NIST DSE Plant Identification with Remote Sensing Evaluation Plan

Version 0.1, Updated on April 25th, 2016

1. Introduction

This document defines the tasks and evaluation plan for the *Data Science for Plant Identification with Remote Sensing* component of the National Institute of Standards and Technology (NIST) Data Science Evaluation (DSE) Series Pilot.

Understanding ecological patterns and processes across geographical scales is crucial to understanding the effects of environmental change on natural systems and human society. However, financial and logistic limitations restrict the scales at which ecological data can be collected by field ecologists on the ground. Recent increases in the availability of remotely sensed imagery from satellites and aircraft provide the potential to observe ecological data at much larger scales than are possible through traditional data collection methods. This DSE focuses on inferring ecological information from remote sensing data.

The DSE focuses on data from the National Science Foundation sponsored National Ecological Observatory Network (NEON). NEON collects continental-scale ecological observations from 81 sites across the United States. Data products range from ground based data on plant and animal locations and traits, to landscape scale airborne remote sensing. Dealing with the volume, velocity, and variety of this data will require interdisciplinary approaches involving ecology, computer science, statistics, and data science. This NIST Data Science Evaluation (DSE) competition is an applied, interdisciplinary, challenge using multiple components of the NEON data stream. Aligning this DSE with NEON has the potential to result in pioneering interdisciplinary approaches and large gains in our understanding of the natural world by engaging the larger Data Science community in this unique data source. In addition, many of the algorithms and techniques are expected to generalize to other problems in data science.

This competition will combine airborne remote sensing data with field-based tree measurements, species identifications, and crown segmentation, to develop and evaluate models to address three common data science tasks (segmentation, alignment, and classification) to estimate tree location, size, and species identity from the remote sensing data.

1. **Segmentation:** Locate objects, or boundaries surrounding objects, in images. Estimate the size, shape, and location of individual tree crowns (the top, sun-exposed portion of the tree visible from above) from remote sensing data.
2. **Alignment:** Associate two or more representations of the same object. Match individual trees measured on the ground with the same tree observed using remotely sensed data.
3. **Classification:** Determine which of a set of categories an object belongs to. Determine the species identity (i.e., the type) of each individual tree from remotely sensed data.

These tasks are designed so that they can be completed independently, but also so that they can be combined to form a data processing pipeline for real world applications. To allow tasks to be completed independently, the release of data and submission of results will occur in three phases, with the ideal output for the previous phase released at the beginning on the next phase. Details of the process for data release and rules for the competition are described in Section 7.

2. Data

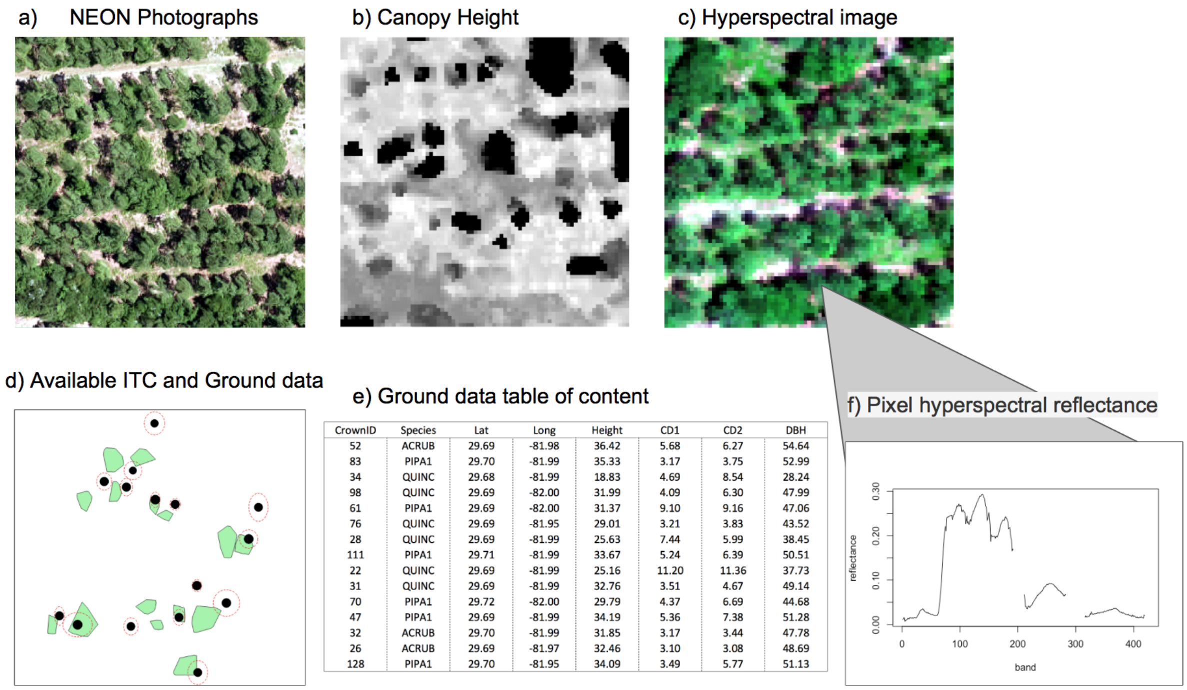
Seven datasets will be used as part of this evaluation:

