

# DIMA Quality Control Report

*29 November, 2016*

## Contents

<b>DIMA Data Summary</b>	<b>1</b>
Report parameters and options . . . . .	1
Data Summary . . . . .	2
<b>Observer data summary</b>	<b>3</b>
Within-crew Variability by Indicator . . . . .	3
<b>Missing Value Checks</b>	<b>7</b>
<b>Data Consistency Checks</b>	<b>8</b>
LPI Values Checks . . . . .	8
Vegetation Height Checks . . . . .	10
Canopy Gap Values Check . . . . .	10
Plot Species Inventory Data . . . . .	11
Temporal Stability of Indicators . . . . .	12
<b>Implications</b>	<b>17</b>

---

It is important to understand the quality of monitoring data in order to be able to successfully and responsibly use it to support management decision making. This report is a **Quality Control** assessment of data contained in a DIMA database. The purpose of this report is to identify statistical patterns that exist in core monitoring data (Line-point Intercept [LPI], Vegetation Height, Canopy Gap, Soil Stability, Species Inventory) and to flag data values that are outside of what is expected (i.e., potential errors or extreme events). This report consists of the following sections:

1. DIMA Data Summary
  2. Within-Crew Variability Checks
  3. Missing Value Checks
  4. Data Consistency Checks
  5. Temporal Stability of Indicators
- 

## DIMA Data Summary

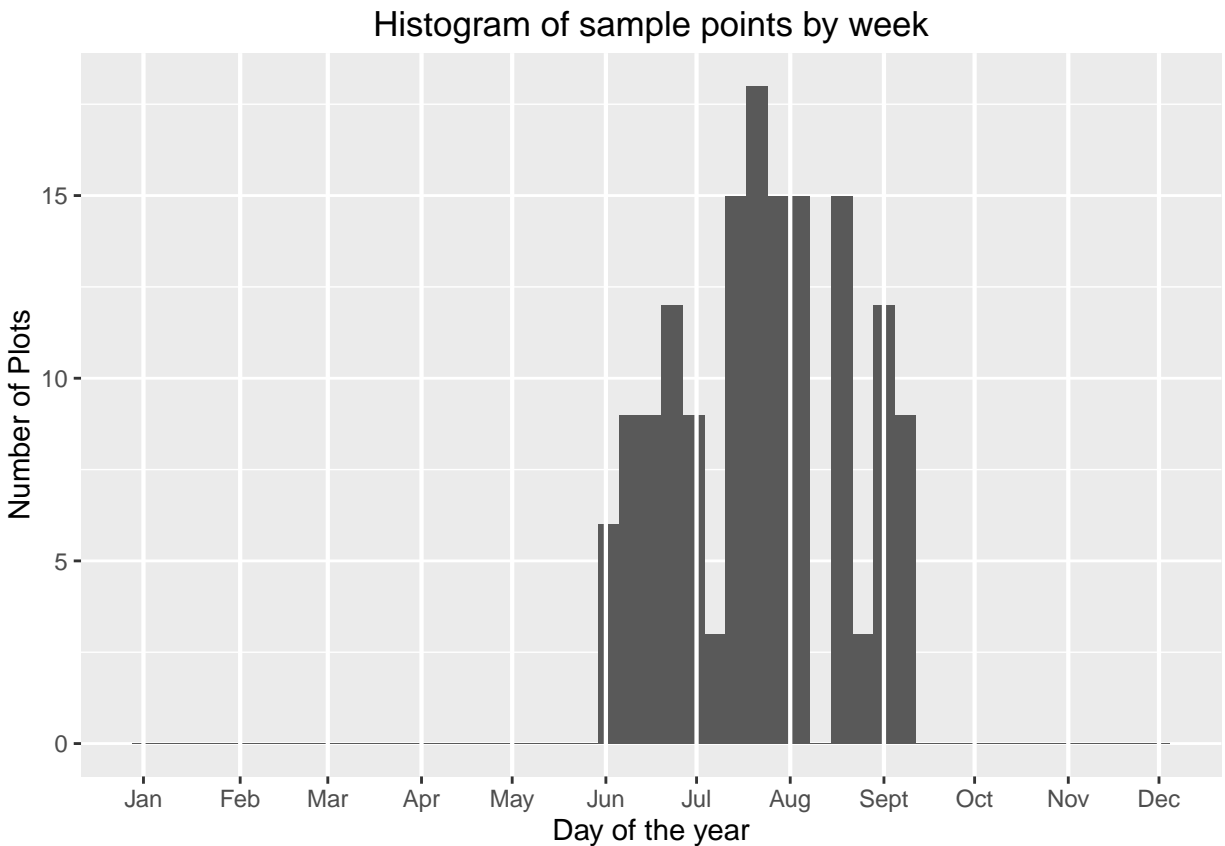
### Report parameters and options

- DIMA Database: WRFO\_2011-2014\_DIMA3.1.mdb
- Site: all
- Year: 2014
- Total number of plots sampled: 50

## Data Summary

### Temporal Sampling Window

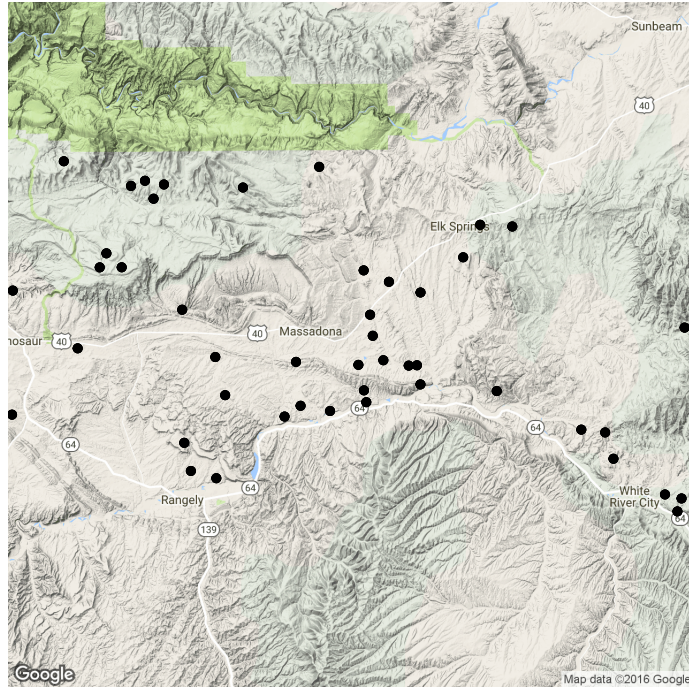
Knowing when during the year that monitoring data were collected is helpful for evaluating whether or not monitoring data are appropriate or useful for a specific purpose, especially with indicators that are sensitive to phenological stage or time of year when they are collected.



### Sample Locations Map

The map below is drawn from latitude/longitude coordinates contained in the DIMA plots table. This is a useful, fast check if there are gross errors in the plot coordinates that were recorded in DIMA (e.g., Are the plots in the right general place? Are there some plots that are way out of position relative to the others?). The map below is not intended as a substitute for comparing the sampled plot locations against the original sample design because plots could still be mislocated or not within stated QC criteria but still look alright on a coarse-scale map.

### Plot sample locations



## Observer data summary

Table 1: Number of plots on which each crew member was an observer.

Observer	Canopy Gap	LPI	Soil Stability	Species Richness
Emma Stewart	47	46	44	46
Stephen White	46	46	2	3
Tim Rogers	3	3	4	1

## Within-crew Variability by Indicator

For data quality, it is important that all members of a field crew observe and record characteristics and attributes of a site in the same manner. Following a method's protocol is a central part of this, but there can also be differences between crew members in how things like litter or canopy gaps are identified and recorded. Training and calibration are designed to maximize the consistency in data recording among crew members.

