

# Development of an FIA dataset to model tree-level changes in the Northern Forest

*Neal Maker*

*December 08, 2019*

## Introduction

The intent of this project is to put together a dataset of remeasured trees from the US Forest Service's Forest Inventory and Analysis (FIA) records in the Northern Forest region, which will allow tree-level changes to be modeled accurately. Specifically, the dataset should support the development of unbiased models of dbh change, height, height change, crown ratio change, and mortality; and the determination as to which variables are the most important predictors of those outcomes. Potential predictors to be kept in the dataset are those that 1) are widely available in the FIA data in the region (to maintain large sample sizes), and 2) could be recorded in forest inventories or remotely without large increases in inventory costs.

The Northern Forest region was chosen because it covers a fairly large geographic extent while still representing a coherent ecological region, in which trees can be expected to follow a similar set of behaviors. Models developed with the dataset should be relatively unbiased for individual forests within the region, but will still allow for streamlined analyses across disparate ownerships. The US Northern Forest is defined here as including Oswego, Oneida, Lewis, Jefferson, Saint Lawrence, Herkimer, Fulton, Hamilton, Franklin, Essex, Clinton, and Warren Counties in New York; Franklin, Orleans, Essex, Chittenden, Lamoille, Caledonia, Washington, Addison, Orange, and Grand Isle Counties in Vermont; Coos, Grafton, and Carroll Counties in New Hampshire; and Oxford, Franklin, Somerset, Androscoggin, Kennebec, Waldo, Hancock, Washington, Penobscot, Piscataquis, and Aroostook Counties in Maine.

```
library("tidyverse")

## -- Attaching packages ----- tidyverse 1.2.1 --
## v ggplot2 3.1.0      v purrr   0.2.5
## v tibble  1.4.2      v dplyr  0.7.8
## v tidyr   0.8.2      v stringr 1.3.1
## v readr   1.3.0      v forcats 0.3.0

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()

library("lubridate")

##
## Attaching package: 'lubridate'
##
## The following object is masked from 'package:base':
##
##     date

library("english")

## Warning: package 'english' was built under R version 3.5.3

# Define States & counties (FIPS codes) in Northern Forest region -----
states <- c("NY", "VT", "NH", "ME")
```

```

NY_counties <- c(75, 65, 49, 45, 89, 43, 35, 41, 33, 31, 19, 113)
VT_counties <- c(11, 19, 9, 7, 15, 5, 23, 1, 17, 13)
NH_counties <- c(7, 9, 3)
ME_counties <- c(17, 7, 25, 1, 11, 27, 9, 29, 19, 21, 3)

```

## Methods

FIA data were downloaded from the FIA DataMart<sup>1</sup> in the form of state-specific csv files, which were generated by the Forest Service from the FIA Oracle database tables. These data are current as of December 08, 2019. In the future, the dataset can be recreated using updated csv files to incorporate new remeasurement data.

```

# Fetch FIA tree, growth, plot, & condition data for Northern Forest states
# and filter to keep only northern forest counties
# (this may take a while; ~140MB of downloads + reading)

temp <- tempfile()

for(state in states){
  download.file(paste("https://apps.fs.usda.gov/fia/datamart/CSV/", state, "_TREE.zip", sep = ""),
               temp, mode = "wb")
  unzip(temp, paste(state, "_TREE.csv", sep = ""))
}

TREE <- lapply(states, function(x){
  read.csv(paste(x, "_TREE.csv", sep = ""), header = T) %>%
    filter(COUNTYCD %in% eval(as.name(paste(x, "_counties", sep = ""))),
           DIAHTCD == 1) %>% # excludes seedlings measured at root collar
  select(CN, PLT_CN, SUBP, PREV_TRE_CN, CONDID, DIA, SPCD, STATUSCD,
         MORTYR, CR, CCLCD, TREECLCD, HT) %>%
    mutate(ba_ac = if_else(DIA >= 5,
                           # poles & larger from 24' radius subplots
                           # saplings from 6.8' radius microplots
                           0.005454*DIA^2*(43560/(pi*24^2)),
                           0.005454*DIA^2*(43560/(pi*6.8^2))))
})

## SUBPLOTS MATTER!!!!
# Subplots may have different sizes depending on the plot design, found in
# PLOT$DESIGNCD (see database guide, appendix i)
# I can just keep DESIGNCD == 1 (the main standard) and lose some data,
# or I can account for the various designs when I calculate TREE$ba_ac
# (above; which would mean calculating ba_ac after combining
# states' data and joining nf_trees to nf_plots).

# DESIGN CODES:
# 1:4 used 1999 - present
# 11:15 used 1994 - 1996
# 100 used 1982 & 1983
# 101:104 used 1991 - 1998
# 101 was continued through 2008

