

CDW Tally Analysis: D02 SERC

Cody Flagg, Courtney Meier

8th January 2017

```
## Windows paths
# inpath <- "C:/Users/cflagg/Documents/GitHub/neonPlantSampling/cdw_tally_analysis/D07/"
# outpath <- inpath

## macOS inputs
if(file.exists("~/Documents/workDocuments")){
  inpath <- "~/Documents/workDocuments/gitRepositories/neonPlantSampling/cdw_tally_analysis/fulcrumData"
  domain <- "D02"
  site <- "SERC"
  outpath <- paste("~/Documents/workDocuments/gitRepositories/neonPlantSampling/cdw_tally_analysis/", domain)
}

if(file.exists("~/Documents/neonScienceDocs")){
  inpath <- "~/Documents/neonScienceDocs/gitRepositories/neonPlantSampling/cdw_tally_analysis/fulcrumData"
  domain <- "D02"
  site <- "SERC"
  outpath <- paste("~/Documents/neonScienceDocs/gitRepositories/neonPlantSampling/cdw_tally_analysis/", domain)
}

## import libraries
library(plyr)
library(dplyr)
library(ggplot2)

## import data
parent_cdw <- read.csv(paste(inpath,"tos_coarse_downed_wood_tally_prod.csv",sep=""), header=TRUE)
child_cdw <- read.csv(paste(inpath,"tos_coarse_downed_wood_tally_prod_per_plot_azimuth_log.csv",sep=""))

## filter out parent fields that are unneeded
parent_cdw_fil <- dplyr::select(parent_cdw,
                                fulcrum_id, domainid, siteid, plotid_parent, tallydate, volume, partid)

## filter out child fields that are unneeded
child_cdw_fil <- dplyr::select(child_cdw,
                                fulcrum_parent_id, lidsazimuth, logdiameter, taxonid, decayclass, logid_inge)

## join parent to child via parent::fulcrum_id to child::fulcrum_parent_id
cdw <- merge(x = parent_cdw_fil, y = child_cdw_fil, by.x = "fulcrum_id", by.y = "fulcrum_parent_id", all=TRUE)

## filter to the selected domain, then site only
cdw <- dplyr::filter(cdw, domainid==domain, siteid==site)

## create diameter class factor
cdw$diameterClass <- ifelse(cdw$logdiameter >= 10, '>=10cm',
                           ifelse(cdw$logdiameter <5, "2-5cm", "5-10cm"))

## simplify decayclass to numeric value wrapped inside apply(decayClassNum, "[", 1) e.g. 'return first
cdw$decayClassNum <- apply(stringr::str_split(cdw$decayclass, pattern = " "), MARGIN=2, FUN=function(x){
```

```
## write file
```

