



Estrategia de muestreo óptima para la planeación del inventario maderable en plantaciones comerciales de *Tectona grandis* L.f.

Optimal sampling strategy for timber inventory planning in commercial plantations of *Tectona grandis* L.f.

Juan Carlos Tamarit-Urias^{1*}, Héctor Manuel de los Santos-Posadas^{2*}, Arnulfo Aldrete², José René Valdez-Lazalde², Hugo Ramírez-Maldonado³, Vidal Guerra-de la Cruz⁴

Resumen

El estado de Quintana Roo ocupa el segundo lugar nacional en producción maderable de especies latifoliadas; de ellas, *Bursera simaruba* y *Metopium brownei* son de relevancia forestal debido al valor de su madera y amplia distribución en la entidad. Para elaborar programas de manejo forestal, las funciones que estiman el volumen fustal y comercial, además del ahusamiento y la razón de volumen son importantes. El objetivo del presente estudio fue ajustar una función no-lineal que estime el volumen comercial para dos taxa en el centro y sur de Quintana Roo, México. Con información dasométrica de 188 y 133 árboles, respectivamente que abarcaron todas las categorías diamétricas observadas en estos ecosistemas forestales; se ajustó un modelo de volumen comercial con efectos aleatorios a nivel de árbol, y se eligieron las mejores combinaciones de inclusión de los efectos aleatorios según la máxima verosimilitud. Las ecuaciones propuestas tienen un sesgo, en promedio, al estimar el volumen comercial para los dos taxones de 0.0045 m³ y una explicación mayor a 90 % de la variabilidad muestral. Las ecuaciones resultantes podrán emplearse en la estimación maderable en los programas de manejo forestal sustentable de los bosques tropicales ubicados en el centro y sur del estado.

Palabras clave: Bosque tropical, distribución de productos, manejo forestal, sistema de cubicación, Quintana Roo, volumen.

Abstract

The state of Quintana Roo occupies the second national place in timber production of broadleaf species, of which *Bursera simaruba* and *Metopium brownei* are relevant, from the value of their wood and wide distribution. The functions that estimate the stem and commercial volume, in addition to the taper and the volume ratio are important in the elaboration of a forest management program. The aim of this study was to fit a nonlinear function that estimates trade volume for two species in central and southern Quintana Roo. A commercial volume model with random effects was fitted at the tree level, using the mensuration data from 188 and 133 trees for each species, respectively, and all the diameter categories found in these forest ecosystems were covered. The best combinations of inclusion of the random effects were chosen according to the maximum likelihood. The proposed equations have an average bias when estimating the commercial volume for the two species of 0.0045 m³ and an explanation greater than 90 % of the sample variability. The resulting equations may be used in the timber estimation of the sustainable forest management programs of tropical forests at the central and southern territories of the state.

Key words: Tropical forest, product distribution, forest management, cubing system, Quintana Roo, volume.

Fecha de recepción/Reception date: 29 de enero de 2021

Fecha de aceptación/Acceptance date: 30 de septiembre de 2021

1Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias. Campo Experimental San Martinito, CIR Golfo Centro. México.

2Posgrado en Ciencias Forestales, Colegio de Postgraduados. México.

3División de Ciencias Forestales, Universidad Autónoma Chapingo. México.

41Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias. Sitio Experimental Tlaxcala, CIR Centro. México.

*Autor para correspondencia; correo-e: hectorm.delossantos@gmail.com

Introduction

Sampling and inventory methods with statistical validity are very useful for generating reliable and scientifically defensible estimates (Schreuder *et al.*, 2004; Fattorini *et al.*, 2015). In a commercial forest plantation (CFP) it is important to accurately estimate timber stocks, which allows having information to plan actions and make informed decisions about technical management and investment (Roldán *et al.*, 2014). When the CFP area is extensive, the inventory is carried out by sampling to obtain parameters of interest from the population in a correct, precise way and at minimal cost (van Laar and Akça, 2007; Köhl and Magnussen, 2016).

The aim of the timber forest inventory is to provide reliable quantitative information for the operational management of the CFP, for which accurate knowledge is demanded at the minimum management unit level on means and totals of the number of trees (N), basimetric area (BA) and volume (V) per hectare. The foregoing makes it necessary to seek a balance between the available financial resources and the required statistical precision (Roldán *et al.*, 2014). In this sense, the best approach to optimize resources is to design a sampling strategy, which must combine the method to select the sample with the procedure to estimate the population parameters (Gregoire and Valentine, 2008; Grafström *et al.*, 2014).

In order to carry out a multipurpose inventory, the single, simple and flexible sampling scheme is normally adopted, as it favors the assessment of all the mensuration variables of interest (Schreuder *et al.*, 1993; Corona and Fattorini, 2006). Whatever the sampling design and sample size, there are implications in terms of cost and efficiency in the reliability of estimating inventory, which can be quantified based on precision, bias, and mean square error (Köhl *et al.*, 2011; Roldán *et al.*, 2014). Precision is a function of the homogeneity of the population and the number of sampling sites surveyed, so the inventory must be carried out to meet the predetermined level of precision and to optimize both, the time and the

costs, to carry it out (Schreuder *et al.*, 2004; van Laar and Akça, 2007; Marchi *et al.*, 2017).

