

# Monitoring cocoa plantations around the Taï National Park in Ivory Coast

Image processing for earth observation - ENV540

Gallardo Joaquin & Spahr Antoine

January 9, 2020

## 1 Introduction

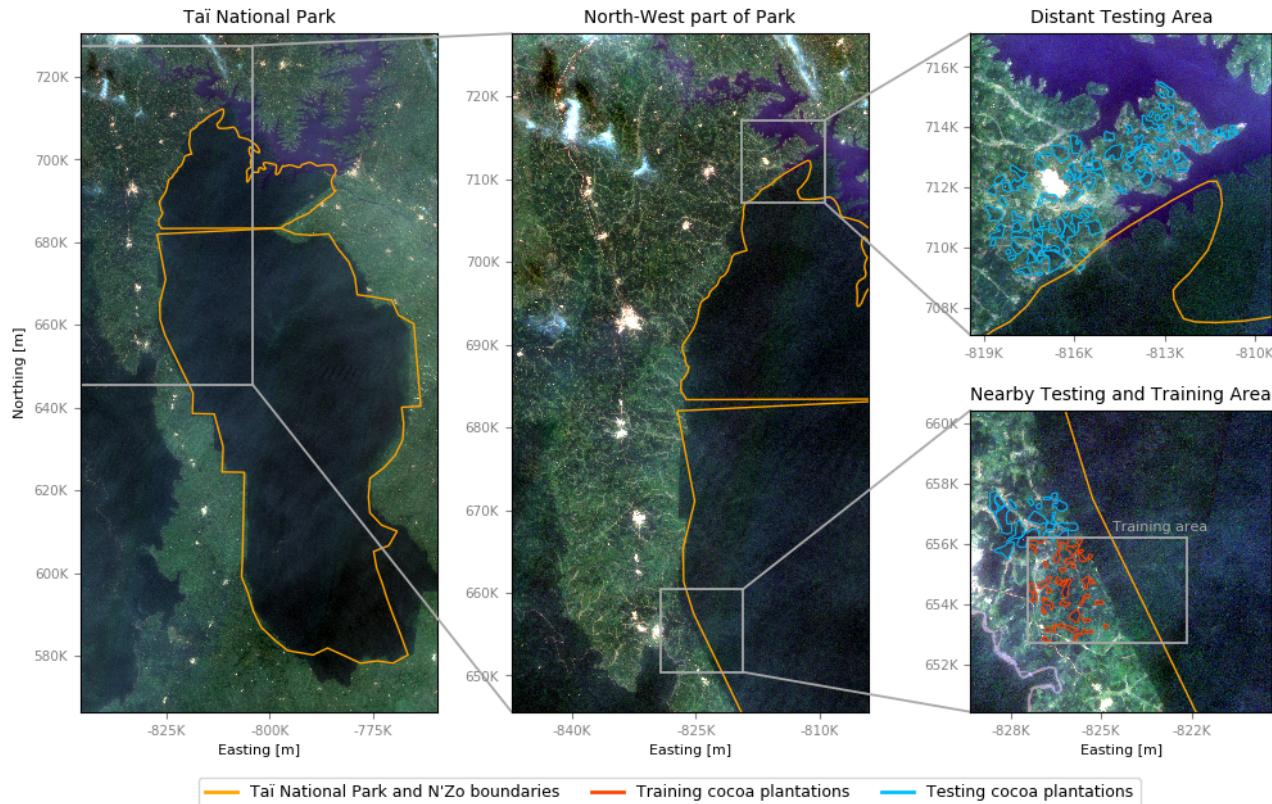
Preserving forest and national parks worldwide is of utterly importance and even more so in the current climate change crisis. Forests are reserves of biodiversity and life, and perform vital roles for the planet such as storing carbon dioxide and aiding in the water and nutrient cycles [1, 2]. In Ivory Coast the forest degradation in the last decades has reached critical levels and the main direct cause is cocoa plantations as explored by Barima *et al.* in their study of the Haut-Sassandra classified forest [3]. Bitty *et al.* showed in their study that the presence of illegal cocoa farming in protected areas is negatively correlated with the presence of primates [4], highlighting the negative impact of plantations on the biodiversity. The evidence of illegal cocoa farming in Ivory Coast [3] brings to light the need of a way to monitor the national parks and nearby cocoa plantations as it can help the authorities to detect violations or risk of future violations. Multiple studies have already explored this need. In their study, Barima *et al.* use a decision tree to classify Landsat images over several years to detect forest, crops and bare soil on the Haut-Sassandra protected forest [3]. Similar studies have also been conducted on the southern part of Ghana [5] and the Metchie-Ngoum Protection Forest Reserve in the Western part of Cameroun [6]. The monitoring of sustainable cacao plantations in Ivory coast has been explored by Hasan *et al.* using a support vector machine (SVM) classifier on high resolution Kompas-3 satellite imagery [7]. Noon *et al.* used a SVM classifier to efficiently map palm oil plantation in Ghana using Landsat images [8]. Tutu Benefoh *et al.* studied cocoa plantations in Ghana with Landsat images using supervised and unsupervised methods [9]. Singh *et al.* interestingly integrate LiDAR measurement with Landsat images to classify different forest and land use in the Phnom Kulen National Park in Cambodia. They tried three different models (SVM, random forest and an artificial neural network) and obtained good classification results [10]. Based on previous studies, remote sensing combined with supervised classification appears to be a powerful tool that has the potential to locate cocoa plantations in a cheap and efficient way. Even though many studies have focused on the western part of Africa to map the land use, none were found that attempt to detect cocoa plantations around the Taï National Park or any other in Ivory coast.

That is why the goal of this study is to detect cocoa plantations with the aid of remote sensing imagery in the surroundings of the Taï National Park. The Taï is the biggest reserve of natural forest left in Ivory coast and preserving it is crucial for the conservation of the hosted biodiversity, especially chimpanzees [11]. For this purpose, different supervised classification algorithms will be evaluated in order to provide a proof of concept for the monitoring of cocoa plantations around the park with free satellite imagery.

