

Qualitätssicherung der Terrestrische Waldschadenserhebung (TWE)

Abstimmung in der Beurteilung der Kronenverlichtung zwischen dem Personal der Forschungsanstalt für Waldökologie und Forstwirtschaft Rheinland-Pfalz (FAWF) sowie Schulung und Abstimmung des Aufnahmepersonals der TWE in der Beurteilung der Kronenverlichtung

Friedrich Engels

Forschungsanstalt für Waldökologie und Forstwirtschaft Rheinland-Pfalz (FAWF)

Abt. Waldschutz

Abstract

Ein wesentlicher Punkt der Waldschadenserhebung ist die Sicherung der Datenqualität. Das für die Waldschadenserhebung eingesetzte Personal stimmt sich zu Beginn jeder Erhebung auf einen einheitlichen Beurteilungsmaßstab ab. Die Ergebnisse der Abstimmungsübung in Rheinland-Pfalz 2001 und 2002 wurden genutzt, um die Streuung der Beurteilung der Kronenverlichtung eines Baumes näher zu betrachten. Am Einzelbaum ist mit einem Schätzfehlerrahmen zu rechnen. Im Schnitt über ein Baumkollektiv ergeben sich Unterschiede im Schätzniveau. Um ein vergleichbares Schätzniveau zu erreichen, bedarf es der gemeinsamen Übung und Abstimmung. Je geübter und erfahrener die Beurteiler sind, je enger wird der Schätzfehlerrahmen und umso geringer sind die Niveauunterschiede. Von langjährig geübten und aufeinander abgestimmten Beurteilern darf erwartet werden, dass sie in ihrem Urteil zur Kronenverlichtung am Einzelbaum in über 90 % der Fälle nicht mehr als 10 Prozentpunkte voneinander abweichen. Über ein gemischtes Kollektiv von etwa 100 Bäumen ergeben sich im Mittleren Verlustprozent Unterschiede von bis zu 2 Prozentpunkten. Die Reihung mehrerer Einzelbäume nach dem Grad ihrer Kronenverlichtung erfolgt, auch von weniger geübten und aufeinander abgestimmten Beurteilern, immer außerordentlich einheitlich.

Schlagworte

Waldschadenserhebung, Waldzustandserfassung, Kronenverlichtung, Kronenzustand, Datenqualität, Beobachterfehler, Schätzgenauigkeit

Einleitung

Ein wesentlicher Punkt der Waldschadenserhebung ist die Sicherung der Datenqualität, in Rheinland-Pfalz wird hierfür ein Bündel von Maßnahmen ergriffen (ENGELS et al.1999). Der Qualifizierung und Abstimmung der an der Erhebung des Kronenzustandes beteiligten Personen kommt dabei eine Schlüsselrolle zu. Abstimmungskurse finden regelmäßig auf europäischer, bundesdeutscher wie rheinland-pfälzischer Ebene statt. Nur durch sorgfältige und stetige Übung wird die, immer auch mit subjektiven Elementen behaftete, qualifizierte Schätzung des Nadel-/Blattverlustes der Probestämme in der Zeitreihe und zwischen verschiedenen Beurteilern untereinander vergleichbar. Die Vergleichbarkeit der Beurteilungsergebnisse ist wesentliche Voraussetzung für die zusammenfassende Auswertung der erhobenen Daten zum Waldzustand.

Methodik der Abstimmung

Mit der fachlichen Durchführung der Waldzustandsüberwachung in Rheinland-Pfalz sind an der FAWF mit der landesweite Übersichtserhebung (TWE) und der Erhebung der intensiv untersuchten Dauerbeobachtungsflächen (DBF) je eine Arbeitsgruppe aus zwei Forstleuten betraut. Diese vier Forstleute stimmen sich vor Beginn der Außenaufnahmen auf ein einheitliches Schätzniveau ab. Hierzu erfolgen zunächst gemeinsame Übungen zur Kronenzustandsansprache, um das Verfahren zu rekapitulieren. Eine erste Abstimmung erfolgt auf dem Abstimmungskurs des Bundes und der Länder, an dem jedoch nicht alle vier des betroffenen Personals der FAWF teilnehmen können. Die endgültige Abstimmung erfolgt auf den zu diesen Zweck in Rheinland-Pfalz angelegten sogenannten Schulungs- und Kontrollparcours. Die hierfür ausgewählten Beispielbäume sind überwiegend ältere, aber auch einige jüngere Fichten, Buchen, Kiefern und Eichen, dazu einzelne Lärchen, Douglasien, Tannen und Birken. Die Kronen dieser Bäume sind relativ gut einsehbar. Die Bäume sind so ausgewählt, dass ein möglichst breites Spektrum der Kronenverlichtung abgedeckt wird und möglichst viele der für die Kronenzustandsansprache wesentlichen Merkmale und Symptome vorgestellt werden können. Die Schulung des Aufnahmepersonals muss zügig durchgeführt werden und sich daher auf das Wesentliche beschränken. Besonderheiten und weniger häufige Baumarten können somit nicht erschöpfend, sondern nur an einzelnen Beispielen abgehandelt werden. Wegen der gezielten Auswahl stellen die Parcoure keine repräsentative Stichprobe der Waldbäume in Rheinland-Pfalz dar.

Das Personal der FAWF geht die Parcoure zur Abstimmung gemeinsam ab. Jeder Beispielsbaum wird von jedem Beurteiler unabhängig bewertet. Als Basis dient ihm dabei die eigene Erfahrung, die auf den Bund-Länder Abstimmungskurs und in der Zusammenarbeit mit den Kollegen gewonnen wurde, sowie die Anweisungen der Arbeitsgemeinschaft Dauerbeobachtungsflächen-Waldschäden (AG DBF 1) mit der Bilderserie (AG DBF 2). Im Jahr 2001 wurden die Ergebnisse dieser Einzelbewertungen unmittelbar nach der Bewertung jedes einzelnen Baumes besprochen. Anschließend wurde für diesen Beispielsbaum dann ein abgestimmter Vorgabewert festgelegt. Auf diese Weise bestand eine unmittelbare Rückkoppelung nach jeder einzelnen Beurteilung. Die Beurteiler werden so während der Abstimmungsübung fortlaufend auf ein einheitliches Schätzniveau getrimmt. Durch diese Vorgehensweise darf ein Höchstmaß an Übereinstimmung erwartet werden. Die auftretenden Abweichungen zwischen den Beurteilern wären damit als Untergrenze der schätzungsbedingten Streuung zu werten. Im Jahr 2002 ist jeder Beurteiler die Parcoure für sich abgegangen und hat sämtliche Bäume alleine bewertet. Die letzte Abstimmung der Beurteiler lag dabei ein Jahr zurück. Es darf damit erwartet werden einen Hinweis auf die Abweichungen zu bekommen, die sich im Laufe der jährlichen Außenaufnahmen ergeben. Erst nach Abschluss wurden die Daten verglichen, die einzelnen Bäume ggf. noch einmal gemeinsam begutachtet, besprochen und ein abgestimmter Vorgabewert festgelegt. So fand im Jahr 2002, anders als 2001, keine unmittelbare Rückkoppelung während der Abstimmung, sondern erst im Anschluss statt.

