

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

Herausgeber: Schweizerischer Forstverein

Band: 167 (2016)

Heft: 3

Artikel: Wiederholungsaufnahmen als Instrument zur Qualitätssicherung im Schweizerischen Landesforstinventar

Autor: Traub, Berthold / Cioldi, Fabrizio / Düggin, Christoph

DOI: <https://doi.org/10.5169/seals-1097429>

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>

Wiederholungsaufnahmen als Instrument zur Qualitätssicherung im Schweizerischen Landesforstinventar

Berthold Traub

Eidgenössische Forschungsanstalt für Wald, Schnee und Landschaft (CH)*

Fabrizio Cioldi

Eidgenössische Forschungsanstalt für Wald, Schnee und Landschaft (CH)

Christoph Duggelin

Eidgenössische Forschungsanstalt für Wald, Schnee und Landschaft (CH)

Wiederholungsaufnahmen als Instrument zur Qualitätssicherung im Schweizerischen Landesforstinventar

Zur Qualitätssicherung der Felderhebungen werden vom Schweizerischen Landesforstinventar (LFI) Wiederholungsaufnahmen durchgeführt. Hierfür werden circa 10% der Probeflächen im Laufe einer Feldsaison ein zweites Mal vollständig aufgenommen. Die Ergebnisse dieser Wiederholungsaufnahmen werden in der vorliegenden Untersuchung genutzt, um zu prüfen, wie genau sich verschiedene Baum- und Bestandesmerkmale erfassen, also reproduzieren lassen. Ausserdem wird geprüft, ob sich die Umstellung von periodischen Felderhebungen (LFI1–LFI3) zu kontinuierlichen (LFI4) positiv auf die Reproduzierbarkeit der Merkmale ausgewirkt hat. Hierfür werden aktuelle Ergebnisse der Wiederholungsaufnahmen des LFI4 (2009/2017) mit denen des LFI3 (2004/2006) verglichen. Neben verschiedenen statistischen Masszahlen dienen vom LFI-Instruktorenteam gutachtlich festgelegte Toleranzbereiche und Qualitätsziele als Referenz zur Beurteilung der Reproduzierbarkeit. Die Ergebnisse der für die Vorratsschätzung wichtigen Baummerkmale sind unterschiedlich. Für den Brusthöhendurchmesser (BHD) entspricht das Ergebnis dem gesetzten Qualitätsziel. Für den Durchmesser auf 7 m Höhe und die Baumhöhen liegt es im LFI4 circa 5% darunter. Bei den sieben ebenfalls analysierten Bestandesmerkmalen wird bei deren vier das Qualitätsziel übertroffen (Bestandesalter, Bestandesstabilität, Deckungsgrad der gesicherten Verjüngung und Entwicklungsstufe). Beim Mischungsgrad, bei der Bestandesstruktur und beim Schlussgrad liegt das Ergebnis 5 bis 18% unter dem Qualitätsziel der Instruktoren. Beim Merkmal «Gehölzartenpräsenz» zeigt sich, dass die Erfassung grösserer Pflanzen (ab 1.30 m Höhe) deutlich besser reproduzierbar ist als die kleineren Pflanzen (40–130 cm Höhe). Im LFI4 hat sich die Reproduzierbarkeit tendenziell für fast alle untersuchten Merkmale verbessert. Die Ergebnisse legen den Schluss nahe, dass die veränderte Organisation der Feldaufnahmen mit nur drei Aufnahmegruppen und kontinuierlichen Feldaufnahmen im LFI4 einen positiven Einfluss auf die Reproduzierbarkeit der erhobenen Merkmale hat.

Keywords: quality assurance, forest inventory, repeat survey, observer agreement, measurement repeatability
doi: 10.3188/szf.2016.0118

* Zürcherstrasse 111, CH-8903 Birmensdorf, E-Mail berthold.traub@wsl.ch

Das Schweizerische Landesforstinventar (LFI) ist ein wichtiges Instrument für die Waldbeobachtung der Schweiz. Die ersten drei Erhebungen (LFI1–LFI3) wurden periodisch in Abständen von zehn Jahren durchgeführt. Mit der vierten Erhebung, die 2009 begann und 2017 enden wird, wurde in einen kontinuierlichen Aufnahme-modus gewechselt. Jedes Jahr wird systematisch über die ganze Schweiz verteilt ein Neuntel der gesamten Stichprobe erhoben (Panel). Das LFI verfügt über eine langjährige Zeitreihe von Daten, die regelmässig analysiert und publiziert werden (LFI3: Brändli 2010, LFI4 [2009/2013]: www.lfi.ch). Die Daten stammen zu einem wesentlichen Teil aus der Aufnahme von über 6500 terrestrischen Probeflächen pro Inventurperiode und einer daran gekoppelten Umfrage

beim lokalen Forstdienst. Weiter werden auch Luftbilder und GIS-Daten genutzt. Die Qualität dieser Zeitreihendaten ist von grosser Bedeutung, weil von ihr mit abhängt, wie genau der Zustand und die Veränderung von Kenngrössen des Schweizer Waldes geschätzt werden können. Das LFI verwendet konzentrische Probekreise mit 200 und 500 m² Grösse als Probeflächen. Eine Probefläche repräsentiert circa 200 ha Wald, ein gemessener Probebaum je nach Brusthöhendurchmesser (BHD) circa 4000 bis 10000 Bäume. Mit diesen «Hochrechnungsfaktoren» werden unweigerlich auch alle Messungenauigkeiten multipliziert. Daher ist es entscheidend, den Anteil an unechten Veränderungen, die beispielsweise durch methodische Anpassungen oder durch Ungenauigkeiten bei der Erfassung der Merkmale entste-

Merkmal	Skala	Beschreibung
BHD (cm)	i	Brusthöhendurchmesser: Durchmesser des Probebaumes in 1.3 m Höhe über Boden (nur stehend lebende Bäume, mit der Kluppe gemessen), auf cm abgerundet.
D7 (cm)	i	Durchmesser in 7 m Höhe (nur Tarifprobabäume, Messbereich bis 60 cm), gemessen.
Baumhöhe (m)	i	Baumhöhe (nur Tarifprobabäume), gemessen.
Bestandesalter (Jahre)	i	Alter des massgebenden Bestandes in Jahren (Schätzung oder Zählung von Jahrringen an Stöcken oder Zählen von Astquirilen).
Kronenlänge (3 Klassen)	o	Die Krone reicht vom ersten grünen Ast, der noch im Zusammenhang mit der Krone steht, bis zum Gipfel.
Schicht (3 Klassen)	o	Schichtzugehörigkeit des einzelnen Probebaumes: Lage der Baumkrone im Bestand in Bezug auf die Oberhöhe (mittlere Höhe der 100 stärksten Bäume pro Hektare). Ohne Überhälter.
Bestandesstabilität (10 Klassen)	o	Die Bestandesstabilität ist die zu erwartende Beständigkeit des massgebenden Bestandes gegenüber störenden Einflüssen für einen Zeitraum von 10 (Mittelland, Jura, Voralpen) bzw. 20 Jahren (Alpen, Alpensüdseite).
Mischungsgrad (4 Klassen)	o	Der Mischungsgrad wird geschätzt aufgrund des Grundflächenverhältnisses von Nadelbäumen zu Laubbäumen.
Deckungsgrad der gesicherten Verjüngung (7 Klassen)	o	Deckungsgrad der gesicherten Verjüngung des massgebenden Bestandes. Als gesicherte Verjüngung gelten alle Pflanzen ab 1.3 m Höhe bis 11.9 cm BHD der Baumarten gemäss LFI-Artenliste.
Entwicklungsstufe (5 Klassen)	o	Die Entwicklungsstufe wird mit dem dominanten Brusthöhendurchmesser (mittlerer BHD der 100 stärksten Bäume pro Hektare) des massgebenden Bestandes bestimmt. Klassen 1–5 (ohne gemischt).
Bestandesstruktur (4 Klassen)	n	Vertikaler Aufbau des massgebenden Bestandes. Die Bestandesstruktur wird durch die Anteile der verschiedenen Schichten definiert.
Schlussgrad (8 Klassen)	n	Mass für die gegenseitige Bedrängung der Baumkronen eines Bestandes (Kronenschluss). Der Schlussgrad bezieht sich auf die herrschende Schicht des massgebenden Bestandes, sofern diese 20% Deckungsgrad erreicht.
Gehölzarten	n	Abgeleitetes Merkmal. Aus den Rohdaten zur Gehölzartenpräsenz wird abgeleitet, wie häufig die Gehölzartenfunde übereinstimmen. Es wird zwischen zwei Höhenklassen (Pflanzen von 40–129 cm und ≥ 130 cm) unterschieden. Mehr Details dazu im Text.

