

# Sample Sizes to Control Error Estimates in Determining Soil Bulk Density in California Forest Soils

## Youzhi Han

College of Forestry  
Shanxi Agricultural Univ.  
Taigu, Shanxi, 030801  
China

## Jianwei Zhang\*

US Forest Service  
Pacific Southwest Research Station  
3644 Avtech Parkway  
Redding, CA 96002

## Kim G. Mattson

Ecosystems Northwest  
189 Shasta Avenue  
Mount Shasta, CA 96067

## Weidong Zhang

Institute of Applied Ecology  
Chinese Academy of Sciences  
Shenyang, Liaoning, 110016  
China

## Thomas A. Weber

US Forest Service  
Northern Research Station  
11 Campus Blvd., Ste. 200  
Newtown Square, PA 19073

Characterizing forest soil properties with high variability is challenging, sometimes requiring large numbers of soil samples. Soil bulk density is a standard variable needed along with element concentrations to calculate nutrient pools. This study aimed to determine the optimal sample size, the number of observation ( $n$ ), for predicting the soil bulk density with a precision of  $\pm 10\%$  at a 95% confidence level among different soil types. We determined soil bulk density samples at three depths at 186 points distributed over three different 1-ha forest sites. We calculated  $n$  needed for estimating means of bulk density using a traditional method. This estimate was compared to a bootstrapping method  $n$  where the variance was estimated by re-sampling our original sample over 500 times. The results showed that patterns of soil bulk density varied by sites. Bootstrapping indicated 3 to 17 samples were needed to estimate mean soil bulk density at  $\pm 10\%$  at a 95% confidence level at the three sites and three depths. Sample sizes determined by the bootstrap method were larger than the numbers estimated by the traditional method. Bootstrapping is considered theoretically to be more robust, especially at a site with more variability or for site measures that are not normally distributed.

Soil bulk density is an indicator of soil compaction. Expressed as the ratio of mass of dry solids to bulk volume of soil, it is an essential variable for estimating soil mass, nutrient pools, and C storage. In addition, it also influences key soil processes and productivity by affecting infiltration, rooting depth, available water capacity, soil porosity and aeration, and the activity of soil microorganisms. Given its spatial variability, an accurate and efficient sampling of bulk density has challenged soil scientists, especially in highly variable forest soils. Determining the properties of forest soils requires more intensive sampling, and they often have less predictive value than agricultural soils for site assessment purposes.

Collecting large numbers of soil samples to estimate the parameters of certain soil properties such as bulk density is not only laborious but also costly. An optimal sample size, the number of observation ( $n$ ), requires an understanding of soil variability. Previous studies have reported high variation in bulk density in forest soils. Mroz and Reed (1991) found that the high spatial variability of soil physical and chemical properties limited accurate assessment of nutrient pools and nutrient cycling in their forest soils. Chaudhuri et al. (2011) concluded that the minimum number of samples required to detect a change in bulk density and soil organic C stock was a site-specific property. Studying soil C and N at second-rotation hoop pine (*Araucaria cunninghamii* Aiton ex D. Don) plantations, Blumfield et al. (2007) found that the sampling sizes were highly dependent on the soil property assessed and the acceptable relative sampling error. Similar results were also reported for estimating key ecosystem characteristics in a tropical *terra firme* rainforest (Metcalf et al., 2008). Sample sizes needed to obtain acceptable variability dif-

## Core Ideas

- Sample sizes varied from 3 to 17 to achieve  $\pm 10\%$  error at a 95% confidence level.
- Bootstrapping is more robust for estimating sample size than the traditional method.
- Soil variability must be considered before sampling.

Soil Sci. Soc. Am. J. 80:756–764

doi:10.2136/sssaj2015.12.0422

Supplemental material available online.

Received 3 Dec. 2015.

Accepted 23 Feb. 2016.

\*Corresponding author (jianweizhang@fs.fed.us).

© Soil Science Society of America, 5585 Guilford Rd., Madison WI 53711 USA. All Rights reserved.

ferred between plantations and natural forests of *Tectona grandis* L. f. (Amponsah et al., 2000). For multiple soil types at the Long-Term Soil Productivity study sites, Page-Dumroese et al. (2006) estimated that between 20 and 62 samples ha<sup>-1</sup> preharvest and 8 to 57 samples ha<sup>-1</sup> postharvest were needed to estimate the bulk density mean within 15% with 90% confidence. As a result, *n* estimation is a process characterized by different degrees of complexity (Confalonieri et al., 2009).

A couple of approaches are used by soil scientists to determine *n* to maximize accuracy and efficiency. A traditional approach is to collect soil samples within a study area, compute the sample variance as an estimate of the population variance, and determine *n* (cf. Snedecor and Cochran, 1967). Because soil properties can be highly variable, estimates of standard errors can also be highly variable. This is particularly true if the population is not normally distributed because the traditional method presumes a normal population. Another way to estimate the errors and thus the needed *n* is to resample the population multiple times and derive multiple estimates and their variances. This is not practical due to the costs of sampling. An alternative is to use the original sample, if collected without bias, as a representation of the actual population. A bootstrapping method, where the original sample is resampled with replacement multiple times, is used to obtain multiple estimates of means and standard errors and thus their confidence intervals. This method can be used on any population of any distribution and is effectively used in sample size calculation (Dane et al., 1986; Johnson et al., 1990). It has also been shown to provide better estimates than normal approximations for means, least square estimates, and many other statistics (Qumsiyeh, 2013).

In this study, we compared these two techniques to analyze soil bulk density variability in ponderosa pine (*Pinus ponderosa* C. Lawson var. *ponderosa*) plantations across three soil types. Because forest soils have a high proportion of rock fragments (>2 mm), soil fine bulk density (i.e., the mass of soil <2 mm per volume of soil <2 mm) was also examined (Flint and Childs 1984; Page-Dumroese et al., 1999). The difference between total and fine bulk density may be small in cultivated soils with

low rock fragment content but could be significant for estimating soil chemical stocks and water content in many forest soils (cf. Flint and Childs, 1984). The specific questions we addressed were: (i) how many samples are needed to accurately predict the soil (total and fine) bulk density within 10% of the mean with 95% confidence; and (ii) do different soil types vary in the number of samples needed?

