# Rapport méthodologique :

# Cartographie de l'occupation du sol sur le site des NIPPES à Haïti

Montpellier, mars 2021

Stéphane Dupuy, Cirad UMR TETIS, Montpellier, France

Camille Lelong, Cirad UMR TETIS, Montpellier, France

Raffaele Gaetano, Cirad UMR TETIS, Montpellier, France

Ce rapport a été élaboré dans le cadre du projet Territoires Productifs Résilients (TPR)

Référence à citer :

Stéphane Dupuy, Camille Lelong, Raffaele Gaetano 2021. Rapport méthodologique - Cartographie de l'occupation du sol sur le site des NIPPES à Haïti, Montpellier (France) : Cirad, 44 pages.

© Cirad, Département ES - UMR TETIS, 2021.

# SOMMAIRE

## Table des matières


1. Introduction ..... 3

2. Les données ..... 4

2.1. Série temporelle d'images à HRS ..... 4

2.1.1. Téléchargement ..... 4

2.1.2. Prétraitements ..... 5

2.2. Images THRS ..... 6

2.2.1. Spot 6/7 ..... 6

2.2.2. Pléiades ..... 8

2.3. Modèle Numérique de Terrain ..... 11

2.4. Base de données de référence et nomenclature ..... 12

3. Cartes d'occupation du sol produites avec la chaine Moringa ..... 15

3.1. Segmentation ..... 16

3.2. Étude des variables mises en œuvre ..... 19

3.3. Post-traitements ..... 24

3.3.1. Routes Open Street Map ..... 24

3.3.2. Lissage ..... 24

3.3.3. Vectorisation ..... 24

3.4. Les cartes obtenues ..... 25

3.5. Validation des cartes ..... 38

4. Conclusion ..... 43

5. Remerciements ..... 44

6. Références ..... 44

# 1. INTRODUCTION

Dans le cadre du projet Territoires Productifs Résilients en Haïti (TPR), il est prévu de produire par télédétection des cartes d'occupation des sols, si possible sur l'ensemble du territoire haïtien, afin de guider l'échantillonnage des protocoles d'études de systèmes de culture innovants.

Pour cartographier avec précision ces espaces agricoles, il est nécessaire de pouvoir utiliser plusieurs images acquises à intervalles réguliers au cours de la saison culturale. Il est ainsi possible d'identifier les cultures grâce à leur comportement au cours de l'année, car cette évolution est différente pour chaque type de culture (semis, croissance de la végétation, récolte et labour pour préparer la culture suivante).

Depuis 2017, avec la mise en service des 2 satellites Sentinel-2 par l'agence spatiale européenne (ESA), l'accès aux images satellites a profondément changé. En effet les images sont acquises tous les 5 jours et couvrent l'ensemble des terres. En plus de la fréquence de revisite, les informations spectrales qu'elles apportent sont particulièrement utiles pour cartographier l'agriculture. De nouveaux outils permettent aussi de les exploiter de façon optimale. On peut citer par exemple MAJA et IOTA2, mis au point par le CESBIO, qui permettent respectivement de prétraiter les données et de produire des cartes d'occupation du sol. L'inconvénient de ces images est la résolution spatiale, qui est comprise entre 10 et 20m en fonction des bandes spectrales. Cette résolution est parfois insuffisante pour cartographier les espaces agricoles des pays du Sud, souvent caractérisés par de fortes variabilités intra-parcellaires, la petite taille des parcelles, la présence de jachères, de cultures associées, d'agroforêts…

Pour répondre à ces problématiques, au sein de l'UMR TETIS, nous avons mis au point le prototype d'une chaine de traitement appelée Moringa (Gaetano et al., 2019) qui fait appel à des fonctions de l'Orfeo Tool Box (OTB), orchestrées par des scripts en python. La différence majeure par rapport à une chaine comme IOTA2, par exemple, est de pouvoir ajouter à la série d'images Sentinel-2 une image à Très Haute Résolution Spatiale (THRS), qui sera utilisée d'une part pour extraire des objets (intéressant pour délimiter des petits parcellaires) et d'autre part pour calculer des indices radiométriques et de texture (intéressant pour identifier l'arboriculture).

L'inconvénient des images THRS est leur emprise spatiale, qui est plus réduite. Mais, surtout, le coût de la donnée peut s'avérer être un frein dans ce type d'approche. En effet, contrairement à Sentinel-2, l'accès à ces images est payant : il faut compter[1] environ 21,25 USD du kilomètre carré pour les images Pléiades (avec une commande minimum de 100km$^2$) et 2,45 USD du kilomètre carré pour SPOT6/7, (avec une commande minimale de 500 km$^2$). Il est donc important de comprendre quel est véritablement le gain en matière de précision de cartographie qu'apporterait l'intégration de ce type d'images dans un projet de développement à l'échelle régionale ou nationale.

Sur le site pilote du TPR, correspondant à trois bassins versants de la région des Nippes à Haïti (Baconois, Rivière Froide et Bondeau, couvrant respectivement 30, 40 et 20 kilomètres carrés), le CNES (en collaboration avec DINAMIS), via le projet RO HAITI, met à disposition des images acquises par des satellites Pléiades et Spot 6/7. Les images Pléiades ont une résolution comprise entre 2m et 50cm alors que pour Spot6/7 elle est comprise entre 6m et 1,5m. Bien sûr ces différences auront un impact sur la cartographie qui sera produite. C'est pour cette raison que nous avons fait le choix de tester la chaine Moringa avec ces 2 types d'images afin que les acteurs locaux puissent avoir des éléments de comparaison au moment de choisir la méthode la plus adaptée au contexte haïtien. Nous développons dans ce rapport les différentes étapes conduisant à la production de ces cartes.

[1]http://www.landinfo.com/satellite-imagery-pricing.html,
https://apollomapping.com/image_downloads/Apollo_Mapping_Imagery_Price_List.pdf

# 2. LES DONNEES

## 2.1. SERIE TEMPORELLE D'IMAGES A HRS

Les images utilisées pour constituer la série temporelle sont des images acquises par la constellation de satellites Sentinel-2 de l'agence spatiale européenne (ESA).

Avec une répétitivité de 5 jours avec 2 satellites (S2A et S2B), toutes les terres émergées et les traits de côte sont couverts. Les images sont gratuites et ont une emprise au sol de 290 km de côté pour 13 bandes spectrales (Figure 1). Ce capteur est particulièrement adapté pour la cartographie des espaces agricoles car il permet de suivre l'évolution de la végétation grâce à la fréquence de revisite mais aussi aux bandes spectrales dans le visible et l'infrarouge.

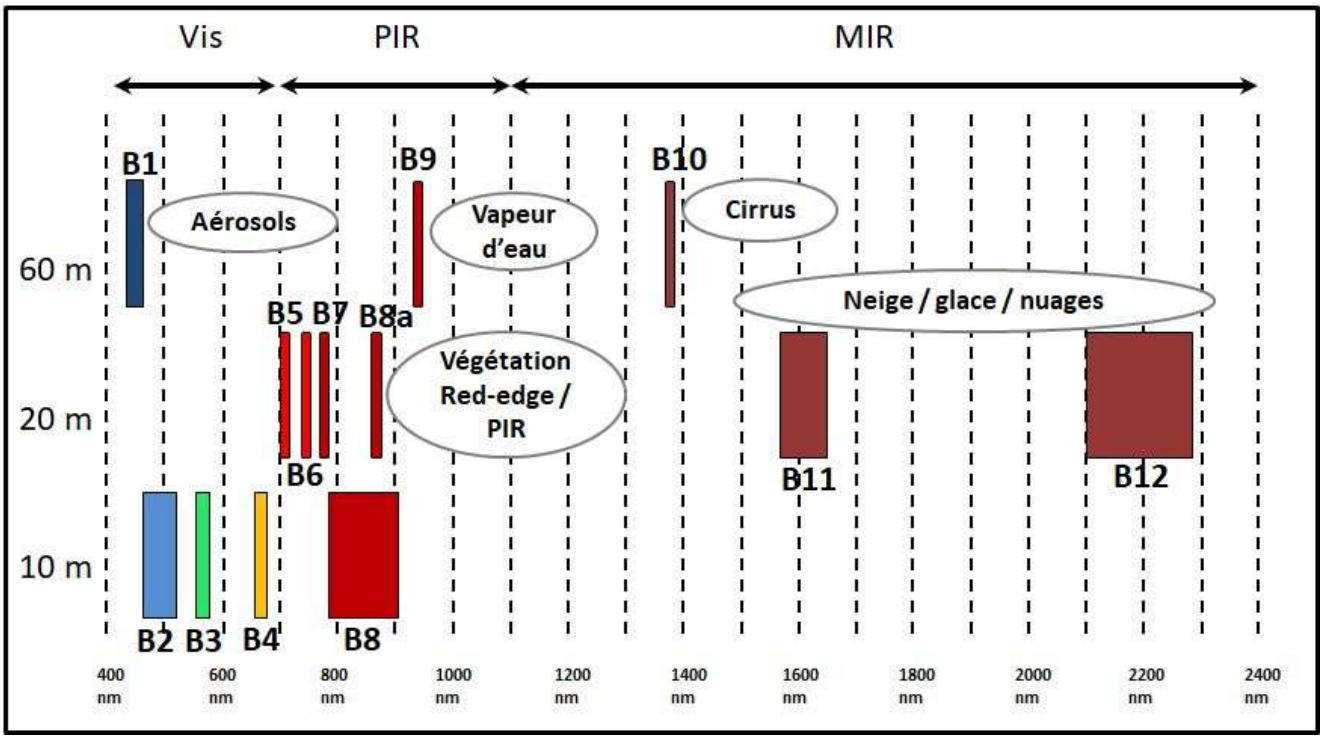


*Figure 1: Bandes spectrales et résolution spatiale des satellites Sentinel-2A et Sentinel-2B*

### 2.1.1. TELECHARGEMENT

Le site étudié est couvert par les produits de niveau 2A distribués par le pôle de données THEIA[2]. Les produits de niveau 2A fournissent des réflectances de surfaces corrigées des effets atmosphériques et sont fournies avec un masque des nuages et de leurs ombres. Le traitement des données est réalisé avec la chaîne MAJA (autrefois appelée MACCS), par le centre de traitement MUSCATE du CNES, pour le compte du centre de données Theia. La documentation sur ce format, ainsi que sur les images, est disponible à cette adresse : https://theia.cnes.fr.

Pour constituer la série temporelle, nous avons tenu compte des dates des images à THRS (décembre 2019 et septembre/octobre 2020) et des relevés de terrain (septembre 2020), et donc téléchargé les images acquises entre le 1er décembre 2019 et le 30 novembre 2020, soit un total de 92 images.

[2] *www.theia.land.fr*

## 2.1.2. PRETRAITEMENTS

Les prétraitements consistent dans un premier temps à ne conserver que les données qui seront utiles et à appliquer aux images les masques de nuages et des ombres de nuages.

L'étape suivante est de générer les images synthétiques qui seront utilisées dans Moringa. Pour rappel, la constellation Sentinel-2 réalise une acquisition tous les 5 jours. La couverture nuageuse implique que la totalité de ces images n'est pas exploitable (surtout en zone tropicale). L'idée est donc de produire des images synthétiques dont on choisira la fréquence : 1 image tous les 10, 15 ou 20 jours, par exemple.

Pour notre cas nous avons choisi une fréquence de 15 jours entre chaque image avec une date de début au 20 décembre 2019 et une date de fin au 15 novembre 2020. Nous avons donc une série temporelle constituée de 23 images synthétiques, illustrée par la Figure 2.

*Figure 2 : Série temporelle d'images Sentinel-2 constituée d'images synthétiques (du 20/12/2019 au 14/11/2020). Données à valeur ajoutée traitées par le CNES pour le pôle de données Theia www.theia.land.fr à partir de données Copernicus. Les traitements utilisent les algorithmes développés par les Centres d'Expertise Scientifique de Theia.*

Ces images synthétiques sont ensuite recalées sur l'image à THRS qui sera utilisée dans la chaine. Cette étape consiste à réaliser une translation des images Sentinel-2 grâce à une sélection automatique de points homologues entre les images.

Enfin, des indices radiométriques sont calculés pour chacune des 23 images. Nous avons choisi des indices permettant d'exploiter au mieux la richesse de l'information spectrale de Sentinel-2 :

- NDVI: normalized difference vegetation index
- NDWI: normalized difference water index
- SWNDVI: short-wave normalized difference vegetation index
- MNDWI: modified normalized difference water index
- NDRE : red-edge NDVI

## 2.2. Images THRS

### 2.2.1. Spot 6/7

Trois images ont été acquises par Airbus DS sur la zone d'étude et mises à disposition sur le site de DINAMIS. Deux images ont été acquises pendant la période des relevés de terrain soit le 26 septembre et le 15 octobre 2020. Comme la couverture nuageuse de ces images est assez importante, une image acquise le 13 décembre 2019 sera aussi utilisée en complément. Pour être exploitables, ces données, délivrées au niveau brut, doivent subir un certain nombre de corrections et prétraitements préalables.

- **Correction radiométrique :** Toutes les images ont été corrigées des effets du capteur et converties en réflectance Top of Atmosphère (TOA) de façon automatique avec le logiciel Orfeo ToolBox (OTB).
- **Correction géométrique :** La correction géométrique des images s'est déroulée en 2 étapes :
  - La première a consisté en une orthorectification utilisant le modèle de capteur et un MNT précis (MNT IGN issu d'acquisition LiDAR). Processus réalisé avec OTB.
  - La seconde a consisté en une correction géométrique de type « polynôme de degrés 2 » qui repose sur la sélection d'une dizaine de points de contrôle par rapport à une image de référence. L'image de référence choisie est une photo aérienne orthorectifiée par l'IGN. Cette image a été utilisée pour choisir les points par rapport à l'image du 13 décembre 2019. Pour les images acquises en 2020, la recherche de point d'appuis a été réalisée à partir de l'image du 13 décembre 2019 préalablement corrigée, afin que la superposition entre les 3 images soit de bonne qualité en vue des étapes suivantes. Processus réalisé avec Erdas Imagine.
- **Pansharpening (fusion multirésolution) :** Le pansharpening consiste à fusionner les images panchromatique et multispectrale d'origine pour obtenir une image multispectrale ayant la résolution de l'image panchromatique, dite « pansharpened ». Cette étape a été réalisée à l'issue de l'orthorectification en utilisant l'algorithme « Bayes » avec le logiciel OTB.
- **Mosaiquage :** Cette étape consiste à assembler (mosaïquer) en une seule image les trois acquisitions pour obtenir une seule image THRS couvrant toute la zone d'étude avec le moins de nuages possible. Cette mosaïque a été produite sous le logiciel OTB, pour chacun des modes : multispectral, panchromatique et pansharpened. Cela a nécessité en outre la création d'un masque en vue de guider le logiciel pour le découpage le plus optimisé possible de l'image. La Figure 3 illustre la carte d'assemblage des images et les Figure 4 et Figure 5 l'image produite.

  Afin de rendre possible l'utilisation de cette mosaïque par les partenaires haïtiens, les images ont été transmises par le Cirad au CNES, qui les a ajoutées au catalogue du projet Recovery Observatory (RO): https://www.recovery-observatory.org . Elles sont donc accessibles sous réserve d'être éligible au projet.

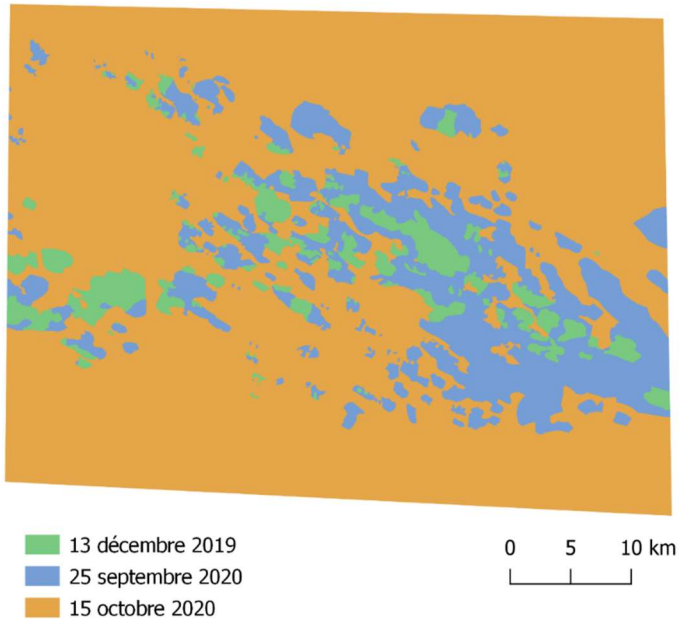


*Figure 3 : Carte d'assemblage des images pour l'étape de mosaïquage pour l'image Spot 6/7.*

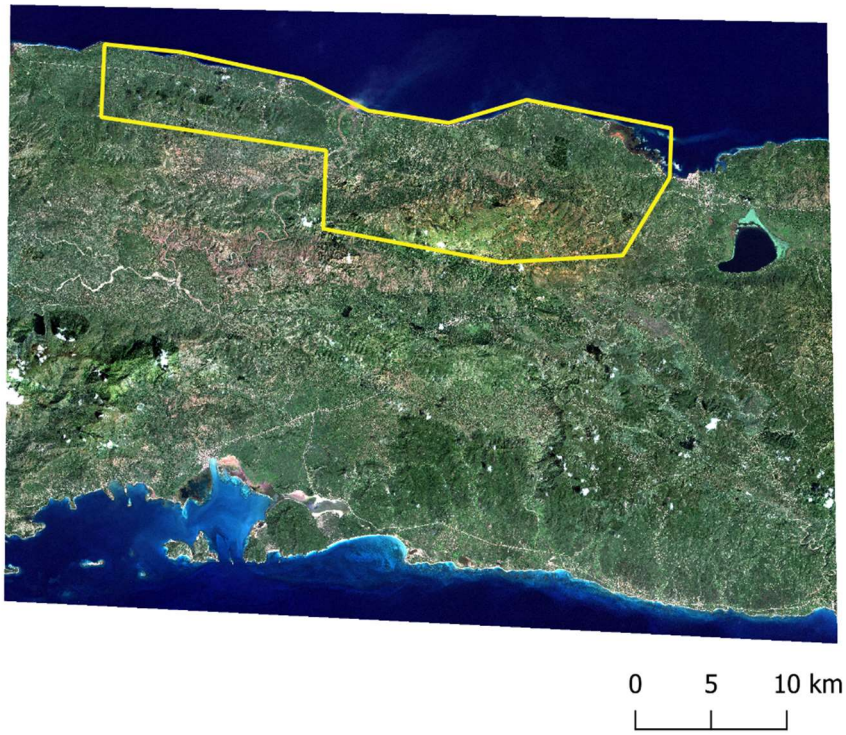


*Figure 4 : Mosaïque Spot6/7 en couleurs naturelles avec en jaune la limite de la zone d'étude de Nippes (© AIRBUS DS 2020 - Distribution GEOSUD / DINAMIS)*

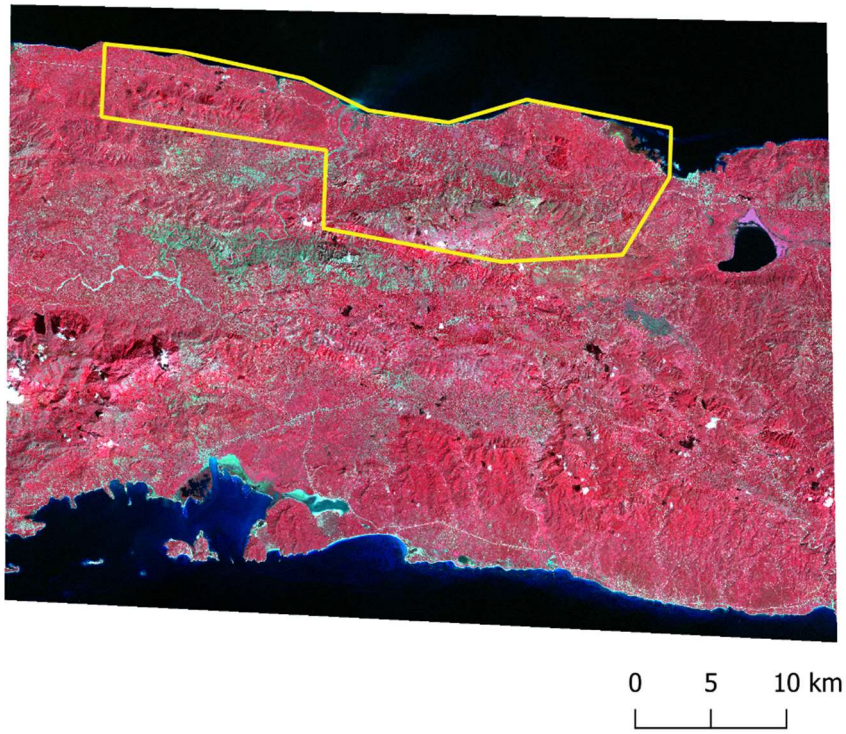


*Figure 5 : Mosaïque Spot6/7 en fausses couleurs avec en jaune la limite de la zone d'étude de Nippes (© AIRBUS DS 2020 - Distribution GEOSUD / DINAMIS)*

- **Calcul des indices de végétation et brillance :** Afin de prendre en compte les informations spectrales de l'image THRS, nous avons calculé des indices à partir des bandes spectrales de l'image fusionnée (pansharpened). Nous avons donc utilisé l'outil « Radiometric Indices » de l'OTB pour calculer les indices suivants :
  - NDVI : Normalized difference vegetation index (Red, NIR)
  - RVI : Ratio vegetation index (Red, NIR)
  - NDWI2 : Normalized difference water index (Mc Feeters 1996) (Green, NIR)
  - BI2 : Brightness index 2 (NIR, Red, Green)
- **Calcul des indices de texture :** Les textures sont importantes pour détecter les motifs visibles sur l'image à THRS, comme par exemple les alignements d'arbres de l'arboriculture. Nous avons utilisé l'outil « Haralick Texture Extraction » de l'OTB appliqué sur l'image panchromatique. Pour l'image Spot6/7 nous avons calculé les indices « Energy », « Contrast », « Correlation » et « Variance » avec des radius de 1 (soit 3x3 pixels), 5 (soit 11x11 pixels) et 11 (soit 23x23 pixels). Soit un total de 12 indices de texture.

