

Téledétection

TÉLÉDÉTECTION

ANDERS KNUDBY



Téledétection Copyright © 2021 by Anders Knudby is licensed under a Licence Creative Commons Attribution 4.0 International, except where otherwise noted.

Sauf avis contraire, ce livre est mis à disposition sous la licence CC BY 4.0.

TABLE DES MATIÈRES

Bienvenue à la télédétection	vii
------------------------------	-----

Partie I. Corps principal

1. Introduction	1
2. Mesures radiométriques	18
3. Calculs de la luminance TOA et de la réflectance TOA	26
4. Interactions d'EMR avec l'atmosphère et la surface de la Terre	32
5. Visualisation et manipulation d'images satellites	43
6. Classification	60
7. Évaluation d'exactitude	84
8. Détection des changements	96
9. Végétation et feu	106
10. Photogrammétrie et Structure-from-Motion	118

BIENVENUE À LA TÉLÉDÉTECTION

Bienvenue dans ce livre d'introduction à la télédétection, qui se veut l'équivalent d'un manuel pour un cours universitaire de premier cycle. Vous êtes libre d'utiliser ce livre comme bon vous-même. Le livre est séparé en « chapitres » qui correspondent aux lectures que mes propres étudiants doivent faire pour préparer les cours en plus d'exercices de laboratoire chaque semaine.

Pour naviguer dans le livre, cliquez sur le bouton « Contents » dans le coin supérieur gauche, puis cliquez sur le signe « + » à côté de l'indication « Corps principal ». Cliquez ensuite sur le chapitre que vous souhaitez lire.

Si vous cherchez un manuel publié pour compléter ce document, je peux vous en recommander plusieurs. Ils contiennent tous des informations supplémentaires qui ne sont pas présentes dans ce texte. Certains sont spécialisés, comme vous pouvez le voir dans leurs titres, d'autres sont généraux pouvant être utilisés comme manuels pour ce cours. La télédétection est un domaine qui, d'une certaine manière, se développe rapidement (de nouveaux capteurs sont lancés dans l'espace en permanence), et d'une autre manière, se développe lentement (de nombreuses techniques de base pour le traitement numérique des images ont été développées il y a plusieurs décennies et sont toujours utilisées). Veuillez garder ce détail à l'esprit si vous lisez un vieux manuel. Voici quelques livres que vous devriez envisager de consulter pour obtenir des informations complémentaires (cette liste n'est pas complète !):

En Anglais:

- J.B. Campbell et R.H. Wynne: *Introduction to Remote Sensing*. 5^{ème} édition, 2011.
- J.R. Jensen: *Introductory Digital Image Processing*. 4^{ème} édition, 2015.
- J.R. Jensen: *Remote Sensing of the Environment: An Earth Resource Perspective*. 2^{ème} édition, 2006.
- T. Lillesand, R.W. Kiefer and J.W. Chipman: *Remote Sensing and Image Interpretation*. 7^{ème} édition, 2015.
- E. Chuvieco: *Fundamentals of Satellite Remote Sensing: An Environmental Approach*. 2^{ème} édition, 2016.
- H.G. Jones and R.A. Vaughan: *Remote Sensing of Vegetation: Principles, Techniques, and Applications*. 1^{ère} édition, 2010.
- I. Woodhouse: *Introduction to Microwave Remote Sensing*. 1^{ère} édition, 2006.

En Français:

- M.-C. Girard and C.M. Girard: *Traitement des données de télédétection*. 2^{ème} édition, 2010.

De plus, le site Web du Centre canadien de télédétection propose des tutoriels faciles à utiliser auxquels vous pouvez accéder gratuitement. N'oubliez pas qu'Internet est votre ami lorsqu'il s'agit d'apprendre la

télédetection — il existe de nombreux sites où les gens ont déjà posé les mêmes questions que celles que vous poserez en suivant ce cours, et d'autres personnes, généralement bien informées, y ont répondu.

1.

INTRODUCTION

Nous espérons que vous avez poursuivi ce cours puisque vous avez déjà une idée de ce qu'est la télédétection en espérant que vous voulez en savoir plus. Mais qu'est-ce que la télédétection exactement ?!?

Qu'est-ce que la télédétection?

Certaines définitions très larges vont jusqu'à dire que la télédétection consiste à étudier quelque chose sans le « toucher ». Bien que c'est une réalité dans son sens le plus fondamental, de telles définitions ne sont pas vraiment très utiles à comprendre l'essence de la télédétection. Sans ce lancer dans des discussions philosophiques sur le sens des mots « étudier » et « toucher », je pense que nous pouvons convenir que vous ne faites pas de la télédétection en ce moment, puisque vous lisez ce document sur un écran d'ordinateur... même si vous apprenez ce que dit ce document sans le toucher. De manière plus appropriée pour ce cours, nous pouvons définir la télédétection comme l'utilisation d'un instrument (le capteur) pour recueillir des informations sur la surface de la Terre (ou d'autres parties de la Terre, ses océans ou son atmosphère, ou d'autres planètes d'ailleurs) sur de grandes étendues et à une certaine distance d'un lieu. Les exemples typiques incluent l'utilisation de l'imagerie satellitaire ou de la photographie aérienne, mais la détection acoustique (par exemple du fond marin) et d'autres technologies qui ne créent pas de données sous la forme d'une « image » font également partie du grand domaine de la télédétection. Ce type de définition signifie qu'un autre terme, « observation de la Terre », est souvent utilisé de manière interchangeable avec « télédétection ».

Revue des technologies de télédétection

Détection passive

Il existe de nombreuses façons de séparer les technologies de télédétection existantes en catégories précises. Certaines technologies s'appuient sur l'énergie ambiante (par exemple, l'énergie naturellement présente dans l'environnement, comme la lumière du soleil). Elles sont appelées technologies de *détection passive*.

Détection active

D'autres fournissent leur propre source d'énergie, en émettant une impulsion énergétique vers une zone cible en enregistrant la partie de cette impulsion qui est réfléchi vers le capteur (comme les sonars utilisés dans les sous-marins). Ces technologies sont appelées technologies de *détection active*.

Détection optique

On peut également distinguer les technologies de télédétection en fonction du type d'énergie qu'elles utilisent. Certaines technologies reposent sur la lumière visible (par exemple, la photographie aérienne traditionnelle), tandis que d'autres technologies étendent la gamme des longueurs d'onde détectées à l'ultraviolet et à l'infrarouge (qui est également le type de « lumière » utilisé par la télécommande de votre téléviseur). Lorsqu'elles mesurent la lumière solaire réfléchi dans ces longueurs d'onde, ces technologies sont appelées technologies de *détection optique passive* – optique parce que les instruments utilisent l'optique classique dans leur réalisation. Les images satellites que l'on voit sur Google Earth, ou ceux servant de fond aux prévisions météorologiques télévisées sont des images créées par télédétection optique passive. Une forme de *télédétection optique active* est appelée lidar (light detection and ranging). Il s'agit d'émettre une impulsion laser (la partie active) vers une cible et de mesurer le temps nécessaire pour que cette impulsion frappe quelque chose et qu'une partie de celle-ci revienne pour être mesurée par le capteur. En général, l'intensité de l'impulsion de retour est mesurée aussi.

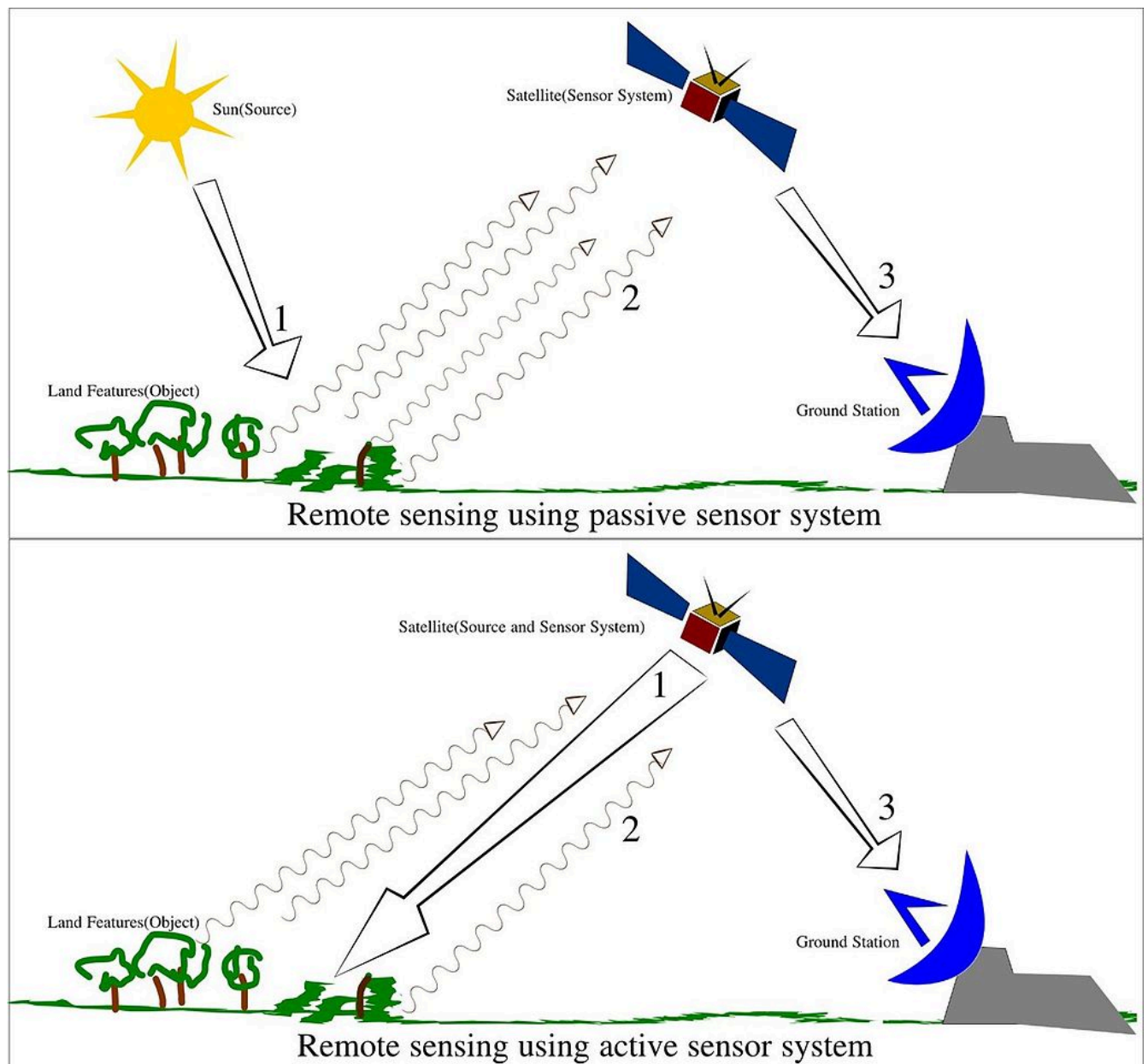


Figure 1: La différence entre les systèmes de télédétection passifs et actifs. Source: Remote Sensing Illustration par Arkarjun, Wikimedia Commons, CC BY-SA 3.0.

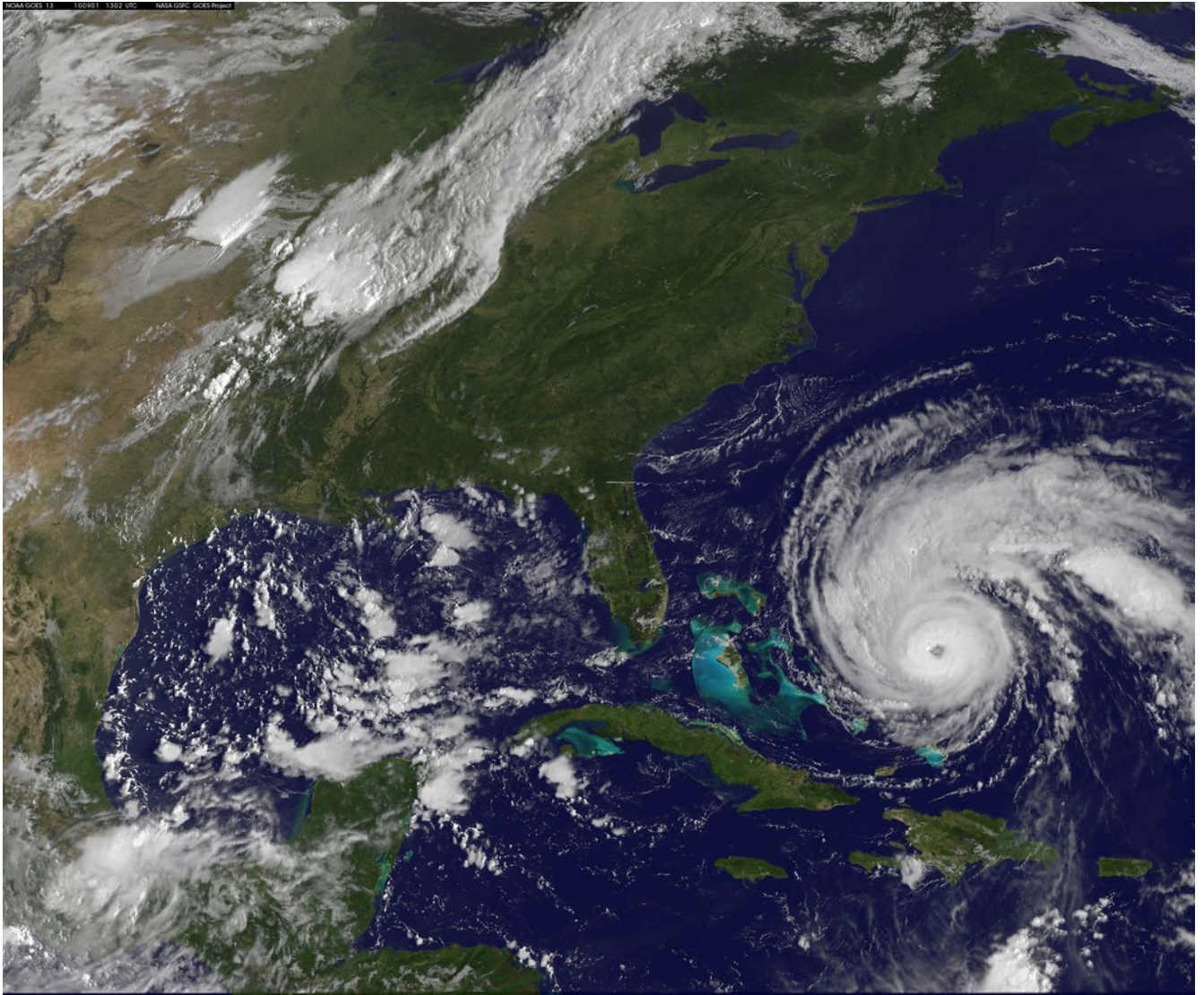


Figure 2: NASA Satellite Captures Hurricane Earl on September 1, 2010 par NASA Goddard Photo and Video, Flickr, CC BY 2.0.

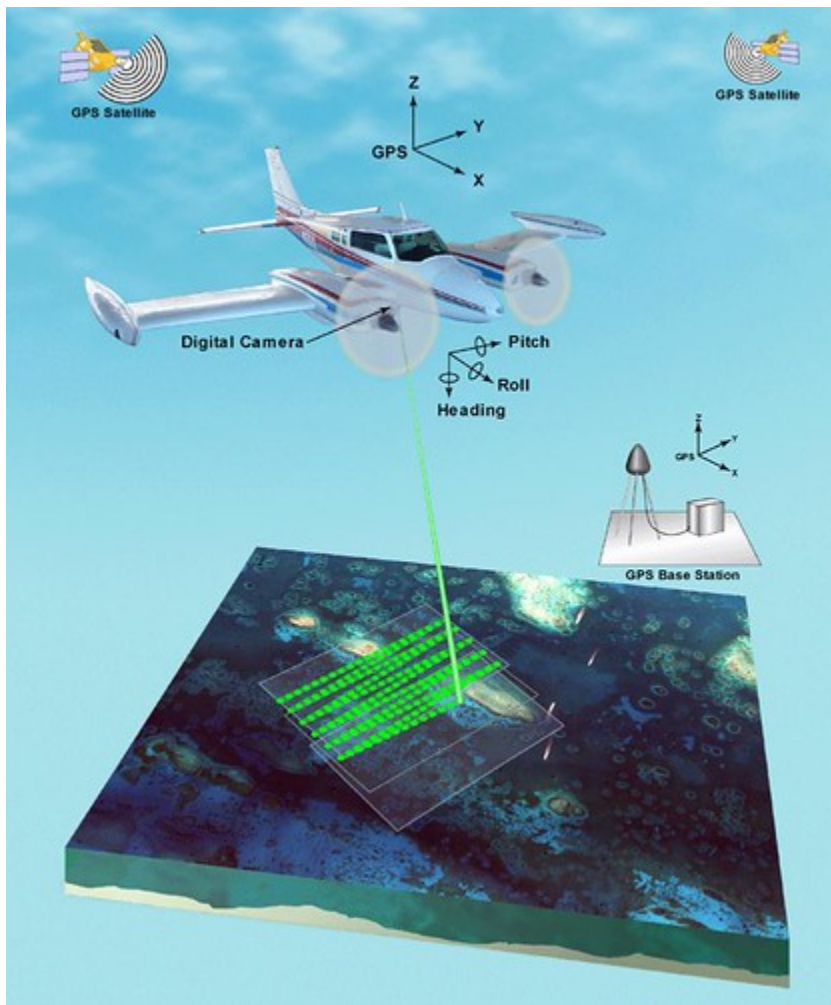


Figure 3: Schéma conceptuel d'un lidar aéroporté. La localisation et l'orientation de l'instrument sont fournies par les instruments GPS et IMU de l'avion. Grâce à ces connaissances, les distances mesurées entre l'instrument et les objets au sol peuvent être converties pour trouver l'emplacement de ces objets, fournissant ainsi une image 3D de la surface de la Terre. fis01335 par NOAA Photo Library, Flickr, CC BY 2.0.

Détection thermique

La détection thermique thermique est semblable à la détection optique passive, sauf que l'énergie mesurée par le capteur se situe dans la « région thermique » du spectre des longueurs d'onde. C'est le genre de détection thermique passive que vous faites vous-même lorsque vous mettez vos mains près du feu de camp pour sentir sa chaleur. Nous pouvons donc sentir le rayonnement thermique, mais nous ne pouvons pas le voir avec nos yeux comme nous pouvons voir la lumière du soleil. Le Soleil produit beaucoup plus de rayonnement thermique que la Terre (il est beaucoup plus chaud, après tout), mais le soleil est également très éloigné, donc si vous construisez un capteur thermique en le plaçant dans un avion ou sur un satellite, en plus de diriger vers la Terre, le capteur va faire en sorte que le rayonnement thermique émis par la terre écrase les petites quantités de

rayonnement thermique qui l'ont atteint du Soleil. Par conséquent, la télédétection thermique est utile pour mesurer la température de la surface de la Terre.

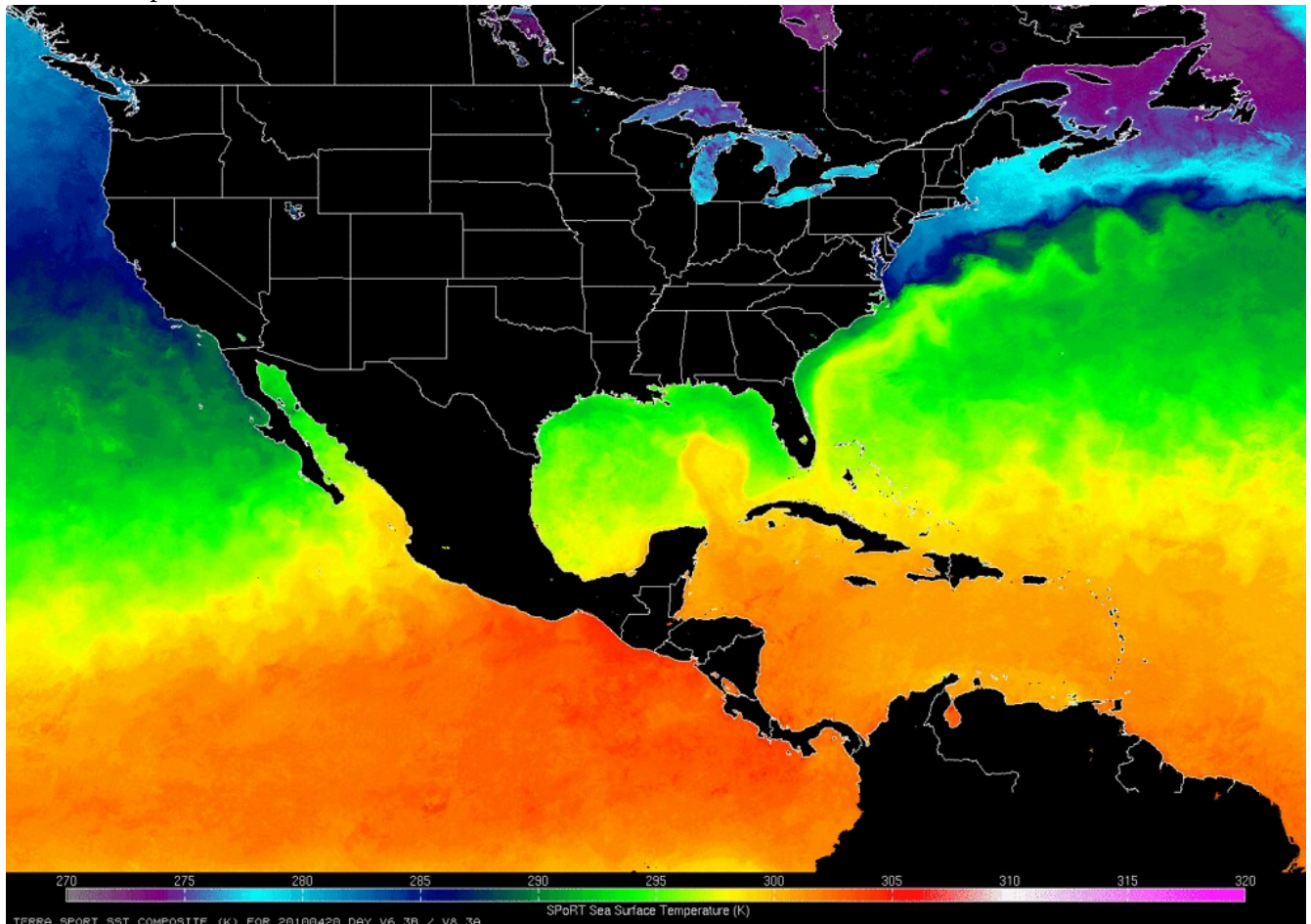


Figure 4: Cette carte montre la distribution de la température de surface de la mer pour une journée en 2010. La température de surface de la mer est mesurée régulièrement par plusieurs satellites, dont MODIS Terra qui a produit les données présentées ici. Sea Surface Temperature Map (NASA, SPoRT, 04/23/10) par NASA's Marshall Space Flight Center, Flickr, CC BY-NC 2.0.

Détection par micro-ondes

En ce qui concerne les types de rayonnement que nous ne pouvons ni voir ni ressentir, la **détection des micro-ondes** est une partie importante de la télédétection. La Terre émet constamment de petites quantités de rayonnement micro-ondes, et la **détection passive des micro-ondes** peut être utile d'une manière quelque peu analogue à la télédétection thermique, pour mesurer les propriétés physiques de la surface de la Terre (par exemple, là où il y a de la neige et de la glace). La **télédétection active par micro-ondes**, également appelée télédétection radar, utilise les impulsions radar émises, tout comme les radars installés sur les navires et dans les aéroports. Les radars sont montés sur des avions ou des satellites et pointés vers la terre, des impulsions radar sont émises et la partie réfléchie de chaque impulsion est mesurée pour fournir des informations sur les aspects de la surface de la Terre, dont la plupart ne peuvent pas être mesurés avec des technologies optiques

ou thermiques. Il s'agit notamment d'informations sur le relief de la Terre (pour la création de modèles numériques d'élévation), l'humidité du sol, la densité et la structure de la végétation, la présence de glace de mer et de navires. La télédétection par radar présente l'avantage que le rayonnement micro-ondes traverse l'atmosphère plus ou moins librement et, comme il s'agit d'une technologie active, elle peut être utilisée partout et à tout moment, en plus de la nuit et par mauvais temps.

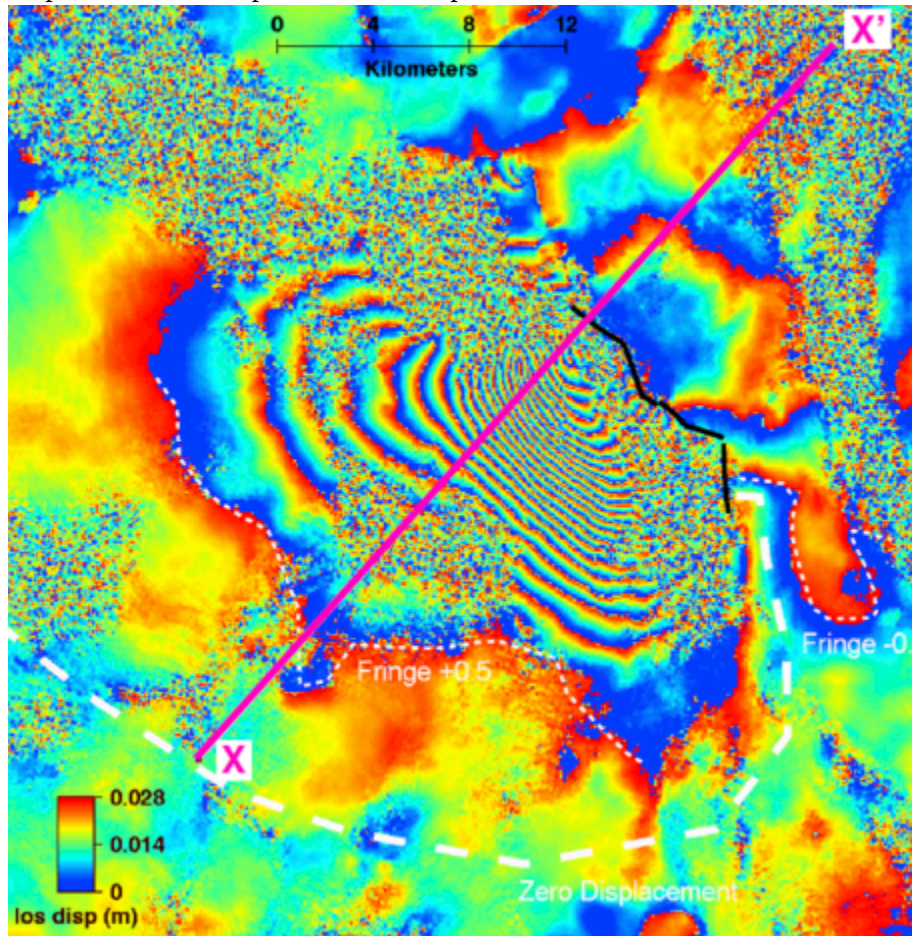


Figure 5: Une application étonnante de la télédétection radar est une technique appelée radar interférométrique à ouverture synthétique, ou InSAR. Lorsqu'un instrument radar passe devant la même zone sur deux orbites différentes, les changements de forme de la surface de la Terre peuvent être cartographiés avec une précision millimétrique. Par exemple, lorsque des tremblements de terre se produisent, le changement de la surface de la Terre peut être cartographié, comme le montre ici le tremblement de terre de 2010 à Dinar, en Turquie. Les changements plus petits qui résultent, par exemple, de l'enfoncement des bâtiments dans le sol ou de l'expansion et de la contraction des ponts par temps chaud ou froid, peuvent également être cartographiés avec InSAR.

Interferogram of the Dinar, Turkey 1995 earthquake par Gareth Funning, GEodesy Tools for Societal Issues (GETSI), CC BY-NC-SA 3.0.

Détection gravimétrique

Un autre type d'instrument, la détection gravimétrique, qui a été placé sur des satellites avec beaucoup de

succès a mesuré le champ de gravité de la Terre et la façon dont celui-ci peut se changer dans l'espace. L'attraction gravitationnelle de la Terre, tel qu'elle est perçue par un satellite en orbite, est plus puissante lorsque le satellite passe au-dessus d'une zone de masse relativement importante (par exemple l'Himalaya), par opposition à une zone de masse moindre (par exemple les océans profonds). Ces différences peuvent être mesurées avec une grande précision et utilisées non seulement pour cartographier la forme générale du champ de gravité de la Terre, mais aussi pour en mesurer les variations dans le temps. Ces changements sont causés par les marées (plus d'eau égale plus de masse égale plus de gravité), la fonte des glaciers, et même la baisse des nappes phréatiques due à une utilisation non durable des eaux souterraines.

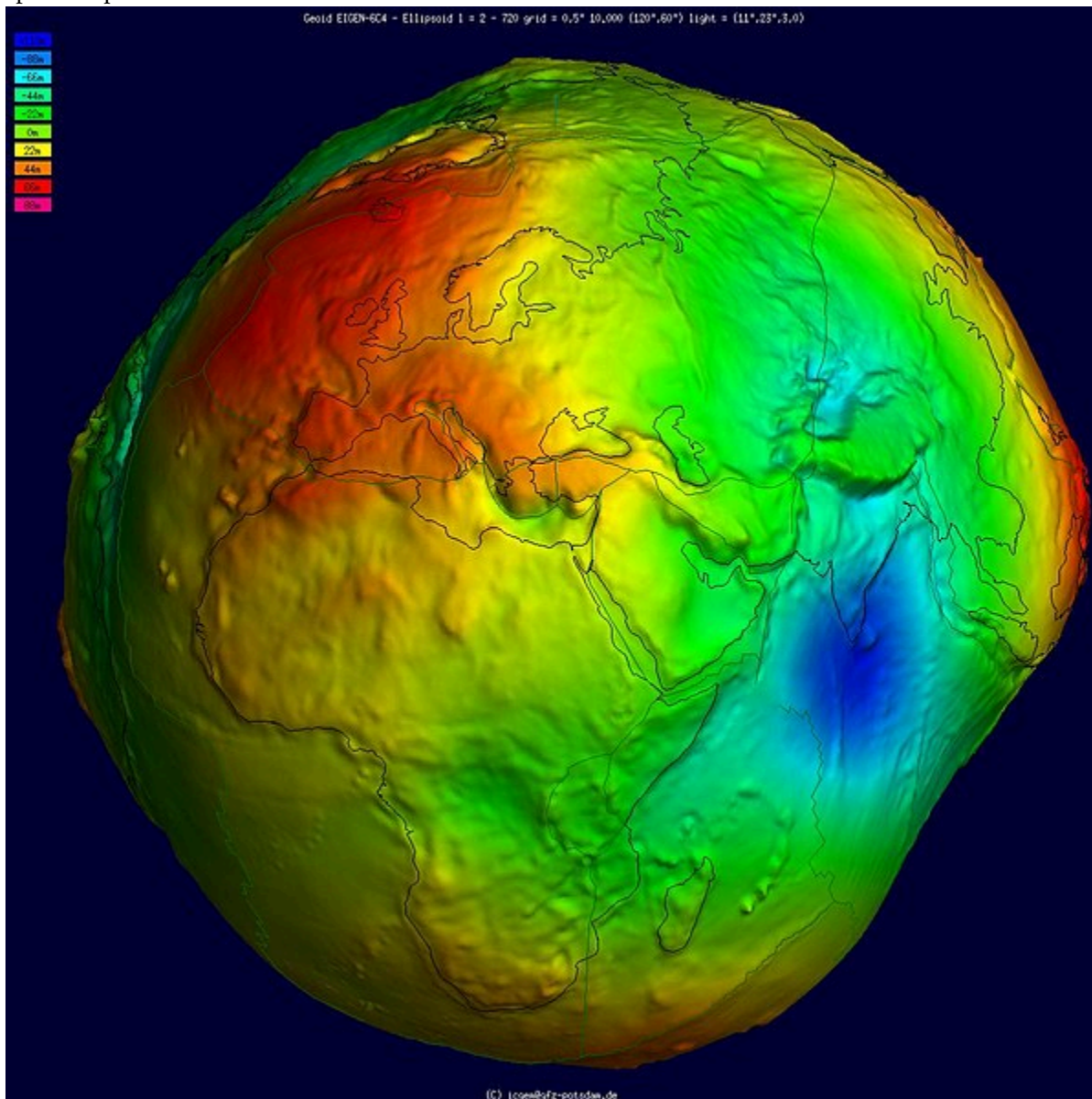


Figure 6: La forme et la couleur de la Terre selon l'attraction gravitationnelle mesurée par le satellite GOCE de l'Agence spatiale européenne. Geoid undulation 10k scale par le Centre international pour les modèles terrestres globaux, Wikimedia Commons, CC

Détection acoustique

La détection acoustique est fondamentalement différente des technologies énumérées précédemment, car elle utilise des ondes acoustiques (c'est-à-dire des ondes qui se forment pour donner suite à la compression et à l'expansion des masses d'eau) alors que toutes les autres utilisent le rayonnement électromagnétique (c'est-à-dire des ondes qui ont une composante magnétique et électrique et qui peuvent se propager dans le vide). Bien que la détection acoustique ait de nombreuses utilisations, principalement la cartographie de la profondeur de l'eau, du plancher océanique et même de la structure du sous-sol, les instruments et leur base physique sont suffisamment différents des technologies énumérées précédemment pour que nous ne les examinons pas davantage dans ce cours.

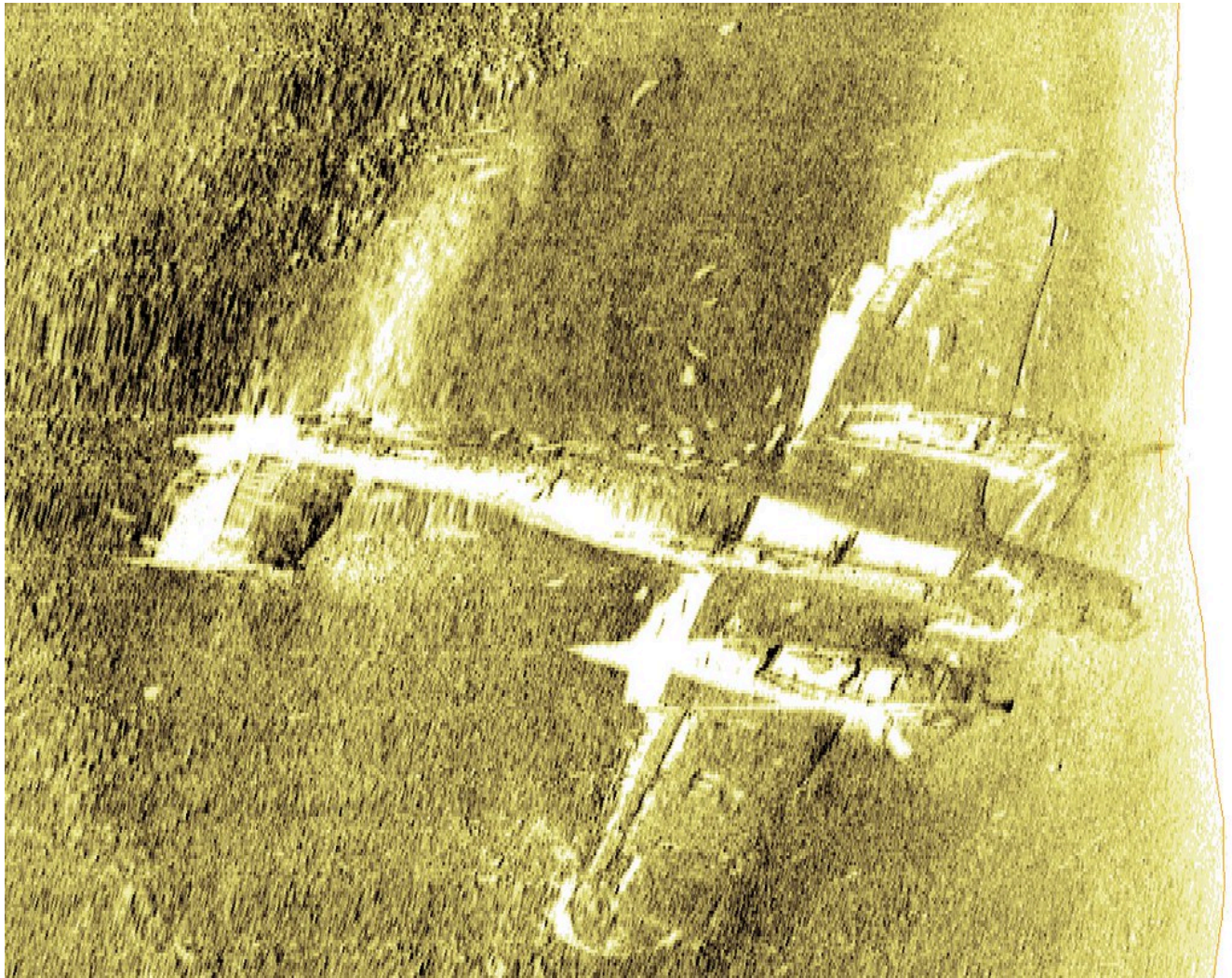


Figure 7: Les capteurs acoustiques peuvent à la fois fournir des informations sur la profondeur de l'eau et la composition du substrat sur de vastes zones, et fournir des informations détaillées sur la structure du plancher océanique dans des zones ciblées plus petites. On voit ici une reconstruction 3D, basée sur des données acoustiques, d'un avion de la Seconde Guerre mondiale situé au large de la côte britannique. Aircraft on the seabed par Wessex Archaeology, Flickr, CC BY-NC-SA 2.0.

Capteurs terrestres

Toutes les technologies de télédétection sont initialement développées et testées dans des laboratoires terrestres, où la science fondamentale est réalisée. Les prototypes sont produits et testés, ensuite les données initiales sont analysées pour enfin contribuer à l'avancement technologique. Bien que certaines soient finalement reconstruites pour être utilisées dans un avion ou sur un satellite, ces technologies maintiennent généralement leurs utilisations sur terre. Parmi les exemples les plus connus, citons les appareils photo des smartphones, les caméras thermiques utilisées par les inspecteurs électriques pour détecter les circuits défectueux, les radars utilisés par la police pour vérifier la vitesse à laquelle vous conduisez et les lidars utilisés par certaines voitures à conduite autonome pour la connaissance de la situation. Si les capteurs terrestres sont utiles pour recueillir des informations très précises sur une petite zone ou un petit objet, ils ne permettent pas d'obtenir les types de données nécessaires à la cartographie. Néanmoins, les données produites par les instruments terrestres sont souvent utiles dans le processus de télédétection (par exemple, en comparant ce que vous percevez sur une image satellite à ce que vous avez mesuré sur le terrain).

Capteurs aéroportés

Un jour, quelqu'un prendra une technologie terrestre et se demandera « ce qui se passerait-il si je prenais cet instrument, le plaçais dans un avion, le pointais vers le bas et le mettais en marche ? » L'avènement de la photographie aérienne a suivi rapidement le développement de l'appareil photo lorsque quelqu'un a décidé de faire monter un appareil photo dans une montgolfière, et les systèmes lidar et radar aéroportés ont également suivi le développement de leurs technologies respectives sans beaucoup de retard. Il y a au moins deux défis à relever pour emporter un instrument de télédétection dans les airs : 1) l'instrument doit pouvoir collecter rapidement de nombreuses données et les stocker pour une utilisation ultérieure (ou les transmettre directement à un dispositif de stockage situé ailleurs), et 2) chaque point de données collecté doit généralement être géoréférencé — en résumé, une paire de coordonnées géographiques (par exemple, la latitude et la longitude) doit être associée à chaque point de données. Ces défis sont loin d'être insurmontables aujourd'hui, alors qu'un smartphone ordinaire peut servir à la fois de stockage de données et de GPS. En plus, l'installation de capteurs sur des plates-formes aériennes est récemment devenue un domaine en pleine croissance, en raison des drones. Ils sont faciles à piloter pouvant désormais remplacer les avions pilotés en tant que plate-forme transportant l'instrument.

Capteurs spatiaux

Une fois qu'un instrument s'est avéré utile lorsqu'il est installé sur un avion, quelqu'un finit par suggérer que « nous devrions absolument prendre cet instrument et le placer sur un satellite, afin de pouvoir collecter des données en continu pendant des années sans devoir atterrir, faire le plein, soumettre des plans de vol, etc. ».

En fait (à ma connaissance) tous les satellites d'observation de la Terre transportent des instruments, dont les prototypes ont été testés sur des avions. Ces derniers permettent non seulement de peaufiner le matériel de l'instrument avant de le lancer dans l'espace, mais aussi de donner aux scientifiques de la télédétection un aperçu de ce à quoi ressembleront les données de l'instrument, afin qu'ils puissent commencer à écrire des programmes qui traitent les données de l'instrument en les convertissant en renseignements pratiques sur la Terre. L'utilisation de satellites présente des avantages et des inconvénients importants par rapport aux drones et aux avions pilotés. Le grand avantage est que les satellites durent généralement de nombreuses années et peuvent recueillir des données pendant cette période. Ils sont donc incroyablement rentables. Imaginez ce qu'il en coûterait de fournir des informations actualisées quotidiennement sur la couverture de glace de mer dans l'Arctique (et l'Antarctique, d'ailleurs) sans les satellites !

Cependant, les satellites ont un carburant limité, en plus les périls de l'espace entraînent des difficultés pour la durabilité des satellites. Les instruments se dégradent donc avec le temps pouvant finir par tomber en panne. Par exemple, en 2012, le satellite européen ENVISAT, qui a coûté plusieurs milliards de dollars, a cessé de communiquer avec son centre de commande et, malgré les efforts de l'Agence spatiale européenne pour rétablir la communication, le satellite et tous ses instruments ont été déclarés « morts » deux semaines plus tard. Landsat 6, un satellite américain polyvalent d'observation des sols, n'a pas réussi à atteindre son orbite après son lancement et n'a jamais produit de données. Un problème important avec les satellites est que lorsque ces difficultés technologiques se produisent, ceux-ci ne peuvent pas atterrir sur terre et être réparés.

Bien que la grande majorité des capteurs spatiaux soient situés sur des satellites, d'autres engins spatiaux ont également transporté quelques instruments de télédétection importants. Un système radar a été placé sur une navette spatiale en 2000 et a été utilisé pour produire un modèle numérique d'élévation quasi mondial avec une résolution spatiale de 30 mètres pendant ses 11 jours de fonctionnement. Récemment, une société canadienne appelée UrtheCast a exploité deux capteurs optiques passifs depuis la Station spatiale internationale — l'un était une caméra vidéo à ultra-haute définition qui produisait des vidéos à trois images par seconde avec une résolution spatiale de 1,1 mètre, l'autre était un instrument d'imagerie plus typique qui produisait des images en couleur avec une résolution spatiale de 5 mètres.

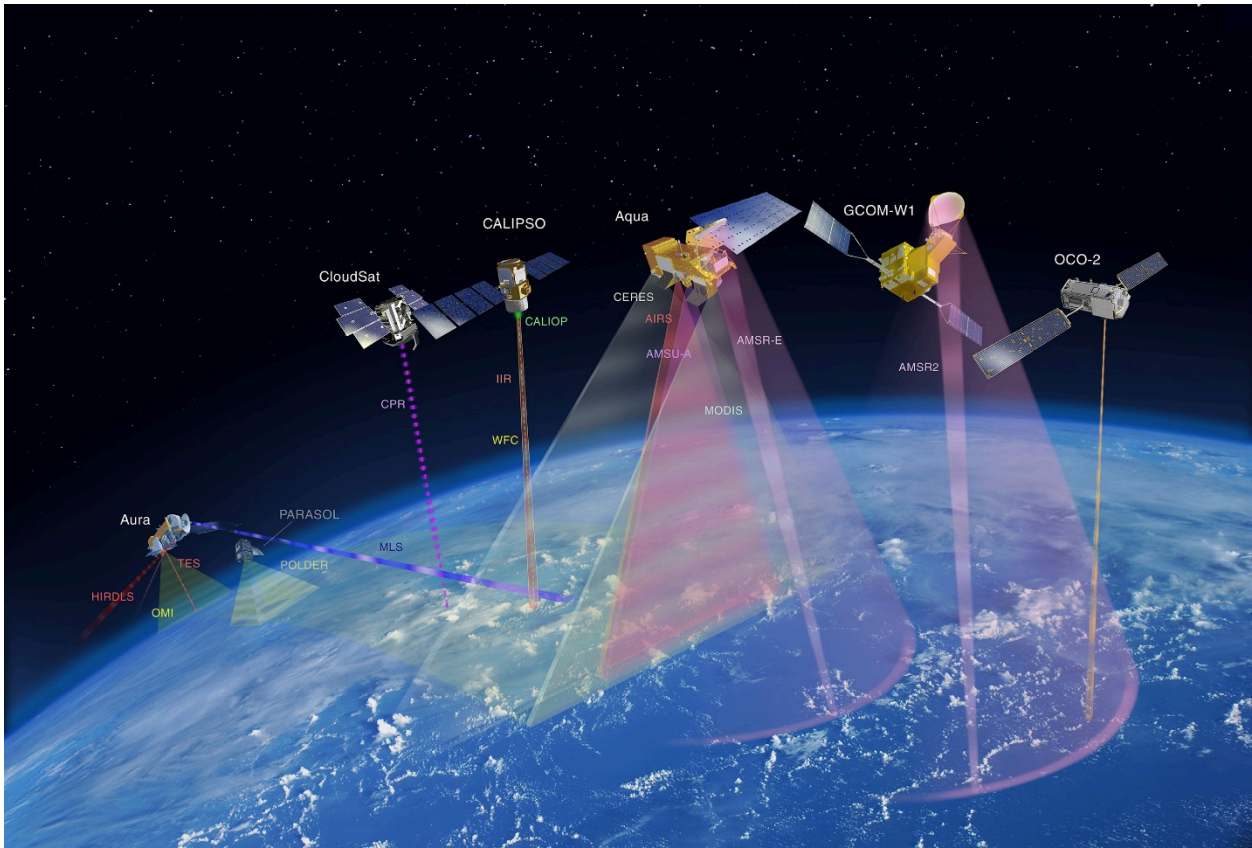


Figure 8: Il existe aujourd'hui un nombre remarquable de satellites d'observation de la Terre en orbite. Cette image montre la constellation de satellites « A-train » de la NASA, une série de satellites dont les orbites se suivent de près et qui passent tous au-dessus de nos têtes en début d'après-midi, heure solaire locale. Les constellations permettent d'utiliser les données d'un satellite pour faciliter l'interprétation des données d'un autre satellite, améliorant ainsi la qualité de nombreux produits satellitaires.

Atrain-879×485 par Shakibul Hasan Win, Wikimedia Commons, CC BY-SA 4.0.

Le rayonnement électromagnétique et ses propriétés

Il est maintenant temps d'examiner de plus près le fonctionnement de la télédétection et sa base physique. Cela vous permettra de mieux comprendre comment les données de télédétection sont créées et comment elles peuvent être utilisées de façon créative pour extraire exactement le type d'information nécessaire à un objectif spécifique.

L'énergie qui est mesurée par un instrument de télédétection (autre que les instruments acoustiques), et qui est utilisée pour produire une image, est appelée rayonnement électromagnétique (souvent abrégé en EMR, « ElectroMagnetic Radiation » en anglais). Il peut être difficile de comprendre exactement ce qu'est le rayonnement électromagnétique, mais une façon utile d'y penser est d'envisager que la lumière est un type spécifique de rayonnement électromagnétique. Un type de rayonnement que nos yeux, donc notre cerveau, est capables de détecter. Parmi les autres types de rayonnement électromagnétique, citons les rayons ultraviolets (UV) nocifs contre lesquels les écrans solaires offrent une certaine protection, le rayonnement thermique que

nous ressentons lorsque nous nous approchons d'un feu de camp et les ondes radar utilisées pour détecter les avions et les navires (et pour cartographier certaines propriétés de la surface de la Terre). Physiquement, le rayonnement électromagnétique peut être visualisé comme des ondes qui se propagent dans l'espace. Les ondes ont deux composantes, l'une électrique et l'autre magnétique, qui forment des angles de 90 degrés entre elles et par rapport à la direction de propagation (figure 9).

