valeurs.

Par exemple, une classe correspondant aux algues vertes peut présenter :

- une réflectance relativement élevée dans le proche infrarouge ;
- une absorption plus forte dans le rouge ;
- et une signature particulière dans la zone Red Edge.

Ces caractéristiques deviennent alors la référence utilisée pour classifier les autres pixels de l'image.

L'hypothèse de distribution normale

Le Maximum Likelihood repose sur une hypothèse mathématique importante :

les signatures spectrales suivent une distribution normale

Autrement dit, les valeurs des pixels d'une même classe sont supposées se répartir selon une courbe statistique classique de type gaussien.

Cette hypothèse fonctionne relativement bien lorsque :

- les classes sont homogènes ;
- les données sont correctement calibrées ;
- les signatures restent stables.

Dans ces conditions, le modèle peut produire des résultats robustes et scientifiquement cohérents.

Une classification probabiliste

Contrairement à certaines méthodes plus simples, le Maximum Likelihood ne se contente pas d'attribuer directement une classe.

Pour chaque pixel, il calcule d'abord :

- une probabilité d'appartenance à chaque classe ;
- puis sélectionne la classe la plus probable.

Cette logique probabiliste constitue l'une des forces historiques de la méthode.

Elle permet une interprétation scientifique relativement claire des résultats.

Les avantages du Maximum Likelihood

Le Maximum Likelihood possède plusieurs qualités importantes.

Il s'agit d'une méthode :

- robuste ;
- bien documentée ;
- historiquement très utilisée en télédétection ;
- relativement facile à interpréter scientifiquement.

Dans des environnements relativement homogènes, cette approche peut encore aujourd'hui produire d'excellents résultats.

Ses limites dans les milieux littoraux

Les environnements intertidaux restent cependant particulièrement complexes.

Les signatures spectrales y sont souvent perturbées par :

- l'humidité ;
- les mélanges de surfaces ;
- les ombres ;
- la présence d'eau résiduelle ;
- les variations de texture.

Dans ce contexte, les hypothèses statistiques simples du Maximum Likelihood deviennent parfois insuffisantes.

Les classes hétérogènes ou les signatures fortement mélangées peuvent alors produire :

- des erreurs de classification ;
 - des confusions entre classes ;
 - des résultats moins stables que les approches modernes de machine learning.
-

La segmentation avec `i.segment`

Passer d'une logique pixel à une logique objet

GRASS GIS propose également des outils de segmentation orientée objet, notamment avec la commande :

`i.segment`

Contrairement aux classifications pixel par pixel, cette approche cherche d'abord à regrouper les pixels voisins présentant des caractéristiques similaires.

L'objectif est de créer des objets homogènes correspondant davantage aux structures réelles du terrain.

Dans les images drone très haute résolution, cette approche devient particulièrement intéressante.

Le principe de la segmentation

L'algorithme fusionne progressivement les pixels voisins en fonction de leur similarité spectrale.

Les régions présentant des comportements proches sont regroupées dans un même segment.

Cette logique permet de transformer l'image en un ensemble d'objets cohérents :

- nappes d'algues ;
- surfaces sableuses ;
- zones d'eau ;
- secteurs de vase humide ;
- structures végétalisées.

Le seuil de similarité

L'un des paramètres les plus importants de `i.segment` correspond au seuil de similarité.

Ce paramètre contrôle le niveau de ressemblance nécessaire pour fusionner deux régions.

Lorsque le seuil est faible, l'algorithme produit :

- de nombreux petits objets ;
- une segmentation très détaillée ;

- une forte sensibilité aux variations locales.

À l'inverse, un seuil plus élevé produit :

- des objets plus grands ;
- une fusion plus importante ;
- une image plus simplifiée.

Le choix du seuil influence donc fortement le résultat final.

La taille minimale des segments

Un autre paramètre important concerne la taille minimale des segments.

Cette contrainte permet de limiter la création de très petits objets souvent associés au bruit ou aux variations locales mineures.

Augmenter cette taille minimale permet généralement :

- d'améliorer la lisibilité ;
- de réduire le bruit ;
- de simplifier la classification.

Cependant, un paramètre trop élevé peut également supprimer certains détails fins importants dans les tapis algaux.

L'influence de la résolution des images drone

Les images issues du DJI Mavic 3 Multispectral possèdent une résolution extrêmement élevée.

Cette caractéristique influence fortement le comportement des algorithmes de segmentation.

Plus la résolution est fine :

- plus le nombre de segments augmente ;
- plus les objets deviennent détaillés ;
- plus les temps de calcul augmentent.

La segmentation des images drone nécessite donc souvent des ajustements spécifiques des paramètres afin d'obtenir des objets cohérents sans produire une fragmentation excessive.

Une chaîne complète de traitement

Combiner segmentation et classification

Dans les workflows modernes de télédétection, la segmentation constitue souvent une première étape avant la classification.

Une chaîne complète de traitement peut alors être mise en œuvre.

La première étape consiste à créer des objets homogènes grâce à des outils comme :

- `LargeScaleMeanShift` dans Orfeo ToolBox ;
- ou `i.segment` dans GRASS GIS.

Une fois les objets créés, différentes caractéristiques peuvent être calculées sur chaque segment :

- statistiques spectrales ;
- indices de végétation ;
- textures ;
- formes ;
- rugosité ;
- informations spatiales.

Ces objets peuvent ensuite être classifiés à l'aide :

- d'une méthode statistique classique ;
- ou d'un algorithme de machine learning comme Random Forest.

Le rôle du post-traitement

Après la classification, plusieurs opérations complémentaires permettent souvent d'améliorer les résultats.

Le post-traitement peut inclure :

- des filtres de nettoyage ;
- la fusion de petits objets ;

- la suppression du bruit ;
- ou encore la vectorisation des segments classifiés.

Ces opérations améliorent généralement la lisibilité et la cohérence des cartes finales.

Une évolution des méthodes de télédétection

Du pixel vers l'objet

Les approches proposées par GRASS GIS illustrent bien l'évolution historique des méthodes de télédétection.

Les premières classifications reposaient principalement sur des approches statistiques pixel par pixel comme le Maximum Likelihood.

Avec l'augmentation de la résolution des images drone, les approches orientées objet sont devenues de plus en plus importantes.

Aujourd'hui, les workflows modernes combinent souvent :

- segmentation ;
- extraction de caractéristiques ;
- machine learning ;
- et post-traitements spatiaux.

Cette évolution reflète la complexité croissante des données de télédétection modernes.

A retenir

Les outils de GRASS GIS montrent qu'il existe plusieurs manières d'aborder une classification multispectrale.

Les méthodes statistiques classiques restent utiles pour comprendre les fondements de la télédétection supervisée.

Les approches orientées objet permettent ensuite d'intégrer :

- la structure spatiale ;
- la texture ;
- la forme ;
- et l'organisation des surfaces.

Dans les environnements littoraux complexes, ces approches complémentaires permettent souvent d'obtenir des classifications plus robustes et plus interprétables.

Cinquième partie — Interpréter, valider et comprendre

Chapitre 43 — Les techniques modernes de deep learning

Une nouvelle génération d'approches en télédétection

Après avoir exploré les approches classiques de classification supervisée avec le plugin Semi-Automatic Classification Plugin, puis les méthodes avancées de segmentation et de machine learning avec Orfeo ToolBox et GRASS GIS, il devient intéressant d'introduire une nouvelle famille de méthodes fondées sur l'intelligence artificielle et le deep learning.

Les approches étudiées jusqu'à présent reposaient principalement sur l'analyse des signatures spectrales, sur des modèles statistiques ou sur des algorithmes de machine learning classiques comme Random Forest. Ces méthodes restent extrêmement efficaces dans de nombreux contextes, notamment lorsque les classes sont relativement bien séparées spectralement.

Cependant, certaines situations deviennent beaucoup plus complexes à traiter. Dans les environnements littoraux, les structures spatiales peuvent être très hétérogènes, les textures difficiles à caractériser et les surfaces fortement mélangées. Les variations d'humidité, les ombres, les biofilms ou les algues partiellement immergées produisent souvent des signatures ambiguës.

Les approches de deep learning cherchent précisément à dépasser ces limites.

Pourquoi le deep learning change la télédétection

Aller au-delà des signatures spectrales

Les méthodes classiques de classification utilisent essentiellement les valeurs spectrales des pixels ou des objets.

Les réseaux de neurones profonds adoptent une logique différente.

Au lieu de s'appuyer uniquement sur des bandes spectrales ou des indices calculés manuellement, ils apprennent automatiquement des structures complexes présentes dans les données.

Ces modèles sont capables d'identifier :

- des textures ;
- des motifs spatiaux ;
- des formes ;
- des structures répétitives ;
- ou encore des relations complexes entre pixels voisins.

Autrement dit, le modèle apprend lui-même quelles caractéristiques sont importantes pour distinguer les classes.

Cette capacité constitue l'une des grandes révolutions du deep learning appliqué à la télédétection.

Une approche particulièrement adaptée aux images drone

Les images acquises avec le DJI Mavic 3 Multispectral possèdent une résolution extrêmement fine.

À cette échelle, les textures et les structures spatiales deviennent très importantes.

Les tapis algaux présentent par exemple :

- des motifs particuliers ;
- des structures spatiales spécifiques ;
- des formes caractéristiques ;
- des continuités ou discontinuités locales.

Les réseaux neuronaux sont particulièrement performants pour apprendre ce type d'information spatiale complexe.

C'est l'une des raisons pour lesquelles le deep learning connaît aujourd'hui un développement très rapide dans la télédétection drone.

Les usages actuels du deep learning en télédétection

Une généralisation rapide des applications

Les approches de deep learning sont aujourd'hui utilisées dans de nombreux domaines liés à la télédétection environnementale et géospatiale.

Elles servent notamment pour :

- la détection automatique d'objets ;
- la segmentation sémantique ;
- la cartographie environnementale ;
- l'analyse automatisée des paysages ;
- ou encore le suivi des milieux naturels.

Dans les projets drone, ces méthodes permettent par exemple :

- de détecter automatiquement des infrastructures ;
- d'identifier des surfaces végétalisées ;
- de cartographier des zones dégradées ;
- ou d'automatiser certaines tâches de suivi environnemental.

Le suivi des algues vertes constitue donc un domaine d'application particulièrement pertinent.

Le plugin Deepness

Une passerelle entre QGIS et l'intelligence artificielle

Dans ce contexte, le plugin Deepness représente une solution particulièrement intéressante.

Deepness — *Deep Neural Remote Sensing* — est une extension de QGIS permettant d'utiliser directement des modèles d'intelligence artificielle dans un environnement SIG, sans devoir programmer en Python.

Cette approche joue un rôle important sur le plan pédagogique.

Elle permet d'introduire les utilisateurs SIG aux techniques modernes de deep learning sans nécessiter une maîtrise avancée des frameworks spécialisés comme :

- TensorFlow ;
- ou PyTorch.

Le plugin agit ainsi comme une passerelle entre les workflows géomatiques classiques et les approches modernes d'intelligence artificielle.

Les principales fonctionnalités de Deepness

La segmentation d'images

L'une des principales fonctions du plugin consiste à réaliser de la segmentation d'images.

Le modèle peut classifier automatiquement les pixels d'une image raster afin de produire une carte thématique.

Selon le modèle utilisé, il devient possible d'identifier automatiquement :

- la végétation ;
- les routes ;
- les bâtiments ;
- l'eau ;
- les cultures agricoles ;
- ou différentes surfaces environnementales.

Dans le cadre d'images drone, cette approche permet par exemple de produire automatiquement des cartes d'occupation du sol à partir d'orthophotos très haute résolution.

La détection d'objets

Deepness permet également la détection automatique d'objets géographiques.

Le plugin peut repérer automatiquement certains objets spécifiques présents dans l'image :

- véhicules ;
- infrastructures ;
- bâtiments ;
- équipements techniques ;
- ou autres structures spatiales.

Les résultats peuvent ensuite être exportés sous forme de couches vectorielles directement exploitables dans QGIS.

La régression spatiale

Le plugin ne se limite pas aux classifications discrètes.

Il peut également produire des surfaces continues à partir de modèles IA.

Cette approche permet par exemple d'estimer :

- une biomasse ;
- des paramètres environnementaux ;
- des niveaux de risque ;
- ou des variables physiques dérivées.

Cette logique ouvre des perspectives intéressantes pour le suivi quantitatif des milieux littoraux.

Compatibilité avec les données géospatiales

Un fonctionnement intégré dans QGIS

L'un des grands avantages de Deepness est sa compatibilité directe avec les données géospatiales classiques utilisées dans QGIS.

Le plugin peut fonctionner avec :

- des orthophotos ;
- des images drone ;
- des images Sentinel-2 ;
- des données satellitaires ;
- des flux WMS ou XYZ ;
- ou encore des rasters locaux.

Le traitement peut être appliqué :

- à l'image entière ;
- uniquement à la zone visible ;
- ou à certaines zones spécifiques définies par l'utilisateur.

Cette flexibilité facilite fortement l'intégration dans les workflows SIG existants.

