AN EXAMPLE OF THE USE OF A TWO-PHASE SAMPLING DESIGN FOR RECONNAISSANCE INVENTORY IN TROPICAL FOREST

I. D. HUTCHINSON*

Abstract

A two-phase stratified random sampling design (as developed by the United States Forest Service) was used in south-eastern Paraguay for a reconnaissance inventory to estimate areas and wood volumes over a region some 6 million ha in size. Details of its use are presented to illustrate the suitability of the design in forest areas where little prior information is available to assist inventory planning.

INTRODUCTION

This paper describes the use in tropical forest of an inventory design for two-phase (or "double") sampling with optimum random allocation of field-plots. The design was developed by the Northeastern Forest Experiment Station (NEFES) of the United States Forest Service for its continuous forest inventory. The two phases of the design are:

Phase I: Photo-interpretation. A relatively large number of points are located at random on aerial photographs. Photo-interpretation allocates each point to one of several forest strata, and thus estimates the relative areas of the strata. As this phase is relatively inexpensive per point, it seeks to provide the maximum amount of information possible.

Phase II: Field enumeration. A subsample of the photo-points is enumerated in the field to derive means and confidence intervals for volume in each stratum.

As pointed out by Bickford (1952), the advantages to be gained by using this design depend on the difference of the cost per point in Phase I from that per plot in Phase II, and upon the extent to which total variation in the forest exceeds the variance in each of the important strata.

The design has proved to be flexible, efficient and economical under a wide range of conditions. It is particularly suited for reconnaissance surveys, and may be used for feasibility studies when stand maps are not required.

---

*FAO Forestry Officer, P.O. Box 240, Kuching, Sarawak, Malaysia.

The example outlined here formed part of a forestry project in Paraguay carried out for the United Nations Development Programme by the Food and Agriculture Organisation during the period 1967-72.

PURPOSE OF THE INVENTORY

Most forest land in south-eastern Paraguay is privately owned, and has been remote and of limited access. For almost a century, hardwood logs have been selectively cut for rafting down the Parana River to an export market in Argentina. At the time of the inventory, Paraguay had neither declared a formal forest policy nor established a forest department. However, it was actively legislating to prohibit log exports and to encourage the establishment of national timber processing industries.

Thus, when planning for an inventory, it was considered that a reconnaissance inventory would be of most service if it provided a basis for national and regional decisions on forest policy; that is, it should estimate areas and volumes over a large region, rather than depict the size and location of commercial stands on selected tracts.

The standards assumed for potential utilisation were those for a modern integrated forest industry producing sawn timber, veneers, and plywood.

STRATIFICATION

The few previous studies made in south-eastern Paraguay, as well as those available from the neighbouring province of Misiones, Argentina, were reviewed in order to form an idea of possible strata and the frequency of their occurrence. One useful source of information proved to be maps by Plan Triangulo (1967), showing vegetative cover and land use for the southern half of the inventory area. These provided the basis for estimating the relative land area of each stratum which was used in intial planning (Table 1). Another was a report by Italconsult (1966) which gave estimates of means and variances for volume per hectare, derived from 200 ha of forest enumerated as part of a proposal for a land development scheme.

From aerial photographs of south-eastern Paraguay (1:50 000 scale taken in 1965), it was considered that five simple strata could be defined in advance. These are listed below and defined in Appendix 1. The strata were:

1. High Forest
2. High Forest Regenerating

3. Low Forest (Scrub)
4. Non-forest
5. High Forest Margins

The "margins" stratum was a catch-all into which were grouped the photo-points which happened to fall on a boundary between high forest and another stratum. Functions of the "margins" stratum include:

(a) Avoiding the complications and bias which arise whenever a point straddling a boundary between strata is moved in order to contain it wholly within one stratum.
(b) Reducing the variance of the high forest stratum by removing from it any photo-points which are only partly covered by high forest.

In south-eastern Paraguay, the frequent intrusion into the forest of long tongues of swampy grassland resulted in 35% of all photo-points falling into the "margins" stratum.

REASONS FOR CHOOSING A TWO-PHASE DESIGN

The alternative to a two-phase design is to delineate strata. When strata are delineated, the entire area under study is classified by photo-interpretation, and lines are drawn on maps to show the boundaries between strata. Delineation provides exact information on areas and, done correctly, avoids the errors of area estimation which are inherent in a design based on plots. On the other hand, delineation is slow and expensive, requires much field checking, and can be done only where a good control base exists for mapping. Furthermore, the objectives of an inventory may not require that strata be delineated.