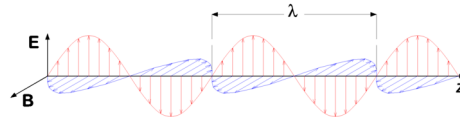


Figure 9: Le rayonnement électromagnétique consiste en une onde transversale avec un champ électrique (rouge, plan vertical) et un champ magnétique (bleu, plan horizontal), se propageant (dans cette figure) dans la direction horizontale. Electromagnetic wave par P.wormer, Wikimedia Commons, CC BY-SA 3.0.

Longueur d'onde

Les ondes électromagnétiques ont certaines propriétés que nous pouvons mesurer et utiliser pour les expliquer. Les ondes électromagnétiques peuvent être caractérisées par leur longueur d'onde, qui sont mesurées comme la distance physique entre un pic d'onde et le suivant, le long de la direction de propagation. L'œil humain est capable de détecter les rayonnements électromagnétiques dont la longueur d'onde se situe approximativement entre 400 et 700 nanomètres (un nanomètre correspond à 10^{-9} mètre, soit un milliardième de mètre), ce que nous appelons donc la lumière visible. Les ondes électromagnétiques se propagent à la vitesse de la lumière, et il existe une relation simple et directe entre la longueur d'onde et la fréquence des ondes électromagnétiques, généralement exprimée par la formule $c = \nu\lambda$, où c est la vitesse de la lumière, ν (la lettre grecque nu) est la fréquence, et λ (la lettre grecque lambda) est la longueur d'onde. La fréquence est définie comme le nombre de pics d'ondes passant par un point fixe par unité de temps (typiquement par seconde). Notez que certaines personnes, notamment les ingénieurs et les physiciens travaillant sur les rayonnements, utilisent parfois le nombre d'onde au lieu de la longueur d'onde ou de la fréquence. Le nombre d'onde ($\tilde{\nu}$) est défini comme $\tilde{\nu} = 1/\lambda$. Nous n'utiliserons pas le nombre d'onde dans la suite de ces notes, mais il est bon de connaître son utilisation pour des études ultérieures.

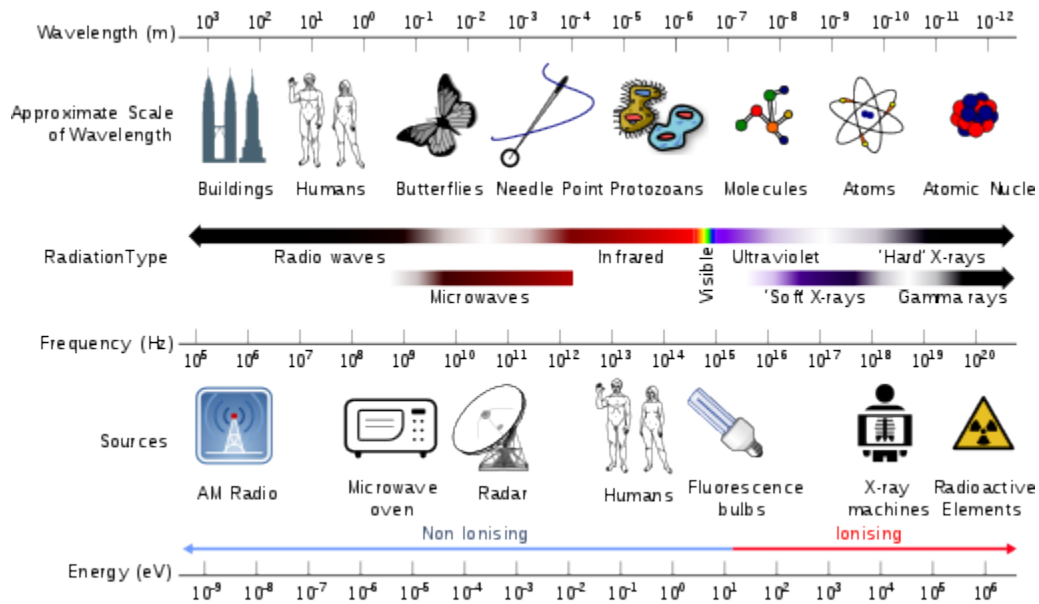


Figure 10: Un aperçu des différents types de EMR, de leurs longueurs d'onde, de leurs fréquences et de certaines sources communes. Electromagnetic spectrum with sources par Dinksbumf, Inductiveload and NASA, Wikimedia Commons, CC BY-SA

3.0.

Selon la théorie quantique, on peut également considérer que l'EMR est constitué d'un flux de paquets d'énergie individuels appelés photons. Chaque photon contient une quantité d'énergie proportionnelle à sa fréquence, une relation exprimée par $E=h\nu$, où h est la constante de Planck (environ $6,6 \cdot 10^{-34}$ Js). La relation entre la longueur d'onde, la fréquence et l'énergie par photon est illustrée à la figure 10. On trouve également dans la figure 10 les noms communs d'EMR d'une certaine longueur d'onde. Vous pouvez voir, par exemple, que les rayons X ont des longueurs d'onde très courtes, des fréquences très élevées et donc un contenu énergétique par photon très élevé, ce qui est une raison importante pour limiter votre exposition à ces rayons.

Polarisation

Outre la longueur d'onde, l'EMR peut être caractérisé par sa polarisation, qui peut être considérée comme l'orientation des oscillations du champ électrique. Par exemple, sur la figure 9, le champ électrique oscille exclusivement dans le plan vertical, l'EMR représenté sur la figure est donc dit polarisé verticalement. De nombreuses sources d'EMR, comme le Soleil, produisent des ondes électromagnétiques dont la polarisation se change si rapidement, lorsqu'on la mesure, la combinaison de plusieurs ondes semble ne pas avoir de polarisation distincte. En d'autres termes, l'orientation des oscillations du champ électrique (et par extension du champ magnétique) est verticale, horizontale, et tout ce qui se situe entre les deux.

Un tel rayonnement électromagnétique est dit non polarisé. D'autres sources de lumière, comme celle créée par la plupart des instruments est polarisée, signifiant ainsi l'orientation des oscillations du champ électrique ne changent pas avec le temps. Un exemple connu de polarisation est la lumière du soleil se reflétant sur une surface d'eau. Lorsque la lumière non polarisée du Soleil frappe la surface de l'eau. Les ondes, dont le champ

électrique est orienté verticalement sont absorbées ou réfractées dans l'eau plus facilement que celles dont le champ électrique est orienté horizontalement, ont plus de chances d'être réfléchies par la surface de l'eau. Lorsque vous regardez une surface d'eau, les parties de la lumière polarisée horizontalement que vous voyez ont donc été principalement réfléchies par la surface de l'eau, tandis que les parties de la lumière polarisée verticalement ont été principalement réfléchies par l'intérieur de l'eau elle-même, ou par le fond si vous regardez une eau peu profonde. Par conséquent, si vous utilisez des lunettes de soleil polarisées qui ont été conçues pour éliminer efficacement la lumière polarisée horizontalement, la lumière qui atteint vos yeux est principalement la partie du champ lumineux qui provient de l'eau elle-même. Un exemple de ce phénomène est illustré à la figure 11, où l'on peut voir que l'utilisation d'un filtre polarisant permet de rendre visibles les détails du fond marin.



Figure 11: Un exemple d'observation d'une eau peu profonde, sans (à gauche) ou avec (à droite) un filtre de polarisation.

Reflection Polarizer2, par Amithshs, Wikimedia Commons, public domain.

Comment le rayonnement électromagnétique est créé

Maintenant que nous savons comment quantifier les différentes propriétés du rayonnement électromagnétique, nous allons examiner brièvement comment il est créé en premier lieu. Les physiciens disent que le rayonnement électromagnétique est produit lorsque des particules chargées sont accélérées. Cela se produit tout le temps, mais ce n'est pas utile dans le cadre de la télédétection. Il est plus utile d'examiner la luminance spectrale créée par une surface, c'est-à-dire la quantité de rayonnement créée et sa distribution sur différentes longueurs d'onde. Cela dépend principalement de la température de la surface, et est décrit de façon claire pour les corps noirs par la loi de Planck :

$$L_{\lambda,T} = \frac{2hc^2}{\lambda^5} \frac{1}{e^{\left(\frac{hc}{\lambda k_B T}\right)} - 1}$$

Dans la loi de Planck, $L_{\lambda,T}$ est la luminance, les indices indiquant qu'elle dépend de la longueur d'onde et de la température (de la surface). h est la constante de Planck, c est la vitesse de la lumière dans le vide, λ est la longueur d'onde, k_B est la constante de Boltzmann (à ne pas confondre avec la constante de Stefan-Boltzmann !), T est la température de la surface en degrés Kelvin. Si la loi de Planck peut sembler intimidante au premier abord, gardez à l'esprit que h , c et k_B sont toutes des constantes, et que tout ce qu'elle dit est que la luminance émise par une surface ($L_{\lambda,T}$) peut être calculée à partir de la température (T) de cette surface pour une longueur d'onde donnée (λ).

Pour examiner le résultat de la loi de Planck, vous pouvez imaginer un objet avec une certaine température de surface, comme le Soleil dont la température est d'environ 5800K, et calculer la luminance qu'il émet à, disons, 400 nm, en branchant simplement la température et la longueur d'onde dans l'équation. Vous pouvez ensuite répéter l'opération pour 401 nm, 402 nm, et ainsi de suite, pour obtenir ce que l'on appelle une courbe de Planck. La figure 12 présente un exemple de plusieurs courbes de Planck. Notez que très peu de rayonnement est produit à des longueurs d'onde très courtes (par exemple, 0-200 nm), et que chaque courbe présente un pic distinct qui indique la longueur d'onde auquel la quantité maximale d'énergie est émise. L'emplacement de ce pic peut être calculé à l'aide de la loi de Wien, qui est basée sur la loi de Planck et qui stipule que la longueur d'onde du pic d'émission (λ_{\max}) peut être calculée comme suit : $\lambda_{\max}=b/T$, où b est la constante de déplacement de Wien : 2 897 767 nm K. Pour le Soleil, cela correspond à environ 500 nm, soit une lumière bleu-vert.

Une remarque importante sur la loi de Planck est qu'elle s'applique aux corps noirs, qui sont des entités imaginaires définies comme des absorbeurs et des émetteurs parfaits. Bien que les objets du monde réel, comme le Soleil et la Terre, ne soient pas des corps noirs, ils s'en approchent suffisamment la plupart du temps pour que la loi de Planck soit une approximation utile qui peut être utilisée pour quantifier leur émission de rayonnement électromagnétique. La proximité ou l'éloignement d'une surface par rapport à un corps noir est défini par son émissivité, qui est la quantité de luminance produite à une longueur d'onde donnée par rapport à celle qui serait produite par un corps noir. Les matériaux naturels ont tendance à avoir des valeurs d'émissivité élevées, de l'ordre de 0,98-0,99 (eau), 0,97 (glace) et 0,95-0,98 (végétation). Cependant, certains matériaux naturels peuvent avoir une faible émissivité, comme la neige sèche (environ 0,8) ou le sable sec (0,7-0,8). Cela signifie, par exemple, que la neige sèche, à une température donnée, ne produit que 80 % de la radiance qu'elle devrait produire selon la loi de Planck. Une autre complication est due au fait que l'émissivité change avec la longueur d'onde, mais c'est un pont à traverser une autre fois.

La loi de Planck peut être utilisée pour déterminer la température d'une surface dont on mesure la luminance spectrale. Pour ce faire, il suffit d'inverser l'équation pour isoler la température :

$$T = \frac{hc}{\ln\left(\frac{2hc^2}{\lambda^5 L} + 1\right) \lambda k_B}$$

Si la détermination effective de la température de la surface de la Terre à partir de la mesure spatiale de la luminance (spectrale) est légèrement plus compliquée, car il faut tenir compte de l'émissivité et de l'influence de l'atmosphère, l'inversion de la loi de Planck, telle que décrite ci-dessus, est le principe de base sur lequel repose cette cartographie.

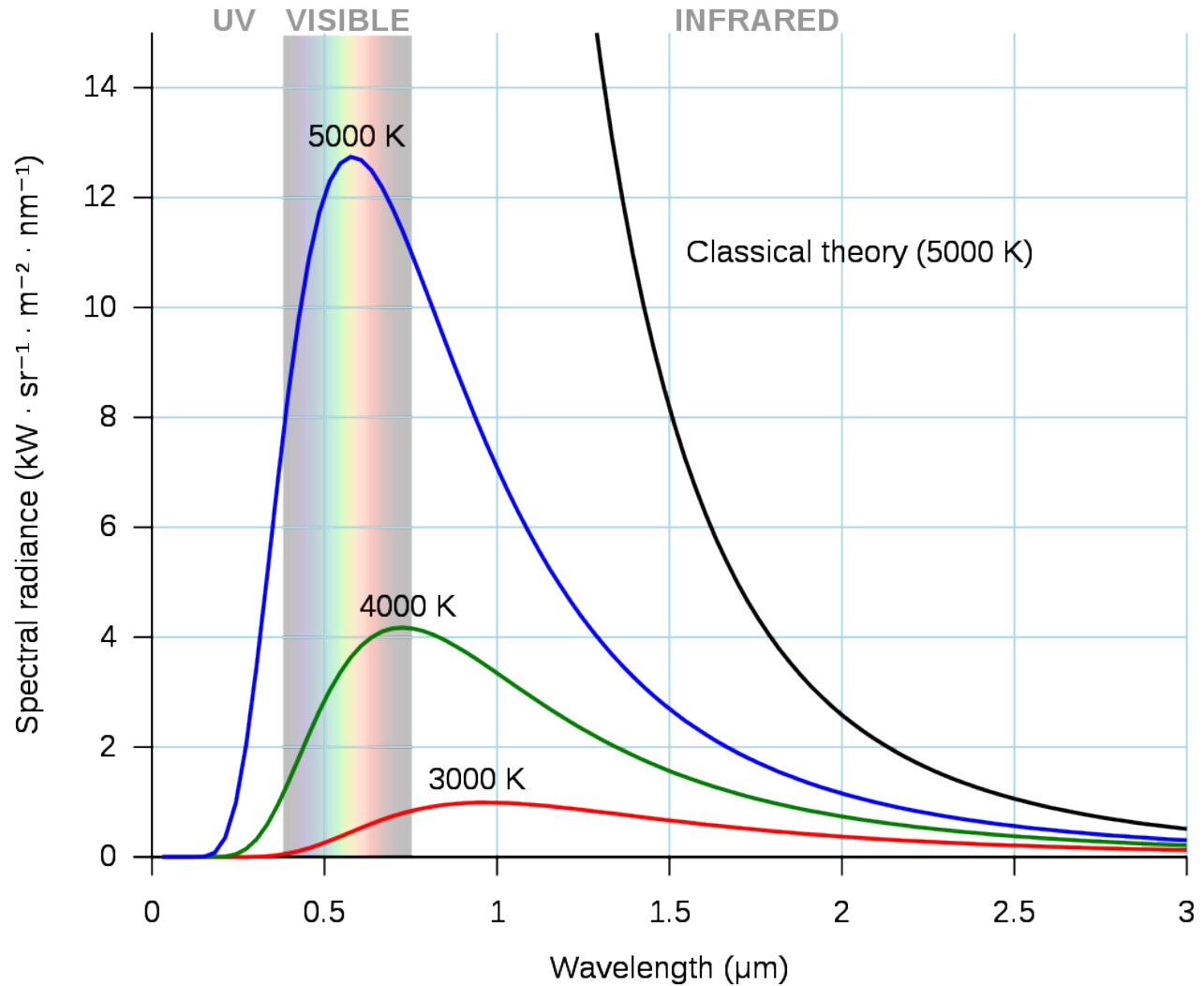


Figure 12: Exemple de trois courbes de Planck, pour des surfaces à 3000-5000 degrés Kelvin. Black body par Darth Kule, Wikimedia Commons, public domain.

2.

MESURES RADIOMÉTRIQUES

Alors que la plupart des données de télédétection peuvent être imaginées sous forme d'images, en plus d'être interprété en l'état actuel des choses pour enfin localiser, des villes, des navires ou des incendies, il faut comprendre que la plupart des capteurs satellitaires peuvent également être considérés comme des instruments précisément calibrés. Ils sont utilisés pour mesurer les propriétés physiques du rayonnement électromagnétique arrivant d'un sens donné au moment de la prise de mesure. Dans ce chapitre, nous nous concentrerons sur cette utilisation de l'imagerie satellitaire. Nous devons d'abord définir certains concepts liés au rayonnement électromagnétique et à la façon dont il est mesuré. Ensuite, nous verrons comment utiliser ces concepts pour extraire des informations utiles de l'imagerie satellitaire.

Concepts liés à l'intensité de l'EMR

Plusieurs autres propriétés des champs électromagnétiques sont liées à leur intensité — la quantité d'énergie qu'ils contiennent. L'unité SI fondamentale de l'énergie est le Joule (J). Comme le rayonnement se déplace lorsque les ondes électromagnétiques se propagent, une mesure importante de l'intensité d'un champ électromagnétique est la quantité d'énergie qui se déplace sur le site, à travers ou à partir d'une surface ou d'un volume *par unité de temps*. C'est ce qu'on appelle le *flux radiatif*, et il est mesuré en joules par seconde, ou en watts (W). En télédétection, nous sommes presque toujours intéressés par la mesure de l'intensité du rayonnement sur une surface finie (par exemple, celle représentée par un pixel dans une image satellite). C'est ce qu'on appelle *la densité du flux radiatif*, et elle est mesurée en tant que flux rayonnant par unité de surface, par exemple W/m^2 . Dans la pratique, le terme *irradiance* est utilisé pour décrire la densité de flux radiative incidente sur une surface, et le terme *exitance* est utilisé pour décrire la densité du flux radiatif quittant une surface (figure 13).

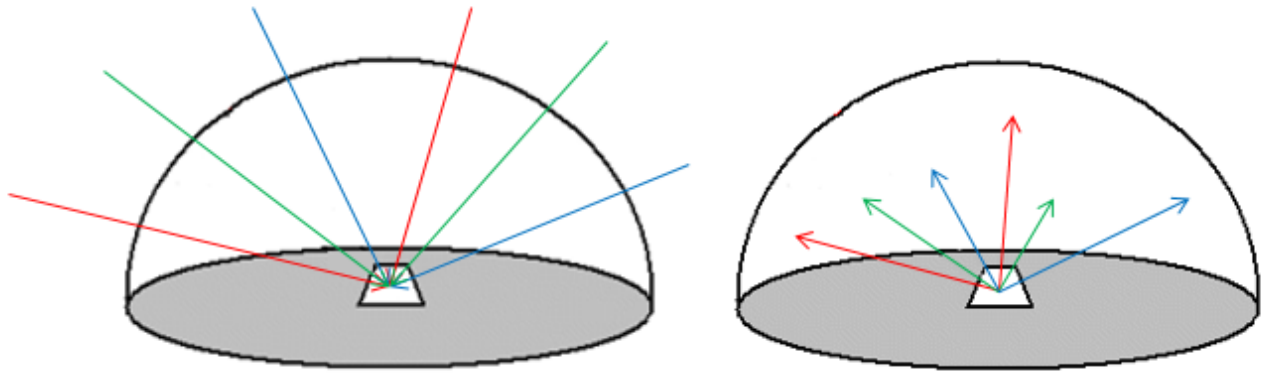


Figure 13: Irradiance et Exitance. À gauche, l'irradiance est le flux radiatif incident sur une surface par unité de surface. À droite, l'exitance est le flux radiatif quittant une surface par unité de surface. Tous deux sont généralement mesurés en W/m^2 . Les flèches de couleur différente indiquent que ces deux densités de flux radiatif sont mesurées pour toutes les longueurs d'onde présentes dans le champ de rayonnement. Par Anders Knudby, CC BY 4.0.

Bien que l'irradiance et l'exitance soient des concepts importants, nous ne pouvons pas les mesurer avec des instruments placés sur des plates-formes volantes, car ils englobent le rayonnement incident ou sortant d'une surface dans toutes les directions. En outre, il faut savoir que pour mesurer l'exitance d'une partie de la surface terrestre de 1 min 2 s, nous devrions placer un capteur hémisphérique au-dessus de la zone en question, mesurant ainsi le rayonnement sortant dans toutes les directions vers le haut (si l'on mesure la lumière solaire réfléchiée par la surface, l'ombrage de l'instrument rendrait évidemment la mesure inutile). Une autre unité d'intensité de rayonnement électromagnétique couramment utilisée en télédétection est donc plus proche de ce qui est réellement mesuré par les instruments. Elle est appelée luminance et est définie comme la densité de flux de rayonnement par unité de surface de source projetée, dans une direction particulière définie par un angle solide (Figure 14). Un angle solide peut être considéré comme un cône — dans la plupart des cas de télédétection, ce cône est incroyablement étroit puisqu'il s'étend de la zone observée à l'élément de détection pertinent du capteur.

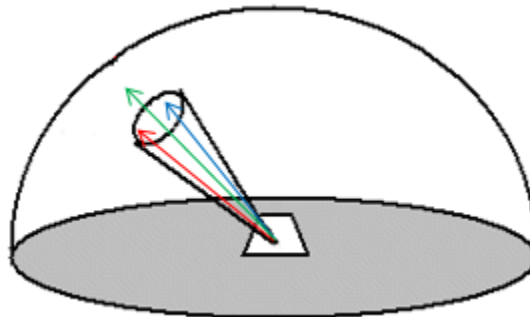


Figure 14: Luminance. Par rapport à l'irradiance et à l'exitance, la luminance est calculée par unité de surface projetée, et dans un angle solide spécifié. La luminance est généralement mesurée en unités de $\text{W m}^{-2} \text{sr}^{-1}$, où sr signifie stéradian, qui est une mesure de l'angle solide. Par Anders Knudby, CC BY 4.0.

Lorsque l'on mesure le rayonnement électromagnétique à l'aide d'instruments installés à bord d'avions ou de satellites, les mesures de la luminance sont généralement effectuées pour des intervalles discrets de longueurs d'onde. Par exemple la « bande 1 » de nombreux satellites a été conçue pour mesurer uniquement la luminance du rayonnement électromagnétique avec certaines longueurs d'onde qui semblent bleues à l'œil humain (entre 420 et 470 nm). En supposant que cette gamme de longueurs d'onde est parfaitement mesurée par la « bande 1 », nous pouvons trouver la luminance spectrale en divisant la luminance mesurée par la gamme de longueurs d'onde mesurée (dans ce cas 50 nm). La *luminance spectrale* est une mesure très couramment utilisée en télédétection, et est souvent directement liée aux valeurs brutes (c'est-à-dire aux nombres numériques) trouvées dans chaque bande des images de télédétection.

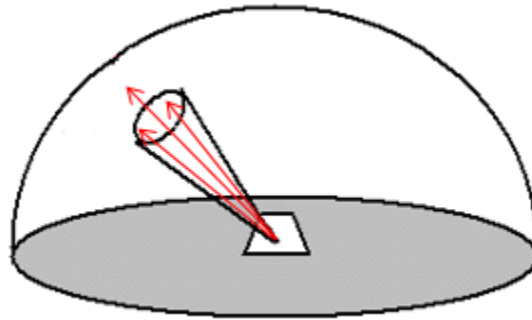


Figure 15: Luminance spectrale. Par rapport à la figure 13 et à la figure 14, les flèches indiquant le rayonnement électromagnétique ont ici été colorées d'une seule couleur, pour souligner que la luminance spectrale est mesurée pour une gamme spécifique de longueurs d'onde. Par Anders Knudby, CC BY 4.0.

Bien que la luminance spectrale soit l'unité d'intensité des CEM la plus proche de ce qui est réellement mesuré par les instruments de télédétection, elle peut être embêtante en raison de ces difficultés de dépendre de l'éclairage de la zone observée. C'est évident si vous imaginez mesurer la quantité de lumière provenant du stationnement de votre quartier. Au cours d'une journée ensoleillée, une grande quantité de lumière se dégage de la surface de l'asphalte, parfois à tel point que vous devez plisser les yeux pour la regarder. Outre, la surface a une luminance spectrale élevée dans les longueurs d'onde visibles. Par temps, la surface reçoit moins de lumière — la luminance spectrale est plus faible. Et pendant la nuit, évidemment, très peu de lumière se dégage de la surface — luminance spectrale très faible. La mesure de la luminance spectrale provenant d'une surface ne peut donc pas nous renseigner directement sur la nature de cette surface, qu'il s'agisse d'un stationnement, d'un lac ou d'une forêt. Ce que nous préférons mesurer, c'est la réflectance d'une surface, c'est-à-dire la quantité de rayonnement qu'elle réfléchit par unité de rayonnement qui la frappe. Il s'agit d'une propriété physique réelle d'un matériau qui, dans la plupart des cas, est totalement indépendante de l'éclairage, et qui peut donc être utilisée pour identifier le matériau. Une réflectance élevée est ce qui rend les surfaces blanches, et une réflectance faible est ce qui rend les surfaces noires. Un type de réflectance couramment utilisé est l'albédo, qui est ce que les spécialistes de la télédétection appellent la réflectance diffuse. Il s'agit d'une mesure sans unité comprise entre 0 (pour les surfaces complètement noires qui n'ont aucun coefficient de sortie) et 1 (pour les surfaces complètement blanches qui réfléchissent tout le rayonnement qui leur est incident).

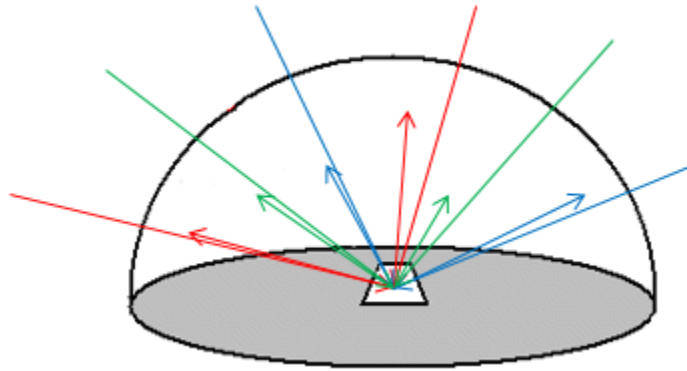


Figure 16: Réflexion diffuse. Définie comme l'exitance (flèches sortantes) divisée par l'irradiance (flèches entrantes). Par Anders Knudby, CC BY 4.0.

La réflectance diffuse ne peut pas être mesurée directement en télédétection — en fait, aucun des deux termes nécessaires ne peut être mesuré. Cependant, elle peut être estimée de manière assez précise au sommet de l'atmosphère (nous y reviendrons plus tard) en utilisant deux astuces. Premièrement, nous connaissons assez bien la quantité de rayonnement produite par le Soleil à différentes longueurs d'onde, et nous pouvons donc estimer l'irradiance avec une bonne précision. Deuxièmement, si nous supposons que la luminance est une fonction connue de la direction de propagation, nous pouvons convertir la luminance mesurée dans une direction en exitance. L'ensemble de ces éléments nous donne les informations nécessaires pour convertir une mesure de radiance en une estimation de la réflectance diffuse.

En effet, la réflectance diffuse est une valeur unique qui nous renseigne sur la luminosité d'une surface, mais elle ne nous dit rien sur sa chromacité (qui est l'aspect de la couleur qui n'inclut pas la luminosité, par exemple, le rouge clair et le rouge foncé peuvent avoir la même chromacité, alors que le rouge clair et le vert clair ont des chromacités différentes). Cependant, comme indiqué ci-dessus dans le paragraphe sur la luminance spectrale, les mesures effectuées avec les instruments de télédétection sont en fait des mesures de la luminance spectrale. Nous pouvons donc convertir ces mesures en estimations de l'exitance spectrale, et les diviser par l'irradiance spectrale, pour obtenir une mesure de la réflectance diffuse spectrale. Si l'on ne tient pas compte des interactions entre le rayonnement électromagnétique et l'atmosphère, c'est le mieux que nous puissions faire pour obtenir cette mesure qui nous renseigne sur ce qui recouvre la surface de la Terre dans le pixel mesuré.

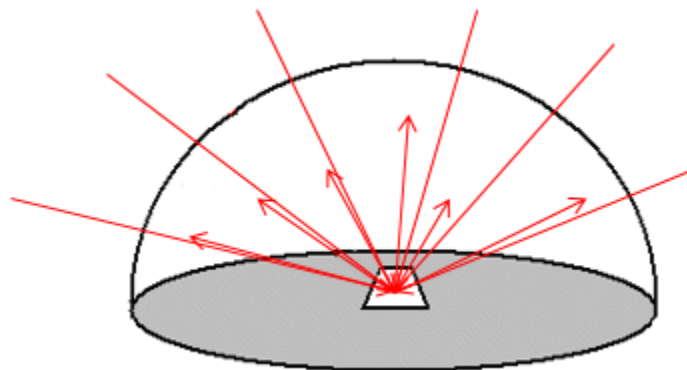


Figure 17: Réflexion diffuse spectrale. Définie comme la sortie spectrale divisée par l'irradiance spectrale. Par Anders Knudby, CC BY 4.0.

Comment les rayonnements électromagnétiques sont détectés et mesurés

Pour mesurer l'intensité du rayonnement électromagnétique entrant, les instruments de télédétection utilisent des matériaux spéciaux qui absorbent efficacement le rayonnement dans la gamme de longueurs d'onde souhaitée et qui génèrent un courant électrique lorsqu'ils sont exposés au rayonnement — un courant qui est ensuite mesuré. Bien que différents matériaux soient utilisés pour absorber et donc détecter le rayonnement à différentes longueurs d'onde dans différents types de capteurs, le principe de base est bien illustré par les instruments de télédétection optique passive, dont la plupart reposent sur des capteurs CCD ou CMOS, tout comme l'appareil photo de votre téléphone.

Imaginez une caméra fixée à un satellite dans l'espace, pointant vers la Terre. Pendant une très courte période, la lentille de l'appareil photo est ouverte, laissant entrer la lumière produite par le Soleil (selon la loi de Planck) et réfléchi par l'atmosphère et la surface de la Terre dans l'espace, exactement dans la direction du satellite. À l'intérieur de l'appareil photo se trouve un CCD bidimensionnel, une sorte de damier composé de 3 x 3 petits éléments de détection individuels, chacun d'entre eux étant capable d'absorber la lumière solaire et de produire ainsi un courant électrique proportionnel à l'intensité de la lumière entrante. Les éléments de détection exposés à une lumière solaire plus intense produisent un courant de plus grande tension, de sorte que la tension est une indication directe de la quantité de lumière solaire à laquelle chaque élément a été exposé. Un exemple est fourni dans le tableau 1.

Tableau 1: Exemple de tensions générées par des éléments de détection individuels dans un CCD logé dans une caméra spatiale.

Par Anders Knudby, CC BY 4.0.

1.2	1.1	0.9
1.3	1.2	1.1
1.1	0.9	0.8

L'optique (système de lentilles) de la caméra garantit que chaque élément de détection individuel est exposé à la lumière provenant d'une direction unique prédéfinie par rapport à la caméra. Sachant où se trouve le satellite et comment il est orienté par rapport à la Terre, chacune de ces orientations peut être convertie en un ensemble de coordonnées géographiques (latitude/longitude) sur la surface de la Terre, à partir laquelle la lumière enregistrée par chaque élément de détection doit avoir été réfléchi. En d'autres termes, le courant

électrique généré par chacun des éléments de détection 3 x 3 peut être ramené à un emplacement sur la surface de la Terre.

Tableau 2: Coordonnées géographiques à partir desquelles la lumière du soleil produisant le courant électrique dans chaque élément du tableau 1 a été réfléchi. Pour cet exemple, ignorez le fait que la taille du pixel implicite dans les coordonnées géographiques n'est pas carrée et très grande. Par Anders Knudby, CC BY 4.0.

45.1 N / 80.5 W	45.1 N / 80.4 W	45.1 N / 80.3 W
45.0 N / 80.5 W	45.0 N / 80.4 W	45.0 N / 80.3 W
44.9 N / 80.5 W	44.9 N / 80.4 W	44.9 N / 80.3 W

Étant donné que la tension du courant électrique généré par chaque élément de détection est proportionnelle à la densité du flux radiant du rayonnement entrant, avec un étalonnage approprié, la tension produite par chaque élément de détection peut être convertie en une mesure de la densité du flux radiant entrant au niveau de la caméra. En outre, si l'on connaît la distance entre le satellite et la surface de la Terre, on peut à son tour la convertir en une mesure de la luminance provenant de la zone observée. Nous obtenons ainsi l'une des mesures fondamentales de la télédétection : la luminance provenant d'une zone bien définie de la surface de la Terre.

Étalonnage

Dans la section précédente, il a été mentionné que la tension générée par un élément de détection peut être convertie en densité de flux rayonnant ou en luminance avec un étalonnage approprié. Un tel étalonnage est effectué pour tous les capteurs avant le lancement en les exposant à une lumière de différents niveaux de luminance connue, en notant la tension créée par chaque exposition. À partir de là, une équation de conversion simple, généralement linéaire ou presque linéaire, peut être créée pour convertir la tension en luminance. La figure 18 présente un exemple du type de configuration nécessaire pour effectuer un tel étalonnage.



Figure 18: Sphère d'intégration utilisée pour l'étalonnage du capteur. Commercial Integrating Sphere par Electro Optical Industries, Wikimedia Commons, CC BY-SA 3.0.

Détection multispectrale

Jusqu'à présent, nous n'avons pas tenu compte du fait que la plupart des capteurs de satellites sont conçus pour détecter la luminance dans plusieurs gammes distinctes de longueurs d'onde, chacune produisant ce que l'on appelle une « bande ». Le champ lumineux entrant qui atteint le capteur est toujours, dans le cas de la lumière solaire réfléchi, un mélange de radiations de différentes longueurs d'onde, de sorte que pour les séparer et mesurer la luminance de gammes distinctes de longueurs d'onde individuellement, il faut une étape supplémentaire au-delà de ce qui a été décrit ci-dessus. Différents types de technologies sont utilisés pour réaliser cette étape, la plus simple d'entre elles reposant sur un séparateur de faisceau. Un diviseur de faisceau peut être considéré comme une sorte de prisme fantaisie qui redirige le rayonnement électromagnétique dans une gamme définie de longueurs d'onde dans une direction, tout en laissant le rayonnement avec d'autres longueurs d'onde se déplacer sans entrave. Plusieurs capteurs individuels peuvent alors être placés aux endroits appropriés du capteur d'image, le CCD destiné à produire des données pour la bande « rouge » étant situé à l'endroit où le séparateur de faisceau redirige le rayonnement de longueurs d'onde comprises entre 600 et 700 nm (typiquement considéré comme de la lumière « rouge »), et ainsi de suite pour chaque CCD individuel. D'autres technologies que les séparateurs de faisceau sont utilisées, chacune accomplissant la même tâche de base qui consiste à permettre l'enregistrement séparé de rayonnements de différentes longueurs d'onde.

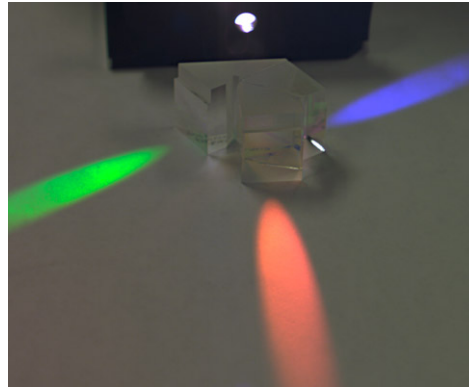


Figure 19: L'idée d'un séparateur de faisceau. Le rayonnement entre par la source, et les photons sont redirigés en fonction de leur longueur d'onde. Trois CCD peuvent alors être positionnés pour mesurer l'intensité de la lumière rouge, verte et bleue séparément.

Color Separation Prism par Dick Lyon, Wikimedia Commons, public domain.

Note: Pour des raisons de simplicité, nous avons utilisé une « caméra » traditionnelle comme modèle pour expliquer comment les capteurs optiques passifs détectent et mesurent le rayonnement électromagnétique entrant. Alors que les caméras traditionnelles restent l'instrument de choix dans l'industrie de la photographie aérienne et dans le domaine en pleine expansion de la télédétection par drones, les capteurs optiques passifs embarqués à bord des satellites se présentent sous deux autres formes : les scanners à balai ou à fouet. Les détails de ces systèmes de détection dépassent le cadre de ces notes, mais des informations de base sont fournies ici :

- Un scanner à ligne est constitué d'une série de CCD unidimensionnels, chacun d'eux étant généralement composé de plusieurs milliers de détecteurs individuels en une rangée. Chaque CCD est positionné avec soin dans l'instrument, « après » un séparateur de faisceau, pour enregistrer le rayonnement dans ce qui devient une bande, plus ou moins comme le montre la figure 19. L'instrument n'a pas d'obturateur, les CCD enregistrent plutôt le rayonnement entrant d'une ligne très large et étroite sur la Terre, allant d'un côté de la bande suborbitale à l'autre. Ces mesures deviennent une ligne (rangée) dans une image satellite. Au fur et à mesure que le satellite se déplace sur son orbite, une autre ligne est enregistrée, puis une autre, et une autre, jusqu'à ce que des milliers de lignes aient été mesurées. Elles sont ensuite assemblées pour former une image.
- Un scanner à balayage repose sur un miroir de balayage rotatif pour diriger le rayonnement de différentes parties de la surface de la Terre sur un plus petit nombre de CCD. La position du miroir à un moment donné peut être utilisée pour calculer la direction depuis laquelle le rayonnement mesuré est arrivé sur le capteur. Ces mesures, ainsi que les informations de géolocalisation correspondantes, sont prises en succession rapide, et finalement assemblées sous forme d'image.

3.

CALCULS DE LA LUMINANCE TOA ET DE LA RÉFLECTANCE TOA

L'étalonnage des capteurs se traduit généralement par une simple conversion linéaire (ou presque) entre la luminance à laquelle le capteur est exposé et la tension créée par les éléments de détection. Cela signifie que si nous enregistrons la tension créée par chaque élément de détection, ainsi que l'emplacement sur la surface de la Terre d'où provient le rayonnement qui a créé le courant, et que nous disposons toutes les tensions mesurées par chacun des éléments de détection dans un schéma spatial en fonction de leur emplacement sur la surface de la Terre, nous pouvons créer une image cartographique dans laquelle les valeurs de chaque pixel sont les tensions mesurées dans cette zone. C'est essentiellement de cette manière que sont créées les images satellites.

Quantification

Cependant, les nombres numériques que nous voyons lorsque nous chargeons une image satellite dans un logiciel ne correspondent pas aux tensions mesurées. Il y a deux raisons liées à cela : réduire la quantité de mémoire informatique nécessaire pour stocker les mesures, tout en fournissant des données avec une précision adéquate. Les tensions mesurées prennent la forme de nombres décimaux, comme 3,560194 mV, qui nécessitent par défaut 32 bits de mémoire informatique pour être stockés (ou 64 bits si un grand nombre de décimales est nécessaire). Mais en réalité, les mesures de tension dans les capteurs des satellites ne sont pas précises à six décimales (comme dans l'exemple ci-dessus) ou à un autre grand nombre de décimales, et il n'est donc pas nécessaire d'utiliser 32 bits de mémoire pour stocker la tension de chaque élément de détection. Ce qui est généralement fait à la place, c'est que pendant, ou même avant, le développement du capteur, la précision nécessaire des mesures est déterminée, et l'instrument est construit pour effectuer des mesures avec cette précision (qui est ensuite testée plus tard lors de l'étalonnage de l'instrument). Une fois que la précision des mesures est connue, et en ajoutant des informations sur la gamme de rayonnements/tensions à laquelle l'instrument doit être sensible, le nombre de bits nécessaires pour contenir les mesures peut alors être déterminé, et la tension peut être convertie en un nombre numérique qui peut être stocké comme un nombre entier, ce qui nécessite beaucoup moins de mémoire informatique pour le stockage.

Imaginons, par exemple, que vous souhaitiez mesurer la taille d'une personne et que vous vouliez la mesurer au cm près. En utilisant les centimètres et sachant que tous les adultes mesurés dans l'histoire de ces mesures mesureraient entre 54 cm et 272 cm, vous réalisez que vous n'avez besoin que d'être capable de mesurer 219

valeurs différentes (les deux extrêmes inclus) : 54 cm, 55 cm, 56 cm..... 272 cm. Pour stocker 219 valeurs différentes, vous n'auriez besoin que de 8 bits, car $2^8 = 256$, ce qui vous laisse même quelques valeurs supplémentaires dont vous n'avez pas besoin. Vous décidez alors de coder une taille de 54 cm avec la valeur « 0 », 55 cm avec la valeur « 1 », et ainsi de suite (les valeurs du code sont simplement 54 de moins que les hauteurs correspondantes en cm). Maintenant, lorsque vous mesurez quelqu'un (avec votre mètre ruban très précis) et que vous estimez qu'il mesure 173,234 cm, vous vous dites que vous n'avez besoin que d'une précision en cm, vous arrondissez donc à 173 cm et vous codez $173 - 54 =$ « 119 ». C'est l'équivalent du nombre numérique que vous voyez dans un pixel d'une bande donnée d'une image. Lorsque vous savez comment le code (« 119 ») se rapporte à la mesure réelle (173 cm), la valeur prend tout son sens. Ce processus de codage est appelé *quantification*.

Numéros numériques à la luminance TOA

Lorsque vous ouvrez une image satellite dans un logiciel, la valeur initiale que vous voyez dans chaque pixel est appelée un numéro numérique (en anglais « Digital Number », DN), qui est équivalent au code (« 119 ») dans l'exemple ci-dessus. Bien que les DN soient suffisants pour produire une belle visualisation de l'image, si vous voulez traiter l'image comme une série de mesures radiométriques, vous devez la convertir en une quantité physique (comme les 173 cm correspondant à la taille de la personne mesurée dans cet exemple). En télédétection, la traduction d'une valeur de nombre numérique en une mesure de luminance est vraiment un processus en deux étapes. Dans la télédétection optique passive, l'étalonnage du capteur a permis d'établir une relation connue entre la luminance à laquelle le capteur est exposé et la tension générée par le rayonnement, et la quantification a ensuite été utilisée pour coder la tension mesurée sous forme de nombre numérique. Pour convertir ce nombre numérique (DN) en luminance, il faut donc en théorie d'abord une conversion en tension, puis une conversion en luminance. Heureusement, il est facile d'effectuer ces deux conversions en même temps. Nous allons voir comment cela est fait pour les données Landsat à titre d'exemple, mais le même principe s'applique aux données de tous les autres capteurs optiques passifs des satellites.

Les métadonnées Landsat contiennent des coefficients d'étalonnage radiométrique (Figure 20) qui vous permettent de convertir directement les valeurs DN en luminance au sommet de l'atmosphère (TOA). Ces coefficients contiennent un terme de « multiplication » et un terme d'« addition », chacun étant unique à chaque bande, et ils sont utilisés comme suit : $L_{\lambda} = \text{mult}_{\lambda} * \text{DN} + \text{add}_{\lambda}$, où L est la luminance et λ , la notation commune pour la longueur d'onde, indique une bande donnée.

```

GROUP = RADIOMETRIC_RESCALING
RADIANCE_MULT_BAND_1 = 1.2168E-02
RADIANCE_MULT_BAND_2 = 1.2460E-02
RADIANCE_MULT_BAND_3 = 1.1482E-02
RADIANCE_MULT_BAND_4 = 9.6820E-03
RADIANCE_MULT_BAND_5 = 5.9249E-03
RADIANCE_MULT_BAND_6 = 1.4735E-03
RADIANCE_MULT_BAND_7 = 4.9664E-04
RADIANCE_MULT_BAND_8 = 1.0957E-02
RADIANCE_MULT_BAND_9 = 2.3156E-03
RADIANCE_MULT_BAND_10 = 3.3420E-04
RADIANCE_MULT_BAND_11 = 3.3420E-04
RADIANCE_ADD_BAND_1 = -60.83886
RADIANCE_ADD_BAND_2 = -62.29974
RADIANCE_ADD_BAND_3 = -57.40868
RADIANCE_ADD_BAND_4 = -48.41023
RADIANCE_ADD_BAND_5 = -29.62465
RADIANCE_ADD_BAND_6 = -7.36738
RADIANCE_ADD_BAND_7 = -2.48320
RADIANCE_ADD_BAND_8 = -54.78706
RADIANCE_ADD_BAND_9 = -11.57798
RADIANCE_ADD_BAND_10 = 0.10000
RADIANCE_ADD_BAND_11 = 0.10000

```

Figure 20: Coefficients de calibrage utilisés pour convertir les valeurs DN en valeurs de luminance TOA pour Landsat 8. Par

Anders Knudby, CC BY 4.0.

Afin de produire une couche de données matricielles dans laquelle chaque valeur de cellule est une mesure de la luminance TOA, vous pouvez donc utiliser la « Calculatrice matricielle » offert dans tous les logiciels de SIG/télétection. Certains logiciels disposent même d’outils dédiés à la « calibration radiométrique Landsat » qui trouvent les informations pertinentes dans les métadonnées et les appliquent pour vous.

Numéros numériques à la réflectance TOA

La conversion des valeurs DN en une mesure de la luminance est une étape importante vers l’utilisation des données de télétection d’une manière quantitative — dans laquelle les valeurs de chaque cellule ont une signification physique réelle. Cependant, il y a un problème principal avec les mesures de luminance : la luminance provenant d’une zone dépend à la fois des caractéristiques physiques de cette zone, comme ce qui se trouve à la surface, mais elle dépend également de la quantité de rayonnement incident sur cette zone. Il est facile de le confirmer par une simple expérience : approchez un interrupteur près de vous, assurez-vous qu’il est allumé et fixez votre regard sur un mur. Notez sa luminosité (une approximation humaine de la quantité de rayonnement qui en émane). Éteignez maintenant la lumière et notez si le mur est devenu plus sombre ou plus clair. Il est devenu plus sombre, bien sûr, ce qui signifie qu’il y a maintenant moins de rayonnement qui en émane. Et cela se produit alors que le mur lui-même n’a pas changé du tout ! Dans le contexte de la télétection, les mesures de luminance sont donc rarement utilisées directement, mais sont typiquement utilisées comme une étape vers une quantité radiométrique qui est plus étroitement liée à la surface d’où il

vient : la réflectance. Pour un rayonnement isotrope, la réflectance (R) d'une surface donnée est le rapport entre la luminance sortante et la luminance entrante, également souvent exprimée sur la base de l'existance (E) et de l'irradiance (I) :

$$R = \frac{L_{up}}{L_{down}} = \frac{E_{up}}{I_{down}}$$

Dans l'exemple du mur, lorsque vous éteignez la lumière, la luminance entrante et sortante est réduite, et la réflectance du mur reste inchangée. Cela s'explique par le fait que la réflectance est une propriété physique du mur qui est indépendante de l'éclairage. Cela en fait un outil idéal pour la télédétection, car il est possible de comparer plusieurs observations avec des éclairagements différents, pour des utilisations telles que le suivi de l'évolution d'une zone dans le temps ou la comparaison de différentes zones sur la Terre.

Si nous convertissons les valeurs de luminance du TOA en valeurs de réflectance, nous obtenons ce que l'on appelle la réflectance du TOA, qui est la réflectance de l'ensemble du système Terre-Atmosphère. Pour comprendre ce que cela signifie, imaginez que vous êtes un astronaute et que vous placez un hula hoop de façon horizontale juste au niveau du sommet de l'atmosphère (si le sommet de l'atmosphère était possible à définir et à trouver !), et que vous mesurez ensuite la quantité de rayonnement passant à travers l'anneau dans les orientations ascendantes et descendantes. En divisant l'un par l'autre, on obtient la réflectance TOA.