## 2 Methods

### 2.1 Data

In order to explore the possibility of monitoring cocoa plantations, Sentinel-2 images are chosen since they are freely accessible and offers the best spatial resolution within free images (10m for the red, green, blue and near infrared [NIR] bands). The spectral resolution is rather high as well with 13 different spectral bands. Finally, Sentinel-2 provides a good coverage of the globe, which



**Figure 1: Data overview.** The Taï national park is situated in the South-West part of Ivory Coast close to the Liberia border. The park border is displayed in orange. North of it is the N'Zo natural reserve which is not part of the Taï national park. The boundaries shapefiles were obtained from protected planet [13, 14]. The part of the park used for training, and testing prediction (nearby and distant) are presented in the context of the whole park.

ensured that the region of interest would be accessible. It is challenging to work on forest imagery near the equator as the scene is often cloudy. That is why the focus has been given to the warm and dry season where it is easier to find a cloudless day throughout the whole region of interest. The selected images are taken on the 29.12.2018 where a clear day enables the observation of the whole park without clouds. A summary of the satellite images is provided on table 1.

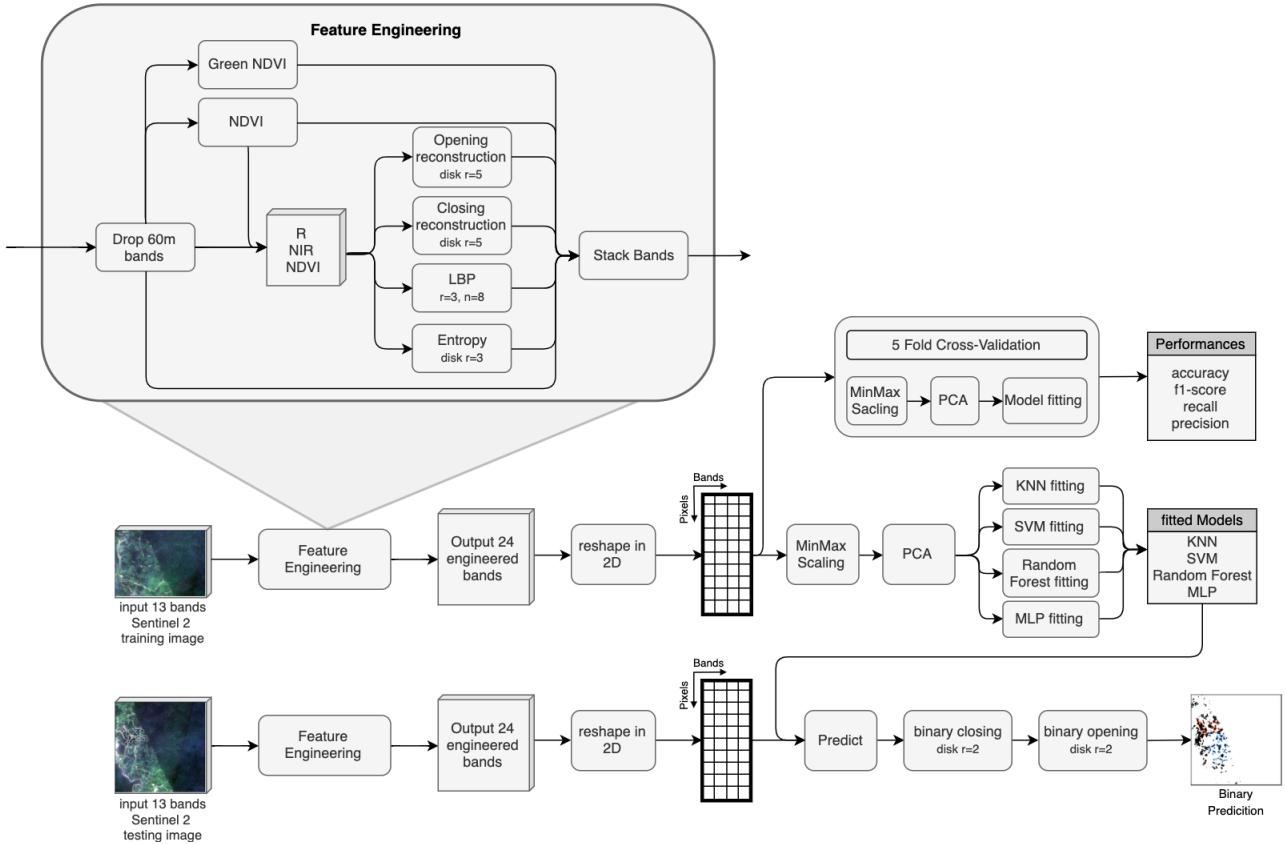
The image for analysis was chosen based on the visual localization of plantations using a map of cocoa cooperatives locations [12] and Google Earth as a reference, which has a very high image resolution due to the inclusion of airborne images. Consequently, a manual ground truth labelling of plantations was performed by creating a shapefile of polygons using QGIS and visually inspecting the located plantations in Google Earth. The output of this process is presented in figure 1.

## 2.2 Model development

In order to detect the cocoa plantation from multi-spectral Sentinel-2 images, different models are developed in a supervised fashion based on the manually labeled plantations. The models construction can be separated in three phases : 1) Feature engineering, 2) Model training and 3) Prediction on new areas. The processing is visually detailed in figure 2.

**Table 1: Data specifications**

Platform	Sensor	source	Acquisition			Resolution			CRS	Type of product
			Method	Date	Spatial	Spectral	Radiometric			
Sentinel-2	L1C	EO-browser	Sentinel-Hub	between 24.10.2019 and 02.12.2019	10m to 60m	13 bands	uint 16 bits	EPSG:4326	Processed Level 1C	



**Figure 2: Processing pipeline.** The flowchart presents the processing steps performed and described in the methods. The pipeline can be split in three main group : the pre-processing, the model development and performance estimations, and the predictions.

