NEON USER GUIDE TO BREEDING LANDBIRD POINT COUNTS (NEON.DP1.10003)

PREPARED BY	ORGANIZATION	DATE
Kate Thibault	FSU	07/18/2017
Sarah Elmendorf	DPS	07/18/2017

CHANGE RECORD

REVISION	DATE	DESCRIPTION OF CHANGE
A	06/29/2017	Initial Release

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, the dry weights of litter functional groups from a single collection event, 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 Breeding landbird point counts - counts, distance from observer, and taxonomic identification of breeding landbirds observed during point counts - 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 Variables for Breeding landbird point counts (NEON.DP1.10003) (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: Breeding Landbird Abundance and Diversity (AD[07]). The raw data that are processed in this document are detailed in the file, NEON Raw Data Validation for Breeding landbird point counts (NEON.DP0.10006) (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., '10033') 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 Level 1, Level 2 and Level 3 Data Products Catalog
AD[04]	NEON.DP0.10006.001 _dataValidation.csv	NEON Raw Data Validation for Breeding landbird point counts (NEON.DP0.10006)
AD[05]	NEON.DP1.10003.001 _variables.csv	NEON Data Variables for Breeding landbird point counts (NEON.DP1.10003)
AD[06]	NEON.DOC.000916	TOS Science Design for Breeding Landbird Abundance and Diversity
AD[07]	NEON.DOC.014041	TOS Protocol and Procedure: Breeding Landbird Abundance and Diversity
AD[08]	NEON.DOC.000913	TOS Science Design for Spatial Sampling
AD[09]	NEON.DOC.000913	TOS Science Design for Plant Diversity
AD[10]	NEON.DOC.000008	NEON Acronym List
AD[11]	NEON.DOC.000243	NEON Glossary of Terms
AD[12]	OS_Generic_Transitions .pdf	NEON Algorithm Theoretical Basis Document: OS Generic Transitions
AD[13]		NEON's Ingest Conversion Language (NICL) specifications

3 DATA PRODUCT DESCRIPTION

The breeding landbird point counts product provides records of species identification, as well as metadata which can be used to model detectability, e.g. weather, distances from observers to birds, and detection methods. Breeding landbirds are defined as "smaller birds (usually exclusive of raptors and upland game birds) not usually associated with aquatic habitats" (Ralph et al. 1993). They are typically censused during the first half of the breeding season, when birds are "most active, paired, on territories, and vocal" (Ralph et al. 1993).

Breeding landbirds are sampled using the point count method. Point counting entails one or more observers going to pre-established points and recording all the birds heard and/or seen during a set period of time (Figure 1). The NEON point count method is adapted from the Integrated Monitoring in Bird Conservation Regions (IMBCR): Field protocol for spatially-balanced sampling of landbird populations. (Hanni et al. 2016; http://bit.ly/17ekDNB). Point counts are six minutes long, with each minute tracked by the observer, following a two-minute settling-in period. All birds are recorded to species and sex, whenever possible, and the distance to each individual or flock is measured with a laser rangefinder, except in the case of flyovers.

Breeding landbird point counts data may be used to inform estimates of abundance, density and diversity of bird populations at plot, site, and continental scales. They also can be used to inform models of mosquito-borne disease dynamics.