Dieselben Parcoure dienen dann der Qualifizierung des durch die Landesforsten für die TWE zur Verfügung gestellten Aufnahmepersonals. Ein wesentliches Ziel ist dabei das Aufnahmepersonal einheitlich auf das von der FAWF vorgegebene Schätzniveau einzueichen. Der Vorgabewert der FAWF richtet sich dabei nach den Ergebnissen des Bund-Länder Abstimmungskurses. Auf den Schulungsparcours werden zunächst die Symptome der Kronenverlichtung und die Vorgabewerte der FAWF vorgestellt und erläutert. Anschließend geht das Aufnahmepersonal den Kontrollparcour selbständig ab und beurteilt jeden Beispielsbaum. Dabei werden bereits die für den gesamten Erhebungszeitraum vorgesehenen

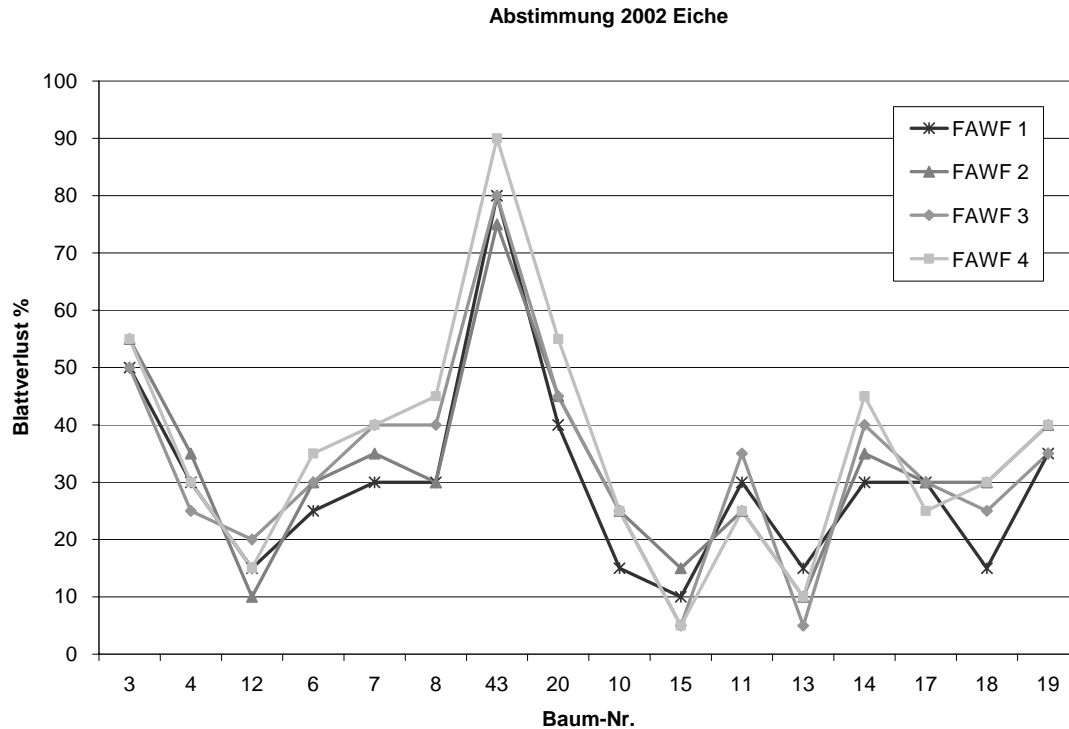
Aufnahmeteams gebildet, so dass die Bereichsleiter jeder für sich und das Aufnahmeteam als Zweiergruppe arbeitet. Die selbstständige Ansprache der Bäume des Kontrollparcours dient dem Aufnahmepersonal als wichtiger Hinweis auf das eigene Niveau der Schätzung der Kronenverlichtung. Im Vergleich zu den Vorgabewerten der FAWF kann jeder erkennen wie seine eigene Ansprache liegt. Die Erhebung des Kontrollparcours ist also als wesentlicher Bestandteil der Qualifizierung und Abstimmung anzusehen und keine Kontrolle des Abstimmungserfolges im engeren Sinne. Alle Teams und Bereichsleiter bearbeiten zur gleichen Zeit den Kontrollparcour, Witterungs- und Lichtverhältnisse sind damit für alle identisch. Es wird zwar versucht die Personen möglichst gleichmäßig über den Kontrollparcour zu verteilen, es kommt jedoch immer wieder zu Berührungen während der Übung, so dass eine gewisse gegenseitige Beeinflussung in Einzelfällen nicht ausgeschlossen werden kann. Gleichwohl werden im Folgenden diese Daten herangezogen, um die Übereinstimmung des Aufnahmepersonals bei der Kronenzustandsansprache zu bewerten.

Im Jahr 2001 nahmen an der Erhebung der Vollstichprobe mit 464 Aufnahmepunkten 29 Forstleute (12 Aufnahmeteams und 5 Bereichsleiter) teil. In 2002 waren für die 163 Aufnahmepunkte der Unterstichprobe 13 Forstleute (5 Aufnahmeteams und 3 Bereichsleiter) ausreichend. Diese waren alle bereits 2001 beteiligt und verfügten somit über wenigstens einjährige Erfahrung in der TWE. Dies wurde ausgenutzt, um mit dem Verfahren vertraute, aber wenig erfahrene Beurteiler in ihrem Schätzverhalten vor und nach der Abstimmungsübung zu vergleichen. Hierfür wurde der Kontrollparcour einmal gleich zu Beginn des Abstimmungskurses und einmal in Anschluss der Übungen auf den Schulungsparcours aufgenommen. In 2001 wurde der Kontrollparcour nur einmal nach erfolgter Abstimmung aufgenommen. Bei der gemeinsamen Besprechung des Kontrollparcours war 2001 wie 2002 von einzelnen Aufnahmeteams auch festzustellen, dass Baumverwechslungen aufgetreten sind. Aus den Daten offensichtlich ist 2001 eine Verwechslung beim Jungwuchs Baum Nr. 1, aber auch für Buche Nr. 10 ist die deutlichste Abweichung auf eine Verwechslung zurückzuführen. Bei den hier vorgenommenen Auswertungen wurde nur die Verwechslung 2001 bei dem Jungwuchsbaum Nr. 1 ausgeklammert, da ein extrem abweichender Wert zustande kam. Die andere bekannte Baumverwechslung verblieb im Auswertungskollektiv, da auch bei den Außenarbeiten mit einzelnen Baumverwechslungen zu rechnen ist. Die Qualifizierung und Abstimmung ist damit keineswegs beendet. Der Aufnahme des Kontrollparcours folgt die gemeinsame Aufnahme einiger TWE-Punkte von Aufnahmeteams und Bereichsleiter. Während der TWE besteht zudem eine fortlaufende Rückkoppelung, falls es erforderlich wird, erfolgt auch eine Nachschulung.

