Canadian Journal of Forest Research

Note

Approximating covariances between nested plot sizes in forest inventory

James A. Westfall

U.S. Forest Service, Northern Research Station, York, PA, USA

Corresponding author: **James A. Westfall** (email: james.westfall@usda.gov)

Abstract

The use of nested plot sizes in forest inventories that encounter a wide range of conditions is relatively common. In straightforward situations, data from the different nested plot sizes are usually combined to create a single plot observation for estimation purposes. However, circumstances may occur where nested plot size estimates are obtained separately and subsequently combined via addition to obtain the final result. In these cases, covariances among the nested plots must be considered in the calculation of estimator variance. However, there are several potential approaches that might be considered given that only partial information is known for all but the smallest nested plot size. In this paper, three approaches to estimating the covariance were examined: (1) a modified form of a partially-dependent sample adjustment factor method, (2) explicit subsetting of trees to the area in common among nested plots, and (3) typical covariance estimation ignoring the lack of common area basis. Although the adjustment factor and subsetting methods showed strong consistency in outcomes, the estimated covariances were much smaller than those from ignoring the area basis issue. A subsequent simulation exercise revealed the most accurate covariances were obtained by ignoring the area issue. Thus, covariance estimation calculations can proceed without the additional complications of accounting for different nested plot sizes.

Key words: partial dependence, adjustment factor, tree location, standard error, cluster plot

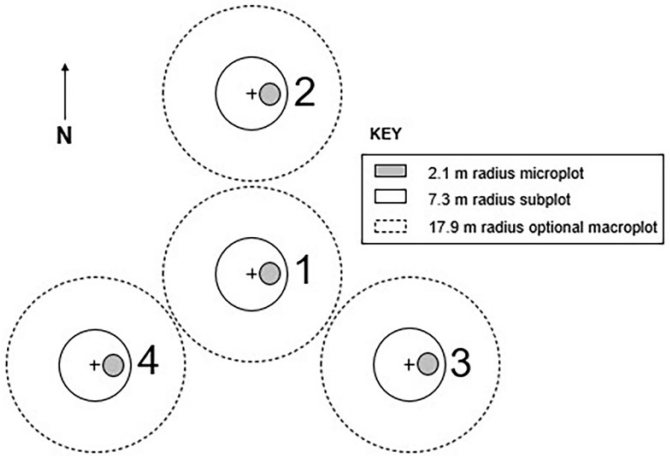
Introduction

A variety of response designs are used when conducting forest inventories (Kangas and Maltamo 2006; Tomppo et al. 2010). Cluster plot designs having a specified spatial arrangement of subplots are often favored for statistical efficiency (Lister and Leites 2022). When a wide range of conditions are likely to be encountered, nested plot designs are parsimonious from a field work perspective considering typical size–density relationships, e.g., reduced plot sizes are used to measure small trees because there may be many stems present (McRoberts et al. 2010). As such, forest inventory response designs consisting of clustered nested plot sizes are frequently encountered (Yim et al. 2015). An attribute that spans different nested plot sizes is usually scaled to a per-unit area basis and combined into a single plot-level observation for estimation purposes (Scott et al. 2005). However, there may exist situations where that approach is not desirable, e.g., when partial plot nonresponse is disproportionate among the nested plots and estimates of population attributes are generated using ratio-to-size estimators (Westfall 2022).

Disproportionate nonresponse can occur when differing conditions are encountered within the plot area. For example, the plot area may be partially on publicly owned land with the remainder on private land. When access to the private land area is not granted, it is possible that, for example, the smallest nested plot size may be fully measured, whereas a larger nested plot size may only be partially measured.

An alternative approach is to calculate separate estimates for each nested plot size and sum the estimates to obtain the desired result. However, this method involves additional complexity for variance calculations as covariances between the nested plot sizes must be accounted for. It is not clear the appropriate method for doing so given the entire suite of information is not known among all plot sizes. It may be considered that only the data occurring within the area common to both plot sizes should be used—which can be accomplished for attributes such as tree volume or biomass when tree locations are known. In the case of unknown locations, an alternative approach to covariance estimation is presented that uses adjustment factors based on nested plot sizes. Lastly, the common area issue could be ignored and covariances calculated using all available data for each nested plot size. The purposes of this study are to (1) outline the details pertaining to covariance estimation for each method, (2) compare the empirical results among the methods using forest inventory data, and (3) conduct a Monte Carlo simulation to determine which method provides the most accurate results.

Fig. 1. Forest Inventory and Analysis plot design implemented in California, U.S.A.



