

Zeitschrift: Schweizerische Zeitschrift für Forstwesen = Swiss forestry journal = Journal forestier suisse

Herausgeber: Schweizerischer Forstverein

Band: 169 (2018)

Heft: 4

Artikel: Neue Hilfsmittel zur Anwendung zweiphasiger Stichprobenverfahren in der Waldinventurpraxis

Autor: Gabriel, Andreas / Hill, Andreas / Breschan, Jochen

DOI: <https://doi.org/10.5169/seals-1097391>

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. [Mehr erfahren](#)

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. [En savoir plus](#)

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. [Find out more](#)

Download PDF: 02.05.2026

ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>

Neue Hilfsmittel zur Anwendung zweiphasiger Stichprobenverfahren in der Waldinventurpraxis

Andreas Gabriel Eidgenössische Technische Hochschule ETH Zürich (CH)*
Andreas Hill Eidgenössische Technische Hochschule ETH Zürich (CH)
Jochen Breschan Eidgenössische Technische Hochschule ETH Zürich (CH)

Neue Hilfsmittel zur Anwendung zweiphasiger Stichprobenverfahren in der Waldinventurpraxis

Um zuverlässige Informationen über den Zustand und die Entwicklung des Waldes zu erhalten, werden Waldinventuren durchgeführt. In der Regel werden hierfür Stichprobenkonzepte angewendet. Bei diesen ist die Anzahl terrestrischer Probepunkte so gewählt, dass die wichtigsten forstlichen Kennzahlen auf der jeweiligen räumlichen Zielebene (Bund, Kanton, Betrieb) mit der geforderten Genauigkeit geschätzt werden können. Um für ein Teilgebiet genaue Informationen zu erhalten, kann die Ergänzung der terrestrischen Inventur mit Fernerkundungsdaten eine kostengünstige Alternative zur Erhöhung der terrestrischen Stichprobendichte sein. Für solche zweiphasige Inventuren wurden an der ETH Zürich Design-basierte Schätzverfahren (sog. Kleingebietsschätzer) entwickelt und als frei verfügbares Softwarepaket *forestinventory* in der Statistiksoftware R implementiert. Im Artikel wird das Grundprinzip zweiphasiger Kleingebietsschätzungen erläutert, und es werden die Arbeitsschritte sowie die Effizienz dieser Methode an einem Fallbeispiel in der Region Embrach (Kanton Zürich) aufgezeigt. Zielgrösse ist die Schätzung der mittleren Vorratsdichte in 22 Kleingebieten des Untersuchungsgebiets (503 ha). Die wichtigsten Arbeitsschritte sind (A) die Erweiterung des einphasigen Inventurdesigns (terrestrisches Stichprobenetz) zu einem zweiphasigen, (B) die Ableitung aussagekräftiger erklärender Variablen aus Fernerkundungsquellen (hier: mittlere Vegetationshöhe, Standardabweichung der Vegetationshöhen und Laubholzanteil), (C) die Verknüpfung der erklärenden Variablen mit der terrestrischen Information durch Formulieren eines Prognosemodells mithilfe der linearen Regression und (D) die Schätzung der Zielgrössen (hier: Vorratsdichte) mithilfe von *forestinventory*. Trotz mässiger Modellgenauigkeit ($\hat{R}^2 = 0.53$) konnte die Varianz der Schätzungen in den 22 Kleingebieten im Mittel um 40% reduziert werden.

Keywords: small area estimation, regression estimators, design-based, methodology, case study
doi: 10.3188/szf.2018.0210

* Universitätstrasse 16, CH-8092 Zürich, E-Mail andreas.gabriel@usys.ethz.ch

Das Ziel von Waldinventuren ist es, zuverlässige Informationen über den Zustand und die Entwicklung des Waldes zu beschaffen. Da Vollkluppierungen ganzer Waldgebiete aufgrund des hohen Aufwandes meistens unmöglich sind, kommen in der Praxis Stichprobenkonzepte zum Einsatz. Bei diesen Verfahren wird an zufällig im Waldgebiet verteilten Stichprobenpunkten die lokale Hektardichte von Holzvorrat, Grundfläche und vielen weiteren Attribute erfasst. Diese Daten werden anschliessend verwendet, um daraus Mittelwerte und Totale für die gesamte Waldfläche zu schätzen, beispielsweise den mittleren oder den gesamten Holzvorrat. Schweizweit stehen hierzu die Daten des Landesforstinventars (LFI) zur Verfügung. Einige Kantone verwenden zusätzlich verdichtete Inventurnetze, um durch eine grössere Anzahl Probepunkte genauere Aussagen auf kantonaler Ebene zu erhal-

ten. Auf Forstbetriebsebene werden in der Schweiz häufig nochmals dichtere Inventurnetze verwendet, meist in der Form von Kontrollstichproben nach der Methode von Schmid-Haas (Schmid-Haas et al 1993).

Der Stichprobenumfang (d.h. die Anzahl terrestrischer Probepunkte) ist bei einer Inventur im Vorhinein so gewählt, dass die wichtigsten forstlichen Kennzahlen auf der jeweiligen räumlichen Zielebene (Bund, Kanton, Betrieb) mit der geforderten Genauigkeit geschätzt werden können. Möchte man allerdings Schätzungen für Teilgebiete mit wenigen Stichprobenpunkten berechnen, kann dies aufgrund des geringen Stichprobenumfangs zu unzureichenden Schätzgenauigkeiten führen. Solche Teilgebiete werden im Inventurjargon auch als Kleingebiete bezeichnet. Zweiphasige Stichprobeninventuren bieten die Möglichkeit, die Schätzgenauigkeit für diese räumlichen Ebenen ohne Erhöhung der terrestri-

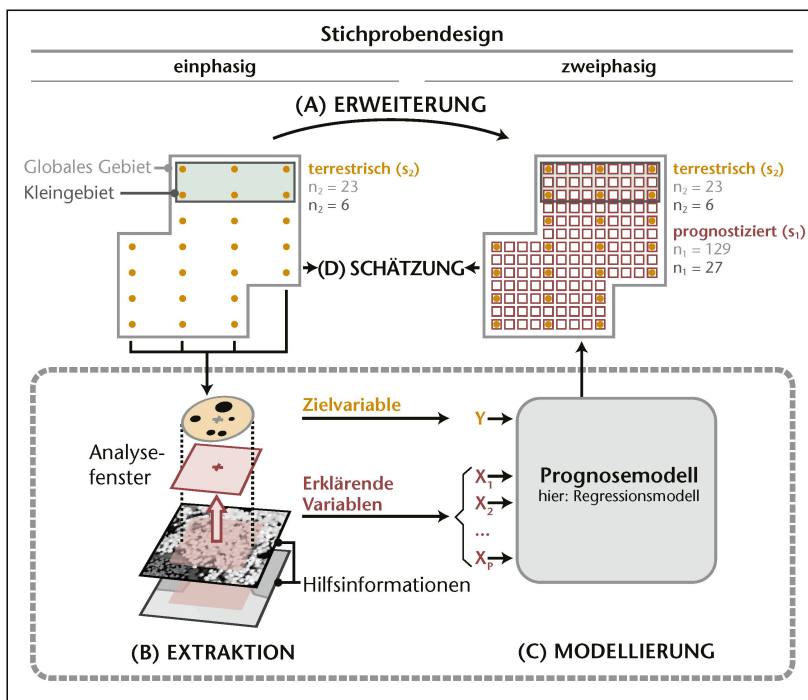


Abb 1 Vorgehen bei Schätzungen auf Basis von zweiphasigen Inventuren. n : Stichprobenumfang, s : Stichprobe.

schen Aufnahmen zu verbessern. Das Prinzip besteht darin, Zielvariablen wie Grundfläche oder Vorrat an zusätzlichen Stichprobenpunkten, an welchen keine terrestrischen Daten vorliegen, mit statistischen Prognosemodellen zu schätzen. Zahlreiche internationale und nationale Studien haben das Potenzial dieser Verfahren in der Vergangenheit bereits aufgezeigt. Beispiele für die Schweiz finden sich u.a. in Steinmann et al (2013), Magnussen et al (2014) und Mandallaz et al (2013).

