Using the Bootstrap to Compare Two Sampling Designs Used by Forest Ecologists

Bernhard Maier
Christian Pfeifer
G.U.H. Seeber

ABSTRACT: Estimates for the average volume of lying dead wood based on conventional plot sampling and line intersect sampling are compared by means of appropriately chosen bootstrap procedures. The latter design is much more easy to apply. In this case study we find that line intersect sampling does not perform considerably poorer than conventional plot sampling, which, however, is found to give rather imprecise estimates.

KEYWORDS: Superpopulation Bootstrap.

1 Introduction

This paper is about estimating the average volume per unit area of lying dead wood.
Dead wood is a characteristic structural element and feature of natural and semi-natural woodlands. In all its various manifestations depending on degree of decomposition, decay mileu, piece form and size, dead wood provides a habitat for an enormous number of highly specialized organisms (e.g. macromycetes, beetles). Dead wood has come to be regarded and appreciated as a useful indicator of biodiversity in forests.
Nature conservationists and forest ecologists want to increase dead wood quantity in intensively managed forests, they want to maximize the quality of dead wood and at the same time ensure sustainable dead wood management of the entire forest area. It might be difficult to reconcile these demands with classical forest management objectives such as yield production, forest protection against pests and the owner's obligation to ensure safe traffic conditions.
In order to assess habitat quality for a great variety of saproxylic species as well as hemeroby, i.e. the proximity of a forest ecosystem to its pristine

state, forest ecologists need to describe dead wood in a quantitative and a qualitative way. Allocating dead wood to different quality classes such as decay classes and agreeing on certain criteria and standards should facilitate comparison between forests. Furthermore it would help to establish a common ground for forest managers and nature conservationists.
For *conventional plot sampling* coordinates of a starting point and an angle are selected at random on a map of the study area. From the starting point a rectangular transect is drawn in direction of the angle. Within the transect middle diameter d_i and length l_i of each piece of dead wood are recorded. For pieces which extend beyond the boundaries only that portion within the transect are to be taken into account. The volume v_i of each item is approximated by treating it as if it were a cylinder, i.e. $v_i = l_i \cdot \pi \cdot d_i^2/4$. An estimate for the average volume is obtained by dividing the sum of volumes of all pieces recorded by the total area covered by transects, i.e. in obvious notation,

$$\widehat{V} = \frac{1}{A^0} \cdot \sum_{j=1}^{t} \sum_{i=1}^{n_i} v_i. \tag{1}$$

Besides conventional plot sampling, *line intersect sampling* (LIS) can be used for the assessment of lying dead wood. LIS has been developed and tested as a quick method to estimate the amount of slash left on a site after logging or the amount of fuel wood that is present on the forest floor in an actively growing stand. As above, coordinates of a starting point and an angle are selected at random. From the starting point a line of fixed length is drawn in direction of the angle. For each piece of lying dead wood crossing the line its diameter at line intersection e_i, say, is recorded. With L being the total length of sample lines and n the total number of pieces of dead wood crossing lines,

$$\tilde{V} = \frac{\pi^2}{8L} \cdot \sum_{i=1}^{n} e_i^2 \tag{2}$$

is calculated as an estimator for the average volume. For a more detailled description see Van Wagner (1968) or Pielou (1985). Kaiser (1983) provides a discussion of the closely related line-intercept sampling.

2 Case Study: Great Monk Wood

The study area 'Great Monk Wood' covers 43 hectares of semi-natural woodland and is part of Epping Forest (51° 40' N, 0° 4' E, about 20 km northeast of London St. Paul's), which has been managed by the Corporation of London since 1978. Great Monk Wood is a structure-rich beech dominated forest with numerous pollards, which testify to its ancient wood-pasture management. The storms of 1987 and 1990 tore some gaps into