```

<sup>1</sup><https://apps.fs.usda.gov/fia/datamart/datamart.html>

```
# 105:120 variously used 1991 - 1993
```

```
for(state in states){
  download.file(paste("https://apps.fs.usda.gov/fia/datamart/CSV/", state,
                      "_PLOT.zip", sep = ""),
               temp, mode = "wb")
  unzip(temp, paste(state, "_PLOT.csv", sep = ""))
}

PLOT <- lapply(states, function(x){
  read.csv(paste(x, "_PLOT.csv", sep = ""), header = T) %>%
    filter(COUNTYCD %in% eval(as.name(paste(x, "_counties", sep = "")))) %>%
    select(CN, PREV_PLT_CN, DESIGNCD, MEASYEAR, MEASMON,
           MEASDAY, LAT, LON, ELEV) %>%
    rename(PLT_CN = CN)
})

for(state in states){
  download.file(paste("https://apps.fs.usda.gov/fia/datamart/CSV/", state,
                      "_COND.zip", sep = ""),
               temp, mode = "wb")
  unzip(temp, paste(state, "_COND.csv", sep = ""))
}

COND <- lapply(states, function(x){
  read.csv(paste(x, "_COND.csv", sep = ""), header = T) %>%
    filter(COUNTYCD %in% eval(as.name(paste(x, "_counties", sep = "")))) %>%
    select(PLT_CN, CONDID, BALIVE, FORTYPCD, ALSTKCD, SITECLCD,
           PHYSCLCD, SLOPE, ASPECT)
})

for(state in states){
  download.file(paste("https://apps.fs.usda.gov/fia/datamart/CSV/", state,
                      "_TREE_GRM_COMPONENT.zip", sep = ""),
               temp, mode = "wb")
  unzip(temp, paste(state, "_TREE_GRM_COMPONENT.csv", sep = ""))
}

GRM <- lapply(states, function(x){
  read.csv(paste(x, "_TREE_GRM_COMPONENT.csv", sep = ""), header = T) %>%
    filter(!is.na(ANN_DIA_GROWTH)) %>%
    select(TRE_CN, STATECD, DIA_BEGIN, DIA_MIDPT, DIA_END, ANN_DIA_GROWTH,
           HT_BEGIN, HT_MIDPT, HT_END, ANN_HT_GROWTH)
})

# Combine states' data

nf_trees <- rbind(TREE[[1]], TREE[[2]], TREE[[3]], TREE[[4]])
nf_plots <- rbind(PLOT[[1]], PLOT[[2]], PLOT[[3]], PLOT[[4]])
nf_conds <- rbind(COND[[1]], COND[[2]], COND[[3]], COND[[4]])
```

```

nf_grms <- rbind(GRM[[1]], GRM[[2]], GRM[[3]], GRM[[4]])

# delete temporary objects and downloaded files

unlink(temp)

remove(TREE, PLOT, COND, GRM, temp, state)

for(state in states){
  file.remove(paste(state, "_TREE.csv", sep = ""))
  file.remove(paste(state, "_PLOT.csv", sep = ""))
  file.remove(paste(state, "_COND.csv", sep = ""))
  file.remove(paste(state, "_TREE_GRM_COMPONENT.csv", sep = ""))
}

```

Variables that were retained from the FIA tables include tree species, dbh, status (alive or dead), class (growing stock, rough cull, or rotten cull), uncompacted crown ratio, crown class, plot basal area, forest type, stocking, site class, physical landscape type (eg. ‘deep sands’ or ‘rolling uplands’), slope, aspect, elevation, latitude, and longitude. In addition, overtopping basal area was calculated for each tree.

```

# Calculates overtopping basal area (BAL) assuming all input trees are in
# same plot and ba is adjusted based on tpa:
pbal <- function(dbh, ba){
  sapply(dbh, function(x){
    index <- dbh > x
    return(sum(ba[index]))
  })
}

# Add BAL
nf_trees <- nf_trees %>%
  mutate(bal = NA)

nf_trees[nf_trees$STATUSCD == 1,] <- nf_trees[nf_trees$STATUSCD == 1,] %>%
  group_by(PLT_CN, SUBP) %>%
  mutate(bal = pbal(DIA, ba_ac))