An der ETH Zürich widmete man sich in den letzten 30 Jahren der Entwicklung, Anwendung und Bereitstellung von Waldinventurmethode. Dabei entstand zuletzt das Softwarepaket forestinventory zur Auswertung zweiphasiger Stichprobeninventuren (R-Package; Hill & Massey 2017) in der statistischen Software R (R Development Core Team 2017). Dieses wurde im Juni 2017 auch Waldplanern der ganzen Schweiz an einem Fortbildungstag vorgestellt. Dieser Artikel erläutert, wie zweiphasige Stichprobeninventuren durchgeführt werden, und illustriert das Vorgehen anhand der Anwendung von forestinventory in einem Fallbeispiel. Bei Letzterem werden auch Möglichkeiten zur Effizienzabschätzung gezeigt. Das Fallbeispiel mit Anleitung steht in der Research Collection der ETH kostenlos zur Verfügung (Breschan et al 2017).

Methode

Mit dem Ziel, ein intuitives Verständnis für die Schätzmethoden und die ihnen zugrunde liegenden Annahmen zu ermöglichen, werden im Folgen-

den die wichtigsten Arbeitsschritte zweiphasiger Stichprobeninventuren und die damit verbundenen Grundprinzipien beschrieben. Die Arbeitsschritte (Abbildung 1) umfassen (A) die Erweiterung des terrestrischen Inventurnetzes (einphasiges Inventurdesign) zu einem zweiphasigen Inventurdesign, (B) die Ableitung von Variablen aus verfügbaren Hilfsinformationen und (C) die Formulierung eines Prognosemodells mit einfachen Techniken der linearen Regression. Diese Schritte erlauben schliesslich (D) die Anwendung zweiphasiger Schätzverfahren zur Ermittlung von Mittelwerten und Totalen.

Erweiterung des einphasigen Stichprobendesigns zu einem zweiphasigen

Bei zweiphasigen Inventuren werden Informationen aus zwei ineinandergeschichteten Stichproben genutzt (Abbildung 1, zweiphasig). Dabei wird ausgehend von der terrestrischen Inventur s_2 (Abbildung 1, gelbe Punkte) ein deutlich dichteres Inventurnetz s_1 (Abbildung 1, rote Quadrate) generiert, welches der künstlichen Erhöhung des Stichprobenumfangs dient. Während an den terrestrischen Probepunkten die lokale Dichte $Y(x)$ (z.B. mittlerer Holzvorrat pro Hektare) ermittelt wird, wird mit der Anwendung eines statistischen Modelles an allen Probepunkten der grossen Stichprobe s_1 eine Prognose $\hat{Y}(x)$ der lokalen Dichte generiert. Das Prognosemodell wird dabei mit den terrestrischen Inventurdaten (Stichprobe s_2) kalibriert und auf die an allen Probepunkten (Stichprobe s_1) vorhandenen Hilfsinformationen angewendet. Mit speziell für zweiphasige Inventuren entwickelten Schätzern werden daraus die Mittelwerte und Totale in den Kleingebieten und der Schätzfehler dieser Kenngrössen ermittelt. Die Idee setzt voraus, dass die Hilfsinformationen die terrestrisch ermittelten lokalen Dichten gut erklären können bzw. mit diesen korrelieren. Den terrestrisch gemessenen Daten kommt dabei eine Doppelfunktion zu: Sie dienen sowohl der Kalibrierung des Prognosemodells als auch zur Ableitung der Schätzfehler.

Extraktion der Hilfsinformationen

Als Hilfsinformationen können prinzipiell jegliche Daten dienen, die über das Inventurgebiet in hoher Anzahl oder gar flächendeckend verfügbar sind und Informationen enthalten, die für die Prognose $\hat{Y}(x)$ der lokalen Dichte $Y(x)$ an jedem Inventurpunkt x nützlich sind. Meist werden diese Hilfsinformationen aus Fernerkundungsdaten wie Höhenmodellen, Luftbildern oder Satellitendaten abgeleitet (Ginzler & Waser 2017). Auch auf Experteneinschätzungen basierende Informationen wie Bestandskarten lassen sich einsetzen. Um die Hilfsinformationen für das Prognosemodell nutzbar zu machen, müssen sie an jedem Inventurpunkt extrahiert und oftmals in erklärende Variablen umgewandelt werden. Dies geschieht im Analysefenster, das

um den jeweiligen Stichprobenpunkt zentriert wird. Innerhalb dieses Analysefensters wird dann die Rohinformation (z.B. die Entwicklungsstufe aus einer Bestandeskarte) oder die bereits abgeleitete Information (z.B. die mittlere Vegetationshöhe) extrahiert (Abbildung 1; B). Für die Herleitung eines geeigneten Prognosemodelles reicht es, wenn die erklärenden Variablen an den Probepunkten der terrestrischen Stichprobe s_2 abgeleitet werden. Ist das Prognosemodell gefunden, sind die im Modell verwendeten erklärenden Variablen jedoch an allen Probepunkten der Stichprobe s_1 zu berechnen.

Grundsätzlich sind die Form und die Grösse des Analysefensters frei wählbar. Die Vermutung liegt nahe, dass auf die Ausdehnung terrestrischer Probepunkte zugeschnittene Analysefenster zu einem besonders guten Zusammenhang zwischen abgeleiteter Hilfsinformation und lokaler Dichte führen. Dieser Zusammenhang hängt jedoch auch von der Art und der räumlichen Auflösung der Hilfsinformation ab (Grafström et al 2017), weshalb die Evaluation verschiedener Analysefenster sinnvoll ist.

Formulierung des Prognosemodells – Regressionsmodellierung

Grundsätzlich stehen zur Prognose von Zielvariablen mittels erklärender Variablen eine Vielzahl von statistischen Vorhersagemodellen zur Verfügung (James et al 2013). Eine häufig eingesetzte Modellierungsmethode stellt die gewöhnliche lineare Regression (OLS-Regression) dar.

Bei der linearen Regression wird eine Zielvariable Y durch die lineare Kombination von p erklärenden Variablen (X_1, X_2, \dots, X_p) beschrieben. Dabei wird der Einfluss jeder erklärenden Variablen X auf die Zielvariable Y mittels eines Regressionskoeffizienten β bestimmt, den man sich als Mass für die Steigung der Geraden vorstellen kann. Sobald anhand von verfügbaren Y - und X -Daten der Zusammenhang für jede X -Variable mit dem Verfahren der kleinsten Quadrate bestimmt ist, lässt sich der Y -Wert für beliebige Realisationen der X -Variablen schätzen. Der Zusammenhang kann dabei für eine einzelne (einfache Regression) oder mehrere erklärende Variablen (multiple Regression) formuliert werden (Modell 1 in Tabelle 1). Letztere sind um Wechselwir-

kungsterme erweiterbar, welche die Berücksichtigung von Abhängigkeiten zwischen Variablen ermöglichen (Modell 2 in Tabelle 1). Die erklärenden Variablen X können dabei als kontinuierliche Variablen, kategoriale Variablen oder als Kombination der beiden Formen vorliegen.

Kontinuierliche Variablen zeichnen sich dadurch aus, dass zwischen zwei beliebigen Skalenwerten unendlich viele Ausprägungen möglich sind. Die mittlere Vegetationshöhe (X_1) und der mittlere Brusthöhendurchmesser (X_2) sind Beispiele kontinuierlicher Variablen, die für die Modellierung der Vorratsdichte (Y) genutzt werden können. Ihr Einfluss auf Y lässt sich analog zu Modell 1 (Tabelle 1) so modellieren, als ob die Ausprägung von X_1 keinen Einfluss auf den Zusammenhang von X_2 und der lokalen Dichte hätte. Die Prognose ergibt sich dabei aus der Summe aller individuellen Einflüsse von X_1, \dots, X_p . Oftmals sind jedoch auch Wechselwirkungen zwischen Variablen bekannt, beispielsweise dass sich der Vorrat bei identischer mittlerer Vegetationshöhe in Abhängigkeit von der Durchmesserklassenverteilung des Bestandes verändert (von Gadow 2005). Dieser Einfluss auf Y lässt sich analog zu Modell 2 (Tabelle 1) berücksichtigen.

