

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

**Herausgeber:** Schweizerischer Forstverein

**Band:** 146 (1995)

**Heft:** 12

**Artikel:** "Le hasard fait bien les choses" : statistische Methoden für die Waldinventur

**Autor:** Mandallaz, Daniel

**DOI:** <https://doi.org/10.5169/seals-767014>

### **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:** 15.07.2025

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

# «Le hasard fait bien les choses»: statistische Methoden für die Waldinventur<sup>1</sup>

Von *Daniel Mandallaz*

*Keywords:* Forest inventories, optimization, inclusion probabilities, geostatistics.

FDK 524.6: 945.4: UDK 519.22

## 1. Einführung

Sicherheit und «Beherrschung» der Natur sind Urbedürfnisse der Menschheit. Die meisten von uns empfinden daher eine tiefe Abneigung gegen den Zufall, von Glücksspielen abgesehen, und dulden die Ungewissheit nur in Ausnahmefällen, zum Beispiel bezüglich des Zeitpunktes des eigenen Todes.

In den Naturwissenschaften wurde der Zufall dank der Wahrscheinlichkeitstheorie salonfähig (au sens du marquis Pierre Simon de Laplace), aber nur sofern er quantifizierbar ist; im Prinzip ist er nur als Spiegelbild unserer immer kleiner werdenden Ignoranz zu deuten: «Gott würfelt nicht», sagte Albert Einstein, als er die so erfolgreiche wahrscheinlichkeitstheoretische Deutung der Quantenmechanik kritisierte. Heute glauben einige namhafte Physiker, dass Gott doch würfelt; für die Medien ist Ordnung «out» und Chaos «in» (für den Mathematiker ist es nicht so einfach: es gibt chaotische deterministische Systeme). Die Diskussion bleibt, gelinde gesagt, offen. Das französische Sprichwort «le hasard fait bien les choses» könnte man demnach als hoffnungsvolle Resignation oder religiöse Beschwörung auffassen.

Das Thema dieses Referats ist selbstverständlich viel bescheidener. Ich möchte die Lesenden (früher: LeserInnen, noch früher: Leser) davon überzeugen, dass der Zufall, und damit die Statistik, für die Waldinventur ein unentbehrliches, nahezu ideales Werkzeug ist. Allerdings müssen dabei die Begriffe «Zufall» und «Waldinventur» genau definiert werden, was im nächsten Kapitel erfolgt.

<sup>1</sup> Leicht abgeänderte Version der Antrittsvorlesung vom 14. November 1994 an der ETH Zürich.

Zur erkenntnistheoretischen Standortbestimmung der Statistik möchte ich noch folgendes anführen: mit Zahlen zu lügen, ist einfach (sogar Politiker schaffen das, jedoch meistens nur vorübergehend), ohne Zahlen zu lügen, ist noch einfacher (die Scharlatane aller Gattungen haben nicht nur erfolgreiche Jahrtausende hinter sich, sondern auch eine rosige Zukunft vor sich).

## 2. Die Ziele der klassischen Waldinventur

Unter dem Begriff der klassischen Waldinventur verstehe ich die Erhebung von Daten, welche an individuellen Bäumen in bestimmten Flächen zu bestimmten Zeitpunkten beobachtet, gemessen oder geschätzt werden (z.B. Stammzahl, Grundfläche, Volumen, Klassifikationsmerkmale wie Baumart, Gesundheitszustand).

Unter dem Begriff der nicht-klassischen Waldinventur versteht man zur Zeit die Erhebung aller relevanten Informationen über Zustand und Entwicklung des Ökosystems «Wald» und dessen Wechselwirkungen zum gesellschaftlichen Umfeld.

Ich werde hier nur die klassische Waldinventur (und dies weitgehend für einen bestimmten Zeitpunkt) behandeln, weil wir für diese heute über klare und einheitliche Konzepte verfügen. Schlecht konzipierte klassische Waldinventuren sind leider, wie die Erfahrung zeigt, vom Aussterben noch nicht bedroht. Dieses Referat soll als kleine Sterbehilfe dienen.

Für die nicht-klassische Waldinventur sind klare einheitliche Konzepte noch nicht verfügbar – und werden es vermutlich nie sein. Die Fragestellungen sind unzählig, oft innerlich vag und komplex, mit den entsprechenden statistischen Nachteilen. Worüber man noch nicht klar reden kann, sollte man, dem Rat von Ludwig Wittgenstein folgend, eher schweigen (allein schon aus Zeit- und Platzgründen).

Mathematisch lässt sich das Problem der klassischen Waldinventur wie folgt formulieren.

Gegeben seien:

Waldfläche  $F$  mit Flächeninhalt  $\lambda(F)$  (in ha)

Wohldefinierte Population von  $N$  Bäumen in  $F$ , mit zugeordneten Zielgrößen  $Y_i^{(k)}$   $i = 1, 2, \dots, N$ ,  $k = 1, 2, \dots, p$

Gesucht sind:

$$T^{(k)} = \sum_{i=1}^N Y_i^{(k)}, \text{ Totale; } \bar{Y}^{(k)} = \frac{1}{\lambda(F)} T^{(k)}, \text{ Hektarendichten} \quad (2.1)$$

$$R_{k,l} = \frac{T^{(k)}}{T^{(l)}}, \text{ Quotienten}$$

Beispiel:

$Y_i^{(1)}$  = Grundfläche des  $i$ -ten Baumes, falls der  $i$ -te Baum krank ist, sonst 0.

$Y_i^{(2)}$  = Grundfläche des  $i$ -ten Baumes.

$R_{1,2}$  = Grundflächenanteil kranker Bäume.

Meistens stellt sich das obige Problem für verschiedene Teilgebiete von  $F$  und für verschiedene Zeitpunkte. Im folgenden werden wir aus Platzgründen nur Hektarendichten für eine einzige Zielgrösse betrachten, womit wir auf den hochgestellten Index verzichten können.

In der überwiegenden Mehrheit der Fälle ist  $N$  so gross, dass eine Vollkluppierung nicht in Frage kommt. Es müssen somit Stichprobenverfahren eingesetzt werden.