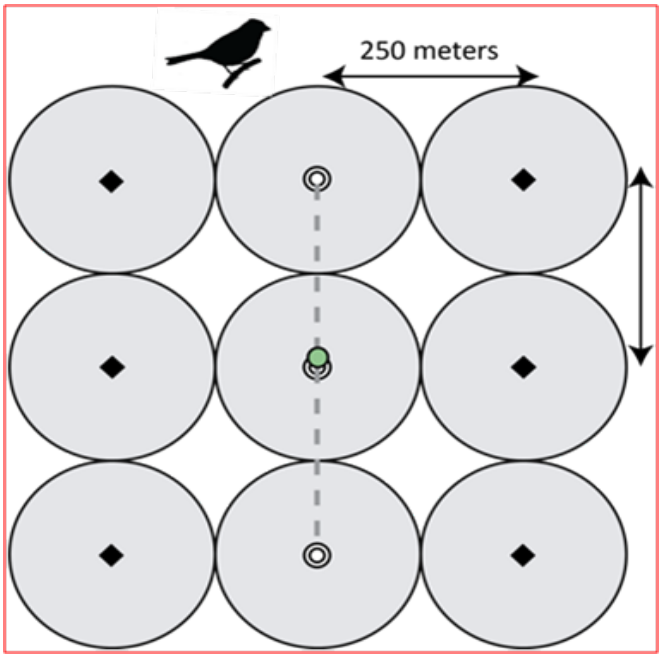

Figure 1: NEON plot (grid of points) for sampling breeding landbirds

3.1 Spatial Sampling Design

Breeding landbird point counts are conducted at terrestrial NEON sites. Depending on the size of the site, sampling for this product occurs either at either points or across grids.

At larger sites, point count sampling occurs across 5-15 9-point grids, with grid centers collocated with distributed plots centers, if possible. The 9-point grids are distributed within the sites proportional to permitted sampling area contained within each National Land Cover Database (NLCD) class, with the restriction that >50% of the points within each grid must fall within the target NLCD class. Grids may be intersected by roads and buildings, but are situated to avoid large bodies of water and to ensure that any point on the grid is at least 10m from any road. Any point on a grid that intersects smaller bodies of water is shifted to dry ground. Bird grids must be separated by a minimum distance of 250m.

At smaller sites (i.e., sites that cannot accommodate a minimum of 5 grids, using the criteria above), point counts occur at the southwest corner (point 21) of 5-25 distributed base plots that are distributed according to a stratified-random design proportional to permitted sampling area contained within each of the National Land Cover Database (NLCD) classes found throughout the entire site. Distributed base plots must be a minimum of 10m from roads and 50m from buildings. Small streams (<1m wide) can bisect plots. If the southwest corner of the plot is in the stream, observations are made from the closest dry location. Distributed base plots selected for bird point count sampling must be separated by a minimum of 250m.

The distance from each observed bird to the location of the observer (pointID) is recorded in the data as part of the protocol for visual detection of birds (observerDistance).

See TOS Science Design for Breeding Landbird Abundance and Diversity (AD[06]), TOS Protocol and Procedure: Breeding Landbird Abundance and Diversity (AD[07]), and TOS Science Design for Plant Diversity (AD[09]) for further details.

3.2 Temporal Sampling Design

Point counts are conducted at each point once or twice per year depending on site size, with grids at larger sites sampled only once per year, during an expert-identified window of breeding activity specific to the local area.

3.3 Variables Reported

All variables reported from the field technician (L0 data) are listed in the file, NEON Raw Data Validation for Breeding landbird point counts (NEON.DP0.10006) (AD[04]). All variables reported in the published data (L1 data) are also provided separately in the file, NEON Data Variables for Breeding landbird point counts (NEON.DP1.10003) (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/nceas/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 Earth Gravitational Model 96 (EGM96) 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.

3.4 Spatial Resolution and Extent

The finest resolution at which spatial data are reported is a single point (Figure 1).

pointID → **plotID** (ID of plot/grid within site) → **siteID** (ID of NEON site) → **domainID** (ID of a NEON domain.

The basic spatial data included in the data downloaded include the latitude, longitude, and elevation of the centroid of the plot/grid where sampling occurred + associated uncertainty due to GPS error and plot width. Shapefiles of all NEON Terrestrial Observation System sampling locations can be found in the Document Library: http://data.neonscience.org/documents.

To derive a more precise estimate of the location of each point, there are two options:

- Use the def.calc.geo.os function from the geoNEON package, available here: https://github.com/NEONScience/NEON-geolocation
- Or follow these steps to perform the same calculation:

