NEON USER GUIDE TO DIGITAL HEMISPHERIC PHOTOS OF PLOT VEGETATION (DP1.10017.001)

PREPARED BY	ORGANIZATION
Courtney Meier	SCI

CHANGE RECORD

REVISION	DATE	DESCRIPTION OF CHANGE
A	02/09/2018	Initial Release
B	06/06/2019	Revision to reflect 2018-19 protocol optimization
C	05/26/2020	Included general statement about usage of neonUtilities R package and statement about possible location changes.
D	04/11/2022	Updated section 4.3 Data Revision with latest information regarding data release. Updated information regarding the geoNEON package
E	08/15/2022	Updated section 3.9.1 with new eventID format
E.1	10/21/2024	Section 3.1: Updated Figure 1 to latest version. Section 3.9: Corrected misnamed 'dhp_perplot' data table, added example of programmatic analysis workflow.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

1 DESCRIPTION

1.1 Purpose

This document provides an overview of the data included in this NEON Level 1 data product, the quality controlled product generated from raw Level 0 data, and associated metadata. In the NEON data products framework, the raw data collected in the field, for example, Digital Hemispheric Photos (DHPs) and associated field metadata, are considered the lowest level (Level 0). Raw data that have been quality checked via the steps detailed herein, as well as simple metrics that emerge from the raw data are considered Level 1 data products.

The text herein provides a discussion of measurement theory and implementation, data product provenance, quality assurance and control methods used, and approximations and/or assumptions made during L1 data creation.

1.2 Scope

This document describes the steps needed to generate the L1 data product Digital Hemispheric Photos of Plot Vegetation (DP1.10017.001), and associated metadata, from input data. This document also provides details relevant to the publication of the data products via the NEON data portal, with additional detail available in the file, NEON Data Publication Workbook for TOS Digital Hemispherical Photos (DP1.10017.001) (AD[05]), provided in the download package for this data product.

This document describes the process for ingesting and performing automated quality assurance and control procedures on the data collected in the field pertaining to TOS Protocol and Procedure: Measurement of Leaf Area Index (AD[07]). The raw data that are processed in this document are detailed in the file, NEON Raw Data Ingest Workbook for TOS Digital Hemispherical Photos (DP0.10017.001) (AD[04]), provided in the download package for this data product. Please note that raw data products (denoted by 'DP0') may not always have the same numbers (e.g., '10017') as the corresponding L1 data product.

2   RELATED DOCUMENTS AND ACRONYMS

2.1   Associated Documents

AD[01]	NEON.DOC.000001	NEON Observatory Design (NOD) Requirements
AD[02]	NEON.DOC.000913	TOS Science Design for Spatial Sampling
AD[03]	NEON.DOC.002652	NEON Data Products Catalog
AD[04]	Available with data download	Validation csv
AD[05]	Available with data download	Variables csv
AD[06]	NEON.DOC.000914	TOS Science Design for Plant Biomass, Productivity, and Leaf Area Index
AD[07]	NEON.DOC.014039	TOS Protocol and Procedure: Measurement of Leaf Area Index
AD[08]	NEON.DOC.000913	TOS Science Design for Spatial Sampling
AD[09]	NEON.DOC.000008	NEON Acronym List
AD[10]	NEON.DOC.000243	NEON Glossary of Terms
AD[11]	NEON.DOC.004825	NEON Algorithm Theoretical Basis Document: OS Generic Transitions
AD[12]	Available on NEON data portal	NEON Ingest Conversion Language Function Library
AD[13]	Available on NEON data portal	NEON Ingest Conversion Language
AD[14]	Available with data download	Categorical Codes csv

2.2   Acronyms

Acronym	Definition
DHP	Digital Hemispheric Photo
DSLR	Digital Single Lens Reflex camera
LAI	Leaf Area Index
NLCD	National Land Cover Database
PAI	Plant Area Index