Le calcul de la réflectance TOA s'effectue un peu différemment selon le type de données. Pour Landsat 8, il existe un autre ensemble de coefficients de calibration que vous pouvez appliquer exactement comme vous pouvez appliquer ceux de la luminance TOA (Figure 21).

```
REFLECTANCE_MULT_BAND_1 = 2.0000E-05
REFLECTANCE_MULT_BAND_2 = 2.0000E-05
REFLECTANCE_MULT_BAND_3 = 2.0000E-05
REFLECTANCE_MULT_BAND_4 = 2.0000E-05
REFLECTANCE_MULT_BAND_5 = 2.0000E-05
REFLECTANCE_MULT_BAND_6 = 2.0000E-05
REFLECTANCE_MULT_BAND_7 = 2.0000E-05
REFLECTANCE_MULT_BAND_8 = 2.0000E-05
REFLECTANCE_MULT_BAND_9 = 2.0000E-05
REFLECTANCE_ADD_BAND_1 = -0.100000
REFLECTANCE_ADD_BAND_2 = -0.100000
REFLECTANCE_ADD_BAND_3 = -0.100000
REFLECTANCE_ADD_BAND_4 = -0.100000
REFLECTANCE_ADD_BAND_5 = -0.100000
REFLECTANCE_ADD_BAND_6 = -0.100000
REFLECTANCE_ADD_BAND_7 = -0.100000
REFLECTANCE_ADD_BAND_8 = -0.100000
REFLECTANCE_ADD_BAND_9 = -0.100000
```

Figure 21: Coefficients de calibrage utilisés pour convertir les valeurs DN en valeurs de réflectance TOA pour Landsat 8. Notez que les valeurs pour les bandes 10 et 11 ne sont pas fournies, car ces bandes ne mesurent pas le rayonnement solaire réfléchi (sauf à un degré négligeable). Notez également que les coefficients sont les mêmes pour toutes les bandes – pour Landsat 8, les valeurs DN

ont été générées spécifiquement pour chaque scène afin de garantir des coefficients de calibration de réflectance TOA constants. Par

Anders Knudby, CC BY 4.0.

Pour les anciennes données Landsat, ainsi que pour plusieurs autres types de capteurs, il est nécessaire de calculer explicitement la quantité de rayonnement solaire entrant au sommet de l'atmosphère, et de diviser la luminance par cette valeur tout en corrigeant la différence entre luminance et irradiance et l'angle du rayonnement solaire entrant. Cela se fait de manière générique avec l'équation suivante (réarrangée pour des raisons pédagogiques à partir de l'équation 4, page 79, du manuel Landsat 7), expliquée plus en détail ci-dessous :

$$R_{TOA} = \frac{\pi * L}{\frac{1}{d^2} * I_{Sun} * \cos\theta_{Sun}}$$

où R_{TOA} est la réflectance au sommet de l'atmosphère, L est la luminance (vers le haut), d est la distance Terre-Soleil en unités astronomiques, I_{Sun} est l'irradiance solaire extraterrestre moyenne, et θ_{Sun} est l'angle zénithal solaire – l'angle entre la direction vers le Soleil et la normale à la surface de la Terre. L'équation est une division entre le rayonnement sortant dans toutes les directions vers le haut (c'est-à-dire l'existance, au numérateur) et le rayonnement entrant dans toutes les directions vers le bas (c'est-à-dire l'irradiance, au dénominateur).

Le numérateur est relativement simple : la luminance ascendante, L , qui est quantifiée comme le rayonnement par unité de surface exposée se déplaçant dans des directions comprises dans un angle solide spécifié, est multipliée par π pour la convertir en une valeur qui quantifie plutôt le rayonnement se déplaçant dans toutes les directions ascendantes, afin de la rendre comparable à l'irradiance solaire entrante au dénominateur.

L'irradiation solaire moyenne entrante au sommet de l'atmosphère. Comme la distance entre le Soleil et la Terre varie tout au long de l'année et que la quantité d'irradiation solaire varie avec elle, nous devons corriger cette valeur par le facteur $1/d^2$, où d est la distance réelle entre le Soleil et la Terre au moment de l'acquisition de l'image. Si le Soleil était au zénith (directement au-dessus de la Terre), cela permettrait de quantifier l'irradiance entrante par unité de surface. Dans le cas où le Soleil n'est pas au zénith, il faut multiplier cette valeur par le cosinus de l'angle zénithal solaire pour obtenir l'éclairement énergétique par unité de surface exposée.

Note: Dans les deux équations ci-dessus, l'indice λ indiquant que ces calculs doivent être effectués bande par bande a été omis pour plus de clarté, mais gardez cela à l'esprit. L varie d'une bande à l'autre, tout comme I_{Sun} , de sorte que les deux doivent être calculés spécifiquement pour chaque bande afin d'être corrects.

Grâce à ces calculs, vous êtes en mesure de convertir les valeurs DN de chaque pixel en une mesure de la réflectance TOA, qui est une propriété physique du système Terre-Atmosphère. Le fait que la réflectance TOA ne soit pas un nombre arbitraire dépendant de l'étalonnage du capteur et de l'éclairage, mais plutôt une

propriété physique des parties de la Terre visibles sur l'image, nous permet d'utiliser des images capturées par différents capteurs, à différents moments et pour différentes zones, et de comparer directement les valeurs de chaque pixel. Le seul problème qui subsiste dans cette approche est que la plupart des chercheurs sont plus intéressés par la réflectance de la surface de la Terre que par la réflectance au sommet de l'atmosphère, mais l'état de l'atmosphère et l'état de la surface influencent tous deux la réflectance TOA. L'étape suivante du traitement des données de chaque pixel consiste donc à supprimer l'influence de l'atmosphère sur ces valeurs, pour arriver à une mesure de la réflectance de la surface !

4.

INTERACTIONS D'EMR AVEC L'ATMOSPHÈRE ET LA SURFACE DE LA TERRE

Après que le rayonnement électromagnétique a été créé par le Soleil, la partie de celui-ci qui a trouvé son chemin à travers le vide de l'espace jusqu'au sommet de l'atmosphère terrestre doit traverser l'atmosphère, être réfléchi par la surface de la Terre, traverser à nouveau l'atmosphère sur son chemin de retour vers l'espace, puis arriver au capteur afin d'être enregistrée. Si rien ne se produit dans le champ de rayonnement lorsqu'il traverse l'espace vide, plusieurs choses se produisent lorsqu'il interagit avec l'atmosphère et la surface de la Terre. C'est grâce à ces interactions que le rayonnement mesuré finit par contenir des informations sur l'environnement terrestre. Il est donc important d'examiner de plus près ce qui se passe exactement dans ces interactions et comment cela affecte le champ de rayonnement.

Interactions avec l'atmosphère

L'interaction entre le rayonnement électromagnétique et l'atmosphère terrestre peut être considérée comme ayant trois composantes : la **réfraction** qui modifie la direction de propagation du champ de rayonnement en raison des différences de densité entre l'espace extra-atmosphérique et l'atmosphère, la **diffusion** qui modifie la direction de propagation des photons individuels lorsqu'ils sont absorbés et réémis par des gaz, des aérosols ou d'autres constituants atmosphériques sans changer de longueur d'onde, et la **absorption** qui convertit les photons en vibrations dans une molécule, énergie qui est (ultérieurement) réémise sous la forme d'un ou plusieurs photons de plus grande longueur d'onde. Chacun de ces phénomènes sera examiné plus en détail ci-dessous.

Réfraction

La réfraction est la courbure (et le ralentissement) de la direction de propagation du rayonnement électromagnétique lorsqu'il se déplace entre deux milieux de densités différentes. Cela se produit lorsque le rayonnement arrive de l'espace (densité ≈ 0) et pénètre dans l'atmosphère (densité > 0). L'angle auquel la direction de la propagation peut se changer est déterminé par les indices de réfraction des deux milieux. L'indice de réfraction d'un milieu (n) est déterminé comme le rapport entre la vitesse du rayonnement

électromagnétique dans le vide (c) et la vitesse similaire dans le milieu (cn) : $n = c/cn$. L'indice de réfraction d'une atmosphère standard est de 1,0003, tandis que l'indice de réfraction de l'eau est de 1,33. En utilisant les indices de réfraction des deux milieux, la quantité de réfraction peut être déterminée avec la loi de Snell : $n_1 * \sin(\Theta_1) = n_2 * \sin(\Theta_2)$, où n sont les indices de réfraction des deux milieux et Θ sont les angles auxquels la direction de propagation coupe la normale de la surface séparant les deux milieux (figure 22). La réfraction est rarement un facteur pertinent dans l'utilisation pratique des données de télédétection. Sa seule influence importante concerne le géoréférencement des images recueillies lorsque le Soleil est proche de l'horizon, et c'est un problème qui est presque toujours traité par le fournisseur d'images. Une situation importante dans laquelle la réfraction est importante et doit être prise en compte est celle où un analyste d'images a besoin de géolocaliser précisément des objets sous-marins (comme les caractéristiques du fond marin dans les zones côtières).

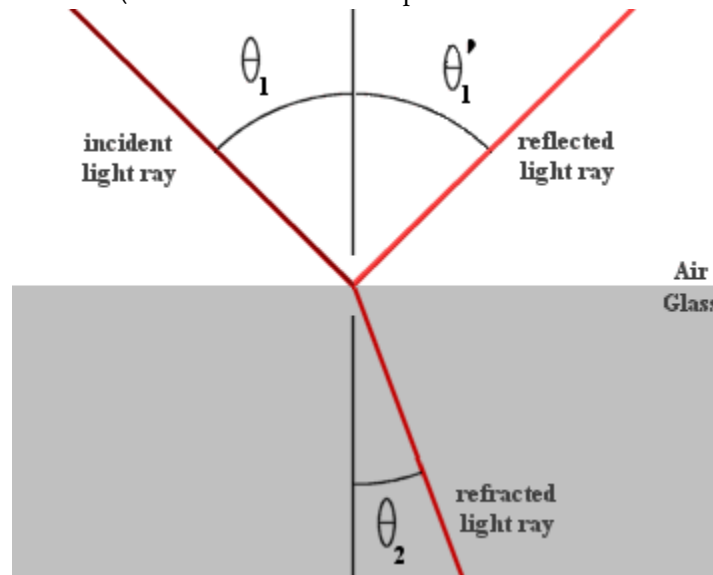


Figure 22: La réfraction selon la loi de Snell. Si cette image illustre la réfraction de la lumière passant de l'air au verre, le même principe s'applique à l'interface air-eau. Exemple snells law par RJHall, Wikimedia Commons, CC BY-SA 2.0 AT.

Diffusion

L'un des deux derniers processus qui influencent le rayonnement électromagnétique lorsqu'il traverse l'atmosphère est la diffusion. La diffusion se produit lorsqu'un photon interagit avec un objet dans l'atmosphère qui le fait changer de direction. En fonction de la taille de l'objet avec lequel le photon interagit, on distingue deux types de diffusion. La diffusion de **Rayleigh** se produit lorsque l'objet est beaucoup plus petit que la longueur d'onde du rayonnement. Dans le cas de la lumière solaire et de l'atmosphère terrestre, cela signifie que la diffusion de Rayleigh est causée par les gaz atmosphériques tels que N_2 , O_2 , CO_2 , etc. La diffusion de **Mie** se produit lorsque la taille de l'objet est semblable à la longueur d'onde du rayonnement, ce qui signifie qu'elle est causée par les aérosols comme les particules de fumée et de poussière. Une diffusion supplémentaire peut se produire si le rayonnement interagit avec des particules dont la taille est supérieure à sa longueur d'onde, comme les gouttelettes d'eau ou les particules de sable. Alors que la réfraction est prévisible

et peut être déterminée par la loi de Snell, la diffusion est un processus intrinsèquement stochastique : ce qui arrive à un photon individuel lorsqu'il traverse l'atmosphère est entièrement imprévisible, y compris s'il subit ou non une diffusion et, le cas échéant, dans quelle direction il est réémis. En revanche, l'ampleur et la direction de la diffusion qui se produit en moyenne pour les nombreux photons d'un champ de rayonnement sont prévisibles.

Diffusion de Rayleigh

Un fait qui revêt une grande importance pour la télédétection de la Terre est que l'ampleur de la diffusion de Rayleigh est inversement liée à la puissance 4 de la longueur d'onde du rayonnement. En d'autres termes, les rayonnements de courte longueur d'onde sont beaucoup plus diffusés par la diffusion de Rayleigh que les rayonnements de grande longueur d'onde. Dans les longueurs d'onde visibles, cela signifie que la lumière bleue est plus diffusée que la lumière verte, qui à son tour est plus diffusée que la lumière rouge. C'est ce processus qui donne aux océans de la Terre leur aspect bleu lorsqu'ils sont vus de l'espace. Ce qui se passe, c'est que sur les surfaces terrestres très sombres, comme les océans, la majorité du rayonnement qui atteint la surface de la Terre est absorbée plutôt que réfléchi par celle-ci. Ce qui est visible depuis l'espace n'est donc pas le rayonnement réfléchi par la surface, mais plutôt le rayonnement diffusé par l'atmosphère. Comme les longueurs d'onde bleues sont celles qui sont le plus fortement diffusées par la diffusion de Rayleigh, l'ensemble de ce rayonnement diffusé paraît bleu (figure 23). Un autre effet de la diffusion Rayleigh est que, indépendamment de ce qui se trouve à la surface de la Terre, un capteur spatial détectera une quantité importante de lumière bleue provenant du système Terre-Atmosphère. Cela peut être un problème, car le « signal bleu » provenant de l'atmosphère écrase les variations de la « réflectance bleue » à la surface. Mais cela peut aussi être un avantage, car les mesures dans les longueurs d'onde bleues peuvent aider à évaluer la force de la diffusion Rayleigh dans le spectre visible et infrarouge, qui peut à son tour être corrigée. C'est la base de la bande « aérosol » qui a été inclus dans Landsat 8 OLI (mais qui n'existait pas sur les instruments précédents), dans Sentinel-2 et dans les capteurs WorldView-2 et -3.



Figure 23: La Terre vue de l'espace. The Blue Marble (globe west) par Reto Stöckli, Visible Earth (NASA), Visible Earth Image Use Policy.

Alors que toute diffusion dans l'atmosphère est une source de bruit (pour ceux qui souhaitent utiliser l'imagerie satellitaire pour caractériser la surface de la Terre), la diffusion Rayleigh est une source de bruit relativement bénigne parce que sa dépendance à la longueur d'onde la rend largement prévisible et parce que les gaz qui en sont responsables ont tendance à avoir des concentrations stables dans l'espace et dans le temps. La diffusion Rayleigh n'est donc pas une source de grande incertitude pour la plupart des applications de télédétection.

Diffusion de Mie

La diffusion de Mie, parce que sa force et sa dépendance à la longueur d'onde dépendent du type et de la densité des particules qui la provoquent, varie considérablement dans le temps et l'espace. C'est pourquoi elle est l'une des plus importantes causes d'incertitude en télédétection, en particulier lorsque l'on utilise des données satellitaires pour étudier les parties sombres de la surface de la Terre à partir desquelles la quantité de rayonnement réfléchi est faible par rapport au signal total de la diffusion atmosphérique. Pour la même raison, il est difficile de généraliser son importance, mais de manière générale, la force de la diffusion de Mie dépasse celle de la diffusion de Rayleigh et, bien qu'elle diminue toujours avec l'augmentation de la longueur d'onde, son influence s'étend plus loin dans le spectre infrarouge. Étant donné que la diffusion de Mie est causée par les particules atmosphériques, elle augmente souvent de façon spectaculaire lors de tempêtes de poussière, de feux de forêt ou d'autres événements ayant entraîné une augmentation de la charge d'aérosols dans l'atmosphère. Un tel exemple est illustré à la figure 24.



Figure 24: Forte augmentation de la diffusion de Mie causée par les feux de forêt en Australie, 2009. Smoke clouds par Warren, Flickr, CC BY-NC-SA 2.0.

Absorption

La dernière chose importante qui arrive au rayonnement électromagnétique lorsqu'il traverse l'atmosphère

est qu'il est partiellement absorbé par les gaz atmosphériques (principalement H_2O , CO_2 et O_3). Bien que l'énergie absorbée soit finalement réémise par ces molécules de gaz, la réémission se produit à des longueurs d'onde typiquement en dehors du spectre pris en compte dans la télédétection optique (mais qui peut être important pour la télédétection thermique), donc pour des raisons pratiques, les photons absorbés peuvent être considérés comme disparus lorsqu'ils sont absorbés. L'intensité de l'absorption dépend fortement de la longueur d'onde, car elle se produit plus facilement lorsque le rayonnement a une longueur d'onde (fréquence) semblable à la fréquence de résonance du gaz qui effectue l'absorption, qui dépend elle-même de sa structure atomique ou moléculaire. Par exemple, en raison de sa structure moléculaire, l' O_2 absorbe particulièrement bien les rayonnements électromagnétiques dont la longueur d'onde se situe autour de 760 nm, mais pas à 750 ou 770 nm. Il existe des longueurs d'onde similaires pour lesquelles d'autres gaz sont efficaces ou non pour absorber le rayonnement électromagnétique et, en combinaison, les gaz atmosphériques laissent certaines longueurs d'onde traverser l'atmosphère sans presque aucune absorption, tandis que d'autres longueurs d'onde sont presque entièrement absorbées avant d'atteindre la surface de la Terre (figure 25 et figure 26). Comme le montre clairement la figure 26, la vapeur d'eau est responsable d'une grande partie de l'absorption gazeuse totale des EMR dans l'atmosphère, y compris dans le spectre visible (non représenté clairement sur cette figure). Il s'agit d'un défi important pour la télédétection, car, alors que les concentrations des autres gaz sont relativement stables dans le temps et l'espace, les concentrations de vapeur d'eau varient fortement dans le temps (jours humides ou secs) et dans l'espace (Arctique sec ou tropical humide).

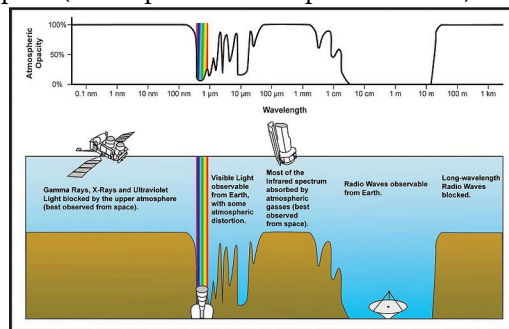


Figure 25: Opacité atmosphérique par longueur d'onde sur l'ensemble du spectre électromagnétique. Atmospheric electromagnetic transmittance or opacity par Edwtie, créé par la NASA, Wikimedia Commons, public domain.

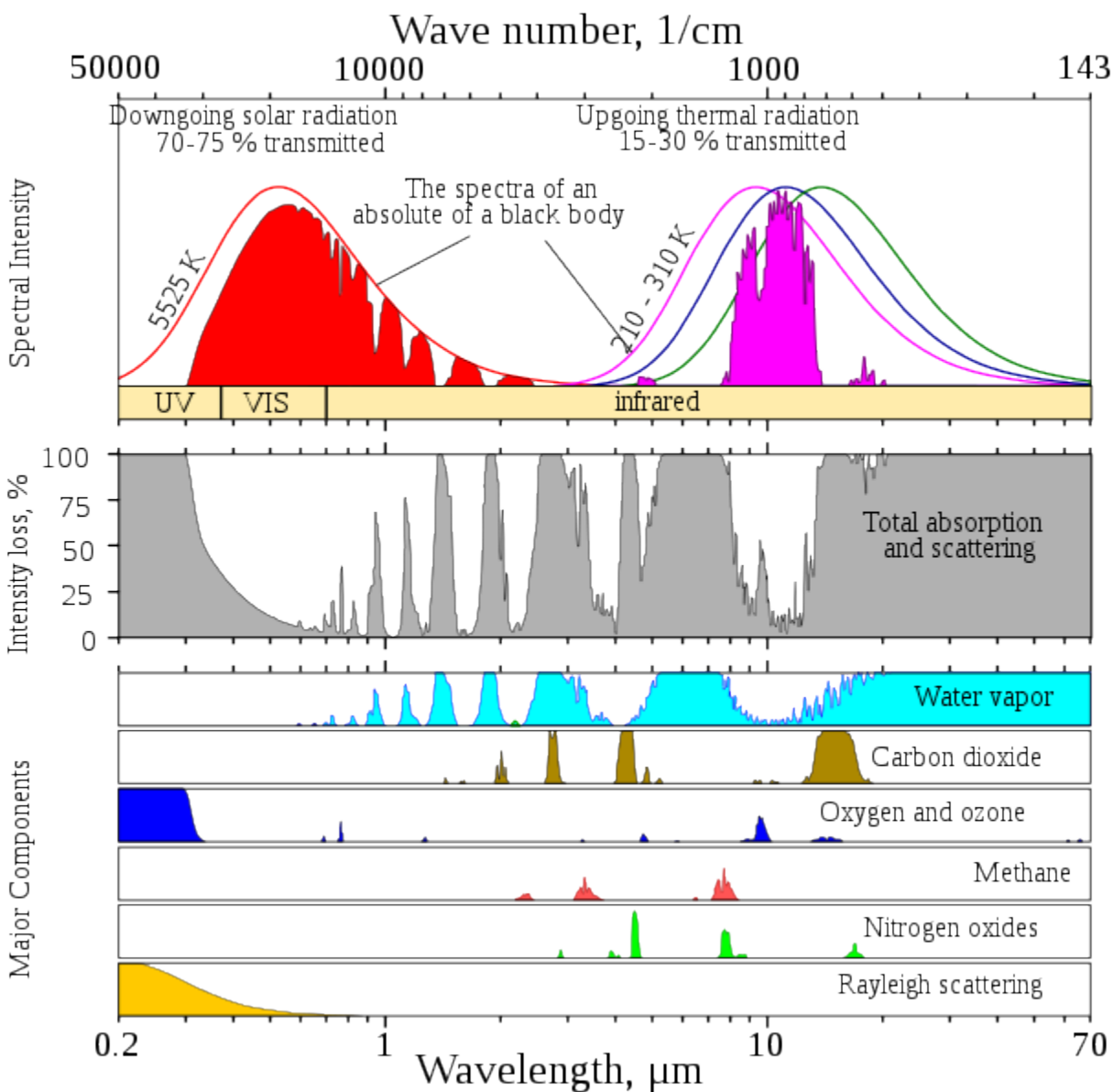


Figure 26: Comme la figure 25, mais couvrant une plus large gamme de longueurs d'onde et montrant quels gaz sont principalement responsables de l'absorption à quelle longueur d'onde. Atmospheric Transmission par Cepheiden, Wikimedia

Commons, CC0 1.0.

Interactions avec la surface

La partie du champ de rayonnement qui a traversé l'atmosphère sans être absorbée ou diffusée vers l'espace atteint maintenant la surface de la Terre. Pour toute longueur d'onde pertinente pour la télédétection, il ne peut arriver que deux choses à chaque photon : il peut être absorbé par la surface de la Terre ou il peut être réfléchi vers l'espace. La probabilité que la réflexion plutôt que l'absorption se produise est appelée la réflectance

de la surface, et elle dépend du matériau de la surface ainsi que de la longueur d'onde du rayonnement entrant. Chaque matériau de surface possède une « signature » unique qui définit la proportion du rayonnement réfléchi pour chaque longueur d'onde. Par exemple, l'eau réfléchit une petite quantité de longueurs d'onde bleues et vertes (généralement autour de 5 à 10 % selon la turbidité), une moindre quantité de longueurs d'onde rouges et presque rien dans les longueurs d'onde infrarouges. La végétation, quant à elle, réfléchit environ la moitié de tous les rayonnements infrarouges entrants, à l'exception de certaines longueurs d'onde précises qui sont effectivement absorbées par l'eau liquide dans les feuilles. Ces signatures spectrales sont généralement représentées sous forme de graphiques, avec les longueurs d'onde sur l'axe des x et la réflectance sur l'axe des y (comme dans la figure 27).

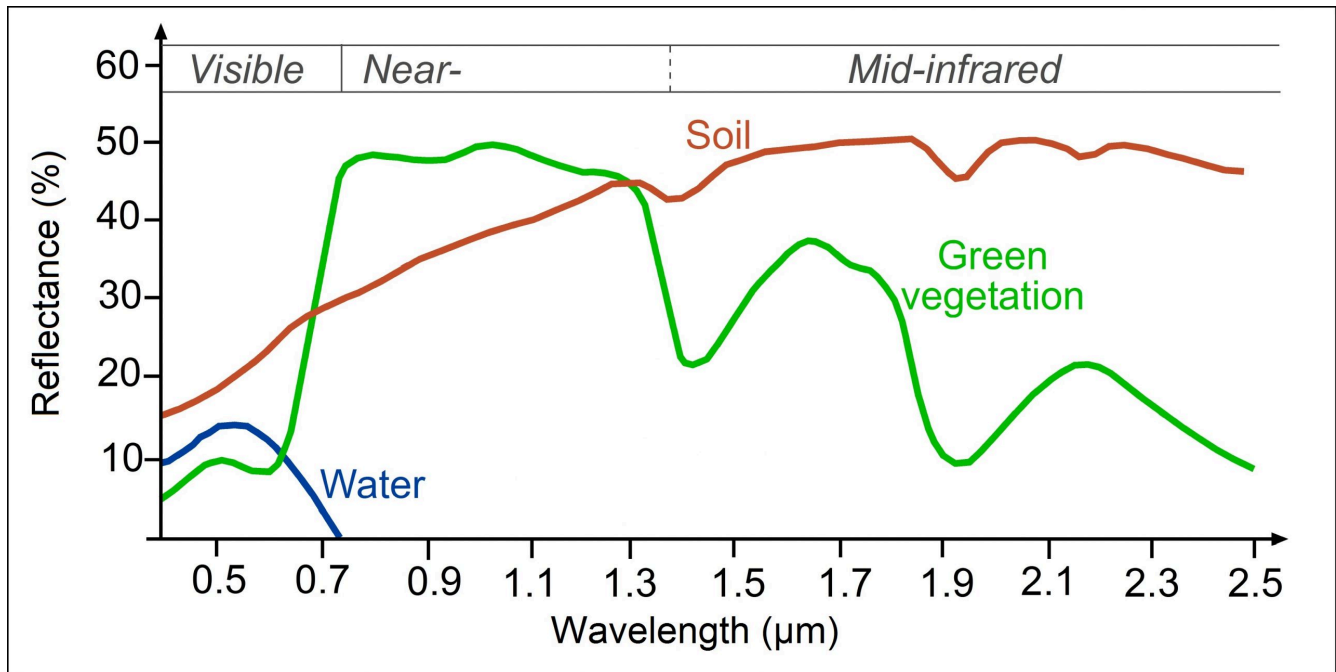


Figure 27: Signatures spectrales de types communs largement définis de matériaux de la surface terrestre. Reflectance of water, soil and vegetation at different wavelengths par Science Education through Earth Observation for High Schools (SEOS), CC BY-NC-SA 2.0.

Les signatures spectrales sont ce qui nous permet de différencier les différents matériaux à la surface de la Terre lorsque nous regardons une image satellite. Comme le montre la figure 27, l'eau a une réflectance quasi nulle aux longueurs d'onde supérieures à 0,7 µm (700 nm), tandis que le sol et la végétation verte ont tous deux des réflectances d'environ 40 % à 1,3 µm. La mesure de la quantité de rayonnement réfléchi par le système Terre-Atmosphère à 1,3 µm sera donc particulièrement utile pour différencier l'eau des deux types de surfaces terrestres. De même, les mesures aux longueurs d'onde autour de 1,4 µm (où l'eau liquide de la végétation est un fort absorbeur) ou 1,9 µm (idem) peuvent être efficaces pour différencier le sol de la végétation verte.

À titre d'exemple plus détaillées, les signatures spectrales ont été efficaces pour la prospection géologique

à grande échelle, car différents minéraux (qui peuvent être caractéristiques de différentes conditions de subsurface) peuvent être identifiés grâce à leurs signatures spectrales uniques (figure 28).

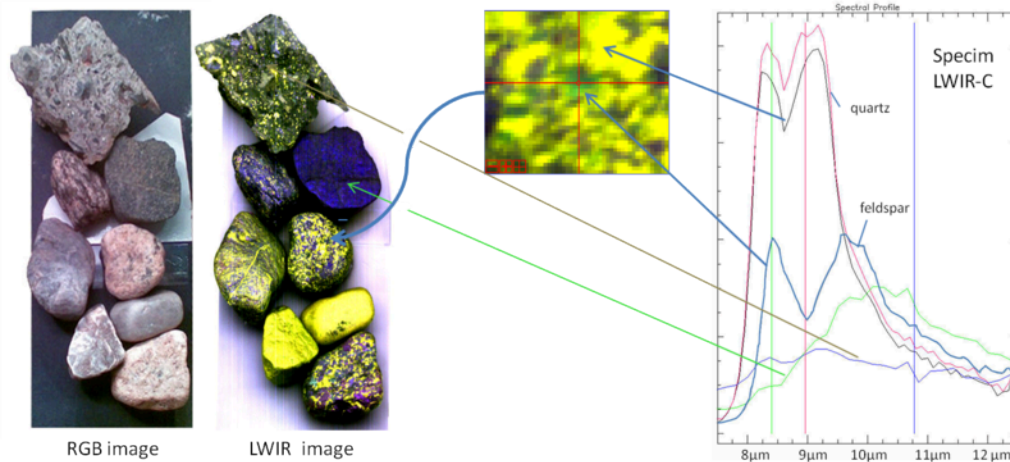


Figure 28: Exemples de signatures spectrales de minéraux. HIS LWIR stones par Aappo Roos, Wikimedia Commons, CC BY-SA

3.0.

La partie du champ de rayonnement qui est réfléchi par la surface de la Terre doit naturellement traverser à nouveau l'atmosphère, avec les phénomènes de réfraction, de diffusion et d'absorption que cela implique, avant de pouvoir être mesurée par un capteur spatial. Bien qu'il y ait de nombreux avantages et inconvénients relatifs aux capteurs aéroportés par rapport aux capteurs spatiaux, la capacité des capteurs aéroportés à mesurer le champ électromagnétique réfléchi avant qu'il n'ait à traverser l'atmosphère une seconde fois est un avantage distinct.

Correction atmosphérique

À moins que vous ne soyez un scientifique de l'atmosphère, les interactions entre le champ de rayonnement électromagnétique et la surface de la Terre sont la seule chose importante qui vous intéresse en télédétection, car c'est ce qui vous permet d'utiliser le rayonnement mesuré pour déduire quelque chose sur l'état de la surface de la Terre. Par exemple, si elle est couverte d'eau ou de terre, ou la composante des minéraux de la terre, la santé de la végétation ou la quantité de sédiments dans l'eau. Le fait que des interactions se produisent également entre le champ électromagnétique et l'atmosphère, et que ces interactions influencent l'intensité du rayonnement mesuré par le capteur à différentes longueurs d'onde constitue une source importante de bruit. L'un des grands et difficiles défis de la télédétection est donc de corriger l'effet de l'atmosphère sur le rayonnement mesuré, convertissant ainsi nos mesures d'une description du système Terre-Atmosphère en une description de la surface de la Terre uniquement. Ce processus est appelé correction atmosphérique. En termes pratiques, la correction atmosphérique est généralement utilisée pour convertir une estimation de la radiance TOA ou de la réflectance TOA en une estimation de la réflectance de la surface. Cette conversion est importante, car elle

permet de faire des comparaisons directes entre les réflectances observées dans différentes bandes pour un pixel et les signatures spectrales mesurées sur Terre, comme celles de la figure 27 et de la figure 28.

Il existe différentes approches de la correction atmosphérique, allant de la plus simple à la plus compliquée. L'approche la plus simple consiste à supposer que, quelque part dans l'image pour laquelle vous souhaitez effectuer une correction atmosphérique, il existe un pixel qui est complètement, parfaitement, noir, dans toutes les bandes disponibles sur l'image. C'est ce qu'on appelle « l'hypothèse du pixel noir ». À proprement parler, il ne doit pas nécessairement s'agir d'un seul pixel, on peut également supposer qu'au moins un pixel a une réflectance nulle dans la bande 1, un pixel (peut-être le même, peut-être un autre) a une réflectance nulle dans la bande 2, et ainsi de suite. En supposant que cela soit vrai, pour chaque bande, le pixel de l'image dont la radiance mesurée provenant de lui est la plus faible (c'est-à-dire dont la valeur DN est la plus basse) doit être le pixel à réflectance nulle, et toute radiance mesurée provenant de lui doit provenir de la diffusion dans l'atmosphère (car si elle avait atteint la surface de la Terre, elle aurait été absorbée par celle-ci). Le pixel hypothétique à réflectance nulle peut donc donner une estimation de premier ordre de la quantité de radiance provenant de la diffusion dans l'atmosphère. En supposant en outre que cette quantité est constante sur l'ensemble de l'image, la « contribution atmosphérique » peut être soustraite de tous les pixels, ne laissant que ce qui a été réfléchi par la surface, appelée « contribution de surface ». Bien sûr, l'hypothèse de l'existence d'un pixel parfaitement noir dans l'image peut ne pas être vraie, et la contribution atmosphérique peut également ne pas être la même partout, mais en tant que méthode pouvant être appliquée sans avoir une connaissance réelle de l'état de l'atmosphère lorsque l'image a été acquise, ce n'est pas une mauvaise méthode.

La méthode du « dark-pixel » est un exemple de ce que l'on appelle les méthodes de correction atmosphérique basées sur l'image. D'autres méthodes de ce type font d'autres hypothèses, comme la présence d'une végétation dense et sombre aux caractéristiques spectrales typiques, ou la présence d'une gamme de types de surface aux signatures spectrales connues. Ces méthodes fonctionnent raisonnablement bien lorsque leurs hypothèses sont respectées, mais peuvent échouer de manière spectaculaire lorsqu'elles ne le sont pas.

Les méthodes plus sophistiquées reposent sur une certaine forme d'information sur l'état de l'atmosphère au moment de l'acquisition de l'image, généralement la quantité et le type d'aérosols et la quantité de vapeur d'eau. Avec ces informations, des modèles numériques qui quantifient la diffusion et l'absorption, et donc le transfert du rayonnement de différentes longueurs d'onde à travers l'atmosphère, peuvent être utilisés pour modéliser la réflectance de surface qui devait exister, en combinaison avec l'atmosphère connue, pour produire la réflectance TOA observée. Certaines de ces méthodes ne nécessitent pas d'informations spécifiques sur les aérosols et la vapeur d'eau, mais plutôt que vous leur indiquiez lesquelles était la visibilité dans la zone au moment de l'acquisition des images — puisque la visibilité est réduite avec l'augmentation des quantités de vapeur d'eau dans l'atmosphère, les méthodes peuvent calculer la vapeur d'eau et partir de là. La correction atmosphérique est un domaine de recherche très actif dans le domaine de la télédétection optique, et les capteurs modernes des satellites sont régulièrement conçus pour inclure des bandes de mesures qui aident à l'estimation de la vapeur d'eau et de la charge d'aérosols en fonction de l'image.

Presque toutes les méthodes de correction atmosphérique partent de l'hypothèse fondamentale que

l'atmosphère est uniforme dans l'image satellite considérée. Bien que cela ne soit évidemment pas parfaitement vrai, il s'agit souvent, mais pas toujours, d'une hypothèse raisonnable. Les capteurs hyperspectraux, qui enregistrent le rayonnement dans de nombreuses bandes contiguës, souvent des centaines, ont la capacité d'estimer les charges de vapeur d'eau et d'aérosols sur une base par pixel, ce qui permet de cartographier efficacement les composantes atmosphériques contribuant à la plus grande incertitude, après quoi une correction atmosphérique par pixel peut être appliquée.

L'importance de la correction atmosphérique et de l'estimation de la réflectance de la surface pour chaque pixel n'est pas seulement de permettre la comparaison avec les signatures spectrales mesurées sur le terrain, ou d'identifier les matériaux de la surface terrestre. Plus fondamentalement, cette importance est due au fait que la réflectance de la surface est une propriété physique fondamentale de la surface, et qu'elle n'est pas influencée par les conditions d'éclairage, la conception du capteur, l'état de l'atmosphère, ou tout autre facteur qui diffèrent entre les différentes images de la même zone. Pour un pixel, une estimation de la réflectance de la surface est donc directement comparable entre des images prises à différents moments, avec différents capteurs, à différents angles zénithaux du soleil, et dans des atmosphères sèches ou humides, claires ou enfumées. Cela signifie qu'une zone peut être surveillée et qu'en l'absence de changement, la réflectance de surface estimée (par bande, c'est-à-dire la signature spectrale) ne devrait pas changer — ce qui signifie que si la signature spectrale change au fil du temps, c'est qu'un changement réel se produit à la surface. Une correction atmosphérique appropriée est donc (à quelques exceptions près) la base d'un vaste domaine de la télédétection appliquée — la détection du changement.

5.

VISUALISATION ET MANIPULATION D'IMAGES SATELLITES

La majeure partie de ce cours portera sur le type typique de « images satellites » que vous pouvez voir aux informations ou sur Google Earth, c'est-à-dire le type d'image produit par des capteurs optiques passifs. Ces images peuvent, à bien des égards, être considérées comme semblables aux photos que vous prenez avec votre téléphone, et vous pouvez donc faire des choses similaires pour les améliorer. Sauf que dans le contexte de la télédétection, « améliorer » ne signifie pas les retoucher pour supprimer les yeux rouges et créer des couleurs plus vives, mais les présenter de manière à mieux les interpréter à des fins déterminées. Il s'agit d'une forme de *visualisation scientifique*, un processus dans lequel l'affichage des données aide à les explorer et à les comprendre. Si l'exploration visuelle des données peut s'avérer plus utile que leur traitement par un algorithme informatique, c'est parce que les êtres humains sont exceptionnellement doués pour l'interprétation des modèles spatiaux. À titre d'exemple, imaginez que vous vous promenez dans un champ de l'est du Canada et que vous voyez la scène ci-dessous. Vous êtes immédiatement capable de réaliser que la partie sombre de l'image ne semble pas à sa place et qu'elle pourrait donc être importante, que la forme vous suggère qu'il s'agit d'un grand animal, et que vous pouvez faire appel à toutes sortes de contextes (la période de l'année, l'environnement, le fait que les ours sont connus pour vivre dans la région) pour vous aider à interpréter ce qui se passe.



Figure 29: Les humains sont exceptionnellement doués pour reconnaître les structures dans les images et, bien que la vision par ordinateur soit en plein développement, le cerveau humain surpasse encore souvent l'intelligence artificielle dans ce domaine. A Bear In The Woods par Linda Tanner, Flickr, CC BY-NC-ND 2.0.

Affichage en niveaux de gris, non amélioré

Il se passe beaucoup de choses en coulisses lorsque vous affichez une image sur un écran d'ordinateur. Pour en examiner les détails, nous prendrons d'abord l'exemple le plus simple possible, une image en niveaux de gris 8 bits, mais n'oubliez pas que les mêmes principes s'appliquent à des données de nature différente. Une image en niveaux de gris de 8 bits (communément, mais à tort, appeler aussi image en noir et blanc) correspond à une seule couche de données bidimensionnelle dont les valeurs vont de 0 à 255. Numériquement, le fichier image (par exemple, un fichier JPG, PNG, tif ou autre) contient un ensemble de chiffres, ainsi que des informations sur les dimensions de l'image. Dans l'exemple ci-dessous, l'image contient 16 pixels disposés en 4 lignes et 4 colonnes, et une valeur comprise entre 0 et 255 est associée à chaque cellule. Il s'agit de la valeur du numéro numérique (DN). La plupart des écrans d'ordinateur sont capables d'afficher 256 intensités lumineuses différentes, et utilisent pour cela trois couleurs (rouge, vert et bleu), ce qui permet de produire $256 * 256 * 256 = 16,8$ millions de couleurs. Imaginez qu'il y a trois petits « canons à lumière colorée » à l'intérieur de chaque pixel de l'écran d'ordinateur, qui déterminent la quantité de lumière à produire, de chaque couleur, dans chaque pixel. Lors de l'affichage d'images en niveaux de gris, l'intensité de la lumière produite en rouge,

vert et bleu est la même, ce qui donne 256 nuances de gris. La plus sombre d'entre elles (0 intensité lumineuse en rouge, 0 en vert et 0 en bleu) produit une couleur noire à l'écran, et la plus claire (255 en rouge, 255 en vert et 255 en bleu) produit du blanc. L'exemple de la figure 30 est donc la visualisation la plus simple (à droite) des valeurs DN du fichier (à gauche).

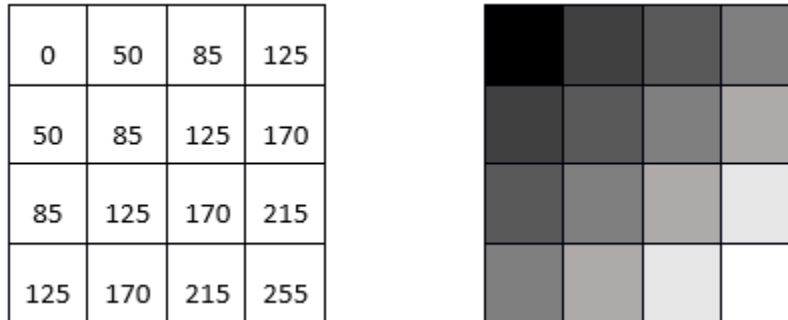


Figure 30: Image monocouche en niveaux de gris, montrant les chiffres numériques à gauche, et leur visualisation à droite. Par

Anders Knudby, CC BY 4.0.

Affichage en couleurs vrai, non amélioré

La plupart des images ont plus d'une couche de données, et peuvent donc être utilisées pour produire des images en couleur. Les appareils photo des téléphones portables enregistrent des fichiers d'images comportant trois couches, correspondant au rouge, au vert et au bleu, de sorte que la visualisation de ces fichiers est facile grâce aux trois « canons à lumière colorée » de votre écran d'ordinateur. Il n'est donc pas surprenant qu'une visualisation en couleurs soit une visualisation dans laquelle vous affichez la couche de données « rouge » à l'aide du « canon » rouge, et ainsi de suite. Par exemple :

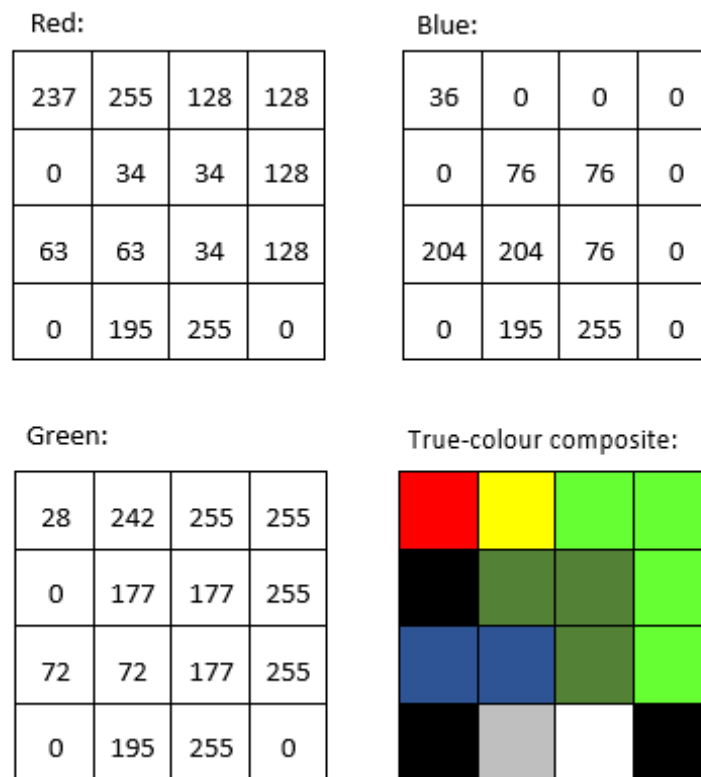


Figure 31: Les valeurs DN dans les bandes rouge, verte et bleue, et leur affichage direct dans un composite de couleurs vraies. Par Anders Knudby, CC BY 4.0.

Affichage en fausses couleurs, non amélioré

La plupart des capteurs optiques passifs des satellites enregistrent la lumière dans plus de trois « bandes » et produisent donc des fichiers d'images comportant plus de trois couches de données. Certaines de ces bandes peuvent être des nuances différentes de la lumière visible (par exemple, le violet), tandis que d'autres, comme l'infrarouge et l'ultraviolet, se situent en dehors du spectre visible par l'homme. Pour montrer de telles données, on utilise des visualisations en fausses couleurs. Le terme « fausses couleurs » indique simplement que ce qui est montré ne représente pas ce à quoi la zone de l'image ressemblerait pour un être humain, sans pour autant impliquer que l'image présente un défaut quelconque (en fait, les affichages en fausses couleurs sont très utiles). La visualisation en fausses couleurs est très variée — imaginez le nombre de combinaisons différentes de trois bandes que vous pouvez faire avec les données d'un capteur qui possède plus de 200 bandes... La visualisation en fausses couleurs la plus courante est souvent appelée « couleur-infrarouge », et elle affiche une bande « proche de l'infrarouge » à l'aide du canon rouge, une bande « rouge » à l'aide du canon vert, et une bande « verte » à l'aide du canon bleu. La figure 32 présente une comparaison côte à côte de l'affichage en couleurs vraies et en fausses couleurs d'une image Landsat contenant la ville de Zanzibar.

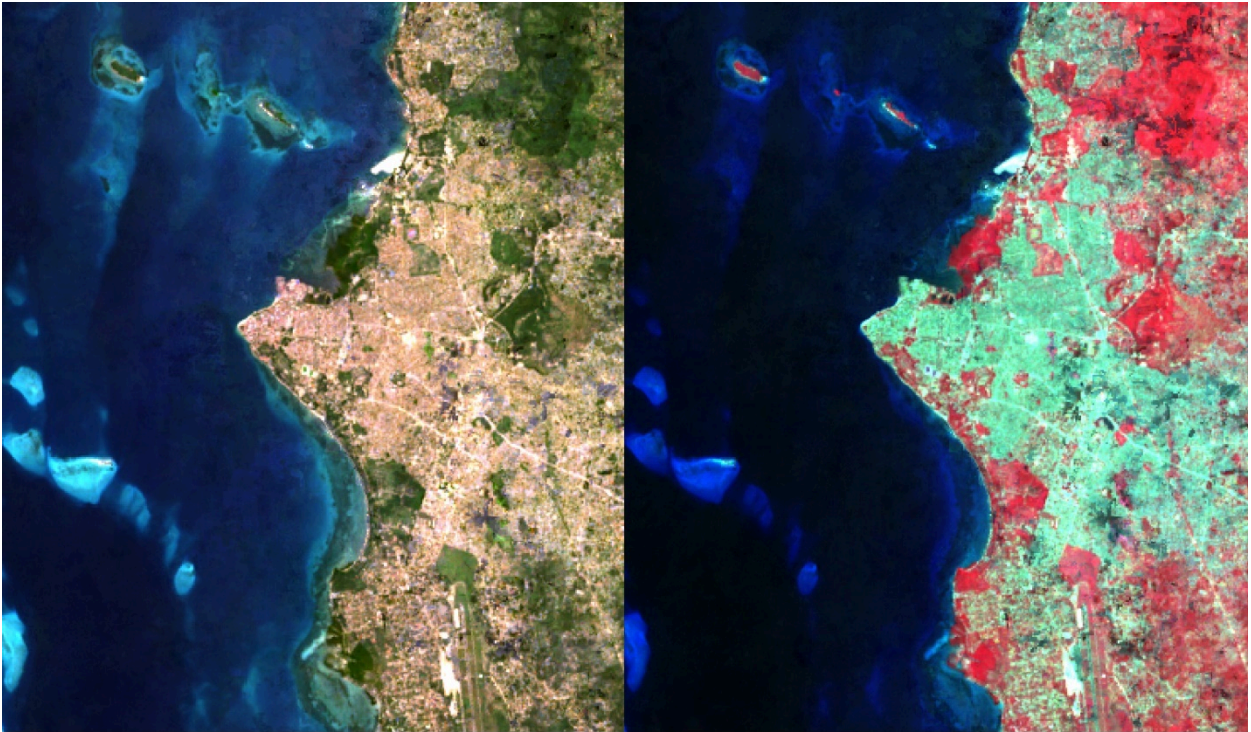


Figure 32: Affichages en couleurs vraies et en couleurs infrarouges d'une image Landsat contenant la ville de Zanzibar. Par Anders Knudby, CC BY 4.0.