Im nächsten Kapitel wird der stichprobentheoretische Ansatz in einer allgemeinen und modernen Darstellung präsentiert, welche es ermöglicht, im Kapitel 4 mathematische Richtlinien für die optimale Planung aufzustellen. Interessant ist vor allem, dass der hier gewählte Ansatz auf Grund von einfachen Prinzipien zu klaren Regeln führt, die zum Teil empirisch schon lange als optimal gelten. Die mathematischen Beweise und die komplizierteren expliziten Formeln werden zu einem späteren Zeitpunkt publiziert.

### 3. Der stichprobentheoretische Ansatz

Wir gehen zunächst von der einfachsten Methode aus. In der Waldfläche  $F$  wird ein Punkt  $x$  zufällig uniform gezogen, d. h.

$$\Pr(x \in B) = \frac{\lambda(B \cap F)}{\lambda(F)}, \text{ für jedes Gebiet } B \quad (3.1)$$

Der Zufall ist somit hier ein menschliches Erzeugnis: die Koordinaten des Punktes werden von einem geeigneten Computerprogramm geliefert.

Um den Punkt  $x$  wird ein Kreis  $K$  gezogen und die in diesem Kreise zur Population gehörenden Bäume aufgenommen (*Abbildung 1*). Diese Auswahlregel kann man ebensogut dual definieren, nämlich: um jeden Baum  $i$  wird ein Kreis  $K_i$  mit konstantem Radius  $R$  gelegt, und der  $i$ -te Baum wird genau dann aufgenommen, wenn der Punkt  $x$  in Kreis  $K_i$  fällt (*Abbildung 2*). Die duale Betrachtung liefert uns sofort eine weitreichende Verallgemeinerung. Wir können nämlich jedem Baum einen Kreis  $K_i$  mit individuellem Radius  $R_i$  zuordnen; der  $i$ -te Baum wird genau dann aufgenommen, wenn der zufällige Punkt  $x$  in  $K_i$  fällt (*Abbildung 3*).

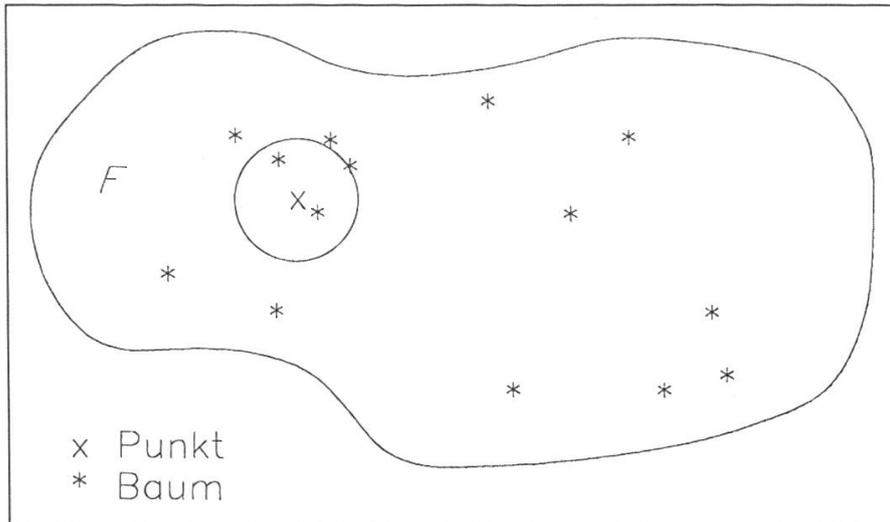


Abbildung 1. Die einfache Kreisprobe.

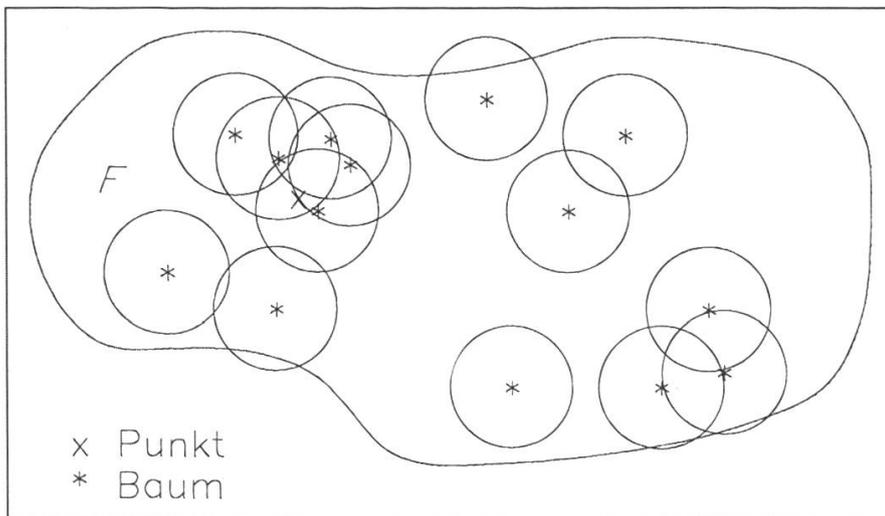


Abbildung 2. Die duale Betrachtung zur einfachen Kreisprobe.

Ein berühmtes Beispiel für individuelle Kreisradien ist die Winkelzählprobe von Bitterlich, bei der die Kreisradien von dem Baumdurchmesser abhängen (mit Zählerfaktor  $k$  beim Grenzwinkel  $\alpha$ ):

$$R_i[\text{in m}] = \frac{D_i[\text{in cm}]}{2\sqrt{k}}, \quad k = 10^4 \sin^2 \frac{\alpha}{2} \quad (3.2)$$

Wir führen nun für jeden Baum die Indikatorvariablen ein:

$I_i(x) = 1,0$  je nachdem ob der Baum vom Punkte  $x$  aus aufgenommen wird oder nicht.