Le format ONNX

Un format standard pour les modèles IA

Deepness utilise des modèles de réseaux neuronaux au format :

ONNX

(*Open Neural Network Exchange*).

Ce format joue un rôle très important dans l'écosystème du deep learning.

Il permet d'utiliser dans QGIS :

- des modèles créés avec PyTorch ;
- des modèles issus de TensorFlow ;
- des modèles pré-entraînés ;
- ou encore des modèles personnalisés convertis au format ONNX.

Cette compatibilité rend le plugin extrêmement flexible.

Le Model Zoo

Une bibliothèque de modèles prêts à l'emploi

Deepness fournit également un *Model Zoo*, c'est-à-dire une bibliothèque de modèles pré-entraînés accessibles directement depuis le plugin.

Ces modèles couvrent différents domaines d'application :

- segmentation de bâtiments ;
- détection de routes ;
- classification d'occupation du sol ;
- segmentation agricole ;
- détection de panneaux photovoltaïques ;
- analyse de dégâts agricoles.

Cette approche facilite énormément l'expérimentation pédagogique, car les utilisateurs peuvent tester rapidement des modèles sans devoir entraîner eux-mêmes un réseau neuronal.

Une interface pensée pour les utilisateurs SIG

Rendre le deep learning accessible

L'interface du plugin a été conçue pour des utilisateurs QGIS ne possédant pas nécessairement une expertise avancée en intelligence artificielle.

L'utilisateur sélectionne simplement :

- le raster d'entrée ;
- le modèle ONNX ;
- la résolution ;
- la taille des tuiles ;
- le recouvrement ;
- et certains paramètres de post-traitement.

Le traitement est ensuite exécuté directement dans QGIS.

Cette simplicité constitue un avantage pédagogique majeur.

Une nouvelle étape dans l'évolution de la télédétection

Des perspectives importantes... mais aussi des limites

Les approches de deep learning ouvrent aujourd'hui des perspectives extrêmement importantes pour la télédétection environnementale.

Elles permettent souvent :

- une meilleure prise en compte des textures ;
- une analyse plus fine des structures spatiales ;
- une meilleure généralisation ;
- et des classifications très performantes sur les images haute résolution.

Cependant, ces méthodes présentent également certaines contraintes importantes :

- besoin de grandes quantités de données ;
- puissance de calcul élevée ;
- difficulté d'interprétation des modèles ;
- dépendance à la qualité des données d'apprentissage.

Le deep learning ne remplace donc pas les approches classiques.

Il constitue plutôt une nouvelle famille d'outils complémentaires dans l'écosystème de la télédétection moderne.

A retenir

Les techniques de deep learning représentent aujourd'hui une évolution majeure de la télédétection drone et environnementale.

Grâce à des outils comme Deepness, il devient possible d'utiliser des réseaux de neurones directement dans QGIS sans programmation complexe.

Ces approches permettent d'exploiter non seulement les signatures spectrales, mais aussi :

- les textures ;
- les formes ;
- les motifs spatiaux ;
- et les structures complexes des paysages.

Dans les environnements littoraux, cette capacité ouvre des perspectives particulièrement prometteuses pour le suivi automatisé des algues vertes et des milieux côtiers.

Chapitre 44 — Amélioration d'un modèle de segmentation par fine-tuning

Une approche beaucoup plus réaliste du deep learning

Dans les chapitres précédents, nous avons découvert les principes généraux de la segmentation d'images et l'utilisation de modèles de deep learning pour classifier automatiquement des surfaces d'algues à partir d'images drone ou aériennes.

Jusqu'à présent, les modèles étaient principalement utilisés comme des "boîtes noires" déjà prêtes à l'emploi.

Dans ce nouveau module, nous allons franchir une étape supplémentaire particulièrement importante :

améliorer un modèle existant grâce au fine-tuning

Cette approche correspond beaucoup mieux à la réalité des projets professionnels en intelligence artificielle appliquée à la télédétection.

Pourquoi ne pas entraîner un modèle depuis zéro ?

Le problème des entraînements complets

En théorie, il est possible de créer un réseau neuronal entièrement nouveau et de l'entraîner à partir de zéro.

Cependant, cette stratégie présente plusieurs difficultés importantes.

Les modèles de deep learning nécessitent souvent :

- des milliers d'images ;
- une grande quantité de données annotées ;
- des temps de calcul importants ;
- des ressources matérielles puissantes.

Dans de nombreux projets environnementaux, il devient difficile de réunir un dataset suffisamment grand pour construire un modèle performant entièrement depuis le début.

Cette difficulté est particulièrement fréquente dans les projets littoraux ou écologiques où les données annotées restent limitées.

Le principe du fine-tuning

Réutiliser un modèle existant

Le fine-tuning repose sur une idée beaucoup plus efficace.

Au lieu de repartir de zéro, on commence avec :

un modèle déjà entraîné sur des données similaires

Ce modèle possède déjà des connaissances utiles :

- reconnaissance de textures ;
- structures spatiales ;
- motifs visuels ;
- formes d'algues ;
- transitions de surfaces.

On lui fournit ensuite un nouveau petit jeu de données afin de l'adapter progressivement à un nouveau contexte.

Le principe général devient alors :

modèle pré-entraîné

- nouveau dataset
→ amélioration du modèle
→ nouveau modèle spécialisé

Cette approche constitue aujourd’hui l’une des méthodes les plus utilisées dans les projets d’intelligence artificielle appliquée à :

- la télédétection environnementale ;
- l’agriculture de précision ;
- la surveillance écologique ;
- le suivi du littoral ;
- la cartographie automatique.

Une approche très utilisée en pratique

Le transfert d’apprentissage

Le fine-tuning repose sur un concept fondamental du deep learning :

le transfert d’apprentissage (*Transfer Learning*)

L’idée consiste à réutiliser les connaissances acquises par un modèle afin d’accélérer l’apprentissage sur un nouveau problème.

Le modèle n’apprend donc plus “à partir de rien”.

Il ajuste simplement ses connaissances existantes pour mieux reconnaître :

- un nouveau site d’étude ;
- de nouvelles textures ;
- des conditions environnementales différentes ;
- ou certaines classes particulières d’algues.

Cette logique rend l’apprentissage beaucoup plus rapide et beaucoup plus réaliste dans les projets opérationnels.

Pour un livre, je supprimerais complètement les références au **TP**, aux **participants** et aux **objectifs pédagogiques**. Le lecteur doit avoir l’impression d’entrer dans un nouveau chapitre qui raconte comment les professionnels travaillent réellement avec l’intelligence artificielle géospatiale.

Je le rédigerais plutôt ainsi :

Améliorer un modèle existant : l'art du fine-tuning

Lorsqu'on découvre le deep learning, on imagine souvent qu'un projet commence par la création d'un réseau neuronal entièrement nouveau. Dans cette vision idéalisée, le chercheur rassemble des milliers d'images, construit une architecture complexe et entraîne son modèle pendant des heures, voire des jours, jusqu'à obtenir des résultats satisfaisants.

La réalité est généralement bien différente.

Dans la plupart des projets opérationnels, qu'il s'agisse de télédétection environnementale, d'agriculture de précision ou d'analyse d'images drone, les spécialistes ne repartent pas de zéro. Ils s'appuient au contraire sur des modèles déjà entraînés, puis les adaptent progressivement à leur propre problématique.

Cette démarche, appelée **fine-tuning**, est aujourd'hui l'une des méthodes les plus utilisées en intelligence artificielle appliquée à la géomatique.

Pourquoi réinventer ce qui existe déjà ?

Un modèle de segmentation entraîné sur plusieurs milliers d'images a déjà appris de nombreuses informations utiles. Il sait reconnaître des contours, des textures, des contrastes ou encore des structures spatiales complexes.

Ces connaissances représentent des centaines d'heures de calcul et parfois plusieurs mois de travail.

Repartir de zéro à chaque nouveau projet serait donc inefficace.

L'approche moderne consiste plutôt à récupérer un modèle existant et à lui apprendre progressivement les particularités d'un nouveau site d'étude.

Dans le cas du suivi des algues vertes, un modèle déjà entraîné possède souvent une capacité de reconnaissance générale des structures algales. Il reste alors à l'adapter aux caractéristiques locales d'un nouveau littoral, à de nouvelles conditions d'éclairage ou à des contextes écologiques particuliers.

Le principe du transfert d'apprentissage

Le fine-tuning repose sur une idée simple : réutiliser les connaissances déjà acquises pour apprendre plus rapidement.

Cette stratégie est connue sous le nom de **transfert d'apprentissage** (*transfer learning*).

Plutôt que d'apprendre à reconnaître les algues à partir de rien, le modèle conserve ses connaissances générales et les affine progressivement grâce à un nouveau jeu de données.

Cette approche présente plusieurs avantages majeurs.

Elle réduit fortement le volume de données nécessaire, diminue le temps d'entraînement et permet d'obtenir des résultats souvent supérieurs à ceux d'un apprentissage réalisé entièrement depuis le début.

C'est également une méthode beaucoup plus proche de la réalité professionnelle, où les données annotées sont souvent rares et coûteuses à produire.

Préparer les données : une étape essentielle

Même lorsqu'un modèle existe déjà, la qualité du résultat dépend directement de la qualité des données utilisées pour son amélioration.

Le nouveau jeu de données doit comporter deux éléments complémentaires : les images à analyser et les masques de segmentation correspondants.

Chaque image doit posséder son masque associé. Les dimensions doivent être identiques, les noms de fichiers cohérents et l'organisation des répertoires rigoureuse.

Cette exigence peut sembler anodine, mais elle conditionne le bon fonctionnement de tout le processus d'apprentissage. Une simple erreur de correspondance entre une image et son masque peut suffire à dégrader fortement les performances du modèle.

Comme souvent en intelligence artificielle, la qualité de la préparation des données joue un rôle au moins aussi important que l'algorithme lui-même.

Utiliser la puissance du cloud

L'entraînement d'un modèle de deep learning nécessite des ressources de calcul importantes.

Pour éviter l'installation complexe d'environnements spécialisés, il est aujourd'hui courant d'utiliser des plateformes cloud telles que Google Colab.

Ces environnements permettent d'exécuter du code Python directement depuis un navigateur web tout en bénéficiant d'accélérateurs matériels comme les GPU.

Cette approche simplifie considérablement l'accès au deep learning et permet de se concentrer sur les concepts plutôt que sur les difficultés d'installation logicielle.

Reprendre l'entraînement d'un modèle existant

Contrairement à un apprentissage classique, où tous les paramètres du réseau sont initialisés aléatoirement, le fine-tuning démarre à partir d'un modèle déjà opérationnel.

Dans notre exemple, un modèle de segmentation des algues est rechargé depuis un espace de stockage partagé avant d'être réentraîné sur un nouveau jeu de données.

Le réseau possède déjà une compréhension générale des structures algales. Il ne s'agit donc plus de lui apprendre ce qu'est une algue, mais de lui permettre de mieux reconnaître les spécificités du nouveau contexte étudié.

Cette nuance est fondamentale.

Le modèle ne recommence pas son apprentissage ; il le poursuit.

Un apprentissage en douceur

L'une des caractéristiques du fine-tuning est l'utilisation d'un taux d'apprentissage relativement faible.

Ce paramètre contrôle la vitesse à laquelle le modèle modifie ses connaissances.

Un taux trop élevé risquerait d'effacer une partie des apprentissages antérieurs. À l'inverse, un taux plus modéré permet d'ajuster progressivement les paramètres du réseau sans détruire les connaissances déjà acquises.

Quelques cycles d'entraînement suffisent généralement pour améliorer significativement les performances sur un nouveau site.

Cette stratégie permet d'obtenir un modèle plus spécialisé tout en conservant la robustesse du modèle d'origine.

De PyTorch à QGIS

Une fois l'entraînement terminé, le modèle doit être exporté afin de pouvoir être utilisé dans un environnement géomatique opérationnel.

Le format natif utilisé pendant l'apprentissage est généralement celui de PyTorch. Cependant, pour une utilisation dans un SIG, il est souvent préférable de convertir le modèle vers un format plus universel.

Le format ONNX s'est progressivement imposé comme une solution standard permettant d'échanger des modèles entre différents logiciels.

Grâce à cette conversion, le modèle peut ensuite être directement utilisé dans le plugin Deepness au sein de QGIS.

Cette étape établit un pont entre le monde du deep learning et celui de la géomatique opérationnelle.

Comprendre plutôt que simplement utiliser

Au-delà de l'aspect technique, l'intérêt principal du fine-tuning réside dans la compréhension des mécanismes modernes de l'intelligence artificielle.

L'objectif n'est pas uniquement de produire un modèle plus performant. Il s'agit également de comprendre comment un réseau neuronal peut être réutilisé, adapté et progressivement amélioré pour répondre à de nouvelles problématiques environnementales.

Cette compréhension constitue aujourd'hui une compétence essentielle pour tout géomaticien souhaitant intégrer l'intelligence artificielle dans ses workflows d'analyse spatiale.

