

SHORT REPORT

Open Access



# Simple random sampling of individual items in the absence of a sampling frame that lists the individuals

P. W. West

## Abstract

**Background:** A 'sampling frame' identifies the sampling units in a population and their locations. It may consist of a listing of sampling units, or it may be based on a map of the population area within which sampling units can be observed. For inventory of large forests or other populations, it is common for no list of individual plants to exist, but it is common to have available a map of the area. When such a map is the only available sampling frame, methods are well established for drawing a simple random sample of fixed area plots. Less well-known methods are available if the sample is to consist of individual population members rather than groups of them in plots. Through simulation studies, the efficacies of two methods devised by Dr. K. Iles are considered for drawing a simple random sample of individuals given a map of the population area.

**Findings:** It is shown that simple random samples of individuals can be drawn satisfactorily using such a map. Further, the estimates obtained from the population mean of individuals, and its precision, are the same as those obtained when a sampling frame consisting of a list of individuals is available. Estimates of the population total can be obtained also, but their precision will be lower than those obtained when a list is available.

**Conclusions:** The absence of a list of individuals in a population does not preclude simple random sampling of individuals as long as a map of the population area is available. However, a preliminary survey of the population must then be made before sampling starts, and it may be necessary to visit many more sampling units to obtain the required sample than is the case when a list is available. The more complex the spatial arrangement of individuals within the population, the greater will be the number of sampling units that must be visited.

**Keywords:** Simple random sample, Inventory, Sampling, Sampling frame, Spatial arrangement

## Findings

### Background

Depending on the nature of a population and the information desired through sampling from it, there are many ways in which the sample may be drawn; these are discussed in texts on sampling techniques (e.g. Schreuder et al. 1993; Cochran 1999; Gregoire and Valentine 2008). Perhaps the most basic method of sampling is 'simple random sampling', where each and every member of a population has the same chance of being included in the sample and where all possible samples of a given size have the same chance of selection. This work is

concerned with difficulties that may be encountered in large populations, such as occur over large forest areas, when it is desired to take a simple random sample consisting of individual items, such as individual trees, from the population.

As discussed by Gregoire and Valentine (2008, p. 8), before a sample can be drawn from a population, it is necessary to have available a 'sampling frame', that is, a mechanism that identifies and locates the sampling units within the population. It may be a 'list sampling frame' whereby a list of each and every sampling unit has been compiled, or it may be an 'area sampling frame' that consists only of a map of the area containing the sampling units. If an area sampling frame only is available, the number of sampling units within the population is unknown and it is impossible to know where to start or

Correspondence: pwest@nor.com.au  
Forest Research Centre, School of Environment, Science and Engineering,  
Southern Cross University, Lismore, NSW 2480, Australia

finish selection of those that are to be included in a sample. In many areas of human endeavour, it is common to have to deal with large populations for which no list sampling frame has been compiled. The items making up such populations could be as varied as households within a city, trees in a forest, pebbles on a beach, objects in a photograph, cells on a microscope slide or any other situation where a large number of items appear on a surface or within a volume.

Forest (or other natural resource) populations may cover very large areas, too large for a list of individual trees or other plants within it to be compiled. However, if a map of the forest area exists, it can be used as an area sampling frame (Gregoire and Valentine 2008, p. 207) as long as plot-based sampling is to be used. There is then little difficulty in taking a simple random sample. Sample plots of a given area are positioned at randomly selected locations across the mapped area (Schreuder et al. 1993, pp. 113–117; Iles 2003, pp. 157–158). After a plot has been selected in the sample, measurements are made of the individuals in the plot to determine a stand value (an amount per unit area) of the variable of interest being measured in the inventory (the ‘target’ variable). To avoid bias in determining these stand values, particular ‘edge overlap’ methods must be used when measuring any plot of which a part falls beyond the forest edge (Schreuder et al. 1993, pp. 297–301; Iles 2003, pp. 621–658; Gregoire and Valentine 2008, pp. 343–355; West 2013, 2015, pp. 126–130).

