1st GEO BON – Microsoft joint call: EBVs on the cloud

AMAZECO: Covering the Amazon with an Ecosystem Structure EBV product combining satellite and airborne LIDAR

Rubén Valbuena (1), Eric B. Görgens (2) & Carlos A. Silva (3)

(1) Bangor University, UK.
(2) Universidade Federal dos Vales do Jequitinhonha e Mucuri, Brazil.
(3) University of Florida, US.

5th June, 2020

1. Summary table

Project title	An Ecosystem Structure EBV product covering the Amazon combining satellite and airborne LIDAR
PI (Team member 1)	Rubén Valbuena. Bangor University School of Natural Sciences. Thoday Building Bangor LL57 2UW. United Kingdom
Co-PI (Team member 2)	Eric B. Görgens. Universidade Federal dos Vales do Jequitinhonha e Mucuri (UFVJM), Brazil
Co-PI (Team member 3)	Carlos A. Silva. University of Florida, USA
Team member 4	PDRA (Post-Doctoral Research Associate). Bangor University, UK (candidate to be selected).
Team member 5	PDRA. Universidade Federal dos Vales do Jequitinhonha e Mucuri, Brazil (candidate to be selected)
Partner collaborator	Jean P. Ometto Brazilian Institute of Space Research (INPE)
Partner collaborator	Michael Keller International Institute of Tropical Forestry (US Forest Service) Jet Propulsion Laboratory (NASA) EMBRAPA – CNPTIA, Brazil
Keywords	Ecosystem structure; ecosystem vertical profile; LIDAR; laser scanning; global ecosystem dynamics investigation (GEDI); ecosystem height; ecosystem vertical complexity; gap size and frequency; ecosystem layering.
Requested budget	US$ 100,000
Requested Azure Credits	US$ 99,971

2. Project abstract

Ecosystem vertical profiles (EVPs) characterize the vertical distribution of sessile biological entities in an ecosystem, which affects the number and variety of potential niches and identifies critical aspects of ecosystem state. In Valbuena et al. (2020) [1] we advocate for a standardization of ecosystem morphological traits derived from EVPs characterized by LIDAR, so that they can become useful to inform ecosystem structure EBVs. These traits should focus on being relevant to the ecosystem, and not on the means for measuring them. Thus, the goal of this project is to demonstrate that we can **deliver platform-independent EVPs from both satellite and airborne LIDAR sensors**, and provide the means for a global ecosystem structure LIDAR product that can be crowdsourced through national BONs. This will be enabled by **high performance computing (HPC) workflows for common satellite/airborne LIDAR derivation of ecosystem traits**, which we will produce and implement into **a first prototype product covering whole of the Amazon with traits produced from combined satellite and airborne LIDAR**. The satellite LIDAR will be obtained from the currently operational global ecosystem dynamics investigation (GEDI) mission. The airborne LIDAR workflow will make use of an unprecedentedly extensive dataset of 906 randomly located transects sizing 375 ha each, from the 'improving biomass estimation methods for the Amazon' (EBA) (Gorgens et al. 2020) [13], plus data from the Sustainable Landscapes Brazil (SLB) project (Longo et al. 2016) [37]. The product will consist of a multilayered raster data product with LIDAR measures of EVP traits – ecosystem height, cover, and structural complexity – [1], including estimations of their uncertainties and a demonstration of how airborne LIDAR can be used to improve those over a satellite product. The code developed will be made publicly available for other GEO BON members and organizations, with procedures incorporated as a function in the rGEDI package (Silva et al. 2020) [33], and HPC pipelines enabling national BONs to compute these EVP traits locally, or nationally, using globally consistent protocols that comply with the standards established for the EBV portal.

3. Team members/collaborators

PI: **Rubén Valbuena (Bangor University, UK)** will lead the overall project coordination, and serve as primary link with GEO BON, ensuring that the resulting deliverables meet the specifications fixed in the Ecosystem Structure WG and GEO BON requirements for the EBV data portal. He will supervise a Post-Doctoral Research Associate (PDRA) who will be contracted to work full time in AMAZECO project and will carry out the bulk of the satellite LIDAR component of the project (see Fig. 4) and the high performance computing (HPC) pipelines in Azure. An unspecified candidate will be recruited for this work, selected on the basis of experience with Azure or similar cloud computing services. Dr Valbuena's involvement with GEO BON set off during a secondment at the World Conservation Monitoring Centre (UNEP-WCMC) in the framework of GlobDiversity project [1]. He has experience in parallel computing on remote servers, specifically for processing LIDAR data, having processed 0.5 pulses/sq-m available for the whole of Finland at 16-m resolution [5-6, 9]. He has a long record on the development of publicly available datasets [3, 12] and open-source software tools [11, 33-35]. He has experience in the management of projects with budget size similar to AMAZECO.

Co-PI: **Eric B. Görgens (Universidade Federal dos Vales do Jequitinhonha e Mucuri, Brazil)** will lead the airborne LIDAR and field support component, and coordinate the liaison with local partners (see J.P. Ometto and M. Keller, below) and stakeholders. He will supervise a PDRA who will be contracted to work full time in AMAZECO project and will carry out the airborne LIDAR component of the project (see Fig. 4) and the satellite/airborne combined upscaling. Prof. Görgens worked on the development, establishing project specifications, and flight planning for the surveying of 906 transects for the EBA project (improving biomass estimation methods for the Amazon). For Sustainable Landscapes Brazil (SLB) project, he is the national contact for the ongoing work preparing and performing the quality control of the newest 2020 campaign.

Co-PI: **Carlos A. Silva (University of Florida, US)** will lead the satellite LIDAR component of the project (see Fig. 4) and rGEDI package component. Dr Silva has been working closely with the GEDI team under the NASA Carbon Monitoring System (NASA-CMS). He lead the development of rGEDI [33], the first open-source tool for GEDI data processing and visualization, and also in 2014 elaborated the first R package devoted to processing airborne LIDAR [36], among any others open-source tools [11,34-35]. Prof. Silva is currently involved in a Brazilian National Council for Scientific and Technological Development (CNPq)-funded project for which he is processing GEDI data for the entire of the Brazilian Cerrado biome (about half the size of the Amazon).

Partner collaborator: **Jean P. Ometto (Brazilian Institute of Space Research – INPE)** will support with permission for the use of data from EBA project, and advising contributor to the research components. Prof. Ometto is the Head of the Earth System Science Center (CCST) at INPE. He has been a nominated contributor of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) on Regional and Sub-Regional Scoping Assessment, and the 5th assessment report of the Intergovernmental Panel on Climate Change (IPCC) for WG-II and Task Force on inventories of greenhouse gas emissions, and also the Clean Air Coalition to Reduce Short-Lived Climate Pollutants (CCAC) in Latin America. He is the Brazilian Representative with the Executive Council and the Conference of the Parties for the InterAmerican institute for Global Change Research (IAI), coordinator of the Cooperation Program for the International Geosphere-Biosphere Program Regional Office (IGBPBRO) and The Global Land Project International project Office (GLP-IPO).

Partner collaborator: **Michael Keller (International Institute of Tropical Forestry (IITF), Forest Service, USDA)** will support with permission for the use of data from Sustainable Landscapes Brazil (SLB) project, and advising contributor to the research components. Prof. Keller is a research physical scientist at the IITF, Distinguished Visiting Scientist at NASA's Jet Propulsion Laboratory (JPL), and visiting researcher at the Brazilian enterprise of agricultural research (EMBRAPA). After being appointed as Chief of Science of NEON Inc. (2007-2010), since 2011 he has coordinated the SLB project [37], supported by the USAID and by the US Dept. of State, aiming at the generation of LIDAR data with the aim to contribute to forest degradation analyses in Brazil by generating locally-based knowledge and new methods.

4. Detailed project description

4.1 Background, problem statement & overall project goal

Biodiversity assessment and monitoring need a set of Essential Biodiversity Variables (EBVs) [38], some of them observed using remote sensing [39], to help identify where conservation efforts are needed and evaluate their efficacy [40-41]. Light detection and ranging (LIDAR) sensors provide detailed 3D measurements of ecosystem structure [6, 16, 26, 42]. As ecosystem structure is a key determinant of niche distribution, LIDAR technologies are regarded as essential for monitoring biodiversity assets [5, 42-46]. However, handling and processing LIDAR datasets requires a high level of specialization, and computing capabilities too. In this proposal we suggest to develop tools for digesting LIDAR datasets, including the consistent use of both satellite and airborne platforms as well as other sources of 3D remote sensing data, into a set of simple and concise indicators of ecosystem structure that may be easily understood and employed by non-specialized conservationists and other interested stakeholders [1, 46].

A standardized framework of morphological traits from ecosystem vertical profiles (EVPs)

3D remote sensing sources airborne can play an important role in the development of EBVs in the near future, involving a great variety of sensors (LIDAR, radar, or stereoscopic restitution such as structure from motion) and platforms (e.g. measurements from ground, drones, airborne or satellites) [42]. Derived information products must be based on their importance of ecosystem traits to biodiversity, not on the means to measure them. For this reason, in Valbuena et al. (2020) [1] we advocated for a sensor and platform-independent set of morphological traits that can be derived from any type of 3D information, including LIDAR sources, ground information of any hitherto unknown technology that may later become relevant, acquired at any type of ecosystem.

Satellite LIDAR must prevail, but it has limitations

A satellite LIDAR onboard the International Space Station, the Global Ecosystem Dynamics Investigation (GEDI) mission is currently completing the first comprehensive global collection of LIDAR data [2]. However, the use of currently available satellite LIDAR data faces critical limitations to practically and effectively provide information useful for EBVs: (1) satellite LIDAR is based on the acquisition of discretely sampled pulses, and thus ineffective to inform EBVs requiring spatially continuous data sets, such as habitat fragmentation; and (2) missions are not designed to acquire laser pulses over any same location twice; and (3) no space agency has plans to launch a platform with a LIDAR sensor amongst the missions planned to 2030, which precludes the delivery of information on progress toward the Convention of Biological Diversity 2030 goals and the UN post-2020 agenda. For these reasons, we advocate for a system flexible enough to incorporate information for disparate data sources [1], arguing that it would permit the incorporation of airborne LIDAR data to EBV products, aiding the overall goals of GEO BON. There are still some limitations on questions on how that can be achieved in practice. Possibly, methods for the measuring ecosystem vertical profiles (EVPs) and derived ecosystem morphological traits from GEDI should be the flagship for others to follow, but we still need the means to incorporate measurements from other data sources, especially airborne LIDAR.

