

Aus dem Institut für Agrar- und Ernährungswissenschaften
(Geschäftsführender Direktor: Prof. Dr. Reinhold Jahn)

der Naturwissenschaftlichen Fakultät III
(Dekan: Prof. Dr. Peter Wycisk)

der Martin-Luther-Universität Halle-Wittenberg

Fachgebiet: Biometrie und Agrarinformatik

„Vergleichende Untersuchungen verschiedener
Auswertungsmethoden für geordnete kategoriale Merkmale bei
Vorliegen räumlicher Kovarianz“

Dissertation

zur Erlangung des akademischen Grades
doctor agriculturarum (Dr. agr.)

von

Diplomagraringenieurin Karen Höttl

Halle/Saale 2012

Aus dem Institut für Agrar- und Ernährungswissenschaften
(Geschäftsführender Direktor: Prof. Dr. Reinhold Jahn)

der Naturwissenschaftlichen Fakultät III
(Dekan: Prof. Dr. Peter Wycisk)

der Martin-Luther-Universität Halle-Wittenberg

„Vergleichende Untersuchungen verschiedener
Auswertungsmethoden für geordnete kategoriale Merkmale bei
Vorliegen räumlicher Kovarianz“

Dissertation

zur Erlangung des akademischen Grades
doctor agriculturarum (Dr. agr.)

vorgelegt von

Diplomagraringenieurin Karen Höttl
geb. am 19.07.1968 in Hamburg

Gutachter: Herr Prof. Dr. J. Spilke

Herr Prof. Dr. H. H. Swalve

Frau Prof. Dr. C. Richter

Verteidigung am: 25. Juni 2012

Halle/Saale 2012

Inhaltsverzeichnis

1	Einleitung und Zielstellung.....	1
2	Statistische Modelle und Auswertungsmethoden	3
2.1	Das lineare gemischte Modell (LMM)	3
2.2	Das Schwellenwertmodell (SWM)	7
2.2.1	Basisidee und Modell für feste Effekte und unabhängige Resteffekte .	7
2.2.2	Schwellenwertmodell mit zufälligen Effekten.....	8
2.2.3	Schwellenwertmodell mit räumlicher Kovarianzstruktur	12
2.2.4	Berechnung von Konfidenzintervallen für kumulative Wahrscheinlichkeiten	12
2.3	Rangverfahren.....	13
2.3.1	Ränge	14
2.3.2	Nichtparametrische Effekte und Hypothesen	14
2.3.3	Hypothesen und Tests	16
2.4	Permutationstests zur Analyse von einer Bonitur je Parzelle	20
2.4.1	Permutationstest für die einfaktorielle vollständig randomisierte Versuchsanlage.....	21
2.4.2	Permutationstest für die einfaktorielle randomisierte Blockanlage mit fixen Blocks	23
2.5	Auswertung der Parzellenmittelwerte der Boniturnoten mittels LMM.....	24
2.6	Auswertung der Boniturnoten als stetiges Merkmal mittels LMM	25
3	Datensimulation und Maßzahlen zur Auswertung der Simulationen	26
3.1	Simulierter Versuchsplan und Datenumfänge.....	26
3.2	Erzeugung der Zufallszahlen	26
3.3	Parameter	29
3.4	Maßzahlen zur Auswertung der Simulationen.....	34
4	Rechentechnische Umsetzung der Auswertungsmethoden.....	36

4.1	Lineares gemischtes Modell	36
4.2	Schwellenwertmodell.....	42
4.3	Rangverfahren.....	44
4.4	Permutationstests	45
5	Ergebnisse	48
5.1	Konvergenzverhalten	48
5.2	Schätzung der Varianzkomponenten.....	49
5.3	Kontraste der fixen Effekte	51
5.3.1	Auswertung der latenten Variablen (LMM).....	51
5.3.2	Auswertung der Bonituren mittels Schwellenwertmodell (SWM).....	52
5.3.3	Auswertung der Parzellenmittelwerte der Bonituren (LMM)	52
5.3.4	Auswertung der Bonituren als stetiges Merkmal (LMM).....	53
5.4	Schätzung kumulativer Wahrscheinlichkeiten	58
5.5	Einhaltung des Fehlers 1. Art.....	65
5.5.1	Auswertung der Simulationsansätze ohne räumliche Kovarianzstrukturen.....	65
5.5.2	Auswertung der Simulationsansätze mit räumlichen Kovarianzstrukturen.....	68
5.5.2.1	Auswertung der latenten Variablen (LMM).....	68
5.5.2.2	Auswertung der Bonituren mittels Schwellenwertmodell (SWM)..	70
5.5.2.3	Auswertung der Bonituren mittels Rangverfahren	74
5.5.2.4	Auswertung der Bonituren mittels Permutationstests.....	75
5.5.2.5	Auswertung der Parzellenmittelwerte der Bonituren (LMM)	77
5.5.2.6	Auswertung der Bonituren als stetiges Merkmal (LMM).....	80
5.6	Zusammenfassende Bewertung der Ergebnisse für Parametersatz 1, mit räumlichen Kovarianzstrukturen.....	82

5.7	Spezielle Vergleiche der Auswertung der Bonituren mittels LMM und SWM	84
5.7.1	Auswertung der Simulationsansätze ohne räumliche Kovarianz für k = 4 und k = 9 Boniturklassen	84
5.7.2	Auswertung der Simulationsansätze mit räumlicher Kovarianz und Simulation heterogener Prüfgliedeffekte	89
6	Diskussion und Schlussfolgerungen	94
6.1	Konvergenzverhalten	94
6.2	Erwartungstreue und Varianz der Parameterschätzungen	96
6.3	Einhaltung des Fehlers 1. Art bei Hypothesentests.....	101
6.4	Konsequenzen für die Auswertungspraxis.....	109
7	Zusammenfassung.....	113
8	Summary	116

Literaturverzeichnis

Verzeichnis der Tabellen und Abbildungen

Anhang

Verzeichnis der verwendeten Abkürzungen

1 Einleitung und Zielstellung

Die Bewertung nicht zählbarer oder messbarer Merkmale wie der Pilzbefall bei Pflanzen oder die Kondition bei Tieren mit Bonituren hat im landwirtschaftlichen Versuchswesen eine große Bedeutung. Bei der Analyse werden Bonituren als geordnete kategoriale Merkmale angesehen. Dabei besteht eine aktuelle Fragestellung aus Sicht der Datenauswertung darin, mit welcher Methode Bonituren zu analysieren sind.

Für die Auswertung von geordneten kategorialen Merkmalen können verschiedene methodische Ansätze verwendet werden.

Dazu zählen:

- das Schwellenwertmodell (McCullagh, 1980),
- die Anwendung von Rangverfahren (Brunner et al., 2001; 2002),
- die Anwendung von Permutationsverfahren (Schumacher et al., 1995 a, b; 1997),
- lineare Modelle und damit Unterstellung von Normalverteilung für alle zufälligen Modelleffekte (Henderson, 1990).

Aus Sicht der praktischen Versuchsauswertung folgt die Frage, welche Vor- und Nachteile die Nutzung der angeführten Methoden hat und von welchen Bedingungen diese abhängen.

Eine besondere Herausforderung liegt vor, wenn neben den im landwirtschaftlichen Versuchswesen meist interessierenden festen Effekten auch zufällige Effekte zu berücksichtigen sind und diese über eine komplexe Kovarianzstruktur verfügen. Das ist beispielsweise bei wiederholten Beobachtungen je Pflanze oder Tier der Fall. Weiterhin bedeutsam sind räumlich bedingte Kovarianzen, die bei speziellen Untersuchungen im Tierversuchswesen, besonders aber im Feldversuchswesen eine Rolle spielen können. Wiederum aus Sicht der praktischen Versuchsauswertung stellt sich dann die Frage nach der Möglichkeit einer Beachtung dieser Kovarianzen bei Anwendung der angeführten Auswertungsverfahren. Weiterhin ist dann auch zu untersuchen, welche Konsequenzen für die Schätzung fester Effekte sowie Inferenzaussagen über die festen Effekte bei einer Vernachlässigung dieser Kovarianzen zu erwarten sind.

Aus den angeführten Problemen leitet sich die Zielstellung der vorliegenden Arbeit ab. Es ist ein umfassender Methodenvergleich bei Existenz einer komplexen Kovarianzstruktur durchzuführen. Dabei erfolgt die Untersuchung für einen im Feldversuchswesen typischen Versuchsplan bei Beachtung einer durch räumliche Abhängigkeiten bedingten Kovarianzstruktur.