```

All trees in the dataset are saplings or larger (with diameters measured at breast height). Records for trees without remeasurement data were discarded, along with records for trees that were already dead at their starting measurement and trees that were incorrectly inventoried during their starting or ending measurements. Also, data from some earlier inventory years were excluded because they were obtained with different plot designs, which would have complicated the data wrangling. These data could be incorporated into the dataset, but the old plot designs would need to be researched and accounted for deliberately to keep basal area per acre calculations for each sampled tree accurate.

```

# get before and after data and remeasurement period for each tree and -----
# add initial cond and plot data -----

nf_end <- nf_trees %>%
  filter(PREV_TRE_CN > 0) %>%
  left_join(nf_plots, by = "PLT_CN") %>%
  left_join(nf_conds, by = c("PLT_CN", "CONDID")) %>%
  rename(cn_e = CN, plt_cn_e = PLT_CN,
         condid_e = CONDID,

```

```

    dbh_e = DIA,
    statuscd_e = STATUSCD,
    mortyr_e = MORTYR,
    cr_e = CR,
    crown_class_e = CCLCD,
    tree_class_e = TREECLCD,
    MEASYEAR_E = MEASYEAR,
    MEASMON_E = MEASMON,
    MEASDAY_E = MEASDAY,
    ba_e = BALIVE,
    bal_e = bal,
    ht_e = HT,
    forest_type_e = FORTYPCD,
    stocking_e = ALSTKCD,
    site_class_e = SITECLCD,
    landscape_e = PHYSCLCD,
    slope_e = SLOPE,
    aspect_e = ASPECT,
    designcd_e = DESIGNCD) %>%
select(-SPCD)

nf_start <- nf_trees %>%
  filter(CN %in% nf_end$PREV_TRE_CN) %>%
  left_join(nf_plots, by = "PLT_CN") %>%
  left_join(nf_conds, by = c("PLT_CN", "CONDID")) %>%
  rename(cn_s = CN, plt_cn_s = PLT_CN, condid_s = CONDID, dbh_s = DIA,
    statuscd_s = STATUSCD, mortyr_s = MORTYR, cr_s = CR,
    crown_class_s = CCLCD, tree_class_s = TREECLCD,
    MEASYEAR_S = MEASYEAR, MEASMON_S = MEASMON, MEASDAY_S = MEASDAY,
    ba_s = BALIVE, bal_s = bal, ht_s = HT,
    forest_type_s = FORTYPCD,
    stocking_s = ALSTKCD, site_class_s = SITECLCD,
    landscape_s = PHYSCLCD, slope_s = SLOPE, aspect_s = ASPECT,
    designcd_s = DESIGNCD) %>%
  select(-PREV_TRE_CN, -PREV_PLT_CN, -LAT, -LON, -ELEV)

nf_fia <- nf_end %>%
  left_join(nf_start, by = c("PREV_TRE_CN" = "cn_s")) %>%
  filter(statuscd_s == 1, # only trees that started live
    statuscd_e != 0, # remove trees that were remeasured incorrectly
    cr_s >= 0, # only trees that had cr at start
    designcd_s == 1, # only those with current plot design
    designcd_e == 1) %>%
  mutate(MEASMON_E = formatC(MEASMON_E, width = 2, format = "d", flag = "0"),
    MEASMON_S = formatC(MEASMON_S, width = 2, format = "d", flag = "0"),
    MEASDAY_E = formatC(MEASDAY_E, width = 2, format = "d", flag = "0"),
    MEASDAY_S = formatC(MEASDAY_S, width = 2, format = "d", flag = "0"),
    #make month and day codes 2 digits
    date_s = ymd(paste(MEASYEAR_S, MEASMON_S, MEASDAY_S, sep = "")),
    date_e = ymd(paste(MEASYEAR_E, MEASMON_E, MEASDAY_E, sep = ""))) %>%
  filter(!is.na(date_e), !is.na(date_s)) %>%
  # remove incorrectly entered dates (eg. Feb 31)
  filter(ba_s < 500, ba_e < 500) %>% # ba's > 500 were found to be mistakes

