

Diss. ETH No. 23025

Multiphase estimation procedures for forest inventories under the design-based Monte Carlo approach

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH Zurich
(Dr. sc. ETH Zurich)

presented by

ALEXANDER FRANCIS MASSEY

MSc. Statistics, ETH Zurich

born 18 May 1983

citizen of the United States of America and Italy

accepted on the recommendation of

PD Dr. D. Mandallaz, supervisor, examiner

Prof. Dr. H. R. Heinimann, co-examiner

Prof. Dr. J. Saborowski, co-examiner

Dr. A. Lanz, co-examiner

2015

Summary

This cumulative dissertation explores methods for improving estimation for forest inventories by incorporating auxiliary data such as those derived from remote sensing. Results are compiled from three articles published in the *Canadian Journal of Forest Research* (see Massey et al. (2014); Massey and Mandallaz (2015a,b)). The proposed methods follow exclusively the design-based Monte Carlo setup with a focus on two- and three-phase regression estimators. The structure of this dissertation can be divided into four main parts: 1) the presentation of the general model-assisted framework in the Monte Carlo approach, 2) an application example for the estimation of state in the Swiss National Forest Inventory (NFI), 3) an extension of regression estimation to nonparametric models such as k-nearest neighbors (kNN), and 4) a reformulation of the proposed approach for the estimation of net change.

Although the methods are applicable in a variety of situations, the majority of the case studies are presented in the context of annual sampling designs. This is because the practical objective of this research was to accommodate the 2009 transition of the Swiss NFI from a periodic to an annual design where only one ninth of the overall sample of permanent plots is measured every year. The transition was primarily motivated by logistical reasons and led to a reduction in sample size that consequently increased the variance when using the standard simple random sampling estimator.

The major result was for the estimation of state where it was demonstrated that a three-phase classical regression estimator could dramatically reduce the increase in variance associated with the transition to the annual design by integrating canopy height information with growth-updated previous measurement information. While kNN regression estimators could probably have also been used effectively in place of the classical version, the most common analytical variance estimator risks potentially severe underestimation of the true variance as is demonstrated in an artificial simulation example. A bootstrap procedure is proposed that avoids this underestimation. Similar improvements are also possible for the estimation of net change, but there is a tradeoff between the estimator's gain in precision and its flexibility to estimate change over any arbitrary time frame.

Zusammenfassung

Die vorliegende kumulative Dissertation befasst sich mit der Verbesserung von Schätzmethoden in der Waldinventur durch die Einbeziehung von Hilfsinformation, u.a. Fernerkundungsdaten. Die Ergebnisse der Arbeit wurden in drei Artikeln im *Canadian Journal of Forest Research* publiziert (Massey et al. (2014); Massey and Mandallaz (2015a,b)). Die vorgeschlagenen zwei- und dreiphasigen Schätzmethoden folgen dabei ausschliesslich dem design-basierten Ansatz. Der Aufbau der Dissertation kann in vier Teile untergliedert werden: 1) eine generelle Einführung in die Methodik des design-basierten Ansatzes, 2) ein Anwendungsbeispiel für die Zustandsschätzung in der schweizerischen Landeswaldinventur, 3) eine Erweiterung der Regressionsschätzung auf nichtparametrische Modelle (u.a. k-nearest neighbors Methoden), und 4) eine Neuformulierung der vorgeschlagenen Methoden für die Schätzung von Zustandsveränderungen.

In den vorgestellten Fallstudien werden die Schätzmethoden meist auf den Fall einer jährlich wiederkehrenden Waldinventur angewendet. Ein Grund hierfür ist, dass es das praxisorientierte Ziel dieser Arbeit war, für die 2009 begonnene Umstellung der schweizerischen Landeswaldinventur von einer periodisch- zu einer jährlich-wiederkehrenden Inventur anwendbare Schätzmethoden zu entwickeln. Durch diese Umstellung, welche durch logistische Gründe motiviert war, werden zukünftig jährlich nur 1/9 aller Stichproben im Land erhoben. Diese Reduzierung der Stichprobengrösse führt folglich zu einer Vergrösserung der Varianz bei Anwendung eines herkömmlichen Schätzers unter der Annahme von zufällig verteilten Stichprobenpunkten.

Es konnte gezeigt werden, dass durch die Anwendung eines dreiphasigen klassischen Regressionsschätzers unter Einbeziehung von Kronenhöheninformationen und einer Aktualisierung früherer Messungen die Varianzerhöhung, welche durch die Umstellung der periodischen zur jährlichen Inventur bedingt ist, erheblich reduziert werden konnte. Während die Anwendung von kNN-basierten Regressionsschätzern eine mögliche Alternative zum klassischen Regressionsschätzer darstellt, konnte in einer Simulationstudie gezeigt werden, dass in diesem Fall ein hohes Risiko besteht, die tatsächliche

Varianz deutlich zu unterschätzen. Um diese Unterschätzung zu verhindern wurde eine entsprechende Bootstrap-Methode vorgestellt. Vergleichbare Verbesserungen sind auch hinsichtlich der Schätzung von Zustandsänderungen möglich. Allerdings zeigte sich hier ein Trade-Off zwischen der realisierbaren Genauigkeitserhöhung der Schätzungen und der Flexibilität, Veränderungen über beliebige Zeiträume zu schätzen.