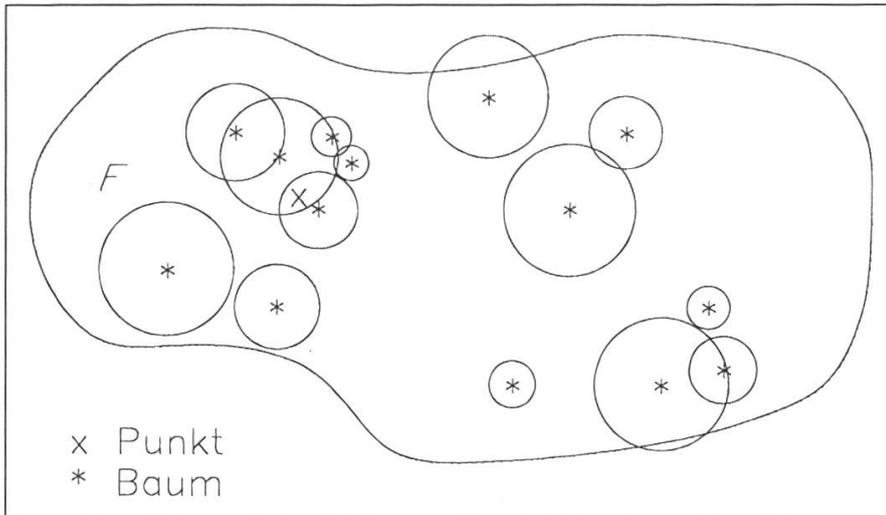


Abbildung 3. Die verallgemeinerte Auswahlregel.

Die Aufnahmewahrscheinlichkeiten sind die Erwartungswerte

$$\Pr(I_i(x) = 1) = \pi_i = E(I_i(x)) = \frac{\lambda(K_i \cap F)}{\lambda(F)} \quad (3.3)$$

Wenn der  $i$ -te Baum kein Randbaum ist gilt  $K_i \subseteq F$  und daher  $\pi_i = \frac{\lambda(K_i)}{\lambda(F)}$

Die erwartete Anzahl aufgenommener Bäume ist dann

$$E\left(\sum_{i=1}^N I_i(x)\right) = \sum_{i=1}^N \pi_i \quad (3.4)$$

Für die Berechnungen von Varianzen spielen die paarweisen Aufnahmewahrscheinlichkeiten eine wesentliche Rolle:

$$\Pr(I_i(x) = 1, I_j(x) = 1) = \pi_{ij} = E(I_i(x)I_j(x)) = \frac{\lambda(K_i \cap K_j \cap F)}{\lambda(F)} \quad (3.5)$$

In dem hier besprochenem Ansatz sind die Aufnahmewahrscheinlichkeiten  $\pi_i$  bekannt (meistens erst nach der Aufnahme der Bäume und der Vermessung des Waldrandes). Wie wir sehen werden, ist dies die *conditio sine qua non* für die Herleitung biasfreier Schätzungen. Gewisse Auswahlverfahren (darunter jene die von einem Punkt aus eine konstante Anzahl Bäume aufnehmen) führen zu unbekanntem  $\pi_i$  und sind mit zahlreichen statistischen Komplikationen und Nachteilen verbunden. Man sollte sie meines Erachtens möglichst vermeiden.

Wir wenden uns jetzt den Schätzverfahren zu. Die lokale Hektarendichte  $Y(x)$  im Punkte  $x$  wird mittels des «Horwitz-Thompson»-Schätzers definiert:

$$Y(x) = \frac{1}{\lambda(F)} \sum_{i=1}^N I_i(x) \frac{Y_i}{\pi_i} \quad (3.6)$$

( $(\lambda(F)\pi_i)^{-1}$  ist der sogenannte Hochrechnungsfaktor).

Es gilt nach Konstruktion

$$E(Y(x)) = \frac{1}{\lambda(F)} \int_F Y(x) dx = \frac{1}{\lambda(F)} \sum_{i=1}^N Y_i = \bar{Y} \quad (3.7)$$

Die lokale Dichte ist wegen der zufälligen Lage des Punktes  $x$  eine Zufallsvariable, dessen Erwartungswert gerade den erwünschten räumlichen Mittelwert liefert. In der Definition der lokalen Dichte werden Randeffekte theoretisch korrekt behandelt, die einzige Schwierigkeit besteht in der Definition und der Erfassung des Waldrandes, und eventuell in der Berechnung von  $\lambda(K_i \cap F)$ . Wir haben dabei stillschweigend angenommen, dass alle topographischen Angaben durch geeignete Projektion in einer Ebene vorliegen (was in der Praxis nicht immer einwandfrei erreicht werden kann, z.B. Hangneigungskorrekturen bei Probeflächen in Kretenlage).

Der Zufall liefert uns somit eine biasfreie Schätzung. Er ist sozusagen der unsichtbare Faden, welcher von den erfassten zu den nicht erfassten Bäumen führt und somit für die «Repräsentativität» sorgt. Die Varianz ergibt sich leicht aus:

$$Var(Y(x)) = \frac{1}{\lambda^2(F)} \left\{ \sum_{i=1}^N Y_i^2 \frac{(1-\pi_i)}{\pi_i} + \sum_{i \neq j} \frac{(\pi_{ij} - \pi_i \pi_j)}{\pi_i \pi_j} Y_i Y_j \right\} \quad (3.8)$$

Im Gegensatz zum Erwartungswert hängt die Varianz, über die  $\pi_{ij}$ , von der räumlichen Verteilung der Bäume im Wald ab.

Der lineare biasfreie Horwitz-Thompson-Schätzer wurde 1952 in die Stichprobentheorie eingeführt. 1955 konnte der indische Statistiker Godambe folgende bahnbrechende Resultate beweisen (übersetzt in die Sprache der Waldinventur):

- (1)  $Y(x)$  ist zulässig, d. h. es gibt keinen anderen linearen biasfreien Schätzer, welcher für alle Wälder eine kleinere Varianz hat ( $Y(x)$  ist keine schlechte Schätzung!).
- (2) Es gibt keinen linearen biasfreien Schätzer, der für alle Wälder die kleinste Varianz hat (es gibt keine beste Schätzung!).