Opportunities for using publicly-available airborne LIDAR data covering large areas

Several national surveying programs are producing airborne LIDAR datasets covering entire countries, many of them with already revisited coverages and future expansion plans with ever increasing densities and frequencies. These datasets are demonstrably useful for ecosystem characterization and ecological applications [5-6, 9, 46-48]. There is general consensus on the methods to derive ecosystem morphological traits from these datasets [1, 47, 50-52], and they are increasingly becoming publicly available along with free tools for data processing. However, little attentions has been put on securing harmonization of workflows for retrieval of ecosystem morphological traits from airborne and satellite LIDAR. These processing workflows should include the determination of uncertainties in the mensuration of morphological traits from the given platform/sensor combination, which should be incorporated in the resulting product [48].

A roadmap to combine satellite and airborne LIDAR

We have a vision for a combined multiplatform EBV product, proposing a clear roadmap for incorporating airborne LIDAR by using GEDI as a standard [1]: the derivation of those same morphological traits from airborne LIDAR following Hancock et al.'s [53] processing steps to derive Ecosystem vertical profiles (EVPs) valid for both satellite and airborne data (Figure 1).

Ecosystem morphological traits can then be derived from those common EVPs, facilitating the homogenization of airborne and satellite acquisition settings, reaching a platform-independent method, which is a key characteristic of any EBV [40, 47]. Members of the proposing team have recently developed the rGEDI package [33], which enables and facilitates the use of GEDI data to larger audience who may be unaware of the specifies of GEDI's formats. Hancock et al.'s [53] algorithms are already incorporated in rGEDI (see *gediFWSimulator* function in Fig. 4), effectively allowing the derivation of morphological traits common from both satellite and airborne LIDAR. AMAZECO will focus in proving that pathway reliable, and in creating the software tools that can empower regional BONs and other local/national stakeholders (using BON in a Box) to generate the same morphological traits locally in a globally-consistent manner. Thus, we have a vision for crowdsourcing the processing of several LIDAR datasets into a same ecosystem structure EBV product, allowing to scale up methods spatially, but also in the temporal span, well beyond the reach of the developers themselves and the extent of this specific project.

Project goals and specific objectives

The overarching goal and long-term vision of this project is to enable in practice the development of a global ecosystem structure EBV product that can be derived and regularly updated from disparate sources of ecosystem 3D structure [1]. We will develop pipelines for high performance computing (HPC) cloud processing which will be made available to the GEO BON community, so that the spatio-temporal domain of such ecosystem structure EBV dataset can be enhanced into global and decadal scales. As part of the project itself, we will use these pipelines to develop a first prototype EBV product covering the entire Amazon basin, based on satellite GEDI data [48] and extensive airborne surveys from current projects spearheading the use of LIDAR in the Amazon [13, 37], which will serve as precedent and tutorial example in BON in a Box. Besides of the HPC pipelines for cloud computing, we will also enable more local processing of ecosystem morphological traits from satellite and airborne LIDAR consistent with the global EBV product, by incorporating the workflows and specifications in an R programming function specifically devoted to ecosystem structure EBVs, and include it in the rGEDI package [33]. The project will also include research components, resolving questions on the specific processing parameters needed to achieve consistency across the satellite and airborne platforms, or means for accounting for spatial resolution-dependent uncertainties in the measurement of each given ecosystem trait.

Specific Objectives, each attached to a given Working Package (WP) are:

- ***Objective 1*** (**O1**)***:*** To set up the Microsoft Azure High Performance Computing (HPC) environment, including containerization of input data and processing software, to enable the common and consistent processing of satellite and airborne LIDAR data.
- ***Objective 2*** (**O2**)***:*** To ensure consistency in the morphological traits derived from both satellite and airborne LIDAR datasets, including the means for evaluating the uncertainty in their measurement at the pixel level.
- ***Objective 3*** (**O3**)***:*** To generate products – i.e. processing workflows (**D2** and **D3**) and maps of ecosystem morphological traits covering the Amazon (**D1**) –, that can set precedent on how information from LIDAR data will be incorporated into the EBV data portal.
- ***Objective 4*** (**O4**)***:*** To secure the uptake of the prototyped methodology by other GEO BON members and organizations, enabling them to employ the derived products for the Amazon and/or empower them to generate similar products at their own or areas of interest.

With objectives **O1**, **O3** and **O4** focused in the generation and delivery of a product of direct application by GEO BON members and organizations, Objective **O2** also incorporates a research component which will answer key questions necessary to envisage the practical delivery of a platform-independent global LIDAR-based ecosystem structure EBV product:

- ***Research question 1*** (**Q1**)***:*** Which LIDAR proxies of ecosystem morphological traits – ecosystem height, cover, and structural complexity – are more robust and consistently produced across LIDAR platforms.
- ***Research question 2*** (**Q2**)***:*** How can the scale dependence of ecosystem morphological traits reveal the relevant resolution at which the LIDAR measurements should be processed?
- ***Research question 3*** (**Q3**)***:*** What is the most sensible means to quantify the uncertainty in the LIDAR measurements and incorporate it in EBV products?

4.2 Data products and derived indicators

Tangible outcomes of this project will be: (1) an R function (*EcoStructEBV*; or a set of many) incorporated rGEDI package which would derive ecosystem traits compliant with the global ecosystem structure EBV; (2) HPC pipelines channelling those same processing workflows for parallel cloud computing of massive LIDAR datasets; and (3) and implementation of these pipelines in a prototype ecosystem structure EBV product covering the Amazon basin, using satellite data from GEDI [48] and airborne data from EBA [13] and SLB [37] projects. The research component will also result in collaborative scientific publications, but these are secondary products, only needed to secure the quality of the main project deliverables:

Deliverable 1* (D1)*: A multilayer raster file (Fig. 2) extending for the entire Amazon containing statistics for values of ecosystem morphological traits – ecosystem height, cover, and structural complexity (Fig. 1) –, derived from EVPs from satellite and airborne LIDAR, and including pixel level uncertainty in the measurement of each trait [1]. The specifications of this multiraster layer will be assured to be in line with those required from the EBV data portal, following GEO BON principles of interoperability of EBV product [55] and EBV metadata standards [56]. This product will be included in the EBV portal, as the first of many that eventually will populate the global ecosystem structure EBV with indicators derived from 3D structural information of ecosystem, from sources such as LIDAR data, in other areas of the world and at other moments in time.

Deliverable 2* (D2)*: R function *EcoStructEBV* incorporated in the rGEDI package [33] deriving ecosystem morphological traits locally in a globally-consistent manner. rGEDI package has already incorporated capabilities for processing satellite LIDAR waveforms, and also for deriving similar waveforms from airborne LIDAR datasets. *EcoStructEBV* function will streamline these capabilities toward the derivation of products suitable for crowdsourcing the global ecosystem structure EBV product, effectively linking the satellite and airborne processing workflows.

Deliverable 3* (D3)*: Processing pipelines for large-scale application of *EcoStructEBV* functionaly in cloud HPC, including tutorials for their use by non-specialized stakeholders and national BONs. These will enable the processing of large amounts of LIDAR data, possibly the entire GEDI mission at some point. Thus, they will not be based just in R, but Jupyter notebooks or C++ (e.g., some rGEDI routines already use rcpp). Algorithms and tutorials for implementing them will be available in BON in a Box. ArcGIS Pro notebooks will also be employed to facilitate visualization and easy use but non HPC-trained stakeholders.

4.3 Technical approach

AMAZECO's approach will be to progressively scale up the dimensions of the work (Fig. 3), to answer the research questions and objectives, reaching milestones (**M1**-**M4.2**; detailed in sections 5 and 6) prior to tackling the next ones. Section 4.5 makes an assessment of the risks (**R1**-**R5**) at each of them, drawing contingency plans which ensure that the final products will be delivered regardless of the results obtained at each step. The work is structured into work packages (WP), corresponding to each of the specific objectives:

WP1. *Data gathering and normalization & HPC environment set up* (see **O1**).

Task 1.1 GEDI data gathering (Silva)**.** Data for the whole Amazon basin will be downloaded from The Land Processes Distributed Active Archive Center (LP DAAC) and stored in MS Azure.

Task 1.2 ALS data gathering and normalization (Görgens). Airborne LIDAR data from the Amazon will be collected and pre-processed into a common standard input dataset: classification of ground returns, digital terrain model generation, calculation of heights above ground, etc.

Task 1.3 Software requirements (Valbuena). Executables and modules need will be identified and set up in the system, ready for processing workflows (see contingency plan to the risk of licensed software to be deemed essential for the initial processing **R2** in Table 1). They will be containerized into kubernetes, ready to de deployed on the Azure cloud services.

In the first month, the whole team of three Co-PIs and two PDRAs will meet and work together to kick-off the project (risk **R1**), hosted at Bangor University for 2-4 weeks (as long as available budget allows). All datasets will be gathered and put in common, setting up GitHub repositories, a project group in Slack for direct interaction, and Azure storage for sharing common documents, making detailed plans and getting everyone ready to proceed with the work in distance (see Milestone **M1** in Sections 5 and 6).

***WP2**. Research component: effective satellite/airborne LIDAR combination* (see **O2**).

Task 2.1 GEDI/ALS geographical coincidence (Valbuena). A subset of the GEDI data will be extracted, including only those satellite LIDAR pulses fired over the airborne LIDAR transects. A join satellite/airborne dataset will be compiled for the research component (Fig. 3).

Task 2.2 Derivation of EVPs from research component subset (Görgens/ Silva). Workflows in Fig. 4 (GEDI & airborne data processing) will be followed to extract ecosystem vertical profiles.

Task 2.3 Research on the convergence of morphological traits products derived from satellite and airborne (Valbuena). Workflows in Fig. 4 (satellite/airborne common processing) will be followed to extract ecosystem morphological traits from the EVPs resulting from Task 2.2.

In Task 2.1, the GEDI pulses coinciding with airborne transects will be extracted (Fig. 3). Then the airborne data at the location of the GEDI pulses will be clipped (Fig. 4), taking into account the uncertainty in the exact positioning of the satellite laser pulse [49]. Then, in Task 2.2., the workflows will differ for the GEDI and airborne LIDAR data, both aimed at obtaining of EVPs from each data source (Fig. 4). The satellite data will follow common rGEDI processing [33], whereas for the airborne LIDAR a GEDI waveform will be simulated [54], substantially deviating from common airborne LIDAR procedures [2, 7, 13, 24, 46-48]. EVPs obtained at same locations will be compared, adjusting the parameters in the simulation of satellite pulses from the airborne data until reaching convergence between the satellite and airborne products (see risk **R3**). In Task 2.3. the satellite and airborne workflows will converge (Fig. 4). Measurements of ecosystem morphological traits will be derived from the resulting EVPs [1] (Fig. 1). Coincidences and discrepancies between the satellite and airborne-derived morphological traits will be investigated. Tasks 2.2. and 2.3 will iterate (see Gantt chart in Fig. 5), determining the most satisfactory combination of processing parameters that secure convergence between the satellite and the airborne workflows (see milestone **M2** in section 5, and contingency plan to the risk of not reaching convergence **R4** in Table 1).