One way to look at within-crew variability in aggregate is to plot the range of values for an indicator by observer for a large set of plots. This gives an idea of how, on average, each observer compares to the other observers in that crew. Large deviations from the rest of the crew can indicate a crew member that is either not well trained or not calibrated to the rest of the crew. Equivalence tests evaluate the null hypothesis that two crew members are, on average, within a specified equivalence region to each other (i.e., if results are, on average, within a specified amount to each other, we conclude that they're equivalent). The equivalence region for these tests is set to 20% of the overall indicator mean. A low p-value for the significance test indicates

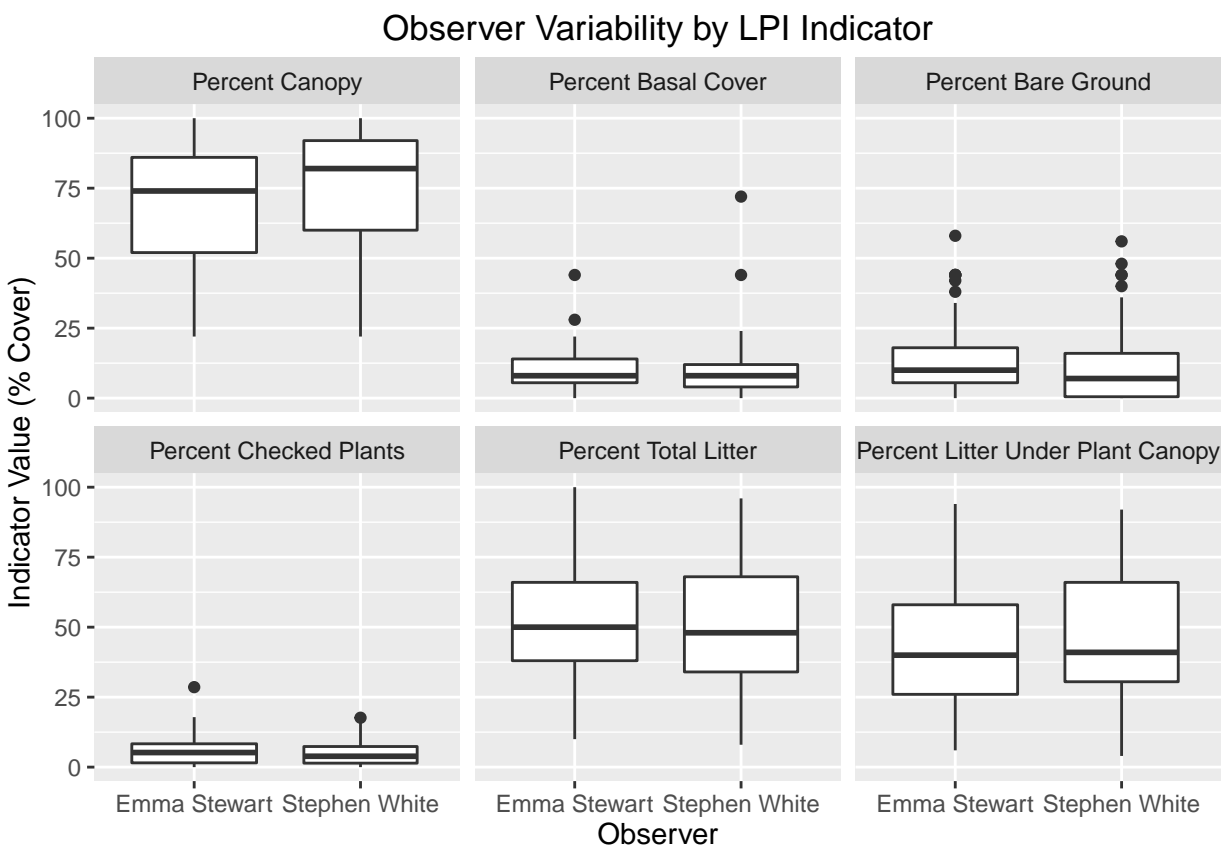
that there may be a significant difference in average indicator values between two observers. Equivalence test results should be evaluated alongside the indicator's boxplot.

Just because an observer's average results are not statistically equivalent to other observers does not necessarily mean that one observer is wrong - other factors may be contributing. For example, it could be that the observer was recording more often for sites of a certain type (e.g., low-productivity sites) and their average results reflect this. Extreme departures from equivalence, though, should be followed up by examination of an observer's data forms to identify what might be causing the issue.

### Within-Crew Equivalence for LPI Indicators

Table 2: Observer equivalence tests by LPI indicator.

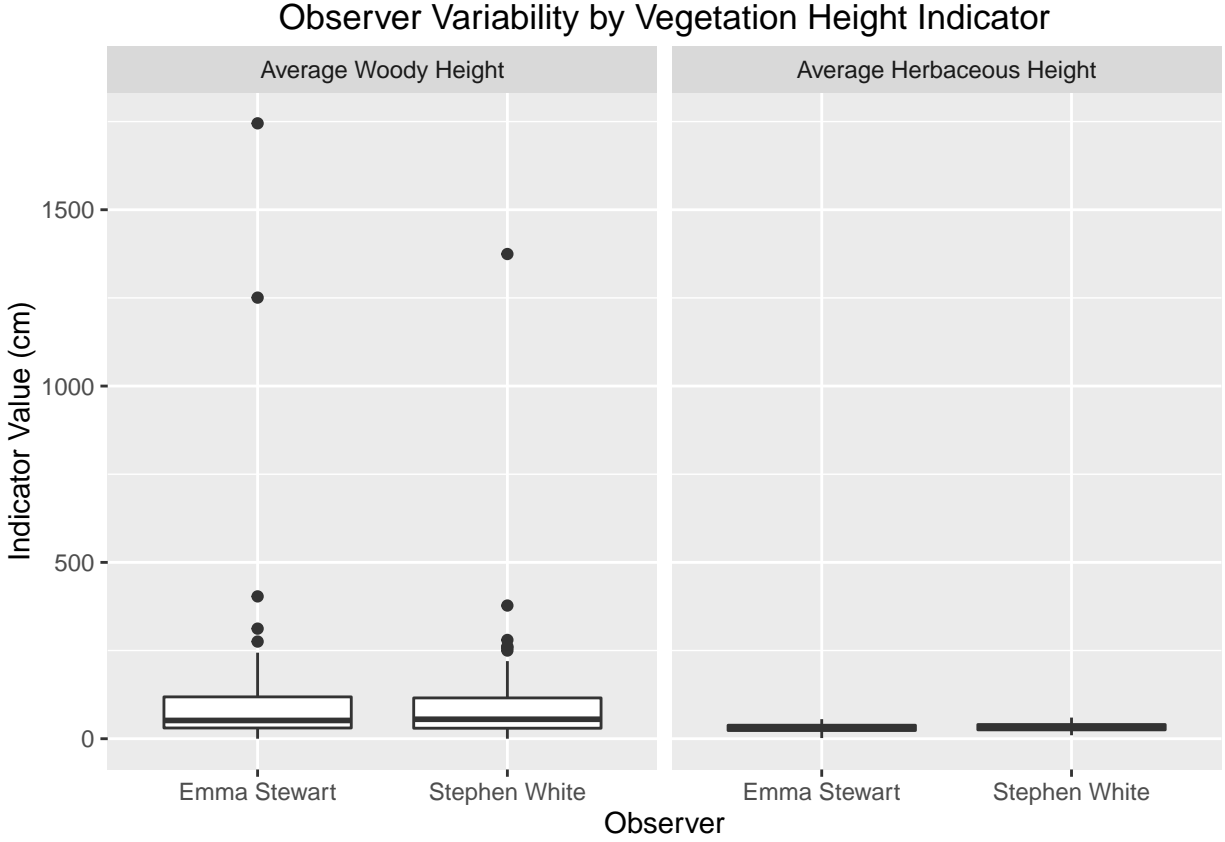
Indicator	Comparisons	Estimates	Std.Err	Eval.Criterion	t.statistic	p.value
pctCanopy	Stephen White - Emma Stewart	5.5591	3.3930	14.3726	-2.5976	0.9948
pctBare	Stephen White - Emma Stewart	-2.3129	2.1589	2.4959	-2.2274	0.9863
pctBasal	Stephen White - Emma Stewart	-0.4545	1.4517	1.9589	-1.6625	0.9507
pctCheckedPlants	Stephen White - Emma Stewart	-0.7703	0.8319	1.1034	-2.2523	0.9871
pctLitterTotal	Stephen White - Emma Stewart	-0.7871	3.3731	10.2438	-3.2703	0.9993
pctLitterUnder	Stephen White - Emma Stewart	2.0508	3.5590	8.6904	-1.8656	0.9679



### Within-Crew Equivalence for Vegetation Height Indicators

Table 3: Observer equivalence tests for Vegetation Height indicators.