### 2.2.1 Feature engineering

The first part of the model development consists in generating the data for the training by creating new features from the various image bands. The input image contains 13 bands (4 at a 10 meters resolution, 6 at 20 meters resolution and 3 at a 60 meters resolution). In the first place, the 60m resolution bands are dropped as their resolution is too low for the pursued application. Then the Red (R) and Near Infrared (NIR) bands are used to compute the NDVI ( $= \frac{NIR-R}{NIR+R}$ ). The NDVI is a proxy for the vegetation content of a pixel as the vegetation reflects highly on NIR. Nuances in NDVI can be imputed to different types of vegetation. It may thus yield different signal between native forest and plantations. In a similar fashion the green NDVI is also computed by using the green band instead of the red and the red instead of the NIR.

The following processing steps are applied to the R, NIR and NDVI images because of their spatial resolution of 10m and their potential discriminatory power for vegetation elements. Each of the three bands are opened by reconstruction with a disk-shaped structuring element of radius 5. This procedure aims to remove small bright elements but still keeping the structure of the original image (contrarily to a simple opening). The three bands are then closed by reconstruction with a disk-shaped structuring element of radius 5. It aims at removing small dark element of the image (dark element are low value of pixel) but also by keeping the structure of the original image. In order to extract information of the texture around a pixel (i.e. add information of neighbors), the entropy is computed over a disk of radius 2 around the given pixel. A high value indicates a rather smooth neighborhood, while a low value reflects a rather random neighborhood. In order to extract other texture information, a local binary pattern (LBP) is applied over 8 neighbors at a distance of 3 pixels from the pixel of interest. This technique should yield different values depending on the neighbors values and reflect edges, corners, etc.

At the end of the process, 14 new features have been created which represent a total of 24 features that will be used for the training procedure.

### 2.2.2 Model training

The model is trained to solve the following binary classification task : classify whether a pixel close to the Taï National Park is part of a cocoa plantation or not. Therefore the models are doing pixel-wise classification. The samples are the pixels from a fully labelled area West of the park (see figure 1).

The image is reshaped in a 2D array (Pixel x Bands) for the training, each pixel is a sample and each band is a feature. There are thus 24 features and 185'500 samples (350x530). The model is composed of three main blocks : a normalizer that scale all features between 0 and 1 in order to give the same power to each of them. It is followed by a PCA in order to generate decorrelated features by linearly combining the 24 original features. And finally the estimator that makes the classification. Four different estimators are explored : K-nearest-neighbors (KNN), support-vector-machine (SVM), random-forest classifier (RF) and a multilayer perceptron (MLP). The hyperparameters of the estimators are chosen as follow. For KNN five neighbors are considered to make the prediction. The SVM L2 regularization strength has been set to 1000. The RF is composed of 150 trees of maximum depth 20. The MLP contains one hidden layer of 100 neurons, and ReLU as activation function. It is trained using an Adam optimizer, a learning rate of 0.001 and maximum 200 iterations.

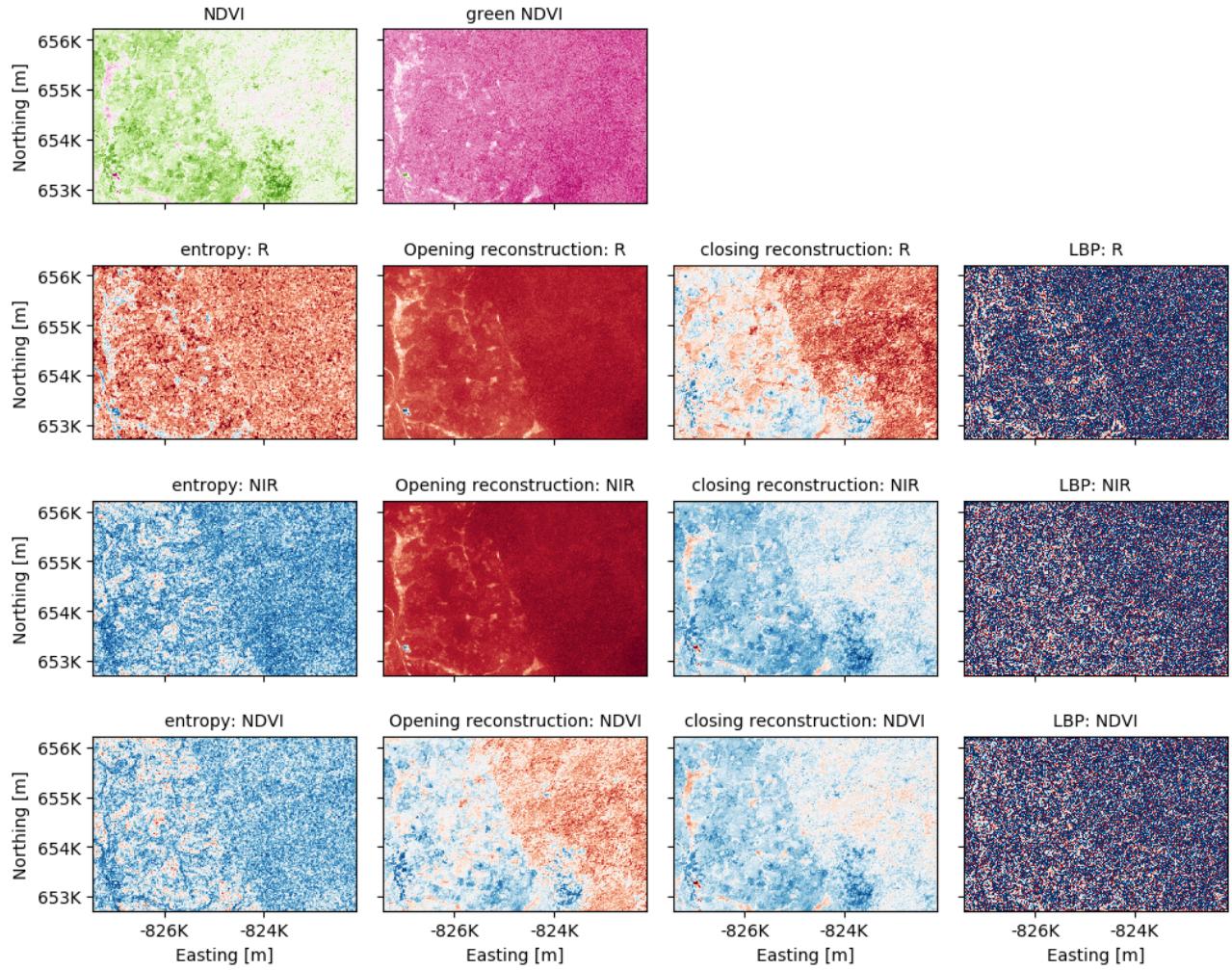
The first step is to obtain the expected performance of each models. It is achieved by a 5-fold cross-validation. Each model is trained on 80% of the data and tested on the 20% of unseen, remaining data. By alternating which fold is the test set, each models is trained on five different training sets and five measures of performance are obtained. This procedure enables to obtain an estimation of performance unbiased by the choice of the train and test set, and a measure of performance with confidence intervals. Note that the normalization and PCA coefficients are estimated on the training set (excluding the test set) and they are applied on the test set to avoid introducing a bias. Indeed, this procedure avoid the training of the model to observe in advance the data it will be tested on.

