

Machine Learning-based Estimation of Forest Carbon Stocks to increase Transparency of Forest Preservation Efforts

Björn Lütjens, Lucas Liebenwein and Katharina Kramer
Massachusetts Institute of Technology

Abstract—An increasing amount of companies and cities plan to become CO₂-neutral, which requires them to invest in renewable energies and carbon emission offsetting solutions. One of the cheapest carbon offsetting solutions is preventing deforestation in developing nations, a major contributor in global greenhouse gas emissions. However, forest preservation projects historically display an issue of trust and transparency, which drives companies to invest in transparent, but expensive air carbon capture facilities. Preservation projects could conduct accurate forest inventories (tree diameter, species, height etc.) to transparently estimate the biomass and amount of stored carbon. However, current rainforest inventories are too inaccurate, because they are often based on a few expensive ground-based samples and/or low-resolution satellite imagery. LiDAR-based solutions, used in US forests, are accurate, but cost-prohibitive, and hardly-accessible in the Amazon rainforest. We propose accurate and cheap forest inventory analyses through Deep Learning-based processing of drone imagery. The more transparent estimation of stored carbon will create higher transparency towards clients and thereby increase trust and investment into forest preservation projects.

I. THE PROBLEM

Deforestation and forest degradation are responsible for ~15% of global greenhouse gas emissions, as burning forest releases stored carbon into the air [1]–[3]. Stopping deforestation and forest degradation and supporting sustainable forestry hence mitigates climate change and also preserves biodiversity, prevents flooding, controls soil erosion, reduces river siltation, and offers a workplace for the local population [1]. Despite the paramount importance of reforestation and preservation efforts, they are far from sufficient, mostly because of a lack of financing [1], [4]. This financial gap is created by a severe lack of trust into reforestation and preservation projects as they are not transparent in their CO₂ impact to client companies that want to offset carbon emissions [4], [5].

Forest inventories are common practice in forestry, account for 25% of the expenses of reforestation and estimate stored carbon [4]. Apart from carbon estimation, forest inventories are also created to identify illegal logging, control pests and diseases, estimate the opportunity cost of deforestation, manage wildfire hazards, and achieve sustainable forestry. Classical forest inventories are created through manually counting and classifying trees in a ~7–15m radius every 0.09–1km [4]. The sparse samples are interpolated, recently with the help of satellite imagery, to create an inventory for the whole forest. Ground-based sampling, however, is prohibitively expensive (~300USD/ha), and time-intensive (2–7 days/20ha) in large-scale rainforests, due to dense vegetation, a large team of experts, and scarcity of roads [4].

Purely satellite-based approaches mostly use publicly



Fig. 1: Medium-resolution drone imagery, collected during pilot flights near San Ramón, Perú.

available RGB-NIR satellite imagery, or radar. As the low-resolution (RGB max. 30cm/px, radar ~250m/px) does not suffice to accurately determine the tree count, species, or height, most satellite-based approaches only measure area covered by forest which leads to rough estimates of carbon sequestering potential with high uncertainties [4], [6]. LiDAR-based approaches, used in US forests, are very accurate, but hardly-accessible and cost-prohibitive for low budget reforestation and preservation projects, because of the expense of the sensor and the bigger carrying drone, or plane [4], [7].

II. THE SOLUTION/INNOVATION

Our goal is to increase investment into reforestation and preservation projects to combat climate change by providing an accurate, cheap, and transparent carbon storage analysis. The analysis is supplied to reforestation and preservation projects that, with the analysis, have sufficient trust to convince their client companies to higher investments.

A. Technical Solution

The proposed forest inventory assessment consists of an on-site data collection and an off-site processing part. During the data collection with the local partner, a low-cost quadrotor (DJI Phantom 4 Pro, 1.5kUSD) and five batteries have to be used to map 100ha in ~5hrs with one operator (~10USD/ha) for Fig. 2. DroneDeploy was used to plan the flight and mosaic the images. The next iteration will be an off-the-shelf, low-cost vertical take off and landing (VTOL) fixed-wing drone to cover up to 250ha in one 60min flight and launch in dense forests. The drone will be equipped with a gimbal, 4k GoPro RGB camera, and a Sentera NDVI-IR camera.

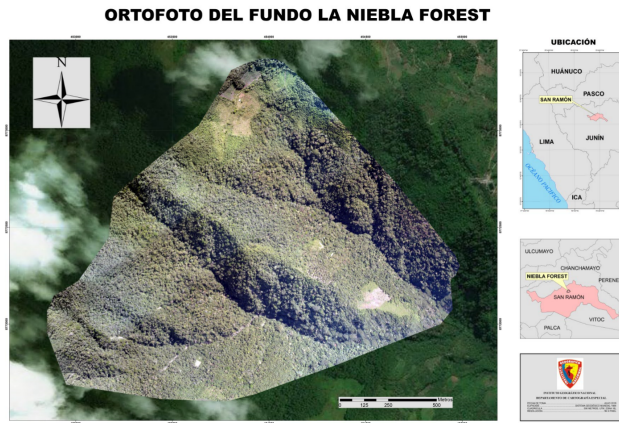


Fig. 2: Collected map from pilot flights with the National Geographic Institute of Peru near San Ramón.