***WP3**.Derivation of gridded products* (see **O3**).

Task 3.1 Small area pilot study (see **Q2** and **Q3**) (Görgens/ Silva**).** The processing workflows will be first developed over a pilot area, to determine the optimal spatial resolution that the final product should have, and the best means to determine the uncertainty. The R function and HPC pipelines for cloud computing will be developed.

Task 3.2 Geographical upscaling (Valbuena). HPC pipelines will be used to implement results from Task 3.1 for the entire Amazon in two steps: first for GEDI only, plus airborne afterwards.

Prior to tackling the entire Amazon, a smaller pilot area will be used to determine optimal processing parameters and test the developed tools (Fig. 3). Once all the data is downloaded, the study area for the pilot (Task 3.1) will be selected according to the most convenient combined availability of GEDI swaths and EBA transects. An approx. area of 10,000,000 ha could contain around 20 transects and intersect circa 300 GEDI swaths. Several gridded products will be generated for the pilot area, using increasingly finer resolutions (0.05, 0.01, 0.005, 0.001, 0.0005, and 0.0002 decimal degrees, which approx. correspond to pixel sizes of around 5,000, 1,000, 500, 100, 50, and 22 m respectively). These gridded products will be suitable proxies of ecosystem morphological traits, and also measures of their uncertainty at pixel level. These proxies are statistics taken from characteristics of the measured LIDAR waveform of pulses obtained within each cell of the grid. Figure 2 shows an example of suitable proxies, since the mean of 100% relative height (rh100) can be used to measure ecosystem height and the standard error of that mean can be used to quantify the uncertainty of that value. The products will be evaluated according to the scale dependence observed in the results, affecting the robustness of the derived morphological traits, their measuring uncertainties (e.g. standard error of mean values at each pixel, see Fig. 2, depending on results in Task 2.3 too), and the requirements of the EBV data portal (Fig. 5). These should enable decisions on the optimal resolution of the final product and the means to account for pixel-level uncertainties within the product itself. Upon ending Task 3.1, at month 10, R functions for small area processing **D2** and workflows for large area processing **D3** will be ready (milestone **M3.1**, see Fig. 5). The larger area processing can be initiated by the PDRA at the end of year 1, but the bulk of it will happen during the following years 2 and 3.

Then, Task 3.2 will concentrate in employing the HPC pipelines to upscale the processing to the entire Amazon **D3** (milestone **M3.2**, see Fig. 5). On a first step, gridded products of ecosystem morphological traits and their uncertainties will be derived from GEDI data only. After that, information from the airborne LIDAR data will be incorporated into that same grid. This two-step approach will be followed to highlight the effects of incorporating airborne LIDAR on top of the satellite grids, allowing to either (i) fill in data gaps in areas where there was no GEDI data available, or (ii) reduce the uncertainty in the measured morphological traits in areas where there was availability of both satellite and airborne data. This will allow us to quantify the gain from incorporating airborne LIDAR information, and also set a precedent on our vision for a satellite-based global product that can become enhanced in areas where higher resolution airborne LIDAR is available (e.g. national airbone LIDAR programmes [1]). EBA transects can enhance the spatial resolution of the product [13], whereas the involvement of SLB data can improve the temporal resolution, given a temporal dimension to the product [25, 37]. New GEDI pulses may be further incorporated as the project goes on over the 3 years while the sensor is operational, which will also set precedent on how data from other satellite LIDAR sensors (e.g. ICESat-2) may be incorporated as well. Thus, our vision for an ecosystem structure EBV is a dynamic one, so it is not only global but can also be updated in time, and the planned workflow seeks to put in relieve how this may work in practice.

WP4. *Outreach* (see **O4**).

Task 4.1 Launching the Amazon product in the EBV data portal (Valbuena)**.** The specifications and data characteristics required by the EBV data portal will shape the development of the final rasterfiles from their inception (see Gantt chart in Fig. 5). The final dataset will be made available in the EBV data portal, and the means for continuous incorporation of datasets from other spatio-temporal domains will be envisaged in liaison with GEO BON secretariat (milestone **M4.1**).

Task 4.2 Incorporating workflows in BON in a Box (Valbuena)**.** Tutorials aiming at non-specialized stakeholders will be generated, guiding on the use of the generated workflows: rGEDI functions and HPC pipelines. These will be part of the deliverables **D2** and **D3** (milestone **M4.1**). These will enable national BONs to process the GEDI and/or airborne LIDAR available in their own areas, and establish the relationships to biodiversity that may be most relevant in their cases.

Microsoft Azure Services

AMAZECO's objectives, especially the Amazon-wide product **D1**, could only very inefficiently be achieved without HPC capabilities. We have estimated that the work for the whole Amazon will involve around 550 Million GEDI pulses, which will occupy a memory of around 275 GB, eventually involving the download of whole-orbit data sizing 65 TB (these not necessarily involved in the processing all at once). We made this estimation as a rough extrapolation from Dr Silva's experience in a CNPq-funded project (see *CV*) for which he has already downloaded six months of GEDI data covering the entire Cerrado biome in Brazil (2,045k sq-km, compared with 5,500k sq-km in the Amazon).Under Dr Valbuena's experience in processing airborne LiDAR for the whole of Finland [6, 9], these are lesser amounts of data and thus more manageable (despite the larger extent of the area covered, being much lower density and resolution). This Amazon project will give us good experience to realize what will be needed for a global GEDI-based ecosystem structure EBV (see impact in section 4.4). In the case of the airborne data, the original .las files occupy a memory of around 2.5 Gb per project, and their derived processed products generate storage requirements reaching around 2Tb in total for both projects. Thus AMAZECO will leverage Microsoft Azure services and infrastructures, making good use of the following:

- **Cloud storage**: Azure will be used for cloud storage in both preparing all the data in the setting up stages (Tasks 1.1 and 1.2), and also in the computation of the final product and preparation for launching in the EBV data portal (Tasks 3.2 and 4.1), also including **Bandwidth** enough for the transference to the cloud of the amount of data estimated.
- **Virtual machines**: We have chosen a system of 2 Windows OS-based virtual machines (in case satellite and airborne require simultaneous processing) with 8 cores for parallel processing. Given the 275 GB estimated above, we have selected them with 400GB of temporary storage, which would be needed during the processing.

- **Azure Kurbernetes Service (AKS)** and **Azure batch**: To containerize processes to input data and executables, manage clusters, and scale and queue the cloud processing.
- **Azure DevOps**: For shared code development along with GitHub.
- **Azure functions**: To assist the deploying and monitoring in the cloud after developing the code and debugging locally and on GitHub, including services for easy deployment of locally processing to cloud processing, such as **Azure Advisor** or **Azure Migrate**.
- **Azure Maps**: For visualization purposes in outreach activities (also using ArcMap Pro).

4.4 Impact of the expected outcomes

The most important impact of AMAZECO will be in taking the **first steps toward realizing an Ecosystem Vertical Profile EBV, globally assessed** with the aid of several remote sensing means [1], which will be a core part of habitat structure within the Ecosystem Structure EBV class, as it is currently being discussed in the Ecosystem Structure working group of GEO BON. The cloud processing using Azure will attain the objective of becoming globally scalable, which eventually can done by completing the deliverable **D3**. As an interesting note, it is estimated that there will be around 10 billion cloud-free observations for the entire earth surface at the end of GEDI mission. Starting this project now in 2020, and having the Amazon experience ready by 2023, will put GEO BON in the position of reaching a GEDI-based EVPs as a habitat structure EBV, plus the means for incorporating airborne LIDAR where/whenever available, as **one of the ten regularly-updated operational products stated among the key GEO BON goals by 2025**.

While it is widely acknowledged that LIDAR data provides great information on ecosystem structure, essential to monitoring habitats and biodiversity, data processing requires specialized capabilities beyond the interest of ecologists and conservation stakeholders. Our approach is to effectively **digest the complexity of LIDAR datasets into conceptually simple ecosystem morphological traits** that any environmentalist can understand: ecosystem height, vegetation and structural complexity [1]. Making the deliverables available through the EBV data portal and BON in a Box will **foster the uptake of LIDAR-derived information by GEO BON members** interested in the relationships between ecosystem structure and biodiversity assets. Besides of participating in GEO BON events, we will participate in Microsoft events and enhance the outreach of the product with presentations among our ecology and remote sensing networks.

The project will **leverage Azure capabilities by creating workflows for processing very large amount of LIDAR data**, including GEDI pulses plus the unprecedented airborne LIDAR dataset from the EBA project, which consistently surveyed random transects across the Amazon [1] and SLB providing the temporal component. A research component will focus on realizing Valbuena et al.'s (2020) [1] vision on converging the satellite/airborne workflows into delivering a same consistent set of morphological traits. **The roadmap proposed in Fig. 1 has ground-breaking nature with a research interest**, putting the focus on converging the airborne and satellite LIDAR in the development of EVPs, as opposed to adapting traditional airborne LIDAR processing steps to the EBV purposes. Thus, following this alternative pathway entails a hypothesis that such workflow would secure a more mechanistically-driven derivation of morphological traits (evaluating measured EVPs, as opposed to following the traditional workflows for satellite and airborne separately, and determine with exploratory analysis which outputs become more similar in a given sample). Furthermore, such roadmap in applying Hancock et al.'s (2019) [54] simulated waveforms in the derivation of traits relevant to EBVs would enhance the application of the approach, which was originally only developed for the calibration/validation of biomass models [23, 49]. The research proposed will clarify the specific workflow parameters (e.g. resolution, most robust waveform descriptors, etc.) that would ensure consistent derivation of ecosystem morphological traits and the uncertainties in measuring them.

This Amazon product will also attain the objectives of **focusing on a geographic area of high current relevance for biodiversity monitoring** and support for conservation policies, and where ecosystem and biodiversity observations are **currently underrepresented**. Given the current environmental and political situation in Brazil, the development of this product would aid conservation in the area on a very critical moment, giving a clear baseline to monitor the ecosystems in the Amazon. AMAZECO could be the **embryo of a Brazil BON**, as it can attract the attention of key players in the area toward GEO BON.

4.5 Anticipated key obstacles and threats

Potential risks have been assessed as marked (**R1**-**R5**) thoughout the proposal. Table 1 provides a contingency plan for eqach of them, which secures that deliverable will be produced even under the worst of circumstances, both in terms of research results and practical capacities.