Methods

Data

The data used for analysis were collected on 13 059 sample plots in the state of California (CA), U.S.A., by the Forest Inventory and Analysis (FIA) program during 2011–2018. The quasi-systematic panelized design has an overall sampling intensity of approximately 1 plot per 2428 ha (5937 acres) (Reams et al. 2005). The field data were collected on plots composed of four circular subplots having 7.32 m (24.0 ft) radius with one subplot located at plot center and the remaining three subplots centered at azimuths 120, 240, and 360 degrees and distance of 36.58 m (120.0 ft) from plot center (Fig. 1; Bechtold and Scott 2005). Each subplot contains a 2.1 m (6.8 ft) radius microplot with center offset 3.7 m (12.0 ft) at 90 degrees azimuth. In CA, 17.95 m (58.9 ft) radius macroplots co-located at subplot centers are also used. FIA implements a mapping protocol within each plot that permits delineation of different stand-level conditions (including demarcation of plot areas that could not be measured). Relevant to this study is the recording of forest type (Burrill et al. 2024) for each measured forested condition present.

In forested conditions, trees having diameter at breast height (dbh) of 2.54 cm (1.0 in.) $\leq$ dbh < 12.70 cm (5.0 in.) were recorded within the microplot, trees having dbh of 12.70 cm (5.0 in.) $\leq$ dbh < 60.96 cm (24.0 in.) were measured on subplots, and trees with dbh $\geq$ 60.96 cm (24.0 in.) were measured on the macroplots. The locations of trees within the nested plot sizes were recorded as azimuth and distance from the respective plot centers as the reference point. The methods described below will pertain to estimating the total number of trees (dbh $\geq$ 2.54 cm) in the population. A summary of the relevant data is provided in Table 1.

To more explicitly describe the known information regarding tree size classes and locations within each of the nested plot structures, the only data pertaining to saplings (2.54 cm $\leq$ dbh < 12.70 cm) are within the microplot area. Saplings may exist outside the microplot within both the subplot and macroplot areas but remain unmeasured. Similarly, the existence of trees (12.70 cm $\leq$ dbh < 60.96 cm) is known for the subplot area (including the nested microplot), but their occurrence within the macroplot area outside the subplot boundary is unknown. The existence and location of macroplot trees with a dbh $\geq$ 60.96 cm are known for the entire area (Fig. 2).

Estimation

To simplify analyses, simple random sample estimators were used to estimate the total number of trees in the population within each of 51 forest types present in CA. Estimates were initially calculated separately for each nested plot size j (=1 for microplot, =2 for subplot, =3 for macroplot): The mean number of trees for plot size j ($\bar{y}_j$) is

$$(1) \qquad \bar{y}_j = \frac{\sum_{i=1}^{n}\sum_{k=1}^{n_{ij}} y_{ijk}}{n} = \frac{\sum_{i=1}^{n} y_{ij}}{n}$$

where y_{ijk} is the expanded trees per ha for tree k measured on plot size j within plot i, n_{ij} is the number of trees measured on plot size j within plot i, and n is the total number of plots. The variance of the mean $v(\bar{y}_j)$ is

$$(2) \qquad v(\bar{y}_j) = \frac{\sum_{i=1}^{n}(y_{ij} - \bar{y}_j)^2}{n(n-1)}$$

The resulting population total $\hat{y}_j$ and its variance $v(\hat{y}_j)$ for plot size j are, respectively, calculated using the population size A:

$$(3) \qquad \hat{y}_j = A\bar{y}_j$$

$$(4) \qquad v(\hat{y}_j) = A^2 v(\bar{y}_j)$$

Subsequently, the estimated population total across all nested plots would be obtained from

$$(5) \qquad \hat{y} = \sum_{j=1}^{3}\hat{y}_j$$

Due to the lack of areal independence, the estimated variance of the population total needs to include the pairwise covariances among the nested plot sizes. Typically, for example, the covariance between $\hat{y}_1$ and $\hat{y}_2$ would have the formulation