- Remote Sensing Data -- 47 plots totalling ~1GB of imagery
  - NEON Photographs - 0.0625 m^2 resolution photographic images (RGB)
  - NEON LiDAR data - 3-dimensional point clouds of the height of trees and ground, which are also provided as 1 m^2 resolution raster images of the estimated height of vegetation above the ground (Canopy Height Model; CHM).
  - NEON Hyperspectral images - 1 m^2 resolution raster images with 426 wavelength bands from 380-2510 nm. These data are georectified and atmospherically corrected, and bands that are strongly influenced by the absorbance by atmospheric water (and should therefore be removed) will be specified in advance.
- Ground Data -- 145 trees
  - NEON Vegetation Structure data - Ground based positions of individual trees, the size of each trunk, and for some trees two measurements of crown size (maximum crown diameter and the diameter of the crown perpendicular to the axis of the maximum diameter)
  - NEON Species Data - the genus and species (type) of each tree
- Individual Tree Crown segmentations (ITC) Data -- 626 stems
  - Ordway-Swisher ITC data - Vector polygon data of tree crown boundaries and vector point data of trunk locations both mapped in the field directly onto remote sensing images. Boundaries and locations are drawn onto remote sensing images from the ground while observing the trees directly and then fine tuned to match clear crown edges in the images. The crown data include confidence scores qualitatively estimating the precision of the tree crown mapping.
  - Correspondence between ITC crowns and ground stems

ll spatial data will be provided in the UTM 17N coordinate system, but this detail can generally be ignored at this scale and *x* and *y values* treated as any standard grid.

Figure 1 illustrates these core data types. All data collected by NEON are publicly available with open licenses. The subset of data needed for this DSE will be posted on the DSE website for download.

ADD CONCEPTUAL FIGURE SHOWING ALL DATA HERE



CrownID	Species	Lat	Long	Height	CD1	CD2	DBH
52	ACRUB	29.69	-81.98	36.42	5.68	6.27	54.64
83	PIPA1	29.70	-81.99	35.33	3.17	3.75	52.99
34	QUINC	29.68	-81.99	18.83	4.69	8.54	28.24
98	QUINC	29.69	-82.00	31.99	4.09	6.30	47.99
61	PIPA1	29.69	-82.00	31.37	9.10	9.16	47.06
76	QUINC	29.69	-81.95	29.01	3.21	3.83	43.52
28	QUINC	29.69	-81.99	25.63	7.44	5.99	38.45
111	PIPA1	29.71	-81.99	33.67	5.24	6.39	50.51
22	QUINC	29.69	-81.99	25.16	11.20	11.36	37.73
31	QUINC	29.69	-81.99	32.76	3.51	4.67	49.14
70	PIPA1	29.72	-82.00	29.79	4.37	6.69	44.68
47	PIPA1	29.69	-81.99	34.19	5.36	7.38	51.28
32	ACRUB	29.70	-81.99	31.85	3.17	3.44	47.78
26	ACRUB	29.69	-81.97	32.46	3.10	3.08	48.69
128	PIPA1	29.70	-81.95	34.09	3.49	5.77	51.13