The specialized literature suggests different sampling estimators to make a CFP timber inventory oriented to technical-operational management. Among them, those that make the inference about the population based on the design are simple random sampling (MSA), systematic sampling and stratified sampling (ME), as well as those in which the inference is made based on to models (Raj, 1980; Scheaffer *et al.*, 1987; Cochran, 1993; Gregoire and Valentine, 2008). The latter use auxiliary information through variables that are highly and positively correlated with the variable of interest; among these types, the ratio and regression estimators that can be applied in MSA and ME stand out (Schreuder *et al.*, 1993; Shiver and Borders, 1996; Pérez, 2005; Grafström *et al.*, 2014). However, there is little recent empirical evidence derived from analysis of specific practical cases that allows determining the best sampling estimators or their possible combinations that lead to the design of an optimal sampling strategy.

In this context, the objectives of this study were: 1) To evaluate the statistical efficiency of six sampling estimators to propose an optimal sampling strategy in terms of precision and time that allows carrying out operational timber inventories that support decision-making aimed at improving the CFP technical handling of *Tectona grandis* L.f. (Teak) established in the state of *Campeche*, Mexico; and 2) To determine the optimal sample size for the mean volume in MSA and ME that ensures a 2.5 % precision and $\alpha = 0.05$.



Materials and Methods

The study was carried out in a commercial forest plantation with an area of 2 207.5 hectares planted with *Tectona grandis* at a 4 m × 2 m spacing (1 250 plants ha⁻¹), which is located in the *Edzná Valley, Campeche* municipality, state of the same name, in southeastern Mexico. The prevailing climate is Aw''0(i')g, which corresponds to the warm subhumid type with rains in summer, average annual precipitation of 1 094.7 mm with six months of drought from December to May; 26.6 °C average annual temperature, with prevailing winds in winter and summer, with maximum gusts of up to 60 km h⁻¹ (Breña, 2004).

A sample size (*n*) was made up of 8 830 rectangular 72 m² sites, each of which represented a coverage of 0.25 hectares. The sampling frame (*N*) was 306 597 units and the sampling intensity was 2.9%. The sites were raised under a systematic pattern on a 50 m × 50 m mesh with which there a uniform distribution was obtained.

At each site, nine strains were evaluated, each one referring to the specific point where a specimen of *Tectona grandis* was planted; in this way, it was possible to count the number of living trees, each of which had their normal diameter (*Dn*) measured with a 283D / 5m-CSE diametric tape. In addition, the age (*E*) of the plantation was recorded in years. The total height (*A*) of each tree in meters was estimated based on *Dn* with a Chapman-Richards model used by Tamarit (2013) that presented the following mathematical structure:



$$A = 19.46016 (1 - \exp(-0.065226 Dn))^{1.100977}$$

Where:

exp = Exponential function

The stem volume (*v*) with bark per tree in m³ was estimated with the Schumacher-Hall model referred by Tamarit *et al.* (2014) expressed as:

$$v = 0.000043 Dn^{1.857931} A^{1.041967}$$

Based on Fierros *et al.* (2018), the *BA* of each tree in m² was estimated with the formula:

$$BA = \frac{\pi}{40,000} D_n^2$$

Where:

π = Constant with a value of 3.1416

The number of living trees, the basimetric area and the total volume estimated by site were projected at the hectare level, in order to obtain the *N*, *BA* and *V* variables scaled to conventional units.

To estimate the mean wood volume, the estimators and parameters of the *MSA* and the *ME* were applied, in which the inference about the population is based on the design (Cochran, 1993; Gregoire and Valentine, 2008). The ratio (*R*) and regression (*Rg*) estimators were also assessed within *MSA*, whose inference is based on

models, in addition to the combined ratio (R_c) and the specific ratio (R_e) within the ME (Shiver and Borders, 1996; Pérez, 2005) (expressions 1 to 14 of Table 1).

Table 1 Sampling estimators and parameters evaluated to determine the timber inventory in a commercial Teak forest plantation.

Estimator	Parameter	Equation
<i>MSA</i>	Mean	$\bar{y} = \frac{\sum_{i=1}^n y_i}{n} \quad (1)$
	Variance of the mean	$S_{\bar{y}}^2 = \frac{S_y^2}{N} f \quad (2)$
	Sample size	$n = \frac{t_{\alpha,gl}^2 NS_{\bar{y}}^2}{t_{\alpha,gl}^2 S_{\bar{y}}^2 + B^2 N} \quad (3)$
<i>ME</i>	Mean	$\bar{y}_{st} = \frac{\sum_{h=1}^L N_h \bar{y}_h}{N} = \sum_{h=1}^L W_h \bar{y}_h \quad (4)$
	Variance of the mean	$V(\bar{y}_{st}) = \sum_{h=1}^L W_h^2 \frac{S_h^2}{n_h} (1 - f_h) \quad (5)$
	Sample size	$n = \frac{t_{\alpha,gl}^2 \sum_{h=1}^L \frac{N_h^2 S_h^2}{W_h}}{B^2 N} \quad (6)$
Ratio under <i>MSA</i>	Mean	$\bar{y}_R = \hat{R} \mu_x = \bar{y} \frac{\mu_x}{\bar{x}} \quad (7)$
	Variance of the mean	$S_{\bar{y}_R}^2 = \frac{S_u^2}{n} f \quad (8)$
Regression under <i>MSA</i>	Mean	$\bar{y}_{RG} = \bar{y} + b(\mu_x + \bar{x}) \quad (9)$
	Variance of the mean	$S_{\bar{y}_{RG}}^2 = \frac{S_{xy}^2}{n} f \quad (10)$ $S_{xy}^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2 - b^2 \sum_{i=1}^n (x_i - \bar{x})^2}{n - 2}$

