

**Analyse und Korrektur des BfN-Schätzers  
für den HNV-Flächenanteil an der landwirtschaftlich genutzten Fläche  
und Vergleich mit dem kombinierten Ratioschätzer**

Gutachten für das Bundesamt für Naturschutz (BfN)

von

Prof. Dr. Joachim Saborowski

Abt. Ökoinformatik, Biometrie und Waldwachstum  
und  
Abt. Ökosystemmodellierung  
Georg-August-Universität Göttingen

Göttingen, 7. Juni 2012

## INHALT

1. Einleitung	2
2. Bias-Korrektur des BfN-Schätzers	3
3. Vergleich des korrigierten BfN-Verfahrens mit dem kombinierten Ratioschätzer	4
4. Fazit	7
5. Direkte Stellungnahme zu den Detailfragen von „Teil C – Leistungsbeschreibung“ der Vergabe- und Vertragsunterlagen	8

## 1. Einleitung

Am Beginn dieses Gutachtens steht eine detaillierte Darstellung der beiden konkurrierenden Hochrechnungsverfahren im Rahmen bekannter Grundlagen der Stichprobentheorie. Dies ist für einen qualifizierten Vergleich der Verfahren notwendig. Zur besseren Vergleichbarkeit und für einen direkten Zugriff auf Lehrbuchformeln wird in dieser Abhandlung die Notation des Lehrbuches von Cochran (1977) benutzt. Der Zusammenhang mit den Bezeichnungen der Leistungsbeschreibung ist aus der folgenden Beschreibung ersichtlich.

Ausgangspunkt ist die stratifizierte Stichprobe mit  $n$  Probeflächen (PF, Quadrate von 1 km<sup>2</sup> Fläche), die auf insgesamt  $L$  Schichten (Straten) verteilt sind, d.h.  $n=n_1+n_2+\dots+n_L$ . Die Zuordnung zu den Schichten  $h=1, \dots, L$  hängt von der Lage des Mittelpunktes einer PF ab. Mit den Bezeichnungen

	$Y(S)$	Gesamte HNV-Fäche in einem Farmlandstratum $S$ (Landnutzungsklassen 1, 2 und 4) der Auswahlgesamtheit $G$
$Fl_{HNV\ S\ G}$	$\hat{Y}(S)$	Schätzung von $Y(S)$ mit Hilfe der ausgewählten PF
	$y_j(S)$	Kartierte HNV-Fläche auf Flächen des Farmlandstratums $S$ innerhalb der PF mit der laufenden Nummer $j$
$Fl_{HNV\ S\ PF}$	$\sum_{j=1}^n y_j(S)$	Summe der auf Flächen des Farmlandstratums $S$ kartierten HNV-Flächen über alle $n$ PF
	$x_j(S)$	Fläche des Farmlandstratums $S$ innerhalb der PF $j$
$Fl_{S\ PF}$	$\sum_{j=1}^n x_j(S)$	Summe der Flächen des Farmlandstratums $S$ in allen PF
$Fl_{S\ G}$	$Fl_{hG}$	Fläche des Farmlandstratums $S$ , bzw. des Stratum $h$ in der Auswahlgesamtheit $G$
$Fl_G$		Fläche der Auswahlgesamtheit $G$ , $Fl_G = \sum_h Fl_{hG}$ (summiert über alle bei der Stichprobenauswahl berücksichtigten Schichten)
$Fl_{LW}$		Farmlandfläche in der Auswahlgesamtheit $G$

erhält Formel 2 der Leistungsbeschreibung

$$Fl_{HNV\ S\ G} = Fl_{HNV\ S\ PF} \times Fl_{S\ G} / Fl_{S\ PF} \quad (\text{Formel 2})$$

für ein Farmlandstratum  $S$  die Form

$$\hat{Y}(S) = \sum_{j=1}^n y_j(S) \cdot \frac{Fl_{SG}}{\sum_{j=1}^n x_j(S)}$$

bzw.

$$\hat{Y}(S) = Fl_{SG} \cdot \frac{\sum_{j=1}^n y_j(S)}{\sum_{j=1}^n x_j(S)}$$

Das "Dach" auf  $Y(S)$  soll symbolisieren, dass  $\hat{Y}(S)$  eine Schätzung („Hochrechnung“) für die wahre HNV-Fläche  $Y(S)$  im Farmlandstratum  $S$  ist. Diese Hochrechnung oder Schätzung wird beim BfN-Verfahren nur für die Straten  $S$  durchgeführt, die zu den 3 Landnutzungsklassen 1, 2 und 4 gehören, also zu Ackerflächen, Grünland und Sonderkulturen, oder kurz zu Farmland. Farmland tritt aber nicht nur in solchen PF auf, die bei der Stichprobenauswahl aufgrund der Lage des PF-Mittelpunktes einem Farmlandstratum zugeordnet wurden.

Für das Verständnis des BfN-Verfahrens ist es deshalb hilfreich, sich klar zu machen, dass die Schichten oder Straten hier in zweierlei Weise eine Rolle spielen. Es gibt zum einen die  $L$  Straten, denen die quadratischen PF zugeordnet werden. Ein Stratum  $h$  ist dabei die Menge aller (unendlich vielen) möglichen Stichprobenpunkte, die in eine Fläche fallen, die zur Schicht  $h$  gehört. Zum anderen gibt es innerhalb jeder PF eines Stratums  $h$  Flächen, die ihrerseits zu unterschiedlichen Schichten gehören können. Dabei ist mit Schicht aber ein bestimmter Flächentyp im Sinne einer Ziel- oder Auswertungskategorie gemeint, auf die man die Kartierung bzw. Auswertung einschränkt. Eine solche Kategorie ist z.B. auch die Kategorie Farmland (Landnutzungsklassen 1, 2 und 4). Diese letzteren „Schichten“ sind also von denen der ursprünglichen Stichprobenerhebung zu unterscheiden. Stichprobeneinheit ist nach wie vor die ausgewählte PF. Um diesen Unterschied deutlich zu machen, habe ich im folgenden für diese Schichten im Sinne von Auswertungskategorien anstelle von  $h$  (Schicht der Stichprobenerhebung) die Bezeichnung  $S$  (Auswertungsschicht) benutzt.

Auch wenn die Lage des Mittelpunktes einer PF in einem Stratum  $S$  noch wenig über die Stratenanteile anderer Straten  $S'$  in der PF aussagt, so muss man doch wenigstens im Mittel davon ausgehen, dass in einer solchen PF relativ hohe Anteile des Stratums  $S$  vorkommen. Damit ist aber zu befürchten, dass das BfN-Verfahren in seiner bisherigen Form zu einem systematischen Fehler, also zu einer Über- oder Unterschätzung der Farmland-HNV-Fläche im Stratum  $S$ , führen kann. Der Grund dafür ist die unterschiedliche Repräsentation der „Erhebungs-Straten“ durch die ausgewählten PF. Da die Landnutzungsklassen nicht mit einer flächenproportionalen Anzahl von PF bestückt wurden, sind PF aus Straten mit höherer Stichprobendichte in der Stichprobe vom Umfang  $n$  überrepräsentiert. Wenn diese Straten auch einen tendenziell höheren Anteil an HNV-Flächen in  $S$  enthalten, muss es zu einer Überschätzung der HNV-Fläche in  $S$  kommen, im umgekehrten Fall zu einer Unterschätzung.