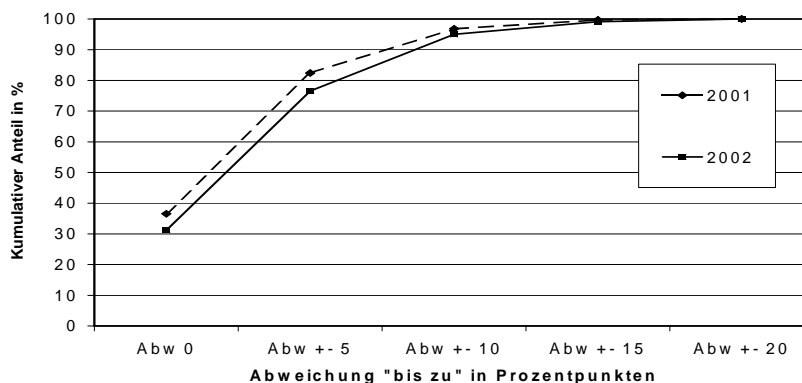
Die vorliegenden Daten wurden mit Hilfe vom MS-Excel und des Statistikprogramms SPSS (DIEHL et al. 2001) ausgewertet. Die graphische Darstellung der Bewertung jedes einzelnen Beispielbaumes gibt einen guten Eindruck darüber, wie die Beurteiler zueinander liegen. Die gewählte Form der Liniengraphik ist zwar nicht korrekt, da zwischen den auf der X-Achse benachbarten Bäumen keine direkte Beziehung besteht, die Werte jedes Beurteilers lassen sich so aber besser als bei einer Punktgraphik verfolgen. Für jeden Beurteiler wurden die Mittelwerte berechnet, untereinander gepaart und die Differenzen mit dem T-Test auf ihre Signifikanz mit 5 % Irrtumswahrscheinlichkeit geprüft. Die Übereinstimmung der Beurteiler hinsichtlich der Reihung nach Stärke der Verlichtung wurde mit dem Kendall-W-Test (Kendalls Konkordanzkoeffizient) geprüft. Die vier Beurteiler der FAWF wurden untereinander, das Aufnahmepersonal im Vergleich zur Vorgabe der FAWF betrachtet.

Ergebnisse der Abstimmung des Personals der FAWF

Im Jahr 2001 wurden 109 Bäume (25 Buchen, 26 Eichen, 39 Fichten und 19 Kiefern) bewertet, im Jahr 2002 waren es 93 Bäume (20 Buchen, 16 Eichen, 27 Fichten, 23 Kiefern und 7 Lärchen).



Die graphische Darstellung der Ergebnisse der Beurteilung am einzelnen Beispielsbaum (oben für Eiche 2002, für alle siehe Anlage) zeigt für die vier Beurteiler bei allen Baumarten einen engen, gleichgerichteten Verlauf. In 2001 liegt keiner der Beurteiler auf einem systematisch niedrigeren oder höheren Schätzniveau. Im Vergleich dazu ist in 2002 eine etwas größere Spreitung der Linien und die in einzelnen Fällen auftretende Neigung zu einem systematisch abweichenden Schätzniveau (bei Eiche und Fichte) sichtbar. Es zeigt sich aber auch, dass diese Abweichungen nicht für alle Beispielsbäume gleichermaßen auftreten.



Abweichung bis zu in Prozentpunkten	± 0	± 5	± 10	± 15	± 20
Kumulativer Anteil 2001	36,5 %	82,4 %	96,8 %	99,7 %	100,0 %
Kumulativer Anteil 2002	31,2 %	76,5 %	95,0 %	99,1 %	100,0 %

Abweichungen in der Bewertung der Kronenverlichtung jedes der Beispielsbäume zwischen den einzelnen Beurteilern der FAWF

Für die Betrachtung der Abweichungen in der Beurteilung der Kronenverlichtung wurden die Werte der vier Beurteiler jeweils wechselseitig verrechnet, was bei den sich ergebenden sechs Fällen pro Beispielsbaum noch praktikabel ist. Auf diese Weise werden die Differenzen der Beurteilung von Bäumen zwischen erfahrenen Beurteilern offengelegt. Es zeigt sich, dass hier in beiden Jahren Abweichungen über 10 Prozentpunkte nur sehr vereinzelt und solche über 20 Prozentpunkte nicht auftreten.

Baumart Anzahl	Alle 109	Buche 25	Eiche 26	Fichte 39	Kiefer 19
FAWF 1	28,76	28,80	33,08	31,28	17,63
FAWF 2	27,94	28,80	31,92	30,26	16,58
FAWF 3	28,81	31,60	30,96	30,77	18,16
FAWF 4	28,39	30,00	31,73	30,77	16,84
Übereinstimmung	0,939	0,935	0,920	0,932	0,914

Mittleres Verlustprozent von Beispielsbäumen nach Bewertung durch die mit Kronenzustandsbonituren betrauten Forstleute der FAWF und ihre Übereinstimmung (Kenndalls Konkordanzkoeffizient) 2001

Das Mittlere Verlustprozent zeigt 2001 Differenzen zwischen den Beurteilern von bis zu 2,8 Prozentpunkten bei der Buche. In etwa der Hälfte der Fälle liegen die Abweichungen unter 1 Prozentpunkt. Die Differenzen zwischen den durch zwei verschiedene Beurteiler festgestellten mittleren Verlustprozenten sind in nur zwei Fällen (Buche Paarungen FAWF 1 / FAWF 3 und FAWF 2 / FAWF 3) signifikant (T-Test für zwei verbundene Stichproben bei 5 % Irrtumswahrscheinlichkeit).