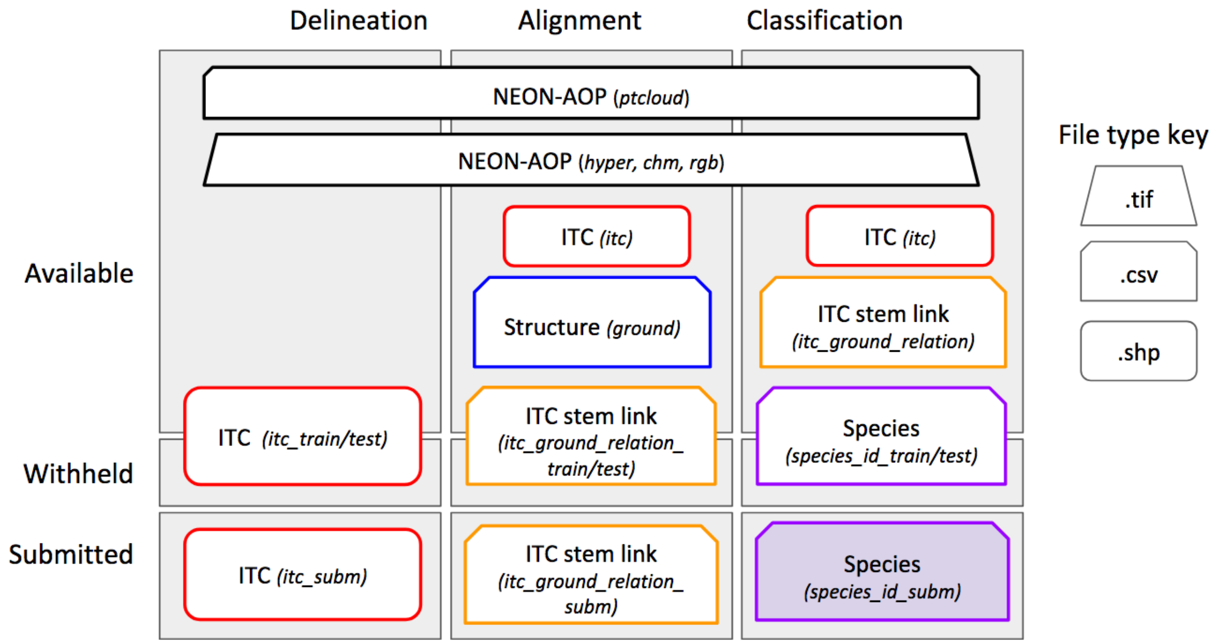
Figure 1. Example data products provided in the contest. Top row represents the remote sensing data: a) is the RGB image of a 80x80m plot, b) a CHM for the same region, c) a representation of the Hyperspectral data in false colors (a mixture of 3 hyperspectral bands; R = band 19, G = band 34, B = band 58). d) represents vectorial spatial data: green polygons ITCs collected for the area, black dots NEON ground data, dots' size is scaled by tree DBH. Red ellipse represent expected crown area, given the widest crown diameter and its orthogonal one. e) is a sample of the attributes associated to ground data. f) is an example of spectral signature for a given hyperspectral pixel.

Location: All data is from the Ordway-Swisher Biological Station (OSBS) NEON site. Field measures and remote sensing data are for 22 OSBS NEON distributed plots (1600 m^2) and 20 OSBS NEON tower plots (1600 m^2), which include hardwood/woody wetlands and upland pine/sandhill ecosystems. Figure 2 shows examples of ITC polygons on a hyperspectral image for a NEON plot from Ordway-Swisher.