## 2. Bias-Korrektur des BfN-Schätzers

Man kann diesen systematischen Fehler, unabhängig davon, ob er nun tatsächlich auftritt oder nicht, aber korrigieren, indem man die kartierten Flächen aller  $n_h$  PF in jedem der  $L$  Straten mittelt

$$\bar{y}_h(S) = \frac{1}{n_h} \sum_{i=1}^{n_h} y_{hi}(S) \quad (0)$$

(Die  $y_{hi}(S)$  und die  $x_{hi}(S)$  mit  $i = 1, \dots, n_h$  sind dieselben Flächen wie die  $y_j(S)$  bzw.  $x_j(S)$ , lediglich neu nummeriert mit  $h$  als laufender Nummer des Stratums, aus dem eine PF nach Lage des PF-Mittelpunktes gezogen wurde, und  $i$  als laufender Nummer der PF im Stratum  $h$ .)

und über die Stratenanteile an der Auswahlgesamtheit gewichtet, d.h.

$$\hat{Y}(S) = Fl_{SG} \cdot \frac{\sum_{h=1}^L \frac{Fl_{hG}}{Fl_G} \bar{y}_h(S)}{\sum_{h=1}^L \frac{Fl_{hG}}{Fl_G} \bar{x}_h(S)} \quad (1)$$

(Sowohl S als auch h werden hier für die Bezeichnung von Straten verwendet, h durchläuft dabei alle Straten, die bei der ursprünglichen stratifizierten Auswahl der PF berücksichtigt wurden, S nur diejenigen, die zu den Landnutzungsklassen 1, 2 und 4 gehören. Siehe dazu auch den vorletzten Absatz von Kapitel 1, sowie Kapitel 5, zu Frage 10 c) )

Ich halte diese Korrektur für absolut notwendig, um potentielle systematische Fehler zu vermeiden. Man kann nicht erst eine geschichtete Stichprobe mit unterschiedlichen Stichprobenintensitäten bzw. Auswahlwahrscheinlichkeiten ziehen und diese Schichtung anschließend bei der Hochrechnung ignorieren. Dies gilt auch für entsprechende Ansätze, die in NRW diskutiert werden (siehe 1a. und 2a. in dem Papier von G. Santora vom 19.04.2011).

Im BfN-Verfahren werden dann nach Formel 3 der Leistungsbeschreibung

$$Fl_{HNVG} = \sum Fl_{HNVSG} \quad (\text{Formel 3})$$

die Schätzungen  $\hat{Y}(S)$  für alle Straten S aufsummiert, die zu den drei Landnutzungsklassen Ackerflächen, Grünland und Sonderkulturen gehören. Dies ist identisch mit

$$\hat{Y} = \sum_{S \in LW} \hat{Y}(S) \quad (2)$$

und führt bei der Fehlerschätzung zu dem Problem, dass die Summanden von (2) als voneinander stochastisch abhängig betrachtet werden müssen. Denn viele PF tragen gleichzeitig zu mehreren Hochrechnungen  $\hat{Y}(S)$  für unterschiedliche S bei; mit weiter reichenden Folgen für die Schätzung der Varianz von  $\hat{Y}(S)$ . Es genügt nämlich nicht, die Varianzen der einzelnen Summanden einfach nur zu addieren, wie in Formel 5 der Leistungsbeschreibung. Man muss zusätzlich auch noch die Kovarianzen zwischen allen Paaren von Summanden in der Varianzformel für  $Fl_{HNVG}$  bzw.  $\hat{Y}$  berücksichtigen. Bei Varianzen und Kovarianzen handelt es sich außerdem um Varianzen und Kovarianzen von Verhältnisschätzern, also nicht von einfachen Mittelwerten wie in Formel 5. Formeln für diese Varianzen und Kovarianzen werden im nun folgenden Abschnitt 3 entwickelt.

### 3. Vergleich des korrigierten BfN-Verfahrens mit dem kombinierten Ratioschätzer

Vor diesem Hintergrund stellt sich die Frage, warum man  $\hat{Y}(S)$  überhaupt erst für unterschiedliche S berechnen und anschließend aufsummieren soll. Man erhält ja einen analog Formel 1 der Leistungsbeschreibung begründeten und viel einfacheren Schätzer, wenn man von Anfang an in jeder PF anstelle von  $y_j(S)$  (HNV-Fläche in PF j auf Flächen des Farmlandstratums S) bzw. anstelle von

$y_{hi}(S)$  die gesamte HNV-Fläche auf Farmland in PF  $j$  bzw.  $hi$  verwendet. Bezeichnet man letztere einfach mit  $y_j$  bzw. mit  $y_{hi}$  (in (0)) und entsprechend mit  $x_j$  bzw.  $x_{hi}$  die Farmlandfläche in PF  $j$  bzw.  $hi$ , so ergibt sich analog (1) die Schätzung (Hochrechnung)

$$\hat{Y} = Fl_{LW} \cdot \frac{\sum_{h=1}^L \frac{Fl_{hG}}{Fl_G} \bar{y}_h}{\sum_{h=1}^L \frac{Fl_{hG}}{Fl_G} \bar{x}_h}$$

für die gesamte HNV-Fläche auf Farmland der Auswahlgesamtheit. Man multipliziert also, analog dem Vorgehen für einzelne Straten  $S$ , einfach die tatsächliche Farmlandfläche in  $G$  mit dem Verhältnis der mittleren HNV-Fläche je PF zu der mittleren Farmlandfläche je PF. Dividiert man schließlich noch durch  $Fl_{LW}$ , so hat man in

$$\hat{R}_c = \frac{\hat{Y}}{Fl_{LW}} = \frac{\sum_{h=1}^L \frac{Fl_{hG}}{Fl_G} \bar{y}_h}{\sum_{h=1}^L \frac{Fl_{hG}}{Fl_G} \bar{x}_h} = \frac{\bar{y}_{st}}{\bar{x}_{st}} \quad (3)$$

einen somit gut begründeten Schätzer für den HNV-Flächenanteil an der Farmlandfläche (der Index  $st$  an den beiden Mittelwerten ganz rechts in Formel 3 steht für „stratifiziert“, um den Unterschied zum einfachen arithmetischen Mittelwert zu verdeutlichen).

Hierbei handelt es sich um einen der beiden bekannten Verhältnisschätzer für stratifizierte Stichproben (Cochran 1977, Kapitel 6.10 Ratio Estimates in Stratified Random Sampling), nämlich den sogenannten kombinierten Verhältnisschätzer (deshalb das  $c$  am  $\hat{R}_c$ ), der nach Cochran dem separaten Verhältnisschätzer vorzuziehen ist, wenn die Stichprobenumfänge in den Straten klein sind, was bei der HNV-Stichprobe gegenwärtig der Fall ist, jedenfalls so lange, wie man sich nicht zu einer deutlichen Aggregation der bisherigen Straten entschließt. Seine Varianz kann mit der Formel

$$\text{var}(\hat{R}_c) = \text{var}\left(\frac{\bar{y}_{st}}{\bar{x}_{st}}\right) = \frac{1}{(\bar{x}_{st})^2} \sum_{h=1}^L \left(\frac{Fl_{hG}}{Fl_G}\right)^2 \left(1 - \frac{n_h \cdot 1km^2}{Fl_{hG}}\right) \frac{s_{uh}^2}{n_h} \quad (4)$$

geschätzt werden (angepasst nach Theorem 6.5 in Cochran 1977). Dabei ist

$$s_{uh}^2 = s_{yh}^2 + \hat{R}_c^2 s_{xh}^2 - 2\hat{R}_c s_{xyh}$$

$$s_{yh}^2 = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h)^2 \quad s_{xh}^2 = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (x_{hi} - \bar{x}_h)^2 \quad s_{xyh} = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (x_{hi} - \bar{x}_h)(y_{hi} - \bar{y}_h)$$

und  $y_{hi}$  ist 0, wenn keine HNV-Fläche auf Farmland in PF  $hi$  vorhanden ist. Ebenso ist  $x_{hi}$  gleich 0, wenn kein Farmland in PF  $hi$  vorhanden ist.