Die Interessierten finden einen ausgezeichneten Überblick über die modernen Grundlagen der Stichprobentheorie, insbesondere auf die obigen Resultate in *Smith, T.F.M., 1976.*

Die Suche nach dem universell besten Schätzer ist somit illusorisch, und wir müssen unsere Ansprüche reduzieren:

erstens, indem wir uns auf die Klasse der Horwitz-Thompson-Schätzer, und somit der damit verbundenen lokalen Dichten, einschränken (sie sind einfach und zulässig).

zweitens, indem wir die Aufnahmewahrscheinlichkeiten  $\pi_i$  so einrichten, dass das Kosten/Varianz-Verhältnis «im Mittel» für eine grosse Klasse von Wäldern «optimal» ist.

Die vage Formulierung «im Mittel» soll nun präzisiert werden. Wir gehen von einem stochastischen Modell  $L$  aus, welches bei vorgegebenen  $N, Y_i, i = 1, 2, \dots, N$  die Lage der Bäume im Wald erzeugt. Die effektive Lage der Bäume wird somit als die Realisierung eines stochastischen Prozesses angesehen. Der Zufall ist jetzt als Gedankenexperiment oder als Glückspiel der Natur zu verstehen. Die zwei einfachsten Lagemodelle sind:

$L_1$ : *Der Poisson-Wald*

Die  $N$  Bäume sind unabhängig uniform verteilt in  $F$ .

$L_2$ : *Der lokale Poisson-Wald*

Der Wald besteht aus Teilgebieten («Straten»), in welchen  $L_1$  gilt.

Ein Plenterwald kann in erster Näherung als ein Poisson-Wald gesehen werden, während ein Wald mit verschiedenen Entwicklungsstufen als lokaler Poisson-Wald gesehen werden kann. Rottenartige Wälder wären Gegenbeispiele. Selbstverständlich sind diese Modelle nur grobe Approximationen der Realität; sie sind jedoch die analytisch einfachsten Modelle, welche vernünftige Ansätze liefern. Die Resultate sind exakt unter  $L_1$  und gelten bis auf die Vernachlässigung der Stratenrandeffekten unter  $L_2$ .

«Im Mittel» soll nun bedeuten, dass wir anstelle der Varianz die *antizipierte Varianz*

(3.9)

$$AV = E_L(\text{Var}(Y(x)))$$

betrachten, d.h. den Erwartungswert der gewöhnlichen Varianz unter allen möglichen Lagen (unter den oben beschriebenen Modellen) der Bäume.

Die Aufgabe besteht nun darin, die  $\pi_i$  so zu bestimmen, dass die antizipierte Varianz bei vorgegebenen Kosten – oder die Kosten bei vorgegebener antizipierter Varianz – minimiert wird.

## 4. Mathematische Richtlinien für die Optimierung

### 4.1 Die einphasige Inventur

Wir gehen von  $n_2$  Punkten  $x_1, x_2, \dots, x_{n_2}$  aus, welche in  $F$  unabhängig uniform verteilt sind. Als biasfreien Schätzer nehmen wir den Mittelwert der lokalen Dichten

$$\hat{Y} = \frac{1}{n_2} \sum_{l=1}^{n_2} Y(x_l) \quad (4.1.1)$$

Varianz und antizipierte Varianz dieses Schätzers sind dann:

$$\begin{aligned} \text{Var}(\hat{Y}) &= \frac{1}{n_2} \text{Var}(Y(x)) \\ AV &= \frac{1}{n_2 \lambda^2(F)} \sum_{i=1}^N \frac{Y_i^2 (1 - \pi_i)}{\pi_i} - \frac{1}{n_2} \phi(Y_1, Y_2, \dots, Y_N) \end{aligned} \quad (4.1.2)$$

wobei  $\phi(Y_1, Y_2, \dots, Y_N)$  nicht von den  $\pi_i$  abhängt und unter Modell  $L_1$  sogar 0 ist. Man merke, dass der Übergang zur antizipierten Varianz die  $\pi_{ij}$ -Terme zum Verschwinden bringt.

Die erwarteten Kosten werden gemäss

$$C = n_2 \left( c_{2,0} + c_{2,1} \sum_{i=1}^N \pi_i \right) + \tau(n_2) \quad (4.1.3)$$

modelliert, wobei:

$c_{2,0}$  : Einrichtungskosten pro Stichprobenpunkt

$c_{2,1}$  : Aufnahmekosten pro Baum zur Bestimmung von  $Y_i$

$\tau(n_2)$  : Transportkosten

Die feine Modellierung der Transportkosten ist äusserst schwierig, wenn nicht illusorisch. Aus geometrischen Gründen nimmt  $\tau(n_2)$  etwa wie  $\sqrt{n_2}$  zu.

In der Praxis genügt es, für eine qualitative Untersuchung eine lineare Approximation zwischen unterer und oberer Schranken für  $n_2$  zu betrachten; für die Optimierung ist nur die Steigung dieser linearen Approximation massgebend. Lineare Transportkosten haben zudem den Vorteil, dass man dieselben Regeln (bis auf Konstanten) bekommt, egal ob man die antizipierte Varianz bei vorgegebenen Kosten oder die Kosten bei vorgegebener antizipierter Varianz minimiert, was im allgemeinen nicht gilt und somit universelle Richtlinien verunmöglicht. Wenn die Transportkostenfunktion numerisch

bekannt wäre (meines Erachtens eine Illusion), so wäre eine numerische Optimierung im Prinzip kein Problem.

Die Minimierung der antizipierten Varianz bei vorgegebenen Kosten führt zu folgenden Regeln:

(1) Für jedes  $n_2$ :  $\pi_i \propto |Y_i|$ , d.h. *Probability Proportional to Size*, *PPS-Regel*

(2)  $n_2$  möglichst klein, weil  $\frac{\partial}{\partial n_2} E_L(\text{Var}\hat{Y}) > 0$

Regel (2) muss vom praktischen Standpunkt aus relativiert werden. Erstens müssen Varianzen immer geschätzt werden, was eine minimale Anzahl Stichprobenpunkte erfordert; zweitens hängt die Proportionalitätskonstante in Regel (1) noch von den vorgegebenen Kosten und  $n_2$  ab; kleine  $n_2$  führen zu grossen Kreisen und daher zu praktischen Schwierigkeiten bei der Feldarbeit. Ferner will man oft Schätzungen für Teilgebiete. Die korrekte Interpretation von (2) ist somit:  $n_2$  so klein wie möglich unter Berücksichtigung der Randbedingungen; es ist nicht sinnvoll, viele sehr kleine «Probeflächen» zu wählen, weil dann die Transportkosten im Vergleich zur Abnahme der Varianz übermässig steigen.