3   DATA PRODUCT DESCRIPTION

The Digital Hemispheric Photos of Plot Vegetation data product (DP1.10017.001) provides 180 degree images and field-collected image metadata, which enable ground-based estimation of Leaf Area Index (LAI) and/or Plant Area Index (PAI) at the plot scale. Leaf Area Index is the one-sided leaf area per unit ground area, and Plant Area Index is similar to LAI but also includes twig and branch area. Photos are acquired with a full-frame DSLR camera equipped with a fisheye lens, and are provided in RAW image format (specifically .NEF format). Within each plot, photos of understory vegetation are acquired, with 'understory' defined as vegetation typically less than 2 meters height. Depending on vegetation density and height, understory photos may be acquired with either a downward-facing or upward-facing camera orientation relative to the ground. In grasslands and short-stature scrublands, the understory is therefore the only vegetation photographed. When an overstory is present, i.e. vegetation greater than 2 m height, upward-facing photographs of overstory are also acquired. Upward-facing understory and overstory images are typically acquired in the crepuscular hours of the day, when direct sunlight is not illuminating the foliage, and may also be acquired when it is overcast such that the sun does not cast a shadow. Downward-facing understory images are typically acquired during daylight hours, to better enable differentiation of plant from non-plant material during image processing.

Leaf area index is a useful proxy variable for numerous other variables of ecological interest including plant biomass, plant productivity, forage quality, carbon balance, ecosystem energy flux, plant density, and the heterogeneity of plant cover. LAI is also used widely as a key input variable to models that seek to predict ecological processes such as carbon cycling. Together with LAI data generated from NEON's airborne remote-sensing platform, LAI data derived from ground-collected Digital Hemispherical Photos can contribute to continental scale estimates of LAI.

3.1   Spatial Sampling Design

Digital Hemispheric Photos of plot vegetation are collected at all NEON terrestrial sites. At most sites, DHPs are collected from a subset of n=20 Distributed Plots. Distributed Plots are randomly positioned within each National Land Cover Database (NLCD) class, with representation within each NLCD class set as proportional to its representation at the site; NLCD classes with less than 5% representation are excluded from sampling. Additionally, Distributed Plots must be separated by a minimum of 55 m, plot centers must be > 50 m from large paved roads and buildings, plot edges must be > 10 m from dirt roads and plots may not be intersected by streams > 1 m wide (see AD[08] for more details). NEON also collects DHPs from n=3 Tower Plots at each site. NEON Field Operations staff are responsible for selecting which 3 Tower plots are sampled, based on logistic feasibility and sensitivity of the vegetation to repeat sampling (Figure 1).
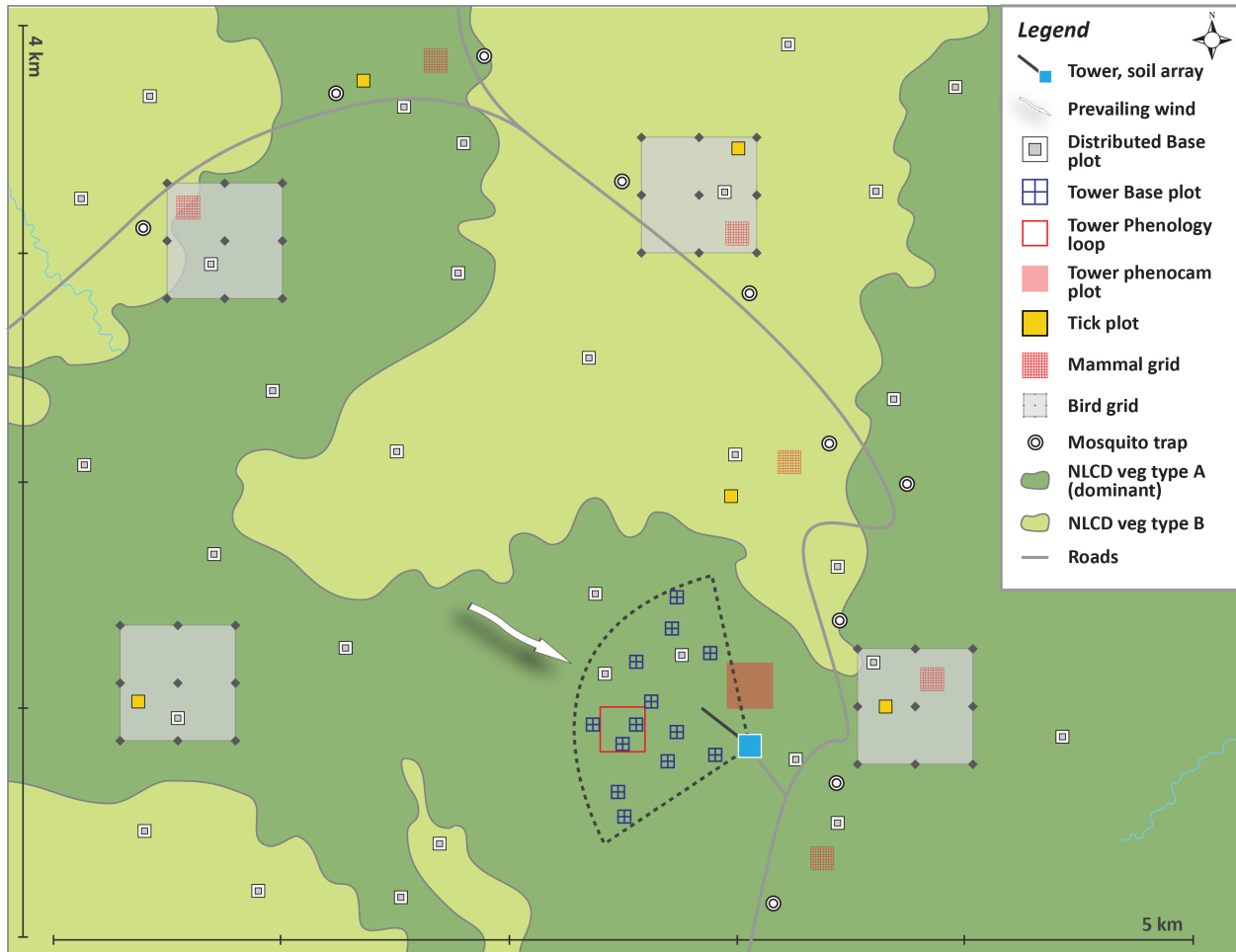
Figure 1: Representation of a NEON site with Distributed and Tower Plots; note that for DHP sampling, the actual Distributed Plot number (n=20, max) and Tower Plot number sampled (n=3) is a subset of those depicted in the schematic.

