Ecological Complexity 7 (2010) 208–211

Contents lists available at ScienceDirect

Ecological Complexity

journal homepage: www.elsevier.com/locate/ecocom

Accuracy in population estimation: A methodological consideration

C. Li [a,*], H.J. Barclay [b], H. Hans [a], J. Liu [c], R. Klos [c], G. Carlson [c]

[a] Canadian Wood Fibre Centre, Canadian Forest Service, Natural Resources Canada, 5320 – 122 Street, Edmonton, Alberta, Canada T6H3S5
[b] Pacific Forestry Centre, Canadian Forest Service, 506 Burnside Road, Victoria, British Columbia, Canada
[c] Forestry Branch, Manitoba Conservation, Box 70, 200 Saulteaux Cres., Winnipeg, Manitoba, Canada

ARTICLE INFO

Article history:
Received 9 January 2009
Received in revised form 18 January 2010
Accepted 4 February 2010
Available online 25 February 2010

Keywords:
Population estimation
Accuracy
Value recovery
Modeling
Manitoba

ABSTRACT

Accuracy in population estimation from individual measurement has been traditionally a research focus in both theoretical and applied ecology. In forest sciences, estimation of productivity and value recovery of forest products is essential for decision-making to achieve the goal of sustainable forest management. In this paper, we review the basic structure of data in forest sciences, describe commonly used statistical procedures in obtaining population estimates, and examine the accuracy associated with the forest products value estimation using forest inventory data of Manitoba, Canada. Our results suggested that simplified statistical procedures could bring about a wide range of bias in estimating lumber value recovery at the stand level, and improved understanding of stand structure and its reconstruction through computer simulation could be essential in reducing the bias involved in the estimation.

1. Introduction

Population measurement or estimation has been traditionally one of the central topics in ecological research (Holt, 1987). The accuracy of the estimation is essential in the expression of population measurement. Its importance is primarily from a practitioner's perspective such as understanding the dynamics of ecosystems and making appropriate decisions on resource management. In forest research, a population can be defined as a collection of individual trees in a forest stand, and the variables associated with this population are the estimates such as wood volume that will determine the quantity of potential forest products. With the increasing challenges facing the forest sector, the valuation of forest product options becomes an important issue in determining optimal wood utilization strategies for the purpose of enhancing market competitiveness and maintaining sustainable resources (Sekot, 2007; Guthrie and Kumareswaran, 2009; Li, 2009). Considerable statistical analysis is involved in this valuation. Owing to the fast development in information technology in recent decades, many complicated statistical methods and procedures can be implemented easily given changing computing speed and availability of software packages. Consequently, practitioners can employ a variety of methods and procedures for population estimation with an improved accuracy.

In this paper, we review the basic structure of data in forest research, describe commonly used statistical procedures for obtaining population estimates, and test a hypothesis that there is no significant difference between the estimated values of lumber recovery from different statistical procedures. We show that the hypothesis did not hold in our case study using forest inventory plot data from the boreal shield of central-east Manitoba, Canada. As a result, the statistical procedure chosen could influence the accuracy of value estimation of forest products. Our results suggested that simplified statistical procedures might not always provide confident answers, and the reconstruction of stand structure from the inventory plot data is probably needed to obtain much reliable value estimation, and this can be realized through computer simulations.

2. Materials and methods

2.1. Basic data structure in forest research

Data in forest research can be categorized into groups of spatial and non-spatial data. The spatial data are usually referred to as spatial forest inventory at regional to national scales. The regional forest inventory is polygon-based and often used in the strategic (up to 200 years), tactical (20 years), and operational (ranging from 1 to 5 years) forest management (harvest) planning performed at the levels of the provinces and Forest Management Licenses (FMLs) (Li et al., 2010). Each polygon has its unique ID associated with its geographical locations, landform strata, and administrative structure expressed as region, district, management unit, etc.

* *Corresponding author.* Tel.: +1 780 435 7240; fax: +1 780 435 7359.
E-mail address: cli@nrcan.gc.ca (C. Li).

1476-945X/$ – see front matter.
doi:10.1016/j.ecocom.2010.02.002

Other variables in each polygon include land cover, tree species composition, stand origin, density expressed as crown closure or density classes, site condition, disturbance history, etc. Most current forest inventories are generated from the interpretation of aerial-photos and validated through ground truthing with plot sampling data.

The non-spatial data are usually obtained from field measurements. Inventory sampling plots are a kind of temporary sample plot (TSP) designed for the development of operational forest inventory and they are chosen in either a systematic or a stratified sampling manner. They are one kind of non-spatial sampling plot data in forest research. Timber cruise data also belong to TSP and they are the randomly selected sites under investigation.

The other kind of non-spatial sampling plot data is from the measurement of permanent sample plots (PSP). The main difference between PSP and TSP data is the repeated measurement in PSP but not in TSP. In British Columbia, Canada, PSP are established in natural stands for high quality and long-term local data on growth of the existing forest for a variety of species, sites, and stand conditions. They are a major part of forest growth and yield program, and provide information on rates of growth, mortality, and changes in stand structure. They are measured repeatedly at a certain intervals, such as 5 or 10 years.