## MATERIALS AND METHODS

### Study Sites

The three sites were located in northern California, where first-rotation experimental ponderosa pine plantations had been established during a previous study of plantation growth responses to fertilization, herbicide, and insecticide applications (Powers and Ferrell, 1996). They represented a range of productivity (from high to low): Feather Falls, Whitmore, and Elkhorn. These sites also represent a range of soil types typically found in the northern California west-side ponderosa pine region (Table 1).

All three field sites are characterized by a Mediterranean climate with dry summers and wet winters. Within-site topographical variability is fairly uniform, with slope variability <10% and aspects within 15°. Before the establishment of the first-rotation plots, land-use history was natural forest at Feather Falls, brush fields at Whitmore, and a sparsely stocked ponderosa pine plantation at Elkhorn. Site preparation included timber harvest, if economical, and/or clearing of brush and logging residues using normal operational practices during summer when soils were dry and less likely to compact. On about 1 ha at each site, 24 plots were established for the first rotation. Each 19.5- by 22.0-m plot was hand planted at a standard 2.4-m square spacing, at Whitmore in 1986 and the other two sites in 1988. The outer 4.8 m of each plot was designated as a buffer strip to minimize edge effects and the influence of adjacent treatments; the measured portion of each plot was a 4 by 5 grid of trees (cf. Powers and Ferrell, 1996).

Following planting, eight combinations of with and without herbicide, fertilizer, and insecticide applications were applied to the 24 plots in a completely randomized design at each site.

**Table 1. Geographic locations and site characteristics of three plantations in northern California.**

Characteristic	Elkhorn	Feather Falls	Whitmore
Location	40°4'57" N, 122°44'32" W	39°37'11" N, 121°11'48" W	40°37'33" N, 121°53'56" W
Elevation, m	1545	1246	755
Geomorphic province	Klamath	Sierra	Cascade
Annual mean max. temp., °C	16.4	18.3	21.4
Annual mean min. temp., °C	2.4	5.7	7.6
Annual precipitation, mm	1015	1780	1140
Soil taxonomy	Sheetiron gravelly loam; loamy-skeletal, mixed, mesic Dystric Xerochrepts	Toadtown loam; fine, parasesquic, mesic Andic Haplohumults	Aiken loam; clayey, oxidic, mesic Xeric Haplohumults
Clay content in top 30 cm, %	18–19	20–29	24–34
Organic matter, %	2.7	6.4	3.4
Typical soil depth, cm	74	200	254
Previous vegetation	plantation	natural stand	brush field
First rotation planted	1988	1988	1986
First rotation harvested	2012	2013	2012
Second rotation planted	2014	2014	2014

Treatments were applied at planting and repeated during the next 6 yr. Fertilization consisted of eight nutrients applied every 2 yr at an exponential rate, ending with a large fertilizer application at the end of Year 6. Spring fertilization, while appropriate in many regions, poses the risk in Mediterranean climates that an early dry season could leave dry salts unsolubilized. Thus, fertilizers were applied following the first fall rains so that dry salts could dissolve and infiltrate the soil profile as the wet season commenced. Competing vegetation control was accomplished by spraying understory plants with glyphosate [*N*-(phosphonomethyl)glycine] each spring for the first 6 yr after planting.

Trees were harvested at Elkhorn and Whitmore in the fall of 2012 and at Feather Falls in the fall of 2013. A feller-buncher with a shear head was used to cut and bundle whole trees. A skidder with a grapple was used to yard the trees to outside the plots. The sites were replanted with ponderosa pine in 2014; 12 plots were planted using seedlings from the same seed sources used for the first rotation and 12 plots were planted with seedlings from superior half-sib families at Elkhorn and full-sib families at Whitmore and Feather Falls. All seedlings for all sites were grown by Cal Forest Nurseries (Etna, CA).

Soil samples at three depths, 0 to 10, 10 to 20, and 20 to 30 cm, were collected in the fall of 2014. To capture the spatial variability of the soil, we systematically selected three points, with one at the plot center and one each 5 m from the center in opposite directions (northwest and southeast) at each of the plots (Supplemental Fig. S1). This produced a total of 72 samples (3 × 24 plots) for each of the three depths at Feather Falls and Whitmore and 42 samples (3 × 14 plots) for each of the three depths at Elkhorn.

### Soil Analysis Procedure

Soil bulk density samples were collected using a custom-designed, hammer-driven, double-wall, soil core sampler. This sampler allowed the extraction of intact soil core of 5.34-cm diameter by 6-cm length in a brass cylinder that is held in the barrel of the sampler. A 1-cm-long cylinder (guard ring) is placed at either end of the core retaining cylinder. Both are placed in the barrel during sample collection. The coring tip (1 cm) is screwed in place on the end of the barrel. The sampler was driven into the soil by a drop hammer to the targeted depth. After a sample had been taken, all of the retaining cylinders were pushed out of the barrel by the slotted core extractor without disturbance of the soil samples. The 1-cm-long guard rings were removed from either end of the assembly to permit trimming of the soil cores. The nominal 0- to 10-, 10- to 20-, and 20- to 30-cm depths were therefore actually composited of 2- to 8-, 12- to 18-, and 22- to 28-cm depths, respectively. As a caveat, soil cores from the top two depths were not sampled in the same hole but within a 30-cm adjacent area. The cores from 20 to 30 cm were collected from either an adjacent spot or the beneath of the 0- to 10-cm depth. The reason for this is that a small soil pit must be dug down to the top level of the deeper soil to allow the sampler to be positioned for the next sample.

Soil samples were returned to the laboratory and dried to a constant weight at 105°C. The samples were next weighed, then sieved through a 2-mm sieve, and the rock fragments >2 mm were weighed. The <2-mm rocks were not separated. Although plant roots were picked, we did not eliminate them from the bulk density calculation because of the difficulty of volume estimation.

Total bulk density ( $D_{bt}$ , Mg m<sup>-3</sup>) was calculated by dividing the oven-dry mass by the sample volume:

$$D_{bt} = \frac{W_s}{V_t} \quad [1]$$

where  $W_s$  is the oven-dry mass of the sample (Mg) and  $V_t$  is the total volume of the sample including pore volume and solid volume (m<sup>3</sup>).