À retenir

Le fine-tuning représente probablement l'une des approches les plus efficaces pour appliquer le deep learning à la télédétection environnementale. Il permet d'exploiter les connaissances déjà acquises par un modèle existant tout en l'adaptant à un nouveau contexte d'étude.

Cette stratégie réduit les besoins en données, accélère l'apprentissage et reflète fidèlement les pratiques actuelles de l'intelligence artificielle géospatiale.

Dans le suivi des algues vertes comme dans de nombreux autres domaines de la télédétection, la capacité à améliorer progressivement un modèle existant est souvent plus précieuse que la création d'un modèle entièrement nouveau.

Les avantages pédagogiques du fine-tuning

Une approche beaucoup plus accessible

Le fine-tuning présente plusieurs avantages pédagogiques très importants.

Il nécessite :

- beaucoup moins de données ;
- beaucoup moins de calcul ;
- et beaucoup moins de temps d'apprentissage.

Les résultats exploitables apparaissent généralement beaucoup plus rapidement qu'avec un entraînement complet.

Cette rapidité facilite fortement l'expérimentation pédagogique.

Une approche plus proche du monde professionnel

Dans les projets réels d'intelligence artificielle, il est relativement rare de repartir entièrement de zéro.

Les équipes utilisent généralement :

- des modèles existants ;
- des modèles pré-entraînés ;
- ou des architectures déjà optimisées.

Le travail consiste ensuite à adapter progressivement ces modèles à des contextes locaux spécifiques.

Le fine-tuning reflète donc beaucoup mieux la réalité opérationnelle des projets modernes de deep learning géospatial.

A retenir

Le fine-tuning représente aujourd'hui l'une des approches les plus efficaces et les plus réalistes pour appliquer le deep learning à la télédétection environnementale.

Plutôt que de construire des modèles entièrement nouveaux, il devient possible :

- de réutiliser des connaissances existantes ;
- d'adapter rapidement un modèle à un nouveau site ;
- et d'améliorer progressivement les performances avec de petits datasets spécialisés.

Cette logique de transfert d'apprentissage constitue désormais un élément central des workflows modernes d'intelligence artificielle géospatiale.

Cinquième partie : Validation et analyse saisonnière-Automatisation

Chapitre 45 — Comparer les méthodes : apprendre à regarder les cartes autrement

Après les premières classifications réalisées dans QGIS, une tentation apparaît presque toujours. Lorsque l'on obtient enfin une carte colorée, nette, remplie de zones identifiées automatiquement, il devient facile de croire que le travail est terminé. La machine a parlé. Les algorithmes ont produit un résultat. Les algues semblent détectées. Les classes paraissent cohérentes. Tout semble fonctionner.

Pourtant, c'est précisément à ce moment-là que commence le véritable travail scientifique.

Car en télédétection, produire une carte ne suffit pas. Il faut encore apprendre à la lire, à la questionner, à la critiquer. Une classification n'est jamais une vérité absolue. C'est une interprétation construite à partir de données imparfaites, de choix méthodologiques, d'hypothèses statistiques et de paramètres techniques.

Comparer plusieurs méthodes devient alors une étape essentielle. Non pas pour désigner un "gagnant", mais pour comprendre comment chaque approche regarde le territoire, ce qu'elle révèle, ce qu'elle simplifie, et parfois ce qu'elle déforme.

Aucune méthode n'est parfaite

Très vite, les participants découvrent une réalité fondamentale : aucune méthode de classification n'est parfaite dans toutes les situations.

Certaines approches produisent des cartes très propres visuellement, presque élégantes, mais passent à côté de petites nappes d'algues pourtant importantes sur le plan écologique. D'autres détectent une grande quantité de détails, mais génèrent aussi énormément de bruit, de pixels isolés ou de faux positifs.

Une classification peut sembler scientifiquement précise tout en étant pratiquement inutilisable sur le terrain. À l'inverse, une carte légèrement simplifiée peut devenir beaucoup plus utile pour un gestionnaire littoral ou un technicien environnemental.

La question cesse alors d'être :

"Quelle est la meilleure méthode ?"

pour devenir :

“Quelle méthode est la plus adaptée à ce contexte précis ?”

Les cartes racontent toujours quelque chose

Lorsque les différentes classifications sont affichées côte à côte dans QGIS, les différences apparaissent immédiatement.

Les classifications pixel par pixel révèlent souvent une image très détaillée, presque nerveuse. Chaque variation locale de l'image semble produire un changement de classe. Les nappes d'algues se fragmentent en petits morceaux. Des pixels isolés apparaissent au milieu du sable. L'image ressemble parfois à une mosaïque instable, un bruit coloré où les frontières deviennent difficiles à interpréter.

C'est ce que l'on appelle souvent l'effet “sel et poivre”.

À l'inverse, les approches orientées objet produisent des surfaces plus homogènes. Les contours deviennent plus continus. Les structures spatiales ressemblent davantage aux formes réellement observées sur le terrain. Une nappe d'algues apparaît comme une entité cohérente plutôt qu'un ensemble de pixels indépendants.

Cette différence transforme profondément la manière dont on perçoit le paysage littoral.

La question du bruit

Le bruit de classification devient rapidement l'un des premiers éléments observés par les participants.

Quelques pixels “algues” perdus au milieu du sable peuvent sembler anecdotiques. Pourtant, lorsqu'ils se multiplient sur l'ensemble de l'image, ils dégradent fortement la lisibilité de la carte. Le regard humain peine alors à distinguer les véritables structures spatiales du simple bruit algorithmique.

Les étudiants découvrent alors que certaines méthodes réduisent naturellement ce phénomène, notamment grâce à la segmentation ou aux post-traitements spatiaux. D'autres restent beaucoup plus sensibles aux variations locales d'illumination, aux ombres ou aux petites différences de texture.

Peu à peu, ils comprennent qu'une classification ne doit pas seulement être exacte. Elle doit aussi être lisible.

Voir le territoire... ou voir l'algorithme ?

Une carte peut parfois révéler davantage la logique de l'algorithme que la réalité du terrain.

Cette idée surprend souvent au début.

Mais lorsqu'on observe certaines classifications très bruitées, certaines limites artificielles ou certaines ruptures incohérentes, une question apparaît :

est-ce que cette structure existe réellement sur le terrain... ou est-ce simplement une conséquence du modèle utilisé ?

Cette interrogation constitue un moment important dans l'apprentissage de la télédétection. Les participants commencent alors à regarder les cartes non plus comme des résultats définitifs, mais comme des objets scientifiques à interpréter avec prudence.

Les limites des signatures spectrales

Le milieu littoral révèle rapidement toute la complexité des classifications multispectrales.

Certaines surfaces présentent des signatures extrêmement proches :

- sable humide ;
- vase ;
- biofilms ;
- végétation halophile ;
- algues partiellement immergées.

Dans certaines zones, les frontières deviennent floues, non seulement pour les algorithmes, mais parfois même pour l'œil humain.

Une fine couche d'algues recouvrant un substrat humide peut produire une signature intermédiaire difficile à interpréter. L'image cesse alors d'être un simple assemblage de classes bien séparées. Elle devient un espace de transitions, de mélanges et d'ambiguïtés.

C'est précisément dans ces situations que les différences entre méthodes deviennent intéressantes.

Les approches statistiques classiques montrent parfois leurs limites lorsque les distributions spectrales se chevauchent. Les méthodes de machine learning parviennent souvent à mieux gérer ces relations complexes, mais au prix d'une plus grande difficulté d'interprétation.

Les petites nappes d'algues : un détail... ou un enjeu majeur ?

Un autre phénomène attire rapidement l'attention : la disparition des petits objets.

Certaines segmentations fusionnent les petites plaques d'algues avec leur environnement voisin. Certains filtrages éliminent les structures jugées trop petites ou trop isolées. Visuellement, la carte devient plus propre. Mais écologiquement, une partie de l'information disparaît.

Or, dans les milieux littoraux, ces petites zones peuvent avoir une importance scientifique considérable.

Elles peuvent représenter :

- des débuts de prolifération ;
- des résidus d'échouage ;
- des zones particulièrement sensibles ;
- ou des indicateurs précoces de changement environnemental.

Les participants découvrent alors qu'améliorer la "beauté" d'une carte peut parfois appauvrir sa valeur scientifique.

Le poids des contraintes techniques

Très vite, les aspects techniques réapparaissent également dans l'analyse.

Les images drone multispectrales sont lourdes, détaillées et exigeantes en calcul. Certaines méthodes fonctionnent rapidement et simplement. D'autres nécessitent davantage de mémoire, des temps de traitement plus longs, voire l'utilisation de GPU dans le cas du deep learning.

Mais la difficulté ne vient pas seulement du temps de calcul.

Certaines approches deviennent aussi plus difficiles à paramétrer, à reproduire ou à expliquer. Une méthode très performante peut devenir problématique si elle produit des résultats impossibles à interpréter ou impossibles à reproduire dans un autre contexte.

Cette question de la robustesse devient centrale.

Une méthode réellement utile doit rester relativement stable :

- sur plusieurs sites ;
- à différentes dates ;
- et sous des conditions environnementales variables.

Le littoral : un environnement instable

Le milieu littoral constitue probablement l'un des environnements les plus difficiles à analyser en télédétection.

Les marées modifient constamment les surfaces visibles. Les zones humides changent rapidement. Les reflets spéculaires apparaissent puis disparaissent selon l'angle du soleil. Les ombres déplacent artificiellement certaines signatures spectrales.

Parfois, deux acquisitions réalisées sur le même site à quelques heures d'intervalle produisent déjà des différences importantes.

Les participants découvrent alors une réalité fondamentale : les variations observées dans les données ne correspondent pas toujours à des changements écologiques réels.

Une partie des différences provient simplement :

- des conditions d'acquisition ;
- de la lumière ;
- de l'humidité ;
- ou du niveau de marée.

Cette prise de conscience transforme profondément la manière d'interpréter les cartes produites.

Les classifications comme outils d'interprétation

Au fil des comparaisons, une idée essentielle finit par émerger.

Les classifications ne représentent pas directement le réel.

Elles représentent une interprétation algorithmique du réel.

Chaque méthode simplifie le territoire à sa manière. Chaque algorithme privilégie certaines structures et en néglige d'autres. Chaque workflow produit une lecture particulière du paysage littoral.

Le rôle du scientifique n'est donc pas seulement d'exécuter des traitements.

Il consiste aussi à comprendre :

- ce que les modèles montrent ;
- ce qu'ils cachent ;
- et dans quelles conditions leurs résultats deviennent fiables.

Conclusion

Cette phase comparative constitue souvent l'un des moments les plus enrichissants d'un projet de télédétection.

À ce stade, le lecteur est invité à dépasser la simple utilisation des outils pour adopter une véritable démarche d'analyse. Les classifications ne sont plus perçues comme des réponses automatiques produites par un logiciel, mais comme des représentations du territoire qu'il convient d'interroger, de comparer et de valider.

Au fil des expérimentations, une réalité s'impose : aucune méthode n'est parfaite, aucun algorithme n'est universel et aucune carte ne doit être interprétée sans recul. Chaque résultat dépend des données utilisées, des paramètres choisis et des hypothèses sous-jacentes au traitement.

La télédétection apparaît alors dans toute sa richesse. Elle ne se résume pas à des calculs, à des indices spectraux ou à des classifications automatiques. Elle est aussi une discipline d'interprétation, de confrontation des points de vue, d'évaluation critique et de compréhension des mécanismes qui façonnent les paysages observés.

Développer ce regard critique constitue sans doute l'un des apprentissages les plus précieux. C'est lui qui permet de passer d'une simple maîtrise des logiciels à une véritable capacité d'analyse environnementale. Derrière chaque carte se cachent des choix méthodologiques, des incertitudes et des limites qu'il est indispensable de comprendre pour produire une information fiable et scientifiquement pertinente.

En définitive, la valeur d'une classification ne réside pas uniquement dans sa précision, mais dans la capacité de l'analyste à en comprendre les forces, les faiblesses et le domaine de validité. C'est cette démarche réflexive qui transforme un traitement d'image en véritable connaissance du territoire.

Chapitre 46 — Apprendre à douter des cartes

La discussion scientifique des résultats de classification.

Après avoir produit plusieurs classifications, une étape essentielle commence : leur interprétation critique.

Les cartes se multiplient, les résultats s'affichent à l'écran et certaines classifications paraissent immédiatement convaincantes. Les surfaces sont bien délimitées, les couleurs homogènes et l'ensemble semble cohérent. D'autres cartes apparaissent plus complexes, plus bruitées ou parfois contradictoires.

C'est souvent à ce moment qu'émerge l'une des illusions les plus fréquentes en télédétection : croire qu'une carte est juste parce qu'elle est visuellement séduisante.

Or, la réalité scientifique est bien plus nuancée.

Une classification peut présenter une apparence irréprochable tout en contenant des erreurs importantes, des biais liés aux données d'apprentissage ou des interprétations discutables. À l'inverse, une carte moins esthétique peut parfois mieux refléter la complexité réelle du terrain.

