Assessment of the Effect of Six Methods of Analysis and Different Sample Sizes for Biomass Estimation in Grasslands of the State of Puebla, Mexico

Efrain Velasco-Bautista 1, Martin Enrique Romero-Sanchez 1,*, David Meza-Juárez 2 and Ramiro Pérez-Miranda 1

Abstract: In the assessment of natural resources, such as forests or grasslands, it is common to apply a two-stage cluster sampling design, the application of which in the field determines the following situations: (a) difficulty in locating secondary sampling units (SSUs) precisely as planned, so that a random pattern of SSUs can be identified; and (b) the possibility that some primary sampling units (PSUs) have fewer SSUs than planned, leading to PSUs of different sizes. In addition, when considering the estimated variance of the various potential estimators for two-stage cluster sampling, the part corresponding to the variation between SSUs tends to be small for large populations, so the estimator’s variance may depend only on the divergence between PSUs. Research on these aspects is incipient in grassland assessment, so this study generated an artificial population of 759 PSUs and examined the effect of six estimation methods, using 15 PSU sample sizes, on unbiased and relative sampling errors when estimating aboveground, belowground, and total biomass of halophytic grassland. The results indicated that methods 1, 2, 4, and 5 achieved unbiased biomass estimates regardless of sample size, while methods 3 and 6 led to slightly biased estimates. Methods 4 and 5 had relative sampling errors of less than 5% with a sample size of 140 when estimating total biomass.

Keywords: two-stage cluster sampling; halophytic grassland biomass; sample size; unbiased estimators

1. Introduction

It is common to use statistical sampling for parameter estimation in natural resource assessment, such as mean, total, or proportion [1,2]. In sampling theory, there are two types of inference: design-based and model-based. In design-based inference, randomisation is used to select the population units to be measured; for statistical validity, estimators are constructed based on this randomisation [3–5]. An immediate implication is that the values of the variables of interest are treated as fixed; furthermore, the randomness in the estimators is only due to the random selection of the samples [6]. The inference is based on the variance among all possible samples selected for the population with a given sampling design. The expected value and variance of the estimators are based on the random variation of the different estimates among all possible samples.

In contrast to design-based inference, the values attached to the population elements are realisations of random variables in model-based inference. Therefore, quantities of interest such as population totals or means are also random variables. A basic assumption of model-based inference is that the random values of the population elements follow some specific model. Therefore the randomness is only due to the model used to describe the population [7].

Simple random, systematic, stratified random, and cluster random sampling presented in classical sampling texts [8,9] belong to the paradigm of design-based inference. In studying natural populations, such as forests or grasslands, applying the cluster sampling
design is expected to be conceptualised in two stages. In this case, primary sampling units (stage 1) are formed by blocks of land easily identifiable in the map base, circles or squares of one hectare. Secondary sampling units (stage 2) are constituted by fractions of land called plots. The shape and size vary depending on the time to plot and make measurements and the variation between plots for each of the variables of interest: circles, squares, or rectangles smaller than one hectare [10].

In Mexico’s National Forest and Soil Inventory (INFyS), the sampling design considers one hectare circular plots as PSUs and subplots of 400 m$^2$ for measuring trees; 12.56 m$^2$ for recording regrowth; and 1 m$^2$ for assessing herbs, ferns, mosses, and lichens as SSUs. In this case, both PSUs and SSUs are distributed in the field in a systematic pattern [11]. Using simple random sampling estimators on systematically drawn samples has been accepted in forest inventories because systematic coverage eliminates clustering of sampling units and loss of precision that could occur with a purely random sample.

In the application of a two-stage cluster design, it is necessary to consider that: (a) there is a random pattern of SSUs even when a fixed or systematic configuration is defined; (b) there is a possibility that when sampling in the field, some PSUs will have fewer SSUs than planned; (c) when considering the estimated variance of the potential estimators for two-stage cluster sampling, the variance of SSUs tends to be small for large populations; (d) for conventional inventory applications, the point estimator refers to the aggregation of the data from the SSUs and the PSU is processed as a single unit—thus, the variance of the estimator considers only the variation between PSUs; and (e) PSUs are the main element in two-stage cluster sampling, so it is of particular interest to assess the number of PSUs in the precision of the estimation [12,13].

Therefore, and considering that research on these aspects is incipient in grassland assessment, this study examined the effect of six estimation methods, using 15 PSU sample sizes, on unbiased and relative sampling errors when estimating aboveground, belowground, and total biomass of halophytic grassland in Puebla, Mexico. Three of these methods strictly considered two-stage cluster sampling for the estimator’s variance. The other three took the PSU as a single unit; consequently, the estimated variance was simplified, and only considered the variation between PSUs. Therefore, the objective of this study was to evaluate the effect of six methods of analysis, using different sample sizes, on the unbiased and relative sampling errors in the estimation of aboveground, belowground, and total biomass in halophytic grassland. Our hypothesis assumed that under different sample sizes, the six analysis methods would not affect the unbiased and relative sampling errors in estimating aboveground, belowground, and total biomass in halophytic grassland.

2. Materials and Methods
2.1. Study Area

The study area was located in the common lands of the south-central portion of the municipality of Tepeyahualco in the state of Puebla, Mexico (Figure 1). The study area had an extension of 746 ha, and the predominant vegetation was halophilic. This type of vegetation inhabits saline, alkaline, and poorly drained soils at the bottom of old lakes, and usually at an altitude of no more than 2250 m. It takes the form of low, dense grassland dominated by grasses that reproduce vegetatively by rhizomes and stolons [14].

2.2. Sampling Design

A two-stage cluster sampling design was implemented in the field [15,16], in which the primary sampling unit (PSU) was a square plot of one hectare (100 m × 100 m) and the secondary sampling units (SSU) were square plots of 1 m$^2$ (1 m × 1 m) for aboveground biomass and square plots of 0.09 m$^2$ (0.3 m × 0.3 m) for belowground biomass. The 1 m$^2$ SSUs, nested within the PSU, were geometrically arranged in the shape of an X: SSU “a” was located at the centre of the PSU; SSUs “b”, “c”, “d”, and “e” were found at the vertices of a hypothetical square centred on SSU “a”. The 0.09 m$^2$ SSUs were located adjacent to the
1 m² SSUs (Figure 2). SSUs nested in PSUs for sampling minor vegetation (herbs, ferns, and mosses) are common in large-scale forest inventories [11].

Considering that in the study, the highest cost was generated at the arrival at the PSUs, and assuming a coefficient of variation of 45% in the aboveground biomass, a sample size of 69 sampling points was determined, which constituted the centre of the PSUs, which were systematically distributed over an area of 746 hectares (Figure 1). In forest inventories, systematic sampling is preferred to random sampling; the uniform dispersion of sampling units in the target population makes it attractive from a theoretical and practical point
of view [17–19]. In Mexico, about 26,000 clusters distributed in a systematic network are assessed in the National Forest and Soil Inventory [19,20].

The variables recorded in the field were the following: dominant species, codominant species, vegetation cover, ground cover, average height, and fresh weight of aboveground biomass and belowground biomass. The dry weight of vegetation collected during the field sampling was obtained in the laboratory for both the above- and belowground parts.

2.3. Analysis Used

Although the SSUs theoretically had a fixed pattern, the SSUs in the field had a random design for geolocation reasons. Even the number of SSUs can be less than five, leading to cluster sampling of different sizes. Thus, six methods of analysis were studied for the estimation of aboveground, belowground, and total biomass. Of these, three methods considered the sampling design strictly in two stages [8,21,22], and three in one step, assuming that the variance between SSUs was negligible [12,23,24].

Next, the six methods of analysis are presented (Table 1), which consisted of the point estimators at the SSU and PSU levels. The respective variances of these estimators, and covariance between the estimator for aboveground biomass and the estimator for belowground biomass, are presented in Appendix A. The covariance between estimators considering the two-stage design was obtained as an extension of the covariance between estimators considering one stage [8,12] and applying the basic principles of error propagation.