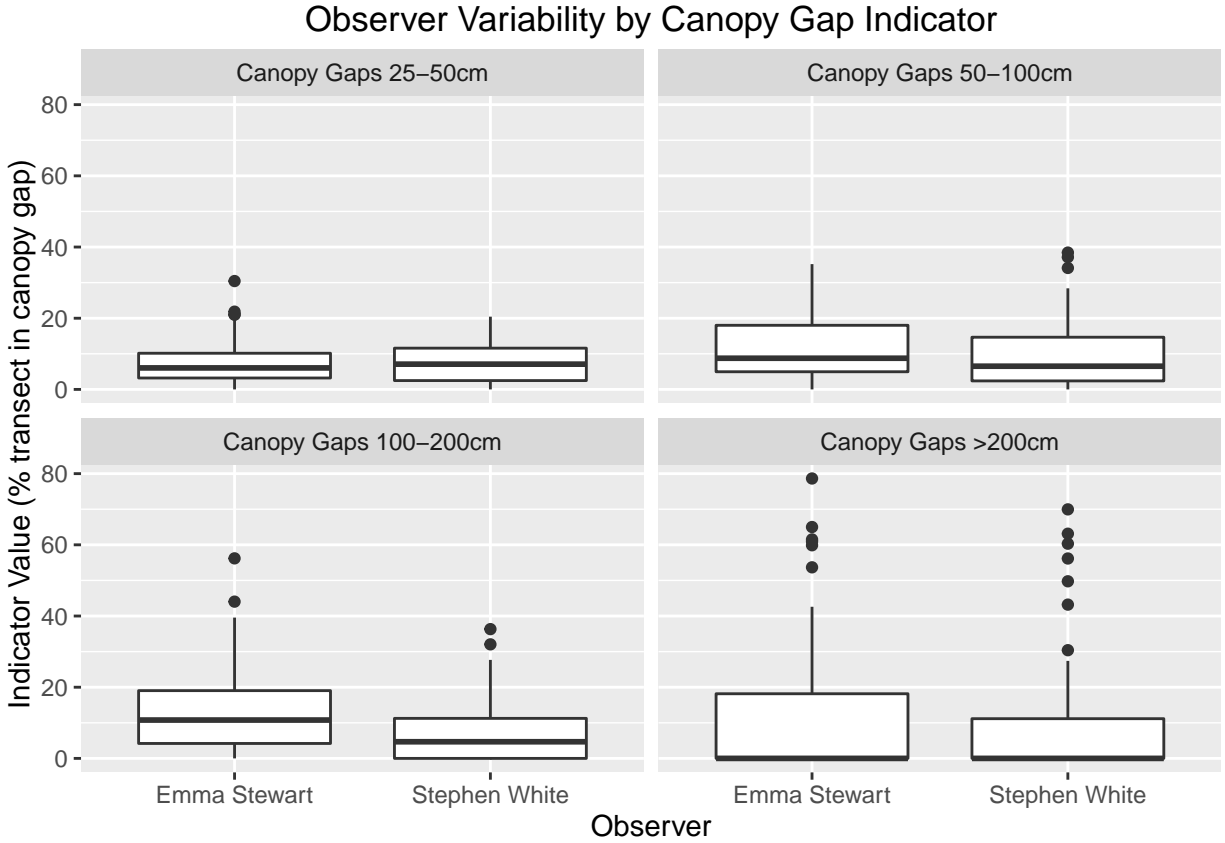
Indicator	Comparisons	Estimates	Std.Err	Eval.Criterion	t.statistic	p.value
Woody Height	Stephen White - Emma Stewart	-10.7351	35.4819	21.8916	-0.9195	0.8203
Herbaceous Height	Stephen White - Emma Stewart	2.2782	1.8756	6.4629	-2.2311	0.9864



#### Within-Crew Equivalence for Canopy Gap Indicators

Table 4: Observer equivalence tests for Canopy Gap indicators.

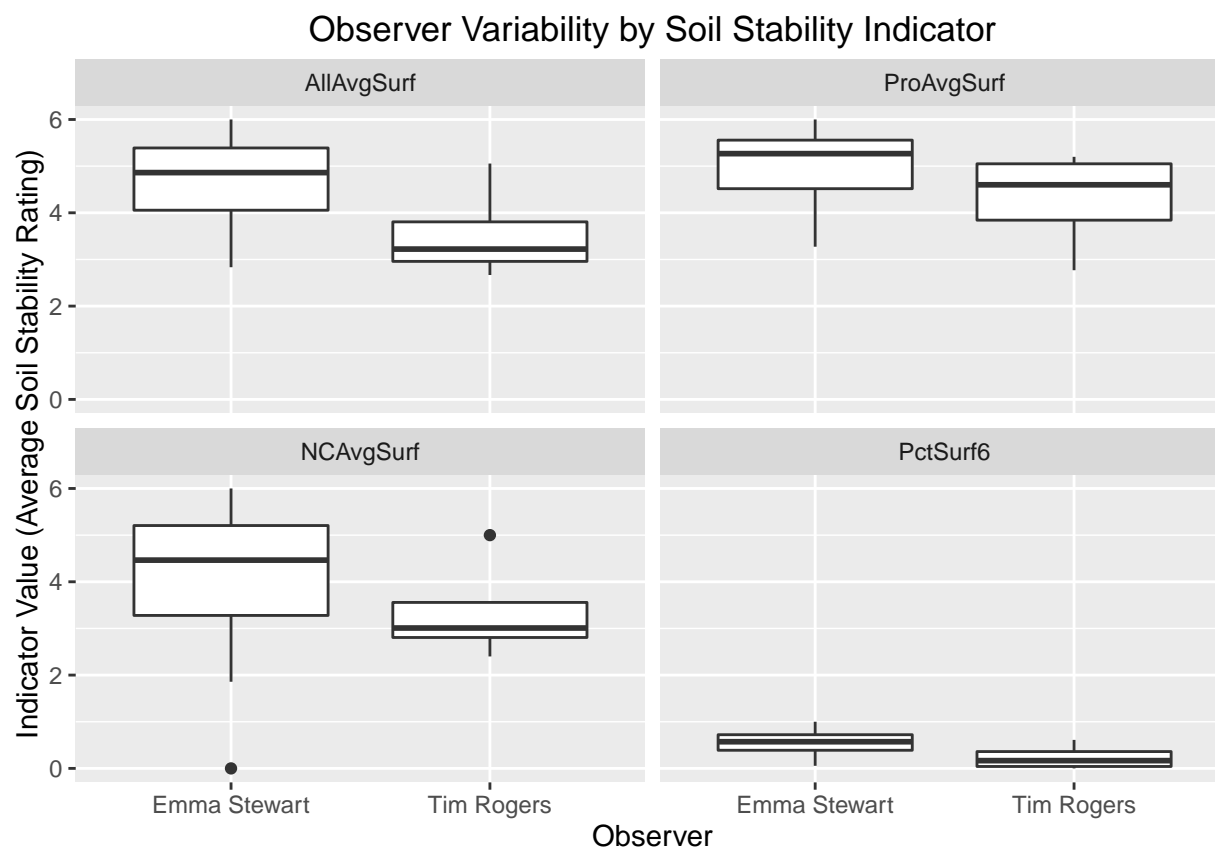
Indicator	Comparisons	Estimates	Std.Err	Eval.Criterion	t.statistic	p.value
pctCanCat1	Stephen White - Emma Stewart	-0.0990	0.9748	1.4974	-1.6377	0.9482
pctCanCat2	Stephen White - Emma Stewart	-1.9579	1.5137	2.1165	-2.6916	0.9960
pctCanCat3	Stephen White - Emma Stewart	-4.8600	1.7821	2.0445	-3.8745	0.9999
pctCanCat4	Stephen White - Emma Stewart	-2.7730	3.0269	2.2321	-1.6535	0.9498



