**Missouri Bladderpod Checklist for NPS Data Package Review**

04/11/2024

Intro text sharing the value of data packages and where to find more information and guidance:

<https://doimspp.sharepoint.com/sites/nps-nrss-imdiv/data-publication/SitePages/Data-Package-Specifications.aspx>

The role of the data and metadata reviewer is to ensure the accuracy, completeness, and usefulness of the entire data package. This process should enlist the help of a colleague who can step through this checklist. Note that this checklist provides general guidelines and can be customized, as necessary, for use in specific NPS programs, parks, or offices. A good reviewer should be knowledgeable about the metadata standard being used, and should be given ample time to review a metadata record. Starting the process early is key.

**Title**: Missouri Bladderpod WICR Data

**DataStore Reference**: https://doi.org/10.57830/2303034

**Author’s Special Instructions**:

**Reviewer’s Name**: Sarah Kelso

| **#** | **Yes** | **No** | **Item** |
| --- | --- | --- | --- |
| 1 |  |  | Many automated checks have been built into the [NPSdataverse DPchecker](https://nationalparkservice.github.io/DPchecker/) tool. Have these been run on the data package and warnings or errors been dealt with? DPchecker version 0.3.4  Comments: Several warnings which may or may not need fixing depending on what you meant to do  ! Metadata does not contain taxonomic coverage information.  ! The data package abstract contains non-standard end of line characters such as \r, \n, or \n\r.  Use `EMLeditor::set\_abstract()` to revise.  ^^everything looks to have printed correctly in Datastore |
| 2 |  |  | Is the data package comprised of only CSV data files and an XML metadata file (thus meeting machine-readable and open format goals)?  Comments: |
| 3 |  |  | Is the title appropriate and sufficiently descriptive (the title conveys the “what, where, when, and (if applicable) scale” of the data)?  Comments: It would be helpful to provide the timeline this study occurred/is occurring in the title itself |
| 4 |  |  | Does the abstract provide a short statement that allows users to quickly evaluate the content of the dataset?  Comments: |
| 5 |  |  | Are QA/QC methods appropriate and adequately described and/or referenced?  Comments: No QA/QC, but I don’t think we decided as a group if we need any here? |
| 6 |  |  | Are the sampling and data-collection methods adequately summarized and referenced such the data can be understood by a 3rd party?  Comments: See additional comments for tracked changes |
| 7 |  |  | Does the metadata contain definitions of the values in each field such that a 3rd party can understand them?  Comments: Count could be changed from “dimensionless” to “number”  Code definition for density class for this file: HTLN\_MoBlad\_CountData1997-2021\_Cleaned\_INTERNAL.csv has non-standard characters for code 7  Definition for Event for this file: HTLN\_MoBlad\_PARData\_Cleaned\_INTERNAL.csv has occurrence misspelled |
| 8 |  |  | If missing values are present (i.e. data are absent), are appropriate missing value codes included and defined in metadata?  Comments: |
| 9 |  |  | Do filenames provide context, are they reasonably clear, and do they use a consistent naming convention (or example, ‘MeasurementA\_Location1\_Date.csv’, ‘MeasurementA\_Location2\_Date.csv’, ‘MeasurementB\_Location1\_Date.csv’,)  Comments: |
| 10 |  |  | Do column labels provide context, can they be interpreted by humans, and are they consistently formatted?  Comments: |
| 11 |  |  | Does each column contain only one data type (excluding missing value codes) or one data format? For example, date formats should be consistent within a column; numeric data only contains numbers, etc.  Comments: |
| 12 |  |  | Across filenames and column names are capitalization and punctuation used consistently (and where applicable follow standard conventions)?  Comments: |
| 13 |  |  | Have unused columns and rows been removed? For example, .csv files do not have trailing commas at the end of a row.  Comments: |
| 14 |  |  | Do all URLs work and point to the expected web page?  Comments: |
| 15 |  |  | Do keywords accurately represent the data and include terms from standard vocabularies whenever possible?  Comments: |
| 16 |  |  | To the best of your knowledge, If the data package is intended for a public audience, does it have protected data that needs to be removed prior to publication?  Comments: |
| 17 |  |  | To the best of your knowledge, does the data package **metadata** (.xml) refrain from sharing any proprietary or protected information, including PII? Metadata files – even for non-public data - will eventually be shared with public audiences.  Comments: |
| 18 |  |  | Do the geographic coordinates in the data and metadata match the locations in titles, keywords, and other text?  Comments:  No coordinates in metadata on purpose |
| 19 |  |  | Does the data package (metadata and data) have an appropriate level of professionalism (comments are relevant to the protocol and the data; no derogatory, negative, or inflammatory language)?  Comments: |

Additional Reviewer Comments:

Awesome data package! Just small changes and suggestions for clarity

Easiest to just track changes, many of these are just suggestions so accept them at your discretion:

Methods:

Beginning the first week in April for each year of the study, the Missouri bladderpod population at grid A is inspected twice weekly to note the phenological status of the plants. Flowering typically begins in early to mid-April. When peak flowering is observed monitoring will begin. At grid A and C, observers mark the corners of 5 m X 5 m cells. During most years at grid A, only cells in the core area are sampled. Every five years, however, the entire grid is sampled. At the other eight grids, observers navigate in 15 m X 15 m (225 m^2) cells using a GNSS (global navigation satellite system) device; corners are not marked. For all sites, Missouri bladderpod density is estimated in grid cells using a density class scale: 0 = no plants, 1 = 1-9 plants, 2 = 10-49 plants, 3 = 50-99 plants, 4 = 100-499 plants, 5 = 500-999 plants, 6 = 1,000-4,999 plants, and 7 = >5,000 plants.

Prior to estimating density, observers calibrate their estimates using the following protocol. Observers estimate abundance using the density class scale (0 = no plants, 1 = 1-9 plants, 2 = 10-49 plants, 3 = 50-99 plants, 4 = 100-499 plants, 5 = 500-999 plants, 6 = 1,000-4,999 plants, 7 = >5,000 plants) and then count the plants in the plot at the ground-level. This process is performed over a range of plot densities until estimates match counts. Observers should also compare results until estimates begin to converge. When estimating density, observers should consider counting up to 100 bladderpods in a 5-X-5m plot before determining estimates. Observers have the discretion to count as many plants as needed to improve confidence in the density estimate. Such counting may encompass all plants in the cell or only the plants in certain section of the plot (e.g., a particularly dense cluster of plants). If needed, plants may be counted in a portion of the plot and then extrapolated to the 5-X-5 m plot.

Accuracy Assessment for Density Class Estimation is done on grid A only. After estimating density for each 5-X-5 m plot at grid A, 60 cells are randomly selected using the rand() function in Excel. Each plot is visited to assess reliability of density class assignments. Additional cells of known density class are selected subjectively for accuracy assessment to bring the total number of cells assessed per density class to 10.

Habitat measurement only takes place on grid A. The percentage of photosynthetically active radiation (PAR) is measured at each plot center using LI-190 Quantum Sensors. PAR is the part of the electromagnetic radiation spectrum (400-700 nm) used in photosynthesis. Field workers also measure dbh of all Eastern red cedar (Juniperus virginiana) stems reaching breast height (1.37 m). Because cedar basal area will not vary greatly in the absence of disturbance, such measurements are only required on a five-year basis. Observers should, however, measure cedar basal area following significant canopy changes (e.g., thinning, fire, ice storms, tornados).