Tab 1 Liste der untersuchten Merkmale. Skala: i = Intervallskala, o = Ordinalskala, n = Nominalskala.

hen, möglichst gering zu halten. Insbesondere bei der Beurteilung von kleinen Veränderungen und deren Signifikanz ist es also unerlässlich, Daten genau und immer auf die gleiche Weise zu erfassen.

Die Bedeutung von Messfehlern bei der Ermittlung des Vorrats wurde bereits im Rahmen der Entwicklung der statistischen Methoden des LFI2 (1993/1995) untersucht (Gertner & Köhl 1992, Köhl 2001). In den sogenannten «error budgets» für die Schätzung des Tarifvolumens wurde dabei unter anderem gezeigt, dass zufällige Messfehler einen relativ geringen Einfluss auf den Gesamtstichprobenfehler haben, während der Einfluss der systematischen Messfehler sehr deutlich ist und mit der Grösse des Stichprobenumfangs zunimmt. Mithilfe umfangreicher Massnahmen zur Sicherung und Verbesserung der Datenqualität versucht man seit dem LFI1 (1983/1985), insbesondere systematische Messfehler zu vermeiden. Dafür müssen die Merkmale und die Methoden zu ihrer Erfassung möglichst präzise definiert und über lange Zeiträume gleich instruiert werden.

Die vorliegende Untersuchung beschäftigt sich mit zwei Fragen zur Qualität der LFI-Daten:

1) Wie gut können die Ergebnisse der Felddaten für ausgewählte Merkmale des LFI reproduziert werden?

2) Hat sich die Reproduzierbarkeit nach dem Wechsel von periodischen Felderhebungen (LFI1–LFI3) zu kontinuierlichen (LFI4) verbessert?

Um diese Fragen zu beantworten, wurden die im Rahmen des LFI3 und des LFI4 durchgeführten Wiederholungsaufnahmen analysiert. Die Ergebnisse erlauben einen Rückschluss auf die Reproduzierbarkeit, indem überprüft wird, wie genau die wiederholte Erfassung eines Merkmals durch zwei verschiedene Aufnahmegruppen übereinstimmt. Dafür wurden Merkmale mit besonderer Relevanz für die Schätzung von Vorrat und Zuwachs sowie wichtige Merkmale zur Bildung von Klassierungen in den LFI-Ergebnistabellen gewählt. Eine Übersicht aller untersuchten Merkmale ist in Tabelle 1 zu finden. Die Ergebnisse dieser Untersuchung sollen zeigen, wo die Instruktion der Feldmitarbeiter noch verbessert werden kann und inwiefern sich die Organisation der Kontroll- und Wiederholungsaufnahmen des LFI optimieren lässt.

Felderhebungen und Qualitätssicherung im LFI

Für die Felddaten im LFI4 arbeiten von April bis November dauerhaft drei Aufnahmegruppen

Analysemethode		Skalenniveau		
		intervall	ordinal	nominal
Berechnung von Differenzen (Wert Wiederholungsaufnahme – Wert reguläre Aufnahme)		x	x	
Paarweiser t-Test/Standardabweichung der Differenzen		x		
Wilcoxon-Rangsummen-Test für paarweise Beobachtungen		x	x	
Symmetrie-Tests (2 Kl. McNemar, >2 Kl. Bowker-Test)			x	x
Korrelation	Pearson'scher Korrelationskoeffizient (r_p)	x		
	Spearman'scher Korrelationskoeffizient (r_s)		x	
Konkordanz	Cohen's Kappa		x	x

Tab 2 Angewendete Analysemethoden in Abhängigkeit des Skalenniveaus, auf dem die Merkmale untersucht werden.

pen (11 im LFI3) à zwei Personen und drei weitere Feldmitarbeiter, die bei Ferien oder Krankheit einspringen. Die Zusammensetzung der Aufnahmegruppen wird dreimal pro Saison eine Woche lang geändert, um den Austausch unter den Gruppen zu fördern. Die drei Aufnahmegruppen arbeiten schwerpunktmässig im Westen, im Osten und in der Mitte der Schweiz, damit die Fahrzeit ins Aufnahmegebiet klein gehalten werden kann. Die Anzahl Probeflächen (ca. 280 pro Jahr und Gruppe) und deren Höhenverteilung sind für die drei Aufnahmegruppen in etwa gleich.

Die Methoden zur Erhebung und die Definition der mittlerweile fast 300 Merkmale (inklusive Waldrand) sind in einer detaillierten Aufnahmeanleitung festgehalten (Keller 2013). Die Aufnahmegruppen erhalten zu Beginn und während der Feldsaison Trainings in Form von mehrtägigen Instruktionkursen, damit ein möglichst homogener Wissensstand und eine einheitliche Anwendung der Aufnahmeanleitung gewährleistet sind. Die LFI-intern entwickelte Software «MAIRA» unterstützt die Dateneingabe in die tragbaren Feldcomputer und enthält umfangreiche Plausibilitätschecks (Meile & Cuhe 2004).

Regelmässige Kontroll- und Wiederholungsaufnahmen dienen zur Überprüfung des Trainingserfolgs der Instruktionkurse und erlauben Rückschlüsse zur Reproduzierbarkeit der Daten. Für Kontrollaufnahmen besuchen LFI-Instruktoren Probeflächen unmittelbar nach der regulären Aufnahme und vergleichen Ergebnisse der Aufnahmegruppen mit ihren eigenen Messungen und Einschätzungen, die als Referenzwerte gelten. Unterschiedliche Interpretationen der Aufnahmeanleitung, Missverständnisse, aber auch bestehende Unklarheiten bei der Merkmalsdefinition oder Gerätefehler werden so schnell erkannt und durch unmittelbare Rückmeldung an die betroffene Gruppe korrigiert. Im LFI4 werden pro Jahr rund fünf Probeflächen pro Feldequipe kontrolliert (ca. 2% des Panels). Die hier analysierten Wiederholungsaufnahmen werden von den drei regulären Aufnahmegruppen durchgeführt und dienen zur Überprüfung, wie gut die Messungen und Beurteilungen von zwei

unabhängigen Aufnahmegruppen mit gleichem Ausbildungsstand übereinstimmen. Sie geben damit Hinweise auf den Erfolg der gesamten Instruktion- und Kontrollmassnahmen. Für die Wiederholungsaufnahmen werden circa 10% der Probeflächen eines Panels zufällig ausgewählt und an drei Terminen pro Saison – in der Regel im Juni, September und November – innerhalb einer Woche ein zweites Mal vollständig erfasst. Die Daten der regulären Aufnahme sind dabei nicht sichtbar.