Kategoriale Variablen (auch Gruppenvariablen genannt) bestehen aus mehreren Kategorien. Wiederum vom Beispiel Vorratsdichte ausgehend wird vermutet, dass diese Zielvariable innerhalb der Kategorien weniger streut als zwischen diesen. Die Entwicklungsstufen einer Bestandeskarte sind ein Beispiel hierfür. Als einfache Regression formuliert, lässt sich die Zielvariable in Abhängigkeit von der Entwicklungsstufe schätzen, wobei die Schätzung dem Mittelwert der Zielvariablen in der jeweiligen Entwicklungsstufe entspricht. Bei Verwendung mehrerer kategorialer Variablen ergibt sich die Vorratsdichte in Funktion der Kategoriekombination. Beispielsweise kann die Entwicklungsstufenkategorie mit der Baumartenkategorie kombiniert werden, wobei auch Wechselwirkungen erlaubt sind. In diesem Fall würde die pro Entwicklungsstufe prognostizierte Vorratsdichte je nach Baumart variieren. Die Verwendung kategorialer Variablen für Modelle in zweiphasigen Regressionssschätzern wird auch als Post-Stratifizierung bezeichnet.

Modell	R-Notation	Mathematische Notation	Bedeutung
1) ohne Wechselwirkung	$Y \sim X_1 + X_2 + \dots + X_p$	$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \epsilon$	Der Wert von Y ist eine lineare Kombination von X_1, X_2, \dots, X_p . Der Zusammenhang zwischen Y und X_1 wird dabei so modelliert, als ob X_1 unabhängig von den andern Werten X_p wäre (und umkehrt).
2) mit Wechselwirkung	$Y \sim X_i \times X_j$	$Y = \beta_0 + \beta_i X_i + \beta_j X_j + \beta_{ij} X_i X_j$	Der Wert von Y ist eine lineare Kombination von X_i und X_j ; der Zusammenhang zwischen Y und X_i ist dabei abhängig vom Wert von X_j , derjenige zwischen Y und X_j vom Wert X_i , wobei ij eine beliebige Kombination erklärender Variablen darstellt.

Tab 1 Spezifizierung eines linearen Regressionsmodelles ohne (1) und mit Wechselwirkungsterm (2).

Ein Beispiel für die Kombination von kontinuierlichen und kategorialen Variablen ist die Modellierung der Vorratsdichte mithilfe der mittleren Vegetationshöhe und der Wuchsregionen. Werden Wechselwirkungen formuliert, dann sagt das resultierende Modell bei gleicher mittlerer Vegetationshöhe je nach Wuchsregion unterschiedliche Vorratsdichten voraus. Anstatt für beispielsweise fünf Wuchsregionen individuelle Modelle zu bestimmen, lassen sich diese somit als kategoriale Variable in einem Regressionsmodell berücksichtigen. In diesem Fall spricht man von einem Spezialfall der Post-Stratifizierung («double sampling for regression within post-strata»).

Schätzverfahren

Die an der ETH Zürich für die Waldinventur entwickelten Schätzer, d.h. die mathematischen Formeln zur Ermittlung von Mittelwerten und Totalen bei zweiphasigen Stichprobenverfahren, gehören zur Gruppe der Design-basierten Schätzer. Sie beruhen auf dem Gedankenexperiment, dass eine Schätzung bei gleicher Datengrundlage (gleicher Wald, gleiche Hilfsinformation), aber variierender Lage der zufällig verteilten Stichprobenpunkte unendlich oft wiederholt würde. Nur Design-unverzerrte Schätzer gewährleisten, dass im Mittel aller theoretisch möglichen Schätzungen der wahre Wert resultiert. Aus dieser Eigenschaft folgt beispielsweise für das 95%-Konfidenzintervall, dass bei imaginärer 100-facher Ausführung einer Inventur bei gleicher Datengrundlage 95 der berechneten 100 Konfidenzintervalle den wahren Wert enthalten. Die Wahrscheinlichkeit ist also hoch, dass Design-unverzerrte Schätzer für die in der Praxis nur einmal ausgeführte Inventur ein Konfidenzintervall liefern, das den wahren Wert enthält. Im Gegensatz dazu gewährleisten potenziell Design-verzerrte Schätzer diese Eigenschaft nicht zwingend, und damit ist auch die Überdeckungsrate durch die Konfidenzintervalle nicht zwingend gegeben.

Für die Auswertung zweiphasiger Stichprobeninventuren werden sogenannte globale Schätzer und Kleingebietsschätzer vorgestellt. Diese sind in Mandallaz (2008), Mandallaz (2013) und Mandallaz et al (2013) sowie Hill & Massey (2017) detailliert beschrieben.

Globale Schätzer werden verwendet, um eine Schätzung für das gesamte betrachtete Inventurgebiet zu erzielen. Kleingebietsschätzer hingegen werden dann verwendet, wenn Aussagen gemacht werden sollen zu einem Teilgebiet, in dem nur wenige terrestrische Stichprobenpunkte liegen. Kleingebietsschätzer kommen also nicht zwingend in flächenmässig kleinen Gebieten zur Anwendung. Vielmehr ist der Begriff Ausdruck dafür, dass nur wenige terrestrische Daten im Teilgebiet liegen. Da sich auf Grundlage einer solch geringen Datenlage eine in-

dividuelle Modellierung des Teilgebietes nicht anbietet, besteht die Idee nun darin, das Modell «global» anhand der gesamten terrestrischen Inventur zu kalibrieren und es dann auf das Kleingebiet anzuwenden. Wo terrestrische Probenpunkte im Kleingebiet vorhanden sind, kann die potenziell systematische Abweichung des global bestimmten Modelles durch das Einbeziehen der Modellabweichungen (Residuen) zwischen den Prognosen und den im Kleingebiet terrestrisch erfassten Daten korrigiert werden. Solche Schätzer sind «Design-unverzerrt». Falls in einem Kleingebiet jedoch keine terrestrischen Daten vorliegen, kann man das globale Modell nur «blind» anwenden, denn allfällige Modellverzerrungen lassen sich nicht korrigieren. Schätzer für solche Fälle werden «synthetisch» genannt. Sie liefern meist sehr kleine Konfidenzintervalle, weil die Modellgenauigkeiten bei der Herleitung der Schätzgenauigkeit nicht berücksichtigt werden.

Datenvorbereitung für zweiphasige Schätzverfahren mit forestinventory

Für die Durchführung zweiphasiger Schätzungen mit forestinventory müssen eine Stichprobentabelle für das erweiterte Stichprobennetz erzeugt und ein Regressionsmodell formuliert werden (Abbildung 2). Die Stichprobentabelle enthält dabei folgende Daten:

- die für jeden Probenpunkt aus den Hilfsinformationen extrahierten erklärenden Variablen,
- die Bezeichnung des Kleingebiets, dem der Probenpunkt angehört,
- die Angabe, ob der Probenpunkt auch terrestrisch erfasst wurde und, wenn ja, den terrestrisch bestimmten Wert der Zielvariable.

Die Stichprobentabelle umfasst damit alle Informationen, die für zweiphasige Schätzungen notwendig sind. Gleichzeitig kann sie als Grundlage zur Identifikation eines Regressionsmodells dienen, falls ein geeigneter Regressionsterm (d.h. die Kombination erklärender Variablen) nicht bereits aus früheren Auswertungen bekannt ist. Mit Übergabe der Stichprobentabelle und des Regressionsterms an forestinventory können die Zielgrößen geschätzt werden. Dabei stehen verschiedene Design-basierte Schätzer sowohl für globale Schätzungen als auch für Kleingebietsschätzungen zur Verfügung. Die Koeffizienten des Regressionsmodells werden automatisch anhand des von den Nutzern angegebenen Regressionsterms und der Stichprobentabelle kalibriert.