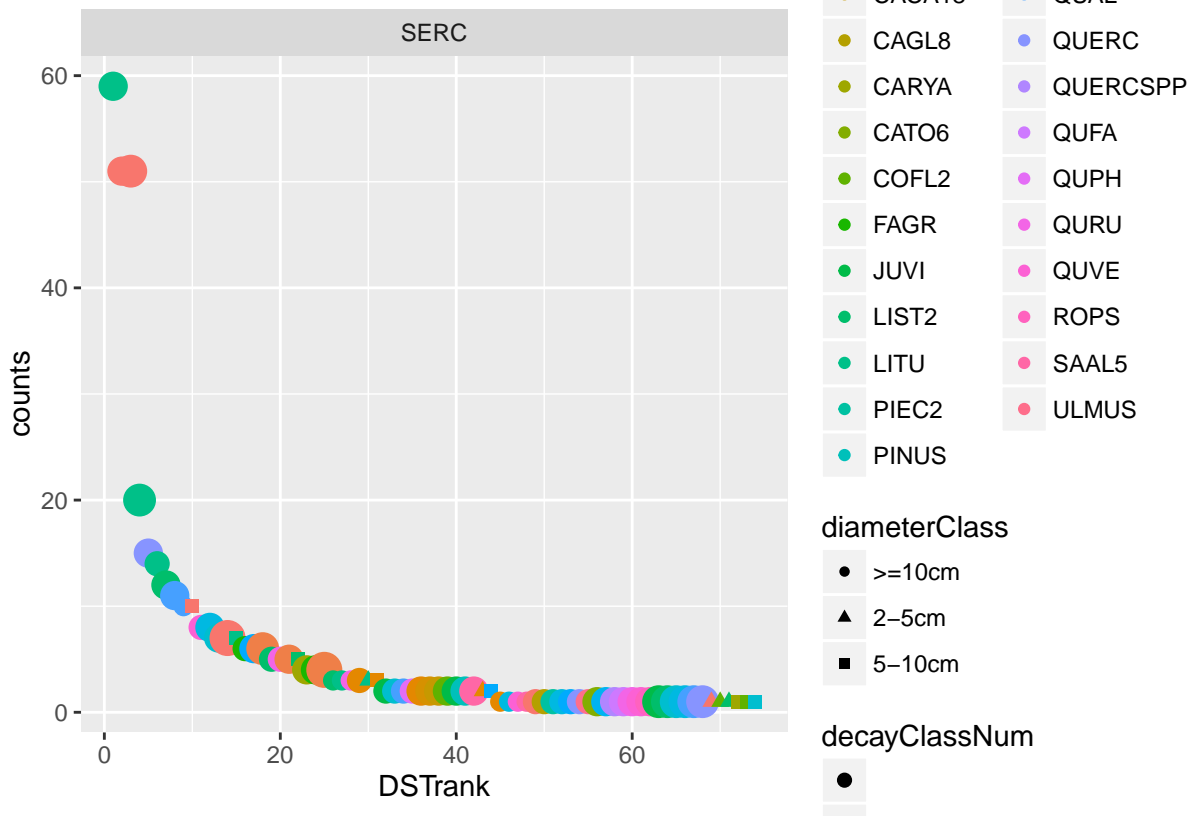
```
write.csv(cdw, file = paste(outpath, paste(domain, site, "merged", "cdw_rawdata.csv", sep="_"), sep = "/"
```

siteid	taxonid	decayClassNum	diameterClass	counts	totalLogs	relativeAbundance	cumulativeAbundance
SERC	LITU	3	>=10cm	59	414	14.25	14.25
SERC	2PLANT-H	3	>=10cm	51	414	12.32	26.57
SERC	2PLANT-H	4	>=10cm	51	414	12.32	38.89
SERC	LITU	4	>=10cm	20	414	4.83	43.72
SERC	QUERC	3	>=10cm	15	414	3.62	47.34
SERC	LITU	2	>=10cm	14	414	3.38	50.72
SERC	LIST2	3	>=10cm	12	414	2.90	53.62
SERC	QUAL	3	>=10cm	11	414	2.66	56.28
SERC	QUAL	1	>=10cm	10	414	2.42	58.70
SERC	2PLANT-H		5-10cm	10	414	2.42	61.12
SERC	QUVE	2	>=10cm	8	414	1.93	63.05
SERC	PIVI2	3	>=10cm	8	414	1.93	64.98
SERC	PIST	3	>=10cm	7	414	1.69	66.67
SERC	2PLANT-H	5	>=10cm	7	414	1.69	68.36
SERC	LITU		5-10cm	7	414	1.69	70.05
SERC	FAGR	2	>=10cm	6	414	1.45	71.50
SERC	PRSE2	3	>=10cm	6	414	1.45	72.95
SERC	2PLANT-S	4	>=10cm	6	414	1.45	74.40
SERC	LIST2	2	>=10cm	5	414	1.21	75.61
SERC	QURU	2	>=10cm	5	414	1.21	76.82
SERC	2PLANT-S	3	>=10cm	5	414	1.21	78.03
SERC	LIST2		5-10cm	5	414	1.21	79.24
SERC	CARYA	3	>=10cm	4	414	0.97	80.21
SERC	FAGR	3	>=10cm	4	414	0.97	81.18
SERC	2PLANT-S	5	>=10cm	4	414	0.97	82.15
SERC	LIST2	1	>=10cm	3	414	0.72	82.87
SERC	LITU	1	>=10cm	3	414	0.72	83.59
SERC	QURU	1	>=10cm	3	414	0.72	84.31
SERC	ACRU	2	>=10cm	3	414	0.72	85.03
SERC	LITU		2-5cm	3	414	0.72	85.75
SERC	ACRU		5-10cm	3	414	0.72	86.47
SERC	JUVI	2	>=10cm	2	414	0.48	86.95
SERC	PIST	2	>=10cm	2	414	0.48	87.43
SERC	QUAL	2	>=10cm	2	414	0.48	87.91
SERC	QUPH	2	>=10cm	2	414	0.48	88.39
SERC	ACRU	3	>=10cm	2	414	0.48	88.87
SERC	AIAL	3	>=10cm	2	414	0.48	89.35
SERC	CACA18	3	>=10cm	2	414	0.48	89.83
SERC	COFL2	3	>=10cm	2	414	0.48	90.31
SERC	JUVI	3	>=10cm	2	414	0.48	90.79
SERC	PINUS	3	>=10cm	2	414	0.48	91.27
SERC	SAAL5	3	>=10cm	2	414	0.48	91.75
SERC	ACRU		2-5cm	2	414	0.48	92.23
SERC	PRSE2		5-10cm	2	414	0.48	92.71
SERC	ACRU	1	>=10cm	1	414	0.24	92.95
SERC	PIVI2	1	>=10cm	1	414	0.24	93.19
SERC	QUVE	1	>=10cm	1	414	0.24	93.43
SERC	SAAL5	1	>=10cm	1	414	0.24	93.67