Für einen Vergleich dieses kombinierten Verhältnisschätzers mit dem korrigierten BfN-Verfahren nach (1) und (2), d.h. mit der Hochrechnung gemäß

$$\hat{R}_{BfN} = \frac{\hat{Y}}{Fl_{LW}} = \frac{1}{Fl_{LW}} \sum_{S \subset LW} Fl_{SG} \frac{\sum_{h=1}^L \frac{Fl_{hG}}{Fl_G} \bar{y}_h(S)}{\sum_{h=1}^L \frac{Fl_{hG}}{Fl_G} \bar{x}_h(S)} \quad (5)$$

benötigt man auch für den letzteren einen Varianzschätzer. Die tatsächliche (aber unbekannte und deshalb zu schätzende) und deshalb hier groß geschriebene Varianz von  $\hat{R}_{BfN}$  ist

$$Var(\hat{R}_{BfN}) = \frac{1}{Fl_{LW}^2} \left[ \sum_{S \subset LW} Fl_{SG}^2 \cdot Var(\hat{R}_c(S)) + \sum_{S \neq S'} Fl_{SG} Fl_{S'G} \cdot Cov(\hat{R}_c(S), \hat{R}_c(S')) \right] \quad (6)$$

Die Summe über  $S \neq S'$  bedeutet, dass über alle Paare von unterschiedlichen Straten S und S' summiert werden soll. Die paarweise identischen Kovarianzen für S und S' bzw. für S' und S sind dabei beide zu berücksichtigen.  $\hat{R}_c(S)$  steht hier für die in (5) ganz rechts auftretenden Quotienten, also für die Schätzung des Anteils der HNV-Flächen im Farmlandstratum S, d.h.

$$\hat{R}_c(S) = \frac{\sum_{h=1}^L \frac{Fl_{hG}}{Fl_G} \bar{y}_h(S)}{\sum_{h=1}^L \frac{Fl_{hG}}{Fl_G} \bar{x}_h(S)}$$

Man benötigt also nur noch geeignete Schätzer für die Varianz von  $\hat{R}_c(S)$  und für die paarweisen Kovarianzen zwischen  $\hat{R}_c(S)$  und  $\hat{R}_c(S')$ , die dann in (6), die Formel der tatsächlichen Varianz, eingesetzt werden. Der Quotient  $\hat{R}_c(S)$  hat aber dieselbe Struktur wie der Verhältnisschätzer  $\hat{R}_c$  aus (3), er beschränkt sich lediglich auf ein Farmlandstratum S anstatt auf die gesamte Farmlandfläche LW. Daher kann (4), eingeschränkt auf S, als Schätzer für  $Var \hat{R}_c(S)$  ganz analog verwendet werden:

$$var(\hat{R}_c(S)) = var\left(\frac{\bar{y}_{st}(S)}{\bar{x}_{st}(S)}\right) = \frac{1}{(\bar{x}_{st}(S))^2} \sum_{h=1}^L \left(\frac{Fl_{hG}}{Fl_G}\right)^2 \left(1 - \frac{n_h \cdot 1 km^2}{Fl_{hG}}\right) \frac{s_{u(S)h}^2}{n_h} \quad (7)$$

mit

$$s_{u(S)h}^2 = s_{y(S)h}^2 + \hat{R}_c^2(S) s_{x(S)h}^2 - 2\hat{R}_c(S) s_{x(S)y(S)h}$$

$$s_{y(S)h}^2 = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (y_{hi}(S) - \bar{y}_h(S))^2 \quad s_{x(S)h}^2 = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (x_{hi}(S) - \bar{x}_h(S))^2$$

$$s_{x(S)y(S)h} = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (x_{hi}(S) - \bar{x}_h(S))(y_{hi}(S) - \bar{y}_h(S))$$

Ein Schätzer für die Kovarianzen der Verhältnisschätzer für zwei unterschiedliche Farmlandstraten  $S$  und  $S'$  lässt sich analog der Herleitung der Varianz des kombinierten Schätzers bei Cochran (1977), Theorem 6.5, entwickeln. Man erhält (vgl. die Analogie mit (7))

$$\text{cov}(\hat{R}_c(S), \hat{R}_c(S')) = \frac{1}{\bar{x}_{st}(S)} \frac{1}{\bar{x}_{st}(S')} \sum_{h=1}^L \left( \frac{Fl_{hG}}{Fl_G} \right)^2 \left( 1 - \frac{n_h \cdot 1 \text{ km}^2}{Fl_{hG}} \right) \frac{s_{u(S)u(S')h}}{n_h} \quad (8)$$

mit

$$\begin{aligned} s_{u(S)u(S')h} &= s_{y(S)y(S')h} + \hat{R}_c(S)\hat{R}_c(S')s_{x(S)x(S')h} - \hat{R}_c(S)s_{x(S)y(S')h} - \hat{R}_c(S')s_{x(S')y(S)h} \\ s_{y(S)y(S')h} &= \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (y_{hi}(S) - \bar{y}_h(S))(y_{hi}(S') - \bar{y}_h(S')) \\ s_{x(S)x(S')h} &= \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (x_{hi}(S) - \bar{x}_h(S))(x_{hi}(S') - \bar{x}_h(S')) \\ s_{x(S)y(S')h} &= \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (x_{hi}(S) - \bar{x}_h(S))(y_{hi}(S') - \bar{y}_h(S')) \\ s_{x(S')y(S)h} &= \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (x_{hi}(S') - \bar{x}_h(S'))(y_{hi}(S) - \bar{y}_h(S)) \end{aligned}$$

Die Varianz des korrigierten BfN-Schätzers (5) kann demnach schließlich mit Hilfe von (7) und (8) durch

$$\text{var}(\hat{R}_{BfN}) = \frac{1}{Fl_{LW}^2} \left[ \sum_{S \in LW} Fl_{SG}^2 \cdot \text{var}(\hat{R}_c(S)) + \sum_{S \neq S'} Fl_{SG} Fl_{S'G} \cdot \text{cov}(\hat{R}_c(S), \hat{R}_c(S')) \right] \quad (9)$$

geschätzt werden. Ebenso wie in (3) und (4) müssen auch hier PF mit  $y_{hi} = 0$  oder  $x_{hi} = 0$  berücksichtigt werden.