Table 1. Methods of analysis used in this study.

| Method | Unbiased Estimator for Biomass: SSU Level | Unbiased Estimator for Biomass: PSU Level (ha) |
|--------|------------------------------------------|-----------------------------------------------|
| 1. Unbiased estimator and PSU of different sizes (two-stage sampling) | $\hat{\mu}_{i} = \frac{1}{n_{i}} \sum_{j=1}^{n_{i}} y_{ij}$ | $\hat{\mu}_{ha} = 10,000 \times \hat{\mu}_{i}$ |
| 2. Ratio estimator and PSU of different sizes (two-stage sampling) | $\hat{\mu}_{i} = \frac{1}{n_{i}} \sum_{j=1}^{n_{i}} \frac{y_{ij}}{M_{ij}}$ | $\hat{\mu}_{ha} = 10,000 \times \hat{\mu}_{i}$ |
| 3. Simple mean estimator and PSU of different sizes (two-stage sampling) | $\hat{\mu}_{i} = \frac{1}{n_{i}} \sum_{j=1}^{n_{i}} y_{ij}$ | $\hat{\mu}_{ha} = 10,000 \times \hat{\mu}_{i}$ |
| 4. Ratio estimator and PSU of different sizes (one-stage sampling) | N/A | N/A |
| 5. VAN DEUSEN * estimator and PSU of different sizes (one-stage sampling) | N/A | N/A |
| 6. Mean ratio estimator and PSU of different sizes (one-stage sampling) | N/A | N/A |

Where: $n$: number of PSUs from the sample; $M_{ij}$: number of SSUs from the $i$-th conglomerate; $M = \sum_{i=1}^{N} M_{i}$: number of SSUs in the population; $\overline{M} = \frac{M}{N}$: mean size of the PSUs in the population; $\overline{m} = \frac{\sum_{i=1}^{N} M_{i}}{n}$: mean size of the PSUs in the sample; $m_{i}$: number of SSUs selected in the sample from the $i$-th conglomerate; $\overline{m} = \overline{m}/N$: mean of the number of SSUs selected in the sample by conglomerate; $y_{ij}$: biomass observation in the $j$-th SSU in the $i$-th PSU; $y_{i}$: biomass sample in the $i$-th PSU; $\overline{y}_{i}$: mean sample biomass for the $i$-th PSU; $x_{i}$: effective sampled area to estimate biomass in the $i$-th PSU. * The VAN DEUSEN estimator calculation had the same structure as the ratio estimator; however, the variance was calculated differently (see Appendix A).

2.4. Generation of an Artificial Population

By considering the sample of 69 PSUs and the total area of 746 ha as representative, an artificial population of 759 PSUs was generated using PROC SURVEYSELECT of the Statistical Analysis System SAS/STAT® v9.3 [25]. This procedure ensured that the population had the same mean and variance as the original sample of 69 sampling units for the aboveground and belowground biomass variables.

From the simulated population of 759 PSUs, random samples of the following sizes were obtained: 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, and 150. Each random selection corresponding to the given sample size was repeated 300 times (replicates); all six estimation methods were applied in each case. The point estimators and their respective estimated variances were programmed directly into the SAS/STAT® v9.3 Statistical Analysis System using statistical procedures that explicitly generated results at the level of SSU and PSU (estimates per hectare). This procedure yielded a database of 27,000 observations with the same number of punctual estimations and interval estimates of aboveground, belowground, and total biomass. This range of sample sizes was determined by considering...
the approximate centre position of the original sample size in the series and oscillation between 20 and 5% of the relative sampling error when estimating total biomass.

The number of replicates was defined while considering at least 100 replicates for variance estimates by resampling [26]; furthermore, the estimator’s normality was evident under these conditions.

The relative estimation bias (REB) was derived according to the expression suggested by Van Deusen [23]:

$$ REB = \frac{\hat{\theta} - \theta}{\theta} $$

where \( \hat{\theta} = \sum_{b=1}^{300} \hat{\theta}_b \), and in every case, \( \hat{\theta}_b \) was obtained according to the estimator of each analysis method; \( \theta = 472.411 \) is the aboveground biomass parameter (kg·ha\(^{-1}\)); \( \theta = 2159.408 \) is the belowground biomass parameter (kg·ha\(^{-1}\)); and \( \theta = 2631.82 \) is the total biomass parameter (kg·ha\(^{-1}\)).

In the expression relative to bias, a value close to 0 means that the estimator is unbiased. A value of 1 shows that the formula predicts the parameter twice, and a value of 2 indicates overestimation by a factor of 3. In the present research, the condition of the unbiased estimator implied relative biases close to zero (less than 0.05). The relative bias was unitless, and was interpreted as an indicator of the number of times the estimator exceeded or contained the true parameter or value. This way of presenting the bias made it easy to identify when there was an under- or overestimation.

The relative sampling error (coefficient of variation) was calculated using the following expression [21]:

$$ \hat{RSE}(\%) = \left( \frac{\sqrt{\hat{\sigma}(\hat{\theta})}}{\hat{\theta}} \right) \times 100 $$

A relative sampling error of 5% was considered as a reference for the analysis.

3. Results

3.1. Aboveground Biomass Estimations

Methods 1, 2, 4, and 5 had the same average estimates of aboveground biomass per hectare for all sample sizes tested. These estimates oscillated around the population value, 472.41 kg·ha\(^{-1}\), making them practically unbiased. The most significant relative bias occurred for sample size 10, with a value of 0.0099, corresponding to an estimate of 477.09 kg·ha\(^{-1}\). The 95% confidence intervals for methods 2, 4, and 5 were similar for all sample sizes; as expected, the widest intervals were for the earlier sample sizes, while the narrowest intervals were for the later ones. For example, method 2 for sample size 10 had estimates from 285.89 to 668.29 kg·ha\(^{-1}\), while for sample size 150, the estimates ranged from 432.62 to 514.09 kg·ha\(^{-1}\). Method 1, on the other hand, presented wider confidence intervals than methods 2, 4, and 5; so that for sample size 10, the estimate ranged from 265.19 to 688.99 kg·ha\(^{-1}\), and for sample size 150, it was from 429.08 to 517.63 kg·ha\(^{-1}\) (Figure 3).

In the different sample sizes studied, methods 3 and 6 led to the same average aboveground biomass results: they slightly underestimated the population value in all cases. The largest relative bias was obtained for sample size 80, which was −0.0384, corresponding to an estimate of 454.29 kg·ha\(^{-1}\). Both methods presented similar confidence intervals and with remarkable amplitude at small sample sizes; for example, for sample size 10 of method 3, the estimate of aboveground biomass was from 266.84 to 657.35 kg·ha\(^{-1}\), while for sample size 150, it was from 417.15 to 500.67 kg·ha\(^{-1}\) (Figure 3).

Although methods 1, 2, 4, and 5 led to unbiased estimates, it is worth noting that method 2 essentially achieved a relative sampling error of 5% at sample size 120; while for the same sample size, methods 1, 4, and 5 obtained values of 5.43, 4.85, and 4.85%, respectively (Table 2). For methods 3 and 6, which were slightly biased, at sample size 120, they reached relative sampling errors of 5.27 and 5.15%, respectively (Table 2). Under these conditions, methods 4 and 5 proved to be the best.
Figure 3. Confidence intervals for estimation of aboveground biomass from six analysis methods (a) M1, (b) M2, (c) M3, (d) M4, (e) M5, and (f) M6 and 15 sample sizes.

Table 2. Relative sampling error (%) in the estimation of aboveground biomass from six analysis methods and 15 sample sizes.

| Method | 10  | 20  | 30  | 40  | 50  | 60  | 70  | 80  | 90  | 100 | 110 | 120 | 130 | 140 | 150 |
|--------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 1      | 19.95 | 14.06 | 11.38 | 9.80 | 8.74 | 7.87 | 7.29 | 6.83 | 6.35 | 5.96 | 5.69 | 5.43 | 5.17 | 4.93 | 4.74 |
| 2      | 17.97 | 12.75 | 10.42 | 8.95 | 7.97 | 7.20 | 6.63 | 6.25 | 5.82 | 5.47 | 5.22 | 4.96 | 4.75 | 4.52 | 4.36 |
| 3      | 18.98 | 13.52 | 10.98 | 9.46 | 8.46 | 7.63 | 7.03 | 6.62 | 6.17 | 5.73 | 5.37 | 5.11 | 4.85 | 4.64 | 4.39 |
| 4      | 17.94 | 12.70 | 10.36 | 8.89 | 7.90 | 7.12 | 6.55 | 6.16 | 5.72 | 5.37 | 5.11 | 4.85 | 4.64 | 4.39 | 4.23 |
| 5      | 18.05 | 12.74 | 10.38 | 8.90 | 7.91 | 7.13 | 6.55 | 6.17 | 5.73 | 5.37 | 5.11 | 4.85 | 4.64 | 4.39 | 4.23 |
| 6      | 18.95 | 13.48 | 10.92 | 9.40 | 8.39 | 7.56 | 6.95 | 6.53 | 6.07 | 5.70 | 5.42 | 5.15 | 4.92 | 4.65 | 4.48 |

Belowground Biomass Estimations

Methods 1, 2, 4, and 5 presented identical belowground biomass estimates per hectare for the 15 sample sizes studied. These values were close to the parameter, 2159.41 kg ha⁻¹; the largest relative bias was 0.0171, corresponding to the estimate of 2196.43 kg ha⁻¹ for sample size 10. Thus, such methods led to unbiased estimates of belowground biomass per hectare. The 95% confidence intervals for methods 2, 4, and 5 resulted in a high degree of similarity; for small sample sizes, the range was wider than for large sample sizes; e.g., for method 2 and sample size 10, the estimate was 1176.46 to 3216.41 kg ha⁻¹. In contrast, for sample size 150, the results ranged from 1939.30 to 2383.40 kg ha⁻¹. Compared to methods 2, 4, and 5, method 1 presented wider confidence intervals; thus, estimates from 1097.18 to 3295.69 kg ha⁻¹ for sample size 10 and 1926.08 to 2396.62 kg ha⁻¹ for sample size 150 were observed (Figure 4).