Fallstudie

Die beschriebenen Konzepte und Arbeitsschritte werden im Folgenden auf ein Untersuchungsgebiet in Embrach (Kanton Zürich) angewendet. Dabei sollen für 22 Kleingebiete der mittlere Holzvorrat

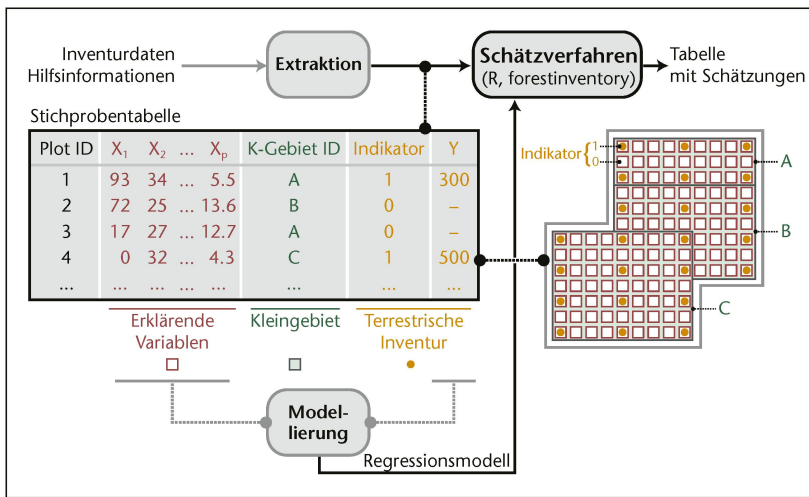


Abb 2 Workflow für zweiphasige Schätzungen. Jede Zeile der Stichprobentabelle entspricht einem Stichprobenpunkt der grossen Stichprobe s_1 . Die erklärenden Variablen (X_1, \dots, X_p) sind als einzelne Attribute abgelegt. Terrestrische Stichprobenpunkte werden zusätzlich mit den Attributen Indikator (1 [für vorhanden]) und Y (lokale Dichte) beschrieben.

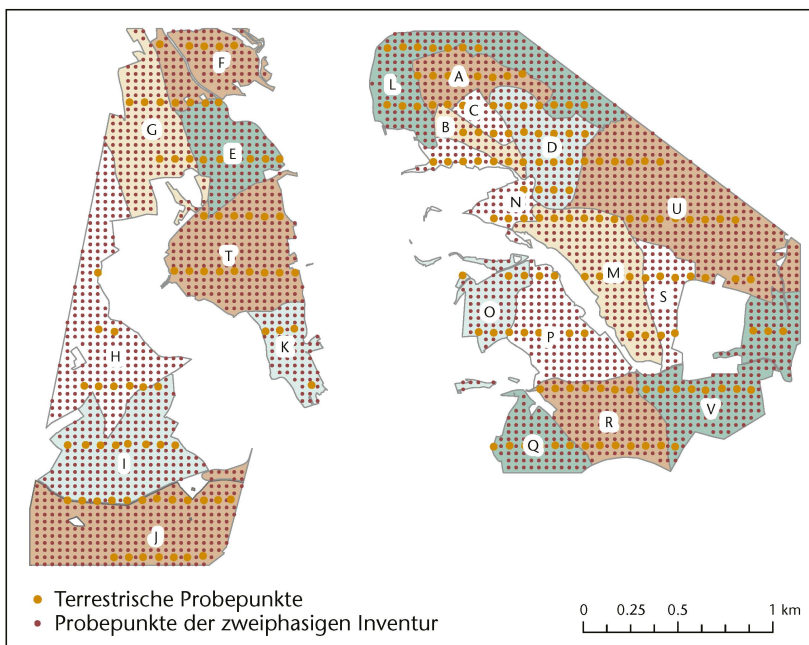


Abb 3 Kleingebiete in Embrach, terrestrisches Stichprobennetz und erweitertes zweiphasiges Stichprobennetz.

mit zwei verschiedenen zweiphasigen Kleingebietschätzern berechnet und die Ergebnisse mit jenen des herkömmlichen, einphasigen Schätzverfahrens unter ausschliesslicher Benutzung der terrestrischen Daten verglichen werden.

Datengrundlage

Das Untersuchungsgebiet besteht aus zwei bewirtschafteten Waldteilen westlich und östlich von Embrach. Es liegt zwischen 420 und 600 m ü. M. und umfasst eine Waldfläche von 503 ha. Gewählt wurde das Gebiet, weil die terrestrische Inventur (2015) zeitnah auf eine LiDAR-Befliegung (2014/2015) folgte, es Laub-, Nadel- und Mischwälder aufweist und die terrestrischen Stichprobenpunkte im Rahmen einer Masterarbeit (Wittwer 2016) mit einem

differenziellen GPS-Gerät (D-GPS) genau eingemessen wurden. Letzteres gewährleistet die bestmögliche räumliche Übereinstimmung zwischen der terrestrischen Information und den Hilfsinformationen.

Im Untersuchungsgebiet liegen 211 terrestrische Probepunkte der Waldinventur des Kantons Zürich, die in einem systematischen Inventurnetz von 80×300 m Maschenweite angeordnet sind. Die theoretische Dichte des terrestrischen Stichprobennetzes beträgt damit 1 Inventurpunkt pro 2.4 ha. Die Erfassung der Probekreise erfolgt innerhalb eines Probekreises mit 9.77 m Radius, was einer Fläche von 300 m^2 entspricht.

Das terrestrische Inventurnetz wurde zu einer zweiphasigen Inventur mit einer Netzdichte von einem Probepunkt pro 0.15 ha erweitert, was einer 16-fachen Verdichtung entspricht. Die 22 fiktiven Kleingebiete wurden so gewählt, dass sie unterschiedlich viele terrestrische Probepunkte (4–15) aufweisen (Abbildung 3).

Extraktion der Hilfsinformationen

Als Hilfsinformationen wurden ein Vegetationshöhenmodell (VHM) und eine Laubholz-/Nadelholz-Klassifikationskarte (Waser et al 2017) verwendet (Abbildung 4, Mitte und rechts). Das VHM wurde aus dem digitalen Oberflächenmodell (DOM) und dem digitalen Terrainmodell (DTM) des Kantons Zürich abgeleitet. Die räumliche Auflösung des Datensatzes beträgt 0.5×0.5 m (Abteilung Geoinformation Kanton Zürich 2015). Aus dem VHM wurden als erklärende Variablen die mittlere Vegetationshöhe (MEAN) und die Standardabweichung der Vegetationshöhen (STD) an jedem Inventurpunkt innerhalb eines quadratischen Analysefensters mit 20 m Kantenlänge abgeleitet. Die Laubholz-/Nadelholz-Klassifikationskarte stand in einer Auflösung von 3×3 m zur Verfügung. Darauf basierend wurde für jeden Inventurpunkt der Laubholzanteil (LHAT) mithilfe des bereits für die Vegetationshöhen verwendeten Analysefensters berechnet.

Regressionsmodellierung

Auf Basis der extrahierten erklärenden Variablen wurde für das gesamte Untersuchungsgebiet das OLS-Regressionsmodell gemäss Formel 1 zur Schätzung der lokalen Vorratsdichte $\hat{Y}(x)$ definiert und mit den Daten der 211 terrestrischen Stichprobenpunkte kalibriert.

$$\hat{Y}(x) = -182.07 + 43.30 \times \text{MEAN} + 16.88 \times \text{STD} - 1.52 \times (\text{MEAN} \times \text{STD}) + 38.85 \times \text{LHAT} - 14.86 \times (\text{MEAN} \times \text{LHAT}) \quad (1)$$

Das Modell konnte 53% (adjusted R^2) der Variation der lokalen Dichte erklären (Abbildung 5). Die kreuzvalidierte mittlere Abweichung der Modellvorhersagen beträgt $165 \text{ m}^3/\text{ha}$. Die Regressionskoeffizienten lassen sich dabei wie folgt interpretieren: Sowohl die mittlere Vegetationshöhe (MEAN) als