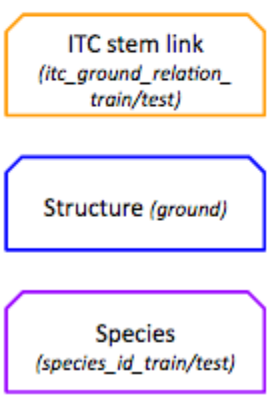
Input/Output data types for Each tasks:

Delineation task uses Neon-AOP dataset and 70% of ITC stem data as Training and 30% of the ITC stem data for testing.
Alignment task uses Neon-AOP dataset, ITC-stem data, NEON ground crown structure data, and 70% ITC-mapping data as Training and 30% of ITC-mapping data as testing.

Classification task uses Neon -AOP dataset, ITC-stem data, ITC-mapping data and 70% NEON ground



species data for Training and 30% of NEON ground species data as testing.
We further note that NEON-AOP data is in tif format, ITC-stem data is in .shp shape files and all other data types are in .csv format.
Figure 3 shows the input/output data types for each three tasks and the training and testing data.



ITC_id	tag_id
002	*391*

tag_id	x	y	dbh	ht	cd1	cd2
391	*32585*	*45963*	*23.3*	*12.2*	*3.5*	*2.1*

ITC_id	sp1	sp2
391	*QUNI*	*QU*

Figure 3: Training/testing Input/output data types for each of the three tasks. Data in both the train and test boxes will be split either at the plot level (segmentation and alignment tasks) on the stem level with 70% of the plots/stems used for training and 30% for testing. Products to be submitted are for the test data.Names in italics correspond to the names of the files provided or to be submitted.

NEON-AOP (black): Spatial data of the 4 types of NEON-AOP data. **ITC (red):** Spatial data of Individual Tree Crowns and a unique crown ID number. **ITC stem link (orange):** Table of unique crown ID number and the NEON ground data stem ID number. **Structure (blue):** Table of NEON ground data with tree stem attributes; unique stem ID, latitude & longitude coordinates, stem diameter, crown dimensions. **Species (purple):** Table of NEON ground data with assignment to two class categories (species class and genus class). Species product is a table of the probability of assignment of each ITC for each class in both class categories.

3. Segmentation

Segmentation is a common Data Science problem that involves locating objects or boundaries in images. In addition to being useful on its own, good segmentation is often important for classification of objects within images.

Summary: Segmentation

Inputs: Hyperspectral images, RGB images, and LiDAR point cloud data.

Output: Shapefiles containing estimates of the position and shape of individual tree crowns (ITC data).

Training Files: itc_train_*plot*.shp, hyper_*plot*.tiff, chm_*plot*.tiff, ptcloud_*plot*.csv, rgb_*plot*.tiff

Test Files: hyper_*plot*.tiff, chm_*plot*.tiff, ptcloud_*plot*.csv, rgb_*plot*.tiff

Submission File: itc_subm_*plot*.shp (including all 4 sub files; .shp, .shx, .dbf, .prj)

3.1 Segmentation in the ecology domain

Identifying the position and size of individual trees from remote sensing is useful for understanding forest structure and an important first step in species classification. It is also a more complex version of the common image segmentation task because trees often overlap each other and look similar, and because the available data is heterogeneous, involving many bands, multiple resolutions, and point cloud height data. Participants will use Hyperspectral, LiDAR and high resolution RGB photographs to locate and segment individual crowns in multiple plots.