Table 1. Risks assessment & Contingency plan

Risk	Level	Solution / Alternative
R1. Coronavirus pandemic impeding overseas travelling	**Medium.** Evolution of pandemic is difficult to evaluate, but overseas travelling possibly resumed in September 2020.	Work will proceed in distance and the travelling budgets will be used for outreach activities at the end of year 1 instead.
R2. Software licenses required for initial steps of airborne LIDAR processing.	**Low.** Open-source tools for processing airborne LIDAR abound. Only very initial processing steps may be bound to licensed software.	In month 1 the WP2 can start regardless of success in WP1, since HPC is not critical for working with the transects only. By month 5 when WP3 starts (Fig. 5), it can be decided whether higher-level LIDAR products would be specified as the input data for the deliverable workflows (**D2** and **D3**): e.g. data already classified into ground/vegetation or with heights above ground calculated.
R3. No good parameter combination attaining reliably similar EVPs from satellite and airborne (systematic biases)	**Very Low.** Hancock et al.'s [53] method has been tested and is being use in current GEDI global biomass estimations [23]	(a) Besides Hancock et al.'s [53], there are other algorithms that can be tested [57]. (b) Separate traditional pathways for satellite and airborne (blue and red colour in Fig. 1) can be followed, and parameters affecting the derivation of morphological traits from them established solely on the basis of convergence (correlation) of the final products (inductive approach, as opposed to the mechanistic approach proposed).
R4. Large differences between satellite and airborne in measuring traits. (large non-systematic errors)	**Medium.** Some level of discrepancy is always likely. The key is on determining the magnitude of the discrepancy, and accounting for it.	If large discrepancies are found in the research component, these will be quantified and incorporated in the uncertainty measure of the gridded products (**D1**). Airborne can be considered as reference to satellite, and field data may be available from SLB, if needed.
R5. Azure credits insufficient to complete the entire Amazon (or eventually the whole world beyond the project)	**Unknown.** Difficult to evaluate beforehand. This is a unprecedent experience on how many Azure credits are spent through the necessary processing steps.	Project PI will liaise closely with GEO BON secretariat under such circumstance. This is a first prototype project that can give a first idea of the amount of processing needed. Tools (e.g. the R package) are being prepared with no direct dependence on Azure itself, which secures broad uptake by GEO BON members independently from Azure credits availability.

5. Monitoring and Evaluation

The **implementation of the project** will be monitored through the set of established milestones to be revised through the project timeline (Fig. 5):

***Milestone 1* (M1)**. Month 2. Repositories and working procedures set up.

***Milestone 2* (M2)**. Month 5. Optimal processing parameters to derive a common set of morphological traits valid for both the satellite and airborne workflows.

***Milestone 3.1* (M3.1)**. Month 10. Deliverables **D2** and **D3** ready: R functions and HPC pipelines.

***Milestone 3.2* (M3.2)**. Year3/Quarter 2. Deliverable **D1** ready: an Amazon-wide multilayer raster.

***Milestone 4.1* (M4.1)**. Year3/Quarter 4. Deliverable **D1** ready in EBV data portal

***Milestone 4.2* (M4.2)**. Year3/Quarter 4. Deliverable **D2** published as an rGEDI update in CRAN [33] and HPC pipelines **D3** ready with tutorials ready in BON in a Box.

Grantees will deliver a mid-year report, and a final report one year from the starting date of the project. They will continue to have access to their sponsored Azure accounts for two additional years to scale their datasets, data products, and tools more broadly.

The **overall impact of the project** will be monitored through the usage of deliverables:
Deliverable 1* (D1)*: It will be provided with a doi in Dryad and/or Zenodo, for download count.
Deliverable 2* (D2)*: rGEDI package download since the deployment of the updated version will
Deliverable 3* (D3)*: Processing pipelines will be provided with a doi in Dryad and/or Zenodo, for download count.
The impact will also be evaluated by the production of scientific publications (mainly from the results of Task 2.3, possibly also 3.1), in which the join support of Microsoft and GEO BON will be acknowledged, and citations and presence in the media stemming from these publications will be monitored.

6. Detailed project timeline

The project timeline is planned to secure that the work of the PDRA concludes at the end of year 1 with the R package functions and HPC pipelines ready along with tutorials making it straightforward to use by any GEO BON member. Then, the accessibility of Azure credits for the remaining 2 years will be employed for the processing of the product covering the Amazon, and its deployment in the EBV portal.

7. Project budget

Salary Costs:

● 1 PDRA at Bangor University for 12 months at 100% (£47,021).	$69,262
● 1 PDRA at UFVJM for 12 months (R$87,600).	$17,888
● Dr Valbuena (Lecturer) at Bangor University. 12 months at 10% fte	$0
● Prof. Görgens (Professor) at UFVJM for 12 months at 8.33% fte	$0
● Dr Silva (Assistant Professor) at University of Florida. 12 months at 8.33% fte	$0
Travelling Costs (remaining of total budget at equal shares) for kick-off in Bangor UK:	
● UFVJM’s PDRA travelling from Brazil to Bangor University in UK.	$4,284
● UFVJM’s Co-I travelling from Brazil to Bangor University in UK.	$4,284
● University of Florida’s Co-I travelling from USA to Bangor University in UK.	$4,284
Total Requested grant	**$100,000**
Azure Costs: see attached Azure calculator cost estimates):	
Total Requested Azure credits	**$99,971**

Given the budget limitations, project PI and Co-PIs decided to request no budget for their time commitments to the project, given the interest of contracting PDRAs to carry out the main bulk of the work and that project objectives are well in line with the strategic research lines of our institutions. Deliverables will remain in the EBV data portal and BON in a Box beyond the extension of the project itself, making the system EBV+pipelines available over time for continuous update. The positions of PI and all Co-PIs at their institutions are permanent, and thus the continued support to the project beyond the grant period is secured. We aim at applying for additional funding making use of the Amazon layers to link ecosystem structure with biodiversity assets and bio-climatic variables, following up our most recent research collaborations [13].

References (use the provided links to access full references)

n.b.! The numbering of references begins at the investigators’ CVs [1-36], and the remaining [37-61] are listed here: **[37]** Longo et al. (2016) *Global Biogeochem.Cycles* 30:1639; **[38]** Pereira et al. (2013) *Science* 339:227; **[39]** Skidmore et al. (2015) *Nature* 523:403; **[40]** Navarro et al. (2017) *Curr.Opin.Environ.Sust.* 29:158; **[41]** Vihervaara et al. (2017) *Global Ecol.Conserv.* 10:43; **[42]** Kissling et al. (2017) *Biol.Rev.* 93:600; **[43]** Hinsley et al. (2008) *Lands.Ecol.* 23:615; **[44]** Davies & Asner (2014) *Trends Ecol.Evol.* 29:681; **[45]** Simonson et al. (2014) *Methods Ecol.Evol.* 5:579; **[46]** Bakx et al. (2019) *Divers.Distrib.* 25:1045; **[47]** Coops et al. (2016) *Ecol.Ind.* 67:346; **[48]** Kissling et al. (2018) *Nat.Ecol.Evol.* 2:1531; **[49]** Dubayah et al. (2020) *Sci.Remote Sens.* 1:100002; **[50]** Guo et al. (2017) *Ecol.Infor.* 38:50; **[51]** Asner et al. (2017) *Science* 355:385; **[52]** Calders et al. (2019) *Trends Ecol.Evol.* 35:6; **[53]** Fahey et al. (2019) *Ecol.Lett.*22:2049;**[54]** Hancock et al. (2014) *Earth Space Sci.6:294*; **[55]** Hardisty et al. (2019) *Ecol.Infor.* 49:22; **[56]** Langer et al. (2019) *EBV-EML-Schema*; **[57]** Popescu et al. (2011) *Remote Sens.Environ.* 115:2786.

8. Optional supporting documents

Examples of additional files you can attach to strengthen your proposal:

- *Technical supporting information*
  - *Charts and Infographics & Technical diagrams*
    - **Figure 1.** The concept: proposed workflow for combining satellite and airborne LIDAR.
    - **Figure 2.** The product: example of gridded product derived from actual GEDI data.
    - **Figure 3.** WP structure: upscaling strategy for the planned work.
    - **Figure 4**. Technical diagram of suggested workflows of rGEDI functions.
    - **Figure 5**. Gantt Chart of project timeline.
- *Organizational supporting information*
  - *CV of PIs and co-PIs (see template below)*
    - **Team member 1.** Rubén Valbuena.
    - **Team member 2.** Eric B. Görgens.
    - **Team member 3.** Carlos A. Silva.
  - *Letters of support*
    - **Partner 1.** Jean P. Ometto.
    - **Partner 2.** Michael Keller.
- *Budgets*
  - Bangor University approved budget for AMAZECO
  - Azure calculator estimated budget

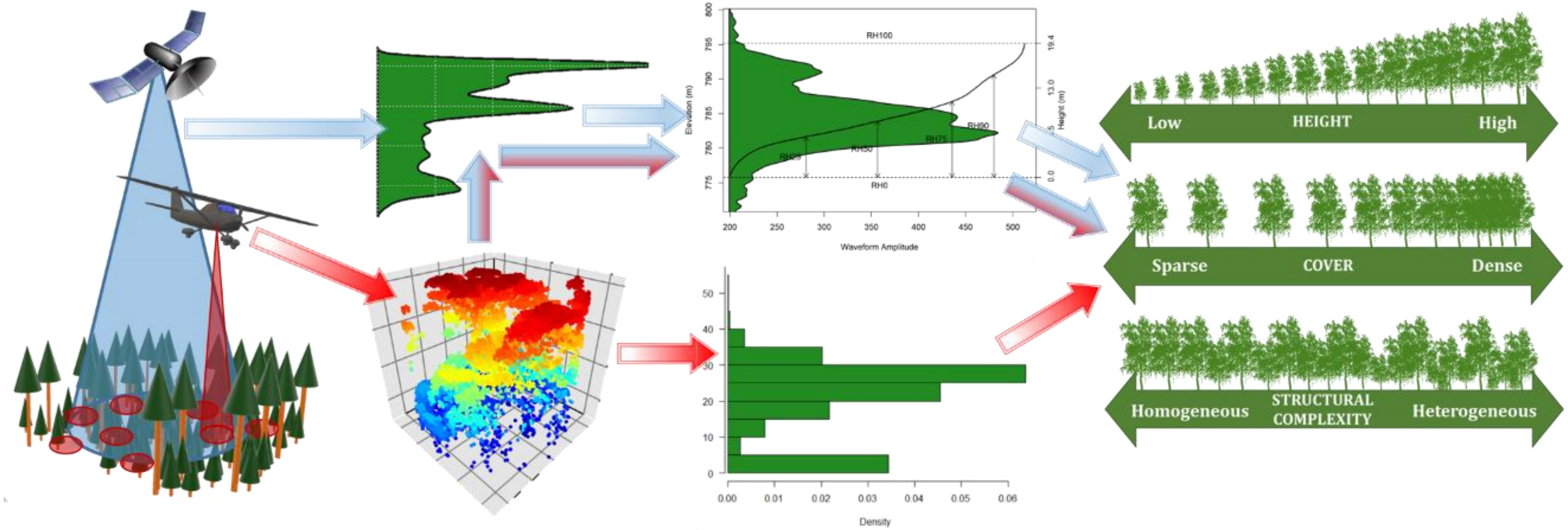


Figure 1. Adaptation from Valbuena et al. (2020) [1] showing approaches to derive EBV products from traditional airborne LIDAR (red arrows) or satellite (blue arrows) and toward standardized morphological traits of ecosystems. The conceptual roadmap proposed in this research is to merge the airborne and satellite workflows (mixed red/blue arrows) in the derivation of EVPs.