Combined ratio in <i>ME</i>	Mean	$\bar{y}_{\hat{R}_C} = \hat{R}_C \mu_x = \frac{\sum_{h=1}^L W_h \bar{y}_h}{\sum_{h=1}^L W_h \bar{x}_h} \mu_x \quad (11)$
	Variance of the mean	$S_{\bar{y}_{RC}}^2 = \sum_{h=1}^L \frac{W_h^2 (1 - f_h)}{n_h (n_h - 1)} \left(\sum_{i=1}^{n_h} (y_{hi} - \hat{R}_C x_{hi})^2 \right) \quad (12)$
Specific ratio in <i>ME</i>	Mean	$\bar{y}_{\hat{R}_S} = \sum_{h=1}^L W_h \hat{R}_h \mu_{x_h} \quad (13)$
	Variance of the mean	$S_{\bar{y}_{RS}}^2 = \sum_{h=1}^L \frac{W_h^2 (1 - f_h)}{n_h (n_h - 1)} \left(\sum_{i=1}^{n_h} (y_{hi} - \hat{R}_h x_{hi})^2 \right) \quad (14)$

y_i = Volume with bark observed at the i -th sampling site and extrapolated to hectare ($m^3 \text{ ha}^{-1}$); f = Correction factor for finitude $1-n/N$; t = Student's t-distribution value at 95 % reliability ($1-\alpha = 0.95$) and with $n-1$ degrees of freedom (g); B = Magnitude of the sampling error acceptable at the specified $1-\alpha$ confidence level and the result of multiplying the required precision (E_μ) by the sample mean; L = Total number of strata in the target population; N_h = Sampling frame in the h -th stratum, with $N = \sum_{h=1}^L N_h$; n_h = Sample size in the h -th stratum, with $n = \sum_{h=1}^L n_h$. The $W_h = N_h/N$ component corresponds weighting by stratum; $f_h = n_h/N_h$ = Sampling

fraction in the h -th stratum; $S_y^2 = \frac{\sum_{i=1}^n y_i^2 - \frac{(\sum_{i=1}^n y_i)^2}{n}}{n-1}$ = estimated variance of the observations of the main variable; $\hat{R} = \frac{\bar{y}}{\bar{x}} = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n x_i}$ = Ratio under *MSA*, x_i represents the auxiliary variable; $S_{\hat{R}}^2 = \frac{1}{\mu_x^2} \frac{S_u^2}{n} f$ = Variance of the ratio under *MSA*, with $S_u^2 = \frac{\sum_{i=1}^n y_i^2 + \hat{R}^2 \sum_{i=1}^n x_i^2 - 2\hat{R} \sum_{i=1}^n y_i x_i}{n-1}$; $b = \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2}$ = Slope of the regression straight line under *MSA*; $\hat{R}_C = \frac{\bar{y}_{st}}{\bar{x}_{st}}$ = Combined ratio under *ME*; $\hat{R}_h = \frac{\bar{y}_h}{\bar{x}_h}$ = estimated ratio in the h -th stratum; $\bar{y}_{\hat{R}_h} = \hat{R}_h \mu_{x_h}$ = Mean by stratum; μ_y = Population mean of the main variable; μ_x = Population mean of the auxiliary variable; $B_{\hat{R}}$, B_{RG} = Acceptable magnitude of the sampling error for Ratio and Regression, respectively; \bar{y}_h , \bar{x}_h = Means of the totals in the h -th stratum.

In all cases, the *MSA* was used as the reference sampling design. To stratify, the age of the plantation was used as an auxiliary variable. Six age classes (*CE*) were defined with one-year intervals (Table 2).

Table 2. Age classes for the stratification, sites and respective surface in a Teak commercial forest plantation.

<i>CE</i>	<i>LI</i>	<i>LS</i>	<i>EP</i>	<i>n</i>	Surface (ha)
3	2	2.99	2.92	100	25.00
4	3	3.99	3.62	897	224.25
5	4	4.99	4.37	3 436	859.00
6	5	5.99	5.46	2 503	625.75
7	6	6.99	6.58	874	218.50
8	7	7.99	7.21	1 020	255.00
Total				8 830	2 207.5

CE = Age class; *LI* = Lower limit; *LS* = Upper limit; *EP* = Average age; *n* = Sample size.

In each stratum of the *ME*, the same estimators referred to for the *MSA* were applied. In the Ratio and Regression estimators, the *BA* and the age of the plantation were used as auxiliary variables, the value of the correlation coefficient (ρ) between the *V* with the *BA* and age was determined, the auxiliary variable is better when ρ is close to 1. In the estimator of *R* under *MSA*, the sample mean of the *BA* obtained under *ME* was assumed as the true population value. With the

estimator of R in ME , the age was known without sampling error because its exact record was present.

The way in which the confidence interval (IC) was calculated for the mean (\bar{y}) of the volume in $\text{m}^3 \text{ ha}^{-1}$, its respective precision (P) in percent and the corresponding inventory (\hat{y}) for the MSA is shown below (Cochran, 1993) and is applicable to the rest of the evaluated estimators.

$$IC = \bar{y} \pm t_{\alpha,gl} \sqrt{S_{\bar{y}}^2} \quad (15)$$

$$P = t_{\alpha,gl} \frac{\sqrt{S_{\bar{y}}^2}}{\bar{y}} 100 \quad (16)$$

$$\hat{y} = N\bar{y} \quad (17)$$

According to Freese (1976), Schreuder *et al.* (1993) and Schreuder *et al.* (2004), the theoretical approach referred to estimate the mean has greater meaning when an IC with a previously established level of reliability is established (in this study $\alpha = 0.05$); this means that under repeated sampling, on average 1 out of every 20 samples has the probability of producing a IC that does not contain the true mean of the volume, so to achieve this level of reliability on average two standard errors of the estimator are needed.