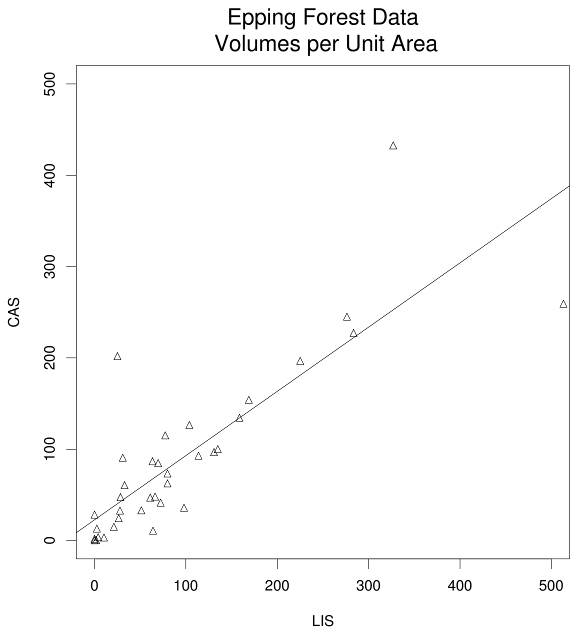

FIGURE 1. Volume per unit area estimated by conventional plot sampling (CAS) and by line intersect sampling (LIS) for each of 36 transects.

the stand texture, in which dense birch- and beech-regeneration has been thriving for the last few years. Since 1960 this area has been untouched by forestry activities and only recently Great Monk Wood was proposed as a minimum intervention area.

In this study the rectangular transect for conventional plot sampling covered 100 m^2 and was 25 m long and 2 m wide on each side of the transect line. With the rectangular sample plot diameter and length of each dead wood piece thicker than 5 cm middle diameter and longer than 0.5 m were recorded. All piece volume totals were multiplied by 100 to give hectare equivalence.

36 rectangular plot with a total of 829 pieces of dead wood were sampled. 256 line intersections were recorded. Average volume is estimated as 89.79 (standard error 15.52) by conventional plot sampling and as 95.28 (standard error 18.40) by line intersect sampling.

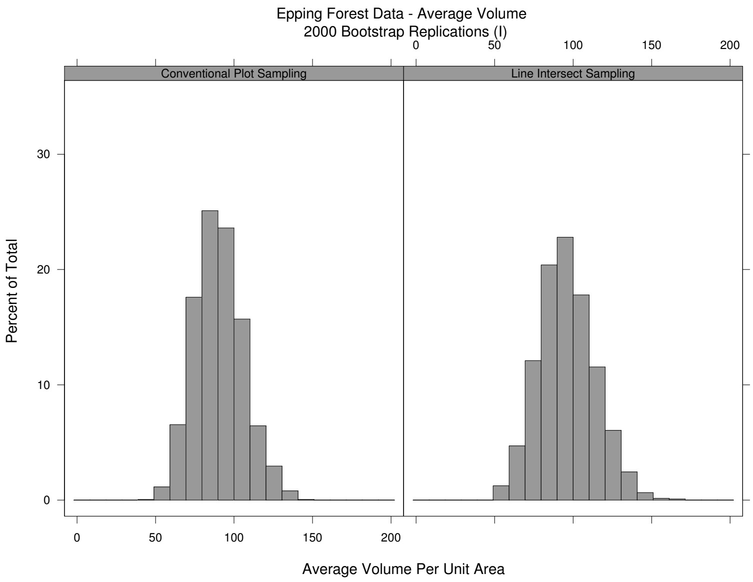

FIGURE 2. Great Monk Wood. Histograms for average volume of lying dead wood estimated by conventional plot sampling and by line intersect sampling, respectively, under resampling transect plots according to procedure I.

3 Bootstrap Procedures

Both sampling designs and their associated estimators are assessed and compared by boostrap procedures. Bootstrapping has become a popular technique in biology (see, e.g., Manly, 1997), it has also been found useful in forestry (see, e.g., Schreuder, Gregoire and Wood, 1993). However, it seems that no procedure adapted for the situation at hand has been proposed in the literature. Modifying suggestions by Presnell and Booth (1994) and Davison and Hinkley (1997) made for the case of sampling finite populations we suggest algorithms with the following general structure:

> Each iteration of the resampling procedures consists of two substeps:
>
> 1. generate a replicate population of the same size as the actual universe, and
> 2. generate a bootstrap sample of the same size as the actual sample from the replicate population, following as closely as possible the original sampling design.