In south-eastern Paraguay, the composition of the forest was similar to that described by Bickford (1952). It comprised many species, with several ages represented on any chosen hectare, in a pattern of small and irregular stands merging from one class to another. Delineation of strata is difficult in such circumstances.

Using the information taken from the reports by Plan Triangulo (for stratum weights) and by Italconsult (means and variances), the number of field-plots which would be required for both (a) delineation of strata (Formula II, Appendix 3) and (b) two-phase sampling (Formulae, V, VI and VII) were calculated. Two-phase sampling required only three field-plots more than did delineation. Consequently it was adopted as the more desirable and economical method.

SAMPLE SIZE

Sample size was calculated according to the method outlined by Bickford *et al.* (1963) and Bickford (1967). Table 1 illustrates the sources of data and the sequence of calculation used to determine the number of photo-points for Phase I and of field-plots for Phase II (Formulae V, VI and VII, Appendix 3).

Formula VI requires an estimate of the ratio between the cost of interpreting a photo-point and the cost of enumerat-

TABLE 1: THE DEVELOPMENT OF BASIC DATA FOR INVENTORY PLANNING

Sequence	*Stratum*	*No. of Photo-points* n'_h	*Relative Surface Area* w_h	*No. of Field-plots* n_h N_h	*Estimated Mean Total Vol./ha* $\bar{y}'_h$	*Standard Deviation* s_h
1st	*Data used to estimate the number of photo-points for Phase I*: Arbitrarily assuming a cost ratio of 1:1000, and equal sampling costs in all strata; values for stratum weights (estimated from the maps by Plan Triangulo), together with means and standard deviations per stratum (calculated from the report by Italconsult), were substituted in Formulae V and VI. This gave a result of 2920 photo-points to be interpreted in Phase I.					
2nd	*Data used to calculate the number of field-plots for Phase II*: (Formula VII)					
	1	1198	0.4014	76	30	30
	2	44	0.0147	1	10	15
	3	218	0.0730	1	2	3
	4	471	0.1578	1	1	1
	5	1054	0.3531	45	15	20
	Σ	2985	1.0000	125		
		(From Phase I)			(From Italconsult)	
	Results: $\bar{y}' = 17.79$ m³/ha (Formula III): $n = 125$.					
3rd	*Results of Phase II*:					
	1	1049	0.3514	85	32.38	16.22
	2	74	0.0248	2	5.72	3.46
	3	307	0.1028	3	1.59	1.00
	4	650	0.2178	2	—	—
	5	905	0.3032	36	8.98	11.49
	Σ	2985	1.0000	128		
		(From photo-interpn. adjusted according to field appraisal.)			(From field enumeration, Phase II.)	
	Results: $\bar{y}' = 16.37$ m³/ha.					

ing a field-plot. The NEFES uses a ratio of 1:200 (Bickford, 1967). A ratio of 1:1000 was arbitrarily assumed for south-eastern Paraguay. Local experience later indicated it to be more correctly in the order of 1:500.

PHASE I. PHOTO-INTERPRETATION

The NEFES uses a photo-point not smaller than the equivalent of approximately half a hectare on the ground (Bickford, 1952). In Paraguay, photo-points were circles of one centimetre in diameter marked on the aerial photographs. Thus, the area of each point was equivalent to approximately 20 ha on the ground.

Photo-interpretation assigned each photo-point to one of the strata. In addition, for each point, the watershed and political subdivision were recorded, enabling estimates to be presented on these bases.

In its own work, NEFES further classifies its photo-points according to the photo-interpreter's estimate of mean standing volume per hectare (Bickford, 1952). This is similar in concept to an Australian technique described by Smith and Webb (1973). In the initial stages of investigating a tropical hardwood forest it is rarely possible to do this type of thing.

PHASE II. FIELD ENUMERATION

Field-plots were selected from the photo-points and distributed among the strata (Formula VII, Appendix 3), using a random start followed by a sampling interval. Anticipating changes in land use since the year in which the aerial photographs were taken, the number of plots to be enumerated in the high forest stratum was increased from 76 to 85. Circumstances did not permit all "margins" plots to be enumerated.

"Bounded" plots were used, making sampling proportional to the frequency of occurrence of tree species within strata. This provided elementary information which could be used later in planning approaches to silviculture and forest management. The field-plot was a grid of transects, each transect 10 metres wide, totalling one hectare in area. The centre of the field-plot purported to coincide with the centre of the corresponding photo-point. Both quantitative and qualitative variables were recorded on each plot.