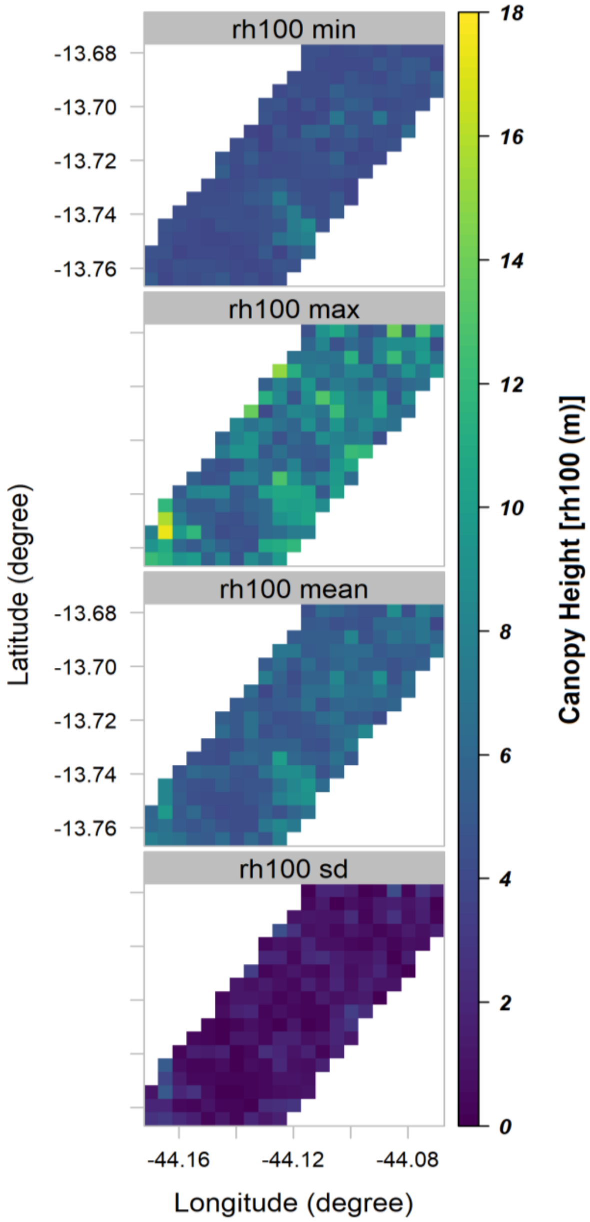


Figure 2. Example of gridded products derived from rGEDI.

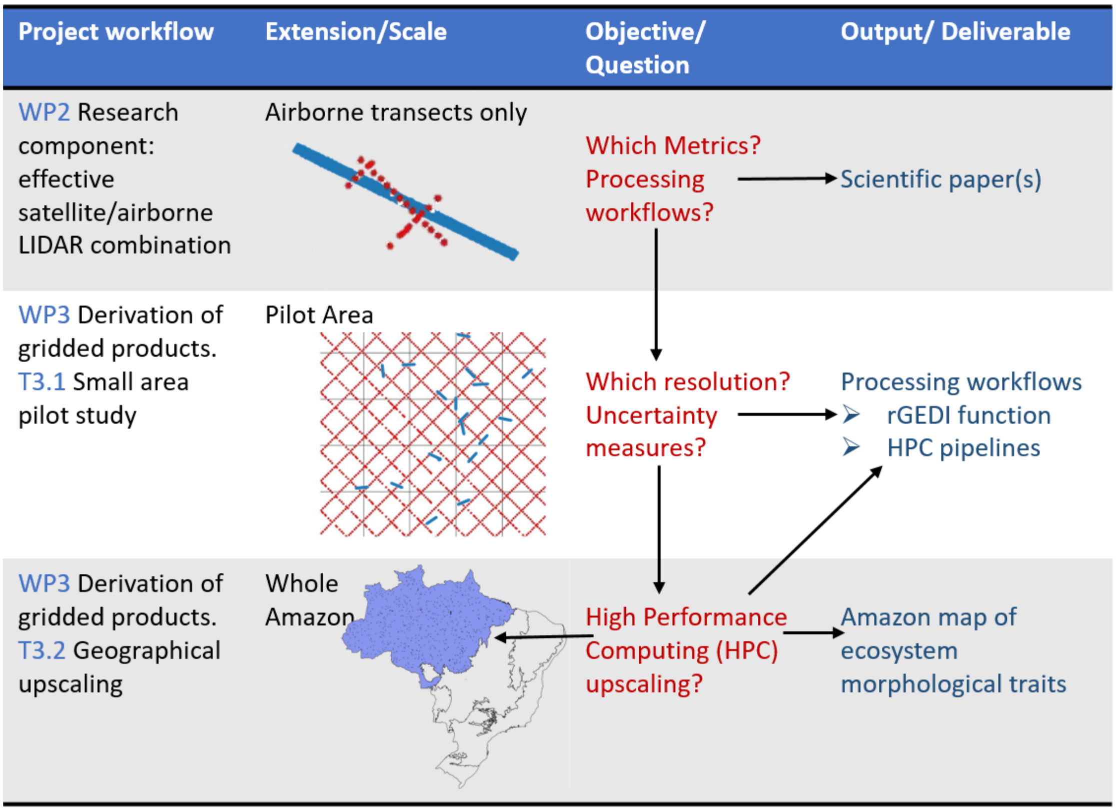


Figure 3. Upscaling strategy for the planned work. Different questions are answered at smaller scales before the work moves on into upscaling, eventually for the whole Amazon.

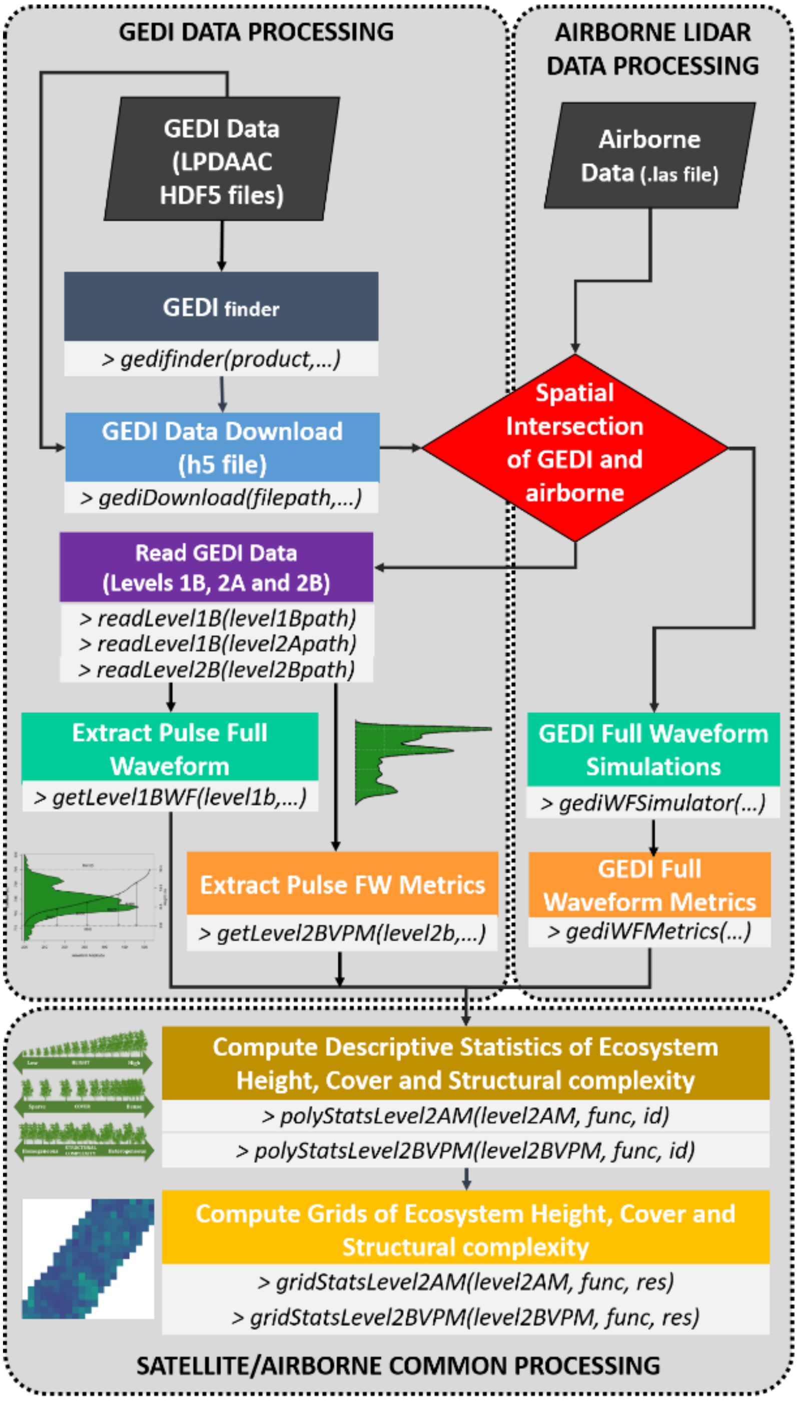
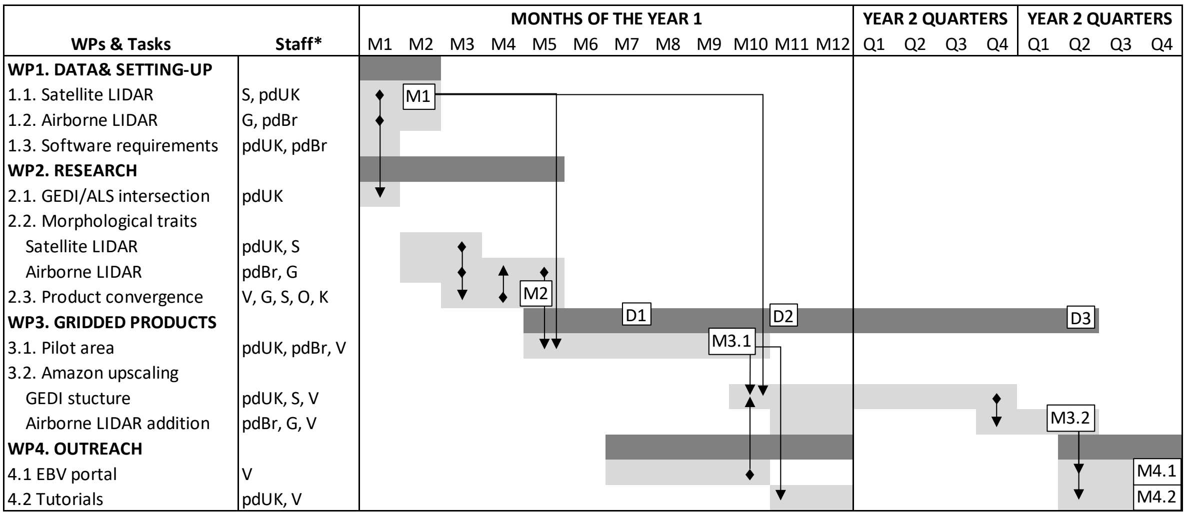


Figure 4. Planned workflow for merging satellite and airborne processing into gridded products for ecosystem height, cover and structural complexity.