The best sampling estimator was selected based on the lowest value of the variance of the mean, the highest precision, and the lowest width of the confidence intervals. The MSA estimators were taken as the reference baseline against which the rest of the estimators were compared (West, 2017). The sampling strategy was formed by combining the sampling design with the best estimator. The calculations to apply the estimators were carried out in the Microsoft Office[®] Excel 2013 spreadsheet.

The optimal sample size for the mean volume in MSA and ME was determined by simulating a repeated sampling using the Bootstrap technique with the free

statistical program *R* version 3.6.2 (<https://www.r-project.org>, R Development Core Team, 2020) for Windows. The total sample (8 830 sites = 100 %) was gradually reduced in 5 % intervals; at each level of reduction 30 samples were obtained without replacement (Marchi *et al.*, 2017) to which the mean, the variance of the mean, the *IC* and the precision, later the average of each parameter was obtained to analyze their graphical behaviors. The optimal sample size was determined with a 2.5 % precision ($\alpha = 0.05$) and it was compared with that calculated by equation 3. For the *ME*, the sample size distribution in each stratum was made according to the respective formulations reported by Cochran (1993) and Pérez (2005) for proportional assignments and Neyman.

Results and Discussion

All the evaluated estimators had accuracies below 2 %. The Specific Ratio estimator in *ME* had the best statistical efficiency followed by the Regression estimator in *MSA* and the combined Ratio in *ME*. These three estimators had the best precisions, the lowest variances, narrow *ICs*, and used *BA* as an auxiliary variable (Table 3).

Table 3. Parameters per sampling estimator assessed in a commercial Teak forest plantation.

Estimator	Parameter					Inventory (m ³)	GP (%)	A
	Mean	VM	LI	LS	P (%)			
<i>MSA</i>	29.98	0.0609	29.49	30.46	1.61	66177	-	0.97
<i>ME</i>	29.98	0.0432	29.56	30.39	1.39	66177	0.23	0.83
<i>R V/BA MSA</i>	29.98	0.0024	29.96	29.99	0.33	66177	1.29	0.03
<i>R V/E MSA</i>	33.32	0.0495	33.24	33.40	1.34	73560	0.28	0.16

<i>Rg V/BA MSA</i>	29.98	0.0014	29.90	30.05	0.25	66177	1.36	0.15
<i>Rg V/E MSA</i>	29.98	0.0475	29.54	30.41	1.45	66177	0.16	0.87
<i>Rc V/BA ME</i>	29.98	0.0024	29.88	30.08	0.33	66177	1.29	0.20
<i>Re V/BA ME</i>	29.98	0.0010	29.92	30.04	0.21	66177	1.40	0.13
<i>Rc V/E ME</i>	29.98	0.0499	29.53	30.43	1.49	66177	0.12	0.89
<i>Re V/E ME</i>	29.98	0.0423	29.57	30.39	1.37	66177	0.24	0.82

VM = Variance of the sample mean in $\text{m}^3 \text{ha}^{-1}$; *LS* and *LI* = Upper and lower limit of the mean; *P* = Precision; *GP* = Gain in precision; *A* = Width of the confidence intervals.

The favorable effect when stratification by age class is made is outstanding. In this regard, Lencinas and Mohr-Bell (2007) point out that a strategy to optimize an inventory is to stratify with some variable strongly related to the variable of interest, as in this case was the age of the CFPs. From the foregoing, it can be deduced that by implementing stratification by age class in practice, the sampling effort can be achieved to be less, since a smaller sample size is required, which coincides with that determined by Roldán *et al.* (2014). Furthermore, the sampling error is favored when estimating the population parameter. In contrast, when age was used only as an auxiliary variable, particularly in the case of the *R* estimator in *MSA*, poor values were obtained in the referred statistical efficiency parameters and the mean was overestimated.

When *BA* was used as an auxiliary variable, the ratio estimators within the *ME* were superior to the ratio estimators in the *MSA*, based on Freese (1976), Scheaffer *et al.* (1987) and Corona and Fattorini (2006). This is due to the fact that by dividing the population into strata, homogeneity is achieved within them, thereby reducing the total variance and thereby increasing the accuracy of the estimator; thus, for the same sampling intensity, the *ME* generates more precise estimates than the *MSA*.

Based on this experience and on Köhl and Magnussen (2016), other variables, in addition to age, that can be considered as logical and feasible to stratify CFPs are the classes by site index, by diameter category or by increment in *BA*.

BA as an auxiliary variable was better than age, a more pronounced effect for the R_c and R_e estimators within the *ME*; in this regard Fattorini *et al.* (2015), Vallée *et al.* (2015) and Adichwal *et al.* (2019) indicate that the auxiliary variables used, in addition to having a high correlation with the variable of interest, must be easy, fast and cheap to measure, which makes them highly efficient. Scheaffer *et al.* (1987) report that, when the correlation coefficient (ρ) between the variable of interest and the auxiliary is greater than 0.5, the estimator of the ratio for the population mean is more precise than the *MSA*; in this case, the superiority of *BA* as an auxiliary variable is due to the fact that the value of ρ was 0.99 and that of age 0.47.

The volume versus the *BA* showed a linear relationship through the origin (Figure 1a), the variance of the volume maintained a proportional trend with respect to the *BA*, favorable conditions for which the ratio and regression estimators had a lower variance for the mean population and consequently greater precision than the simple mean obtained from the *MSA*. This coincides with what was reported by Roldán *et al.* (2014) for *Eucalyptus urophylla* S. T. Blake CFPs, who for the volume mean determined that the *AB* with $\rho = 0.97$ was more accurate for the R_c estimator in *ME*. It also agrees with that recorded by Fierros *et al.* (2018) when noting that the R_g estimator, with *BA* as an auxiliary variable, was the most accurate to estimate the timber inventory in CFPs of *Pinus chiapensis* (Martínez) Andersen.