Amélioration de l'image

Dans les exemples précédents, nous avons imaginé que les images contenaient toujours des valeurs comprises entre 0 et 255, fournissant ainsi les 256 différentes valeurs possibles que les canons à couleurs d'un écran d'ordinateur typique peuvent afficher. Il existe deux situations dans lesquelles l'utilisation directe de la valeur DN trouvée dans le fichier image pour déterminer l'intensité de la lumière produite par un pistolet à couleurs n'est pas la meilleure option :

- Tous les capteurs n'enregistrent pas les données sur 8 bits par pixel (par couleur). Pour les capteurs (anciens) qui utilisent moins de 8 bits, ou les capteurs (plus récents) qui utilisent plus de 8 bits, une conversion entre la valeur DN et l'intensité avec laquelle elle doit être affichée est généralement nécessaire.
- Si le capteur a produit une image sur une zone particulièrement claire ou sombre, ou du moins si une telle zone présente un intérêt spécifique, les valeurs trouvées sur cette zone (par exemple entre 50 et 80) peuvent être affichées en utilisant une gamme d'intensités plus large sur l'écran (par exemple entre 20 et 110), améliorant ainsi le contraste visuel sur la zone d'intérêt.

Pour mieux comprendre ces questions, nous devons examiner comment les valeurs DN sont stockées dans la mémoire de l'ordinateur. Nous allons considérer un seul pixel et sa valeur DN pour une seule bande. Pour un

capteur 8 bits, cette valeur DN est stockée sur 8 bits, chacun d'entre eux pouvant être ON (1) ou OFF (0). La valeur est codée comme illustré à la Figure 33. Dans cet exemple, les bits représentant les valeurs 128, 32, 16, 8 et 2 sont tous activés, tandis que les autres sont désactivés. La valeur DN du pixel est donc calculée comme suit : $128 + 32 + 16 + 8 + 2 = 186$. La gamme des valeurs possibles avec 8 bits va de 0 (tous les bits sont OFF) à 255 (tous les bits sont ON), soit un total de 256 valeurs possibles. En fait, le nombre de valeurs possibles, étant donné le nombre de bits, est facilement calculé comme 2^n , où n est le nombre de bits.

Value if ON	128	64	32	16	8	4	2	1
ON / OFF	1	0	0	1	1	0	1	0

Figure 33: Cet exemple illustre un pixel de 8 bits avec la valeur 186. Par Anders Knudby, CC BY 4.0.

C'est la raison pour laquelle les capteurs qui utilisent moins de 8 bits pour stocker le DN de chaque pixel ne peuvent pas produire 256 valeurs différentes. Toutefois, à titre d'exemple, si nous devons afficher directement les valeurs DN de 4 bits (valeurs DN comprises entre 0 et 15) sans changement, nous ne verrions que des nuances de gris très sombres (aucun capteur actuel n'enregistre d'informations en utilisant seulement 4 bits, mais les premiers capteurs Landsat n'utilisaient que 6 bits, et le problème est le même).

D'autre part, les capteurs qui utilisent 12 bits pour stocker chaque valeur DN (comme le font de nombreux capteurs modernes), contiennent des valeurs allant de 0 à $2^{12}-1 = 4095$. Pour afficher cette gamme de valeurs, nous devons les transformer pour qu'elles soient toutes comprises entre 0 et 255, ce qui est plus facile à faire en divisant simplement par 16 et en arrondissant vers le bas. Notez que dans ce cas, bien que l'écran de l'ordinateur ne soit pas capable d'afficher les données dans tous leurs détails, les algorithmes de traitement d'images pourront tout de même les utiliser.

Les choses deviennent plus intéressantes lorsque nous avons une image dans laquelle une gamme spécifique de valeurs est la plus intéressante. Un exemple courant est le type d'image de la figure 32, qui contient de l'eau (très sombre), de la terre (relativement sombre et largement verte) et des nuages (très clairs). À moins que nous ne soyons spécifiquement concernés par l'étude des nuages dans cette image, nous n'avons pas vraiment besoin de pouvoir distinguer les pixels dont les valeurs DN sont 245, 250 ou 255. Ce que nous pouvons faire, alors, c'est afficher tous les pixels « brillants » dont nous ne nous soucions pas vraiment avec la même couleur (blanc), et ainsi « réserver » plus de luminosités différentes avec lesquelles nous pouvons afficher les pixels pour les valeurs DN qui nous intéressent plus. Mathématiquement, cela signifie créer un tableau qui traduit chaque valeur DN, pour chaque bande, en une valeur de luminosité équivalente avec laquelle elle va être affichée à l'écran. Sur le plan graphique, nous pouvons représenter cela comme le fait la figure 34 ci-dessous, où les valeurs DN sont représentées sur l'axe des x , et la luminosité avec laquelle elles sont affichées est représentée sur l'axe des y . La ligne bleue représente l'utilisation directe de la valeur DN, sans amélioration du contraste. La ligne orange représente une amélioration linéaire du contraste, dans laquelle toutes les valeurs DN inférieures à 50 sont affichées avec une luminosité de 0, et toutes les valeurs DN supérieures à 240 sont affichées avec une luminosité de 255. Cela nous permet d'utiliser les 254 valeurs de luminosité restantes pour mieux afficher les

petites différences entre les pixels dont les valeurs DN sont comprises entre 50 et 240. La ligne grise est une transformation de racine carrée qui conserve un certain contraste sur toute la gamme des valeurs DN, mais qui se concentre sur l'affichage des différences entre les pixels ayant des valeurs DN faibles, au détriment de ceux ayant des valeurs DN élevées. Les figures 35 à 37 présentent des exemples de ce que peut donner cette amélioration du contraste dans la pratique :

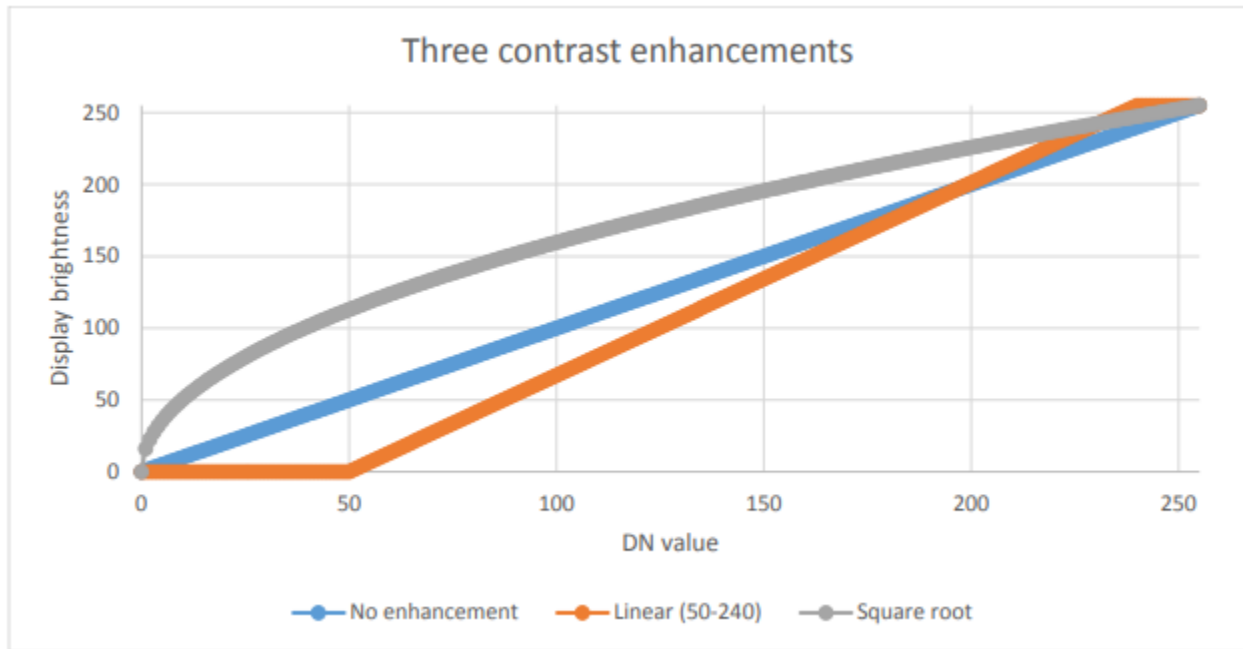


Figure 34: Trois options pour l'affichage des données 8 bits. Par Anders Knudby, CC BY 4.0.



Figure 35: Pas d'amélioration du contraste. Par Anders Knudby, CC BY 4.0.



Figure 36: Renforcement linéaire du contraste en utilisant les valeurs minimales et maximales de l'image à la place de 50 et 240 dans l'exemple ci-dessus. Par Anders Knudby, CC BY 4.0.

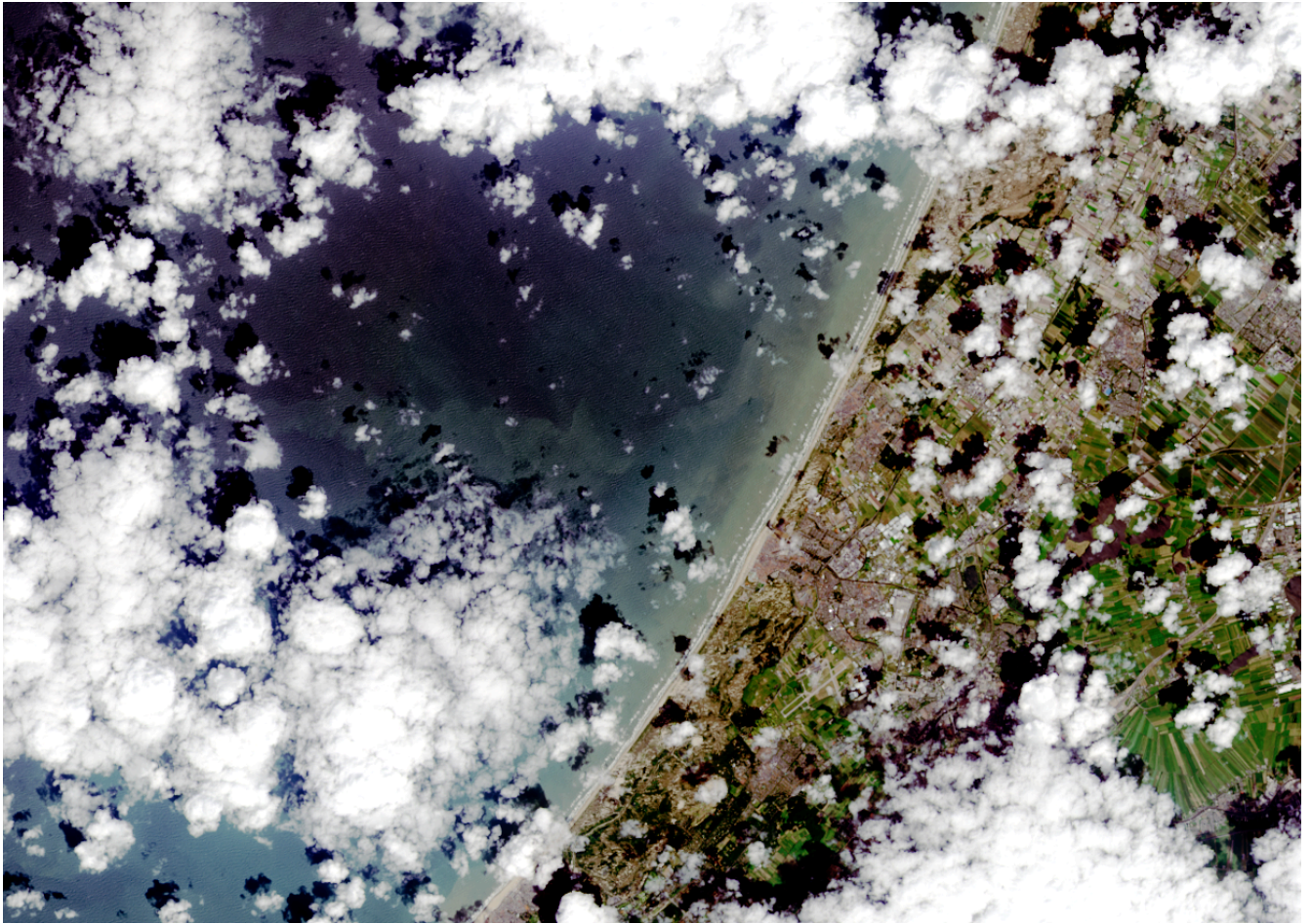


Figure 37: Amélioration du contraste, connue sous le nom d'égalisation d'histogramme. Cette technique est mathématiquement plus complexe que les exemples présentés ci-dessus, mais donne souvent de bons résultats. Par Anders Knudby, CC BY 4.0.

« Pan-sharpening »

Lorsque vous regardez des films hollywoodiens et que l'analyste de la CIA ou le policier étonnamment beau clique sur un bouton pour « améliorer » les images d'une caméra de surveillance, révélant ainsi les traits du visage de ce qui n'était auparavant qu'une tache floue, eh bien, c'est Hollywood. Le problème avec les images, c'est qu'elles ne sont que des données et qu'il est impossible de faire apparaître des détails (de nouvelles données) là où il n'y en a pas dans les données d'origines. Vous trouverez une explication amusante de cette idée générale ici. Cependant, avec certaines images satellites, il existe un type d'amélioration qui fonctionne (bien qu'il provoque des dégâts subtils sur la radiométrie de l'image). La technique en question est appelée « pan-sharpening », ce qui fait référence au fait que la plupart des capteurs optiques passifs des satellites ont une bande « pan- chromatique » qui enregistre des données sur la luminosité dans toutes les bandes visibles et à une résolution spatiale plus élevée que les bandes de couleur individuelles. Par exemple, les capteurs des satellites Landsat 4-7 ont des bandes bleues, vertes et rouges, tous avec une résolution spatiale de 30 mètres, et une bande pan- chromatique qui enregistre la lumière bleue et verte ainsi que la lumière rouge, avec une

résolution spatiale de 15 mètres. Pour les capteurs modernes à haute résolution, la différence de résolution spatiale est encore plus grande, WorldView-2 ayant des bandes de couleur à une résolution de 2 mètres et une bande panchromatique à une résolution de 50 cm. En gros, nous pouvons donc combiner les informations de couleur des bandes de couleur individuelles avec les détails spatiaux de la bande panchromatique, pour obtenir quelque chose comme ce que vous voyez à la figure 38.

Plusieurs algorithmes peuvent être utilisés pour effectuer cette opération, et chacun d'eux utilise un ensemble différent d'hypothèses et de règles pour générer des informations de couleur pour chacun des pixels de l'image panchromatique. Bien que le contexte détermine quel algorithme est supérieur pour une image donnée, les résultats de la plupart des algorithmes sont impressionnants et peuvent grandement aider à l'interprétation visuelle des images satellites. La plupart des logiciels commerciaux de traitement d'images disposent d'un ou plusieurs algorithmes de pansharpening. Ainsi, bien que le processus en lui-même soit assez complexe, son application à une image se fait en un clic.



Figure 38: Illustration du pan-sharpening avec une image QuickBird. L'image en couleurs (à gauche, résolution de 2,4 m) est combinée avec l'image pan-chromatique (au centre, résolution de 60 cm) pour produire l'image pan-sharpened (à droite). Two-Path Network with Feedback Connections for Pan-Sharpning in Remote Sensing (Figure 7) par Fu et al, CC BY 4.0.

Rapports de bande et autres mathématiques de bande

Les sections précédentes ont principalement traité des différentes manières d'afficher les valeurs DN trouvées dans chaque bande — en niveaux de gris ou en couleur, directement ou transformées par l'amélioration du contraste, avec ou sans pansharpening. Une autre solution consiste à créer de nouvelles bandes à l'aide d'un processus communément appelé « Band Math ».

Par exemple, on peut vouloir produire une bande qui quantifie le degré de « rouge » d'un objet. Le problème de la bande « rouge » originale d'une image satellite typique est visible en comparant les valeurs DN « rouges » de la partie supérieure gauche de la figure 31 avec les couleurs indiquées dans la partie inférieure droite de la même figure. On peut dire que le pixel qui se trouve dans le coin supérieur gauche est ce que la plupart des gens considèrent comme « rouge » mais les deux pixels ayant les valeurs DN les plus élevées dans la bande « rouge » sont les pixels jaune et blanc (tous deux ont des valeurs DN « rouges » de 255). Malgré

leur valeur DN « rouge » élevée, ces pixels n'apparaissent pas « rouges », car ils ont *également* des valeurs DN élevées dans une ou plusieurs autres couleurs. Pour mieux quantifier l'apparence « rouge » de chaque pixel, nous pouvons calculer ce que l'on appelle la « coordonnée chromatique rouge » (CCR). Pour chaque pixel, cette coordonnée est calculée comme suit : $RCC = \text{Rouge} / (\text{Rouge} + \text{Vert} + \text{Bleu})$, où Rouge, Vert et Bleu représentent les valeurs DN respectives de chaque pixel. Comme vous pouvez le voir dans les calculs de la figure 39, le pixel rouge dans le coin supérieur gauche a une valeur RCC plus élevée que tous les autres pixels, suivi par le pixel jaune à côté (parce que le jaune est « proche » du rouge). Notez que les mathématiques de bande opèrent sur chaque pixel individuellement, et effectuent la même opération mathématique sur tous les pixels d'une image. Ce faisant, elle produit une nouvelle bande qui est une fonction des bandes d'origine. D'autres utilisations courantes de la mathématique des bandes consistent à calculer des « rapports de bande », comme l'indice de végétation appelé « rapport simple », calculé comme suit : $SR = \text{NIR} / \text{Rouge}$, où NIR est une valeur (idéalement la réflectance TOA) dans la bande du proche infrarouge, et Rouge est la même valeur dans la bande rouge. Étant donné que les plantes saines réfléchissent très efficacement la lumière proche infrarouge, alors qu'elles absorbent la majorité de la lumière rouge (et reflètent donc très peu de lumière rouge à détecter par le capteur), les zones avec végétation ont tendance à avoir des valeurs de réflectance et de DN élevées dans la bande proche infrarouge et des valeurs faibles dans la bande rouge. Les zones avec végétation ont donc des valeurs de rapport simple très élevées, et ce rapport de bande peut donc être utilisé pour produire rapidement et facilement une visualisation qui aide les gens à distinguer les zones avec et sans végétation. Un indice de végétation encore plus populaire est l'indice de végétation par différence normalisée (NDVI), qui est calculé comme suit : $NDVI = (\text{NIR} - \text{Rouge}) / (\text{NIR} + \text{Rouge})$

Red:				Blue:			
237	255	128	128	36	0	0	0
0	34	34	128	0	76	76	0
63	63	34	128	204	204	76	0
0	195	255	0	0	195	255	0

Green:				RCC:			
28	242	255	255	0.8	0.5	0.3	0.3
0	177	177	255	NA	0.1	0.1	0.3
72	72	177	255	0.2	0.2	0.1	0.3
0	195	255	0	NA	0.3	0.3	NA

Figure 39: Les valeurs de rouge, vert et bleu sont les mêmes que dans la figure 31. La coordonnée chromatique rouge (CCR) est

calculée comme suit : $CCR = \text{Rouge} / (\text{Rouge} + \text{Vert} + \text{Bleu})$. Les pixels ayant des valeurs DN de 0 dans toutes les bandes (c'est-à-dire les pixels noirs) ont reçu des valeurs NoData (NA). Par Anders Knudby, CC BY 4.0.

Les mathématiques des bandes peuvent être utilisées pour calculer de nouvelles bandes avec n'importe quelle combinaison de bandes d'entrées à laquelle vous pouvez penser. Différents types de mathématiques des bandes sont utilisées, parmi de nombreuses autres applications, pour surveiller la déforestation mondiale, suivre les changements saisonniers de la couverture neigeuse, cartographier la profondeur des eaux et surveiller l'expansion urbaine.

« Tasseled cap transformation (TCT) »

Une application particulière de la mathématique des bandes, appelée « tasseled cap transformation », a été développée pour les données Landsat et est maintenant couramment utilisée pour produire un ensemble de bandes normalisé et facilement interprétable à partir des données d'origine. Elle a été développée à l'origine pour un instrument appelé Landsat MSS (qui ne fonctionne plus), et a depuis été mise à jour pour Landsat TM (les capteurs de Landsat 4 et 5), ETM+ (le capteur de Landsat 7), et OLI (le capteur optique passif de Landsat 8). L'idée qui sous-tend la transformation de la calotte est de produire trois nouvelles bandes, « verdure », « luminosité » et « humidité », qui décrivent les variations réelles de ces paramètres sur l'image. Dans une image satellite typique d'un paysage rural, les zones recouvertes de sols de couleur variable formeront une « ligne de sols » dans l'espace bidimensionnel (NIR/ROUGE), tandis que les zones de végétation se trouveront à l'écart de la ligne et les différents stades de croissance des plantes formeront ensemble une forme ressemblant grossièrement à un chapeau à tasse (figure 40). Les trois nouvelles bandes sont créées par une rotation effective des axes du système de coordonnées, de sorte que la « luminosité » est déterminée le long de la « ligne des sols » de la figure 40, la « verdure » est déterminée le long d'un axe qui s'éloigne de la ligne des sols et dont la valeur augmente à mesure qu'il se déplace vers le sommet du « chapeau à glands », et l'« humidité » est déterminée le long d'un troisième axe (non représenté sur la figure 40) perpendiculaire aux deux premiers. Un exemple des résultats, basé sur une image IKONOS de l'Idaho, aux États-Unis, est présenté à la figure 41. Concrètement, chaque composante du chapeau à glands est calculée comme une somme pondérée des valeurs des bandes d'origine. Par exemple, pour les données Landsat TM, la « luminosité » est calculée comme suit : $BR = 0,3037 * BLEU + 0,2793 * VERT + 0,4343 * ROUGE + 0,5585 * NIR + 0,5082 * SWIR1 + 0,1863 * SWIR2$ (SWIR1 et SWIR2 sont deux bandes infrarouges à ondes courtes également présentes sur les capteurs Landsat TM).

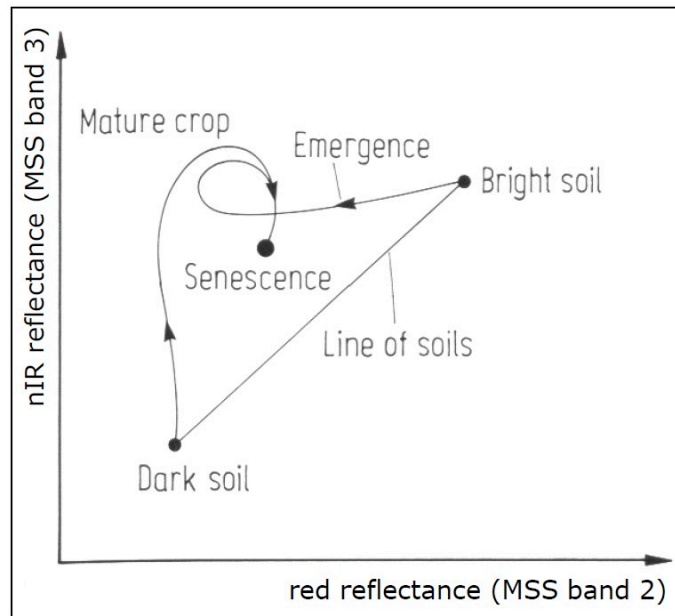


Figure 40: Le principe de la « tasseled cap transformation ». Des sols nus de luminosité variable tombent le long d'une ligne lorsque leurs valeurs sont tracées dans les bandes ROUGE et NIR, tandis que la végétation de densité croissante se trouve à l'intérieur du « chapeau », en haut et à gauche de la ligne des sols. La végétation à différents stades de croissance, poussant sur des sols de luminosité différente, forme ainsi une forme qui ressemble à un chapeau à glands. La formulation originale du transformation incluait une troisième dimension, en utilisant la bande verte. Tasseled Cap Transformation for bright and dark soil par Gurkengräber (adapted from Richards and Jia), Wikimedia Commons, CC BY-SA 4.0.

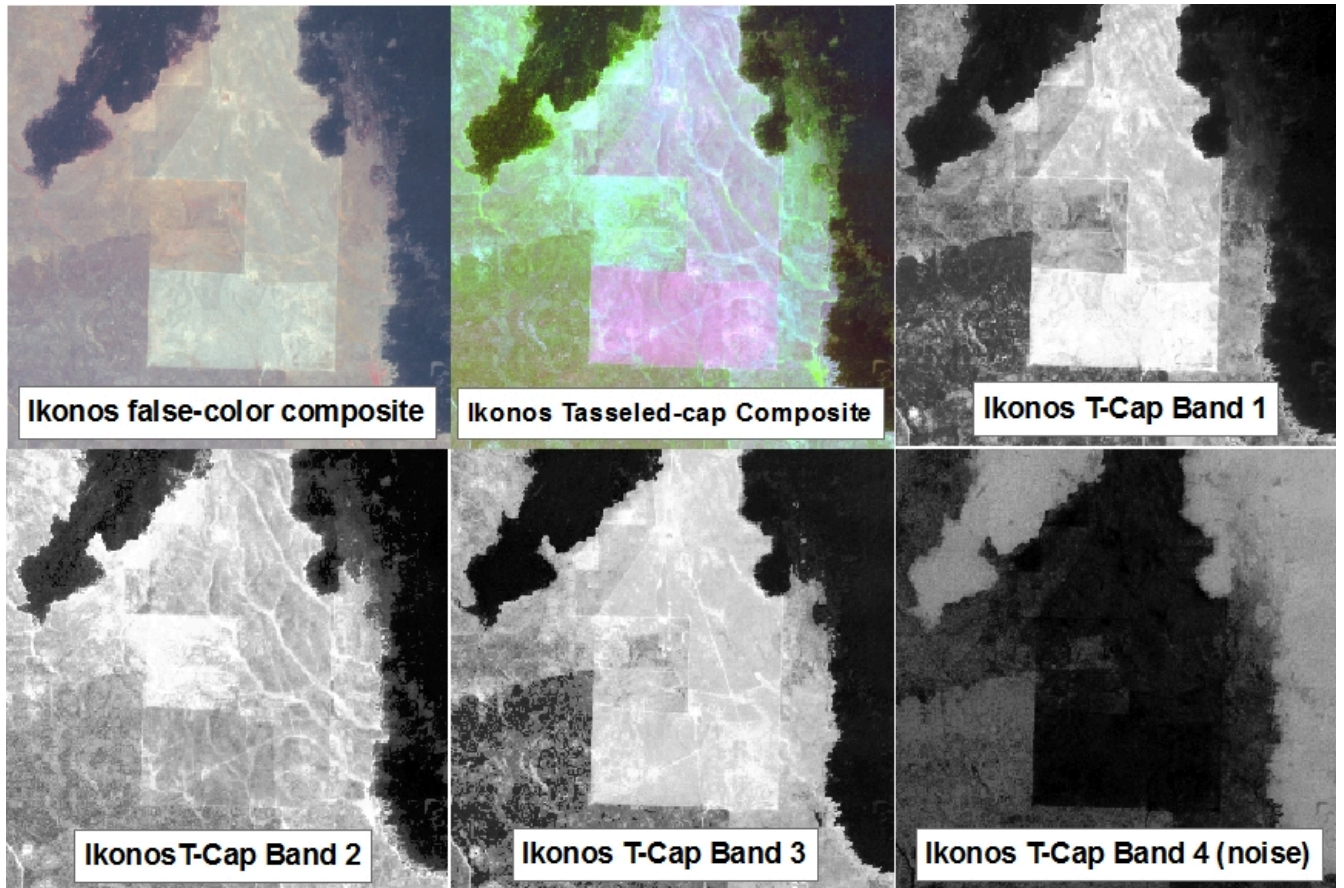


Figure 41: Un exemple de la « tasseled cap transformation » appliquée à une image IKONOS de l'Idaho, USA. tcap_example par Jason Karl, The Landscape Toolbox, CC0 1.0.

Analyse en composantes principales (ACP)

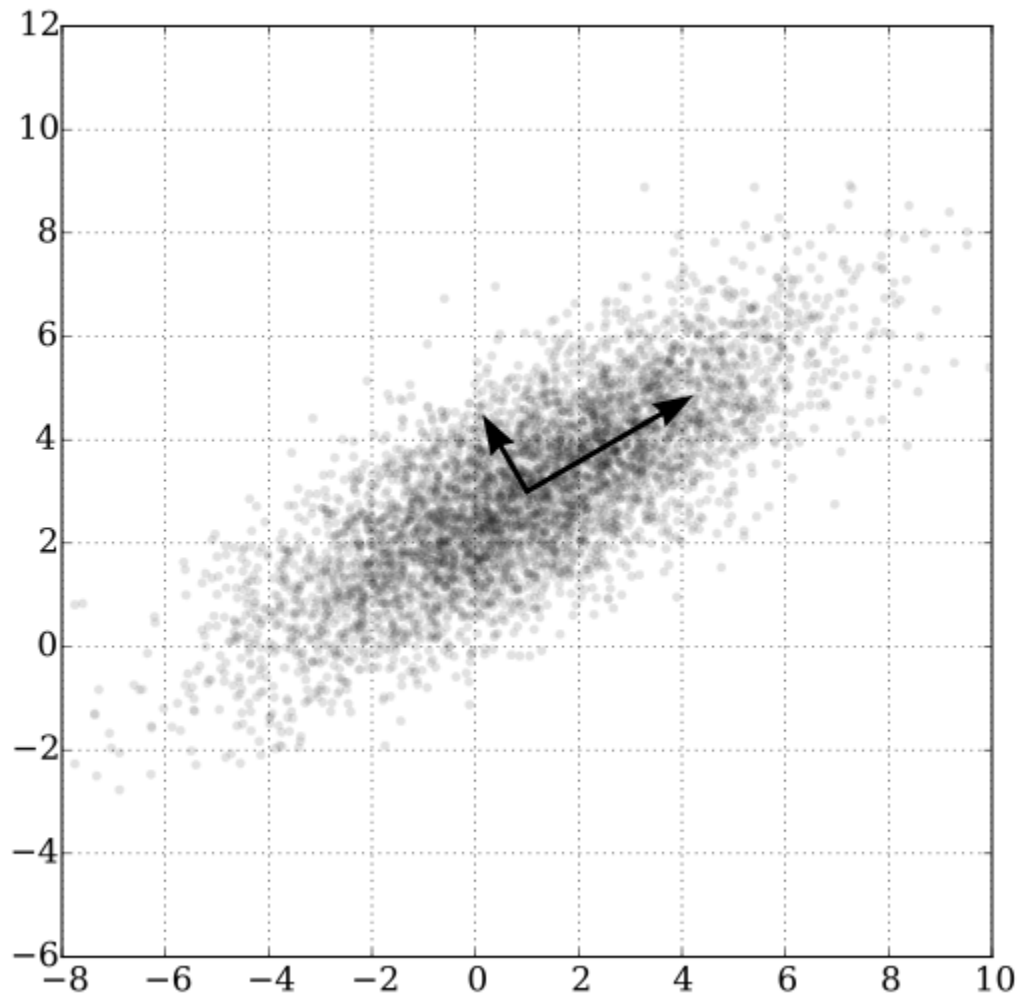


Figure 42: Un exemple bidimensionnel d'analyse en composantes principales. GaussianScatterPCA par Nicoguardo, Wikimedia Commons, CC BY 4.0.

Le dernier type de mathématique des bandes que nous allons examiner dans ce chapitre est semblable à la « tasseled cap transformation » dans la mesure où le calcul implique une somme pondérée des bandes originales, mais il est différent dans le sens où les pondérations sont déterminées à partir des propriétés de l'image elle-même — même, et en tant que telles, elles varient d'une image à l'autre. La figure 42 illustre bien le fonctionnement de cette méthode. Nous pouvons imaginer que chaque point du nuage de points est un pixel de notre image, et qu'il est représenté en fonction de sa valeur dans une bande sur l'axe des x et une autre bande sur l'axe des y. Il est clair, d'après la distribution des points, qu'il s'agit d'un pixel de l'image. Il ressort clairement

de la distribution des points de la figure 42 que les points ayant des valeurs élevées dans une bande ont tendance à avoir également des valeurs élevées dans l'autre bande, et que la majorité de la distribution des points suit une ligne imaginaire allant du bas à gauche au haut à droite du graphique. L'orientation de cette ligne imaginaire est appelée la première composante principale (PC_1) et est illustrée par la flèche la plus longue. Mais il existe également une variation orthogonale (perpendiculaire) à la PC_1 . C'est l'orientation de la deuxième composante principale, PC_2 , illustrée par la flèche la plus courte. Selon le nombre de bandes dans une image, on pourrait continuer à trouver de nouvelles lignes de ce type dans les troisièmes, quatrièmes, etc. dimensions, chacune étant orientée selon la plus grande distribution de points dans toute orientation perpendiculaire à toutes les composantes principales précédentes. Dans la figure 42, l'ensemble de données original n'avait que deux dimensions, et deux composantes principales ont donc été créées. La longueur de chaque flèche correspond à la quantité de distribution totale capturée par chaque composante principale, et il est clair que la PC_2 a capturé assez peu de la distribution des points originaux. Cela peut être interprété comme le fait que la PC_2 contient très peu d'informations, et souvent les composantes principales de rang supérieur peuvent être entièrement supprimées avec une faible perte d'informations. Avec cette connaissance, de nouvelles bandes peuvent être créées, dans lesquelles la valeur de chaque pixel n'est plus sa valeur DN dans les bandes originales, mais plutôt sa valeur le long de chacun des axes des composantes principales. C'est là le principal atout de l'ACP : elle permet de transformer les données en un plus petit nombre de bandes (occupant ainsi moins d'espace sur le disque dur, et permettant une lecture et un traitement plus rapides) sans perdre beaucoup d'informations. Des recherches ont même montré que pour certains types d'images (par exemple, une image Landsat sans nuages d'une zone agricole), des composantes principales spécifiques tendent à correspondre à des variables environnementales spécifiques, telles que l'humidité du sol, la couverture végétale, etc. Cependant, une mise en garde s'impose ici. Comme l'orientation des composantes principales (les flèches de la figure 42) dépend de la distribution de la totalité des valeurs des pixels dans chaque image individuelle, les valeurs des composantes principales d'un pixel ne sont pas directement interprétables et peuvent varier considérablement entre les images, même pour des pixels identiques. L'ACP peut donc vous aider à réduire le volume de données d'une image, mais elle détruit (ou du moins, complique) l'interprétation dans le processus. Une autre très bonne explication de l'ACP, y compris un nuage de points animé qui, selon moi, explique très bien les choses, se trouve ici.

6.

CLASSIFICATION

Si l'interprétation visuelle d'une image satellite peut suffire à certains usages, elle présente également des lacunes importantes. Par exemple, si vous voulez savoir où se trouvent toutes les zones urbaines d'une image, a) il faudra beaucoup de temps, même à un analyste d'images expérimenté, pour examiner une image entière et déterminer quels pixels sont « urbains » et lesquels ne le sont pas, et b) la carte des zones urbaines qui en résulte sera nécessairement quelque peu subjective, car elle est basée sur l'interprétation par l'analyste individuel de ce que signifie « urbain » et de la façon dont cela est susceptible d'apparaître sur l'image. Une solution très courante consiste à utiliser un algorithme de classification pour traduire la couleur observée dans chaque pixel en une classe thématique qui décrit sa couverture terrestre dominante, transformant ainsi l'image en une carte de couverture terrestre. Ce processus est appelé classification d'images.

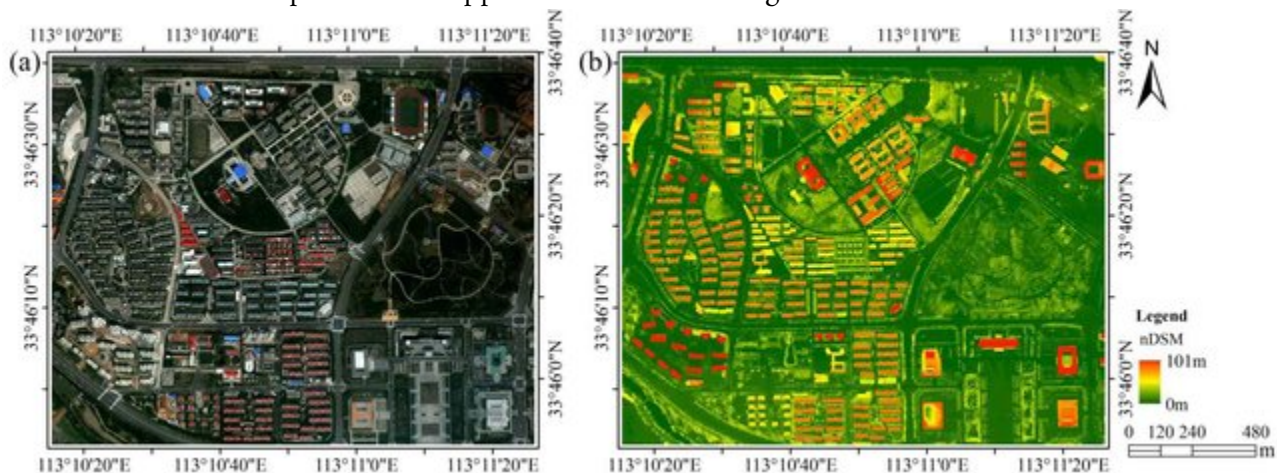


Figure 43: La classification des images traduit les couleurs de l'image originale (à gauche) en classes thématiques, généralement définies comme la couverture terrestre dominante ou l'utilisation des sols de chaque zone. WorldView-2 and LiDAR nDSM image of the study area (Figure 1) par Wu et al., CC BY 3.0.

On peut distinguer deux catégories d'approches de la classification des images. La méthode traditionnelle et la plus simple consistent à examiner chaque pixel individuellement et à déterminer la classe thématique correspondant à sa couleur. C'est ce qu'on appelle la classification par pixel, et c'est ce que nous allons voir en premier. Une méthode plus récente et de plus en plus populaire consiste à diviser l'image en segments homogènes, puis à déterminer la classe thématique correspondant aux attributs de chaque segment. Ces attributs peuvent être la couleur du segment ainsi que d'autres éléments comme la forme, la taille, la texture et l'emplacement. C'est ce qu'on appelle généralement l'analyse d'image basée sur l'objet, et nous y reviendrons dans la seconde moitié de ce chapitre.

Même dans la catégorie de la classification par pixel, deux approches différentes sont disponibles. La première est appelée classification « supervisée », car l'analyste d'images « supervise » la classification en fournissant des informations supplémentaires dans les premières étapes. L'autre est appelée classification « non supervisée », car un algorithme effectue la majeure partie du travail (presque) sans aide, et l'analyste d'images n'a plus qu'à intervenir à la fin pour terminer le travail. Chacune de ces méthodes présente des avantages et des inconvénients, qui sont décrits ci-après.

Classification supervisée par pixel

L'idée derrière la classification supervisée est que l'analyste de l'image fournit à l'ordinateur certaines informations qui permettent de calibrer un algorithme de classification. Cet algorithme est ensuite appliqué à chaque pixel de l'image pour produire la carte requise. La meilleure façon d'expliquer ce fonctionnement est de prendre un exemple. L'image de la figure 44 provient de Californie, et nous voulons traduire cette image en une classification comportant les trois (grandes) classes suivantes : « Urbain », « Végétation » et « Eau ». Le nombre de classes et la définition de chaque classe peuvent avoir un impact important sur le succès de la classification. Dans notre exemple, nous ne tenons pas compte du fait qu'une grande partie de la zone est constituée de sol nu (et n'entre donc pas vraiment dans l'une de nos trois classes). Nous pouvons également nous rendre compte que l'eau dans l'image a une couleur très différente selon son degré de turbidité, et que certaines zones urbaines sont très claires alors que d'autres sont d'un gris plus foncé, mais nous ignorons ces questions pour l'instant.



Figure 44: Composite en couleurs vraies montrant une partie d'une image Landsat de Californie. Une zone urbaine avec de l'eau trouble est visible dans le coin supérieur droit, bordée par un mélange de zones de végétation et de zones stériles et un lac avec de l'eau claire. Une zone montagneuse largement boisée est visible dans la partie inférieure gauche de l'image, avec quelques zones claires près du centre gauche. Par Anders Knudby, CC BY 4.0, basé sur une image Landsat 5 (USGS).

La « supervision » dans la classification supervisée se présente presque toujours sous la forme d'un ensemble de données de calibrage, qui consiste en un ensemble de points et/ou de polygones dont on sait (ou croit) qu'ils appartiennent à chaque classe. Dans la figure 45, un tel ensemble de données a été fourni sous la forme de trois polygones. Le polygone rouge délimite une zone connue pour être « urbaine », de même le polygone bleu est « eau » et le polygone vert est « végétation ». Notez que l'exemple de la figure 45 *n'est pas* un exemple de bonne pratique — il est préférable d'avoir des polygones plus nombreux et plus petits pour chaque classe, répartis sur l'ensemble de l'image, car cela permet aux polygones de couvrir que les pixels de la classe voulue, et aussi d'intégrer les variations spatiales, par exemple la densité de la végétation, la qualité de l'eau, etc.

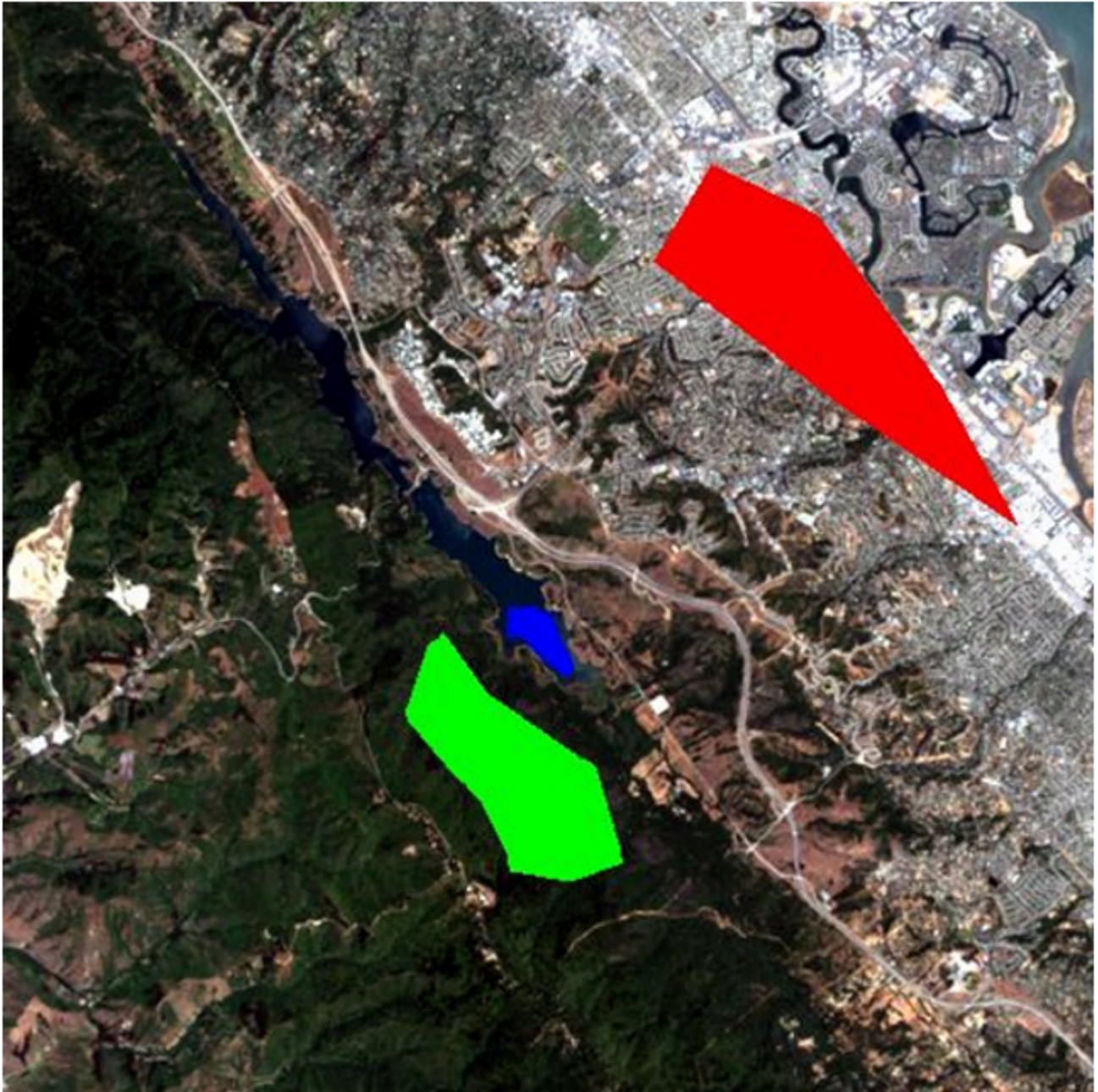


Figure 45: La même image que dans la Figure 44, avec trois polygones superposés. Le polygone rouge représente une zone connue par l'analyste de l'image comme étant « urbaine », le polygone bleu est « eau », et le polygone vert est « végétation ».

Anders Knudby, CC BY 4.0.

Regardons maintenant comment ces polygones nous aident à transformer l'image en une carte des trois classes. En gros, les polygones disent à l'ordinateur « regardez les pixels sous le polygone rouge – c'est à cela que ressemblent les pixels "urbains" », et l'ordinateur peut alors trouver tous les autres pixels de l'image qui ressemblent aussi à cela, et les étiqueter « urbains ». Et ainsi de suite pour les autres classes. Cependant, certains pixels peuvent ressembler un peu à « urbain » et un peu à « végétation », il faut donc trouver un moyen mathématique de déterminer à quelle classe chaque pixel ressemble le plus. Nous avons besoin d'un algorithme de classification.

Si nous prenons toutes les valeurs de tous les pixels des bandes 3 et 4 de Landsat et les montrons sur un nuage de points, nous obtenons quelque chose comme la figure 46. Cette image a une résolution radiométrique de 8 bits, donc les valeurs de chaque bande vont théoriquement de 0 à 255, nous observons que les plus petites valeurs de l'image sont supérieures à 0. Les valeurs de la bande 3 sont représentées sur l'axe des x, et celles de la bande 4 sur l'axe des y.

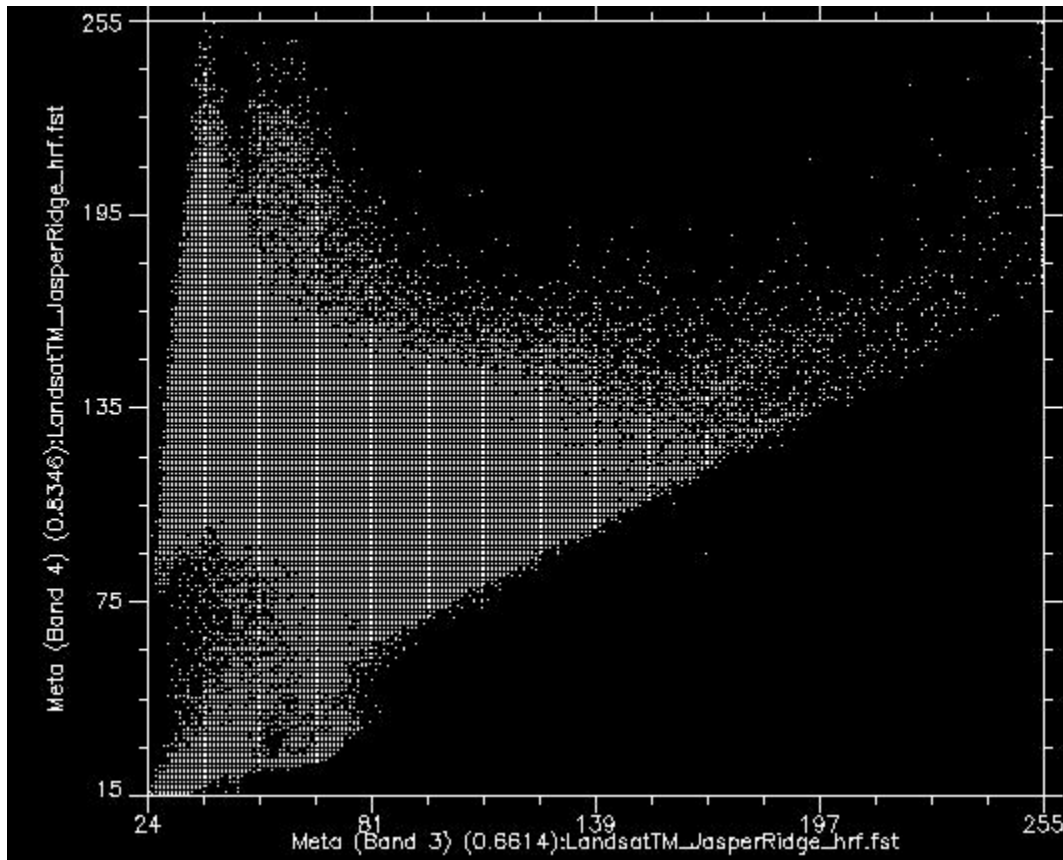


Figure 46: Nuage de points montrant toutes les valeurs de pixels dans les bandes 3 et 4 pour l'image de la Figure 44. Nuage de points créé à l'aide du logiciel ENVI. Par Anders Knudby, CC BY 4.0.