Erste Ergebnisse zur Reproduzierbarkeit der Felderhebungen des LFI wurden von Kaufmann & Schwyzer (2001) und Herold & Ulmer (2001) veröffentlicht. Kontroll- und Wiederholungsaufnahmen als wichtiger Bestandteil von Qualitätssicherungsmassnahmen werden auch in vielen anderen nationalen Forstinventuren durchgeführt. In der Regel werden zwischen 5% und 10% der Probeflächen, mehrheitlich von speziell ausgebildeten «control teams» oder besonders erfahrenen Aufnahmegruppen, wiederholt aufgenommen (Tomppo et al 2010, Berger et al 2012, Kitahara et al 2009, Pollard et al 2006, Gasparini et al 2009).

Auswertungsmethoden für die Wiederholungsaufnahmen

Zur Beurteilung der allgemeinen Reproduzierbarkeit der Merkmale werden die Variabilität und Hinweise auf systematische Abweichungen zwischen den Ergebnissen der regulären Aufnahme und der Wiederholungsaufnahme untersucht. Verursacht werden die Abweichungen zwischen beiden Erhebungen in aller Regel durch zufällig oder systematisch ungenaue Messungen oder durch den Interpretationsspielraum, der in der Definition der Merkmale steckt. Allerdings können sich Baum- oder Probeflächenmerkmale zwischen den beiden Erhebungszeitpunkten auch tatsächlich verändern und somit zum Unterschied beitragen. Die Wahrscheinlichkeit hierfür wächst mit zunehmendem Zeitabstand zwischen der regulären Aufnahme und der Wiederholungsaufnahme. Dies gilt insbesondere für die Durchmesser- und Höhenmessungen, wenn die reguläre Mes-

sung vor den zuwachsstarken Monaten Juni und Juli stattfindet.

Aus statistischer Sicht handelt es sich bei der Analyse der Unterschiede um eine verbundene Stichprobe. Bestimmte Merkmale werden an identischen Objekten (Bäumen, Probeflächen) von zwei verschiedenen Beobachtern (den beiden Aufnahmegruppen) erfasst. Die Wahl der statistischen Analysemethoden ist mit dem Skalenniveau (intervall-, ordinal und nominal skaliert) verbunden, auf dem die Merkmale erhoben werden. Tabelle 2 zeigt die in dieser Untersuchung angewendeten Methoden im Zusammenhang mit dem Skalenniveau.

In Anlehnung an eine Methode der nationalen Forstinventur der USA (FIA) wird die Variabilität der Differenzwerte über den Anteil von Beobachtungen (Qualitätsziel QZ) in einem vorgegebenen Toleranzbereich analysiert (Pollard et al 2006, Kitahara et al 2009). Der Toleranzbereich und das Qualitätsziel für die hier vorliegende Untersuchung wurden gutachtlich durch das LFI-Instruktorenteam ausgehend von der Definition des Merkmals festgelegt. So wird beispielsweise für den BHD eine Genauigkeit von ± 1 cm festgelegt und erwartet, dass 98% der Differenzwerte in diesem Toleranzbereich liegen. Das Qualitätsergebnis (QE) ergibt sich aus der Analyse des vorliegenden Datenmaterials für die einzelnen Merkmale.

Der Pearsonsche (r_p) und der Spearmansche (r_s) Korrelationskoeffizient beschreiben die Stärke des Zusammenhangs zwischen den Werten der regulären Aufnahme und denjenigen der Wiederholungsaufnahme. Die Werte der Koeffizienten bewegen sich zwischen -1 (negative Korrelation) und 1 (positive Korrelation).

Ein weiteres Mass für die Beobachterübereinstimmung beziehungsweise Urteilskonkordanz ist das Cohen's Kappa (Formel 1). Der Anteil konkordanter Urteile (identische Klassifizierungen) wird mit dem Anteil der zufällig zu erwartenden, konkordanten Urteile verglichen.

$$\text{Kappa} = \frac{(P_o - P_e)}{(1 - P_e)} \quad 1)$$

Im Falle einer Vierfeldertafel ($i = 2; j = 2$, mit $i =$ Kategorie der regulären Aufnahme und $j =$ Kategorie der Wiederholungsaufnahme) ist der Anteil der beobachteten Übereinstimmung als $P_o = (a+d)/N$ definiert, mit $a = n_{11}$ und $d = n_{22}$ (n keine Veränderungen). Kappa = 1 ergibt sich, falls ausschliesslich konkordante Urteile auftreten ($P_o = 1$). Maximale Diskordanz (Kappa = -1) ergibt sich, falls alle Urteile diskordant ($P_o = 0$) und die vier Randsummen identisch sind ($P_e = 0.5$). Wenn die auftretenden Konkordanzen exakt dem Zufall entsprechen ($P_o = P_e$), ergibt sich ein Wert von Kappa = 0 (Bortz et al 2008).

Für die Interpretation der Kappa-Werte gilt es zu beachten, dass sie stark von der Anzahl Katego-

rien des untersuchten Merkmals abhängen und mehr Kategorien zu tendenziell kleineren Werten führen. Gleiches gilt für schiefe Verteilungen der Randsummen, also wenn die meisten Beobachtungen auf eine oder wenige Kategorien entfallen; diese ergeben ebenfalls tendenziell kleinere Kappa-Werte (Grouven et al 2007).

Für die Untersuchung systematischer Abweichungen wird die Hypothese geprüft, ob sich das arithmetische Mittel beziehungsweise der Median der Differenzen signifikant von null unterscheidet. Zur Beurteilung der Ergebnisse wird ausserdem das 95%-Konfidenzintervall des Mittelwerts angegeben. Ist der Wert 0 mit eingeschlossen, wird angenommen, dass der festgestellte Unterschied nicht signifikant ist. Für die Berechnung der Wilcoxon-Statistik werden Rangsummen der positiven und negativen Paar-Differenzen aus den Werten der regulären Aufnahme und der Wiederholungsaufnahme gebildet (Stahel 2009). Eine weitere Möglichkeit zur Überprüfung der symmetrischen Verteilung abweichender Beurteilungen ist der Bowker-Test (Walker 2002, Bortz et al 2008). Die Funktionsweise des Tests soll anhand des McNemar-Tests erläutert werden, der ein Spezialfall des Bowker-Tests für Vierfeldertafeln ist. Der Test berücksichtigt alle Fälle, in denen Veränderungen aufgetreten sind, und prüft die Nullhypothese, dass die Anteile der Veränderungen in beide Richtungen gleich beziehungsweise symmetrisch zur Hauptdiagonale ad sind. Die Nullhypothese wird anhand der Chi-Quadrat-Statistik nach Formel 2 bewertet.

$$\chi^2 = \frac{(b-c)^2}{b+c} \text{ mit } FG = 1 \quad 2)$$

mit $b = n_{12}$ und $c = n_{21}$ (n Veränderungen).

Da bei der Berechnung von Kappa nur die Hauptdiagonale berücksichtigt wird, ergänzen sich beide Ansätze zu einem gesamten Bild, denn Symmetrie bezüglich der Abweichungen zur Hauptdiagonale ist nicht gleichbedeutend mit einer hohen Konkordanz auf der Hauptdiagonale.