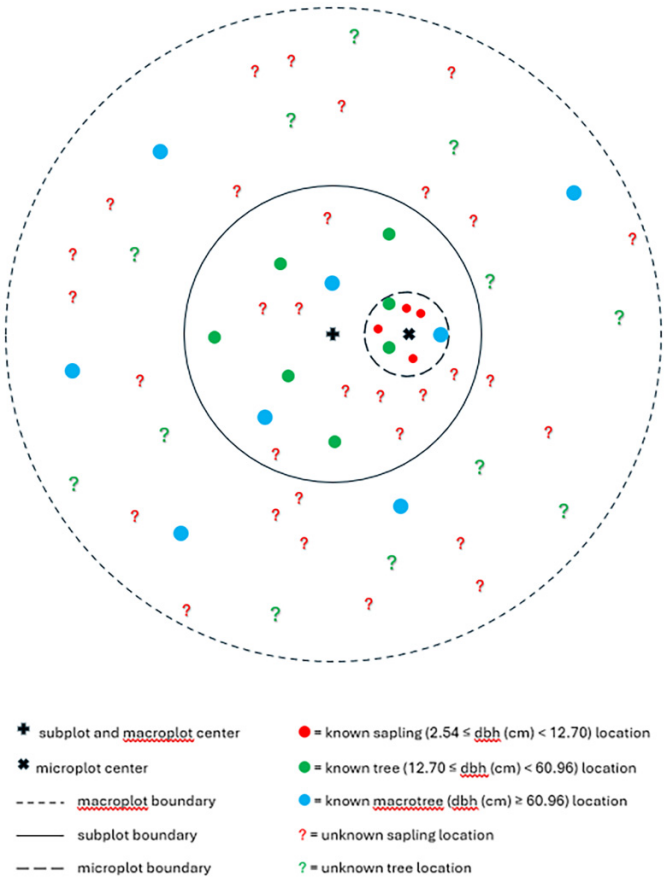
$$(6) \qquad \mathrm{cov}(\hat{y}_1, \hat{y}_2) = A^2 \frac{\sum_{i=1}^{n}(\bar{y}_1 - y_{i1})(\bar{y}_2 - y_{i2})}{n(n-1)}$$

However, a potential issue may be the covariance should only be calculated based on the in-common area between the respective plot sizes where the observed data for each tree size are known, i.e., the subplot area for cov $(\hat{y}_2, \hat{y}_3)$ and the microplot area for cov $(\hat{y}_1, \hat{y}_2)$ and cov $(\hat{y}_1, \hat{y}_3)$. In the example of cov $(\hat{y}_1, \hat{y}_2)$ assuming tree location information is available, the location information is based on different nested plot center locations (Fig. 1) and thus some mathematical manipulation is required to assess which subplot trees (12.70 cm

Table 1. Summary of Forest Inventory and Analysis tree data collected on 13 059 plots in California, U.S.A., from 2011 to 2018 by nested plot size.

		Trees per ha				Diameter distribution percentiles				
Plot size	Tree size	Min.	Mean	Max.	Std. dev.	5th	25th	50th	75th	95th
Microplot	2.54 cm ≤ dbh < 12.70 cm	0.0	146.5	10558.9	515.2	2.5	3.6	5.3	7.9	11.4
Subplot	12.70 cm ≤ dbh < 60.96 cm	0.0	82.9	1858.9	179.1	13.5	16.8	22.4	31.8	49.8
Macroplot	dbh ≥ 60.96 cm	0.0	5.2	167.9	14.6	62.2	67.3	75.9	90.7	123.4

Fig. 2. Illustration of nested plot structure with known and unknown data by tree size class and plot area.



≤ dbh < 60.96 cm) are located within the microplot area. Continuing with the cov $(\hat{y}_1, \hat{y}_2)$ scenario, once the subset of trees has been identified, y_{i2} and $\bar{y}_2$ need to be recalculated prior to calculating

$$(7) \quad \text{cov}(\hat{y}_1, \hat{y}_2^*) = A^2 \frac{\sum_{i=1}^{n} (\bar{y}_1 - y_{i1})(\bar{y}_2^* - y_{i2}^*)}{n(n-1)}$$

where the * superscript indicates values based only on the smaller of the two nested plot sizes being considered. A similar exercise must be performed prior to calculating cov $(\hat{y}_1, \hat{y}_3^*)$ and cov $(\hat{y}_2, \hat{y}_3^*)$. The estimated variance of the population total $v(\hat{y}^*)$ is

$$(8) \quad v(\hat{y}^*) = v(\hat{y}_1) + v(\hat{y}_2) + v(\hat{y}_3) + 2\text{cov}(\hat{y}_1, \hat{y}_2^*) + 2\text{cov}(\hat{y}_1, \hat{y}_3^*) + 2\text{cov}(\hat{y}_2, \hat{y}_3^*)$$

An alternative scenario to consider is when the tree locations are unknown and therefore performing the areal subsetting for covariance calculations is impossible. A partially-dependent areas approach is offered based on an adaptation of the sample size method given in Zheng (2004). Specifically, in the sample size formulation, the covariance is calculated from the sample units in common to both samples and adjustment factors are applied to account for the disparate sample units. A specification in the context of this study for two generic random variables x and y is shown as

$$(9) \quad \text{cov}(\hat{y}, \hat{x}) = A^2 \frac{\sum_{i=1}^{t} (\bar{y} - y_i)(\bar{x} - x_i)}{t(t-1)} \times \frac{t}{m} \times \frac{t}{n}$$

