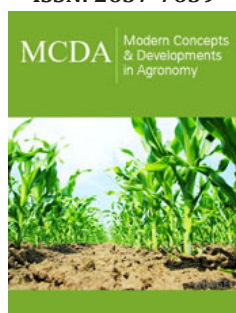


Sampling Design Contribution to Small Area Estimation Procedure in Forest Inventories

Aristeidis Georgakis* and Georgios Stamatellos

Laboratory of Forest Biometrics, School of Forestry and Natural Environment, Aristotle University of Thessaloniki, Thessaloniki, Greece

ISSN: 2637-7659



***Corresponding author:** Aristeidis Georgakis, Laboratory of Forest Biometrics, School of Forestry and Natural Environment, Aristotle University of Thessaloniki, Thessaloniki, Greece

Submission: 📅 June 23, 2020

Published: 📅 August 18, 2020

Volume 7 - Issue 1

How to cite this article: Aristeidis Georgakis, Georgios Stamatellos. Sampling Design Contribution to Small Area Estimation Procedure in Forest Inventories. *Mod Concep Dev Agrono.* 7(1).MCDA.000654.2020. DOI: [10.31031/MCDA.2020.07.000654](https://doi.org/10.31031/MCDA.2020.07.000654)

Copyright@ Aristeidis Georgakis, This article is distributed under the terms of the Creative Commons Attribution 4.0 International License, which permits unrestricted use and redistribution provided that the original author and source are credited.

Abstract

The sampling design is a crucial topic that would be considered in Small Area Estimation (SAE). Applications of sampling designs presented in Forest Inventories (FIs) for SAE, with the two-phase sampling to have the most references. Eventually, FIs that are applied for SAE is an open research topic. An important contribution to this topic would be the comparison and the optimization of sampling designs that aims to improve SAE in FIs.

Keywords: Survey sampling; SAE procedure; Domain; Forest management unit; Auxiliary information

Introduction

Forest inventories (FIs), based on a geographical scale, can be distinguished in National Forest Inventories (NFIs) and Management Forest Inventories (MFIs), that provide information for policy-making or local decision-making correspondingly. The initial objective of sampling design in FIs is to produce information (estimates) for one or more population parameters (variables of interest) of a targeted population, after selecting the proper formulas (estimators) [1,2] and the second aim is to provide suitable statistics for subpopulations, the so-called “domains” or “small areas” [3]. The last objective of sampling design (survey sampling) can be achieved through small area estimation (SAE) techniques.

There is an increasing need to use national or regional inventories for local estimations [4], particularly reliable forest attribute information is needed at different geographical scales with different requirements per scale. “SAE techniques address the situation where the number of samples within a small area is too small to provide reliable estimates for that unit” [5]. A small area characterized by small or even null sample size [6]. In the case of a small area, where direct estimations are not possible and when the sample size cannot be increased, indirect estimators (SAE technique) can be applied, “borrowing strength” from other domains or periods and combining the terrestrial information with the extensive use of auxiliary information such as derived from remotely sensed variables [4,6-8]. Borrowing strength is the basic idea of SAE, where models are fitted globally and applied locally, albeit with minor modifications [9].

Although FIs depicts the state of forests through a plethora of target variables, in SAE the most important quantitative variables of interest are the growing stock volume and the aboveground forest biomass. The basic prerequisite of SAE implementation is the acquisition of auxiliary variables (Figure 1). The main auxiliary data/information are satellite imagery and 3-D data from LiDAR or airborne laser scanning (ALS) and photogrammetry. The most critical step for having small area statistics is the selection of suitable estimation procedure under the existing (usually) sampling design.

The problem of small area statistics starts when the original sample design aims to the estimations of population totals (mean and variance) for a variable of interest and not in the small area of interest such as management units (eg. forest stands or compartments). What sample design can be used for SAE of small domains in the design phase, is an open question and a basic issue that should be considered [3]. In this paper, we will present existing sampling designs that support effectively the SAE procedure and we will discuss restrictions and opportunities about the implementation in FIs.

Sampling designs in SAE for FIs

Knowing the variable of interest, having defined the small area of interest and having available suitable auxiliary information with existing terrestrial data, the last “two steps” (Figure 1) for an effective “small area estimation strategy” are the sampling design and the selection of proper statistical modelling (estimation design) [3,10,11]. From another perspective, the last steps of design and estimation can be considered inseparable [6]. The research of SAE literature is broader out of the forestry borders, as well as, on sampling designs for SAE purposes. In socioeconomic fields, various sampling designs examined parallelly with different types of estimation strategies for SAE implementation [3,10-12]. Generally, there is a gap of this kind of research in forestry literature. Some exception is the work of [13] who compared and tested different sizes of sampling grids for SAE of forest area and the growing stock volume of temperate mixed forests.