### 2.2.2. PLEIADES

Cinq images ont été acquises par Airbus DS sur la zone d'étude et mises à disposition par le CNES sur le site du projet Recovery Observatory (RO) : https://www.recovery-observatory.org. Trois d'entre elles ont été acquises pendant la période des relevés de terrain soit les 19 et 24 septembre 2020 et le 15 octobre 2020. La partie Sud Est du site d'étude des Nippes n'a pas été couverte en totalité en 2020 par Pléiades donc, pour compléter cette partie ainsi que les zones couvertes par des nuages sur les images de 2020, nous avons utilisé l'image acquise le 10 octobre 2019. Cette image ayant été acquise pendant la saison des pluies, comme celle de 2020, les différences de stades de végétation seront limitées. Pour être exploitables, ces données aussi doivent subir un certain nombre de corrections et prétraitements préalables

- **Correction radiométrique :** Toutes les images Pléiades ont été corrigées des effets du capteur et converties en réflectance Top of Atmosphère (TOA) de façon automatique avec le logiciel Orfeo ToolBox (OTB).
- **Correction géométrique :** Toutes les images Pléiades ont été orthorectifiées en utilisant le modèle de capteur et le MNT IGN issu d'acquisition LiDAR. Processus réalisé avec OTB.
- **Pansharpening (fusion multirésolution) :** Toutes les images Pléiades ont été fusionnées en utilisant l'algorithme « Bayes » avec le logiciel OTB, comme les images SPOT6/7. Une autre version du pansharpening a aussi été réalisée avec l'algorithme Gram-Schmidt proposé par le logiciel ENVI. Cette version présente plus de contrastes, dans les zones de végétation notamment.
- **Mosaïquage :** Cette étape a consisté à mosaïquer les cinq images Pléiades pour en obtenir une seule couvrant toute la zone d'étude avec le moins de nuages possible. Ce processus est réalisé avec le logiciel OTB. Il a nécessité la création d'un masque en vue de guider le logiciel pour le découpage le plus optimisé possible de l'image. La Figure 6 illustre la carte d'assemblage des images et les Figure 7 et Figure 8 l'image produite.

  Afin de rendre possible l'utilisation de cette mosaïque par les partenaires haïtiens, les images ont été transmises par le Cirad au CNES, qui les a ajoutées au catalogue du projet Recovery Observatory (RO): https://www.recovery-observatory.org . Elles sont donc accessibles sous réserve d'être éligible au projet.

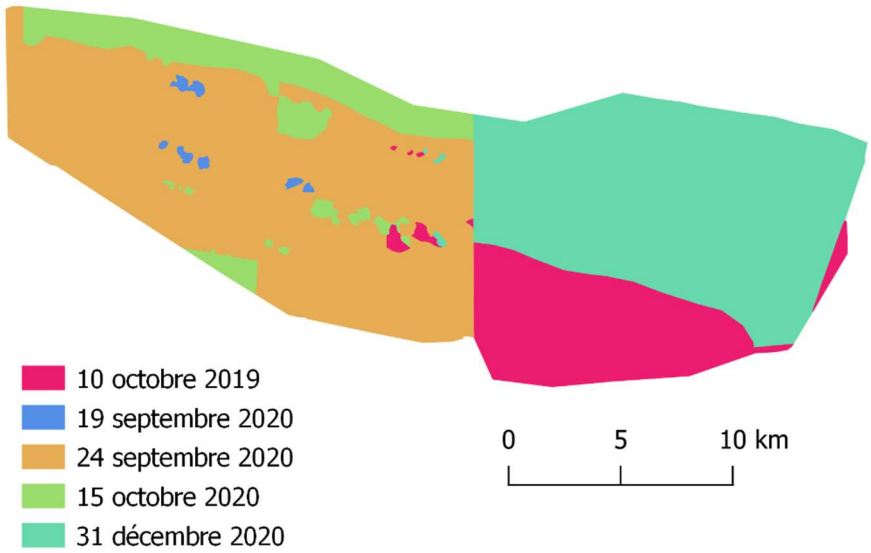


*Figure 6 : Carte d'assemblage des images pour l'étape de mosaïque pour l'image Pléiades.*

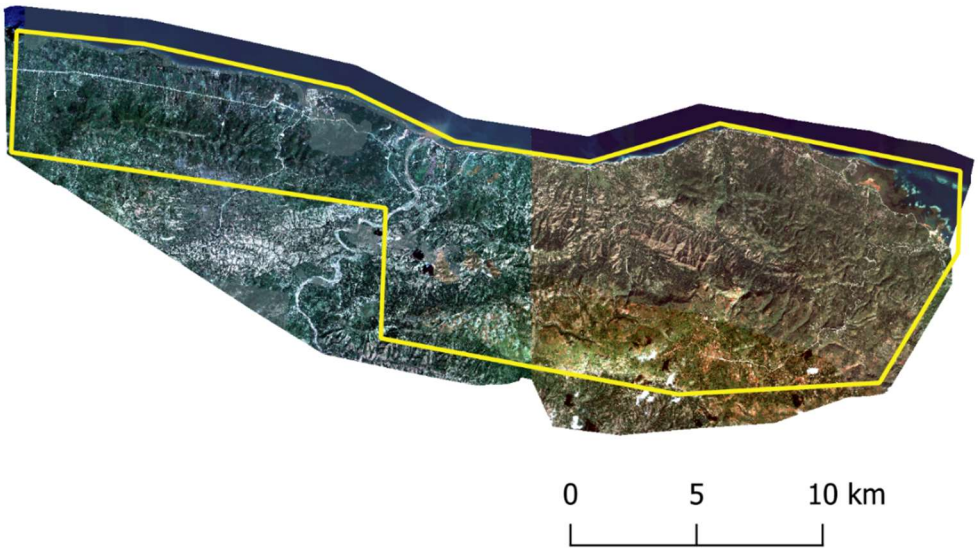


*Figure 7 : Mosaïque Pléiades en couleurs naturelles avec en jaune la limite de la zone d'étude de Nippes (© CNES 2020 - Distribution AIRBUS DS)*

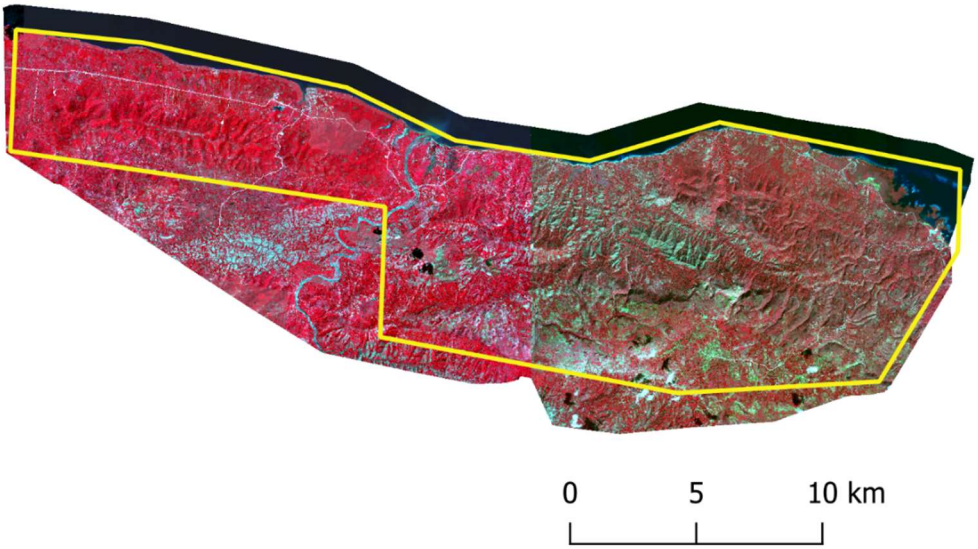


*Figure 8 : Mosaïque Pléiades en fausses couleurs avec en jaune la limite de la zone d'étude de Nippes (© CNES 2020 - Distribution AIRBUS DS)*

- **Calcul des indices de végétation et brillance :** Afin de prendre en compte les informations spectrales de l'image à THRS, nous avons calculé des indices à partir des bandes spectrales de l'image fusionnée (pansharpened). Nous avons donc utilisé l'outil « Radiometric Indices » de l'OTB pour calculer les indices suivants :
  - NDVI : Normalized difference vegetation index (Red, NIR)
  - RVI : Ratio vegetation index (Red, NIR)
  - NDWI2 : Normalized difference water index (Mc Feeters 1996) (Green, NIR)
  - BI2 : Brightness index 2 (NIR, Red, Green)
- **Calcul des indices de texture :** Les textures sont importantes pour détecter les motifs visibles sur l'image à THRS, comme par exemple les alignements d'arbres de l'arboriculture. Nous avons utilisé l'outil « Haralick Texture Extraction » de l'OTB appliqué sur l'image panchromatique. Pour l'image Pléiades nous avons calculé les indices « Energy », « Contrast », « Correlation » et « Variance » avec des radius de 1 (soit 3x3 pixels), 5 (soit 11x11 pixels), 11 (soit 23x23 pixels) et 21 (soit 43x43 pixels). Soit un total de 16 indices de texture.

# 2.3. MODELE NUMERIQUE DE TERRAIN

Le Modèle Numérique de Terrain (MNT) est une donnée informant sur l'altitude d'un territoire. L'intégrer dans un processus de classification permet de prendre en compte l'altitude dans les variables de la classification. Pour notre zone d'étude nous avons utilisé le MNT issu d'une acquisition LiDAR de 2014 réalisé par IGN FI, ayant mesuré l'altitude avec une résolution altimétrique de 20cm à la résolution spatiale de 1,5m.

Une couche représentant les pentes a également été dérivée de ce MNT. Le calcul des pentes a été réalisé avec le logiciel QGIS. Ces pentes seront donc également ajoutées dans les variables de la classification.

La Figure 9 illustre ces deux types de produits topographiques (altitude et pentes) sur un extrait de la zone d'étude situé dans les environs de la Grande Rivière des Nippes.

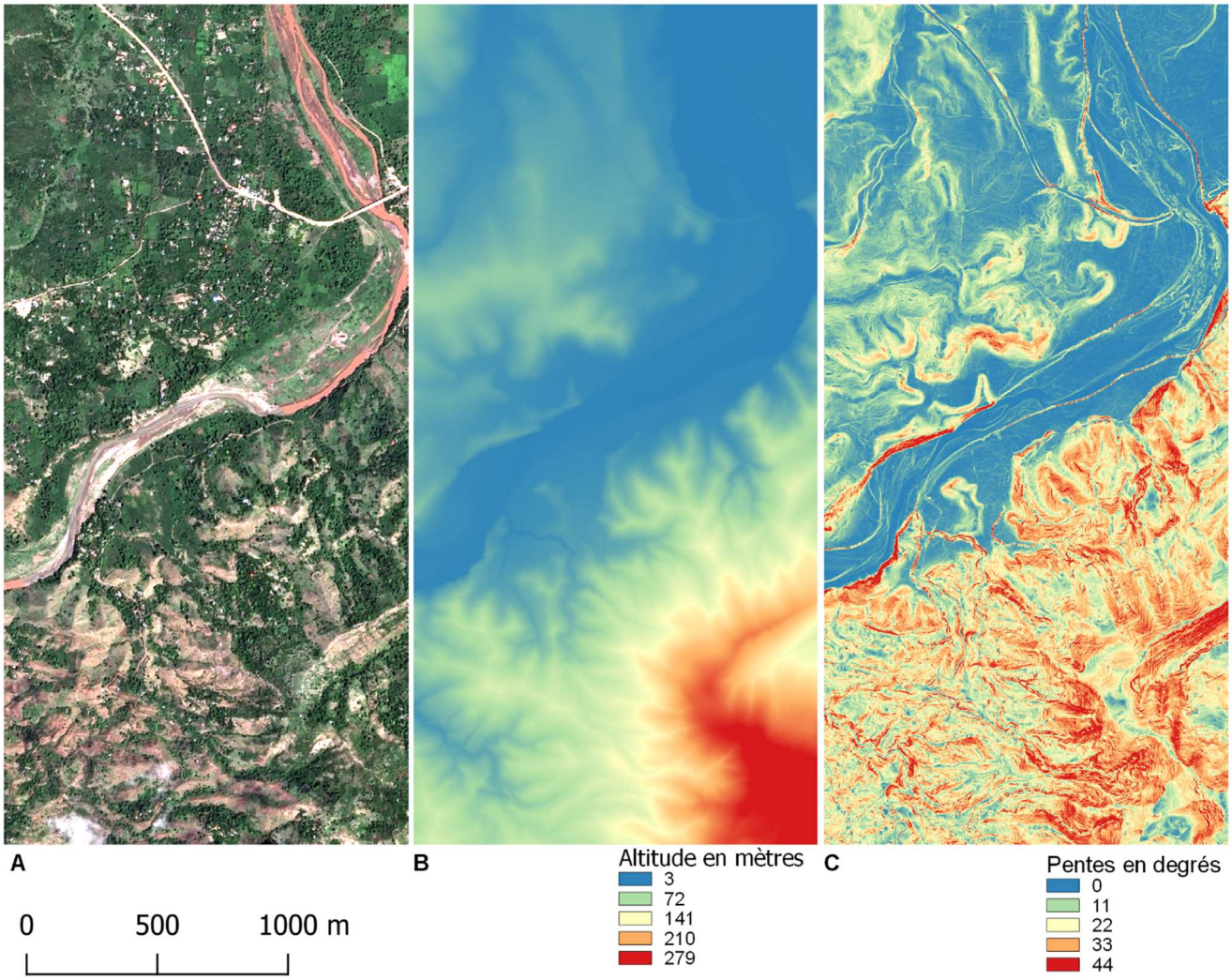


*Figure 9 : Extrait sur la même zone, localisée près de la Grande Rivière des Nippes, de **A** : l'image Spot 6/7, **B** : l'altitude issue du Modèle Numérique de Terrain (MNT) LIDAR IGN FI, et **C** : les pentes dérivées de ce même MNT (contient des données © AIRBUS DS 2020 - Distribution GEOSUD / DINAMIS)*

# 2.4. BASE DE DONNEES DE REFERENCE ET NOMENCLATURE

Dans le cadre d'un processus de classification supervisé, la base de données de références guide l'algorithme afin qu'il reconnaisse dans les données issues des images, au cours de sa phase d'apprentissage, les classes d'occupation du sol qui nous intéressent. Il s'agit donc d'un élément essentiel dans le processus. Cette base de données est également utilisée pour valider le résultat obtenu en utilisant la procédure de validation croisée qui produit les statistiques de précision globale et par classe. Pour répondre à cette exigence, il existe un certain nombre de règles à suivre pendant la constitution de ladite base. Le document « Guide Technique : Constitution d'une base de données de référence dédiée à l'apprentissage et la validation des classifications d'images satellites : la vérité-terrain. » (Lelong Camille et Dupuy Stéphane. 2021. Montpellier : CIRAD-ES-UMR TETIS, 16 p.), livré aux différents partenaires du projet en date du 1er mars 2021 et disponible dans la base documentaire Agritrop, présente dans ses détails ce protocole.

Une première base de données de référence a été constituée à partir de relevés de train réalisés par les équipes du CNIGS et du projet TPR en septembre et octobre 2020. Les points relevés sur le terrain ont ensuite été complétés par photo-interprétation par le CNIGS. La dernière version envoyée par le CNIGS le 18 janvier 2021 présente un total de 683 polygones, représentant une superficie de 1620 ha. La Figure 10 illustre la répartition de ces polygones. On constate que ces relevés sont bien répartis sur l'ensemble des 3 bassins versants de l'étude.

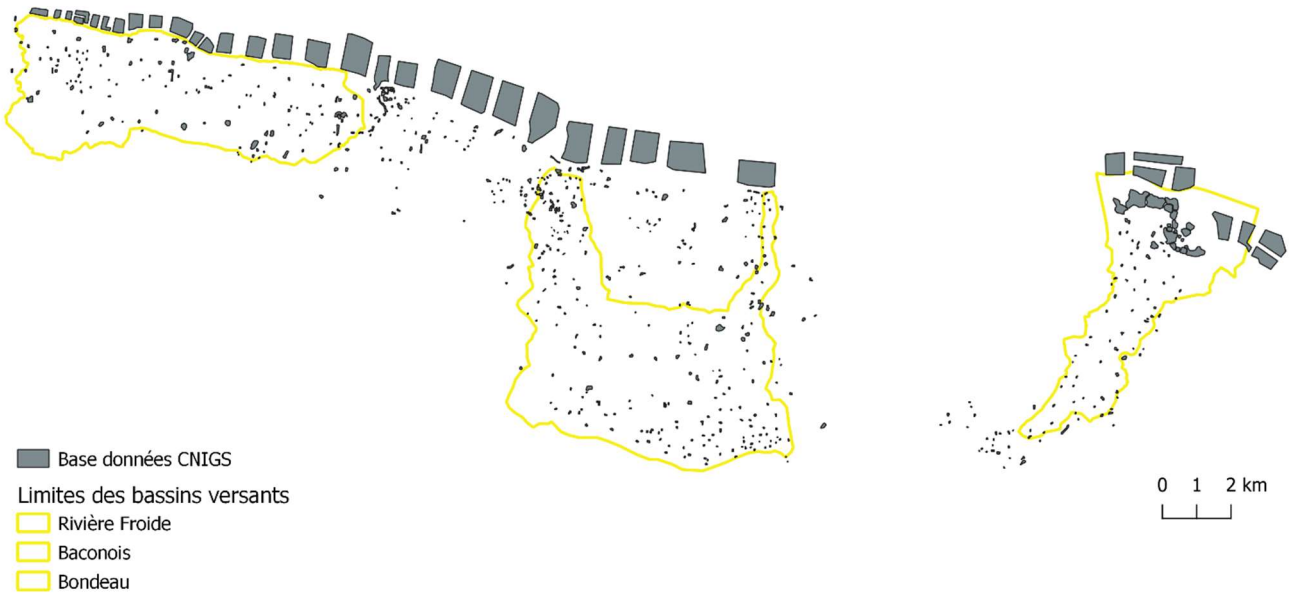


*Figure 10 : Répartition des polygones de la base de données du CNIGS transmise le 18 janvier 2021*

Toutefois, on remarque (cf. Tableau 1) un déséquilibre important entre les différentes classes en termes de nombre de polygones (trop faible par exemple pour l'arboriculture et l'eau continentale) mais aussi de superficie (trop faible pour toutes les classes très variables comme les cultures sous couvert arboré, l'arboriculture, les espaces peu ou pas végétalisés, les eaux continentales, mais au contraire beaucoup trop grande en proportion pour les zones humides et l'eau maritime).