Der Beweis von Regel (1) erfolgt über die Cauchy-Schwartz-Ungleichung. Wenn die Zielgrösse die Grundfläche ist, so führt Regel (1) direkt zur Winkelzählprobe, deren «Optimalität» somit theoretisch gerechtfertigt wäre.

Für andere Zielgrössen, z.B. Volumen, ist Regel (1) schwierig zu implementieren. Zudem sind die Kosten für die genau Bestimmung von  $Y_i$  oft recht hoch (z.B. zwei Durchmesser- und eine Höhenmessung). Aus diesem Grunde erscheint es sinnvoll, die  $\pi_i$  proportional zu einer billigen Approximation (Prognose)  $\hat{Y}_i$  von  $Y_i$  zu wählen (z.B. Volumen über BHD-Einwegtarif), was im Abschnitt 4.2 skizziert wird.

#### 4.2 Die einphasige-zweistufige Inventur

Wie in 4.1 gehen wir von  $n_2$  Punkten  $x_1, x_2, \dots, x_{n_2}$  aus, welche in  $F$  unabhängig uniform verteilt sind. In jedem dieser Punkte wird die lokale Dichte (Horwitz-Thompson-Schätzer) mit den Prognosen  $\hat{Y}_i$  mittels den  $\pi_i$  bestimmt. Unter allen im Punkte  $x$  erfassten Bäumen werden einige, unabhängig voneinander und mit individuellen Aufnahmewahrscheinlichkeiten  $p_i$ , zur Bestimmung der exakten Zielgrösse  $Y_i$  ausgewählt; die entsprechende Indikatorvariable soll mit  $J_i$  bezeichnet werden. Die Abweichung zwischen exaktem Wert und Prognose, kurz Residuum, ist per definitionem  $R_i = Y_i - \hat{Y}_i$ .

Der verallgemeinerte Horwitz-Thompson-Schätzer im Punkte  $x$  ist wie folgt definiert:

$$Y^*(x) = \frac{1}{\lambda(F)} \sum_{i=1}^N \frac{I_i(x) \hat{Y}_i}{\pi_i} + \frac{1}{\lambda(F)} \sum_{i=1}^N \frac{I_i(x) J_i(x) R_i}{\pi_i p_i} \quad (4.2.1)$$

und

$$\hat{Y}^* = \frac{1}{n_2} \sum_{l=1}^{n_2} Y^*(x_l) \quad (4.2.2)$$

Der verallgemeinerte Horwitz-Thompson-Schätzer hat den Vorteil, dass er für die exakte Zielgrösse biasfrei ist. Die Varianz ist dann:

$$\text{Var} \hat{Y}^* = \frac{1}{n_2} \text{Var} Y(x) + \frac{1}{n_2 \lambda(F)} \int_F V(x) dx \quad (4.2.3)$$

wobei  $V(x) = \text{Var} Y^*(x)$  die Varianz der zweistufigen Ziehung in  $x$  bezeichnet.

Wir modellieren die erwarteten Kosten gemäss:

$$C = n_2 \left( c_{2,0} + c_{2,1} \sum_{i=1}^N \pi_i + c_{2,2} \sum_{i=1}^N \pi_i p_i \right) + \tau(n_2) \quad (4.2.4)$$

wobei:

$c_{2,0}$  : Einrichtungskosten pro Stichprobenpunkt

$c_{2,1}$  : Aufnahmekosten pro Baum zur Bestimmung von  $\hat{Y}_i$  (erste Stufe)

$c_{2,2}$  : Aufnahmekosten pro Baum zur Bestimmung von  $Y_i$  (zweite Stufe)

$\tau(n_2)$  : Transportkosten

Die Minimierung der antizipierten Varianz führt nun zu den Regeln:

- (1) Für jedes  $n_2$ :  $\pi_i \propto |\hat{Y}_i|$ , d.h. *Probability Proportional to Prediction*,  
PPP-Regel
- (2) Für jedes  $n_2$ :  $\pi_i p_i \propto |R_i|$ , d.h. *Probability Proportional to Error*, PPE-Regel
- (3)  $n_2$  möglichst klein, weil  $\frac{\partial}{\partial n_2} E_L(\text{Var} \hat{Y}^*) > 0$

Regel (2) ist auch gültig, wenn die  $\pi_i$  vorgegeben sind.

Somit hätten wir die Pseudo-Gleichung  $PPS \approx PPP + PPE$  (weil eben  $S = P + E$ )

Die Approximation von Regel (1) durch Treppenfunktionen führt, am Beispiel des Volumens, zu den bekannten konzentrischen Kreisproben oder zur Winkelzählprobe.

Der verallgemeinerte Horwitz-Thompson-Schätzer wurde 1993 für die zweite schweizerische Landesforstinventur bereits erfolgreich implementiert

(verglichen mit der ersten Landesinventur kann man, bei gleicher Genauigkeit und mit rund 10 000 Probeflächen, die Anzahl der (3-Weg-)Tarifbäume von rund 40 000 auf 10 000 reduzieren, und dazu noch den Bias stark verkleinern).

Das optimale Verhältnis der Anzahl Bäume in der ersten und jener in der zweiten Stufe ergibt sich aus:

$$\frac{\sum_{i=1}^N \pi_i p_i}{\sum_{i=1}^N \pi_i} = \sqrt{\frac{c_{2,1}}{c_{2,2}}} \frac{\sum_{i=1}^N |R_i|}{\sum_{i=1}^N |\hat{Y}_i|} \quad (4.2.5)$$

Bis jetzt haben wir die rein terrestrischen Inventuren betrachtet. Bekanntlich kann mit der Verwendung von Hilfsinformation (durch Fernerkundung, Bestandskarten, frühere Inventuren, okulare Schätzung usw.) die Effizienz noch verbessert werden, oft durch einfache Stratifizierung. Wir werden im Abschnitt 4.3 den verallgemeinerten Horwitz-Thompson-Schätzer auf die zweiphasigen Inventuren erweitern, und dies in einer sehr allgemeinen Form.