Deep Learning algorithms are proposed to extract crown diameter, species, and count of emergent and canopy trees. Specifically, a pixel-wise segmentation algorithm, based on DeepLabv3+ [8], a Convolutional Neural Network (CNN) architecture, will classify the tree species at each pixel of the collected RGB-NIR imagery and extract crown diameter and tree count. The expected success of the algorithm assumes that a canopy’s RGB-NIR spectrum and shape strongly correlate with the tree species. The correlation is shown for high-resolution sensors in [9], [10], but needs to be validated with the available low-cost sensors in future results. Additionally, a Bayesian regression model with spatial random effects [11] with the same in- and outputs is being developed to increase overall accuracy via model ensembling, and counteract the inaccuracy of the CNN model on novel data.

In addition to crown diameter and species, the estimation of forest carbon stocks requires canopy heights (distance from ground to canopy). Canopy heights cannot be accurately inferred from drone imagery, because visibility of the forest floor is prohibited by dense vegetative cover. Hence, a digital surface model (DSM) of the surface heights (distance from sea level to canopy), based on GPS, IMU, and structure from motion was created with the DroneDeploy software. A satellite-based digital elevation/terrain model (DEM) (distance from sea level to ground) will be subtracted from the DSM to obtain the canopy height model (CHM). The accuracy of the approach will be benchmarked on ground-based inventories.

Allometric equations can be used to calculate forest biomass and carbon stocks, from canopy height, crown diameter, and species [6], [12]. The accuracy of multiple allometric equations for tropical rainforest, and Andean rainforests that do or do not contain information about the tree species [13] will be evaluated.

An accurate, but small dataset [14] with tree height, species and crown segmentation is used. A larger dataset will be created by fusing ground-based and remotely sensed inventories of well studied forests (e.g., US national forests [16]).

B. Partnerships

- A very close connection to a local community partner, which offers 100 hectares of rainforest in San Ramón, Perú as testing ground has been established. The com-

munity partner visits local mayors, and schools, and creates social media initiatives to reduce deforestation. The partner has started a small-scale reforestation project.

- NGOs and ministries have been visited to access data, co-develop software, and deploy it at scale
- We are continuously reaching out to gain knowledge in Forestry, Citizen Science, and Remote Sensing.

C. Scalability

As the approach is scaled to larger areas of forest, the local communities will be involved in the monitoring of preservation projects to make them feel responsible and technologically capable to protect their forest. To do so, an app will be developed that allows locals to map forests and scale up the data collection nationally. The app will be rolled out to the community partners’ network of volunteers and local municipalities that possess a drone.

For the long-term, the cheap, and accurate ML-based carbon inventories are proposed to be embedded as standard in the cap-and-trade carbon market. The California Air Resources Board currently considers a bill to integrate CO₂ offsets from tropical reforestation. This would allow reforestation and preservation projects to earn 10–15 USD per ton of sequestered CO₂ and incentivize locals, strongly concerned about monetary aspects, to sustain primary forests. Forests would be a competitive carbon offsetting choice, because they store a ton of CO₂ at roughly 20–25USD (6–8 trees; one tree costs ~3 USD (30% seedling, 45% labour, 25% monitoring)) [4], whereas carbon capturing plants convert CO₂ at a price of 94–232 USD/tCO₂ [17].

The proposed method to infer forest inventories can also help reduce illegal logging. Timber companies are allotted internationally salable trees based on forest inventories of their land. The inventories, however, can be untruthfully overestimated, and companies sell rare and valuable trees from outside of their land. The proposed method can be used to cheaply verify the reported inventories of tree species.

III. IMPACT

Although mitigating climate change is this project’s main goal, success is measured via the UN sustainable development goal 15.1.1, the “ratio of total land covered by forest“, to incorporate the beneficial side effects of forest cover. As this project is trying to increase the amount of trust and understanding that people have for carbon offsetting initiatives, e.g. reforestation, it is trying to change the bigger system. While at the beginning, it would be a success to increase investment into one offsetting project, the project aims for a large scale impact where people are more aware of how much effort it takes to offset their emissions, make them more environmentally conscious, and make investments into reforestation for carbon offsetting a standard.