siteid	taxonid	decayClassNum	diameterClass	counts	totalLogs	relativeAbundance	cumulativeAbundance
SERC	2PLANT-H	2	>=10cm	1	414	0.24	93.91
SERC	CAGL8	2	>=10cm	1	414	0.24	94.15
SERC	PIEC2	2	>=10cm	1	414	0.24	94.39
SERC	PIVI2	2	>=10cm	1	414	0.24	94.63
SERC	PRSE2	2	>=10cm	1	414	0.24	94.87
SERC	QUERC	2	>=10cm	1	414	0.24	95.11
SERC	ULMUS	2	>=10cm	1	414	0.24	95.35
SERC	CATO6	3	>=10cm	1	414	0.24	95.59
SERC	PLOC	3	>=10cm	1	414	0.24	95.83
SERC	QUERCSPP	3	>=10cm	1	414	0.24	96.07
SERC	QUFA	3	>=10cm	1	414	0.24	96.31
SERC	QURU	3	>=10cm	1	414	0.24	96.55
SERC	QUVE	3	>=10cm	1	414	0.24	96.79
SERC	ROPS	3	>=10cm	1	414	0.24	97.03
SERC	JUVI	4	>=10cm	1	414	0.24	97.27
SERC	LIST2	4	>=10cm	1	414	0.24	97.51
SERC	PIST	4	>=10cm	1	414	0.24	97.75
SERC	PIVI2	4	>=10cm	1	414	0.24	97.99
SERC	QUAL	4	>=10cm	1	414	0.24	98.23
SERC	QUERC	4	>=10cm	1	414	0.24	98.47
SERC	2PLANT-H		2-5cm	1	414	0.24	98.71
SERC	COFL2		2-5cm	1	414	0.24	98.95
SERC	LIST2		2-5cm	1	414	0.24	99.19
SERC	CARYA		5-10cm	1	414	0.24	99.43
SERC	COFL2		5-10cm	1	414	0.24	99.67
SERC	PIST		5-10cm	1	414	0.24	99.91

Warning: Using size for a discrete variable is not advised.