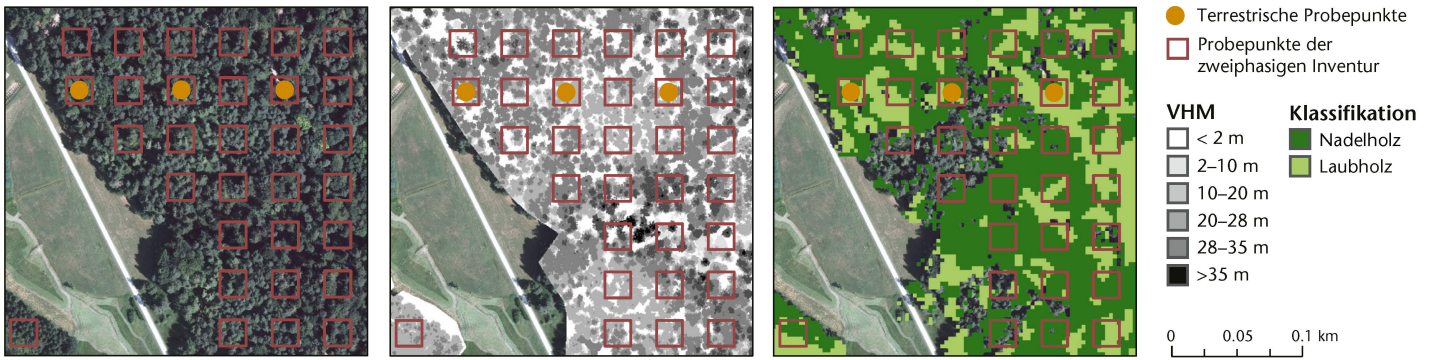


Abb 4 Ausschnitt aus dem Untersuchungsgebiet mit Überlagerung des terrestrischen Stichprobennetzes mit dem der zweiphasigen Inventur. Links: Orthofoto (GIS-ZH), Mitte: Vegetationshöhenmodell (GIS-ZH), rechts: Laubholz-/Nadelholz-Klassifikationskarte (Waser et al 2017).

auch die Standardabweichung der Vegetationshöhen (STD) haben einen positiven Effekt auf die Vorratsdichte. Die Wechselwirkung zwischen MEAN und STD hat demgegenüber einen dämpfenden Effekt. Die Einbeziehung des Laubholzanteiles (LHAT) inklusive seiner Wechselwirkung mit der mittleren Vegetationshöhe impliziert, dass bei gleichbleibender Vegetationshöhe sich ein sinkender Laubholzanteil positiv auf die Vorratsdichte auswirkt. Dieser Effekt wird mit steigender Vegetationshöhe verstärkt. Im Regressionsmodell ist der Y-Achsenabschnitt mit -182 stark negativ, verursacht aber nur in wenigen Fällen negative Vorräte. Der negative Achsenabschnitt kommt wohl daher, dass Flächen im Jungwuchs-/Dickungsstadium aufgrund der geltenden Kluppschwelle nur mangelhaft beschrieben sind.

Schätzverfahren

Als Referenz für die Schätzung der mittleren Vorratsdichte \hat{Y}_G für jedes Kleingebiet wurde der einphasige Schätzer verwendet. Bei diesem wird die mittlere Vorratsdichte \hat{Y}_G als einfacher Mittelwert der Vorratsdichten $Y(x)$ aller Stichprobenpunkte, die im jeweiligen Kleingebiet terrestrisch erfasst wurden, berechnet (Formel 2). Die geschätzte Varianz

$\hat{V}(\hat{Y}_G)$ (Formel 3) quantifiziert die Unsicherheit hinter dieser Schätzung. Aus der Varianzformel lässt sich ableiten, dass bei gleichem Stichprobenumfang der Schätzfehler in homogenen Wäldern in der Regel kleiner ist als in heterogenen Wäldern, weil die lokale Dichte weniger stark variiert.

$$\hat{Y}_G = \frac{1}{n_{2,G}} \sum_{x \in S_{2,G}} Y(x) \quad (2)$$

$$\hat{V}(\hat{Y}_G) = \frac{1}{n_{2,G}} \frac{1}{n_{2,G}-1} \sum_{x \in S_{2,G}} (Y(x) - \hat{Y}_G)^2 = \frac{1}{n_{2,G}} \hat{V}_{s_{2,G}}(Y(x)) \quad (3)$$

wobei $n_{2,G}$ die Anzahl terrestrischer Probepunkte der Stichprobe $s_{2,G}$ im Kleingebiet G ist.

Als erster zweiphasiger Kleingebietsschätzer wurde der Design-unverzerrte Pseudo Small Area Estimator (Formel 4; Mandallaz 2013) angewendet. Dieser setzt sich aus dem Mittelwert der Modellprognosen $\frac{1}{n_{1,G}} \hat{Y}(x)$ in der ersten Phase und dem Mittelwert der Modellabweichung $\frac{1}{n_{2,G}} \hat{R}(x)$ in der zweiten Phase zusammen, potenzielle systematische Abweichungen der Modellvorhersage werden also durch die Addition der mittleren Modellabweichung ausgeglichen.

$$\hat{Y}_{G,Ps\text{mall}} = \hat{Z}_G^T \hat{\beta}_{s_2} + \frac{1}{n_{2,G}} \hat{R}(x) = \frac{1}{n_{1,G}} \hat{Y}(x) + \frac{1}{n_{2,G}} \hat{R}(x) \quad (4)$$

$$\hat{V}(\hat{Y}_{G,Ps\text{mall}}) = \hat{Z}_G^T \hat{\Sigma}_{\hat{\beta}_{s_2}} \hat{Z}_G + \hat{\beta}_{s_2}^T \hat{\Sigma}_{\hat{Z}_G} \hat{\beta}_{s_2} + \frac{1}{n_{2,G}} \hat{V}_{s_{2,G}}(\hat{R}(x)) \quad (5)$$

wobei folgende Werte mit dem Regressionsmodell (Formel 1) in Beziehung stehen:
 \hat{Z}_G : Vektor der Mittelwerte jeder erklärenden Variable der verdichteten Stichprobe $s_{1,G}$ im Kleingebiet G
 $\hat{\beta}_{s_2}$: Vektor der Regressionskoeffizienten
 $\hat{R}(x) = \hat{Y}(x) - Y(x)$: Modellabweichung (Residuum) zwischen beobachteter lokaler Dichte $Y(x)$ und deren Prognose $\hat{Y}(x)$ mittels Regressionsmodell

Die geschätzte Varianz $\hat{V}(\hat{Y}_{G,Ps\text{mall}})$ (Formel 5) der Kleingebietsschätzung ist teilweise in Matrixnotation angegeben, weil diese aufgrund des vorausgesetzten linearen Regressionsmodells zwischen terrestrischen Inventurdaten und Hilfsinformationen eine besonders kompakte Darstellung des Schätzers

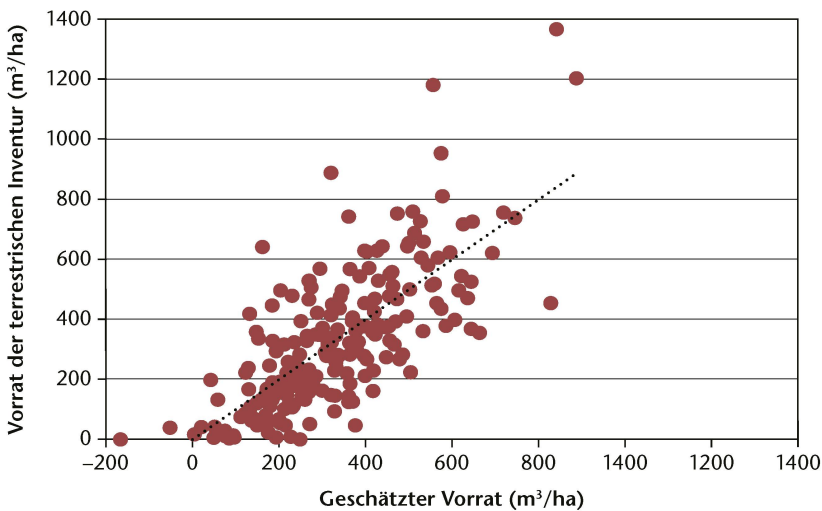


Abb 5 Gegenüberstellung der mit dem Regressionsmodell berechneten Vorräte mit den Werten der terrestrischen Aufnahme (lokale Dichte).

erlaubt. Die Varianz setzt sich aus drei additiven Termen zusammen. Der erste Term bezieht die Ungenauigkeiten bei der Schätzung der Regressionskoeffizienten ein, was durch die Varianz-Kovarianz-Matrix $\hat{\Sigma}_{\beta_{s_2}}$ ausgedrückt ist. Solche Ungenauigkeiten entstehen, weil die Regressionskoeffizienten je nach Realisation der terrestrischen Stichprobe s_2 variieren. Der zweite Term berücksichtigt durch die Varianz-Kovarianz-Matrix $\hat{\Sigma}_{Z_G}$ den Umstand, dass auch die Hilfsvariablen $Z(x) = (z(x)_1, \dots, z(x)_p)^T$ mittels einer Stichprobe $s_{1,G}$ erhoben werden. Der dritte Term stellt die Varianz der Residuen im Kleingebiet dar. Durch seine Einbeziehung in die geschätzte Varianz werden die Prognoseungenauigkeiten des Regressionsmodelles berücksichtigt.