```

```

mutate(interval = as.double(as.period(date_e - date_s), unit = "years"),
       cr_rate = (cr_e - cr_s)/interval,
       cr_mid = (cr_e + cr_s)/2,
       dbh_rate = (dbh_e - dbh_s)/interval,
       dbh_mid = (dbh_e + dbh_s)/2,
       ba_mid = (ba_e + ba_s)/2,
       bal_mid = (bal_e + bal_s)/2,
       ht_mid = (ht_e + ht_s)/2,
       ht_rate = (ht_e - ht_s)/interval,
       status_change = case_when(statuscd_e == 1 ~ "lived",
                                statuscd_e == 2 ~ "died",
                                statuscd_e == 3 ~ "cut",
                                TRUE ~ "error"),
       status_change = as.factor(status_change),
       SPCD = as.factor(SPCD),
       plt_cn_e = as.factor(plt_cn_e)) %>%
select(cn_e, spp = SPCD, dbh_e, cr_s, cr_mid,
       cr_e, cr_rate, crown_class_s, crown_class_e, tree_class_s,
       tree_class_e, ba_s, ba_mid, ba_e,
       bal_s, bal_mid, bal_e, ht_s, ht_mid, ht_e, ht_rate,
       forest_type_s, forest_type_e, stocking_s, stocking_e,
       landscape_s, landscape_e, site_class_s, site_class_e,
       slope_s, slope_e, aspect_s, aspect_e, lat = LAT, lon = LON,
       elev = ELEV, date_s, date_e, interval, status_change,
       plot = plt_cn_e) %>% # mortality year was all null and was removed
inner_join(nf_grms, by = c("cn_e" = "TRE_CN")) %>%
rename(dbh_s = DIA_BEGIN, dbh_mid = DIA_MIDPT,
       dbh_rate = ANN_DIA_GROWTH, state = STATECD)

```

```
## Warning: 3675 failed to parse.
```

```

remove(nf_start, nf_end, nf_conds, nf_plots, nf_trees, nf_grms, states,
       VT_counties, NH_counties, NY_counties, ME_counties, pbal, state)

```

For variables that naturally change from one measurement to another, like crown class and stocking, starting and ending values are recorded. Midpoint values are also recorded for dbh, crown ratio, basal area, and overtopping basal area; to better reflect average conditions during the remeasurement period. They were calculated by averaging the starting and ending values. In addition, annual rates of change are recorded for dbh and crown ratio, to be used as the dependent (outcome) variables in those respective models.

Some of the ostensibly fixed variables (like slope, aspect, and site class) were found to change from one measurement to another in a minority of instances. For example, aspect was recorded differently in nine percent of remeasurements, slope was recorded differently in 17 percent of remeasurements, and site class was recorded differently in four percent of remeasurements, despite the fact that they were measured on the same plots and should have remained constant. The differences between starting and ending values are generally small, however, and are probably measurement errors. In the case of slope, the mean absolute difference of deviating measurements is only four percent. Among erroneous site class measures, the average is only one site class. Aspect errors tend to be higher, averaging 87 degrees, but they can be attributed to the difficulty of determining aspects in relatively flat terrain. If only plots with slopes over 20 percent are considered, the mean absolute aspect error falls to 34 degrees. All these discontinuities were assumed to be random measurement errors, and starting values were arbitrarily retained in the dataset while ending values were discarded.

```

# Keep starting values for fixed variables & -----
# only keep records with all necessary fields -----

nf_fia <- nf_fia %>%
  rename(landscape = landscape_s,
         site_class = site_class_s,
         slope = slope_s,
         aspect = aspect_s) %>%
  select(-landscape_e, -site_class_e, -slope_e, -aspect_e) %>%
  filter(!is.na(spp), # only keep records with all necessary fields
         !is.na(dbh_s),
         !is.na(dbh_mid),
         !is.na(dbh_e),
         !is.na(crown_class_s),
         !is.na(tree_class_s),
         !is.na(ba_s),
         !is.na(ba_mid),
         !is.na(ba_e),
         !is.na(bal_s),
         !is.na(forest_type_s),
         !is.na(forest_type_e),
         !is.na(stocking_s),
         !is.na(stocking_e),
         !is.na(landscape),
         !is.na(site_class),
         !is.na(slope),
         !is.na(aspect),
         !is.na(lat),
         !is.na(lon),
         !is.na(elev),
         !is.na(status_change),
         xor(is.na(bal_e), status_change == "lived"))