Aussi, cette base de données ne présente pas un nombre suffisant de polygones par classe pour être représentative des classes d'occupation du sol présentes sur le terrain, ni pour couvrir la variabilité rencontrée dans chaque classe au travers de la zone d'étude.

Enfin, l'observation des contours digitalisés a mis en évidence de nombreuses erreurs dans la délimitation des zones de référence, recouvrant parfois plusieurs types de classes au lieu d'être bien homogènes.

*Tableau 1: Statistiques calculées sur la base de données du CNIGS du 18 janvier 2021*

| Code2 | Nom | Nombre de polygones | Surface (Ha) |
|---|---|---|---|
| **1** | Bâti | 33 | 14.50 |
| **2** | Culture sans couvert arboré | 121 | 20.18 |
| **3** | Culture avec présence d'arbres | 66 | 12.52 |
| **4** | Système agro-forestier | 104 | 33.39 |
| **5** | Arboriculture | **19** | 6.43 |
| **6** | Végétation arborée | 49 | 25.95 |
| **7** | Végétation arbustive | 102 | 44.34 |
| **8** | Pâturage, prairie et savane herbacée | 93 | 27.99 |
| **9** | Zone humide côtière (mangrove) | 26 | **111.32** |
| **10** | Rocher, sol nu et plage | 31 | 8.86 |
| **11** | Eau continentale | **2** | 3.23 |
| **12** | Eau marine | 37 | **1311.38** |
| | **Total** | **683** | **1620.10** |

Le Cirad a alors corrigé les polygones existants en redessinant leurs contours, puis vérifié et corrigé si besoin les affectations de classe de chacun de ces polygones. Puis le Cirad a enrichi la base en y ajoutant de nouveaux polygones, en s'appuyant sur son expérience en photo-interprétation des images Pléiades dans les milieux ruraux tropicaux de même faciès qu'en Haïti, sur la connaissance apportée par les relevés de terrain effectués par le CNIGS et le TPR, ainsi que sur les photos fournies par le TPR, acquises lors de la campagne de terrain et ultérieurement (février 2021). Cette **base de référence validée** a été transmise à tous les partenaires du projet en date du 17 février 2021 (*Dupuy, S.; Lelong, C.; Manasse, M.-E.; Rambao, J.; Mondésir, J. P.; Gaetano, R., 2021, "Nippes - Haïti - 2020, Reference Spatial Database", doi:10.18167/DVN1/29RZMQ, CIRAD Dataverse, V1*).

La Figure 11 illustre la nouvelle répartition des polygones digitalisés dans la zone d'étude ; le Tableau 2 présente les évolutions apportées par le Cirad en termes de nombres de polygones et de superficies couvertes par ces polygones.

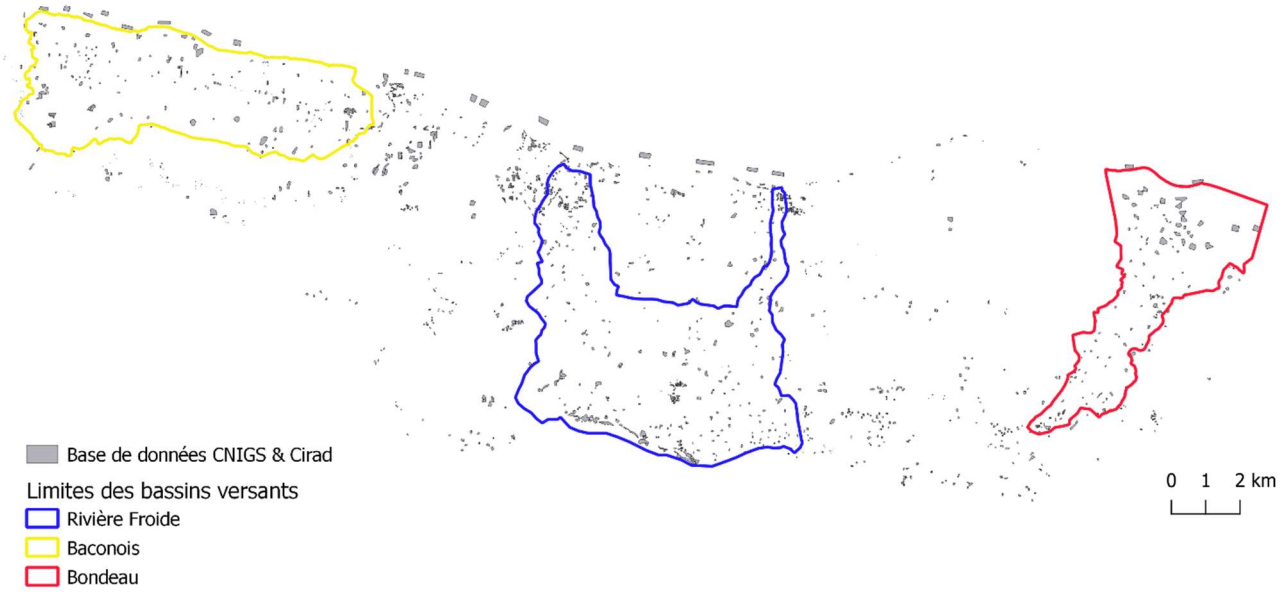


*Figure 11: Répartition des polygones de la base de données de référence complétée et validée par le Cirad, transmise au TPR le 17 février 2021. (Dupuy, S.; Lelong, C.; Manassé, M.-E.; Rambao, J.; Mondésir, J. P., Gaetano, R., 2021, "Nippes - Haïti - 2020, Reference Spatial Database", doi:10.18167/DVN1/29RZMQ, CIRAD Dataverse, V1)*

*Tableau 2: Statistiques calculées sur la base de données du Cirad du 17 février 2021, et variations par rapport à la base de données d'origine du CNIGS.*

| Code2 | Nom | Nombre de polygones | Variation nombre | Surface (Ha) | Variation surface |
|---|---|---|---|---|---|
| **1** | Bâti | 268 | +235 | 9.91 | -4.59 |
| **2** | Culture sans couvert arboré | 368 | +247 | 60.20 | 40.01 |
| **3** | Culture avec présence d'arbres | 90 | +24 | 41.86 | 29.34 |
| **4** | Système agro-forestier | 181 | +77 | 16.63 | -16.76 |
| **5** | Arboriculture | 34 | +15 | 35.90 | 29.47 |
| **6** | Végétation arborée | 109 | +60 | 61.13 | 35.18 |
| **7** | Végétation arbustive | 148 | +46 | 22.64 | -21.71 |
| **8** | Pâturage, prairie et savane herbacée | 153 | +60 | 9.55 | -18.44 |
| **9** | Zone humide côtière (mangrove) | 34 | +8 | 4.75 | -106.57 |
| **10** | Rocher, sol nu et plage | 149 | +118 | 74.82 | 65.96 |
| **11** | Eau continentale | 32 | +30 | 76.58 | 73.35 |
| **12** | Eau marine | 33 | -4 | 84.09 | -1227.29 |
| | **Total** | **1599** | **+916** | **498.06** | **-1122.03** |

La nomenclature finale retenue est présentée dans le Tableau 3. Elle est constituée d'un niveau détaillé (dit « niveau 2 ») comprenant 12 classes, correspondant aux types d'occupation du sol potentiellement distinguables sur les images Pléiades et utiles au projet TPR. Au-dessus se trouve un niveau moins détaillé (dit « Niveau 1 ») comprenant 8 classes seulement, plutôt destiné à produire une classification avec la chaine Iota2 qui n'utilise pas d'image THRS. En effet, il y a peu de chances de distinguer entre elles l'arboriculture et l'agroforesterie à la résolution de Sentinel-2, ou encore les végétations naturelles arborées et arbustives, ni même la présence d'arbres dans une culture de base. On considère généralement qu'il vaut mieux produire une carte à la nomenclature moins détaillée mais plus robuste et plus valide. En plus de ces 2 niveaux, nous proposons un troisième niveau de nomenclature (dit « Niveau 0 ») qui regroupe tous les couverts ligneux cultivés ou non (arboriculture/agroforesterie vs. espaces boisés/arbustifs). En effet, une distinction entre ces différents types d'occupation du sol est assez peu probable à 10m de résolution spatiale.

*Tableau 3 : Nomenclature finale à 3 niveaux utilisée pour la production des cartes d'occupation du sol avec Moringa*

| Code0 | Nom de classe niveau 0 |
|---|---|
| 1 | Bâti |
| 2 | Culture non ligneuse |
| 3 | Végétation ligneuse (cultivée et naturelle) |
| 4 | Pâturage, prairie et savane herbacée |
| 6 | Zone humide côtière (mangrove) |
| 5 | Rocher, sol nu et plage |
| 7 | Eau |
| 8 | Routes (origine OSM) |

| Code1 | Nom de classe niveau 1 |
|---|---|
| 1 | Bâti |
| 2 | Culture non ligneuse |
| 3 | Agroforesterie et arboriculture |
| 4 | Végétation naturelle ligneuse |
| 5 | Pâturage, prairie et savane herbacée |
| 6 | Zone humide côtière (mangrove) |
| 7 | Rocher, sol nu et plage |
| 8 | Eau |
| 9 | Routes (origine OSM) |

| Code2 | Nom de classe niveau 2 |
|---|---|
| 1 | Bâti |
| 2 | Culture sans couvert arboré |
| 3 | Culture avec présence d'arbres |
| 4 | Système agro-forestier |
| 5 | Arboriculture |
| 6 | Végétation arborée |
| 7 | Végétation arbustive |
| 8 | Pâturage, prairie et savane herbacée |
| 9 | Zone humide côtière (mangrove) |
| 10 | Rocher, sol nu et plage |
| 11 | Eau continentale |
| 12 | Eau marine |
| 13 | Routes (origine OSM) |

Les cartes d'occupation du sol ont ainsi été produites sous Moringa avec ces 3 niveaux de précision de la nomenclature et dans 2 versions distinctes : l'une utilisant Spot6/7 et l'autre utilisant Pléiades.

Au final, il restera à identifier dans la carte toute l'infrastructure routière, dont la délimitation est difficile à obtenir avec une grande précision par classification d'images. Nous avons donc incorporé aux cartes produites avec Moringa le réseau des routes issu de la base de données collaborative Open Street Map (OSM).

# 3. Cartes d'occupation du sol produites avec la chaine Moringa

La chaine Moringa (cf. Gaetano et al. 2019) repose sur la combinaison d'une segmentation et d'une classification supervisée orientée objet utilisant l'algorithme Random Forest, entrainée par une base de données d'apprentissage. Elle utilise conjointement :

- 1 image à Très Haute Résolution Spatiale (THRS), de type Spot6/7 ou Pléiades,
- 1 ou plusieurs séries temporelles d'images optiques à Haute Résolution Spatiale (HRS), de type Sentinel-2,
- 1 Modèle Numérique de Terrain (MNT) et ses produits dérivés tels que les pentes,
- 1 base de données de référence (la vérité-terrain).

La Figure 12 présente le schéma général de la méthode, dont les différentes étapes sont les suivantes:

1. la segmentation, qui extrait des objets homogènes de l'image à THRS grâce à l'algorithme de segmentation LSGRM (Large Scale Generic Region Merging) de l'OTB.
2. l'extraction des variables, qui attribue à chaque objet de la segmentation (et de la base de données de référence) la valeur moyenne pour chacune des variables issues de la série temporelle, des images THRS et du MNT.
3. la modélisation, qui analyse avec l'algorithme Random Forest (RF, ou forêt d'arbres de décisions) les variables en entrée pour identifier celles qui seront les plus discriminantes pour caractériser chaque classe de la base de données de référence. Nous avons choisi RF pour sa capacité à traiter un grand nombre de variables. À la fin de cette étape, à chaque objet de la segmentation sera attribué une probabilité d'appartenance à une classe pour produire la classification.
4. la validation, qui consiste à estimer les erreurs de classification dans la carte produite, avec une méthode validation croisée.

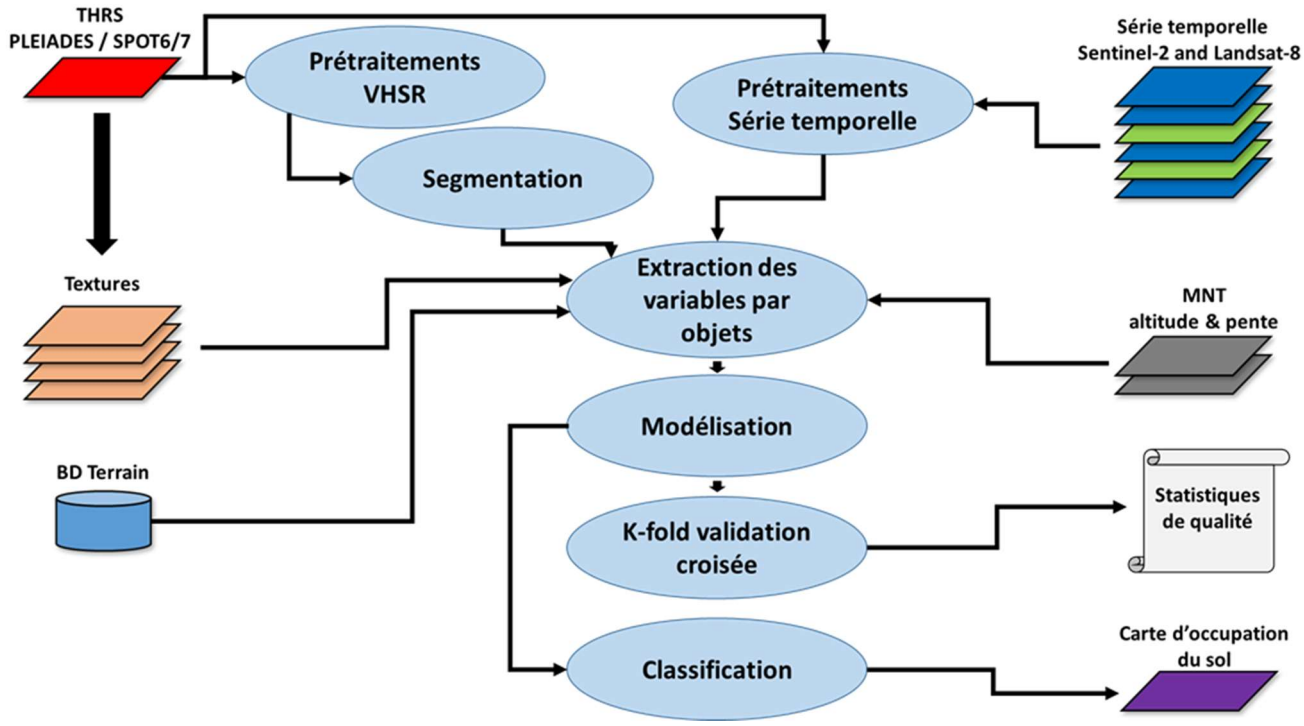


*Figure 12: Schéma général de la chaine Moringa*

# 3.1. SEGMENTATION

La segmentation est une étape importante dans le processus de classification. Elle demande des ressources de calcul importantes. Il s'agit en effet d'extraire de l'image THRS des objets correspondant à des groupes de pixels homogènes. Nous utilisons la méthode décrite dans Baatz (2000). Il s'agit d'une technique ascendante de croissance par paires de régions qui débute par un (ou plusieurs) objet(s) constitué(s) d'un seul pixel de l'image. Les plus petits objets sont ensuite fusionnés de façon itérative dans de plus grands en fonction des critères d'homogénéité utilisant l'information des bandes spectrales de l'image. Le processus s'arrête lorsque l'hétérogénéité des objets candidats à la fusion dépasse l'hétérogénéité maximale acceptée. Pour obtenir un résultat de segmentation adapté à l'étude, l'opérateur agit à la fois sur les critères d'homogénéité et sur le seuil d'hétérogénéité maximale. Les critères finaux sont choisis généralement à l'issue de plusieurs tests les faisant varier indépendamment les uns les autres pour affiner le résultat. La Figure 13 illustre les paramètres à faire varier.

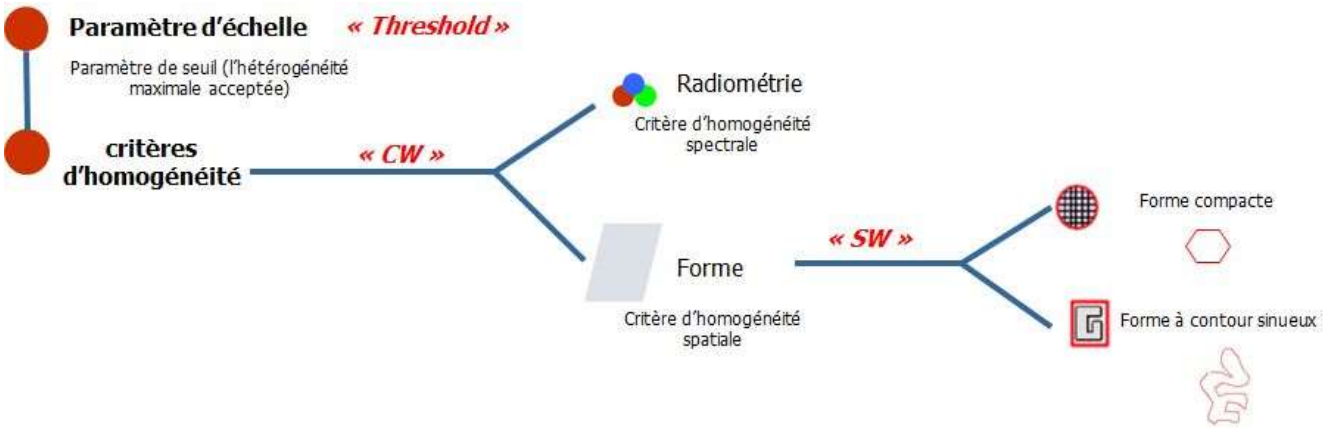


*Figure 13 : Schéma décrivant l'algorithme de segmentation. En rouge et en italique les paramètres à renseigner sur OrfeoToolBox.*

Pour produire cette segmentation, nous utilisons l'outil LSGRM (Generic Region Merging) de l'OTB (adapté de Lassalle et al. (2014) et disponible sur https://gitlab.irstea.fr/remi.cresson/LSGRM).

Le choix des paramètres de segmentation se fait sur des extraits d'images situés dans des zones représentatives afin qu'ils correspondent au mieux à toutes les configurations du paysage. Les 3 paramètres à régler sont les suivants :

- Paramètre d'échelle (« Threshold for the criterion » dans OTB) : ce paramètre joue sur la taille des objets car il correspond à l'hétérogénéité maximale qui est acceptée ;
- Paramètre de poids sur la radiométrie (« CW[3] : weight for the spectral homogeneity » dans OTB) : il s'agit ici de privilégier la radiométrie contre la forme des objets ;
- Paramètre de poids sur la compacité (« SW weight for the spatial homogeneity » sur OTB) : il s'agit ici de privilégier des formes compactes contre les contours sinueux.

Les paramètres finalement retenus pour produire la segmentation respectivement sur l'image Spot6/7 et l'image Pléiades sont indiqués au Tableau 4. Les résultats de la segmentation correspondante sont illustrés respectivement à la Figure 14 et à la Figure 15.

*Tableau 4 : Paramètres retenus pour la segmentation des images à THRS dans Moringa.*

| | Threshold | CW | SW |
|---|---|---|---|
| SPOT 6/7 | 230 | 0,7 (soit 0,3 dans Moringa) | 0,7 |
| Pléiades | 300 | 0,7 (soit 0,3 dans Moringa) | 0.7 |

[3] Dans Moringa, le paramètre à spécifier est le poids de la forme contre la radiométrie, ce qui correspond à 1-CW par rapport au paramètre CW de l'application OTB LSGRM.

*Figure 14: Illustration des objets issus de la segmentation. A gauche : extrait de l'image Spot 6/7 seule ; à droite : la même image sur laquelle sont superposés les polygones obtenus lors de la segmentation. Contient des informations © Airbus DS 2020, © Production IRD, Inrae, IGN, tous droits réservés.*

*Figure 15: Illustration des objets issus de la segmentation. À gauche : extrait de l'image Pléiades seule ; à droite : la même image sur laquelle sont superposés les polygones obtenus lors de la segmentation. Contient des informations © CNES 2020 - distribution AIRBUS DS.*

# 3.2. ÉTUDE DES VARIABLES MISES EN ŒUVRE

Les variables d'entrée de la chaine Moringa sont donc les réflectances et indices spectraux de la série temporelle Sentinel-2 et des images pansharpened Spot6/7 ou Pléiades, les indices de texture calculés sur l'image panchromatique Spot6/7 ou Pléiades et, enfin, l'altitude et les pentes issues du MNT. Au total 391 variables ont été utilisées pour produire les classifications à partir de l'image Spot6/7, 395 pour Pléiades (4 indices de texture supplémentaires), listées dans le Tableau 5.