Für die Interpretation der Signifikanztests ist generell zu beachten, dass mit zunehmendem Stichprobenumfang immer kleinere Unterschiede als signifikant entdeckt werden können. Daher müssen signifikante Ergebnisse immer auf ihre Relevanz aus fachlicher Sicht überprüft werden. Signifikanztests für intervallskalierte Daten reagieren sensitiv auf Extremwerte beziehungsweise Ausreisser und auf Abweichungen von der Normalverteilung (insbesondere für Stichprobenumfänge mit $n < 30$). Daher ist es wichtig für intervallskalierte Daten, den robusteren Wilcoxon-Test mit in die Interpretation einzubeziehen. Welcher Stichprobenumfang notwendig ist, um einen bestimmten Unterschied als signifikant zu entdecken, und die Wahrscheinlich-

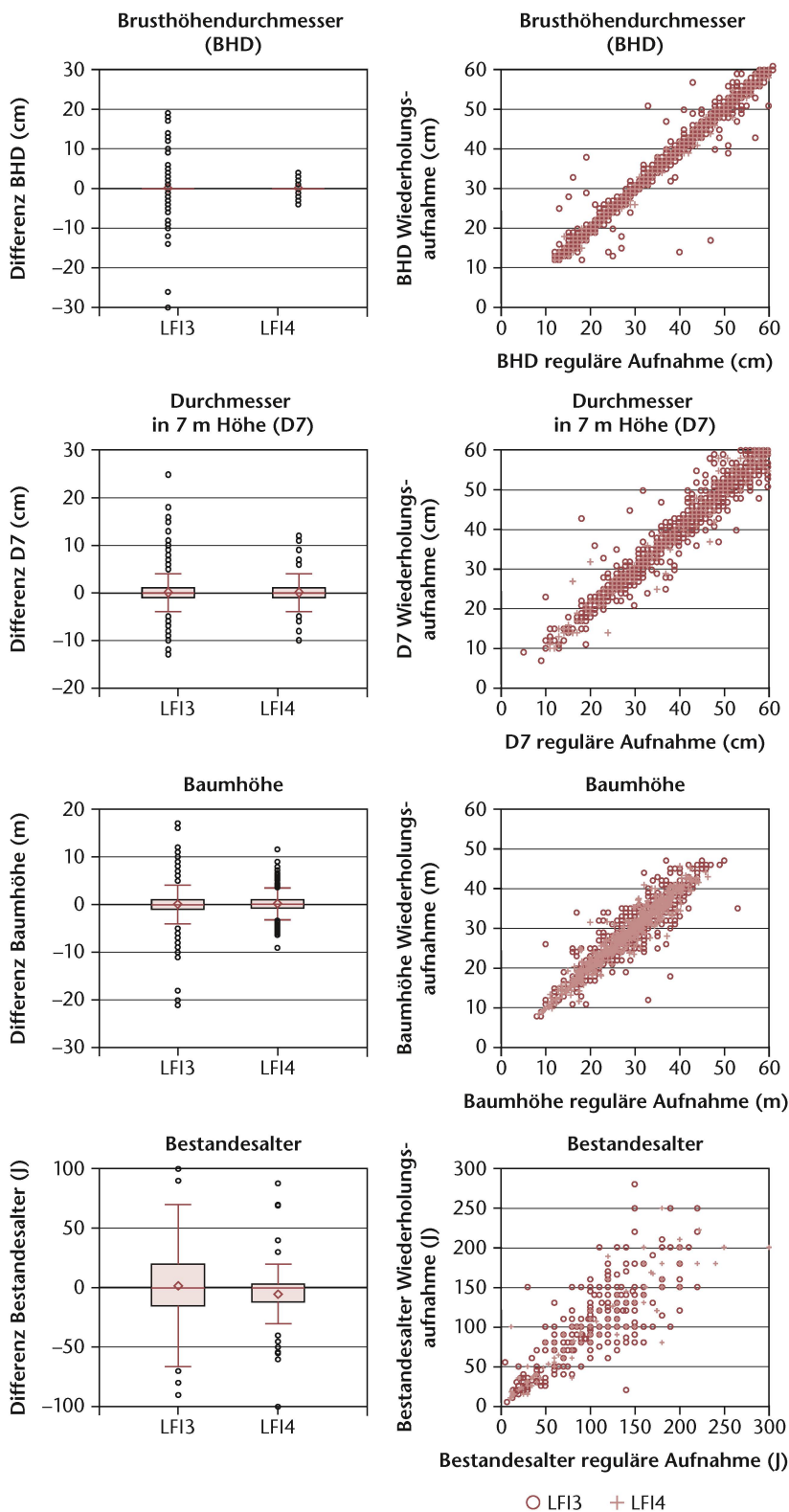


Abb 1 Box-Plots und Streudiagramme für die Merkmale Brusthöhendurchmesser (BHD), Durchmesser in 7 m Höhe (D7), Baumhöhe und Bestandalter. Differenz (bei Box-Plots): Wert Wiederholungsaufnahme minus Wert reguläre Aufnahme.

keit, dass ein solcher Unterschied bei gegebenem Stichprobenumfang auch entdeckt wird, kann im Rahmen einer Post-hoc-Poweranalyse untersucht werden (Cohen 1988). Alle statistischen Analysen wurden mit SAS Release 9.4 berechnet (SAS 2013, Walker 2002).

Daten

Die Felderhebungen des LFI3 wurden in den Monaten April bis November der Jahre 2004 bis 2006 durchgeführt. Die Daten des LFI4 stammen aus den Aufnahmejahren 2010 bis 2015. Im Jahr 2009 wurden aus organisatorischen Gründen keine Wiederholungsaufnahmen durchgeführt. Der Median des zeitlichen Abstandes der regulären Aufnahmen und der Wiederholungsaufnahmen liegt im LFI3 bei 37 Tagen und im LFI4 bei 43 Tagen. In die vorliegende Untersuchung wurden alle Probeflächen im zugänglichen Wald und im Gebüschwald einbezogen (621 im LFI3, 329 im LFI4). Da die Probeflächen eines Panels im LFI4 jeweils systematisch über die ganze Schweiz verteilt sind, können diese Daten mit denen des LFI3 verglichen werden, obwohl die Erhebungen des LFI4 noch nicht abgeschlossen sind.

Für die vorliegende Untersuchung wurden die Baummerkmale Brusthöhendurchmesser (BHD), Durchmesser auf 7 m Höhe (D7) und Baumhöhe auf der Intervallskala sowie die Kronenlänge und das Merkmal Schicht auf der Ordinalskala analysiert. Von den Merkmalen zur Bestandesbeschreibung wurden das Bestandesalter auf der Intervallskala, die Bestandesstabilität, der Mischungsgrad, der Deckungsgrad der gesicherten Verjüngung und die Entwicklungsstufe auf der Ordinalskala sowie die Bestandesstruktur und der Schlussgrad auf der Nominalskala untersucht. Als weiteres Merkmal wurde die Präsenz von Gehölzarten untersucht, die auf dem Zwei-Aren-Kreis der Probefläche in den Grössenklassen 9 (40–129 cm Höhe) und 10 (ab 130 cm Höhe und mit einem BHD von 0.1–11.9 cm) anhand der LFI-Artenliste erfasst werden. Für die Beurteilung der Merkmale zur Bestandesbeschreibung werden ausschliesslich Probeflächen verwendet, die übereinstimmend keine Bestandesgrenze aufweisen. Da sich die genannten Merkmale immer auf den massgebenden Bestand, also den Bestand, in dem das Probeflächenzentrum liegt, beziehen, können sich aufgrund unterschiedlicher Grenzfestlegung grosse Unterschiede ergeben, die im Rahmen dieser Untersuchung störend wirken würden. Die Datenpaare aller Merkmale wurden ohne Korrektur von Ausreissern oder Extremwerten übernommen, weil diese Teil der Realität sind, die untersucht werden soll. Lediglich einige wenige offensichtliche Zuordnungsfehler oder Verwechslungen wurden von der Datenanalyse ausgeschlossen.