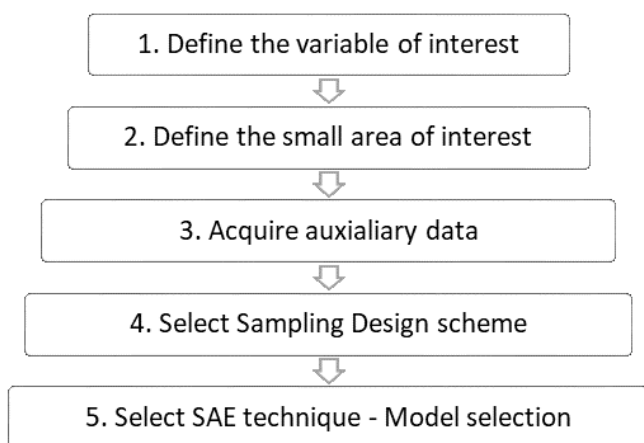


Figure 1: Procedure for SAE implementation.

The following sampling designs have been applied to SAE in FIs. The common component of all SAE applications is the use of auxiliary information that is exhaustive or partial exhaustive (for the whole population). Double-Sampling or two-phase is one of the most frequently used sampling design, characterized by its cost-efficiency for inventories in large remote forest areas [4,5,7,9,14,15] and [16] (section 6.3), three-phase sampling in smaller extend [5,17,18], stratified systematic (cluster) sampling [19], stratified random sampling [20], and post-stratification [15,21-23] for design-unbiased estimates (mean and variance) when a reasonable amount of field plots is needed in a small area [24]. Systematic or grid (sample locations on a regular grid) is one of the most common sampling (including cluster) scheme in MFIs and especially in NFIs. Correspondingly the majority of SAE bibliography utilizes NFI data to downscale the estimates to finer resolutions like territories, forest districts, or domains [5,13,25]. In small scale MFIs, systematic sampling design has comparatively less references in SAE literature [26,27,28] and aims for local estimations of forest management units such as forest stands or compartments.

Having exhaustive (wall-to-wall) auxiliary information (usually ALS), we can select beforehand more representative field samples

(well-spread), using the balanced sampling [29]. Considering that imputation methods are well suited for SAE such as nearest neighbour method [30-33], further improvements expected to reveal after the application of balanced sampling [29] or the nearest centroid [34]. Efficiency gains in the SAE also have been explored from Nearest centroid [34,35].

Double-sampling or two-phase sampling seems to be one of the major sampling design schemes in the applications of SAE in FIs. The advantage of two-phase sampling, compared to the two-stage sampling, relies on the very large sample units/points [4,9,18] of the first phase with high correlated variables of Remote sensing (ex. ALS) data that covers (nearly) the whole population. In the second phase, rationally we draw a smaller sample of terrestrial data. The sample unit is the same in both phases. In the first phase Mandallaz [4] introduces the infinite population or Monte Carlo approach on the design-based model-assisted estimations as more appropriate in the forest inventory context than the finite approach [12] of design-based inference. New regression estimators in FIs with two-phase sampling also have been proposed [9].

Another approach that looks similar to the double-sampling is the following: in the first step a sample survey is drawn, after that, regression models are fitted with dependent variables (ex, mean height, basal area or volume) and the auxiliary metrics (ex. ALS) as independent variables, and finally, the predictions of the units/pixels aggregated to larger areas (ex. forest stands) [19]. Initially, this estimation procedure characterized as “two-phase sampling” or “a two-step procedure”, wanting to characterize the more appropriate “synthetic regression estimation for small areas” [36] as referred from [19]. It is also well known that synthetic estimators have the property to provide estimations in a small area such as a forest stand without sample plots within [37]. An important model-based approach in FIs is the selection of unit-level models in the area-based approach (pixel/plot) [37]. Generally speaking, for pursuing design efficiency and cost reductions many forest surveys have adopted systematic sampling designs aided by remotely sensed auxiliary variables [8].

