The Canopy Project

An application for better understanding forest structure and health

Maria Granroth, Robert Plueger, Dongyu Xia

College of Engineering, Informatics, and Applied Sciences
NAU Northern Arizona University

Introduction

The goal of this project is to develop an application for better understanding forest structure and health. We can do this by utilizing GEDI data!

The *Global Ecosystem Dynamics Investigation (GEDI)* has been performing high resolution laser ranging of Earth's forests and topography from the International Space Station.

Our client, Dr. Jantz, is a member of the *Vegetation Structure as an Essential Biodiversity Variable (VSEBV)* project based at NAU and is also a team member funded by the *NASA Applied Sciences* program. He uses the data from GEDI to develop vegetation structure essential biodiversity variables (EBVs). Scientists can use EBVs to improve land use decisions and conserve the biodiversity in tropical landscapes.

Unfortunately, GEDI data is hard to process and analyze for most people. Currently, the user needs to be able to use the R language, and it is inconvenient and time consuming. They need to download the GEDI data, process the files and find the intersections of data that they need for a given area, organize the relevant labels, and analyze the processed files. The result of the analysis can also be hard to comprehend. This can take several hours of the user's time and effort.

We have designed an easy-to-use web application to help overcome these problems and make GEDI data more accessible to a wider audience. Using our application, the user just needs to choose the area that they want to analyze and the statistical functions that they want performed on that area. Our application will automatically generate the result and send it to them via email.

Our Solution and its Key Features

Diagram 1 shows the detailed user process.

Choose the area - Choose a South American country via the dropdown menu or upload your own shapefile that describes the area to analyze.
Choose the statistical functions - Select the statistical function(s) you wish to have performed on the area. Here, you can also choose to ask for the raw, unanalyzed data.
Enter your email - to receive the resulting files, enter your email.
Hit submit and wait! Our application will generate csv files from relevant analysis results and send them to users via email. If the user chooses to also download the raw data, the application will also send them the files generated from processing the GEDI files for that area.

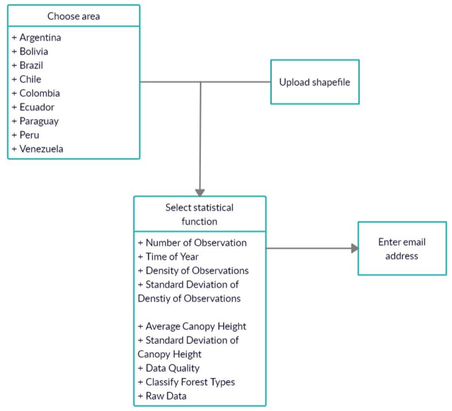


Diagram 1 – User Process

Architecture

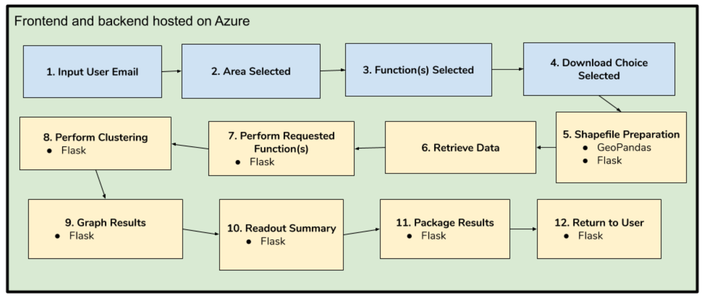
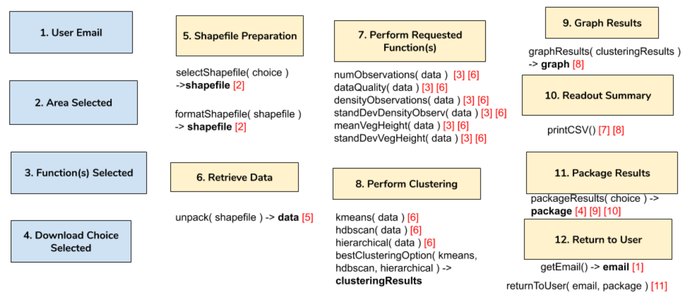


Diagram 2 – Architecture Workflow. Blue boxes indicate what modules the user will directly interact with through the website. Yellow boxes indicate what the application will do for them.

Diagram 3 – Module Interaction. Base functions required for each module are listed below each one. The bolded variable name is the output of a function. The red numbers in brackets denote the number of the module from which it requires data.



Outcomes

Our system has several main impacts. The web application we produced takes the strain off the user. They will no longer have to manually work through the process of gathering and analyzing data themselves, as our product will offer to do that for them. Additionally, our product will make it easier for users to access the data they need. More importantly, it will make it easier for researchers and policy makers to gather information and evaluate environmental variables. This will allow for a better understanding of the earth's ecosystems, and we will be able to redirect resources accordingly.

Technologies

Flask - Python web framework that allows the web interface to interact and communicate with external servers.
GeoPandas - extension of the Python library, "Pandas," that works with shapefiles. Shapefiles are files that store a location. We are using this library to verify that the shapefiles our system receives from the user are valid shapefiles.
Hd5py - Python package which has the purpose of unpacking and reading .hdf5 files, which is the form of GEDI files, into files we can use for analysis.
Microsoft Azure - web hosting service that is able to provide a reachable public domain for our web application.

Future Work

Due to the COVID-19 pandemic, our project expectations have changed. Our client has two primary goals for version 2 of this application. Module 8 in *Diagrams 3* and *4* mentions clustering, which would be used to classify forest structures. Module 9 would graph the results of that clustering. Our application sets a strong base for our client to add many more statistical functions alongside the clustering. The second goal would be to make the website public by successfully hosting it on Microsoft Azure or something similar, such as Amazon Web Services. From that point, further goals may include adding more countries for users to choose from, or other options for improved user customization.

Acknowledgements

We would like to thank Scooter Nowak for being our mentor throughout this whole process, and Dr. Patrick Jantz for giving us the opportunity to work on this project. Special thanks to Dr. Doerry for offering and assisting us with the NAU Capstone experience.