3.2 Training Data

Input for the training phase will be data for a subset of plots at Ordway Swisher and will include shapefiles with polygons for each individual tree crown (ITC) along with associated hyperspectral, LiDAR, and RGB data. LiDAR data provides information on the spatial variation in canopy height that may allow partitioning of crowns of neighboring trees with similar spectral signatures. Hyperspectral data allows development of spectral signatures to identify object categories (e.g. by assigning spectrally-similar categories to the same cluster). RGB photographs provide finer resolution information (0.25x0.25 m as opposed to 1.0x1.0 m for the hyperspectral data), which may be helpful to separate trees that are close to one another and to refine boundary placement. Using the remote sensing data, participants will build models to estimate crown boundaries in plots where ITC polygons are not available. The ITC data can be considered as ground truth, since polygons of individual crowns were manually delineated in the field by field experts and drawn directly on the remote sensing images.

The input data consists of two types of data: a shape file containing ITC data for 70% of the plots and remote sensing data for all plots, including one of the following for each plot: Hyperspectral (geotiff format), LIDAR (geotiff and csv), and RGB (geotiff). File naming will be of the form itc_train_*plot*.shp, hyper_*plot*.tiff, chm_*plot*.tiff, ptcloud_*plot*.csv, rgb_*plot*.tiff where *plot* indications the numeric code of the plot of each plot.

- In each ITC shape file, each attribute will be an individual crown and its associated crown ID number. Each “shape file” is composed of four files (.shp, .shx, .dbf, .prj), which can be read with widely-available packages (including most GIS software and GDAL).
- Each of the geotiff files contain the values for each band, the dimensions of the 2D or 3D space, and the spatial location information.
    - Hyperspectral - 426 bands on an 80 x 80 pixel lattice at 1 m^2 resolution
    - CHM (canopy height model)- 1 band on an 80 x 80 pixel lattice at 1 m^2 resolution
    - RGB - 3 bands on a 320 x 320 pixel lattice at 0.0625 m^2 resolution
- The point cloud data will be provided as a 3 column csv file with the x, y, and z coordinates of of each LIDAR return/point.
    - *x*: utm coordinates *m*
    - *y*: utm coordinates in *m*
    - *z*: height in *m* above sea level

3.3 Test Data

Data provided for testing will be the hyperspectral, LiDAR, and RGB data described above for the plots not provided in the training dataset. ITC data will not be provided for the test set, as the goal is for participants to predict its contents from the remote sensing data.

3.4 Submission Data

The submission files will include one shape file (including all 4 subfiles) for each test plot named as itc_subm_*plot*.shp.

- In the ITC shape file, each attribute will be an individual crown and a unique crown ID.

3.5 Performance Metrics

Participants’ performance on each plot will be calculated as the mean pairwise Jaccard Coefficient, J(A,B), calculated for each predicted polygon with the corresponding ground truth polygon. When there is ambiguity with regard to which submitted polygon corresponds to which ITC, we will choose the mapping that gives the best score to the participant.

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}.$$

The final score is the average of the plot level scores, which are themselves the average scores of the polygons within each plot..

This method is simple, does not require assignment of predicted crowns to specific ITCs by the participants, and has good continuity since the Jaccard coefficient has good continuity.

The output for the crown delineation algorithm is expected to produce polygons with no overlap, but if there are cases of overlapping crowns in the submitted polygons, we would be disambiguating by disregarding the intersection in the numerator of the Jaccard Coefficient. Thus, it is in the participant's best interest to clean up areas of intersection in a manner that best fits their expectations.

4. Alignment

Alignment is a common Data Science problem that involves associating two or more representations of the same object. One common instance of this problem is identifying the same object in data collected from different instruments or data sources.

Summary: Alignment

Inputs: ITC data, Ground data. *Optional:* Hyperspectral images, RGB images, LiDAR data

Output: Labels relating the ITC crown ID to the Ground data stem ID for each tree

Training Files: itc_train_*plot*.shp, ground_*plot*.csv, itc_ground_relation_train_*plot*.csv, *Optional:* hyper_*plot*.tiff, chm_*plot*.tiff, ptcloud_*plot*.csv, rgb_*plot*.tiff

Test Files: itc_test_*plot*.shp, ground_*plot*.csv, *Optional:* hyper_*plot*.tiff, chm_*plot*.tiff, ptcloud_*plot*.csv, rgb_*plot*.tiff