We consider three variations of the procedure:

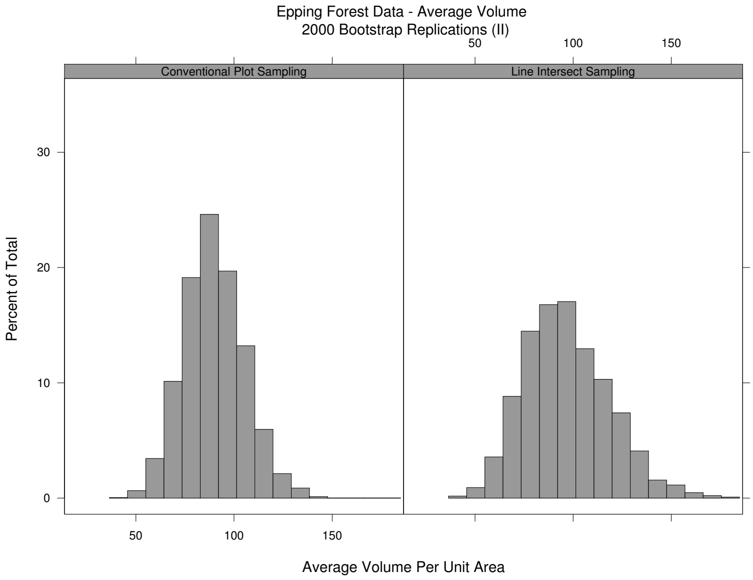

FIGURE 3. Great Monk Wood. Histograms for average volume of lying dead wood estimated by conventional plot sampling and by line intersect sampling, respectively, under resampling pieces of lying dead wood according to procedure II.

Procedure I: *Substep 1:* Sample with replacement from the set of rectangular plots in the actual sample until the area of resampled plots equals the total study area.

Substep 2: Sample without replacement as many transect plots or lines as have been in the actual sample. For each bootstrap sample calculate $\widehat{V}^*$ and $\tilde{V}^*$ according to (1) and (2), respectively.

Procedure II: *Substep 1:* Sample with replacement from the set of ractangular plots in the actual sample until the area of resampled plots equals the total study area.

Substep 2: Sample without replacement as many transect plots as have been in the actual sample. Use these resampled rectangular plots to calculate $\widehat{V}^*$.

Take each piece of dead wood in one of the resampled plots and 'throw' it randomly into a large rectangle of the same width as the rectangular plots and length equal to the total length of rectangular plots. To be more precise: The midpoint of each piece of dead wood is randomly assigned a coordinate in the large rectangle as well as an angle. Then it is determined whether the piece entirely lies within the boundaries of the rectangle. If not, it will be assigned new coordinates

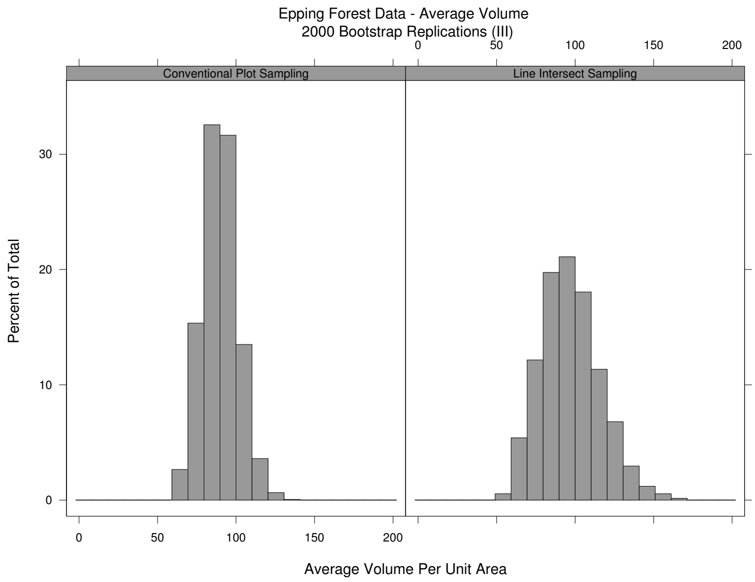

FIGURE 4. Great Monk Wood. Histograms for average volume of lying dead wood estimated by conventional plot sampling and by line intersect sampling, respectively, under resampling pieces of lying dead wood according to procedure III.