Baumart Anzahl	Alle 93	Buche 20	Eiche 16	Fichte 27	Kiefer 23	Lärche 7
FAWF 1	28,01	32,75	30,00	30,56	22,61	17,86
FAWF 2	28,12	32,25	32,81	31,67	20,22	17,86
FAWF 3	30,22	32,75	33,13	36,30	21,96	20,00
FAWF 4	28,82	31,50	35,63	33,89	19,13	17,86
Übereinstimmung	0,955	0,956	0,914	0,955	0,918	0,932

Mittleres Verlustprozent von Beispielsbäumen nach Bewertung durch die mit Kronenzustandsbonituren betrauten Forstleute der FAWF und ihre Übereinstimmung (Kenndalls Konkordanzkoeffizient) 2002

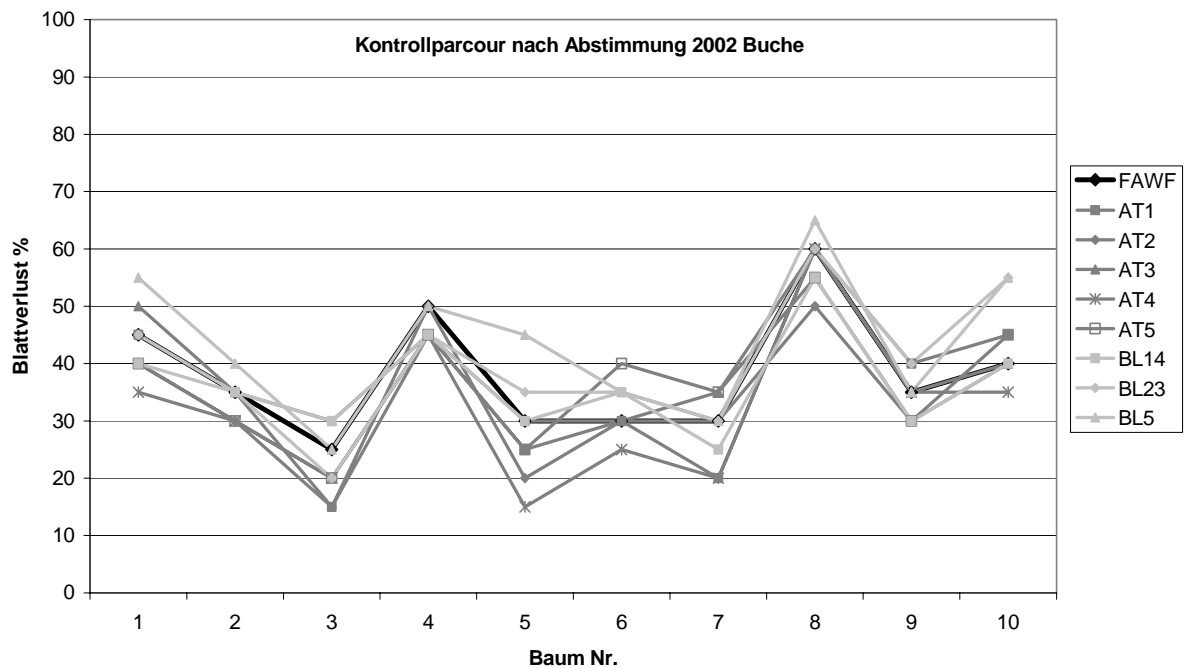
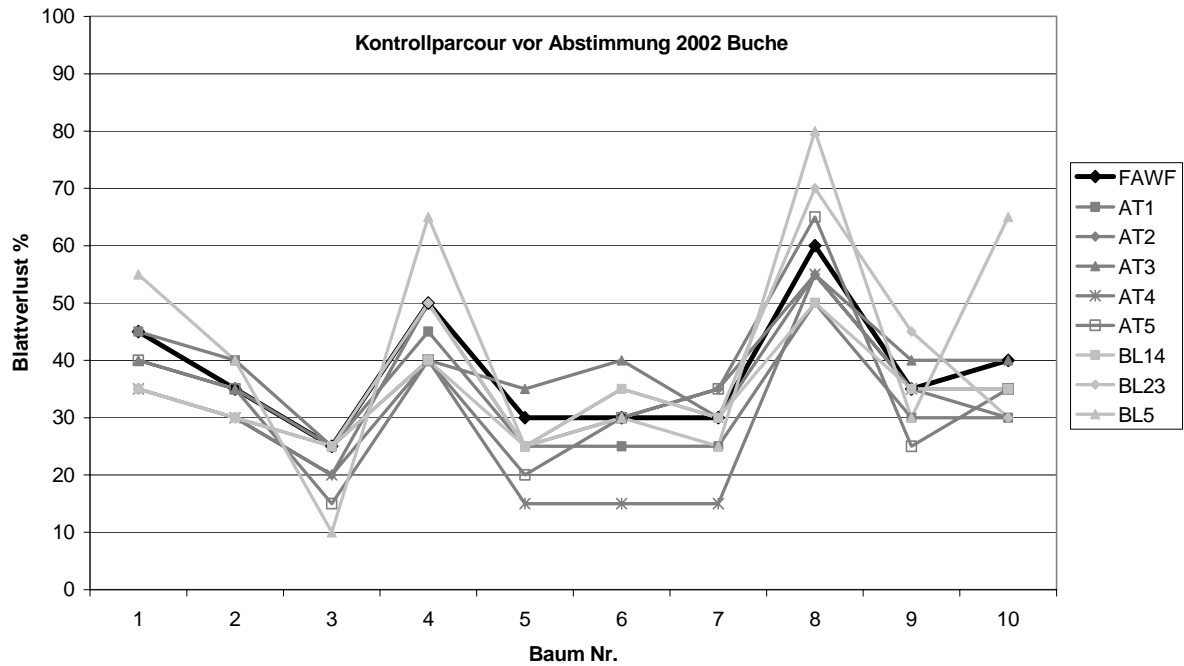
Im Jahre 2002 beträgt die Differenz des zwischen zwei verschiedenen Beurteilern festgestellten Mittleren Verlustprozent bis zu 5,63 Prozentpunkten. In 50 % der Fälle liegt sie unter 2 in 80 % unter 3 Prozentpunkten. Die festgestellten Differenzen sind über alle Baumarten ab 1 Prozentpunkt signifikant (T-Test für zwei verbundene Stichproben bei 5 % Irrtumswahrscheinlichkeit). Bei Fichte und Kiefer bestehen ab 2 Prozentpunkten signifikante Differenzen, für Eiche nur bei der höchsten Differenz von über 5 Prozentpunkten. Bei Buche und Lärche ergaben sich keine signifikanten Differenzen. Die 2002 aufgetretenen Differenzen im Mittleren Verlustprozent sind damit höher und häufiger signifikant als 2001.

Der Übereinstimmungskoeffizient liegt in beiden Jahren immer über 0,9 und ist damit außerordentlich hoch. Bei einer rein zufälligen Reihung ist bei vier Beurteilern ein Wert von bis zu 0,25 zu erwarten, der Maximalwert bei völliger Übereinstimmung wäre 1,0.

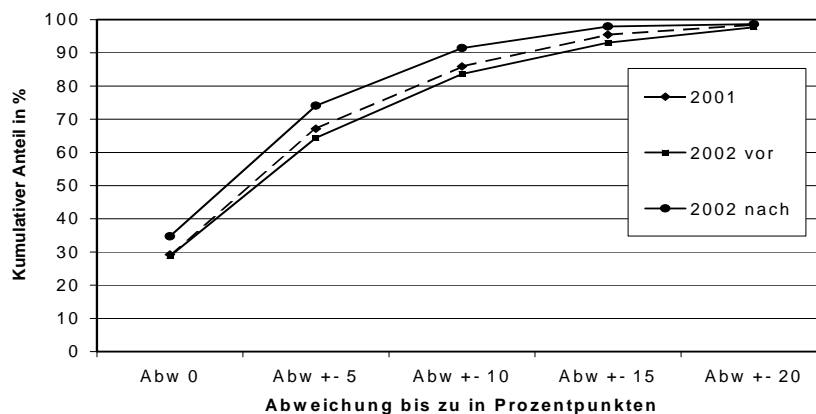
Wird für die gepaarten Stichproben eine Korrelation nach Pearson berechnet, so ergeben sich 2001 wie 2002 in allen Fällen hochsignifikante Korrelationen mit einem Bestimmtheitsmaß von wenigstens 79 % (2002) bzw. 83 % (2001), meist aber über 90 %.