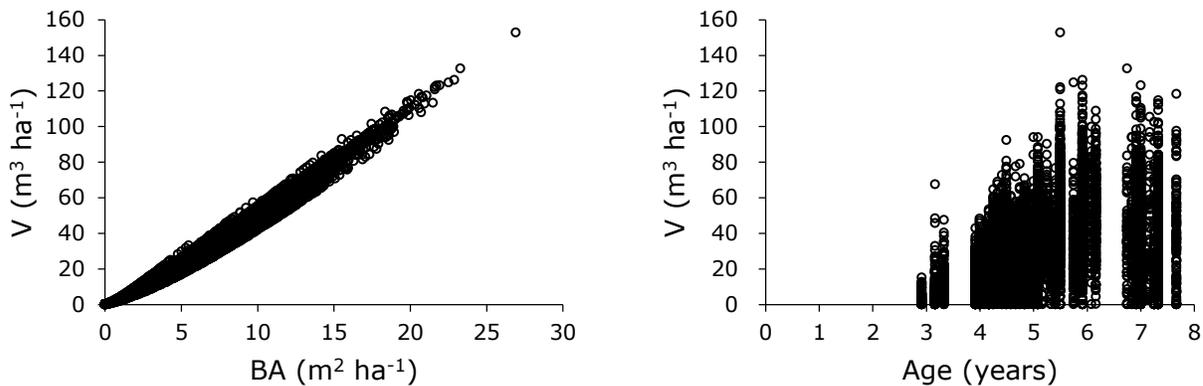


Figure 1. Behavior of the volume compared to the basimetric area (a) and age (b) used as auxiliary variables in the ratio and regression estimators in commercial Teak forest plantations.

Ratio estimators have the additional advantage that the auxiliary variable offers practical implication when estimating inventory. In this case, the value of the ratio indicates that for each m^2 of BA there is a volume of $4.7 \text{ m}^3 \text{ ha}^{-1}$, which allows estimating the inventory in volume immediately, just by knowing precisely the population BA , which it could be determined by surveying sampling points with a relascope. When age was the auxiliary variable, it was determined that for each passing year, the volume increases on average $5.8 \text{ m}^3 \text{ ha}^{-1}$; this makes it possible for the estimator to function as a simplified growth and yield system, because by knowing the age weighted by strata (5.7 years), the inventory can be estimated in an annualized way as the plantation grows; thus, it is assumed that the planted area will remain constant over time and that plantation areas are not incorporated or removed.

The V/AB ratio value in this study was slightly lower than that determined with a growth system developed by Tamarit *et al.* (2019) for the same species and study region, which was $5.3 \text{ m}^3 \text{ ha}^{-1}$ for the average condition (site index 15 m). Meanwhile, the value of the V/age ratio was less than the $9.3 \text{ m}^3 \text{ ha}^{-1}$ estimated with the growth system for the same average condition referred to.

Even though the ratio estimators are biased, the bias is minimal and can be considered null as long as the sample size per stratum is equal to or greater than 30 sampling units (Freese, 1976; Scheaffer *et al.*, 1987; Velasco *et al.*, 2003), a condition that is fulfilled in this analysis for each stratum because the smallest stratum was 100 sites. This helps to explain the greater efficiency of the estimator of Re , which has the additional advantage over that of Rc , the fact that when obtaining the estimator for each stratum and then adding them, it causes the ratio to vary between strata, which gives separate estimations and information on the population at the stratum level.

When the relationship between the main variable and the auxiliary variable is linear through the origin (Figure 1a), the estimators of Rc and Re by stratum are practically unbiased and in practice the bias can be considered null. In this context, according to Raj (1980) and Pérez (2005), the estimator of Re was higher than that of Rc (in both the BA was an auxiliary variable), because the number of strata was small and each stratum was relatively large, which implied having few addends and the accumulation of bias was minimal; furthermore, the ratios per stratum were different and ascending as the age class increased.

With reference to the estimator of Rc in the ME with the AB as an auxiliary variable, Pérez (2005) indicates that its advantage over that of Re is that it does not present accumulation of biases in the strata, so that the total bias, when it is present, is reduced to the minimum. However, it has the disadvantage that it does not offer estimates separated by strata; in practice, the Rc should be used when the number of strata is large and the size of each one is small.

The analysis of the behavior of the sample size (n) when varying the precision showed that to achieve high precision (less than 2 %) a greater number of sites is required (Figure 2a), which implies a greater sampling effort and that the costs increase. It was detected that the MSA is more demanding in terms of sample size than the ME , a logical situation because the former does not rely on the auxiliary information of strata in the population, in which the units that compose them are

internally as homogeneous as possible and the more heterogeneous among them. This allows the sampling error to be reduced and the precision to improve (Pérez, 2005). Therefore, for the same sample size, the *ME* has a higher precision gain than the *MSA*.