La qualité d'un résultat ne se mesure donc pas à son apparence graphique, mais à sa capacité à représenter fidèlement les phénomènes observés.

Cette étape marque un changement de perspective important. L'objectif n'est plus seulement de produire des cartes, mais de les questionner. Chaque classification devient une hypothèse de travail qu'il convient d'examiner avec rigueur.

Les résultats ne doivent pas être considérés comme des vérités produites automatiquement par un logiciel. Ils constituent une interprétation du territoire fondée sur des données, des choix méthodologiques et des modèles mathématiques. Comme toute interprétation scientifique, ils doivent être analysés, confrontés à d'autres sources d'information et, lorsque cela est nécessaire, remis en question.

Cette démarche critique est au cœur de la télédétection moderne. Elle permet de dépasser la simple production de cartes pour accéder à une véritable compréhension des processus environnementaux observés. Apprendre à douter des cartes n'est pas un signe de méfiance envers les méthodes de classification ; c'est au contraire une condition indispensable pour en faire un usage scientifique rigoureux.

La télédétection comme discipline d'interprétation

Au début d'un projet, il est tentant de considérer les algorithmes comme des outils capables de révéler directement la réalité du terrain.

Mais plus les classifications deviennent complexes, plus une évidence apparaît :

les algorithmes ne voient pas le terrain comme un humain

Ils analysent des signatures spectrales, des textures, des distributions statistiques ou des relations mathématiques entre variables. Ils ne "comprennent" pas réellement les algues, les marées ou les dynamiques écologiques.

Ce décalage est fondamental.

Deux méthodes appliquées à la même image peuvent produire des cartes très différentes. Une classification peut faire apparaître des nappes d'algues très étendues tandis qu'une autre les

réduit fortement. Certaines méthodes fragmentent les surfaces. D'autres les homogénéisent excessivement.

Peu à peu, les participants découvrent que chaque algorithme construit en réalité une lecture particulière du territoire.

Les cartes révèlent aussi les limites des méthodes

Lorsque les différentes classifications sont comparées côte à côte dans QGIS, les contrastes deviennent immédiatement visibles.

Les classifications pixel par pixel révèlent souvent une grande quantité de détails. Chaque variation locale semble détectée. Les petites nappes apparaissent avec précision. Mais cette finesse produit aussi du bruit. Des pixels isolés apparaissent au milieu du sable. Les contours deviennent irréguliers. L'image donne parfois l'impression d'un territoire instable, fragmenté, presque agité.

Les approches orientées objet racontent une autre histoire. Les surfaces deviennent plus continues. Les nappes d'algues semblent plus cohérentes spatialement. Les formes paraissent plus proches de ce que l'on observerait réellement sur le terrain.

Mais cette stabilité a parfois un prix : certaines petites structures disparaissent, absorbées par des segments plus grands.

Les participants comprennent alors qu'aucune méthode ne révèle "la vérité". Chaque approche simplifie le réel à sa manière.

L'importance des erreurs

Un moment important survient lorsque l'on commence à analyser les erreurs plutôt qu'à les éviter.

Car en télédétection, les erreurs ne sont pas simplement des défauts techniques. Elles sont souvent riches d'enseignements.

Une zone de vase humide classée comme algue révèle parfois une proximité spectrale réelle entre les deux surfaces. Une nappe d'algues non détectée peut signaler une humidité excessive, une ombre locale ou un mélange spectral difficile à interpréter.

Les faux positifs et les faux négatifs cessent alors d'être de simples "problèmes". Ils deviennent des indices sur le comportement des données et sur les limites des modèles utilisés.

Cette étape transforme profondément la manière d'aborder les classifications.

Le regard ne cherche plus seulement ce qui fonctionne.

Il cherche aussi à comprendre pourquoi certaines choses échouent.

Le littoral : un environnement profondément ambigu

Le milieu littoral joue ici un rôle essentiel.

Contrairement à certains paysages plus homogènes, l'estran est un espace de transition permanente. L'eau, le sable, la vase, les algues, les biofilms et la végétation halophile se mélangent continuellement.

Les frontières deviennent floues.

Un pixel peut contenir simultanément :

- un film d'eau ;
- du sable humide ;
- des fragments d'algues ;
- et des micro-organismes.

Les signatures spectrales deviennent alors ambiguës, intermédiaires, parfois impossibles à séparer clairement.

On découvre que certaines difficultés de classification ne proviennent pas d'un mauvais algorithme, mais simplement de la complexité réelle du milieu observé.

Les paramètres changent le paysage

Une autre découverte importante concerne l'influence des paramètres.

Au début, beaucoup imaginent qu'un algorithme produit un résultat "objectif". Pourtant, quelques modifications suffisent parfois à transformer complètement une classification.

Un seuil de segmentation légèrement différent peut fusionner des nappes entières ou, au contraire, fragmenter l'image en centaines de petits objets. Quelques arbres supplémentaires dans un modèle Random Forest peuvent améliorer la stabilité... ou parfois ne rien changer du tout.

On prend alors conscience qu'une classification n'est jamais indépendante des choix réalisés par l'analyste.

Chaque paramètre influence la manière dont le territoire sera représenté.

Le piège des algorithmes sophistiqués

L'arrivée du machine learning et du deep learning suscite souvent beaucoup d'enthousiasme.

Les modèles modernes semblent capables de résoudre des problèmes très complexes. Ils gèrent mieux les signatures non linéaires, les textures, les interactions entre bandes spectrales. Dans certains cas, ils produisent effectivement des résultats impressionnants.

Mais cette puissance peut aussi devenir trompeuse.

On découvre progressivement qu'un algorithme sophistiqué ne compense jamais :

- des données médiocres ;
- des images mal acquises ;
- des ROI mal choisies ;
- ou une mauvaise connaissance du terrain.

Cette idée devient centrale.

La qualité scientifique d'une classification dépend souvent davantage :

- de la qualité des données ;
- de la rigueur méthodologique ;
- et de la compréhension écologique du site

que de la complexité de l'algorithme lui-même.

Le rôle du terrain

Au fil des discussions, une évidence finit par s'imposer : aucune classification ne peut être interprétée correctement sans connaissance du terrain.

Les observations GPS, les photographies drone, les visites de site ou les relevés écologiques deviennent indispensables pour comprendre certaines anomalies.

Une zone que l'algorithme considère comme "algue" peut en réalité correspondre :

- à une vase très humide ;
- à un biofilm ;
- ou à une végétation halophile.

La télédétection cesse alors d'être une simple discipline informatique.

Elle redevient une science de l'observation du territoire.

La matrice de confusion : regarder les erreurs en face

La matrice de confusion joue ici un rôle pédagogique très fort.

Elle nous oblige à regarder concrètement les erreurs du modèle. Certaines classes sont très bien détectées. D'autres sont régulièrement confondues. Certaines disparaissent presque complètement malgré une excellente précision globale.

Ce moment est souvent marquant.

On comprend qu'une classification ne peut pas être évaluée uniquement "à l'œil". Une carte visuellement convaincante peut cacher des déséquilibres importants ou des biais statistiques majeurs.

La validation devient alors une partie intégrante du raisonnement scientifique.

Une réflexion sur le choix des méthodes

Progressivement, les discussions dépassent les simples questions techniques.

Les participants commencent à réfléchir au choix des méthodes lui-même.

Quand faut-il privilégier une approche simple et rapide ? Quand devient-il nécessaire d'utiliser une segmentation ? À partir de quel moment le machine learning apporte-t-il réellement une amélioration ?

Les réponses deviennent nuancées.

Dans certains projets simples, une classification pixel par pixel reste parfaitement suffisante. Dans des images drone très haute résolution, les approches orientées objet deviennent souvent beaucoup plus pertinentes. Les méthodes de machine learning apportent parfois des gains importants, mais pas systématiquement.

Peu à peu, les participants découvrent qu'il n'existe pas de méthode universelle.

Le choix dépend toujours :

- du terrain ;
 - des objectifs ;
 - des données disponibles ;
 - du niveau de précision recherché ;
 - et des moyens techniques accessibles.
 -
-

Une transformation du regard

Au-delà des techniques de classification, des indices spectraux ou des algorithmes étudiés dans cet ouvrage, l'apprentissage de la télédétection conduit souvent à une évolution plus profonde : une transformation de la manière de regarder les données géographiques.

Lors des premières analyses, il est naturel de considérer les classifications comme des résultats produits par un logiciel. Une image entre dans la chaîne de traitement, un algorithme est exécuté et une carte est générée. Le résultat apparaît alors comme une réponse objective à une question posée.

Avec l'expérience, cette vision évolue progressivement.

Les classifications ne sont plus perçues comme de simples produits techniques, mais comme des constructions méthodologiques résultant d'une succession de choix : choix des données, des paramètres, des échantillons d'apprentissage, des algorithmes ou encore des méthodes de validation.

Elles deviennent également des interprétations du territoire. Derrière chaque classe cartographiée se cache une traduction simplifiée d'une réalité souvent beaucoup plus complexe. Les frontières entre les objets observés ne sont pas toujours aussi nettes que celles dessinées sur la carte, et certaines ambiguïtés du terrain ne peuvent être totalement résolues par les modèles.

Peu à peu, les classifications apparaissent pour ce qu'elles sont réellement : des représentations imparfaites du monde, utiles et souvent très performantes, mais qu'il convient toujours de comprendre, d'évaluer et de questionner.

Cette évolution du regard constitue sans doute l'un des apports les plus importants de la télédétection. Elle marque le passage d'une utilisation essentiellement technique des outils à une démarche d'analyse fondée sur l'esprit critique, la compréhension des méthodes et l'interprétation raisonnée des résultats.

C'est précisément cette capacité à interroger les cartes plutôt qu'à les accepter aveuglément qui permet de transformer des données géospatiales en connaissances scientifiques.

Conclusion

Cette discussion scientifique finale montre que la télédétection ne se résume jamais à l'utilisation d'outils ou d'algorithmes.

Elle repose avant tout sur :

- l'analyse ;
- l'interprétation ;
- la validation ;
- et la compréhension du milieu observé.

Les classifications deviennent alors non pas des réponses définitives, mais des outils d'exploration scientifique du territoire.

Et c'est précisément cette capacité à interpréter les résultats avec recul, nuance et esprit critique qui distingue un simple utilisateur de logiciel d'un véritable analyste en télédétection environnementale.

Chapitre 47 — Du classement des pixels à la compréhension des dynamiques littorales

Après avoir appris à acquérir des images, à les prétraiter, à calculer des indices spectraux et à réaliser différentes classifications, une question essentielle demeure :

que peut-on réellement apprendre de ces cartes ?

La classification constitue souvent l'étape la plus visible d'un projet de télédétection. Pourtant, elle ne représente qu'un point de départ. Une carte classifiée n'est pas une fin en soi. Elle devient véritablement utile lorsqu'elle permet de produire des connaissances sur le fonctionnement du milieu étudié.

Dans le cadre du suivi des algues vertes, l'objectif n'est pas seulement de savoir où se trouvent les algues à une date donnée. Il s'agit également de comprendre comment leur répartition évolue dans le temps, quelles zones sont les plus sensibles aux proliférations et quels facteurs environnementaux semblent influencer leur développement.

Cette dernière étape du parcours transforme donc progressivement le géomaticien en analyste de l'environnement.

La nécessité de valider les résultats

Avant toute interprétation écologique, une règle fondamentale doit être respectée : vérifier la qualité des résultats produits.

Cette exigence peut sembler évidente, mais elle est souvent sous-estimée. Une carte visuellement convaincante n'est pas nécessairement correcte. Des couleurs harmonieuses, des contours nets ou une apparente cohérence spatiale peuvent masquer des erreurs importantes.

Dans toute démarche scientifique, la confiance accordée à une observation dépend de sa capacité à être vérifiée.

La validation consiste précisément à confronter la classification à une référence indépendante. Cette référence peut provenir d'observations de terrain, de photographies aériennes, de relevés GPS ou d'une expertise visuelle soigneusement documentée.

L'objectif n'est pas de démontrer que la classification est parfaite. Aucune méthode ne l'est. Il s'agit plutôt de mesurer son degré de fiabilité et de comprendre quelles classes sont correctement détectées et lesquelles posent davantage de difficultés.

Apprendre à mesurer l'erreur

L'erreur joue un rôle central dans toute démarche scientifique.

Dans un contexte littoral, certaines confusions apparaissent fréquemment. Les algues humides peuvent être confondues avec des substrats très sombres. Une eau chargée en sédiments peut présenter une signature proche de certaines surfaces végétalisées. Des zones d'ombre peuvent être interprétées comme de l'eau ou de la vase.

L'analyse de ces erreurs apporte souvent davantage d'informations que les succès de la classification.

Comprendre pourquoi une zone a été mal classée permet d'identifier les limites du capteur, de la méthode utilisée ou parfois même de mieux comprendre le fonctionnement écologique du site.

Cette réflexion conduit naturellement à l'utilisation de la matrice de confusion, véritable outil d'autopsie de la classification.