Methods 3 and 6 led to the same point values of belowground biomass. For all sample sizes, a slight underestimation was observed. The maximum relative bias was −0.0329, which corresponded to 2088.27 kg ha⁻¹ for sample size 80. The confidence intervals of both methods were similar and wide at small sample sizes; e.g., for sample size 10 of method 3, the estimate was from 1129.42 to 3160.56 kg ha⁻¹, and for sample size 150, from 1892.25 to 2330.47 kg ha⁻¹ (Figure 4).
Belowground Biomass Estimations

Methods 1, 2, 4, and 5 presented identical belowground biomass estimates per hectare for the 15 sample sizes studied. These values were close to the parameter, 2159.41 kg ha\(^{-1}\); the largest relative bias was 0.0171, corresponding to the estimate of 2196.43 kg ha\(^{-1}\) for sample size 10. Thus, such methods led to unbiased estimates of belowground biomass per hectare. The 95% confidence intervals for methods 2, 4, and 5 resulted in a high degree of similarity; for small sample sizes, the range was wider than for large sample sizes; e.g., for method 2 and sample size 10, the estimate was 1176.46 to 3216.41 kg ha\(^{-1}\). In contrast, for sample size 150, the results ranged from 1939.30 to 2383.40 kg ha\(^{-1}\).

Compared to methods 2, 4, and 5, method 1 presented wider confidence intervals; thus, estimates from 1097.18 to 3295.69 kg ha\(^{-1}\) for sample size 10 and 1926.08 to 2396.62 kg ha\(^{-1}\) for sample size 150 were observed (Figure 4).

Methods 3 and 6 led to the same point values of belowground biomass. For all sample sizes, a slight underestimation was observed. The maximum relative bias was \(-0.0329\), which corresponded to 2088.27 kg ha\(^{-1}\) for sample size 80. The confidence intervals of both methods were similar and wide at small sample sizes; e.g., for sample size 10 of method 3, the estimate was from 1129.42 to 3160.56 kg ha\(^{-1}\), and for sample size 150, from 1892.25 to 2330.47 kg ha\(^{-1}\) (Figure 4).

Although methods 1, 2, 4, and 5 were unbiased, and on average equal to the population value, it was interesting that methods 4 and 5 achieved a relative sampling error of almost 5\% at the sample size of 150. For this same number of observations, methods 1 and 2 had relative sampling errors of 5.50 and 5.20\%. Likewise, when considering the slightly biased methods, for the sample size of 150, method 3 had a relative sampling error of 5.25\%, and method 6 of 5.12\% (Table 3). Thus, methods 4 and 5 proved to be the best in terms of precision.

**Table 3. Relative sampling error (%) in estimating belowground biomass from six analysis methods and 15 sample sizes.**

| Method | Sample Size |
|--------|-------------|
|        | 10  | 20  | 30  | 40  | 50  | 60  | 70  | 80  | 90  | 100 | 110 | 120 | 130 | 140 | 150 |
| 1      | 22.03 | 15.85 | 13.05 | 11.26 | 9.20 | 8.56 | 7.91 | 7.43 | 6.97 | 6.63 | 6.34 | 6.06 | 5.78 | 5.50 |
| 2      | 20.43 | 14.78 | 12.25 | 10.58 | 9.48 | 8.64 | 8.03 | 7.43 | 7.00 | 6.57 | 6.24 | 5.96 | 5.70 | 5.45 | 5.20 |
| 3      | 21.00 | 15.11 | 12.47 | 10.73 | 9.65 | 8.77 | 8.13 | 7.54 | 7.08 | 6.65 | 6.31 | 6.04 | 5.77 | 5.51 | 5.25 |
| 4      | 20.39 | 14.74 | 12.20 | 10.53 | 9.42 | 8.57 | 7.95 | 7.34 | 6.91 | 6.47 | 6.14 | 5.86 | 5.59 | 5.33 | 5.07 |
| 5      | 20.51 | 14.78 | 12.22 | 10.54 | 9.43 | 8.58 | 7.96 | 7.35 | 6.92 | 6.47 | 6.14 | 5.86 | 5.60 | 5.33 | 5.08 |
| 6      | 20.96 | 15.06 | 12.42 | 10.67 | 9.58 | 8.69 | 8.05 | 7.45 | 6.99 | 6.55 | 6.20 | 5.93 | 5.65 | 5.39 | 5.12 |

3.2. Total Biomass Estimation

Methods 1, 2, 4, and 5 presented equal estimates of total biomass per hectare. Regardless of sample size, on average, these estimates were equal to the parameter, 2631.82 kg ha\(^{-1}\). The largest relative bias was observed for sample size 10, with a value of 0.0158, corresponding to an estimate of 2673.52 kg ha\(^{-1}\). The 95\% confidence intervals for methods 2, 4, and 5 were similar and wide at small sample sizes. Thus, for example, using method 2, for sample size 10, the total biomass was estimated from 1528.05 to 819.00 kg ha\(^{-1}\), while for sample size 150, the values ranged from 2386.14 to 2883.27 kg ha\(^{-1}\). Method 1, compared to methods 2, 4, and 5, presented relatively wider confidence intervals: from 1423.56 to
3923.49 kg ha$^{-1}$ for sample size 10, and from 2368.57 to 2900.83 kg ha$^{-1}$ for sample size 150 (Figure 5).

Table 4. Relative sampling error (%) in estimating total biomass from six analysis methods and 15 sample sizes.

| Method | Sample Size |
|--------|-------------|
| 10     | 20          | 30          | 40          | 50          | 60          | 70          | 80          | 90          | 100         | 110         | 120         | 130         | 140         | 150         |
| 1      | 20.60       | 14.80       | 12.14       | 10.48       | 9.41        | 8.54        | 7.94        | 7.35        | 6.89        | 6.46        | 6.13        | 5.88        | 5.62        | 5.36        | 5.11        |
| 2      | 18.84       | 13.64       | 11.27       | 9.74        | 8.71        | 7.93        | 7.37        | 6.83        | 6.43        | 6.02        | 5.73        | 5.47        | 5.24        | 4.99        | 4.77        |
| 3      | 19.49       | 14.03       | 11.55       | 9.95        | 8.93        | 8.11        | 7.52        | 6.98        | 6.54        | 6.14        | 5.83        | 5.59        | 5.33        | 5.08        | 4.86        |
| 4      | 18.81       | 13.60       | 11.22       | 9.69        | 8.65        | 7.87        | 7.30        | 6.75        | 6.35        | 5.94        | 5.64        | 5.38        | 5.14        | 4.89        | 4.66        |
| 5      | 18.92       | 13.63       | 11.24       | 9.70        | 8.66        | 7.87        | 7.30        | 6.76        | 6.35        | 5.94        | 5.64        | 5.38        | 5.14        | 4.89        | 4.66        |
| 6      | 19.46       | 13.99       | 11.50       | 9.89        | 8.87        | 8.04        | 7.44        | 6.90        | 6.46        | 6.05        | 5.74        | 5.49        | 5.23        | 4.98        | 4.74        |
4. Discussion

In all study methods and for all three biomass sources, significant changes in relative sampling errors were observed for the first 10 sample sizes. In contrast, for the remaining five sample sizes, the relative sampling errors remained almost stable. In methods 4 and 5 for estimating total biomass, differences in relative sampling errors of about 13% were recorded when moving from sample size 10 to 100, and 1% when moving from sample size 110 to 150.

For total, aboveground, and belowground biomass estimation, methods 1, 2, 4, and 5 were found to be in the group of unbiased estimators, while methods 3 and 6 were in the class of slightly biased estimators. The unbiasedness or slight bias occurred independently of the sample sizes studied. Within each group, the same estimates were obtained. In the original sample of 69 PSUs, there were 18 PSUs smaller than 5 SSUs. The analysis was conceptualised under the theory of cluster sampling of unequal sizes [22,27,28]. Even the application of method 3 considered the average of the means of PSUs of different sizes. Under these considerations and through algebraic manipulations, the equivalence of these estimators can be demonstrated. In Appendices B and C, the demonstrations for aboveground and belowground biomass, respectively, are presented; in each case, adjustments were made for the size of the SSU and the number of SSUs in the PSU.