The performance of the models are assessed through four different metrics. The accuracy which quantifies the global correct classification. The precision which quantifies the goodness of detection (the rate of correct cacao plantation detection among the predicted cacao plantation). The recall which quantifies the rate of detection (the percentage of cacao plantation detected). And the F1-score which is the harmonic mean of the recall and the precision. A high F1-score means that the classifier detects most of the cacao plantations and predicts wrongly only a few pixel as cacao plantation.

Once the performance of the models have been defined, the four models are trained with the whole training data (red polygons in figure 1).

### 2.2.3 Model predictions generalization

Because the models are trained on a single area where the plantation positions are known, they might learn properties specific to that area (the sampling of training sample is limited to a single zone and is thus not uniform over the area to monitor). In consequence, the models might not generalize to further areas. In order to check whether the models generalize over space, the fitted models are used to predict the pixel class of two regions outside of the training area where there are known plantations. One is a *nearby prediction* and is situated just alongside the training area. The other one is a *distant prediction* and is situated north of the Taï national park (see figure 1). In both cases the fitted model generates a binary image of where it thinks plantations are. Black pixels



**Figure 3: Engineered Features.** The engineered features are presented over the training area. The top row presents the NDVI and the green NDVI. A green color indicates a high positive value while pink indicate a negative value. The second, third and fourth row present the feature engineering respectively on the red (R), the near-infrared (NIR) and NDVI bands. For each of those bands red indicates a positive value and blue a negative one. For each band (R, NIR and NDVI), the entropy, the opening and closing by reconstruction, and the LBP are computed. In total, 14 new features are generated.

represent a plantation and are labelled with a value of 1 and white pixels are anything else and have a value of 0. Those images are usually noisy predictions as there are single pixel of predicted cocoa plantation alone and there are holes within cluster of predicted cocoa plantations. Because the plantation are usually homogeneous piece of lands bigger than 10 meters (single pixel), the model's predictions are smoothed using a morphological closing with a disk-shaped structuring element of radius 2 (disk of diameter 50 meters) to fill holes in plantation patch, followed by a morphological opening with a similar structuring element to remove small isolated plantation prediction.

The quality of the detection is assessed through the recall only (*i.e.* detection rate). Indeed, only some known *nearby* and *distant* plantations on the images have been manually labelled. The whole extent of this images have not been inspected, so the recall is preferred to overcome the effect of unlabelled plantations. The recall is calculated as the percentage of the labelled polygons that have been detected by the model (*i.e.* among the labelled plantations, how much have been detected).

### 2.3 National park predictions

Now that the performances and predictive capabilities of the models have been explored, the best performing one can be used to predict unlabeled areas around the park. Predicting the presence of plantations over the whole park would be computationally too intensive as the park spread over around 150x50km which represents roughly 75 millions pixel to predict (10 meters resolution).

Therefore only 10x10km samples of the park are predicted to explore whether the cocoa plantations violate the national park boundaries.

## 3 Results

### 3.1 Model performances

#### 3.1.1 Feature Engineering

The features engineered on the training image are presented on figure 3. On the training image, most of the engineered features present different values between the native forest of the national park and the space used for cocoa plantations. The most discriminative ones seem to be the closing by reconstruction on the R band and the opening by reconstruction on the NDVI. The local entropy on the NDVI images seems to extract the patches that could correspond to the cocoa plantations.