Ergebnisse der Schulung und Abstimmung des Aufnahmepersonals der TWE

Der Kontrollparcour umfasste 53 Bäume in 2001 und 54 in 2002. Die Datengrundlage ist daher besonders bei der Betrachtung der einzelnen Baumarten nur gering (10 Buchen, 11 Eichen, 12 Fichten, 10 bzw. 11 Kiefern und 10 Jungwuchsbäume verschiedener Arten). Die Ergebnisse können daher auch schon durch eine einzelne extreme Abweichungen stark beeinflusst werden.



Die graphische Darstellung der Ergebnisse der Beurteilung am einzelnen Beispielsbaum (oben für Buche 2002 vor und nach der Abstimmung, für alle siehe Anlage) zeigt für das Aufnahmepersonal bei allen Baumarten eine gewisse Bandbreite der Kronenverlichtungsansprache, aber einen gleichgerichteten Verlauf. Im Jahr 2001 ist mit der höheren Anzahl an Beurteilern die Bandbreite der Schätzungen höher als 2002. In 2002 wiederum zeigt sich vor der Abstimmung eine höhere Bandbreite als danach. Aufgezeigt werden dabei die Niveauunterschiede in der Beurteilung der Kronenverlichtung. Die Abweichungen der einzelnen Aufnahmeteams / Bereichsleiter von der Vorgabe der FAWF werden verdeutlicht. So zeigt sich, ob eine Abweichung vom mittlerem Niveau durch eine geringe aber gleichgerichtete Abweichung bei mehreren Probenbäumen oder durch eine starke Abweichung einzelner Probebäume verursacht wird. Auch wird bei übereinstimmenden mittlerem Niveau erkennbar, ob tatsächlich nur geringe ungerichtete Abweichungen vorliegen oder starke Abweichungen in beide Richtungen sich zufälligerweise im Mittel gegenseitig ausgleichen.



Abweichung bis zu in Prozentpunkten	± 0	± 5	± 10	± 15	± 20
Kumulativer Anteil 2001	29,2 %	67,1 %	85,9 %	95,4 %	98,4 %
Kumulativer Anteil 2002 vor Abstimmung	28,7 %	64,4 %	83,6 %	93,1 %	97,7 %
Kumulativer Anteil 2002 nach Abstimmung	34,7 %	74,1 %	91,4 %	97,9 %	98,6 %

Abweichungen zwischen den Urteilen über die Kronenverlichtung der Bäume des Kontrollparcour der Aufnahmeteams bzw. Bereichsleiter und der Vorgabe der FAWF

Die Abweichungen jedes Aufnahmeteams bzw. Bereichsleiters zur Vorgabe der FAWF wurde für jeden einzelnen Beispielbaum berechnet. Auf diese Weise werden die Differenzen des Aufnahmepersonals zu dem vorgegebenen Maßstab der FAWF offengelegt. In Einzelfällen treten auch Abweichungen über 20 Prozentpunkte auf. Ein hoher Anteil bewegt sich jedoch im Bereich bis ±10 Prozentpunkten. Für 2002 wird deutlich, dass nach der Abstimmungsübung die Anteile in den Bereichen geringerer Abweichungen deutlich größer sind als davor. 2001 sind die Anteile im Bereich bis ± 10 Prozentpunkten geringer als 2002.

Baumart Anzahl	Alle 53	Buche 10	Eiche 11	Fichte 12	Kiefer 10	Jungwuchs 10
Vorgabe FAWF	25,47	28,50	33,64	37,08	23,00	2,00
Aufnahmeteam 1	26,51	22,50	37,73	37,92	25,00	6,00
Aufnahmeteam 2	25,47	26,00	30,45	37,08	21,50	9,50
Aufnahmeteam 3	21,60	20,00	28,18	31,67	21,00	4,50
Aufnahmeteam 4	20,96	23,89	29,55	27,08	20,50	2,00
Aufnahmeteam 5	23,40	25,50	32,27	35,00	19,00	2,00
Aufnahmeteam 6	23,96	20,50	31,82	36,67	26,50	1,00
Aufnahmeteam 7	23,11	21,00	36,82	29,58	23,00	2,50
Aufnahmeteam 8	22,26	19,50	28,18	32,92	25,50	2,50
Aufnahmeteam 9	24,53	19,50	33,64	40,42	23,50	1,50
Aufnahmeteam 10	27,26	25,00	34,09	37,92	31,00	5,50
Aufnahmeteam 11	27,17	28,50	35,45	32,08	31,50	6,50
Aufnahmeteam 12	26,73	22,00	30,45	30,83	33,00	15,00
Bereichsleiter 1	27,26	28,00	33,64	40,83	24,00	6,50
Bereichsleiter 2	28,87	25,55	47,27	35,83	23,50	9,00
Bereichsleiter 3	25,57	23,50	32,73	38,75	25,50	4,00
Bereichsleiter 4	24,25	26,00	28,18	35,42	23,00	6,00
Bereichsleiter 5	25,57	31,00	30,91	34,58	22,50	6,50
Übereinstimmung	0,855	0,709	0,853	0,913	0,782	0,329 *

Mittleres Verlustprozent der Bäume des „Kontrollparcour“ nach Bewertung durch das Aufnahmepersonal der TWE und ihre Übereinstimmung (Kenndalls Konkordanzkoeffizient) 2001

* hier ohne Aussagekraft, Erläuterung siehe Text

Bei dem mittleren Verlustprozent treten 2001 trotz vorheriger Abstimmungsübung noch Differenzen auf. Zwischen den Werten eines Aufnahmeteams und der Vorgabe der FAWF betragen sie über alle Baumarten bis zu 4,5 Prozentpunkte, liegen jedoch in zwei Drittel der Fälle unter 2 Prozentpunkten. Bei den einzelnen Baumarten erreicht die höchste Differenz 19,1 Prozentpunkte, liegt jedoch auch hier überwiegend unter 2 Prozentpunkten. Die Unterschiede im mittleren Verlustprozent über alle Baumarten sind bei 7 von 17 Fällen signifikant (T-Test für zwei verbundene Stichproben bei 5 % Irrtumswahrscheinlichkeit), dies sind in der Regel die Differenzen ab 2 Prozentpunkten. Bei den einzelnen Baumarten treten in 27 von 85 Fällen signifikante Differenzen auf, diese betragen dann über 4 Prozentpunkte.