```

In addition to renaming the variables in the dataset, FIA codes for the levels of categorical variables were replaced with descriptive strings to make them more intuitive and user-friendly. Tree species were also grouped into species groups, and FIA forest types were grouped into more general forest types so they match common inventory protocols. For example, most species in the genus *Populus* are combined into a single “aspen” group to make data collection easier, although cottonwoods (*Populus deltoides*) are kept in their own group because they exhibit very different growth characteristics. Similarly, the FIA forest types “balsam fir”, “white spruce”, “red spruce”, “red spruce/balsam fir”, and “black spruce” were combined into a single “Spruce-fir” group, but “northern white-cedar” was kept in its own “Cedar” group.

```

# Make names and factor levels more intuitive -----

species_codes <-
  c(12, 43, 68, 70, 71, 91, 94, 95, 96, 97, 105, 123, 125, 126, 129,
    130, 136, 202, 221, 241, 261, 310, 313, 314, 315, 316, 317, 318,
    319, 320, 331, 341, 355, 356, 357, 367, 370, 371, 372, 373, 375,
    379, 391, 400, 402, 403, 407, 409, 421, 462, 491, 500, 531, 540,
    541, 543, 544, 546, 552, 601, 602, 621, 651, 655, 660, 661, 663,
    680, 693, 701, 712, 731, 741, 742, 743, 744, 746, 760, 761, 762,
    763, 764, 771, 802, 804, 806, 816, 823, 832, 833, 837, 901, 920,
    922, 923, 926, 934, 935, 936, 937, 950, 951, 970, 972, 975, 977,
    999)

```

```

species <-
  c("fir", "other softwood", "cedar", "tamarack", "tamarack",
    "norway spruce", "spruce", "spruce", "spruce", "spruce", "other softwood",
    "other softwood", "red pine", "other softwood", "white pine", "scots pine",
    "other softwood", "other softwood", "other softwood", "cedar", "hemlock",
    "other hardwood", "soft maple", "hard maple", "striped maple", "soft maple",
    "soft maple", "hard maple", "other hardwood", "hard maple",
    "other hardwood", "other hardwood", "other hardwood", "other hardwood",
    "other hardwood", "other hardwood", "other hardwood", "yellow birch",
    "other hardwood", "other hardwood", "paper birch", "other hardwood",
    "other hardwood", "hickory", "hickory", "hickory", "hickory", "hickory",
    "other hardwood", "other hardwood", "other hardwood", "other hardwood",
    "beech", "ash", "ash", "ash", "ash", "ash", "other hardwood", "butternut",
    "other hardwood", "other hardwood", "other hardwood", "other hardwood",
    "other hardwood", "other hardwood", "other hardwood", "other hardwood",
    "other hardwood", "hophornbeam", "other hardwood", "other hardwood",
    "aspen", "cottonwood", "aspen", "cottonwood", "aspen", "other hardwood",
    "other hardwood", "black cherry", "other hardwood", "other hardwood",
    "other hardwood", "white oak", "white oak", "red oak", "white oak",
    "white oak", "white oak", "red oak", "red oak", "other hardwood",
    "other hardwood", "other hardwood", "other hardwood", "other hardwood",
    "other hardwood", "other hardwood", "other hardwood", "other hardwood",
    "basswood", "basswood", "elm", "elm", "elm", "elm", "other hardwood")

names(species) <- as.character(species_codes)

nf_fia$spp <- factor(unname(species[as.character(nf_fia$spp)]),
  levels = levels(factor(species))) # standardize levels

#-----

forest_type_codes <-
  c(101, 102, 103, 104, 105, 121, 122, 123, 124, 125, 126, 127,
    167, 171, 381, 384, 385, 401, 402, 409, 503, 505, 509, 512,
    513, 515, 516, 517, 519, 520, 701, 702, 703, 704, 705, 706,
    707, 708, 709, 801, 802, 805, 809, 901, 902, 903, 904, 905,
    962, 995, 999)

forest_types <-
  c("Red pine", "Red pine", "White pine", "Mixed softwood", "Hemlock",
    "Spruce-fir", "Spruce-fir", "Spruce-fir", "Spruce-fir", "Spruce-fir",
    "Larch", "Cedar", "Mixed softwood", "Mixed softwood", "Scots pine",
    "Norway spruce", "Larch", "Pine-hardwood", "Mixedwood",
    "Pine-hardwood", "Oak-hickory", "Oak-hickory", "Oak-hickory",
    "Transition hardwood", "Transition hardwood", "Oak-hickory",
    "Transition hardwood", "Transition hardwood", "Northern hardwood",
    "Northern hardwood", "Northern hardwood", "Transition hardwood",
    "Cottonwood", "Other", "Other", "Other", "Northern hardwood",
    "Northern hardwood", "Cottonwood", "Northern hardwood",
    "Northern hardwood", "Northern hardwood", "Northern hardwood",
    "Northern hardwood", "Northern hardwood", "Northern hardwood",
    "Northern hardwood", "Northern hardwood", "Other", "Other",
    "Nonstocked")