#### 3.1.2 Model Performances

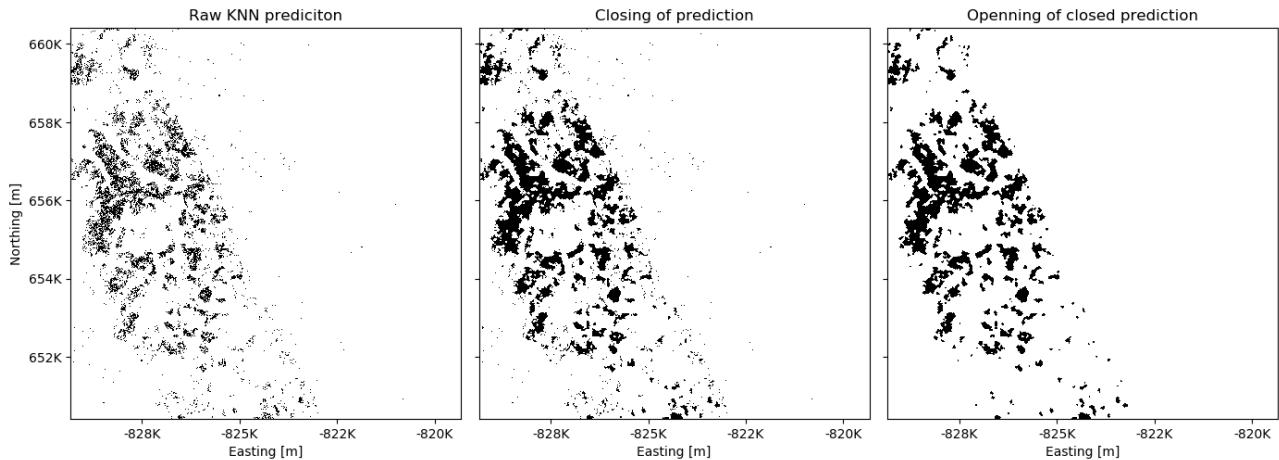
The expected performances of the four models considered are presented on table 2 for the four metrics considered, namely the accuracy, the f1-score, the precision and the recall. Both the random forest classifier and the MLP yield the highest mean accuracy with  $94.14\% \pm 1.38\%$  ( $\pm 2.62\%$ ) respectively. The highest recall have been obtained with the MLP ( $54.00\% \pm 10.22\%$ ). However note that the KNN's recall falls in the two standard deviation range of the MLP's one. They are thus not as different as they look like due to the high standard deviation. The model displaying the highest precision is the random forest classifier, but again with a high standard deviation. Finally the f1-score is higher for the MLP ( $57.64\% \pm 9.06\%$ ).

The random forest classifier presents some good performances on the test set, but those are really different from the train performances. This fact indicates the presence of overfitting : the model has become too specific to the training data and has thus trouble to generalize. On the other hand, the three others models show a lesser overfitting effect.

Globally the accuracy is much higher than the recall, precision and f1-score. This is likely due to the imbalance of the data-set. Indeed there are more non-plantation pixels than plantation pixels. In consequence, if the models has a poor detection rate (low detection of plantation) it would still reach a high accuracy. In addition, manually labelling the plantations is likely to have introduced wrong labeling : some plantation pixels may have been omitted and some non-plantation pixels may be labeled as plantation. In consequence, the models may predict correctly those few wrongly labeled pixels but this good behavior will not be reflected in the scores. Therefore, they must be analyzed with caution.

**Table 2: Models Performances** The performances of the four models estimated through 5-fold cross-validation are presented. For each model, the accuracy, f1-score, precision and recall are displayed for both the train and test folds. The performances are presented as the mean and standard deviation over the five permutations of the cross-validation. The best test performance for each metric is highlighted in pale green.

		KNN		SVM		RF		MLP	
		mean	std	mean	std	mean	std	mean	std
accuracy	test	92.93%	0.94%	89.96%	7.68%	94.14%	0.69%	94.14%	1.31%
	train	95.75%	0.15%	91.27%	4.13%	99.86%	0.02%	95.46%	0.21%
f1	test	47.77%	3.46%	42.45%	17.90%	51.61%	3.09%	57.64%	4.63%
	train	68.49%	1.33%	44.24%	14.59%	99.04%	0.17%	67.30%	1.63%
precision	test	52.97%	7.99%	57.30%	23.11%	67.82%	11.47%	63.84%	12.71%
	train	74.56%	1.06%	54.16%	16.25%	98.71%	0.14%	70.87%	1.76%
recall	test	44.18%	4.35%	53.57%	36.47%	43.19%	7.47%	54.00%	5.11%
	train	63.34%	1.75%	54.06%	34.96%	99.36%	0.27%	64.10%	2.22%



**Figure 4: Morphological post-processing.** On the left is shown the raw prediction of the model (here the one from KNN on the *nearby* image). In the middle is shown the intermediate closed images to remove the 'salt' noise. On the right is shown the final post-processed image after subsequent opening to remove the 'pepper' noise.