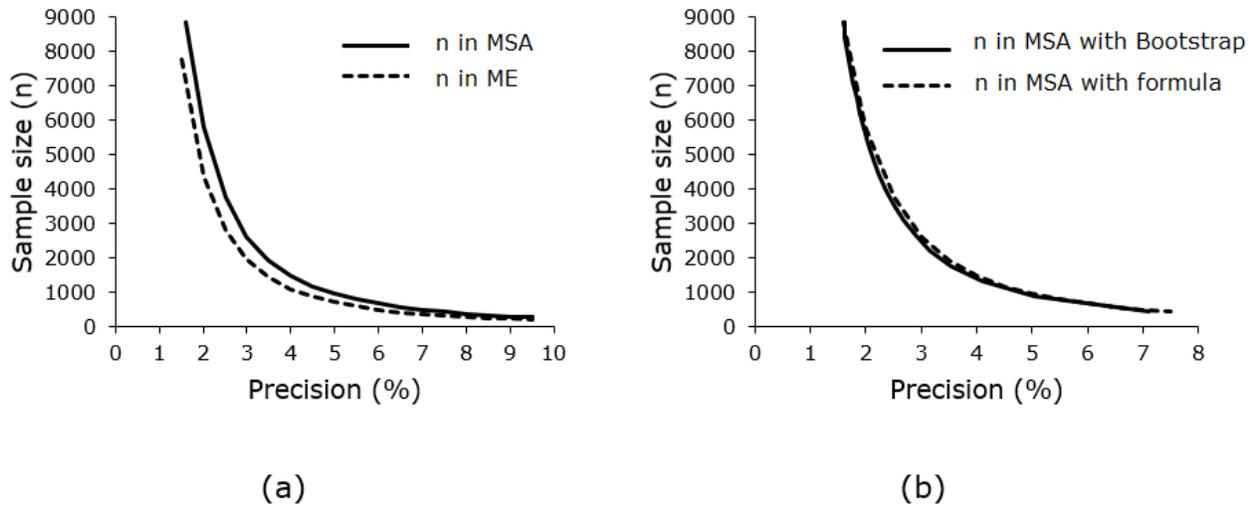


Figure 2. Sample size calculated by varying the precision in *MSA* and *ME* (a) and sample size in *MSA* calculated with the Bootstrap technique and by formula (b).

The trend of the averages of the estimated precisions, when varying the sample size, was very similar when it was determined with the Bootstrap technique for the *MSA* than when it was obtained by the formula (Figure 2b). This behavior leads to the assumption that similar patterns of correspondence would be defined in the rest of the assessed estimators.

To achieve the precision set at 2.5 %, it was determined that the optimal n calculated with the respective formulas was $n = 3\,744$ for the *MSA* and $n = 2\,799$ for the *ME*; which implies that the original sample size can be reduced by 57.6 % and 68.3 %, respectively. This means less sampling effort to collect the information

in the field and, consequently, the time is shortened and the cost of investing to carry out the inventory decreases.

The Bootstrap procedure also allowed a better appreciation of the effect of the sample size reduction on the variance, which increases up to a certain limit as the sample decreases (Figure 3a). Furthermore, the upper and lower limits of the *IC* s of the mean tend to widen, which is accentuated when the mean is reduced above 70 % (Figure 3b).

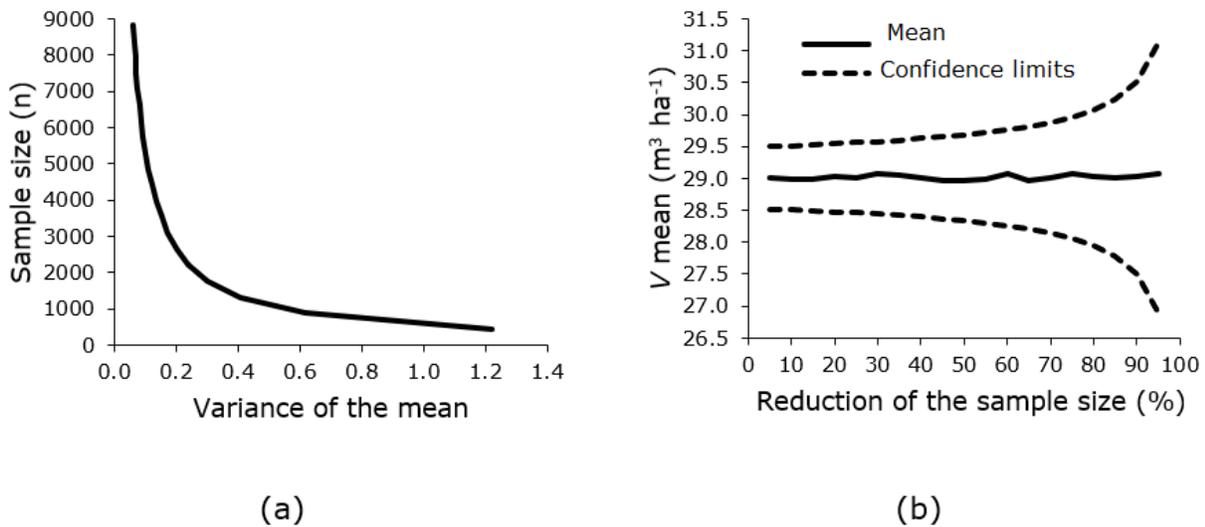


Figure 3. Effect of the reduction of the sample size in *MSA* on the variance of the mean in volume (a) and on the respective *IC* (b).

The sample size could be further reduced as precision relaxes, especially in *ME*; however, it must be borne in mind that in order to put into practice the best referred estimators, it must be ensured that the sample has at least 30 sites per stratum; thus, the population parameters will have the best precision. In this regard, the distribution of the sample size calculated by stratum and total for the 2.5% precision ($\alpha = 0.05$) in the proportional allocation and the Neyman is shown in Table 4. From a total of n sampling units, Cochran (1993) and Pérez (2005) refer that in the proportional allocation the number of sample units in each stratum is

proportional to its size and in the Neyman it is guaranteed that the sample size in each stratum has minimal variance and, therefore, the estimate of the population mean is more accurate. The increase in the number of sites in both types of assignments when passing from stratum 7 to 8 is possibly due to the fact that the growth conditions, given by the quality of the site, for age class 8 are comparatively more heterogeneous, in addition to the fact that the surface area is greater in 36.5 ha (Table 2).

