Design-based treatment of missing data in forest inventories using canopy heights from aerial laser scanning

Piermaria Corona, Gherardo Chirici, Sara Franceschi, Daniela Maffei, Marzia Marcheselli, Caterina Pisani, and Lorenzo Fattorini

Abstract: Nonresponse is often a problem in forest inventories. It may arise when sample plots are inaccessible because of hazardous terrain. To account for this problem, the use of nonresponse calibration weighting is considered in a complete design-based framework, i.e., both nonresponse and survey variables are viewed as fixed characteristics of the plots. Information derived from remotely sensed data is exploited to compensate for the missing plots. Calibration is performed adopting canopy height from airborne laser scanning as an auxiliary variable. Conditions for approximate unbiasedness of the calibration estimator in two-phase inventories are derived, and some estimators of the sampling variance are proposed. Results from one-phase inventories are achieved as a particular case. Dummy variables are introduced in the presence of different forest types. Monte Carlo results support the validity of the procedure. An application to a forest survey carried out in Central Italy is performed.

Key words: auxiliary information, nonresponse calibration weighting, unbiasedness, variance estimation, Monte Carlo study.

Résumé : La non-réponse est souvent un problème en inventaire forestier. Cela se produit lorsque des placettes-échantillons sont inaccessibles en raison de conditions de terrain périlleuses. Pour tenir compte de ce problème, l’utilisation de la pondération par calage fait partie d’un plan d’échantillonnage complet, c’est-à-dire qu’à la fois les variables de non-réponse et les variables d’inventaire sont considérées comme des caractéristiques fixes des placettes. L’information qui provient de données de télédétection est utilisée afin de compenser pour les placettes manquantes. L’étalonnage est fait en utilisant la hauteur du couvert forestier, obtenue avec un laser à balayage aéroporté, comme variable auxiliaire. Les conditions pour une absence approximative de biais de l’estimateur par calage dans un échantillonnage à deux phases sont démontrées et certains estimateurs de la variance d’échantillonnage sont proposés. Les résultats d’un échantillonnage à une seule phase sont considérés comme un cas particulier. Des variables nominales sont introduites pour représenter différents types forestiers. Les résultats d’une simulation Monte-Carlo démontrent la validité de la méthode. La méthode est appliquée à un inventaire forestier réalisé en Italie centrale.

Mots-clés : information auxiliaire, nonresponse calibration weighting, unbiasedness, variance estimation, Monte Carlo study.

1. Introduction

Many forest surveys performed over large scales involve two phases of sampling. Relevant examples are the national forest inventories (NFI) of Canada (Gillis 2001), the USA (Scott et al. 2004), the former Soviet Union (Gäbler and Schadauer 2007), New Zealand (Stephens et al. 2012), and Romania and Switzerland (Tomppo et al. 2010). Other countries such as Austria, Belgium, Brazil, Poland, and Sweden (Tomppo et al. 2010) adopt one-phase inventories. In few cases (e.g., the Italian NFI), a third phase is adopted to further reduce survey costs and time (Fattorini et al. 2006).

Usually, satellite imagery or aerial photo-based information is collected by means of an intensive first-phase sampling, while the second phase is performed by ground inspections to combine aerial and field data (Corona et al. 2011). In the first phase, the area is often partitioned into regular polygons of the same size, and points are randomly or systematically selected in each polygon. In the second phase, a sample of first-phase points is selected in accordance with a probabilistic sampling scheme. Second-phase points are visited on the ground, and several variables of qualitative and quantitative nature are recorded. In particular, biophysical attributes are recorded within plots of prefixed size centred at these points. One-phase inventories can be viewed as particular cases of two-phase ones when all of the first-phase points are selected in the second phase. The design-based treatment of two-phase inventories, viewing the survey variable as a fixed characteristic of plots, appears in many familiar textbooks (e.g de Vries 1986; Schreuder et al. 1993; Gregoire and Valentine 2008; Mandallaz 2008). Unbiased or nearly unbiased estimators for the total of biophysical attributes are provided together with the expressions of their variances and of the corresponding estimators.

Nonresponse occurs in forest inventories for two main reasons: (i) plots selected in the second phase are located in difficult terrains and cannot be reached by survey crews or, even if reached, the steep slope of the terrain does not allow the recording activities...