La matrice de confusion : un miroir de la classification

Parmi les outils utilisés pour évaluer une classification, la matrice de confusion occupe une place particulière. Contrairement aux cartes produites par les algorithmes, elle ne met pas en valeur les réussites du modèle. Elle met en évidence ses erreurs.

C'est précisément cette capacité à révéler les imperfections qui en fait un outil aussi précieux.

En confrontant les classes prédites par le modèle aux observations de référence, la matrice de confusion offre une lecture détaillée des performances de la classification. Elle permet d'identifier les classes correctement reconnues, mais aussi celles qui sont régulièrement confondues avec d'autres. Là où la carte donne une vision globale du résultat, la matrice révèle les mécanismes qui ont conduit à ce résultat.

Cette analyse conduit souvent à remettre en question certaines impressions initiales. Une classification peut sembler très satisfaisante lorsqu'elle est observée à l'écran, tout en présentant des faiblesses importantes sur certaines catégories de surfaces. À l'inverse, une carte visuellement imparfaite peut parfois se révéler plus robuste qu'il n'y paraît.

L'un des enseignements majeurs de la matrice de confusion concerne la prudence nécessaire dans l'interprétation des indicateurs globaux. Une précision élevée ne garantit pas toujours une classification de qualité. Lorsqu'une classe est largement dominante dans le paysage, elle

peut fortement influencer les statistiques et masquer les difficultés rencontrées sur des classes plus rares.

Dans le cas du suivi environnemental, cette situation est particulièrement importante. Des éléments écologiquement significatifs, comme de petites nappes d'algues, des zones humides localisées ou des habitats peu représentés, peuvent être mal détectés tout en ayant un impact limité sur la précision globale. Une lecture superficielle des résultats conduirait alors à surestimer la qualité réelle du modèle.

La matrice de confusion rappelle ainsi une idée essentielle : évaluer une classification ne consiste pas seulement à mesurer ce qui fonctionne, mais aussi à comprendre ce qui échoue. Les erreurs ne sont pas des anomalies à ignorer ; elles constituent souvent la source d'information la plus riche pour améliorer les méthodes, affiner les données d'apprentissage et mieux comprendre les limites du modèle.

À bien des égards, la matrice de confusion agit comme un miroir de la classification. Elle reflète non seulement ses performances, mais aussi ses faiblesses, ses biais et les incertitudes inhérentes à toute représentation simplifiée du territoire. C'est pourquoi elle demeure l'un des outils les plus importants de la démarche scientifique en télédétection.

Transformer une carte en indicateur environnemental

Une fois la classification validée, une nouvelle étape commence.

Les pixels cessent d'être uniquement des catégories spatiales. Ils deviennent des mesures.

La présence ou l'absence d'algues peut désormais être quantifiée. Les surfaces occupées par les peuplements algaux peuvent être calculées avec précision. Des comparaisons peuvent être réalisées entre différents secteurs du littoral ou entre différentes périodes de l'année.

Cette transformation est essentielle.

On passe d'une information qualitative à une information quantitative.

Au lieu d'affirmer que les algues semblent nombreuses sur une plage, il devient possible de mesurer qu'elles occupent par exemple 340 mètres carrés dans un secteur donné et seulement 125 mètres carrés dans un autre.

Les cartes deviennent alors des indicateurs environnementaux.

Lire le littoral à travers les chiffres

Les surfaces calculées ne sont pas de simples statistiques.

Elles racontent une histoire.

Une augmentation rapide de la surface occupée par les algues peut révéler une prolifération saisonnière. Une diminution brutale peut résulter d'un épisode météorologique particulier ou d'une modification des conditions hydrodynamiques. Une stabilité prolongée peut signaler un équilibre écologique spécifique.

Chaque valeur doit donc être replacée dans son contexte environnemental.

Le chiffre seul n'a pas de sens. C'est son interprétation qui lui donne sa valeur scientifique.

Observer le temps à travers les cartes

L'un des grands intérêts des drones et de la télédétection réside dans leur capacité à répéter les observations.

Une image isolée fournit un instantané du territoire. Une série d'images révèle une dynamique.

Lorsque plusieurs classifications sont comparées à différentes dates, les cartes commencent à raconter l'évolution du milieu. Les surfaces algales augmentent, diminuent ou se déplacent. Certaines zones restent stables alors que d'autres connaissent des changements rapides.

Le paysage cesse d'être figé.

Il devient un système en mouvement.

Comprendre les changements

Comparer deux dates ne consiste pas simplement à mesurer une différence de surface.

Il s'agit également d'identifier la nature du changement.

Certaines zones montrent une extension des peuplements algaux. D'autres révèlent une régression. D'autres encore demeurent stables au fil du temps.

La cartographie des changements permet alors de distinguer plusieurs comportements écologiques et de localiser précisément les secteurs où les dynamiques sont les plus marquées.

Cette approche apporte une information beaucoup plus riche qu'une simple comparaison de surfaces globales.

Les rythmes saisonniers du littoral

À mesure que les observations s'accumulent, une autre dimension apparaît : la saisonnalité.

Les algues ne se développent pas de manière uniforme tout au long de l'année. Leur croissance dépend de nombreux facteurs tels que la température, l'ensoleillement, la disponibilité en nutriments, les marées ou encore les conditions météorologiques.

L'analyse de plusieurs campagnes d'acquisition permet alors de mettre en évidence :

- les périodes de prolifération maximale ;
- les périodes de régression ;
- les zones de persistance ;
- les secteurs les plus dynamiques.

Le suivi temporel devient ainsi un véritable outil de compréhension écologique.

Une classification n'est jamais une fin

Au terme de ce parcours, une idée essentielle s'impose progressivement : la classification n'est pas l'objectif final de la télédétection.

Les cartes produites, aussi détaillées et précises soient-elles, ne constituent pas une finalité en elles-mêmes. Elles représentent une étape intermédiaire dans un processus plus vaste de compréhension du territoire. Leur véritable intérêt réside dans ce qu'elles permettent de révéler, de mesurer et d'interpréter.

Une classification n'est finalement qu'une traduction simplifiée de la réalité observée. Elle organise l'information, met en évidence certaines structures spatiales et facilite l'analyse de phénomènes parfois complexes. Mais sa valeur ne se limite pas à l'identification de classes ou à la production d'une carte thématique.

À partir de ces résultats, il devient possible de construire des indicateurs, de quantifier des phénomènes, de suivre des évolutions dans le temps et d'identifier des tendances parfois invisibles à l'œil nu. Les cartes cessent alors d'être de simples représentations visuelles pour devenir de véritables outils d'observation et d'analyse.

Dans le contexte du suivi des algues vertes, par exemple, la classification permet d'estimer des surfaces, de comparer différentes périodes, d'évaluer l'efficacité de mesures de gestion ou encore de mieux comprendre les dynamiques écologiques du littoral. Chaque carte devient ainsi un instantané d'un système en évolution.

Cette capacité à transformer des images en connaissances constitue sans doute l'une des plus grandes forces de la télédétection. Derrière chaque pixel classé se cache la possibilité de mieux comprendre le fonctionnement d'un environnement, d'anticiper certaines évolutions et d'éclairer les décisions de gestion.

La télédétection ne se résume donc pas à l'analyse d'images ou à l'application d'algorithmes. Elle s'inscrit dans une démarche plus large d'observation scientifique du territoire. Les classifications, les indices et les modèles ne sont que des moyens au service d'un objectif plus ambitieux : produire une connaissance utile du monde qui nous entoure.

C'est précisément à ce moment que la télédétection dépasse sa dimension technique pour devenir un véritable outil de compréhension écologique, capable de relier les données, les processus environnementaux et les enjeux de gestion des espaces naturels.

Chapitre 48 — Automatiser la validation : quand Python devient assistant du géomaticien

Au fil des chapitres précédents, de nombreuses opérations ont été réalisées dans QGIS : préparation des données, calcul d'indices spectraux, classifications, validation des résultats et production d'indicateurs environnementaux. Ces manipulations sont indispensables pour comprendre les principes de la télédétection, appréhender les mécanismes des algorithmes et développer un regard critique sur les cartes produites.

Cependant, lorsqu'un projet dépasse le stade de l'expérimentation pour entrer dans une phase opérationnelle, une nouvelle problématique apparaît. Les mêmes traitements doivent être répétés à chaque nouvelle acquisition, parfois sur plusieurs sites et pendant plusieurs années.

Créer des points de validation, extraire les valeurs de classification, calculer des matrices de confusion, produire des indicateurs statistiques ou générer des rapports de qualité sont autant d'opérations qui deviennent rapidement répétitives. Réalisées manuellement, elles mobilisent un temps considérable et augmentent le risque d'erreurs ou d'incohérences entre campagnes d'analyse.

C'est dans ce contexte que l'automatisation prend toute son importance.

L'objectif n'est pas de remplacer l'expertise du géomaticien ni le travail d'interprétation du terrain. Au contraire, l'automatisation vise à prendre en charge les tâches répétitives afin de libérer du temps pour les activités à forte valeur ajoutée : l'analyse des résultats, la validation scientifique et la compréhension des phénomènes observés.

Python occupe aujourd'hui une place centrale dans cette démarche. Grâce à sa simplicité d'utilisation et à son intégration dans QGIS, il permet de transformer une succession de manipulations manuelles en chaînes de traitement reproductibles, documentées et facilement réutilisables.

Cette évolution marque souvent une étape importante dans la pratique de la géomatique. Le géomaticien ne se contente plus d'utiliser des outils ; il construit progressivement ses propres méthodes d'analyse, adaptées à ses besoins et à ses problématiques environnementales.

Dans ce chapitre, nous verrons comment automatiser une étape essentielle de tout projet de classification : la validation des résultats. À travers l'exemple de la matrice de confusion, nous découvrirons comment quelques scripts Python peuvent transformer une procédure fastidieuse en un workflow robuste, reproductible et directement exploitable dans le cadre d'un suivi environnemental à long terme.

Le rôle irremplaçable de l'expert

Une idée revient souvent lorsque l'on découvre Python et l'automatisation :

« Peut-on tout automatiser ? »

La réponse est généralement non.

Dans le cas de la validation d'une classification, une étape reste fondamentalement humaine : la détermination de la vérité terrain.

Aucun script ne peut décider à la place de l'observateur si un point correspond réellement à une nappe d'algues, à du sable humide ou à une zone d'ombre.

Cette décision repose sur :

- l'analyse des orthophotographies ;
- les observations de terrain ;
- l'expérience de l'opérateur ;
- la connaissance écologique du site.

L'automatisation intervient donc autour de cette expertise, mais ne la remplace pas.

Première étape : préparer automatiquement les points de validation

Toute validation de classification repose sur un principe simple : confronter les résultats produits par l'algorithme à une référence indépendante représentant le plus fidèlement possible la réalité du terrain.

Cette étape est indispensable, mais elle peut rapidement devenir fastidieuse lorsque le nombre de points de contrôle augmente. Créer manuellement des centaines de points, extraire les valeurs de classification et préparer les données nécessaires à l'évaluation constitue un travail répétitif qui mobilise un temps important sans apporter de réelle valeur scientifique.

L'automatisation permet précisément de simplifier cette phase préparatoire.

À partir d'un raster classifié, un script Python peut générer automatiquement un ensemble de points répartis sur l'emprise de l'image. Ces points constituent les futurs échantillons de validation qui serviront à évaluer la qualité de la classification.

Le programme crée également les champs attributaires nécessaires à l'analyse et extrait automatiquement, pour chaque point, la classe attribuée par le modèle de classification. En quelques secondes, il produit ainsi une couche prête à être examinée et interprétée.

Cette automatisation présente plusieurs avantages. Elle réduit considérablement le temps de préparation, garantit une procédure identique d'une campagne à l'autre et limite les risques d'erreurs liés aux manipulations manuelles. Surtout, elle permet de consacrer davantage de temps à ce qui constitue le cœur de la validation : l'analyse critique des observations et la détermination de la vérité terrain.

Le rôle du géomaticien ne disparaît pas pour autant. Au contraire, il se recentre sur les tâches qui nécessitent une expertise humaine. Le script prépare les données, mais c'est toujours l'observateur qui interprète les images, confronte les résultats aux connaissances du terrain et attribue la classe de référence à chaque point.

L'automatisation ne remplace donc pas l'expertise ; elle la valorise en prenant en charge les tâches répétitives pour laisser davantage de place à l'analyse et à la réflexion scientifique.

Le premier script présenté a pour objectif de préparer l'ensemble du travail de validation.

À partir d'un raster classifié, le programme génère automatiquement plusieurs centaines de points répartis sur l'emprise de l'image. Ces points constituent les futurs échantillons de validation.