Table 4. Proportional and Neyman distribution of sample size by stratum and total in commercial Teak forest plantation.

Distribution	Stratum (age class)						Total
	3	4	5	6	7	8	
Proportional	32	284	1 089	793	277	323	2 799
Neyman	6	130	906	975	349	434	2 799

When the objective is to obtain separate estimators for each stratum and the global estimator is considered to be of secondary importance, Freese (1976) points out that intense sampling should be carried out in the strata that have material of high commercial value, then the referred assignments go to background. Under this scenario, the number of necessary sites must be raised in such a way as to achieve the desired degree of precision for the strata of interest. This is relevant to apply in CFPs whose ages are close to the shift and in which commercial thinning will be applied.

One additional factor that must be considered in determining the sample size is what the current forest legislation establishes in terms of precision and reliability. In this regard, the Official Mexican Standard NOM-152-SEMARNAT-2006 demands that in the preparation of timber inventories through sampling, a minimum reliability of

95 % ($\alpha = 0.05$) at the farm level and a maximum sampling error of 10 % must be obtained (Semarnat, 2008).

The efficient quantification of timber stocks gives certainty to face forest pledge processes, thereby contributing to the capitalization of the CFP as a business and forest company, because it is possible to place the value of the plantation as collateral for the banking transaction or good to improve securitization and market conditions.

Conclusions

The optimal sampling strategy in terms of precision to carry out the operational inventory and estimate the mean volume of PFC of Teak in *Campeche, Mexico*, was formed by associating the sampling design under simple random sampling with the estimator of specific ratio within sampling stratified, with basal area as an auxiliary variable and when stratifying by age class.

In the stratified sampling, the initial sample size could be reduced to 68.3 % to maintain the precision of 2.5 % (reliability of $\alpha = 0.05$), which implies less time and effort, with the consequent reduction of costs in the timber inventory.

The age of the CFP was more efficient when it was used for stratification than when it was used as an auxiliary variable in the ratio and regression estimators. To obtain the highest precision in the population parameters, it is recommended to stratify by age classes with intervals of one year.

Conflict of interests

The authors declare no conflict of interest.

Contribution by author

Juan Carlos Tamarit-Urias and Héctor Manuel de los Santos-Posadas: conceptualization and organization of the research, creation of databases, statistical analysis and writing of the document; Arnulfo Aldrete, José René Valdez-Lazalde, Hugo Ramírez-Maldonado and Vidal Guerra-de la Cruz: contribution of documentary references and review of the document. All authors participated in the proofreading of the document.

References

- Adichwal, N. K., P. Sharma, Y. R. Singh, and R. Singh. 2019. A class of estimators for population mean utilizing information on auxiliary variables using two phase sampling schemes in the presence of non-response when study variable is an attribute. *Pakistan Journal of Statistics* 35(3): 187-196. <http://www.pakjs.com/wp-content/uploads/2019/10/35301.pdf> (23 de diciembre de 2020).
- Breña P., A. F. 2004. *Precipitación y recursos hidráulicos en México*. Universidad Autónoma Metropolitana. México, D. F.: <https://uamenlinea.uam.mx/materiales/licenciatura/hidrologia/precipitacion/PRHM04-libro.pdf> (25 de mayo de 2020).
- Cochran, W. G. 1993. *Técnicas de muestreo*. Compañía Editorial Continental. México, D. F., México. 513 p.

Corona, P. and L. Fattorini. 2006. The assessment of tree row attributes by stratified two-stage sampling. *European Journal of Forest Research* 125(1): 57-66. Doi:10.1007/s10342-005-0078-2.

Fattorini, L., P. Corona, G. Chirici and M. C. Pagliarella. 2015. Design-based strategies for sampling spatial units from regular grids with applications to forest surveys, land use, and land cover estimation. *Environmetrics* 26(3): 216-228. Doi:10.1002/env.2332.

Fierros M., R., H. M. De los Santos P., A. M. Fierros G., F. Cruz C., L. Martínez Á., E. Velasco B. y G. Quiñonez B. 2018. Estimadores de muestreo para inventario de plantaciones de *Pinus chiapensis* (Martínez) Andresen. *Revista Mexicana de Ciencias Forestales* 9(47): 131-148. Doi: 10.29298/rmcf.v9i47.155.

Freese, F. 1976. Elementary forest sampling. USDA FS Agriculture Handbook Num. 232. Washington, DC, USA. 91 p.

Grafström, A., S. Saarela and L. T. Ene. 2014. Efficient sampling strategies for forest inventories by spreading the sample in auxiliary space. *Canadian Journal of Forest Research* 44(10): 1156-1164. Doi: 10.1139/cjfr-2014-0202.

Gregoire, T. G. and T. H. Valentine. 2008. Sampling strategies for natural resources and the environment. Chapman & Hall/CRC. Boca Raton, FL, USA. 474 p.

Köhl, M. and S. Magnussen. 2016. Sampling in forest inventories. *In*: Pancel, L. and Köhl M. (eds). *Tropical Forestry Handbook*. Heidelberg, Germany. pp. 777-837.