Submission Files: itc_ground_relation_subm_*plot*.csv

4.1 Alignment in the ecology domain

To build models relating remote sensing to field data, it is necessary to align the ground data to the remote sensing images. The three remote sensing datasets have been georeferenced and pre-aligned with one another by NEON. The NEON Ground data are georeferenced separately from the remote sensing data using a combination of GPS locations for fixed points and surveying methods to relate tree locations to those fixed points. Both the remote sensing and ground data coordinates involve measurement error, with the error likely being greater in the ground data. In addition, the ground data provides a location for the trunk of the tree, which is not directly visible in the remote sensing data. Furthermore, the trunk is not always located at the center of the crown because crowns are positioned in part to maximize light capture and avoid other tree crowns. As a consequence, apparent positions of trees can differ between the two the ground and remote sensing data sources with offsets of up to several meters even when no other errors have occurred with geo-referencing. Such offsets make it more difficult to develop models that related the ground and remote sensing data.

The goal of the task is to correctly align each tree crown polygon to a single tree from the ground data, thus allowing data collected on the ground (e.g. species identity, height, stem diameter) to be accurately associated with remote sensing data. This can be accomplished by considering the position and size of stems and crowns. Other information from the remote sensing images may also be helpful.

In real world situations this task will typically be conducted in an unsupervised manner matching segmented crowns to ground data using tree location and size without the use of ITC data for either training or evaluation. In this DSE we have collected ITC data to allow us to: 1) make this task independent of the segmentation task by providing the ITC segmentation as a near perfect segmentation of the image; 2) evaluate the performance of the algorithms against the known mapping between the ITC crowns and the ground data stems; and 3) provide data that allows for supervised rather than unsupervised algorithm development.

4.2 Training Data

Input data will be for a subset of plots and will include ITC polygon data, NEON ground data (including the stem ID, location of the stem, diameter of the stem, and some of the plots the maximum crown radius and the radius perpendicular to the axis of the maximum radius). Only ITC polygon data for which ground data is available will be included. The stem diameter of trees is positively related to the crown diameter, which should allow the size of the crown to be estimated from the diameter of the stem in cases where the crown diameter is not measured directly. The remote sensing data (Hyperspectral, RGB, and LiDAR) will also be available for optional use.

The primary input data consists of the ground and ITC data for 70% of the plots and will include: 1) one csv file per plot (ground_*plot*.csv) containing information on the location, stem size, crown size, and stem ID of each tree; 2) one shapefile per plot (itc_train_*plot*.shp) with polygons for each individual tree crown (ITC); and 3) one csv file per plot containing the stem ID and associated crown ID for each tree. The hyperspectral, canopy height, and camera data described in 3.2 will also be provided for optional use. The goal is to match stem ID from the ground data to the crown ID for the ITC tree crowns.

- Each ground plot csv file contains these comma separated fields
    - *stem_id*: a unique for each stem
    - *x*: utm coordinates in *m*
    - *y*: utm coordinates in *m*
    - *diameter*: diameter of the stem in *cm* at a standard height above the ground (~1.3 m)
    - *crown_radius_max*: maximum radius of the crown in *m*
    - *crown_radius_perp*: radius of the crown in *m* in the direction perpendicular to the maximum radius
- In each ITC shape file each attribute will be an individual crown and its associated crown ID stored as `crown_id`. Each “shape file” is composed of four files (.shp, .shx, .dbf, .prj).
- Each ground-itc relation file contains these comma separated fields
    - *stem_id*: the stem ID from the ground data
    - *crown_id*: the matching crown ID from the ITC data

4.3 Test Data

Data provided for testing will the ground and ITC data (and optionally the remote sensing data) for the plots not provided in the training dataset. Relationship data will not be provided for the test set.