Within each plot, DHPs are collected from photo points that are arranged according to the shape of a square cross, oriented in alignment with the cardinal axes, and with points spaced every 4 meters (Figure 2). Cross-shaped and square-shaped arrangements of points produce LAI estimates with the lowest coefficients of variation (Majasalmi *et al.* 2012), and cross-shaped arrangements of points are also employed by the European VALERI project to validate satellite derived LAI products (Garrigues *et al.* 2002). Regardless of plot type (Distributed vs. Tower), plot size (20m x 20m or 40m x 40m) or vegetation stature, the spacing of NEON photo points is consistently 4 m along the cardinal axes, and the cross layout of points is always centered on the plot centroid. When acquiring DHPs, technicians extend the camera approximately 1 m perpendicular to the cardinal axis at a given photo point, and the exact photo location is chosen to maintain a minimum distance of 50 cm between the lens and any foliage elements.
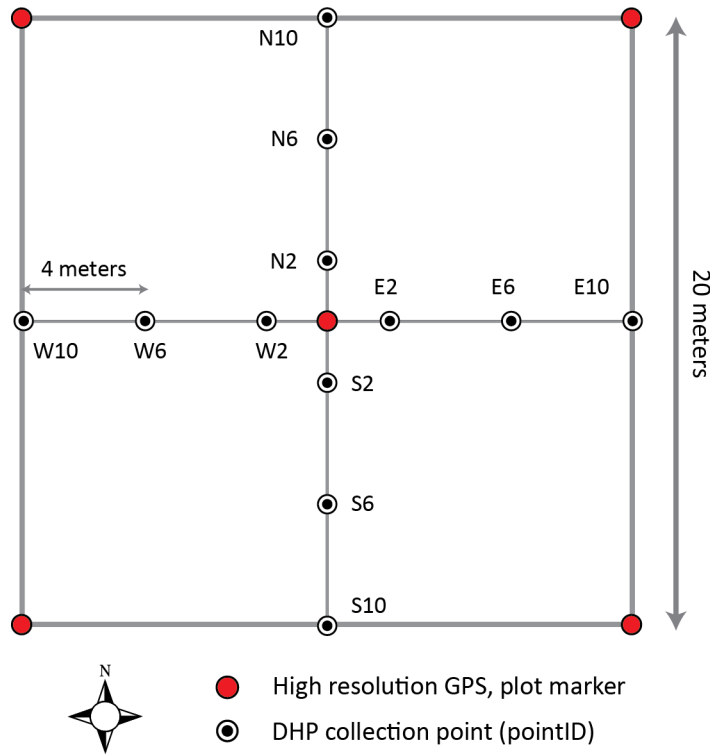
Figure 2: NEON Distributed Plot showing the location and spacing of DHP collection points relative to plot center. Spacing and number of points is the same in NEON Tower Plots, regardless of Tower Plot size.