The situation is rather different if no list sampling frame of the individuals in a population exists, but a sample is desired that consists of individuals rather than plots. Such a sample would be needed where the target variable is a characteristic of individuals rather than stands. Forest examples might involve estimation of the average height of the suppressed trees in the forest, the average number of mistletoe infections on individual trees, the average thickness of tree stem bark, the average number of habitat holes on old-growth trees in mature forest or the average clump size of an understorey grass species. Examples in other than forest populations might be the average size of pieces of gravel in a large pile, the average number of people living in individual dwellings in Siberia or the average femur length of skeletons in a mass grave in an archaeological dig.

Cluster sampling is one method that has been used in these circumstances. It involves dividing the population into recognisable ‘clusters’ (say, streets within a city suburb) and then sampling from the clusters (say, the houses in a street) (Cochran 1999, Chaps. 9–10; Schreuder et al. 1993, Chap. 3; Gregoire and Valentine 2008, Chap. 12). However, this is inappropriate if it is desired to take a simple random sample of individuals from the population; cluster sampling generally involves varying

chances of selection for the population members, rather than equal chances as required for a simple random sample. Pinkham (1987) described a simple method to take a simple random sample of individuals from a population without a list sampling frame. However, it involves visiting each and every individual in the population, generally an impossibility for very large populations and a process through which a list sampling frame could be compiled in any case.

In his works on forest inventory practice, Iles (1979, 2003, pp. 165–167) suggested methods to select a simple random sample of individuals from a population which has an area sampling frame available but no list sampling frame. These methods do not seem to have had any substantial appreciation previously in the statistical or forestry literature. This work describes two of Iles’ methods and considers their efficacy in inventory through simulation studies.

### Sampling methods

The two sampling methods considered here were termed Method 1 and Method 2 in Iles (2003, pp. 165–166) and the ‘plot reduction method’ and the ‘elimination technique’, respectively, in Iles (1979); for the present work, the names Method 1 and Method 2 will be used. The two methods are closely related. Both involve the random location of sample plots of a fixed area, counting the number of individuals in each plot, the selection of no more than one individual from the plot to become a member of the required simple random sample and then measurement of the target variable on that selected individual. Iles (1979) suggested other methods also, but they involve rather more measurement of individuals and will not be considered further here. The bias inherent in one of those other methods has been quantified recently (Lynch 2015).

Suppose the objective of the sampling process is to obtain a simple random sample containing  $n$  individuals. The sampling is done using sample plots, each of area  $a$ , with their centres positioned at randomly chosen locations across the forest area. For geometric ease, it is most practical to use square, rectangular or circular plots although other plot shapes could be used. Plot centres are allowed to be located within an area slightly larger than the forest area, an area that is then considered to constitute the entire forest population being sampled. This area may be defined as a rectangle completely surrounding the forest area (Iles 2003, Fig. 5.02, p. 166). For circular plots, no side of the outer rectangle should be, anywhere, closer to the forest edge than the plot radius and for square or rectangular plots, no closer than half the length of the plot diagonal. Suppose the total area defined by the forest and its surrounding rectangle is then  $A$ . Locating plot centres at random within such a rectangle allows sampling near

the forest edge to be done without use of the edge overlap methods mentioned earlier; any part of a sample plot that falls beyond the forest edge is then simply part of the total population being sampled as defined by the area  $A$ .

Sampling proceeds by choosing locations of plot centres at random points within the forest area and its surrounding rectangle. The plots become members of a set  $V$ , the  $i$ th ( $i \in V$ ) member of which contains  $p_i$  ( $\geq 0$ ) individuals. If a plot contains one or more individuals, selection methods described below are used to choose no more than one individual from the plot. That individual then becomes a member of a set  $S$  which will become the required simple random sample. The target variable is then measured on that individual to have the value  $y_i$  ( $i \in S$ ). The sampling process then continues until  $S$  contains  $n$  individuals, the sample size required, by which time  $V$  will contain  $\nu$  ( $\geq n$ ) plots.

Methods 1 and 2 differ in the way in which an individual is selected from a plot to become a member of  $S$ . Method 1 requires using a plot area small enough so that no more than one individual ever occurs in a plot. No selection process is then required, and if a plot contains an individual, that individual immediately becomes a member of  $S$ . This procedure may work well in circumstances where individuals are reasonably well spaced. However, if some individuals occur in close proximity to each other or if the individuals are, say, small understorey plants that occur in large numbers in clumps, the plot size would need to be impractically small. In those cases, a large number of selected plots would contain no individuals, so slowing the sampling process.