While methods 1, 2, 4, and 5 presented unbiased estimates for all sample sizes studied, it was interesting that in the estimation of total, aboveground, and belowground biomass, methods 4 and 5 led to the lowest relative errors. The ratio estimator considered in these methods used the area sampled in the field in each PSU as an auxiliary variable. The objective of the ratio estimator was to obtain increased precision by taking advantage of the correlation between the variable of interest and the auxiliary variable [8,29]. In addition to its statistical advantages, the one-stage ratio estimator is simple in terms of computational applicability, and has the advantage of direct estimation per hectare of biomass or carbon when the area sampled in the field is used as an auxiliary variable. This estimator has been recommended for the assessment of forest parameters in forests. In the Forest Inventory Analysis Program (FIA) and the National Forest Health Monitoring Program (NFHM) in the United States, the ratio estimator can be applied in the estimation of: (a) attributes per hectare, (b) attributes at tree level, and (c) attributes at stand level [12]. Similarly, it is recommended for the estimation of tree density in urban areas (trees per kilometre) [30] and the estimation of attributes or health indicators measured at the tree crown level [31,32]. In this case, the auxiliary variable refers to the number of trees recorded per plot or sampling unit. On the other hand, recent research indicates that the Van Deusen mean of ratios and ratio of means estimators perform well in estimating unbiasedness and variance of timber volume and density as the degree of completeness of the clusters decreases [19,27,33]. A close inspection of the variances of estimators 1, 2, and 3 allows observing that the second term contains the expression \( \frac{1}{N} = \left( \frac{N}{n} \right) \left( \frac{1}{n} \right) \), so when \( \frac{N}{n} \) is negligible, the estimated variances can be calculated from the knowledge of the variability of the PSUs only [8]. This result is useful when the subsampling is systematic because, in this case, it is not possible to obtain an unbiased estimator of the variance between SSUs. This theoretical argument helps explain why the variances, and consequently, the confidence intervals and relative sampling errors of the two-stage estimators (1, 2, and 3), were relatively like the one-stage estimators (4, 5, and 6). The fact that the relative sampling error of estimator 1 was slightly larger than that of estimator 2 can be explained by the fact that the variance between PSUs of the former considered the average cluster size, while the latter included the size of the cluster as such. As the PSUs became incomplete, the difference in relative errors of the two estimators became more noticeable. Furthermore, in simulations of small sample sizes where it was possible to achieve complete PSUs, all six estimators led to the same relative sampling errors for aboveground, belowground, and total biomass. In the first term of the variance of the two-stage estimators or the single term of the single-stage estimators, the factor \( \frac{1}{n} \) or \( \frac{1}{\sum \frac{1}{x_i}} \) (method 5) was observed, which, as can be seen, decreased as the sample size increased. For this reason, the relative
sampling errors of the six methods studied for the estimation of total, aboveground, and belowground biomass decreased as the sample size increased. For the same reason, the width of the confidence intervals decreased as the sample size increased from 10 to 150. The shapes of the relative sampling error curves obtained in this research were congruent with those obtained in other studies of this nature [19,28,29,34].

5. Conclusions

Methods 1, 2, 4, and 5 presented unbiased estimates of halophytic grassland’s total aboveground and belowground biomass irrespective of sample size. Methods 3 and 6 had slightly downward-biased estimates of halophytic grassland’s total, aboveground, and belowground biomass, regardless of sample size. In no case did the bias exceed 0.05. Methods 4 and 5 were the most precise. They had relative sampling errors equal to 5% in the estimation of belowground biomass and less than 5% in the estimation of aboveground and total biomass.

Author Contributions: Conceptualization, E.V.-B. and M.E.R.-S.; methodology, E.V.-B. and M.E.R.-S.; validation, D.M.-J.; formal analysis, E.V.-B.; data curation, M.E.R.-S. and D.M.-J.; writing—original draft preparation, E.V.-B.; writing—review and editing, E.V.-B., M.E.R.-S., D.M.-J. and R.P.-M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author.

Acknowledgments: Thanks to the National Institute of Forestry, Agriculture and Livestock Research for partial support in the development of this research. The authors are deeply grateful for comments and suggestions during the review process, which were helpful in improving the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Six Estimators and Their Respective Variances. In the Case of the Total Biomass Estimation, the Covariance Between the Estimator for Aboveground Biomass and the Estimator for Belowground Biomass are also Presented.

Method 1. Unbiased estimator and PSU of different sizes (two-stage sampling)

1.1. Unbiased estimator for biomass: SSU level (plot)

\[ \hat{\mu}^1 = \frac{1}{M} \sum_{i=1}^{n} M_i \bar{y}_i \]

where \( \overline{M} \) can be estimated by \( \bar{m} \) and \( \hat{\mu}^1 \) identifies the estimator by method 1:

\[ \hat{\sigma}^2(\hat{\mu}^1) = \left(1 - \frac{n}{N}\right) \left(\frac{1}{nM^2}\right) \sum_{i=1}^{n} (M_i \bar{y}_i - \overline{M \hat{\mu}^1})^2 + \left( \frac{1}{nNM^2} \right) \sum_{i=1}^{n} M_i^2 \left(1 - \frac{m_i}{M_i}\right) \left(\frac{1}{m_i}\right) \sum_{j=1}^{m_i} (y_{ij} - \bar{y}_i)^2 \]

\[ \hat{\sigma}^2(\hat{\mu}^1) = \left(1 - \frac{n}{N}\right) \left(\frac{1}{nM^2}\right) s_b^2 + \left( \frac{1}{nNM^2} \right) \sum_{i=1}^{n} M_i^2 \left(1 - \frac{m_i}{M_i}\right) \left(\frac{1}{m_i}\right) s_i^2 \]

where:

\[ s_b^2 = \frac{\sum_{i=1}^{n} (M_i \bar{y}_i - \overline{M \hat{\mu}^1})^2}{n-1} \]

\[ s_i^2 = \frac{\sum_{j=1}^{m_i} (y_{ij} - \bar{y}_i)^2}{m_i - 1}, \quad i = 1, 2, \ldots, n \]
1.2. Unbiased estimator for aboveground biomass: PSU level (hectare)

\[ \hat{\mu}_{ha}^l = 10,000 \times \hat{\mu}^l \]

\[ \sigma(\hat{\mu}_{ha}^l) = 10,000^2 \times \hat{\sigma}(\hat{\mu}^l) \]

1.3. Unbiased estimator for belowground biomass: SSU level (plot)

\[ \hat{\mu}^{1'} = \frac{1}{M} \sum_{i=1}^{n} M_i \bar{y}_i' \]

\[ \sigma(\hat{\mu}^{1'}) = \left(1 - \frac{n}{N}\right) \left(\frac{1}{nM^2}\right) \sum_{i=1}^{n} \frac{(M_i \bar{y}_i' - \bar{M} \hat{\mu}^{1'})^2}{n - 1} + \left(\frac{1}{nNM^2}\right) \sum_{i=1}^{n} M_i^2 \left(1 - \frac{m_i}{M_i}\right) \left(\frac{1}{m_i}\right) \sum_{j=1}^{m_i} (y_{ij}' - \bar{y}_i')^2 \]

\[ \sigma(\hat{\mu}^{1'}) = \left(1 - \frac{n}{N}\right) \left(\frac{1}{nM^2}\right) s_b^{2v} + \left(\frac{1}{nNM^2}\right) \sum_{i=1}^{n} M_i^2 \left(1 - \frac{m_i}{M_i}\right) \left(\frac{1}{m_i}\right) s_i^{2v} \]

where:

\[ s_b^{2v} = \sum_{i=1}^{n} \frac{(M_i \bar{y}_i' - \bar{M} \hat{\mu}^{1'})^2}{n - 1} \]

\[ s_i^{2v} = \frac{\sum_{j=1}^{m_i} (y_{ij}' - \bar{y}_i')^2}{m_i - 1}, \quad i = 1, 2, \ldots, n \]

1.4. Unbiased estimator for belowground biomass: PSU level (hectare)

\[ \hat{\mu}_{ha}' = 111,111 \times \hat{\mu}^{1'} \]

\[ \sigma(\hat{\mu}_{ha}') = 111,111^2 \times \hat{\sigma}(\hat{\mu}^{1'}) \]

