Identification of native vegetation species by linking UAV imagery with terrestrial spectroradiometers

SGSD lead supervisor: Dr Fernando Benitez-Paez
External supervisors: Dr Marcela Camacho, Universidad de Colombia, Dr Urška Demšar, St Andrews.
PhD programme: PhD Geography (Science)

Introduction

After nearly sixty years of conflict, in November 2016 a peace agreement was signed by the Colombian government and the largest and longest armed group in the region (Revolutionary Armed Forces of Colombia-FARC). This agreement includes a comprehensive rural reform to improve the access and use for one million hectares of public land registered in the National Lands Fund. To identify new uses of land and define how local farmers can contribute to this new rural reform, national governments created what is called the agricultural frontier. This is a process of identification and classification of the so-called micro-areas, which are suitable to be croplands with emphasis on native vegetation. Traditional agricultural surveying techniques include the use of satellite imagery, in-situ surveys, as well as integration of multi-scale data sources in which national or regional administrations define the required agriculture frontier. However, none of the existing methods are suitable for identification and classification of small cropland regions with native plants.

Conventionally, satellite imagery missions have been one of the most reliable sources of information for study of the characteristics of vegetation. However, the use of satellite imagery for agriculture comes with two significant challenges. The first problem is the cloud cover, which, as our study area is in equatorial region, is often present in the area. Optical remote sensing instruments cannot see through the clouds, which therefore create a barrier, restricting the viewing of the earth surface from space. A second issue is related to the spatial resolutions of available data, that are too coarse to resolve features that are important to agriculture. This challenge is particularly significant when the causes of variation in vegetation performance can only be identified using spatial patterns, or textures that become visible at relatively fine spatial resolutions. Satellite imagery, therefore, does not provide a complete remote sensing solution when it comes to the fine-grain analysis required to identify native species. Thus, other remote sensing platforms, like the so-called close-in remote sensing (e.g., Unnamed Aerial Vehicles-UAV or Drones) are an adequate, and relative low-cost approach to fill the mentioned gaps. In recent years, drones have become a practical and widely applied remote sensing tool in precision agriculture and studies have shown the use of UAV systems for biological observation capable of filling critical spatial and spectral observation gaps in mapping plant species and plant functional types that have been difficult to measure from space-borne sensors.

Although remote sensing from drones is a powerful tool, there are some caveats to be considered, the accurate interpretation still requires a certain level of expertise. A sample measured directly on the ground by an expert will often be the most accurate method of identifying the causes of variations in aerial data, but that will limit the study area and usually is the most expensive method. Another semi-automatic process is the classification of the collected spectral signatures by using a collection of laboratory-based spectral signatures used as proxy to provide a certain level of automatization in the characterization of an aerial map. The problem is, those spectral signatures are rarely available and only calculated for the usual type of croplands present in large areas (e.g., rice, corns, grasslands), but that is usually an expensive process defined in a controlled environment that might vary from the external conditions and therefore increase the uncertainty of any automatic classification process. Overall, field spectroradiometers are traditionally used in biology for DNA sampling of specific species and small areas. At the same time, drone imagery is typically used to classify crop rows and, more recently, to identify vegetation characteristics with the support of post-processing methods in larger areas.

To date however, no one has combined spectral signatures collected in the field using field spectroradiometers and UAV multispectral imagery, which would be required to create an open collection of spatial signatures that serve as proxy for semi-automatic classification methods of croplands. This is particularly challenging not only due to the lack of availability of the proper equipment but also for native species that only grow under certain conditions, and which are otherwise not present. We propose an interdisciplinary PhD project where spatial data science meets biology and sustainable development and employs a new data-driven approach to study the reflectance response of native species under particular geographical and environmental conditions (figure 1).
With the support of the recently launched GIScience and Remote Sensing laboratory at the National University located at the northern of Colombia in the department of Cesar, we will collect a set of different spectral signatures under different environmental conditions defined by its tropical rainforest characteristics (Sierra Nevada de Santa Marta). These spectral signatures will be collected in different environmental conditions and periods of time to provide different profiles of reflectance measures for each species, with a focus on native species like palm oil, coffee, platano (wild banana) and cacao, which are typical of this region.

In parallel with the collection of the spectral signatures from field spectroscopy, a set of UAV multispectral imagery will be collected, using a wing-fixed drone eBee with a multispectral camera. Different areas will be mapped in different periods of the rainfall and dry sessions, to capture a whole set of imagery for the post-processing and classification process. A data fusion method will be deployed to integrate the multispectral imagery from UAV and the ground-truth spectroscopy collected for the study areas and native species. Finally, as the last step a classification and identification process will be applied to validate or calibrate the processed spectral signatures that identify the reflectance response for each study species, in a particular environmental condition. This will be later be integrated within an open and accessible collection of spectral signatures for the use of other researchers or biologists interested in classification and identification using multispectral imagery.

Methodology. The project will be organized in three steps (figure 1):

Step 1: Definition of the study areas for the selected native species in a controlled environment where we can study the environmental conditions and gather the spectral signatures using the on-field spectroradiometers to get the most precise measure of the reflectance for the study species.

Step 2: Using the suggested drone, capture the multispectral imagery required as input for the data fusion method. This data collection process will be aligned with the environmental conditions presented in the study areas as well as the collection of multiple images for several periods of the year to have a heterogeneous sample of spectral UAV imagery for classification and calibration purposes. Step 2.1: Create a data fusion method to integrate drone data and the spectral signatures collected in the study areas. An extensive analysis of the reflectance response in both data sets will be considered in this step, likewise the integration of machine learning models to predict the values captured from the UAV imagery.

Step 3: Once the data has been properly integrated, we will use other UAV multispectral imagery to validate and calibrate a classification and identification process and employ this over larger areas. A collection of curated and available spectral signatures from field-spectrometry as ground truth and UAV imagery as use case will provide the confidence required to let other researchers or biologists classify an identify these native species in other multispectral imagery.

Student profile, training, and resources

Candidate should have a solid technical background in spatial data science, geoinformatics or remote sensing, and an interest in modelling of biological processes. Student should also have coding experience and (preferably R or Python), as well have been collecting or processing UAV imagery, would be ideal. Student will receive training in methodological research, relevant biological aspects, publishing and other-research-related skills. This will be delivered both through supervision and through courses on transferable skills (presentation, project management, outreach) from the Centre for Educational Enhancement and Development (CEED) of the University of St Andrews. The student will be expected to present their work at national and international conferences throughout their study and take external courses as needed and will receive a Research and Training Support Grant (RTSG) for this purpose (£2000/year, see below). By the end of the project, the work is expected to result in three journal publications in spatial data science or biological journals, which will also serve as the basis for the PhD thesis.

RTSG: This project will include field trips and visits to Colombia to collaborate with researchers there and obtain data for methodological development (which will take place in St Andrews). We therefore budget for an annual visit to Colombia as per this: travel (flight+local travel) £1500, accommodation £300 and subsistence £200, for a total of £2000 per year.