where the random variable y was observed on m sample units, the random variable x was observed on n sample units, and both y and x were observed on a subset of t sample units in common to both samples. Applying similar logic to the nested plot scenario suggests that the covariance is calculated using the original nested plot values and the adjustment factors are based on nested plot size differences. As such, t in the adjustment factors would now represent the area of the smaller of the two plot sizes involved, e.g., the microplot when considering cov $(\hat{y}_1, \hat{y}_2)$. Similarly, m and n would represent the respective nested plot sizes, e.g., 0.0014 ha for the microplot and 0.0168 ha for the subplot. A remaining adaptation is that there are always n plots in the sample, such that

$$(10) \quad \text{cov}(\hat{y}_1^a, \hat{y}_2^a) = A^2 \frac{\sum_{i=1}^{n} (\bar{y}_1 - y_{i1})(\bar{y}_2 - y_{i2})}{n(n-1)} \times \frac{0.0014}{0.0014} \times \frac{0.0014}{0.0168} = A^2 \frac{\sum_{i=1}^{n} (\bar{y}_1 - y_{i1})(\bar{y}_2 - y_{i2})}{n(n-1)} \times 0.0833$$

For completeness

$$(11) \quad \text{cov}(\hat{y}_1^a, \hat{y}_3^a) = A^2 \frac{\sum_{i=1}^{n} (\bar{y}_1 - y_{i1})(\bar{y}_3 - y_{i3})}{n(n-1)} \times 0.0138$$

$$(12) \quad \text{cov}(\hat{y}_2^a, \hat{y}_3^a) = A^2 \frac{\sum_{i=1}^{n} (\bar{y}_2 - y_{i2})(\bar{y}_3 - y_{i3})}{n(n-1)} \times 0.1660$$

The final variance of the estimate $\hat{y}^a$ is

$$(13) \quad v(\hat{y}^a) = v(\hat{y}_1) + v(\hat{y}_2) + v(\hat{y}_3) + 2\text{cov}(\hat{y}_1^a, \hat{y}_2^a) + 2\text{cov}(\hat{y}_1^a, \hat{y}_3^a) + 2\text{cov}(\hat{y}_2^a, \hat{y}_3^a)$$

A third option that could be used regardless of tree location information would be to use all the data and ignore the

nested plot size differences. In this case, the covariance would be calculated as shown in 6. To reiterate, the covariance between subplots and microplots would be

$$\text{(14)}\quad \operatorname{cov}(\hat{y}_1, \hat{y}_2) = A^2 \frac{\sum_{i=1}^{n} (\bar{y}_1 - y_{i1})(\bar{y}_2 - y_{i2})}{n(n-1)}$$

and analogously for the other nested plot size combinations. As would be expected, the variance is then specified as

$$\text{(15)}\quad v(\hat{y}) = v(\hat{y}_1) + v(\hat{y}_2) + v(\hat{y}_3) + 2\operatorname{cov}(\hat{y}_1, \hat{y}_2) + 2\operatorname{cov}(\hat{y}_1, \hat{y}_3) + 2\operatorname{cov}(\hat{y}_2, \hat{y}_3)$$

Methods comparison

A comparison among covariance methods was conducted using the FIA data from CA (Table 1) and the estimation formulae expressed above. The comparison of covariance methods was conducted via linear regression analysis. No intercept was specified as zero covariance would be present in cases where one of the nested plot sizes had no trees measured. This phenomenon may occur for rare forest types having few sample plots. For example, when assessing the subplot:microplot covariances the models may be specified as

$$\text{(16)}\quad \operatorname{cov}(\hat{y}_1, \hat{y}_2^*) = \beta_1 \operatorname{cov}(\hat{y}_1, \hat{y}_2) + \varepsilon$$

$$\text{(17)}\quad \operatorname{cov}(\hat{y}_1, \hat{y}_2^*) = \beta_1 \operatorname{cov}(\hat{y}_1^a, \hat{y}_2^a) + \varepsilon$$

$$\text{(18)}\quad \operatorname{cov}(\hat{y}_1, \hat{y}_2) = \beta_1 \operatorname{cov}(\hat{y}_1^a, \hat{y}_2^a) + \varepsilon$$

In 16–18, β_1 is an estimated slope parameter and ε represents random residual error. In this context, strong agreement between the two covariance methods would be indicated when the estimated parameter is $\beta_1 \approx 1$. The effect of the restriction (no intercept) on model error sum of squares was tested to evaluate adverse consequences on fit statistics (SAS Institute Inc. 2017).