Le script ajoute ensuite les champs nécessaires à la validation et extrait automatiquement, pour chaque point, la classe attribuée par la classification.

```

import pandas as pd
from sklearn.metrics import confusion_matrix, accuracy_score, cohen_kappa_score
import os
# =====
# PARAMÈTRES (A mettre à jour avant chaque exécution)
# =====
# Raster classifié
raster_path = r"path\mage_classée.tif"
# Dossier de sortie
output_folder = r"path\automatisation"
# Nombre de points
n_points = 200
# Nom du champ vérité terrain
reference_field = "classe_ref"
# =====
# CRÉATION DOSSIER
# =====
os.makedirs(output_folder, exist_ok=True)
# =====
# CHARGEMENT RASTER
# =====
raster_layer = iface.addRasterLayer(raster_path, "classification")
extent = raster_layer.extent()
# =====
# 1. CRÉER POINTS ALÉATOIRES
# =====
points_path = os.path.join(output_folder, "validation_points.gpkg")
random_points = processing.run(
    "native:randompointsinextent",
    {
        'EXTENT': extent,
        'MIN_DISTANCE': 1,
        'POINTS_NUMBER': n_points,
        'TARGET_CRS': raster_layer.crs(),
        'OUTPUT': points_path
    }
)
points_layer = QgsVectorLayer(points_path, "validation_points", "ogr")
# =====
# 2. AJOUTER CHAMPS
# =====
provider = points_layer.dataProvider()
provider.addAttributes([
    QgsField(reference_field, QVariant.Int)
])
points_layer.updateFields()
# =====
# 3. EXTRAIRE VALEURS RASTER
# =====
sampled_path = os.path.join(output_folder, "validation_sampled.gpkg")
processing.run(
    "native:rastersampling",
    {
        'INPUT': points_layer,
        'RASTERCOPY': raster_layer,
        'COLUMN_PREFIX': 'r_',
        'OUTPUT': sampled_path
    }
)
sampled_layer = QgsVectorLayer(sampled_path, "sampled", "ogr")
QgsProject.instance().addMapLayer(sampled_layer)
print("=====")
print("REMPLIR MAINTENANT LE CHAMP :")
print(reference_field)
print("avec les classes de référence")
print("=====")

```

Le moment où l'humain reprend la main

Une fois la couche créée, l'automatisation s'arrête volontairement.

L'utilisateur examine alors chaque point à partir :

- de l'orthophotographie ;
- des images drone ;
- des observations de terrain ;
- de sa connaissance du site.

Il renseigne le champ **classe_ref**, qui représente la classe réellement observée.

Cette étape est essentielle.

Elle rappelle que la validation n'est pas une simple opération informatique. Elle constitue un véritable travail d'expertise.

Le script prépare le terrain, mais c'est l'observateur qui apporte la connaissance du réel.

Transformer les observations en indicateurs

Une fois les points renseignés, le second script entre en jeu.

Son rôle consiste à comparer automatiquement :

- la classe observée sur le terrain ;
- la classe prédite par la classification.

À partir de cette comparaison, le programme construit immédiatement la matrice de confusion.

Mais il va beaucoup plus loin.

Il calcule également :

- l'accuracy globale ;
- la précision ;
- le rappel ;
- le F-score ;
- le coefficient Kappa ;
- les statistiques détaillées pour chaque classe.

Quelques secondes suffisent pour produire ce qui nécessiterait autrement plusieurs calculs manuels.

```

from sklearn.metrics import (
    accuracy_score,
    precision_score,
    recall_score,
    f1_score,
    cohen_kappa_score
)
import pandas as pd
from sklearn.metrics import confusion_matrix, accuracy_score, cohen_kappa_score
import os
from qgis.PyQt.QtCore import QVariant
# =====
# 4. EXPORT ATTRIBUTS VERS DATAFRAME
# =====
features = []
for feat in sampled_layer.getFeatures():
    raster_val = feat["r_1"]
    ref_val = feat[reference_field]
    if not raster_val:
        continue
    if not ref_val:
        continue
    try:
        features.append({
            "raster_val": int(float(raster_val)),
            "classe_ref": int(float(ref_val))
        })
    except Exception:
        continue
df = pd.DataFrame(features)
# =====
# 5. CLASSES AUTOMATIQUES
# =====
classes = sorted(
    list(
        set(df["raster_val"]).union(
            set(df["classe_ref"])
        )
    )
)
# =====
# 6. MATRICE DE CONFUSION
# =====
cm = confusion_matrix(
    df["classe_ref"],
    df["raster_val"],
    labels=classes
)
cm_df = pd.DataFrame(
    cm,
    index=[f"Ref_{c}" for c in classes],
    columns=[f"Pred_{c}" for c in classes]
)
# =====
# 7. STATISTIQUES
# =====
y_true = df["classe_ref"]
y_pred = df["raster_val"]
accuracy = accuracy_score(
    y_true,
    y_pred
)
precision = precision_score(
    y_true,
    y_pred,
    average="weighted",
    zero_division=0
)
recall = recall_score(
    y_true,
    y_pred,
    average="weighted",
    zero_division=0
)
f1 = f1_score(

```

```

    y_true,
    y_pred,
    average="weighted",
    zero_division=0
)
kappa = cohen_kappa_score(
    y_true,
    y_pred
)
stats_df = pd.DataFrame({
    "Metric": [
        "Accuracy",
        "Precision",
        "Recall",
        "F1-score",
        "Kappa"
    ],
    "Value": [
        accuracy,
        precision,
        recall,
        f1,
        kappa
    ]
})
print("\n===== STATISTIQUES =====")
print(f"Accuracy : {accuracy:.4f}")
print(f"Precision : {precision:.4f}")
print(f"Recall : {recall:.4f}")
print(f"F1-score : {f1:.4f}")
print(f"Kappa : {kappa:.4f}")
# =====
# 8. EXPORT EXCEL
# =====
from sklearn.metrics import classification_report
report = classification_report(
    y_true,
    y_pred,
    output_dict=True,
    zero_division=0
)
report_df = pd.DataFrame(report).transpose()
excel_path = os.path.join(
    output_folder,
    "confusion_matrix.xlsx"
)

cm_df.to_csv(
    os.path.join(
        output_folder,
        "ConfusionMatrix.csv"
    ),
    sep=";",
    decimal=","
)
stats_df.to_csv(
    os.path.join(
        output_folder,
        "Statistics.csv"
    ),
    sep=";",
    decimal=","
    index=False
)
report_df.to_csv(
    os.path.join(
        output_folder,
        "PerClassMetrics.csv"
    ),
    sep=";",
    decimal=","
)
df.to_csv(
    os.path.join(
        output_folder,

```

```

    "Points.csv"
),
sep=";",
decimal=".",
index=False
)
html_path = os.path.join(
    output_folder,
    "rapport_validation.html"
)
with open(html_path, "w", encoding="utf-8") as f:
    f.write("""
<html>
<head>
<meta charset="utf-8">
<style>
body {font-family: Arial, sans-serif; margin: 20px;}
table {
    border-collapse: collapse;
    margin-bottom: 20px;
}
th, td {
    border: 1px solid #999;
    padding: 6px;
    text-align: center;
}
th {
    background-color: #dddddd;
}
</style>
</head>
<body>
""")
    f.write("<h1>Rapport de validation</h1>")
    f.write("<h2>Matrice de confusion</h2>")
    f.write(cm_df.to_html())
    f.write("<h2>Statistiques globales</h2>")
    f.write(stats_df.to_html(index=False))
    f.write("<h2>Statistiques par classe</h2>")
    f.write(report_df.to_html())
    f.write("</body></html>")
print("Rapport HTML créé :")
print(html_path)
print("=====")
print("MATRICE DE CONFUSION CRÉÉE")
print(excel_path)
print("=====")

```

Produire automatiquement un rapport scientifique

L'un des aspects les plus intéressants du script est la génération automatique de livrables.

À partir des résultats calculés, le programme crée :

- un tableau des points validés ;
- une matrice de confusion ;
- un tableau des indicateurs statistiques ;
- un rapport HTML complet ;
- plusieurs fichiers CSV directement exploitables dans Excel ou LibreOffice.

L'utilisateur obtient ainsi un véritable dossier de validation prêt à être archivé ou intégré dans un rapport d'étude.

Rapport de validation

Matrice de confusion

	Pred_1	Pred_2	Pred_3	Pred_4
Ref_1	25	0	1	0
Ref_2	1	18	1	0
Ref_3	1	1	88	2
Ref_4	0	1	3	23

Statistiques globales

Metric	Value
Accuracy	0.933333
Precision	0.933138
Recall	0.933333
F1-score	0.932953
Kappa	0.892553

Statistiques par classe

	precision	recall	f1-score	support
1	0.925926	0.961538	0.943396	26.000000
2	0.900000	0.900000	0.900000	20.000000
3	0.946237	0.956522	0.951351	92.000000
4	0.920000	0.851852	0.884615	27.000000
accuracy	0.933333	0.933333	0.933333	0.933333
macro avg	0.923041	0.917478	0.919841	165.000000
weighted avg	0.933138	0.933333	0.932953	165.000000

Cette capacité à produire automatiquement des documents reproductibles constitue l'un des grands avantages de Python dans un contexte scientifique.

La reproductibilité avant tout

L'intérêt principal de ces scripts n'est finalement pas le gain de temps.

Le véritable enjeu est la reproductibilité.

Chaque campagne de validation suit exactement le même protocole :

- même génération des points ;

- même extraction des valeurs ;
- mêmes indicateurs ;
- mêmes rapports.

Les résultats obtenus deviennent alors directement comparables d'une année à l'autre.

Cette cohérence est indispensable dans le cadre d'un observatoire environnemental ou d'un programme de suivi à long terme.

Du traitement ponctuel à l'observatoire environnemental

À petite échelle, ces scripts permettent simplement d'éviter des tâches répétitives.

Mais à plus grande échelle, ils deviennent les briques d'un véritable système de monitoring.

Le même principe pourrait être étendu à :

- plusieurs plages ;
- plusieurs dates ;
- plusieurs années ;
- différents types de classifications.

La validation devient alors une procédure normalisée capable de produire automatiquement des indicateurs comparables dans le temps.

Conclusion

L'automatisation ne consiste pas à remplacer le géomaticien par un programme.

Elle consiste à automatiser les tâches répétitives afin de laisser davantage de place à l'analyse, à l'interprétation et à la réflexion scientifique.

Les deux scripts développés présentés ici illustrent parfaitement cette philosophie. Le premier prépare les données de validation. Le second transforme les observations de terrain en indicateurs objectifs et reproductibles.

En combinant l'expertise humaine et l'automatisation, il devient possible de construire des chaînes de traitement robustes, documentées et adaptées au suivi environnemental à long terme.

C'est souvent à ce stade que la géomatique cesse d'être une simple utilisation d'outils pour devenir une véritable démarche scientifique appuyée sur des méthodes reproductibles.

Table des matières

Introduction.....	2
Comment est organisé cet ouvrage ?.....	3
Première partie — Les bases scientifiques et les données drone.....	3
Deuxième partie — Pré-traitements géomatiques sous QGIS.....	3
Troisième partie — SCP : le cœur de la chaîne de classification.....	4
Quatrième partie — Comparaison des méthodes de classification.....	4
Cinquième partie — Intelligence artificielle et amélioration des modèles.....	4
Sixième partie — Interpréter, valider et comprendre	5
Septième partie — Automatiser les traitements avec Python.....	5
Du pixel à la connaissance du territoire	5
Chapitre 1 — Les drones au service du suivi des algues vertes	6
Observer le littoral autrement	6
Une résolution spatiale très élevée.....	6
Une grande flexibilité d’acquisition.....	7
Une observation locale et ciblée	7
Chapitre 2 — Drone et satellite : des approches complémentaires	7
Deux échelles d’observation différentes.....	7
Sentinel-2 : l’observation à grande échelle.....	8
Le drone : l’observation à très haute résolution.....	8
Une complémentarité essentielle	8
Chapitre 3 — La chaîne de traitement des données drone	9
De l’acquisition à l’analyse	9
L’acquisition des images.....	9
La photogrammétrie et la production d’orthomosaïques	9
L’analyse dans un SIG	10
La classification des surfaces.....	10
La production d’indicateurs.....	10
Chapitre 4 — Comprendre les capteurs du drone multispectral	11
Deux systèmes de caméras complémentaires	11
La caméra RGB : une vision proche de l’œil humain.....	11
Les caméras multispectrales : l’analyse scientifique des surfaces.....	11
Pourquoi le multispectral ne comporte-t-il pas de bande bleue ?	12
Signature spectrale des algues vertes.....	13
Les particularités du milieu intertidal.....	13
Chapitre 5 — Les contraintes du milieu intertidal	14