### Within-Crew Equivalence for Soil Stability Indicators

Table 5: Observer equivalence tests for Soil Stability Indicators

Indicator	Comparisons	Estimates	Std.Err	Eval.Criterion	t.statistic	p.value
AllAvgSurf	Tim Rogers - Emma Stewart	-1.1700	0.4402	0.9228	-4.7542	1.0000
ProAvgSurf	Tim Rogers - Emma Stewart	-1.1700	0.4402	0.9887	-4.9038	1.0000
NCAvgSurf	Tim Rogers - Emma Stewart	-0.8100	0.6581	0.8194	-2.4759	0.9915
PctSurf6	Tim Rogers - Emma Stewart	-0.3089	0.1221	0.1039	-3.3812	0.9993

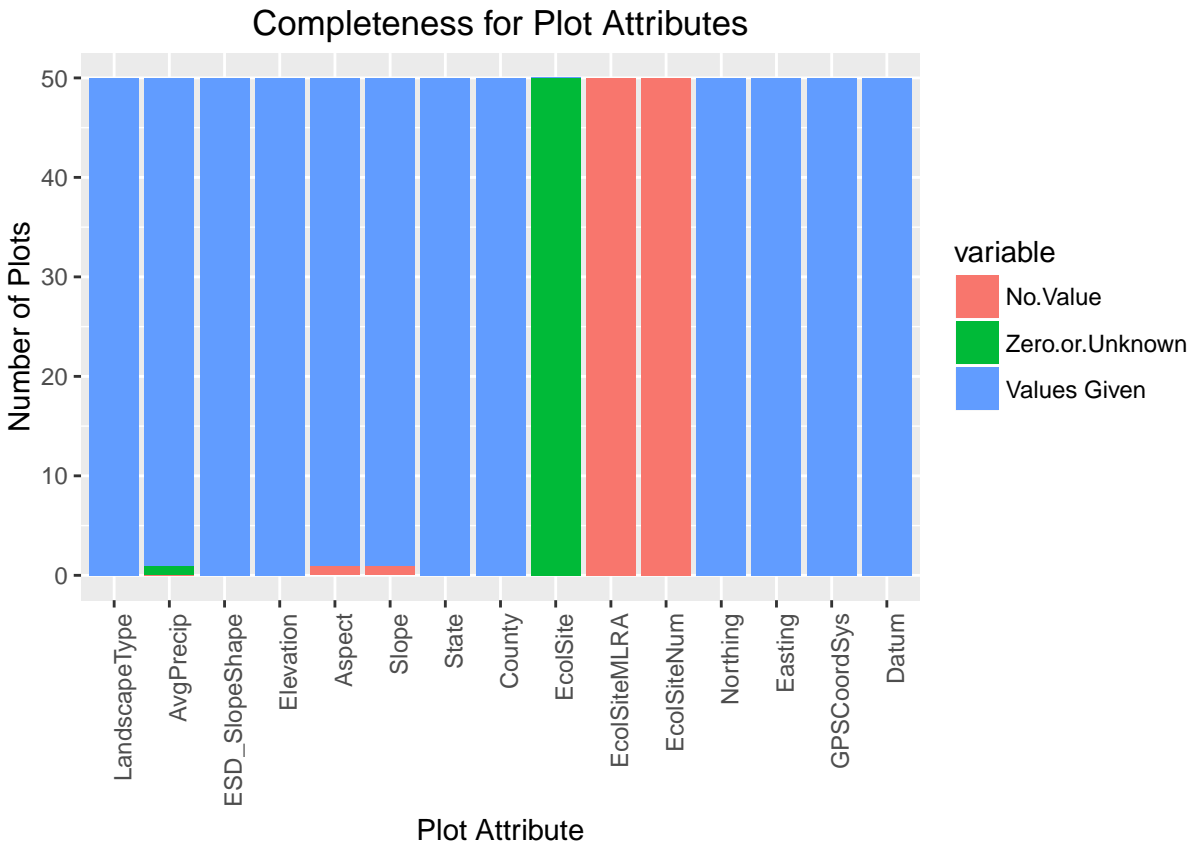


## Missing Value Checks

Queries/results of plots that have missing information

### Plot Characterization Information Completeness

Plot characterization information is important for describing the general setting where monitoring data were collected and provide valuable covariate information for analyses. Below is a quick check of how many plots are missing valid values for a set of plot attributes. To find specifically which plots are missing these plot attribute values, check the tblPlots table in DIMA.



## Data Consistency Checks

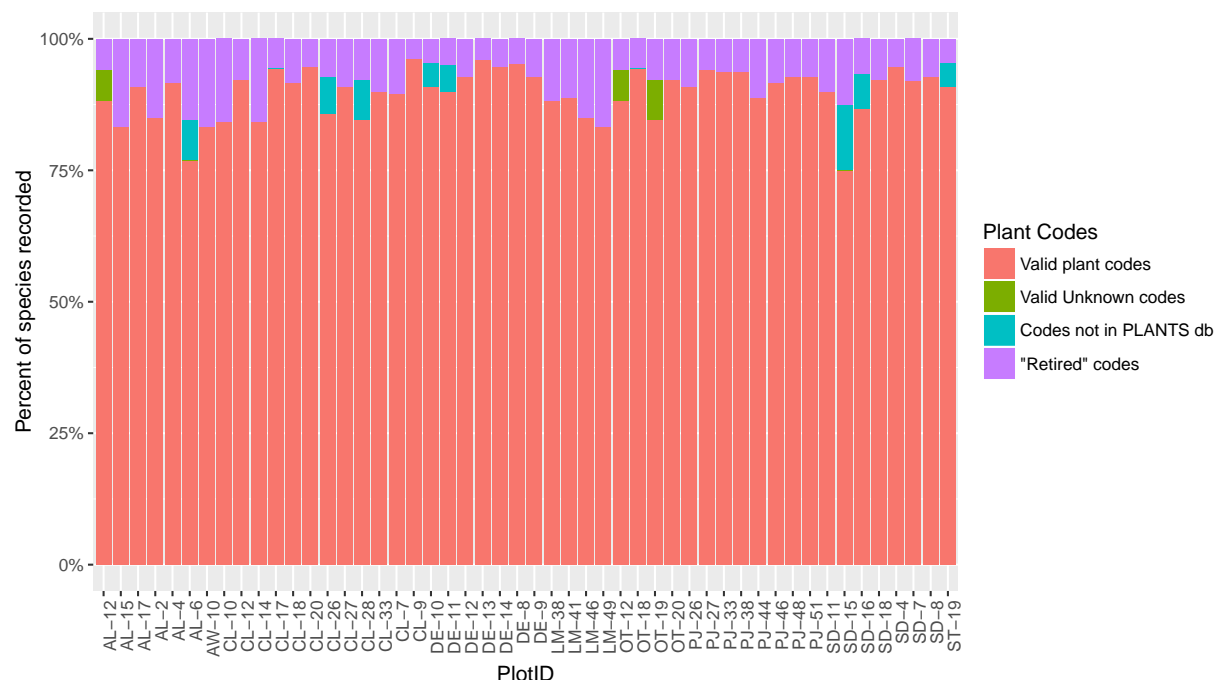
Data consistency checks consist of graphs and/or statistical analyses of plot and transect-level indicators to flag possible errant data values. It is important to note that results of the data consistency checks do not necessarily mean that the data are wrong, but that it falls outside of what is expected for an indicator. The appropriate steps to take for data that are flagged in a data consistency check is to go back to the original DIMA forms or paper data sheets, field notes, and plot photos to see if something was possibly recorded incorrectly or if there is another explanation for the values.

