

# INGROWTH IMPUTATION IN EVEN-AGED PERMANENT SAMPLE PLOTS

OSCAR GARCÍA

*Dasometrics, Concón, Chile*

---

**ABSTRACT.** In forest sample plots it is common to measure only trees that are larger than some minimum size. This fact needs to be considered when using such truncated samples. I describe here a procedure for imputing the presence and sizes of the missing trees. Computer code is included.

**Keywords:** Missing data, truncation, forest growth and yield, data processing, tree measurements, *R* language.

---

## INTRODUCTION

It is common to only measure trees above some minimum diameter threshold, especially in stands originated by natural regeneration or seeding where a large number of small trees at early ages would make full measurement expensive. Ingrowth, unobserved trees that appear in later measurements, complicate data processing and can cause important biases if ignored (Curtis and Marshall 2005, Searle and Chen 2017).

There are several alternatives for dealing with ingrowth when developing even-aged growth models from permanent sample plot data. Sometimes, smaller trees have been sampled in subplots, but often these subplots are small, and the information is imprecise and difficult to use (Flewellling 1981, Curtis and Marshall 2005). One may omit measurements likely to contain substantial ingrowth, at the cost of reduced model reliability for younger stands. Another common strategy is to model the growth of the trees above the threshold and develop a separate model for ingrowth (Weiskittel et al. 2011). One must then accept “unnatural” empirical equations describing, for instance, an increasing number of trees at young ages. Another possibility is to somehow estimate the number and size of ingrowth trees and add them to the plot measurements. McGarrigle et al. (2011) proposed a method for estimating ingrowth with information from a single measurement.

In this note, I describe a method for estimating ingrowth using repeated permanent sample plot observations. The procedure was briefly explained in García (2013). The principles and description are easy to understand, but the implementation can be error-prone and not so obvious. Therefore, generic *R* computer code (R Development Core Team 2009) is included in the Appendix and can be downloaded from <https://github.com/ogarciav/ingrowth>. It is assumed that all measured trees have been tagged with identification numbers.

## IDENTIFYING INGROWTH TREES

At any measurement occasion  $i$ , we add as ingrowth any missing trees that appear at some future measurement occasion  $j$ . This might overestimate ingrowth at  $i$  if a tree observed at  $j$  is the result of late regeneration that occurred at some time between  $i$  and  $j$ . That should be rare in even-aged stands. On the other hand, ingrowth could be underestimated if trees reach the threshold after the last plot measurement. To guard against this, it seems advisable to exclude plots where at the time of the oldest measurement there could still be substantial ingrowth. E. g., ensure that in the last measurement the smallest observed diameters are far from the threshold.

Although not “ingrowth” by our definition, the procedure also misses trees that die or are removed before reaching the minimum diameter threshold, or

between that time and the following measurement. Sometimes that could be a problem, even though one might expect such trees to have little influence on stand dynamics.

## DIAMETERS

Once an ingrowth tree in measurement  $i$  is identified, it is necessary to assign a diameter to it. The actual diameter  $d_i$  is bounded between 0 and the minimum diameter threshold  $d_{\min}$ , so that in many instances the exact value is not critical. A simple ratio estimator seems sufficient: Compute the ratio  $r$  between the sum for the diameters of trees observed at measurement  $i$  and the sum of those observed at measurement  $j$ , where  $j$  is the measurement occasion where the tree first appears with a diameter  $d_j \geq d_{\min}$ . The sums are over the trees that appear in both measurements. Then we estimate  $d_i$  as the smallest of  $rd_j$  and  $d_{\min}$ . In mixtures, separate ratios could be computed for each species.

To limit the effect of imputation errors it may be advisable not to use measurements where the estimated ingrowth basal area exceeds some fraction of the total, such as 5%, 10%, or 20%. The cutoff point may be chosen as a compromise between the potential error and the number of available measurements.

## CONCLUSIONS

Ingrowth imputation can make usable many measurements that otherwise would have to be discarded or be treated with more cumbersome modeling approaches. The technique has been used successfully in García (2013, 2011, 2017). The software facilitates its application