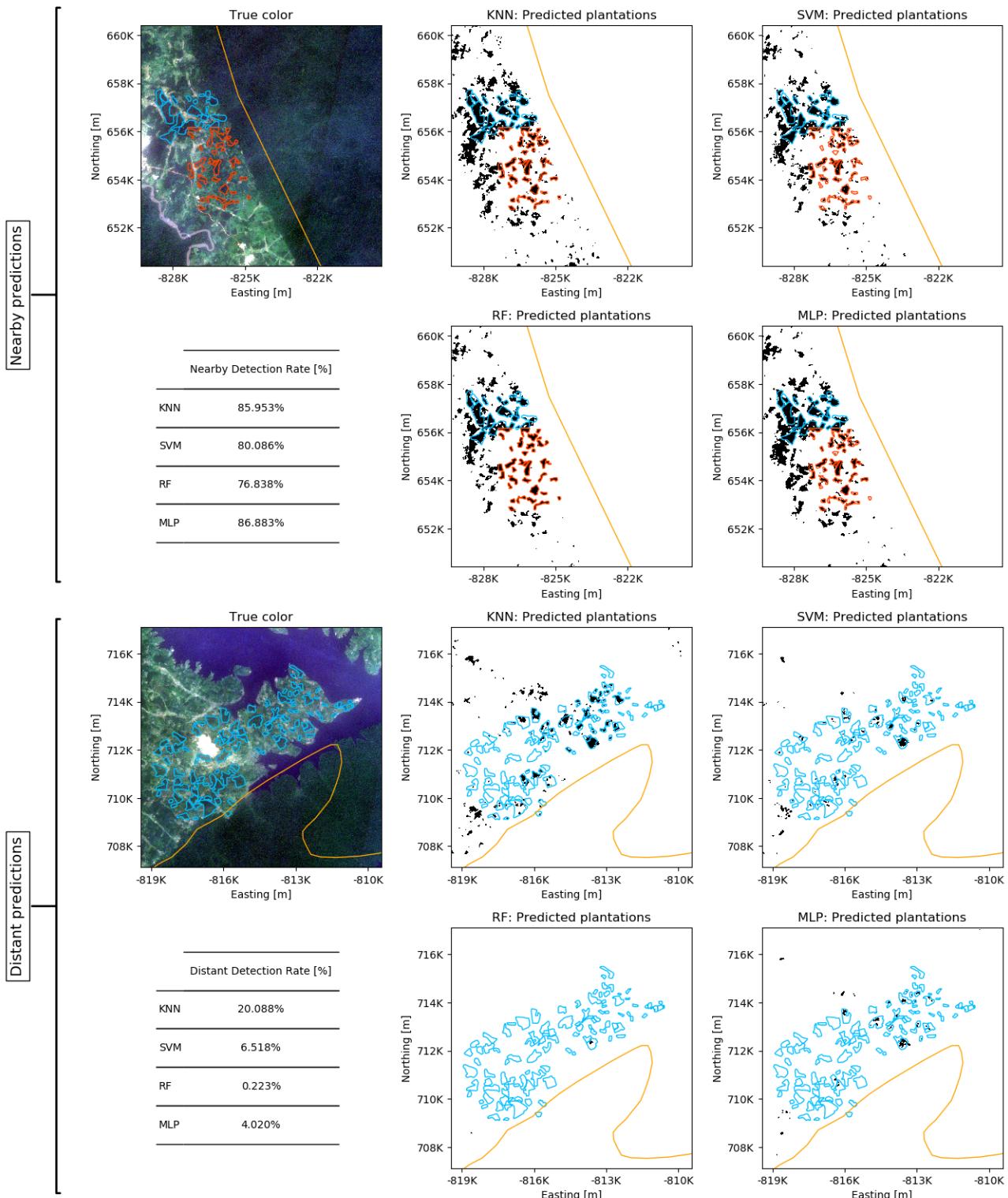
### 3.1.3 Model Generalization

The generalization over space is assessed qualitatively by visually controlling if the models predictions away from the training area detect known plantations and quantitatively by computing the recall. For this task, the model's predictions are post-processed with a morphological closing and opening. The effect of this post processing is presented on figure 4 for the KNN prediction on the *nearby* area. The left image shows the raw KNN prediction. Main area of predicted plantations filled with 'salt' noise (*i.e.* presence of white pixels within the black areas) can be observed. In addition single black pixels are also present in white area ('pepper' noise). The middle image presents the closed prediction, and shows that the 'salt' noise has been efficiently removed and the predicted areas are now uniform to patches. The image on the right presents the post-processed image after the subsequent opening. Only the main detected objects remains (larger than a disk of 50 meters radius) and the 'pepper' noise has been successfully removed.

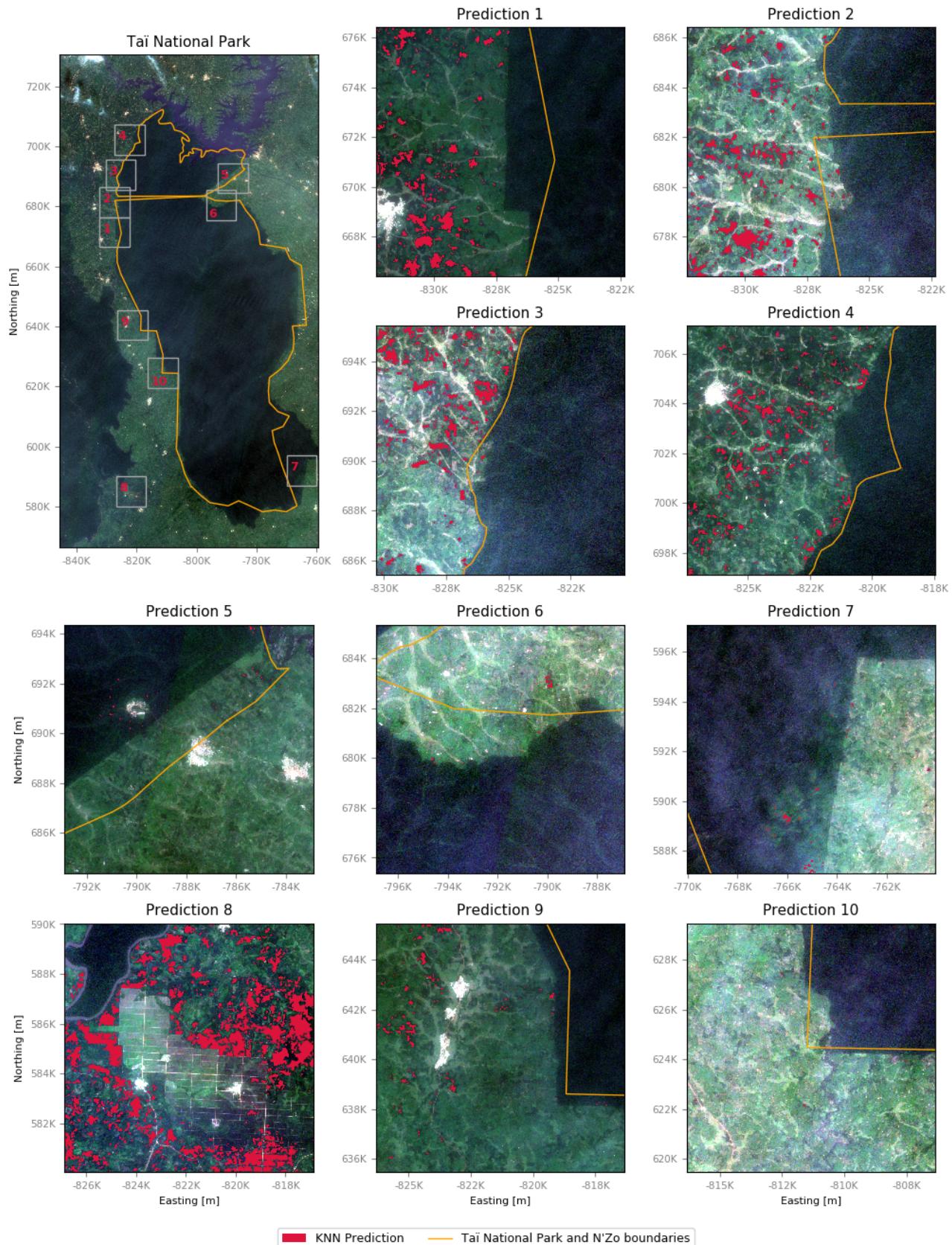
Two areas are used to test the models : a *nearby* area around the training area and a *distant* area on the north of the park (around 60km north of the training area. See figure 1). The results of the four models on both areas are presented on figure 5. The detection rates for each model can be found on the tables of the same figure.