A l'issue de la classification, Moringa classe les variables d'entrée en fonction de leur importance dans le processus d'extraction des différentes classes de la nomenclature. Les figures suivantes représentent cette information sous forme de graphiques qui permettent de mieux comprendre cette information :

- Il apparait clairement que les variables de topographie (altitudes et pentes, en noir sur les graphiques) apportent une information importante quel que soit le niveau de la nomenclature.
- Dans la classification au niveau de nomenclature « 0 » (Figure 16 et Figure 17), ce sont les variables issues de la série temporelle Sentinel-2 qui ont le plus d'influence dans le processus, et en priorité les indices spectraux. Ceci s'explique par la grande différence spectrale entre les différentes occupations du sol structurant le paysage, que sont l'eau, le sol nu et la végétation.

- Pour les niveaux de nomenclature plus précis, c'est-à-dire les niveaux « 1 » (Figure 18 et Figure 19) puis « 2 » (Figure 20 et Figure 21), l'importance de Sentinel-2 diminue très rapidement, alors que les variables issues des images THRS prennent de plus en plus d'importance. Cela confirme l'intérêt de la THRS pour discriminer des classes d'occupation du sol plus fines. Il s'agit principalement de distinguer des espèces végétales entre elles, en particulier ligneuses/non ligneuses.
- Ce sont essentiellement les textures qui apportent l'information issue des données à THRS. Elles dépassent même l'importance de l'information des données sentinel-2 pour SPOT6/7 dès le niveau de nomenclature « 1». Ceci exprime bien le caractère essentiel de l'information structurelle intra-parcellaire pour distinguer différentes occupation végétales ligneuses qui ne se différencient que par leur agencement dans l'espace, donc par leur texture dans l'image.

*Tableau 5 : Liste des 391 (pour SPOT6/7) / 395 (pour Pléiades) variables utilisées dans le processus de classification. [1] B2: bleu; B3: vert; B4: rouge; B5, B6 et B7: red -edge; B8: PIR; B8A: PIR étroit; B11 et B12: MIR*

| | HRS (Sentinel-2) | THRS (Spot 6/7 ou Pléiades) | MNT | Total |
|---|---|---|---|---|
| **Réflectances multitemporelles (23 images)** | B2, B3, B4, B5, B6, B7, B8, B8A, B11, B12[1] | | | 230 |
| **Indices radiométriques multitemporels (23 images)** | NDVI, NDWI, Brightness, SWNDVI, MNDWI, NDRE | | | 138 |
| **Réflectances** | | B, V, R, PIR, PAN | | 5 |
| **Indices radiométriques** | | NDVI, RVI, NDWI2, BI2 | | 4 |
| **Indices de texture** | | Energy, Contrast, Correlation et Variance avec radius de 1, 5, 11 (+ 21 /Pléiades) | | 12 (16 /Pléiades) |
| **Topographie** | | | altitude, pentes | 2 |

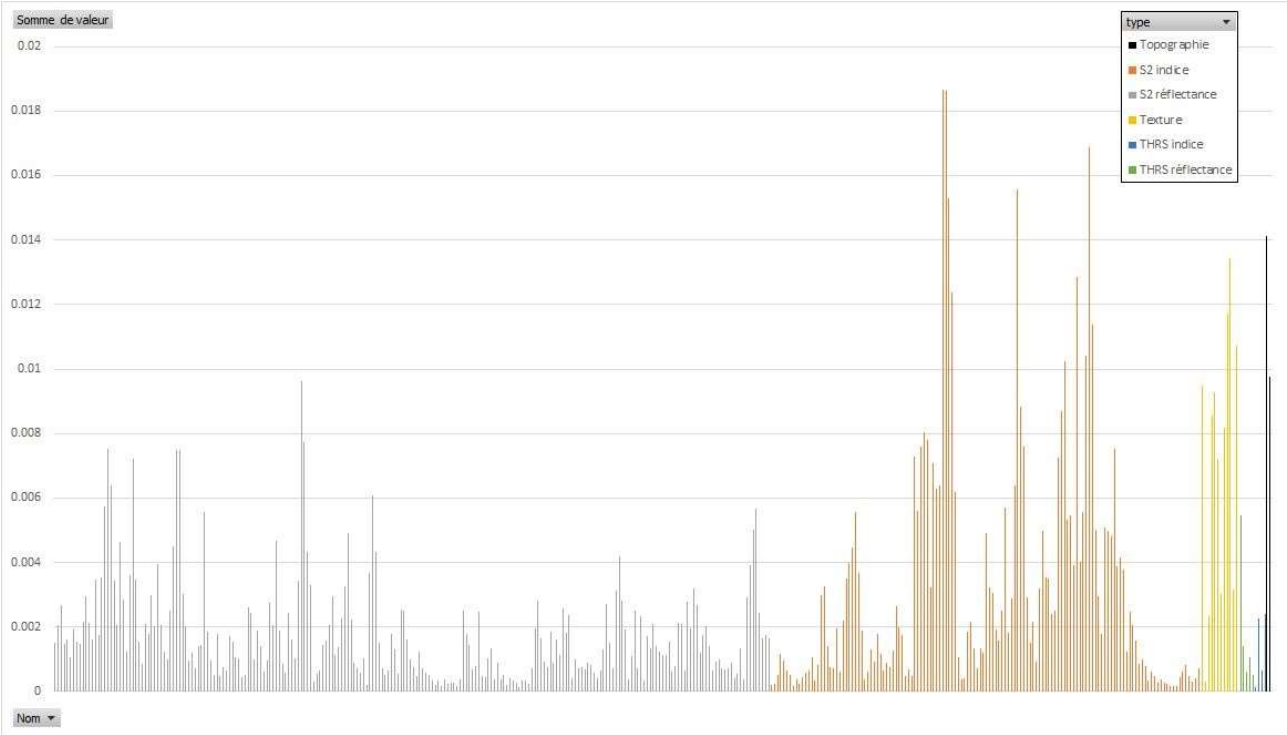


*Figure 16 : Graphique de l'importance des variables pour la classification au niveau « 0 » produite avec Spot 6/7. Les variables sont classées par type.*

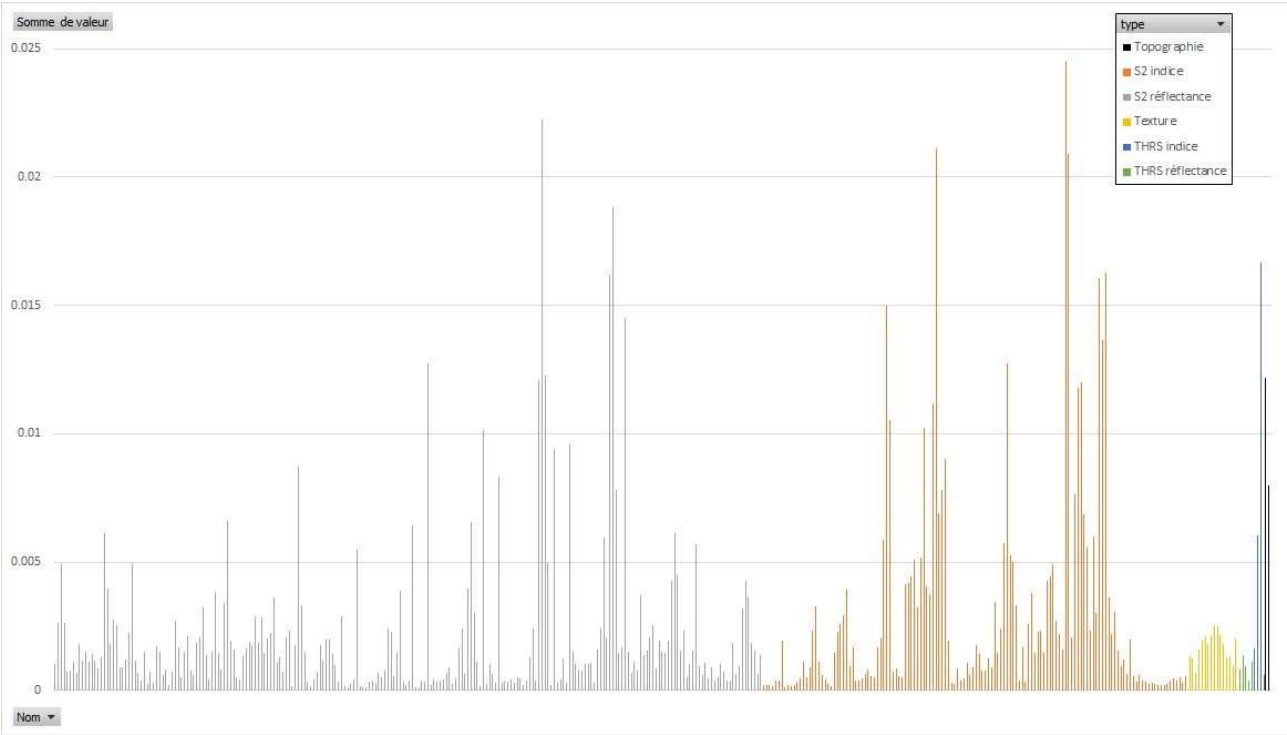


*Figure 17 : Graphique de l'importance des variables pour la classification au niveau « 0 » produite avec Pléiades. Les variables sont classées par type.*

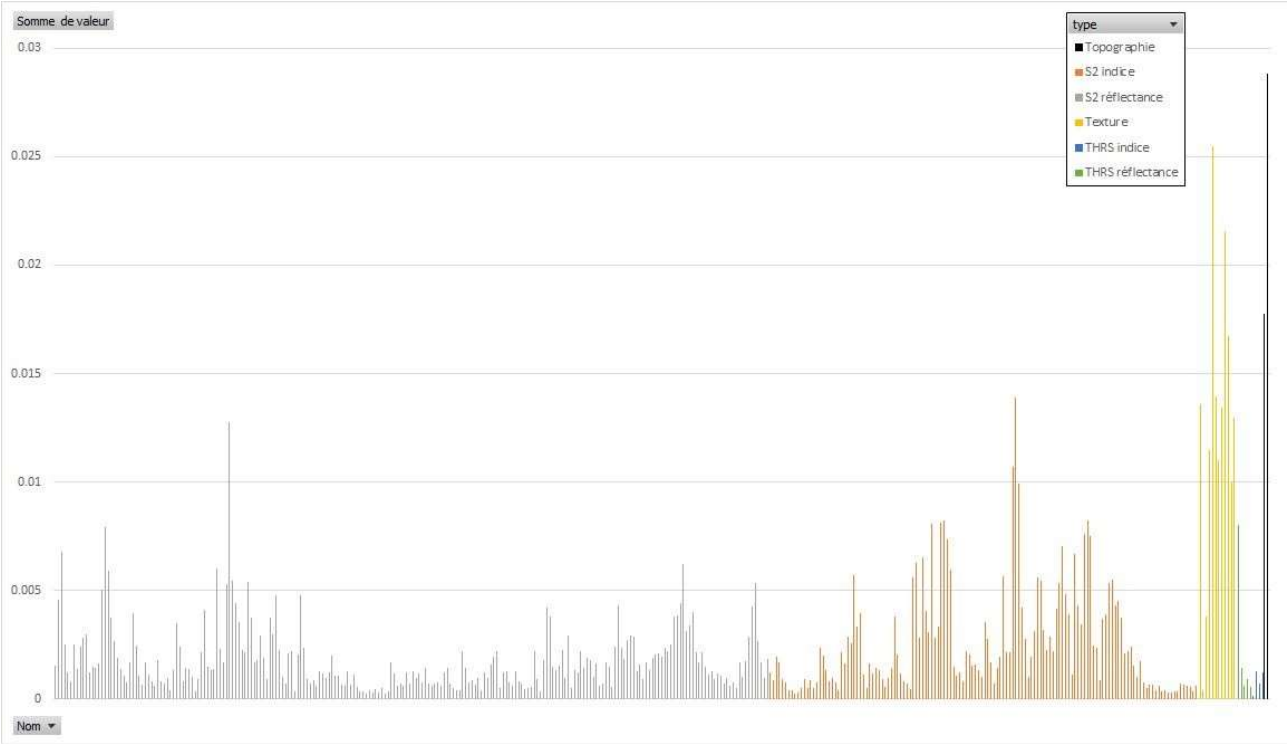


*Figure 18 : Graphique de l'importance des variables pour la classification au niveau « 1 » produite avec Spot 6/7. Les variables sont classées par type.*

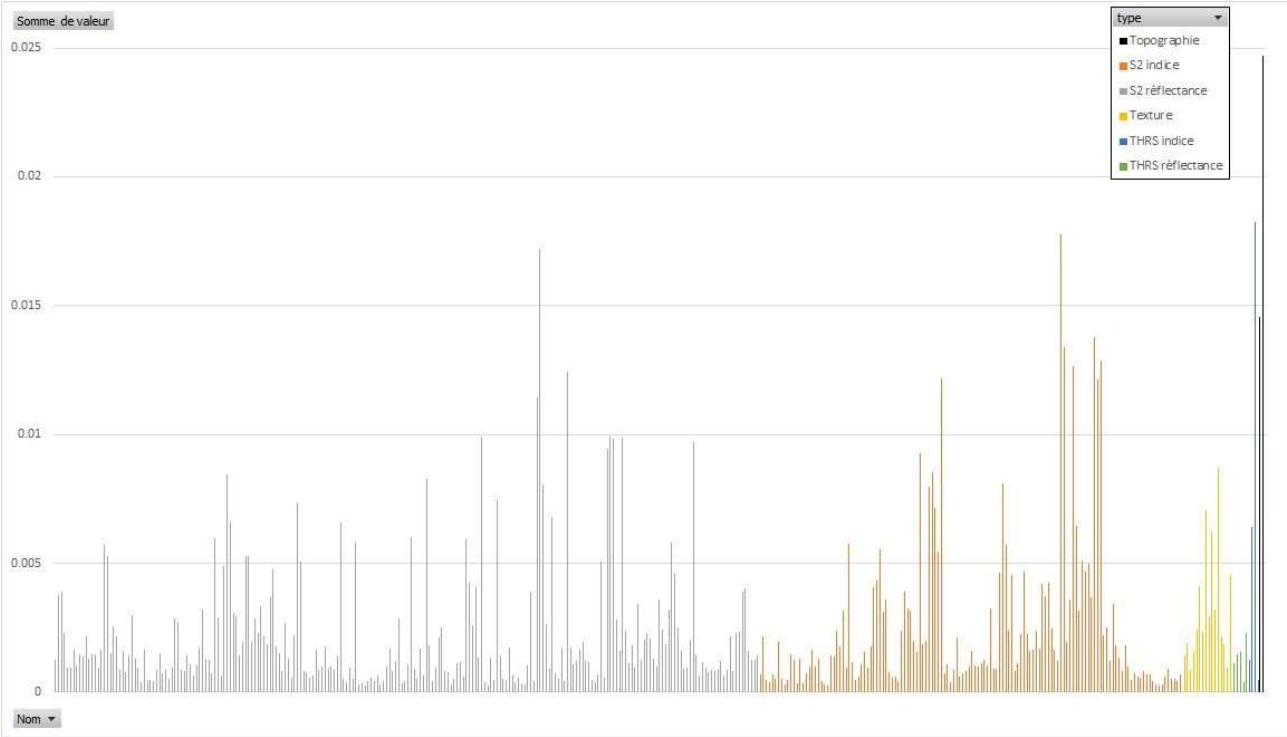


*Figure 19 : Graphique de l'importance des variables pour la classification au niveau « 1 » produite avec Pléiades. Les variables sont classées par type.*

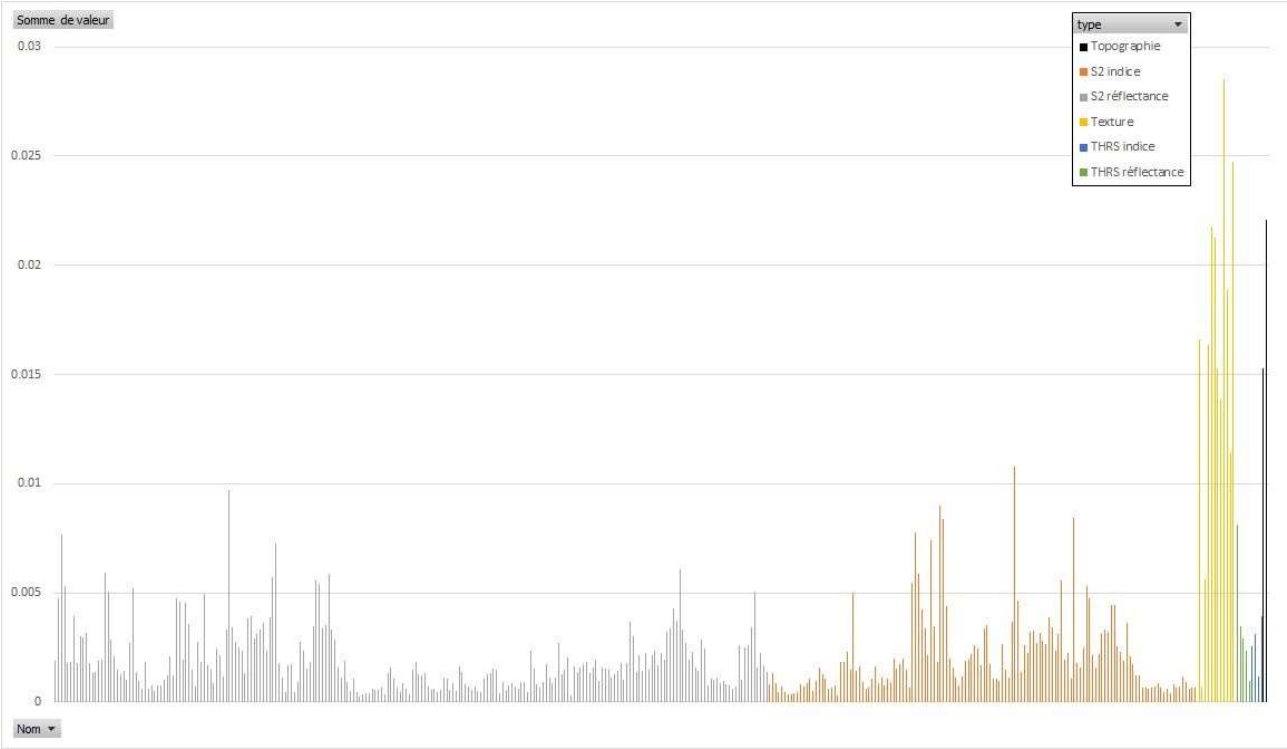


*Figure 20 : Graphique de l'importance des variables pour la classification au niveau « 2 » produite avec Spot 6/7. Les variables sont classées par type.*

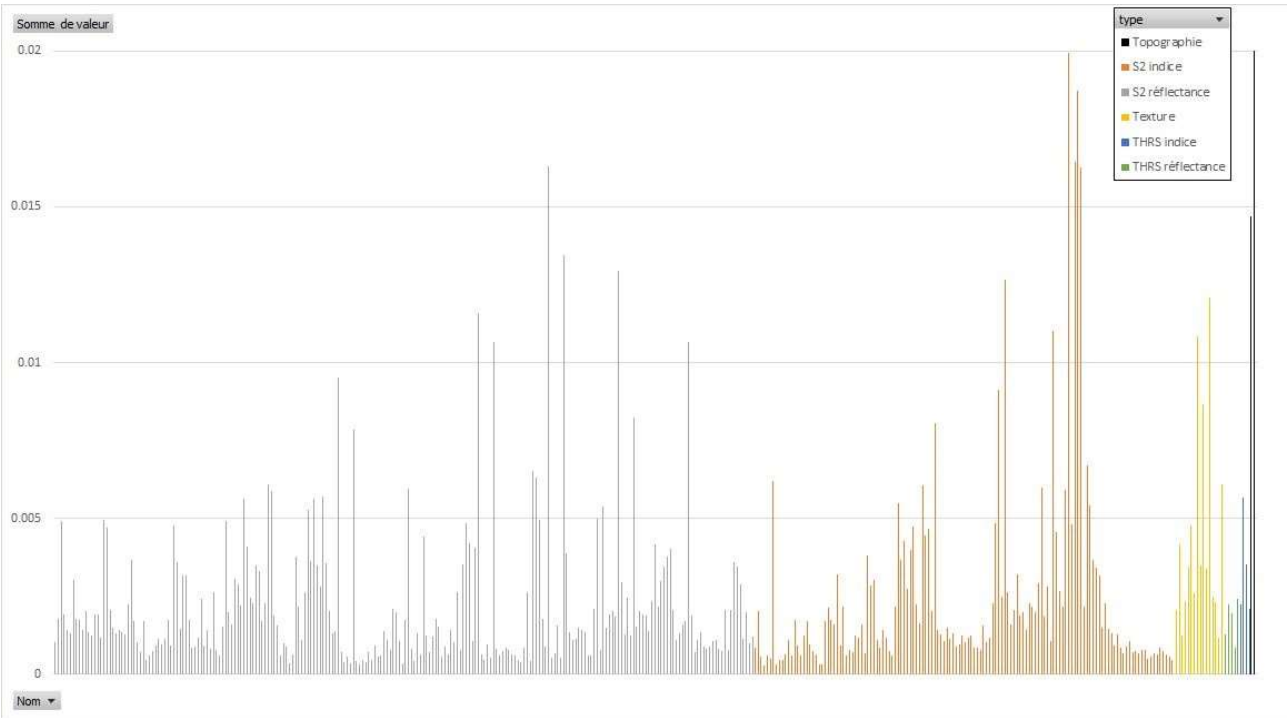


*Figure 21 : Graphique de l'importance des variables pour la classification au niveau « 2 » produite avec Pléiades. Les variables sont classées par type.*

## 3.3. Post-traitements

### 3.3.1. Routes Open Street Map

Les classifications d'images ne permettent pas d'extraire de façon fiable les routes. Bien sûr, celles-ci sont visibles sur les images à THRS comme Spot6/7 ou Pléiades mais, quand elles sont bordées de bâtiments ou d'arbres, les polygones sont interrompus. La résolution de Sentinel-2 ajoute une difficulté pour les classer convenablement car un pixel « pur » de route n'existe pas (les routes sont plus petites que la taille du pixel). Pour intégrer une classe « routes » aux cartes produites par la chaine Moringa, nous avons donc décidé d'y incorporer les routes issues de la base de données participative Open Street Map (OSM) : https://www.openstreetmap.org.