Basierend auf einer stochastischen Simulation ergeben sich die folgenden Teilziele:

- Untersuchung der Genauigkeit geschätzter Effekte und deren Differenzen,
- Untersuchung der Einhaltung vorgegebener Konfidenzniveaus bzw. Fehler 1. Art für feste Effekte und deren Differenzen,
- Untersuchung der Einhaltung vorgegebener Wahrscheinlichkeiten für Boniturnoten und kumulativer Wahrscheinlichkeiten,
- Untersuchung der Abhängigkeit der zu treffende Aussagen von der Höhe der räumlichen Korrelation und dem Stichprobenumfang,
- Untersuchung des Konvergenzverhaltens der betrachteten Analysemethoden.

Die Bearbeitung der formulierten Teilziele erfordert die Erarbeitung eines flexiblen Simulationsprogramms für Bonituren und dessen Kombination mit verschiedenen Auswertungsmethoden. Die rechentechnische Umsetzung erfolgt mit der Statistiksoftware SAS, Version 9.2.

Die zu treffenden Aussagen beziehen sich im engeren Sinne wegen des betrachteten Versuchsplanes und der Beachtung räumlicher Kovarianzen auf das Feldversuchswesen, erlauben in ihren grundlegenden Schlussfolgerungen aber eine Übertragung auf alle Anwendungsbereiche, bei denen geordnete kategoriale Merkmale vorliegen, wie im Tierversuchswesen, aber auch in der Biologie oder Medizin.

2 Statistische Modelle und Auswertungsmethoden

In diesem Kapitel werden die verwendeten statistischen Modelle und Auswertungsmethoden beschrieben. Das in Abschnitt 2.1 beschriebene lineare gemischte Modell bildet die Grundlage der Datensimulation (Abschnitt 3). Dieses Modell wird verwendet und den übrigen Modellansätzen gegenübergestellt, wenn für Datensätze mit einer oder mehr als einer Beobachtung pro Parzelle die Bonituren oder deren Parzellenmittelwerte als kontinuierliche Variable aufgefasst und dementsprechend ausgewertet werden (Abschnitt 2.5 und 2.6). In Abschnitt 2.2 wird der derzeitige Kenntnisstand zur Auswertung von geordneten kategorialen Daten bei Nutzung des Schwellenwertmodells sowie seine Erweiterung bei zusätzlichen zufälligen Effekten und komplexen Kovarianzstrukturen beschrieben. Weitere Möglichkeiten der Auswertung geordneter kategorialer Daten bestehen in der Nutzung von Rangverfahren oder Permutationstests. Diese Modelle werden in Abschnitt 2.3 und 2.4 dargestellt.

2.1 Das lineare gemischte Modell (LMM)

Das lineare gemischte Modell bildet in dieser Arbeit unter anderem die Grundlage der Simulation der latenten Variablen und wird im Folgenden zur thematischen Einführung, insbesondere in das Konzept räumlicher Kovarianzen, dargestellt.

Die Analyse von stetigen, normalverteilten Daten mit festen und zufälligen Einflussgrößen führt zu einem linearen gemischten Modell (Searle, 1971). Im Englischen wird es als „linear mixed model“, abgekürzt mit LMM, bezeichnet. Die Abkürzung LMM wird in den folgenden Abschnitten verwendet. Innerhalb der gut ausgebauten Theorie für dieses Modell können neben festen sowie zufälligen Effekten unterschiedliche Kovarianzstrukturen der Resteffekte und der zufälligen Effekte modelliert werden. Sei Y ein Vektor von stetigen, normalverteilten Zufallsvariablen, für den das folgende Modell gilt:

$$(2.1-1) \quad Y = X\beta + Zu + e$$

mit: β = Vektor der unbekanntesten festen Effekte

u = Vektor der unbekanntesten zufälligen Effekte

e = Vektor der unbekanntesten zufälligen Resteffekte

X, Z = bekannte Versuchsplanmatrizen der festen bzw. zufälligen Effekte

mit den Annahmen

$$u \sim N(0, G), \quad e \sim N(0, R), \quad E(Y) = X\beta, \quad \text{Var}(Y) = ZGZ' + R = V,$$

und

$$(2.1-2) \quad \text{Var} \begin{pmatrix} Y \\ u \\ e \end{pmatrix} = \begin{bmatrix} V & ZG & R \\ GZ' & G & 0 \\ R & 0 & R \end{bmatrix},$$

wobei in der Diagonalen Varianzen stehen und an den anderen Positionen Kovarianzen. Sind G und R und dadurch V bekannt, so können für die Linearkombinationen der festen Effekte lineare, erwartungstreue Schätzungen mit kleinster Varianz („best linear unbiased estimation“ = BLUE) aus

$$(2.1-3) \quad K' \hat{\beta} = K'(X'V^{-1}X)^{-1} X'V^{-1} y$$

mit

$$(2.1-4) \quad \text{var}(K' \hat{\beta}) = K'(X'V^{-1}X)^{-1} K$$

und K = Matrix zur Beschreibung der gewünschten Linearkombinationen der festen Effekte, gewonnen werden.

Bei praktischen Anwendungen ist jedoch V unbekannt und wird durch eine Schätzung \hat{V} aus den Daten ersetzt. Für die Bereitstellung der erforderlichen Schätzwerte hat sich die „Restricted Maximum Likelihood“ (REML) Schätzung durchgesetzt (Searle et al., 1992, p. 254). Dies hat zur Folge, dass die wünschenswerten Eigenschaften (Erwartungstreue und minimale Varianz) der Schätzungen aus (2.1-3) nur noch asymptotisch, also für großen Stichprobenumfang, gelten. Im Allgemeinen existieren dann keine exakten Tests über die festen Effekte und es werden approximative Tests, so genannte Wald-F- und t-Tests, eingesetzt (Khuri et al., 1998). Für eindimensionale Nullhypothesen der Form $H_0: K' \beta = 0$ mit $\text{Rang}(K) = 1$ stellt K keine Matrix, sondern einen Vektor k dar. Diese eindimensionalen Hypothesen werden getestet mit der Prüfgröße

$$(2.1-5) \quad t = \frac{k' \hat{\beta}}{\sqrt{k'(X' \hat{V}^{-1} X)^{-1} k}} \sim t_{FG},$$

mit k' = Zeilenvektor zur Beschreibung der Linearkombination der festen Effekte.

Im mehrdimensionalen Fall ($\text{Rang}(K) > 1$) werden die Hypothesen getestet mit den Prüfgrößen

$$(2.1-6) \quad F = \frac{\hat{\beta}' K (K'(X' \hat{V}^{-1} X)^{-1} K)^{-1} K' \hat{\beta}}{\text{Rang}(K)} \sim F_{\text{Rang}(K), FG}.$$

und

$$(2.1-7) \quad \chi^2 = (\mathbf{K}'\hat{\beta})'[\mathbf{K}'(\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X})^{-1}\mathbf{K}]^{-1}(\mathbf{K}'\hat{\beta}) \sim \chi_{\text{FG}=\text{Rg}(\mathbf{K})}^2$$

wobei in allen Prüfgrößen $\mathbf{K}'(\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X})^{-1}\mathbf{K} = \widehat{\text{Var}}(\mathbf{K}'\hat{\beta})$.

Die unter praktischen Bedingungen stets vorliegende Situation der Verwendung von $\hat{\mathbf{V}}$ bedingt die Notwendigkeit der Bestimmung der Freiheitsgrade (abgekürzt mit FG) der Prüfgrößen. Die exakten Freiheitsgrade der Prüfgrößen sind nicht bekannt. In dieser Arbeit wird für das lineare gemischte Modell die Freiheitsgradapproximation von Kenward und Roger (2009) genutzt (siehe auch Abschnitt 4.1). Zusätzlich wird durch die von Kenward und Roger (1997; 2009) vorgeschlagene Methode die durch den Einsatz von $\hat{\mathbf{V}}$ verursachte Unterschätzung der Standardfehler in (2.1-4) korrigiert.