Method 2 allows larger plots to be used that may contain none or one or more individuals. If a plot contains no individuals, it becomes a member of  $V$  and sampling moves on to the next plot. Before sampling starts, a value  $M$  must have been chosen that is slightly in excess of the maximum number of individuals that will be found in any plot. When the  $i$ th ( $i \in V$ ) sample plot contains  $p_i$  individuals ( $1 \leq p_i < M$ ), they are numbered  $1 \dots p_i$  in an arbitrary order and a uniform  $(1, M)$  random number  $r$  is generated. If  $r > p_i$ , no individual is selected from that plot and sampling moves on to the next plot. If  $r \leq p_i$ , the individual in the plot assigned that number becomes the next member of  $S$  and is measured for the target variable. Before sampling starts, it is important that an appropriate value be determined for  $M$  by initial reconnaissance of the population. If a plot is encountered containing  $M$  or more individuals, the sampling process will be frustrated. On the other hand, the number of plots included in  $V$  will increase directly with the value chosen for  $M$ , so minimisation of the value of  $M$  will minimise the sampling effort.

Because both methods allow only one individual to be selected from any one randomly located plot, every

individual in the population has an equal chance of being included in the sample. Also, the random plot locations allow any one plot to overlap any other plot. This means that inclusion of any one tree selected in the sample does not preclude any other tree from being selected, even though both may have been included together in different plots. Thus, the final sample may consist of any possible combination of individuals from the population meaning that the requirements for a simple random sample are satisfied.

As described here, this process involves sampling with replacement. If, as is often the case, sampling without replacement is required, then no individual is permitted to enter the sample more than once. This may then require some additional plot selection to achieve the desired sample size.

For both Methods 1 and 2, estimates of the population mean of the target variable for the individuals,  $\bar{y}$ , may be determined with the usual simple random sample estimator as

$$\bar{y} = [\sum_{i \in S} y_i] / n \quad (1)$$

When sampling with replacement, an estimate of its standard error,  $\hat{\sigma}(\bar{y})$ , may be determined as

$$\hat{\sigma}(\bar{y}) = \{[\sum_{i \in S} (y_i - \bar{y})^2] / [n(n-1)]\}^{1/2} \quad (2)$$

and when sampling without replacement, as

$$\hat{\sigma}(\bar{y}) = \{[1 - n/N][\sum_{i \in S} (y_i - \bar{y})^2] / [n(n-1)]\}^{1/2} \quad (3)$$

where  $N$  is the total number of individuals in the population. In the absence of a list sampling frame,  $N$  is unknown. When relatively small samples are taken from relatively large populations, such as are common in the absence of a list sampling frame, the term  $n/N$  approaches zero and Eq. (3) reduces to Eq. (2). However, if sampling is without replacement and the term  $n/N$  is likely to be of a magnitude sufficient to affect  $\hat{\sigma}(\bar{y})$  appreciably, then  $\hat{\sigma}(\bar{y})$  might need to be determined using bootstrapping.

An estimator of the population total of the target variable,  $\hat{Y}_T$ , is

$$\hat{Y}_T = \hat{N}\bar{y} \quad (4)$$

where  $\hat{N}$  is an estimate of the population size that is unknown in the absence of a list sampling frame.  $\hat{N}$  may be determined from the information collected for either method as

$$\hat{N} = [A/a][\sum_{i \in V} p_i] / \nu \quad (5)$$

The target variable value of individuals selected in the sample ( $y_i$ ) is not necessarily uncorrelated with the number of individuals in the plot from which each was selected ( $p_i$ ). Furthermore, the number of individuals selected ( $n$ )

