Regional scale mapping and characterization of the poplar resource using two remote sensing data mining approaches

Auteur : Lacaille, Baptiste
Promoteur(s) : Lejeune, Philippe; Latte, Nicolas
Faculté : Gembloux Agro-Bio Tech (GxABT)
Diplôme : Master en bioingénieur : gestion des forêts et des espaces naturels, à finalité spécialisée
Année académique : 2021-2022
URI/URL : http://hdl.handle.net/2268.2/16312

Errata for MSc Thesis

Lacaille Baptiste

August 29, 2022

This document lists errors found in the master's thesis submitted by Baptiste Lacaille on August 16, 2022, entitled "Regional scale mapping and characterization of the poplar resource using two remote sensing data mining approaches", along with corrections as applicable.

Corrections to the text

Location	Original text	Correction
Page 2; § 1	The wood of low density, soft, creamy white, with a uniform texture and diffuse pores, is versatile and traditionally used for peeling (light packaging or plywood), sawing (pallets and packaging boxes), pulpwood or energy wood (**Peuplier** 2022).	The wood of low density, soft, creamy white, with a uniform texture and diffuse pores, is versatile and traditionally used for peeling (light packaging or plywood), sawing (pallets and packaging boxes), pulpwood or energy wood (**Hout Info Bois** 2022).
Page 5; § 5	There are several algorithms commonly used for classification of remote sensing data, such as Support Vector Machines (SVMs), Naïve Bayes classifier, k-nearest neighbor classifier, **remote sensing**, as cited in (Kishore et al., 2016; C. Zhang et al., 2017)	There are several algorithms commonly used for classification of remote sensing data, such as Support Vector Machines (SVMs), Naïve Bayes classifier, k-nearest neighbor classifier, **Random Forests (RFs),** as cited in (Kishore et al., 2016; C. Zhang et al., 2017)
Page 9; § 1	The super-resolution images return the 10 bands from the 10 m and **10** m resolution S2 images, but with an enhanced pixel size at 2.5 m.	The super-resolution images return the 10 bands from the 10 m and **20** m resolution S2 images, but with an enhanced pixel size at 2.5 m.
Page 10; § 2	Two classes were considered in the DNF GIS data to form the reference polygons, "**poplars**" and "**others**";	Two classes were considered in the DNF GIS data to form the reference polygons, "**poplar**" and "**other**";
Page 10; § 3	[…] and **15,638** ha for the "other" label from the DNF GIS data.	[…] and **15,628** ha for the "other" label from the DNF GIS data.
Page 15; § 1	The three main hyperparameters to set before training the RFs are: the maximum number of trees in the forest, the minimum node size for a split and the number of features tested at each node, which have been set with the following values, respectively **100**, **25** and **12**. These values were tuned using the R packages "randomForest" and "caret", **training was then performed using the Shark Random Forests classification algorithm within OTB. The latter includes another hyperparameter to set** the fraction of the training dataset for the calculation of the out of the bag (OOB) **value** which is a type of RF cross-validation method. **The default value of 0.66 was kept, meaning that** 0.66% of each bootstrap at the origin of the decision trees are	The three main hyperparameters to set before training the RFs are: the maximum number of trees in the forest, the minimum node size for a split and the number of features tested at each node, which have been set with the following values, respectively **300**, **30** and **10.** These values were tuned using the R packages "randomForest" and "caret" **and then refined according to the other hyperparameters of the Random Forest classifier (OpenCV) in OTB used to train the model. The last ones are the maximum depth of the tree being set at 20 and the sufficient precision (OOB error) of 0.01.** The fraction of the training dataset for the calculation of the out of the bag (OOB) **error** which is a type of RF cross-validation method **was imposed with a value of 1/3. In other words, about 33.33%** of each bootstrap at the origin of the decision trees