Fine soil (i.e., <2-mm fraction) bulk density ( $D_{bf}$ ) was calculated according to Andraski (1991):

$$D_{bf} = \frac{D_{bt}(1-g_r)}{1-v_r} \quad [2]$$

where  $g_r$  is the gravimetric rock-fragment content that was calculated by dividing the mass of rock fragments by the total sample mass. Volumetric rock-fragment content ( $v_r$ ) was calculated as

$$v_r = D_{bt} \frac{g_r}{D_{br}} \quad [3]$$

where the rock-fragment density ( $D_{br}$ ) was assumed to be 2.65 Mg m<sup>-3</sup> (Page-Dumroese et al., 1999).

### Data Analysis

To test whether there were any effects on bulk density of the first-rotation treatments, the depths, or the sites, PROC MIXED (Supplemental Fig. S2) in SAS was used, with treatment as the fixed effect and plot within site as the random effect (SAS Institute, 2012). Following the analyses of Gbur et al. (2012, p. 199–236), we set both site and depth as the doubly repeated measures in space. Both are treated as fixed effects with an unrestricted covariance (un) for site effect and a first-order autoregressive [ar(1)] for depth effect. The covariance structures and distribution selection were conducted using the conditional Studentized residuals and the corrected Akaike information criteria in PROC GLIMMIX (Gbur et al., 2012, p. 199–236). In those combinations where no significant effects were detected, the number of samples necessary to characterize soil total ( $D_{bt}$ ) and fine ( $D_{bf}$ ) bulk density within a user-defined absolute difference was calculated using the following traditional equation:

$$\hat{n} = \frac{\sigma^2 t_{\alpha/2}^2}{E^2} \quad [4]$$

where  $\hat{n}$  is the number of samples necessary,  $t_{\alpha/2}^2$  is the value of the Student's  $t$  distribution with  $n - 1$  degrees of freedom at the  $\alpha$  probability level (0.05 in this study), where  $n$  is the number of samples used to calculate the sample variance,  $\sigma^2$  is the population variance that was assumed to be the same as the variance of the samples, and  $E$  is the allowable error. In this study, we

set  $E$  equal to 10% of the sample population means (Blyth and MacLeod, 1978).

In our bootstrapping method, we resampled 3 to 30 samples from our two original samples of 72 from Feather Falls or Whitmore or our sample of 42 from Elkhorn 500 times using SAS (Supplemental Fig. S2). The decision to use 500 bootstraps was based on Johnson et al. (1990). In fact, we found that larger resampling numbers affected only the thousandths digit of bulk density. From these resamples, means, variances, and other statistics were calculated as if we had gone to the field and resampled it 500 times.

## RESULTS

We found that both measures of bulk density were normally distributed within site and within depths. Volumetric and gravimetric rock content were lognormally distributed and were logarithmically transformed before conducting analyses. The conditional Studentized residuals for the entire data set are shown in Supplemental Fig. S3.

No treatment effects ( $p > 0.89$ ) were detected on our measures of soil bulk densities ( $D_{bt}$  and  $D_{bf}$ ) and rock-fragment contents ( $g_r$  and  $v_r$ ) (Table 2), nor were there interactions between treatment and depth ( $p > 0.43$ ) or treatment and site ( $p > 0.60$ ). Site differences were highly significant ( $p < 0.001$ ), with the greatest total (1.4–1.5) and fine (0.9–1.0) bulk densities at Elkhorn and about 1/3 lower densities at Whitmore and Feather Falls (Fig. 1), which was not surprising because soil types differ among sites (Table 1). The same trends were observed for rock-fragment contents (Fig. 2). Significant site  $\times$  depth interactions were found for  $D_{bf}$ ,  $v_r$ , and  $g_r$  ( $p < 0.001$ ) but not for  $D_{bt}$  ( $p = 0.42$ ). Significant differences in  $D_{bt}$  and  $g_r$  were detected among depths ( $p < 0.001$ ) but not for  $D_{bf}$  and  $v_r$  ( $p > 0.24$ ). Both  $D_{bt}$  and  $D_{bf}$  increased with depth at Feather Falls and Whitmore (Fig. 1B, 1C, 1E, and 1F), but there was no trend in  $D_{bt}$  while  $D_{bf}$  decreased with depth at Elkhorn (Fig. 1A and 1D). These bulk density trends were in the opposite direction of the trends for rock content, as both  $g_r$  and  $v_r$  decreased with depth at Feather Falls and Whitmore and increased considerably at Elkhorn (Fig. 2).

Figure 3 illustrates how the number of samples ( $n$ ) varies with the desired magnitude of the allowable error at the  $\pm 10\%$  level of means of  $D_{bt}$  and  $D_{bf}$ . At Elkhorn, where total bulk den-

sities were relatively high (e.g., 1.4–1.5  $\text{Mg m}^{-3}$ ) but variation low, sample sizes of four, three, and five were sufficient for total bulk densities at soil depths at 0 to 10, 10 to 20, and 20 to 30 cm, respectively, to achieve a  $\pm 10\%$  error at a 95% confidence level (Fig. 3A). For fine bulk densities, sample sizes two to three times higher (e.g., 9, 12, and 17) were needed for the respective depths (Fig. 3D). At Feather Falls, where lower bulk densities were observed (Fig. 3B and 3E;  $D_{bt}$  of 0.9–1.0  $\text{Mg m}^{-3}$  and  $D_{bf}$  of 0.6–0.8  $\text{Mg m}^{-3}$ ), sample sizes of seven to eight were needed for  $D_{bt}$  and 12 to 13 for  $D_{bf}$ . The lowest variation within depths for both bulk densities were found at Whitmore ( $D_{bt}$  of 1.0–1.1  $\text{Mg m}^{-3}$  and  $D_{bf}$  of 0.9–1.0  $\text{Mg m}^{-3}$ ; Fig. 3C and 3F). Here, sample sizes of about five and six were needed for  $D_{bt}$  and  $D_{bf}$  respectively.