## REFERENCES

- Curtis, R. O., and D. D. Marshall, 2005. Permanent-plot procedures for silvicultural and yield re-  
search. Gen. Tech. Rep. PNW-GTR-634, U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Portland, OR. 86 p.
- Flewelling, J. W., 1981. Compatible estimates of basal area and basal area growth from remeasured point samples. *Forest Science* 27(1):191–203.
- García, O., 2011. A parsimonious dynamic stand model for interior spruce in British Columbia. *Forest Science* 57(4):265–280.
- García, O., 2013. Building a dynamic growth model for trembling aspen in western Canada without age data. *Canadian Journal of Forest Research* 43(3):256–265.
- García, O., 2017. Cohort aggregation modelling for complex forest stands: Spruce-aspen mixtures in British Columbia. *Ecological Modelling* 343:109–122.
- McGarrigle, E., J. A. Kershaw, Jr, M. B. Lavigne, A. R. Weiskittel, and M. Ducey, 2011. Predicting the number of trees in small diameter classes using predictions from a two-parameter Weibull distribution. *Forestry* 84(4):431–439.
- R Development Core Team, 2009. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0 (<http://www.R-project.org>).
- Searle, E. B., and H. Y. Chen, 2017. Tree size thresholds produce biased estimates of forest biomass dynamics. *Forest Ecology and Management* 400:468–474.
- Weiskittel, A. R., D. W. Hann, J. John A. Kershaw, and J. K. Vanclay, 2011. *Forest Growth and Yield Modeling*. Wiley-Blackwell. 430 p.

## APPENDIX — COMPUTER CODE

The following R function prioritizes clarity over efficiency or elegance. The code should be easy to translate into other computer languages, and to be adapted as needed, for instance, to handle multiple species.

```
ingrowth <- function(treedata, plotdata, plot, meas, tree, diam, thresh){
  ingr <- treedata[NULL,] # ingrowth, initially empty
  for(plt in unique(treedata[, plot])){ # one plot at a time
    data <- droplevels(treedata[treedata[, plot] == plt,]) # data for a plot
    dmin <- plotdata[match(plt, plotdata[, plot]), thresh] # diameter threshold
    msrmnts <- sort(unique(data[, meas])) # measurement IDs
    for(i in msrmnts[-length(msrmnts)]){ # current measurement, exclude last
```

```

data_i <- data[data[, meas] == i, ]
trees_i <- data_i[, tree] # trees in measurement i
trees_iplus <- trees_i # including ingrowth (none so far)
for(j in msrmnts[msrmnts > i]){ # future measurement
  data_j <- data[data[, meas] == j, ]
  trees_j <- data_j[, tree] # trees in measurement j
  new <- !(trees_j %in% trees_iplus) # in j but not in i yet
  if(any(new)){ # add ingrowth
    ingr_ij <- data_j[new,] # new ingrowth for i
    ingr_ij[, meas] <- i # fix it
    dsumi <- sum(data_i[trees_i %in% trees_j, diam]) # over the shared
    dsumj <- sum(data_j[trees_j %in% trees_i, diam]) # observations
    ingr_ij[, diam] <- pmin((dsumi / dsumj) * ingr_ij[, diam],
                           dmin) # estimated diameters
    ingr <- rbind(ingr, ingr_ij) # cummulated ingrowth
    trees_iplus <- c(trees_iplus, trees_j[new]) # plus added ingrowth
  } # end if
} # j
} # i
} # plt
treedata$ingrowth <- FALSE # observed
if(nrow(ingr) > 0){
  ingr$ingrowth <- TRUE # imputed
  treedata <- rbind(treedata, ingr)
}
return(treedata)
}

```

Usually, variables other than the identifiers and diameters in the ingrowth records will be meaningless and could have been changed to NA's. However, there may be valid unchanging tree characteristics like species codes and location coordinates.

The function documentation follows.

## Description

Adds ingrowth records to a tree data frame

## Usage

```
ingrowth(treedata, plotdata, plot, meas, tree, diam, thresh)
```

## Arguments

<code>treedata</code>	Tree data frame. Contains at least plot, measurement, and tree identifiers, and diameter values.
<code>plotdata</code>	Plot data frame. Contains at least a plot identifier, and the minimum diameter threshold for the plot.
<code>plot</code>	Name of the plot identifier in <code>treedata</code> and <code>plotdata</code> .
<code>meas</code>	Name of the measurement identifier in <code>treedata</code> .
<code>tree</code>	Name of the tree identifier in <code>treedata</code> .
<code>diam</code>	Name of the diameter values in <code>treedata</code> .
<code>thresh</code>	Name of the minimum diameter threshold in <code>treedata</code> .

## Value

A data frame with `treedata` followed by the imputed ingrowth tree records. A logical variable named `ingrowth` is appended, containing `FALSE` for the observed trees and `TRUE` for the ingrowth. The values in the ingrowth records for variables other than the identifiers and diameters are inherited from later measurements and are often meaningless.

## Details

The data should not include plots with substantial ingrowth at the oldest measurement. The measurement identifier values must be capable of being sorted in time, e.g., they could be consecutive integer indices, years, or dates. The other identifiers can be of any type (numbers, strings, or factors). It is advisable to discard measurements where the imputed ingrowth basal area is a large fraction of the total.

## Examples

```
withIngrowth <- ingrowth(treeData, plotData, plot="PlotId", meas = "MeasId",  
                          tree = "TreeId", diam = "DBH", thresh = "Thresh")
```

The software is available in *Github* at <https://github.com/ogarciav/ingrowth>.