## LPI Values Checks

### Prevalence of unknown plant codes, “retired” plant codes, and codes not in the PLANTS database

The graph below shows what proportion of the species recorded on a plot were valid plant codes, valid unknown codes, “retired” plant codes, or codes that were not in the PLANTS database. A “retired” plant code is one that used to be a valid PLANTS Database code, but due to taxonomic changes is no longer valid. For example, bluebunch wheatgrass used to have the scientific name of *Agropyron spicatum*, but is now known as *Pseudoroegneria spicata*. Thus the old code of AGSP is no longer valid, and PSSP6 is to be used for data entry. Luckily, the PLANTS Database maintains a listing of the “retired” codes as synonyms of the current codes.





## Check LPI Species Codes

### Species codes that are not currently valid codes in the PLANTS database

The following plant codes were not found as current codes in the PLANTS database, but were found as a synonym of a current code. The current code and species information is also presented. Please verify that this is the species that was intended before changing any plant codes.

Code in DIMA	Current PLANTS Code	Scientific Name	Common Name
AGSP	PSSPS	Agropyron spicatum Pursh	
ASGL3	EUGL19	Aster glaucodes S.F. Blake	
BERE	MARE11	Berberis repens Lindl.	
BRJA	BRAR5	Bromus japonicus Thunb.	
ERCO24	ARCOC4	Eremogone congesta (Nutt.) Ikonn.	
HIJA	PLJA	Hilaria jamesii (Torr.) Benth.	
LARE	LAOCO	Lappula redowskii auct. non (Hornem.) Greene	
POAG	POPRP2	Poa agassizensis B. Boivin & D. LÃ¶ve	
RATE	CETE5	Ranunculus testiculatus Crantz	
STCO3	ACNED	Stipa columbiana auct. non Macoun	

### Species codes not recognized by the PLANTS database

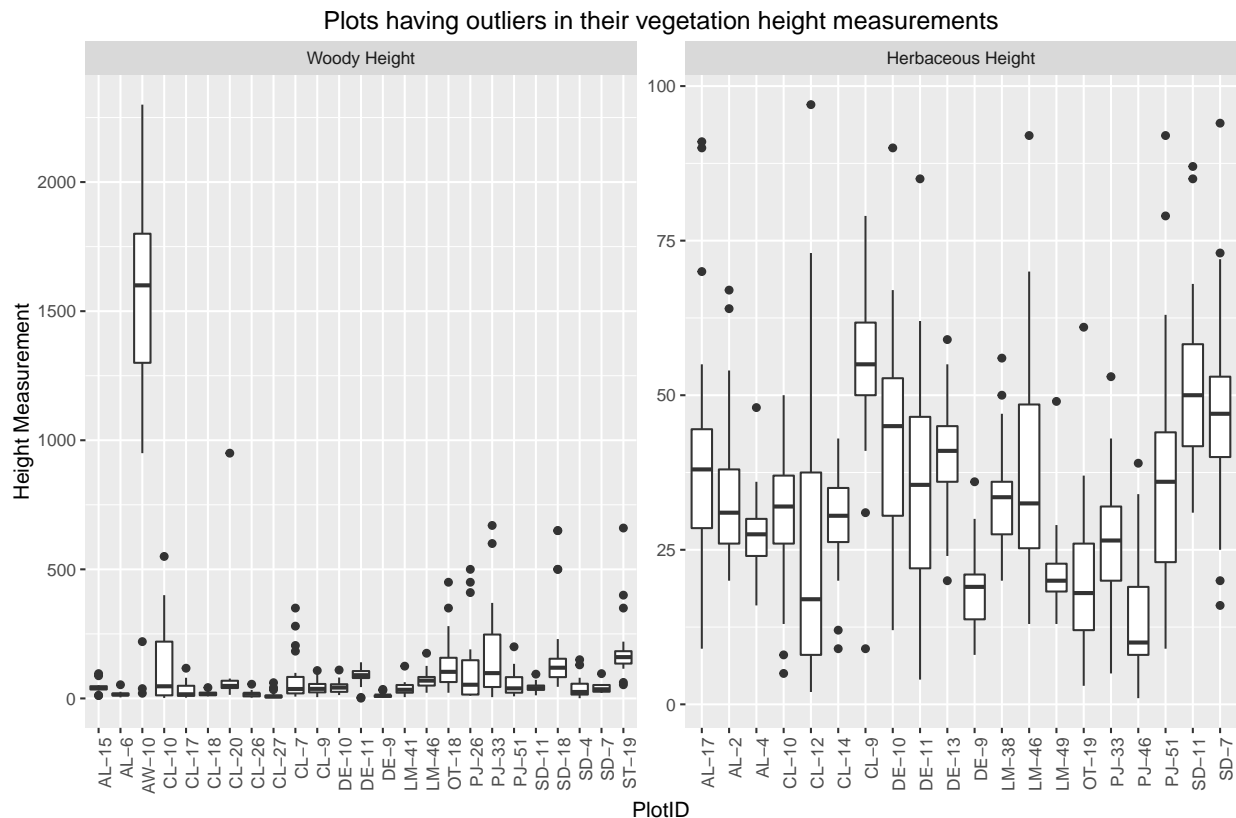
The following plant codes are not recognized as either a legitimate/current plant code in the USDA PLANTS Database, a synonym of a current code, or an AIM unknown plant code.

PlotID	Code not in PLANTS list
CL-26	AAFF
CL-28	PPGG
AL-6	UUFF

PlotID	Code not in PLANTS list
ST-19	AAFF
DE-10	AAFF
DE-11	AAFF
SD-15	CRYPTAA
SD-16	ERIOGAA

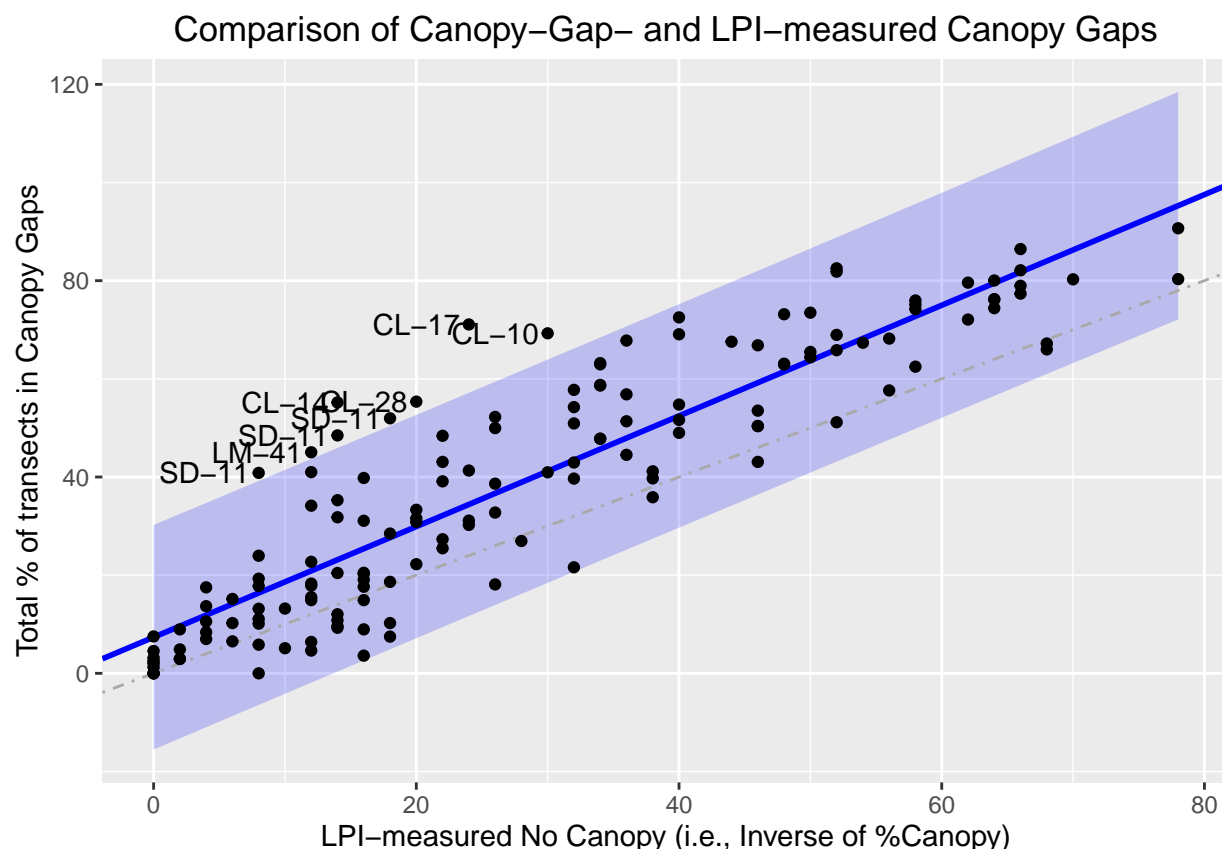
## Vegetation Height Checks

The plots shown in the graphs below were flagged as having vegetation heights considered to be outliers (i.e., more than 1.5 times the inner quartile range of all heights measured on the plot). Extreme values for vegetation height do not necessarily mean that an error has occurred, but height values for these plots should be double-checked to ensure accuracy. If a suspicious height value or an error is identified, please make note of it in the QC tracking worksheet.