The bootstrapped estimates of  $D_{bt}$  means and the number of samples within our allowable error showed slightly different results at both Elkhorn and Feather Falls but the same results at Whitmore when compared with the traditional calculation (Fig. 4). At Elkhorn, slightly more (i.e., five) samples were needed for 10 to 20 cm, but about double the number of samples (about eight) was required to detect the difference in the  $D_{bt}$  at other depths (Fig. 4A–4C). At Feather Falls, the same number of samples (i.e., eight) seemed sufficient for soil bulk density at the 0- to 10-cm depth (Fig. 4D), but slightly fewer (i.e., about 10) samples would be required at an expected accuracy of 95% with a precision level of 10% to estimate means of  $D_{bt}$  in the deeper soil. The variation among depths was relative uniform across the entire Whitmore site (Fig. 4G–4I); bootstrapping estimated that the same number of samples (i.e., five) was sufficient. Similar trends were also found for fine soil bulk density (Supplemental Fig. S4). At Elkhorn, not only were more samples required than the other sites for fine bulk density, but the deeper soils required more samples than the shallower soils. More samples were also needed for  $D_{bf}$  than for  $D_{bt}$  at Feather Falls. It appeared that five samples was sufficient at Whitmore.

## DISCUSSION

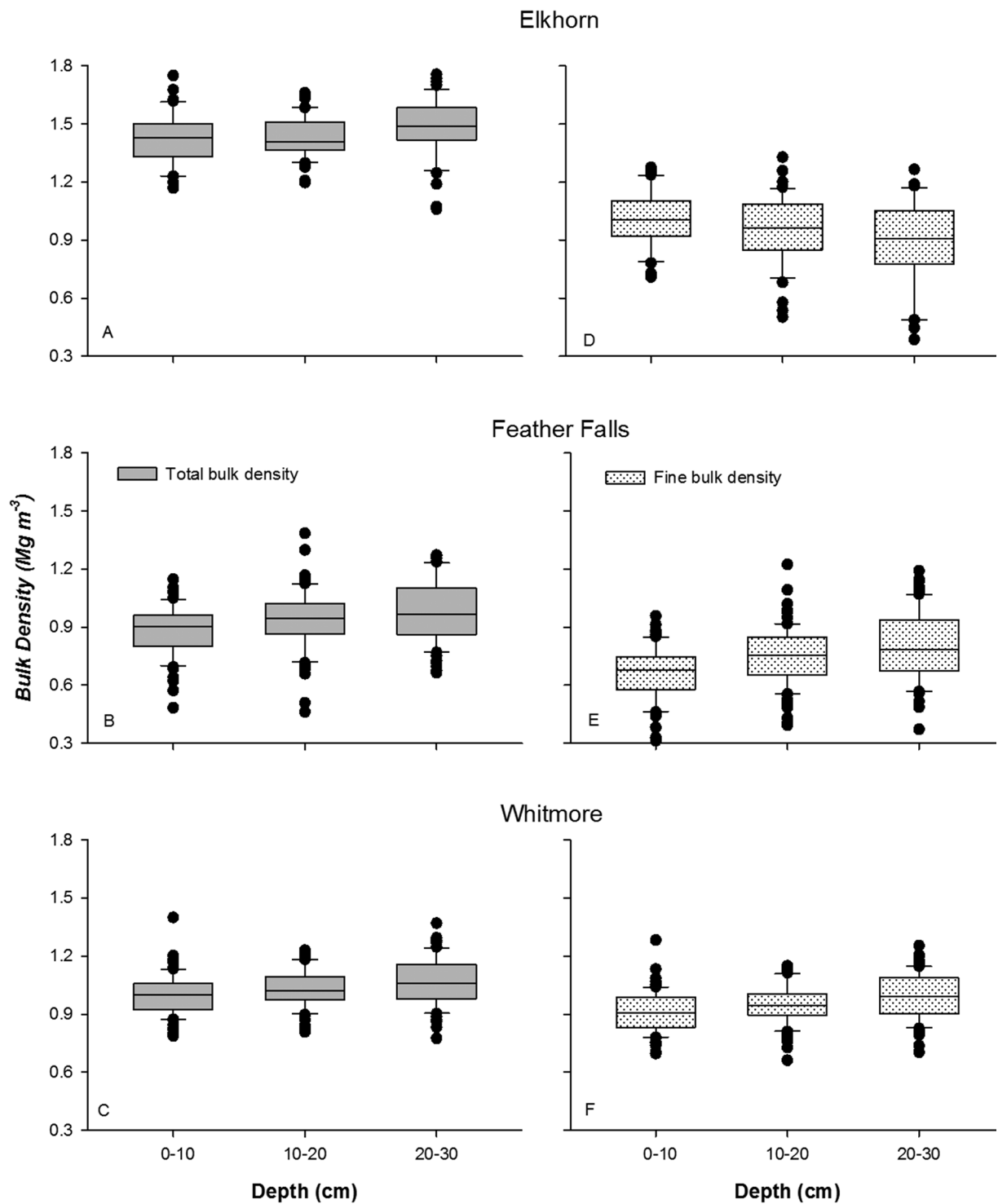
The results of this study showed that both the traditional method (Fig. 3) and bootstrapped estimates (Fig. 4 and Supplemental Fig. S2) required fewer samples than what we had collected for detecting the difference in soil bulk density with an allowable error of  $\pm 10\%$  of the population mean at 95% confi-

**Table 2. The  $P$  values of fixed effects for the treatment of first rotation, site, depth, and their interactions on soil bulk density and rock-fragment content at three sites of ponderosa pine plantations.**

Source of variation	Num dft	Den dft	Bulk density		Rock-fragment content	
			Total	Fine soil	Gravimetric	Volumetric
			$\text{Mg m}^{-3}$		%	
Treatment (TRT)	7	22	0.955	0.959	0.898	0.896
Depth	2	42	<0.001	0.239	<0.001	0.689
Depth $\times$ TRT	14	42	0.828	0.764	0.869	0.436
Site	2	26	<0.001	<0.001	<0.001	<0.001
Site $\times$ TRT	12	28	0.605	0.770	0.834	0.921
Site $\times$ depth	4	46	0.420	<0.001	<0.001	<0.001

† Numerator degrees of freedom.

‡ Denominator degrees of freedom, which changed slightly among variables.