1.5. Unbiased estimator for total biomass (aboveground + belowground): PSU level (hectare)

\[ \hat{\mu}_h = \hat{\mu}_{ha} + \hat{\mu}_{ha}' \]

\[ \sigma(\hat{\mu}_h) = \sigma(\hat{\mu}_{ha}) + \sigma(\hat{\mu}_{ha}') + 2 \sigma(\hat{\mu}_{ha}) \sigma(\hat{\mu}_{ha}') \]

\[ \sigma(\hat{\mu}_{ha}) = \sigma(\hat{\mu}_{ha}) + \sigma(\hat{\mu}_{ha}') + 2 \sigma(\hat{\mu}_{ha}) \left(10,000 \times \hat{\mu}^l, 111,111 \times \hat{\mu}^{1'}\right) \]

\[ \sigma(\hat{\mu}_{ha}') = \sigma(\hat{\mu}_{ha}) + \sigma(\hat{\mu}_{ha}') + 2 \sigma(\hat{\mu}_{ha}') \left(10,000 \times \hat{\mu}^l, 111,111 \times \hat{\mu}^{1'}\right) \]

where:

\[ \sigma(\hat{\mu}_h, \hat{\mu}_h') \cong \left(1 - \frac{n}{N}\right) \left(\frac{1}{nM^2}\right) \sum_{i=1}^{n} \frac{(M_i \bar{y}_i' - \bar{M} \hat{\mu}^l)^2(M_i \bar{y}_i' - \bar{M} \hat{\mu}^{1'})}{(M_i + M_i^{1/2})} \left(\frac{1}{M_i}\right) \sum_{j=1}^{m_i} (y_{ij}' - \bar{y}_i')^2 \]

\[ + \left(\frac{1}{nNM^2}\right) \sum_{i=1}^{n} M_i^2 \bar{y}_i' \left(1 - \frac{m_i}{M_i}\right) \left(\frac{1}{m_i}\right) \sum_{j=1}^{m_i} (y_{ij}' - \bar{y}_i') \]

Method 2. Ratio estimator and PSU of different sizes (two-stage sampling)

2.1. Unbiased estimator for biomass: SSU level (plot)

\[ \hat{\mu}_p = \frac{\sum_{i=1}^{n} M_i \bar{y}_i}{\sum_{i=1}^{n} M_i} \]
where \( \hat{\mu}^2 \) identifies the estimator by method 2:

\[
\hat{\sigma}(\hat{\mu}^2) = \left(1 - \frac{n}{N}\right) \left(\frac{1}{nM^2}\right) \sum_{i=1}^{n} \left(\frac{M_i y_i - M^2 \hat{\mu}^2}{n-1}\right)^2 + \left(\frac{1}{nNM^2}\right) \sum_{i=1}^{n} M_i^2 \left(1 - \frac{m_i}{M_i}\right) \left(\frac{1}{m_i}\right) \sum_{j=1}^{m_i} (y_{ij} - \bar{y}_i)^2
\]

\[
\hat{\sigma}(\hat{\mu}^2) = \left(1 - \frac{n}{N}\right) \left(\frac{1}{nM^2}\right) s^2 + \left(\frac{1}{nNM^2}\right) \sum_{i=1}^{n} M_i^2 \left(1 - \frac{m_i}{M_i}\right) \left(\frac{1}{m_i}\right) s_i^2
\]

where:

\[
s^2 = \frac{\sum_{i=1}^{n} (M_i y_i - M \hat{\mu}^2)^2}{n-1} \quad s_i^2 = \frac{\sum_{j=1}^{m_i} (y_{ij} - \bar{y}_i)^2}{m_i - 1}, \quad i = 1, 2, \ldots, n
\]

2.2. Ratio estimator for aboveground biomass: PSU level (hectare)

\[
\hat{\mu}_{ha}^2 = 10,000 \ast \hat{\mu}^2
\]

\[
\hat{\sigma}(\hat{\mu}_{ha}^2) = 10,000^2 \ast \hat{\sigma}(\hat{\mu}^2)
\]

2.3. Ratio estimator for belowground biomass: SSU level (plot)

\[
\hat{\mu}^2 \prime = \frac{\sum_{i=1}^{n} M_i^\prime y_i^\prime}{\sum_{i=1}^{n} M_i^\prime}
\]

\[
\hat{\sigma}(\hat{\mu}^2 \prime) = \left(1 - \frac{n}{N}\right) \left(\frac{1}{nM^2}\right) \sum_{i=1}^{n} \left(\frac{M_i y_i - M_i^\prime \hat{\mu}^2 \prime}{n-1}\right)^2 + \left(\frac{1}{nNM^2}\right) \sum_{i=1}^{n} M_i^2 \left(1 - \frac{m_i}{M_i^\prime}\right) \left(\frac{1}{m_i}\right) \sum_{j=1}^{m_i} (y_{ij} - \bar{y}_i)^2
\]

\[
\hat{\sigma}(\hat{\mu}^2 \prime) = \left(1 - \frac{n}{N}\right) \left(\frac{1}{nM^2}\right) s^2 \prime + \left(\frac{1}{nNM^2}\right) \sum_{i=1}^{n} M_i^2 \left(1 - \frac{m_i}{M_i^\prime}\right) \left(\frac{1}{m_i}\right)
\]

\[
s^2 \prime = \frac{\sum_{i=1}^{n} (M_i^\prime y_i^\prime - M_i^\prime \hat{\mu}^2 \prime)^2}{n-1} \quad s_i^2 \prime = \frac{\sum_{j=1}^{m_i} (y_{ij} - \bar{y}_i)^2}{m_i - 1}
\]

2.4. Ratio estimator for belowground biomass: PSU level (hectare)

\[
\hat{\mu}_{ha}^2 = 111,111 \ast \hat{\mu}^2 \prime
\]

\[
\hat{\sigma}(\hat{\mu}_{ha}^2) = 111,111^2 \ast \hat{\sigma}(\hat{\mu}^2 \prime)
\]

2.5. Ratio estimator for total biomass (above and belowground): PSU (hectare)

\[
\hat{\mu}_{ha}^{2L} = \hat{\mu}_{ha}^2 \ast \hat{\mu}_{ha}^2
\]

\[
\hat{\sigma}(\hat{\mu}_{ha}^{2L}) = \hat{\sigma}(\hat{\mu}_{ha}^2) + \hat{\sigma}(\hat{\mu}_{ha}^2) + 2c\sigma (\hat{\mu}_{ha}^2, \hat{\mu}_{ha}^2)
\]

\[
\hat{\sigma}(\hat{\mu}_{ha}^{2L}) = \hat{\sigma}(\hat{\mu}_{ha}^2) + \hat{\sigma}(\hat{\mu}_{ha}^2) + 2c\sigma (10,000 \ast \hat{\mu}^2, 111,111 \ast \hat{\mu}^2 \prime)
\]

\[
\hat{\sigma}(\hat{\mu}_{ha}^{2L}) = \hat{\sigma}(\hat{\mu}_{ha}^2) + \hat{\sigma}(\hat{\mu}_{ha}^2) \ast 2 \ast 10,000 \ast 111,111c\sigma (\hat{\mu}^2, \hat{\mu}^2 \prime)
\]

where:

\[
c\sigma (\hat{\mu}^2, \hat{\mu}^2 \prime) \equiv \left(1 - \frac{n}{N}\right) \left(\frac{1}{nM^2}\right) \sum_{i=1}^{n} (M_i y_i - M_i \hat{\mu}^2)(M_i y_i - M_i \hat{\mu}^2 \prime)
\]

\[
+ \left(\frac{1}{nNM^2}\right) \sum_{i=1}^{n} M_i M_i^\prime \left(1 - \frac{m_i}{M_i^\prime}\right) \left(\frac{1}{m_i}\right) \sum_{j=1}^{m_i} (y_{ij} - \bar{y}_i)(y_{ij} - \bar{y}_i)^\prime
\]
Method 3. Simple mean estimator and PSU of different sizes (two-stage sampling)

3.1. Mean estimator for aboveground biomass: (plot)

\[
\tilde{\mu}^3 = \frac{1}{nm_i} \sum_{i=1}^{n_m} \sum_{j=1}^{m_i} y_{ij} = \frac{1}{n} \sum_{i=1}^{n} \bar{y}_i
\]

where \(\tilde{\mu}^3\) identifies the estimator by method 3:

\[
\hat{\nu}_{\bar{y}} = \left(1 - \frac{n}{N}\right) \left(\frac{1}{nm}\right) \left(\frac{m}{n-1}\right) \sum_{i=1}^{n} \left(y_{ij} - \mu^3\right)^2 + \left(1 - \frac{m}{M}\right) \left(\frac{1}{Nm}\right) \sum_{i=1}^{n_m} \sum_{j=1}^{m_i} (y_{ij} - \bar{y}_i)^2
\]

\[
\hat{\nu}_{\mu^3} = \left(1 - \frac{n}{N}\right) \left(\frac{1}{nm}\right) MSB + \left(1 - \frac{m}{M}\right) \left(\frac{1}{Nm}\right) MSW
\]

\[
MSB = \left(\frac{m}{n-1}\right) \sum_{i=1}^{n} \left(y_{ij} - \mu^3\right)^2 y MSW = \left(\frac{1}{n(m_i - 1)}\right) \sum_{i=1}^{n_m} \sum_{j=1}^{m_i} (y_{ij} - \bar{y}_i)^2
\]