Un environnement complexe à observer.....	14
Une forte variabilité spatiale.....	14
La présence d'eau résiduelle.....	15
Les effets du substrat humide.....	15
Les contraintes liées à l'éclairage.....	15
Une interprétation qui nécessite de la prudence.....	16
Chapitre 6 — L'importance des conditions d'acquisition.....	16
Pourquoi le moment de l'acquisition est-il crucial ?.....	16
L'importance des fenêtres de marée.....	17
Observer l'estran à marée basse.....	17
Les avantages des marées de vives-eaux.....	17
Une surface observable très variable.....	17
La question de la reproductibilité temporelle.....	18
Les conditions d'éclairage.....	18
Un facteur majeur en télédétection.....	18
Les avantages de la lumière diffuse.....	18
Les limites d'un fort ensoleillement.....	18
Les effets des ombres.....	19
Les reflets spéculaires.....	19
L'influence du vent.....	19
Un facteur souvent sous-estimé.....	19
Les conditions les plus favorables.....	20
Les conséquences d'une mauvaise acquisition.....	20
Des impacts directs sur les traitements.....	20
Un principe fondamental en télédétection.....	20
Conclusion de la séquence.....	20
Chapitre 7 — La calibration radiométrique : rendre les images scientifiquement exploitables.....	21
Une image "belle" n'est pas forcément une bonne donnée scientifique.....	21
DN et réflectance : deux notions différentes.....	21
Les DN : des valeurs brutes dépendantes des conditions d'acquisition.....	21
La réflectance : une mesure physique comparable.....	22
Pourquoi les DN ne suffisent pas.....	22
Le rôle du panneau de calibration.....	23
Convertir les DN en réflectance.....	23
Pourquoi le panneau est-il si important ?.....	23
Le capteur de lumière incident.....	23

Les problèmes fréquemment rencontrés	24
L'absence de calibration.....	24
Les variations d'éclairage pendant le vol	24
Les ombres et les nuages	24
Les conséquences pour les comparaisons temporelles	25
A retenir.....	25
Conclusion	25
Chapitre 8 — Les limites de l'interprétation en milieu intertidal	26
Comprendre les limites des observations multispectrales	26
Les limites physiques.....	26
La présence permanente d'eau résiduelle	26
Les effets de la turbidité.....	26
L'influence de la profondeur d'eau	27
Les limites biologiques	27
La présence de biofilms.....	27
Les algues partiellement immergées.....	27
Les limites radiométriques	28
Les effets du substrat humide	28
Les perturbations liées aux ombres	28
Les conséquences pour l'interprétation.....	29
Ce que représentent réellement les cartes multispectrales	29
A retenir.....	29
Les implications scientifiques.....	29
Chapitre 9 — Pourquoi les pré-traitements sont essentiels	30
La partie invisible des projets de télédétection	30
Pourquoi cette étape est-elle si importante ?	30
Les principales sources d'erreurs	31
Les problèmes d'alignement entre bandes multispectrales	31
Les problèmes radiométriques.....	31
Les surfaces perturbatrices	32
Chapitre 10 — Les produits issus de la photogrammétrie	32
Des données multiples pour des usages différents.....	32
L'orthomosaïque RGB.....	32
L'orthomosaïque multispectrale	33
Les bandes spectrales individuelles.....	34
Le modèle numérique de surface.....	34

Chapitre 11 — Organiser correctement un projet QGIS	35
Une organisation indispensable	35
Le dossier des orthomosaïques	35
Le dossier multispectral	35
Le dossier DSM	36
Le dossier des indices	36
Le dossier des scripts.....	36
Chapitre 12 — Importer et vérifier les rasters dans QGIS.....	37
Charger les données dans le projet.....	37
Vérifier la cohérence géométrique des données	37
Vérifier le système de coordonnées.....	37
Vérifier la résolution spatiale	37
Vérifier l'étendue spatiale	38
La vérification visuelle	38
Chapitre 13 — Vérification des valeurs radiométriques	38
Pourquoi vérifier les valeurs des rasters ?	38
Les plages de valeurs attendues.....	39
Deux formats couramment rencontrés.....	39
Vérifier les valeurs dans QGIS	39
Les propriétés du raster	39
L'inspection directe des pixels.....	40
Le problème des données 8 bits	40
Des images adaptées à la visualisation, mais pas à l'analyse scientifique	40
Les conséquences de la compression.....	40
Chapitre 14 — Comprendre les signatures spectrales des surfaces littorales.....	41
Les bandes du drone multispectral	41
Les signatures spectrales typiques	41
L'eau	41
Le sable	42
Les algues vertes.....	42
Une astuce de visualisation multispectrale.....	42
Les compositions colorées dans QGIS	42
Chapitre 15 — Utilisation du modèle numérique de surface (DSM)	43
Une information complémentaire aux données spectrales.....	43
L'intérêt du DSM pour l'analyse de l'estran.....	43
Comprendre les micro-structures du terrain	43

Visualiser le relief dans QGIS.....	44
Le calcul d'ombrage du relief	44
Une approche multi-source de l'interprétation	44
Combiner données spectrales et topographiques	44
Chapitre 16 — Comprendre les histogrammes des rasters	45
Pourquoi analyser les histogrammes ?	45
Comprendre la forme d'un histogramme	46
Une représentation de la distribution des pixels	46
Détecter les phénomènes de saturation.....	46
Quand le capteur atteint ses limites	46
Identifier les valeurs aberrantes	47
Détecter les anomalies radiométriques	47
Détecter des problèmes de calibration radiométrique.....	47
Une image visuellement correcte peut être scientifiquement incohérente.....	47
Interpréter les histogrammes des images littorales	48
Une distribution liée à la diversité des surfaces.....	48
Les histogrammes comme outil de diagnostic	48
Une étape essentielle avant toute analyse	48
Chapitre 17 — Inspection des pixels et compréhension des signatures spectrales	49
Observer directement les valeurs des pixels.....	49
Utilisation de l'outil Value Tool	49
Explorer les surfaces de l'image	49
Comparer différentes surfaces littorales	50
L'eau	50
Le sable	50
Les algues vertes.....	50
Pourquoi cette inspection est-elle importante ?	51
Comprendre les données avant les traitements	51
Les signatures spectrales typiques des algues	51
Le comportement des différentes bandes	51
Chapitre 18 — L'importance de l'alignement des bandes	52
Une condition essentielle des analyses multispectrales	52
Pourquoi cette hypothèse est-elle si importante ?.....	52
Les indices reposent sur plusieurs bandes	52
Les particularités des capteurs multispectraux.....	52
Plusieurs capteurs pour plusieurs bandes.....	52

Les origines possibles des décalages	53
Des erreurs parfois très faibles mais importantes	53
Les conséquences d'un mauvais alignement	53
Des indices incorrects et du bruit spatial	53
Une sensibilité encore plus forte avec les drones	54
L'effet de la très haute résolution	54
Vérifier l'alignement dans QGIS	54
Une vérification visuelle indispensable	54
Chapitre 19 — Alignement raster avec GDAL	55
Harmoniser les rasters avant les analyses	55
Pourquoi cet alignement est-il nécessaire ?	55
Des différences parfois invisibles mais problématiques	55
Fonctionnement de l'outil Align Raster	55
Utiliser un raster de référence	55
Harmoniser la résolution spatiale	56
Une étape essentielle avant les indices	56
Résultat final	56
Des rasters parfaitement comparables	56
Chapitre 20 — Création des indices spectraux	57
Transformer les bandes spectrales en indicateurs interprétables	57
Le NDVI : l'indice de végétation le plus utilisé	57
Principe du NDVI	57
Le principe physique du NDVI	57
Interprétation des valeurs	58
Calcul du NDVI dans QGIS	58
Les limites du NDVI	58
Le NDRE : exploiter la bande Red Edge	59
Un indice plus sensible à la chlorophylle	59
Différences entre NDVI et NDRE	59
Calcul du NDRE dans QGIS	60
Le GNDVI : utiliser la bande verte	60
Une variante sensible à la chlorophylle	60
Intérêt du GNDVI	60
Calcul du GNDVI	61
Le NDWI : détecter l'eau	61
Un indice dédié aux surfaces humides	61

Le comportement spectral de l'eau	61
Utilité du NDWI sur l'estran	61
Calcul du NDWI.....	62
Le Simple Ratio (SR).....	62
Un indice historique de végétation	62
Caractéristiques du Simple Ratio	62
Le Chlorophyll Index (CI).....	63
Estimer la chlorophylle	63
Intérêt du CI pour les algues	63
Comparer plusieurs indices	63
Aucun indice n'est parfait.....	63
Une interprétation toujours critique.....	64
Les indices ne sont pas des mesures directes	64
Chapitre 21 — Création de masques raster	64
Pourquoi utiliser des masques ?	64
Le principe d'un masque raster	65
Isoler les zones utiles.....	65
Le masque d'eau.....	65
Pourquoi l'eau pose problème	65
Détection de l'eau avec le NDWI.....	65
Utiliser le contraste Green / NIR	66
Création du masque dans QGIS.....	66
Utiliser la Raster Calculator	66
Application du masque.....	67
Exclure les pixels perturbateurs	67
Les principaux types de masques.....	67
Des masques adaptés aux différentes perturbations	67
Chapitre 22 — Le masque des ombres.....	68
Pourquoi les ombres perturbent les analyses.....	68
Détection simple des ombres.....	68
Utiliser un seuil sur le NIR.....	68
Une approche plus robuste	68
Utiliser plusieurs bandes	68
Les limites du masque d'ombre	69
Des confusions possibles avec l'eau	69
Chapitre 23 — Distinguer l'ombre et l'eau	69

Deux surfaces sombres mais des signatures différentes	69
Une approche combinée recommandée.....	70
Combiner NDWI et NIR.....	70
Pourquoi cette méthode fonctionne-t-elle ?	70
Différence entre structure spectrale et absence de lumière	70
Combiner plusieurs masques	71
Conserver uniquement les zones exploitables.....	71
Une étape essentielle des pré-traitements.....	71
Nettoyer les données avant les analyses	71
Chapitre 24 — Inspection des signatures spectrales avec SCP	71
Comprendre le concept de signature spectrale	71
Le plugin SCP dans QGIS.....	72
Un outil dédié à l'analyse multispectrale	72
Création des zones d'intérêt (ROI)	72
Définir des surfaces représentatives.....	72
L'importance du choix des ROI.....	73
Extraction des signatures spectrales.....	73
Transformer les pixels en courbes spectrales	73
Visualisation des courbes spectrales.....	73
Lire les signatures comme des profils radiométriques	73
Exemples de signatures spectrales.....	74
L'eau	74
Le sable.....	74
Les algues vertes.....	74
Interpréter les courbes spectrales	75
Comparer les classes	75
Pourquoi cette étape est essentielle.....	75
Préparer correctement une classification	75
Une idée fondamentale en télédétection	75
On ne classe pas directement des objets visibles	75
Chapitre 25 — Les limites de l'interprétation spectrale	76
Des signatures parfois ambiguës.....	76
Une distinction essentielle	76
Cartographier une signature, pas directement les algues	76
Chapitre 26 — Les 10 erreurs fréquentes en pré-traitement sous QGIS.....	77
Pourquoi les erreurs de pré-traitement sont critiques	77

Des bandes multispectrales mal alignées	77
Une erreur fréquente mais souvent invisible.....	77
Utiliser des données 8 bits	78
Des images “jolies” mais scientifiquement limitées	78
Mélanger plusieurs projections.....	78
Un problème classique dans les SIG	78
Ignorer la taille des pixels.....	79
La résolution influence directement les calculs	79
Calculer des indices sur des zones d’eau.....	79
Une source importante de bruit.....	79
Ne pas inspecter les valeurs des pixels	80
Une erreur fréquente chez les débutants	80
Ignorer les zones saturées.....	80
Quand le capteur atteint ses limites	80
Interpréter directement un indice spectral.....	80
Un indice n’est pas un objet réel.....	80
Oublier les conditions d’acquisition	81
Des variations parfois indépendantes du phénomène étudié	81
Ne pas vérifier les signatures spectrales	81
Une étape indispensable avant la classification.....	81
Conclusion	82
Ce que nous allons apprendre dans cette partie	83
Chapitre 27 — Pourquoi utiliser SCP ?	83
Du raster multispectral à la carte thématique	83
Les limites des outils raster classiques	84
Ce qu’il est possible de faire sans SCP.....	84
Les difficultés rencontrées sans outil spécialisé.....	84
Ce que SCP apporte	85
Une plateforme complète de classification	85
Les Band Sets.....	85
Organiser les bandes multispectrales	85
Les indices spectraux dans SCP	86
Calculer rapidement les indicateurs.....	86
Les ROI et les signatures spectrales	86
Construire les classes d’apprentissage.....	86
L’intérêt de SCP pour le suivi des algues vertes	87

Produire des cartes objectives	87
Pourquoi cette approche est importante scientifiquement	87
Une méthode reproductible.....	87
A retenir.....	88
Chapitre 28 — Préparer les bandes multispectrales dans Semi-Automatic Classification Plugin	88
Le rôle central des <i>Band Sets</i>	88
Une condition indispensable : la cohérence géométrique	88
Les bandes du DJI Mavic 3 Multispectral	89
Création du <i>Band Set</i> dans SCP	89
Les erreurs fréquentes lors de la création du <i>Band Set</i>	90
Le résultat attendu	90
Chapitre 29 — Calcul des indices spectraux dans SCP	90
Transformer les bandes en indicateurs exploitables	90
Où calculer les indices dans SCP ?	91
Le calcul du NDVI.....	91
Le calcul du NDRE	91
Le calcul du GNDVI	92
Les étapes pratiques dans SCP	92
Interpréter les indices avec prudence.....	92
Chapitre 30 — Création d'indices personnalisés pour le suivi des algues vertes	93
Les limites des indices spectraux classiques	93
Pourquoi les algues vertes sont difficiles à détecter.....	94
L'idée des indices personnalisés.....	94
Combiner plusieurs informations spectrales	95
L'importance des conditions logiques.....	95
Une démarche expérimentale et scientifique.....	96
Le rôle de QGIS et de SCP	96
A retenir.....	97
Chapitre 31 — Les ROI et la stratégie d'échantillonnage : le cœur de la classification supervisée	97
Pourquoi les ROI sont si importantes.....	97
Le principe des ROI	98
La notion d'homogénéité	98
L'importance de la répartition spatiale	98
Capturer la variabilité des classes	99
La taille des ROI	99
Les erreurs fréquentes	100