Once field-plots had been selected, the logistics of the inventory became clear. It became possible to begin planning for the types of terrain and distances involved, and for the problems of organisation and transport. As no ordnance maps existed for the inventory area. sketch maps to guide field access were made from the photo-indices. These were continually amended with information gathered by field parties during the course of the inventory.

An incidental advantage of the design was that field parties travelled throughout almost all parts of the inventory area on their way from one field-plot to another. This enabled staff to form a subjective impression of changes in land use, of the types of field checks required, and of any strata in which it might be advisable to increase the number of field-plots.

Plots were located in the field by means of an access line from a point identifiable on the aerial photographs, this point having been chosen as close to the plot centre as possible. From this point, a baseline was established by line of sight to a second identifiable point. From the baseline, a compass bearing was calculated for the access line from the first point to the plot centre.

The longest access line was in the order of 10 km. Most were quite short because field parties took advantage of the intrusions of natural grassland, and also made a practice of employing guides from nearby farmhouses detected on the aerial photographs. As a result, hunters' paths, not visible on the photographs, were frequently used to arrive close to the centre of a field-plot.

Slope readings were recorded along the length of each access line to assemble information on the operability of the terrain.

INVENTORY RESULTS

Data Processing

Data from the field enumeration were processed at NEFES by the computer program, FINSYS, a development from that described by Wilson and Peters (1967). FINSYS is a series of flexible programs which include subsystems for editing, compiling tables, and output. Variance and sampling error are calculated for each estimate. Applications of this program have been discussed by a number of authors, including Barnard and Letourneau (1974) and Cunia (1975).

Discrepancies in Stratification

In a number of instances, a field-plot was found to belong to a stratum different from that in which it had been placed by photo-interpretation. Such differences arose either from errors in photo-interpretation, or from changes in land use since the aerial photographs were taken.

Where necessary for estimates of area, the relative land areas estimated from photo-interpretation were adjusted proportionally according to the stratification of field-plots. Adjusted relative areas, reflecting changes in land use, formed a useful part of the information from the inventory. The "margins" stratum was valuable in this respect, being the most sensitive to changes of the type wrought by the incursion of agriculture or the abandonment of cultivation at the forest's edge. Assuming the proportion of change observed did actually represent change which had taken place, changes were converted to hectares and prorated according to the time elapsed between photography and enumeration.

For volume calculations, all field-plots, without exception, remained in the stratum into which they were placed by photo-interpretation. Any increase in variance, due to discrepancies observed in the field, was accepted.

Estimates of Land Area

Aeronautical charts were the best maps available to show the gross area of the tract under study. This area was measured by planimeter as being 6 401 000 ha.

The relative land area of each stratum was converted to hectares (Formula VIII, Appendix 3). Areas of strata were shown for both the physical and political subdivisions of the inventory area. Using the results of field enumeration, estimates of land area were presented for simple soil types, operability, and land ownership.

The actual disposition of the field-plots in the terrain was used as a basis for sketch maps showing the geographic distribution of tree species throughout the inventory area, and the location of tracts in which the advance growth of desirable species was sufficient to justify a beginning with silviculture and management.

Estimates of Wood Volume

Mean total wood volume per hectare was calculated from the results of photo-stratification and field enumeration (Formula X, Appendix 3).

Estimates of wood volume were presented for the physical and political subdivisions of the inventory area; for classes of land ownership; for simple soil types and forest exploitation classes; for tree species, wood quality groups, and dbh classes; and for log grades and classes of log length.

Sampling Errors

The allowable standard error was arbitrarily set at $\pm$ 10% of the overall mean volume (Formula III, Appendix 3) at the

67% confidence level. Lacking further information, no margins of error were specified for other variables (Bickford, 1967).

The initial planning for the inventory suggested a mean total volume of 17.79 m^3/ha (Table 1). This signified an allowable error of 1.78 m^3.

Field enumeration revealed the sample mean to be 16.37 m^3 (Formula III). The actual standard error of the volume estimate was calculated to be 1.01 m^3 (Formula XII), that is, it fell within the limit of 10%.

Information for Planning Later Inventories

Means and variances for each stratum were calculated by the computer program. This completed the basic information required for planning any later inventories in similar forest (Table 1).