3.2. Mean estimator for aboveground biomass: PSU level (hectare)

\[
\tilde{\mu}^3_{ha} = 10,000 * \tilde{\mu}^3
\]

\[
\hat{\nu} (\tilde{\mu}^3_{ha}) = 10,000^2 * \hat{\nu} (\tilde{\mu}^3)
\]

3.3. Mean estimator for belowground biomass: SSU (plot)

\[
\tilde{\mu}^3' = \frac{1}{nm_i} \sum_{i=1}^{n_m} \sum_{j=1}^{m_i} y_{ij}' = \frac{1}{n} \sum_{i=1}^{n} \bar{y}_i'
\]

\[
\hat{\nu}_{\bar{y}}' = \left(1 - \frac{n}{N}\right) \left(\frac{1}{nm}\right) \left(\frac{m}{n-1}\right) \sum_{i=1}^{n} \left(y_{ij}' - \mu^3\right)^2 + \left(1 - \frac{m}{M}\right) \left(\frac{1}{Nm}\right) \sum_{i=1}^{n_m} \sum_{j=1}^{m_i} (y_{ij}' - \bar{y}_i')^2
\]

\[
\hat{\nu}_{\mu^3'} = \left(1 - \frac{n}{N}\right) \left(\frac{1}{nm}\right) MSB' + \left(1 - \frac{m}{M}\right) \left(\frac{1}{Nm}\right) MSW'
\]

\[
MSB' = \left(\frac{m}{n-1}\right) \sum_{i=1}^{n} \left(y_{ij}' - \mu^3\right)^2 y MSW' = \left(\frac{1}{n(m_i - 1)}\right) \sum_{i=1}^{n_m} \sum_{j=1}^{m_i} (y_{ij}' - \bar{y}_i')^2
\]

3.4. Mean estimator for belowground biomass: PSU (hectare)

\[
\tilde{\mu}^3_{ha}' = 111,111 * \tilde{\mu}^3'
\]

\[
\hat{\nu} (\tilde{\mu}^3_{ha}') = 111,111^2 * \hat{\nu} (\tilde{\mu}^3')
\]

3.5. Mean estimator for total biomass: PSU (hectare)

\[
\tilde{\mu}^{3l} = \tilde{\mu}^3_{ha} + \tilde{\mu}^3_{ha}'
\]

\[
\hat{\nu} (\tilde{\mu}^{3l}) = \hat{\nu} (\tilde{\mu}^3_{ha}) + \hat{\nu} (\tilde{\mu}^3_{ha}') + 2c\hat{\nu} (\tilde{\mu}^3_{ha}, \tilde{\mu}^3_{ha}')
\]

\[
\hat{\nu} (\tilde{\mu}^{3l}_{ha}) = \hat{\nu} (\tilde{\mu}^3_{ha}) + \hat{\nu} (\tilde{\mu}^3_{ha}') + 2\cdot 10,000 * 111,111 * \hat{\nu} (\tilde{\mu}^3_{ha}) + 2 * 10,000 * 111,111c\hat{\nu} (\tilde{\mu}^3_{ha}, \tilde{\mu}^3_{ha}')
\]
where:

\[ \sigma^2(\hat{\mu}^3, \hat{\mu}^3') \cong (1 - \frac{n}{N}) \left( \frac{1}{nM} \right) \sum_{i=1}^{m} (y_i - \hat{\mu}^3)(y_i' - \hat{\mu}^3') \]

\[ + \left( 1 - \frac{m}{M+M'} \right) \left( \frac{1}{N} \right) \left( \frac{1}{n(m-1)} \right) \sum_{i=1}^{n} \sum_{j=1}^{m} (y_{ij} - \bar{y}_i)(y_{ij}' - \bar{y}_i') \]

Method 4. Ratio estimator and PSU of different sizes (one-stage sampling)

4.1. Ratio estimator for aboveground biomass: PSU (hectare)

\[ \hat{\mu}_n^4 = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i} = \frac{\bar{y}}{\bar{x}} \]

where \( \hat{\mu}_n^4 \) identifies the estimator by method 4:

\[ \hat{\phi}(\hat{\mu}_n^4) = (1 - \frac{n}{N}) \left( \frac{1}{nN^2} \right) \frac{\sum_{i=1}^{n} (y_i - \hat{\mu}_n^4 x_i)^2}{n-1} \]

\[ \hat{\phi}(\hat{\mu}_n^4') = (1 - \frac{n}{N}) \left( \frac{1}{nN^2} \right) s_r^2 \]

where

\[ s_r^2 = \frac{\sum_{i=1}^{n} (y_i - \hat{\mu}_n^4 x_i)^2}{n-1} \]

4.2. Ratio estimator for belowground biomass: PSU (hectare)

\[ \hat{\mu}_n^4' = \frac{\sum_{i=1}^{n} y_i'}{\sum_{i=1}^{n} x_i'} = \frac{\bar{y}'}{\bar{x}'} \]

\[ \hat{\phi}(\hat{\mu}_n^4') = (1 - \frac{n}{N}) \left( \frac{1}{nN^2} \right) \frac{\sum_{i=1}^{n} (y_i' - \hat{\mu}_n^4' x_i')^2}{n-1} \]

\[ \hat{\phi}(\hat{\mu}_n^4') = (1 - \frac{n}{N}) \left( \frac{1}{nN^2} \right) s_r^2, \text{ donde } s_r^2 = \frac{\sum_{i=1}^{n} (y_i' - \hat{\mu}_n^4' x_i')^2}{n-1} \]

4.3. Ratio estimator for total biomass (above and belowground): PSU (hectare)

\[ \hat{\mu}_n^{4t} = \hat{\mu}_n^4 + \hat{\mu}_n^4' \]

\[ \hat{\phi}(\hat{\mu}_n^{4t}) = \hat{\phi}(\hat{\mu}_n^4) + \hat{\phi}(\hat{\mu}_n^4') + 2\sigma \phi(\hat{\mu}_n^4, \hat{\mu}_n^4') \]

where:

\[ c\sigma \phi(\hat{\mu}_n^4, \hat{\mu}_n^4') \cong (1 - \frac{n}{N}) \left( \frac{1}{nN^2} \right) \frac{\sum_{i=1}^{n} (y_i - \hat{\mu}_n^4 x_i)(y_i' - \hat{\mu}_n^4' x_i')}{n-1} \]

Method 5. VAN DEUSEN estimator and PSU of different sizes (one-stage sampling)

5.1. VAN DEUSEN estimator for aboveground biomass: PSU (hectare)

\[ \bar{\mu}_n^5 = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i} \]

where \( \bar{\mu}_n^5 \) identifies the estimator by method 5:

\[ \hat{\phi}(\hat{\mu}_n^5) = (1 - \frac{n}{N}) \left( \frac{1}{\sum_{i=1}^{n} x_i} \right) \frac{\sum_{i=1}^{n} (y_i - \hat{\mu}_n^5 x_i)^2}{\left( \sum_{i=1}^{n} x_i - x_0 \right)^2}, \text{ donde } x_0 = \left( \frac{1}{10,000} \right) \times 5 \]
\[ \hat{\rho}^5_{\text{ha}} = \left(1 - \frac{n}{N}\right) \left(\frac{1}{\sum_{i=1}^{n} x_i}\right) s^2_{r,\text{VAN}}, \text{ donde } s^2_{r,\text{VAN}} = \frac{\sum_{i=1}^{n} (y_i - \hat{\rho}^5_{\text{ha}} x_i)^2}{\left(\sum_{i=1}^{n} x_i - x_0\right)^2} \]

5.2. VAN DEUSEN estimator for belowground biomass: PSU (hectare)

\[ \hat{\rho}^5_{\text{ha}}' = \frac{\sum_{i=1}^{n} y_i'}{\sum_{i=1}^{n} x_i'} \]

\[ \hat{\sigma}^2 \left(\hat{\rho}^5_{\text{ha}}', \hat{\rho}^5_{\text{ha}}''\right) = \left(1 - \frac{n}{N}\right) \left(\frac{1}{\sum_{i=1}^{n} x_i'}\right) \left(\frac{\sum_{i=1}^{n} (y_i' - \hat{\rho}^5_{\text{ha}} x_i')^2}{\left(\sum_{i=1}^{n} x_i' - x_0'\right)^2}\right) \]

where:

\[ \hat{\sigma}^2 \left(\hat{\rho}^5_{\text{ha}}', \hat{\rho}^5_{\text{ha}}''\right) = \hat{\sigma}^2 \left(\hat{\rho}^5_{\text{ha}}''\right) + 2 \hat{\sigma} \left(\hat{\rho}^5_{\text{ha}}', \hat{\rho}^5_{\text{ha}}''\right) \]