Maintenant, si nous colorions tous les points qui proviennent des pixels sous le polygone rouge (c'est-à-dire les pixels que nous « savons » être « Urbains »), et que nous faisons de même avec les pixels sous les polygones bleus et verts, nous obtenons quelque chose comme la figure 47. Il y a quelques éléments importants à noter dans la figure 47. Tous les points bleus (« Eau ») se trouvent dans le coin inférieur gauche de la figure, sous le cercle jaune, avec de faibles valeurs dans la bande 3 et de faibles valeurs dans la bande 4. C'est en effet typique de l'eau, car l'eau absorbe très efficacement le rayonnement entrant dans les longueurs d'onde du rouge (bande 3) et du proche infrarouge (bande 4), de sorte que très peu sont réfléchis pour être détectés par le capteur. Les points verts (« végétation ») forment une longue zone sur le côté gauche de la figure, avec des valeurs faibles dans la bande 3 et des valeurs modérées à élever dans la bande 4. Là encore, cela semble raisonnable, car la végétation absorbe efficacement le rayonnement entrant dans la bande rouge (en l'utilisant pour la photosynthèse) tout en réfléchissant le rayonnement entrant dans la bande du proche infrarouge. Les points

rouges (« Urbain ») forment une zone plus large près du centre de la figure, et couvrent une gamme de valeurs beaucoup plus large que les deux autres classes. Si leurs valeurs sont similaires à celles de la « Végétation » dans la bande 4, elles sont généralement plus élevées dans la bande 3.

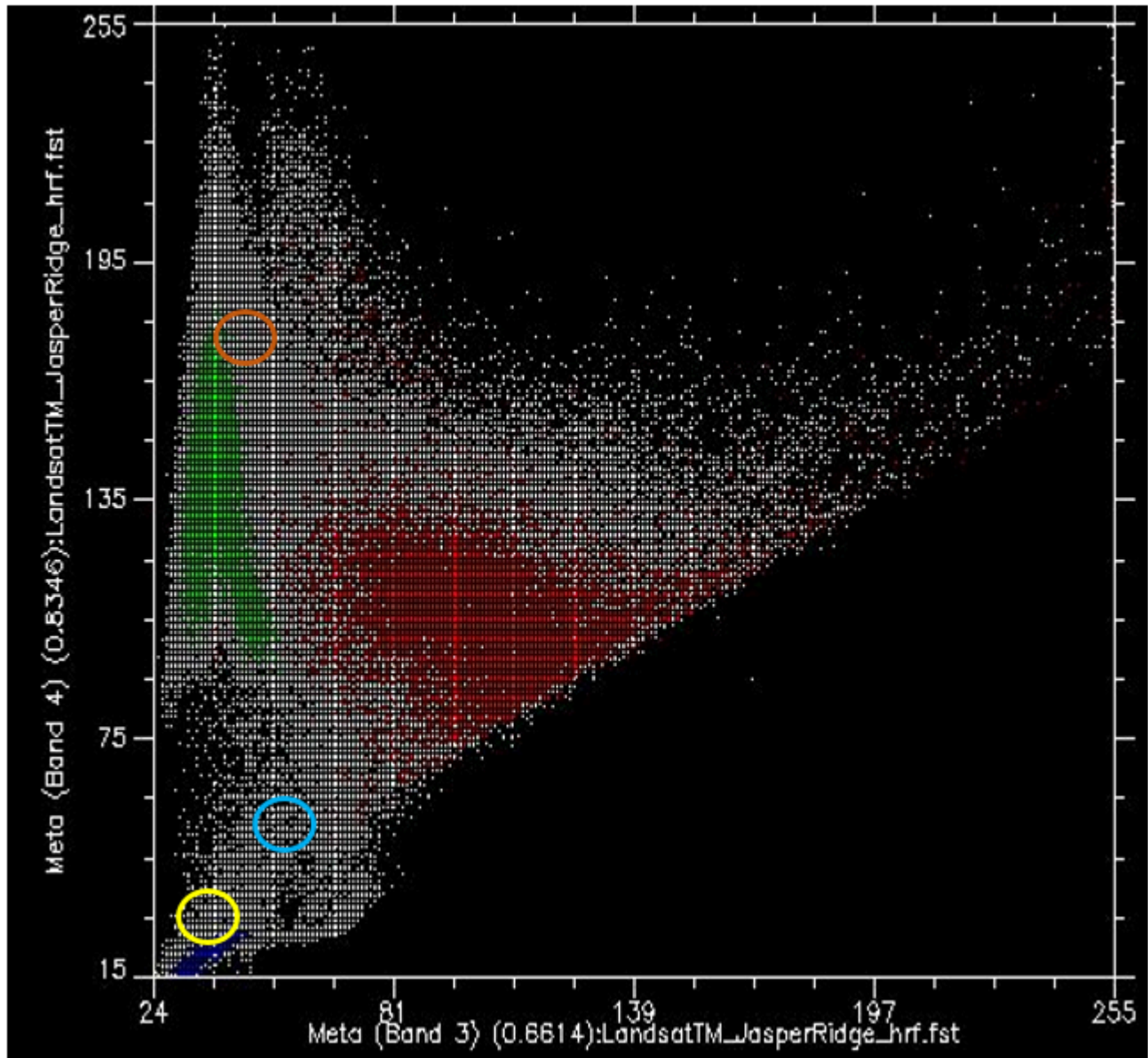


Figure 47: Comme la Figure 46, mais avec chaque point coloré par le polygone sous lequel il se trouve. Nuage de points créé à l'aide du logiciel ENVI. Par Anders Knudby, CC BY 4.0.

Ce que nous voulons que l'algorithme de classification supervisée fasse maintenant, c'est prendre tous les autres pixels de l'image (c'est-à-dire tous les points blancs sur le nuage de points) et les affecter à l'une des trois classes en fonction de leur couleur. Par exemple, à quelle classe pensez-vous que les points blancs dans le cercle jaune de la figure 47 devraient être affectés ? L'eau, probablement. Et qu'en est-il de ceux qui se trouvent dans le cercle marron clair ? Végétation, probablement. Mais qu'en est-il de ceux dans le cercle bleu clair ? Pas si facile à déterminer.

Remarque : l'algorithme de classification peut utiliser toutes les bandes de l'image Landsat, ainsi que

toute autre information que nous fournissons pour l'ensemble de l'image (comme un modèle numérique d'altitude), mais comme il est plus facile de continuer à représenter cette image en deux dimensions en utilisant uniquement les bandes 3 et 4, nous allons continuer à le faire. Gardez à l'esprit que le nuage de points est en réalité un graphique à n dimensions, où n est égal au nombre de bandes (et autres couches de données) que nous voulons utiliser dans la classification.

Classificateur à distance minimale

Une façon d'estimer à quelle classe appartient chaque pixel est de calculer la « distance » entre le pixel et le centre de tous les pixels connus pour appartenir à chaque classe, puis de l'affecter à la classe la plus proche. Par « distance », nous entendons ici la distance dans « l'espace caractéristique », dans lequel les dimensions sont définies par chacune des variables que nous considérons (dans notre cas, les bandes 3 et 4), par opposition à la distance physique. Notre espace caractéristique est donc bidimensionnel, et les distances peuvent être calculées à l'aide d'une distance euclidienne standard.

Par exemple, pour les points de la figure 48, nous avons calculé les valeurs moyennes de tous les pixels verts, rouges et bleus pour les bandes 3 et 4, et que nous avons indiquées par de gros points. Disons qu'ils ont les valeurs suivantes :

Tableau 3: Les valeurs moyennes dans les bandes 3 et 4 pour les classes « Urbain », « Végétation » et « Eau » sont présentées dans la Figure 49.

Valeurs moyennes	Points rouges ("Urbain")	Points verts ("Végétation")	Points Bleus ("Eau")
Bande 3	100	40	35
Bande 4	105	135	20

Disons ensuite qu'un pixel indiqué par le point jaune de la figure 48 a une valeur de 55 dans la bande 3, et de 61 dans la bande 4. Nous pouvons alors calculer la distance euclidienne entre ce point et la valeur moyenne de chaque classe :

$$\text{Distance à la moyenne rouge : } (100-55)^2 + (105-61)^2 = 62.9$$

$$\text{Distance à la moyenne verte : } (40-55)^2 + (135-61)^2 = 75.5$$

$$\text{Distance à la moyenne bleue : } (35-55)^2 + (20-61)^2 = 45.6$$

minimale affectera ce point particulier à la classe « bleu ». Bien que le classificateur de distance minimale soit très simple et rapide et qu'il donne souvent de bons résultats, cet exemple illustre une faiblesse importante : dans notre exemple, la distribution des valeurs pour la classe « Eau » est très faible — l'eau est en principe toujours foncé et bleu vert, et même l'eau trouble ou l'eau contenant beaucoup d'algues a toujours l'air foncée et bleue — verte. La distribution des valeurs pour la classe « Végétation » est beaucoup plus importante, surtout

dans la bande 4, parce que certaines végétations sont denses et d'autres ne le sont pas, certaines végétations sont saines et d'autres ne le sont pas, certaines végétations peuvent être mélangées avec des sols sombres, des sols clairs, ou même des éléments urbains comme une route. Il en va de même pour la classe « urbaine », qui présente une large distribution de valeurs dans les bandes 3 et 4. En réalité, le point jaune de la figure 48 n'est probablement pas de l'eau, car l'eau qui a des valeurs aussi élevées dans les bandes 3 et 4 n'existe pas. Il est beaucoup plus probable qu'il s'agisse d'un type de végétation inhabituel, d'une zone urbaine inhabituelle ou (plus probablement encore) d'un mélange de ces deux classes. Le prochain classificateur que nous allons examiner prend explicitement en compte la distribution des valeurs dans chaque classe, pour remédier à ce problème.

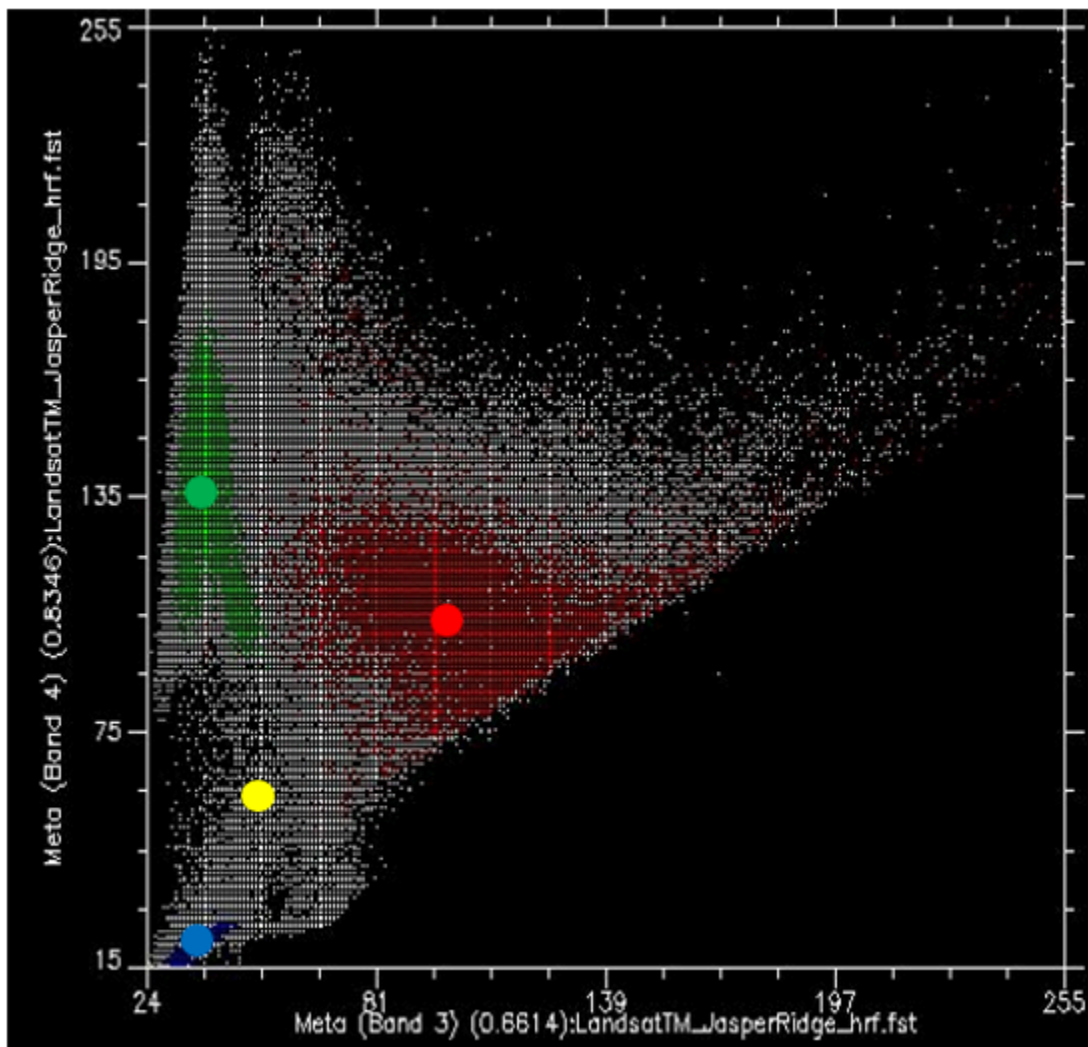


Figure 48: Le classificateur de distance minimale attribue la classe dont le centre est le plus proche (dans l'espace des caractéristiques) de chaque pixel. La valeur moyenne de tous les points rouges, dans les bandes 3 et 4, est indiquée par le gros point rouge, de même pour les points verts et bleus. Le point jaune indique un pixel que nous souhaitons attribuer à l'une des trois classes.

Diagramme de dispersion créé à l'aide du logiciel ENVI. Par Anders Knudby, CC BY 4.0.

Classificateur à maximum de vraisemblance

Jusqu'à il y a une dizaine d'années, le classificateur de vraisemblance maximale était l'algorithme de référence pour la classification d'images. Il est toujours populaire, implémenté dans tous les logiciels de télédétection sérieux et figure généralement parmi les algorithmes les plus performants pour une tâche donnée. Les descriptions mathématiques de son fonctionnement peuvent sembler compliquées, car elles reposent sur des statistiques bayésiennes appliquées à plusieurs dimensions, mais le principe est relativement simple : au lieu de calculer la distance au centre de chaque classe (dans l'espace des caractéristiques) et donc de trouver la classe la plus proche, nous allons calculer la probabilité que le pixel appartienne à chaque classe, et donc trouver la classe la plus probable. Pour que ce calcul fonctionne, il faut faire quelques hypothèses.

- Nous supposons qu'avant de connaître la couleur du pixel, la probabilité qu'il appartienne à une classe est la même que la probabilité qu'il appartienne à toute autre classe. Cela semble assez raisonnable (bien que dans notre image il y ait clairement beaucoup plus de « végétation » que d' « eau », on pourrait donc dire qu'un pixel de couleur inconnue a plus de chances d'être de la végétation que de l'eau... cela peut être incorporé dans le classificateur, mais c'est très rare, et nous l'ignorerons pour l'instant).
- Nous supposons que la distribution des valeurs dans chaque bande et pour chaque classe est gaussienne, c'est-à-dire qu'elle suit une distribution normale (une courbe en cloche).

Pour commencer par un exemple unidimensionnel, notre situation pourrait ressembler à ceci si nous n'avions que deux classes :

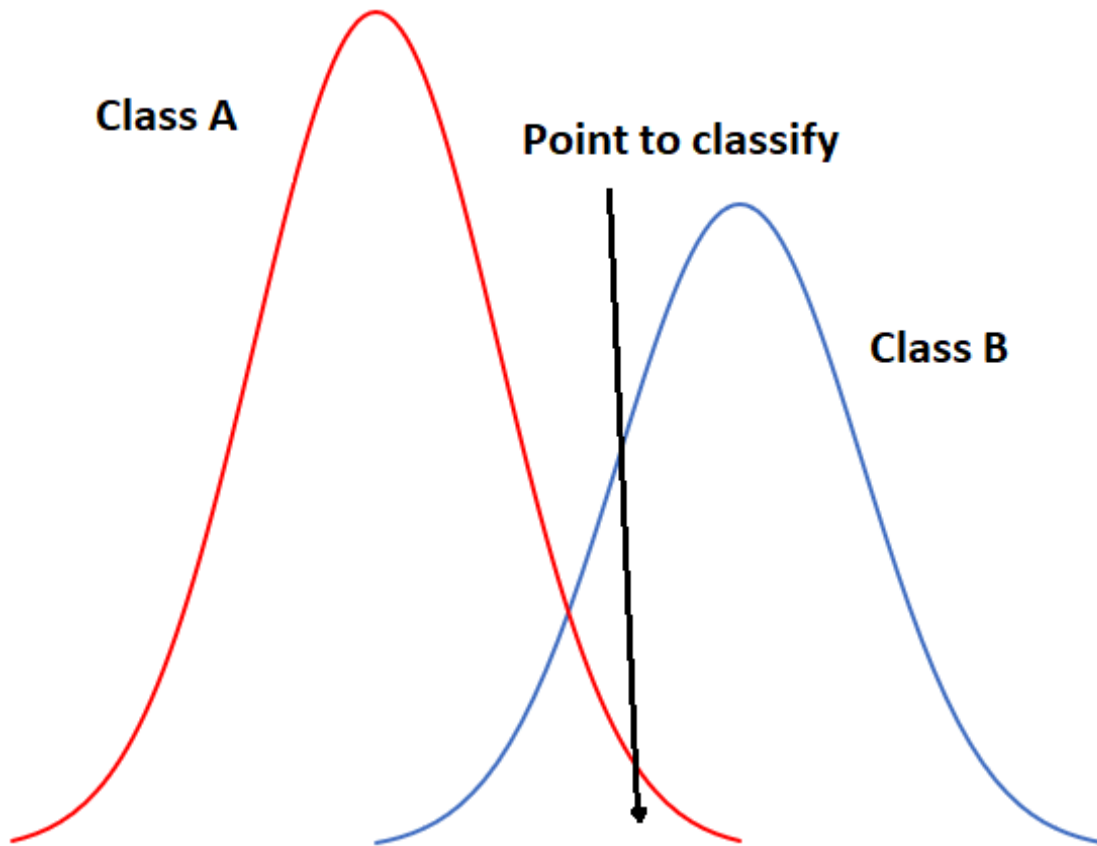


Figure 49: Exemple unidimensionnel de classification par maximum de vraisemblance avec deux classes. Par Anders Knudby, CC BY 4.0.

Dans la figure 49, l'axe des x représente les valeurs dans une bande d'images, et l'axe des y montre le nombre de pixels dans chaque classe qui a une valeur donnée dans cette bande. Il est clair que la classe A présente généralement des valeurs faibles, et la classe B des valeurs élevées, mais la distribution des valeurs dans chaque bande est suffisamment importante pour qu'il y ait un certain chevauchement entre les deux. Les deux distributions étant gaussiennes, nous pouvons calculer à la fois la moyenne et l'écart type pour chaque classe, puis calculer le z-score (combien d'écarts types nous séparent de la moyenne). Dans la figure 49, les deux classes ont le même écart type (les « cloches » ont la même « largeur »), et parce que le point est situé un peu plus près de la moyenne de la classe B que de la classe A, son score z serait le plus faible pour la classe B et il serait affecté à cette classe. Un exemple légèrement plus réaliste est fourni ci-dessous dans la figure 50, où nous avons deux dimensions et trois classes. Les écarts types de la bande 4 (axe des x) et de la bande 3 (axe des y) sont représentés sous forme de contours d'équiprobabilité. Dans ce cas, le défi pour le classificateur à maximum de vraisemblance est de trouver la classe pour laquelle le point se trouve dans le contour d'équiprobabilité au plus proche du centre de la classe. Voyez par exemple que les contours des classes A et B se chevauchent, et que les écarts types de la classe A sont plus grands que ceux de la classe B. Par conséquent, le point rouge est plus proche (dans l'espace des caractéristiques) du centre de la classe B que du centre de la classe A, mais il se trouve sur le troisième contour d'équiprobabilité de la classe B et sur le deuxième de la classe A. Le classificateur de distance minimale classerait ce point dans la « classe B » sur la base de la distance euclidienne la plus courte, tandis

que le classificateur de vraisemblance maximale le classerait dans la « classe A » en raison de sa plus grande probabilité d'appartenir à cette classe (selon les hypothèses utilisées). Qu'est-ce qui a le plus de chances d'être correct ? La plupart des comparaisons entre ces classificateurs suggèrent que le classificateur de vraisemblance maximale a tendance à produire des résultats plus précis, mais cela ne garantit pas qu'il soit toujours supérieur.

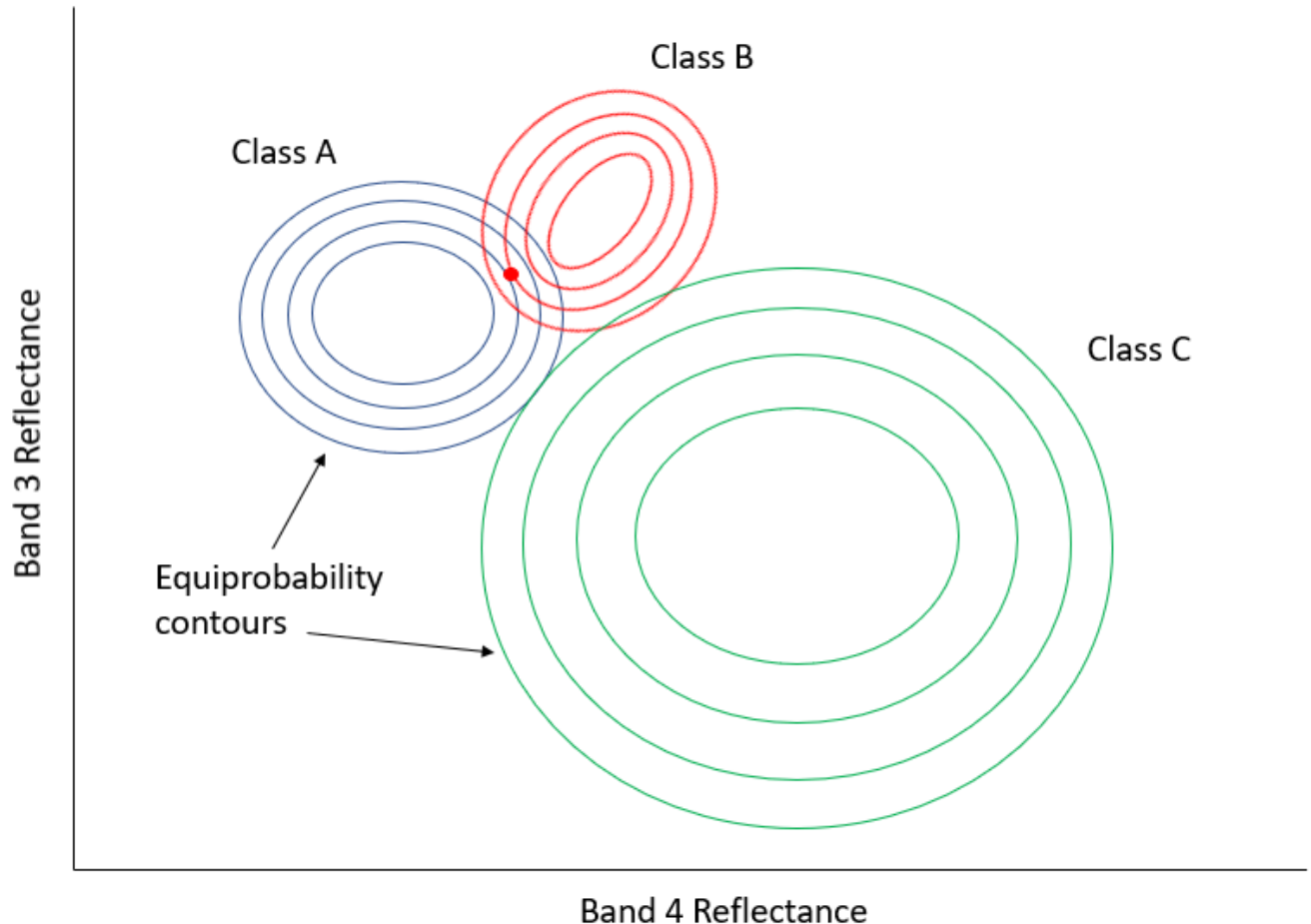


Figure 50: Exemple bidimensionnel de la situation de classification par maximum de vraisemblance, avec six classes qui ont des distributions standard inégales. Par Anders Knudby, CC BY 4.0.

Classificateurs non paramétriques

Au cours des dernières années, les scientifiques de la télédétection se sont de plus en plus tournés vers le domaine de l'apprentissage automatique pour adopter de nouvelles techniques de classification. L'idée de la classification est fondamentalement très générique : vous disposez de données sur quelque chose (dans notre cas, les valeurs des bandes d'un pixel) et vous voulez savoir de quoi il s'agit (dans notre cas, quelle est la couverture terrestre). Un problème peut difficilement être plus générique, si bien que l'on en trouve des versions partout : une banque dispose de certaines informations sur un client (âge, sexe, adresse, revenus, historique de remboursement des prêts) et veut savoir s'il doit être considéré comme un « faible risque », un « risque moyen » ou un « risque élevé » pour un nouveau prêt de 100 000 dollars. Un météorologue dispose

d'informations sur le temps actuel (« pluie, 5 °C ») et sur les variables atmosphériques (« 1003 mb, 10 m/s de vent du NW »), et doit déterminer s'il pleuvra ou non dans trois heures. Un ordinateur dispose de certaines informations sur les empreintes digitales trouvées sur une scène de crime (longueur, courbure, position relative de chaque ligne) et doit déterminer s'il s'agit des vôtres ou de celles d'un tiers.) Parce que la tâche est générique, et parce que les utilisateurs en dehors du domaine de la télédétection disposent de grandes quantités d'argent et peuvent utiliser les algorithmes de classification pour générer des profits, les informaticiens ont développé de nombreuses techniques pour résoudre cette tâche générique, et certaines de ces techniques ont été adoptées en télédétection. Nous allons examiner un seul exemple ici, mais n'oubliez pas qu'il existe de nombreux autres algorithmes de classification génériques qui peuvent être utilisés en télédétection.

Celui que nous allons examiner s'appelle un classificateur à arbre de décision. Comme les autres classificateurs, le classificateur à arbre de décision fonctionne selon un processus en deux étapes : 1) calibrer l'algorithme de classification, et 2) l'appliquer à tous les pixels de l'image. Un classificateur à arbre de décision est calibré en divisant récursivement l'ensemble des données (tous les pixels sous les polygones de la figure 45) pour maximiser l'homogénéité des deux parties (appelées nœuds). Une petite illustration : disons que nous avons 7 points de données (vous ne devriez *jamais* avoir seulement sept points de données lorsque vous calibrez un classificateur, ce petit nombre est utilisé uniquement à des fins d'illustration !)

Tableau 4: Sept points de données utilisés pour développer un classificateur à arbre de décision.

Data point	Band 1 value	Band 2 value	Actual class
1	10	30	A
2	20	40	A
3	30	40	A
4	15	55	B
5	35	40	B
6	40	35	B
7	45	35	B

La première tâche consiste à trouver une valeur, soit dans la bande 1, soit dans la bande 2, qui peut être utilisée pour diviser l'ensemble de données en deux nœuds de sorte que, dans la mesure du possible, tous les points de la classe A se trouvent dans un nœud et tous les points de la classe B dans l'autre. D'un point de vue algorithmique, cela se fait en testant toutes les valeurs possibles et en quantifiant l'homogénéité des classes résultantes. Ainsi, nous pouvons observer que la plus petite valeur dans la bande 1 est 10, et la plus grande est 45. Si nous divisons l'ensemble de données selon la règle suivante : « Tous les points dont la bande 1 est

inférieure à 11 vont au nœud X, et tous les autres au nœud y », nous obtiendrons des points divisés comme suit :

Data point	Band 1 value	Band 2 value	Actual class
1	10	30	A
2	20	40	A
3	30	40	A
4	15	55	B
5	35	40	B
6	40	35	B
7	45	35	B

Figure 51: Répartition des points selon la valeur seuil 11 dans la bande 1. Par Anders Knudby, CC BY 4.0.

Comme nous pouvons le voir, cela nous laisse avec un seul point « A » dans un nœud (X), et deux points « A » et quatre points « B » dans l'autre nœud (Y). Pour savoir si nous pouvons faire mieux, nous pouvons essayer d'utiliser la valeur 12 au lieu de 11 (ce qui nous donne le même résultat), 13 (toujours le même), et ainsi de suite, et lorsque nous avons testé toutes les valeurs de la bande 1, nous continuons avec toutes les valeurs de la bande 2. Finalement, nous trouverons qu'utiliser la valeur 31 dans la bande 1 nous donne le résultat suivant :

Data point	Band 1 value	Band 2 value	Actual class
1	10	30	A
2	20	40	A
3	30	40	A
4	15	55	B
5	35	40	B
6	40	35	B
7	45	35	B

Figure 52: Points crachés selon la valeur seuil 31 dans la bande 1. Par Anders Knudby, CC BY 4.0.

C'est presque parfait, sauf que nous avons un seul « B » dans le nœud X. Mais bon, c'est plutôt bien pour un premier partage. Nous pouvons représenter cela sous forme d'un « arbre » comme ceci :

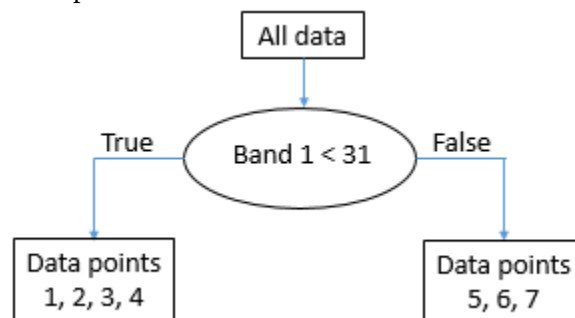


Figure 53: La structure « arborescente » qui émerge de la division des données sur la valeur seuil 31 dans la bande 1. Chaque ensemble de points de données est appelé un nœud. Le « nœud racine » contient tous les points de données. Les « nœuds feuilles », également appelés « nœuds terminaux », sont les points finaux. Par Anders Knudby, CC BY 4.0.

Le nœud avec les points de données 5, 6 et 7 (qui ont tous des valeurs de bande 1 supérieures à 31) est

maintenant ce qu'on appelle un « nœud pur » — il est composé de points de données d'une seule classe, nous n'avons donc plus besoin de le diviser. Les nœuds qui sont des points terminaux sont également appelés « feuilles ». Le nœud avec les points de données 1, 2, 3 et 4 n'est pas « pur », car il contient un mélange de points de classe A et de classe B. Nous recommençons donc à tester toutes les différentes valeurs possibles que nous pouvons utiliser comme seuil pour diviser ce nœud (et seulement ce nœud), dans les deux bandes. Il se trouve que le point de la classe B dans ce nœud a une valeur dans la bande 2 qui est plus élevée que tous les autres points, donc une valeur de séparation de 45 fonctionne bien, et nous pouvons mettre à jour l'arbre comme ceci :

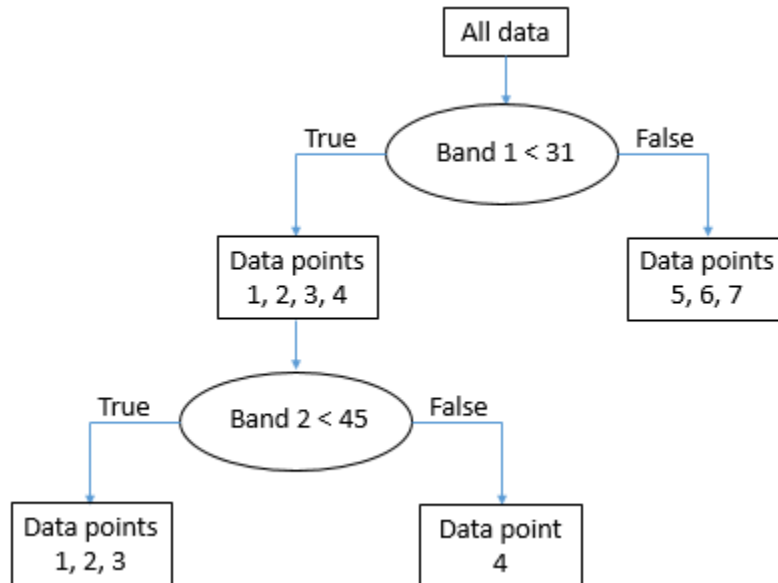


Figure 54: La structure « arborescente » finale. Tous les nœuds (parties finales de l'ensemble de données) sont maintenant purs.

Par Anders Knudby, CC BY 4.0.

Une fois que « l'arbre » est en place, nous pouvons maintenant prendre tous les autres pixels de l'image et les faire « tomber » dans l'arbre pour voir dans quelle feuille ils atterrissent. Par exemple, un pixel ayant des valeurs de 35 dans la bande 1 et de 25 dans la bande 2 « ira à droite » au premier test et atterrira donc dans la feuille qui contient les points de données 5, 6 et 7. Comme tous ces points appartiennent à la classe B, ce pixel sera classé dans la classe B. Et ainsi de suite.

Notez que les feuilles ne doivent pas nécessairement être « pures », certains arbres cessent de diviser les nœuds lorsqu'ils sont inférieurs à une certaine taille, ou en utilisant un autre critère. Dans ce cas, un pixel atterrissant dans une telle feuille se verra attribuer la classe qui compte le plus de points dans cette feuille (et l'information selon laquelle il ne s'agissait pas d'une feuille pure peut même être utilisée pour indiquer que la classification de ce pixel particulier est sujette à une certaine incertitude).

Le classificateur à arbre de décision n'est qu'un des nombreux exemples possibles de classificateurs non — paramétriques. Il est rarement utilisé directement sous la forme présentée ci-dessus, mais il constitue la base de certains des algorithmes de classification les plus performants utilisés aujourd'hui. D'autres algorithmes de classification non paramétriques populaires incluent les réseaux neuronaux et les machines à vecteurs de support, qui sont tous deux implémentés dans de nombreux logiciels de télédétection.

Classification non supervisée par pixel

Que faire si nous ne disposons pas des données nécessaires pour calibrer un algorithme de classification ? Si nous n'avons pas les polygones illustrés à la figure 45, ou les points de données indiqués au tableau 4 ? Que faisons-nous alors ? Nous utilisons une classification non supervisée à la place !

La classification non supervisée consiste à laisser un algorithme diviser les pixels d'une image en « clusters naturels », c'est-à-dire en combinaisons de valeurs de bandes que l'on retrouve fréquemment dans l'image. Une fois ces groupes naturels identifiés, l'analyste d'images peut alors les étiqueter, en se basant généralement sur une analyse visuelle de l'emplacement de ces groupes dans l'image. Le regroupement est en grande partie automatique, bien que l'analyste fournisse quelques paramètres initiaux. L'un des algorithmes les plus couramment utilisés pour trouver des groupes naturels dans une image est l'algorithme K-Means, qui fonctionne comme suit :

1) L'analyste détermine le nombre souhaité de classes. Fondamentalement, si vous voulez une carte avec un haut niveau de détail thématique, vous pouvez définir un grand nombre de classes. Notez également que les classes peuvent être combinées ultérieurement, il est donc souvent judicieux de fixer le nombre de classes souhaitées à un niveau légèrement supérieur à celui que vous pensez serait votre résultat de la fin. Un nombre de points « d'amorçage » égal au nombre de classes souhaité est ensuite placé de manière aléatoire dans l'espace des caractéristiques.

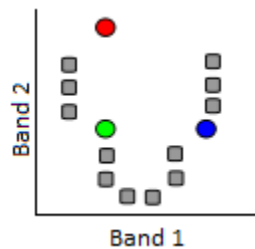


Figure 55: Classification K-Means étape 1. Un certain nombre de points « d'amorçage » (points colorés) sont distribués aléatoirement dans l'espace des caractéristiques. Les points gris représentent ici les pixels à regrouper. Modifié à partir de K Means

Example Step 1 par Weston.pace, Wikimedia Commons, CC BY-SA 3.0.

2) Des clusters sont ensuite générés autour des points « graines » en attribuant tous les autres points à la graine la plus proche.

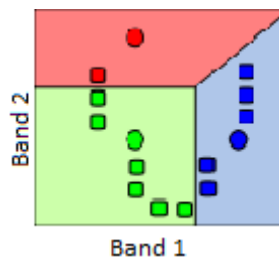


Figure 56: Un cluster est formé autour de chaque graine en attribuant tous les points à la graine la plus proche. Modifié à partir de K Means Example Step 2 par Weston.pace, Wikimedia Commons, CC BY-SA 3.0.

3) Le centroïde (centre géographique) des points de chaque groupe devient la nouvelle « graine ».

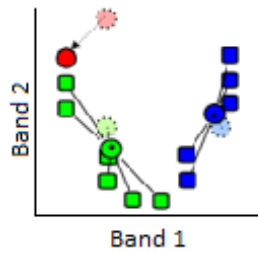


Figure 57: Les graines sont déplacées vers le centroïde de chaque grappe. Le centroïde est calculé comme le centre géographique de chaque grappe, c'est-à-dire qu'il est situé à la valeur x moyenne de tous les points de la grappe, et à la valeur y moyenne de tous les points de la grappe. Modifié à partir de K Means Example Step 3 par Weston.pace, Wikimedia Commons, CC BY-SA 3.0.

4) Répétez les étapes 2 et 3 jusqu'au critère d'arrêt. Le critère d'arrêt peut être qu'aucun point ne se déplace vers un autre cluster, ou que le centroïde de chaque cluster se déplace de moins d'une distance prédéfinie, ou qu'un certain nombre d'itérations ont été complétées.

D'autres algorithmes de classification non supervisée effectuent le regroupement de manière légèrement différente. Par exemple, un algorithme populaire appelé ISODATA permet également la division de grands clusters pendant le processus de clustering, et de la même manière la fusion de petits clusters proches. Néanmoins, le résultat de l'algorithme de clustering est que chaque pixel de l'image entière fait partie d'un cluster. L'espoir est donc que chaque cluster représente un type d'occupation du sol qui peut être identifié par l'analyste d'images, par exemple en superposant l'emplacement de tous les pixels du cluster sur l'image originale pour identifier visuellement ce à quoi correspond ce cluster. C'est l'étape finale de la classification non supervisée — l'étiquetage de chacun des clusters qui ont été produits. C'est l'étape où il peut être pratique de fusionner des clusters, si par exemple vous avez un cluster qui correspond à une eau trouble et un autre qui correspond à une eau claire. À moins que vous ne soyez spécifiquement intéressé par la qualité de l'eau, différencier les deux n'est probablement pas important, et les fusionner fournira un produit cartographique plus clair. De même, il se peut que vous ayez simplement deux grappes qui semblent toutes deux correspondre à une forêt feuillue saine. Même si vous travaillez pour un service forestier, à moins que vous ne puissiez déterminer avec certitude quelle est la différence entre ces deux groupements, vous pouvez les fusionner en un seul et les appeler « forêt de feuillus ».

À titre d'exemple, l'image ci-dessous montre l'image originale en arrière-plan, et les pixels centraux colorés selon le produit d'une classification non supervisée. Il est clair que la zone « bleue » correspond aux pixels recouverts d'eau, et que la zone verte correspond en grande partie à la végétation. Une analyse plus détaillée de l'image serait nécessaire pour étiqueter chaque zone, en particulier les zones rouges et grises, de manière appropriée.

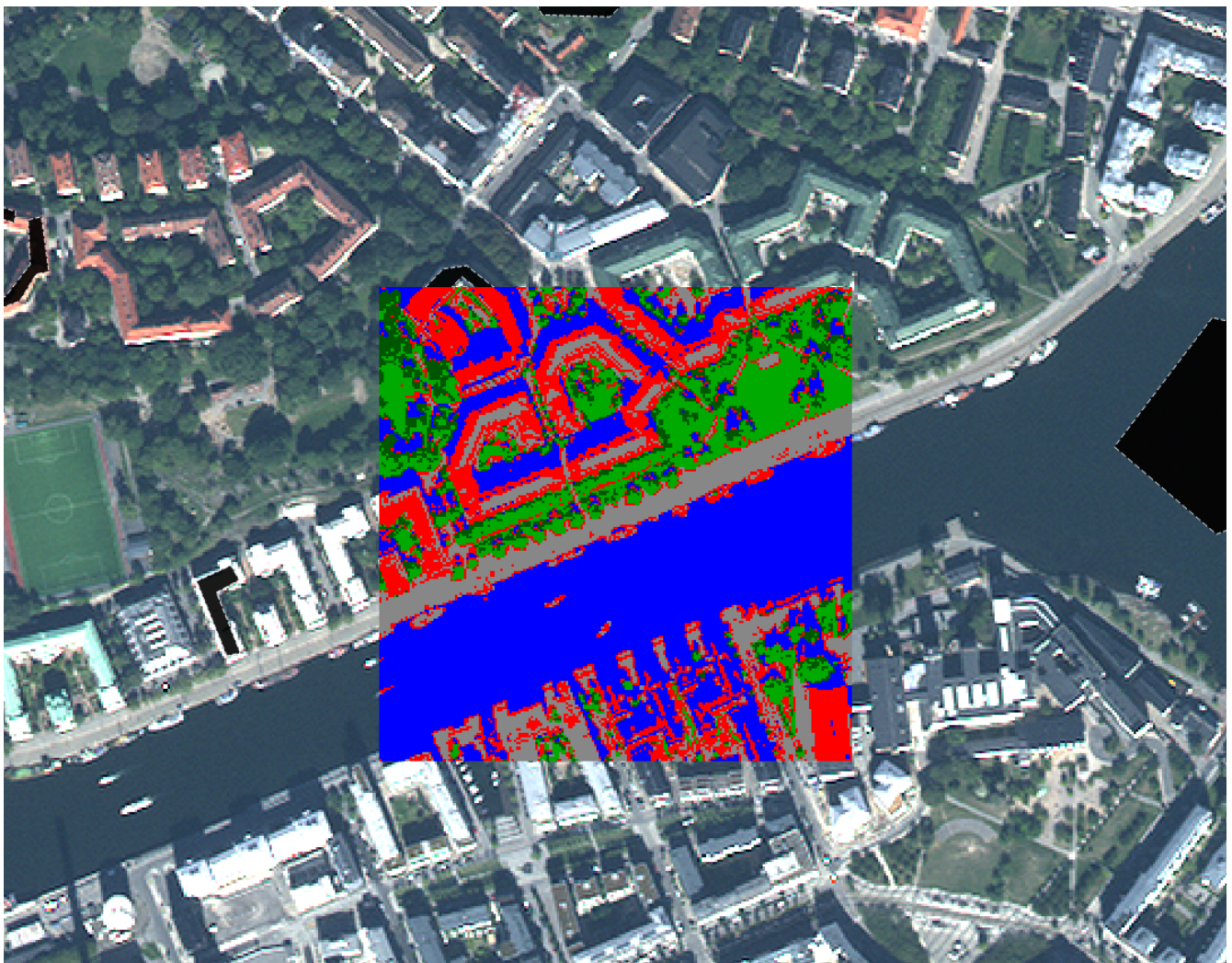


Figure 58: Exemple de correspondance entre l'image originale et les clusters formés dans un processus de classification non supervisé. Par Anders Knudby, CC BY 4.0.

La classification de l'occupation du sol est l'une des plus anciennes utilisations de la télédétection, et c'est quelque chose que de nombreux gouvernements nationaux font régulièrement pour leur territoire. Par exemple, au Canada, le Centre canadien de télédétection travaille avec des partenaires des États-Unis et du Mexique pour créer une carte de la couverture terrestre nord-américaine. Des cartes mondiales d'occupation du sol sont également produites par divers établissements, comme l'USGS, l'Université du Maryland, l'ESA et la Chine, pour n'en citer que quelques-unes.

L'un des défauts typiques des systèmes de classification d'images qui fonctionnent au niveau du pixel par pixel est que les images sont bruyantes, et que les cartes d'occupation du sol réalisées à partir d'images héritent de ce bruit. Un autre défaut plus important est qu'il y a de l'information dans une image au-delà de ce que l'on trouve dans les pixels individuels. Une image est l'illustration parfaite du dicton « le tout est plus grand que la somme de ses parties », car les images ont une structure, et la structure n'est pas prise en compte lorsqu'on regarde chaque pixel indépendamment du contexte fourni par tous les pixels voisins. Par exemple, même sans

connaître la couleur d'un pixel, si je sais que tous ses pixels voisins sont classés comme « eau », je peux dire avec une grande confiance que le pixel en question est aussi « eau ». J'aurai tort à l'occasion, mais j'aurai raison la plupart du temps. Nous allons maintenant nous pencher sur une technique appelée analyse d'image basée sur l'objet, qui tient compte du contexte lorsqu'elle génère des classifications d'images. Cet avantage lui permet souvent de surpasser les méthodes plus traditionnelles de classification pixel par pixel.

Analyse d'images basée sur les objets (OBIA)

Beaucoup de nouvelles avancées dans le monde de la télédétection proviennent du côté matériel des choses. Un nouveau capteur est lancé, et il a une meilleure résolution spatiale ou spectrale que les capteurs précédents, ou il produit des images moins bruyantes, ou il est mis gratuitement à disposition alors que les autres solutions étaient payantes. Les drones sont un autre exemple : le type d'images qu'ils produisent n'est pas très différent de ce qui était auparavant disponible avec les caméras des avions pilotés — en fait, leur qualité est souvent inférieure — mais le faible coût des drones et la facilité avec laquelle ils peuvent être déployés par des non-experts ont créé une révolution en matière de quantité d'images disponibles à basse altitude et de coût d'obtention d'images à haute résolution pour un petit site.

Le développement de l'analyse d'images basée sur les objets (en anglais « Object-Based Image Analysis », OBIA) est l'une des rares avancées substantielles du côté des logiciels. Le principe fondamental de l'OBIA est de considérer qu'une image est composée d'objets plutôt que de pixels. L'un des avantages de ce principe est que les gens ont tendance à voir le monde comme étant composés d'objets, et non de pixels, de sorte qu'une analyse d'image qui adopte le même point de vue produit des résultats qui sont plus facilement interprétés par les gens. Par exemple, lorsque vous regardez la figure 59, vous pouvez probablement voir le visage d'un homme (si vous le connaissez, vous reconnaîtrez également qui est cet homme).