```



```

names(forest_types) <- as.character(forest_type_codes)

nf_fia$forest_type_s <-
  factor(unnname(forest_types[as.character(nf_fia$forest_type_s)]),
    levels = levels(factor(forest_types)))

nf_fia$forest_type_e <-
  factor(unnname(forest_types[as.character(nf_fia$forest_type_e)]),
    levels = levels(factor(forest_types)))

#-----

landscape_codes <- # add 19 & 33
  c(11, 12, 13, 19, 21, 22, 23, 24, 25, 29, 31, 32, 33, 34, 39)

landscapes <-
  c("dry tops", "dry slopes", "deep sands", "other xeric", "flatwoods",
    "rolling uplands", "moist slopes & coves", "narrow floodplains/bottomlands",
    "broad floodplains/bottomlands", "other mesic", "swamps/bogs",
    "small drains", "small drains", "beaver ponds", "other hydric")

names(landscapes) <- as.character(landscape_codes)

nf_fia$landscape <- factor(unnname(landscapes[as.character(nf_fia$landscape)]),
  levels = levels(factor(landscapes)))

#-----

remove(forest_type_codes, forest_types, landscape_codes,
  landscapes, species, species_codes)

# put columns in order
nf_fia <- nf_fia %>%
  select(spp, dbh_s, dbh_mid, dbh_e, dbh_rate, cr_s, cr_mid, cr_e, cr_rate,
    crown_class_s, crown_class_e, tree_class_s, tree_class_e,
    ba_s, ba_mid, ba_e, bal_s, bal_mid, bal_e, ht_s, ht_mid, ht_e,
    ht_rate, forest_type_s, forest_type_e,
    stocking_s, stocking_e, landscape, site_class, slope, aspect,
    lat, lon, elev, state, date_s, date_e, interval, status_change,
    plot)

save(nf_fia, file = "rda/nf-fia.rda")

```

## Organization

The final dataset contains 379,768 unique tree records, which were tallied across 10,318 plots evenly distributed throughout the region. Tallied trees belong to 28 different species groups and were located in 17 different forest types in 14 different physiographic (landscape) positions. Remeasurement periods ranged from 2.91 to 7.67 years and averaged 5.11 years. Ninety six percent of tallied trees lived through the remeasurement period and the remaining four percent died.

A description of each variable in the final dataset is provided below. Variable names ammended with “\_s” are measurements taken at the start of the remeasurement period; those ammended with “\_e” are measures

taken at the end of the remeasurement period; those ammended with “\_mid” are estimates of mid-period measures, calculated by averaging the starting and ending measures; and those ammended with “\_rate” are annual rates of change, averaged over the remeasurement period. Positive rates are increasing measures, and negative rates are decreasing measures.

**spp**

Species or species group. Similar species are grouped to ease data collection and interpretation.

**dbh**

Diameter at breast height (4.5’ above ground), measured in inches. Dbh rates are calculated by FIA program.