* Staff:
V. Valbuena
G. Görgens
S. Silva
pdUK. PDRA UK
pdBr. PDRA Brazil
O. Ometto
K. Keller

Figure 5. Gantt chart of project timeline.

Curriculum Vitae **Valbuena, Rubén**

1. General information:

Valbuena, Rubén, DSc, PhD. DoB: 25/06/1979, male
Bangor University, School of Natural Sciences, +44 12 4838 2445, r.valbuena@bangor.ac.uk
Lecturer in Forest Sciences
Web: GoogleScholar, ResearchGate, Publons

2. Academic qualifications:

Doctorate	DSc (Agriculture & Forest Science), University of Eastern Finland, 2015, Supervisors: Profs. Matti Maltamo, Petteri Packalen & Gert-Jan Nabuurs
Doctorate	PhD (Remote Sensing), Technological University of Madrid, 2013, Supervisors: Profs. José Antonio Manzanera & Susana Martín

3. Professional career:

2018-present	Lecturer in Forest Science, Bangor University, School of Natural Sciences
2017-present	Adjunct Professor in Biodiversity Indicators & Remote Sensing, University of Eastern Finland (UEF), School of Forest Sciences
2016-2018	Marie S. Curie Research Fellow, University of Cambridge, Department of Plant Sciences
2017-2018	Researcher (secondment), UN Environment – World Conservation Monitoring Centre (UNEP-WCMC)
2014-2016	Post-Doctoral Researcher, UEF, School of Forest Sciences
2010-2014	Researcher, European Forest Institute (EFI)
2006-2010	PhD student / Research Assistant, Technological University of Madrid
2008-2009	Guest Researcher, UK Forestry Commission. Forest Research Agency

4. Publications: (See also collaborative publications in other partner's CVs.)

[1] **Valbuena R.**, O'Connor B., Zellweger F., Simonson W., Vihervaara P., Maltamo M., **Silva C.A.**, Almeida D.R.A., Danks F., Morsdorf F., Chirici G., Coomes D.A. & Coops N.C. (2020) Standardising Ecosystem Morphological Traits from 3D Information Sources. ***Trends in Ecology and Evolution*** [in press].

[2] Adhikari H., **Valbuena R.**, Heiskanen J. & Pelikka P. (2020) Mapping Forest Structural Heterogeneity of Tropical Montane Forest Remnants from Airborne Laser Scanning and Landsat Time Series. ***Ecological Indicators*** 108: 105739.

[3] Schepaschenko D., Chave J., Phillips O.L., Lewis S.L., Davies S.J., Réjou-Méchain M et al. (148 authors in alphabetical order; see 'Author Contributions' for details) (2019) The Forest Observation System, Building a Global Reference Data Set for Remote Sensing of Forest Biomass. ***Scientific Data (Nature)*** 6: 198. See **Forest Observation System Repository**

[4] Adnan S., Maltamo M., Coomes D.A., García-Abril A., Malhi Y., Manzanera J.A., Butt N., Morecroft M. & **Valbuena R.** (2019) A Simple Approach to Forest Structure Classification Using Airborne Laser Scanning that Can Be Adopted across Bioregions. ***Forest Ecology and Management*** 433: 111-121.

[5] Mononen L., Vihervaara P, Packalen P, Virkkala R, Auvinen A, **Valbuena R.**, Bohlin I. & Valkama J. (2018) Usability of Citizen Science Observations together with Airborne Laser Scanning Data in Determining Habitat Preferences of Forest Birds. ***Forest Ecology and Management*** 430: 498-508

[6] **Valbuena R.**, Maltamo M., Mehtätalo L. & Packalen P. (2017) Key Structural Features of Boreal Forests May Be Detected Directly Using L-Moments from Airborne Lidar Data. ***Remote Sensing of Environment*** 194: 437-446.

[7] **Valbuena R.**, Hernando A., Manzanera J.A., **Görgens E.B.**, Almeida D.R.A., Mauro F., García-Abril A. & Coomes D.A. (2017) Enhancing of Accuracy Assessment for Forest Above-Ground Biomass Estimates Obtained from Remote Sensing via Hypothesis Testing and Overfitting Evaluation. ***Ecological Modelling*** 366: 15-26.

[8] Hernando A., Velázquez J., **Valbuena R.**, Legrand M. & García-Abril A. (2017) Influence of the Resolution of Forest Cover Maps in Evaluating Fragmentation and Connectivity to Assess Habitat Conservation Status. ***Ecological Indicators*** 79: 295-302.

[9] Vihervaara P., Mononen L., Auvinen A.P., Virkkala R., Lü Y., Pippuri I., Packalen P., **Valbuena R.** & Valkama J. (2015) How to Integrate Remotely Sensed Data and Biodiversity for Ecosystem Assessments at Landscape Scale. ***Landscape Ecology*** 30 (3): 501-516.

[10] **Valbuena R.**, Packalén P., Martín-Fernández S. & Maltamo M. (2012) Diversity and Equitability Ordering Profiles Applied to the Study of Forest Structure. ***Forest Ecology and Management*** 276: 185-195.

5. Projects:

- 2019–2022: **NERC Envision DTP project** (Extending global forest carbon models based on LIDAR from tropics to other bioregions). Bangor University (£83k). Role: PI.
- 2016–2018: **H2020-MSCA-IF project LORENZLIDAR** (Classification of forest structural types with LiDAR applied to tree size-density scaling theories). Grant No. 658180, Cambridge University (€195k). Role: Post-Doctoral Research Associate.
- 2010–2014: **Foundation for European Forest Research's scholarship** (Pan-European Indicators of Forest Structure from LIDAR). European Forest Institute (€80k). Role: PI.
- 2017: European Space Agency project **GlobDiversity** (RS-enabled EBVs for terrestrial ecosystems). UNEP-WCMC. Role: Participation in vegetation height component: assessing the feasibility of LIDAR to retrieve EBVs.
- 2013: **EC FP7-SPACE project HELM** (Harmonised European Land Monitoring). European Forest Institute. Role: Project Manager for EFI after a substitution.

6. Other:

- Member of GEO BON for ecosystem structure working group and remote sensing task force.
- International Union of Forest Researchers (IUFRO) Outstanding Doctoral Research Award.
- Editorial work: Editorial Board Member of *Forests*, and Guest Editor for special issues of *Remote Sensing*.
- Publicly-available software development and datasets: (See also collaborative projects in other partner's CVs.)

[11] de Almeida D.R.A., Stark S.C., **Silva C.A.**, Hammamura C. & **Valbuena R.** (2019) leafR: Calculates the Leaf Area Index (LAD) and Other Related Functions. R package. Version 0.3. **CRAN - The Comprehensive R Archive Network**. Downloads: 2,322 (18/01/2020)

[12] Verkerk P.J., Levers C., Kuemmerle T., Lindner M., **Valbuena R.**, Verburg P.H. & Zudin S. (2015) Data from: Mapping Wood Production in European Forests. **Dryad Digital Repository**. Downloads: 510 (13/05/2020)

Curriculum Vitae **Görgens, Eric Bastos**

1. General information:

Görgens, Eric Bastos, Dr., DoB: 05/06/1983, male
Universidade Federal dos Vales do Jequitinhonha e Mucuri, Departament de Engenharia Florestal, Diamantina, Minas Gerais, Brazil, +55 38 998 440 480, eric.gorgens@ufvjm.edu.br
Professor
Google Scholar: https://scholar.google.com.br/citations?user=5y_CUfQAAAAJ&hl=pt-BR
ORCID: https://orcid.org/0000-0003-2517-0279

2. Academic qualifications:

Doctorate	Forest Resources, University of Sao Paulo, 2014, Prof. Luiz Carlos Estraviz Rodriguez
Other advanced qualification	Forest Sciences, Universidade Federal de Viçosa, 2006, Prof. Helio Garcia Leite

3. Professional career:

2013 – 2015	FSC External Auditor, IMAFLORA, São Paulo, Brazil
2008 – 2011	Research and Planning Coordinator, Companhia do Vale do Araguaia, Mato Grosso, Brazil
2006 – 2008	Junior Engineer, Votorantim Metais, Minas Gerais, Brazil

4. Publications: (See also collaborative publications in other partner's CVs.)

[13] **Görgens E.B.**, Nunes M.H., Jackson T., Reis C.R., Almeida D.R.A., Coomes D.A., **Keller M.**, Gimenez B., **Valbuena R.**, Rosette J., Cantinho R.Z., Motta A.Z., Assis M., Pereira F.S., Spanner G., Higuchi N. & **Ometto J.P.** (2020) Resource Availability and Disturbance Shape Maximum Tree Height across the Amazon. ***bioRxiv*** [pre-print].

[14] **Görgens E.B.**, Mund J.P., Cremer T., de Conto T., Krause S., **Valbuena R.** & Rodríguez L.C. (2020) Automated Operational Logging Plan Considering Multi-Criteria Optimization. ***Computers and Electronics in Agriculture*** 170: 105253.

[15] **Görgens E.B.**, Motta A.Z., Assis M., Nunes M.H., Jackson T., Coomes D.A., Rosette J., Aragão L.E.O.C. & **Ometto J.P.** (2019) The Giant Trees of the Amazon Basin. ***Frontiers in Ecology and the Environment*** 17 (7), 373-374.

[16] Almeida D.R.A., Stark S.C., Schietti, J., Camargo J.L.C., Amazonas N.T., **Görgens E.B.**, Rosa D.M., Smith M.N., **Valbuena R.**, Saleska S., Nelson B.W., Andrade A., Mesquita R., Laurance S.G., Laurance W.F., Lovejoy T.E. & Brancalion P.H.S (2019) Persistent Effects of Fragmentation on Tropical Rainforest Canopy Structure after 20 Years of Isolation. ***Ecological Applications*** 29 (6): e01952. (**Biological Dynamics of Forest Fragments Project (BDFFP)** technical series number 758).