Räumliche Kovarianzen im linearen gemischten Modell

Ziel der Aufnahme von räumlichen Kovarianzen in das Modell ist die flexible Berücksichtigung von räumlich abhängigen Effekten bei der Auswertung von Versuchen. Die Annahme einer räumlichen Kovarianzstruktur folgt der Grundüberlegung, wonach in der Fläche nah beieinander liegende Messwerte ähnlicher sind als solche, die weiter voneinander entfernt liegen. Zur Analyse des Zusammenhangs zwischen zwei Zufallsvariablen wird in der Statistik die Kovarianz genutzt. Dieses Konzept kann auf räumliche Prozesse ausgeweitet werden (Cressie, 1993; Webster and Oliver, 2001). Im räumlichen Modell wird jeder Zufallsvariable zusätzlich eine x- und y-Koordinate auf der Fläche zugeordnet. Die Kovarianzen werden in Abhängigkeit der räumlichen Distanz formuliert, wodurch die Varianz-Kovarianz-Matrix \mathbf{R} (im Folgenden abgekürzt mit VC-Matrix) für die Resteffekte die folgende Struktur annimmt:

$$\mathbf{R} = \begin{bmatrix} \sigma^2 & \sigma_{12} & \cdots & \sigma_{1j} & \cdots & \sigma_{1n-1} & \sigma_{1n} \\ \sigma_{21} & \sigma^2 & & & & & \sigma_{2n} \\ \vdots & & \ddots & & & & \vdots \\ \sigma_{i1} & & & \sigma^2 & & & \sigma_{in} \\ \vdots & & & & \ddots & & \vdots \\ \sigma_{n-11} & & & & & \sigma^2 & \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nj} & \cdots & \sigma_{nn-1} & \sigma^2 \end{bmatrix}.$$

In der Diagonale steht der so genannte *Sill* aus dem räumlichen Modell $\sigma^2 = \sigma_{\text{R}}^2 + \sigma_0^2$, wobei σ_{R}^2 der *Nugget*-Effekt und σ_0^2 der so genannte „*Partielle Sill*“ sind. Der Sill stellt damit die Summe der räumlich strukturierten sowie der nicht räumlich strukturierten Varianz dar. Die Nichtdiagonalelemente $\sigma_{ij} = \sigma_0^2 f(d_{ij})$ entsprechen den Kovarianzen zwi-

schen zwei Beobachtungen, wobei d_{ij} die Euklidische Distanz für den räumlichen Abstand zwischen zwei Beobachtungen mit den Koordinaten (x_i, y_i) und (x_j, y_j) darstellt mit:

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

(x = horizontale und y = vertikale Koordinate z. B. zweier Beobachtungen)

und $f(d_{ij})$ vom räumlichen Modell abhängt. Wird für die räumliche Kovarianzstruktur beispielsweise das sphärische Modell angenommen, berechnen sich die Elemente von R durch Nutzung der Kovarianzfunktion $C(d_{ij})$, gegeben mit:

$$(2.1-8) \quad C(d_{ij}) = \begin{cases} \sigma_R^2 + \sigma_0^2 & d_{ij} = 0 \\ \sigma_0^2 \left[1 - \frac{3d_{ij}}{2\rho} + \frac{d_{ij}^3}{2\rho^3} \right] & 0 < d_{ij} \leq \rho \\ 0 & d_{ij} > \rho \end{cases}$$

mit $\sigma^2 = \sigma_R^2 + \sigma_0^2$: „Sill“,

σ_R^2 : „Nugget“-Effekt ,

d_{ij} : Euklidische Distanz zwischen zwei Messpunkten

und ρ : Reichweiten („Range“) - Parameter.

Im sphärischen Modell ist für $d_{ij} > \rho$ die Kovarianz gleich 0. Das angeführte räumliche Modell geht von Isotropie aus. Dies bedeutet, dass die Kovarianz nur vom Abstand, aber nicht von der Richtung abhängt (Cressie, 1993, p. 61 - 62).

Liegt eine räumliche Kovarianzstruktur vor, so kann diese im gemischten linearen Modell auf unterschiedliche Art berücksichtigt werden. Beispielsweise kann sie, wie oben beschrieben, durch eine entsprechende Struktur der VC-Matrix R der Resteffekte formuliert werden. Es kann aber auch versucht werden, die vorhandenen Kovarianzstrukturen über die Einbeziehung eines Blockeffektes zumindest teilweise abzubilden (Stroup, 2002). Der Blockeffekt kann als fixer Effekt innerhalb β oder als zufälliger Effekt innerhalb u berücksichtigt werden. Zusätzlich kann bei mehreren Beobachtungen pro Parzelle auch ein Parzelleneffekt formuliert werden. Wird keine dieser Möglichkeiten genutzt und nach einem Modell mit vollständiger Randomisation ausgewertet, so wird die existierende Kovarianz vernachlässigt.

2.2 Das Schwellenwertmodell (SWM)

2.2.1 Basisidee und Modell für feste Effekte und unabhängige Resteffekte

Eine Möglichkeit zur Analyse von Boniturnoten besteht in der Anwendung der Multinomialverteilung und kann durch die Nutzung des Schwellenwertmodells (McCullagh, 1980) erfolgen. Dieses Modell ist den generalisierten linearen Modellen (engl.: „generalized linear models“) zuzuordnen und wird im Folgenden mit SWM abgekürzt. Das Schwellenwertmodell unterstellt, dass dem kategorialen Merkmal Y eine stetige, aber nicht beobachtbare (latente) Zufallsgröße Z zugrunde liegt. Für die latente Zufallsgröße Z wird ein lineares Modell angenommen und ihr Erwartungswert wird als linearer Prädiktor bezeichnet. Die Funktion, die den Zusammenhang zwischen den kumulativen Wahrscheinlichkeiten in der Boniturskala und dem linearen Prädiktor herstellt, wird als inverse Linkfunktion bezeichnet. Als Verteilung für die latente Zufallsgröße kann beispielsweise die Normalverteilung oder die logistische Verteilung angenommen werden. Der Zusammenhang zwischen dem kategorialen Merkmal Y und der latenten Zufallsgröße Z wird über so genannte Schwellenwerte hergestellt.

Sei Y_{ij} die Zufallsgröße zugehörig zur j -ten Beobachtung ($j = 1, \dots, r$) innerhalb der Behandlung i ($i = 1, \dots, a$).

Dann gilt $Y_{ij} = k$ genau dann, falls $Z_{ij} \in [\theta_{k-1}, \theta_k]$, $k = 1, \dots, m$ ($m =$ Anzahl der verwendeten Boniturstufen)

Für die Schwellenwerte θ_k gilt: $\theta_k > \theta_{k-1}$ für $k = 1, \dots, m$ mit $\theta_0 = -\infty$ und $\theta_m = +\infty$.

Für die latente Variable gelte die Annahme eines linearen Modells:

$$(2.2-1) \quad Z_{ij} = x'_{ij} \beta + e_{ij}$$

mit $\beta =$ Vektor der unbekanntesten festen Effekte,

$x'_{ij} =$ bekannter Versuchsplanvektor der festen Effekte.

Es sei $e = (e_{11}, \dots, e_{ar})'$ der Vektor der Resteffekte, dann gilt:

$\text{Var}(e) = R$, $E(e) = 0$ wobei R die Kovarianzmatrix der zufälligen Resteffekte darstellt. Dabei wird hier noch keine Annahme über die Verteilung der Resteffekte getroffen. Neben den Resteffekten liegen keine weiteren zufälligen Effekte vor.

Für den Erwartungswert von Z_{ij} gilt beispielsweise:

$$(2.2-2) \quad E(Z_{ij}) = \eta_i \quad \text{mit} \quad \eta_i = \mu + \beta_i,$$

mit μ = allgemeines Mittel und β_i = fester Effekt der i-ten Behandlung. In Darstellung (2.2-2) steht η_i für den so genannten linearen Prädiktor.

a) latente Variable normalverteilt

Wird für Z_{ij} eine Normalverteilung mit $E(Z_{ij}) = \eta_i$ und $\text{Var}(Z_{ij}) = \sigma_e^2$ unterstellt, so folgt für die kumulativen Wahrscheinlichkeiten innerhalb der Boniturklassen:

$$(2.2-3) \quad q_{ik} = \Pr(Y_{ij} \leq k) = \Pr(Z_{ij} \leq \theta_k) = \Pr\left(\frac{Z_{ij} - \eta_i}{\sigma_e} \leq \frac{\theta_k - \eta_i}{\sigma_e}\right) = \Phi\left(\frac{\theta_k - \eta_i}{\sigma_e}\right)$$

mit $\Phi(z)$ = Verteilungsfunktion der Standardnormalverteilung, deren inverse Funktion die so genannte „probit“ Linkfunktion ergibt. Da der Quotient $\frac{\theta_k - \eta_i}{\sigma_e}$ nicht eindeutig identifizierbar ist, wird gefordert: $\text{Var}(Z_{ij}) = \sigma_e^2 = 1$.

b) latente Variable logistisch verteilt

Unterstellt man für Z_{ij} eine logistische Verteilung mit $E(Z_{ij}) = \eta_i$ und $\text{Var}(Z_{ij}) = \pi^2 / 3$ so folgt für die kumulativen Wahrscheinlichkeiten innerhalb der Boniturklassen:

$$(2.2-4) \quad q_{ik} = \Pr(Y_{ij} \leq k) = \Pr(Z_{ij} \leq \theta_k) = \Psi(\theta_k - \eta_i)$$

mit der logistischen Verteilungsfunktion $\Psi(z) = \frac{e^z}{1 + e^z}$, deren inverse Funktion die so genannte „logit“ Linkfunktion ergibt.