**Fig. 1.** Boxplots of soil total and fine bulk density for soils collected from three depths at three sites.

dence. Our estimates of optimal sample sizes (i.e., 17 or less) for bulk density were at the lower end of numbers estimated by Page-Dumroese et al. (2006), who found that between 8 and 62 samples  $\text{ha}^{-1}$  were needed to estimate the bulk density mean within 15% with 90% confidence. This demonstrates the value of pre-sample assessments of variability. Because traditional methods of finding sample sizes depend on knowing or assuming the underlying distribution (Qumsiyeh, 2013), the bootstrap method can be used without the need to know the distribution (Dane et al., 1986;

Johnson et al., 1990). For example, we found that the bootstrap method estimated larger sample sizes than the traditional method for both  $D_{\text{bt}}$  and  $D_{\text{bf}}$  at Elkhorn and Feather Falls regardless of soil depth (Fig. 3 and 4). However, both methods yielded the same sample sizes at Whitmore, where variability was lowest (Fig. 1). The fact that bootstrapping predicted larger sample sizes than the traditional method, in spite of our large number of samples used in the traditional method, suggests the value of bootstrapping as a more robust method to assess actual population variation.

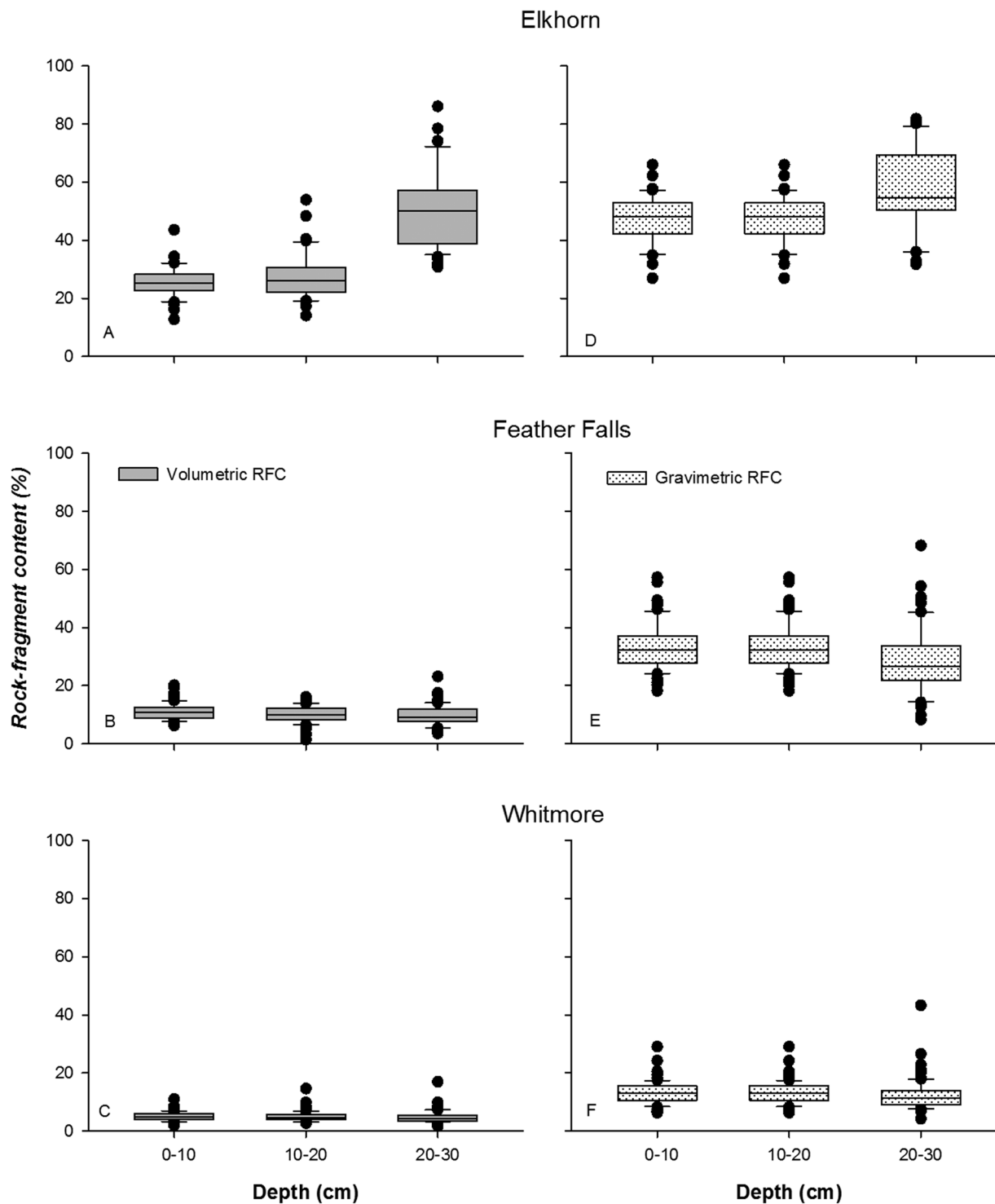
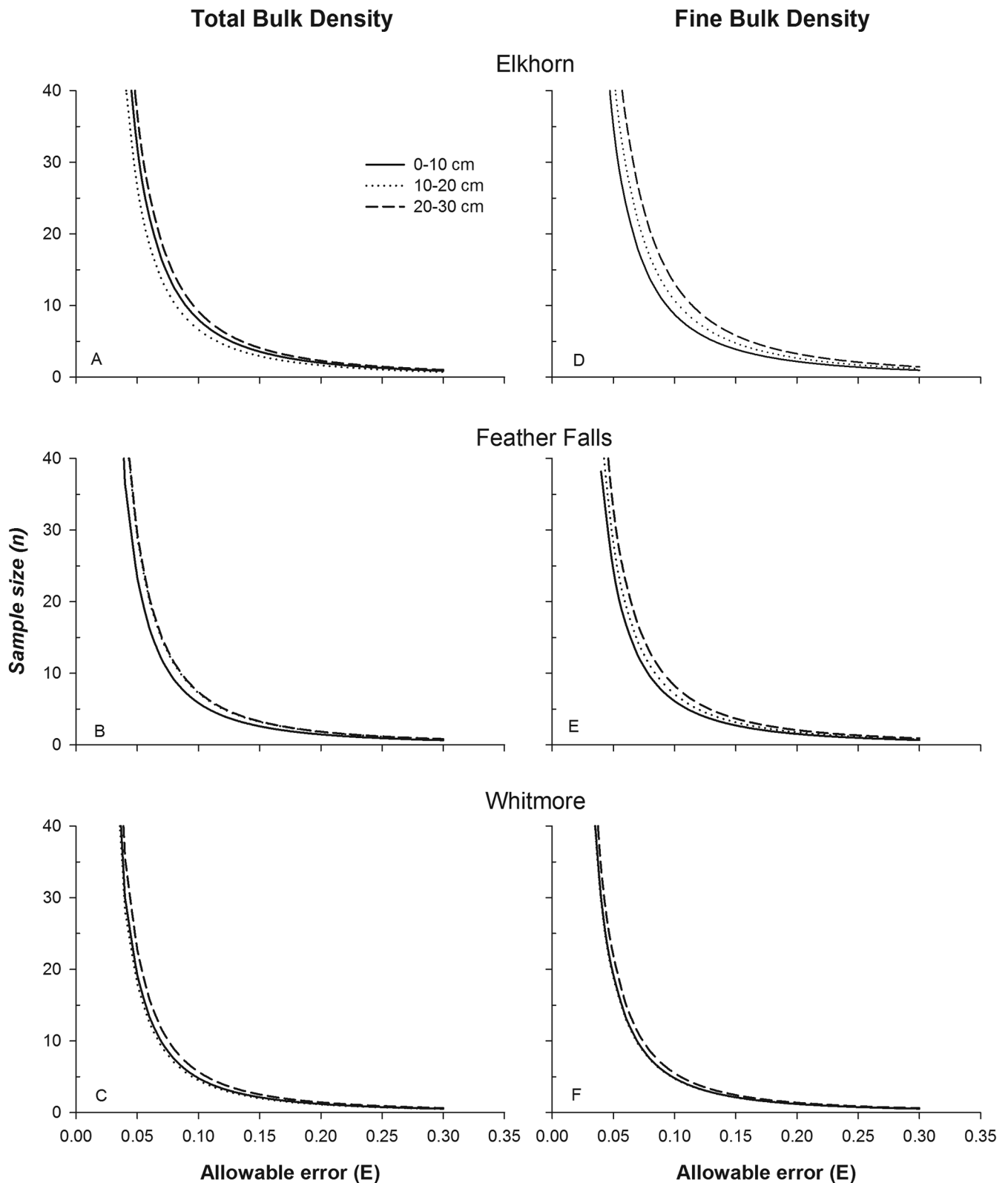