Discussion

The majority of SAE bibliography (including FIs) referred to the heart of SAE which is the statistical modelling or the selection of suitable estimator. Obviously, we cannot rely only on a traditional sample survey if we have only a few or no plots. When design-based it is not always feasible, then a model-based or model-dependent approach is one solution. If we select a model-based approach, then the sampling design can be ignored [27]. Magnussen et al. [8,38,39] demonstrate that an effective sampling design of a small area, considers possible domain or area effects (random effects) through measuring at least two plots per forest stand to avoid “a serious risk of a gross underestimation of uncertainty in a synthetic estimate of a stand mean”. When the sample size is kept low, a remaining challenge is to optimize the allocation of sample units [8].

Sampling design possibly cannot aim for both total and domain estimations. For example, in systematic sampling we cannot select

a priori the SAE technique, design, or model-based, “due to a sparse or nonexisting degree of replicated sampling within domains” [8]. A practical solution for applying designed-based and model-assisted estimators, instead of model-based is to extend the small area via post-stratification and thus to increase the sample size within.

In conclusion, the sampling design applied for SAE is an open research area. Questions like: “How the sampling design (sample size, plot size, plot allocation) affect the SAE?” or “What estimators we can use under existing sampling design in SAE?” are open. An important contribution to this topic would be the comparison and the optimization of sampling designs that aims to improve SAE in FIs.

Acknowledgement

This research has been financially supported by the General Secretariat for Research and Technology (GSRT) and the Hellenic Foundation for Research and Innovation (HFRI) (Scholarship Code: 1319).

References

- Köhl M, Magnussen S, Marchetti M (2006) Sampling methods, remote sensing and GIS multiresource forest inventory: Springer, Heidelberg, Berlin, Germany.
- McRoberts RE, Andersen HE, Næsset E (2014) Using airborne laser scanning data to support forest sample surveys. In Maltamo M, Næsset E, Vauhkonen J (Eds.), *Forestry applications of airborne laser scanning: concepts and case studies*, Springer, Dordrecht, Netherlands, pp. 269-292.
- Molefe WB (2011) Sample design for small area estimation. (Doctor of Philosophy thesis), University of Wollongong, Australia.
- Mandallaz D (2013a) Design-based properties of some small-area estimators in forest inventory with two-phase sampling. *Canadian Journal of Forest Research* 43(5): 441-449.
- Hill A, Mandallaz D, Langshausen J (2018) A double-sampling extension of the german national forest inventory for design-based small area estimation on forest district levels. *Remote Sensing* 10(7): 1052.
- Fabrizi E, Źądło T (2018) Survey sampling and small-area estimation. *Mathematical Population Studies* 25(4): 181-183.
- Gabriel A, Hill A, Breschan J (2018) Neue Hilfsmittel zur Anwendung zweiphasiger Stichprobenverfahren in der Waldinventurpraxis. *Schweizerische Zeitschrift für Forstwesen* 169(4): 210-219.
- Magnussen S (2016) A new mean squared error estimator for a synthetic domain mean. *Forest Science* 63(1): 1-9.
- Mandallaz D, Breschan J, Hill A (2013) New regression estimators in forest inventories with two-phase sampling and partially exhaustive information: a design-based Monte Carlo approach with applications to small-area estimation. *Canadian Journal of Forest Research* 43(11): 1023-1031.
- Nekrašaitė Liegė V (2012) Small area estimation (Doctoral Dissertation), Vilnius Gediminas Technical University, Vilnius, Lithuania.
- Zimmermann T (2018) The interplay between sampling design and statistical modelling in small area estimation. (Doctoral Thesis), Trier University, Trier, Germany.
- Rao JN, Molina I (2015) *Small area estimation*: John Wiley & Sons, Inc., Hoboken, New Jersey, USA.
- Steinmann K, Mandallaz D, Ginzler C, Lanz A (2013) Small area estimations of proportion of forest and timber volume combining Lidar data and stereo aerial images with terrestrial data. *Scandinavian Journal of Forest Research* 28(4): 373-385.
- Andersen HE, Breidenbach J (2007) Statistical properties of mean stand biomass estimators in a LIDAR-based double sampling forest survey design. Paper presented at the Proceedings of the ISPRS workshop laser scanning.
- Mandallaz D, Hill A, Massey A (2016) Design-based properties of some small-area estimators in forest inventory with two-phase sampling-revised version. Retrieved from ETH Zürich.
- Mandallaz D (2008) *Sampling techniques for forest inventories*. CRC Press, Boca Raton, Florida, USA.
- Hill A, Massey A (2017) The R package forestinventory: design-based global and small area estimations for multi-phase forest inventories.
- Mandallaz D (2013b) A three-phase sampling extension of the generalized regression estimator with partially exhaustive information. *Canadian Journal of Forest Research* 44(4): 383-388.
- Næsset E (2014) Area-based inventory in norway-from innovation to an operational reality. In Maltamo M, Næsset E, Vauhkonen J (Eds.), *Forestry applications of airborne laser scanning: concepts and case studies*, Dordrecht, Netherlands, pp. 215-240.
- Reich RM, Aguirre Bravo C (2009) Small-area estimation of forest stand structure in Jalisco, Mexico. *Journal of Forestry Research* 20(4): 285.
- Haakana H, Heikkinen J, Katila M, Kangas A (2019a) Efficiency of post-stratification for a large-scale forest inventory-case Finnish NFI. *Annals of Forest Science* 76(1): 9.
- Haakana H, Heikkinen J, Katila M, Kangas A (2019b) Precision of exogenous post-stratification in small-area estimation based on a continuous national forest inventory. *Canadian Journal of Forest Research* 50(4): 359-370.
- Strand GH, Aune Lundberg L (2012) Small-area estimation of land cover statistics by post-stratification of a national area frame survey. *Applied Geography* 32(2): 546-555.
- Kangas A, Rätty M, Korhonen KT, Vauhkonen J, Packalen T (2019) Catering information needs from global to local scales-potential and challenges with national forest inventories. *Forests* 10(9): 800.
- Breidenbach J, Astrup R (2012) Small area estimation of forest attributes in the Norwegian National Forest Inventory. *European Journal of Forest Research* 131(4): 1255-1267.
- Goerndt ME, Monleon VJ, Temesgen H (2011) A comparison of small-area estimation techniques to estimate selected stand attributes using LIDAR-derived auxiliary variables. *Canadian Journal of Forest Research* 41(6): 1189-1201.
- Magnussen S (2015) Arguments for a model-dependent inference? *Forestry: An International Journal of Forest Research* 88(3): 317-325.
- Mauro F, Molina I, García Abril A, Valbuena R, Ayuga Téllez E (2016) Remote sensing estimates and measures of uncertainty for forest variables at different aggregation levels. *Environmetrics* 27(4): 225-238.
- Grafström A, Saarela S, Ene LT (2014) Efficient sampling strategies for forest inventories by spreading the sample in auxiliary space. *Canadian Journal of Forest Research* 44(10): 1156-1164.
- Breidenbach J, Nothdurft A, Kändler G (2010) Comparison of nearest neighbour approaches for small area estimation of tree species-specific forest inventory attributes in central Europe using airborne laser scanner data. *European Journal of Forest Research* 129(5): 833-846.
- Latifi H, Koch B (2012) Evaluation of most similar neighbour and random forest methods for imputing forest inventory variables using