#### 4. Fazit

Mit (3) und (4) (kombinierter Ratioschätzer) und (5) und (9) (korrigierter BfN-Schätzer) stehen nun zwei Hochrechnungsverfahren mit zugehörigen Stichprobenfehlern zur Verfügung, die beide wohlbegründet sind. Die Frage ist, ob sie sich wirklich wesentlich in ihrer Präzision unterscheiden. Dazu müssten für einige Auswahlgesamtheiten die Varianzen (4) und (9) verglichen werden (nicht die Variationskoeffizienten!). Beide Schätzer und die zugehörigen Varianzen werden sich natürlich unterscheiden, vielleicht aber nicht so wesentlich, dass es eine praktische Relevanz hat. Möglicherweise gibt es auch keinen klaren Trend, in dem Sinne, dass immer, also für alle betrachteten Auswahlgesamtheiten, eines der beiden Verfahren die kleinere Varianz liefert. Außerdem sind beide Varianzschätzer auch nur Näherungen. Es bleibt also zu untersuchen, ob eines der beiden Verfahren sich als klar besser erweist. Die Varianz ist das entscheidende Qualitätskriterium zum Vergleich zweier Schätzer, wenn man davon ausgeht, dass beide nur kleine, vernachlässigbare systematische Fehler haben, was hier unterstellt wird.

Im Kern handelt es sich nämlich bei beiden Verfahren um kombinierte Ratioschätzer. Der Unterschied liegt nur darin, dass beim korrigierten BfN-Verfahren zunächst mit dem kombinierten Ratioschätzer  $\hat{R}_c(S)$  die HNV-Anteile für viele kleine Auswertungsstraten  $S$  geschätzt und später, in (5), nochmals gewichtet gemittelt werden, während der direkte kombinierte Ratioschätzer (3) die HNV-Anteile sofort auf die Farmlandfläche bezieht. Letzteres halte ich für sehr naheliegend, und es führt sicher zu einer weniger fehleranfälligen Programmierung. In jedem Fall möchte ich empfehlen, die programmiertechnische Umsetzung der Formeln an zwei voneinander unabhängigen Stellen durchzuführen, die Ergebnisse zu vergleichen und bei Differenzen nach den Ursachen zu forschen, um Programmierfehler zu vermeiden.

## **5. Direkte Stellungnahme zu den Detailfragen von „Teil C – Leistungsbeschreibung“ der Vergabe- und Vertragsunterlagen**

**Zu Frage 1:** (zusätzliche Probeflächen in einzelnen Bundesländern, Nichtberücksichtigung von PF)

*Das im Rahmen des HNV-Monitorings verwendete Stichprobennetz entspricht nicht dem ursprünglichen Design. Durch Nichtberücksichtigung von Stichprobenflächen mit niedrigen Anteilen an Landwirtschaftsfläche oder hohen Anteilen an Biotopfläche und der Aufnahme zusätzlicher Flächen aus dem Vertiefungsprogramm wird letztlich die Annahme einer Zufallsverteilung der Stichprobenflächen in einer Schicht verletzt. Dürfen bei der Hochrechnung trotzdem alle kartierten Stichprobenflächen berücksichtigt werden? Mehrere Bundesländer kartieren die Flächen des Vertiefungsprogramms vollständig und verdichten ihre Stichprobe so, wie es im ursprünglichen Gesamtkonzept vorgesehen war. Wie ist aber damit umzugehen, wenn Bundesländer nicht das gesamte Vertiefungsprogramm, erfassen sondern neben dem Grundprogramm eine Auswahl an Flächen des Vertiefungsprogramms kartieren, die nach unterschiedlichen Kriterien ausgewählt wurden?*

a) Wenn die zusätzlichen PF aus dem Vertiefungsprogramm, die in einigen Bundesländern über das Grundprogramm hinaus kartiert wurden, bei der bundesweiten Auswertung berücksichtigt werden sollen, müssen die Bundesländer als weiteres Stratifizierungsmerkmal neben den Standorttypen und Landnutzungsklassen berücksichtigt werden. Die betreffenden Bundesländer können dabei entweder zusammen eine gemeinsame Klasse bilden, falls sie dieselben Stichprobendichten verwendet haben, oder sie sind als einzelne Klassen zu betrachten, falls sie unterschiedliche Dichten verwendet haben. Alle übrigen können auch zusammen als eine Klasse betrachtet werden, da sie alle, wie ich vermute, die vorgegebenen Dichten des Grundprogramms verwendet haben.

Die Berücksichtigung aller kartierten PF aus Grund- und Vertiefungsprogramm führt zu einer Verringerung des Stichprobenfehlers und ist daher zu empfehlen. Sie erhöht aber auch die Anzahl der Straten, mindestens um den Faktor zwei.

Grundsätzlich ist bei allen Verdichtungen sicherzustellen, dass die Auswahlwahrscheinlichkeiten für alle Punkte bzw. PF eines Auswahlstratums identisch sind. In einzelnen Bundesländern individuell vorgenommene Verdichtungen (z. B. zum Auffüllen schwach besetzter Straten) können daher mit berücksichtigt werden, wenn sie eine zusätzliche, d. h. über das Grund- oder Verdichtungsprogramm hinaus gehende, Zufallsstichprobe oder wenigstens systematische Stichprobe innerhalb jedes verdichteten Stratums darstellen. Die so verdichteten Straten können allerdings bei eventuellen

Zusammenlegungen von Straten nur mit Straten gleicher Stichprobendichte zusammengelegt werden.

Nicht zulässig ist dagegen, nur solche PF für eine Verdichtung auszuwählen, die besondere Kriterien erfüllen, welche erst nach der näheren „Betrachtung“ der PF eingeschätzt werden. Dazu gehört z. B., dass eine zufällig aus einem Stratum ausgewählte PF nur dann zur Verdichtung verwendet wird, wenn sie „besonders gut“ in dieses Stratum passt, wenn also beispielsweise ihr Offenlandanteil besonders groß ist. Dadurch kann eine systematische Verzerrung für dieses Stratum und damit für die gesamte Hochrechnung hervorgerufen werden.

Bei einer Verdichtung oberhalb einer bestimmten Höhe über NN muss beachtet werden, dass dadurch ein neues Stratifizierungskriterium eingeführt wird, das durch die Verschneidung mit anderen Straten zu mindestens einem oder mehreren neuen Straten führt. Solche neuen Kriterien können das Zusammenlegen von Straten im Falle zu kleiner Stichprobenanzahlen ebenfalls einschränken, da nur Straten mit gleicher Stichprobendichte zusammengelegt werden sollten.

**b)** Nach telefonischer Auskunft von Herrn Benzler und Herrn Fuchs (gemeinsames Telefonat am 19. März 2012) können große, reine Naturschutzflächen, die von der Kartierung ausgenommen wurden, als Flächen betrachtet werden, die nicht zur landwirtschaftlichen Nutzfläche gehören. Dann gehören sie aber definitionsgemäß auch nicht zu den zu kartierenden HNV-Flächen.

**c)** Die Nichtberücksichtigung von PF mit bestimmten Eigenschaften ist im übrigen immer eine potentielle Quelle für systematische Fehler der HNV-Flächenschätzung.