Figure 59: Robert De Niro. Ou, si vous utilisez un logiciel d'analyse d'images basé sur les pixels, une trame à trois bandes comportant 1556 colonnes et 2247 lignes, chaque pixel et chaque bande affichant un niveau de luminosité variable. Robert De Niro KVIFF portrait par Petr Novák (che), Wikimedia Commons, CC BY-SA 2.5.
Comme il s'agit d'une image numérique, nous savons qu'elle est en fait composée d'un certain nombre de

pixels disposés proprement en colonnes et en rangées, et que la luminosité (c'est-à-dire l'intensité de la couleur rouge, verte et bleue de chaque pixel) peut être représentée par trois nombres. Nous pourrions donc classer les parties claires de l'image comme « peau » et les parties moins claires comme « autres », une classe mixte comprenant les yeux, les cheveux, les ombres et le fond. Mais ce n'est pas une classification particulièrement utile ou importante ! Ce qui serait plus important serait de classer l'image en classes telles que « œil », « main », « cheveux », « nez », etc.

Un exemple un peu plus pertinent pour la télédétection est présenté ci-dessous à la figure 60, dans laquelle une zone urbaine a été classée en objets, notamment un stade facilement reconnaissable, des rues, des bâtiments individuels, de la végétation, etc.

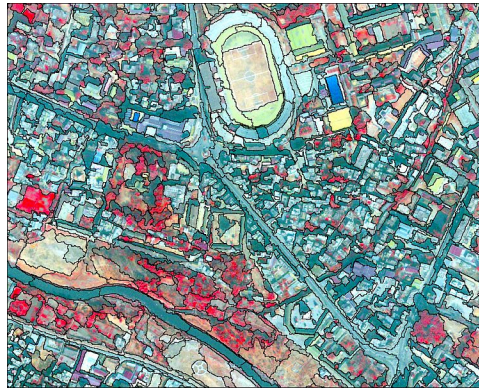


Figure 60: Classification d'une zone urbaine à l'aide d'une analyse d'images basée sur les objets. Object based image analysis par Uddinkabir, Wikimedia Commons, CC BY-SA 4.0.

La segmentation d'image

Le but de la segmentation d'image est de prendre tous les pixels de l'image et de les diviser en segments — des parties contiguës de l'image qui ont une couleur similaire. La segmentation d'image est utile parce qu'elle nous évite d'analyser l'image pixel par pixel et nous permet d'analyser les segments individuels. Cette méthode présente plusieurs avantages très importants. Tout d'abord, nous pouvons examiner la « couleur moyenne » d'un segment et l'utiliser pour classer ce segment, plutôt que d'utiliser la couleur de chaque pixel individuel pour classer ce pixel. Si nous avons affaire à une imagerie bruyante (et c'est toujours le cas !), l'utilisation des moyennes de segments réduit l'influence du bruit sur la classification. Deuxièmement, les segments ont un certain nombre d'attributs qui peuvent être considérables et peuvent être utilisés pour aider à les classer — des attributs que les pixels n'ont pas. Par exemple, un segment d'image est constitué d'un certain nombre de pixels, et nous pouvons donc quantifier sa taille. Comme tous les segments n'ont pas la même taille, l'attribut « taille du segment » contient des informations qui peuvent être utilisées pour aider à classer les segments. Par exemple, dans la figure 60, remarquez que le segment couvrant le lac est assez grand par rapport à tous les segments terrestres. Cela s'explique par le fait que le lac est une partie très homogène de l'image. La même différence est observée entre les différentes parties de la zone terrestre — les segments situés juste à l'ouest du lac

sont généralement plus grands que ceux situés au sud-ouest du lac, là encore parce qu'ils sont plus homogènes, et que cette homogénéité peut nous renseigner sur le type de couverture terrestre que l'on y trouve. En dehors de la taille, les segments ont un grand nombre d'autres attributs qui peuvent ou non être utiles dans une classification. Chaque segment a un nombre spécifique de segments voisins, et chaque segment est également plus foncé ou plus clair ou quelque part au milieu par rapport à ses voisins. Voir cette page pour une liste des attributs que vous pouvez calculer pour les segments dans le module d'extraction de caractéristiques d'ENVI (qui n'est *pas* le module le plus complet, et de loin). La possibilité d'obtenir toutes ces informations sur les segments peut aider à les classer, et ces informations ne sont pas toutes disponibles pour les pixels (par exemple, tous les pixels ont exactement quatre pixels voisins, sauf s'ils sont situés sur un bord de l'image, donc le nombre de pixels voisins n'est pas utile pour classer un pixel).

Il existe différents types d'algorithmes de segmentation, et tous sont compliqués à calculer. Certains sont code ouvert (open source), tandis que d'autres sont propriétaires, de sorte que nous ne savons pas vraiment comment ils fonctionnent. Nous n'allons donc pas entrer dans le détail des spécificités des algorithmes de segmentation, mais ils ont quelques points communs que nous pouvons prendre en compte.

- Échelle : Tous les algorithmes de segmentation requièrent un « facteur d'échelle », que l'utilisateur définit pour déterminer la taille qu'il souhaite donner aux segments résultants. Le facteur d'échelle ne correspond pas nécessairement à un certain nombre de pixels ou à un certain nombre de segments, mais doit généralement être considéré comme un nombre relatif. Ce qu'il signifie réellement en matière de taille des segments résultants est généralement déterminé par un processus d'essai et d'erreur.
- Couleur vs forme : Tous les algorithmes de segmentation doivent faire des choix quant à l'endroit où tracer les limites de chaque segment. Comme il est généralement souhaitable d'avoir des segments qui ne sont pas trop bizarres, cela implique souvent un compromis entre l'ajout d'un pixel dans un segment existant si a) ce pixel rend la couleur du segment plus homogène, mais donne aussi une forme plus bizarre, ou b) si ce pixel rend la couleur du segment moins homogène, mais donne une forme plus compacte. Un ou plusieurs paramètres contrôlent généralement ce compromis, et comme pour le paramètre « échelle », trouver le meilleur réglage est une question d'essais et d'erreurs.

Dans un monde idéal, la segmentation de l'image devrait produire des segments qui correspondent chacun à un, et exactement un, objet du monde réel. Par exemple, si vous avez l'image d'une zone urbaine et que vous souhaitez cartographier tous les bâtiments, l'étape de segmentation de l'image devrait idéalement aboutir à un segment distinct pour chaque bâtiment. Dans la pratique, c'est généralement impossible, car l'algorithme de segmentation ne sait pas que vous cherchez des bâtiments... si vous cherchez à cartographier des carreaux, il serait inutile d'avoir des segments correspondant à un toit entier, comme ce serait le cas si vous cherchiez à cartographier des pâtés de maisons. « Mais je pourrais définir le facteur d'échelle en conséquence », direz-vous, et c'est vrai dans une certaine mesure. Mais tous les toits n'ont pas la même taille ! Si vous cartographiez des bâtiments dans une image qui contient à la fois votre propre maison et le Pentagone, il est peu probable

que vous trouviez un facteur d'échelle qui vous donne exactement un segment couvrant chaque bâtiment... La solution à ce problème est généralement une intervention manuelle, dans laquelle les segments produits par la segmentation initiale sont modifiés selon des règles spécifiques. Par exemple, vous pouvez sélectionner un facteur d'échelle qui fonctionne pour votre maison et laisser le Pentagone divisé en 1000 segments, puis fusionner ultérieurement tous les segments voisins qui ont des couleurs très similaires. En supposant que le toit du Pentagone soit assez homogène, tous ces segments seraient fusionnés, et en supposant que votre propre maison soit entourée de quelque chose d'apparence différente, comme une rue, une cour ou une allée, votre propre toit ne serait pas fusionné avec ses segments voisins. Cette possibilité de modifier manuellement le processus pour obtenir les résultats souhaités est à la fois une grande force et une importante faiblesse de l'analyse d'images basée sur les objets. C'est une force parce qu'elle permet à un analyste d'images de produire des résultats extrêmement précis, mais une faiblesse parce qu'il faut beaucoup de temps, même aux analystes experts, pour faire cela pour chaque image. Un exemple d'une bonne segmentation est présenté dans la figure 61, qui illustre également l'influence des changements dans les paramètres de segmentation (comparez les images en bas à gauche et en bas à droite)

Il convient de noter ici que l'analyse d'images par objet a été initialement développée et largement utilisée dans le domaine de l'imagerie médicale, pour analyser des images de rayons X, des cellules vues au microscope, etc. Deux choses sont très différentes entre l'imagerie médicale et la télédétection : dans le domaine de l'imagerie médicale, la santé des personnes est très directement affectée par l'analyse de l'image, de sorte que le fait qu'il faille plus de temps pour obtenir un résultat précis est moins contraignant que dans le domaine de la télédétection, où les impacts humains d'une mauvaise analyse d'image sont très difficiles à évaluer (bien qu'ils puissent finalement être tout aussi importants). L'autre problème est que les images étudiées par les médecins sont créées dans des environnements hautement contrôlés, avec pratiquement aucun bruit de fond, la cible de l'image toujours au point, et peuvent être refaites si elles sont difficiles à analyser. En télédétection, si la brume, un mauvais éclairage, la fumée ou d'autres facteurs environnementaux se combinent pour créer une image bruyante, notre seule option est d'attendre le prochain passage du satellite au-dessus de la zone.

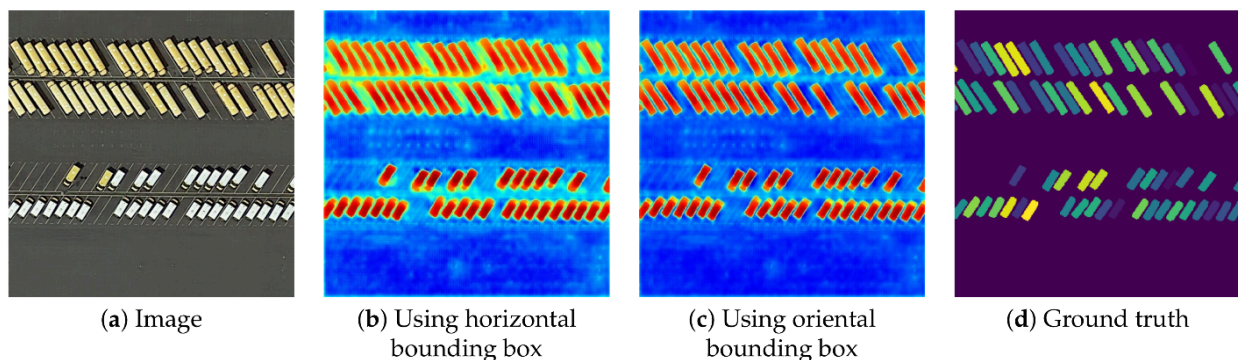


Figure 61: Exemple de segmentation d'image utilisée pour cartographier des véhicules individuels sur un parking à l'aide d'une imagerie aérienne. Une segmentation d'image est appliquée à l'image, et les segments qui correspondent aux caractéristiques d'un

véhicule (forme, taille, couleur) sont ensuite identifiés. Efficient Hybrid Supervision for Instance Segmentation in Aerial Images

(Figure 6) par Chen et al., CC BY 4.0.

Classification des segments

La classification des segments pourrait en théorie suivre les approches supervisées et non supervisées décrites dans le chapitre sur la classification par pixel. En d'autres termes, des zones dont la couverture terrestre est connue pourraient être utilisées pour calibrer un classificateur basé sur un ensemble prédéterminé d'attributs de segments (approche supervisée), ou un ensemble d'attributs prédéterminés pourrait être utilisé dans un algorithme de regroupement pour définir des groupes naturels de segments dans l'image, qui pourraient ensuite être étiquetés par l'analyste d'images (approche non supervisée). Cependant, dans la pratique, l'analyse d'images basée sur les objets se déroule souvent de manière plus interactive. Une approche courante consiste pour l'analyste à développer un ensemble de règles structuré sous forme d'arbre de décision, à voir le résultat produit par cet ensemble de règles, à le modifier ou à le compléter, puis à vérifier à nouveau, le tout dans un processus itératif très intensif pour l'utilisateur. C'est l'un des domaines où la télédétection ressemble plus à un art qu'à une science, parce que les analystes d'images acquièrent de l'expérience avec ce processus et deviennent de plus en plus compétents, développant essentiellement leur propre « style » de développement des ensembles de règles nécessaires pour obtenir une bonne classification. Par exemple, après avoir segmenté une image, vous pouvez vouloir séparer les surfaces artificielles (routes, trottoirs, stationnements, toits) des surfaces naturelles, ce qui est typiquement possible, car les premières ont tendance à être grises et les secondes non. Vous développez donc une variable qui quantifie le degré de gris d'un segment (par exemple en utilisant la valeur de « saturation » HSV, plus d'informations ici), et vous trouvez manuellement une valeur seuil qui sépare efficacement les surfaces naturelles des surfaces artificielles dans votre image. Si vous faites des études urbaines, vous voudrez peut-être aussi différencier les différents types de surfaces artificielles. Elles sont toutes grises, donc avec un classificateur basé sur les pixels, vous n'auriez pas de chance. Cependant, vous traitez des segments, et non des pixels, et vous quantifiez donc l'allongement de chaque segment, et définissez une valeur seuil qui vous permet de séparer les routes et les trottoirs des toits et des stationnements. Ensuite, vous utilisez la largeur des segments allongés pour différencier les routes des trottoirs, et enfin vous utilisez le fait que les stationnements ont des segments voisins qui sont des routes, alors que les toits n'en ont pas, pour les séparer. Trouver toutes les bonnes variables à utiliser (valeur de saturation HSV, allongement, largeur, classe de segments voisins) est un exercice intrinsèquement subjectif qui implique des essais et des erreurs et qui s'améliore avec l'expérience. Et une fois que vous avez créé une première structure d'arbre de décision (ou similaire) pour votre ensemble de règles, il est très probable que vos yeux soient attirés par un ou deux segments qui, malgré tous vos efforts, ont été mal classés (peut — être y a-t-il le toit occasionnel qui s'étend si loin au-dessus de la maison que son segment voisin est en fait une route, donc il a été mal classé comme un stationnement). Vous pouvez maintenant créer des règles encore plus sophistiquées pour remédier à ces problèmes spécifiques, dans un processus qui ne s'achève que lorsque votre classification est parfaite. C'est

tendant si vous êtes fier de votre travail, mais cela prend du temps, et cela finit par aboutir à un ensemble de règles si spécifiques qu'elles ne fonctionnent jamais que pour l'image pour laquelle elles ont été créées. Une possibilité est de conserver l'ensemble de règles plus générique qui était imparfait, mais fonctionnait assez bien, puis de modifier manuellement la classification pour les segments que vous savez avoir été mal classés.

En résumé, la classification d'images basée sur les objets est une méthodologie relativement nouvelle qui repose sur deux étapes : 1) la division de l'image en segments contigus et homogènes, suivie par 2) une classification de ces segments. Elle peut être utilisée pour produire des classifications d'images qui surpassent presque toujours les classifications par pixel en termes en matière de précision, mais qui demandent beaucoup plus de temps à un analyste d'images. Le premier, le meilleur et le plus utilisé des logiciels d'OBIA dans le domaine de la télédétection sont eCognition, mais d'autres logiciels commerciaux comme ENVI et Geomatica ont également développé leurs propres modules d'OBIA. Les logiciels SIG tels que ArcGIS et QGIS fournissent des outils de segmentation d'images, ainsi que des outils qui peuvent être combinés pour effectuer la classification des segments, mais offrent généralement des flux de travail OBIA moins rationalisés. En outre, au moins un logiciel libre a été spécifiquement conçu pour l'OBIA, et des outils de segmentation d'images sont disponibles dans certaines bibliothèques de traitement d'images, comme OTB.

7.

ÉVALUATION D'EXACTITUDE

Une fois que nous avons produit une classification de l'occupation du sol (ou autre) à partir d'une image de télédétection, une question évidente est « quelle est l'exactitude de cette carte ? ». Il est important de répondre à cette question, car nous voulons que les utilisateurs de la carte aient un degré de confiance approprié en elle. Si la carte est parfaite, nous voulons que les gens le sachent afin qu'ils puissent en tirer le maximum d'avantages. Et si la carte n'est pas plus précise qu'une affectation aléatoire des classes aux pixels ne l'aurait été, nous voulons aussi que les gens le sachent, afin qu'ils ne l'utilisent pas pour quoi que ce soit (sauf peut-être pour l'accrocher au mur ou la montrer aux étudiants comme exemple de ce qu'il ne faut *pas* faire...).

Le sujet de l'évaluation de l'exactitude va également au-delà des classifications et concerne les cartes de variables continues, telles que la température de la surface de la Terre, la concentration de CO₂ près de la surface, la santé de la végétation, ou d'autres variables se présentant sous la forme de variables continues plutôt que discrètes. Quel que soit le contenu de votre carte, vous voudrez que les gens sachent à quel point elle est bonne, à quel point ils peuvent lui faire confiance. Bien qu'il y ait des similitudes entre l'évaluation de cartes de variables catégorielles et continues, les mesures spécifiques utilisées pour quantifier l'exactitude sont différentes dans les deux cas, c'est pourquoi nous les traiterons tour à tour dans ce chapitre.

Évaluation de l'exactitude des classifications

Le principe de base pour toute évaluation de l'exactitude est de comparer les estimations avec la réalité, et de quantifier la différence entre les deux. Dans le contexte des classifications de l'occupation du sol basées sur la télédétection, les « estimations » sont les classes cartographiées pour chaque pixel, et la « réalité » est l'occupation réelle du sol dans les zones correspondant à chaque pixel. Étant donné que l'algorithme de classification nous a déjà fourni les « estimations », le premier défi de l'évaluation de l'exactitude est de trouver des données sur la « réalité ». Ces données sont souvent appelées « données de terrain » et consistent généralement en des observations de terrain géoréférencées de l'occupation du sol. Une technique souvent utilisée consiste à se rendre physiquement dans la zone d'étude avec un GPS et un appareil photo, et à prendre des photos géoréférencées qui permettent ensuite de déterminer visuellement la couverture du sol à partir de chaque photo. Étant donné que les gens peuvent visuellement distinguer les différents types d'occupation du sol avec une grande exactitude, ces données peuvent raisonnablement être considérées comme représentant la « réalité ». Dans de nombreux cas, cependant, le terme « vérité terrain » surestime l'exactitude de ce type d'information. Les gens peuvent être capables de distinguer le « désert » de la « forêt » sur une photo, mais

ils sont nettement moins bons pour distinguer la « forêt à haute densité » de la « forêt à moyenne densité ». En particulier, si la différence entre deux classes est basée sur le pourcentage de couverture (par exemple, la différence entre une forêt de densité moyenne et une forêt de haute densité peut résider dans le fait que les arbres couvrent plus ou moins de 50 % de la surface), les observations sur le terrain ne conduisent pas toujours à une description parfaite de la réalité. De nombreux spécialistes de la télédétection aiment donc mieux le terme de « données de validation », suggérant que ces données sont appropriées comme base de comparaison avec les classifications basées sur la télédétection, tout en reconnaissant la possibilité qu'elles ne correspondent pas parfaitement à la « vérité ».

Création de données de validation

Si vous voulez produire une évaluation honnête et impartiale de l'exactitude de votre carte d'occupation du sol (et je suppose que c'est le cas !), il y a quelques éléments à prendre en compte lors de la création de votre ensemble de données de validation :

- Vous devez disposer de données de validation couvrant toutes les différentes classes d'occupation du sol de votre carte. Si ce n'est pas le cas, vous ne pourrez vraiment évaluer l'exactitude que des parties de la carte couvertes par les classes pour lesquelles vous disposez de données.
- Idéalement, vous devriez aussi avoir des données de validation qui sont distribuées de façon aléatoire (ou plus ou moins égale) dans votre zone d'étude. Pour produire un ensemble de données de validation qui couvre toutes les classes et qui présente une bonne distribution spatiale dans votre zone d'étude, une sélection aléatoire stratifiée de points de validation est souvent utilisée (c'est-à-dire qu'elle inclut un certain nombre de points de chaque classe, les points appartenant à chaque classe étant distribués aléatoirement dans la zone couverte par cette classe).
- Le nombre de points de données utilisés pour chaque classe doit être soit le même, soit refléter l'étendue relative de chaque classe sur votre carte. La première approche est la plus appropriée si vous souhaitez comparer des classes et déterminer lesquelles sont mieux cartographiées que d'autres. La seconde approche est la plus adaptée si vous souhaitez produire une seule estimation de l'exactitude pour l'ensemble de la carte.
- Plus il y a de données de validation, mieux c'est. Cependant, la création de données de validation peut prendre du temps et de l'argent, donc obtenir « suffisamment » de données est souvent un objectif raisonnable. Il existe des règles empiriques sur ce qui constitue une quantité « suffisante » de données (par exemple, 50 par classe), mais il existe de nombreuses exceptions à ces règles. Si vous utilisez des observations sur le terrain pour créer vos données de validation, il est important de se rappeler que les données de validation doivent être comparables aux classes dérivées de votre image, de plusieurs façons.

Si vous utilisez des observations sur le terrain pour créer vos données de validation, il est important de se

rappeler que les données de validation doivent être comparables aux classes dérivées de votre image, de plusieurs façons :

- Les définitions utilisées pour chaque classe doivent être les mêmes entre la classification et les données de validation. Par exemple, si dans votre classification vous considérez qu'une « masse d'eau » doit avoir une taille d'au moins $0,1 \text{ km}^2$, vous devez garder cela à l'esprit lorsque vous créez vos données de validation. Ainsi, lorsque vous allez sur le terrain et que l'un de vos points de données se trouve dans une flaque d'eau, vous ne le considérez pas comme une « masse d'eau », mais vous déterminez plutôt quelle est la couverture terrestre autour de la flaque.
- En rapport avec le point précédent, gardez à l'esprit la résolution spatiale de l'image utilisée pour produire votre classification. Si vous avez basé votre classification sur l'imagerie Landsat (TM, ETM+, OLI) sans pansharpning, chaque pixel correspond à une zone d'environ 30 x 30 mètres sur le terrain. Ainsi, lorsque vous allez sur le terrain, vous devriez documenter la couverture terrestre dominante dans des zones de 30 x 30 mètres, plutôt que la couverture terrestre aux coordonnées exactes du point de données.

Une autre approche pour créer des données de validation, utile lorsque se rendre dans la zone d'étude et recueillir des observations sur le terrain est trop coûteux, est l'inspection visuelle d'images de télédétection à haute résolution. Si vous choisissez cette approche, vous devez être sûr de pouvoir distinguer visuellement toutes les différentes classes, à partir de l'image, avec une grande exactitude. Les gens utilisent parfois l'imagerie de Google Earth pour la validation, ou ils utilisent l'interprétation visuelle de la même image utilisée pour la classification. Cette dernière option semble un peu circulaire — comme dans « pourquoi utiliser un classificateur en premier lieu, si vous pouvez en toute confiance attribuer des classes sur la base de l'interprétation visuelle de l'image ? ». Cependant, l'interprétation visuelle peut être tout à fait appropriée pour définir avec exactitude la couverture du sol pour un certain nombre de points de données de validation, tandis que l'interprétation visuelle d'une image entière pourrait être une tâche extrêmement laborieuse. Les mêmes considérations décrites dans les points ci-dessus s'appliquent que les données de validation soient créées à partir d'observations sur le terrain ou d'une interprétation visuelle de l'imagerie.

Une nouvelle approche intéressante pour créer des données de validation consiste à utiliser des photos géolocalisées accessibles au public, telles que celles disponibles sur Flickr ou d'autres sites où les gens partagent leurs photos. En particulier pour les villes et les sites touristiques populaires, l'Internet contient un vaste référentiel de photos géolocalisées qui peuvent être utilisées par n'importe qui comme observations sur le terrain. Un certain contrôle de la qualité est cependant nécessaire, car toutes les photos disponibles en ligne ne sont pas géolocalisées automatiquement avec le GPS (certaines sont « géolocalisées » manuellement lorsqu'elles sont mises en ligne), et la plupart des photos montrent les conditions de la couverture terrestre à un moment différent de celui où l'image de télédétection a été acquise (par exemple, en hiver ou en été).

La matrice de confusion

Une fois que vous avez créé un ensemble de données de validation auxquelles vous faites confiance, vous pouvez utiliser leur géoréférence pour les associer à la couverture terrestre correspondante cartographiée dans la classification. Vous pouvez considérer la comparaison résultante comme un tableau qui ressemble à ceci :

Tableau 5: Comparaison des estimations dérivées d'un algorithme de classification (à gauche) avec les données de validation (à droite).

Mapped land cover (estimate)	Validation data (reality)
Forest	Forest
Water	Water
Forest	Grassland
Grassland	Grassland
Grassland	Bare soil
Bare soil	Bare soil
...	...

Avec de nombreux points de données de validation, une méthode est nécessaire pour résumer toute cette information. En télédétection, la méthode utilisée universellement, et ce depuis des décennies, s'appelle la matrice de confusion (également appelée « matrice d'erreur » ou « tableau de contingence »). En utilisant les quatre classes énumérées dans l'exemple ci-dessus, le cadre de la matrice de confusion ressemblerait à ceci :

Tableau 6: Exemple de matrice de confusion, avant l'introduction des valeurs.

		Validation data				Total
		Class	Forest	Water	Grassland	
Classification	Forest					
	Water					
	Grassland					
	Bare soil					
	Total					

Lisez le long des lignes, chaque ligne vous indique ce que sont en réalité les pixels classés dans une classe donnée

selon les données de validation. Lisez le long des colonnes, chaque colonne vous indique ce que les données de validation connues pour être une classe donnée ont été classées comme. Par exemple :

Tableau 7: Matrice de confusion avec les valeurs saisies.

		Validation data				Total
		Forest	Water	Grassland	Bare soil	
Classification	Class					
	Forest	56	0	4	2	62
	Water	1	67	1	0	69
	Grassland	5	0	34	7	46
	Bare soil	2	0	9	42	53
Total		64	67	48	51	230

En lisant les lignes, le tableau ci-dessus vous indique que 56 pixels classés comme « forêt » étaient également considérés comme « forêt » dans les données de validation, que 0 pixel classé comme « forêt » était considéré comme « eau » dans les données de validation, que 4 pixels classés comme « forêt » étaient considérés comme « prairie » dans les données de validation et que 2 pixels classés comme « forêt » étaient considérés comme « sol nu » dans les données de validation, soit un total de 62 pixels classés comme forêt. Et ainsi de suite.

Exactitude de l'utilisateur, du producteur, et globale

En utilisant les informations de la matrice de confusion, nous pouvons trouver des réponses à des questions raisonnables concernant l'exactitude de la carte d'occupation du sol produite par la classification. Il y a trois types de questions typiquement posées et répondues avec la matrice de confusion.

L'exactitude de l'utilisateur répond à une question du type suivant : « Si je dispose de votre carte et que je me rends sur un pixel que votre carte indique comme étant de classe "x", quelle est la probabilité que j'y trouve effectivement la classe "x" ? ». En utilisant l'exemple des « prairies » du tableau ci-dessus, nous pouvons voir qu'un total de 46 pixels classés comme « prairies » ont été vérifiés par rapport aux données de validation. Sur ces 46 pixels, 34 ont été considérés comme des « prairies » dans les données de validation. En d'autres termes, 34 pixels, sur les 46 pixels classés comme « prairie », sont réellement des « prairies ». 34 sur 46 est 74 %, donc l'exactitude de l'utilisateur de la classification, pour la classe « prairie », est de 74 %. L'exactitude de l'utilisateur varie selon les classes, car certaines classes sont plus faciles à distinguer des autres que d'autres. Les caractéristiques de l'eau ont tendance à être faciles à cartographier parce qu'elles sont sombres et bleuâtres et que peu de caractéristiques terrestres leur ressemblent. Dans l'exemple ci-dessus, l'exactitude de l'utilisateur pour la classe « eau » est de 67 sur 69, soit 97 %.

L'exactitude du producteur répond à une question du type suivant : « Si une zone est effectivement de

la classe “x”, quelle est la probabilité qu’elle ait également été cartographiée en tant que telle ? En utilisant à nouveau l’exemple des “prairies”, nous constatons qu’un total de 48 points de données de validation ont été considérés comme des “prairies”, et que 34 d’entre eux ont également été classés comme tels. 34 sur 48 est 71 %, donc l’exactitude du producteur pour la classe “prairie” est 71 %.

Alors que les exactitudes de l’utilisateur et du producteur se concentrent sur les classes individuelles, *l’exactitude globale* répond à la question suivante : » quelle proportion de la carte est correctement classée. « , ce qui peut souvent être interprété simplement comme » quelle est l’exactitude de la carte ? « . En regardant les valeurs dans la diagonale de la matrice de confusion dans l’exemple ci-dessus, nous pouvons voir que 56 pixels étaient considérés comme “forêt” dans les données de validation et avaient également été classés comme “forêt”, et nous pouvons observer des nombres similaires de 67 pour “eau”, 34 pour “prairie”, et 42 pour “sol nu”. Le total est de $56 + 67 + 34 + 42 = 199$, sur un total de 230 pixels dans l’ensemble de données de validation. 199 sur 230 est 87 %, donc sur la base des données de validation nous estimons que 87 % de la carte est correctement classée.

L’exactitude globale doit être rapportée avec précaution, comme l’illustre l’exemple suivant. Imaginons que l’image utilisée pour la classification couvre une zone côtière et que la trajectoire suborbitale du satellite ait été un peu au large, de sorte que 80 % de l’image est couverte par de l’eau. Les 20 % restants de l’image étaient recouverts de “sol nu” ou de “végétation”. Si vous réfléchiez cette distribution inégale dans la création de vos données de validation, 80 % de vos données de validation se trouveraient au-dessus de l’eau, et comme l’eau est relativement facile à distinguer d’autres types de surface, votre matrice de confusion pourrait ressembler à ceci :

Tableau 8: Un autre exemple de matrice de confusion. Notez le nombre très différent de points de validation utilisés pour chaque classe.

		Validation data			Total
		Water	Vegetation	Bare soil	
Classification	Class				
	Water	82	1	0	83
	Vegetation	0	12	2	14
	Bare soil	0	2	9	11
	Total	82	15	11	108

Alors que les exactitudes de l’utilisateur et du producteur pour la “végétation” et le “sol nu” ne sont pas impressionnantes dans ce scénario, comme prévu, l’« eau » a été classée presque parfaitement. La prédominance des pixels d’« eau » influence le calcul de l’exactitude globale, qui est de $82 + 12 + 9 = 103$ sur 108, soit une exactitude globale de 95 %. Si le but de la carte est de déterminer où se trouve le littoral, ou quelque chose d’autre qui ne nécessite que de séparer l’eau de la terre, cela peut être acceptable comme estimation de la

qualité de la carte. Mais si vous avez réalisé la carte pour une agence gouvernementale locale chargée de surveiller la végétation côtière, l'exactitude globale de 95 % peut faussement donner l'idée que la carte doit être utilisée en toute confiance à cette fin, ce qui nécessite en grande partie de séparer la « végétation » du « sol nu ».

En général, tant que vous rapportez a) comment vous avez produit la carte, b) comment vous avez produit les données de validation, et c) l'ensemble de la matrice de confusion ainsi que toute mesure de précision supplémentaire qui en découle, un lecteur intelligent sera en mesure de juger si la carte est appropriée pour un but donné, ou non.

Évaluation de l'exactitude des classifications lorsque vous essayez de cartographier une seule chose.

Un cas particulier d'évaluation de la précision se présente lorsque vous faites une carte d'un seul type d'objet, comme des maisons, des piscines, etc. Si cette situation est encore rare en télédétection, elle devient de plus en plus nécessaire avec l'analyse d'images basée sur l'objet, qui est un moyen efficace de cartographier des types d'objets spécifiques. Nous allons utiliser les piscines comme exemple. Imaginez que vous ayez créé un flux de travail d'analyse d'image basé sur l'objet qui prend une image satellite haute résolution et tente de détecter toutes les piscines dans la zone couverte par l'image. Le produit de ce flux de travail est un ensemble de polygones qui délimitent toutes les piscines identifiées sur l'image. De même, vos données de validation sont constituées d'un ensemble de polygones décrivant toutes les piscines identifiées manuellement sur l'image, pour une petite partie de l'image que vous utilisez pour la validation. Vous avez donc maintenant deux ensembles de polygones à comparer, l'un étant votre « estimation », l'autre la « réalité ». Votre matrice de confusion peut être configurée de la manière suivante (explication ci-dessous) :

Tableau 9: Matrice de confusion utilisée pour calculer le score F1 lorsque la précision est évaluée pour la détection d'un seul type d'objets.

		Validation data		Total
		Presence	Absence	
Classification	Class			
	Presence	TP	FP	Precision = $TP / (TP + FP)$
Classification	Absence	FN	TN	
	Total	Recall = $TP / (TP + FN)$		

Dans ce tableau, « presence » indique la présence d'une piscine (dans l'un ou l'autre des ensembles de données) et « absence » indique l'absence d'une piscine (également dans l'un ou l'autre des ensembles de données). TP est le nombre de Vrais Positifs (en anglais, « True Presences ») — piscines qui existent dans les données de

validation, et qui ont été correctement identifiées dans votre carte comme étant des piscines. FP est le nombre de Fausses Présences (en anglais, « False Presences ») — objets identifiés dans votre carte comme étant des piscines, mais qui sont en réalité autre chose. FN est le nombre de faux négatifs — des piscines qui existent en réalité, mais que votre carte n'a pas détectées. Notez que dans ce tableau, il n'y a pas de Vrais Négatifs (objets qui ne sont en réalité pas des piscines, et qui n'ont pas non plus été identifiés dans l'image comme des piscines). Ceci a été omis, car, dans le cas d'une analyse d'image qui ne vise à trouver qu'une seule chose, aucun autre objet n'est identifié dans l'image ni dans les données de validation.

L'objectif d'une bonne analyse d'image est, bien sûr, d'obtenir un grand nombre de vraies présences, un petit nombre de fausses présences et un petit nombre de faux négatifs. Pour quantifier la réussite de l'analyse d'image, la valeur généralement calculée est appelée le score F1, qui se calcule comme suit : $F1 = (2 * \text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})$. Le score F1 a l'avantage d'avoir des valeurs allant de 0 (pire) à 1 (meilleur), ce qui le rend facile à interpréter.

Évaluation de l'exactitude pour les variables continues

Lorsqu'il s'agit de variables continues, la comparaison entre les « estimations » et la « réalité » ne consiste plus à vérifier si elles sont identiques ou non, car lorsqu'elles sont mesurées avec suffisamment de détails, elles ne le sont jamais. Par exemple, vous pouvez avoir cartographié un pixel comme ayant une température de surface de 31,546 °C alors que votre observation de terrain correspondante indique qu'elle est en réalité de 31,543 °C. Bien que les deux valeurs ne soient pas identiques, vous ne voudriez probablement pas que ce pixel soit simplement considéré comme « non conforme ». Ce que nous devons faire, c'est donner aux utilisateurs de la carte une idée de la *différence typique* entre l'estimation cartographiée et la réalité.

Création de données de validation

Comme pour l'évaluation de la précision de la classification, vous avez besoin d'un ensemble de données de validation considérées comme représentatives de la réalité. Celles-ci proviennent presque universellement de mesures sur le terrain, et il est important de se rappeler que, comme pour l'évaluation de la précision des classifications, les données de validation doivent être comparables aux mesures dérivées de votre image. La question de la résolution spatiale peut être particulièrement problématique ici, car il est difficile d'effectuer des mesures précises sur de grandes surfaces avec la plupart des équipements de terrain. Prenons le cas de la température de surface, qui est généralement mesurée avec un thermomètre infrarouge portatif (Figure 62).

Un thermomètre infrarouge (comme les thermomètres auriculaires utilisés pour vérifier si vous avez de la fièvre ou non) mesure le rayonnement provenant d'une petite zone circulaire de la surface de la Terre, là où le thermomètre est pointé. Les satellites mesurent essentiellement le même rayonnement et estiment la température de la même manière qu'un thermomètre infrarouge portatif, sauf qu'ils intègrent le rayonnement mesuré sur une plus grande surface (100 x 100 mètres, dans le cas de TIRS sur Landsat 8) alors que la version

portative mesure une très petite surface (par exemple 0,5 x 0,5 mètre). Étant donné que la température de surface varie notamment en fonction de l'humidité du sol, de la couverture végétale et des matériaux de surface, prendre une mesure sur le terrain qui soit représentatif de la température de surface moyenne dans une zone de 100 x 100 m est un défi presque insurmontable dans des environnements hétérogènes. Une façon de contourner ce problème est de créer toutes vos données de validation dans des zones aussi homogènes que possible, une autre est de prendre plusieurs mesures et d'utiliser les valeurs moyennes sur chaque site pour produire un seul point de données de validation. Quelle que soit votre stratégie, il est essentiel de faire correspondre les données de validation à la résolution spatiale de l'image pour produire une évaluation de précision importante.



Figure 62: Thermomètre infrarouge portatif du type utilisé pour mesurer la température de surface d'une petite zone. Isometric Medical Digital Non-Contact Infrared Thermometer Temperature Measurement Device blue white par Jernej Furman, Flickr, CC BY 2.0.

Une fois que vous disposez d'un bon ensemble de données de validation, le tableau qui sert de base à l'évaluation de la précision est un tableau qui compare des valeurs numériques, comme dans l'exemple ci-dessous de la biomasse ligneuse par ha dans une forêt :

Tableau 10: Exemple de tableau utilisé pour comparer les estimations cartographiées d'une variable continue avec les données de validation provenant d'observations sur le terrain :

Mapped biomass, tons per ha (estimate)	Observed biomass, tons per ha (reality)
10.1	9.2
5.7	4.8
3.2	4.0
6.7	6.6
7.8	7.1
9.3	9.1
...	...

L'une des premières choses que nous pouvons quantifier avec ce type de données est de savoir si les estimations cartographiées sont typiquement sous-estimées ou surestimées, lorsqu'elles sont comparées aux données de validation. Si c'est le cas, le flux de travail (algorithme) qui a produit les estimations est considéré comme biaisé, et il peut être utile de le réviser. Le biais peut être calculé comme *l'erreur moyenne*, c'est-à-dire la valeur moyenne de l'estimation moins la validation. Nous pouvons le calculer en mettant à jour le tableau 10 pour obtenir le tableau 11 :

Tableau 11: Le biais est calculé comme l'erreur moyenne des estimations.

Mapped biomass, tons per ha (estimate)	Observed biomass, tons per ha (reality)	Error (estimate – reality)
10.1	9.2	0.9
5.7	4.8	0.9
3.2	4.0	-0.8
6.7	6.6	0.1
7.8	7.1	0.7
9.3	9.1	0.2
...	...	
Mean Error:		0.33

Dans le tableau 11, nous pouvons voir que les estimations ont une erreur moyenne de 0,33. En d'autres termes, en moyenne, les estimations de la biomasse dans notre carte surestiment la réalité de 0,33 tonne par ha. Cela peut ou non justifier le perfectionnement de l'algorithme utilisé pour produire les estimations, mais c'est au

moins quelque chose qui mérite d'être signalé aux utilisateurs potentiels de la carte de la biomasse qui contient les surestimations.

Un autre élément crucial à quantifier est le degré d'erreur typique des estimations. C'est important, car on peut imaginer avoir une carte non biaisée qui surestime fortement la biomasse dans une partie et la sous-estime fortement ailleurs. Une telle carte, non biaisée, serait toujours fautive partout ! La manière la plus simple de rendre compte de cette situation est d'utiliser *l'erreur absolue moyenne*, comme le montre le tableau 12 :

Tableau 12: L'erreur absolue moyenne (en Anglais « Mean Absolute Error », MAE) est généralement indiquée pour montrer la « précision » de la carte.

Mapped biomass, tons per ha (estimate)	Observed biomass, tons per ha (reality)	Error (estimate – reality)	Absolute Error
10.1	9.2	0.9	0.9
5.7	4.8	0.9	0.9
3.2	4.0	-0.8	0.8
6.7	6.6	0.1	0.1
7.8	7.1	0.7	0.7
9.3	9.1	0.2	0.2
...	...		
Mean Absolute Error:			0.6

La MAE vous indique la différence à laquelle vous devez vous attendre entre l'estimation et la réalité, mais pas si cette différence prend la forme d'une surestimation ou d'une sous-estimation. Comme la plupart des algorithmes utilisés pour cartographier les variables continues sont très efficaces pour minimiser les biais, la MAE est souvent utilisée pour comparer différentes approches de cartographie d'une variable continue, en considérant comme l'approche qui conduit à la MAE le plus faible comme la « meilleure ».

Une autre mesure de la « précision » pour les variables continues est l'erreur quadratique moyenne (en anglais « Root Mean Square Error », RMSE), qui est calculée comme dans le tableau 13 :

Table 13: Calcul de l'erreur quadratique moyenne.

Mapped biomass, tons per ha (estimate)	Observed biomass, tons per ha (reality)	Error (estimate – reality)	Squared Error
10.1	9.2	0.9	0.81
5.7	4.8	0.9	0.81
3.2	4.0	-0.8	0.64
6.7	6.6	0.1	0.01
7.8	7.1	0.7	0.49
9.3	9.1	0.2	0.04
...	...		
Mean Squared Error:			0.47
Root Mean Squared Error:			0.68

La RMSE est plus grande que la MAE (ou égale à celle-ci, dans des cas extrêmement rares), car elle pèse plus lourdement sur les grosses erreurs dans son calcul. Les algorithmes qui évitent les erreurs importantes sont donc avantagés lorsqu'ils sont comparés à l'aide de la RMSE, tandis que les algorithmes qui ont souvent des erreurs très faibles, mais occasionnellement des erreurs très importantes ne le sont pas.

Enfin, une autre valeur généralement rapportée lors de la cartographie d'une variable continue est le coefficient de détermination, ou R^2 , qui indique dans quelle mesure la variance des valeurs des données de validation (« réalité ») est prévisible à partir des estimations.

Bien qu'il soit important de comprendre pourquoi et comment rendre compte de l'exactitude de toute carte que vous produisez, en termes pratiques, c'est plus facile que ce que l'on pourrait croire dans les sections ci-dessus. La plupart des logiciels utilisés pour le traitement d'images ont des fonctions intégrées pour l'évaluation de l'exactitude, et offrent même souvent une fonctionnalité pour vous aider à créer les données de validation. Si le logiciel que vous utilisez ne vous permet pas de calculer une mesure spécifique (par exemple la MAE, ou le score F1), il vous aidera probablement à exporter les données nécessaires vers un fichier facilement lisible dans Excel, R, ou tout autre logiciel que vous connaissez et que vous pouvez utiliser pour effectuer le reste des calculs manuellement.

8.

DÉTECTION DES CHANGEMENTS

De nombreux changements environnementaux se produisent à des échelles temporelles et spatiales qui rendent leur étude difficile et potentiellement coûteuse. Parmi les exemples évidents, citons certaines des manifestations du changement climatique : l'augmentation des températures et la fonte des glaces dans l'Arctique, ainsi que les changements subtils de la végétation dans tout le Canada, pour n'en citer que quelques-uns. Surveiller ces phénomènes exclusivement à l'aide de mesures sur le terrain serait incroyablement coûteux, ou fortement biaisé vers les zones peuplées et, lorsque des mesures historiques ne sont pas disponibles, impossible. Les données de télédétection viennent ici à la rescousse, car la photographie aérienne et l'imagerie satellitaire sont régulièrement stockées et cataloguées et elles peuvent donc fonctionner comme des données du passé et du présent — pourvu que nous puissions en extraire l'information nécessaire. Imaginez simplement répondre à ces questions en toute confiance sans avoir accès aux données de télédétection :

- La superficie annuelle brûlée par les feux de forêt est-elle en hausse ou en baisse au Canada ?
- Quelle partie du Canada se réchauffe le plus rapidement ?
- Où y a-t-il encore de la glace pluriannuelle dans l'Arctique, et à quelle vitesse cette superficie se diminue-t-elle ?

Avant de commencer à utiliser les données de télédétection pour détecter les changements environnementaux, il est utile d'examiner ce que nous entendons par là — qu'est-ce que ***un changement environnemental*** ? L'environnement change en tout temps — il neige dehors au moment où j'écris ces lignes alors qu'il n'y avait pas de neige il y a deux jours. C'est un type de changement environnemental, mais pas celui que la plupart des gens veulent détecter et cartographier avec des données de télédétection. Néanmoins, c'est exactement ce type de changement qui est visible sur les images de la Terre prises depuis l'espace, donc, que nous le voulions ou non, nous détectons beaucoup de changements non pertinents avec les données de télédétection. Pour mieux réfléchir aux changements que nous voulons détecter et à ceux que nous ne voulons pas détecter, une distinction utile à faire est celle entre les changements à court terme/instantanés et les changements à long terme/graduels. Parfois, nous souhaitons à trouver le changement à court terme, alors qu'à d'autres moments, lorsque nous sommes plus concernés par les tendances graduelles à long terme, nous considérons qu'il s'agit de bruit.

Changement instantané

La télédétection peut être efficace pour détecter un changement instantané, c'est-à-dire un changement qui se produit entre une acquisition d'image et la suivante. Ce changement n'est pas nécessairement instantané au sens strict du terme, mais pour des raisons pratiques de télédétection, un changement qui se produit entre deux acquisitions d'images peut être considéré comme « instantané ». La détection d'un tel changement se fait souvent assez facilement par l'interprétation visuelle de deux images, en les comparant côte à côte pour identifier facilement les zones de changement. Un exemple de ce type de comparaison est présenté à la figure 63, qui montre les changements survenus dans une zone urbaine de Taïwan entre février (à gauche) et décembre 2002 (à droite) 2002. La végétation (zones rouges) a clairement changé à certains endroits, de même que certaines autres zones. De nombreux algorithmes de traitement d'images ont également été développés à cette fin, afin d'automatiser le processus de recherche des zones qui ont changé ; nous les examinerons plus loin dans ce chapitre.



Figure 63: Exemple de comparaison côte à côte de deux images montrant la ville de Zanzibar en 1995 (à gauche) et en 2020 (à droite). L'expansion urbaine qui s'est produite dans la période entre l'acquisition des deux images peut être identifiée visuellement ou algorithmiquement. Par Anders Knudby, CC BY 4.0.

Changement progressif

La comparaison visuelle ou autre entre images peut être efficace pour mettre en évidence des changements

abrupts évidents tels que ceux causés par le développement urbain, les glissements de terrain, les feux de forêt, etc. Cependant, les changements plus subtils, tels que les variations annuelles ou décennales de l'humidité du sol ou de la santé de la végétation, ou les changements dans l'épaisseur du manteau neigeux en hiver, nécessitent souvent une approche différente de la détection des changements, une approche qui repose sur l'analyse quantitative des changements observés sur de nombreuses images. C'est particulièrement le cas lorsque des changements individuels d'une image à l'autre peuvent masquer la tendance à long terme plus graduelle en raison de schémas météorologiques ou saisonniers. Par exemple, pour étudier comment le changement climatique affecte les modèles de végétation dans l'Arctique canadien, la comparaison de deux images fournit, au mieux, un instantané des années dont ces images sont issues, et au pire, elle nous dit simplement qu'il y a plus de végétation en 2002 (août) qu'en 1994 (février), ce qui n'a rien à voir avec un changement entre les deux années, mais est plutôt une fonction du moment de l'année où chaque image a été acquise. Quoiqu'il en soit, ces comparaisons d'image à image ne nous renseignent pas beaucoup sur la tendance à long terme. Pour détecter de telles tendances, il faut plutôt quantifier une mesure de la variable d'intérêt (par exemple, la densité ou la santé de la végétation) à intervalles réguliers sur une longue période. Un exemple est fourni par la figure 64, dans laquelle l'étendue de la glace de mer arctique a été quantifiée sur une base quasi quotidienne pour la période 1979 – 2018.