Baumart Anzahl	Alle 54		Buche 10		Eiche 11		Fichte 12		Kiefer 11		Jungwuchs 10	
Vorgabe FAWF	28,24		38,00		34,55		40,42		22,27		3,50	
	vor	nach	vor	nach	vor	nach	vor	nach	vor	nach	vor	nach
Aufnahmeteam 1	26,57	27,59	34,00	36,00	35,91	34,09	38,33	40,83	19,55	22,27	2,50	2,00
Aufnahmeteam 2	22,78	23,52	34,50	33,50	27,73	32,27	25,00	24,58	19,55	19,55	6,50	7,00
Aufnahmeteam 3	24,63	26,11	38,00	39,00	30,00	32,27	33,75	32,92	16,36	20,00	3,50	5,00
Aufnahmeteam 4	21,39	24,26	29,50	32,00	26,36	31,82	25,83	35,42	20,91	18,18	3,00	1,50
Aufnahmeteam 5	23,61	29,63	34,00	36,50	25,00	40,00	32,92	36,25	21,82	27,73	2,50	5,50
Bereichsleiter 1	24,72	26,94	34,00	36,50	28,18	32,73	32,92	37,50	22,27	20,91	4,50	5,00
Bereichsleiter 2	26,76	27,78	37,50	40,00	30,00	32,27	36,25	38,75	22,73	20,91	5,50	5,00
Bereichsleiter 3	24,81	28,06	42,50	43,50	19,09	29,09	35,00	37,08	20,00	20,45	6,50	9,00
Übereinstimmung	0,909	0,923	0,826	0,880	0,895	0,888	0,899	0,908	0,799	0,877	0,503	0,346

Mittleres Verlustprozent der Bäume des „Kontrollparcour“ nach Bewertung durch das Aufnahmepersonal der TWE und ihre Übereinstimmung (Kenndalls Konkordanzkoeffizient, bei Jungwuchs ohne Aussagekraft, siehe Text) vor und nach dem Abstimmungskurs an der FAWF 2002

Auch 2002 treten bei dem mittleren Verlustprozent Differenzen auf. Zwischen den Werten eines Aufnahmeteams und der Vorgabe der FAWF betragen sie vor der Abstimmung in 29 % der Fälle bis zu 2, in 31 % über 5 Prozentpunkte, nach der Abstimmung liegen 50 % der Fälle im Bereich bis zu 2 und 15 % über 5 Prozentpunkte. Die Unterschiede im mittleren

Verlustprozent über alle Baumarten sind vor der Abstimmung in 7, nach der Abstimmung in 3 von 8 Fällen signifikant (T-Test für zwei verbundene Stichproben bei 5 % Irrtumswahrscheinlichkeit). Bei den einzelnen Baumarten treten vor der Abstimmung in 17, nach der Abstimmung in 10 von 50 Fällen signifikante Differenzen auf.

Der Übereinstimmungskoeffizient liegt in beiden Jahren mit Ausnahme des Jungwuchses über 0,7, meist um 0,9. Damit wird die These einer rein zufälligen Reihung der Probebäume hinsichtlich ihrer Kronenverlichtung klar abgelehnt. Der Übereinstimmungskoeffizient ist 2002 vor und nach der Abstimmung fast gleich hoch. Für eine rein zufälligen Reihung hätte der Übereinstimmungskoeffizient bei 18 Beurteilern (2001) einen Wert von 0,06 bei 9 Beurteilern (2002) einen Wert von 0,11 nicht überschreiten dürfen. Je näher der Wert bei 1,0 liegt, umso einheitlicher ist die Reihung durch die verschiedenen Beurteiler. Die deutlich niedrigeren Werte beim Jungwuchs waren zu erwarten, da hier etliche der Probebäume hinsichtlich der Kronenverlichtung einheitlich sind, also keine klare Reihung aufweisen.

Wird für die mit der Vorgabe der FAWF gepaarte Stichprobe des einzelnen Aufnahmeteams bzw. Bereichsleiters eine Korrelation nach Pearson berechnet, so ergeben sich meist signifikante Korrelationen. Auch hier ist der Jungwuchs wieder auszunehmen. Über alle Baumarten ist die Korrelation immer hochsignifikant. Das Bestimmtheitsmaß liegt 2001 bei mindestens 66 %, in der Regel um 85 %, 2002 bei mindestens 78 % in der Regel um 88 %. Im Jahr 2001 sind für die einzelnen Baumarten die Korrelationen bei Eiche und Fichte durchweg signifikant, bei Buche und Kiefer treten bei jeweils 2 von 17 Fällen nicht signifikante Korrelationen auf. Das Bestimmtheitsmaß ist allgemein niedriger und beginnt bei 50 % übersteigt aber in zwei Drittel der Fälle 66 %. Im Jahr 2002, mit weniger und erfahreneren Beurteilern, sind für die einzelnen Baumarten die Korrelationen ohne Ausnahme hochsignifikant; das Bestimmtheitsmaß liegt bei mindestens 63 %, meist aber über 80 %.

Wertung und Schlussfolgerungen

In der Betrachtung der Schätzung am Einzelbaum zeigen die Abweichungen der Beurteiler untereinander, respektive gegenüber der Vorgabe, dass eine absolut gleiche Bewertung auch bei engster Abstimmung nur in etwa einem Drittel der Fälle erreicht wird. Der Schätzfehlerahmen von ± 5 Prozentpunkten am Einzelbaum deckt aber rund zwei Drittel aller Fälle ab und Abweichungen über ± 10 Prozentpunkten treten nur in etwa einem Zehntel der Fälle auf. In Einzelfällen waren auch Ausreißer mit Abweichungen von über ± 20 Prozentpunkten zu verzeichnen. Diese grobe Bandbreite der Streuung der Ansprache am Einzelbaum deckt sich mit früheren Auswertungen anhand der Abweichungen zwischen Aufnahmeteams und Bereichsleiter im Zuge der Außenaufnahmen der TWE. Auch eigene Erfahrungen aus der wiederholten Ansprache des selben Einzelbaumes durch den selben Beurteiler zeigen, dass die Ansprache am Einzelbaum diesem Streurahmen unterworfen ist. Eine absolut idente Beurteilung der Kronenverlichtung durch verschiedene Beurteiler oder zu verschiedenen Zeitpunkten darf nur bei voll belaubten oder toten Bäumen erwartet werden. Ein gewisser Streurahmen muss akzeptiert und bei der Datenauswertung und Interpretation berücksichtigt werden. Naturgemäß sind höhere Abweichungen besonders im mittlerem Feld der Kronenverlichtung möglich. Es ist daher durchaus sinnvoll, dass die Schadstufe 2 „mittelstark geschädigt“ (30 % – 65 % Kronenverlichtung) breiter gefasst ist als die Schadstufen 0 und 1. Bei den Untersuchungen zeigt sich deutlich, dass die Abweichungen am Einzelbaum zwischen den erfahrenen Beurteilern geringer sind als beim Aufnahmepersonal und wiederum nach der Abstimmungsübung geringer sind als davor.