Ergebnisse

Baumdaten

Das Qualitätsergebnis (QE-Wert) der BHD-Messungen entspricht den erwarteten Werten, beim D7 und der Baumhöhe liegt es im LFI4 etwa 5% un-

Merkmal	Toleranzbereich/ Qualitätsziel	LFI	QE	s	r _p	Mittelwert (KI)	T	W	n
Baummerkmale									
BHD	±1 cm/98%	3	96.97%	1.09	0.99	0.06 (0.03;0.09)	***	***	5973
		4	98.31%	0.57	0.99	0.05 (0.03;0.07)	***	***	3200
D7	±2 cm/95%	3	74.90%	3.09	0.97	0.12 (-0.10;0.34)			789
		4	89.74%	2.03	0.99	0.13 (-0.08;0.35)			351
Baumhöhe	±2 m/85%	3	80.86%	2.70	0.94	0.05 (-0.11;0.21)			1097
		4	80.03%	1.95	0.97	0.14 (-0.02;0.29)		*	641
Bestandesmerkmale									
Bestandesalter	±20 Jahre/75%	3	64.55%	32.52	0.79	1.49 (-2.43;5.40)			268
		4	80.51%	26.28	0.90	-5.38 (-10.17;-0.59)	*	*	118

Tab 3 Ergebnisse der statistischen Auswertungen für intervallskalierte Baum- und Bestandesmerkmale des dritten (LFI3) und vierten Landesforstinventars (LFI4). QE: Qualitätsergebnis im bezeichneten Toleranzbereich (fett = Ergebnis schlechter als Qualitätsziel); s: Standardabweichung der Differenzen; r_p: Pearsonscher Korrelationskoeffizient; Mittelwert (KI): arithmetisches Mittel der Differenzen mit 95%-Konfidenzintervall; T: t-Test; W: Wilcoxon Rangsummentest, Signifikanzniveau beidseitig: *** α < 0.001, ** α < 0.01, * α < 0.05; n: Anzahl Bäume bzw. Probeflächen.

ter den Erwartungen (Tabelle 3). Im Vergleich der beiden Inventuren zeigt sich für diese Merkmale ein Trend zu weniger Ausreißern und einer geringeren Streuung beziehungsweise Standardabweichung der Differenzen im LFI4 (Abbildung 1 und Tabelle 3). Insbesondere die Extremwerte bei den BHD-Differenzen haben deutlich abgenommen. Die Korrelationskoeffizienten und Streudiagramme zeigen den generell starken Zusammenhang zwischen den re-

gulären Messungen und den Wiederholungsmessungen. Die Mittelwerte der Differenzen des BHD, des D7 und der Baumhöhe sind alle positiv, die Werte der Wiederholungsaufnahme sind also tendenziell etwas höher. Signifikant ist dieser Unterschied aber nur beim BHD (Tabelle 3). Die tendenziell positiven Mittelwerte der Differenzen könnten mit dem Höhen- und Durchmesserwachstum der Bäume während der Vegetationsperiode erklärt werden. Der zeit-

Merkmal (Klassen)	Toleranzbereich/ Qualitätsziel	LFI	QE	r _s	Kappa (KI)	W	Sym	n
Baummerkmale								
Kronenlänge (3)	±0 Kl./85%	3	82.70%	0.67	0.62 (0.60;0.64)	*		6418
		4	86.29%	0.60	0.57 (0.54;0.61)			3508
Schicht (3)	±0 Kl./85%	3	88.83%	0.78	0.74 (0.72;0.75)	***	***	6634
		4	89.17%	0.83	0.78 (0.76;0.80)	*		3630
Bestandesmerkmale								
Bestandesstabilität (10)	±1 Kl./75%	3	75.38%	0.25	0.16 (0.10;0.23)			394
		4	77.55%	0.43	0.26 (0.18;0.34)			196
Mischungsgrad (4)	±0 Kl./90%	3	84.37%	0.93	0.87 (0.84;0.90)			364
		4	85.71%	0.93	0.88 (0.84;0.93)			196
Deckungsgrad der gesicherten Verjüngung (7)	±1 Kl./80%	3	84.03%	0.61	0.47 (0.41;0.54)			357
		4	89.44%	0.76	0.56 (0.49;0.63)			180
Entwicklungsstufe (5)	±1 Kl./90%	3	98.79%	0.84	0.75 (0.70;0.81)			247
		4	98.30%	0.94	0.88 (0.83;0.93)			118
Bestandesstruktur (4)	±0 Kl./75%	3	58.55%	–	0.36 (0.29;0.43)	–	–	386
		4	67.35%	–	0.46 (0.36;0.57)	–	–	196
Schlussgrad (8)	±0 Kl./65%	3	38.86%	–	0.27 (0.21;0.33)	–	–	386
		4	46.94%	–	0.37 (0.28;0.44)	–	–	196

Tab 4 Ergebnisse der statistischen Auswertung für ordinal und nominal skalierte Baum- und Bestandesmerkmale des dritten (LFI3) und vierten Landesforstinventars (LFI4). QE: Qualitätsergebnis im bezeichneten Toleranzbereich (fett = Ergebnis schlechter als Qualitätsziel); r_s: Spearmanscher Korrelationskoeffizient; Kappa (KI): Kappa mit 95%-Konfidenzintervall; W: Wilcoxon-Rangsummentest, Sym: Ergebnisse des McNemar- bzw. Bowker-Tests, Signifikanzniveau beidseitig: *** α < 0.001, ** α < 0.01, * α < 0.05; n: Anzahl Bäume bzw. Probeflächen.

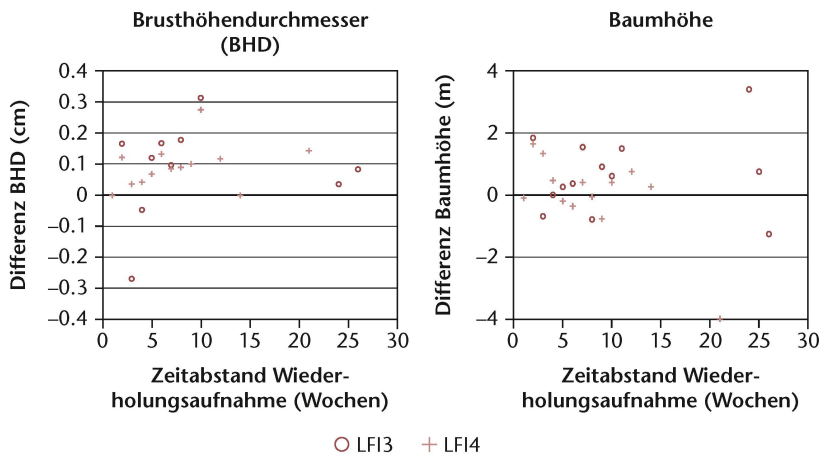


Abb 2 Differenzen der BHD- und Höhenmessung (Wochenmittel) in Abhängigkeit der Zeit zwischen den Messungen in Wochen. Datengrundlage: Bäume mit Erstmessung in den Monaten März bis Mai und Wiederholungsmessung in den Monaten Juni bis November. Positive Werte ergeben sich, wenn der Wert der Wiederholungsmessung grösser ist.

liche Abstand zwischen regulären Aufnahmen und Wiederholungsaufnahmen wäre dann positiv mit der Grösse der Differenzen korreliert. Für den BHD ist dieser Effekt sichtbar (Abbildung 2) und bestätigt die Ergebnisse des LFI2 (Kaufmann & Schwyzer 2001). Für die Baumhöhe kann für das LFI3 und das LFI4 hingegen kein bedeutender Effekt festgestellt werden (Abbildung 2).