Validation of results was accomplished via a Monte Carlo simulation where the population to be sampled was assembled using FIA data collected on 13 059 plots in CA from 2011 to 2018. For the subsetting method, the subplot:microplot and macroplot: microplot covariance calculations were conducted using the microplot area only. Similarly, the macroplot:subplot covariance was based on trees within the subplot area. The adjustment factor and ignore methods used all observed trees within each plot size in the covariance calculation. In each of $r = 1$ to 20 000 iterations, a sample of size 200 plots was drawn and estimated numbers of trees for each of the three nested plot sizes were calculated along with the pairwise covariances associated with each of the three candidate methods (e.g., $\operatorname{cov}(\hat{y}_1, \hat{y}_2^*)$, $\operatorname{cov}(\hat{y}_1, \hat{y}_3^*)$, and $\operatorname{cov}(\hat{y}_2, \hat{y}_3^*)$ for the subsetting method). Upon completion of the simulation, estimates of the number of trees for each nested plot size were used to determine the Monte Carlo covariances (indicated by the MC subscript) for each plot size combination, e.g., for the subplot:microplot assessment:

$$\text{(19)}\quad \operatorname{cov}_{\text{MC}}(\hat{y}_1, \hat{y}_2) = \frac{\sum_{r=1}^{20000} (\bar{y}_{r1} - \hat{y}_{r1})(\bar{y}_{r2} - \hat{y}_{r2})}{(20000 - 1)}$$

and likewise for the other nested plot combinations to obtain $\operatorname{cov}_{\text{MC}}(\hat{y}_1, \hat{y}_3)$ and $\operatorname{cov}_{\text{MC}}(\hat{y}_2, \hat{y}_3)$. The outcomes represent the covariances between the estimates obtained from repeated samples of the population. These values were compared to the mean of the calculated covariance estimates for each method, denoted as $\overline{\operatorname{cov}}(\hat{y}_1^a, \hat{y}_2^a)$, $\overline{\operatorname{cov}}(\hat{y}_1, \hat{y}_2^*)$, $\overline{\operatorname{cov}}(\hat{y}_1, \hat{y}_2)$ for the subplot:microplot and correspondingly for the other nested plot size combinations. The veracity of each method was assessed via percentage differences. For example, the differences for the subplot:microplot with the subsetting, adjustment factor, and ignore methods, would, respectively, be

$$\text{(20)}\quad \%D_{\text{MC}}^* = \frac{\left[\operatorname{cov}_{\text{MC}}(\hat{y}_1, \hat{y}_2) - \overline{\operatorname{cov}}(\hat{y}_1, \hat{y}_2^*)\right]}{\operatorname{cov}_{\text{MC}}(\hat{y}_1, \hat{y}_2)} \times 100$$

$$\text{(21)}\quad \%D_{\text{MC}}^a = \frac{\left[\operatorname{cov}_{\text{MC}}(\hat{y}_1, \hat{y}_2) - \overline{\operatorname{cov}}(\hat{y}_1^a, \hat{y}_2^a)\right]}{\operatorname{cov}_{\text{MC}}(\hat{y}_1, \hat{y}_2)} \times 100$$

$$\text{(22)}\quad \%D_{\text{MC}} = \frac{\left[\operatorname{cov}_{\text{MC}}(\hat{y}_1, \hat{y}_2) - \overline{\operatorname{cov}}(\hat{y}_1, \hat{y}_2)\right]}{\operatorname{cov}_{\text{MC}}(\hat{y}_1, \hat{y}_2)} \times 100$$

Results

Nested plot covariance estimations showed that the areal partial-dependence adjustment method used when tree locations are unknown provided reasonable approximations when compared to covariances generated from the subsetting method. When considering the subplot:microplot pairing, there was a slight tendency for the adjustment method to produce smaller covariances, as indicated by the slope ($\beta_1 = 0.9139$) being significantly less than 1 (Table 2). The results for the macroplot:microplot scenario were similar to those for the subplot:microplot in that modestly smaller covariances were attributed to the adjustment factor method, i.e., $\beta_1 = 0.9367$ was significantly less than 1 (Fig. 3; Table 2). The best agreement between nested plot size covariance estimates occurred for the macroplot:subplot combination. In this case, the parameter estimate was not statistically different than the desired outcome of $\beta_1 = 1$ (Fig. 3; Table 2). Despite the lack of strict correspondence between the two methods, the relationship was strongly linear as evidenced by the R^2 statistic being $\geq$0.99 (Table 2). The specification of no intercept had minimal effect on the R^2 calculation as indicated by the nonsignificant p-values associated with the model restriction (Table 2).

Due to the similar outcomes for the adjustment factor and subsetting methods, comparisons with the covariances calculated using the ignore method were only performed against the subsetting approach. In this case, substantial differences were found between the two methods where the subsetting covariances were consistently larger (Fig. 4). As with the comparison between subsetting and adjustment factor methods, there was a strong linear correlation between subsetting and ignore methods. However, the estimated β_1 parameters ($\beta_1 < 0.20$) were notably distant and statistically significantly different from the $\beta_1 = 1$ outcome that suggests agreement between the methods (Table 3).