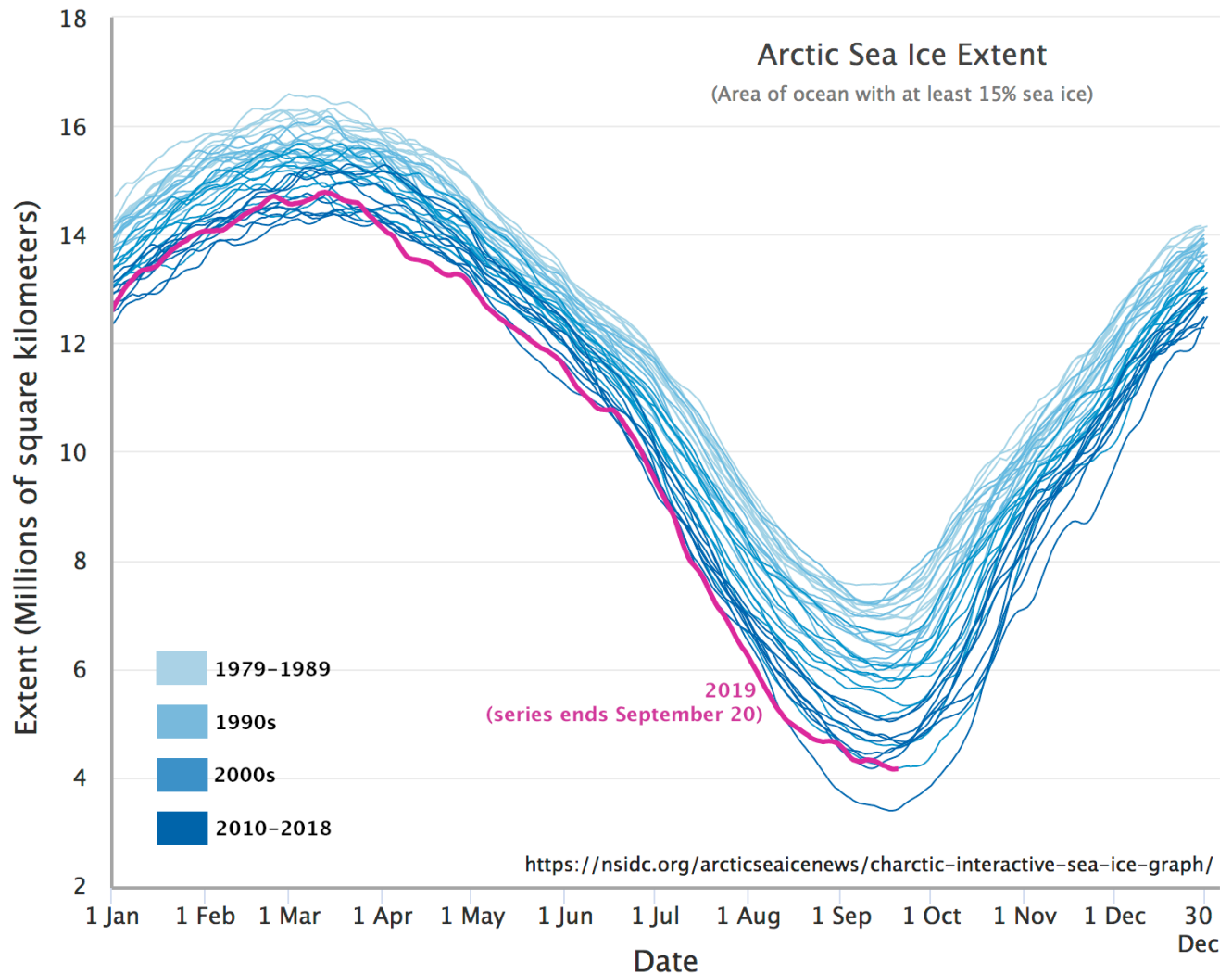


Figure 64: Estimation quasi quotidienne de l'étendue de la glace de mer arctique. Arctic Sea Ice Extent par M. Scott, National Snow & Ice Data Center (NSIDC), NSIDC Use and Copyright.

La méthodologie spécifique utilisée pour détecter un changement graduel à long terme dépend entièrement du type de changement en question (par exemple, la végétation par rapport au changement de la glace de mer), il est donc difficile de fournir des détails sur la façon de procéder pour détecter un changement dans ce cas. Cependant, l'une des principales préoccupations de ce type de « détection des tendances » est que l'estimation (par exemple, l'étendue de la glace de mer) doit être relativement non biaisée tout au long de la période d'observation. En d'autres termes, l'étendue de la glace de mer ne doit pas être sous-estimée ou surestimée au cours de la période 1979-1989, et ne doit pas non plus être sous-estimée ou surestimée au cours des années 2010 ni à aucun moment entre les deux. Ceci est important pour éviter de « détecter » des tendances qui sont dues à un biais dans l'acquisition ou le traitement des données, mais qui n'existent pas en réalité.

Séparer le bruit du changement instantané réel

Lorsque l'on compare des images dans le but de détecter des changements entre elles, le principal défi consiste à détecter un changement environnemental réel tout en évitant de détecter un changement qui ne s'est pas produit. Comme c'est toujours le cas dans la classification binaire (dont ceci est un exemple), il existe quatre combinaisons possibles de la réalité (changement ou pas de changement) et de l'estimation (changement détecté vs pas de changement détecté, tableau 14) :

Tableau 14: Une classification binaire (qui implique deux options catégorielles) peut être décrite à l'aide d'un tableau comme celui-ci. Pour tout système de traitement d'images, le défi de la classification binaire est d'optimiser le nombre/taux de vrais positifs et de vrais négatifs.

	Change detected	No change detected
Change	True positive (TP)	False negative (FN)
No change	False positive (FP)	True negative (TN)

L'objectif est donc d'optimiser le taux de détection des vrais positifs et des vrais négatifs. La plupart des algorithmes de détection des changements fonctionnent pixel par pixel, ce qui signifie qu'il faut détecter correctement les pixels qui ont réellement changé sans « détecter » incorrectement les changements dans les pixels qui n'ont pas changé. Pour ce faire, nous avons besoin d'un moyen de séparer trois situations différentes, chacune pouvant se produire dans un pixel donné :

- Pas de changement : le pixel est exactement le même dans chaque image.
- Bruit : Le pixel semble différent dans les deux images, mais la différence est suffisamment faible pour qu'elle soit probablement causée par des facteurs non liés à un **changement environnemental réel**. Il peut s'agir de différences dans les conditions atmosphériques entre les deux images, de bruit aléatoire dans les images, d'un géoréférencement imparfait et d'autres problèmes.
- Changement réel : Un **changement environnemental réel** s'est produit dans le pixel, et il se manifeste par une différence substantielle entre l'aspect du pixel dans la première image et dans la deuxième image.

La première situation, « aucun changement », se produit très rarement, car il existe un bruit inhérent au processus de création de l'image, et il est peu probable que ce bruit soit identique entre deux images différentes. Le véritable défi consiste donc à séparer les situations 2) et 3). Pour ce faire, deux principes sont en jeu.

Réduction du bruit

Tout d'abord, il est important d'éliminer autant de sources de bruit que possible avant de comparer les données des deux images. Voici quelques stratégies souvent employées à cette fin :

- Utilisez les « dates anniversaires ». Choisissez deux images qui ont été capturées à la même date (ou presque) dans des années différentes. C'est un bon moyen d'éliminer les grandes différences entre chaque image liée aux changements saisonniers de l'humidité du sol, de l'état de la végétation, de la couverture neigeuse et d'autres facteurs environnementaux qui changent avec les saisons.
- Utilisez des images capturées par le même capteur. Étant donné qu'aucun capteur n'est parfaitement calibré, l'utilisation d'images provenant de deux capteurs différents peut potentiellement introduire une différence qui est liée au capteur plutôt qu'à l'environnement. Par exemple, si la radiance dans le spectre 500-600 nm est légèrement surestimée par un capteur et légèrement sous-estimée par un autre, la comparaison des images entre les deux peut « détecter un changement » là où il n'y en a pas.
- Utilisez des images capturées dans des conditions atmosphériques similaires. Cela est difficile, car nous ne disposons souvent pas d'informations précises sur les charges d'aérosols, la vapeur d'eau, la vitesse du vent, etc., mais évitez absolument de comparer des images présentant des différences évidentes de brume, de visibilité et d'autres facteurs atmosphériques visibles.
- Comparez les images en fonction de leur réflectance de surface plutôt que de la radiance ou de la réflectance du TOA. En effet, la réflectance de la surface est un attribut physique de la surface et est, au moins en principe, indépendant de l'état atmosphérique et de l'illumination.
- Pour les capteurs inclinables, comparez si possible des images prises à peu près à la même géométrie.

Identification d'un seuil

Une fois que le plus grand nombre possible de sources de bruit a été éliminé et que les images ont été converties en réflectance de surface pour la comparaison, il y aura toujours un certain bruit qui fera que les valeurs de réflectance de surface dans les deux images différeront légèrement entre les images, même pour les pixels qui n'ont pas subi le type de changement que nous voulons détecter. Pour séparer les situations 2) et 3) ci-dessus, il est donc important de considérer ce que signifie un changement environnemental « réel » et de définir un seuil en dessous duquel tout « changement » observé dans la comparaison d'images est considéré comme du « bruit » plutôt que comme un « changement réel ». Par exemple, imaginez que vous étudiez les changements dans la végétation et que vous disposez de deux images de la même forêt. Le changement réel qui s'est produit entre ces deux images est qu'une feuille est tombée d'un des arbres de la forêt. Il s'agit d'un changement réel observable dans l'environnement, mais qui ne justifie en aucun cas l'appellation de « déforestation » ! Une condition préalable à la séparation des situations 2) et 3) ci-dessus est donc de définir le degré de changement nécessaire pour que vous puissiez considérer la zone comme « modifiée ». À moins de disposer d'observations

sur le terrain (et ce n'est généralement pas le cas, car le changement est détecté à rebours dans le temps et il est impossible d'aller chercher des données sur le terrain dans le passé), il faut définir le seuil sur la base des images elles-mêmes, ce qui est généralement un processus interactif et subjectif.

Méthodes de détection d'un changement instantané

Certaines des approches les plus simples pour détecter *l'ampleur* d'un changement sont basées sur les « mathématiques de l'image ».

Différence de bande

Par exemple, l'ampleur du changement peut être définie comme la différence entre les valeurs de réflectance de surface dans la bande 1 des deux images, plus la différence dans la bande 2, etc. :

$$\text{Changement} = (B_{1I2} - B_{1I1}) + (B_{2I2} - B_{2I1}) \dots$$

Ratio de bande

On peut aussi utiliser des ratios à la place des différences :

$$\text{Changement} = (B_{1I2}/B_{1I1}) + (B_{2I2}/B_{2I1}) \dots$$

Distance euclidienne

Une méthode plus couramment utilisée consiste à calculer la « distance euclidienne », en utilisant chaque bande comme une dimension et chaque image comme un point. Imaginez que vous tracez les réflectances de surface de chaque image comme des points dans un système de coordonnées. La distance euclidienne entre eux serait alors égale à :

$$\text{Changement} = (B_{1I2} - B_{1I1})^2 + (B_{2I2} - B_{2I1})^2 \dots$$

Pour chacune de ces approches, afin de séparer le bruit des véritables changements environnementaux, une valeur seuil devra être définie. Les pixels qui n'ont subi aucun changement ou un changement inférieur à la valeur seuil seront alors considérés comme effectivement « inchangés », tandis que ceux qui ont subi un changement plus important seront considérés comme contenant un « changement réel ». Notez également que pour chacune de ces approches, le nombre de bandes impliquées dans le calcul a été limité à deux dans les

équations ci-dessus, mais peut être étendu pour inclure n'importe quel nombre de bandes présentes dans les images.

Méthodes pour détecter et attribuer un changement instantané

Les trois approches ci-dessus peuvent être efficaces pour *détecter* les changements, mais elles présentent également l'inconvénient de nous renseigner très peu sur *le type* de changement qui s'est produit dans un pixel donné. Si la détection d'un changement est une excellente première étape, il n'est pas facile de fonder des décisions de gestion sur cette information. Il est donc utile d'en savoir un peu plus sur le type de changement qui a été détecté, ce que l'on appelle *l'attribution du changement*.

Vecteurs de changement

Les vecteurs de changement peuvent y contribuer. L'analyse des vecteurs de changement est une prolongation du calcul de la distance euclidienne, dans laquelle la direction du changement est également calculée. La figure 65 donne un exemple avec seulement deux dimensions. Dans l'image « avant », la valeur de réflectance de la surface dans la bande 1 est de 0,1, et la valeur NDVI est de 0,5. Dans l'image « après », ces valeurs sont passées respectivement à 0,3 et -0,2. Ceci illustre deux choses : 1) il n'est pas nécessaire d'utiliser les valeurs réelles des bandes comme entrées dans les analyses de détection des changements, et 2) la direction du changement – augmentation de la réflectance de la bande 1, diminution du NDVI, peut être trouvée.

Dans l'analyse du vecteur de changement, l'ampleur du changement est calculée comme la distance euclidienne, selon l'équation ci-dessus. La direction du changement peut être calculée en degrés (par exemple, la direction de la boussole) ou, pour faciliter l'interprétation, en directions principales : A) haut et droite, B) bas et droite, C) bas et gauche, D) haut et gauche. Si les “bandes” utilisées peuvent être interprétées de manière importantes, l'attribution du changement peut être facile. Les changements qui entraînent une diminution de l'indice de végétation impliquent une perte de végétation, tandis que ceux qui conduisent à une augmentation, par exemple, de la température de surface impliquent, eh bien, une augmentation de la température de surface (ces deux phénomènes vont souvent de pair, car la végétation contribue à maintenir une surface fraîche). Si le calcul de la distance euclidienne est facile lorsque plus de deux dimensions sont utilisées dans l'analyse du vecteur de changement — il suffit d'étendre l'équation pour qu'elle contienne davantage de termes — la direction du changement devient plus difficile à définir et les catégories sont généralement développées pour une application spécifique. Comme pour les méthodes utilisées simplement pour détecter le changement, un seuil peut être appliqué à la magnitude du vecteur de changement, en dessous duquel “aucun changement” n'est détecté.

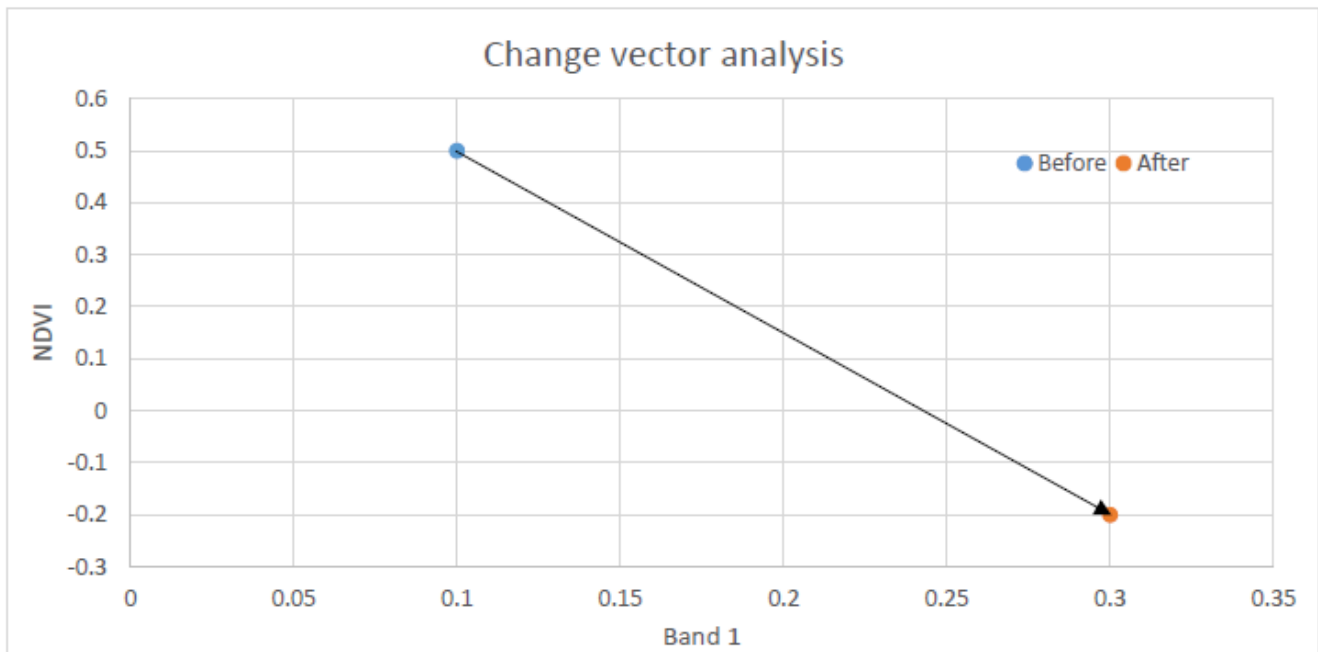


Figure 65: Exemple d'analyse du vecteur de changement avec seulement deux bandes. NDVI est l'indice de végétation par différence normalisée, décrit plus en détail dans le chapitre suivant. Anders Knudby, CC BY 4.0.

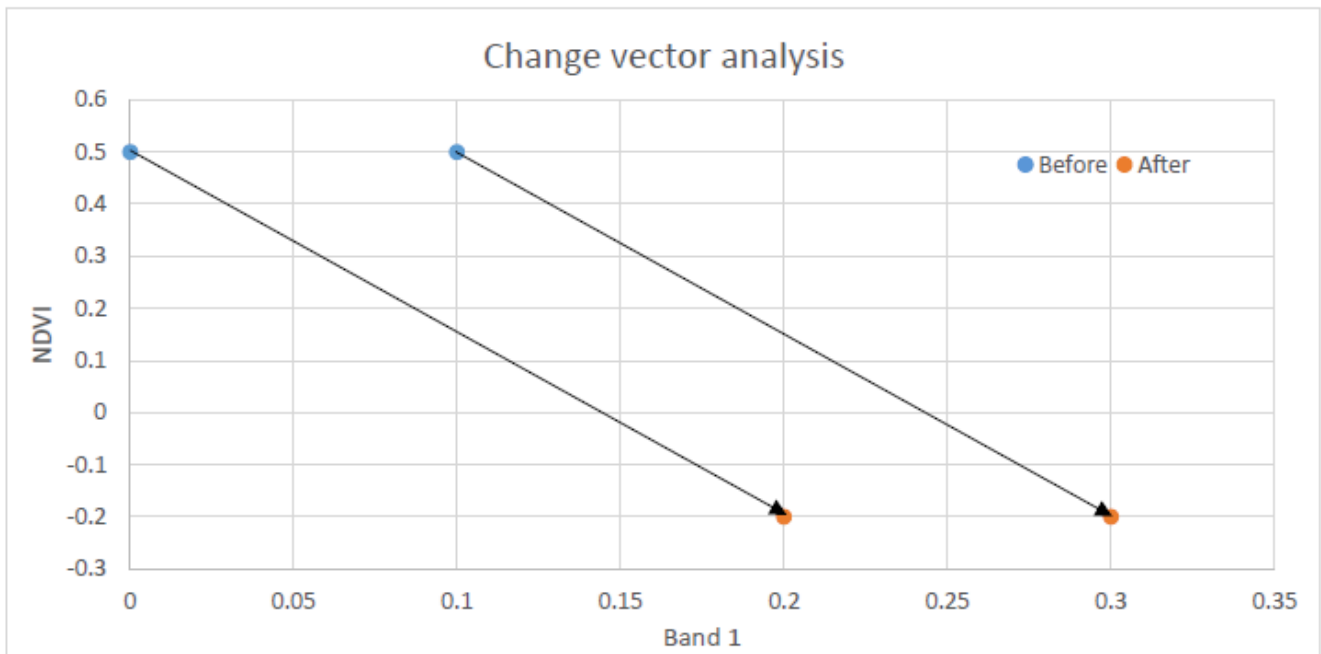


Figure 66: Un inconvénient de l'analyse des vecteurs de changement est que deux vecteurs de changement identiques peuvent représenter des types de changement différents. Anders Knudby, CC BY 4.0.

L'un des inconvénients de l'analyse du vecteur de changement est que deux changements différents peuvent avoir exactement le même vecteur de changement. Par exemple, regardez les deux flèches de la figure 66.

Bien que les deux vecteurs soient identiques — ils ont la même direction et la même magnitude — ils sont susceptibles de représenter différents types de changement, car ils commencent et se terminent à des endroits différents. Il est donc difficile, voire impossible dans la pratique, d'utiliser l'analyse des vecteurs de changement pour déterminer spécifiquement ce qu'était la surface avant et après le changement.

Détection des changements post-classification

La façon la plus simple de détecter et d'attribuer un changement dans une zone est probablement de réaliser une classification de l'occupation du sol sur l'image "avant", de réaliser une autre classification de l'occupation du sol sur l'image "après", puis de trouver les pixels qui ont été classés différemment dans les deux images. Bien que cette méthode soit simple et attrayante, et qu'elle puisse fonctionner occasionnellement, elle présente un inconvénient majeur : il est très difficile de produire des résultats précis de cette manière. La raison en est qu'aucune classification n'est parfaite, et que lorsque l'on compare deux classifications imparfaites, les erreurs se combinent. Je suggère donc fortement d'éviter cette approche.

Classification des changements

Si vous voulez vraiment savoir quelle était la zone modifiée avant et après le changement, une autre possibilité à la détection de changement post-classification consiste à combiner toutes les bandes des deux images en une seule image. Par exemple, si vous avez deux images, chacune avec six bandes, vous pouvez les « empiler » pour obtenir une seule image à 12 bandes. Cela vous permet d'effectuer une seule classification sur l'image à 12 bandes. Avec de bonnes données de terrain pour calibrer cette classification, chaque classe peut être définie en fonction de la combinaison des couvertures terrestres avant et après. La classification peut être supervisée ou non supervisée, basée sur les pixels ou sur les segments. Cette approche permet aux deux vecteurs de la figure 66 de se retrouver dans deux classes différentes, l'une pouvant correspondre au passage d'une forêt de conifères à un marais et l'autre à un passage d'une forêt de feuillus à une tourbière.

La classification des changements vous permet également de supprimer les combinaisons de classes qui sont impossibles ou hautement improbables, comme le passage de « océan » à « forêt de conifères », ou de « industrie » à « zone humide ». Décider quelles combinaisons de classes avant-après sont probables et lesquelles sont improbables dans une zone donnée requiert une certaine expertise, et les systèmes complexes de détection des changements sont souvent construits en combinant l'analyse des vecteurs de changement, la classification des changements et l'apport de ces experts.

9.

VÉGÉTATION ET FEU

Les premiers chapitres de ces notes ont largement couvert des sujets génériques de la télédétection optique passive — des questions d'importance générale et des techniques de traitement d'images qui ne sont pas spécifiques à un domaine d'application ou à un autre. Dans ce chapitre, nous allons examiner certaines techniques développées et largement utilisées spécifiquement pour la télédétection de deux choses : la végétation et le feu. Des chapitres similaires pourraient être écrits sur de nombreuses autres applications spécifiques de la télédétection — la végétation et les incendies sont utilisés ici parce que l'utilisation de l'imagerie satellitaire pour les cartographier et les surveiller est bien développée et utilisés de façon opérationnelle à l'échelle nationale et mondiale.

Végétation

Étant donné que la Terre est couverte d'une certaine forme de végétation verte, il n'est pas surprenant que de nombreux efforts de recherche aient été consacrés à l'utilisation de la télédétection pour cartographier et évaluer divers aspects de la végétation. Ceci est important pour les industries de base qui soutiennent les moyens de subsistance des humains à travers le monde, comme la foresterie et l'agriculture, ainsi que pour presque toute évaluation environnementale des zones terrestres. En fait, les recherches ont été si nombreuses que des livres entiers ont été écrits sur la télédétection de la végétation.

En termes simples, la capacité de détecter la végétation en télédétection optique passive est basée sur sa signature spectrale (Figure 67), qui est très différente des autres types de surface.

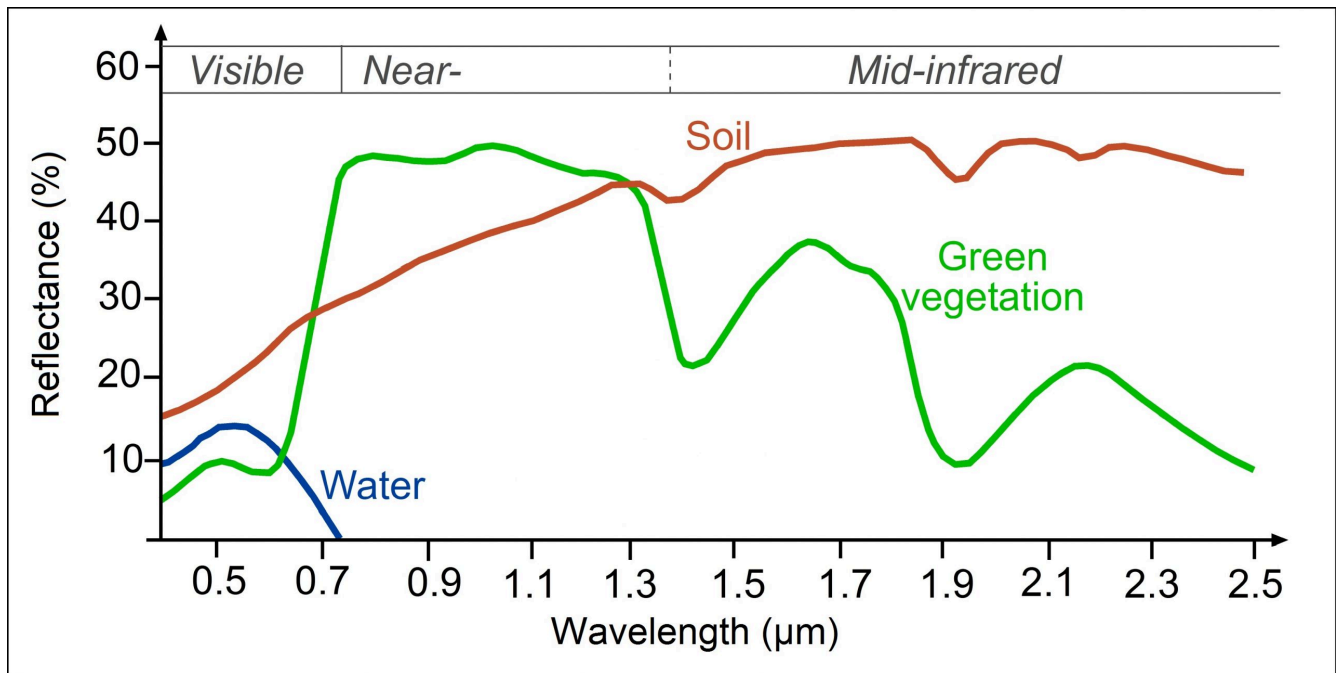


Figure 67: Signatures spectrales généralisées de la végétation verte, du sol et de l'eau. Identique à la figure 27. Reflectance of water, soil and vegetation at different wavelengths par Science Education through Earth Observation for High Schools (SEOS), CC BY-NC-SA 2.0.

Région des longueurs d'onde visibles

Une végétation verte et saine absorbe la majeure partie de la lumière visible qui la frappe et l'utilise pour se développer grâce au processus de photosynthèse. La photosynthèse est un processus biochimique relativement complexe qui commence par l'absorption d'un photon par une molécule de pigment, laquelle déclenche une transition électronique (le déplacement d'un électron d'un niveau d'énergie à un niveau supérieur) qui amorce une réaction en chaîne conduisant à la production de matière végétale à partir des composants chimiques de base que sont le CO₂ et l'eau. Une forme de chlorophylle est le pigment dominant de la quasi-totalité de la végétation. La chlorophylle absorbe très bien la lumière visible, ce qui explique pourquoi la végétation présente de faibles réflectances dans les longueurs d'onde visibles. La chlorophylle absorbe un peu mieux les longueurs d'onde bleues et rouges que les longueurs d'onde vertes, de sorte qu'une part relativement plus importante du rayonnement vert est réfléchi, ce qui donne à la végétation son aspect vert.

Région des longueurs d'onde infrarouges

Les photons entrants à des longueurs d'onde dans le proche infrarouge contiennent individuellement moins d'énergie (rappelons que l'énergie d'un photon est proportionnelle à sa fréquence, et donc inversement proportionnelle à sa longueur d'onde). Ces photons sont donc incapables de provoquer une transition

électronique ; les pigments utilisés pour la photosynthèse ne peuvent donc pas les utiliser, et en général ne les absorbent pas. D'autres parties des plantes, en particulier l'eau liquide présente dans les feuilles, absorbent ces photons, qui font tourner et étirent les liaisons chimiques dans l'eau ou dans les cellules végétales, ce qui a pour effet de chauffer le matériau absorbant. En général, les plantes n'ont pas besoin de se réchauffer et, à l'exception des longueurs d'onde autour de 1,4 μm et 1,9 μm où l'eau est un fort absorbeur, les plantes réfléchissent une grande partie du rayonnement infrarouge proche entrant.

La combinaison d'une faible réflectance dans le visible et d'une réflectance élevée dans les longueurs d'onde du proche infrarouge est la signature la plus caractéristique de la végétation utilisée en télédétection, et elle est utilisée pour une cartographie rapide de la végétation sur Terre. Il faut savoir que la plupart des zones terrestres contiennent un mélange de zones végétalisées et non végétalisées. Parfois, ce « mélange » est très unilatéral, comme en Amazonie, dans le désert du Sahara ou dans l'Arctique canadien, et parfois il est vraiment plus mélangé, comme dans la savane africaine (figure 68).

Indices de végétation

Le fait d'utiliser uniquement la réflectance dans les longueurs d'onde visibles ne nous renseigne pas sur la quantité de végétation présente dans une zone, car l'eau a également une faible réflectance dans cette région de longueurs d'onde. Et l'utilisation de la réflectance dans la région du proche infrarouge ne nous renseigne pas non plus sur la végétation, car d'autres types de surface, comme le sable brillant et la neige brillante représentés sur la figure 68, ont également une réflectance plus élevée dans le proche infrarouge. Cependant, le rapport entre les réflectances visibles et proche infrarouges est un indicateur utile de la quantité de végétation présente dans une zone, car aucun autre type de surface n'a une réflectance proche infrarouge aussi élevée et visible aussi faible que la végétation. Cette observation a conduit initialement au développement de ce que l'on appelle le rapport simple (SR) :

$$\text{SR} = \text{NIR} / \text{RED}$$

où NIR est la réflectance de la surface dans la région des longueurs d'onde du proche infrarouge (généralement autour de 700-1000 nm), et RED est la même chose pour la région des longueurs d'onde du rouge (600-700 nm).



Figure 68: Quatre zones présentant des caractéristiques de végétation assez différentes. En haut à gauche : La savane africaine contient un mélange d'arbres et de couverture herbeuse. Trees On The African Savanna par Lynn Greyling, PublicDomainPictures.net, CC0 1.0. En haut à droite : Le désert du Sahara est dans la plupart des endroits dépourvu de végétation. Sahara Desert par Azer Koçulu, Wikimedia Commons, CC0 1.0. En bas à gauche : La forêt amazonienne est couverte d'une végétation dense. Amazon Rainforest par CIFOR (Neil Palmer/CIAT), Flickr, CC BY-NC-ND 2.0. En bas à droite : L'Arctique canadien comporte de vastes zones dépourvues de végétation. Patterned ground in Canadian tundra par Raymond M. Coveney, Wikimedia Commons, CC BY-SA 3.0.

Alors que les valeurs de réflectance de la surface devraient en principe toujours être utilisées pour calculer cet indice et tout autre indice de végétation, dans la pratique, ces calculs sont souvent effectués à la place sur la base de la réflectance TOA ou même de la luminance TOA (ou même des valeurs DN !). Le rapport simple présente un inconvénient qui fait qu'il est rarement utilisé : ses valeurs approchent l'infini lorsque la réflectance rouge s'approche de zéro, ce qui rend les différences entre les grandes valeurs difficiles à interpréter. Une amélioration simple appelée « indice de végétation » par différence normalisée (NDVI), basée sur les deux mêmes mesures, est donc plus couramment utilisée :

$$\text{NDVI} = (\text{NIR} - \text{RED}) / (\text{NIR} + \text{RED})$$

Les valeurs NDVI vont d'un minimum de -1 à un maximum de +1. Les valeurs typiques pour le sol nu se situent autour de 0, pour l'eau autour de -0,2, et pour la végétation dans la gamme 0,1 – 0,7 selon la santé et la densité de la végétation. Comme le NDVI ne repose que sur deux mesures, dans le proche infrarouge et dans le

rouge, il peut être calculé avec les données des premiers capteurs multispectraux en orbite, notamment les séries Landsat (depuis 1972) et AVHRR (depuis 1978). Cela permet de visualiser facilement la couverture végétale, à petite ou grande échelle, comme celle de la figure 69.

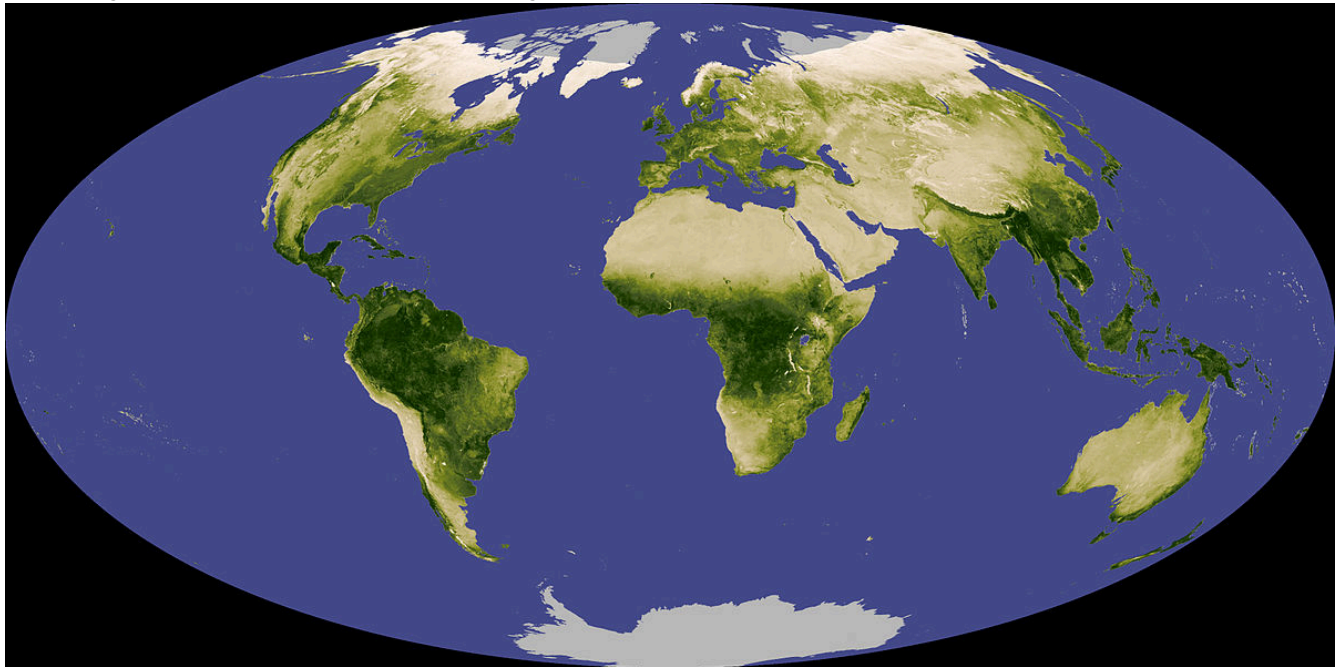


Figure 69: Composite NDVI de la Terre, montrant les différences relatives de couverture végétale. Globalndvi tmo 200711 lrg par Reto Stockli and Jesse Allen (NASA), Wikimedia Commons, public domain.

Les indices de végétation peuvent servir à autre chose qu'à produire de jolies images de la répartition de la végétation mondiale. En raison du réchauffement de la planète, qui est particulièrement prononcé dans la région arctique, les parties septentrionales du Canada subissent des changements rapides dans la distribution et les modèles de croissance de la végétation. Le suivi de ces changements dans cette vaste région très peu peuplée est basé sur les tendances à long terme de l'indice NDVI pendant les mois d'été, ce qui démontre entre autres que les régions qui subissent un réchauffement plus rapide connaissent également une croissance plus rapide de la densité de la végétation. Une autre utilisation importante des indices de végétation est la prévision du rendement des cultures : en suivant le NDVI tout au long d'une saison de croissance agricole, il est possible de quantifier le développement des cultures et de prédire statistiquement leurs rendements éventuels. L'utilisation la plus importante de ces modèles de prévision des récoltes est probablement la prévision des sécheresses, des mauvaises récoltes et des famines dans les régions du monde où les populations dépendent de l'agriculture locale pour leur approvisionnement alimentaire. C'est ce que fait systématiquement le réseau FEWS (Famine Early Warning Systems Network), une organisation basée aux États-Unis, pour les pays sujets à la famine, et ce que font également de nombreux gouvernements nationaux et organisations régionales/mondiales. Les indices de végétation ont donc littéralement le potentiel de sauver des vies – si la communauté mondiale réagit aux prévisions qu'ils génèrent.

De nombreux perfectionnements ont été apportés depuis le développement du NDVI, afin de créer des indices de végétation plus performants dans diverses conditions. Ces indices comprennent l'indice de

végétation ajusté au sol, l'indice de végétation amélioré utilisé avec les données MODIS et l'indice de végétation global utilisé avec les données MERIS. Ces indices de végétation plus récents ont été spécifiquement conçus pour être plus robustes face à des conditions environnementales variables, telles que les variations de la luminosité du sol et des conditions atmosphériques. Malgré ces améliorations, tous les indices de végétation souffrent d'un inconvénient fondamental : si leur interprétation est assez claire en termes relatifs (des valeurs plus élevées indiquent une végétation plus verte), leur relation avec les caractéristiques réelles de la végétation mesurables utilisées par des personnes extérieures au domaine de la télédétection n'est pas claire.

Un indice de végétation qui n'est pas destiné à représenter la densité de la végétation, mais plutôt les différences relatives de contenu en eau de la végétation, a été développé spécifiquement pour tirer parti de la caractéristique d'absorption de l'eau autour de $1,4 \mu\text{m}$ montrée sur la figure 67. Il s'agit de l'indice de différence d'eau normalisée (NDWI), dont l'équation générique est la suivante :

$$\text{NDWI} = (\text{NIR} - \text{SWIR}) / (\text{NIR} + \text{SWIR})$$

où SWIR désigne une mesure dans la région infrarouge à ondes courtes ($1,4 \mu\text{m}$). La végétation sèche contient moins d'eau et absorbe donc moins le rayonnement SWIR, ce qui réduit la valeur de l'indice NDWI. Pour tout indice de végétation, les bandes spécifiques utilisées comme bandes « RED », « NIR » et « SWIR » dépend du capteur en question, et tous les capteurs ne disposent pas de bandes appropriées pour le calcul de tel ou tel indice.

Cartographie des attributs de la végétation réelle

Les attributs de la végétation les plus utilisés par les personnes qui ne font pas de télédétection sont, entre autres, la biomasse aérienne, le rayonnement photosynthétique actif absorbé (APAR) et l'indice de surface foliaire (en anglais « Leaf Area Index », LAI). La biomasse aérienne peut être mesurée (en la coupant et en la pesant) en kg/ha, et présente un intérêt fondamental pour le suivi des tendances de la végétation. L'APAR est la quantité de rayonnement absorbée par la végétation (par unité de surface), et est fortement liée à la croissance de la végétation et au rendement des cultures. Le LAI est la surface (unilatérale) des feuilles divisée par la surface du sol, il s'agit donc d'une mesure sans unité de la densité des feuilles. Étant donné que les échanges (d'eau et de gaz) entre la végétation et l'atmosphère se font à travers les feuilles, le LAI peut être utilisé comme un indicateur de cet échange, qui est important pour la fonction de la végétation en tant que puits de carbone, entre autres choses. Tous ces attributs de la végétation sont donc plus directement liés à la fonction de la végétation dans l'environnement, et certainement plus directement interprétables que ne l'est un indice de végétation. Mais ils sont aussi beaucoup plus difficiles à cartographier avec des données de télédétection.

Approche empirique

Une approche simple de la cartographie du LAI est illustrée à la figure 70, dans laquelle les valeurs NDVI ont régressé par rapport au LAI mesuré sur le terrain. Comme le montre la figure 70, ce que nous mesurons dans les

données de télédétection est lié, mais pas de manière très étroite, et certainement pas de manière linéaire, au LAI réel mesuré sur le terrain. Néanmoins, les lignes de régression peuvent être utilisées pour produire une première approximation du LAI à partir de ces données de télédétection. Une approche similaire peut être adoptée pour cartographier d'autres attributs de la végétation.

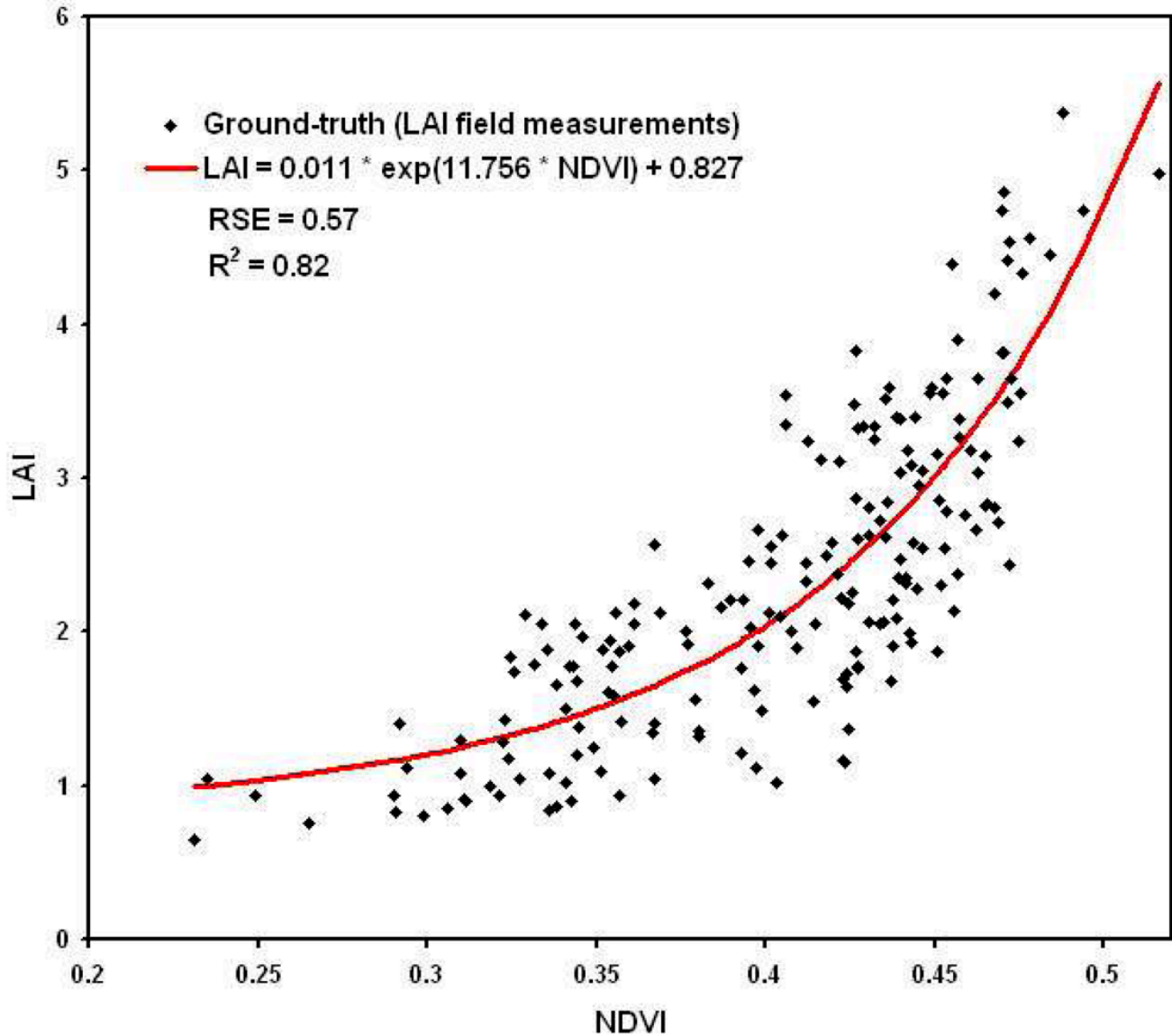


Figure 70: Relations empiriques entre quatre indices de végétation et le LAI mesuré sur le terrain. Assessment of Unmanned Aerial Vehicles Imagery for Quantitative Monitoring of Wheat Crop in Small Plots (Figure 8) par LeLong et al., CC BY 3.0.

Approche par transfert radiatif

Des approches plus sophistiquées de la cartographie des attributs de la végétation reposent sur des modèles de transfert radiatif. Les modèles de transfert radiatif construisent un monde 3D simple composé a) du sol,

b) d'une partie de la végétation au sol, c) de l'atmosphère et d) du soleil. Chacun de ces éléments possède des attributs spécifiques qui définissent la manière dont ils interagissent avec le rayonnement : Le Soleil émet une certaine quantité de rayonnement à chaque longueur d'onde modélisée, l'atmosphère absorbe et diffuse ce rayonnement dans une certaine mesure, et la végétation et le sol absorbent et reflètent le rayonnement. Une fois spécifié, le modèle peut alors être « exécuté » pour déterminer ce que serait la luminance TOA (ou tout autre paramètre de rayonnement mesurable) pour une combinaison donnée de définitions des éléments. Ces éléments peuvent ensuite être modifiés et le modèle peut être exécuté à nouveau pour voir comment le changement affecte la luminance TOA. Par exemple, différentes densités de végétation peuvent être incluses, représentant différentes valeurs de LAI, chacune produisant une combinaison unique de luminances TOA dans différentes longueurs d'onde. Pour obtenir une image solide de la façon dont le LAI influence la luminance TOA, les autres éléments doivent également être variés, de sorte que l'influence de la couleur du sol et des constituants atmosphériques est également prise en compte dans le processus. Plus important encore, il faut également tenir compte de la variation de la végétation en dehors du LAI. Les feuilles individuelles peuvent avoir des signatures spectrales différentes en fonction de leur composition pigmentaire, les feuilles de formes et d'orientations différentes interagissent différemment avec le rayonnement solaire, et la structure de la canopée exerce également sa propre influence (par exemple, les feuilles sont-elles réparties de manière aléatoire, uniforme ou en touffes ?). En faisant varier tous ces paramètres et en exécutant chaque fois le modèle de transfert radiatif pour déterminer ce que leur combinaison spécifique donnerait au sommet de l'atmosphère, vous pouvez établir une table de consultation avec a) les conditions environnementales, y compris le paramètre d'intérêt (par exemple LAI) et b) la luminance TOA (spectrale) résultante. Lors de la cartographie du LAI à l'aide d'une image satellite, la luminance TOA observée de chaque pixel peut alors être mise en correspondance avec l'entrée de la table à laquelle elle correspond le mieux afin de déterminer la valeur probable du LAI dans ce pixel.