The data structures of PSP and TSP are the same or very similar. In each plot, only those trees with a diameter at breast height (*DBH*) larger than a certain standard are measured, as is often the case, i.e., not all the trees in the plot are measured. For all the measured trees, species, age, *DBH*, and height (*H*) are recorded. Other variables may also be recorded such as site condition, disturbance, etc.

This data structure can provide within plot species composition, distributions of diameter and height, relationship between height and diameter, tree density or stocking, age, total volume and merchantable volume per unit area. They are the essential information for modeling tree-based growth, describing stand structure, and calibrating and validating growth and yield models that are unit area-based for inventory projection.

2.2. Case study area and forest inventory

The Forest Management License (FML) #1 of Manitoba, Canada, located in central-eastern Manitoba has been chosen as our study area (see http://www.gov.mb.ca/conservation/forestry/forest-practices/companies/fml1.html). The total size of this FML is 889,471 ha, and about two-thirds of it is considered to be productive and potentially productive forest land, under the management of the Pine Falls Operations, Tembec Industries Inc. Fig. 1 shows the location of our study area.

The operational forest inventory was generated based on the 1997 aerial-photos, and took 5 years to complete (Manitoba Conservation, 2006). The sizes of the polygons can be from less than 1 ha to as large as 390 ha for forested lands. There are 163,699 polygons with a size larger than 1 ha, and the mean size of the polygons is about 6.9 ha. During the 5 years more than 700 polygons were selected for field plot sampling and they contain three 100 m^2 plots per polygon. In each plot all the trees with a *DBH* larger than 7.1 cm were measured and that resulted in about 26,000 trees in total for developing this forest inventory. *DBH* and *H* are among the measured variables. Other variables were also recorded in the field sampling such as site conditions, understory vegetation, and disturbance type and history, etc. Through enormous investment of time and work, this forest inventory provides the best information on forest conditions to support the provincial forest management agency and forest industry to make sound strategic and operational management planning for the region. Since April 2009, the Nopiming Provincial Park has been excluded from this FML for harvest thus the total size of the FML is reduced.

A test plot (located at easting 750,156 and northing 5,646,985 of the UTM Zone 14) dataset was used in this study first for a detailed examination of whether the standard procedure of estimating stand volume can be directly used in the value recovery estimate at the stand level. The plot is within an almost pure black spruce (*Picea mariana* (Mill.) B.S.P.) stand with an age of 68 years. The plot data were collected on April 16, 2002, and contained the measurements of 116 individual black spruce trees. Tree taper was calculated from the diameter inside the bark (*DIB*) between 15 cm above ground and the tree height at 5 cm *DIB* from the top based on the method of Klos et al. (2007).

The same analytical procedure was applied to all the polygon plots in the study area with a black spruce component of at least 10% (499 plots in total) to investigate the potential influence of the standard procedure on the accuracy of lumber value recovery estimates.

2.3. Methods of valuation of lumber recovery

A general approach of estimating stand level forest wood volume from measurements of individual trees is to develop individual tree-based volume equations first and estimate stand level volume according to the stand structure through using these volume equations. However, it is usually simplified in operation to a simple procedure as (1) obtain the mean values of measured *DBH* and *H* of all sampled trees for a given tree species under investigation, and (2) use these mean values as representative of all trees in the stand and multiplied by the total number of trees of the stand to calculate the mean wood volume per unit area using the volume equations. This procedure has become more or less standard in the forest inventory construction and is used in many forest growth and yield modeling practices.

With the shift of forest management paradigm from being volume-based to value-based, an assumption that value creation potential is proportional to wood volume may have been used in

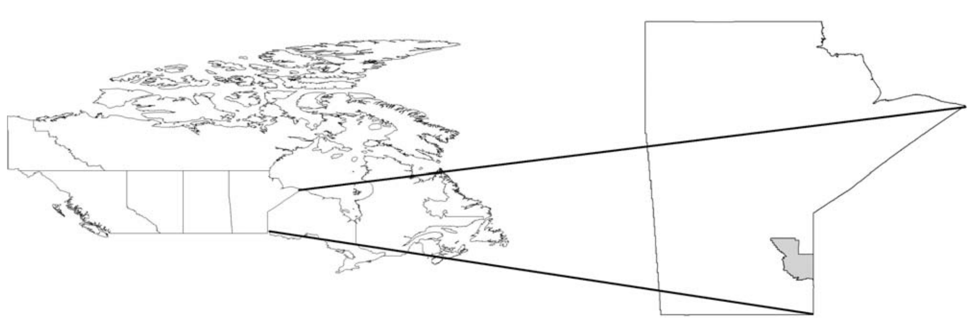

Fig. 1. Study area location indicated by the grey coloured area (data as March 2009).