Das Modell mit Normalverteilung der latenten Variablen wird in der Literatur auch als „cumulative probit model“ (CPM) und dasjenige mit logistischer Verteilung der latenten Variablen auch als „cumulative logit model“ bezeichnet (Agresti, 2002, p. 275, 283). In Anlehnung an die innerhalb von SAS genutzten Bezeichnungen werden die Linkfunktionen im Folgenden mit „cumprobit“ und „cumlogit“ bezeichnet.

2.2.2 Schwellenwertmodell mit zufälligen Effekten

Werden im Schwellenwertmodell außer fixen Effekten und neben den Resteffekten weitere zufällige Effekte betrachtet, hat dies Konsequenzen für die Schätzung der kumulativen

Wahrscheinlichkeiten der Boniturklassen. Wird ein zufälliger Effekt wie beispielsweise ein Parzelleneffekt im Falle mehrerer Beobachtungen je Parzelle in das Schwellenwertmodell aufgenommen, so muss das Modell (2.2-1) auf ein generalisiertes lineares gemischtes Modell erweitert werden (McCulloch and Searle, 2001). Im Englischen werden diese Modelle als „generalized linear mixed model“, abgekürzt als GLMM, bezeichnet. Diese Abkürzung wird nachfolgend verwendet. Die Einführung zufälliger Effekte führt dazu, dass die Schätzung für die kumulativen Wahrscheinlichkeiten in Abhängigkeit der Schätzung des zufälligen Effektes erfolgt. Es werden also bedingte Wahrscheinlichkeiten pro Parzelle geschätzt. Sollen so erhaltene Ergebnisse mit denjenigen von Modellen ohne zufällige Effekte verglichen werden, so müssen die bedingten Wahrscheinlichkeiten in marginale Wahrscheinlichkeiten umgerechnet werden. Die Aussagen sind dann gültig für eine Parzelle, zufällig herausgegriffen aus einer größeren Grundgesamtheit. Diese Grundgesamtheit ist charakterisiert durch die Parzellenvarianz und durch die Verteilungsannahme hinsichtlich der zufälligen Parzelleneffekte. Die Transformation in bedingte Wahrscheinlichkeiten erfolgt gewöhnlich unter der Annahme, dass der Parzelleneffekt Null ist. Es wird also eine Parzelle betrachtet, die dem Mittel aller Parzellen im Versuch entspricht.

Sei y_{ij} (Realisation der Zufallsvariable Y_{ij}) die j -te Beobachtung ($j = 1, \dots, n$) einer Parzelle mit Behandlung i ($i = 1, \dots, a$) innerhalb Block l ($l = 1, \dots, r$).

Dann gilt für die latente Variable

$$(2.2-5) \quad Z_{ij} = \eta_{li} + u_{li} + e_{lij}$$

mit $e_{lij} \sim N(0, \sigma_e^2)$

und

$$(2.2-6) \quad \eta_{li} = \mu_{li} + u_{li} \quad \text{mit} \quad \mu_{li} = \mu + \alpha_l + \beta_i \quad \text{und} \quad u_{li} \sim N(0, \sigma_u^2)$$

mit $\mu =$ allgemeines Mittel

$$\alpha_l = \text{fester Effekt des } l\text{-ten Blocks}$$

$$\beta_i = \text{fester Effekt der } i\text{-ten Behandlung}$$

$$u_{li} = \text{zufälliger Effekt der Parzelle mit Behandlung } i \text{ innerhalb Block } l$$

Die bedingte Wahrscheinlichkeit einer Boniturklasse für Parzelle i innerhalb Block l ist unter Annahme der Normalverteilung für die latente Variable gegeben mit:

$$(2.2-7) \quad q_{lik} = \Pr(Y_{lij} \leq k \mid u_{li}) = \Pr(Z_{lij} \leq \theta_k) = \Pr\left(\frac{Z_{lij} - \eta_{li}}{\sigma_e} \leq \frac{\theta_k - \eta_{li}}{\sigma_e}\right) = \Phi\left(\frac{\theta_k - \eta_{li}}{\sigma_e}\right)$$

Wegen $u_{li} + e_{lij} \sim N(0, 1 + \sigma_u^2)$ folgt für die marginalen Wahrscheinlichkeiten (McCulloch and Searle, 2001, p. 244) die Darstellung:

$$(2.2-8) \quad q_{lik}^* = \Pr\left(\frac{u_{li} + e_{lij}}{\sqrt{1 + \sigma_u^2}} \leq \frac{\theta_k - \mu_{li}}{\sqrt{1 + \sigma_u^2}}\right) = \Phi\left(\frac{\theta_k - \mu_{li}}{\sqrt{1 + \sigma_u^2}}\right) = \Phi(c \cdot (\theta_k - \mu_{li}))$$

mit $c = (1 + \sigma_u^2)^{-\frac{1}{2}}$.

Für die Annahme der logistischen Verteilung der latenten Variablen ist die bedingte Wahrscheinlichkeit einer Boniturklasse gegeben mit:

$$(2.2-9) \quad q_{lik} = \Pr(Y_{lij} \leq k \mid u_{li}) = \Pr(Z_{lij} \leq \theta_k) = \Psi(\theta_k - \eta_{li})$$

Die marginale Wahrscheinlichkeit kann für das Modell der logistisch verteilten latenten Variablen nur näherungsweise angegeben werden, indem die logistische Verteilung durch die Normalverteilung approximiert wird. Die marginale Wahrscheinlichkeit berechnet sich dann nach:

$$(2.2-10) \quad q_{lik}^* = \Psi(a \cdot (\theta_k - \mu_{li}))$$

mit $a = \frac{1}{\sqrt{1 + \tilde{c}^2 \cdot \sigma_u^2}}$ und $\tilde{c} = \frac{16}{15} \cdot \frac{\sqrt{3}}{\pi}$.

Wobei \tilde{c} der Faktor zur Approximation der logistischen Verteilung durch die Normalverteilung ist.

Natürlich können auch mehrere zufällige Faktoren, wie beispielsweise zufällige Block- und Parzelleneffekte, in das Modell aufgenommen werden.

Dann gelte für die latente Variable

$$(2.2-11) \quad Z_{lij} = \mu_i + a_1 + u_{li} + e_{lij}$$

mit $a_1 \sim N(0, \sigma_{\text{Block}}^2)$, $u_{li} \sim N(0, \sigma_u^2)$, $e_{lij} \sim N(0, \sigma_e^2)$.

Für den bedingten Erwartungswert des linearen Prädiktors gelte

$$(2.2-12) \quad \eta_{li} = \mu_i + a_1 + u_{li} \quad \text{mit} \quad \mu_i = \mu + \beta_i$$

mit $a_1 =$ zufälliger Effekt des l-ten Blocks

$\beta_i =$ fester Effekt der i-ten Behandlung

$u_{li} =$ zufälliger Effekt der Parzelle mit Behandlung i innerhalb Block l.

Für die marginalen Wahrscheinlichkeiten werden unter der Annahme voneinander unabhängiger Parzellen- und Blockeffekte die Konstanten c in (2.2-8) und a in (2.2-10) wie folgt modifiziert:

$$(2.2-13) \quad c = (1 + \sigma_{\text{Block}}^2 + \sigma_u^2)^{-\frac{1}{2}} \quad \text{und} \quad a = \frac{1}{\sqrt{1 + \tilde{c}^2 \cdot (\sigma_{\text{Block}}^2 + \sigma_u^2)}}.$$

Die Einführung zufälliger Faktoren hat Konsequenzen für die Interpretation der Ergebnisse, wenn im Modell mehr als ein fixer Faktor vorhanden ist. Wird über einen fixen Faktor akkumuliert, beispielsweise über alle Blocks in Modell (2.2-6), so kann dies für bedingte und marginale Wahrscheinlichkeiten wie nachfolgend dargestellt auf unterschiedlichen Ebenen erfolgen.

a) Berechnung von bedingten Wahrscheinlichkeiten für eine Behandlung bei Akkumulation im linearen Prädiktor je Block

Sei $\eta_{li} = \mu + \alpha_1 + \beta_i + u_{li}$ der lineare Prädiktor für eine Parzelle u_{li} der Behandlung i innerhalb des fixen Block l, dann ergibt die Berechnung von bedingten Wahrscheinlichkeiten für eine Behandlung i durch Akkumulation über alle Parzellen der Behandlung i im linearen Prädiktor:

$$(2.2-14) \quad \bar{\eta}_i = \mu + \frac{\sum_{l=1}^r \alpha_1}{r} + \beta_i + \frac{\sum_{l=1}^r u_{li}}{r}.$$

Unter der Annahme $\sum_{l=1}^r u_{li} = 0$ ergibt sich formal folgende bedingte Wahrscheinlichkeit für Note k:

$$(2.2-15) \quad q_{ik} = \Phi(\theta_k - \bar{\eta}_i) \quad \text{mit} \quad \bar{\eta}_i = \mu + \beta_i + \frac{\sum_{l=1}^r \alpha_1}{r}.$$