Zwischen den Beurteilern der FAWF bestehen trotz langjähriger Zusammenarbeit Unterschiede im Ansprachemaßstab, die in unterschiedlichen Mittleren Verlustprozenten ihren Ausdruck finden. Auch bei sehr enger Rückkoppelung während der Kronenzustandsansprache bleiben Differenzen bestehen, die als Untergrenze des Schätzfehlers zwischen einzelnen Beurteilern bewertet werden können. Liegt die Abstimmung schon länger zurück, wie es bei den Außenaufnahmen der TWE und an den DBF der Fall ist, und arbeiten die Beurteiler für sich alleine, treten höhere Unterschiede auf. Noch höhere Unterschiede im Ansprachemaßstab zeigen sich für das Aufnahmepersonal. Die Forstleute lernen zwar recht schnell die Symptomatik zur Kronenzustandsansprache der Waldbäume, es bedarf aber einiger eigener Übung und Erfahrung einen einheitlichen, vorgegebenen Maßstab des Blatt-/Nadelverlustes anzuwenden. Die Untersuchung des Jahres 2002 zeigt den Erfolg der Abstimmungsübung. Die Unterschiede lassen sich zwar nicht eliminieren und auch nicht auf ein so geringes Niveau, wie bei den Beurteilern der FAWF untereinander, drücken, sind in der Summe aber deutlich geringer geworden. Da der Stichprobenumfang für das Aufnahmepersonal mit rund 10 Bäumen jeder Art nur sehr gering ist, sind die hier gewonnenen Ergebnisse entsprechend vorsichtig zu bewerten. Auch nur ein einzelner Ausreißer kann hier einen starken Einfluss auf das Mittlere Verlustprozent haben.

Wie der Übereinstimmungskoeffizient nach KENDALL und die Korrelationen nach PEARSON zeigen, wird die Reihung der Beispielpflanzen nach den Grad ihrer Kronenverlichtung immer außerordentlich einheitlich vorgenommen. Für alle Baumarten, mit Ausnahme des Jungwuchses, zeigt sich eine sehr hohe Übereinstimmung. Das gilt nicht nur für die erfahrenen Beurteiler der FAWF, sondern auch für das Aufnahmepersonal, egal ob es nach der Schulung erstmalig an einer Waldschadenserhebung teilnimmt, wie im Jahr 2001, oder, wie 2002, mit einjähriger TWE-Erfahrung ohne vorherige Abstimmung. Nach der Abstimmung 2002 ist die Übereinstimmung nur unwesentlich besser. Relative Unterschiede in der Kronenverlichtung zweier Bäume werden also immer sehr zuverlässig wahrgenommen. Der Grund für die schlechtere Übereinstimmung bei dem Jungwuchs ist, dass die

ausgewählten Beispielbäume, anders als bei den anderen Baumarten, sehr wenig hinsichtlich ihrer Kronenverlichtung differenziert sind. Es bestehen entweder keine oder nur sehr geringe Unterschiede in der Kronenverlichtung, die damit vom Schätzfehler am Einzelbaum überdeckt werden. Eine Reihung ist so nicht sinnvoll, Übereinstimmungskoeffizient und Korrelation nur von geringem Aussagewert. Die deutlich niedrigeren Werte waren beim Jungwuchs zu erwarten.

Die Abstimmungskurse des mit Waldschadenserhebungen betrauten Fachpersonals finden zuerst auf der Ebene Bundesrepublik und danach Bundesland statt. Ihre Bedeutung für die Vergleichbarkeit der erhobenen Daten und die Sicherung der Qualität der Waldschadenserhebung kann nicht genug hervorgehoben werden. Unabhängige Zweiterhebungen dienen einerseits zur Abschätzung der Schätzfehler des Aufnahmepersonals und sind gleichzeitig während der Erhebung, bei Rückkoppelung zum Ersterheber, ein wichtiges Instrument den Schätzfehler zu minimieren.

Des weiteren zeigt sich, wie wichtig es ist, die mit der Kronenzustandsansprache betrauten Forstleute sorgfältig zu schulen und dabei auch eine selbstständige Ansprache im Vergleich mit anderen Beurteilern bzw. dem Vorgabemaßstab vorzunehmen. Die Qualifizierung und Abstimmung des Aufnahmepersonals ist in Rheinland-Pfalz mit der Aufnahme des sogenannten Kontrollparcours nicht beendet, sondern baut darauf auf. Jeder einzelne Beurteiler muss seine Schwächen und Unsicherheiten hinsichtlich der Kronenzustandsansprache unter Anwendung des Vorgabemaßstabes erkennen und durch weitere Übung abbauen. Hierfür sind eigenständige Ansprachen und Vergleiche mit den Ergebnissen der Beurteilung mit anderen in der Kronenzustandsansprache qualifizierten Forstleuten erforderlich. Nach der Ansprache des Kontrollparcours werden die Bäume, bei den gehäuft größere Abweichungen auftraten, gemeinsam besprochen. Meist sind es die Bäume bei denen eine Besonderheit des Kronenaufbaus die Ansprache erschwert. Auch betrachten die einzelnen Teams selbstständig die Bäume noch einmal, bei denen ihr Urteil eine größere Abweichung zur Vorgabe aufweist.

Bei der vergleichsweise geringen Anzahl an Beispielbäumen bei dem Abstimmungskurs sind in dieser Betrachtung erst größere Differenzen ab 2 bis 4 Prozentpunkten im Mittleren Verlustprozent signifikant. Bei den Außenaufnahmen ist die Zahl der Probebäume pro Hauptbaumart mit 30 bis über 100 auf einer Dauerbeobachtungsfläche und in der Übersichtserhebung TWE mit mehreren 100 bis zu 1000 pro Baumart deutlich höher. Rein stichprobentechnisch ist dann schon bei relativ geringen Differenzen von um ± 1 Prozentpunkt eine statistische Signifikanz erreicht. Die Nicht-Stichprobenfehler übersteigen damit den Stichprobenfehler und sind daher von wesentlicher Bedeutung. Das betont die Wichtigkeit der in Rheinland-Pfalz üblichen Maßnahmen zur Qualitätssicherung während der Erhebung. Grundsätzlich werden alle Kronenzustandsansprachen, egal ob auf Dauerbeobachtungsflächen oder bei der TWE in Zweierteams durchgeführt. So werden Fehleinschätzungen, die in Einzelfällen auch bei sehr geübten Beurteilern immer wieder vorkommen, durch den Teampartner weitgehend erkannt und vermieden. Individuelle Niveauunterschiede zwischen den beiden Teampartnern werden ausgeglichen. Die Aufnahme der Dauerbeobachtungsflächen findet nur durch langjährig erfahrenes und geübtes Personal der FAWF statt. Während der TWE wird durch die fortlaufende Betreuung der Aufnahmeteams und die unabhängige Zweiterhebung eines großen Teiles der Stichprobenbäume durch Forstleute mit mehrjähriger Erfahrung in der Waldschadenserhebung (Bereichsleiter) ein stetiger Abgleich zwischen den Beurteilern gesichert. Diese stetige Rückkoppelung erlaubt es ein Abgleiten vom vorgegeben Maßstab rechtzeitig zu erkennen und zu korrigieren. So können nicht nur umfangreich Nacharbeiten