and angle until a position is reached, where the whole piece lies within the boundaries. For each piece check whether it intersects the transect line in the middle of the large rectangle. If it does, add it to the set of pieces used to calculate $\tilde{V}^*$.

This procedure is motivated by the derivation of the solution to the famous Buffon's needle problem, which states that the probability for a randomly 'thrown' piece if length l intersects a line of length L in an area of size A equals $(2 \cdot L \cdot l)/(\pi \cdot A)$.

Procedure III: *Substep 1:* Sample $\widehat{n}$ times with replacement from the set of pieces of dead wood, where $\widehat{N}$ is a reasonable estimate for the number of pieces of dead wood lying in the study ares.

Substep 2: Sample without replacement as many pieces of dead wood as found in the original sample. Use these pieces to calculate $\widehat{V}^*$.

Use substep 2 of procedure II to calculate $\tilde{V}^*$.

	$\widehat{V}^*$		$\tilde{V}^*$	
Procedure	mean	st.dev.	mean	std.dev.
I	90.40	15.30	95.88	17.88
II	89.82	15.34	96.74	21.62
III	90.03	11.22	96.99	18.94

TABLE 1.1. Great Monk Wood. Results for 2000 bootstrap replications using the three procedures described in the text.

4 Discussion

Table 1.1 summarizes the results for 2000 bootstrap replications using the procedures described above. No surprise, means and standard deviations for conventional plot sampling are close to those found in the sample, as $\widehat{V}$ is a mean. For procedure III standard deviations seem to be too small – due to the fact of ignoring some sort of clustering in the original sampling design.

In a simulations study designed to assess line intersect sampling in the specific situation of forest residue of tractor-logged and cable-logged populations Hazard and Pickford (1986) find that the line intersect method requires considerable sampling efforts to satisfy high levels of precision. The Epping Forest data suggest this to be true for natural environments as well. Standard errors for LIS are rather high for the actual sample size of 36 lines of length 25 m. However, conventional plot sampling does not perform significantly better as regards precision.

Table 1.1 provides evidence that the estimator $\tilde{V}$ based on line intersect sampling has a positive bias.

References

Davison, A., and Hinkley, D. (1997). *Bootstrap Methods and Their Application.* Cambridge: Cambridge University Press.

De Vries, P.G. (1986). *Sampling Theory for Forest Inventory: A Teach-yourself Course.* Berlin: Springer.

Hazard, J.W., and Pickford, S.G. (1986). Simulation studies on line intersect sampling of forest residue, part II. *Forest Science* **32**, 447–470.

Kaiser, L. (1983). Unbiased estimation in line-intercept sampling. *Biometrics* **39**, 956–976.

Manly, B.F.J. (1997). *Randomization, Bootstrap, and Monte Carlo Methods in Biology. Second Edition.* London: Chapman & Hall.

Pielou, E.C. (1985). Line intersect sampling. In: *Encyclopedia of Statistics, Vol. 5* (N.L. Johnson and S. Kotz, eds.), 71–72. New York: Wiley.

Presnell, B., and Booth, J. (1994). Resampling methods for sample surveys. Technical report no. 470, Dept. of Statistics, University of Florida.

Schreuder, H.T., Gregoire, T.G., and Wood, G.B. (1993). *Sampling Methods for Multiresource Forest Inventory.* New York: Wiley.

Van Wagner, C.E. (1968). The line intersect method in forest fuel sampling. *Forest Science* **14**, 20–26.