Als zweiter zweiphasiger Kleingebietsschätzer wird der Pseudo Synthetic Estimator (Mandallaz 2013), also ein synthetischer und damit potenziell Design-verzerrter Schätzer verwendet (Formel 6). Dieser beruht ausschliesslich auf den Prognosen des global kalibrierten Regressionsmodelles und bezieht allfällige Modellabweichungen im Kleingebiet nicht mit ein.

$$\hat{Y}_{G,Psynth} = \hat{Z}_G^T \hat{\beta}_{s_2} = \frac{1}{n_{1,G}} \hat{Y}(x) \quad (6)$$

$$\hat{V}(\hat{Y}_{G,Psynth}) = \hat{Z}_G^T \hat{\Sigma}_{\beta_{s_2}} \hat{Z}_G + \hat{\beta}_{s_2}^T \hat{\Sigma}_{Z_G} \hat{\beta}_{s_2} \quad (7)$$

Folglich kann bei der geschätzten Varianz (Formel 7) die Modellungenauigkeit in Form der Residuen (dritter Term von Formel 5) nicht mehr berücksichtigt werden. Die Varianz wird also verglichen mit jener des Design-unverzerrten Kleingebietsschätzers in der Regel deutlich kleiner sein. Man sollte sich jedoch bewusst sein, dass die hohen Schätzgenauigkeiten durch die Vernachlässigung der Modellungenauigkeiten erreicht werden.

Metriken zur Beschreibung der Schätzgenauigkeit

Die Umwandlung der Varianzen in Konfidenzintervalle ermöglicht eine intuitivere Interpretation der Schätzgenauigkeit. Das 95%-Konfidenzintervall für den einphasigen Schätzer (Formel 2) wird mit der geschätzten Varianz $\hat{V}(\hat{Y}_G)$ (Formel 3) auf Basis der t-Verteilung mit Formel 8 berechnet ($\alpha=0.05$):

$$CI_{1-\alpha}(\hat{Y}_G) = \left[\hat{Y}_G - t_{n_{2,G}-1, 1-\frac{\alpha}{2}} \sqrt{\hat{V}(\hat{Y}_G)}, \hat{Y}_G + t_{n_{2,G}-1, 1-\frac{\alpha}{2}} \sqrt{\hat{V}(\hat{Y}_G)} \right] \quad (8)$$

Durch Ersetzen von \hat{Y}_G mit $\hat{Y}_{G,Psynth}$ kann Formel 8 analog zur Berechnung des 95%-Konfidenzintervalls für den Pseudo Small Area Estimator (Formel 4) benützt werden. Beim Pseudo Synthetic Estimator berechnet sich das 95%-Konfidenzintervall hingegen mit Formel 9:

$$CI_{1-\alpha}(\hat{V}(\hat{Y}_{G,Psynth})) = \left[\hat{Y}_{G,Psynth} - t_{n_{2-P,1}-\frac{\alpha}{2}} \sqrt{\hat{V}(\hat{Y}_{G,Psynth})}, \hat{Y}_{G,Psynth} + t_{n_{2-P,1}-\frac{\alpha}{2}} \sqrt{\hat{V}(\hat{Y}_{G,Psynth})} \right] \quad (9)$$

Der Schätzfehler (in %) wird dabei aus dem Standardfehler $\sqrt{\hat{V}(\hat{Y})}$ und der Schätzung \hat{Y} abgeleitet.

$$\text{Schätzfehler [\%]} = \frac{\sqrt{\hat{V}(\hat{Y})}}{\hat{Y}} \times 100 \quad (10)$$

Um den Mehrwert von zweiphasigen Inventuren zu quantifizieren, wird für den Pseudo Small Area Estimator die relative Effizienz berechnet. Diese leitet sich aus dem Verhältnis der Varianzen des einphasigen Schätzers und des zweiphasigen Schätzers ab (Formel 11). Die relative Effizienz kann als Faktor interpretiert werden, um welchen der terrestrische Stichprobenumfang erhöht werden müsste, um mit der einphasigen Schätzung die Genauigkeit der zweiphasigen Schätzung zu erreichen.

$$\text{Relative Effizienz} = \frac{\hat{V}_{\text{einphasig}}(\hat{Y})}{\hat{V}_{\text{zweiphasig}}(\hat{Y})} \quad (11)$$

Schätzergebnisse

Der zweiphasige Pseudo Small Area Estimator führte in 18 der 22 Kleingebiete (80%) zu deutlich kleineren Schätzfehlern als der einphasige Schätzer, der ausschliesslich auf den terrestrischen Inventurdaten beruht (Abbildung 6, links). Im Mittel konnte die Varianz um 40% reduziert werden. In vier Fällen kam es allerdings zu einer Erhöhung des Schätzfehlers. Die Ursache hierfür ist nicht abschliessend klar. Möglich ist, dass das Prognosemodell der globalen Schätzung in diesen Kleingebieten hohe Fehler liefert. Eine andere Möglichkeit ist, dass die terrestrischen Probepunkte des Kleingebiets die tatsächliche Heterogenität des Waldes nicht abbilden, d.h. in diesem Fall nicht repräsentativ sind. Bei genauerer Betrachtung des Kleingebietes K ist beispielsweise ersichtlich, dass die einphasige Schätzung hier nur auf vier terrestrischen Probepunkten beruht. Damit ist die Datenbasis für die Berechnung einer einphasigen Schätzung sehr gering. Trotzdem ist der Schätzfehler mit ca. 11% relativ klein. Bei der Analyse der VHM-Daten und des Luftbilds stellte sich heraus, dass alle vier terrestrischen Probepunkte zufällig in sehr ähnlichen Bestandesstrukturen liegen. Die Holzvorräte der terrestrischen Probepunkte liegen zwischen 257 und 454 m³/ha und variieren daher relativ wenig. Vermutlich weist das Kleingebiet eine höhere Variabilität auf, und das verdichtete Inventurnetz s_1 erfasst diese deutlich besser. Einen gegenteiligen Effekt zeigt sich im Kleingebiet O, in dem ebenfalls sehr wenige terrestrische Probepunkte ($n_2=5$) zur Verfügung standen. Der Vorrat der terrestrischen Probepunkte variiert hier allerdings zwischen 84 und 1365 m³/ha, also sehr stark. Dies hat zur Folge, dass der Fehler der einphasigen Schätzung mit 47% sehr hoch wird. Durch das detailliertere Abtasten der Waldfläche mit der verdichteten Stichprobe kann der Schätzfehler hier auf 26% und damit deutlich gesenkt werden.

Erwartungsgemäss liefert der Pseudo Synthetic Estimator (Abbildung 6) in der Fallstudie in allen Kleingebieten deutlich kleinere Schätzfehler als der Pseudo Small Area Estimator. Es muss jedoch beachtet werden, dass hier die Modellungenauigkeiten nicht berücksichtigt werden.

Die Schätzungenauigkeiten können auch in der Form von Konfidenzintervallen zusammen mit den berechneten Mittelwerten abgebildet werden (Abbildung 6, rechts). Kleinere Schätzfehler (Abbildung 6, links) führen dementsprechend zu kleineren Konfidenzintervallen.

Im Mittel beträgt die relative Effizienz des Pseudo Small Area Estimators 1.82 (Tabelle 2). Zwischen den Kleingebieten variiert sie stark mit Werten von 0.25 bis 4.50. In wenigen Kleingebieten liegt die relative Effizienz unter dem Wert 1. Dies kann

dadurch bedingt sein, dass die terrestrische Stichprobe s_2 die Variabilität des Kleingebiets nicht erfasst und der einphasige Schätzer somit zu geringe Varianzen liefert. In diesem Fall würde die zweiphasige Inventur somit nicht ein schlechteres, sondern ein plausibleres Resultat erzielen.