- data from target and auxiliary stands. *International Journal of Remote Sensing* 33(21): 6668-6694.
32. McRoberts RE (2012) Estimating forest attribute parameters for small areas using nearest neighbors techniques. *Forest Ecology and Management* 272: 3-12.
33. Nothdurft A, Saborowski J, Breidenbach J (2009) Spatial prediction of forest stand variables. *European Journal of Forest Research* 128(3): 241-251.
34. Melville G, Stone C (2016) Optimising nearest neighbour information-a simple, efficient sampling strategy for forestry plot imputation using remotely sensed data. *Australian Forestry* 79(3): 217-228.
35. Melville G, Stone C, Rombouts J (2016) Survey designs which maximize efficiency gains in ALS-based forestry plot imputation. *Proceedings of Spatial Accuracy* pp. 1-8.
36. Särndal CE, Swensson B, Wretman JH (1992) *Model assisted survey sampling*: Springer, Verlag, Berlin, Germany.
37. Breidenbach J, Magnussen S, Rahlf J, Astrup R (2018) Unit-level and area-level small area estimation under heteroscedasticity using digital aerial photogrammetry data. *Remote Sensing of Environment* 212: 199-211.
38. Magnussen S, Breidenbach J (2017) Model-dependent forest stand-level inference with and without estimates of stand-effects. *Forestry: An International Journal of Forest Research* 90(5): 675-685.
39. Magnussen S, Mandallaz D, Breidenbach J, Lanz A, Ginzler C (2014) National forest inventories in the service of small area estimation of stem volume. *Canadian Journal of Forest Research* 44(9): 1079-1090.

For possible submissions Click below:

[Submit Article](#)