**cr**

Uncompacted crown ratio (percent of tree height supporting live crown)

**crown\_class**

Tree canopy position:

- 1 Open grown (crown has received full light for most or all of its life)
- 2 Dominant (crown extends above main canopy and receives full light from above and partly from sides)
- 3 Codominant (crown in main canopy and receives full light from above, but little from sides)
- 4 Intermediate (crown extends into main canopy, but receives little direct light)
- 5 Overtopped (crown entirely below main canopy level, receiving no direct light)

**tree\_class**

General quality of a live tree:

- 2 Growing-stock (of commercial species and meeting minimum merchantability standards)
- 3 Rough-cull (sound wood, but does not meet minimum merchantability standards)
- 4 Rotten-cull (does not meet minimum merchantability standards and more than half of cull is rotten)

**ba**

Plot basal area, measured in square feet per acre of all live trees, 1" dbh or greater

**bal**

Overtopping basal area, measured in square feet per acre (basal area of all trees in plot with larger diameter than subject tree)

**ht**

Total tree height, measured in feet. For trees with broken tops, heights are estimated by FIA program.

**forest\_type**

Adapted from FIA forest types and defined by species dominating the plot’s stocking. Note that FIA does not recognize a “mixedwood” forest type, so plots with greater than half of their basal area in softwood species are generally considered softwood types, and those with greater than half of their stocking in hardwoods are considered hardwood types. The exceptions are the “Pine-hardwood” and “Cedar-hardwood” types. The forest types used here do not always coincide well with available stocking charts. In the Northern Forest region, the types include:

*Northern hardwood*  
*Transition hardwood*  
*Oak-hickory*  
*Cottonwood*  
*Pine-hardwood*  
*Cedar-hardwood*  
*Spruce-fir*  
*Cedar*  
*Hemlock*  
*Larch* (includes tamarack)  
*Norway spruce*

*White pine*  
*Red pine*  
*Scots pine*  
*Mixed softwood*  
*Other*  
*Nonstocked*

**stocking**

Plot-level stocking of all live trees 1" dbh and larger:

- 1 Overstocked
- 2 Fully stocked
- 3 Medium stocked
- 4 Poorly stocked
- 5 Nonstocked

**landscape**

Physiography—depends on land form, topographic position, and soil type. Classes include:

*dry tops*  
*dry slopes*  
*deep sands*  
*other xeric*  
*flatwoods*  
*rolling uplands*  
*moist slopes & coves*  
*narrow floodplains/bottomlands*  
*broad floodplains/bottomlands*  
*other mesic*  
*swamps/bogs*  
*small drains*  
*beaver ponds*  
*other hydric*

**site\_class**

Site productivity class, defined by potential wood growth in cubic feet per acre per year:

- 1 225+ ft<sup>3</sup>ac<sup>-1</sup>yr<sup>-1</sup>
- 2 165-224 ft<sup>3</sup>ac<sup>-1</sup>yr<sup>-1</sup>
- 3 120-164 ft<sup>3</sup>ac<sup>-1</sup>yr<sup>-1</sup>
- 4 85-119 ft<sup>3</sup>ac<sup>-1</sup>yr<sup>-1</sup> (equivalent to class I in VT)
- 5 50-84 ft<sup>3</sup>ac<sup>-1</sup>yr<sup>-1</sup> (equivalent to class II in VT)
- 6 20-49 ft<sup>3</sup>ac<sup>-1</sup>yr<sup>-1</sup> (equivalent to class III in VT)
- 7 0-19 ft<sup>3</sup>ac<sup>-1</sup>yr<sup>-1</sup> (equivalent to class IV in VT)

**slope**

Slope in percent

**aspect**

Aspect in degrees

**lat**

Plot latitude in decimal degrees (random offset is applied to protect landowners' privacy)

**lon**

Plot longitude in decimal degrees (random offset is applied to protect landowners' privacy)

**elev**

Plot elevation in feet above mean sea level

**date**

Inventory date

**interval**

Length of remeasurement period in years ( $date\_e - date\_s$ )

**status\_change**

Change in tree status during remeasurement period. One of:

*lived*

*died* (natural mortality)

*cut*

**plot**

A unique identifier for the plot the tree was recorded on. Corresponds to PLOT.CN attribute for the ending inventory in the FIA Oracle database.