Nous avons donc téléchargé la couche « routes », puis modifié la projection pour la passer en UTM 18N (code epsg 32618). Puis, nous avons sélectionné les routes principales en conservant uniquement les types suivants : « primary », « residential », « secondary » et « tertiary » de la colonne « Highway ».

Ensuite, sous le logiciel QGIS, nous avons converti en polygones cette couche des routes, en calculant une zone tampon (buffer) de 5m de part et d'autre de la ligne (soit 10m au total), ordre de grandeur d'une largeur de route carrossable. Puis les polygones ont été convertis au format image en utilisant la résolution correspondante sur la carte (50cm pour Pléiades et 1,5m pour Spot6/7).

Enfin, l'incorporation finale des routes dans chaque carte a été réalisée avec l'OTB.

### 3.3.2. Lissage

Une commande dite de régularisation des cartes a été appliquée afin de lisser les transitions entre chaque classe. Ce lissage permet d'atténuer les excroissances souvent visibles entre les unités de classes différentes, effet produit par la segmentation à THRS. Cette étape permet de diminuer le nombre de nœuds de la couche vecteur lors de l'étape de vectorisation de donc diminuer la taille de la couche vecteur.

Afin de ne pas affecter l'information produite, c'est-à-dire d'altérer la précision de classification, ce lissage est effectué avec un radius de 1 (soit une fenêtre glissante de 3x3 pixels) avec la commande « Classification Map Regularization » de l'OTB.

### 3.3.3. Vectorisation

Le passage du format raster (image) au format vecteur est réalisé avec le logiciel Arc Gis car il propose de simplifier les polygones pour encore une fois limiter le nombre de nœuds sur les lignes et ainsi diminuer la taille de la couche finale. Cette opération engendre peu de modification dans la précision des contours d'unités cartographiée mais allège la taille de la couche vectorielle produite.

Enfin une jointure permet d'intégrer dans la table finale les données attributaires telles que les codes et les noms des classes.

# 3.4. LES CARTES OBTENUES

Toutes les couches produites sont disponibles sur le Dataverse du Cirad. Elles sont en accès libre sous licence Licence Creative Commons (LCC, BY : http://creativecommons.org/licenses/by/4.0 ).

Voici la référence et le lien pour télécharger la carte :

**Dupuy, Stéphane; Lelong, Camille; Gaetano, Raffaele, 2021, "Nippes - Haïti - 2020, Land cover map", doi:10.18167/DVN1/QOLTTE, CIRAD Dataverse, V1**

Ces cartes sont illustrées sur les douze figures suivantes, pour l'ensemble de la zone traitée dans cette étude et zoomées sur les trois bassins versant d'intervention du projet TPR (Baconois, Rivière Froide et bondeau), pour les deux jeux de données à THRS différents et pour les trois niveaux de nomenclature proposés :

| | | Source THRS | |
|---|---|---|---|
| | | Pléiades | SPOT6/7 |
| Niveau de nomenclature | niveau 2 | Figure 22, Figure 23 | Figure 28, Figure 29 |
| | niveau 1 | Figure 24, Figure 25 | Figure 30, Figure 31 |
| | niveau 0 | Figure 26, Figure 27 | Figure 32, Figure 33 |

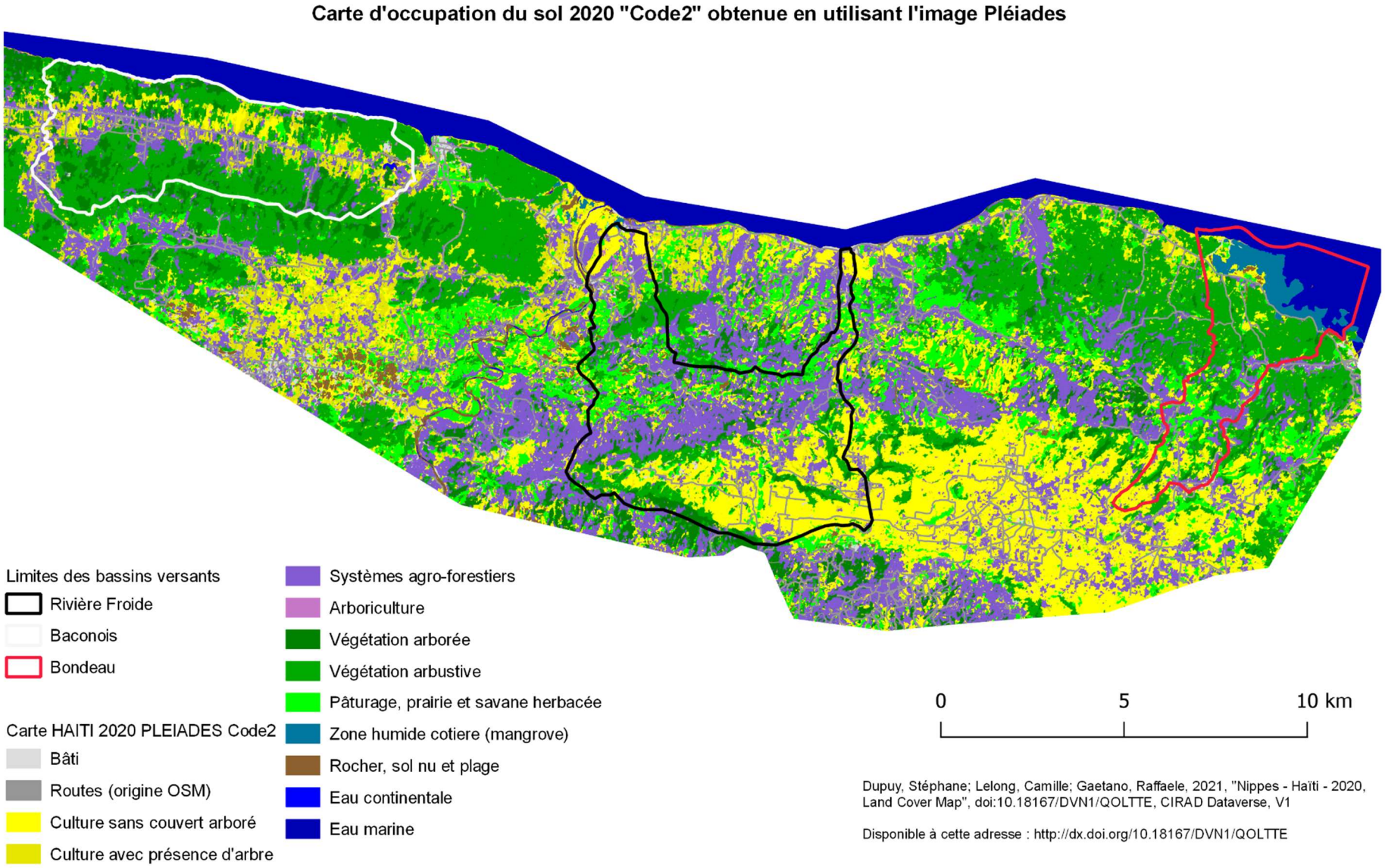


*Figure 22 : Carte d'occupation du sol en 2020, obtenue en utilisant l'image Pléiades au niveau de nomenclature « 2 ». Référence : Dupuy, S.; Lelong, C.; Gaetano, R., 2021, "Nippes - Haïti - 2020, Land Cover Map", doi:10.18167/DVN1/QOLTTE, CIRAD Dataverse,V1.*

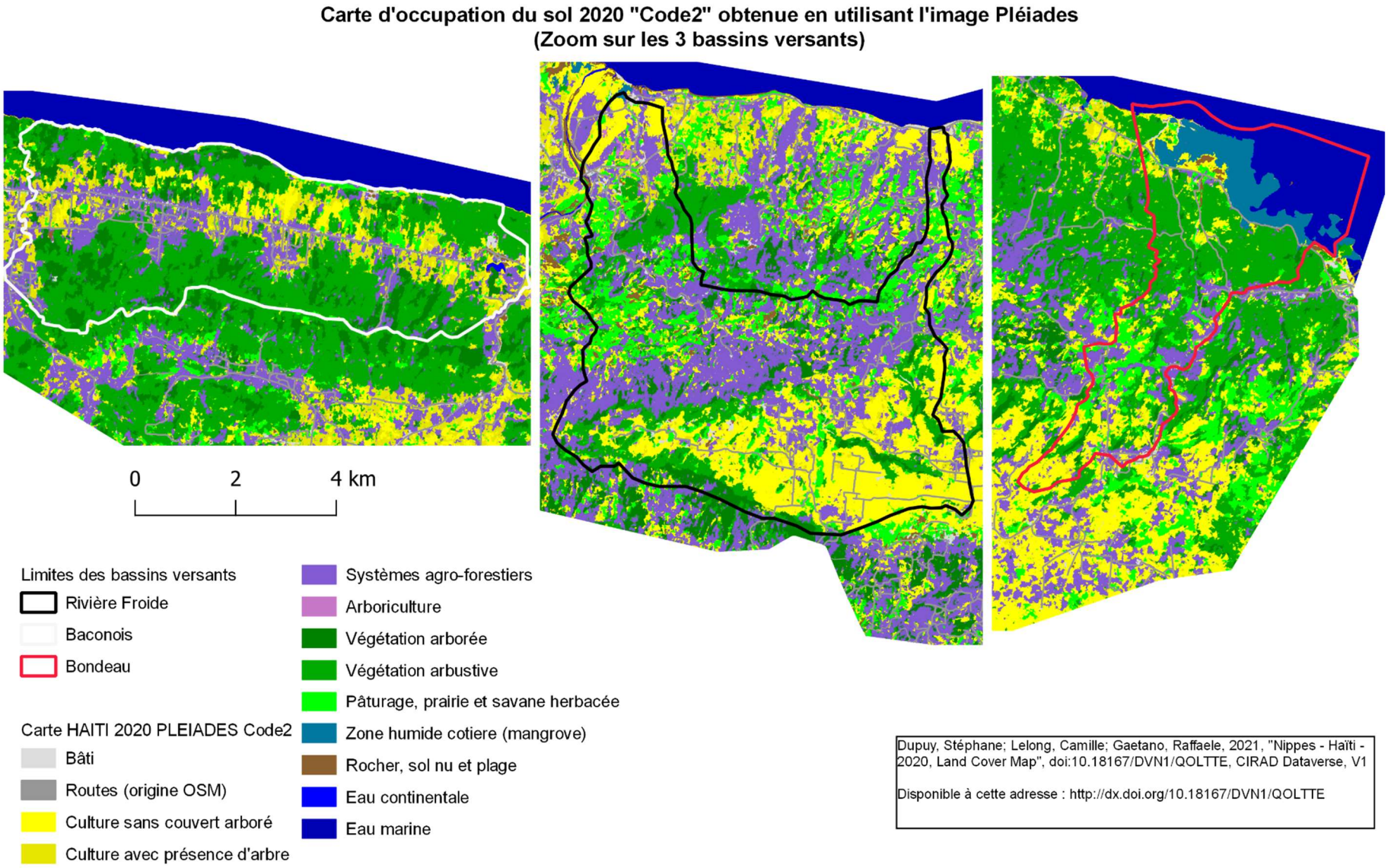


*Figure 23 : Carte d'occupation du sol en 2020, obtenue en utilisant l'image Pléiades au niveau de nomenclature « 2 ». Zooms sur les 3 bassins versant étudiés avec, de gauche à droite : Baconois, Rivière Froide et Bondeau. Référence : Dupuy, S.; Lelong, C.; Gaetano, R., 2021, "Nippes - Haïti - 2020, Land Cover Map", doi:10.18167/DVN1/QOLT*

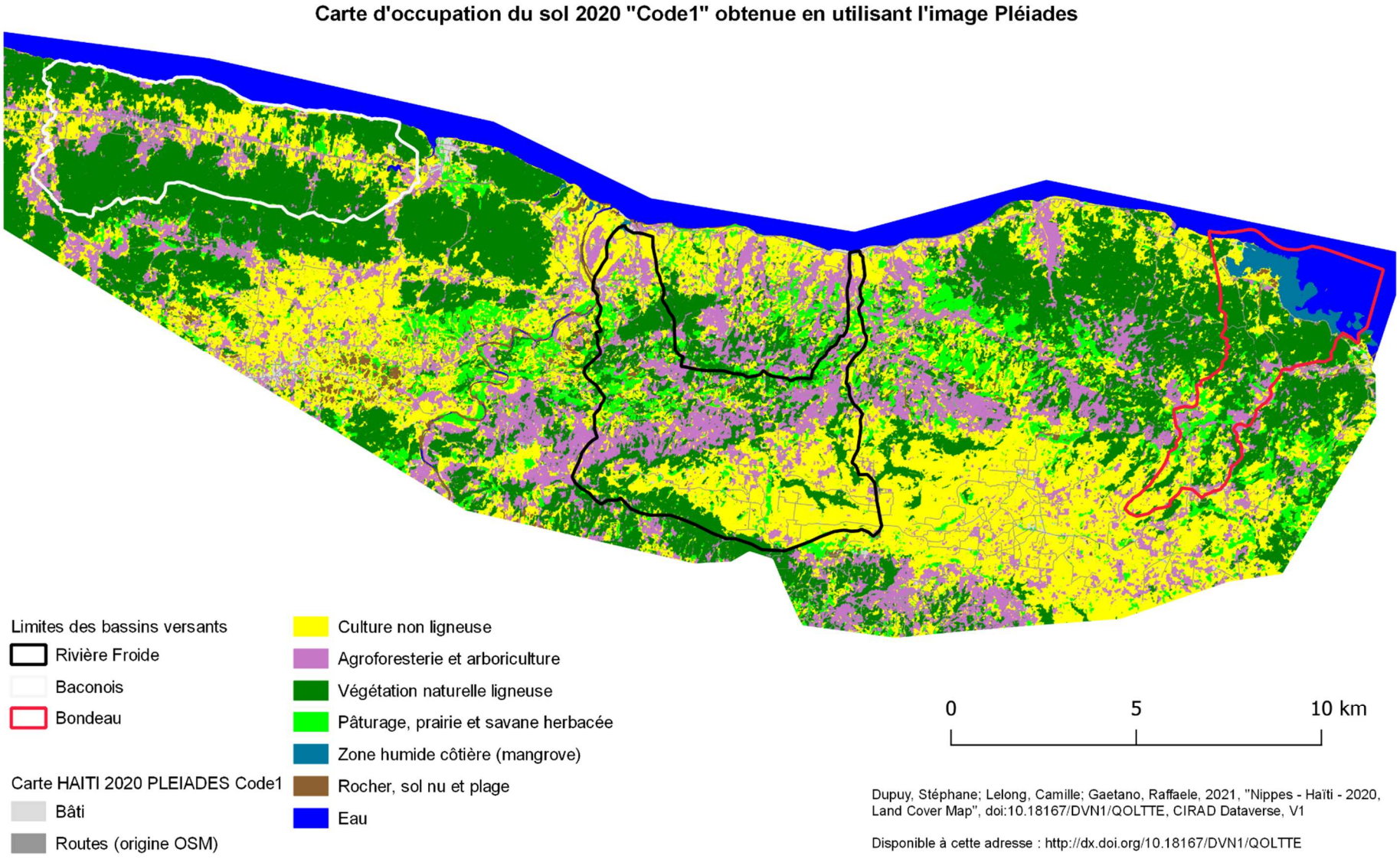


*Figure 24 : Carte d'occupation du sol en 2020, obtenue en utilisant l'image Pléiades au niveau de nomenclature « 1 ». Référence : Dupuy, S.; Lelong, C.; Gaetano, R., 2021, "Nippes - Haïti - 2020, Land Cover Map", doi:10.18167/DVN1/QOLTTE, CIRAD Dataverse,V1*

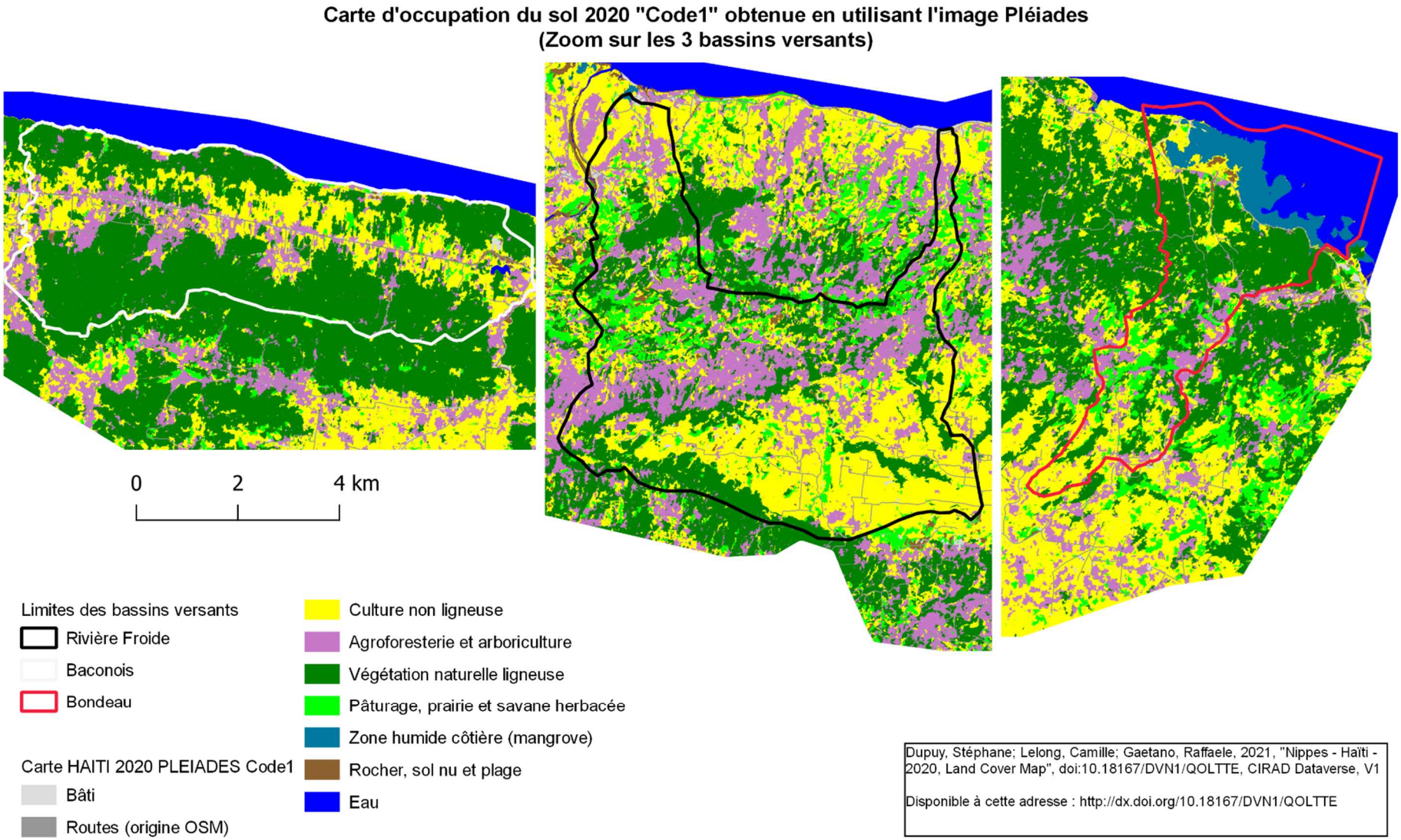


*Figure 25 : Carte d'occupation du sol en 2020, obtenue en utilisant l'image Pléiades au niveau de nomenclature « 1 ». Zooms sur les 3 bassins versant étudiés avec, de gauche à droite : Baconois, Rivière Froide et Bondeau. Référence : Dupuy, S.; Lelong, C.; Gaetano, R., 2021, "Nippes - Haïti - 2020, Land Cover Map", doi:10.18167/DVN1/QOLTTE, CIRAD Dataverse,V1*