## Canopy Gap Values Check

The total amount of canopy gaps recorded on an AIM plot should correlate well to the plot's LPI estimates of bare ground and non-plant ground cover (e.g., rocks, litter) not under a plant canopy. Extreme differences between these two measures may indicate that there is a problem with the canopy gap estimates for the plot.



Estimates of total amount of “no canopy” along transects based on Canopy Gap and LPI should be a close reflection of each other. The blue shaded region in the graph is the 95% prediction interval for a regression between the LPI and canopy gap measures of “no canopy.” Labeled points are extreme values (outside the 95% prediction interval) and should be more closely examined to determine if they represent errors. Points above the blue-shaded region have a significantly larger estimate of “no canopy” from Canopy Gap than was recorded in LPI. Those points below the blue-shaded region have significantly lower estimates of “no canopy” from Canopy Gap than was recorded in LPI.

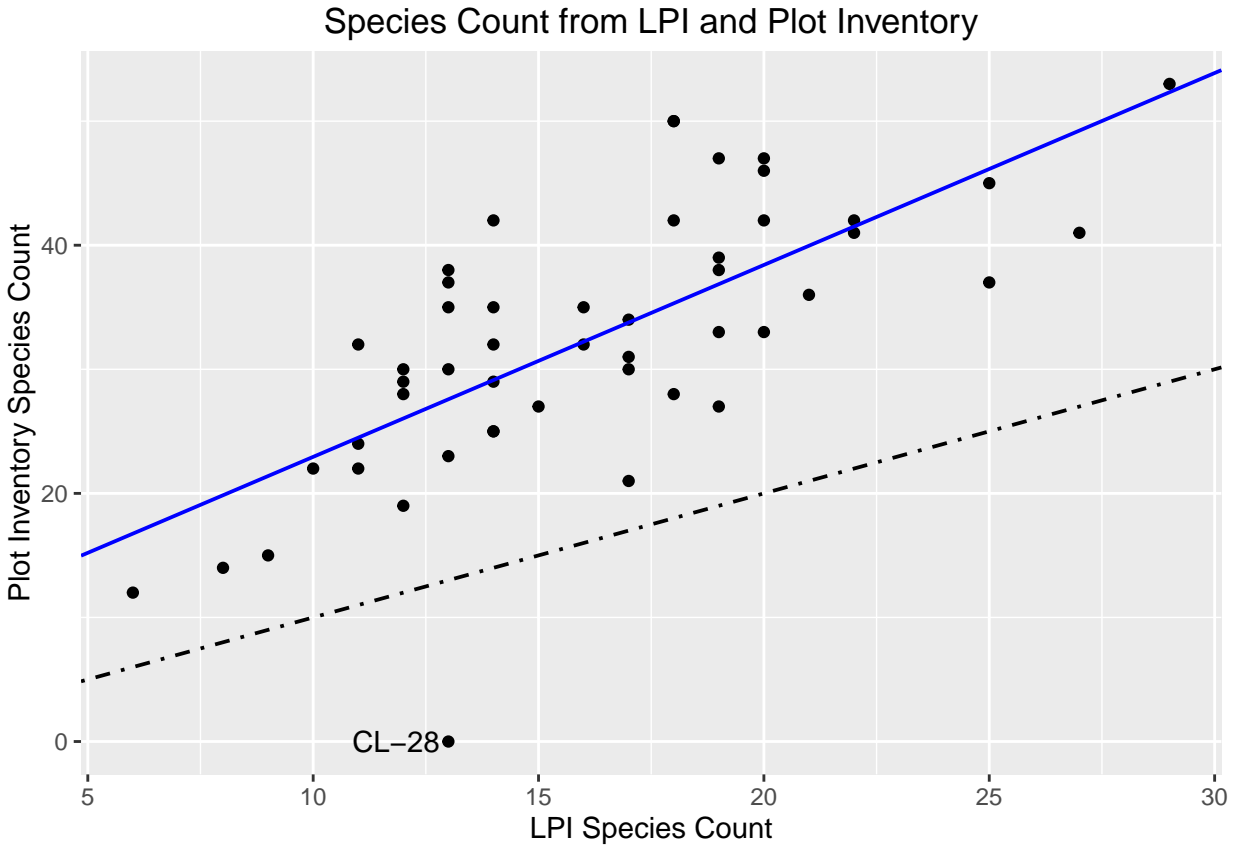
## Plot Species Inventory Data

The following plots have species codes listed in their species inventory that are not in the DIMA master species table (tblSpecies) list. This is uncommon, but can happen if you override the DIMA check for a species not being in the table when entering data and then fail to enter the species afterward.

## All plant codes in the species inventory are in the DIMA Species Table. Check above for plant codes r

## Comparison of number of species in species inventory to number from LPI.

The graph below plots number of species recorded in a plot’s LPI forms against the number of species recorded in the plot-level species inventory. The number of species in the plot inventory should always be greater than or equal to the number of species encountered in LPI. If it is not (i.e., plots below the dashed 1:1 line), then this signifies an error in the plot species inventory and should be noted in the data QC worksheet.