A. Ethical considerations

- An accurate forest inventory must be stored securely to prevent misuse for finding and logging rare trees
- Best practices for wildlife monitoring are respected [18]
- Drone flights must be restricted via GPS to only fly over approved government or private land

IV. ACKNOWLEDGEMENTS

The authors want to thank La Niebla Forest for hospitality and support in the local community; World Wildlife Fund (WWF) Peru, Peru Ministry of Agriculture - National Wildlife and Forest Service (SERFOR), Peru Ministry of Environment - National Forest Conservation Program (BOSQUES), Peru National Geographics Institute (IGN), VividEconomics, WeRobotics, and UAV Peru for helpful discussions about the difficulties of reforestation and forest conservation; Prof. Newman, Prof. Wood, Prof. Fernandez, Prof. How, and Prof. Rus for their advice on remote sensing, UN politics, carbon sequestration, and robotics; MIT Sandbox Innovation Fund, MIT PKG IDEAS Global Challenge, MIT Legatum Seed Travel Grant, Microsoft AI for Earth Grant, and NASA Space Grant for their support. The work is conducted by the Sustainable AI Initiative {sustainable-ai.mit.edu}, at the Massachusetts Institute of Technology, 77 Mass. Ave., Cambridge, MA, USA. lutjens@mit.edu.

REFERENCES

- [1] Forest Carbon Partnership Facility (FCPF), "2018 annual report."
- [2] UN-REDD, "Reducing emissions from deforestation and forest degradation and the role of conservation, sustainable management of forests and enhancement of forest carbon stocks in developing countries (redd+)." [Online]. Available: <https://www.un-redd.org/>
- [3] IPCC, "Climate change 2014: Synthesis report. contribution of working groups i, ii and iii to the fifth assessment report of the intergovernmental panel on climate change [core writing team, r.k. pachauri and l.a. meyer (eds.)]" 2014.
- [4] "Interviews with la niebla forest, world wildlife fund (wwf) peru, peru ministry of agriculture and irrigation - national forest and wildlife service (minagri - serfor), peru ministry of the environment - national forest conservation program (minam - bosques), national institute of geographics (ign) peru, werobotics, vivideconomics, hartree, and weforest," 2018-19.
- [5] Alcoa, "2017 alcoa sustainability report," 2017.
- [6] H. Gibbs, S. Brown, J. O Niles, and J. A Foley, "Monitoring and estimating tropical forest carbon stocks: Making redd a reality," *Environmental Research Letters*, vol. 2, p. 045023, 12 2007.
- [7] S. Zolkos, S. Goetz, and R. Dubayah, "A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing," *Remote Sensing of Environment*, vol. 128, pp. 289 – 298, 2013.
- [8] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *Computer Vision – ECCV 2018*. Springer International Publishing, 2018, pp. 833–851.
- [9] M. A. Cochrane, "Using vegetation reflectance variability for species level classification of hyperspectral data," *International Journal of Remote Sensing*, vol. 21, no. 10, pp. 2075–2087, 2000.
- [10] J. Vauhkonen, T. Tokola, P. Packalen, and M. Maltamo, "Identification of scandinavian commercial species of individual trees from airborne laser scanning data using alpha shape metrics," *Forest Science*, vol. 55, no. 1, p. 37–47, 2009.
- [11] A. O. Finley, "A bayesian approach to multisource forest area estimation," *Proceedings of the seventh annual forest inventory and analysis symposium*, pp. 261–264, 2007.
- [12] Gold Standard, "Gold standard afforestation/reforestation (a/r) ghg emissions reduction and sequestration methodology."
- [13] J. Chave, M. Réjou-Méchain, A. Búrquez, E. Chidumayo, M. S. Colgan, W. B. Delitti, A. Duque, T. Eid, P. M. Fearnside, R. C. Goodman, M. Henry, A. Martínez-Yrizar, W. A. Mugasha, H. C. Muller-Landau, M. Mencuccini, B. W. Nelson, A. Ngomanda, E. M. Nogueira, E. Ortiz-Malavassi, R. Pélissier, P. Ploton, C. M. Ryan, J. G. Saldarriaga, and G. Vieilledent, "Improved allometric models to estimate the aboveground biomass of tropical trees," *Global Change Biology*, vol. 20, no. 10, pp. 3177–3190, 2014.
- [14] "Neon data challenge: Identifying trees using remote sensing data," 2017.
- [15] Harvard, "Tree inventories for validating terrestrial lidar measurements at harvard forest since 2007."
- [16] NSF-NEON, "Nsf neon woody plant vegetation structure dataset."
- [17] D. W.Keith, G. Holmes, D. S. Angelo, and K. Heidel, "A process for capturing co2 from the atmosphere," *Joule*, vol. 2, pp. 1573–1594, 2018.
- [18] J. C. Hodgson and L. P. Koh, "Best practice for minimising unmanned aerial vehicle disturbance to wildlife in biological field research," *Current Biology*, vol. 26, no. 10, pp. R404 – R405, 2016.