5.3. VAN DEUSEN estimator for total biomass (above and belowground): PSU level (hectare)

\[ \hat{\rho}^5_{\text{st}} = \hat{\rho}^5_{\text{ha}} + \hat{\rho}^5_{\text{ha}}' \]

\[ \hat{\sigma} \left(\hat{\rho}^5_{\text{st}}\right) = \hat{\sigma} \left(\hat{\rho}^5_{\text{ha}}\right) + \hat{\sigma} \left(\hat{\rho}^5_{\text{ha}}'\right) + 2 \hat{\sigma}_V \left(\hat{\rho}^5_{\text{ha}}, \hat{\rho}^5_{\text{ha}}'\right) \]

Method 6. Mean ratio estimator and PSU of different sizes (one-stage sampling)

6.1. Mean ratio estimator for aboveground biomass: PSU level (hectare)

\[ \hat{\rho}^6_{\text{ha}} = \frac{1}{n} \sum_{i=1}^{n} \frac{y_i}{x_i} = \frac{1}{n} \sum_{i=1}^{n} \frac{y_i^*}{x_i} \]

where \( \hat{\rho}^6 \) identifies the estimator by method 6:

\[ \hat{\sigma} \left(\hat{\rho}^6_{\text{ha}}\right) = \left(1 - \frac{n}{N}\right) \left(\frac{1}{n}\right) \left(\frac{\sum_{i=1}^{n} (y_i^* - \hat{\rho}^6_{\text{ha}})^2}{n - 1}\right) \]

\[ \hat{\sigma} \left(\hat{\rho}^6_{\text{ha}}'\right) = \left(1 - \frac{n}{N}\right) \left(\frac{1}{n}\right) s^2_{m}, \text{ donde } s^2_{m} = \frac{\sum_{i=1}^{n} (y_i^* - \hat{\rho}^6_{\text{ha}})^2}{n - 1} \]

6.2. Mean ratio estimator for belowground biomass: PSU (hectare)

\[ \hat{\rho}^6_{\text{ha}}' = \frac{1}{n} \sum_{i=1}^{n} \frac{y_i'}{x_i'} = \frac{1}{n} \sum_{i=1}^{n} \frac{y_i'^*}{x_i'^*} \]

\[ \hat{\sigma} \left(\hat{\rho}^6_{\text{ha}}'\right) = \left(1 - \frac{n}{N}\right) \left(\frac{1}{n}\right) \left(\frac{\sum_{i=1}^{n} (y_i' - \hat{\rho}^6_{\text{ha}}')^2}{n - 1}\right) \]

\[ \hat{\sigma} \left(\hat{\rho}^6_{\text{ha}}''\right) = \left(1 - \frac{n}{N}\right) \left(\frac{1}{n}\right) s^2_{m'}, \text{ donde } s^2_{m'} = \frac{\sum_{i=1}^{n} (y_i'^* - \hat{\rho}^6_{\text{ha}}')^2}{n - 1} \]

6.3. Mean ratio estimator for total biomass: PSU level (hectare)

\[ \hat{\rho}^{6t}_{\text{ha}} = \hat{\rho}^6_{\text{ha}} + \hat{\rho}^6_{\text{ha}}' \]

\[ \hat{\sigma} \left(\hat{\rho}^{6t}_{\text{ha}}\right) = \hat{\sigma} \left(\hat{\rho}^6_{\text{ha}}\right) + \hat{\sigma} \left(\hat{\rho}^6_{\text{ha}}'\right) + 2 \hat{\sigma}_V \left(\hat{\rho}^6_{\text{ha}}, \hat{\rho}^6_{\text{ha}}'\right) \]
\[ c_{Ov}^\prime (\hat{\mu}_{ha}^\prime) = \left(1 - \frac{n}{N}\right) \left(\frac{1}{n}\right) \frac{\sum_{i=1}^{n} (y_i^* - \hat{\mu}^\prime_{ha}) (y_i^* - \hat{\mu}^\prime_{ha})}{n - 1} \]

Appendix B. Estimator’s Equivalence for Aboveground Biomass

\[ \hat{\mu}^1 = \frac{1}{\bar{m}} \sum_{i=1}^{n} M_i \bar{y}_i \implies \frac{1}{\bar{m}} \sum_{i=1}^{n} M_i \bar{y}_i = \frac{1}{\bar{m'}} \sum_{i=1}^{n} M_i \bar{y}_i = \frac{1}{\bar{m}} \sum_{i=1}^{n} M_i \bar{y}_i = \hat{\mu}^2 \]

Since

\[ \bar{m} = \frac{\sum_{i=1}^{n} M_i}{n} \]

therefore:

\[ \hat{\mu}^1_{ha} = \hat{\mu}^2_{ha} = 10,000 \frac{\sum_{i=1}^{n} M_i \bar{y}_i}{\sum_{i=1}^{n} M_i} \]

\[ \hat{\mu}^2 = \frac{\sum_{i=1}^{n} M_i \bar{y}_i}{\sum_{i=1}^{n} M_i} = \frac{\sum_{i=1}^{n} (m_i / 5) 10,000 \bar{y}_i}{\sum_{i=1}^{n} m_i} = \frac{\sum_{i=1}^{n} m_i \bar{y}_i}{\sum_{i=1}^{n} m_i} \]

Since

\[ \bar{y}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} y_{ij} \implies \bar{y}_i m_i = \sum_{j=1}^{m_i} y_{ij} = y_i \]

thus:

\[ \hat{\mu}^2_{ha} = 10,000 \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} m_i} \]

Now:

\[ \hat{\mu}^4_{ha} = \hat{\mu}^5_{ha} = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i} = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} \left(\frac{1}{10,000}\right) \cdot m_i} = 10,000 \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} m_i} = \hat{\mu}^2_{ha} \]

Since

\[ x_i = \left(\frac{1}{10,000}\right) \cdot m_i \]

on the other hand,

\[ \hat{\mu}^3 = \frac{1}{n m_i} \sum_{i=1}^{n} \sum_{j=1}^{m_i} y_{ij} = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{m_i} \sum_{j=1}^{m_i} y_{ij} = \frac{1}{n} \sum_{i=1}^{n} \frac{y_i}{m_i} \]

therefore:

\[ \hat{\mu}^3_{ha} = 10,000 \frac{1}{n} \sum_{i=1}^{n} \frac{y_i}{m_i} \]

Now:

\[ \hat{\mu}^6_{ha} = \frac{1}{n} \sum_{i=1}^{n} \frac{y_i}{x_i} = \frac{1}{n} \sum_{i=1}^{n} \frac{y_i}{\left(\frac{1}{10,000}\right) \cdot m_i} = 10,000 \frac{1}{n} \sum_{i=1}^{n} \frac{y_i}{m_i} = \hat{\mu}^3_{ha} \]

Appendix C. Estimator’s Equivalence for Belowground Biomass

\[ \hat{\mu}^1 = \frac{1}{\bar{M}} \sum_{i=1}^{n} M_i \bar{y}_i \implies \frac{1}{\bar{M}} \sum_{i=1}^{n} M_i \bar{y}_i = \frac{1}{\bar{M'}} \sum_{i=1}^{n} M_i \bar{y}_i = \frac{1}{\bar{M}} \sum_{i=1}^{n} M_i \bar{y}_i = \hat{\mu}^2 \]

Since

\[ \bar{m} = \frac{\sum_{i=1}^{n} M_i}{n} \]

(A1)
therefore:
\[ \hat{\mu}^1_{ha} = \hat{\mu}^2_{ha} = 111,111 \frac{\sum^n_{i=1} M_i \bar{y}_i}{\sum^n_{i=1} M_i} \]
\[ \hat{\mu}^2 = \frac{\sum^n_{i=1} M_i \bar{y}_i}{\sum^n_{i=1} M_i} = \frac{\sum^n_{i=1} (m_i / 5) 111,111 \bar{y}_i}{\sum^n_{i=1} m_i} = \frac{\sum^n_{i=1} m_i \bar{y}_i}{\sum^n_{i=1} m_i} = \frac{\sum^n_{i=1} y_i}{\sum^n_{i=1} m_i} \]

Given that
\[ \bar{y}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} y_{ij} \rightarrow \bar{y}_i m_i = \sum_{j=1}^{m_i} y_{ij} = y_i \]

thus:
\[ \hat{\mu}^2_{ha} = 111,111 \frac{\sum^n_{i=1} y_i}{\sum^n_{i=1} m_i} \]

Now:
\[ \hat{\mu}^4 = \hat{\mu}^5 = \frac{\sum^n_{i=1} y_i}{\sum^n_{i=1} x_i} = \frac{\sum^n_{i=1} y_i}{\sum^n_{i=1} \left( \frac{0.09}{10,000} \right) m_i} = 111,111 \frac{\sum^n_{i=1} y_i}{\sum^n_{i=1} m_i} = \hat{\mu}^2_{ha} \]