Die Einschätzung des Merkmals Kronenlänge hat sich bezüglich der Korrelation und des Kappa-

Wertes tendenziell verschlechtert. Dennoch hat sich der QE-Wert im LFI4 um circa 3% verbessert und liegt jetzt bei den Erwartungen der Instruktoren. Im Falle des Merkmals Schicht haben sich alle statistischen Werte im LFI4 etwas verbessert, der QE-Wert dieses Merkmals liegt in beiden LFI jeweils über den Erwartungen der Instruktoren (Tabelle 4). Die Interpretation der Kappa-Werte ist für beide Merkmale wegen der extrem ungleichen Verteilung der Beobachtungen in den jeweiligen Kategorien problematisch. Für beide Merkmale liegen mehr als 80% der Beobachtungen in Klasse 1. Die Klasse 3 enthält jeweils weniger als 5% der Beobachtungen. Aus diesem Grund sind die Kappa-Werte lediglich für den Vergleich zwischen den LFI heranzuziehen. Da sich die Konfidenzintervalle von Kappa für das LFI3 und LFI4 überschneiden, kann nicht auf eine signifikante Verbesserung geschlossen werden. Die Tendenz zu systematischen Abweichungen hat sich im LFI4 für beide Merkmale verringert.

Bestandesmerkmale

Der QE-Wert für das intervallskalierte Merkmal Bestandesalter hat sich um 15% verbessert und liegt im LFI4 circa 5% über den Erwartungen der Instruktoren (Tabelle 3). Die Differenzen zeigen aber im LFI4 eine signifikante systematische Abweichung, die Werte der regulären Aufnahme sind tendenziell um fast fünf Jahre höher als diejenigen der Wiederholungsaufnahme. Die Werte des LFI3 zeigen keine systematische Abweichung, aber deutlich grössere Streuungen, mehr Extremwerte und einen weniger engen Zusammenhang ($r = 0.79$) als die Werte des LFI4 ($r = 0.90$; Abbildung 1 und Tabelle 3).

Im Falle der ordinal skalierten Merkmale Bestandesstabilität, Mischungsgrad, Deckungsgrad der gesicherten Verjüngung und Entwicklungsstufe liegt nur der QE-Wert für den Mischungsgrad unter dem erwarteten Wert (um ca. 5%; Tabelle 4). Bezüglich der Korrelation und Konkordanz der beiden Einschätzungen schneiden die Merkmale Mischungsgrad und Entwicklungsstufe mit Abstand am besten ab, gefolgt vom Deckungsgrad der gesicherten Verjüngung und dem Merkmal Bestandesstabilität (Tabelle 4 und Abbildung 3). Die Kappa-Werte aller untersuchten Bestandesmerkmale haben sich im LFI4 tendenziell verbessert, im Falle des Merkmals Entwicklungsstufe ist die Veränderung signifikant. Beim Vergleich der Ergebnisse zwischen den einzelnen Merkmalen muss beachtet werden, dass die vier Merkmale eine jeweils unterschiedliche Anzahl von Klassen aufweisen, also das Potenzial für abweichende Beurteilungen auch unterschiedlich ist. Der Algorithmus zur Berechnung der Kappa-Werte berücksichtigt diesen Umstand zu einem gewissen Grad.

Die QE-Werte für die nominal skalierten Merkmale Bestandesstruktur und Schlussgrad liegen mit

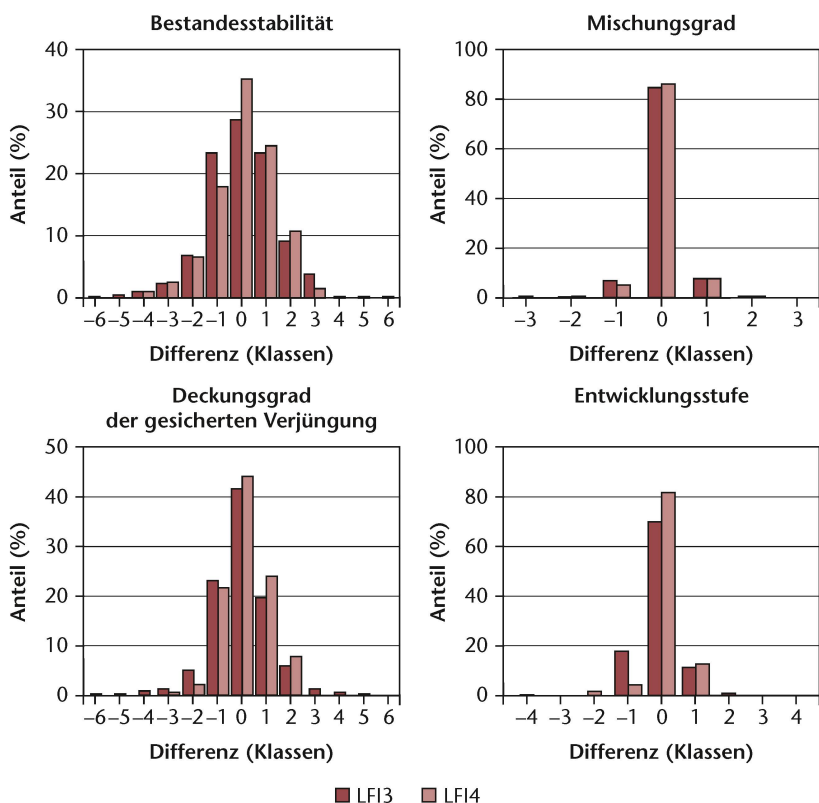


Abb 3 Verteilung der Differenzen der ordinal skalierten Merkmale Bestandesstabilität, Mischungsgrad, Deckungsgrad der gesicherten Verjüngung und Entwicklungsstufe. Differenz: Klasse Wiederholungsaufnahme minus Klasse reguläre Aufnahme.

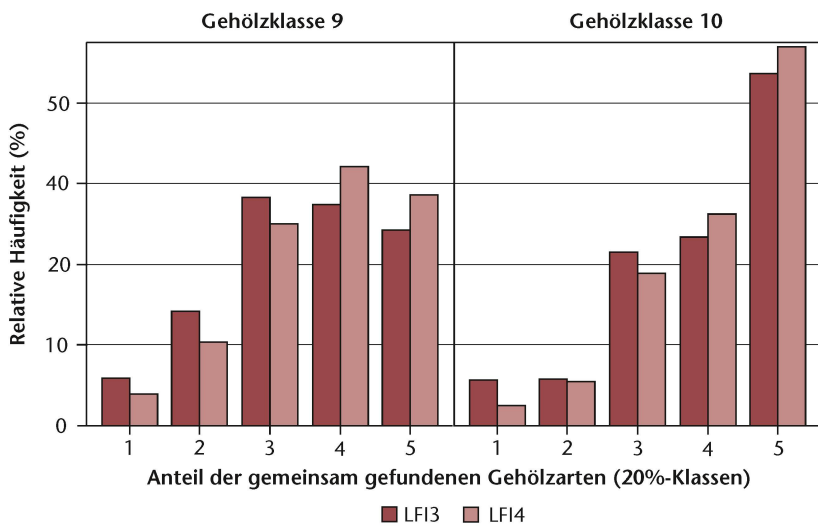


Abb 4 Häufigkeit der Anteile der gemeinsam gefundenen Pflanzenarten. Gehölzklasse 9: Höhe = 40–129 cm, Gehölzklasse 10: Höhe ≥ 130 cm und BHD zwischen 0.1 und 11.9 cm. Anteil der gemeinsam gefundenen Gehölzarten: 1 = [0–20%], 2 = [20–40%], 3 = [40–60%], 4 = [60–80%], 5 = [80–100%].