The *nearby* predictions of the four models yield satisfying detection of the few known plantations (blue ones). More precisely, the SVM did not detect some training plantation (red), and present thus a higher qualitative false negative rate. Oppositely, the RF classifier shows a perfect detection of the training plantations, and a still good *nearby* detection but with the lowest detection rate of the four models (76.838%). The perfect detection of the training plantations by the RF is a sign of potential overfitting. The MLP has the highest detection rate (86.883%) and appears to provide a slightly better detection of plantations as there are (qualitatively) less false negative detection in the training area and in the testing plantation (*less white in the polygons*). Finally, KNN seems to qualitatively yield the best detection results (with the post-processing). Indeed, almost all the training plantations have been detected (with some false positive), and the testing plantations are well detected (detection rate of 85.953%).

The four models detect less *distant* plantations as presented on the bottom part of figure 5 and the detection is sparser. More precisely, the RF detected almost no plantations (0.2230%), it thus further highlights the presence of overfitting of this model. Conversely, the MLP does not provide a good detection as, visually, only a few small patches are detected and the detection rate is rather low (4.020%). Similarly, SVM provides also a low detection rate (6.518%). KNN appears to be the best model of the four to detect plantations further away, as the detection is way higher (20.088%) and visually most satisfying.



**Figure 5: Test predictions.** The prediction of the nearby (top) and distant areas (bottom) by the four models are presented. The national park boundary is displayed in orange, the training plantation are presented in red while the control (test) plantations are highlighted in blue. Those test plantation enables to assess whether the models can generalize over space. Each of the predictions presented have been post-processed with a morphological closing followed by an opening. The true color composition is also presented on the left to provide the context of the image. The boundaries shapefiles were obtained from protected planet website [13, 14]



**Figure 6: Taï national park predictions.** The top left global view presents the monitored area. The number inside the area correspond to the number on the titles of the other images. The places have been selected by hand where interesting elements seemed to be present. The red patches represent the prediction made by the KNN model. The boundaries shapefiles were obtained from protected planet [13, 14]

Therefore KNN appears to be the best model of the four which is a different conclusion than from the cross-validation performances (table 2). This discrepancy may be due to the morphological post-processing that favor KNN in a stronger fashion. The raw predictions may have a lot of noise, thus reducing the performances reported by the cross-validation. The morphological operation that removes this noise is likely improving the classification results. In any case, it can be noted that KNN is the most reliable model for monitoring plantations that are far away from the training area.

### 3.2 National park predictions

The KNN model have been used to make some predictions at ten 10x10km places around the park. The results and locations of those places are presented on figure 6. The results from figure 5 suggest that prediction further away from the training area might miss some plantations. This effect should be considered when interpreting them. Predictions 1, 2 and 3 are closer to the training area and are supposed to be more trustworthy. Globally KNN has detected multiple plantations on those images, especially the number 1, 2, 3, 4 and 8. The detection do not appear to be present within the park (here considered as the dark green area). Only prediction 5 and 7 present some small detection inside the dark green area. In the image 5, the detection gathered around a building structure. On image 7, the detection seems to be present on a brighter region of the park.

## 4 Discussion

The engineered features yields results that seems to discriminate the plantations from the rest. The goal of those features (expect the NDVIs) is to introduce local information about the neighboring pixels as the arrangement of the pixels carry meaningful information. Even though the engineered features seems relevant for the task, many other alternatives can be imagined and may likely improve the models performances.

The four models built on the training area yield good classification accuracies but the recall is lower, meaning that the models detect less of the plantations. This is likely due to the presence of 'salt' noise within patches (as in figure 4). A possible source could be that what has been labeled as patch may contain pixels that may not exactly be plantations and are therefore not classified as such. Note that the models performances are computed on the raw prediction (without the morphological post-processing). Therefore the performances presented in table 2 are likely underestimated values. Additionally, the manual labelling performed may have introduced wrongly labelled pixel in the data-set due to human error and approximate regions drawing. Moreover, the labelling from Google Earth images may suffer from temporal gap : it is not ensured that a plantation on Google Earth is already present on the Sentinel images of the 29.12.2018. In consequence, some prediction may thus be meaningful but considered as wrong reducing the performances measured by cross-validation.

In addition, it seems that the RF classifier is overfitting as the train scores are clearly larger than the test ones. The classifier has become too specific to the training set and has trouble to generalize. This is confirmed by the testing prediction in which the RF has in both the *nearby* and *distant* the lowest detection rate. In the *distant* case, the RF detects barely nothing. The models parameters should thus be fine tuned to reduce the overfitting effect. The amount of trees and their maximum depth could be reduced, but this could hinder the power of the model; alternatively, more features could be computed and data augmentation could be performed to enlarge the data-set and thus try to avoid overfitting.

Based on the *distant* testing, KNN appears to be the best model. However, only a single distant location with plantations have been found, labelled and used for testing. In order to better grasp the space generalization, other *distant* locations with plantations should be tested as well.

Indeed the image used for testing has some elements (a water body and a city) that may influence the models capabilities. The presence of water can change the atmospheric humidity, or reflect the light in a different fashion. Those small changes can change what is measured by the remote sensors and the models will be presented with an altered signal for the same object type that could lead to misclassification.

Another dimension that has not been explored in this study is the generalization over time. Indeed, the whole work has been performed on cloudless images of the 29.12.2018 since the goal is mainly to provide a proof of concept rather than an end-product. In consequence, the models developed here may not be applicable for other dates, when the seasons may change the vegetation characteristics. If an actual detector is to be developed, the trained model should be able to generalize over space and time. To achieve this stability, the models should be trained simultaneously on different areas around the park at different dates and seasons.