Die Wahrscheinlichkeit q_{ik} stellt eine bedingte Wahrscheinlichkeit dar unter der Annahme, dass die Parzelleneffekte mit Behandlung i im Mittel Null ergeben.

b) Berechnung von marginalen Wahrscheinlichkeiten für eine Behandlung bei Akkumulation der geschätzten kumulativen Wahrscheinlichkeiten je Block

Sei die marginale kumulative Wahrscheinlichkeit für eine Note $\leq k$ einer Behandlung i im fixen Block l

$$q_{lik}^* = \Phi(c \cdot (\theta_k - \eta_{li}^*))$$

mit $\eta_{li}^* = \mu + \alpha_l + \beta_i$ als Erwartungswert des linearen Prädiktors

$$\text{und } c = (1 + \sigma_u^2)^{-\frac{1}{2}}.$$

Dann erhält man mittlere marginale Wahrscheinlichkeiten wie folgt:

$$(2.2-16) \quad q_{ik}^* = \frac{\sum_{l=1}^r q_{lik}^*}{r}.$$

2.2.3 Schwellenwertmodell mit räumlicher Kovarianzstruktur

Im Schwellenwertmodell werden die räumlichen Kovarianzen zwischen allen Beobachtungen in Abhängigkeit der räumlichen Distanz formuliert, indem die Kovarianzmatrix R der Resteffekte der latenten Variablen modifiziert wird. Dies wird analog zu Abschnitt 2.1 durchgeführt. Allerdings erlauben nicht alle Programme die rechentechnische Umsetzung dieser Modifizierung der R -Matrix. Beispielsweise ist in der Statistiksoftware SAS innerhalb der Prozedur GLIMMIX, die zur Lösung des Schwellenwertmodells genutzt werden kann, die Umsetzung von Kovarianzstrukturen nur über die Modifizierung der Varianz-Kovarianz-Matrix G der zufälligen Effekte gegeben (Schabenberger, 2005). Nähere Erläuterungen zur rechentechnischen Umsetzung der räumlichen Kovarianzstrukturen im Schwellenwertmodell werden in Abschnitt 4.2 gegeben.

2.2.4 Berechnung von Konfidenzintervallen für kumulative Wahrscheinlichkeiten

Bei Nutzung des Schwellenwertmodells können Konfidenzintervalle für die kumulativen Wahrscheinlichkeiten der Boniturnoten berechnet werden. Diese werden beispielsweise in der SAS Prozedur GLIMMIX für das „cumulative probit model“ nach

$$(2.2-17) \quad \text{KI}_{\text{probit}} = [\phi(\hat{\eta} - d); \phi(\hat{\eta} + d)],$$

$\phi(z)$ = Verteilungsfunktion der Standardnormalverteilung

und für das „cumulative logit model“ nach

$$(2.2-18) \quad \text{KI}_{\text{logit}} = [\psi(\hat{\eta} - d); \psi(\hat{\eta} + d)], \quad \psi(z) = \frac{e^z}{1 + e^z}$$

und in beiden Fällen mit $d = t(1 - \frac{\alpha}{2}; \text{FG}) \cdot \text{SE}(\hat{\eta})$ berechnet.

Mit dieser Vorgehensweise wird das Konfidenzintervall im linearen Prädiktor gebildet und die Grenzen werden zurück transformiert. Somit ist gesichert, dass die Intervallgrenzen zwischen 0 und 1 liegen. Alternativ zur Vorgehensweise in SAS können die Konfidenzintervalle direkt für die kumulativen Wahrscheinlichkeiten nach

$$(2.2-19) \quad \text{KI} = [(\hat{p} - d); (\hat{p} + d)] \quad \text{mit} \quad d = t(1 - \frac{\alpha}{2}; \text{FG}) \cdot \text{SE}(\hat{p})$$

berechnet werden, womit Grenzen kleiner Null und größer Eins entstehen können. In dieser Arbeit werden beide Varianten verglichen. Für die Berechnung der Konfidenzintervalle werden die Freiheitsgrade der verwendeten t-Verteilung wie folgt bestimmt:

$\text{FG} = N - \text{Rang}(X)$, mit N = Gesamtstichprobenumfang und $\text{Rang}(X)$ = Rang der Versuchspanmatrix der festen Effekte.

Sind im Modell zufällige Faktoren enthalten, so gilt die Berechnung der Konfidenzintervalle nach (2.2-19) nur für die bedingten Wahrscheinlichkeiten. Für die marginalen Wahrscheinlichkeiten (siehe Abschnitt 2.2.2) wird die Varianz (und damit der Standardfehler) über die sogenannte Delta-Methode (Greene, 2008, p. 1056) bestimmt. Die Berechnung der Varianzen nach dieser Methode für unterschiedliche Modelle findet sich in Anhang 1.

2.3 Rangverfahren

Eine weitere Alternative zur Auswertung von Boniturdaten stellen nichtparametrische Methoden dar. Eine Möglichkeit innerhalb dieser Methodenklasse stellen die sog. „Rangverfahren“ dar. Nachfolgend werden die Rangverfahren erläutert, wobei die Darstellungsweisen und Modifikationen von Brunner et al. (1997; 2001; 2002) benutzt werden. Die Einbeziehung von räumlichen Kovarianzstrukturen ist für die Rangverfahren noch nicht ausgearbeitet. Es kann nur versucht werden, die räumliche Kovarianz über die Einbeziehung eines Blockeffektes abzubilden. Nähere Erläuterungen hierzu werden in Abschnitt 4.3 gegeben.

2.3.1 Ränge

Ränge stellen eine Transformation der ordinalen Daten in Rangzahlen dar. Gegeben eine Beobachtung y in einem Datensatz, ist der Rang von y die Anzahl der Beobachtungen kleiner oder gleich y . Ränge repräsentieren die zugrunde liegende Ordnung der Beobachtungen und sind invariant unter monotonen Transformationen. D. h., wenn beispielsweise eine Skala 0 bis 4 oder 0 bis 120 verwendet wird oder die Wurzel dieser Werte, bleibt der Rang der Beobachtungen unverändert. Differenzen von Rängen sind einfach zu interpretieren. Falls eine Rangdifferenz von 3 zwischen zwei Beobachtungen y_1 und y_2 auftritt, so folgt bei Datenreihen ohne Bindungen, dass 2 Beobachtungen zwischen den betrachteten Beobachtungen liegen.

2.3.2 Nichtparametrische Effekte und Hypothesen

Im Rahmen nichtparametrischer Analysen existieren keine Parameter wie beispielsweise Erwartungswert und Varianz für eine Normalverteilung, auf die Behandlungseffekte und Hypothesen bezogen werden können. Ein für viele Anwendungen verallgemeinerungsfähiger Ansatz besteht aber darin, Behandlungseffekte in Bezug auf die Verteilung der im jeweiligen Versuch ermittelten Daten zu beziehen. Angenommen, wir haben einen Versuch mit a Behandlungen ($i = 1, \dots, a$) und n_i unabhängigen Wiederholungen je Behandlung ($k = 1, \dots, n_i$). Die Beobachtung der k -ten Wiederholung von Behandlung i repräsentiert eine Realisation der Zufallsvariablen Y_{ik} mit der normalisierten Version der Verteilungsfunktion $F_i(y)$.

Die normalisierte Version der Verteilungsfunktion von Y_{ik} leitet sich von der links-stetigen und rechts-stetigen Form ab (Brunner und Munzel, 2002 p. 14ff). Danach nennt man $F(y) = P(Y < y)$ die links-stetige Version, $F^+(y) = P(Y \leq y)$ die rechts-stetige Version der Verteilungsfunktion. Für eine kontinuierliche Zufallsvariable wird ein einzelner Wert y mit der Wahrscheinlichkeit null angenommen. Somit gilt $P(Y = y) = 0$ und daher $F^+(y) = F(y)$. D. h., die rechts- und links-stetige Version sind in diesem Fall identisch. Brunner und Munzel (2002, p. 14/15) argumentieren, dass zur gemeinsamen Darstellung sowohl stetiger als auch diskreter Zufallsvariablen eine einheitliche Form vorteilhaft ist und verwenden die so genannte „normalisierte Version“ der Verteilungsfunktion nach Ruymgart (1980). Diese Version stellt den Mittelwert der rechts- und links-stetigen Version dar: $F(y) = \frac{1}{2}[F^+(y) + F^-(y)]$.

