

ReadMe for Data

The Nature Conservancy's LANDFIRE team

10/13/2021

This document serves to document changes made to datasets (will be repeated on website) and provide links to datasets.

Goals and deliverables Phase 1

In general the aim of Phase 1 for ABRP is to get the data wrangled and to provide input that guides future monitoring, and eventually management.

Deliverables

1. Cleaned datasets. All datasets located in [this folder](#)
2. Charts of pine stems and wiregrass
3. LANDFIRE Maps and charts
4. Website with downloadable code

Pine Data Wrangling

- worked from data delivered by Beatriz **ABRP_bioplots_1995-2018**
- added “B-P” column, concatenated BU and Plot numbers to generate unique value for charts
- added MONYEAR and LASTBURNYEAR columns for charts (also, some burns only had years recorded). Used “text to columns” tool in Excel with “/” as the delimiter to create new YEAR-based columns where possible. Manually entered years for many records.
- Bull-dozing happened (instead of fire?) for MONYEAR 2007, B-Ps 7_8, 7_7, 7_6, and 7_4. Did not transfer this information to LASTBURNYEAR (but is in LASTBURN column). Swaty did not want to mix numerical and text data types in the LASTBURNYEAR column.
- B-Ps 7_3, 7_5 was missing MONDATE information for 2019. Swaty manually entered “2019” based on exploring neighboring plots in the dataset.
- Found many typos and case issues (e.g., “Seedling” instead of “SEEDLING”)
- In pine.csv:
 - deleted 21_4 data for 2001
 - deleted 40_10 data for 1998

Note-deletions only done in “pine.csv”. ABRP_bioplots_1995-2018 was left intact.

Pine Data

There are multiple files with “pine” in the name. Most were for charts.

1. pine.csv **main pine dataset**. Raw, but note processing steps above. Some data deleted.
2. for comparing stem counts to reference sites (BU 5)
 - averagePineStemsBU
 - pineBU5all
 - pineStemsBU5average
3. For heatmaps

- pineDiverge (intermediate dataset. Not used in charts.)
- fiveStems (intermediate dataset. Not used in charts.)
- comparePineClean

Grouping of Pine Data

David and Chaz provided a grouping scheme. This was added to the pine data as a new column, then due to time was formatted for the final dataset.

Datasets with names pineClassAonly, pineSeedlingsOnly, pineNoBPH were taken from the pineAll.csv by filtering, copy/paste

Wiregrass data wrangling

In general worked with ABRP_bioplots_1995-2018 herb tab, then saved that to herb.WU.csv. Wiregrass (perGrass.csv) data is a refinement of herb data.

- All cleaned Herbaceous data saved as “herbWU.csv”
- For Perennial Grasses, that data was copied and pasted into a new spreadsheet, “perGrass.csv”. Used Pivot Table in herbWU.csv to create this new spreadsheet.
- Assumed “99999” meant “no data”. Changed to “NA”. Looks like (at least for some years) only one quadrat was sampled for % PG, % WP, % OP, % LI AND % BG.
- Did nothing with blank cells.
- Deleted rows with AVERAGES
- – Bull-dozing happened (instead of fire?) for MONYEAR 2007, B-Ps 7_8, 7_7, 7_6, and 7_4. Did not transfer this information to LASTBURNYEAR (but is in LASTBURN column). Swaty did not want to mix numerical and text data types in the LASTBURNYEAR column.
- In perGrass.csv:
 - deleted 21_4 data for 2001
 - deleted 40_10 data for 1998
 - recalculated average and standard dev for 2010, BP 7_1
 - fixed missing years for BP 7_1

Wiregrass (perinnieal grass) datasets:

- pergrass.csv **main wiregrass dataset**. Raw, but note processing steps above. Some data deleted.
- herbWU.csv spreadsheet for working up the wiregrass data
- Beatriz provided additional cleaning. Her notes and cleaned data are in an Excel file named [Copy of Herbaceous data - BPA. This data replaced all other data and was pasted into other herb datasets.