## 5 Conclusion

The present work aims to explore the use of supervised machine learning algorithms with free satellite imagery for detecting cocoa plantations in the Taï National Park of Ivory Coast. Sentinel-2 images of a single date were used throughout the project and KNN, SVM, RF and MLP algorithms were evaluated for a binary classification task on these images. Manual labelling of plantations was done in two locations around the park to provide training and testing examples. Image processing techniques were applied to perform feature engineering, that included the computation of vegetation indexes, such as NDVI and green NDVI, and texture indexes such as the entropy and LBP. Morphological operators such as opening and closing by reconstruction were also used for feature engineering and image post-processing.

The results obtained show that all algorithms have a good capacity (above 75% of detection rate) for predicting plantations that are nearby the training area, with only RF showing clear signs of overfitting. However, the algorithms show a very poor performance when tested on a distant area with plantations, with only KNN showing a detection rate over 20%. This results prove that it is possible to detect plantations with free satellite imagery within an acceptable accuracy, provided that they are close to the original ones.

Overfitting might be the obvious reason of the poor results when evaluating on distant plantations. Therefore measures like tuning the hyperparameters of the models and the computation of other relevant features could be done. Likewise, enlarging the data-set with labelled examples of other dates and locations would be advisable. Additionally, the evaluation of stronger models such as neural networks deeper than the MLP would be an interesting option.

Overall, the results are promising even when used for locating plantations around the park and they prove the viability of the approach selected for the task. Nevertheless, further work is needed to improve its reliability for generalization, before being able to trust the trained models to monitor the full extension of the park and others around the country.

## Ressources

The code used in this study is available on github at : [https://github.com/antoine-spahr/Cocoa\\_plantations\\_detection](https://github.com/antoine-spahr/Cocoa_plantations_detection).

## References

- [1] Giacomo Grassi, Jo House, Frank Dentener, Sandro Federici, Michel den Elzen, and Jim Penman. The key role of forests in meeting climate targets requires science for credible mitigation. *Nature Climate Change*, 7(3):220–226, March 2017.
- [2] R. A. Houghton, Brett Byers, and Alexander A. Nassikas. A role for tropical forests in stabilizing atmospheric CO<sub>2</sub>. *Nature Climate Change*, 5(12):1022–1023, December 2015.
- [3] Yao Sadaiou Sabas Barima, Akoua Tamia Madeleine Kouakou, Issouf Bamba, Yao Charles Sangne, Michel Godron, Julien Andrieu, and Jan Bogaert. Cocoa crops are destroying the forest reserves of the classified forest of Haut-Sassandra (Ivory Coast). *Global Ecology and Conservation*, 8:85–98, October 2016.
- [4] E. Anderson Bitty, Sery Gonedele Bi, Jean-Claude Koffi Bene, Philippe K. Kouassi, and W. Scott McGraw. Cocoa Farming and Primate Extirpation Inside Cote D'Ivoire's Protected Areas. *Tropical Conservation Science*, 8(1):95–113, March 2015.
- [5] Addo Koranteng, Tomasz Zawila-Niedzwiecki, and Isaac Adu-Poku. Remote Sensing Study of Land Use/Cover Change in West Africa. page 16, October 2016.
- [6] Reeves Meli Fokeng, Walter Gadinga Forje, Vivien Meli Meli, and Bernard Nyuyki Bodzemo. Multi-temporal forest cover change detection in the Metchie-Ngoum Protection Forest Reserve, West Region of Cameroon. *The Egyptian Journal of Remote Sensing and Space Science*, January 2019.
- [7] Fares Al Hasan. Remote sensing for sustainable farming. *EGUGA*, page 13562, April 2018.
- [8] I. K. Nooni, A. A. Duker, I. Van Duren, L. Addae-Wireko, and E. M. Osei Jnr. Support vector machine to map oil palm in a heterogeneous environment. *International Journal of Remote Sensing*, 35(13):4778–4794, July 2014.
- [9] Daniel Tutu Benefoh, Grace B. Villamor, Meine van Noordwijk, Christian Borgemeister, Winston A. Asante, and Kwabena O. Asubonteng. Assessing land-use typologies and change intensities in a structurally complex Ghanaian cocoa landscape. *Applied Geography*, 99:109–119, October 2018.
- [10] Minerva Singh, Damian Evans, Jean-Baptiste Chevance, Boun Suy Tan, Nicholas Wiggins, Leaksmy Kong, and Sakada Sakhoeun. Evaluating remote sensing datasets and machine learning algorithms for mapping plantations and successional forests in Phnom Kulen National Park of Cambodia. *PeerJ*, 7:e7841, October 2019.
- [11] Roman M. Wittig. Taï Chimpanzees. In Jennifer Vonk and Todd Shackelford, editors, *Encyclopedia of Animal Cognition and Behavior*, pages 1–7. Springer International Publishing, Cham, 2018.
- [12] MapHubs. Seven maps that explain the ivory coast cocoa crisis. <https://www.maphubs.com/single-post/2017/09/14/Seven-maps-that-explain-the-Ivory-Coast-Cocoa-Crisis>, 2017. [On-line; accessed 27-12-2019].
- [13] UNEP-WCMC and IUCN. Protected Planet: Taï National Park ID721; The World Database on Protected Areas (WDPA)/The Global Database on Protected Areas Management Effectiveness (GD-PAME). [https://www.protectedplanet.net/green\\_list/tai-national-park-national-park](https://www.protectedplanet.net/green_list/tai-national-park-national-park). [On-line; accessed 13-10-2019].
- [14] UNEP-WCMC and IUCN. Protected Planet: N'Zo Fauna Reserve ID2293 ; The World Database on Protected Areas (WDPA)/The Global Database on Protected Areas Management Effectiveness (GD-PAME). <https://www.protectedplanet.net/2293>. [On-line; accessed 21-12-2019].