In einem Versuch ergibt sich das gewichtete Mittel $H(y)$ aller $F_i(y)$ aus

$$(2.3-1) \quad H(y) = \frac{1}{N} \sum_{i=1}^a n_i F_i(y),$$

wobei N der Gesamtanzahl der Beobachtungen entspricht $N = \sum_{i=1}^a n_i$ (Shah and Madden, 2004). Werden vereinfachend $H(y)$ mit H und $F_i(y)$ mit F_i bezeichnet, so ergibt sich der so genannte „relative Effekt“ p_i der i -ten Behandlung als:

$$(2.3-2) \quad p_i = \frac{1}{N} \sum_{j=1}^a \sum_{k=1}^{n_j} [P(Y_{jk} < Y_{i1}) + \frac{1}{2} P(Y_{jk} = Y_{i1})] = \int H dF_i$$

mit dF_i als 1. Ableitung von F_i . Der Wert p_i beschreibt allgemein die so genannte „stochastische Tendenz“ von F_i in Bezug auf H (Brunner und Munzel, 2002 p.15, 29; Shah and Madden, 2004).

Für den Fall $p_i > 1/2$ bedeutet das, dass die nach F_i verteilte Zufallsvariable Y_i tendenziell größere Werte annimmt als die nach einer mittleren Verteilung H verteilte Zufallsvariable Z . Umgekehrt gilt für $p_i < 1/2$, dass Y_i tendenziell kleinere Werte annimmt. Der Fall $p_i = 1/2$ beschreibt die Situation, wonach keine Tendenz für Y_i besteht, gegenüber der mittleren Verteilung H größer oder kleiner zu sein.

Dieser Sachverhalt wird für den Vergleich von zwei unabhängigen Zufallsvariablen Y_i und Y_j ausgenutzt, indem der Vergleich über die jeweiligen relativen Effekte p_i und p_j erfolgt. Dann bedeuten $p_i > p_j$, dass Y_i gegenüber Y_j zu größeren Werten tendiert. Entsprechend sind $p_i < p_j$ und $p_i = p_j$ zu interpretieren.

Die Bereitstellung von Schätzwerten \hat{p}_i für den relativen Effekt p_i gelingt bei Nutzung von Rängen. Sei R_{ik} der Rang von y_{ik} innerhalb aller N Beobachtungen, dann ergibt sich der mittlere Rang von Behandlung i aus:

$$(2.3-3) \quad \bar{R}_i = \frac{1}{n_i} \sum_{k=1}^{n_i} R_{ik}$$

und

$$(2.3-4) \quad \hat{p}_i = \frac{1}{n_i} \sum_{k=1}^{n_i} \frac{1}{N} \left(R_{ik} - \frac{1}{2} \right) = \frac{1}{N} \left(\bar{R}_i - \frac{1}{2} \right).$$

Der Beweis ist bei Brunner und Munzel (2002, p. 35) angegeben (vgl. auch Shah and Madden, 2004). Brunner und Munzel (2002, p. 43, 195) zeigen weiter, dass \hat{p}_i ein erwartungstreuer und konsistenter Schätzer für p_i ist. D. h., es gibt einen direkten Zusammenhang zwischen den mittleren Rängen und dem relativen Effekt. Liegen beispielsweise zwei Behandlungen mit $n_1=15$ und $n_2=15$ Beobachtungen vor. Ergeben sich weiter die

mittleren Ränge $\bar{R}_1 = 9.5$ und $\bar{R}_2 = 21.5$, so ergeben sich die relativen Effekte der Behandlungen nach (2.3-4) zu $\hat{p}_1 = \frac{1}{30} \left(9.5 - \frac{1}{2} \right) = 0.3$ und $\hat{p}_2 = \frac{1}{30} \left(21.5 - \frac{1}{2} \right) = 0.7$.

Basierend auf diesen Schätzungen lassen sich zwei Aussagen treffen:

- Beobachtungen von Behandlung 2 tendieren zu größeren Werten gegenüber einer Zufallsvariablen mit der Verteilung H;
- Beobachtungen von Behandlung 2 tendieren zu größeren Werten gegenüber einer kombinierten Verteilung aus beiden Behandlungen.

Das entspricht den Aussagen für Zufallsvariable mit Normalverteilung, für die man die Aussage treffen würde, dass der Erwartungswert von Behandlung 2 größer ist als das Gesamtmittel beider Behandlungen. Es ist zu betonen, dass die Schätzungen des relativen Effekts nicht von monotonen Transformationen abhängen.

Nach Brunner und Munzel (2002, p. 26, 52) kann für den speziellen Fall von zwei Behandlungen die Interpretation des relativen Effekts auch direkt vorgenommen werden. Anstelle der Nutzung von H kann der relative Effekt der Verteilungsfunktion F_2 in Bezug auf die Verteilungsfunktion F_1 auch mit

$$(2.3-5) \quad p = p_2 - p_1 + \frac{1}{2}$$

dargestellt werden (hier: $0.7 - 0.3 + 0.5 = 0.9$).

Wie bereits oben angeführt, bedeuten Schätzwerte $\hat{p} > \frac{1}{2}$ bei Beachtung der genutzten Berechnungsvorschrift, dass Beobachtungen von Behandlung 2 zu größeren Werten als Beobachtungen von Behandlung 1 tendieren und umgekehrt, keine Differenzen der Verteilungen führen zu $\hat{p}_1 = \hat{p}_2 = \frac{1}{2}$.

Abgeleitet von $\hat{p} > \frac{1}{2}$ gilt für das Beispiel, dass die Verteilung \hat{F}_1 links von \hat{F}_2 angeordnet ist.

2.3.3 Hypothesen und Tests

Statistische Hypothesen für die Behandlungseffekte sind formulierbar bei Nutzung der normalisierten Version der Verteilungsfunktion F. Die einfachste Hypothese H_0 besteht

in der Annahme von keinen Behandlungseffekten. Diese Hypothese kann geschrieben werden als

$$H_0^F : F_1 = F_2 = \dots = F_a \quad \text{oder} \quad H_0^p : p_1 = p_2 = \dots = p_a$$

(entspricht dem F-Test in der ANOVA $H_0 : \mu_1 = \mu_2 = \dots = \mu_a$).

Die Alternativhypothese besteht darin, dass sich mindestens ein relativer Effekt p_i von den übrigen relativen Effekten unterscheidet. D. h., wenn die Behandlungen ungleiche Verteilungen aufweisen, dann werden diese Verteilungsunterschiede durch Unterschiede in den relativen Effekten ausgedrückt.

Da die Datenanalyse auf den beobachteten Rängen (bspw. R_{ik}) basiert und die Ränge eine nichtlineare Transformation darstellen, werden die Varianzen der Ränge im Allgemeinen zwischen den Behandlungen variieren (Brunner and Puri, 2001). Das ist der wesentliche Grund, weshalb bei ordinalen Daten die Rangtransformation und Auswertung mit parametrischen Verfahren, wie von Conover und Iman (1981) vorgeschlagen, nicht geeignet ist.

Zwei Typen von Statistiken können verwendet werden, siehe beispielsweise Brunner und Munzel (2002) sowie Shah and Madden (2004):

- Wald-Typ Statistik (WTS)

- ANOVA-Typ Statistik (ATS).

Betrachten wir zur Erläuterung ein zweifaktorielles Modell (Zweifachklassifikation). Die Daten können als Realisation einer unabhängigen Zufallsvariablen angesehen werden:

$$Y_{ijk} \sim F_{ij}(y)$$

$$i = 1, \dots, a; j = 1, \dots, b; k = 1, \dots, n_{ij}.$$

Dabei ist $F_{ij}(y)$ die oben eingeführte normalisierte Version der Verteilungsfunktion F . Der Vektor der $a \cdot b$ Verteilungen wird mit $F = (F_{11}, \dots, F_{1b}, \dots, F_{a1}, \dots, F_{ab})'$ bezeichnet.

Seien weiter $\bar{F}_{i.} = \frac{1}{b} \sum_{j=1}^b F_{ij}$ und $\bar{F}_{.j} = \frac{1}{a} \sum_{i=1}^a F_{ij}$ die zugehörigen Spalten- und Zeilenmittel, so ergeben sich ähnlich der linearen Modelle die folgenden Effekte:

$$\text{Haupteffekt A: } A_i = \bar{F}_{i.} - \bar{F}_{..}$$

oder

Haupteffekt B: $B_j = \bar{F}_j - \bar{F}_{..}$.

Die Darstellung der Haupteffekte in Matrixform ergibt beispielsweise:

$$(2.3-6) \quad \begin{pmatrix} A_1 \\ \vdots \\ A_a \end{pmatrix} = \begin{pmatrix} \bar{F}_{1.} - \bar{F}_{..} \\ \vdots \\ \bar{F}_{a.} - \bar{F}_{..} \end{pmatrix} = (P_a \otimes \frac{1}{b} 1_b') F \quad ,$$

mit P_a = zentrierende Matrix für den Haupteffekt A und $1_b'$ = Eins-(Spalten-) Vektor der Dimension b (Brunner and Munzel, 2002, p. 140).