[17] Tejada G., **Görgens E.B.**, Espírito-Santo F.D.B., Cantinho R.Z. & **Ometto J.P.** (2019) Evaluating Spatial Coverage of Data on the Aboveground Biomass in Undisturbed Forests in the Brazilian Amazon. ***Carbon balance and management*** 14 (1), 11.

[18] Almeida D.R.A., Mendes A.F., Nelson B.W., **Görgens E.B.**, Schietti, J., Duarte M.M., Gonçalves N.B., Amazonas N.T., Meli P., Brancalion P.H.S, Chazdon R., **Valbuena R.**, Stark S.C. & Moreno V. (2019) The Effectiveness of Lidar Remote Sensing for Monitoring Tree Cover Attributes in Forest and Landscape Restoration. ***Forest Ecology and Management*** 438: 34-43.

[19] Almeida D.R.A., Broadbent E., Zambrano A., Wilkinson B., Ferreira M.A., Chazdon R., Meli P., **Görgens E.B.**, **Silva C.A.**, Stark S.C., **Valbuena R.**, Papa D.A. & Brancalion P.H.S. (2019) Monitoring the Structure of Forest Restoration Plantations with a Drone-Lidar System. ***International Journal of Applied Earth Observation and Geoinformation*** 79: 192-198.

[20] Andrade M.S., **Görgens E.B.**, Cantinho R.Z., Assis M., Sato L. & **Ometto J.P.** (2018) Airborne Laser Scanning for Terrain Modeling in the Amazon Forest. ***Acta Amazonica*** 48 (4), 271-279.

[21] **Görgens E.B.**, **Valbuena R.** & Rodríguez L.C. (2017) A Method for Optimizing Height Threshold when Computing Airborne Laser Scanning Metrics. ***Photogrammetric Engineering and Remote Sensing*** 83(5): 343-350.

[22] Almeida D.R.A., Nelson B.W., Schietti, J., **Görgens E.B.**, Resende A.F., Stark S.C. & **Valbuena R.** (2016) Contrasting Fire Susceptibility and Fire Damage Between Seasonally Flooded Forest and Upland Forest in the Central Amazon Using Portable Terrestrial Profiling Lidar. ***Remote Sensing of Environment*** 184: 153-160.

5. Projects:

2014 – 2018	**EBA - Improving Biomass Estimation for the Amazon** - Coordinated by Dr. Jean Ometto (INPE). Funded by Amazon Fund.
2020 - 2021	**Paisagens Sustentáveis Brasil**. Coordinated by Michael Keller. Funded by US-FS.
2017 – 2020	Mapping operational and environmental constraints, and logging potential for an Amazon site. Coordinated by E. B. Gorgens. Funded by The Brazilian National Council for Scientific and Technological Development (CNPq).
2020 - 2021	**The giant trees of RDS Iratapuru**. Coordinated by Diego Armando Silva. Funded by Fundo Iratapuru, Fundação Natura.
2012	EU FP7-PEOPLE-2009-IRSES project ForEAdapt: Knowledge exchange between Europe and America on forest growth models and optimisation for adaptive forestry. University of São Paulo.

6. Other:

- Member of GEO BON for ecosystem structure working group.
- **Leading the expedition to locate, identify and measure the giant trees of the Amazon**. This expedition attracted media coverage, following up the publication of Görgens et al. (2019) in *Frontiers in Ecology and the Environment*:
  Video in English:
  https://www.cam.ac.uk/research/news/expedition-finds-tallest-tree-in-the-amazon
  Video in Portuguese:
  https://globoplay.globo.com/v/7924940/

Curriculum Vitae **Silva, Carlos Alberto**

1. General information:

Silva, Carlos A., PhD. DoB: 24/07/1987, male
University of Florida, School of Forest Res. & Conservation, +55 31996630407, c.silva@ufl.edu
Assistant Professor of Quantitative Forest Science
ORCID: https://orcid.org/0000-0002-7844-3560

2. Academic qualifications:

Doctorate	Ph.D., Natural Resources, University of Idaho, College of Natural Resources, 2018, Supervisors: Lee A. Vierling, Ph.D.; Andrew T. Hudak, Ph.D.
MSc qualification	Forest Resources , University of São Paulo – "Luiz de Queiroz" college of Agriculture –ESALQ, 2013, Luiz Carlos Estraviz Rodriguez, Ph.D.

3. Professional career:

2020 - present	Assistant Professor of Quantitative Forest Science. University of Florida, School of Forest Res. & Conservation, USA.
2018-2020	Postdoctoral Researcher. University of Maryland / NASA Goddard Space Flight Center, USA.
2016-2017	Research Visitor, NASA - Jet Propulsion Laboratory (JPL), USA
2013-2014	Research Visitor, USDA Forest Service - Rocky Mountain Research Station (RMRS), USA

4. Publications: (See also collaborative publications in other partner's CVs.)

[23] Duncanson L., Neuenschwander A., Hancock S., Thomas N., Fatoyinbo T., Simard M., Luthcke S., **Silva C. A.**, Armston J., Hofton M. & Dubayah R. (2020) Biomass Estimation from Simulated GEDI, ICESat-2 and NISAR Data Across Environmental Gradients in Sonoma County, California. ***Remote Sensing of Environment*** 242: 111779.

[24] Qu Y., Shaker A., Korhonen L. **Silva C.A.**, Jia K., Tian L. & Song J. (2020) Direct Estimation of Forest Leaf Area Index based on Spectrally Corrected Airborne LiDAR Pulse Penetration Ratio. ***Remote Sensing*** 12: 217.

[25] Rex F.E., **Silva C.A.**, Dalla Corte A.P., Klauberg C., Mohan M., Cardil A., Sousa da Silva V., Almeida D.R.A., García M., Broadbent E.N., **Valbuena R.**, Stoddart J., Merrick T. & Hudak A.T. (2019) Comparison of Statistical Modelling Approaches for Aboveground Biomass Change Estimation in a Selectively Logged Tropical Forest from multi-temporal LiDAR data. ***Remote Sensing*** 12: 1498.

[26] Almeida D.R.A., Broadbent E., Wendt A., Wilkinson B., Papa D.A., Chazdon R., Meli P., **Görgens E.B.**, **Silva C.A.**, Stark S.C., **Valbuena R.**, Almeyda A., Fagan M., Foster P. & Brancalion P.H.S. (2020) Detecting Successional Changes In Tropical Forest Structure Using Gatoreye Drone-Borne Lidar. ***Biotropica*** [in press].

[27] Papa D., de Almeida D.R.A., **Silva C.A.**, Figueiredo E., Stark S., **Valbuena R.**, Rodriguez L.C. & d'Oliveira M.V.N. (2020) Evaluating Tropical Forest Classification and Field Sampling Stratification from Lidar to Reduce Effort and Enable Landscape Monitoring. ***Forest Ecology and Management*** 457: 117634.

[28] **Silva C.A.**, **Valbuena R.**, Pinagé E.R., Mohan M., Almeida D.R.A., Broadbent E.N., Jaafar W.S.W.M., Papa D.A., Cardil A. & Klauberg C. (2019) ForestGapR: An R Package for Forest Gap Analysis from Canopy Height Models. ***Methods in Ecology and Evolution*** 10: 1347-1356.

[29] Almeida D.R.A., Stark S.C., **Valbuena R.**, Broadbent E.N., Silva T.S.F., Resende A.F., Ferreira M.P., Cardil A., **Silva C.A.**, Amazonas N., Zambrano A.M. & Brancalion P.H. (2019) A New Era in Forest Restoration Monitoring. ***Restoration Ecology*** 28: 8-11.

[30] Klauberg C., Hudak A., **Silva C.A.**, Lewis S., Robichaud P. & Jain T. (2019) Characterizing Fire Effects on Conifers at Tree Level from Airborne Laser Scanning and High-Resolution, Multispectral Satellite Data. ***Ecological Modeling*** 412: 108820.

[31] Eitel J., Maguire A., Boelman N., Vierling L.A., Griffin K., Jensen J., Mahoney P., Meddens A., **Silva C.A.** & Sonnentag O. (2019) Proximal Remote Sensing of Tree Physiology at Northern Treeline: Do Late-Season Changes in the Photochemical Reflectance Index (PRI) Respond to Climate or Photoperiod? ***Remote Sensing of Environment*** 221: 340-350.

[32] Almeida D.R.A., Stark S.C., Shao G., Schietti J., Nelson B.W., **Silva C.A.**, **Görgens E.B.**, **Valbuena R.**, Papa D.A. & Brancalion P.H.S (2019) Optimizing the Remote Detection of Tropical Rainforest Structure with Airborne Lidar: Leaf Area Profile Sensitivity to Pulse Density and Spatial Resolution. ***Remote Sensing*** 11(1): 92-107

<u>6. Projects:</u>

- 2019: The **Strategic Environmental Research and Development Program (SERDP) -** 3D fuel characterization for evaluating physics-based fire behavior, fire effects, and smoke models on US Department of Defense military lands. (Collaborator). US$ 2.6M.
- 2019: SERDP **-** Object-based aggregation of fuel structures, physics-based fire behavior and self organizing smoke plumes for improved fuel, fire, and smoke management on military lands (Collaborator). US$ 2.4M.
- 2018: The **Brazilian National Council for Scientific and Technological Development (CNPq) -** Mapping fuel load and simulation of fire behavior and spread in the Cerrado biome using modeling and remote sensing technologies. (Co-PI) US$45k.
- 2018: **NASA Carbon Monitoring System (NASA-CMS) -** Future Mission Fusion for High Biomass Forest Carbon Accounting (Collaborator).