Rank Abundance of decayClass x sizeCategory x taxonID (DST)



Code

```
## Windows paths
# inpath <- "C:/Users/cflagg/Documents/GitHub/neonPlantSampling/cdw_tally_analysis/D07/"
# outpath <- inpath

## macOS inputs
if(file.exists("~/Documents/workDocuments")){
  inpath <- "~/Documents/workDocuments/gitRepositories/neonPlantSampling/cdw_tally_analysis/fulcrumData"
  domain <- "D02"
  site <- "SERC"
  outpath <- paste("~/Documents/workDocuments/gitRepositories/neonPlantSampling/cdw_tally_analysis/", domain, site)
}

if(file.exists("~/Documents/neonScienceDocs")){
  inpath <- "~/Documents/neonScienceDocs/gitRepositories/neonPlantSampling/cdw_tally_analysis/fulcrumData"
  domain <- "D02"
  site <- "SERC"
  outpath <- paste("~/Documents/neonScienceDocs/gitRepositories/neonPlantSampling/cdw_tally_analysis/", domain, site)
}

## import libraries
library(plyr)
library(dplyr)
```

```

library(ggplot2)

## import data
parent_cdw <- read.csv(paste(inpath,"tos_coarse_downed_wood_tally_prod.csv",sep=""), header=TRUE)
child_cdw <- read.csv(paste(inpath,"tos_coarse_downed_wood_tally_prod_per_plot_azimuth_log.csv",sep=""))

## filter out parent fields that are unneeded
parent_cdw_fil <- dplyr::select(parent_cdw,
                                fulcrum_id, domainid, siteid, plotid_parent, tallydate, volumefactor, parti
## filter out child fields that are unneeded
child_cdw_fil <- dplyr::select(child_cdw,
                                fulcrum_parent_id, lidsazimuth, logdiameter, taxonid, decayclass, logid_ingre

## join parent to child via parent::fulcrum_id to child::fulcrum_parent_id
cdw <- merge(x = parent_cdw_fil, y = child_cdw_fil, by.x = "fulcrum_id", by.y = "fulcrum_parent_id", all

## filter to the selected domain, then site only
cdw <- dplyr::filter(cdw, domainid==domain, siteid==site)

## create diameter class factor
cdw$diameterClass <- ifelse(cdw$logdiameter >= 10, '>=10cm',
                           ifelse(cdw$logdiameter <5, "2-5cm", "5-10cm"))

## simplify decayclass to numeric value wrapped inside sapply(decayClassNum, "[", 1) e.g. 'return first
cdw$decayClassNum <- sapply(stringr::str_split(cdw$decayclass, pattern = " "), "[", 1)

## write file
write.csv(cdw, file = paste(outpath, paste(domain, site, "merged", "cdw_rawdata.csv", sep="_"), sep = "/"

#names(cdw)
## table 1 -- count of taxonIDs, report top 10?
## order ascending output by totalTaxa length() will count length of a vector

## Filter out 'blank' transects where TTP = NO
cdw_fil <- dplyr::filter(cdw, targettaxapresent == "Y")
t1 <- ddply(cdw_fil, ~siteid+taxonid+decayClassNum+diameterClass, summarize, counts = length(taxonid))

# an alternative to ddply call, returns similar result but output column name is defaulted
# plyr::count(cdw, c("siteid", "taxonid", "decayClassNum", "diameterClass"))

## Sort rows by siteID then by descending number of counts
t1 <- dplyr::arrange(t1, siteid, desc(counts), diameterClass, decayClassNum, taxonid)

## Calculate relative abundances
t1$totalLogs <- nrow(cdw_fil)
t1$relativeAbundance <- round((t1$counts/t1$totalLogs)*100,2)
t1$cumulativeAbundance <- round(cumsum(t1$relativeAbundance),2)

## determine raw ranking by count of DST combo within siteid
t1 <- plyr::ddply(t1, ~siteid, mutate, DSTrank = as.numeric(paste(1:length(counts))), DSTcombo = paste0
t2 <- ddply(cdw, ~siteid+taxonid, summarize, counts = length(taxonid))

## Determine required sample size per DST combo, and cumulative sample size

```

```
t1$sampledDiskNum <- ifelse(t1$diameterClass == ">=10cm", 10,5)
t1$cumulativeDiskNum <- cumsum(t1$sampledDiskNum)
```

```
## species abundance/total abundance
```

```
ggplot(data = t1, aes(x = DStrank, y = counts, color = taxonid, shape = diameterClass, size = decayClass
```