Damit lassen sich Hypothesen beispielsweise für die Haupteffekt wie folgt angeben:

$$H_0^F(A) : A_1 = A_2 = \dots = A_a = 0$$

und äquivalent:

$$H_0^F(A) : (P_a \otimes \frac{1}{b} 1_b') F = 0 \quad \text{bzw. verkürzt: } H_0^F(A) : C_a F = 0.$$

Bei Nutzung dieser Bezeichnungen ergeben sich nach Brunner und Munzel (2002, p 141ff) die folgenden Teststatistiken:

Wald-Typ Statistik (WTS)

Die Prüfgröße

$$(2.3-7) \quad Q(C_a) = N \cdot \hat{p}' C_a' (C_a \hat{V} C_a')^{-1} C_a \hat{p}$$

folgt unter $H_0^F(A) : C_a F = 0$ asymptotisch einer zentralen χ^2 -Verteilung mit den Freiheitsgraden $FG = \text{Rang}(C_a)$.

Dabei bedeuten:

C_a = Kontrastmatrix,

$$\hat{\mathbf{p}} = \frac{1}{N} \begin{pmatrix} \bar{R}_{11.} - \frac{1}{2} \\ \vdots \\ \bar{R}_{ab.} - \frac{1}{2} \end{pmatrix},$$

$$\hat{\mathbf{V}} = \text{VC-Matrix} \quad \text{mit} \quad \hat{\mathbf{V}} = N \cdot \text{diag} \left(\frac{\hat{\sigma}_{11}^2}{n_{11}}, \dots, \frac{\hat{\sigma}_{ab}^2}{n_{ab}} \right)$$

$$\text{mit} \quad \hat{\sigma}_{ij}^2 = \frac{1}{N^2 (n_{ij} - 1)} \sum_{k=1}^{n_{ij}} (R_{ijk} - \bar{R}_{ij.})^2 \quad \text{und} \quad N = \sum_{i=1, j=1}^{a,b} n_{ij}.$$

Allerdings ergaben Untersuchungen von Brunner et al. (1997), dass die Approximation durch die χ^2 -Verteilung bei kleinen Stichprobenumfängen schlecht ist und daher nominale Risiken nicht eingehalten werden. Dabei ist besonders bedeutsam, dass die Approximation umso schlechter ist, je größer $\text{Rang}(C_a)$ ausfällt. Daher stellt die Nutzung der Wald-Typ Statistik besondere Anforderungen an den Stichprobenumfang, bei geringem Stichprobenumfang ist die ANOVA-Typ Statistik zu bevorzugen.

ANOVA-Typ Statistik (ATS)

Brunner und Munzel (2002, p. 147) geben als Grund der schlechten Approximation der χ^2 -Verteilung die schlechte Schätzung der VC-Matrix $\hat{\mathbf{V}}$ an. Bei der ANOVA-Typ Statistik wird $\hat{\mathbf{V}}$ in etwas anderer Form berücksichtigt. Die Prüfgröße lautet zunächst:

$$(2.3-8) \quad Q^*(C_a) = N \cdot \hat{\mathbf{p}}' C_a' (C_a C_a')^{-1} C_a \hat{\mathbf{p}} = N \cdot \hat{\mathbf{p}}' T \hat{\mathbf{p}}$$

Die Verteilung von Q^* kann durch eine gestreckte χ^2 -Verteilung approximiert werden. Brunner und Munzel (2002, p 148) schlagen die Prüfgröße

$$(2.3-9) \quad F = \frac{1}{\text{Sp}(T\hat{\mathbf{V}})} N \cdot \hat{\mathbf{p}}' T \hat{\mathbf{p}} \quad \text{mit} \quad T = C_a' [C_a C_a']^{-1} C_a$$

vor.

Die Verteilung von F wird durch eine zentrale F-Verteilung approximiert.

$$\text{Dabei sind die Zähler-Freiheitsgrade} \quad \text{FG}_z = \frac{[\text{Sp}(T\hat{\mathbf{V}})]^2}{\text{Sp}(T\hat{\mathbf{V}} \cdot T\hat{\mathbf{V}})},$$

die Nenner-Freitsgrade $FG_N = \frac{[SP(D\hat{V})]^2}{Sp(D^2\hat{V}^2 \cdot \Lambda)}$

und es bedeuten $D = \text{diag}(T)$, $\Lambda = N \cdot \text{diag}((n_{11} - 1)^{-1}, \dots, (n_{ab} - 1)^{-1})$ sowie $SP = \text{Spur}$ der Matrix.

Simulationsuntersuchungen zeigen, dass F gut für Fälle $n_{ij} > 7$ geeignet ist, gegenüber der Wald-Typ Statistik jedoch mit einem Effizienzverlust gerechnet werden muss (Brunner und Munzel, 2002, p 147ff). D. h., in diesem Fall ist mit einem Verlust an Power zu rechnen.

2.4 Permutationstests zur Analyse von einer Bonitur je Parzelle

Einen weiteren nichtparametrischen Ansatz zur Analyse geordneter kategorialer Daten stellen die sog. „Permutationstests“ (auch „Randomisierungstests“) dar (Bortz et al., 2000). Wenn die Voraussetzungen der Normalverteilung und der Homoskedastizität der Daten nicht gegeben sind, sind Permutationstests eine verteilungsfreie Alternative zur klassischen Varianzanalyse. Synonym wird für „Permutationstest“ auch der Begriff „Randomisationstest“ oder „Rerandomisationstest“ verwendet. Für die Auswertung von Boniturdaten mit nur einer Beobachtung je Parzelle schlagen Schumacher et al. (1995a; 1995b; 1997) die Nutzung von Permutationsverfahren vor. Die Autoren haben diese Verfahren für bestimmte Versuchsanlagen innerhalb der Prozedur IML von SAS über Makros umgesetzt.

Die Permutationstests beruhen auf folgendem Prinzip: Aus den Daten der Stichprobe wird eine Teststatistik berechnet, danach werden die Daten wiederholt permutiert, d. h., sie werden den Behandlungen zufällig neu zugeordnet. Anschließend wird die Teststatistik der realisierten Daten mit den Werten der Teststatistik bei allen möglichen Permutationen der Daten (inklusive derjenigen der Stichprobe) verglichen (Edgington, 1995). Der Anteil dieser Teststatistiken, die einen Wert größer oder gleich der Teststatistik aus der Stichprobe haben, stellt die Überschreitungswahrscheinlichkeit dar.

Der große Vorteil der Permutationstests besteht darin, dass jede mögliche Statistik gewählt werden kann, die für das vorliegende Problem geeignet ist. Das bedeutet im vorliegenden Fall, die Teststatistik muss in der Lage sein, zwischen der Null- und der Alternativhypothese zu unterscheiden (Good, 1995, p. 4). Als Teststatistik kann beispielsweise die Differenz von arithmetischen Mittelwerten (Reiczigel, 2005), deren Spannweite (Petrondas and Gabriel, 1983), oder auch die Summe der Beobachtungswerte dienen (Good, 1995). Bei den von Schumacher et al. (1995a; 1995b; 1997) entwickelten Verfahren für die Analyse von Einzel-Bonituren wird als Teststatistik ein Maß für die Ungleichheit der Zuordnung der Beobachtungen zu den Prüfgliedern vorgestellt, das auf der so genannten „Behandlungstotalen“ basiert (siehe Abschnitt 2.4.1 und 2.4.2). Die Verteilung dieser Teststatistik unter der Nullhypothese kann exakt bestimmt werden, indem sie für

alle möglichen Zuordnungen der Beobachtungen zu den Prüfgliedern (Permutationen) berechnet wird. Da die Rechenkapazität im Allgemeinen nicht ausreicht, diese Verteilung exakt zu bestimmen, wird sie über Monte-Carlo-Simulationen geschätzt und für eine bestimmte Anzahl von Rerandomisationen die Teststatistik berechnet. Die von Schuhmacher et al. (1995a; 1995b; 1997) entwickelten Makros berechnen dann Überschreitungswahrscheinlichkeiten für die beobachtete Realisierung der Prüfgröße. Mit Hilfe dieser Verfahren können die Globalhypothese getestet und paarweise Vergleiche durchgeführt werden. Im Folgenden wird die Anwendung von Permutationsverfahren für einfaktorielles vollständig randomisierte Anlagen und einfaktorielles randomisierte Blockanlagen beschrieben. Die Einbeziehung von räumlichen Kovarianzstrukturen ist innerhalb der Permutationsverfahren nicht möglich. Wie bei den Rangverfahren, kann nur versucht werden die räumliche Kovarianz über die Einbeziehung eines Blockeffektes abzubilden. Nähere Erläuterungen hierzu werden in Abschnitt 4.4 gegeben.