Eine grobe Abschätzung dieses systematischen Fehlers, etwa durch die Nichtberücksichtigung von PF mit niedrigen Anteilen an Landwirtschaftsfläche, wäre möglich, wenn die beroffenen PF, oder zumindest ihre Anzahl je Stratum, noch bekannt sind. Man könnte dann neben der Hochrechnung ohne diese Flächen zwei Vergleichsrechnungen mit den beiden Extremen durchführen, nämlich eine unter der Annahme, dass diese PF alle genau 0 ha HNV-Fläche und 0,05 ha Farmlandfläche haben, sowie eine mit 0,05 ha HNV-Fläche und 0,05 ha Farmlandfläche. Damit ließe sich dieser Fehler eingrenzen, und man würde auch zwei extreme Varianzschätzungen erhalten, die den Einfluss dieser nicht berücksichtigten PF auf die Varianzschätzung deutlich machen. Wenn man Glück hat, sind die Auswirkungen vernachlässigbar klein.

Eine Alternative wäre, für diese von der Kartierung ausgenommenen PF die HNV-Fläche (je nachdem, ob BfN-Verfahren oder Ratioschätzer verwendet wird  $y_{hi}(S)$  bzw.  $y_{hi}$ ) und die Farmlandfläche ( $x_{hi}(S)$  bzw.  $x_{hi}$ ) zu simulieren und die simulierten Zahlen bei der Hochrechnung und Varianzschätzung zu verwenden. Man könnte dafür z.B. eine Gleichverteilung der  $x_{hi}$  zwischen 0 und 0,05 ha annehmen und für die  $y_{hi}$  eine Verteilung zwischen 0 und dem zuvor simulierten  $x_{hi}$ , die dieselbe Form hat wie die zwischen 0 und 1 liegende Verteilung von  $y_{hi}/x_{hi}$  der übrigen PF (hierfür könnte eine Beta-Verteilung geeignet sein).

#### **Zu Frage 2:** (Schichtzugehörigkeit)

*Bei der Hochrechnung der HNV-Fläche nach Formel 2 werden die Parameter  $Fl_{HNV S PF}$  (kartierte HNV-Fläche in der betrachteten Schicht S in allen Probeflächen PF) und  $Fl_{S PF}$  (Fläche der Schicht S in allen Probeflächen PF) aus den tatsächlichen Schichtflächen in jeder Probefläche ermittelt. Dieses Verfahren trägt unserer Ansicht nach am besten der Tatsache Rechnung, dass die*

*Stichprobengesamtfläche in keiner Schicht flächenproportional zur Schichtfläche ist. Nach dem ursprünglichen Stichprobendesign hätten die beiden Parameter aber über die (durch den Stichprobenflächen-Mittelpunkt determinierten) Schichtzugehörigkeit der gesamten Stichprobenfläche ermittelt werden sollen. Die Frage ist, ob diese Abweichung vom ursprünglichen Design a) aus stichprobentheoretischer Sicht zulässig ist und b) wie angenommen zu einer höheren Genauigkeit des ermittelten Schätzwerts führt.*

Die Hochrechnung nach Formel 2 ignoriert die stratifizierte Stichprobenauswahl völlig. Deshalb habe ich sie auch in Kapitel 2 dieses Gutachtens korrigiert, indem ich die Zugehörigkeit zu den Straten  $h = 1, \dots, L$  nach Mittelpunktzuordnung eingebaut habe, um die unterschiedliche Repräsentation durch PF der unterschiedlichen „Mittelpunktstraten“ berücksichtigen zu können. Die Zuordnung nach Mittelpunkten ist sicher nicht die vernünftigste, aber sie stellt dennoch eine Stratifizierung dar, von der man erwarten kann, dass PF eines solchen Stratums sich tendenziell ähnlich sind. Sie muss und kann, wie im Kapitel 2 dieses Gutachtens gezeigt wurde, bei der Hochrechnung berücksichtigt werden, um systematische Fehler zu vermeiden.

Eine effektivere Stratifizierung wäre z.B. möglich, indem man G durch ein Quadratraster mit 1 km Maschenweite lückenlos und überschneidungsfrei überdeckt (Pflasterung), danach jedes Quadrat abhängig von der darin dominierenden Schicht einem Stratum zuordnet (dies könnte man im GIS flächenscharf oder aber durch ein dichtes Punktraster im Quadrat machen) und schließlich aus den so stratifizierten, endlich vielen Quadraten eine stratifizierte Stichprobe zieht. Eine solche Stratifizierung hätte wahrscheinlich wesentlich homogenere Straten erzeugt als die tatsächlich verwendete, wo zwischen den PF eines Stratums wegen der alleinigen Zuordnung nach Lage der Mittelpunkte vermutlich große Streuungen entstehen. Aber dieses Rad kann man wohl jetzt nicht mehr zurückdrehen, da die PF-Auswahl erfolgt ist und bei Wiederholungsinventuren wohl nicht erneut vorgenommen werden wird.

Mit dem BfN-Verfahren wird der Versuch unternommen, durch eine separate Schätzung von HNV-Flächen innerhalb der einzelnen Farmlandstratenflächen und ihrer anschließenden Summierung eine höhere Präzision, also einen kleineren Stichprobenfehler, zu erzielen. Dies ist zulässig, entlässt einen aber nicht aus der Notwendigkeit, die ursprüngliche stratifizierte Stichprobenauswahl zu berücksichtigen. Rein argumentativ kann ich im Augenblick nicht begründen, warum dieses Verfahren aufgrund seiner Konstruktion eine höhere oder auch geringere Präzision haben müsste als der einfache kombinierte Ratioschätzer. Die Frage der zugehörigen Varianzschätzung wurde mit diesem Gutachten beantwortet, so dass ein rechnerischer Vergleich der Varianzen beider Hochrechnungsverfahren möglich ist.

**Zu Frage 3:** (Bezugswert Landwirtschaftsfläche)

*Als Bezugswert für den HNV-Indikator wird die Landwirtschaftsfläche verwendet. Unserer Ansicht nach sollte dieselbe Flächengrundlage für die Ermittlung dieses Werts verwendet werden wie für die Hochrechnung, nämlich die Flächen der Ebenen 1, 2 und 4 des Basis-DLM (Acker, Grünland und Sonderkulturen). Es ist aber bekannt, dass die landwirtschaftliche Nutzfläche von dieser Zahl in allen Bundesländern um einige Prozentpunkte abweicht, weil sie mit genaueren Methoden ermittelt wird. Für die politische Verankerung des Indikators wäre es einfacher, die landwirtschaftliche Nutzfläche als Bezugswert zu verwenden. Zu dieser liegen aber keine Geodaten vor, so dass sie nicht für ein HNV-Element oder eine Stichprobe ermittelt*

*werden kann, sondern nur für ein Bundesland und damit ganz Deutschland. Wäre es zulässig, die landwirtschaftliche Nutzfläche als Bezugsfläche zu verwenden?*

Die Frage verstehe ich so, dass man in (3) bzw. in (5) durch die in den Ländern ermittelte, präzisere landwirtschaftliche Nutzfläche statt durch die aus dem DLM ermittelte Fläche  $Fl_{LW}$  teilen möchte. Dies halte ich in Strenge nicht für zulässig, da die kartierten HNV-Flächen immer Teilflächen von Flächen der DLM-Landnutzungsklassen 1,2 und 4 sind. D.h. bei genauerer Betrachtung könnten kartierte HNV-Flächen teilweise außerhalb der präziser ermittelten landwirtschaftlichen Nutzfläche liegen, und es könnte umgekehrt auch HNV-Flächen geben, die innerhalb der präziseren landwirtschaftlichen Nutzfläche liegen, aber nicht kartiert wurden, weil sie außerhalb der DLM-Farmlandklassen liegen. Auch wenn sich diese beiden gegenläufigen Effekte zufällig ausgleichen könnten, sind sie eine weitere Quelle potentieller systematischer Fehler, die man nicht ohne Not eröffnen sollte.