### 4.3 Zweiphasige/zweistufige Inventuren

Die Hilfsinformation wird in  $n_1$  Punkten  $x \in s_1$  erhoben, welche wir als unabhängig uniform verteilt in  $F$  annehmen. Für jeden Punkt  $x \in s_1$  wird mittels eines Modells die Hilfsinformation zur Berechnung einer Prognose  $\hat{Y}(x)$  der lokalen Dichte  $Y(x)$  verwendet («wie» soll uns hier nicht kümmern!). In jedem Punkt  $x \in s_1$  wird mit Wahrscheinlichkeit  $p(x)$  und unabhängig von jedem anderen Punkt entschieden, ob in diesem Punkt Aufnahmen zur Ermittlung des verallgemeinerten Horwitz-Thompson-Schätzers durchgeführt werden;  $p(x)$  soll nur von der Hilfsinformation abhängen, z. B. von  $\hat{Y}(x)$ . Die so ausgewählten  $n_2$  Punkte bilden die Unterstichprobe  $s_2$ . Man merke, dass  $n_2$  eine Zufallsvariable ist.

Der verallgemeinerte zweiphasige/zweistufige Horwitz-Thompson-Schätzer wird dann wie folgt definiert:

$$\hat{Y}_{reg}^* = \frac{1}{n_1} \sum_{x \in s_1} \hat{Y}(x) + \frac{1}{n_1} \sum_{x \in s_2} \frac{Y^*(x) - \hat{Y}(x)}{p(x)} \quad (4.3.1)$$

mit Varianz

$$Var \hat{Y}_{reg}^* = \frac{1}{n_1 \lambda(F)} \int_F \frac{r^2(x)(1-p(x))dx}{p(x)} + \frac{1}{n_1 \lambda(F)} \int_F \frac{V(x)dx}{p(x)} + \frac{1}{n_1} Var Y(x) \quad (4.3.2)$$

wobei  $r(x) = Y(x) - \hat{Y}(x)$  die wahren Residuen (nicht beobachtbar) sind.

Die erwarteten Kosten werden recht allgemein durch

$$n_1 \left\{ \frac{1}{\lambda(F)} \int_F c_1(x) dx + \frac{1}{\lambda(F)} \int_F c_2(x) p(x) dx \right\} \quad (4.3.3)$$

modelliert, wobei:

$c_1(x)$  : Erhebungskosten der Hilfsinformation im Punkte  $x$

$$c_2(x) = c_{2,0}(x) + c_{2,1}(x) \sum_{i=1}^N \pi_i + c_{2,2}(x) \sum_{i=1}^N \pi_i p_i \quad (4.3.4)$$

$c_{2,0}(x)$  : Einrichtungskosten samt linearisierter Transportkosten

$c_{2,1}(x)$  : Aufnahmekosten der ersten Stufe, pro Baum

$c_{2,2}(x)$  : Aufnahmekosten der zweiten Stufe, pro Baum

Die globale Optimierung nach allen Parametern scheint, zur Zeit, ausser Reichweite. Wenn  $p(x)$  konstant ist, kann man zeigen, dass die Regeln (1) bis (3) vom Abschnitt 4.2 weiterhin gelten. Wir gehen daher davon aus, dass  $\pi_i, p_i$  vorgegeben sind, und optimieren dann nur nach  $p(x)$  und bekommen die optimale Lösung als:

$$p_o(x) = \sqrt{\frac{\bar{c}_1}{c_{2,2}(x)}} \sqrt{\frac{r^2(x) + V(x)}{\text{Var}\hat{Y}(x)}} \quad (4.3.5)$$

wobei  $\bar{c}_1(x) = \frac{1}{\lambda(F)} \int_F c_1(x) dx$ .

Mit anderen Worten ist das optimale Verfahren wieder eine Art *PPE*. Obige Formel verallgemeinert die bekannte optimale Stratifizierung für beliebige Prognosemodelle und zweistufige Verfahren.

Der Vollständigkeit halber sollte erwähnt werden, dass biasfreie Varianzschätzungen für 4.1., 4.2 und 4.3 erhältlich sind und dass die Theorie auf Traktstichproben erweitert werden kann.

Anhand dieser drei Beispiele können wir im stichprobentheoretischen Ansatz die vier Gebote der Forstinventur etwa wie folgt formulieren:

- I. Messe nur das, was schlecht modelliert werden kann und nicht zuviel kostet!
- II. Traue nur bedingt deinen Modellen und korrigiere mit den Residuen!
- III. Beschränke die Feldaufnahmen auf möglichst wenige «grosse» Probenflächen, soweit es die Randbedingungen erlauben!
- IV. Verwende *PPS*-, *PPP*-, *PPE*- Verfahren!

Unter gleichzeitiger Verwendung dieser vier Regeln können die Kosten, bei gleicher Genauigkeit, nicht selten um mehr als 50% im Vergleich zu einer einfachen terrestrischen Inventur reduziert werden, was den Titel dieses Referats rechtfertigt. Allerdings sollte nun klar sein, dass «zufällig» gar nicht «irgendwie» bedeutet.

Der stichprobentheoretische Ansatz beruht wesentlich auf der Annahme, dass die Beobachtungspunkte unabhängig voneinander verteilt sind. In der Praxis liegen aber die Punkte meistens auf einem systematischen Gitter, dessen Startpunkt und Orientierung zufällig, wenn überhaupt, gewählt sind, was die Unabhängigkeitsannahme zunichte macht. Theorie und Praxis sind somit in einem wesentlichen Punkt unversöhnlich. Dieses Problem ist seit «Urzeiten» bekannt, und die empirische Erfahrung weist darauf hin, dass die Auswertung einer systematischen Stichprobe, als ob sie zufällig wäre, in der Regel zu überschätzten Varianzen führt. Einige «Zwitteransätze» (z.B. Paardifferenzformel) werden gelegentlich verwendet, sind aber vom theoretischen Standpunkt aus nicht überzeugend.