As much as possible, sampling occurs in the same locations over the lifetime of the Observatory. However, over time some sampling locations may become impossible to sample, due to disturbance or other local changes. When this occurs, the location and its location ID are retired. A location may also shift to slightly different coordinates. Refer to the locations endpoint of the NEON API for details about locations that have been moved or retired: https://data.neonscience.org/data-api/endpoints/locations/

3.2   Temporal Sampling Design

Distributed Plots are sampled every 5 years, and within an 'on' year, DHP collection occurs within a one month window targeted to coincide with a NEON AOP remote-sensing flight over the site. Ground collection of DHPs and the AOP flight are both timed to co-occur with the window of peak greenness, according to a 10 y average of MODIS-EVI phenology index timecourse data at each site. Tower Plots are sampled annually on a 2 week bout interval, from leaf-out to senescence.

3.3   Sampling Design Changes

- 2016-2018: During the first 3 sampling years, distributed plots were scheduled on a 3-year sampling rotation, rather than 5 years.

3.4   Variables Reported

All variables reported from the field or laboratory technician (L0 data) are listed in the file, NEON Raw Data Ingest Workbook for TOS Digital Hemispherical Photos (DP0.10017.001) (AD[04]). All variables reported in the published data (L1 data) are also provided separately in the file, NEON Data Publication Workbook for TOS Digital Hemispherical Photos (DP1.10017.001) (AD[05]).

Field names have been standardized with Darwin Core terms (http://rs.tdwg.org/dwc/; accessed 16 February 2014), the Global Biodiversity Information Facility vocabularies (http://rs.gbif.org/vocabulary /gbif/; accessed 16 February 2014), the VegCore data dictionary (https://projects.nceas.ucsb.edu/nc eas/projects/bien/wiki/VegCore; accessed 16 February 2014), where applicable. NEON TOS spatial data employs the World Geodetic System 1984 (WGS84) for its fundamental reference datum and GEOID09 for its reference gravitational ellipsoid. Latitudes and longitudes are denoted in decimal notation to six decimal places, with longitudes indicated as negative west of the Greenwich meridian.

Some variables described in this document may be for NEON internal use only and will not appear in downloaded data.

3.5   Spatial Resolution and Extent

The finest spatial resolution at which Digital Hemispheric Photos of Plot Vegetation will be tracked is that of the **pointID** within a given **plotID** (Figure 2), resulting in the following spatial hierarchy from finest to coarsest resolution:

> **pointID** (ID of point within plot) → **plotID** (ID of plot within site) → **siteID** (ID of NEON site) → **domainID** (ID of NEON domain)

The basic spatial data included in the data download include the latitude, longitude, and elevation of the centroid of the plot where sampling occurred, plus associated uncertainty due to GPS error and plot width. Shapefiles of all NEON Terrestrial Observation System sampling locations can be found here: NEON Terrestrial Field Sites.

To derive a more precise estimate at which each DHP was acquired within a plot, there are two options:

1. Use the getLocTOS function from the geoNEON package, available here: https://github.com/NEONS cience/NEON-geolocation, or
2. Follow these steps to perform the same calculation:
    a. Add the appropriate lat/long or easting/northing offset associated with the pointID to the coordinates of the plot centroid based on the (Table 2).

Table 1: Offset of pointIDs from the plot centroid

pointID	Easting offset (m)	Northing offset (m)
N2	0	2
N6	0	6
N10	0	10
E2	2	0
E6	6	0
E10	10	0
S2	0	-2
S6	0	-6
S10	0	-10
W2	-2	0
W6	-6	0
W10	-10	0

b. Increase coordinateUncertainty by an appropriate amount (suggested amount of 2 m) to account for the fact that:
   i. Technicians typically place the camera 1 m perpendicular to the cardinal axis, in either direction, at a given pointID; and
   ii. Technicians use calibrated pacing to identify pointID locations within the plot relative to the plot centroid.

