Mapping cocoa in Côte d'Ivoire and Ghana with deep learning

Nikolai Kalischek[1,*], Nico Lang[1], Rodrigo Caye Daudt[1], Thomas Addoah[3], William Thompson[4], Wilma J. Blaser-Hart[5], Rachael Garrett[3], Konrad Schindler[1], and Jan Dirk Wegner[1,2]

[1]EcoVision Lab, Photogrammetry and Remote Sensing, ETH Zurich, 8092, Zurich, Switzerland
[2]Data Science for Sciences, Institute for Computational Science, University of Zurich, 8092, Zurich, Switzerland
[3]Environmental Policy Lab, Department of Humanities, Social and Political Sciences, ETH Zurich, 8092, Zurich, Switzerland
[4]Department of Biology, University of Oxford, Oxford, OX1 2JD, Great Britain
[5]School of Biological Sciences, The University of Queensland, St Lucia, Brisbane, QLD 4072, Australia
[*]nikolai.kalischek@geod.baug.ethz.ch

ABSTRACT

Cocoa production has been a main driver of deforestation in West Africa for decades, leading to a substantial loss of forest cover. Côte d'Ivoire and Ghana, the the world's largest producer of cocoa, experienced the highest rates of deforestation in recent years. Yet scalable and transparent procedures for mapping cocoa farming are missing, and neither the public sector nor the industry have accurate maps at their disposal. Here, we combine cocoa farming data with publicly available satellite imagery in a deep learning framework and create high-resolution maps of cocoa plantations for both countries. Our cocoa map is the first to be validated in situ, and our model can detect full-sun as well as agroforestry cocoa farms across entire countries with high precision, including generalization beyond the training regions. Our results suggest that official figures significantly underreport the harvested areas, and provide evidence for both undeclared plantings and an overestimated yield per hectare.

In particular for this work, we have trained a neural network on a dataset of $>$100,000 geo-referenced cocoa farms to map cocoa plantations at country scale. We leverage publicly available Sentinel-2 satellite imagery as input in a twofold way. First, we train a neural network to predict canopy height in the sub-Saharan region utilizing ground truth acquired from the GEDI LIDAR mission. Second, we train a deep neural network on the same satellite imagery and a large corpus of polygons delineating cocoa farms, using the canopy height map as an additional input for the network. With the help of a team in Côte d'Ivoire, we validate our map *in situ* in a three months process accompanied by further verification with a hand-labelled test set in Ghana. Our map suggests an estimated planted area of 4.3 million hectare and 2.7 million hectare in Côte d'Ivoire and Ghana, respectively, which is dramatically higher than official figures in Ghana. Our map has direct implications, not only on harvested area, but also on farming practices and sustainability efforts on reducing deforestation and highlights the need for land cover mapping independent of farmers, industry and governments. We illustrate the utility of our map in further downstream tasks to identify regional areas that are exposed to poor growing conditions.

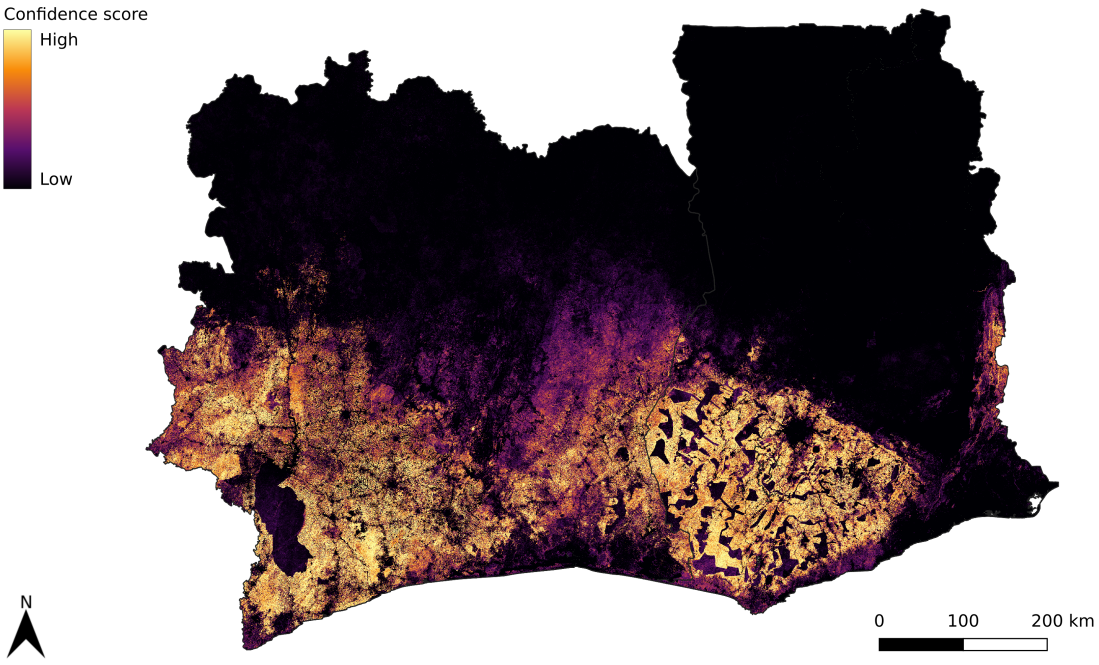

Figure 1. Cocoa map for Côte d'Ivoire and Ghana. Confidence map with 10×10 m ground sampling distance. The map indicates detection confidence on a range $[0\ldots1]$, i.e., values near 1 indicate that model predictions across most time steps agree on the presence of cocoa.