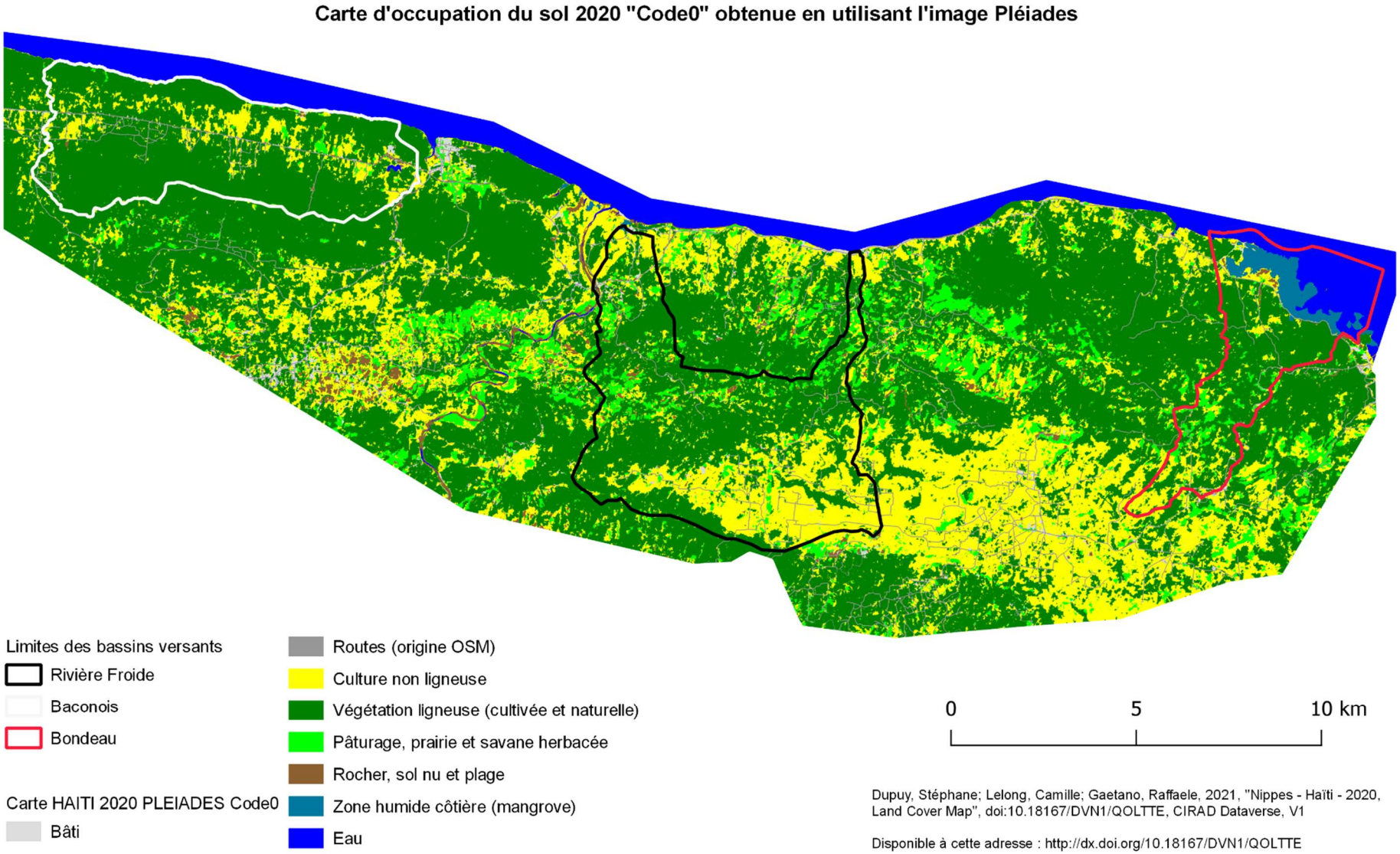


*Figure 26 : Carte d'occupation du sol en 2020, obtenue en utilisant l'image Pléiades au niveau de nomenclature « 0 ». Référence : Dupuy, S.; Lelong, C.; Gaetano, R., 2021, "Nippes - Haïti - 2020, Land Cover Map", doi:10.18167/DVN1/QOLTTE, CIRAD Dataverse,V1*

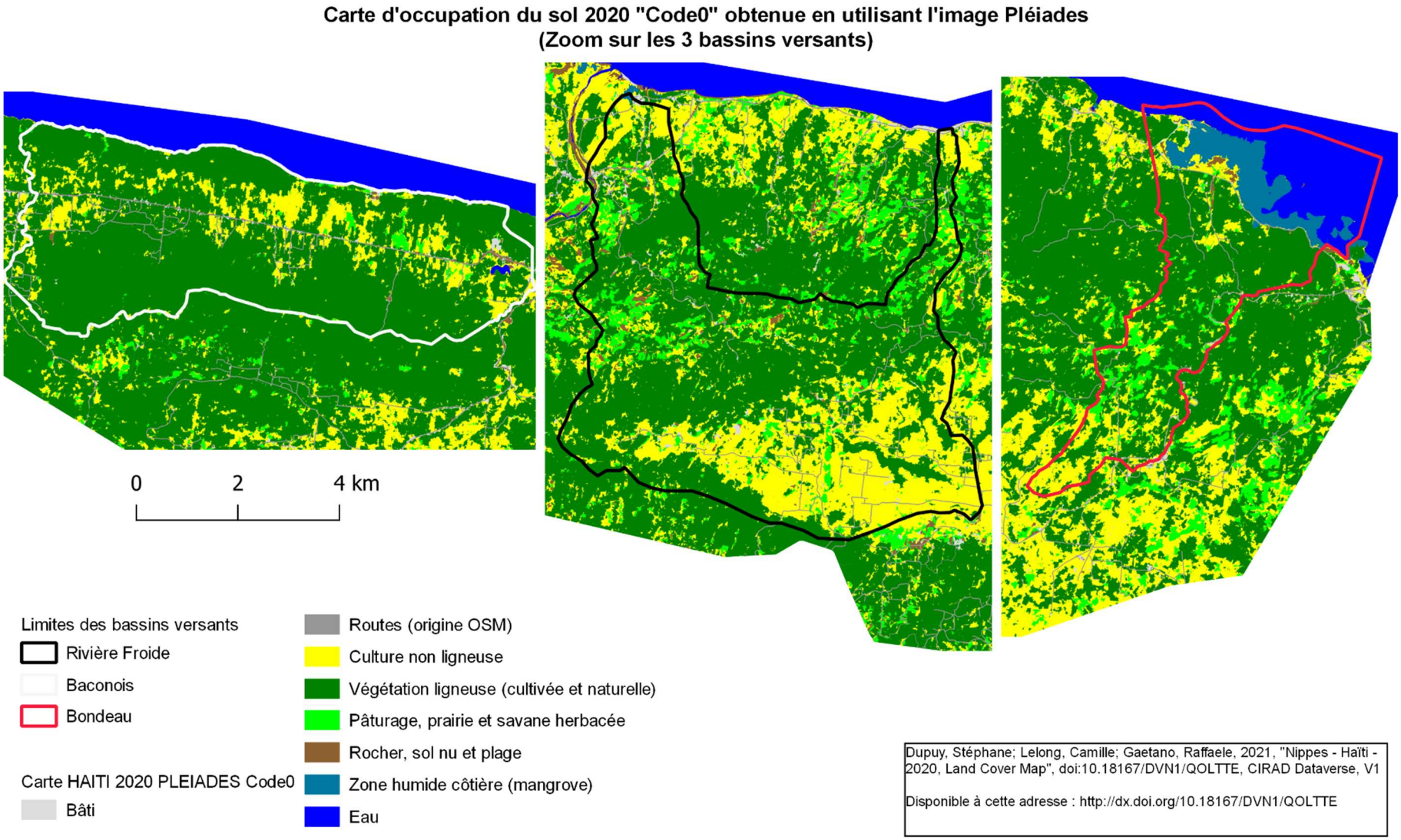


*Figure 27 : Carte d'occupation du sol en 2020, obtenue en utilisant l'image Pléiades au niveau de nomenclature « 1 ». Zooms sur les 3 bassins versant étudiés avec, de gauche à droite : Baconois, Rivière Froide et Bondeau. Référence : Dupuy, S.; Lelong, C.; Gaetano, R., 2021, "Nippes - Haïti - 2020, Land Cover Map", doi:10.18167/DVN1/QOLTTE, CIRAD Dataverse,V1*

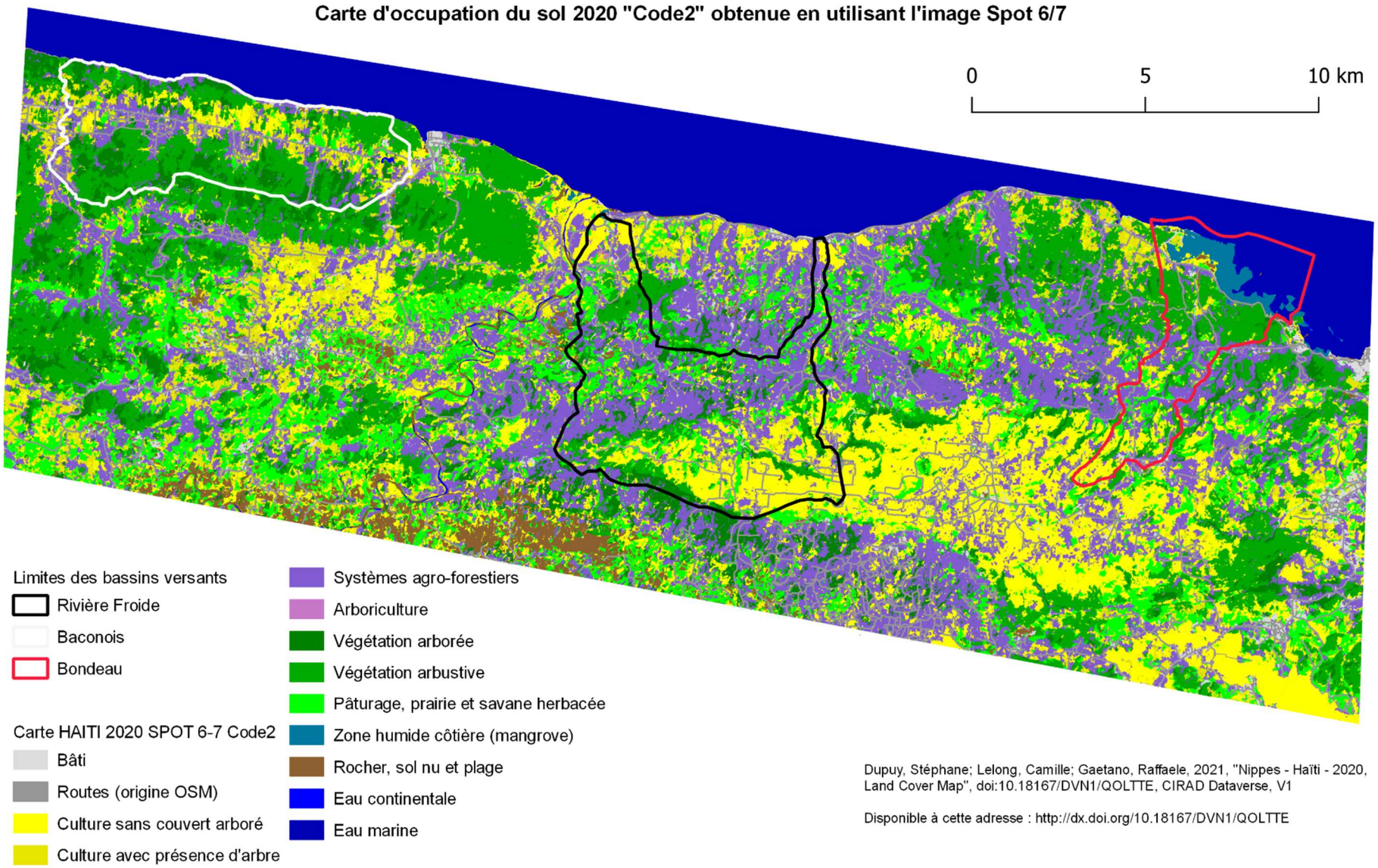


*Figure 28 : Carte d'occupation 2020 du sol obtenue en utilisant l'image Spot 6/7 au niveau de nomenclature « 2 ». Référence : Dupuy, Stéphane; Lelong, Camille; Gaetano, Raffaele, 2021, "Nippes - Haïti - 2020, Land Cover Map", doi:10.18167/DVN1/QOLTTE, CIRAD Dataverse,V1*

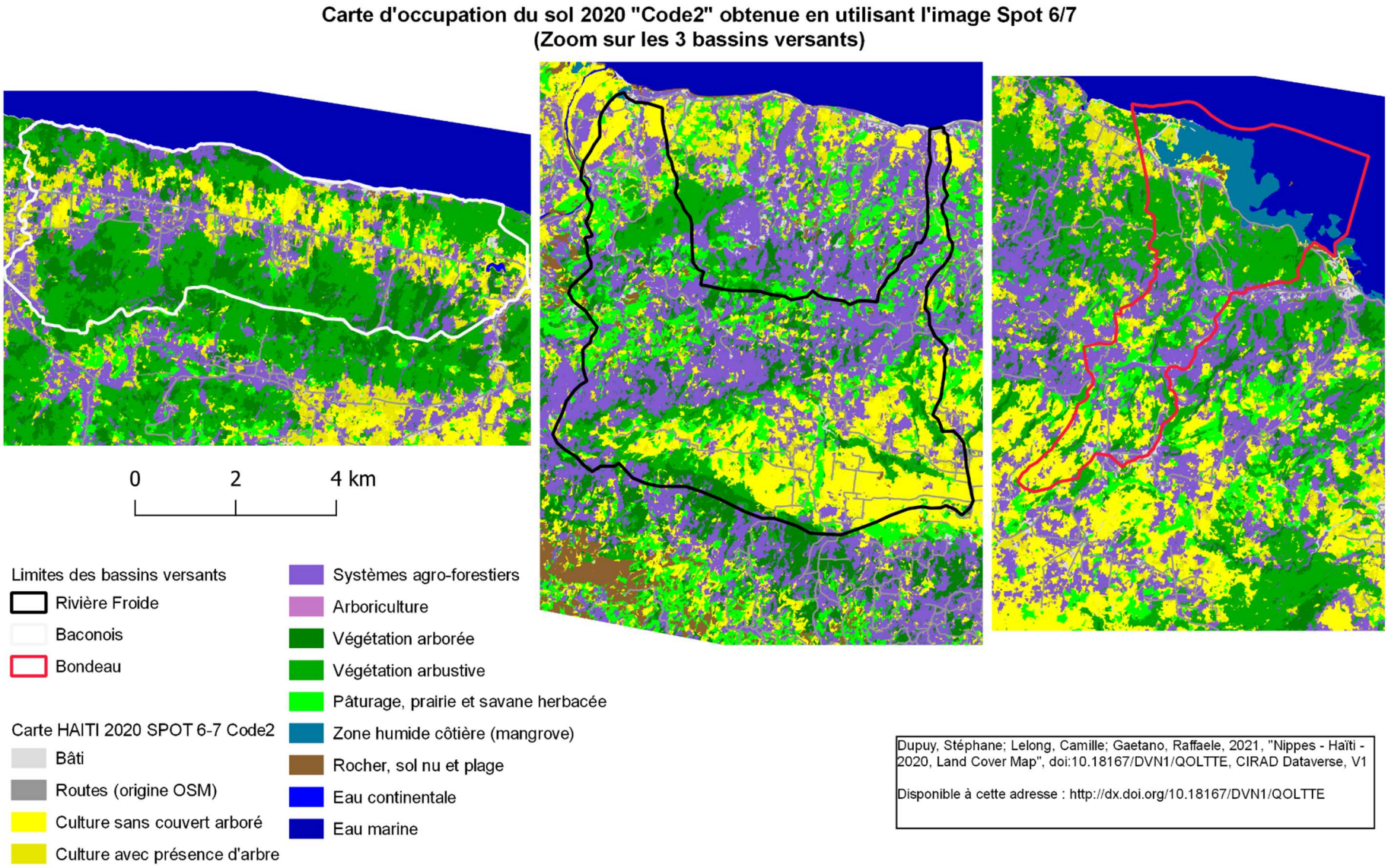


*Figure 29 : Carte d'occupation 2020 du sol, obtenue en utilisant l'image Spot 6/7 au niveau de nomenclature « 2 » Zooms sur les 3 bassins versant étudiés avec, de gauche à droite : Baconois, Rivière Froide et Bondeau. Référence : Dupuy, S.; Lelong, C.; Gaetano, R., 2021, "Nippes - Haïti - 2020, Land Cover Map", doi:10.18167/DVN1/QOLTTE, CIRAD Dataverse,V1.*