Table 2. Linear regression between adjustment and subsetting covariances with resulting parameter estimates, standard errors, 95% confidence intervals, goodness of fit (R^2), and model restriction hypothesis test *p*-values.

Nested plots	Parameter	Estimate	Standard error	Confidence interval (95%)	R^2	Restrict[a]
Subplot:microplot	β_1	0.9139	0.0061	0.9019, 0.9258	0.998	0.369
Macroplot:microplot	β_1	1.0106	0.0144	0.9823, 1.0389	0.990	0.792
Macroplot:Subplot	β_1	0.9367	0.0040	0.9288, 0.9446	0.999	0.887

[a]*p*-value for test of significant effect on error sum of squares due to no intercept.

Fig. 3. Linear regression comparison of subsetting and adjustment covariance estimation methods based on 51 forest types in California for nested plot combinations subplot:microplot, macroplot:microplot, and macroplot:subplot.

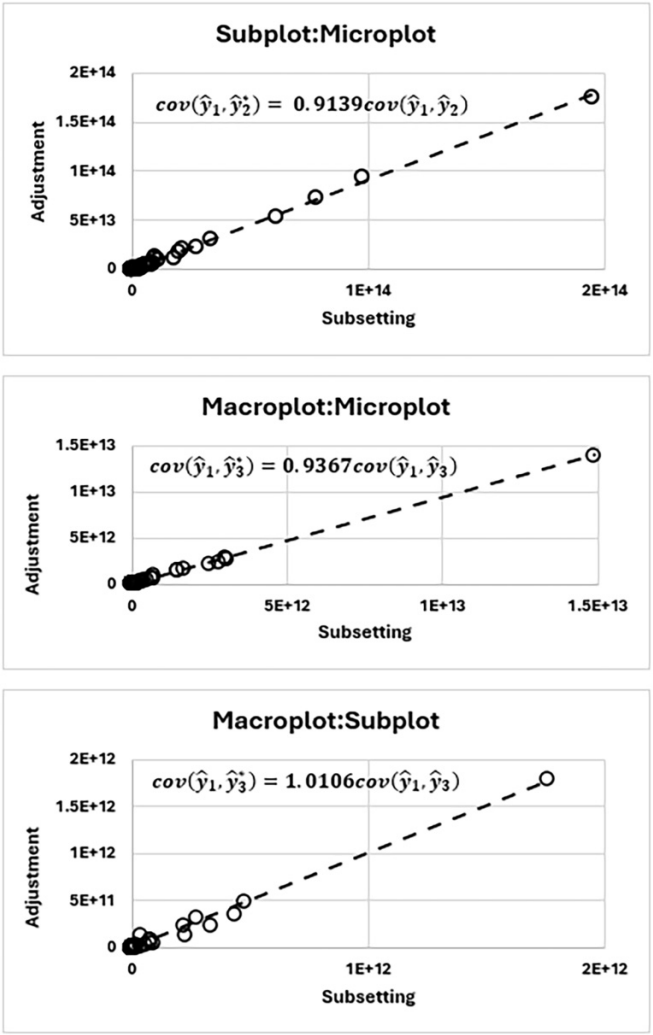
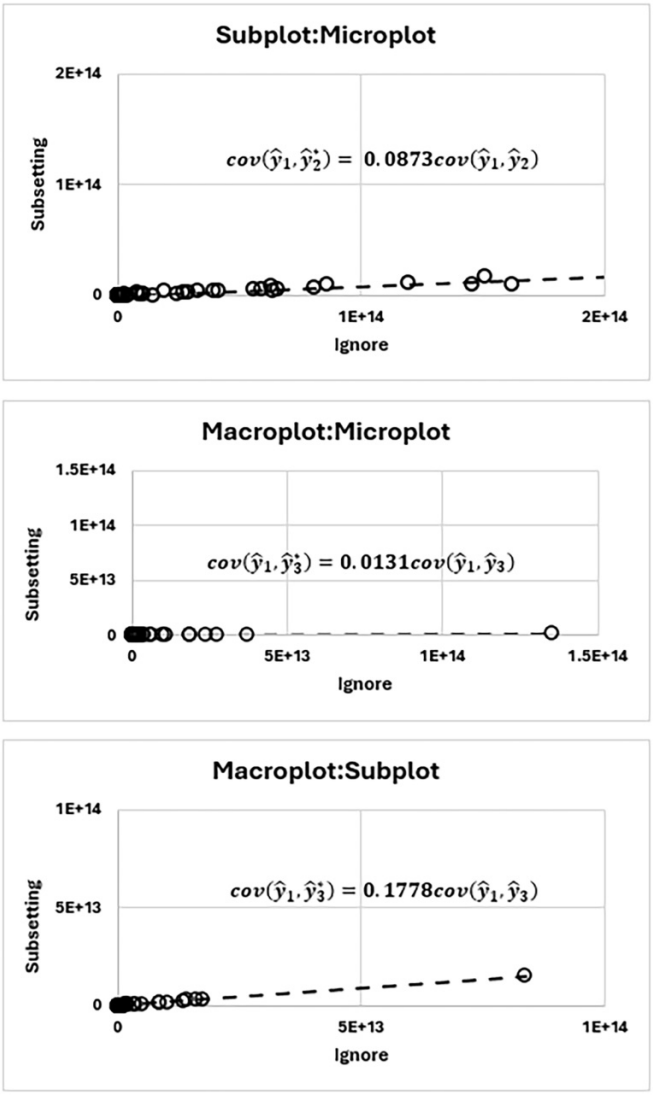