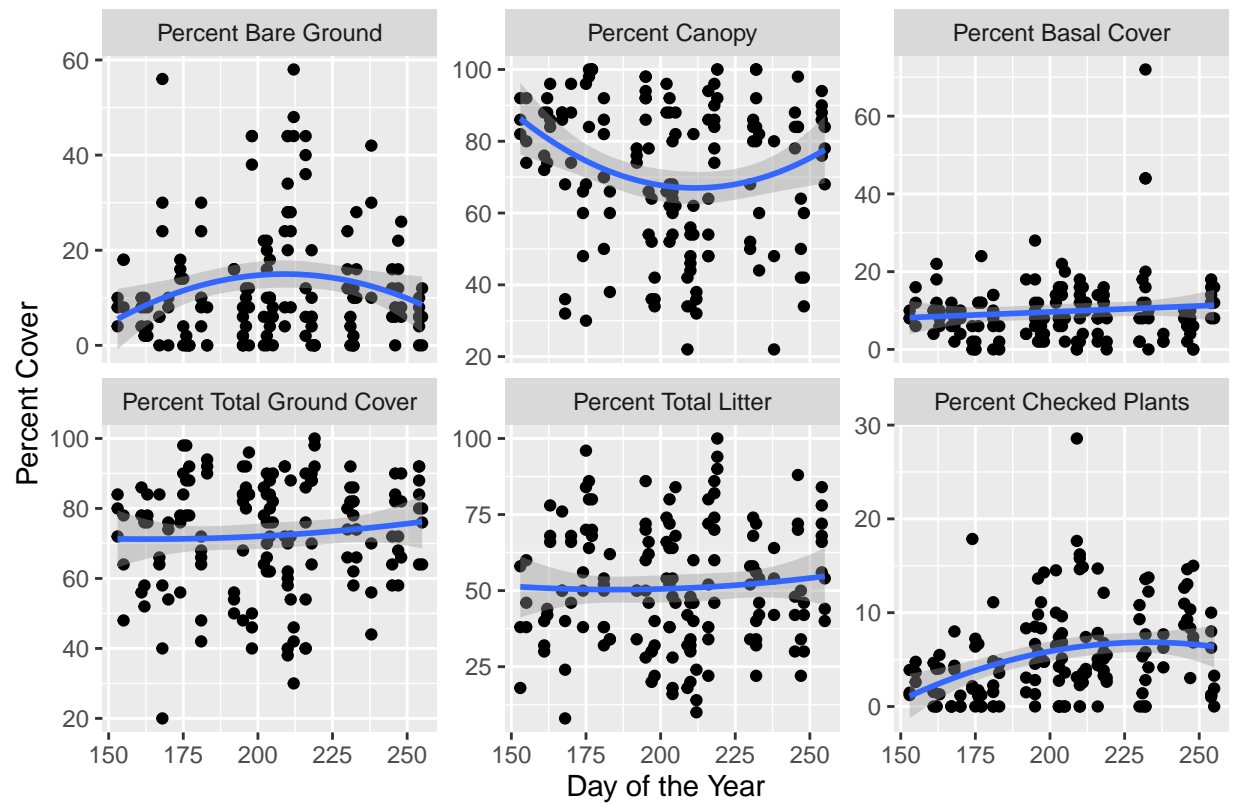


## Temporal Stability of Indicators

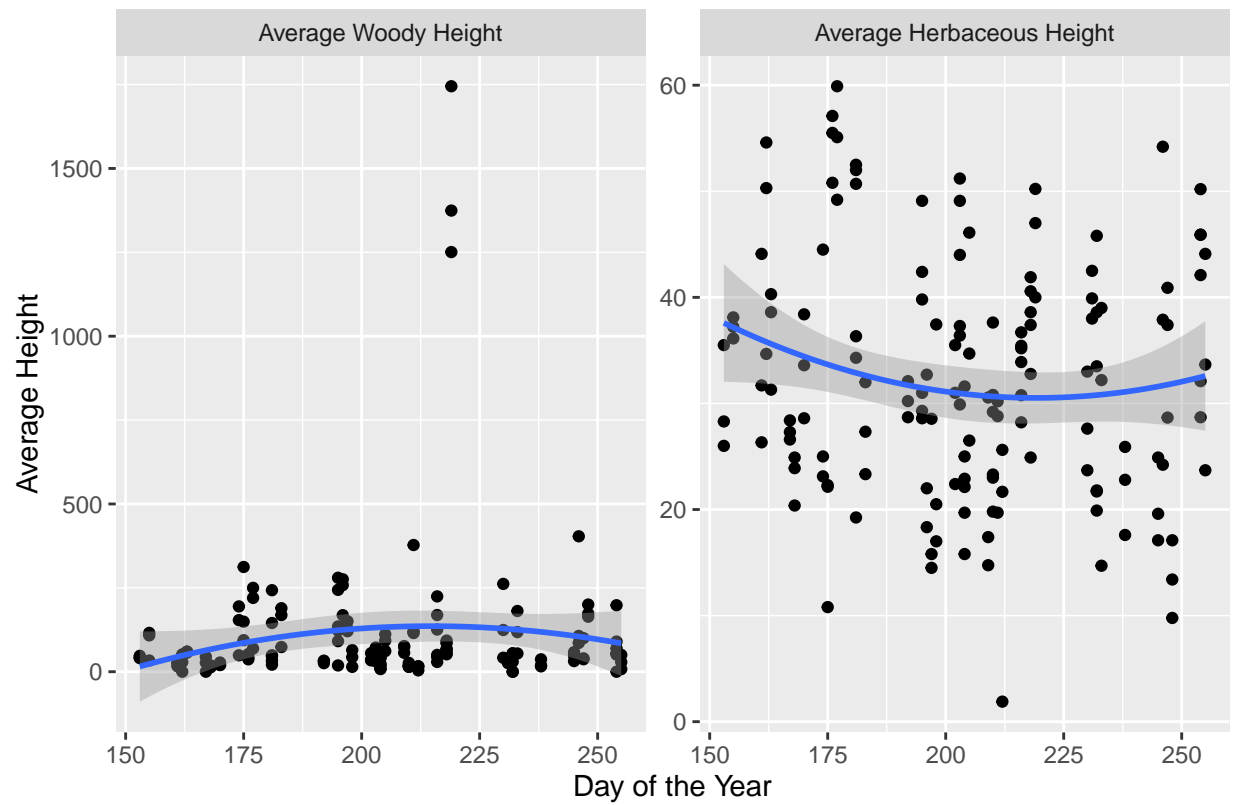
Examining the variability of indicators by day across the sampling period can be a useful tool for understanding if there are significant trends due to changes in phenology. For example, values for certain indicators (e.g., forb cover) are believed to be sensitive to sampling date. The graphs below look at how DIMA's standard, pre-calculated indicators for each of the core methods change over the course of the year. A 2nd-order polynomial trend line is fit to the data to help highlight if there is an overall seasonal pattern to the indicator values. Be sure to interpret the trend line in the light of the amount, spread, and pattern of the underlying points (i.e., plot data).

While this analysis can be valuable for understanding factors that might be influencing monitoring data, the results below should be interpreted with caution - i.e., it is easy to over-interpret these graphs. Many different factors may contribute to a pattern of indicator values by time. For instance, if monitoring plots are visited in such an order that low-productivity sites are visited early in the year and higher-productivity sites visited later, this will manifest as a trend of increasing cover and species diversity with time. In this case, the pattern is a reflection of hitch planning and logistics, and not necessarily due to sampling sites at non-optimal times. Conversely, if data from many different ecosystem types are included in this analysis, then patterns in indicators (e.g., a peak in cover corresponding to peak growing season) may be obscured by data from other systems. Use your knowledge of the data to filter plots in the Shiny tool to include in the analyses and to interpret visible patterns.

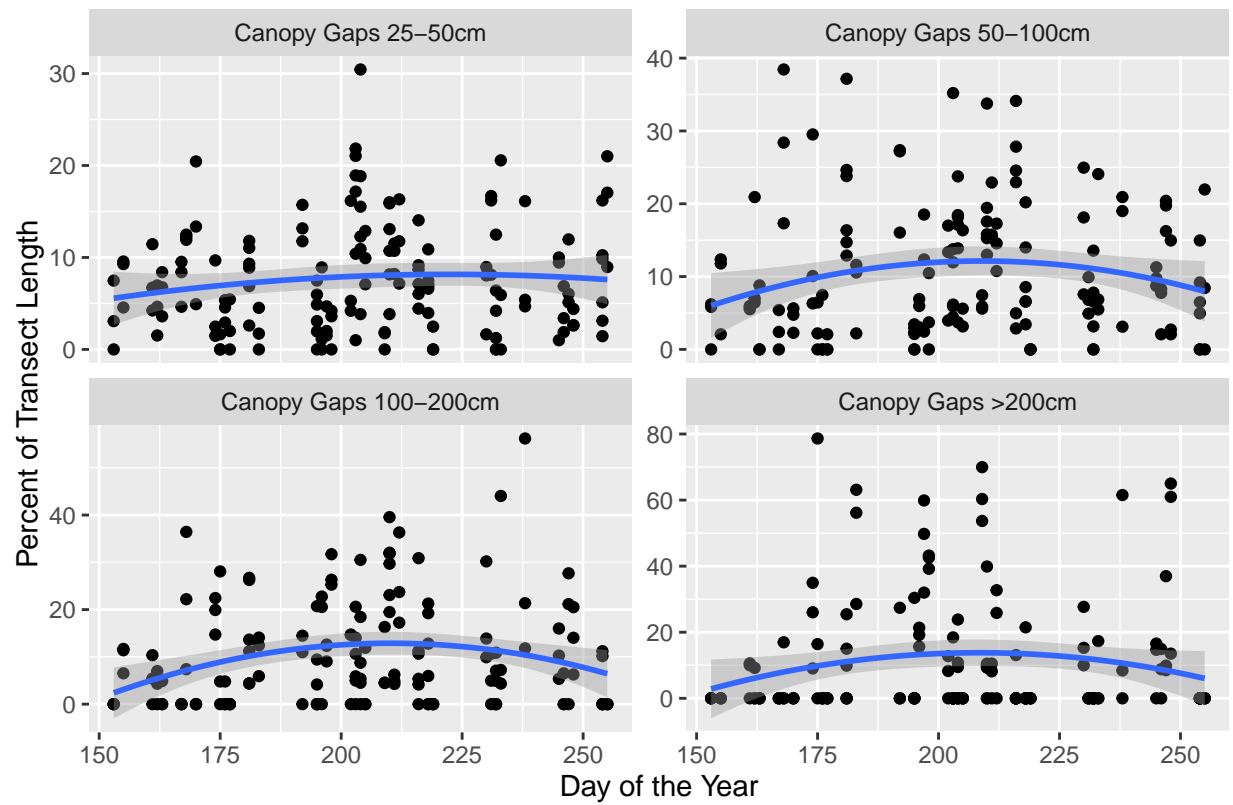
## Variability of LPI Cover Indicators by Day of the Year



