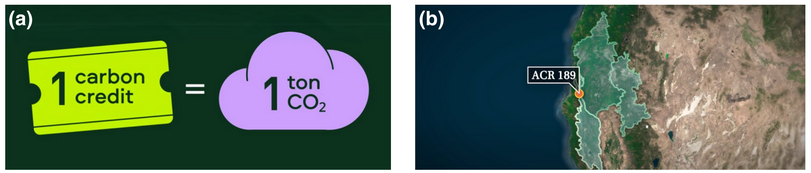
Advancing Carbon Credit Accuracy: A Machine Learning Approach for Precise Carbon Sequestration Estimation

Eric Mao[1]†, Tom Zheng[1]†, Chengtai (Richard) Li[1], Ken Cheng[1], Hudson Haas[1]

[1]Crescent School

†Authors contributed equally to this work

Introduction

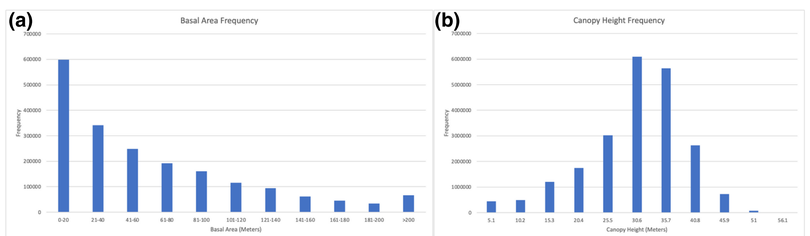


(a) CO_2 sequestration for 1 carbon credit (b) Super-sectioned eco sectors causing error baselines (ex. ACR 189) [1]

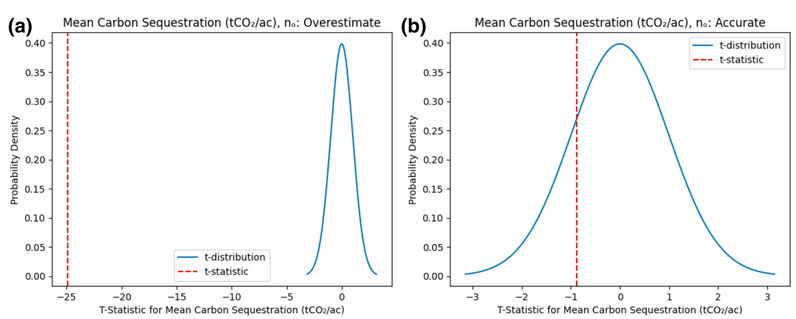
- Forests have **absorbed 30%** of all anthropogenic emissions in the last few decades [2]
- Carbon credits preserve forests for CO_2 offsets
  - **29.4%** over-crediting error due to crude baseline [1]
  - Highly time-consuming recrediting

Data Acquisition

- Global Ecosystem Dynamics Investigation Global Canopy Data
- U.S. Forestry Service Tree Species Data
- USDA + USFS + Esri ArcGIS Tree Species Density Data



(a) Frequency chart of basal area in CAR1190 (b) Frequency chart of canopy height in CAR1190



(a) T-test of modeled data against n_o=overestimate (b) T-test of modeled data against n_o=accurate

Highlights

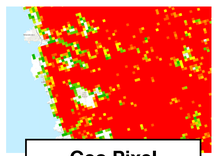
- **Effective and scalable carbon crediting solution**
- **Accurate machine learning model**
- **Correct over-crediting errors by 29.4%**

System Overview

What is the Carbon Sequestration Rate of...

Geo Data → **Post Processing** → **Mathematical**

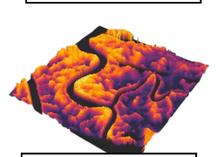
USDA Tree Frequency



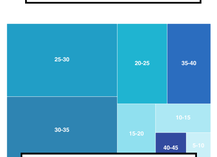
Geo Pixel Analysis

$$\sum_{i=1}^{n} \left(33.204 a H_i \sqrt{A_i}\right)$$

Sum of all Carbon Sequestration

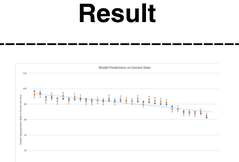


LiDAR Canopy Heights



Processed Pixel Data

Result



Model Prediction

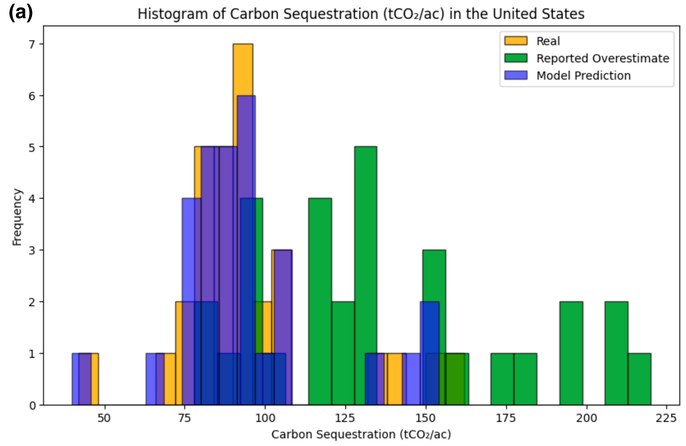
Sampled 30 conservatories, 150 million trees

Performance Comparison

	Plot Extrapolation	Lidar Analysis
Speed	Real-Time	Post Processing
Error Rate	±100%	±5%

Lidar Analysis provides an **20x improvement** in Error Bound in return for extra compute and time resources for post processing

Results



(a) Machine learning model prediction (blue) against overestimate (green) and accurate (real) data for 30 forests

A T-test against accurate mean yielded a p-value of p = 0.864

Summary, Analysis, & Future Work

- Processed scalable geographic data utilizing software programs
- Machine learning to calculate carbon sequestration in selected regions

In **future,** we plan to 1. Refine plots to include more complex geographic aggregation for more robust carbon calculations; 2. Implement model for *in situ* forest carbon sinks to better determine carbon sequestration over time anywhere in the world

Acknowledgements

This work was made possible through the support from Mr. David Lesbarreres, Mr. Michael Stephens, and Crescent Alum Adam Omarali

References

[1] Grayson Badgley, Jeremy Freeman, Joseph J Hamman, Barbara Haya, Anna T Trugman, William RL Anderegg, and Danny Cullenward. Systematic over-crediting in california's forest carbon offsets program. Global Change Biology, 28(4):1433–1445, 2022.

[2] Pan, Y. et al. Science 333, 988–993 (2011).