Diskussion und Schlussfolgerungen

Zweiphasige Schätzungen von Waldparametern erweisen sich dann als besonders effektiv zur Reduktion der absoluten Schätzfehler, wenn der Umfang der terrestrischen Stichprobe in den betreffenden Inventurgebieten gering ist. Ein Vorteil der vorliegend beschriebenen Design-basierten Schätzer ist, dass sie keine spezifischen Anforderungen an die

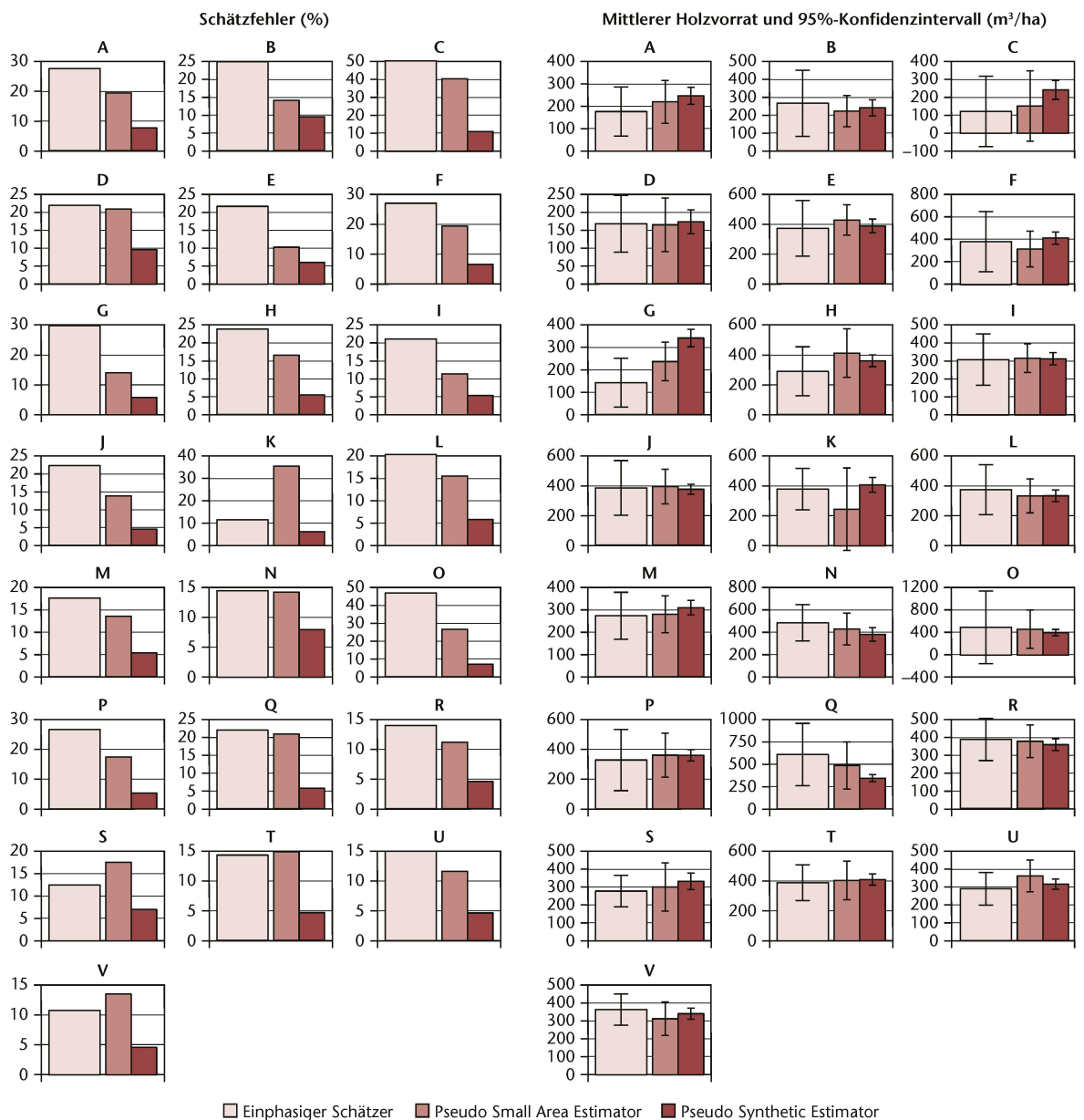


Abb 6 Schätzfehler (links) und Schätzung der mittleren Holzvorräte unter Angabe des 95%-Konfidenzintervalls (rechts) der drei angewendeten Schätzverfahren: 1) einphasiger Schätzer, 2) Pseudo Small Area Estimator (Design-unverzerrter, zweiphasiger Kleingebietsschätzer) und 3) Pseudo Synthetic Estimator (potenziell Design-verzerrter, zweiphasiger Kleingebietsschätzer).

Kleingebiet	Fläche (ha)	Einphasige Schätzung				Zweiphasige Schätzung				Relative Effizienz
		Schätzung (m ³ /ha)	95%-Konfidenzintervall (m ³ /ha)	Standardfehler (%)	Anzahl terrestrischer Probepunkte (n ₂)	Schätzung (m ³ /ha)	95%-Konfidenzintervall (m ³ /ha)	Standardfehler (%)	Anzahl Probepunkte (n ₁)	
A	13.5	176.6	109.4	27.8	11	220.0	95.9	19.6	86	1.30
B	6.2	266.8	185.7	25.1	5	222.5	87.6	14.2	40	4.50
C	4.3	123.4	198.4	50.5	4	154.2	198.9	40.5	30	1.00
D	18.9	168.9	80.1	22.1	15	166.1	75.2	21.1	123	1.13
E	17.9	374.8	188.7	21.8	9	431.0	103.0	10.4	122	3.35
F	17.9	378.7	264.9	27.2	6	314.2	157.3	19.5	105	2.83
G	26.5	141.8	109.0	29.9	6	236.3	85.7	14.1	167	1.62
H	36.3	287.5	163.2	24.0	8	410.3	161.9	16.7	241	1.02
I	33.2	307.7	142.6	21.3	13	315.6	79.2	11.5	213	3.24
J	41.1	384.0	183.8	22.3	15	393.1	116.9	13.9	272	2.47
K	11.5	378.9	137.9	11.4	4	244.7	274.9	35.3	73	0.25
L	35.8	372.4	167.2	20.4	12	330.9	113.5	15.6	207	2.17
M	31.3	273.9	105.5	17.5	12	280.7	82.9	13.4	168	1.62
N	56.1	488.7	163.1	14.5	9	433.3	142.2	14.2	72	1.32
O	27.4	488.6	636.9	47.0	5	454.0	335.5	26.6	75	3.60
P	11.2	330.2	207.0	26.5	8	362.9	148.7	17.3	144	1.94
Q	12.9	608.9	348.7	22.3	6	484.9	264.1	21.2	82	1.74
R	22.5	388.7	118.7	14.0	13	378.6	92.2	11.2	144	1.66
S	13.4	277.3	87.8	12.3	6	300.5	133.8	17.3	70	0.43
T	22.5	389.9	119.5	14.3	15	405.3	129.1	14.9	234	0.86
U	31.7	290.7	91.6	14.9	17	363.1	88.8	11.5	371	1.06
V	10.4	361.9	86.1	10.8	12	310.7	92.8	13.6	208	0.86
Mittel	22.8	330.0	177.1	22.6	9.6	327.9	139.1	17.9	147.6	1.82

Tab 2 Resultate der einphasigen und der zweiphasigen Schätzung (nur Pseudo Small Area Estimator) für die verschiedenen Kleingebiete und relative Effizienz der zweiphasigen Schätzung.

Modelle stellen, die zur Prognose der terrestrisch ermittelten lokalen Dichten verwendet werden. Die Modelle werden hier als Mittel zur Effizienzsteigerung der Schätzungen benutzt, ohne dass man sich dabei auf die Richtigkeit der Modellvorhersagen verlassen muss. Wie die Fallstudie zeigt, können die Schätzgenauigkeiten schon mit einfachen Modellen deutlich verbessert werden. Für die Anwendung der Schätzer reichen bereits Grundkenntnisse in der linearen Regressionsmodellierung aus.