CONCLUSION

In the tropical forest in which it was employed, the two-phase sampling design functioned successfully, in spite of the sparse information which existed prior to the inventory. The design proved to be economical, producing results of good precision from relatively few field-plots. It shows promise for extensive inventories designed to monitor changes in forest resources.

ACKNOWLEDGEMENT

The author acknowledges the generous assistance of Professor C. A. Bickford, consultant to the Paraguayan inventory; of the Director of the Northeastern Forest Experiment Station of the U.S. Forest Service, and of the officers of the NEFES Forest Survey Project. Permission was granted by the Assitant Director-General, Forestry Department, FAO, to use material from FAO reports. The author thanks Drs A. G. D. Whyte and C. J. Goulding for their comments.

REFERENCES

Barnard, Joseph E.; Letourneau, Larry, 1974. *Data-processing Systems for Large Inventories.* Inventory Design and Analysis. Proceedings of Forest Inventory Workshop at Colorado State University, Society of American Foresters, Washington, D.C., pp. 137-49.

Bickford, C. A., 1952. The sampling design used in the forest survey of the northeast. *J. For., 50* (4): 290-3.

——— 1967. Sampling Plan Proposed for the Reconnaissance Inventory UN Special Fund Project No. 290 to Develop Forestry and Forest Industries, Asuncion, Paraguay. 37 pp. (mimeo).

Bickford, C. A.; Mayer, C. E.; Ware, K. D., 1963. An efficient sampling design for Forest Inventory: the northeastern forest resurvey. *J. For.*, *61* (11): 826-33.

Cochran, W. G., 1963. *Sampling Techniques.* Wiley, New York. 412 pp.

Cunia, T., 1975. *Inventory Designs in Europe and the United States, Canadian Forest Inventory Methods.* University of Toronto Press, Toronto, pp. 48-61.

Freese, F., 1962. Elementary forest sampling. *Agricultural Handbook No. 232,* USDA Forest Service, Washington, D.C., 91 pp.

Hutchinson, I. D., 1974. Inventario forestal de reconocimiento. *Informe Tecnico No. 1, Proyecto de Desarrollo Forestal y de Industrias Forestales, PNUD/FAO PAR/66/515.* FAO, Rome, 353 pp.

Italconsult, Argentina S.A., 1966: Proyecto de Colonizacion, Alto Parana. *Anexo No. 6, Instituto de Bienestar Rural del Paraguay, Buenos Aires.* 35 pp.

Plan Triangulo, 1967. *Borrador del Informe Final, Apendice C, Asuntos Forestales.* Secretaria de Planificacion Economica y Social, Asuncion, Paraguay. 35 pp.

Smith, R. B.; Webb, A. W., 1973: *Regional Timber Resources Survey.* Forest Commission, Victoria, IUFRO, Nancy. 9 pp. (mimeo).

Wilson, Robert W.; Peters, Robert C., 1967. The Northeastern Forest Inventory Data Processing System. *USDA Forest Service, Research Papers No. NE-61 and NE 70-80.*

APPENDIX 1

Definitions of the Photo-strata

1. *High Forest*: Not less than 90% of the area contained within the photo-point is composed of forest considered to be of economic value (that is, high forest or high forest regenerating). The area of high forest within the limits of the point exceeds that of high forest regenerating. In high forest the minimum height of dominant and codominant trees is 15 metres.

2. *High Forest Regenerating*: A minimum of 90% of the area at the photo-point is composed of forest of current economic value. A minimum of 50% of the area was originally high forest but now appears on the aerial photographs in a state of natural regeneration.

3. *Low Forest*: No more than 10% of the area at the photo-point is composed of high forest and/or high forest regenerating. The total height of dominant and codominant trees is estimated to be less than 15 metres. Within the bounds of the photo-point, the area of low forest is greater than that of non-forest.

4. *Non-forest*: Not more than 10% of the area within the bounds of the photo-point carries high forest and/or high forest regenerating. The area of non-forest is greater than that of low forest.

5. *High Forest Margins*: Between 10 and 90% of the area within the bounds of the photo-point carries high forest and/or high forest regenerating. The balance of the area of the photo-point may be composed of low forest, non-forest, or a combination of both.