	used for performance evaluation and selection of variables of interest.	are used for performance evaluation and selection of variables of interest.
Page 21; Table 9 (Title)	Evaluation metrics with the defining equation used in this project, where κ is the confidence score threshold, […]	Evaluation metrics with the defining equation used in this project, where τ is the confidence score threshold, […]
Page 25; § 1	The distribution of labels within the **24** WFI sample plots **of 24,429 m²** is detailed in Table 11.	The distribution of labels within the **22** WFI sample plots **for 3423 pixels** is detailed in Table 11.
Page 25; § 3	The estimated area for the "poplar" label from the image interpretation points was 10,236 ha, with a distribution of **83.52**% for the "other" label and **16.48**% for the "poplar" label.	The estimated area for the "poplar" label from the image interpretation points was 10,236 ha, with a distribution of **84.14**% for the "other" label and **15.86**% for the "poplar" label.
Page 30; § 4	These are particularly noticed at the Sapling stage, as shown at point C and on the left side of the canal at point **E** in Figure 11.	These are particularly noticed at the Sapling stage, as shown at point C and on the left side of the canal at point **F** in Figure 11.
Page 31; § 2	[…] are sometimes partially detected (see point **B** in Figure 12), they are rarely fully detected.	[…] are sometimes partially detected (see point **D** in Figure 12), they are rarely fully detected.
Page 34; § 3	[…] and at the edges of poplar plantations (see point **D** in Figure 14).	[…] and at the edges of poplar plantations (see point **B** in Figure 14).
Page 36; § 2	but some imprecision of the pCHM is still observed on alignments as shown in point **C** in Figure 16.	but some imprecision of the pCHM is still observed on alignments as shown in point **D** in Figure 16.
Page 42; § 6	ESA (2022a). Newcomers Earth Observation Guide | ESA Business Applications.	ESA (2022a). Newcomers Earth Observation Guide | ESA Business Applications. **URL: https://business.esa.int/newcomers-earth-observation-guide (visited on 07/28/2022).**
Page 42; § 7	– (2022b). Sentinel Online.	– (2022b). Sentinel Online. **URL: https: //business.esa.int/newcomers-earth-observation-guide (visited on 07/28/2022).**
Page 42; § 8	– (2022c). Sentinel-2 MSI Technical Guide.	– (2022c). Sentinel-2 MSI Technical Guide. **URL: https://sentinel.esa.int/web/sentinel/ (visited on 07/11/2022).**
Page 45; § 1	Peuplier (2022). fr-FR. Publication Title: Hout Info Bois.	**Hout Info Bois** (2022). Publication Title: **Peuplier**. **URL: https://www.houtinfobois.be /essences/peuplier/ (visited on 07/05/2022).**
Page 45; § 2	Planet Labs PBC (Sept. 2021). Satellite Imagery and Archive. Publication Title: Planet.	Planet Labs PBC (Sept. 2021). Satellite Imagery and Archive. Publication Title: Planet. **URL: https://planet.com/products/ planet-imagery/ (visited on 06/27/2022).**
Page 45; § 4	PSW (2022a). Nuage de points LIDAR 2013-2014. en.	PSW (2022a). Nuage de points LIDAR 2013-2014. **URL: http://geoportail.wallonie.be/ catalogue/cd7578ef-c726-46cb-a29e-a90b3d4cd368.html (visited on 07/11/2022).**
Page 45; § 5	– (2022b). Orthophotos : la Wallonie vue du ciel. en.	– (2022b). Orthophotos : la Wallonie vue du ciel. **URL: http://geoportail.wallonie.be/ georeferentiel/orthophotos (visited on 07/11/2022).**

Correction of Figure 2 (Page 12)

Original version

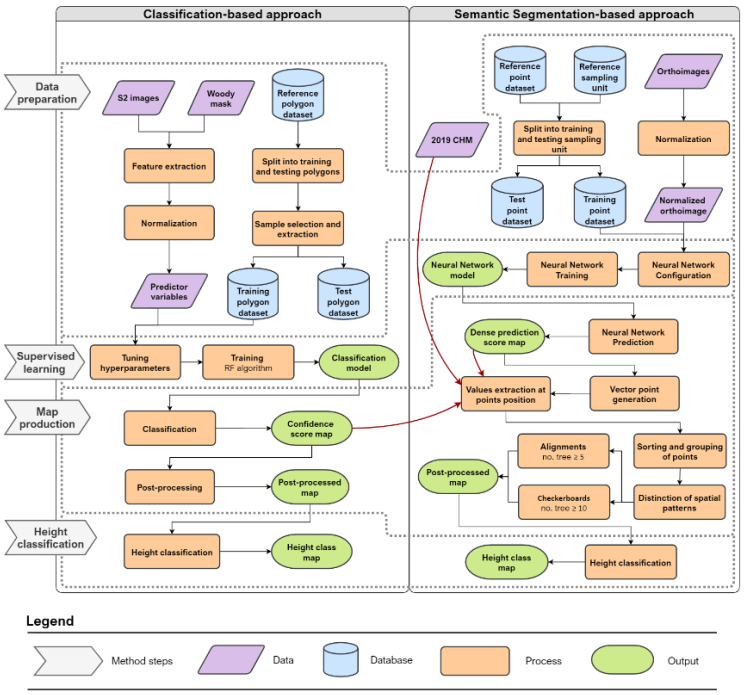

Figure 2: Process flowchart of the methods used for the mapping and characterization of the poplar resource in the province of Hainaut using remote sensing data, from data preparation to height classification.