Fig. 2. Boxplots of soil volumetric and gravimetric rock fragment content (RFC) for soils collected from three depths at three sites.

Variation among sites was also substantial. Elkhorn showed the greatest soil bulk density ( $D_{bt} > 1.42 \text{ Mg m}^{-3}$  and  $D_{bf} > 0.89 \text{ Mg m}^{-3}$ ). This site has a sandy loam soil with parent material of metasediment. Although the bulk density values are close to the sandy loam bulk densities reported by Page-Dumroese et al. (2006), these values are much higher than previously estimated at the same site (McFarlane et al., 2009). This discrepancy could be the result of soil compaction caused by the heavy machines used for harvesting trees and removing slash in preparation for the sec-

ond-rotation plantings in 2012 (Greacen and Sands, 1980; Cambi et al., 2015). Soil compaction at the surface is also suggested by the observation that soil in the top 10 cm had greater bulk densities than that in the lower depths. At the other two sites, bulk density was intermediate at Whitmore ( $D_{bt} = 1.00\text{--}1.07 \text{ Mg m}^{-3}$  and  $D_{bf} = 0.91\text{--}0.99 \text{ Mg m}^{-3}$ ) and lowest at Feather Falls ( $D_{bt} = 0.88\text{--}0.98 \text{ Mg m}^{-3}$  and  $D_{bf} = 0.66\text{--}0.80 \text{ Mg m}^{-3}$ ). These sites are both derived from volcanic parent materials but have different soil texture; soils are clay at Whitmore and loam at Feather