**Zu Frage 4:** (Stichprobenfehler und Normalverteilungsannahme)

*Wichtige Grundannahmen für die verwendete Methode (z. B. die Annahme einer Normalverteilung der Stichprobenmittelwerte bzw. die notwendigen Mindestzahl von Stichproben in jeder Schicht für die Gültigkeit des zentralen Grenzwertsatzes) treffen nicht zu. Ist die vorgeschlagene Methode zur Ermittlung des Stichprobenfehlers dennoch grundsätzlich zulässig? Wahrscheinlich muss die Frage getrennt betrachtet werden, je nachdem ob (siehe Frage 2) die Schichtzuordnung über die tatsächliche Flächenverteilung erfolgt oder über die Mittelpunkte der Stichprobenflächen. Wenn die Schichtzuordnung über Flächenverteilung erfolgt, tritt ein weiteres Problem auf: da dann auf einer Stichprobenfläche mehrere Schichten vertreten sind, können die Schichten in den Stichproben nicht als unabhängig voneinander betrachtet werden, sondern sind geklumpt.*

**a)** Die Schätzung des Stichprobenfehlers als Wurzel aus den Varianzschätzungen gemäß (4) bzw. (9) ist auch ohne die Gültigkeit der Normalverteilungsannahme zulässig, ebenso die Schätzung von Variationskoeffizienten. Verteilungsannahmen spielen erst eine Rolle, wenn Konfidenzintervalle geschätzt werden sollen. Bei hinreichend großen Stichprobenumfängen gehen Sie mit Recht davon aus, dass man für die Verteilung von arithmetischen Mittelwerten nach dem zentralen Grenzwertsatz annähernd Normalverteilung unterstellen kann. Solche Stichprobenumfänge sind aber für viele Straten im vorliegenden Fall nicht gegeben. Was kann man also tun?

Eine sinnvolle Maßnahme ist in jedem Fall die Zusammenlegung von Straten. Dies ist sowieso notwendig, um in allen Straten ein absolutes Minimum von  $n_h = 2$  PF für die Varianzschätzung zu bekommen. Zusammenlegen darf man aber nur solche Straten, die dieselbe Stichprobendichte (Auswahlwahrscheinlichkeit) haben. Da innerhalb der Landnutzungsklassen (zumindest annähernd) flächenproportional aus den Standortstypen ausgewählt wurde, können vor diesem Hintergrund ohne weiteres unterschiedliche Standortstypen innerhalb einer Landnutzungsklasse zu einem größeren Stratum aggregiert werden. Sinnvoll ist es dabei, solche Standortstypen zusammenzufassen, die ähnliche HNV-Anteile erwarten lassen, sie sollte aber nicht anhand der tatsächlich beobachteten HNV-Anteile erfolgen. Letzteres würde die Stratifizierung zufällig machen und eine Variationsquelle darstellen, die bei der Varianzschätzung nicht berücksichtigt wird.

**b)** Beim BfN-Verfahren erfolgt keine Schichtzuordnung über die Flächenverteilung. Dies ist meiner Ansicht nach ein Missverständnis. Die PF sind nach wie vor über ihren Mittelpunkt einer Schicht zugeordnet. Es wird lediglich auf jeder PF separat für jedes Farmlandstratum die HNV-Fläche ermittelt, hochgerechnet und hinterher wieder über alle Farmlandstraten zusammengefasst. Das hat nichts mit einer andersartig stratifizierten Stichprobe zu tun. Man erkennt dies bei genauerer Betrachtung des korrigierten BfN-Schätzers (5). Das Problem der Abhängigkeit zwischen den separaten Hochrechnungen ist aber in Form der Varianzschätzung gemäß (9) gelöst.

**Zu Frage 5:** (Bootstrapping)

*Unabhängig von der Antwort auf Frage 4: wäre es sinnvoll, eine andere Methode zur Fehlerermittlung zu verwenden, die keine solchen Beschränkungen aufweist, wie z. B. ein Bootstrap-Verfahren?*

Da für beide zur Diskussion stehenden Hochrechnungsverfahren geeignete Varianzschätzer existieren und das Problem der Normalverteilungsannahme für  $\hat{R}_{BfN}$  und  $\hat{R}_c$  über die Aggregation von Straten angegangen werden kann, würde ich das Bootstrapping nicht als Verfahren der Wahl ansehen, zumal Bootstrapping bei sehr schiefen Verteilungen auch unbefriedigend funktionieren kann. Die Aggregation muss meiner Ansicht nach die Stichprobenumfänge nicht unbedingt in allen Straten auf  $n_h = 30$  anheben, ohne dass ich hierfür eine klare Regel angeben könnte. Für die Güte der Normalverteilungsapproximation spielt auch die große Zahl von PF insgesamt eine Rolle. Wenn man dann noch statt des kritischen Wertes der Standardnormalverteilung den einer t-Verteilung mit Freiheitsgrad  $n_h - 1$  wählt ( $n_h$ : kleinster in einem Stratum vorkommender Stichprobenumfang), also

$$\hat{R}_c \pm t_{n_h - 1; 0,025} \sqrt{\text{var}(\hat{R}_c)}$$

dann sollte man ausreichend zuverlässige Konfidenzintervalle bekommen.

**Zu Frage 6:** ( $n_h = 1$ )

*Die Ermittlung der Varianz als wichtigem Parameter bei der Fehlerermittlung ist bei Schichten mit nur einer Stichprobenfläche unmöglich. Natürlich kann auch eine Bootstrap-Methode bei einer Stichprobenzahl von 1 keine Varianz ermitteln. Diese Schichten gehen bisher mit einem Stichprobenfehler von 0 in die Fehlerermittlung ein. Ist es möglich und sinnvoll, über andere Methoden auch diese Schichten in die Fehlerberechnung einzubeziehen?*

Wurde bereits im Zusammenhang mit Frage 5 beantwortet.

**Zu Frage 7:** (Zusammenlegung von Schichten, Verletzung der Zufallsauswahl)

*In verschiedenen Diskussionsbeiträgen seitens der Landesbehörden wurde vorgeschlagen, auf eine der Schichtungsebenen zu verzichten, und die Stichproben nur noch nach den Standorttypen oder den Landnutzungsklassen nach Basis-DLM zu schichten. Abgesehen davon, dass dies das ursprüngliche Stichprobendesign noch stärker verletzen würde als die ohnehin durchgeführten Maßnahmen, erscheint uns diese Vereinfachung nicht zulässig, weil in keiner der neu entstandenen Schichten annähernd von Zufallsverteilung der Stichprobenflächen und/oder von Flächenproportionalität ausgegangen werden kann. Ist das richtig?*

Ich habe bereits zu Frage 4 darauf hingewiesen, dass eine Zusammenlegung von Straten unterschiedlicher Landnutzungsklassen zu vermeiden ist. Innerhalb der Landnutzungsklassen ist das aber möglich. Die Auswahl bleibt dadurch nach wie vor (annähernd) flächenproportional. Die Zufallsauswahl der PF ist sowieso nicht gegeben (siehe Bericht Heidrich-Riske), sondern sie ist systematisch, was erfahrungsgemäß tendenziell zu einer Überschätzung der Stichprobenfehler führt, also zu einer konservativen Fehlerschätzung, d.h. die Inventur ist im Mittel präziser als der geschätzte Stichprobenfehler angibt. Durch die Zusammenlegung zweier oder mehrerer Straten einer Landnutzungsklasse werden also unabhängig voneinander gezogene (siehe Heidrich-Riske S. 11, 4.1.1), systematische Stichproben mit identischen Auswahlwahrscheinlichkeiten zusammengefasst. Dagegen ist aus meiner Sicht nichts einzuwenden.