8% bei der Bestandesstruktur und 18% im Falle des Schlussgrades deutlich unter den erwarteten Werten der Instruktoren (Tabelle 4). Die Kappa-Werte für das LFI4 sind tendenziell etwas höher als für das LFI3, die Unterschiede sind aber aufgrund der sich überschneidenden Konfidenzintervalle als nicht signifikant zu beurteilen. Wie bereits bei den Merkmalen Kronenlänge und Schicht erwähnt, ist auch für diese Merkmale zu berücksichtigen, dass die Häufigkeit der Beobachtungen sehr ungleich auf die einzelnen Kategorien des jeweiligen Merkmals verteilt ist.

Die Ergebnisse der Symmetrietests sind für alle ordinal skalierten Bestandesmerkmale nicht signifikant. Somit ergibt sich für diese Bestandesmerkmale kein Hinweis auf systematische Abweichungen.

Für die Analyse der Gehölzartenpräsenz wurde für jede Probefläche das Verhältnis der An-

zahl gemeinsam gefundener Arten zur Anzahl insgesamt von beiden Aufnahmegruppen gefundener Arten berechnet. Diese (Übereinstimmungs-)Anteile wurden in 20%-Klassen eingeteilt und die prozentuale Häufigkeitsverteilung dieser Klassen untersucht (Abbildung 4). Für die Gehölzklasse 9 (40–129 cm Höhe) wird auf der Hälfte aller Probeflächen eine Übereinstimmungsrate von besser als 65% erzielt (Median = 66.6 im LFI4), für die Gehölzklasse 10 (ab 130 cm Höhe und mit BHD von 0.1–11.9 cm) ergeben sich Werte von besser als 75% auf der Hälfte der Probeflächen (Median = 77.7 im LFI4). Im Vergleich der beiden LFI ist der Median im LFI4 tendenziell etwas höher, die Übereinstimmung also etwas besser.

Diskussion und Ausblick

In Bezug auf die Reproduzierbarkeit wurden für das LFI4 zufriedenstellende Ergebnisse gefunden. Bei den gemessenen Baummerkmalen divergieren die BHD-Messungen der regulären Aufnahmen und der Wiederholungsaufnahmen nur in geringem Mass. Die Reproduzierbarkeit des D7 hat sich im LFI4 zwar deutlich verbessert, liegt aber wie das Merkmal Baumhöhe immer noch 5% unter dem erwarteten Qualitätsziel. Bei den Bestandesmerkmalen, die auf Einschätzungen beziehungsweise Klassifizierungen beruhen, ist die Übereinstimmung erwartungsgemäss etwas geringer als bei den gemessenen Baummerkmalen. Auffallend sind die systematischen Abweichungen beim Merkmal Bestandesalter und die Ergebnisse für die Merkmale Bestandesstruktur und Schlussgrad, bei denen das Qualitätsergebnis mehr als 10% unter den Erwartungen der Instruktoren liegt. Die Reproduzierbarkeit bei der Präsenz von Gehölzarten liegt im Mittelfeld. Die Übereinstimmung



Abb 5 Beispiel für Veränderungen im Vegetationszustand zwischen regulärer Aufnahme am 23. April 2012 (links) und der Wiederholungsaufnahme am 5. Juni 2012 (rechts).

bei den grösseren Pflanzen ab 1.30 m Höhe ist deutlich besser als bei den kleineren Pflanzen. Möglicherweise werden grössere Pflanzen eher gefunden und eindeutiger identifiziert.

Zusammenfassend zeigen die Untersuchungsergebnisse, dass der Wechsel von periodischen zu kontinuierlichen Felderhebungen mit der Reduzierung von elf auf drei Aufnahmegruppen offenbar zu einer tendenziell verbesserten Reproduzierbarkeit beigetragen hat. Die kontinuierlich durchgeführte Arbeit von Mitarbeitern mit langjähriger Felderfahrung, verbunden mit einer geringen Fluktuation sowie angepasste Instruktionen und Kontrollmassnahmen sind wichtige Faktoren in diesem Zusammenhang.

Wiederholungsaufnahmen sind für das LFI ein wichtiges Instrument zur gesamthaften Überprüfung aller Instruktionen und Kontrollmassnahmen. Die Ergebnisse können als «Bruttomass» für die Reproduzierbarkeit interpretiert werden, indem sie zeigen, wie homogen gemessen beziehungsweise geschätzt wird. Eindeutige Rückschlüsse auf systematische Abweichungen einzelner Gruppen oder Personen sind allerdings nur mit Kontrollaufnahmen durch Instruktorenteams möglich, da sie den Vergleich der Ergebnisse der Feldteams mit einem Referenzwert (Kontrollaufnahme) erlauben. Eine Ursache für Abweichungen zwischen den regulären Aufnahmen und den Wiederholungs- bzw. Kontrollaufnahmen ist die tatsächliche Veränderung von Merkmalen zwischen den beiden Aufnahmezeitpunkten (Abbildung 5). Diese sollte möglichst ausgeschlossen werden, indem die Wiederholungs- oder Kontrollaufnahmen direkt im Anschluss an die reguläre Aufnahme durchgeführt würden.

Die gefundenen Ergebnisse bedürfen einer sorgfältigen Analyse. Beispielsweise ergeben sich im Falle des in zehn Klassen eingeteilten Merkmals «Bestandesstabilität» niedrige Korrelations- und Konkordanzwerte, obschon – wie von den Instruktoren erwartet – über 75% der Abweichungen in einem Bereich von lediglich +/- einer Klasse liegen. Möglicherweise lässt die Definition des Merkmals eine höhere Präzision gar nicht zu. An der Einteilung in zehn Klassen sollte aber festgehalten werden, denn eine Reduzierung auf beispielsweise die Hälfte der Klassen würde lediglich zu einer Entdifferenzierung der Beurteilung führen. Die Entscheidung für eine bestimmte Klasse würde somit eher erschwert als erleichtert.

Wie an diesem und anderen Beispielen zu sehen ist, haben Merkmalsdefinitionen und deren Instruktion auch ihre Grenzen. Entscheidend für die Analyse der LFI-Ergebnisse ist letzten Endes die Kenntnis, mit welcher Präzision ein Merkmal geschätzt oder gemessen werden kann, und ob eine Tendenz zu systematischen Abweichungen vorliegt. Der Wert der Zeitreihe sollte nicht durch wiederholte

Anpassung der Merkmalsdefinition infrage gestellt werden.

Abschliessend ist festzustellen, dass die Ergebnisse der Wiederholungsaufnahmen substanziell zur Beurteilung der Qualität der Zeitreihendaten des LFI beitragen. Sie liefern zudem wertvolle Informationen für die Gestaltung der Instruktionkurse und für die Weiterentwicklung der Definitionen und Erhebungsvorschriften des Merkmalskatalogs. Von Bedeutung ist auch, dass die Wiederholungsaufnahmen die Feldteams motivieren, die Baum- und Bestandesmerkmale möglichst präzise gemäss Definition und Instruktion zu erfassen. Da nur Kontrollaufnahmen den Vergleich mit Referenzwerten ermöglichen, ist gleichwohl eine Reorganisation des heutigen Systems von Wiederholungs- und Kontrollaufnahmen in Betracht zu ziehen; so könnten eventuell die Wiederholungsaufnahmen zugunsten der Kontrollaufnahmen reduziert werden. Auf jeden Fall sollten die Ergebnisse der Kontrollaufnahmen systematisch und regelmässig analysiert werden. ■