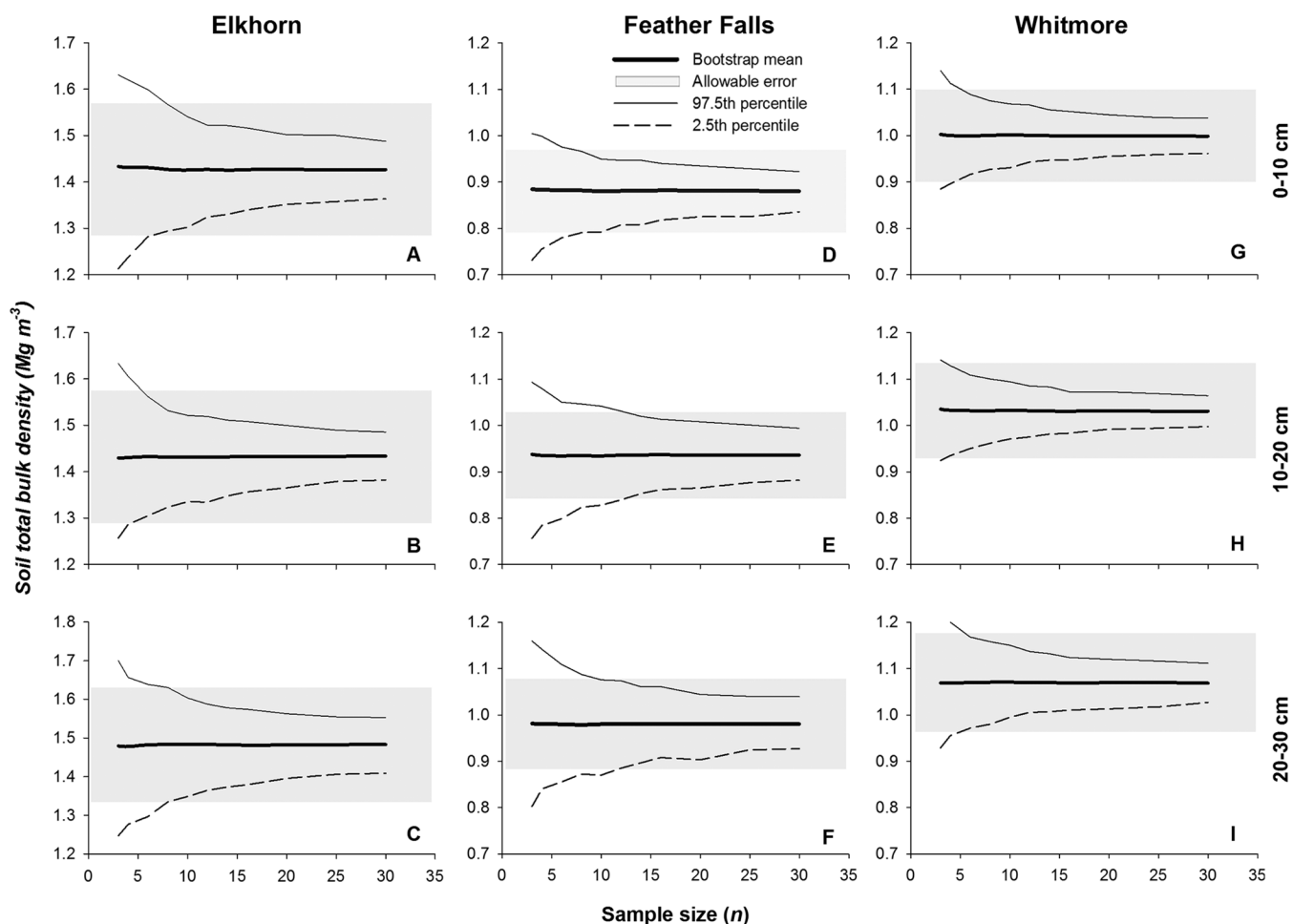


**Fig. 3.** Sample size ( $n$ ) required as calculated from the traditional method of Eq. [4] to achieve means with 95% confidence with  $\pm 10\%$  allowable error for (A,B,C) total bulk density and (D,E,F) fine bulk density at three depths of three sites.

Falls. Clays tend to show higher bulk density than loams (Corns, 1988; Page-Dumroese et al., 2006). Similar results were obtained by McFarlane et al. (2009) using a different method.

The results also suggest that neither the treatments imposed in the first rotation (i.e., fertilization, understory vegetation con-

trol, and insecticide application) nor the process of harvesting trees affected the soil bulk density. Soil bulk density increased with depth at Feather Falls and Whitmore, which was observed in previous studies (cf. Page-Dumroese et al., 2006). Low bulk density in topsoils is thought to be related to the activity of vegetation roots



**Fig. 4.** Mean,  $\pm 10\%$  allowable error, and 2.5 and 97.5 percentile values of total bulk density estimated by bootstrapping across different sample sizes (3–30) at three soil depths of three sites. The points at which the 95% confidence lines just enter the 10% error range are the sample sizes.

and an increase in organic matter (Adams, 1973; Federer et al., 1993). The lack of depth difference or slightly lower soil fine bulk density at Elkhorn may be due to the increase in rock content with depth (Fig. 2), which is consistent with the studies that reported decreased fine soil bulk density with increased rock content but failed to explain why (Andraski, 1991; Torri et al., 1994). It may be that the soil has higher organic matter concentrations in the non-rock portions because there is less soil per unit volume or there are extra voids in the soil associated with rock fragments. Also, there is the possibility of harvest compaction, as mentioned above.

Although this study aimed only to determine the number of samples needed to achieve the mean soil bulk density at  $\pm 10\%$  at a 95% confidence level, we have to acknowledge that the sampling apparatus (core method) could have introduced some biased errors such as changing the sampling spot if we came across a stone larger than the core diameter (Andraski, 1991; Flint and Childs, 1984; Page-Dumroese et al., 1999; Throop et al., 2012). Therefore, this type of sampler may not be appropriate for large-gravel forest soils.

Comparing the two methods, we posit that the advantages of the traditional method are that it is readily understood and easy to use. It is certainly better than not doing any assessment of site variability before initiating a field study. The disadvantage

of the traditional method is that it is probably done with fewer samples (e.g., six) and, at sites with high variability or with interactions, this would not capture the variability very well and would probably underestimate the number of samples needed. The bootstrapping approach is a superior approach because it appeared to capture a slightly higher estimate of spatial variation than the traditional method despite the fact that we collected a very high number of samples and should have captured the range of variation. A second advantage of bootstrapping is that it does not depend on the population being normal, and soil field data are often not normal. Third, the bootstrapping method provides an estimate of the confidence interval for  $n$  while the traditional method does not. The disadvantage of the bootstrapping method is that it is conceptually more complex, requiring computer software and some understanding of statistics. It also needs a relatively accurate sample that may not be particularly large but is unbiased.

Regardless of the methods that are used, dealing with spatial variability in the field is difficult and costly. Fewer samples without compromising precision is always preferred, but determining the appropriate sample size requires collection and analysis of many samples. Here, we presented  $n$  determination for soil bulk density at three relative small (1-ha) plantations. Because



variation may be greater for other soil properties such as soil nutrients and C (Amponsah et al., 2000; Blyth and MacLeod, 1978; Metcalfe et al., 2008), more samples might be required. We focused on soil bulk density because it is the most difficult procedure in the field, for which the volume of soil must be accurately determined.

## CONCLUSIONS

Significant site variation suggests that our sampling strategies for soil bulk density and other related ecosystem variables should be site specific. At each site, a relatively uniform bulk density across the 1-ha area yielded a much smaller sample size than what we previously used. In general, the number of samples needed as determined by the bootstrap method was larger than the number estimated by the traditional method. Bootstrapping provides an estimate of the confidence interval for  $n$  and is considered to be more robust, especially at sites with more variability or for site measures that are not normally distributed.

## ACKNOWLEDGMENTS

Youzhi Han and Weidong Zhang were visiting scholars when this study was performed at the Pacific Southwest Research Station in Redding, CA. Youzhi Han was supported by Shanxi Agricultural University through the Shanxi Scholarship Council of China in 2015. Weidong Zhang was supported by the Institute of Applied Ecology, Chinese Academy of Sciences, in 2014. Thomas A. Weber is a former forest technician. Partial financial support came from the Sierra Cascade Intensive Forest Management Research Cooperative. We thank Crane Mills, Sierra Pacific Industries, and W.B. Beatty & Associates for managing these plantations on their land. The comments from Dr. John Brown, Dr. Linda Heath, and four anonymous reviewers for improving the manuscript are greatly appreciated. The use of trade names and commercial products in this article is solely for the purpose of providing specific information and does not constitute endorsement by the US Forest Service.