Weiter reichende Modellierungskennnisse können aber genutzt werden, um die Modellgenauigkeiten zu verbessern und so die Schätzgenauigkeiten zu erhöhen. Mögliche Ansätze sind hier beispielsweise 1) die Berücksichtigung weiterer erklärender Variablen und Modellterme im Modell, 2) die Verbesserung der Qualität der verwendeten Hilfsinformationen und 3) die möglichst präzise Einmessung der Stichprobenzentren (Wittwer 2016).

Die vorgestellte Methodik kann neben der Vorratsschätzung analog auf beliebige andere kontinuierliche forstliche Kennzahlen angewendet werden. Wichtig bei der praktischen Anwendung ist die Unterscheidung zwischen Design-unverzerrten (Pseudo Small Area Estimator) und synthetischen, also potenziell Design-verzerrten Schätzern (Pseudo Synthetic Estimator). Trotz den höheren Schätzfehlern

sollte man sich nach Möglichkeit auf die Resultate der Design-unverzerrten Schätzer berufen, da diese die Modellgenauigkeiten miteinbeziehen und damit generell vertrauenswürdiger Schätzergebnisse liefern. Nach ersten Erfahrungen setzt dies jedoch eine Mindestanzahl von etwa sechs terrestrischen Probepunkte in einem Kleingebiet voraus. ■

Eingereicht: 15. Dezember 2017, akzeptiert (mit Review): 26. April 2018

Literatur

- BRESCHAN J, HILL A, GINZLER C, GABRIEL A (2017) Kombination Forstinventur und Fernerkundungsdaten. Zürich: ETH Zürich, Research Collection, Educational Material. <http://hdl.handle.net/20.500.11850/224913> (24.3.2018)
- GINZLER C, WASER LT (2017) Entwicklungen im Bereich der Fernerkundung für forstliche Anwendungen. Schweiz Z Forstwes 168: 118–126. doi: 10.3188/szf.2017.0118
- GRAFSTRÖM A, SCHNELL S, SAARELA S, HUBBELL S, CONDIT R (2017) The continuous population approach to forest inventories and use of information in the design. Environmetrics 28 (8): e2480.
- HILL A, MASSEY A (2017) forestinventory: Design-based global and small-area estimations for multiphase forest inventories. (Version R package version 0.3.1). <https://cran.r-project.org/web/packages/forestinventory/index.html> (24.3.2018)
- JAMES G, WITTEN D, HASTIE T, TIBSHIRANI R (2013) An introduction to statistical learning (Vol. 6). New York: Springer. 426 p.

- MAGNUSSEN S, MANDALLAZ D, BREIDENBACH J, LANZ A, GINZLER C (2014) National forest inventories in the service of small area estimation of stem volume. *Can J For Res* 44: 1079–1090.
- MANDALLAZ D (2008) Sampling techniques for forest inventories. Boca Raton: CRC Press. 256 p.
- MANDALLAZ D (2013) Design-based properties of some small-area estimators in forest inventory with two-phase sampling. *Can J For Res* 43: 441–449.
- 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. *Can J For Res* 43: 1023–1031.
- R DEVELOPMENT CORE TEAM (2017) R: A language and environment for statistical computing (Version 3.3.3). Vienna: R Foundation for Statistical Computing. www.R-project.org (24.3.2018)
- SCHMID-HAAS P, BAUMANN E, WERNER J (1993) Kontrollstichproben: Aufnahmeinstruktion. Birmensdorf: Eidgenöss. Forschungsanstalt WSL. 143 p.
- 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. *Scand Journal For Res* 28: 373–385.
- VON GADOW K (2005) Forsteinrichtung: Analyse und Entwurf der Waldentwicklung. Göttingen: Universitätsverlag Göttingen. 355 p.
- WASER LT, GINZLER C, REHUSH N (2017) Wall-to-wall tree type mapping from countrywide airborne remote sensing surveys. *Remote Sens* 9: 766.
- WITTWER R (2016) Auswirkungen der Lagegenauigkeit von Stichprobeninventuren auf die Schätzung der Grundfläche mit LiDAR-Daten. Zürich: ETH Zürich, Master Thesis. 68 p.

Quellen

- GEOGRAPHISCHES INFORMATIONSSYSTEM DES KANTONS ZÜRICH (GIS-ZH) Digitales Oberflächenmodell (DOM), <http://geolion.zh.ch/geodatensatz/2606> (18.4.2018)
- GEOGRAPHISCHES INFORMATIONSSYSTEM DES KANTONS ZÜRICH (GIS-ZH) Digitales Terrainmodell (DTM), <http://geolion.zh.ch/geodatensatz/2605> (18.4.2018)
- GEOGRAPHISCHES INFORMATIONSSYSTEM DES KANTONS ZÜRICH (GIS-ZH) Orthofoto Sommer RGB 2014/15, <http://geolion.zh.ch/geodatensatz/2831> (18.4.2018)

Nouveaux outils pour l'utilisation des échantillonnages deux phases dans la pratique de l'inventaire forestier

Afin d'obtenir des informations fiables sur l'état et l'évolution de la forêt, il est nécessaire d'effectuer des inventaires forestiers basés en général sur des méthodes par échantillonnage. Le nombre de placettes terrestres échantillonnées est déterminé en fonction de la précision requise pour les grandeurs forestières d'intérêt, et ce dans différents domaines (e.g. Confédération, cantons, entreprise). Des informations fiables au niveau de sous-domaines (aussi dénommés petits domaines) peuvent être obtenues en utilisant en complément des données issues de la télédétection, ce qui offre une alternative économique à une augmentation du nombre de placettes terrestres. L'EPFZ a développé dans ce but le logiciel forestinventory dans le logiciel de statistique R (libre d'accès) permettant, entre autres, l'utilisation d'estimateurs pour des petits domaines dans le cadre d'inventaires à deux phases. L'article présente les principes fondamentaux des estimateurs à deux phases pour les petits domaines et illustre les étapes de calcul ainsi que l'efficacité de la méthode par une étude de cas effectuée dans la région d'Embrach (canton de Zurich). La variable cible à estimer est la densité moyenne de matériel sur pied dans le domaine entier de 503 ha ainsi que dans 22 petits domaines ci-inclus. Les étapes principales sont (A) l'élargissement de l'inventaire monophasé à un inventaire biphasé, (B) l'identification de variables explicatives obtenues par télédétection (dans le cas présent la hauteur moyenne et l'écart type de la hauteur de la canopée ainsi que la proportion de feuillus) impactant significativement la variable cible, (C) l'élaboration d'un modèle de régression multiple reliant la variable cible aux variables explicatives et (D) l'utilisation du logiciel forestinventory pour l'estimation des petits domaines. En dépit d'un coefficient de détermination plutôt modeste ($\hat{R}^2 = 0.53$), la variance des estimations a pu être réduite pour les 22 petits domaines considérés en moyenne de 40%.

New tools for application of two-phase sampling in forest inventory practice

Sample-based forest inventories provide reliable information on the status and development of the forest. The sample sizes are thereby chosen in order to meet high estimation precisions for forest attributes on the national, cantonal or forest district level. In order to provide sufficient precision for regions where the sample size is limited, the combination of the terrestrial information with auxiliary data such as derived from remote sensing sources can constitute a cost-saving alternative compared to increasing the terrestrial sample size. Design-based estimation methods developed at ETH Zurich which implement such two-phase inventory concepts have lately been provided by the open-source software package "forestinventory" in the statistical software R. This article explains the principle of a two-phase inventory of small-area estimation and illustrates the workflow by the example of a case study in Embrach (Canton of Zurich). Target variable to be estimated for 22 subregions (small areas) located in the study area (503 ha) was the average timber volume density. The main work steps were (A) the enlargement of the one-phase inventory to a two-phase inventory, (B) the derivation of explanatory variables from remote sensing data (mean vegetation height, standard deviation of the vegetation heights and broadleaf tree proportion), (C) the formulation of a regression model to link these explanatory variables to the terrestrial information, and (D) the calculation of small area estimates (average timber volume density) using forestinventory. Despite average model accuracy ($\hat{R}^2 = 0.533$), it was possible to reduce the variance (as a measure of the estimation accuracy) in the 22 small areas by 40% on average.