Fig. 4. Linear regression comparison of subsetting and ignore covariance estimation methods based on 51 forest types in California for nested plot combinations subplot:microplot, macroplot:microplot, and macroplot:subplot.



Given the disparity in results for the ignore method when compared to the subsetting and adjustment factor approaches, comparisons with the Monte Carlo simulation outcomes were needed to determine the most appropriate method. It was clearly shown that $\%D_{MC}$ was much closer to zero than either $\%D^a_{MC}$ or $\%D^*_{MC}$ (Table 4). Considering the Monte Carlo results as being close approximations of the true covariances, the adjustment factor and the subsetting methods for covariance calculations produce values that are considerably too small. Subsequently, use of these values in calculations of variance will cause underestimation and portray an overly enthusiastic indication of estimate reliability.

Discussion

When considering the subsetting and adjustment factor approaches that explicitly address the nonoverlapping areas of nested plots, the outcomes from both methods were quite

Table 3. Linear regression between subsetting and ignore method covariances with resulting parameter estimates, standard errors, 95% confidence intervals, goodness of fit (R^2), and model restriction hypothesis test p-values.

Nested plots	Parameter	Estimate	Standard error	Confidence interval (95%)	R^2	Restrict[a]
Subplot:microplot	β_1	0.0873	0.0006	0.0862, 0.0884	0.998	0.369
Macroplot:microplot	β_1	0.0131	0.0002	0.0127, 0.0135	0.983	0.792
Macroplot:subplot	β_1	0.1778	0.0008	0.1763, 0.1793	0.999	0.887

[a]p-value for test of significant effect on error sum of squares due to no intercept.

Table 4. Percent differences in covariance as compared to Monte Carlo results for the adjustment factor, subsetting, and ignore methods.

Nested plots	$\%D^{a}_{MC}$	$\%D^{*}_{MC}$	$\%D_{MC}$
Subplot:microplot	91.9%	90.9%	−1.0%
Subplot:macroplot	82.9%	81.7%	−2.7%
Microplot:macroplot	98.6%	98.7%	−3.3%

similar with the adjustment factor method results being up to approximately 10% smaller than those from subsetting (Fig. 3; Table 2). However, it was shown that both methods considerably underestimated the covariance when compared to Monte Carlo simulation results. In the common application where population estimates arise as a sum over different nested plot sizes and covariances are generally positive, a concern is warranted that variances (and associated standard errors) of estimates will be too small and a false impression of precision will be imparted.

The differences in estimated covariances between the ignore and subsetting methods (and the adjustment factor method by extension) were stark (Fig. 4; Table 3). Given the consistency between subsetting and adjustment factor methods, the ignore method appeared to provide questionable results. However, consultation with Monte Carlo estimations suggested the ignore method gives reasonable approximations to the covariances among nested plot sizes (Table 4). Presuming that the variances for each of the nested plot size estimates (i.e., $v(\widehat{y}_1)$, $v(\widehat{y}_2)$, and $v(\widehat{y}_3)$) can be established accurately, this leads to assurance that the overall variance $v(\widehat{y})$ is appropriately estimated as well. A secondary benefit of the ignore method is that the availability of tree location information does not need consideration.

Theoretically, the method could be applied to any number of nested plot configurations and sizes. More complicated scenarios would likely require additional investigation, i.e., covariances between nested plots and transects that intersect nested plot boundaries (Woodall et al. 2019). To provide a more holistic understanding of ignore method performance, further research is warranted under a broader range of forest conditions and nested plot configurations.