## REFERENCES

Adams, W.A. 1973. The effect of organic matter on the bulk and true densities of some uncultivated podzolic soils. *J. Soil Sci.* 24:10–17. doi:10.1111/j.1365-2389.1973.tb00737.x

Amponsah, G.I., W.L. Meyer, and H.G. Nurchison. 2000. Soil sampling size estimates for soils under teak (*Tectona grandis* Linn. F) plantations and natural forests in Ashanti Region, Ghana. *Can. J. Soil Sci.* 80:327–336. doi:10.4141/S99-066

Andraski, B.J. 1991. Balloon and core sampling for determining bulk density of alluvial desert soil. *Soil Sci. Soc. Am. J.* 55:1188–1190. doi:10.2136/sssaj1991.03615995005500040048x

Blumfield, T.J., Z.-H. Xu, and N.V. Prasolova. 2007. Sampling size required for determining soil carbon and nitrogen properties at early establishment of second rotation hoop pine plantations in subtropical Australia. *Pedosphere* 17:706–711. doi:10.1016/S1002-0160(07)60085-X

Blyth, J.E., and D.A. MacLeod. 1978. The significance of soil variability for forest soil studies in north-east Scotland. *J. Soil Sci.* 29:419–430. doi:10.1111/j.1365-2389.1978.tb00791.x

Cambi, M., G. Certini, F. Neri, and E. Marchi. 2015. The impact of heavy traffic

on forest soils: A review. *For. Ecol. Manage.* 338:124–138. doi:10.1016/j.foreco.2014.11.022

Chaudhuri, S.E., M. Pena-Yewtukhiv, L.M. McDonald, J. Skousen, and M. Sperow. 2011. Land use effects on sample size requirements for soil organic carbon stock estimations. *Soil Sci.* 176:110–114. doi:10.1097/SS.0b013e31820a0fe2

Confalonieri, R., A. Perego, M.E. Chiodini, B. Scaglia, A.S. Rosenmund, and M. Acutis. 2009. Analysis of sample size for variables related to plant, soil, and soil microbial respiration in a paddy rice field. *Field Crops Res.* 113:125–130. doi:10.1016/j.fcr.2009.04.014

Corns, I.G.W. 1988. Compaction by forestry equipment and effects on coniferous seedling growth on four soils in the Alberta foothills. *Can. J. For. Res.* 18:75–84. doi:10.1139/x88-012

Dane, J.H., R.B. Reed, and J.W. Hopmans. 1986. Estimating soil parameters and sample size by bootstrapping. *Soil Sci. Soc. Am. J.* 50:283–287. doi:10.2136/sssaj1986.03615995005000020005x

Federer, C.A., D.E. Turcotte, and C.T. Smith. 1993. The organic fraction–bulk density relationship and the expression of nutrient content in forest soils. *Can. J. For. Res.* 23:1026–1032. doi:10.1139/x93-131

Flint, A.L., and S. Childs. 1984. Development and calibration of an irregular hole bulk density sampler. *Soil Sci. Soc. Am. J.* 48:374–378. doi:10.2136/sssaj1984.03615995004800020030x

Gbur, E.E., W.W. Stroup, K.S. McCarter, S. Durham, L.J. Young, M. Christman, et al. 2012. Analysis of generalized linear mixed models in the agricultural and natural resources sciences. ASA, CSSA, and SSSA, Madison, WI.

Greacen, E.L., and R. Sands. 1980. Compaction of forest soils: A review. *Aust. J. Soil Res.* 18:163–189. doi:10.1071/SR9800163

Johnson, C.E., A.H. Johnson, and T.G. Huntington. 1990. Sample size requirements for the determination of changes in soil nutrient pools. *Soil Sci. Plant Nutr.* 150:637–644.

McFarlane, K.J., S.H. Schoenholtz, and R.F. Powers. 2009. Plantation management intensity affects belowground carbon and nitrogen storage in northern California. *Soil Sci. Soc. Am. J.* 73:1020–1032. doi:10.2136/sssaj2008.0158

Metcalfe, D., P. Meir, L. Eduardo, O.C. Aragao, A. da Costa, S. Almeida, et al. 2008. Sample sizes for estimating key ecosystem characteristics in a tropical *terra firme* rainforest. *For. Ecol. Manage.* 255:558–566. doi:10.1016/j.foreco.2007.09.026

Mroz, G.D., and D.D. Reed. 1991. Forest soil sampling efficiency: Matching laboratory analyses and field sampling procedures. *Soil Sci. Soc. Am. J.* 55:1413–1416. doi:10.2136/sssaj1991.03615995005500050035x

Page-Dumroese, D.S., M.F. Jurgensen, R.E. Brown, and G.D. Mroz. 1999. Comparison of methods for determining bulk densities of rocky forest soils. *Soil Sci. Soc. Am. J.* 63:379–383. doi:10.2136/sssaj1999.03615995006300020016x

Page-Dumroese, D.S., M.F. Jurgensen, A.E. Tiarks, F. Ponder, F.G. Sanchez, Jr., R.L. Fleming, et al. 2006. Soil physical property changes at the North American Long-Term Soil Productivity Study sites: 1 and 5 years after compaction. *Can. J. For. Res.* 36:551–564. doi:10.1139/x05-273

Powers, R.F., and G.T. Ferrell. 1996. Moisture, nutrient, and insect constraints on plantation growth: The “Garden of Eden” experiment. *N.Z. J. For. Sci.* 26:126–144.

Qumsiyeh, M. 2013. Using the bootstrap for estimating the sample size in statistical experiments. *J. Mod. Appl. Stat. Methods* 12:45–53.

SAS Institute. 2012. SAS user’s guide. SAS Inst., Cary, NC.

Snedecor, G.W., and W.G. Cochran. 1967. *Statistical methods*. 6th ed. Iowa State Univ. Press, Ames.

Throop, H.L., S.R. Archer, H.C. Monger, and S. Waltman. 2012. When bulk density methods matter: Implications for estimating soil organic carbon pools in rocky soils. *J. Arid Environ.* 77:66–71.

Torri, D., J. Poesen, F. Monaci, and E. Busoni. 1994. Rock fragment content and fine soil bulk density. *Catena* 23:65–71. doi:10.1016/0341-8162(94)90053-1