APPENDIX 2

Definition of Statistical Symbols

a_h	estimated area of the "*h*th" stratum, hectares. Formula VIII.
A	total gross area of the inventory tract, hectares.
C_n	the cost of enumerating a field-plot. Formulae V and VI.
$C_{n'}$	the cost of interpreting a photo-point. Formulae V and VI.
e	the specified allowable error, cubic metres per hectare. See Formula III.
h	a subscript identifying one particular stratum, the "*h*th".
n'	the number of photo-points. Phase I. Formula V.
n	the number of field-plots. Formulae II and VI.
n_h	the number of field-plots in the "*h*th" stratum. Formula VII.
N_h	the number of field-plots actually enumerated in the "*h*th" stratum. Phase II.
s_h	the estimated standard deviation of the "*h*th" stratum, cubic metres per hectare. See Formula IV.
w_h	the estimated relative surface area of the "*h*th" stratum. The stratum weight. Formula I.
$\bar{y}_h$	the estimated mean volume of the population of the "*h*th" stratum, cubic metres per hectare. Prior sources, and Phase II. See Formula X.
$\bar{Y}'$	the estimated total mean volume of the population of all strata, cubic metres per hectare. Formula III.
Y'	the estimated total volume of the population of all strata, cubic metres. Formula XI.

APPENDIX 3

Formulae

Formula I: Relative surface area of a stratum. The "stratum weight".

(a) $w_h = n'_h/n$ (b) $w_h = a_h/A$

Formula II: The number of field-plots required for optimum allocation with complete delineation of strata. Sampling costs are assumed to be equal for all strata — Freese (1962), Cochran (1963), Bickford (1967).

$$n = (\Sigma\, w_h \cdot s_h)^2/e^2$$

Formula III: The allowable error, "e", was specified as 10% of the estimated mean total volume for the inventory area, $\bar{Y}'$ m³/ha, within 67% confidence limits.

$$\bar{Y}' = \Sigma\, w_h \cdot \bar{y}_h$$

Formula IV: Variance of the estimated mean total volume per stratum, in cubic metres per hectare.

$$s^2_h = \frac{n\,(\Sigma\, y^2) - (\Sigma\, y)^2}{n_h(n_h - 1)}$$

Formula V: The number of photo-points required for optimum allocation with two-phase (or "double") sampling. (Derived by Bickford (1967) from 12.8, 12.8′, and 12.9 in Cochran (1963).)

$$n' = \frac{n \cdot \sqrt{[\Sigma\, w_h\,(\bar{y}_h - \bar{Y}')^2 \cdot (C_n/C_{n'})]}}{\Sigma\, w_h \cdot s_h}$$

Formula VI: The number of field-plots. This calculation must be made twice, once to provide an approximate value of "n" to be substituted in Formula V, and a second time, using the relative surface areas provided by photo-interpretation, to calculate the number of plots to be enumerated in the field. Optimum allocation. The costs of field sampling are assumed to be equal in all strata. (Derived by Bickford (1967) from 12.8, 12.8′, and 12.9 in Cochran (1963).)

$$n = 1/e^2 \cdot z^2 \cdot (\alpha/2)\ [\Sigma\, w_h \cdot s_h)^2 + (\Sigma\, w_h \cdot s_h)$$
$$\sqrt{[\Sigma\, w_h(\bar{y}_h - \bar{Y}')^2 \cdot (C_{n'} / C_n)]}$$

For south-eastern Paraguay, the chosen confidence interval was one standard deviation, 67%. At this level, the expression "$z^2 \cdot$(alpha/2)" equals unity. At a confidence level of 95%, it equals 1.96.

Formula VII: The number of field-plots to be enumerated per stratum (Cochran, 1963). Allocation is optimal. Sampling costs are assumed to be equal in all strata.

$$n_h = \frac{n \cdot (w_h \cdot s_h)}{(\Sigma\, w_h \cdot s_h)}$$

Formula VIII: The estimated land area per stratum, hectares.

$$a_h = A \cdot w_h$$

Formula IX: The surface area represented by one field-plot in the "*h*th" stratum equals,

$$a_h / N_h$$

Formula X: Estimated total volume per stratum, cubic metres,

$$Y'_h = a_h \cdot \bar{y}_h$$

Formula XI: Estimated total volume for the inventory area, cubic metres,

$$Y = A \cdot (\Sigma w_h \cdot \bar{y}_h)$$

Formula XII: Standard error of the estimated mean total volume for the inventory area, cubic metres,

$$s_{\bar{Y}'} = \sqrt{\left[\frac{(\Sigma\, w_h \cdot s_h)^2}{n} + \frac{\Sigma\, w_h\, (\bar{y}_h - \bar{Y}')^2}{n'}\right]}$$