Köhl, M., A. Lister, C. T. Scott, T. Baldauf and D. Plugge. 2011. Implications of sampling design and sample size for national carbon accounting systems. *Carbon Balance and Management* 6(1): 10. Doi:10.1186/1750-0680-6-10.

Lencinas, J. D. y D. Mohr-Bell. 2007. Estimación de clases de edad de las plantaciones de la provincia de Corrientes, Argentina, con base en datos satelitales Landsat. *Bosque* 28(2): 106-118. Doi:10.4067/S0717-92002007000200003.

Marchi, M., C. Ferrara, G. Bertini, S. Fares and L. Salvati. 2017. A sampling design strategy to reduce survey costs in forest monitoring. *Ecological Indicators* 81: 182-191. Doi:10.1016/j.ecolind.2017.05.011.

Pérez, L. C. 2005. *Muestreo estadístico. Conceptos y problemas resueltos*. Pearson Prentice Hall. Madrid, España. 374 p.

R Core Team. 2020. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. <http://www.r-project.org/> (28 de febrero de 2020).

Raj, D. 1980. *Teoría del muestreo*. Fondo de Cultura Económica. México, D. F., México. 305 p.

Roldán, C. M. A., H. M. De los Santos P., H. Ramírez M., J. R. Valdez L., G. Ángeles P. y A. Velázquez M. 2014. Estimadores de muestreo para inventario de plantaciones forestales comerciales de eucalipto en el sureste Mexicano. *Revista Mexicana de Ciencias Forestales* 5(26): 38-56. Doi:10.29298/rmcf.v5i26.289.

Scheaffer, R. L., W. Mendenhall y L. Ott. 1987. *Elementos de muestreo*. Grupo Editorial Iberoamérica. México, D. F., México. 321 p.

Schreuder, H. T., G. T. Gregoire and G. B. Wood. 1993. *Sampling methods for multiresource forest inventory*. John Wiley & Sons, Inc. New York, NY, USA. 446 p.

Schreuder, H. T., R. Ernst and H. Ramírez M. 2004. *Statistical techniques for sampling and monitoring natural resources*. USDA F S. General Technical Report Num. 126. Fort Collins, CO, USA. 111 p.

Secretaría del Medio Ambiente y Recursos Naturales (Semarnat). 2008. Norma Oficial Mexicana NOM-152-SEMARNAT-2006, Que establece los lineamientos, criterios y especificaciones de los contenidos de los programas de manejo forestal para el aprovechamiento de recursos forestales maderables en bosques, selvas y vegetación de zonas áridas. Diario Oficial de la Federación. Secretaría de Gobernación. México, D. F., México.

http://dof.gob.mx/nota_detalle.php?codigo=5064731&fecha=17/10/2008 (2 de abril de 2021).

Shiver, B. D. and E. B. Borders. 1996. Sampling techniques for forest resources inventory. John Wiley & Sons, Inc. New York, NY, USA. 356 p.

Tamarit U., J. C. 2013. Cubicación, crecimiento y rendimiento maderable e inventario operativo para *Tectona grandis* en el Sureste de México. Tesis Doctoral. Programa Forestal. Colegio de Postgraduados. Texcoco, Edo. de Méx., México. 97 p.

Tamarit U., J. C., H. M. De los Santos P., A. Aldrete, J. R. Valdez L., H. Ramírez M. y V. Guerra D. 2014. Sistema de cubicación para árboles individuales de *Tectona grandis* L. f. mediante funciones compatibles de ahusamiento-volumen. Revista Mexicana de Ciencias Forestales 5(21): 58-74. Doi:10.29298/rmcf.v5i21.358.

Tamarit U., J. C., H. M. De los Santos P., A. Aldrete, J. R. Valdez L., H. Ramírez M. y V. Guerra D. 2019. Sistema de crecimiento y rendimiento maderable para plantaciones de teca (*Tectona grandis* L. f.) en Campeche, México. Madera y Bosques 25(3):e2531908. Doi:10.21829/myb.2019.2531908.

Vallée, A. A., F. R. Bastien, L. P. Rivest and Y. Tillé. 2015. Incorporating spatial and operational constraints in the sampling designs for forest inventories. Environmetrics 26(8): 557-570. Doi:10.1002/env.2366.

van Laar, A. and A. Akça. 2007. Forest mensuration. Springer. Dordrecht, The Netherlands. 383 p.

Velasco B., E., H. Ramírez M., F. Moreno S. y A. de la Rosa V. 2003. Estimadores de razón para el inventario nacional forestal de México. Ciencia Forestal en México 28(94): 23-43.

<https://cienciasforestales.inifap.gob.mx/editorial/index.php/forestales/article/view/882/2145> (4 de mayo de 2020).

West, P. W. 2017. Simulation studies to examine bias and precision of some estimators that use auxiliary information in design-based sampling in forest inventory. *New Zealand Journal of Forestry Science* 47: 22. Doi:10.1186/s40490-017-0101-7. Tapia, J. y J. Návar. 2011. Ajuste de modelos de volumen y funciones de ahusamiento para *Pinus pseudostrobus* Lindl. en bosques de pino de la sierra madre oriental de Nuevo León, México. *Foresta Veracruzana* 13(2): 19-28. <https://www.redalyc.org/pdf/497/49721457004.pdf> (30 de septiembre de 2021).



All the texts published by **Revista Mexicana de Ciencias Forestales** –with no exception– are distributed under a *Creative Commons* License [Attribution-NonCommercial 4.0 International \(CC BY-NC 4.0\)](https://creativecommons.org/licenses/by-nc/4.0/), which allows third parties to use the publication as long as the work’s authorship and its first publication in this journal are mentioned.