1. The namedLocation field in the data is the named location of the plot; more precise geographic data require the named location of the point. **pointCountMinute**s of '88' reflect incidental observations of birds after the count period has ended or while traversing between plots. These observations are certain to have occurred at the sites, somewhere near the plot, but coordinate uncertainty is unknown. For all **pointCountMinute**s other than '88', construct the named location of the point of each record in brd_perpoint and/or brd_countdata by concatenating the fields for namedLocation and pointID as: namedLocation + '.' + pointID, e.g. subplotID 'A1' of namedLocation 'GRSM_013.birdGrid.brd' has a complete named location of 'GRSM_013.birdGrid.brd.A1'.
2. Use the API (http://data.neonscience.org/api; e.g. http://data.neonscience.org/api/v0/locations/GRSM_013.birdGrid.brd.A1) to query for elevation("locationElevation"), easting("locationUtmEasting"), northing("locationUtmNorthing"), coordinateUncertainty ("Value for Coordinate uncertainty"), elevationUncertainty ("Value for Elevation uncertainty"), and utmZone ("locationUtmZone") as inputs to the next step.
3. Increase coordinateUncertainty by an appropriate amount to account for error introduced by navigating within plots. Specifically, pointIDs of '21' and 'B2' are monumented; the realized coordinate uncertainty off the high resolution GPS units used in plot establishment are recorded in the coordinate uncertainty value in the spatial data. **pointCountMinute**s of '88' have an indeterminate amount of coordinate uncertainty. Other pointIDs are not monumented and are located using recreational-grade GPS units, with a realized uncertainty of ~15m.

3.5 Temporal Resolution and Extent

The finest temporal resolution at which point count data will be recorded is the minute at which an observation occurred within a six-minute sampling period, with start time and date specified. This resolution allows analyses that involve temporal subsetting of the sampling period, and also allows aggregation to coarser scales. An **eventID** is used to uniquely identify the point and time at which a point count took place, and is comprised of the plotID, pointID, and date of sampling.

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. Code to stack files across months is available here: https://github.com/NEONScience/NEON-utilities

3.6 Product Instances

There are a maximum of 15 9-point bird grids per site (large sites), and a minimum of 10 distributed points per site. Each point is surveyed 1-2x year, yielding a range of 20-135 6-minute point counts per site per year. The number of birds observed during each point count varies with site-specific detection probability and abundance.

3.7 Data Relationships

Each technician takes up to 3 bird quizzes over the course of the season, generating up to 3 records per technicianID per observerInstitutionName per year in brd_personnel.

Each field crew reports the source or sources used to identify birds (identificationReferences) per site per year. There is one record per site per year in brd_references, with multiple sources concatenated and separated by a semi-colon.

The protocol dictates that each pointID within a grid is counted once or twice per year (one or two expected records per pointID per plotID per year in brd_perpoint). Each combination of plotID*pointID*date generates a single eventID, which should uniquely identify each record of brd_perpoint. A record from brd_perpoint may have one or more child records in brd_countdata depending on the number of birds observed; records from these two tables can be joined on eventID.

Records from brd_countdata where pointID is '88' are not expected to have a corresponding record in brd_perpoint; observations with this pointID are incidental observations of rare or unusual birds outside of the formal point count locations and/or times.

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.

brd_personnel.csv - > One record expected per technicianID per observerInstitutionName per quiz per year.

brd_references.csv - > One record expected per siteID per year.

brd_perpoint.csv - > One record expected per date (startDate, date and time) per pointID per plotID per siteID.

brd_countdata.csv - > One record expected per date (startDate, date and time) per observation per plotID per siteID. Observations are recorded separately by taxonID, sexOrAge, clusterSize and observerDistance to facilitate distance sampling analyses.

4 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 Validation for Breeding landbird point counts (NEON.DP0.10006), 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[13]).

5 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[12]).

6 REFERENCES

Hanni, D. J., C. M. White, N. J. Van Lanen, J. J. Birek, J. M. Berven, and M. F. McLaren. 2016. Integrated Monitoring in Bird Conservation Regions (IMBCR): Field protocol for spatially-balanced sampling of landbird populations. Unpublished report. Bird Conservancy of the Rockies, Brighton, Colorado, USA.

Ralph, C. J., T. E. Martin, G. R. Geupel, D. F. Desante, and P. Pyle. 1993. Handbook of Field Methods for Monitoring Landbirds. Gen. Tech. Rep. PSW-GTR-144-www. Page 41. Pacific Southwest Research Station, Forest Service, U.S. Department of Agriculture, Albany, CA.