and the number of plots sampled ( $\nu$ ) are not usually the same. Under these circumstances, no analytical estimator is available to determine the standard error of this estimate of the population total,  $\hat{\sigma}(\hat{Y}_T)$ , which would involve the product of the variances of the two terms of Eq. (4) (Goodman 1960). Instead, this standard error might be estimated using bootstrapping. To do this, the members of the set  $V$  would be re-sampled with replacement until a new set  $V'$ , consisting of  $\nu'$  plots of which the  $i$ th ( $i \in V'$ ) originally contained  $p'_i$  individuals and in which a subset  $S'$  of size  $n$  is made up from plots that contained a sample individual. A new estimate of  $\bar{y}$  would then be obtained, using the data from  $S'$  with Eq. (1), and also of  $\bar{N}$ , by replacing  $\nu$  with  $\nu'$  and the  $p_i$  with the  $p'_i$  in Eq. (5). These new estimates would be used with Eq. (4) to get a new estimate of  $\hat{Y}_T$ . This process would be repeated many times and the standard error of the many estimates of  $\hat{Y}_T$  used as the estimate of  $\hat{\sigma}(\hat{Y}_T)$ .

Of course, other, plot-based, methods could be used to estimate the population size, perhaps with greater precision. The method described above was mentioned here only because it uses the data already collected in determining the population mean and requires no additional sampling effort.

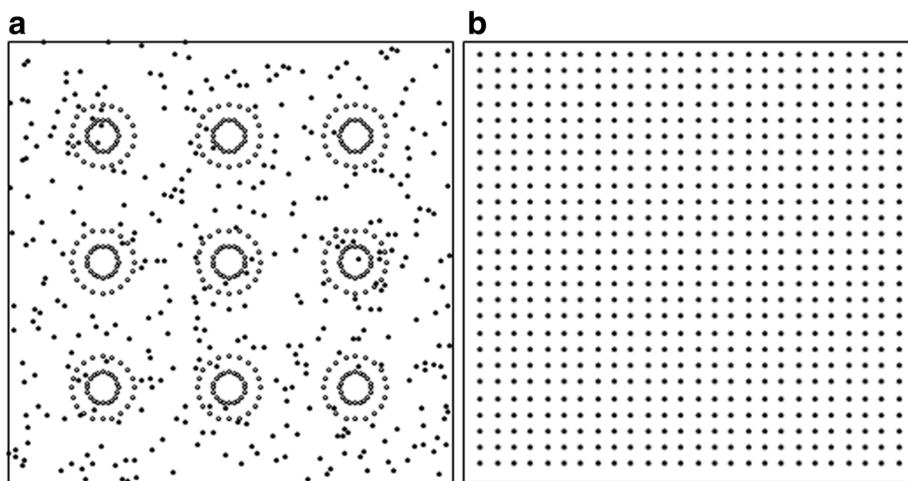
### Simulations

Simulations using both Methods 1 and 2 were conducted using two artificially constructed forest populations from which simple random samples of individual trees were to be selected. The populations were contained within square 49-ha forest areas, and the locations of the individuals within the two populations are shown in Fig. 1. The first population (Fig. 1a) had a

rather complex arrangement of 720 individuals. Half of the individuals were arranged in circular patches whilst half were scattered randomly across the entire area. The individuals in the circular patches had stem diameters at breast height selected at random in the range 60–90 cm. The randomly scattered individuals were all larger with random diameters from the range 90–120 cm. Over the whole forest area, the average tree stem basal area was  $0.662 \text{ m}^2$  and the total basal area was  $476 \text{ m}^2$ . This arrangement was such that no individual was closer to any other than 7.8 m. Such a complex arrangement of individuals leads to considerable small-scale variation in tree stocking density; it was felt that problems encountered in selecting samples should become evident with such complexity. In the second population, (Fig. 1b), 676 individuals were positioned in a 26.05-m square arrangement with stem diameters in the range 60–120 cm. In that case, the average tree stem basal area was  $0.646 \text{ m}^2$  and their total basal area was  $473 \text{ m}^2$ .

Simulations were done for both example populations, by sampling with replacement and using both Methods 1 and 2, to estimate the population tree average basal area and the total basal area of all trees, together with their standard errors. In obtaining bootstrap estimates of the standard errors of each simulation estimate of the population total basal area, 100 bootstrap samples were used; more bootstrap samples might have been used, but some studies of sampling in forest circumstances have suggested that 100 is sufficient (Schreuder et al. 1992; West 2016).