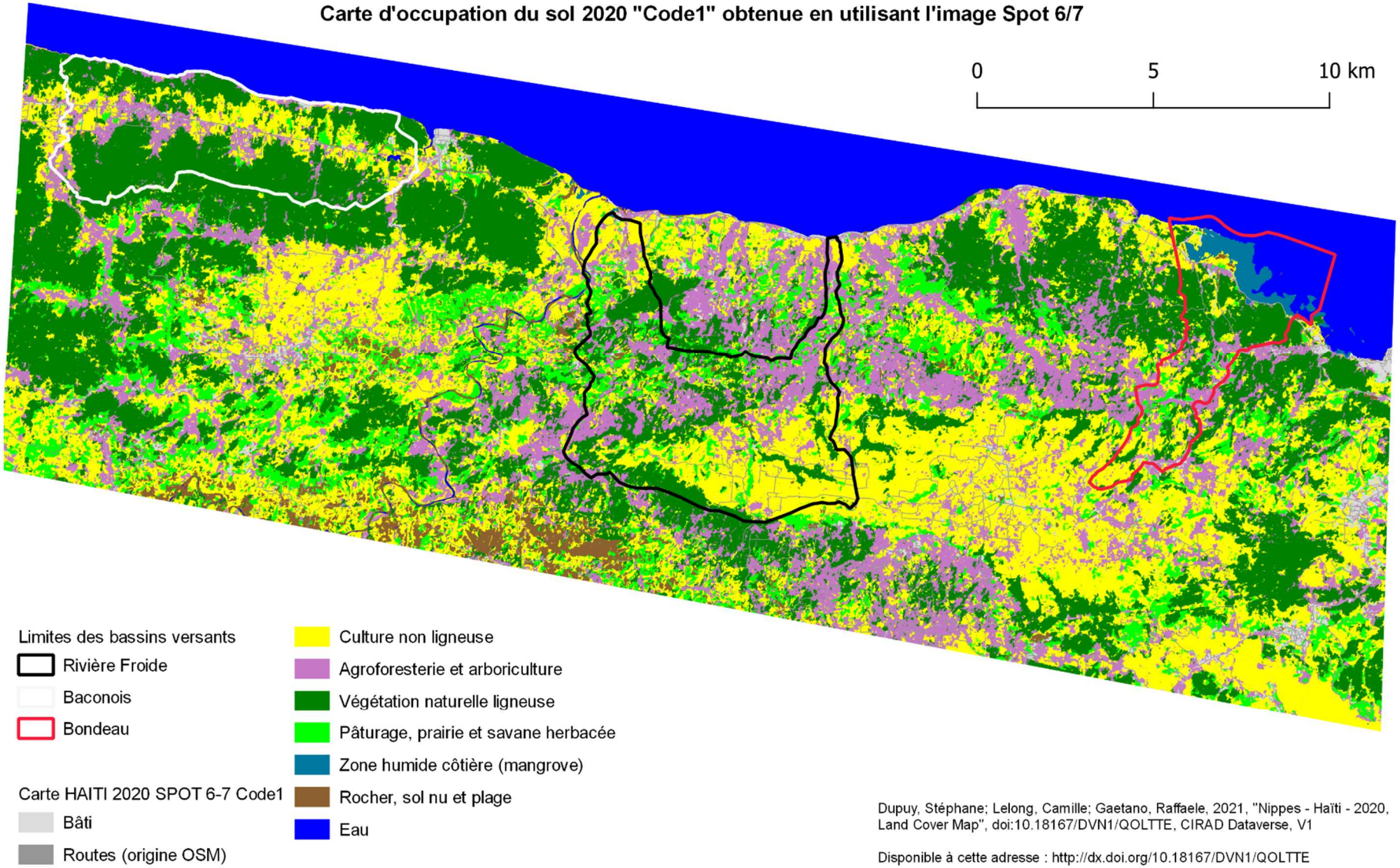


*Figure 30 : Carte d'occupation 2020 du sol, obtenue en utilisant l'image Spot 6/7 au niveau de nomenclature « 1 ». Référence : Dupuy, S.; Lelong, C.; Gaetano, R., 2021, "Nippes - Haïti - 2020, Land Cover Map", doi:10.18167/DVN1/QOLTTE, CIRAD Dataverse,V1.*

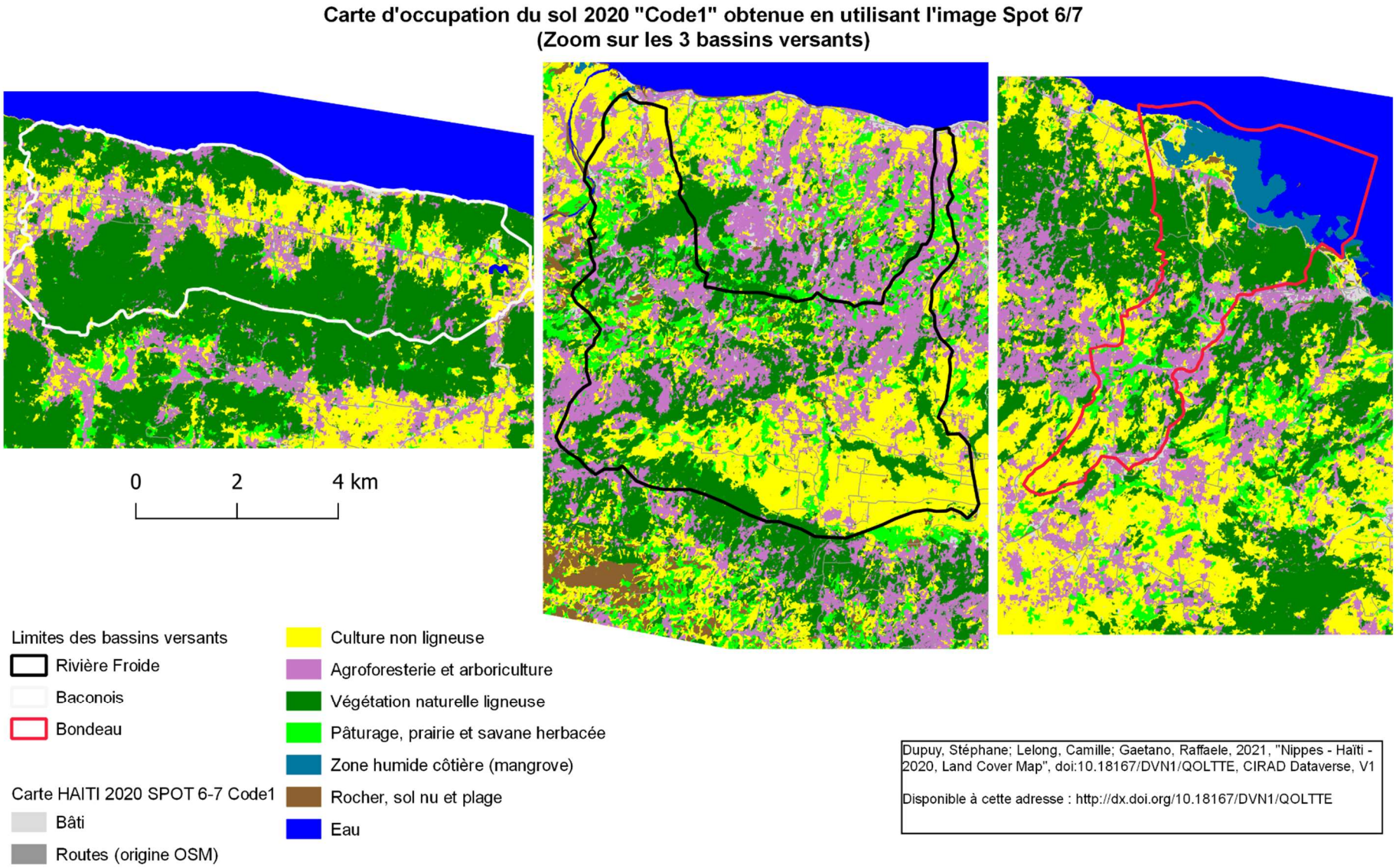


*Figure 31 : Carte d'occupation 2020 du sol, obtenue en utilisant l'image Spot 6/7 au niveau de nomenclature « 1 ». Zooms sur les 3 bassins versant étudiés avec, de gauche à droite : Baconois, Rivière Froide et Bondeau. Référence : Dupuy, S.; Lelong, C.; Gaetano, R., 2021, "Nippes - Haïti - 2020, Land Cover Map", doi:10.18167/DVN1/QOLTTE, CIRAD Dataverse,V1.*

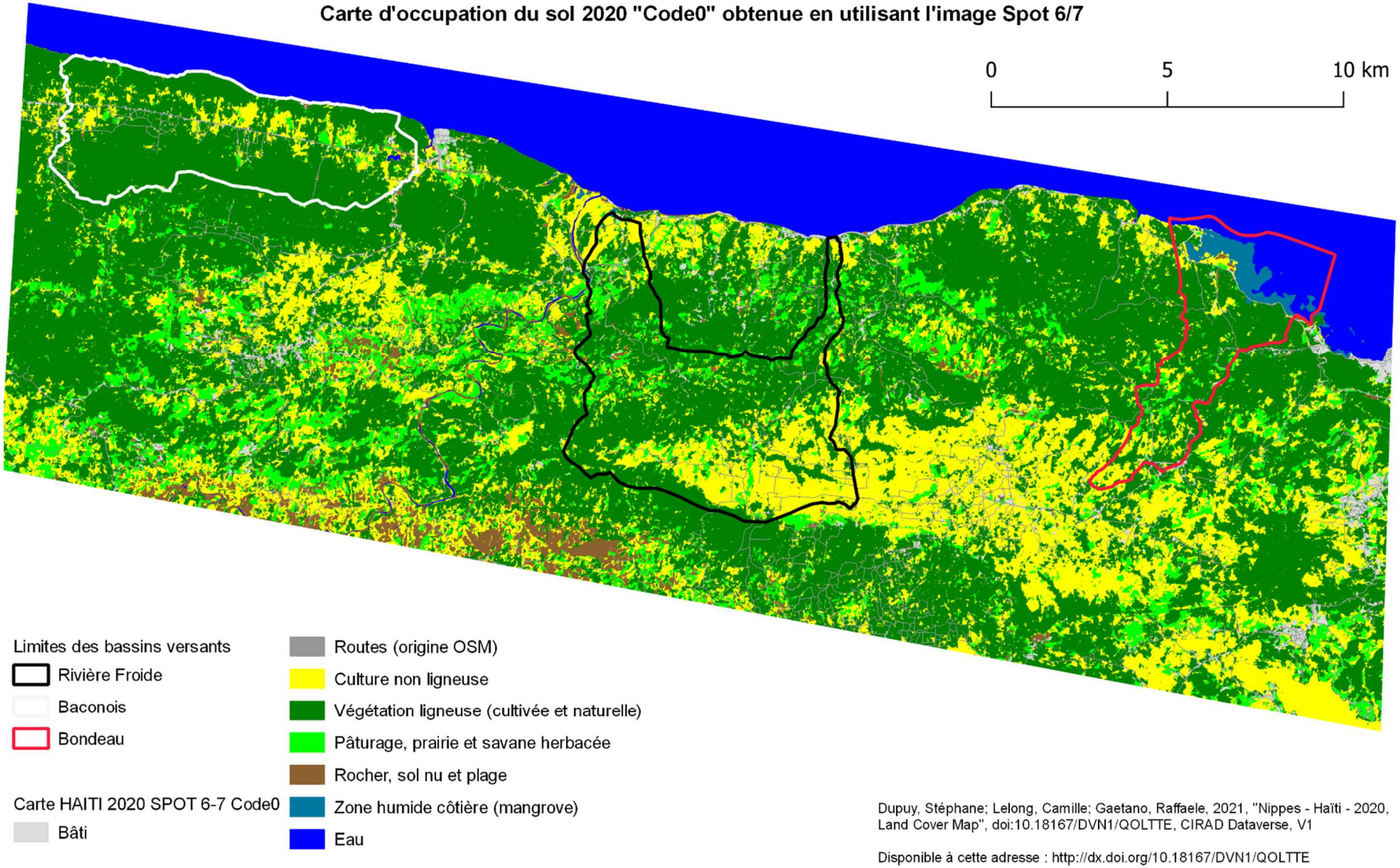


*Figure 32 : Carte d'occupation 2020 du sol, obtenue en utilisant l'image Spot 6/7 au niveau de nomenclature « 0 ». Référence : Dupuy, S.; Lelong, C.; Gaetano, R., 2021, "Nippes - Haïti - 2020, Land Cover Map", doi:10.18167/DVN1/QOLTTE, CIRAD Dataverse,V1.*

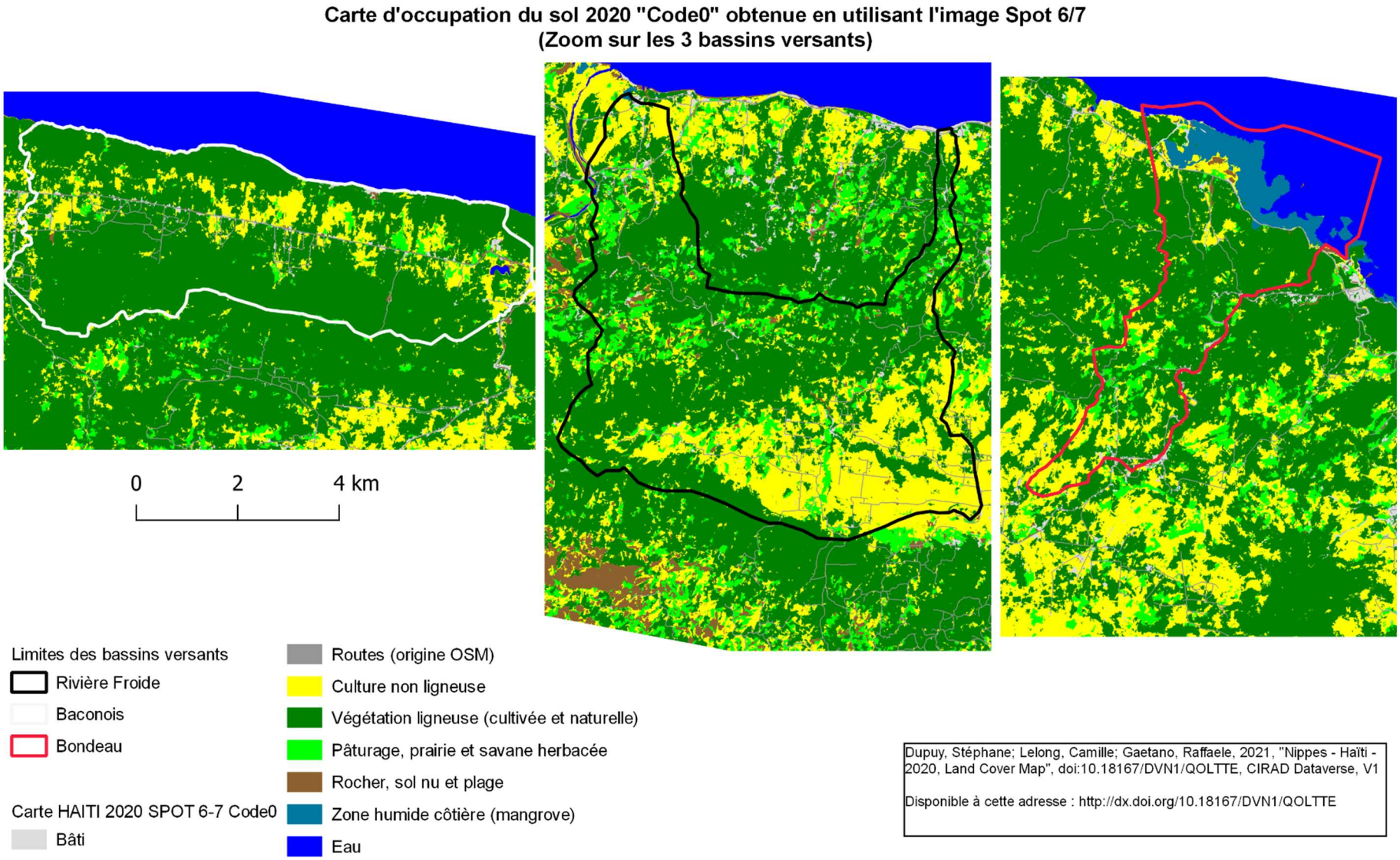


*Figure 33 : Carte d'occupation 2020 du sol, obtenue en utilisant l'image Spot 6/7 au niveau de nomenclature « 0 ». Zooms sur les 3 bassins versant étudiés avec, de gauche à droite : Baconois, Rivière Froide et Bondeau. Référence : Dupuy, S.; Lelong, C.; Gaetano, R., 2021, "Nippes - Haïti - 2020, Land Cover Map", doi:10.18167/DVN1/QOLTTE, CIRAD Dataverse,V1.*

# 3.5. VALIDATION DES CARTES

L'évaluation quantitative des produits obtenus est réalisée avec une méthode de validation croisée mobilisant à nouveau l'algorithme Randon Forest (RF). La base de données de référence est séparée en cinq sous-ensembles relativement égaux en nombre d'entités. RF est alors exécuté cinq fois, en entrainant sur quatre sous-ensembles et en validant avec le sous-ensemble restant, de sorte que chaque sous-ensemble soit utilisé une fois pour la validation. Cette procédure a été utilisée pour chaque niveau de la nomenclature.

Les indicateurs obtenus sont issus de la moyenne des précisions obtenues à chacune des 5 validations croisées. Ils sont de 2 ordres :

- Au niveau global, 2 indices sont obtenus :
  - La précision globale qui correspond à la surface totale des objets correctement classés divisée par la surface totale des objets testés.
  - L'indice de Cohen Kappa qui reflète la différence entre l'accord réel et l'accord attendu par un traitement réalisé au hasard.
- Au niveau d'une classe spécifique, la précision a été résumée en utilisant le F1-score, qui est calculé en faisant la moyenne harmonique des précisions du producteur et de l'utilisateur. La marge d'erreur correspond à l'écart type de ces précisions sur l'ensemble des 5 validations croisées.

Les précisions ainsi calculées sont données dans le Tableau 6 pour les cartes produites avec la mosaïque Pléiades et dans le Tableau 7 pour celles produites avec la mosaïque Spot6/7. La classe « route » n'est pas présente dans ces tableaux puisque les routes sont issues de la base de données OSM. Plus précisément, cette validation est faite sur les cartes à la sortie de la chaîne Moringa, donc avant tout post-traitement (lissage, etc.).

Globalement, **le niveau de précision de ces différentes classifications est très bon, voire excellent** (plus de 90%) pour le niveau « 0 » de nomenclature. Comme attendu théoriquement, cette précision diminue progressivement à mesure qu'on affine la nomenclature, mais reste proche de 80% pour le niveau le plus fin (niveau « 2 »). Aussi, les classes inchangées d'un niveau de nomenclature à l'autre restent stables, et gagnent même en précision : seules les classes qui sont affinées et séparées en plusieurs sous-classes, à savoir les cultures non ligneuses et les différents types de végétation ligneuse, perdent en précision. On constate aussi que la part de hasard dans la détermination des classes augmente aussi dans le niveau « 2 », le plus fin de nomenclature (kappa inférieur à 0,78/0,75). C'est là où le compromis est à discuter par chaque utilisateur selon l'utilisation qu'il ferait de ce type de cartes : préserver la fiabilité de l'information sur une nomenclature moins précise, ou augmenter la précision des classes au risque de perdre en validité.

Pour les classes les plus fines, il est important de souligner que la végétation arborée reste une problématique compliquée en télédétection et que la présente étude ne résout pas les conflits de discrimination habituellement rencontrés avec ce type de données et dans ce type de paysage. La distinction entre espèces ligneuses cultivées et naturelles, et même la distinction entre espèces arborées et arbustives, passeraient par l'utilisation de données à THRS ayant une bien plus grande richesse spectrale (de type WV3), qui ont des coûts assez prohibitifs à l'échelle régionale (plus de 60 euros du km2) et pour les projets de développement agricole en général. L'arboriculture en milieu tropical et agriculture paysanne est difficilement distinguable de l'agroforesterie sans des analyses de texture poussées à des résolutions parfois plus fines que celle de Pléiades (capacités du drone ou de l'aéroporté). Cependant, les systèmes agroforestiers sont ici assez bien classés, avec une précision de 78,9% avec du Pléiades et 74,7% avec du SPOT6/7, ce qui est assez encourageant. Une meilleure construction de la base de référence, avec des relevés plus nombreux, plus variés et plus diversifiés sur le terrain dans ces classes très hétérogènes et complexes, permettraient aussi d'améliorer ces résultats.