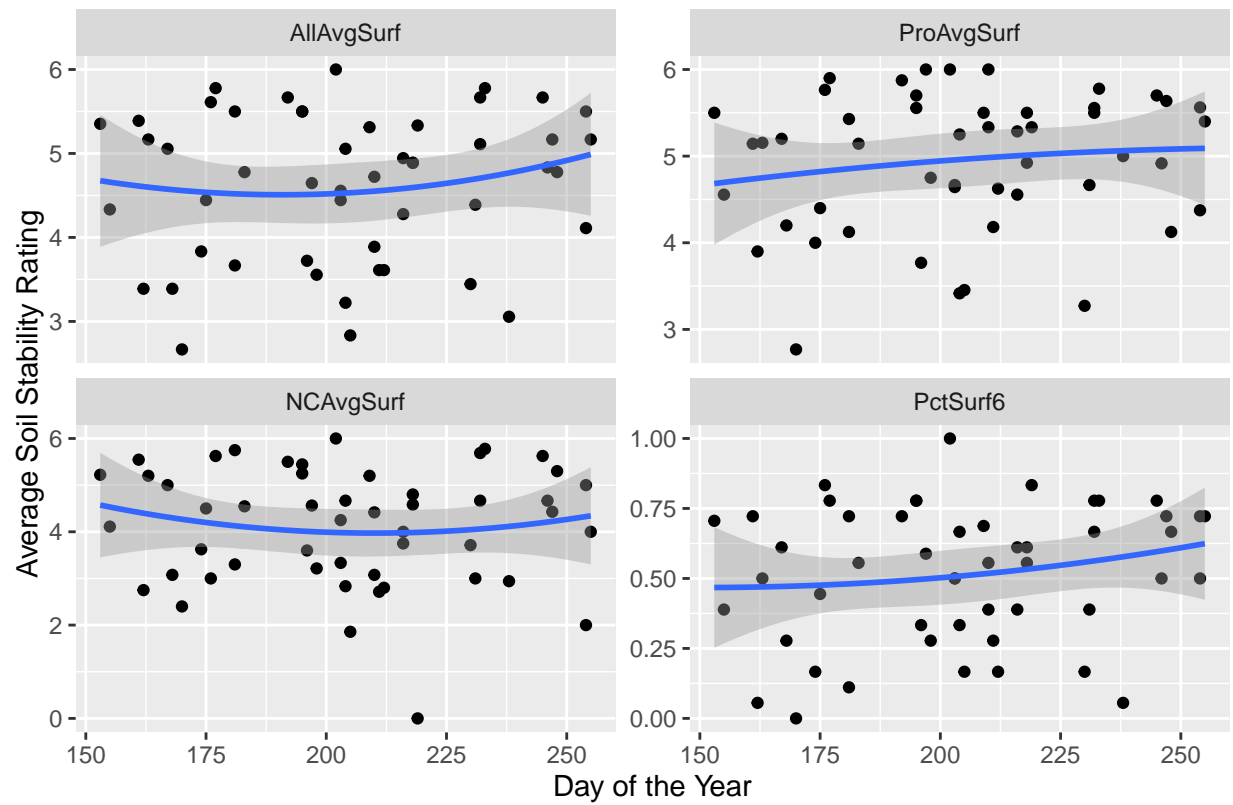
## Variability of Height Indicators by Day of the Year



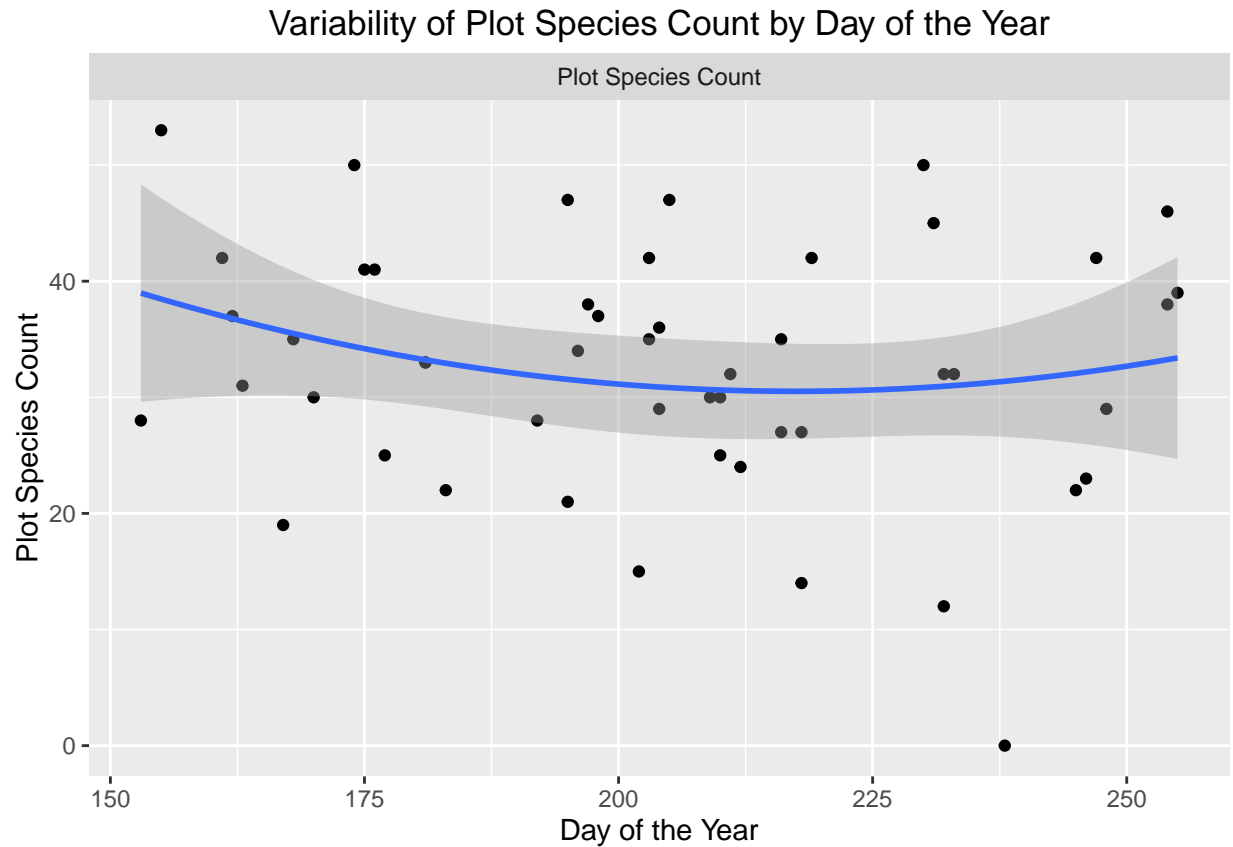
## Variability of Canopy Gap Indicators by Day of the Year



## Variability of Soil Stability Indicators by Day of the Year







## Implications

Fill out a data quality workbook that summarizes the significant findings from this QC report and identifies any missing data and/or suspect or aberrant plot indicator values.