Eingereicht: 19. Oktober 2015, akzeptiert (mit Review): 29. März 2016

Literatur

- BERGER A, GSCHWANTNER T, GABLER K, SCHADAUER K (2012)** Analysis of tree measurement errors in the Austrian National Forest Inventory. *Austrian J For Sci* 129: 153–181.
- BORTZ J, LIENERT GA, BOEHNKE K (2008)** Verteilungsfreie Methoden in der Biostatistik. Heidelberg: Springer. 929 p.
- BRÄNDLI UB, EDITOR (2010)** Schweizerisches Landesforstinventar. Ergebnisse der dritten Erhebung 2004–2006. Birmensdorf: Eidgenöss Forsch.anstalt WSL. 312 p.
- COHEN J (1988)** Statistical power analysis for the behavioral sciences. New York: L. Erlbaum Associates. 567 p.
- GASPARINI P, BERTANI R, DE NATALE F, DI COSMO L, POMPEI EP (2009)** Quality control procedures in the Italian National Forest Inventory. *J Environ Monitor* 11: 761–768.
- GERTNER G, KÖHL M (1992)** An assessment of some nonsampling errors in a national survey using an error budget. *For Sci* 38: 525–538.
- GROUVEN U, BENDER R, ZIEGLER A, LANGE S (2007)** Der Kappa-Koeffizient. *Deutsche Medizinische Wochenschrift* 132 (S 01): 65–68.
- HEROLD A, ULMER U (2001)** Stand stability in the Swiss National Forest Inventory: assessment technique, reproducibility and relevance. *For Ecol Manage* 145: 29–42.
- KÄUFMANN E, SCHWYZER A (2001)** Control survey of the terrestrial inventory. In: Brassel P, Lischke H, editors. *Swiss National Forest Inventory: Methods and models of the second assessment*. Birmensdorf: Swiss Federal Research Institute WSL. pp. 114–124.
- KELLER M (2013)** Schweizerisches Landesforstinventar. Feldaufnahme-Anleitung 2013. Birmensdorf: Eidgenöss Forsch.anstalt WSL. 214 p.
- KITAHARA F, MIZOUE N, YOSHIDA S (2009)** Evaluation of data quality in Japanese National Forest Inventory. *Environ Monit Assess* 159: 331–340.
- KÖHL M (2001)** Error sources and their Influence on the NFI Results. In: Brassel P, Lischke H, editors. *Swiss National Forest Inventory: Methods and models of the second assessment*. Birmensdorf: Swiss Federal Research Institute WSL. pp. 297–305.

- MEILE R, CUCHE L (2004) Neue Konzepte für Feldapplikation, Synchronisierung und Datenbanken beim dritten Schweizerischen Landesforstinventar. In: Wunn U, Quednau HD, editors. Die Grüne Reihe. Deutscher Verband Forstlicher Forschungsanstalten. Sektion Forstliche Biometrie und Informatik. pp. 245–248.
- POLLARD JE, WESTFALL JE, PATTERSON PL, GARTNER DL, HANSEN M ET AL (2006) Forest inventory and analysis national data quality assessment Report for 2000 to 2003. Fort Collins/CO: USDA Forest Service, General Technical Report RMRS-GTR-181. 43 p.
- SAS (2013) Base SAS 9.4 procedures guide: Statistical procedures. Cary/NC: SAS Institute Inc, 2 ed. 550 p.
- STAHEL WA (2009) Statistische Datenanalyse: eine Einführung für Naturwissenschaftler. Wiesbaden: Vieweg. 418 p.

- TOMPPO E, GSCHWANTNER T, LAWRENCE M, MCROBERTS RE, EDITORS (2010) National Forest Inventories. Dordrecht: Springer Netherlands. 612 p.
- WALKER GA (2002) Common statistical methods for clinical research with SAS examples. Cary/NC: SAS Institute Inc. 464 p.

Quellen

- Abegg M, Brändli UB, Cioldi F, Fischer C, Herold-Bonardi A et al (2014) Viertes Schweizerisches Landesforstinventar – Ergebnistabellen und Karten im Internet zum LFI 2009–2013 (LFI4b). Birmensdorf: Eidgenöss. Forschungsanstalt WSL. www.lfi.ch/resultate (16.3.2016)

La répétition d'une partie des relevés est la base du contrôle de qualité dans l'Inventaire forestier national suisse

Pour contrôler la qualité des données, environ 10% des surfaces échantillonnées par l'Inventaire forestier national suisse (IFN) sont relevés entièrement une deuxième fois au cours de la même année. Les résultats comparatifs de ces relevés doubles sont utilisés dans la présente étude pour examiner la fiabilité du recensement de différents critères d'inventaire caractérisant soit les arbres soit les peuplements, autrement dit leur reproductibilité. Par ailleurs, elle examine si le passage de relevés de terrain périodiques (IFN1–IFN3) à des relevés continus (IFN4) a amélioré la reproductibilité des critères d'inventaire. Pour ceci, les résultats comparatifs des relevés répétés de l'IFN4 (2009/2017) sont mis en regard de ceux de l'IFN3 (2004/2006). Différentes mesures statistiques, ainsi que des plages de tolérance et des objectifs de qualités validés par l'équipe d'instructeurs IFN servent de référence pour l'évaluation de la reproductibilité. Les résultats des trois mesures faites sur les arbres, fondamentales pour l'estimation du volume de bois, diffèrent entre eux. Pour le diamètre à hauteur de poitrine (DHP), le résultat correspond à l'objectif de qualité fixé. Pour le diamètre à une hauteur de 7 m et pour la hauteur des arbres, il se situe environ 5% en dessous. Parmi les sept attributs de peuplement analysés, quatre ont dépassé l'objectif de qualité (âge du peuplement, stabilité du peuplement, degré de couverture du rajeunissement garanti et stade de développement). Pour le degré de mélange, la structure du peuplement et le degré de fermeture, le résultat se situe de 5 à 18% en dessous de l'objectif de qualité des instructeurs. Pour le paramètre «présence d'espèces ligneuses», il apparaît que la détermination des plantes d'une certaine taille (à partir de 1.30 m) est mieux reproductible que celle des petites plantes (40–130 cm). Dans l'IFN4, la reproductibilité s'est tendanciellement améliorée pour presque toutes les caractéristiques étudiées. Les résultats permettent de conclure que la modification de l'organisation des relevés de terrain, avec uniquement trois équipes de terrain et une campagne de relevés continue sur neuf ans, exerce un impact positif sur la reproductibilité des caractéristiques collectées.

Repeat surveys as a quality assurance tool in the Swiss National Forest Inventory

The Swiss National Forest Inventory (NFI) repeats surveys to guarantee the quality of fieldwork. To this end, approximately 10% of sample plots are completely surveyed a second time over a field season. Based on the results of the repeat survey, the current investigation focuses on the assessment precision, i.e. the reproducibility of various tree and stand attributes in NFI4. It also investigates whether the change from periodic (NFI1–NFI3) to continuous (NFI4) fieldwork has had a positive effect on the reproducibility of the attributes. The current results of the repeat surveys for NFI4 (2009/2017) are compared with those for NFI3 (2004/2006) to this end. We used statistical measures as well as measurement quality objectives (MQO) set by the NFI instructor team as a reference for evaluating reproducibility. The results vary for tree attributes which are vital for estimating stock. The result for the diameter at breast height (dbh) corresponds to the expected values, while that for upper stem diameter at seven meters height and tree height were approximately 5% below the expected values. With regard to the seven stand attributes also analyzed, four of them exceeded the quality goals (stand age, stand stability, the degree of cover of secured regeneration, and stage of development). The results for the mixture proportion, the stand structure and crown closure were between 5 and 18% below MQO. The result for presence of woody species shows that the recording of larger plants (above 130 cm) is clearly more reproducible than for smaller plants (40–130 cm). In NFI4, the reproducibility for almost all studied attributes was improved. The results suggest that the modified structure of fieldwork (with only three field teams and continuous fieldwork in NFI4) has a positive influence on the reproducibility of the included attributes.