Hinzu kommt, dass der stichprobentheoretische Ansatz (design-based inference), auch wenn alle Annahmen erfüllt wären, für die Schätzung von kleinen Gebieten, wegen der kleinen Anzahl Beobachtungspunkte, oft ungeeignet ist. Design-based-Inferenz ist zwar im Prinzip immer gültig (bis auf den oben erwähnten Vorbehalt für systematische Stichproben), egal ob die verwendeten Prognosemodelle richtig oder falsch sind (die Residuen korrigieren den Modellbias). Diese «Modell-Unabhängigkeit» muss jedoch bei lokalen Schätzproblemen mit grossen Varianzen teuer bezahlt werden. Der Ausweg aus diesem Dilemma erfordert einen grundsätzlich neuen Ansatz, und daher einen neuen Zufallsbegriff, der im Kapitel 5 kurz erläutert wird<sup>2</sup>.

## 5. Der geostatistische Ansatz

Im stichprobentheoretischen Ansatz ist der Wald fest: die lokale Dichte  $Y(x)$  ist eine Zufallsvariable, weil der Punkt  $x$  zufällig ist.

Im geostatistischen Ansatz (model-dependent inference) sind die Beobachtungspunkte  $x_i$  fest (wie sie erzeugt werden, spielt bei der Auswertung keine Rolle, bei der Planung schon); und dafür wird die lokale Dichte  $Y(x)$  als die Realisierung eines stochastischen Prozesses aufgefasst. Der gegebene Wald ist somit ein zufälliges Erzeugnis der Natur. Um weiter zu kommen, muss vorausgesetzt werden, dass dieses Glückspiel der Natur eine minimale Gesetzmässigkeit aufweist. Die Grundfrage bleibt dieselbe: wie kann von den beob-

<sup>2</sup> Für eine detaillierte Abhandlung der hier skizzierten Resultate siehe *D. Mandallaz*, 1993, für eine ausgezeichnete, aber anspruchsvolle Einführung in die Geostatistik siehe *N. Cressie*, 1991.

achteten Punkten auf die nicht beobachteten geschlossen werden? In der design-based-Inferenz geschieht dies über die «Repräsentativität», induziert durch die zufällige Ziehung, während es in der model-dependent-Inferenz über die «Stationarität» erfolgt, nämlich:

Für alle Ortsvektoren  $x$  und Verschiebungsvektoren  $h$  gelten

$$E(Y(x+h)-Y(x)) = 0 \tag{5.1}$$

$$E(Y(x+h)-Y(x))^2 = 2\gamma(h)$$

Die Funktion  $\gamma(h)$  heisst das (semi-)Variogramm. Falls die stationäre räumliche Autokovarianzfunktion  $c(h) = cov(Y(x+h), Y(x))$  existiert, so gilt:

$$\gamma(h) = c(0)-c(h) \tag{5.2}$$

In der Regel werden benachbarte Punkte stärker korrelierte lokale Dichten haben als weit entfernte Punkte. In der Praxis wird oft angenommen, dass die räumliche Korrelation ab einem gewissen Abstand (Reichweite) verschwindet. Wir betrachten nun ein Teilgebiet  $F_0 \subset F$ . Wegen der Biasfreiheit des Horwitz-Thompson Schätzers gilt:

$$\frac{1}{\lambda(F_0)} \sum_{i \in F_0} Y_i = \frac{1}{\lambda(F_0)} \int_{F_0} Y(x) dx =: Y_0 \tag{5.3}$$

Das Schätzen einer endlichen Summe wird somit auf das Schätzen eines Integrals zurückgeführt.

Für die einphasige Inventur geschieht dies über «ordinary kriging». Man bestimmt Koeffizienten  $\lambda_i$ , so dass der Schätzer

$$\hat{Y}_0 = \sum_{i \in S_2} \lambda_i Y(x_i) \tag{5.4}$$

biasfrei ist und einen minimalen quadratischen Fehler hat, d.h.

$$E(Y_0 - \hat{Y}_0) = 0 \tag{5.5}$$

$$E(Y_0 - \hat{Y}_0)^2 = \min$$

Die optimalen Gewichte sind dann Lösungen eines linearen Gleichungssystems, dessen Koeffizienten numerisch relativ aufwendig sind (Integrale des Variogrammes über das Gebiet). Der quadratische Fehler kann ebenfalls geschätzt werden. Die Verallgemeinerung auf die zweiphasige Inventur erfolgt über «double kriging»:

(5.6)

$$\hat{Y}_o = \sum_{j \in S_1} \mu_j \hat{Y}(x_j) + \sum_{i \in S_2} \lambda_i r(x_i) \quad , \quad r(x_i) = Y(x_i) - \hat{Y}(x_i)$$

Die  $\mu_j$  erhält man durch ordinary kriging der Prognosen und die  $\lambda_i$  durch ordinary kriging der Residuen. Man merke, dass bei «double kriging» das Prognosemodell nicht «richtig» zu sein braucht (die Korrektur erfolgt automatisch über die Residuen). Falls die Hilfsinformation in allen Punkten des Gebietes erhältlich ist und das Prognosemodell «richtig» ist (d.h.: das räumliche Mittel der Residuen ist Null), geht «double kriging» in «universal kriging» über. Die Theorie kann auf zweistufige Verfahren erweitert werden.

Im Prinzip sind Kriging-Verfahren, bis auf den numerischen Aufwand, einfache Prozeduren, wenn die entsprechenden Variogramme bekannt sind. In der Praxis müssen diese zunächst geschätzt werden, was, vor allem für die Residuen, eine anspruchsvolle statistische und numerische Aufgabe ist.

Zur Illustration geben wir die Resultate einer Betriebsinventur (mit Trakten zu je nominal 5 Punkten). Das ganze Gebiet umfasst etwa 218 ha, davon wurde ein kleines Teilgebiet von 17 ha vollkluppiert. Die Hilfsinformation bestand aus einer vereinfachten Bestandskarte, das Prognosemodell war eine multiple lineare Regression mit sechs (0/1)-Indikatorvariablen (*Tabellen 1 und 2*).

*Tabelle 1.* Ganzes Gebiet, 218 ha. Punktschätzung/(Standardfehler).