Les particularités du site de Guéthary.....	100
Procédure détaillée pour créer une ROI “algues”	101
Préparer correctement l’affichage	101
Identifier des zones fiables.....	101
Dessiner les ROI dans SCP	101
Multiplier les ROI pour une même classe	102
Vérifier les signatures spectrales.....	102
Contrôler les ROI avant la classification	102
Une stratégie avancée : les sous-types de classes	102
Sauvegarder les ROI.....	103
A retenir.....	103
Chapitre 32 — La matrice de confusion : évaluer la qualité d’une classification	104
Pourquoi valider une classification ?	104
Le principe général	104
Comprendre la structure de la matrice	105
Interpréter les erreurs.....	105
La précision globale.....	106
Les limites de la validation	106
Une règle fondamentale : ne jamais utiliser les ROI	106
Création des points de validation dans QGIS	107
Une méthode simple et manuelle	107
Une approche plus scientifique : les points aléatoires.....	107
Récupérer la classe prédite	108
Construire la matrice.....	108
Les règles importantes pour une bonne validation	109
Une idée fondamentale.....	109
Utilisation d’un fichier Excel de validation	109
Automatiser les calculs.....	109
Chapitre 33 — Peut-on distinguer les algues vertes des algues rouges ?	110
Une question fréquente en télédétection littorale.....	110
Pourquoi la séparation est difficile.....	110
Une limite liée au capteur	111
Tester des indices plus spécifiques	112
Explorer les différences spectrales subtiles	112
Utiliser le GNDVI.....	112
Tester des ratios simples.....	112

Les indices personnalisés	113
Une approche souvent plus intéressante.....	113
L'approche multi-indices	113
Comparer plusieurs représentations simultanément	113
Ajouter davantage d'informations à la classification	114
Les indices comme nouvelles variables.....	114
Le rôle des algorithmes	114
Certains modèles gèrent mieux les signatures proches.....	114
L'importance de l'information spatiale	115
Une approche souvent sous-estimée.....	115
La véritable solution scientifique	115
Quand les bandes multispectrales ne suffisent plus.....	115
A retenir.....	116
Chapitre 34 — Peut-on utiliser la bande bleue du RGB ?	116
Une question logique... mais plus complexe qu'elle n'en a l'air	116
Ce que contient réellement l'orthomosaïque RGB	116
Extraire la bande bleue dans QGIS	117
Vérifier la cohérence des valeurs	117
Le problème fondamental : le RGB n'est pas calibré	118
Une différence majeure avec les bandes multispectrales	118
Pourquoi cela pose problème	118
Peut-on malgré tout utiliser cette bande ?	119
Oui... mais avec prudence	119
Les cas où cette approche peut être utile.....	119
Les situations où il faut rester prudent	119
Les bonnes pratiques.....	120
Normaliser les valeurs	120
Comparer les résultats	120
Garder une approche critique	120
A retenir.....	121
Chapitre 35 — Les trois grandes familles d'algorithmes de classification	121
Comprendre la logique des classifications supervisées	121
Les méthodes classiques : comparer les pixels aux signatures moyennes	122
Une approche statistique historique.....	122
Le principe du Maximum Likelihood	122
Le principe du Minimum Distance.....	122

Les avantages des méthodes classiques	123
Leurs limites.....	123
Les méthodes spectrales : analyser la forme des signatures	124
Une approche plus “physique”	124
Le principe du SAM.....	124
Les avantages du SAM.....	124
Les limites du SAM	125
Le machine learning : apprendre des relations complexes.....	125
Une approche moderne basée sur l’apprentissage	125
Le Random Forest.....	125
Le SVM.....	126
Les avantages du machine learning	126
Les limites du machine learning	126
Comparer les trois familles.....	127
Trois philosophies différentes	127
Une idée essentielle	127
A retenir.....	127
Chapitre 36 — Introduction à Random Forest en télédétection	128
Pourquoi Random Forest est devenu si populaire.....	128
Qu’est-ce que Random Forest ?	129
Une “forêt” d’arbres de décision	129
Le principe du vote	129
Pourquoi “Random” ?	129
Le rôle de l’aléatoire.....	129
Pourquoi Random Forest fonctionne bien en télédétection	130
Gérer des données complexes	130
Les variables utilisées par Random Forest	130
Bien plus que les bandes spectrales.....	130
En analyse orientée objet.....	131
Les paramètres importants	131
Le nombre d’arbres	131
La profondeur des arbres	132
Le nombre de variables utilisées	132
Les limites de Random Forest	132
La qualité des ROI reste fondamentale	132
Le déséquilibre des classes.....	132

Le sur-apprentissage	133
Évaluer un modèle Random Forest	133
Pourquoi la précision globale ne suffit pas	133
La précision par classe	134
Le rappel (<i>Recall</i>)	134
Le F-score	134
L'indice Kappa	135
Un exemple typique en télédétection	135
Quand l'eau domine les résultats	135
Workflow typique avec Random Forest	136
Une démarche complète	136
A retenir	136
Chapitre 37 — Mise en œuvre de classifications alternatives	137
Aller au-delà des classifications pixel par pixel	137
Pourquoi comparer plusieurs méthodes ?	137
Les outils utilisés	138
Chapitre 38 — Présentation de Orfeo ToolBox	138
Un outil conçu pour la télédétection avancée	138
Pourquoi OTB est intéressant pour les images drone	139
Intégration dans QGIS	139
Chapitre 39 — Introduction à la segmentation orientée objet	139
Pourquoi segmenter une image ?	139
Le principe de la segmentation	140
L'algorithme LargeScaleMeanShift	140
Une méthode particulièrement adaptée au suivi des algues	140
Une nouvelle manière de voir l'image	141
A retenir	141
Chapitre 40 — Classification Random Forest avec Orfeo ToolBox	141
Une méthode devenue incontournable en télédétection	141
Le principe général du Random Forest	142
Pourquoi cette méthode fonctionne bien en télédétection	143
Étape 1 — L'apprentissage supervisé	143
Fournir des exemples au modèle	143
L'importance des données d'apprentissage	144
Étape 2 — Construction du modèle	144
Une forêt d'arbres de décision	144

Les variables utilisées par le modèle	145
Étape 3 — Production de la carte classifiée	145
Appliquer le modèle à toute l'image	145
Les probabilités de confiance	146
Une méthode puissante... mais pas magique	146
Les limites du Random Forest	146
A retenir.....	146
Chapitre 41 — Évaluer un modèle Random Forest : comment interpréter les résultats ?.....	147
Pourquoi l'évaluation est indispensable	147
La précision globale (<i>Overall Accuracy</i>).....	148
L'indicateur le plus connu	148
Pourquoi la précision globale peut être trompeuse	148
La précision par classe (<i>Precision</i>)	149
Une question fondamentale	149
Interpréter la précision.....	149
Le rappel (<i>Recall</i>)	149
Une autre question essentielle	149
Comprendre le rappel	150
La différence entre précision et rappel	150
Deux indicateurs complémentaires.....	150
Le F-score.....	151
Combiner précision et rappel	151
L'indice Kappa.....	151
Une métrique très importante en télédétection	151
Interpréter le Kappa	151
Un exemple typique de piège statistique.....	152
Exemple simplifié en télédétection littorale	152
Quand les algues deviennent difficiles à détecter	152
Pourquoi les classes rares sont difficiles	152
Le rôle central de la matrice de confusion	153
Comprendre les erreurs du modèle	153
A retenir.....	153
Chapitre 42 — Classification avec GRASS GIS	154
Un environnement historique de la télédétection.....	154
La classification supervisée avec <i>i.maxlik</i>	154
Une approche statistique historique.....	154

Le rôle des signatures spectrales	155
L'hypothèse de distribution normale	155
Une classification probabiliste	155
Les avantages du Maximum Likelihood	156
Ses limites dans les milieux littoraux.....	156
La segmentation avec <code>i.segment</code>	157
Passer d'une logique pixel à une logique objet.....	157
Le principe de la segmentation	157
Le seuil de similarité	157
La taille minimale des segments	158
L'influence de la résolution des images drone.....	158
Une chaîne complète de traitement	159
Combiner segmentation et classification	159
Le rôle du post-traitement	159
Une évolution des méthodes de télédétection.....	160
Du pixel vers l'objet.....	160
A retenir.....	160
Chapitre 43 — Les techniques modernes de deep learning	161
Une nouvelle génération d'approches en télédétection	161
Pourquoi le deep learning change la télédétection	161
Aller au-delà des signatures spectrales.....	161
Une approche particulièrement adaptée aux images drone	162
Les usages actuels du deep learning en télédétection.....	163
Une généralisation rapide des applications	163
Le plugin Deepness.....	163
Une passerelle entre QGIS et l'intelligence artificielle.....	163
Les principales fonctionnalités de Deepness	164
La segmentation d'images.....	164
La détection d'objets.....	164
La régression spatiale	164
Compatibilité avec les données géospatiales.....	165
Un fonctionnement intégré dans QGIS	165
Le format ONNX	166
Un format standard pour les modèles IA.....	166
Le Model Zoo	166
Une bibliothèque de modèles prêts à l'emploi	166

Une interface pensée pour les utilisateurs SIG	167
Rendre le deep learning accessible	167
Une nouvelle étape dans l'évolution de la télédétection	167
Des perspectives importantes... mais aussi des limites	167
A retenir.....	168
Chapitre 44 — Amélioration d'un modèle de segmentation par fine-tuning.....	168
Une approche beaucoup plus réaliste du deep learning	168
Pourquoi ne pas entraîner un modèle depuis zéro ?	169
Le problème des entraînements complets.....	169
Le principe du fine-tuning	169
Réutiliser un modèle existant.....	169
Une approche très utilisée en pratique.....	170
Le transfert d'apprentissage	170
Améliorer un modèle existant : l'art du fine-tuning.....	171
Pourquoi réinventer ce qui existe déjà ?.....	171
Le principe du transfert d'apprentissage	172
Préparer les données : une étape essentielle	172
Utiliser la puissance du cloud.....	172
Reprendre l'entraînement d'un modèle existant.....	173
Un apprentissage en douceur	173
De PyTorch à QGIS.....	174
Comprendre plutôt que simplement utiliser	174
À retenir.....	174
Les avantages pédagogiques du fine-tuning	175
Une approche beaucoup plus accessible	175
Une approche plus proche du monde professionnel.....	175
A retenir.....	175
Chapitre 45 — Comparer les méthodes : apprendre à regarder les cartes autrement.....	177
Aucune méthode n'est parfaite	177
Les cartes racontent toujours quelque chose	178
La question du bruit	178
Voir le territoire... ou voir l'algorithme ?.....	179
Les limites des signatures spectrales	179
Les petites nappes d'algues : un détail... ou un enjeu majeur ?	180
Le poids des contraintes techniques	180
Le littoral : un environnement instable.....	181

Les classifications comme outils d'interprétation	181
Conclusion	182
Chapitre 46 — Apprendre à douter des cartes	182
La discussion scientifique des résultats de classification.	182
La télédétection comme discipline d'interprétation	183
Les cartes révèlent aussi les limites des méthodes	184
L'importance des erreurs	184
Le littoral : un environnement profondément ambigu	185
Les paramètres changent le paysage	185
Le piège des algorithmes sophistiqués.....	186
Le rôle du terrain	186
La matrice de confusion : regarder les erreurs en face.....	187
Une réflexion sur le choix des méthodes	187
Une transformation du regard	188
Conclusion	188
Chapitre 47 — Du classement des pixels à la compréhension des dynamiques littorales	189
La nécessité de valider les résultats	189
Apprendre à mesurer l'erreur	190
La matrice de confusion : un miroir de la classification	190
Transformer une carte en indicateur environnemental	191
Lire le littoral à travers les chiffres	191
Observer le temps à travers les cartes	192
Comprendre les changements	192
Les rythmes saisonniers du littoral.....	193
Une classification n'est jamais une fin	193
Chapitre 48 — Automatiser la validation : quand Python devient assistant du géomaticien	194
Le rôle irremplaçable de l'expert	195
Première étape : préparer automatiquement les points de validation	195
Le moment où l'humain reprend la main.....	198
Transformer les observations en indicateurs.....	198
Produire automatiquement un rapport scientifique	201
La reproductibilité avant tout	202
Du traitement ponctuel à l'observatoire environnemental	203
Conclusion	203