3.6 Temporal Resolution and Extent

The finest resolution at which temporal data are recorded is the date and time at which DHP collection commenced (**startDate**) and was completed (**endDate**) within a plot for a given sampling bout (**eventID**). The temporal hierarchy within the data is therefore:

date (date a plot was photographed) → **eventID** (bout-level identifier)

The exact time of image acquisition is also recorded by the DSLR and embedded in Exif data for each image, but NEON cannot guarantee that each camera's clock was accurately and consistently set to local time at the time of image acquisition.

The NEON Data Portal provides data in monthly files for query and download efficiency. Queries including any part of a month will return data from the entire month.

3.7 Associated Data Streams

Measurements of LAI from NEON's remote sensing platform can be found in the LAI - spectrometer - flightline (DP2.30012.001) and the LAI - spectrometer - mosaic (DP3.30012.001) data products.

3.8 Product Instances

For Distributed Plots, DHPs are collected one bout per site x 9-10 sites per year x 20 plots per site x 12 points per plot. Approximately 2/3 of NEON sites are forested, and at these sites both understory and overstory DHPs are collected at each point. The remaining 1/3 of NEON sites are short-stature vegetation, and only support collection of understory images. This sampling scheme generates approximately 3600 DHPs from Distributed Plots per year across the entire Observatory.

For Tower Plots, DHPs are collected annually on a 2 week interval from leaf-out to senescence, resulting in approximately 14 bouts per site per year x 3 plots per site x 12 points per plot. Assuming Tower Plots at a given site are similar to Distributed Plots in terms of supporting collection of understory and overstory images, this sampling scheme generates approximately 39,816 DHPs from Tower Plots per year across the entire Observatory.

In some years, sampling frequency and density may have differed from these numbers. See Sampling Design Changes section above for details.

3.9 Data Relationships

A unique **sampleID** is generated in the 'dhp_perplot' table each time DHPs are collected from a plot within a bout. Because 12 images per plot per bout (understory only) or 24 images per plot per bout (understory + overstory) may be collected, we expect either 12 or 24 child records with a unique **subsampleID** in the 'dhp_perimagefile' table for each unique **sampleID**. There may be fewer than the expected number of DHPs per plot per bout if all images did not meet protocol requirements, or could not be matched with field-collected metadata. To link each DHP in 'dhp_perimagefile' with data collected at the plot level, users should join 'dhp_perplot' and 'dhp_perimagefile' on **sampleID**. Duplicates and/or missing data may exist where protocol and/or data entry abberations have occurred; users should check data carefully for anomalies before joining tables.

dhp_perplot → One record (**sampleID**) expected per bout (**eventID**)

dhp_perimagefile → Either 12 or 24 records (**subsampleIDs**) expected per **sampleID**

Data downloaded from the NEON Data Portal are provided in separate data files for each site and month requested. The neonUtilities R package contains functions to merge these files across sites and months into a single file for each table described above. The neonUtilities package is available from the Comprehensive R Archive Network (CRAN; https://cran.r-project.org/web/packages/neonUtilities/index.html) and can be installed using the install.packages() function in R. For instructions on using neonUtilities to merge NEON data files, see the Download and Explore NEON Data tutorial on the NEON website: https://www.neonscience.org/download-explore-neon-data

3.9.1 Understanding 'eventID'

The eventID includes the year, week bout began, siteID, and plot type (TOWER for Tower plots collected on a 2 week bout interval, or DISTRIBUTED for Distributed plots collected every 5 years). The 'week bout began' portion of the eventID is automatically assigned to be the week of the year that the bout began.

For example: 'dhp.2022.16.BART.TOWER' identifies DHPs collected from Tower plots during the 16th week of 2022 at the NEON D01 BART site.

Prior to March 1, 2022, DHPs collected from Distributed plots had a bout number of '00', and DHPs collected from Tower plots had a bout number that incremented by one for each bout within a year, starting at '01'.

3.9.2   Downloading DHPs

There are two mechanisms end-users may employ to download DHPs:

1. In the 'basic' download package, links to the cloud-storage location for each DHP are provided in the **imageFileUrl** field.
    a. From the 'dhp_perimagefile' table, determine which DHPs are desired.
    b. Use a script to programmatically read and save desired images from each **imageFileUrl** location.
    c. DHPs downloaded directly via the **imageFileUrl** link are named according to **imageFileName**.
2. In the 'expanded' download package, DHPs are provided in a .zip file for each selected record. Please note that:
    a. Packaged .zip files will likely be quite large, and require significant time to download.
    b. To facilitate parsing of downloaded files, DHPs are re-named when downloaded via this mechanism to the value specified in 'dhp_perimagefile' **downloadFileName**.
    c. File names specified in **imageFileName** are no longer relevant.