4.4 Submission Data

The submission files will include one csv file for each test plot named itc_ground_relation_subm_*plot*.csv. Each file must contain these comma separated fields

- *stem_id*: the stem ID from the ground data
- *crown_id*: the matching crown ID from the ITC data

4.5 Performance Metrics

Performance of the pairing of field stems to ITC crowns will be evaluated using a scoring function below. In the testing stage, suppose we have a set of probe data (ITC) denoted as $\{p_n|n=1,..,N\}$, and ground truth data denoted as $\{g_n|n=1,..,N\}$,. We know in advance that there is a unique one-to-one mapping between P and G sets. Without loss of generality, let's assume p_n should be mapped to g_n for $n=1,..,N$. For each probe data point p_i, a program predicts a non-negative condence score that should be aligned with ground truth data point g_j, which forms a prediction matrix $M = (m_{i,j})$ where $i,j = 1,...,N$. Then, the quality of prediction can be measured by the following scoring function:

$$score = \frac{trace(M)}{\Sigma_{i,j} m_{i,j}}$$

where *trace(•)* represents trace of a matrix and M represents the prediction matrix which has been aligned in the order which matches the ground truth.

5 Classification

Classification is a common Data Science problem that involves determining which of a set of categories an object belongs to.

Summary: Classification

Inputs: ITC data, Ground data with species and genus IDs, Hyperspectral images, RGB images, LiDAR data

Output: A probability that each tree in the ITC data belongs to each possible species

Training Files: species_id_train_*plot*.csv, itc_*plot*.shp, ground_*plot*.csv, itc_ground_relation_*plot*.csv, hyper_*plot*.tiff, chm_*plot*.tiff, ptcloud_*plot*.csv, rgb_*plot*.tiff

Test Files: species_id_test_*plot*.csv, itc_*plot*.shp, ground_*plot*.csv, hyper_*plot*.tiff, chm_*plot*.tiff, ptcloud_*plot*.csv, rgb_*plot*.tiff

Submission Files: species_id_subm.csv

5.1 Classification in the ecology domain

A large number of ecological, environmental, and conservation oriented questions depend on species identification. This includes efforts to conserve individual species, understand and maintain biodiversity, and incorporating the biosphere into global circulation models. Being able to describe the density and distribution of different species using remote sensing would allow these efforts to occur more readily and at larger scales than field sampling. The goal of this task is to classify trees in remote sensing data into species and genera (higher level groupings of species). In addition to its utility for the domain, this task represents a challenging version of the more general classification problem because it involves classifying different species with very similar spectral signatures and categorizing data where some categories (species) have only small samples in the training set (i.e. rare species).

This task involves using the Hyperspectral, LiDAR, and RGB data to build classifiers for the ground based species identifications. To make this task independent of the segmentation and alignment task pre-aligned ITC data will be provided for all trees to be classified. Participants will determine probability that each tree with an ITC polygon belongs to each species and genus. Spectral reflectance signatures are the most commonly used data type for building species classifiers. Since classification is at the level of the crown, any pixel-level classification models should be upscaled to the crown.

5.2 Training Data

Input data will be for 70% of trees and will include the ITC crown data, species identities for each crown, and the hyperspectral remote sensing data. It will include: 1) one shapefile per plot (itc_*plot*.shp) with polygons for each individual tree crown (ITC); 2) one geotiff file per plot with the Hyperspectral remote sensing data (hyperspectral_*plot*.tiff) and 3) one csv file per plot containing the crown ID and associated genus and species for each tree. The ground data on stem location, stem size, and crown size, and the LiDAR and RGB remote sensing data will also be provided for optional use. The goal is to match species and genus from the ground data to the crown ID for the ITC tree crowns.

- In each ITC shape file each attribute will be an individual crown and its associated crown ID. Each "shape file" is composed of four files (.shp, .shx, .dbf, .prj).
- Each species ID file contains these comma separated fields
  - *crown_id*: the matching crown ID from the ITC data
  - *genus*: the genus of the tree
  - *species*: the species of the tree
  - *genus_id*: a unique two letter code for each genus
  - *species_id*: a unique four letter code for each species
- Each of the geotiff files contain the values for each band, the dimensions of the 2D or 3D space, and the spatial location information.
  - Hyperspectral - 426 bands on an 80 x 80 pixel lattice at 1 m^2 resolution
  - CHM - 1 band on an 80 x 80 pixel lattice at 1 m^2 resolution
  - RGB - 3 bands on a 320 x 320 pixel lattice at 0.0625 m^2 resolution

5.3 Test Data

Test data will include non-training ITC data along with associated remote sensing data.