Corrected version (2018 pCHM added)

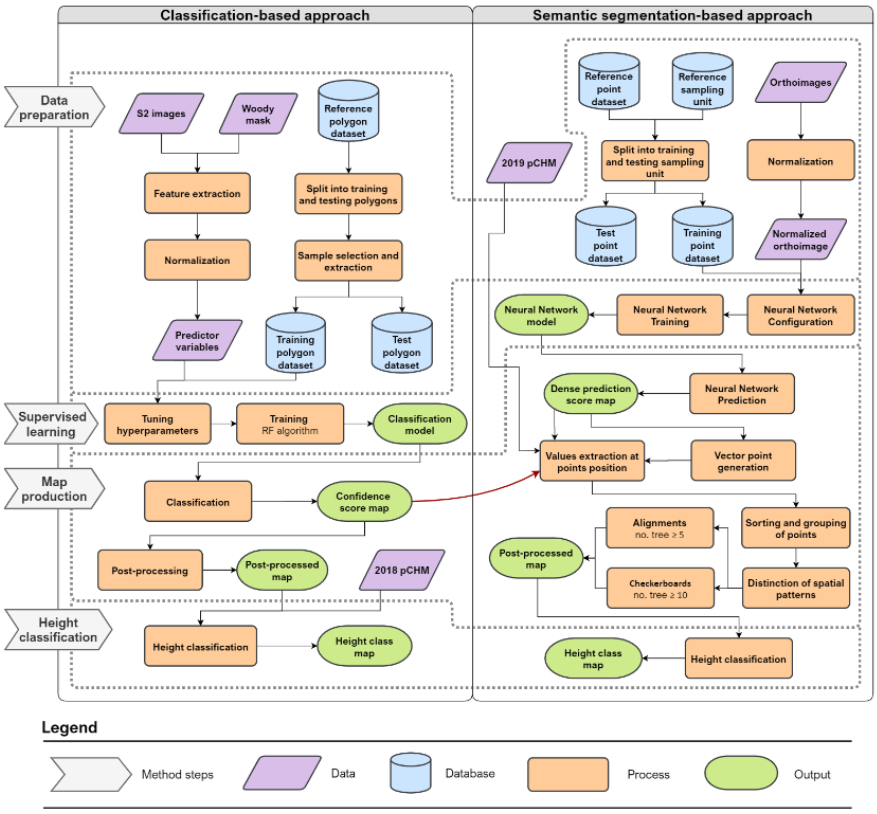

Figure 2: Process flowchart of the methods used for the mapping and characterization of the poplar resource in the province of Hainaut using remote sensing data, from data preparation to height classification.

Correction of Table 5 (Page 18)

Original version

Table 5: Confidence score (c) and dense prediction score (d) used to retain a point according to the average confidence score value per tile (a).

For an average confidence score	Values used to retain a point
$a < 0.1$	$(c \geq 0.9 \,\&\, d \geq 0.6)$
$0.1 \leq a < 0.4$	$(c \geq 0.9 \,\&\, d \geq 0.5)$
$0.4 \leq a < 0.65$	$(c \geq 0.8 \,\&\, d \geq 0.4)$
$0.65 \leq a < 0.85$	$(c \geq 0.7 \,\&\, d \geq 0.35)$
$a \geq 0.85$	$(c \geq 0.9 \,\&\, d \geq 0.4)$

Corrected version

Table 5: Confidence score (c) and dense prediction score (d) used to retain a point according to the average confidence score value per tile (a).

For an average confidence score	Values used to retain a point
$a < 0.1$	$(c \geq 0.9 \mid d \geq 0.6)$
$0.1 \leq a < 0.4$	$(c \geq 0.9 \mid d \geq 0.5)$
$0.4 \leq a < 0.65$	$(c \geq 0.8 \mid d \geq 0.4)$
$0.65 \leq a < 0.85$	$(c \geq 0.7 \mid d \geq 0.35)$
$a \geq 0.85$	$(c \geq 0.9 \mid d \geq 0.4)$