*Tableau 6 : **Précisions** des cartes d'occupation du sol produites avec la chaine **Moringa** en utilisant les données à THRS **Pléiades**. Les cases colorées représentent trois niveaux de précision par classe, évaluée au travers du F1 Score : fort en vert (supérieur à 80%), acceptable en jaune (compris entre 70 et 80%), et limité en rouge (inférieur à 70%).*

| **Niveau 0** *Classe* | *Score* | *Marge d'erreur* |
|---|---|---|
| Bâti | **86,8%** | ±3,7% |
| Culture non ligneuse | **81,2%** | ±5,9% |
| Végétation ligneuse (cultivée et naturelle) | **94,9%** | ±0,8% |
| Pâturage, prairie et savane herbacée | **78,4%** | ±9,1% |
| Zone humide côtière (mangrove) | **94,9%** | ±4,3% |
| Rocher, sol nu et plage | **83,2%** | ±6,4% |
| Eau | **99,9%** | ±0,1% |
| **Précision globale** | **91,8%** | ±1,6% |
| **Indice de Kappa** | **0,89** | ±0,02% |

| **Niveau 1** *Classe* | *Score* | *Marge d'erreur* |
|---|---|---|
| Bâti | **86,7%** | ±3,2% |
| Culture non ligneuse | **81,4%** | ±3,2% |
| Agroforesterie et arboriculture | **79,2%** | ±3,7% |
| Végétation naturelle ligneuse | **91,4%** | ±2,9% |
| Pâturage, prairie et savane herbacée | **78,4%** | ±6,2% |
| Zone humide côtière (mangrove) | **96,3%** | ±5,2% |
| Rocher, sol nu et plage | **84,4%** | ±6,1% |
| Eau | **99,9%** | ±0,1% |
| | **88,5%** | ±1,7% |
| | **0,86** | ±0,02% |

| **Niveau 2** *Classe* | *Score* | *Marge d'erreur* |
|---|---|---|
| Bâti | **86,5%** | ±3,6% |
| Culture sans couvert arboré | **76,9%** | ±2,8% |
| Culture avec présence d'arbres | **40,1%** | ±11,3% |
| Systèmes agro-forestiers | **78,9%** | ±3,8% |
| Arboriculture | **64,1%** | ±13,2% |
| Végétation arborée | **71,1%** | ±6,4% |
| Végétation arbustive | **72,7%** | ±4,5% |
| Pâturage, prairie et savane herbacée | **78,9%** | ±7,4% |
| Zone humide côtière (mangrove) | **97,1%** | ±3,6% |
| Rocher, sol nu et plage | **84,5%** | ±6,6% |
| Eau continentale | **97,3%** | ±3,8% |
| Eau marine | **100,0%** | ±0,0% |
| | **80,4%** | ±2,5% |
| | **0,78** | ±0,03% |

*Tableau 7 : **Précisions** des cartes d'occupation du sol produites avec la chaine **Moringa** en utilisant les données à THRS **Spot 6**/7. Les cases colorées représentent trois niveaux de précision : fort en vert (supérieur à 80%), acceptable en jaune (compris entre 70 et 80%), et limité en rouge (inférieur à 70%).*

| **Niveau 0** *Classe* | *Score* | *Marge d'erreur* |
|---|---|---|
| Bâti | 83,9% | ±4,6% |
| Culture non ligneuse | 77,6% | ±3,2% |
| Végétation ligneuse (cultivée et naturelle) | 93,8% | ±2,1% |
| Pâturage, prairie et savane herbacée | 76,4% | ±3,3% |
| Zone humide côtière (mangrove) | 96,9% | ±4,4% |
| Rocher, sol nu et plage | 84,2% | ±0,8% |
| Eau | 99,6% | ±0,4% |
| **Précision globale** | **90,3%** | ±1,4% |
| **Indice de Kappa** | **0,86** | ±0,02% |

| **Niveau 1** *Classe* | *Score* | *Marge d'erreur* |
|---|---|---|
| Bâti | 80,8% | ±12,6% |
| Culture non ligneuse | 78,5% | ±4,1% |
| Agroforesterie et arboriculture | 75,3% | ±5,6% |
| Végétation naturelle ligneuse | 90,1% | ±3,1% |
| Pâturage, prairie et savane herbacée | 77,1% | ±3,4% |
| Zone humide côtière (mangrove) | 98,0% | ±1,6% |
| Rocher, sol nu et plage | 82,7% | ±0,9% |
| Eau | 99,7% | ±0,2% |
| | **86,9%** | ±2,1% |
| | **0,84** | ±0,02% |

| **Niveau 2** *Classe* | *Score* | *Marge d'erreur* |
|---|---|---|
| Bâti | 80,6% | ±13,1% |
| Culture sans couvert arboré | 73,5% | ±2,1% |
| Culture avec présence d'arbres | 37,1% | ±1,1% |
| Systèmes agro-forestiers | 74,7 | ±6,3% |
| Arboriculture | 62,0% | ±13,6% |
| Végétation arborée | 68,7% | ±3,7% |
| Végétation arbustive | 70,6% | ±1,19% |
| Pâturage, prairie et savane herbacée | 76,8% | ±3,6% |
| Zone humide côtière (mangrove) | 98,2% | ±1,5% |
| Rocher, sol nu et plage | 82,4% | ±9,0% |
| Eau continentale | 92,2% | ±8.8% |
| Eau marine | 99,9% | ±0,1% |
| | **78,2%** | ±1,18% |
| | **0,75** | ±0,01% |

**Cartes d'occupation du sol 2020 "Code1" et "Code2" obtenues en utilisant l'image Pléiades**

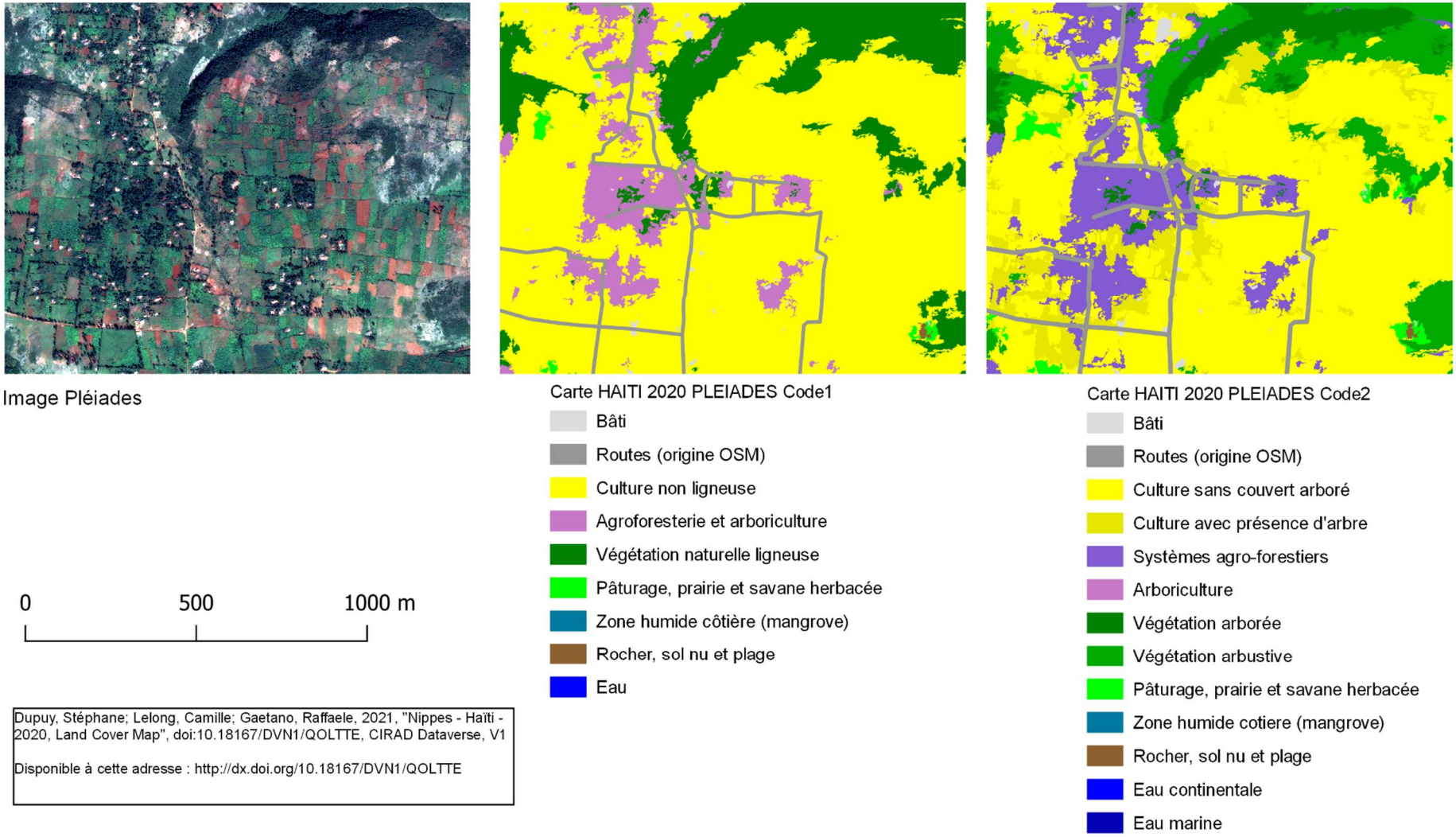


*Figure 34 : Extrait de la carte d'occupation du sol en 2020, obtenue en utilisant l'image Pléiades aux niveaux de nomenclature « 1 » et « 2 ». Zoom sur une zone située au Sud-Est du bassin versant de Rivière Froide, région des Nippes. Référence : Dupuy, S.; Lelong, C.; Gaetano, R., 2021, "Nippes - Haïti - 2020, Land Cover Map", doi:10.18167/DVN1/QOLTTE, CIRAD Dataverse,V1. Image © CNES 2020 - distribution AIRBUS DS.*

**Cartes d'occupation du sol 2020 "Code1" et "Code2" obtenues en utilisant l'image Pléiades**

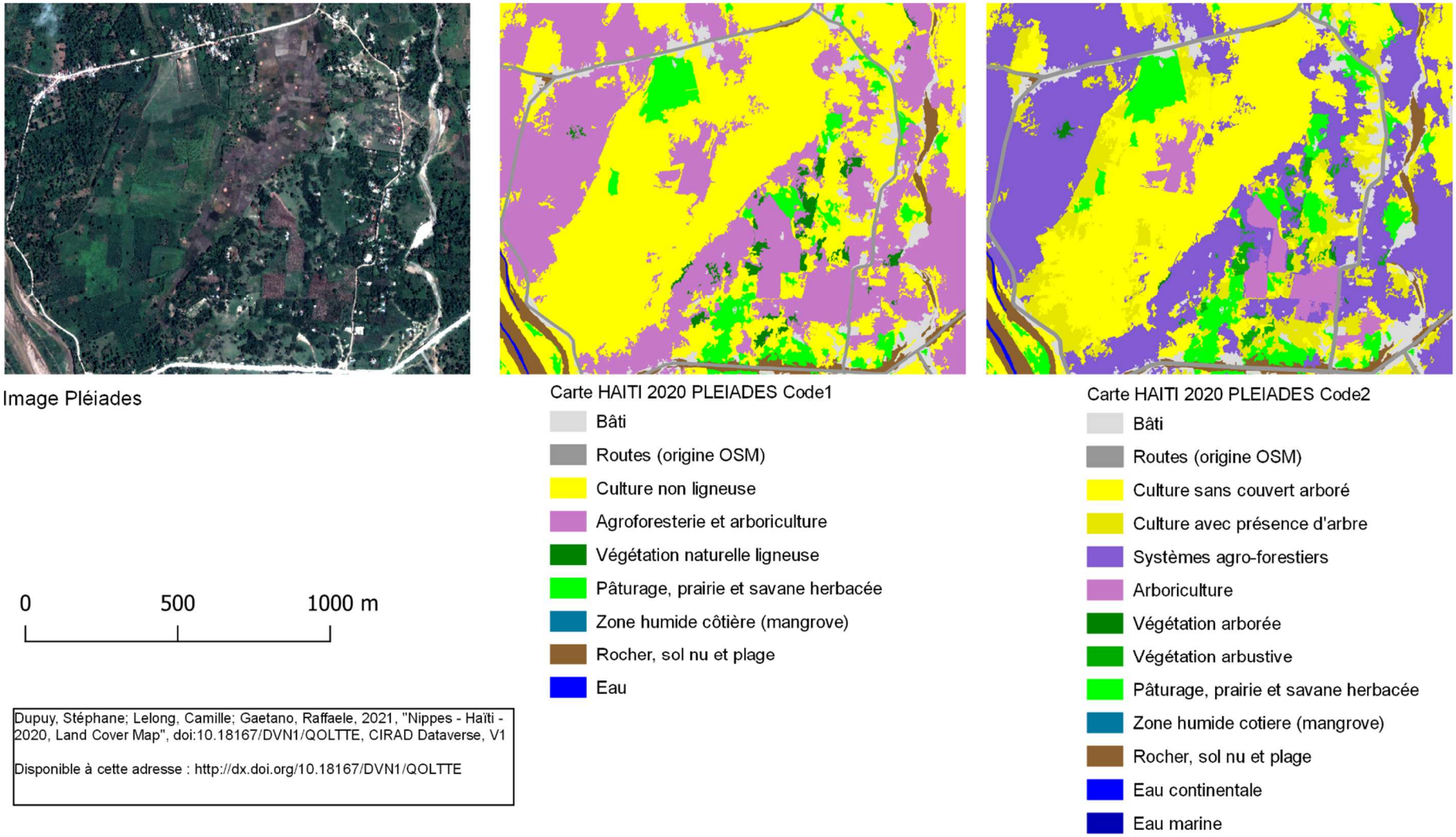


*Figure 35 : Extrait de la carte d'occupation du sol en 2020, obtenue en utilisant l'image Pléiades aux niveaux de nomenclature « 1 » et « 2 ». Zoom sur une zone située au Nord-Ouest du bassin versant de Rivière Froide, région des Nippes. Référence : Dupuy, S.; Lelong, C.; Gaetano, R., 2021, "Nippes - Haïti - 2020, Land Cover Map", doi:10.18167/DVN1/QOLTTE, CIRAD Dataverse,V1. Image © CNES 2020 - distribution AIRBUS DS*

Nous souhaitons aussi attirer l'attention sur la classe définie par le CNIGS comme « culture sous couvert arboré », qui est très mal classée (seulement 37,1% de précision avec SPOT6/7 et 40,1% avec Pléiades) et dont la distinction engendre une baisse dans la performance des classifications à identifier les cultures non ligneuses en général. Si on peut comprendre l'intérêt de cartographier ces systèmes de culture, il apparait cependant dangereux de le faire sans une vraie définition de ladite classe (à partir de combien d'arbres considère-t-on leur présence dans la parcelle ? à partir de quelle densité d'arbres la considère-t-on comme agroforestière ?), mais aussi sans voir analysé la taille de l'arbre au regard de la capacité de la donnée à le détecter et à l'inclure quand même dans l'objet « parcelle » et non comme objet à part entière. Là aussi il y a une vraie question de recherches en télédétection.

On constate par ailleurs que les classifications obtenues avec Pléiades d'une part et SPOT6/7 d'autre part sont assez comparables, avec des précisions globales comme par classe légèrement meilleures pour Pléiades. Cette équivalence en termes de résultats doit, elle aussi, amener l'utilisateur à mettre en balance le budget alloué pour l'achat des images à THRS, la disponibilité de ces données, l'emprise de la zone à cartographier, la complexité de mise en œuvre des prétraitements et des extractions de variables, le niveau de nomenclature exigé et le gain en fiabilité attendu de ce type de carte.

Les Figure 34 et Figure 35 présentent, pour exemple, deux extraits des cartes obtenues en utilisant l'image Pléiades comme image à THRS. On peut y observer une bonne précision dans la délimitation des parcelles de différentes occupations du sol, y compris dans des détails assez fins. La validité spatiale de ces produits de classification n'est pas chiffrable et seule l'analyse comparée des cartes et des images peut donner une information sur cette précision. Globalement, cette cartographie apparait ici de très bonne qualité. Le niveau « 2 » de nomenclature permet même d'appréhender les différents types d'usage ou de systèmes de culture. On remarque qu'il y a aussi très peu de modification dans l'affectation des classes générales en passant d'un niveau à l'autre, ce qui est un gage de bonne stabilité des algorithmes. Le gain en précision dans la qualification des classes est notable, et les confusions rencontrées sont plutôt de l'ordre de la distinction entre classes affinées au niveau « 2 » de nomenclature, très peu au niveau « 1 ». L'arboriculture et l'agroforesterie sont plutôt bien distinguées, comme la plupart des arbres inclus dans les parcelles de culture non ligneuse. Il reste cependant une assez forte confusion entre les milieux arborés et arbustifs, entre milieux arborés et agroforestiers, ou encore entre la végétation naturelle herbacée et les cultures non ligneuses. Aussi, le bâti se confond parfois avec le sol nu. Mais la délimitation des objets reste toujours plus précise au niveau « 2 ». Ces observations d'ordre qualitatif confirment la tendance observée dans les statistiques quantitatives, à savoir que les niveaux de nomenclature sont atteints avec une forte validité mais proposent des renseignements différents et complémentaires, avec des compromis à analyser suivant l'usage final de la carte.

# 4. Conclusion

Dans cette étude nous avons montré comment combiner une série temporelle d'images à Haute Résolution Spatiale (HRS) acquise par Sentinel-2 à une image à Très Haute Résolution Spatiale (THRS) pour produire des cartes d'occupation du sol sur la zone d'étude du projet TPR dans le département des Nippes en Haïti.

Nous avons produit ces cartes à trois niveaux de précision de la nomenclature, en utilisant pour chaque niveau 2 types d'images à THRS : Spot6/7 d'une part et Pléiades d'autre part. Nous avons donc produit au total 6 cartes indépendantes, qui sont toutes de très bonne qualité, tant du point de vue quantitatif que qualitatif.

Les précisions globales des cartes issues des traitements utilisant Pléiades sont plus élevées, avec des valeurs de 91% (niveau « 0 »), 85% (niveau « 1 ») et 80% (niveau « 2 »), que celles obtenues avec Spot 6/7 : 90% (niveau « 0 »), 86% (niveau « 1 ») et 78% (niveau « 2 »).

Les précisions par classe sont aussi très fortes, généralement supérieures à 80% et pouvant atteindre 100%. Les classes les plus détailles ont tendance à voir leur précision diminuer quand on affine le niveau de nomenclature, alors que les classes plus générales voient leur précision augmenter. Il est donc aussi possible de se baser sur la production de la carte au niveau le plus précis (niveau « 2 ») et fusionner en post-production certaines des classes les plus détaillées si nécessaire, afin d'obtenir le bon compromis entre la capacité à discriminer les différents usages du sol et la validité de l'information extraite. Nous avons aussi constaté qu'en fonction du niveau de précision de la nomenclature les images à THRS ont une importance plus ou moins grande dans le processus d'extraction des classes, mais apportent dans tous les cas l'information spatiale détaillée essentielle à la bonne classification des objets.

Ainsi, l'apport de la THRS est tout à fait évident pour obtenir une cartographie précise et surtout atteindre des niveaux de nomenclature détaillés, en particulier en rapport avec des types d'usage du sol (distinction agroforesterie/arboriculture par exemple). Mais, en fonction des besoins en termes de validité et de fiabilité, il faudra accepter des compromis sur cette précision de la nomenclature, et choisir les données les plus adaptées à la problématique nécessitant l'utilisation de ces cartes.

# 5. Remerciements

Nous tenons à remercier le CNES pour la mise à disposition des images Pléiades et Spot6/7 acquises dans le cadre du projet Recovery Observatory (RO) Haïti.

Nous remercions également le CNES pour les données Sentinel-2 à valeur ajoutée traitées pour le pôle de données Theia (www.theia.land.fr) à partir de données Copernicus. Les traitements utilisent les algorithmes développés par les Centres d'Expertise Scientifique de Theia.

# 6. References

**Baatz, M., Schäpe, A**. (2000) : "Multiresolution Segmentation: an optimization approach for high quality multi-scale image segmentation." Journal of Photogrammetry and Remote Sensing 58(3-4): 12–23.

**Dupuy, S.; Lelong, C.; Manasse, M.-E.; Rambao, J.; Mondésir, J. P.; Gaetano, R**. (2021) : "Nippes - Haïti - 2020, Reference Spatial Database", doi:10.18167/DVN1/29RZMQ, CIRAD Dataverse, V1.

**Dupuy, S., Lelong, C., Gaetano, R.** (2021) : "Nippes - Haïti - 2020, Land Cover Map", doi:10.18167/DVN1/QOLTTE, CIRAD Dataverse,V1.

**Gaetano R., Dupuy S., Lebourgeois V., Le Maire G., Tran A., Jolivot A., Bégué A**. (2019) : The MORINGA processing chain: Automatic object-based land cover classification of tropical agrosystems using multi-sensor satellite imagery. Milan : Italian Space Agency, 3 p. 2019 Living Planet Symposium, 2019-05-13/2019-05-17, Milan (Italie).

**Lassalle, P.; Inglada, J.; Michel, J.; Grizonnet, M.; Malik, J.** (2014): Large Scale Region-merging Segmentation Using the Local Mutual Best Fitting Concept. In Proceedings of the 2014 IEEE Geoscience and Remote Sensing Symposium, Quebec City, QC, Canada, 13–18 July 2014; pp. 4887–4890.

**Lelong, C. et Dupuy S.** (2021): Guide technique - Constitution d'une base de données de référence dédiée à l'apprentissage et la validation des classifications d'images satellites : la vérité-terrain, Montpellier (France) : Cirad, 16 pages.

Département Environnement et Société (ES)
Unité Mixte de Recherche Territoire, Environnement, Télédétection et Information Spatiale (TETIS), 2021
Maison de la Télédétection - 500 rue Jean-François Breton - TA C-91/MTD 34093 Montpellier cedex 5
https://umr-tetis.fr

www.cirad.fr