	<i>einphasig</i> <i>design-based/kriging</i>		<i>zweiphasig</i> <i>design-based/kriging</i>	
	Stammzahldichte	321 (18)	325 (15)	325 (12)
Grundflächendichte	31,9 (1,1)	31,9 (0,7)	31,2 (0,9)	31,4 (0,7)
<i>Anzahl Trakte</i>	$n_2 = 73$ Trakte		$n_1 = 298$ Trakte	$n_2 = 73$ Trakte

*Tabelle 2.* Kleines Gebiet, 17 ha. Punktschätzung/(Standardfehler).

	<i>einphasig</i> <i>design-based/kriging</i>		<i>zweiphasig</i> <i>design-based/kriging</i>	
	Stammzahldichte wahrer Wert: 280	246 (66)	294 (41)	257 (49)
Grundflächendichte wahrer Wert: 29,6	24,5 (3,6)	27,7 (2,1)	24,0 (3,7)	29,5 (1,4)
<i>Anzahl Trakte</i>	$n_2 = 19$ Trakte		$n_1 = 92$ Trakte	$n_2 = 19$ Trakte

N.B. Beim Kriging wurden immer alle Datenpunkte verwendet.

Dieses Beispiel und qualitative theoretische Überlegungen zeigen, dass für globale Schätzprobleme die design-based- und die model-dependent-Inferenz

im wesentlichen dieselben Punktschätzungen liefern, mit etwas kleineren Fehlern für die letztere. Für lokale Schätzprobleme liefert die Geostatistik bessere Schätzungen. Der Vorteil der zweiphasigen Verfahren ist ebenfalls unbestritten.

Der Nachteil der Geostatistik ist, dass es zur Zeit noch nicht möglich ist, qualitative Richtlinien für die Optimierung anzugeben. Weil Optimierung jedoch eine globale Angelegenheit ist, scheint es vernünftig, diese Aufgabe mit dem einfacheren stichprobentheoretischen Ansatz anzupacken. Eine sorgfältige Abklärung dieser Frage bedarf noch weiterer, anspruchsvoller Forschung. Bei der Auswertung kann die Geostatistik wertvolle Hilfe leisten, vor allem bei lokalen Schätzproblemen oder bei Folgeinventuren mit abgeänderten Stichprobenplänen (die Verteilung der Punkte spielt ja keine Rolle!).

## 6. Ausblick

Die knapperen finanziellen Mittel einerseits sowie die immer grössere Bedeutung der Fernerkundung und der geographischen Informationssysteme andererseits führen meines Erachtens zu folgenden Entwicklungen in der klassischen Waldinventur:

- 1.) Bessere Koordination und Synergien (in Planung und Auswertung) zwischen nationalen, regionalen und lokalen Inventuren.
- 2.) Zunehmendes Bedürfnis nach modellabhängigen Schätzverfahren.

In der nichtklassischen Waldinventur (multifunctional resource assessment) wird man vermutlich mit zahlreichen Ad-hoc-Fragestellungen und -Schätzverfahren leben müssen. Ich möchte es dennoch hier nicht unterlassen vor «all in one» Inventuren zu warnen, wo «alles» untersucht wird, mit der Konsequenz, dass «nichts» auch nur annähernd optimal geschätzt werden kann. Getrennte kleinere Studien scheinen hierzu vernünftiger, wenn eine Vernetzung der Daten durch eine intelligente Koordination und Planung von Anfang ermöglicht wird.

Diejenigen Leser, die von der Wichtigkeit effizienter und daher anspruchsvoller statistischer Methoden in den Forstwissenschaften im allgemeinen und in der Waldinventur insbesondere immer noch nicht überzeugt sind, werden vielleicht durch folgenden Spruch von Mark Twain ihre Meinung ändern:

«If all you have got is a hammer then everything looks like a nail.»

## **Zusammenfassung**

Dieser Aufsatz gibt einen Überblick über die wahrscheinlichkeitstheoretischen Grundlagen der modernen statistischen Verfahren, welche für die Waldinventur verwendet werden. Aufnahmewahrscheinlichkeiten proportional zur Zielgrösse, zu einer Prognose der Zielgrössen bzw. zum Fehler dieser Prognose erlauben eine Optimierung der Stichprobenverfahren, sowohl der einphasigen oder zweiphasigen wie auch der einstufigen oder zweistufigen und deren Kombinationen. Die geostatistischen Verfahren werden ebenfalls kurz beschrieben, und ein Beispiel zeigt ihre Überlegenheit für lokale Schätzprobleme.

## **Résumé**

### **«Le hasard fait bien les choses»: méthodes statistiques pour l'inventaire forestier**

Cet article présente un survol des principes probabilistes fondamentaux pour les techniques statistiques modernes utilisées dans l'inventaire forestier. Les probabilités d'inclusion proportionnelles soit à la grandeur cible, soit à une prévision de celle-ci, soit enfin à l'erreur de cette prévision, permettent à l'inventoriste d'optimiser les plans de sondage à une ou deux phases, et avec un ou deux degrés. Les techniques géostatistiques sont brièvement décrites et un exemple démontre leur supériorité pour l'estimation locale.

## **Summary**

### **«Le hasard fait bien les choses»: statistical methods for forest inventory**

This paper gives an overview of the probabilistic foundations of modern statistical techniques used in forest inventory. Inclusion probabilities proportional to size, to a prediction thereof or to the error of this prediction give the inventorist the possibility to optimize one-phase or two-phases sampling schemes, possibly combined with one-stage or two-stages procedures. A brief account of geostatistical techniques is also given and an example illustrates their superiority for small area estimation.

### *Literatur*

- Cressie, N.* (1991) *Statistics for spatial data*, John Wiley & Sons, Inc., New York.
- Mandallaz, D.* (1993) *Geostatistical methods for double sampling schemes: application to combined forest inventories*. Habilitationsschrift, Chair of Forest Inventory and Planning, ETH Zürich, 133 p.
- Smith, T.F.M.* (1976) The foundations of survey sampling: a review, *Journal Royal Statistical Society, series A*, 139, 183–204.

*Verfasser:* PD Dr. Daniel Mandallaz, Professur für Forsteinrichtung und Waldwachstum, Departement Wald- und Holzforschung der ETH Zürich, CH-8092 Zürich.