Of course, in these example populations, the number and locations of all individuals in the populations were actually known, so that a list sampling frame did in fact exist for them, although it was being assumed there was



**Fig. 1** Locations of individual trees in artificially constructed forest populations contained in 700-m square areas. In **a**, there were 720 individuals, with 360 trees of smaller stem diameters (*open circle*) arranged in circular groups whilst 360 larger diameter trees (*filled circle*) were scattered randomly. In **b**, 676 individuals were arranged on a 26.05-m square grid

only an area sampling frame available when applying Methods 1 and 2. Thus, simulations were done also using simple random sampling with replacement from the list sampling frame; these results could then be compared with those using Methods 1 and 2 to determine if the lack of a list was disadvantageous.

### Results and discussion

Table 1 shows results for the complex population (Fig. 1a) of 50,000 simulations of sampling with a sample size ( $n$ ) of 50 individuals. For Method 1, 5-m square sample plots were used ( $a = 25 \text{ m}^2$ ), and for Method 2, 25-m square plots were used ( $a = 625 \text{ m}^2$ ).

With an area sampling frame for both Methods 1 and 2 and when sampling from a list sampling frame, the mean of the simulation estimates of the individual tree mean basal area was the same (at least to the nearest  $0.001 \text{ m}^2$ ) as the true mean, supporting the contention that all three methods were unbiased estimators as expected for simple random samples. In all three cases, the average of the simulation estimates of the standard error of the tree mean basal area was close to that of the actual standard error of the 50,000 simulation estimates of the mean. They were close also to the true standard error of the population mean for a sample size of  $n = 50$ , determined as  $\left\{ \left[ \sum_{i=1 \dots N} (y_i - \bar{Y})^2 \right] / [nN] \right\}^{1/2}$ , where  $N$  was the population size,  $y_i$  was the value of the target variable of the  $i$ th member of the population and  $\bar{Y} = (\sum_{i=1 \dots N} y_i) / N$ , that is, all three sampling methods appear

to have operated satisfactorily in estimating the mean of the target variable and its standard error.

However, whilst sampling with an area sampling frame yielded similar results for estimates of the mean of the individuals as did sampling from a list sampling frame, there was a price to pay for the absence of the list. In the case of Method 1, an average of 1388 of the 5-m square sample plots had to be visited to find 50 that contained an individual to make up the required sample. For Method 2, a larger, 25-m square, plot size was used with a value of  $M = 8$ ; that value of  $M$  was one more than the maximum number of trees that occurred in any sample plot. In that case, an average of 480 plots still had to be visited to find the required sample of 50 individuals; if a higher value had been used for  $M$ , an even greater number of plots would have to have been visited. Where there was a list sampling frame, there was no such cost and the sampler needed to visit only the 50 individuals selected from the list. Inevitably, there will always be such a price to pay in the absence of a list sampling frame; this is considered further below.

To estimate the population total basal area, an estimate was required of the total number of individuals in the population when Methods 1 and 2 were applied. This estimate was made for any one sample using Eq. (5). The results of Table 1 show that the true value (720 individuals) was estimated quite closely with Method 2 (averaging 723 individuals), but rather less closely with Method 1 (averaging 734 individuals). The poorer result with Method 1 no doubt reflects the fact that an average

**Table 1** Results from 50,000 simulations of inventory of the complexly arranged population of individual trees (Fig. 1a) to estimate individual tree mean basal area, number of individuals in the population and population total basal area, together with their standard errors. In some simulations it was assumed that a list sampling frame was available and, in others, it was assumed that only an area sampling frame list was available. Sample size ( $s$ ) was 50 individuals. Sample plot size for Method 1 was 5-m square plots and for Method 2 was 25-m square plots