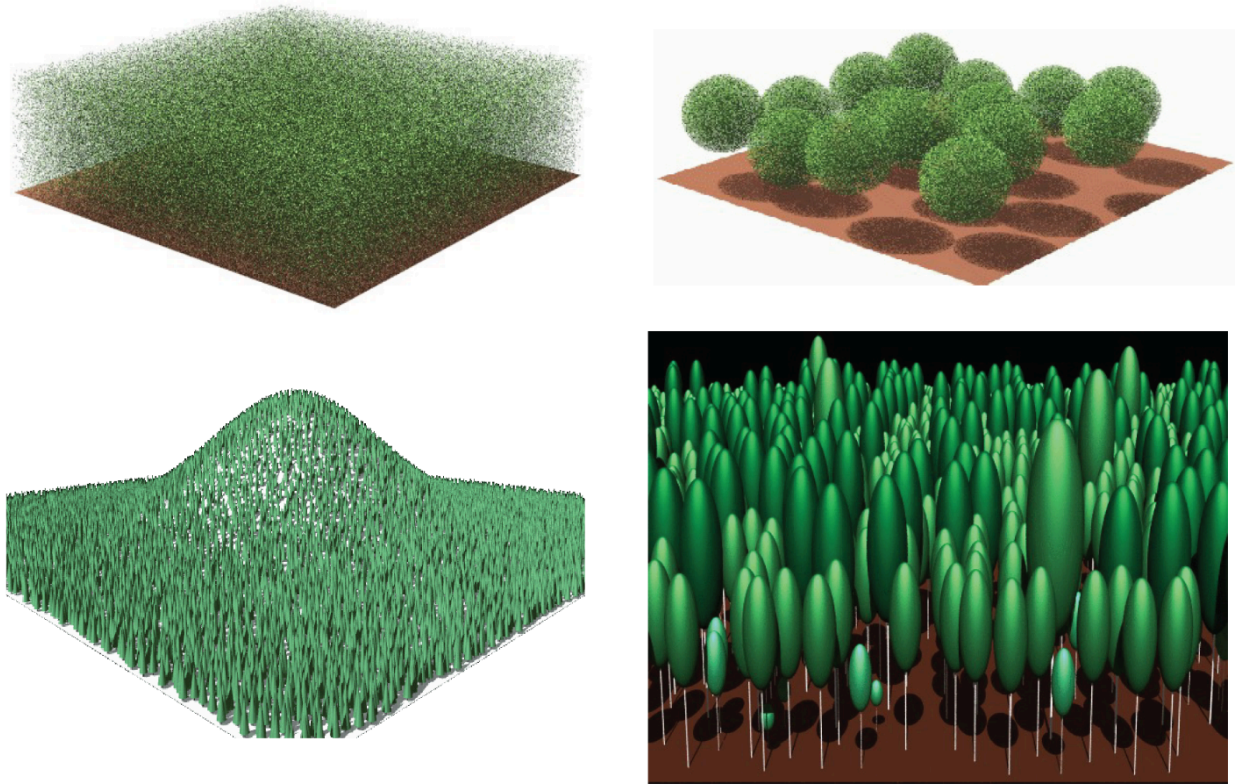


Figure 71: Exemple d'un modèle de transfert radiatif de la végétation. Modifié à partir de Accelerated RAPID Model Using Heterogeneous Porous Objects (Figure 7) par Huaguo Huang, CC BY 4.0.

Feu

Un autre sous-domaine de la télédétection concerne la cartographie des incendies. Un feu actif est relativement facile à détecter parce qu'il est chaud, et la chaleur peut être détectée par des capteurs thermiques, et parce qu'il crée de la fumée, et la fumée est visible sur les images satellites. La détection des feux actifs à l'aide de données satellitaires est si bien développée qu'un « produit feu actif » quasi instantané est réalisé chaque fois qu'un capteur MODIS ou VIIRS passe au-dessus d'une zone.

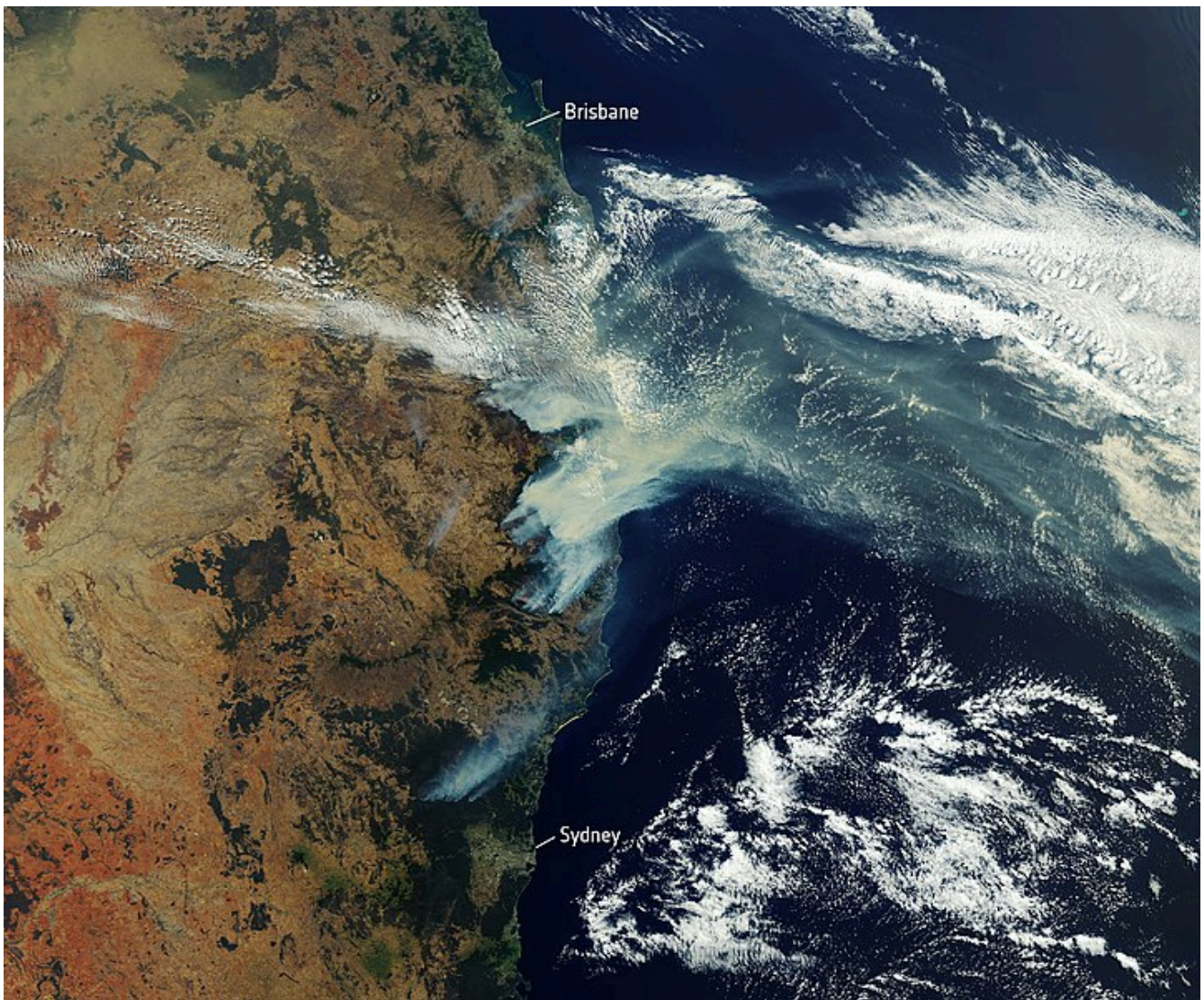


Figure 72: Exemple de détection active des incendies depuis l'espace pendant les feux de forêt australiens de 2019. Satellite image of bushfire smoke over Eastern Australia par l'ESA, Flickr, CC BY-SA 2.0.

Une fois l'incendie terminé, il peut être important d'évaluer la superficie brûlée et l'intensité du feu. Cela a des répercussions importantes sur la régénération de la végétation et du biote associé dans la région, et cela peut également nous aider à répondre à des questions importantes liées au changement climatique, comme l'évolution de la fréquence, de la durée, de la taille et de l'intensité des incendies de forêt à mesure que le climat se réchauffe. Il existe différentes façons d'évaluer ces paramètres, mais l'approche la plus couramment utilisée consiste à quantifier ce que l'on appelle le rapport de brûlage normalisée (NBR). Le NBR est calculé comme suit :

$$\text{NBR} = (\text{NIR} - \text{SWIR}) / (\text{NIR} + \text{SWIR})$$

ce qui, me direz-vous, est identique à la définition de NDWI ! Et c'est proche, mais le terme SWIR couvre une gamme assez large de longueurs d'onde, et alors que pour NDWI la définition originale de « SWIR » se référait à des longueurs d'onde autour de 1,24 μm , pour NBR « SWIR » a été spécifiquement

conçu pour la bande 7 sur Landsat 4 et 5, qui est centré autour de $2,2 \mu\text{m}$. Néanmoins, le NBR n'est en réalité qu'un indice de végétation, et bien qu'il ne soit pas le seul indice que l'on puisse utiliser pour cartographier les impacts du feu, ni même nécessairement le meilleur, il est le plus couramment utilisé. Pour l'utiliser afin de quantifier la gravité d'un incendie, il faut comparer les valeurs de NBR dans deux images, l'une avant et l'autre après l'incendie, ce qui conduit au calcul de δNBR :

$$\delta\text{NBR} = \text{NBR}_{\text{avant}} - \text{NBR}_{\text{après}}$$

Pour minimiser les effets d'autres facteurs environnementaux sur ce calcul, il est idéal d'utiliser des « dates anniversaires » pour calculer le δNBR , bien que cela ne soit pas nécessaire si deux images sont disponibles pour une petite période couvrant juste avant et juste après l'incendie (Figure 73).

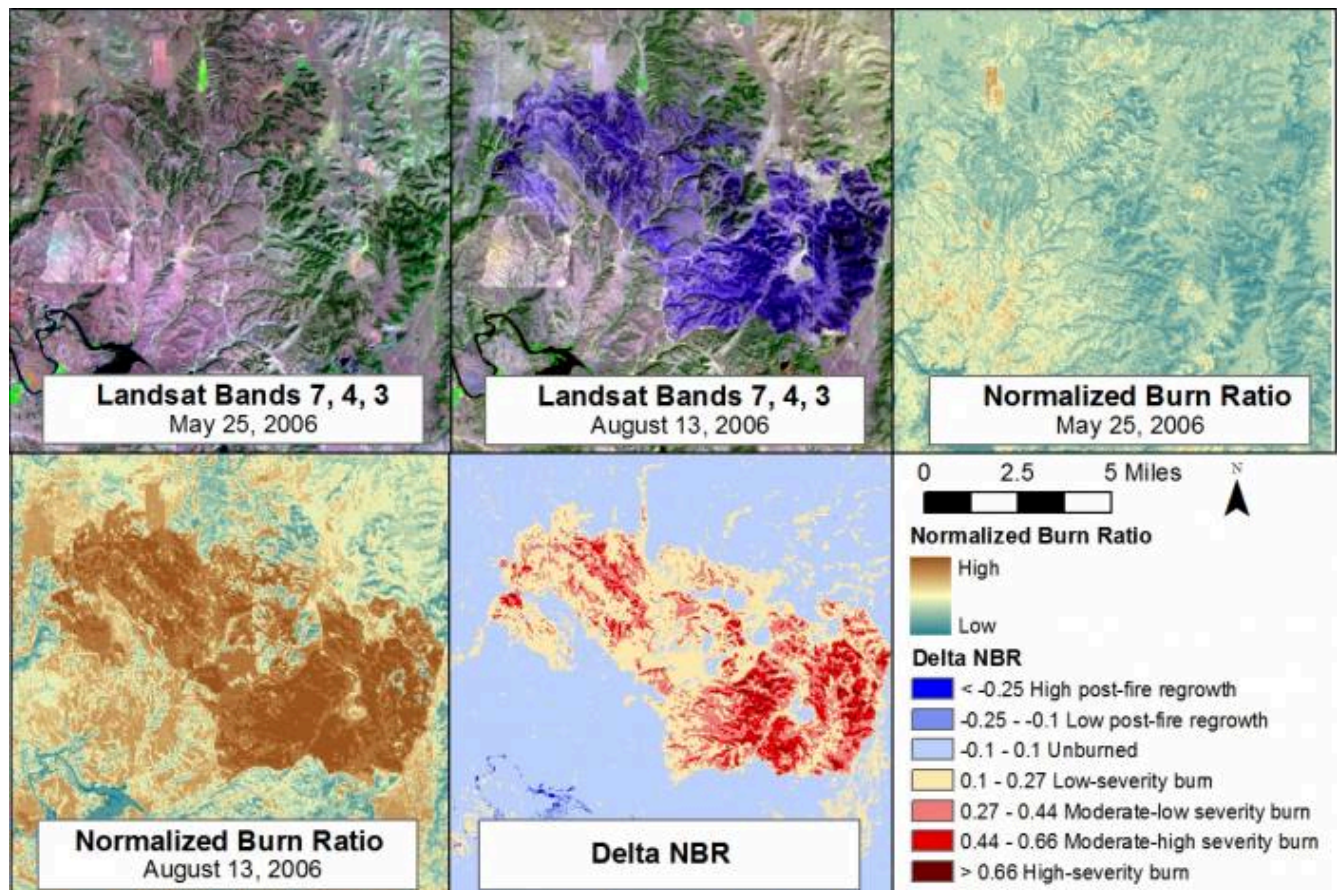


Figure 73: Exemple d'une image avant incendie (en haut à gauche), une image après incendie (en haut au milieu), NBR avant incendie (en haut à droite), NBR après incendie (en bas à gauche), et δNBR (en bas au milieu). Alors que l'impact de l'incendie est évident dans l'imagerie originale, l'utilisation de δNBR permet de souligner la différence entre les deux images. Example of normalized burn ratio from a fire at Camp Gurnsey, Wyoming in 2006 par Jason Karl, The Landscape Toolbox, CC0 1.0.

Parce que NBR est calculé comme une différence normalisée, ses valeurs varient entre -1 et +1, donc δNBR varie, en théorie, de -2 à +2, avec des valeurs plus élevées indiquant une plus grande sévérité du feu. Comme pour les autres indices de végétation, les valeurs extrêmes sont cependant rares, de sorte que même des valeurs de 0,2-0,3 indiquent des impacts d'incendie relativement sévères.



Figure 74: Un exemple de végétation et de feu, qui peuvent tous deux être bien cartographiés par la télédétection. Forest Fire, par YIvers, Pixabay, Pixabay License.

10.

PHOTOGRAMMÉTRIE ET STRUCTURE-FROM-MOTION

Toutes les applications des données de télédétection couvertes dans les chapitres précédents ont été basées soit sur des images individuelles (par exemple, la classification), soit sur des séries temporelles d'images (par exemple, la détection de changements). Cependant, lorsque deux images (ou plus) de la même zone sont acquises, en succession rapide et sous différents angles, il est possible d'extraire des informations géométriques sur la zone cartographiée, tout comme nos yeux et notre cerveau extraient des informations 3D sur le monde qui nous entoure lorsque nous le regardons avec nos deux yeux. C'est le sous-domaine de la télédétection appelé photogrammétrie.

Photogrammétrie traditionnelle (image unique et stéréo)

La photogrammétrie consiste à utiliser des informations sur la géométrie de l'image (distance entre le capteur et l'objet, position et orientation du capteur, et géométrie interne du capteur) pour en déduire des informations sur la position, la forme et la taille des objets vus sur l'image. La photogrammétrie est utilisée depuis les premières utilisations de la photographie aérienne, et l'idée de base dans ce domaine est de permettre la mesure des distances dans l'image. Ainsi, la photogrammétrie vous permet de déterminer la longueur d'un segment de route, la distance entre un point A et un point B et la taille réelle de votre patinoire locale (en supposant que vous en ayez une image bien calibrée). Vous pouvez imaginer que vous prenez une photographie aérienne et que vous la manipulez de manière à pouvoir la superposer dans une SIG ou sur Google Earth. Une fois les coordonnées géographiques associées à l'image, la mesure des distances devient possible.

Image verticale unique, terrain plat

Le scénario le plus simple possible en photogrammétrie est que vous avez une image produite par un capteur pointant directement vers la Terre, et que l'image contient un terrain parfaitement plat. Si nous supposons que l'image a été prise par un capteur situé soit sur un avion volant à très haute altitude, soit sur un satellite, et que l'image ne couvre qu'une très petite zone, nous pouvons faire l'hypothèse réductrice que chaque partie de l'image a été imagée exactement depuis le zénith, comme ceci, et que par conséquent la distance entre le capteur et la surface de la Terre est la même sur toute l'image (Figure 75).

Dans ce cas, l'échelle de l'image est égale au rapport entre la distance focale de la caméra (f , ou une mesure équivalente pour les types de capteurs autres que les caméras à sténopé) et la distance entre le capteur et le sol (H) :

$$\text{Échelle} = f / H$$

Si vous connaissez l'échelle de votre image, vous pouvez alors mesurer les distances entre les objets du monde réel en mesurant leur distance dans l'image et en divisant par l'échelle (notez que l'échelle est une fraction dont le numérateur est 1, donc diviser par l'échelle revient à multiplier par le dénominateur).

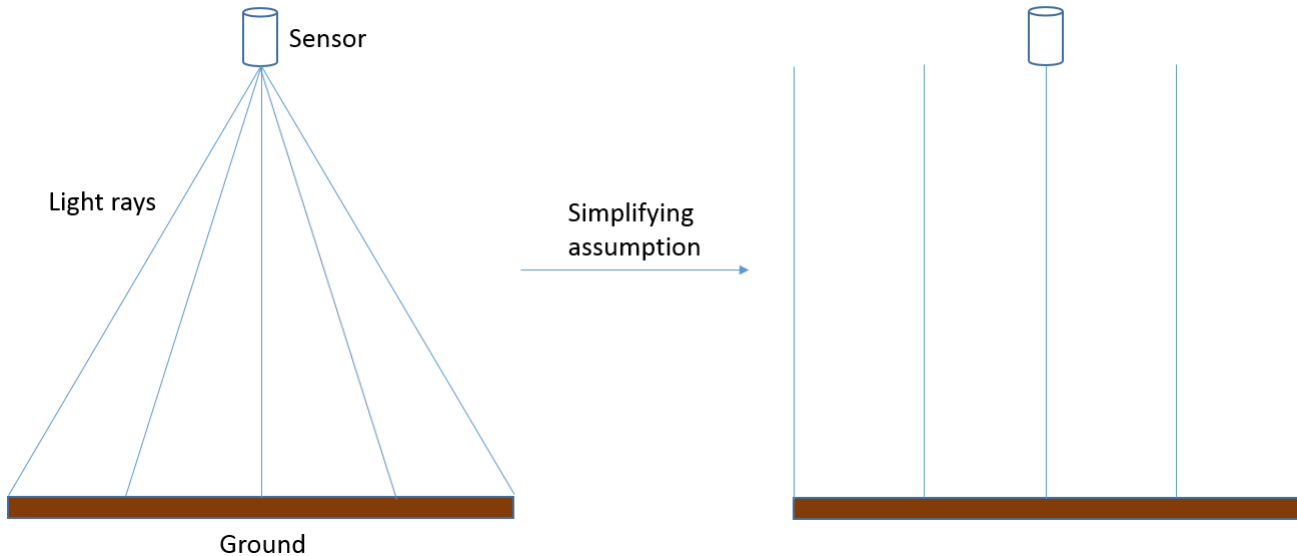


Figure 75: En réalité, une image est créée à partir de la lumière qui atteint le capteur à différents angles depuis différentes parties de la surface de la Terre. Pour les besoins actuels, nous pouvons supposer que ces angles ne sont en fait pas différents et que, par conséquent, la distance entre le capteur et la surface de la Terre est la même sur toute l'image. En réalité, cette hypothèse n'est pas correcte, et l'échelle/résolution spatiale change en fait du centre de l'image vers les bords de l'image. Anders Knudby, CC BY 4.0.

Cette définition de l'échelle n'est vraiment pertinente que pour la photographie aérienne analogique, où l'image est imprimée sur papier et où la relation entre la longueur des éléments de l'image et la longueur de ces mêmes éléments dans le monde réel est fixe. Pour la photographie aérienne numérique, ou lorsque l'imagerie satellitaire est utilisée en photogrammétrie, l'« échelle » de l'image dépend du niveau de zoom, que l'utilisateur peut faire varier, de sorte que la mesure la plus pertinente du détail spatial est la résolution spatiale que nous avons déjà examinée dans les chapitres précédents. La résolution spatiale est définie comme la longueur sur le terrain d'un côté d'un pixel, et elle ne change pas en fonction du niveau de zoom (ce qui change lorsque vous faites un zoom avant ou arrière est simplement le grain avec lequel l'image apparaît sur votre écran). Nous pouvons toutefois définir une résolution spatiale « native » d'une manière similaire à celle utilisée pour l'imagerie analogique, en nous basant sur la distance focale (f), la taille des éléments de détection individuels sur le capteur de la caméra (d) et la distance entre le capteur et le sol (H) :

$$\text{Résolution} = (f \cdot d) / H$$

Image verticale unique, terrain vallonné

La situation se complique lorsqu'une image est prise sur un terrain qui ne peut raisonnablement pas être considéré comme parfaitement plat, ou lorsque la distance entre le capteur et le sol est suffisamment petite, ou la zone couverte par l'image suffisamment grande, pour que la distance réelle entre le capteur et le sol change sensiblement entre les différentes parties de l'image :

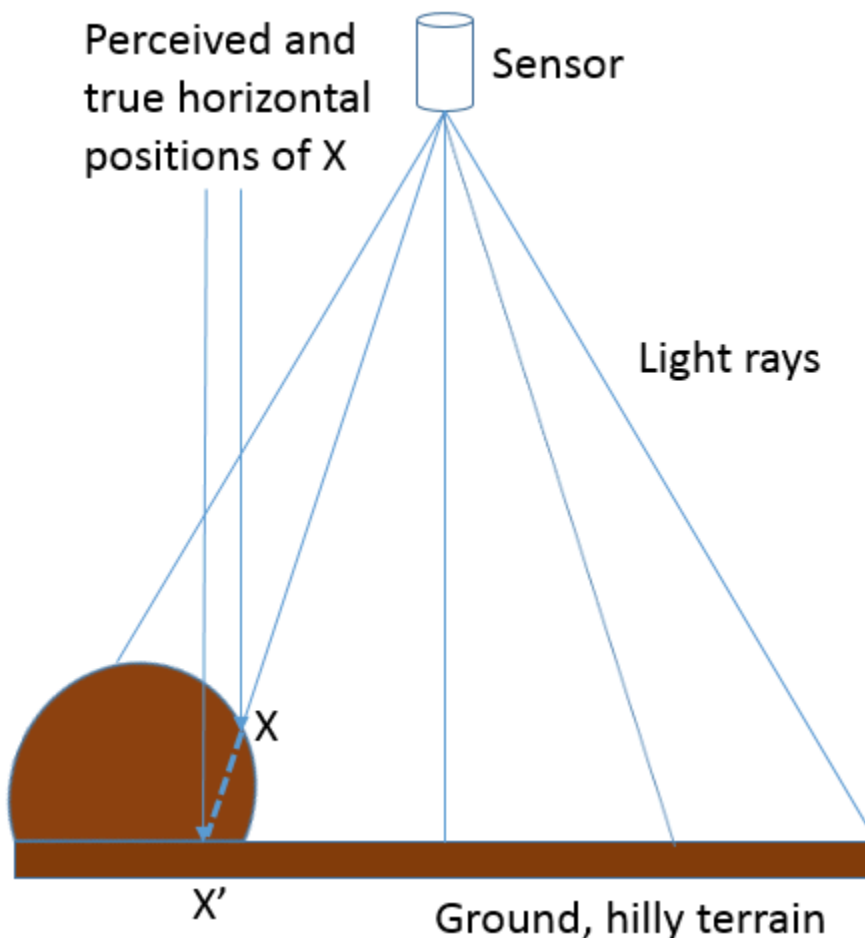


Figure 76: Lorsqu'une image est prise plus près de la surface de la Terre, ou sur un terrain plus accidenté, la distance entre le capteur et la surface de la Terre ne peut pas être considérée comme constante sur toute l'image. Si l'on n'en tient pas compte, cela entraîne un déplacement vertical et horizontal des objets imagés. Anders Knudby, CC BY 4.0.

Les terrains accidentés posent un certain nombre de problèmes en photogrammétrie, le plus important étant le déplacement vertical et horizontal des objets imagés. Comme l'illustre la figure 76, si l'on suppose que le point X se trouve sur la surface plane (et non sur la colline où il se trouve réellement), sa position horizontale sera censée se situer en X », car c'est là que le « rayon lumineux » coupe le sol plat. En réalité, bien sûr, sa position horizontale n'est pas à X ». Si le terrain de la zone imagée est connu, par exemple grâce à un modèle numérique d'élévation de haute qualité, il peut être pris en compte lors du géoréférencement de l'image et l'intersection correcte entre le « rayon lumineux » et le sol (sur la colline) peut être identifiée. En revanche, s'il

n'existe pas de modèle numérique d'élévation, l'effet du terrain — connu sous le nom de déplacement du relief — ne peut pas être explicitement pris en compte, même s'il est évident dans l'imagerie résultante.

Imagerie stéréo

Jusqu'à il y a dix ans, l'utilisation la plus avancée de la photogrammétrie consistait probablement à utiliser un couple stéréoscopique (deux images couvrant la même zone à partir de points de vue différents) pour dériver directement un modèle numérique d'élévation à partir de l'imagerie, ce qui permettait également de mesurer la hauteur d'objets tels que des bâtiments, des arbres, des voitures, etc. Le principe de base de la stéréoscopie, comme on appelle cette technique, est la reproduction de ce que vous faites tout le temps avec vos deux yeux : en observant un objet ou un paysage sous deux angles différents, il est possible d'en reconstruire un modèle tridimensionnel. Bien qu'il existe un art sophistiqué pour réaliser cette technique avec une grande précision (un art que nous ne pouvons pas explorer dans le cadre de ce cours), nous allons passer en revue l'idée de base ici.

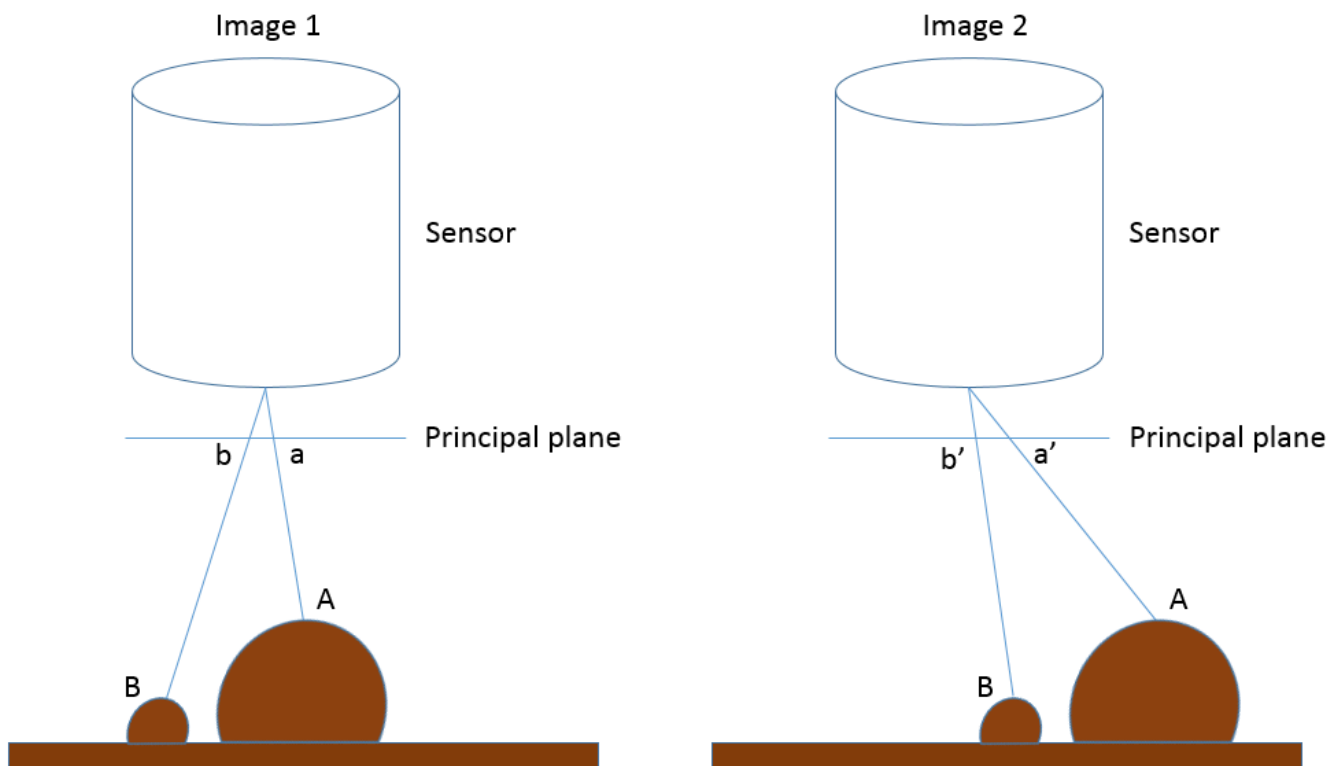


Figure 77: Illustration de la façon dont les positions horizontales apparentes des points A et B changent entre les deux images qui forment une vue stéréo de la zone. La taille du capteur a été fortement exagérée pour montrer l'emplacement de chaque point sur le plan principal. Anders Knudby, CC BY 4.0.

Dans la figure 77, l'image 1 et l'image 2 représentent une vue stéréo du paysage qui contient un terrain plat et deux collines, avec les points A et B. Sur le côté gauche de la figure (image 1), a et b indiquent les emplacements

de chaque point sur le plan principal de cette image, ce qui correspond à l'emplacement des points dans l'image résultante. Sur le côté droit de la figure, la caméra s'est déplacée vers la gauche (en suivant la trajectoire de vol de l'avion ou le mouvement du satellite sur son orbite), et là encore, b' et a' désignent les emplacements où chaque point serait enregistré sur le plan principal de cette image. Il est important de noter qu' a est différent de a' parce que le capteur s'est déplacé entre les deux images, et de la même manière, b est différent de b' . Plus précisément, comme l'illustre la figure 78, la distance entre a et a' est plus grande que la distance entre b et b' , parce que A est plus proche du capteur que B . Ces distances, connues sous le nom de parallaxe, sont celles qui transmettent l'information sur la distance entre le capteur et chaque point, et donc sur l'altitude à laquelle chaque point est situé. Pour calculer la hauteur de A par rapport au niveau du sol plat, au niveau relatif de la mer ou à une référence verticale, ou encore mieux, pour obtenir un modèle numérique d'élévation de toute la zone, il faut positionner soigneusement les deux images (généralement en localisant de nombreux éléments de surface pouvant être reconnus sur les deux images), connaître la position et l'orientation du capteur au moment de l'acquisition des deux images (généralement à l'aide d'informations basées sur le GPS), et enfin connaître la géométrie interne du capteur (généralement, ces informations sont fournies par le fabricant du capteur).

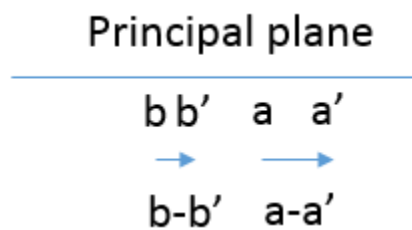


Figure 78: Mouvement de a à a' , et de b à b' , entre l'image 1 et l'image 2 de la figure 77. Anders Knudby, CC BY 4.0.

Les mathématiques appliquées à la stéréoscopie sont assez compliquées, mais en pratique, même un novice peut produire des modèles numériques d'élévation utiles à partir d'images stéréoscopiques avec un logiciel commercial. Bien qu'il existe des logiciels spécialisés en photogrammétrie, les grands progiciels commerciaux de télédétection comprennent généralement un « module de photogrammétrie » qui peut vous aider à produire des modèles numériques d'élévation, même si vous n'êtes pas un expert en photogrammétrie. Dans ce qui précède, nous avons supposé que l'imagerie était orientée au nadir (c'est-à-dire que le capteur était pointé directement vers la Terre pour chaque image), car les concepts deviennent plus compliqués lorsque ce n'est pas le cas, mais des solutions mathématiques existent également pour l'imagerie oblique. En fait, l'un des attributs utiles de plusieurs satellites modernes équipés de capteurs d'images est leur capacité à s'incliner, ce qui permet à deux images d'être capturées par le même capteur au cours de la même orbite. Lorsque le satellite s'approche de sa cible, il peut être pointé vers l'avant pour capturer une image, et plus tard sur l'orbite, il peut être pointé vers l'arrière pour capturer à nouveau la même zone sous un angle différent.

Structure-from-Motion (photogrammétrie multi-image)

Comme il arrive parfois dans le domaine de la télédétection, des techniques développées initialement à d'autres

fins (par exemple l'imagerie médicale, l'apprentissage automatique et la vision par ordinateur) présentent un potentiel pour le traitement de l'imagerie de télédétection, et parfois elles finissent par révolutionner une partie de la discipline. C'est le cas de la technique connue sous le nom de Structure-from-Motion (SfM, oui le « f » est censé être minuscule), dont le développement a été influencé par la photogrammétrie, mais a été largement basé dans la communauté de la vision par ordinateur/intelligence artificielle. L'idée de base de la SfM peut être expliquée par un autre parallèlement avec la vision humaine : si la stéréoscopie reproduit la capacité de votre système oculocérébral à utiliser deux vues d'une zone acquises sous des angles différents pour en déduire des informations en 3D, la SfM reproduit votre capacité à fermer un œil et à courir autour d'une zone, en déduisant des informations en 3D à mesure que vous vous déplacez (n'essayez pas cela en classe, mais imaginez comment cela pourrait fonctionner !) En d'autres termes, la SfM est capable de reconstruire un modèle 3D d'une zone à partir de plusieurs images superposées prises avec la même caméra sous différents angles (ou avec des caméras différentes, mais cela devient plus compliqué et est rarement pertinent pour les applications de télédétection).

Il y a plusieurs raisons pour lesquelles la SfM a rapidement révolutionné le domaine de la photogrammétrie, au point que la stéréoscopie traditionnelle pourrait rapidement appartenir au passé. Certaines des différences importantes entre les deux sont énumérées ci-dessous :

- La stéréoscopie traditionnelle (à deux images) exige que la géométrie interne de la caméra soit bien caractérisée. Bien qu'il existe des processus établis pour déterminer cette géométrie, ils exigent un niveau d'expertise moyen et obligent l'utilisateur à calibrer la caméra. La SfM effectue ce calibrage de la caméra en utilisant les mêmes images que celles utilisées pour cartographier la zone d'intérêt.
- Dans le même ordre d'idées, la stéréoscopie traditionnelle exige que la position et l'orientation (le pointage) de la caméra soient connues avec une grande précision pour chaque image. La SfM bénéficie d'informations approximatives sur certains ou tous ces paramètres, mais n'en a pas strictement besoin et est capable de les déduire très précisément des images elles-mêmes. Comme de nombreux appareils photo modernes (comme ceux de votre téléphone) possèdent des récepteurs GPS internes qui fournissent des informations approximatives sur la localisation, les appareils photo grand public peuvent être utilisés pour dériver des modèles numériques d'élévation remarquablement précis.
- Pour les deux techniques, si la position des caméras n'est pas connue, le modèle 3D reconstruit à partir des images peut être géoréférencé à l'aide de points imagés dont la position est connue. De cette façon, les deux techniques sont similaires.
- L'évolution rapide des drones a permis à des personnes ordinaires de disposer d'une plate-forme permettant de déplacer un appareil photo grand public d'une manière qui était auparavant réservée aux personnes ayant accès à des avions pilotés. Les images prises par des caméras grand public basées sur des drones sont exactement le type d'entrée nécessaire dans un flux de travail SfM, alors qu'elles ne fonctionnent pas bien dans le flux de travail traditionnel de stéréoscopie.

Comme nous l'avons mentionné plus haut, SfM utilise plusieurs photos qui se chevauchent pour produire un modèle 3D de la zone imagée sur les photos. Cette opération s'effectue en une série d'étapes, comme indiquée ci-dessous et illustrée à la figure 79 :

Étape 1) identifier les mêmes caractéristiques dans plusieurs images. Pour pouvoir identifier les mêmes caractéristiques dans des photos prises sous différents angles, à différentes distances, avec des appareils photo qui ne sont pas toujours orientés de la même manière, éventuellement dans des conditions d'éclairage différentes, voire dans des cas extrêmes avec des appareils photo différents, SfM commence par identifier les caractéristiques à l'aide de l'algorithme SIFT. Ces caractéristiques sont généralement celles qui se distinguent de leur environnement, comme les coins ou les bords, ou les objets qui sont plus sombres ou plus clairs que leur environnement. Bien que de nombreuses caractéristiques de ce type soient faciles à identifier pour les humains, et que d'autres ne le soient pas, l'algorithme SIFT est spécifiquement conçu pour permettre l'identification des mêmes caractéristiques malgré les difficultés décrites ci-dessus. Cela lui permet de fonctionner particulièrement bien dans le contexte de la SfM.

Étape 2) utiliser ces caractéristiques pour dériver les informations de la caméra intérieure et extérieure ainsi que les positions 3D des caractéristiques. Une fois que plusieurs caractéristiques ont été identifiées dans plusieurs images se chevauchant, un ensemble d'équations est résolu pour minimiser l'erreur de position relative globale des points dans toutes les images. Par exemple, considérons les trois images suivantes (ou plus), dans lesquelles les mêmes points multiples ont été identifiés (représentés par des points de couleur) :

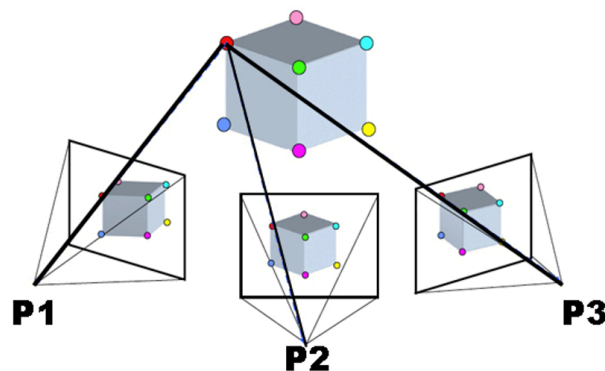


Figure 79: Exemple d'ajustement des blocs de regroupement dans SfM. Sfm1 par Maiteng, Wikimedia Commons, CC BY-SA

4.0.

Sachant que le point rouge de P1 est le même que le point rouge de P2 et P3 et de toutes les autres caméras concernées, et de même pour les autres points, on peut déduire la position relative des deux points et les emplacements et orientations des caméras (ainsi que les paramètres optiques impliqués dans le processus de formation d'image de la caméra). (Notez que le terme « caméra » utilisé ici est identique au terme « image »).

Le résultat initial de ce processus est un emplacement relatif de tous les points et caméras – en d'autres termes, l'algorithme a défini l'emplacement de chaque élément et caméra dans un système de coordonnées

tridimensionnel arbitraire. Deux informations supplémentaires peuvent être utilisées pour transformer les coordonnées de ce système arbitraire en coordonnées du monde réel, comme la latitude et la longitude.

Si la géolocalisation de chaque caméra est connue avant ce processus, par exemple si la caméra a produit des images géolocalisées, les emplacements des caméras dans le système de coordonnées arbitraire peuvent être comparés à leurs emplacements réels dans le monde réel, et une transformation des coordonnées peut être définie et également appliquée aux emplacements des éléments identifiés. En outre, si l'emplacement de certaines des caractéristiques identifiées est connu (par exemple, à partir de mesures GPS sur le terrain), une transformation de coordonnées similaire peut être utilisée. Pour les besoins de la télédétection, l'une de ces deux sources d'informations de géolocalisation du monde réel est nécessaire pour que le produit résultant soit utile.

Étape 3) sur la base des caractéristiques maintenant géolocalisées, identifiez d'autres caractéristiques dans plusieurs images et reconstruisez leur position 3D. Imaginez qu'il y ait un petit point gris près du point rouge de la figure 79, disons $\frac{1}{4}$ du chemin vers le point vert, dans cette direction (il n'y en a pas, c'est pourquoi vous devez l'imaginer !). Ce petit point gris était trop faible pour être utilisé comme une caractéristique SIFT, et n'a donc pas été utilisé pour dériver l'emplacement initial des caractéristiques et les orientations des caméras. Cependant, maintenant que nous savons où trouver le point rouge dans chaque image, et que nous connaissons la position relative de chacun des autres points, nous pouvons estimer où ce point gris devrait se trouver dans chacune des images. Sur la base de cette estimation, nous pourrions effectivement être en mesure de trouver le point gris dans chaque image, et donc d'estimer sa position dans l'espace. Une fois identifiés et localisés dans l'espace, nous pouvons également attribuer une couleur au point, sur la base de sa couleur dans chacune des images où il a été trouvé. Cette opération, qui concerne des millions d'éléments trouvés dans l'ensemble des images, est appelée densification du nuage de points et constitue une étape facultative, mais presque universellement adoptée dans la SfM. Dans les zones où il existe une différence suffisante entre les couleurs des pixels voisins dans les images pertinentes, cette étape peut souvent conduire à la dérivation de l'emplacement de chaque pixel individuel dans une petite zone d'une image. Le résultat de cette étape est donc un « nuage de points » incroyablement dense, dans lequel des millions de points sont localisés dans le monde réel. Un exemple est présenté ci-dessous :

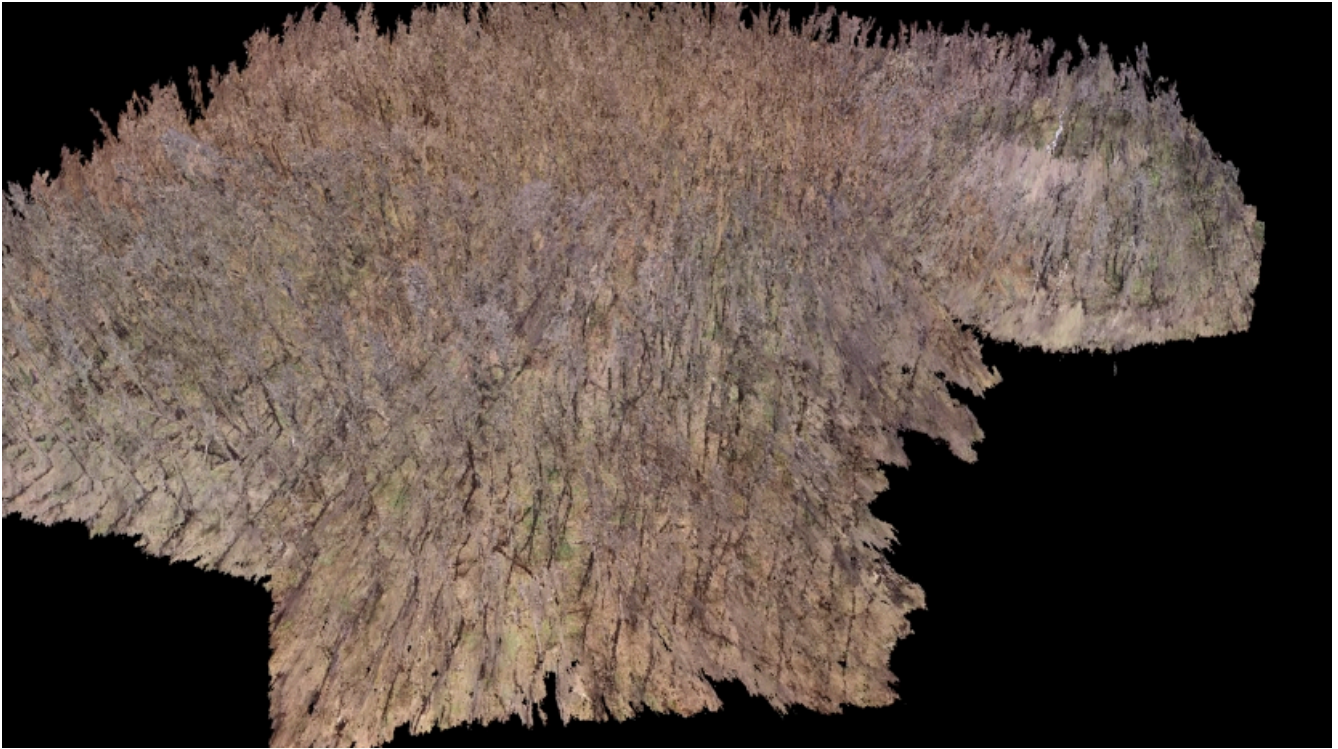


Figure 80: Exemple de nuage de points dense, créé pour une partie du parc de la Gatineau à l'aide d'images de drone. Anders Knudby, CC BY 4.0.

Étape 4) en option, utilisez toutes ces caractéristiques positionnées en 3D pour produire une orthomosaique, un modèle numérique d'élévation ou tout autre produit utile. Le nuage de points dense peut être utilisé à différentes fins, dont les plus courantes sont la création d'une orthomosaique et d'un modèle numérique d'élévation. Une **orthomosaique** est un type de données matricielles, dans lequel la zone est divisée en cellules (pixels) et la couleur de chaque cellule correspond à ce à quoi elle ressemblerait si elle était vue directement du dessus. Une orthomosaique est donc semblable à une carte, sauf qu'elle contient des informations sur la couleur de la surface, et représente ce à quoi ressemblerait la zone dans une image prise à une distance infiniment grande (mais avec une très bonne résolution spatiale). Les orthomosaiques sont souvent utilisées seules, car elles donnent aux gens une impression visuelle rapide d'une zone, et elles constituent également la base de la cartographie par la numérisation, la classification et d'autres approches. En outre, comme l'altitude est connue pour chaque point du nuage de points, l'altitude de chaque cellule peut être identifiée afin de créer un **modèle numérique de terrain**. Différents algorithmes peuvent être utilisés à cette fin, par exemple en identifiant le point le plus bas de chaque cellule pour créer un modèle numérique de terrain « terre nue », ou en identifiant le point le plus haut de chaque cellule pour créer un modèle numérique de surface. Des produits encore plus avancés peuvent être réalisés, par exemple en soustrayant le modèle de terrain nu du modèle de surface pour obtenir des estimations d'éléments tels que la hauteur des arbres et des bâtiments.

Détection des changements à l'aide de la

photogrammétrie

Si la création de modèles numériques d'élévation ou d'autres produits à l'aide de la stéréoscopie ou de la méthode de la structure à partir du mouvement peut être utile en soi, une application de plus en plus courante de ces techniques est la détection des changements de surface. Si, par exemple, un modèle numérique d'élévation est créé pour un glacier une année, et un nouveau deux ans plus tard, les changements de forme de la surface du glacier peuvent être cartographiés avec une grande précision. Les glissements de terrain, les effondrements dus au dégel, les changements de végétation, les variations de l'épaisseur de la neige au cours d'une saison — de nombreux changements dans l'environnement physique peuvent être détectés et quantifiés en mesurant les changements de la position horizontale ou verticale de la surface de la Terre. Étant donné le développement relativement récent de la technologie SfM, et les capacités encore en développement et le coût des drones équipés de caméras, il s'agit d'un domaine en développement rapide susceptible d'être utile dans des parties surprenantes et insoupçonnées du domaine de la télédétection. La figure 81 illustre ce principe en comparant le profil du glacier et les changements d'élévation de la surface à deux moments différents.

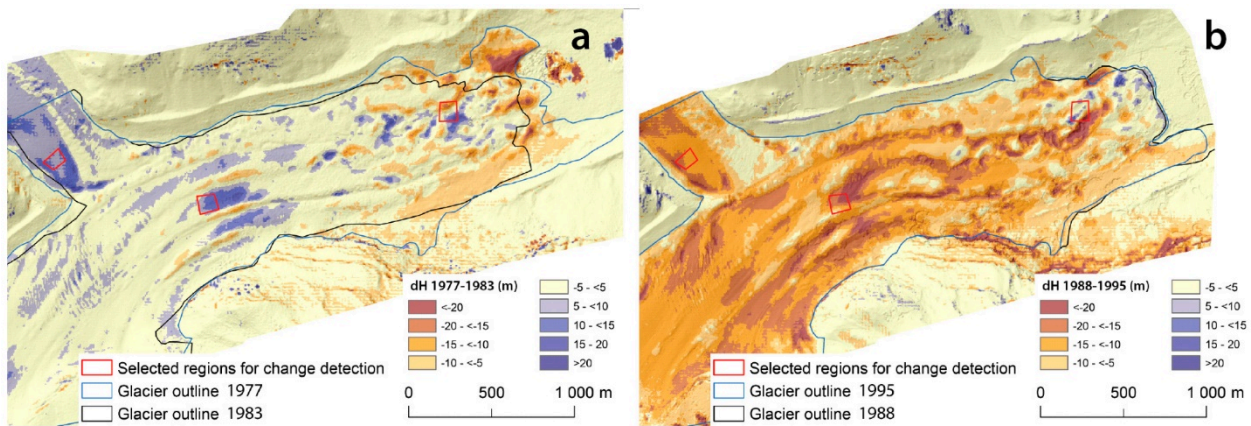


Figure 81: Détection de changement 3D basée sur la SfM répétée. Structure-from-Motion Using Historical Aerial Images to Analyse Changes in Glacier Surface Elevation (Figure 7) par Mölg and Bolch, CC BY 4.0.