3.9.3   DHP Processing

All NEON DHPs require additional processing to generate Leaf Area Index or Plant Area Index data. DHPs typically appear very dark because exposure in the field is optimized to avoid blowing out the highlights, which can result in significant underestimation of LAI in upward-facing images. As such, processing steps may include exposure optimization in the 'shadow' and 'black' region of the histogram, followed by conversion to another image format (e.g., JPG). To estimate LAI, optimized DHPs are typically classified by pixel into 'plant' or 'leaf', and 'non-plant' material. This step may be accomplished either programmatically or manually via specialty software.

- Example of programmatic classification software: HemiPy (Brown *et al.* 2023).
- Example software to optimize Nikon RAW image exposure and convert to JPG: Nikon ViewNX 2
- Example software to analyze DHPs for LAI: Can-Eye

4   DATA QUALITY

4.1   Data Entry Constraint and Validation

Many quality control measures are implemented at the point of data entry within a mobile data entry application or web user interface (UI). For example, data formats are constrained and data values controlled through the provision of dropdown options, which reduces the number of processing steps necessary to prepare the raw data for publication. An additional set of constraints are implemented during the process

of ingest into the NEON database. The product-specific data constraint and validation requirements built into data entry applications and database ingest are described in the document NEON Raw Data Ingest Workbook for TOS Digital Hemispherical Photos (DP0.10017.001), provided with every download of this data product. Contained within this file is a field named **entryValidationRulesForm**, which describes syntactically the validation rules for each field built into the data entry application. Data entry constraints are described in Nicl syntax in the validation file provided with every data download, and the Nicl language is described in NEON's Ingest Conversion Language (NICL) specifications (AD[12]).

Data collected prior to 2017 were processed using a paper-based workflow that did not implement the full suite of quality control features associated with the interactive digital workflow.

4.2   Automated Data Processing Steps

Following data entry into a mobile application of web user interface, the steps used to process the data through to publication on the NEON Data Portal are detailed in the NEON Algorithm Theoretical Basis Document: OS Generic Transitions (AD[11]).

4.3   Data Revision

All data are provisional until a numbered version is released. Annually, NEON releases a static version of all or almost all data products, annotated with digital object identifiers (DOIs). The first data Release was made in 2021. During the provisional period, QA/QC is an active process, as opposed to a discrete activity performed once, and records are updated on a rolling basis as a result of scheduled tests or feedback from data users. The Issue Log section of the data product landing page contains a history of major known errors and revisions.

4.4   Quality Flagging

The **dataQF** field in each data record is a quality flag for known errors applying to the record. Please see below for an explanation of **dataQF** codes specific to this product.

Table 2: Descriptions of the dataQF codes for quality flagging

fieldName	value	definition
dataQF	legacyData	Data recorded using a paper-based workflow that did not implement the full suite of quality control features associated with the interactive digital workflow

Records of land management activities, disturbances, and other incidents of ecological note that may have a potential impact are found in the Site Management and Event Reporting data product (DP1.10111.001)

5   REFERENCES

Brown, L.A., Morris, H., Leblanc, S., Bai, G., Lanconelli, C., Gobron, N., Meier, C. and J. Dash. 2023. HemiPy: A Python module for automated estimation of forest biophysical variables and uncertainties from digital hemispherical photographs. Methods in Ecology and Evolution. Vol 14, Issue 9, pp. 2329-2340, DOI: https://doi.org/10.1111/2041-210X.14199

Garrigues, S., Allard, D., Weiss, M. and F. Baret. 2002. Comparing VALERI sampling schemes to better represent high spatial resolution statellite pixel from ground measurements: How to characterize an ESU. Institut National de la Recherche Agronomique, pp. 1-5, http://w3.avignon.inra.fr/valeri/

Majasalmi, T., Rautiainen, M., Stenberg, P. and H. Rita. 2012. Optimizing the sampling scheme for LAI-2000 measurements in a boreal forest. Agricultural and Forest Meteorology. Vol 154-155, pp. 38-43, DOI: https://doi.org/10.1016/j.agrformet.2011.10.002