Finegayan Baseline Survey

Section 1: Research Data Management Documentation Table

Name of Principle Investigator(s) or Contractor(s) and ORCiD number

Haldre Rogers, Martin Kastner

Current Project Title, and all previous project titles

Iowa DOT Project Manager

Iowa DOT Project number

Other contract or grant numbers

Iowa DOT Research-assigned project Digital Object Identifier (DOI), or researcher acquired DOI

TRB Research in Progress (RiP) Title, Accession Number, and URL

Project Duration (projected)   Start Date: End date:

Do the data management requirements of the US DOT "Plan to Increase Public Access to the Results of Federally-Funded Scientific Research" apply to this project   Yes or No; and if No, why not:

Name(s) of Federal funder(s), Funding Program Name(s), Agency Code(s) and/or Contract/Grant numbers

DMP Version

Date DMP amended, if any

Name and ORCiD number of each author

Persistent links or identifiers assigned to this project, datasets, reports, or peer reviewed publications generated by this project

Name and URL of all peer reviewed publications which have been generated from this project

Format files

Raw data - excel spreadsheets, proofed and scanned copies of original data sheets stored in Cloud Box Drive under EBL Longterm Repository. Data points saved as coordinates in Projected coordinate system: WGS 1984 UTM 55N

Merged and Tidy data - .csv files, tidyed and wrangled in R. R acquisition codes will be saved in Cloud Box Drive

Forest transect points will be saved and sent in WGS 1984 UTM 55N Projected Coordinate System final for the Navy in Navy Data Model 3.1

Abiotic and biotic conditions within Forest Enhancement Area using belt transect survey method following EBL Data Gathering Protocols created by Martin Kastner. Randomly selected transects were placed within each of the following forest types; Native Limestone Forest, Degraded Limestone Forest, Vitex Forest,

We will then conduct 25-meter-long, 2-m wide belt transects at random points across North and South Finegayan forest enhancement areas, stratified by the habitat types identified above. Points will be at least 50 meters apart, but at a density sufficient to capture the heterogeneity within each forest type.

Along each transect, we will record trees, shrubs, vines, and herbaceous plants, substrate/soil type, ungulate damage, canopy cover, and ground cover. For each tree, we will record height, diameter at breast height (dbh), and species identity. For each shrub, we will record only species identity. We will record presence/absence of each vine species along the transect, and whether the vine covers 0-25, 25-50, 50-75 or 75-100% of the transect. We will subsample for herbaceous plants and woody seedlings, in 1 x 1m plots at 0, 5, 10, 15, 20, and 25m along the transect. In each of those plots, we will also record the substrate, light, % canopy cover, and % vegetation cover; if there is soil, we will take a soil sample for nutrient analysis, combining soil from each of the subplots. For canopy cover, we will use a densiometer, taken once in each of four directions by the same person on all transects. Data collected will be on hard copy datasheets, entered, proofed and scanned into Box Cloud Drive.

Using multivariate approaches with our survey results, we will cluster transects into landcover types, then revisit our original landcover categorizations, and amend as necessary.

Pre-existing data

An in depth literature review of Northern Guam will be done to get a historic overview of previous habitat types, species composition, and past disturbances within FEA that contains forest transect survey.

94 belt transect data with canopy height, understory, vine cover, and additional species collected per belt transect. data sheet is ~30 kb, 3MB of raw data sheets, scanned images of datasheets at 30 MB.

Stored data in Longterm Data Repository stored in Iowa State University Box iCloud Drive, and backed up with 1 Terrabyte hard drive.

All on the project, PI Haldre Rogers verifies proper storage of data.

Data is unique in the time it was collected, methods will be reproducible, findings may not due to stochastic events, and future management actions (ungulate fencing).

If data was lost aor became unusuable later, data will have to be recollected in the field. However, there are several places that data is stored to prevent that from happening.

Natural resource managers will use the summarized data from forest transects to compare baseline surveys with future vegetation monitoring surveys to determine if management actions are effective.

Raw data format follows Hadly Wickam recommended data sheet format. Every observation is a row.

Final data will be stored in EBL Longterm Data Repository. Each folder of data will have metadata form describing file layers.

Final summarized data may also be located as appendices to final Finegayan Baseline Monitoring Report.

Final GPS transect points will be saved in Navy Data Model (NDM) 3.1 as dictated by contract requirements. NDM requires metadata to be filled out for all layers to be able to provide context of coordinate points.

Standards

NDM 3.1,Darwin Core for biological data, ISO 15836:2006.

Descriptive metadata will be found in GPS coordinate points using ArcGIS.

Research protocols are archived to capture data collection methods and within the data sheet excel for how to fill out data forms.

Each folder containing data will include a metadata document, providing context for each type of file and/or layer.

All data in various forms will be saved in one folder with subfolders seperated by raw data, data wrangling, and tidy data. Raw data is entered, proofed and scanned datasheets. Data wrangling folder contains data acquisition code that uses data from raw data folder to create .csv files that will be used for data analysis. Tidy data folder will include tidy data summary.csvs and all R code to conduct data analysis. All data folders will contain metadata document to provide context of files.

Raw data files will include Title of project, Transect # and date taken (Finegayan Veg Transect_2_30Apr2019)

Tidy data files will include type of data collected and date merged. (VegUNDERSTORYmerged30Aug2019)

Navy Data Model 3.1

Data collected is does not contain any personally identifiable information, but it is the property of the Department of Navy and all papers, articles, and presentations generated fromt the following information must receive DON approval prior to publishing information.

Raw data will be stored and only accessible to EBL employees therefore restricting access until proper approval is received. If data is to be accessed a request must be submitted to the PI Dr.Haldre Rogers for use.

Funder and PI control the data. Proper permission must be acquired by funder prior to using in other capacities seperate from stated scope in contract.

Audience for reuse will be future contractors working for the DON and will have access through the DON as they will receive all raw data collected from forest transects.

Funder data sharing policy requires to be acknowledged in any approved published or publicy presented data.

Data will be published to the DON in a final report for their use.

Software needed to work with data is Microsoft Office, ArcGIS, and R Studio or Console.

The archive that will data will be stored in will be a cloud drive controlled and monitored by the main PI to protect intellectual property rights.

Data will be stored for long-term accesss through the PI.

Data should be retained permanently for the detailed information collected in a region that will receive heavy management actions.

Files will be saved in the following formats doc., .csv, .gdb, .xls, .Rproj, .txt pdf files of scanned images which are not proprietary.

EBL will maintain data long-term and Navy will have a copy of portions of complete and summarized data.