**Zu Frage 8:** (6 Landschaftstypen statt 21 Standorttypen)

*Alternativ könnte man statt der 21 Standorttypen die höher aggregierten 6 Landschaftstypen (siehe Abb. 1) verwenden, so dass sich statt 110 Schichten nur noch 33 ergeben. Allerdings wären auch bei dieser Auswahl nur 14 der neuen Schichten mit 30 oder mehr Stichprobenflächen belegt, und eine Schicht mit nur einer Stichprobenfläche (Schichtzuordnung nach Mittelpunktdaten). Wichtig ist, dass die Landschaftstypen nicht einfach einer höheren Hierarchieebene bei der Ermittlung der Standorttypen entsprechen, sondern hier auch benachbarte Standorttypen, die aus verschiedenen Zweigen des Entscheidungsbaums stammen, zu einem Landschaftstyp zusammengefasst wurden. Unter welchen Bedingungen wäre diese Vorgehensweise zulässig bzw. von Vorteil?*

Prinzipiell kann man statt der 21 Standorttypen auch nur die 6 Landschaftstypen verwenden, da dies eine Zusammenlegung innerhalb einer Landnutzungsklasse darstellt (siehe Zu Frage 4). Wenn dadurch immer noch eine Schicht nur eine PF hat, ist ein Ziel der Zusammenlegung noch nicht erreicht worden. Die andere Frage ist, ob diese Zusammenlegung zu einer effektiven Stratifizierung führt. Eine effektive Stratifizierung liegt dann vor, wenn die HNV-Anteile innerhalb der Straten sehr homogen sind. Dies wird durch die vorgeschlagene Zusammenlegung möglicherweise nicht erreicht.

Die Strategie bei einer Zusammenlegung sollte deshalb sein, kleinere Straten (insbesondere solche mit  $n_h=1$ ) mit Straten zusammenzulegen, die ähnliche HNV-Anteile erwarten lassen. Eine erste, vorsichtige Zusammenlegung könnte dabei zunächst dafür sorgen, dass überall mindestens  $n_h=2$  oder mehr PF vorliegen, so dass eine Varianzschätzung in allen Straten möglich wird. Dann schätzt man mit (4) oder (9) die Varianz der HNV-Anteilschätzung auf der Basis dieser vorsichtigen Zusammenlegung. Sie dient als Vergleich zu den Varianzen, die in weiteren Schritten mit höher aggregierten Straten erzielt werden. Anhand dieser Varianzen können unterschiedliche Aggregationen beurteilt werden, und man kann sich für diejenige Stratifizierung entscheiden, die zum einen möglichst hohe Stichprobenumfänge innerhalb der Straten erzeugt, aber auch zu einem möglichst kleinen Anstieg der Varianz führt. Da nur Varianzen geschätzt werden, muss für diese Untersuchung auch nicht die Normalverteilung der Verhältnisschätzer vorausgesetzt werden.

**Zu Frage 9:** (HNV-Indikator für (kleine) Bundesländer)

*Bei der Einführung des Monitoringprogramms wurde beschlossen, aus Kostengründen nur die Stichprobenflächen des Grundprogramms (und auch von diesen nicht alle) einzubeziehen. Dabei wurde davon ausgegangen, dass mit diesem Stichprobenprogramm auch zuverlässige Werte des HNV-Indikators für jedes einzelne Bundesland berechnet werden können (unter der Maßgabe,*

*dass das Saarland als kleinster Flächenstaat statt 7 Probeflächen des Grundprogramms 30 Probeflächen des Vertiefungsprogramms untersuchen lässt. Die Stadtstaaten sind nicht in das Monitoring einbezogen). Bei Berechnung eines Werts auf Landesebene verschärfen sich natürlich zahlreiche Probleme, weil die Stichprobenzahl je Schicht noch geringer ist, und ebenso noch weniger von Zufallsverteilung und Flächenproportionalität der Stichproben ausgegangen werden kann. Unter welchen Maßgaben kann trotzdem ein Landeswert für den HNV-Indikator auch in den Ländern ermittelt werden, die nur die Stichprobenflächen des Grundprogramms untersuchen lassen (siehe Tab. 1)?*

Bei sehr geringen Stichprobenumfängen ist eine Hochrechnung weiterhin ohne weiteres möglich, und man könnte auf die Angabe von Konfidenzintervallen verzichten und lediglich die Stichprobenfehler (Wurzel aus der Varianz) als Maß für die Präzision der Hochrechnung angeben. Dafür ist, wie schon zu Frage 4 a) ausgeführt, keine Normalverteilung notwendig. Es wäre aber gut, wenn es auch in kleinen Bundesländern mindestens  $n_h=2$  PF in jedem Stratum gäbe (Aggregation). Andernfalls könnte man in Straten mit  $n_h=1$  z.B. als Näherung die Varianz aus demselben Stratum eines anderen Bundeslandes einsetzen. Die Grundfrage, ob es besser ist, die Bundesländer als Stratifizierungskriterium beizubehalten und die Straten innerhalb der Bundesländer stärker zu aggregieren, oder auf die Bundesländer als Kriterium zu verzichten (dann nur noch das Grundprogramm auswertbar) und dafür die übrigen Straten etwas weniger stark zu aggregieren, kann letztlich nur durch den Varianzvergleich qualifiziert beurteilt werden, der schon zu Frage 8 beschrieben wurde.

Ob die Präzision für einzelne Bundesländer am Ende „ausreicht“, muss anhand der Konfidenzintervalle bzw. der Varianzen beurteilt werden.

**Zu Frage 10:** (Bedeutung der Schichten, Simulation und Validität)

*Als das Stichprobennetz mit der doppelten Schichtung nach Standorttyp und Landnutzungs-kategorie entwickelt wurde, bestand keine Möglichkeit, die Validität und Bedeutung der Schichten empirisch zu überprüfen. Dies liegt im Wesentlichen daran, dass Daten aus flächendeckenden Umweltbeobachtungsprogrammen meist auf Länderebene entwickelt und umgesetzt werden und nicht nach bundesweit einheitlichen Programmen und Methoden. Dies ist im Wesentlichen auch heute noch der Fall. Es wäre aus unserer Sicht dennoch wünschenswert, eine solche empirische Überprüfung vorzunehmen. Auf Bundesebene müsste wahrscheinlich eine Simulation verwendet werden, indem z. B. Zufallspunkte erzeugt werden, deren Dichte je Fläche je nach Schicht unterschiedlich ist. Ist eine derartige Validierung sinnvoll zu realisieren und welche Bedingungen müssten solche Simulationen erfüllen, damit sie zur Überprüfung der Hoch- und Fehlerrechnung verwendet werden können? Auf Landesebene könnten für solche Validierungen (zumindest in einigen Bundesländern) die Geodaten der Kartierungen geschützter Biotopflächen verwendet werden, die jeweils flächendeckend vorliegen. Da ein Teil der HNV-Flächen auch geschützte Biotopflächen sind, und da ein großer Teil der Biotopflächen zumindest ehemals landwirtschaftlich genutzte Flächen sind, kann in erster Näherung von einer ähnlichen Verteilung beider Flächentypen ausgegangen werden. Testweise wurde eine solche Validierung bereits mit den Daten der Kartierung geschützter Biotopflächen für Baden-Württemberg und des Biotopkatasters Nordrhein-Westfalen durchgeführt. Die Biotopdaten wurden mit den Daten der Schichten und den im HNV-Monitoring verwendeten Stichprobenflächen verschnitten. Dabei wurden jeweils zwei Stichprobensätze verwendet: die für das HNV-Monitoring*