<u>7. Other:</u>

- Editorial work: Guest Editor for special issues of *Remote Sensing* and *Forests*.
- <u>Publicly-available software development and datasets</u>, including **R packages** for processing of satellite LIDAR data (rGEDI) and the first R package for airborne LIDAR (rLIDAR):

[33] **Silva C.A.**, Hammamura C., **Valbuena R.**, Hancock S., Cardil A., Broadbent E.N., de Almeida D.R.A., Silva C.H.L., & Klauberg C. (2020) rGEDI: An R Package for NASA's Global Ecosystem Dynamics Investigation (GEDI) Data Visualization and Processing. R package. Version 0.0.6. **CRAN - The Comprehensive R Archive Network** Downloads: 1,723 (28/05/2020)

[34] **Silva C.A.**, Hudak A.T., Crookston N. L., Vierling L.A., Klauberg C.A., & **Valbuena R.** (2019) TreeTop: a web-LiDAR application for forest inventory. **Shiny App** . Visits: 8,343 (13/05/2020)

[35] **Silva C.A.**, Pinagé E.R., Mohan M., de Almeida D.R.A., Broadbent E.N., Jaafar W.S.W.M., Papa D.A., Cardil A., **Valbuena R.**, & Klauberg C. (2018) ForestGapR: Tropical Forest Gaps Analysis. R package. Version 0.0.02. **CRAN** Downloads: 4,713 (18/01/2020)

[36] **Silva C.A.**, Crookston, N.L., Hudak, A.T., and Vierling, L.A. (2014) rLiDAR: LiDAR Data Processing and Visualization. Version 0.1.1. **CRAN** Downloads: 23,000 (28/25/2020)

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÃO
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

Dr. Jean Pierre Ometto
Head of Earth System Science Center
Brazilian National Institute for Space Research (CCST/INPE)
Av dos Astronautas, 1.758
CEP: 12227-010
São José dos Campos – SP
Tel: 55 (12) 3945-7109
jean.ometto@inpe.br

Dr Rubén Valbuena
Lecturer in Forest Sciences

School of Natural Sciences,
Bangor University,
Gwynedd, LL57 2UW, UK.
r.valbuena@bangor.ac.uk

Ref.: Letter of support for AMAZECO (Covering the Amazon with an Ecosystem Structure EBV product combining satellite and airborne LIDAR) proposal

04/06/2020

To whom it may concern,

I am pleased to support the project proposal lead by Dr Rubén Valbuena, Prof. Eric Görgens and Dr. Carlos A. Silva, which aims to develop tools for compute EBV combining satellite and airborne LIDAR in the Amazon, bidding for support from Microsoft- GEO BON funding.

The Earth System Science Center (CCST) is part of the Brazilian National Institute for Space Research (Instituto Nacional de Pesquisas Espaciais, INPE). The main goals of which are fostering scientific research and technological applications and qualifying personnel in the fields of space and atmospheric sciences, space engineering, and space technology.

CCST will contribute to AMAZECO by securing the accessibility of airborne LIDAR data from the improving biomass estimation methods for the Amazon (EBA) project. Prof. Görgens has been collaborating with the EBA project since its inception, and I will liaise very closely with him to provide support for the use of these datasets in the framework of AMAZECO. With the intention to attain the combination with GEDI data, the AMAZECO project suggests processing workflows which differ substantially from the standard airborne LIDAR workflows we currently employ with those datasets. Thus, there is no overlap between AMAZECO objectives and research opportunities with the current goals and ongoing research using EBA data. We are thus pleased to contribute with EBA data and benefit from the synergies pursued in AMAZECO, both in the combination of with GEDI but also in the derivation of Essential Biodiversity Variables (EBVs) for the Amazon basin.

I am looking forward to see AMAZECO project funded, follow up its development closely and provide suitable advice upon request, and contribute to derived research components with my experience. It will be a great opportunity to start participating with GEO BON, which eventually could derive in the development of a national Brazilian BON.

Yours sincerely,

Jean Pierre Ometto
Senior Research Scientist
Earth system Science Centre, Head
National Institute of Space Research

USDA United States Department of Agriculture | Forest Service | International Programs | 1099 14th Street, NW
Suite 5500W
Washington, DC 20005-3402

June 2, 2020

Rubén Valbuena, D.Sc., Ph.D.
School of Natural Sciences,
Bangor University,

To Whom It May Concern:

This is a letter of support for Dr Rubén Valbuena's proposal for AMAZECO proposal to be submitted for application to Microsoft – GEO BON funding.

The AMAZECO project will develop products relevant to Essential Biodiversity Variables (EBV) in the Amazon. In particular, it will combine satellite LiDAR (Light Detection and Ranging) from NASA's global ecosystem dynamics investigation (GEDI) mission with airborne LiDAR data locally surveyed in Brazil. The development of EBV products can be of key importance in preserving and monitoring Amazon's biodiversity. Thus, I am interested in the objectives and outputs proposed in AMAZECO project. I will support the project through my with knowledge of airborne LiDAR from the Sustainable Landscapes project in Brazil. These extensive data provide a temporal dimension to on forest dynamics with time series of both pristine and managed forests throughout Brazil.

The Sustainable Landscapes project focuses in the monitoring and investigating dynamics of ecosystems after anthropogenic interference, essential for mitigating the increase of global levels of greenhouse gases. We aim to develop local technical capabilities in Brazil to act on reduction of emissions from deforestation and forest degradation REDD+ initiatives, with emphasis on monitoring, reporting and verification actions. The USDA Forest Service works on this initiative with Embrapa Agricultural Informatics, which aims at developing methods for extrapolating local measurements of forest carbon and other greenhouse gases to the national level. We use LiDAR data to increase Brazil's capacity for forest carbon inventories.

The synergies between AMAZECO and Sustainable Landscapes are strong and reinforce our goal to develop local technical capabilities. Making accessible tools for the scientific community can speed up findings and improve shareable knowledge among research groups. EBV products from LiDAR will enhance the study of biodiversity in the Amazon, a topic of outmost importance, given the current regional and global challenges.

Yours sincerely,

Michael Keller, Ph.D.
Research Physical Scientist

FEC AND SPONSOR PRICE

BANGOR UNIVERSITY

Bangor Reference:	**F009137**	**PROJECT TITLE:**	**AMAZECO - An Ecosystem Structure EBV product covering the Amazon combining satelite and airborne LIDAR**
Project Start Date:	**01 Sep 20**	**FUNDER:**	**Microsoft**
Project End Date:	**31 Aug 21**	**LEAD APPLICANT:**	**Bangor University**
Project Duration:	**12.0 mnths**		
Date of Issue:	**03 Jun 20**		Show All

						FEC				**PRICE**		
						Year 1 01/09/20 to 31/08/21	**TOTALS (£)**	Eligible Cost	% Allowable	**Year 1** 01/09/20 to 31/08/21	**TOTALS (£)**	**Budget in $**
DIRECTLY INCURRED COSTS												
STAFF												
Name/Job Title	Grade and Point	Start Date	End Date	Duration (months)	% FTE on project				**100.0%**			
PDRA	Grade 7 Point 30	01 Sep 20	31 Aug 21	12	100.0%	47,021	47,021	**YES**	100.0%	47,021	47,021	**69,262**
PDRA in UFVJM, Brazil		01 Sep 20	31 Aug 21	12	100.0%	12,142	12,142	**YES**	100.0%	12,142	12,142	**17,885**
Sub-Total Directly Incurred Staff						**59,163**	**59,163**			**59,163**	**59,163**	**87,148**
TRAVEL AND SUBSISTENCE												
US Co-I in Bangor. Travel & Acoomodation						2,908	2,908	**YES**	100.0%	2,908	2,908	**4,284**
Brazil Co-I in Bangor. Travel & Acoomodation						2,908	2,908	**YES**	100.0%	2,908	2,908	**4,284**
Brazil PDRA in Bangor. Travel & Acoomodation						2,908	2,908	**YES**	100.0%	2,908	2,908	**4,284**
Sub-Total Travel & Subsistence						**8,724**	**8,724**			**8,724**	**8,724**	**12,851**
DIRECTLY ALLOCATED COSTS												
STAFF - COST OF TIME ON PROJECT												
Name/Job Title		Start Date	End Date	Duration (months)	% FTE on project							
Ruben Valbuena		01 Sep 20	31 Aug 21	12	10.0%	6,949	6,949	**NO**				
Sub-Total Directly Allocated Staff Costs						**6,949**	**6,949**					
ESTATES COSTS												
Estates costs related to Bangor University Project Staff FTEs						5,040	5,040	**NO**				
Sub-Total Estates Costs						**5,040**	**5,040**					
INDIRECT COSTS												
Indirect costs related to Bangor University Project Staff FTEs						48,201	48,201	**NO**				
Sub-Total Indirect Costs						**48,201**	**48,201**					
TOTAL COST OF PROJECT AND TOTAL PRICE TO BE PAID BY FUNDER						**128,077**	**128,077**			**67,887**	**67,887**	**$99,998**

Microsoft Azure Estimate

Your Estimate

Service type	Custom name	Region	Description	Estimated monthly cost	Estimated upfront cost
Storage Accounts		UK West	Block Blob Storage, Blob Storage, LRS Redundancy, Hot Access Tier, 80 TB Capacity - Pay as you go, 100,000 Write operations, 100,000 List and Create Container Operations, 100,000 Read operations, 100,000 Archive High Priority Read, 1 Other operations. 1,000 GB Data Retrieval, 1,000 GB Archive High Priority Retrieval, 1,000 GB Data Write	$1,550.50	$0.00
Virtual Machines		UK West	2 D13 v2 (8 vCPU(s), 56 GB RAM); Windows – (OS Only); 3 year reserved; 0 managed OS disks – S4, 100 transaction units	$948.38	$0.00
Azure Functions		UK West	Consumption tier, 1536 MB memory, 100 milliseconds execution time, 1,000 executions/mo	$0.00	$0.00
Cost Management + Billing			No charge for managed Azure spend. Additional premium capabilities are available at no cost through December 2018 when they will become paid features. 1% of managed spend for AWS and Google Cloud Platform.	$0.00	$0.00
Batch		UK West	Cloud Services: 1 A1 (1 Cores, 1.75 GB RAM) x 730 Hours; Virtual Machines: 1 A1 (1 Cores, 1.75 GB RAM) x 700 Hours	$92.24	$0.00
Bandwidth		UK West	Zone 1: North America, Europe, 1 TB	$88.65	$0.00
Azure Maps		West US	Standard S0 Tier, 100,000 Maps transactions, 1,000 Other Services transactions	$12.50	$0.00
Azure DevOps			5 Basic Plan license users, 0 Basic + Test Plans license users, Free tier - 1 Microsoft Hosted Pipeline(s), 1 Self Hosted Pipeline(s), 0 GB Artifacts, 0 VUMs	$0.00	$0.00
Azure Advisor			There are no charges to use Azure Advisor.	$0.00	$0.00
Azure Migrate		West Central US	There is no charge to use Azure Migrate.	$0.00	$0.00
Azure Kubernetes Service (AKS)		UK West	1 D2 v3 (2 vCPU(s), 8 GB RAM) nodes x 730 Hours; Pay as you go; 0 managed OS disks – S4, 0 clusters	$84.68	$0.00
Support			Support	$0.00	$0.00
			Licensing Program	Microsoft Online Services Agreement	
			Total	$2,776.95	$0.00

Total in 36 months	$99,970.18

Disclaimer

All prices shown are in US Dollar ($). This is a summary estimate, not a quote. For up to date pricing information please visit https://azure.microsoft.com/pricing/calculator/

This estimate was created at 6/5/2020 9:40:37 PM UTC.