5.4 Submission Data

Submissions will be two csv files named species_id_subm.csv and genus_id_subm.csv containing information on the crown ID and the species and genus identification probabilities for each crown. These files must contain these comma separated fields

- *crown_id*: the matching crown ID from the ITC data
- *ID*: the predicted genus or species ID
- *probability*: the probability that the crown belongs to the associated genus or species. The probabilities for a given *crown_id* will be normalized to sum to 1 if the submitted values do not already.

The files should contain one row for each possible *crown_id* and *genus/species* combination(i.e., the number of rows should be equal to the number of *crown_id*'s in the test set times the number of unique *ID*s (either species or genus) present in the training set). The results should be sorted in ascending order by *crown_id* and then *ID*. The dictionary of genus/species classes and their unique IDs will be provided. The performance metrics will be computed separately for genus and species classes.

5.5 Performance Metrics

We will evaluate performance using two metrics. The first one, rank-1 accuracy, is the fraction of crowns in the test set whose ground truth *species_id* (or *genus_id*) was assigned the highest probability by the participant. It is calculated as mean(argmax_*k*(*p_{ik}*)==*g_i*), where *g_i* is the ground-truth class of crown *i*, and *p_{ik}* is the probability assigned by the participant that crown *i* belongs to class *k* This metric only considers whether the correct class has the highest probability, not whether the probabilities are well-calibrated..

The second metric, average cross-entropy, is defined as

$$\text{cost} = \frac{-\sum_{n,k} ln(p_{nk}) * \delta(g_n, k)}{N}$$

given that $p_{nk} \neq 0$, to avoid the singularity. The $\delta(x, y)$ is an indicator function that takes value 1 when $x = y$. This metric rewards participants for submitting well-calibrated probabilities that reflect their uncertainty about which crowns belong to which class.

Performance metrics will be evaluated on the two different class labels (genus and species) separately resulting in four separate scores.

6. System Descriptions

For each submission to a task, a brief description of each system is requested. Each system description should provide:

1. *System Name.* The name of the system.
2. *Task.* The task(s) the system(s) was used for: cleaning, alignment, prediction, or forecasting.
3. *Team Affiliation.* The names of the submitting team.
4. *System Summary.* A brief summary the system in a few sentences (at most a paragraph).
5. *Algorithmic Information.* A brief description of additional details about the methods and algorithms used (1-2 paragraphs).
6. *Data Sources.* A brief description of the data sources used. This data includes the training data sources as well as additional data fed to the algorithms. Mention which of the provided data sets were used as well as whether there were any external data sets used. Cite any external sources.
7. *Development Data.* (optional) a brief description of any development sets that were generated and any brief results.

7. Schedule

The key dates for the NIST evaluation will follow the scheme summarized in Table 1.

Registration Opens	July 3, 2017
Register to the evaluation by	July 30, 2017

Release of NEON AOP data, and training ITC set	July 31st, 2017
Release of the rest of the test data and NEON Field data. First task submission Deadline	October 30th, 2017
Second task submission deadline	December 3rd, 2017
Release of the rest of the test NEON Field data	December 10th, 2017
Third task submission deadline	December 17th, 2017
Release of initial results	January 14th, 2018
Publication of Reports from the Participants and Evaluation	March 2018

8. Rules

- Code must be Free Open Source Software (FOSS) i.e. No Tools
- No. of Submissions will be limited to 2 / week
- One account per team
- No Merging of teams after the fact
- No limits on team size
- Decision of UF/NIST is final

Appendix: terminology glossary