vermieden werden, zu bedenken ist auch, dass zu spät erkannte Abweichungen gar nicht mehr zu korrigieren sind, da die Kronenzustandsansprache schon ab Ende August wegen der beginnenden Verfärbung einiger Baumarten nicht mehr sinnvoll möglich ist. Die Daten würden dann zweifelhaft bleiben und damit auch die daraus gewonnenen Ergebnisse. Die Bereichsleiter können als Einzelpersonen arbeiten, da hier der Vergleich mit dem Ergebnis des Aufnahmeteams Fehleinschätzungen des allein arbeitenden Beurteilers aufdeckt und deren Korrektur erlaubt. Bei der Auswertung der Daten der landsweiten TWE wird die räumliche Verteilung gesondert beachtet. Festgestellte Veränderungen werden auch dahingehend geprüft, ob einzelne Aufnahmeteams in auffälliger Weise von einem allgemeinen Trend abweichen. In einem solchen Fall wird weiter geprüft, inwieweit diese Besonderheit in dem Aufnahmebereich des Teams durch den Bereichsleiter bestätigt wird. In Zweifelsfällen werden zusätzliche Nacherhebungen durch die Inventurleitung vorgenommen. Auch diese Prüfung muss daher möglichst unmittelbar erfolgen.

Anhand der in dieser Untersuchung gewonnenen Daten wird auch deutlich, dass die Schätzgenauigkeit eines jeden Beurteilers je nach Baumart unterschiedlich ist. Eine Inhomogenität der Varianzen gleicher Beobachter auf verschiedenen Parcoursen und bei verschiedenen Baumarten stellt auch schon KÖHL 1993 fest, wobei er auf die Daten von über 400 Bäumen aus 4 verschiedenen Parcoursen von 9 Beobachtern zurückgreifen konnte. Im Vorfeld dieser Untersuchung fand er 1991 eine Vielzahl von Faktoren (z.B. Baumart, Altersstufe, Verlichtungsstufe, Wetter, Höhenlage), die die Ansprache während der Erhebung der Aufnahmepunkte beeinflussen. KÖHL (1991) lehnt eine Korrektur der Verlustschätzung ab, sofern nicht neben dem Beurteilereffekt auch alle anderen Einflussgrößen einfließen. Eine Eliminierung aller von ihm beschriebenen Einflussfaktoren durch Schulung oder Schätztechnik hält er für unmöglich. Es wird daher kaum möglich sein, anhand von Vergleichsbeurteilungen Korrekturfaktoren für einzelne Aufnahmeteams oder Länder zu bestimmen. Vielmehr ist es wichtiger, die Schätzgenauigkeit der Beurteiler, ihre Abweichungen und die damit verbundenen Unsicherheiten zu kennen, um die gewonnenen Daten entsprechend vorsichtig zu interpretieren. Bei der Auswertung hinsichtlich eines Bezugsgebietes darf erwartet werden, dass die Schätzfehler sich um so mehr gegenseitig ausgleichen, je bunter die Mischung nicht nur der Baumarten und Altersklassen ist, sondern auch wenn mehrere Aufnahmeteams das Bezugsgebiet bearbeitet haben. Je mehr bei der Auswertung stratifiziert wird, umso eher ist damit zu rechnen, dass Schätzfehler einzelner Aufnahmeteams auf das Ergebnis durchschlagen. Bei der Bewertung festgestellter Unterschiede einer räumlichen Verteilung oder in der zeitlichen Entwicklung dürfen Unterschiede von einem oder wenigen Prozentpunkten nicht zu hoch bewertet werden, sondern sind in einer längeren Zeitreihe zu bestätigen. Bei der Auswertung von bestimmten Altersstufen einer Baumart, seltenerer Baumarten, einzelner Aufnahmepunkten oder gar von Einzelbäumen ist die Zeitreihe besonders wichtig. Je länger und geschlossener diese Zeitreihe ist, um so sicherere Aussagen sind abzuleiten.

Literaturhinweise:

AG DBF (1) (Arbeitsgemeinschaft Dauerbeobachtungsflächen – Waldschäden, Bund-Länder-Arbeitsgruppe) (ohne Jahr): Kronenansprache auf Dauerbeobachtungsflächen, Arbeitsanweisung, nicht veröffentlicht.

AG DBF (2) (Arbeitsgemeinschaft Dauerbeobachtungsflächen – Waldschäden, Bund-Länder-Arbeitsgruppe) (ohne Jahr): Waldbäume, Bilderserien zur Einschätzung von Kronenverlichtung bei Waldbäumen. Verlag M. Faste.

DIEHL J.M., STAUFENBIEL T. (2001): Statistik mit SPSS, Version 10.0. Verlag Dietmar Kloetz.

ENGELS F., BLOCK J., WUNN U. (1999): Methodenbeschreibung Terrestrische Waldschadenserhebung (TWE) in Rheinland-Pfalz. Forschungsanstalt für Waldökologie und Forstwirtschaft Rheinland-Pfalz Selbstverlag 13 S.

KÖHL M. (1991): Waldschadensinventuren: mögliche Ursachen der Variation der Nadel-/Blattverlustschätzung zwischen Beobachtern und Folgerungen für Kontrollaufnahmen. Allgemeine Forst- und Jagd-Zeitung, 162. Jg., Heft 11/12, Seiten 210 bis 221.

KÖHL M. (1993): Quantifizierung der Beobachterfehler bei der Nadel-/Blattverlustschätzung. Allgemeine Forst- und Jagd-Zeitung, 164. Jg., Heft 5, Seiten 83 bis 92.

Anhang :

(Als gesonderte Datei abgelegt)

Kronenverlichtung von Beispielbäumen der Arten Buche, Eiche, Kiefer und Fichte aus der Ansprache der Schulungsparcours und des Kontrollparcours an der Forschungsanstalt für Waldökologie und Forstwirtschaft Rheinland-Pfalz im Jahr 2001 und 2002 dargestellt als Liniengraphik der 5-%-Stufen des Blatt- / Nadelverlustes über der Baumnummer und die Abweichungen der Beurteiler als Summenkurve.

- Des mit Kronenzustandsbonituren betrauten Fachpersonals der FAWF untereinander.
- Des von den Landesforsten für die terrestrische Waldschadenserhebung zur Verfügung gestellten Aufnahmepersonals zum Vorgabewert der FAWF.