Conclusion

Among the three potential methods considered for estimating covariances among nested plot sizes, the subsetting and adjustment factor approaches considered the need to account for nonoverlapping areas. Indeed, it seemed initially intuitive when estimating covariances to only use pairs of observations based on the same measured area. Yet, the method that ignored the potential issue was shown to provide the most accurate estimated covariances when assessed against Monte Carlo simulation results. Appropriate estimations of covariance are imperative for obtaining empirically unbiased estimates of overall variance, which are often translated into typical indicators of uncertainty such as standard errors and associated confidence intervals that data users rely upon when making management or policy decisions. While the results of this study clearly indicated the superior performance of the ignore method, further studies should be conducted using inventory data for forests having different population characteristics and/or nested plot configurations.

Acknowledgements

The author is indebted to the associate editor and anonymous reviewers for their diligence and expert contributions to improve the original manuscript.

Article information

History dates

Received: 25 April 2025
Accepted: 13 June 2025
Accepted manuscript online: 16 June 2025
Version of record online: 18 July 2025

Data availability

Data analyzed during this study are available in the FIA Data Mart: https://research.fs.usda.gov/products/dataandtools/fia-datamart.

Author information

Author ORCIDs

James A. Westfall https://orcid.org/0009-0007-4123-0709

Author contributions

Conceptualization: JAW
Formal analysis: JAW

Methodology: JAW
Writing – original draft: JAW
Writing – review & editing: JAW

Competing interests

The author declares there are no competing interests.

References

Bechtold, W.A., and Scott, C.T. 2005. The Forest Inventory and Analysis plot design. *In* The enhanced Forest Inventory and Analysis program–national sampling design and estimation procedures. *Edited by* W.A. Bechtold and P.L. Patterson. USDA For. Serv. Gen. Tech. Rep. SRS-80. pp. 37–52.

Burrill, E.A., DiTommaso, A.M., Turner, J.A., Pugh, S.A., Christensen, G., Perry, C.J., et al. 2024. The Forest Inventory and Analysis Database, FIADB user guides, volume database description (version 9.2), nationwide forest inventory (NFI). Available from https://www.fs.usda.gov/research/products/dataandtools/datasets/fia-datamart [accessed June 2025].

Kangas, A., and Maltamo, M. (*Editors*). 2006. Forest inventory: methodology and applications. Vol. 10. Springer Science & Business Media.

Lister, A.J., and Leites, L.P. 2022. Cost implications of cluster plot design choices for precise estimation of forest attributes in landscapes and forests of varying heterogeneity. Can. J. For. Res. **52**(2): 188–200. doi:10.1139/cjfr-2020-0509.

McRoberts, R.E., Tomppo, E.O., and Næsset, E. 2010. Advances and emerging issues in national forest inventories. Scand. J. For. Res. **25**(4): 368–381. doi:10.1080/02827581.2010.496739.

Reams, G.A., Smith, W.D., Hansen, M.H., Bechtold, W.A., Roesch, F.A., and Moisen, G.G. 2005. The forest inventory and analysis sampling frame. *In* The enhanced Forest Inventory and Analysis program–national sampling design and estimation procedures. *Edited by* W.A. Bechtold and P.L. Patterson. USDA For. Serv. Gen. Tech. Rep. SRS-80. pp. 21–36.

SAS Institute Inc. 2017. SAS/STAT® 14.3 User's Guide, REG Procedure. SAS Institute Inc, Cary, NC.

Scott, C.T., Bechtold, W.A., Reams, G.A., Smith, W.D., Westfall, J.A., Hansen, M.H., and Moisen, G.G. 2005. Sample-based estimators used by the forest inventory and analysis national information management system. *In* The enhanced Forest Inventory and Analysis program–national sampling design and estimation procedures. *Edited by* W.A. Bechtold and P.L. Patterson. USDA For. Serv. Gen. Tech. Rep. SRS-80. pp. 53–77.

Tomppo, E., Gschwantner, T., Lawrence, M., McRoberts, R.E., Gabler, K., Schadauer, K., et al. 2010. National forest inventories. Pathways for common reporting. Springer, Dordrecht. doi:10.1007/978-90-481-3233-1.

Westfall, J.A. 2022. An estimation method to reduce complete and partial nonresponse bias in forest inventory. Eur. J. For. Res. **141**: 901–907. doi:10.1007/s10342-022-01480-6.

Woodall, C.W., Monleon, V.J., Fraver, S., Russell, M.B., Hatfield, M.H., Campbell, J.L., and Domke, G.M. 2019. The downed and dead wood inventory of forests in the United States. Sci. Data, **6**(1): 1–13. doi:10.1038/sdata.2018.303. PMID: 30647409.

Yim, J.-S., Shin, M.-Y., Son, Y., and Kleinn, C. 2015. Cluster plot optimization for a large area forest resource inventory in Korea. For. Sci. Tech. **11**(3): 139–146.

Zheng, B. 2004. Poverty comparisons with dependent samples. J. Appl. Econ. **19**: 419–428. doi:10.1002/jae.779.