Since
\[ x_i = \left( \frac{0.09}{10,000} \right) m_i \]
on the other hand,
\[ \hat{\mu}^3 = \frac{1}{n M_i} \sum_{i=1}^{n} \sum_{j=1}^{m_i} y_{ij} = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{m_i} \sum_{j=1}^{m_i} y_{ij} = \frac{1}{n} \sum_{i=1}^{n} \frac{y_i}{m_i} \]

therefore:
\[ \hat{\mu}^3_{ha} = 111,111 \frac{1}{n} \sum_{i=1}^{n} \frac{y_i}{m_i} \]

Now:
\[ \hat{\mu}^6_{ha} = \frac{1}{n} \sum_{i=1}^{n} \frac{y_i}{x_i} = \frac{1}{n} \sum_{i=1}^{n} \frac{y_i}{\left( \frac{0.09}{10,000} \right) m_i} = 111,111 \frac{1}{n} \sum_{i=1}^{n} \frac{y_i}{m_i} = \hat{\mu}^3_{ha} \]

References
1. Arieira, J.; Karssenberg, D.; de Jong, S.M.; Addink, E.A.; Couto, E.G.; da Cunha, C.N.; Skøien, J.O. Integrating field sampling, geostatistics and remote sensing to map wetland vegetation in the Pantanal, Brazil. *Böögescience* 2011, 8, 667–686. [CrossRef]
2. Barabesi, L.; Pisani, C. Steady-state ranked set sampling for replicated environmental sampling designs. *Environmetrics* 2004, 15, 45–56. [CrossRef]
3. Magnussen, S.; Næsset, E.; Gobakken, T. An Estimator of Variance for Two-Stage Ratio Regression Estimators. *For. Sci.* 2014, 60, 663–676. [CrossRef]
4. Mandallaz, D.; Massey, A. Comparison of estimators in one-phase two-stage Poisson sampling in forest inventories. *Can. J. For. Res.* 2012, 42, 1865–1871. [CrossRef]
5. McRoberts, R.E.; Walters, B.F. Statistical inference for remote sensing-based estimates of net deforestation. *Remote Sens. Environ.* 2012, 124, 394–401. [CrossRef]
6. Breidt, F.J.; Opsomer, J.D. Model-Assisted Survey Estimation with Modern Prediction Techniques. *Stat. Sci.* 2017, 32, 190–205. [CrossRef]
7. Stähli, G.; Saarela, S.; Schnell, S.; Holm, S.; Breidenbach, J.; Healey, S.P.; Patterson, P.L.; Magnussen, S.; Næsset, E.; McRoberts, R.E.; et al. Use of models in large-area forest surveys: Comparing model-assisted, model-based and hybrid estimation. *For. Ecosyst.* 2016, 3, 5. [CrossRef]
8. Gregoire, T.G.; Valentine, H.T. *Sampling Strategies for Natural Resources and Environment*; Chapman and Hall/CRC: Boca Raton, FL, USA, 2008; ISBN 9781584883708.
9. Sing, A.S.; Masuku, M.B. Sampling Techniques and Determination of sample size in applied statistics research: An overview. *Int. J. Econ. Commn. Manag.* 2014, 2, 32–33.
10. Henttonen, H.M.; Kangas, A. Optimal plot design in a multipurpose forest inventory. *For. Ecosyst.* 2015, 2, 31. [CrossRef]
11. CONAFOR National System of Forest Information. National Forest and Soils Inventory; CONAFOR: Zapopan, Mexico, 2012.
12. Zarnoch, S.J.; Bechtold, W.A. Estimating mapped-plot forest attributes with ratios of means. Can. J. For. Res. 2000, 30, 688–697. [CrossRef]
13. Velasco-Bautista, E.; de los Santos-Posadas, H.M.; Ramirez-Maldonado, H.; Rendon-Sanchez, G. Comparación de estimadores para volumen maderable en selva mediana del sureste de México. Agrociencia 2016, 50, 119–132.
14. Rzedowski, J. Vegetacion de Mexico, 1st ed.; Comision Nacional para el Uso y Conocimiento de la Biodiversidad: Tlalpan, Mexico, 2006.
15. Maltamo, M. Inventories of Vegetation, Wild Berries and Mushrooms. In Forest Inventory: Methodology and Applications; Kangas, A., Maltamo, M., Eds.; Springer: Dordrecht, The Netherlands, 2006; pp. 141–153. ISBN 978-1-4020-4381-9.
16. Schreuder, H.; Ernst, R.; Ramirez, H. Statistical Techniques for Sampling and Monitoring Natural Resources; U.S. Department of Agriculture Forest Service: Fort Collins, CO, USA, 2004.
17. McRoberts, R.; Tomppo, E.; Schadauer, K.; Vidal, C.; Stahl, G.; Chirici, G.; Lanz, A.; Cienciala, E.; Winter, S.; Smith, B. Harmonizing National Forest Inventories. J. For. 2012, 58, 189–190. [CrossRef]
18. Grafström, A.; Saarela, S.; Ene, L.T. Efficient sampling strategies for forest inventories by spreading the sample in auxiliary space. Can. J. For. Res. 2014, 44, 1156–1164. [CrossRef]
19. Bautista, E.V.; Posadas, H.D.L.S.; Maldonado, H.R.; Sánchez, G.R.; Lazalde, J.R.V.; Mireles, M.A. Estimadores para atributos forestales en una población simulada de selva mediana del sur de México. Bosque 2020, 41, 307–320. [CrossRef]
20. CONAFOR Inventario Nacional Forestal y de Suelos. Informe de Resultados 2009–2014; CONAFOR: Zapopan, Mexico, 2018.
21. Lohr, S. Sampling: Design and Analysis, 2nd ed.; Brooks/Cole: Boston, MA, USA, 2010.
22. Scheaffer, R.; Mendenhall, W.; Lyman, R.; Gerow, K. Elementary Survey Sampling, 7th ed.; Brooks/Cole: Boston, MA, USA, 2012.
23. Van Deusen, P. Comparing Mapped Plot Estimators. In Proceedings of the Sixth Annual Forest Inventory and Analysis Symposium, Denver, CO, USA, 21–24 September 2004; McRoberts, R., Reams, G., Van Deusen, P., McWilliams, W., Eds.; U.S. Department of Agriculture Forest Service: Fort Collins, CO, USA, 2005; pp. 93–97.
24. Šmelko, Š.; Merganiˇc, J. Some methodological aspects of the National Forest Inventory and Monitoring in Slovakia. J. For. Sci. 2008, 54, 476–483. [CrossRef]
25. SAS Institute Inc. SAS/STAT User’s Guide; SAS Institute Inc.: Cary, NC, USA, 2017.
26. Kangas, A. Design-Based Sampling and Interference. In Forest Inventory: Methodology and Applications; Kangas, A., Maltamo, M., Eds.; Springer: Dordrecht, The Netherlands, 2006; pp. 13–38.
27. Van Deuse, P.C.; Roesch, F.A. Application of mapped plots for single-owner forest surveys. J. For. 2009, 107, 414–418. [CrossRef]
28. Fierros-Mateo, R.; Santos-Posadas, H.M.D.L.; Fierros-González, A.M.; Cruz-Cobos, F.; Angel, L.M.; Bautista, E.V.; Barraza, G.Q. Estimadores de muestreo para inventario de plantaciones de Pinus chiapensis (Martínez) Andresen. Rev. Mex. Cien. For. 2018, 9, 131–150. [CrossRef]
29. Lynch, T.; Hamlin, D.; Ducey, M.; Borders, B. Design-Unbiased Estimation and Alternatives for Sampling Trees on Plantation Rows. Forests 2018, 9, 362. [CrossRef]
30. Alvarez, I.A.; Del Nero Velasco, G.; Barbin, H.S.; Lima, A.M.L.P.; do Couto, H.T.Z. Comparison of two sampling methods for estimating urban tree density. J. Arboric. 2005, 31, 209–213. [CrossRef]
31. Woodall, C.W.; Amacher, M.C.; Bechtold, W.A.; Coulston, J.W.; Jovan, S.; Perry, C.H.; Randolph, K.C.; Schulz, B.K.; Smith, G.C.; Tkacz, B.; et al. Status and future of the forest health indicators program of the USA. Environ. Monit. Assess. 2011, 177, 419–436. [CrossRef]
32. Bechtold, W.; Randolph, K. FIA Crown Analysis Guide. Version 2.0; United States Department of Agriculture: Washington, DC, USA, 2018.
33. Zhu, S.; Zhang, H.; Liu, R.; Cao, Y.; Zhang, G. Comparison of Sampling Designs for Estimating Deforestation from Landsat TM and MODIS Imagery: A Case Study in Mato Grosso, Brazil. Sci. World J. 2014, 2014, 1–10. [CrossRef]
34. Köhl, M.; Lister, A.; Scott, C.T.; Baldauf, T.; Pluge, D. Implications of sampling design and sample size for national carbon accounting systems. Carbon Balance Manag. 2011, 6, 10. [CrossRef]