CTFS R FORMATTED DATA TABLES

R. CONDIT

STEM TABLES

These tables are stored in files *Plotname*.stem#.rdata, where # refers to the census number. The R object within is a dataframe of the same name. There is one file and one dataframe for each census.

All the stem dataframes from one plot have exactly the same number of rows, with one row for every stem ever found in any census, in exactly the same order. Stems thus appear before they recruit, while they are alive, and after they die. See the description of the full table, which has a similar structure, but for trees not stems.

0.1. **Census 5 from BCI data.** Attach the Rdata set. (The folder defined by 'mydatapath' is where it is stored on my computer.)

```
> mystemdatapath='/home/fullplotdata/stem/'
> if(!exists('bci.stem5'))
+   attach(paste(mystemdatapath,'bci.stem5.rdata',sep=''))
> str(bci.stem5)
'data.frame':        770887 obs. of  20 variables:
 $ treeID   : int  1 2 3 4 5 6 7 8 9 10 ...
 $ stemID   : int  1 1 1 1 1 1 1 1 1 1 ...
 $ tag      : chr  "-05599" "-22851" "-24362" "-26589" ...
 $ StemTag  : chr  "a" "a" "a" "a" ...
 $ sp       : chr  "swars1" "hybapr" "aegipa" "beilpe" ...
 $ quadrat  : chr  "4007" "0718" "0417" "0007" ...
 $ gx       : num  800.2 151.5 95.2 11.7 7.7 ...
 $ gy       : num  152.2 378.8 357.5 151.1 96.2 ...
 $ MeasureID: int  NA NA NA NA NA NA NA NA NA 1 ...
 $ CensusID : int  NA NA NA NA NA NA NA NA NA 5 ...
 $ dbh      : num  NA NA NA NA NA NA NA NA NA NA ...
 $ pom      : chr  NA NA NA NA ...
 $ hom      : num  NA NA NA NA NA NA NA NA NA NA ...
 $ ExactDate: chr  NA NA NA NA ...
 $ DFstatus : chr  "dead" "dead" "dead" "dead" ...
 $ codes    : chr  NA NA NA NA ...
 $ countPOM : num  NA NA NA NA NA NA NA NA NA 1 ...
 $ status   : chr  "D" "D" "D" "D" ...
 $ date     : num  14801 14656 14666 14636 14629 ...
 $ agb      : num  NA NA NA NA NA NA NA NA NA NA ...
```

Date: September 20, 2010.

Description of columns in the *stem* R format tables. Each row is data from a single stem. The biomass in this table is for the individual stem.

Column name	Description
treeID	The unique tree identifier in CTFS database. Useful to be certain in matching trees.
stemID	The unique stem identifier in CTFS database. Useful to be certain in matching stems.
tag	Tag number used in the field.
StemTag	Tag number on the individual stem, if present.
sp	The species mnemonic. To get full species names, the taxonomy table must be downloaded from the CTFS database: http://ctfs.arnarb.harvard.edu/CTFSReports
quadrat	Quadrat designation.
gx	The x coordinate within the plot, relative to one edge of the plot.
gy	The y plot coordinate.
MeasureID	The unique identifier of a single measurement in the CTFS database.
CensusID	The numeric identifier of the census.
dbh	Diameter of the stem.
pom	The point-of-measure, where the diameter was taken, identical to hom, but a character variable with only 2 decimal places.
hom	The height-of-measure, identical to pom but a numeric variable with full precision.
ExactDate	The date on which the stem was measured.
DFstatus	The status taken from the DFtemp table of the CTFS database: alive, dead, lost_stem, missing, or prior[1]
codes	The codes for the measurement as recorded in the field.
countPOM	The number of POMs (HOMs) for the same stem in this census.
status	An abbreviated version of status, for compatibility with earlier versions of functions in the CTSF R package: A, D, M, P
date	The julian date, for date arithmetic.
agb	Above-ground-biomass of the stem, in Mg (= metric tons or 10^6 g). Some are NA.

[1] Status of stems is not as clear-cut as for trees. Stems have status dead (D) only if the entire tree is dead. Stems are alive (A) only if the tree is alive and the stem has a dbh. Status = 'lost_stem' indicates that was broken in the given census, while the tree has no other stem. Status = 'missing' (M) are cases where dbh and codes for a tree were not recorded (or lost), so it is not certain whether the tree was alive or dead. Status = 'prior' (P) indicates a tree had not yet recruited into the plot during the census. An additional status for stems is gone (G), in cases where a tree is alive and a stem which formerly had a measurement does not in this census. This happens in some plots, where individual stems do not have tags and cannot be tracked.

Sample queries of tree tables from two censuses, using BCI census 5 and 6 to illustrate.

```
> if(!exists('bci.stem6'))
+   attach(paste(mystemdatapath,'bci.stem6.rdata',sep=''))
```

Because rows match, it is possible to check the status in two censuses of all trees. This shows that of all trees alive in census 5 at BCI, 162655 were still alive in census 6, while 28432 had died.

```
> table(bci.stem5$status,bci.stem6$status,exclude=NULL)
```

	A	D	G	M	<NA>
A	162655	28432	72102	0	0
D	67	117278	0	0	0
G	0	0	294004	0	0
M	0	1	0	52	0
P	96296	0	0	0	0
<NA>	0	0	0	0	0

The number of stems with given diameters:

```
> head(table(subset(bci.stem5,status=='A')$dbh),12)
```

10	11	12	13	14	15	16	17	18	19	20	21
6432	11957	12876	12061	11097	10755	9308	8764	8203	7085	7112	6399

Again, with matching rows, growth is easy to calculate (multiplying by the number of days in a year, since the date field is in days.) This is biomass growth, and the fastest is 7.5 Mg/year (and an obvious error).

```
> growth=round(365.25*(bci.stem6$agb-bci.stem5$agb)/(bci.stem6$date-bci.stem5$date),2)
> head(sort(growth,decreasing=TRUE),10)
 [1] 7.47 7.15 5.57 2.68 2.52 2.09 1.88 1.76 1.46 1.39
> subset(bci.stem5,growth>7.4,
+       select=c('tag','stemID','sp','dbh','agb','pom','ExactDate','status'))
       tag stemID    sp dbh      agb pom ExactDate status
8877 005143      8 ficuc2 1066 10.87051 3.80 2000-08-14      A
> subset(bci.stem6,growth>7.4,
+       select=c('tag','stemID','sp','dbh','agb','pom','ExactDate','status'))
       tag stemID    sp dbh     agb pom ExactDate status
8877 005143      8 ficuc2 2177 48.542 3.80 2005-08-31      A
```