	Individual tree mean basal area ( $\text{m}^2$ )	Number of trees in the population	Population total basal area ( $\text{m}^2$ )
True value	0.662	720	476
True standard error of a sample	0.036	–	–
Sampling with an area sampling frame—Method 1			
Mean of all simulations	0.662	734	486
Standard error of simulation mean estimates	0.036	104	74
Mean of estimates of standard error of simulation means	0.036	102	75
Sampling with an area sampling frame—Method 2			
Mean of all simulations	0.662	723	478
Standard error of simulation mean estimates	0.036	43	35
Mean of estimates of standard error of simulation means	0.036	42	35
Sampling with a list sampling frame			
Mean of all simulations	0.662	–	476
Standard error of simulation mean estimates	0.035	–	25
Mean of estimates of standard error of simulation means	0.035	–	25

of 1388 of the small 5-m square sample plots contained no individuals in taking any sample with that method whilst only 50 (the sample size) contained an individual. This accounts also for the rather large standard error that accompanied these estimates (an average of 102 individuals). The many fewer plots visited and the many more individuals counted in the case of Method 2 ensured its much smaller standard error of the estimates (an average of 42 individuals) and much closer approach on average to the true value. The corresponding results for the estimates of the population total basal area (obtained using Eq. 4) were close to the true value for Method 2 and rather more deviant, with a much larger standard error, for Method 1.

The average of the estimates from bootstrapping, used in both Methods 1 and 2 to estimate the standard error of the estimate of the total basal area, agreed closely with the actual standard error of the 50,000 simulation estimates, that is, bootstrapping as applied here appeared to be an appropriate method to estimate the standard error for any one sample.

The results of the simulations done with the regularly arranged population depicted in Fig. 1b are given in Table 2. It is evident that the regular arrangement has yielded very satisfactory results for Methods 1 and 2, with the averages of their simulations close to the population true values and with much reduced standard errors of estimates for the population total basal area. Because much larger, 25-m square, sample plots could be used in this case with Method 1, an average of only 64 plots had to be visited to find the 50 sample individuals, rather than the 1388 5-m square plots that had to be visited with the

complexly arranged population. For Method 2, an average of 88 50-m square plots had to be visited rather than the 480 25-m square plots required in the previous example. These reductions in sampling effort would make both methods practically more feasible than was the case for the complexly arranged population. They led also to substantial reductions in the standard errors of the estimates of both population numbers of individuals and total basal area, with Method 2 only slightly superior in this regard to Method 1.

It is interesting to compare these results with those of Gordon and Pont (2015), who tested a number of different sampling methods to estimate the mean stem wood volume of individual trees in five populations of plantation *Pinus radiata* in New Zealand. Four of the populations varied in area over the range 1.3–2.3 ha, whilst the fifth was 22.8 ha in area. A list sampling frame was available for each population with a detailed list of the trees with their stem wood volumes and spatial locations. Through simulation studies, they tested five sampling methods that could be applied in the absence of a list sampling frame to provide samples that aimed to mimic simple random samples. They compared these results with actual simple random sampling that used the list sampling frame. In effect, their five alternative methods involved forms of cluster sampling where sample points were located randomly across the population area and the trees that occurred in a plot surrounding those points were included in the sample; that contrasts with Iles' approach where no more than one tree from any such plot is included in the sample. Whilst the methods Gordon and Pont used are not necessarily unbiased

**Table 2** Results, in the same form as Table 1, from 50,000 simulations of inventory of the squarely arranged population of individual trees illustrated in Fig. 1b. In this case, sample plot size for Method 1 was 25-m square plots and for Method 2 was 50-m square plots

	Individual tree mean basal area (m <sup>2</sup> )	Number of trees in the population	Population total basal area (m <sup>2</sup> )
True value	0.646	676	437
True standard error of a sample	0.036	–	–
Sampling with an area sampling frame—Method 1			
Mean of all simulations	0.646	679	439
Standard error of simulation mean estimates	0.036	44	38
Mean of estimates of standard error of simulation means	0.036	44	37
Sampling with an area sampling frame—Method 2			
Mean of all simulations	0.646	678	438
Standard error of simulation mean estimates	0.036	39	35
Mean of estimates of standard error of simulation means	0.036	39	35
Sampling with a list sampling frame			
Mean of all simulations	0.646	–	437
Standard error of simulation mean estimates	0.034	–	23
Mean of estimates of standard error of simulation means	0.034	–	23

estimators, a number of the methods they tested gave estimates of population means with only a slight bias and often with precision close to that of simple random sampling. However, none gave results that were completely consistent with simple random sampling across the several populations they considered. Their methods certainly reduce the effort involved in taking a sample of any particular size when compared with the methods considered here. However, if the sampler is willing to live with the possible risk of some bias and less certain estimates of precision when compared with actual simple random sampling, Gordon and Pont's methods may prove satisfactory.