verwendete Teilstichprobe des Grundprogramms und die Stichprobenflächen des Vertiefungsprogramms. Aus dem Ergebnis dieser Verschneidungen wurden für jeden Datensatz folgende Werte ermittelt:

a) Der reale Landesanteil der Biotopfläche an der Landwirtschaftsfläche (gemessen als Gesamtfläche der Ebenen 1 – Acker und 2 – Grünland des Basis-DLM).

b) Die Hochrechnung des Landesanteils der Biotopflächen an der Landwirtschaftsfläche (jede Biotopfläche wird exakt der Schicht zugeordnet, in der sie liegt).

c) Die Hochrechnung des Landesanteils der Biotopflächen an der Landwirtschaftsfläche über eine Modifikation, bei der jede Biotopfläche der Schicht zugeordnet wird, in der der Mittelpunkt der Stichprobenfläche liegt (wie in Frage 2 beschrieben).

d) Wie oben, aber mit vereinfachter Schichtzuweisung: der Standorttyp wird vernachlässigt und nur mehr die DLM-Ebene wird als Schicht verwendet (diese Methode geht auf einen Vorschlag des LANUV Nordrhein-Westfalen zurück).

Die Ergebnisse dieser vorläufigen Validierung sind in der folgenden Übersicht zusammengefasst:

Anteil Biotopfläche an Landwirtschaftsfläche:					
	n =	Wert a)	Wert b)	Wert c)	Wert d)
BW	97	2,42 %	2,56 %	1,79 %	1,99 %
	400		2,46 %	2,06 %	2,35 %
NW	73	11,37 %	11,44 %	8,31 %	11,91 %
	170		12,82 %	13,18 %	15,12 %

Es wird klar, dass der mit der bisher verwendeten Hochrechnungsmethode ermittelte Wert jeweils näher am realen Wert liegt als die Werte nach Methode c) und d). Vertiefte Analysen zum Vergleich der Methoden sollten nach Klärung der sinnvollsten Methode zur Ermittlung des Stichprobenfehlers durchgeführt werden.

a) Die Bedeutung der Schichten, d.h. ob wirklich alle Schichten notwendig sind, um zu einer präzisen Schätzung der HNV-Flächenanteile zu kommen, lässt sich am besten durch den Vergleich der aus unterschiedlichen Schichtungen resultierenden Stichprobenfehler beurteilen (siehe zu Frage 8). Dazu ist keine Simulation notwendig.

Die Hoch- und Fehlerrechnung muss meiner Ansicht nach nicht weiter überprüft werden, da sie nun mit (3) und (4) bzw. (5) und (9) auf stichprobentheoretisch recht sicheren Füßen steht. Es muss aber die Entscheidung für eines der beiden Verfahren getroffen werden. Wie dies gemacht werden kann, wurde in Kapitel 4 (Fazit) beschrieben.

Will man dennoch simulieren, um den Bias der Hochrechnung und der Varianzschätzung, sowie die wahre Überdeckungswahrscheinlichkeit der Konfidenzintervalle abzuschätzen, so muss für einen

bestimmten HNV-Flächentyp eine landes- oder bundesweite Kartierung vorliegen, so wie dies bei den Fallstudien Baden-Württemberg und Nordrhein-Westfalen offenbar der Fall ist. Für diese Auswahlgesamtheit müsste man dann nach meinen Erfahrungen etwa hundert- bis zweihundertmal (für die Überprüfung der Konfidenzintervallschätzungen etwa zweitausendmal) die Stichprobenziehung nach der Beschreibung von Heidrich-Riske durchführen und für jede Ziehung die Schätzer (3) und (4) sowie (5) und (9) und die zugehörigen Konfidenzintervalle berechnen. Die Mittelwerte aller Wiederholungen von (3) und (5) minus dem wahren HNV-Anteil ergeben dann eine Näherung für den systematischen Fehler, und die beiden Varianzen der Wiederholungen von (3) und (5) stehen für die wahre Varianz der beiden Schätzer (3) und (5). Diese wahre Varianz jedes der beiden Schätzer ist zu vergleichen mit dem zugehörigen Mittelwert der Wiederholungen von (4) bzw. (9), um die Verzerrung der Varianzschätzer zu prüfen. Weiterhin kann (bei etwa 2000 Wiederholungen) ausgezählt werden, wieviel Prozent der Konfidenzintervalle den Mittelwert der Wiederholungen von (3) bzw. (4) enthalten, um zu prüfen, ob die vorgegebene Überdeckungswahrscheinlichkeit von z.B. 95% eingehalten wird.

**b)** Die Ergebnisse der bisherigen Validierungen sind, wenn überhaupt, nur mit großer Vorsicht zu interpretieren. Erstens deshalb, weil die bisherige Varianzschätzung nach Formel 5 nicht korrekt ist (siehe Kapitel 2) und zweitens weil die Konfidenzintervalle der vier Varianten a) bis d), wenn man einmal annimmt, dass zumindest die Größenordnung der Varianz stimmt, sich mehr oder weniger stark überlappen. Solche Fallstudien können einen Methodenvergleich mit Hilfe der geschätzten Varianzen, wie er in Kapitel 4 beschrieben wurde, nur ergänzen, dann aber unter Verwendung theoretisch abgesicherter Verfahren, wie sie durch (3) und (5) mit den zugehörigen Varianzschätzern beschrieben wurden.

**c)** Biotopflächen nicht der Schicht zuzuordnen, in der sie liegen, sondern der Schicht, in der der Mittelpunkt der PF liegt (Variante c) der Fallstudie auf Seite 19 der Leistungsbeschreibung) macht für mich überhaupt keinen Sinn. Eine Biotopfläche gehört zunächst zu der Schicht, in der sie auch liegt; nur so wird ja entschieden, ob sie überhaupt kartiert wird, nämlich nur dann, wenn sie in einer der Farmlandschichten liegt. Zusätzlich gehört sie immer zu einer bestimmten PF, die wiederum aufgrund ihres Mittelpunktes ebenfalls einer Schicht, möglicherweise einer anderen, zugeordnet ist. Beides muss man klar auseinanderhalten, einerseits die Zugehörigkeit einer HNV-Fläche zu einer Schicht (in der sie liegt), andererseits die Zugehörigkeit der ganzen PF zu einer Schicht aufgrund der Lage ihres Mittelpunktes. Aus diesem Grund habe ich in allen Formeln im ersteren Fall die Schicht mit S bezeichnet, im letzteren mit h.