### Conclusions

The results showed that it is possible to take simple random samples of individuals from a population that has only an area sampling frame available. This may be done using the methods of Iles (1979) that involve selecting sample individuals from randomly located plots. Such samples yielded results for population estimates of the mean of individuals that were the same as those obtained when a list sampling frame was available.

However, when Iles' methods are used, a preliminary survey of the population must be made before sampling starts and it may be necessary to visit many more sampling units to obtain the required sample than is the case when a list is available. The more complex the spatial arrangement of individuals within the population, the greater will be the number of sampling units that must be visited.

Not only estimates of the population mean but also estimates of the population total can be obtained when sampling using Iles' methods (Eq. 4). However, there will be a reduction in the precision of those estimates because of the need to estimate, from the sample data, the total number of individuals in the population. Estimates of the standard error of these estimates may be obtained satisfactorily using bootstrapping.

Because it is impossible to predict before sampling how many sample plots will have to be visited to obtain the required sample size, it must be recognised that the methods used here preclude systematic sampling. In that case, the number of sampling units to be visited at pre-determined, regular intervals throughout the population must be known in advance of sampling.

### Acknowledgements

I was stimulated to undertake this work through discussions with Professor Tim Gregoire. Dr. Kim Iles, whose methods I have used, gave invaluable advice in developing the present manuscript. Journal reviewers provided useful suggestions for improvement.

### Competing interests

The author declares that he has no competing interests.

Received: 19 April 2016 Accepted: 19 July 2016

Published online: 21 August 2016

### References

- Cochran, WG (1999). *Sampling techniques* (3rd ed.). New York: Wiley.
- Goodman, LA (1960). On the exact variance of products. *Journal of the American Statistical Association*, 55(4), 708–713.
- Gordon, AD, & Pont, D (2015). Inventory estimates of stem volume using nine sampling methods in thinned *Pinus radiata* stands, New Zealand. *New Zealand Journal of Forestry Science*, 45, 8.
- Gregoire, TG, & Valentine, HT (2008). *Sampling strategies for natural resources and the environment*. Boca Raton: Chapman & Hall/CRC.
- Iles, K (1979). *Systems for the selection of truly random samples from tree populations and the extension of variable plot sampling to the third dimension*. Vancouver: Ph.D. thesis, Department of Forestry, The University of British Columbia. Available from [https://circle.ubc.ca/bitstream/handle/2429/22171/UBC\\_1979\\_A1%20I44.pdf?sequence=1](https://circle.ubc.ca/bitstream/handle/2429/22171/UBC_1979_A1%20I44.pdf?sequence=1). Accessed 26 July 2016.
- Iles, K (2003). *A sampler of inventory topics*. Nanimo: Kim Iles & Associates Ltd.
- Lynch, TB (2015). Voronoi polygons quantify bias when sampling the nearest plant. *Canadian Journal of Forest Research*, 45(12), 1853–1859.
- Pinkham, RS (1987). An efficient algorithm for drawing a simple random sample. *Applied Statistics*, 36(3), 370–372.
- Schreuder, HT, Gregoire, TG, & Wood, GB (1993). *Sampling methods for multiresource forest inventory*. New York: Wiley.
- Schreuder, HT, Ouyang, Z, & Williams, M (1992). Point-Poisson, point-pps, and modified point-pps sampling: efficiency and variance estimation. *Canadian Journal of Forest Research*, 22(8), 1071–1078.
- West, PW (2013). Precision of inventory using different edge overlap methods. *Canadian Journal of Forest Research*, 43(11), 1081–1083.
- West, PW (2015). *Tree and forest measurement* (3rd ed.). Switzerland: Springer International Publishing.
- West, PW (2016). Population structure and correlation between auxiliary and target variables may affect precision of estimates in forest inventory. *Communications in Statistics—Simulation and Computation*, in press, doi:10.1080/03610918.2016.1139128.

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](http://springeropen.com)