2.4.1 Permutationstest für die einfaktorielles vollständig randomisierte Versuchsanlage

Sei y_{ij} der Boniturwert der j -ten Beobachtung einer Behandlung i ($i=1, \dots, a$), n_i die Anzahl der Beobachtungen auf der i -ten Stufe und $N = \sum n_i$ der Gesamtstichprobenumfang. Voraussetzung für den Einsatz des Permutationstests ist die Anlage des Versuchs entsprechend der vollständigen Randomisation. Dann ergibt sich die folgende Vorgehensweise:

Liegen keine Unterschiede zwischen den a Stufen des Einflussfaktors vor, dann sind alle Permutationen des N -dimensionalen Datenvektors

$$y' = (y_{11}, y_{12}, \dots, y_{1n}, y_{21}, \dots, y_{2n}, \dots, y_{a1}, \dots, y_{an}) \text{ gleichwahrscheinlich.}$$

Die Hypothesen für den Globaltest auf Prüfgliedunterschiede lauten:

H_0 : Die Verteilungen der Beobachtungsgrößen sind für alle Behandlungen gleich, d. h. keine unterschiedlichen Behandlungseffekte.

H_A : Mindestens zwei Behandlungen haben unterschiedliche Effekte.

Als Teststatistik führt Schumacher (2010) für die Auswertung von Einzel-Bonituren in einer einfaktorielles vollständig randomisierten Anlage die Größe T^2 ein. Dabei ist T^2 die so genannte gewichtete Behandlungstotale.

$$(2.4-1) \quad T^2 = \sum_{i=1}^a \frac{1}{n_i} T_i^2,$$

mit $T_i^2 = \text{Quadrat der Behandlungssumme je Prüfglied}$. Die Wichtung mit $1/n_i$ dient dazu, ungleiche Stichprobenumfänge der Behandlungsstufen zu berücksichtigen.

Für die Stichprobe wird aus den a Behandlungssummen t_1, t_2, \dots, t_a die Realisation der Teststatistik gebildet:

$$(2.4-2) \quad t_{\text{beob}}^2 = \sum_{i=1}^a \frac{1}{n_i} t_i^2, \quad \text{mit } t_i = \sum_{j=1}^{n_i} y_{ij} \text{ für alle } t_i (i=1, \dots, a).$$

Treten keine Bindungen zwischen den Beobachtungswerten auf, so sind unter Gültigkeit von H_0 alle $N!$ Permutationen des Beobachtungsvektors gleichwahrscheinlich. Zu jeder dieser Permutationen wird die Teststatistik T^2 berechnet. Da die Permutationen, in denen die Daten nur innerhalb ihrer Behandlung vertauscht wurden, zu demselben Ergebnis für die Teststatistik führen, können maximal $N!/(n_1! \cdot n_2! \cdot \dots \cdot n_a!)$ unterschiedliche Realisationen der Teststatistik gebildet werden. Aus ihnen kann die Verteilungsfunktion der Teststatistik exakt ermittelt werden. Treten Bindungen in den Beobachtungswerten auf, reduziert sich die Gesamtzahl der möglichen Permutationen.

Unter Gültigkeit von H_0 ist die Wahrscheinlichkeit für alle Realisationen von T^2 gleich groß:

$$(2.4-3) \quad P(T^2 = t_s^2) = \frac{n_1! \cdot n_2! \cdot \dots \cdot n_a!}{N!}$$

für alle $s = 1, 2, \dots, N!/(n_1! \cdot n_2! \cdot \dots \cdot n_a!)$ Realisationen von T^2 .

Für den Testentscheid wird nun der Anteil derjenigen Realisationen der Teststatistik ermittelt, für die $T^2 \geq t_{\text{beob}}^2$ gilt, und die Nullhypothese wird verworfen, falls $q = P(T^2 \geq t_{\text{beob}}^2) \leq \alpha$. Die Existenz von Bindungen in den Beobachtungswerten reduziert die Testgüte, da eine Realisation der Teststatistik mit $T^2 \geq t_{\text{beob}}^2$ durch mehrere Permutationen erhalten werden kann bzw. die Gesamtanzahl aller möglichen unterschiedlichen Realisationen der Teststatistik durch vorhandene Bindungen reduziert wird.

Da die Werte der n_i und a für eine exakte Bestimmung von q häufig zu groß sind, werden bei praktischen Anwendungen oft Anzahlen von kleiner $N!/(n_1! \cdot n_2! \cdot \dots \cdot n_a!)$ Permutationen berücksichtigt und somit die Überschreitungswahrscheinlichkeit q approximativ ermittelt.

Für paarweise Vergleiche wird der oben besprochene Randomisationstest für $a = 2$ durchgeführt, und es werden nur die Daten der betreffenden Prüfglieder permutiert. Falls keine Bindungen auftreten, können nur $(n_1! + n_2!)$ mögliche Permutationen mit maximal $(n_1 + n_2)!/(n_1! \cdot n_2!)$ verschiedenen Werten der Teststatistik T^2 auftreten. Die Anzahlen n_1 und n_2 sollten mindestens 3 und $n_1 + n_2$ mindestens 7 sein, da ansonsten beim zweiseiti-

gen Test keine Signifikanz auf dem 0.05-Niveau möglich ist. Das exakte Niveau ist beispielsweise bei $n_1 = 3, n_2 = 4$ mindestens gleich $1/35 = 0.0286$.

2.4.2 Permutationstest für die einfaktorielle randomisierte Blockanlage mit fixen Blocks

Sei y_{ij} der Boniturwert der i -ten Behandlung ($i = 1, \dots, a$) im j -ten Block ($j = 1, \dots, n$) und in jedem Block seien die a Behandlungen genau einmal vertreten. Voraussetzung für den Einsatz der Permutationstests in einfaktoriellen, randomisierten Blockanlagen ist die Unabhängigkeit zwischen den Blocks, d. h. die Zufallsvektoren (Y_{ij}, Y_{ij^*}) mit $j \neq j^*$ seien unabhängig. Dann können die Hypothesen analog wie in Abschnitt 2.4.1 formuliert werden:

H_0 : Die Verteilungen der Beobachtungsgrößen sind für alle Behandlungen gleich, d. h. keine unterschiedlichen Behandlungseffekte.

H_A : Mindestens zwei Behandlungen haben unterschiedliche Effekte.

Für die Blockanlagen führen Schumacher et al. (1995; 1995; 1997) wiederum die so genannte Behandlungstotale T^2 als Teststatistik ein. Aufgrund der Anwendung auf vollständige Blockanlagen ist hier eine Wichtung mit der Stichprobengröße nicht erforderlich (vgl. Abschnitt 2.4.1).

$$(2.4-4) \quad T^2 = \sum_{i=1}^a T_i^2, \quad \text{mit } T_i^2 \text{ Behandlungssumme je Prüfglied (bei vorausgesetztem gleichen Stichprobenumfang, vgl. zu 2.4-1).}$$

Aus den a Behandlungssummen $t_i = y_{i1} + y_{i2} + \dots + y_{in}, i = 1, \dots, a$ wird die Realisation

$$(2.4-5) \quad t_{\text{beob}}^2 = \sum_{i=1}^a t_i^2 \quad \text{der Teststatistik gebildet.}$$

Treten keine Bindungen zwischen den Beobachtungswerten auf, so sind unter der Nullhypothese (in jedem Block) alle $a!$ Permutationen des Beobachtungsvektors gleichwahrscheinlich, und die Gesamtzahl der möglichen Permutationen führt auf $(a!)^{n-1}$ mögliche Realisationen der Teststatistik T^2 . Diese Anzahl reduziert sich aber bei Vorhandensein von Bindungen. Aus diesen Realisationen kann die Wahrscheinlichkeitsverteilung der Teststatistik unter der Nullhypothese bestimmt werden.

Auch in der Blockanlage ist unter H_0 die Auftretenswahrscheinlichkeit für alle Realisationen der Prüfgröße dieselbe, gegeben mit:

$$(2.4-6) \quad P(T^2 = t_s^2) = \frac{1}{(a!)^{n-1}} \quad \text{für alle} \quad s = 1, 2, \dots, (a!)^{n-1} .$$

Analog zum Permutationstest bei vollständiger Randomisation (Abschnitt 2.4.1) wird die Nullhypothese verworfen, falls $q = P(T^2 \geq t_{\text{beob}}^2) \leq \alpha$.

Bei größeren Werten für a und n wird wiederum ein approximativer Test durchgeführt. Für die paarweisen Vergleiche werden nur die $a = 2$ Prüfglieder permutiert. Für Letztere sollte die Anzahl der Blocks wenigstens 6 sein, um beim zweiseitigen Test das 0.05-Niveau erreichen zu können. Im Falle kleiner Blockzahlen ($n < 7$) sind die multiplen Testverfahren für paarweise Vergleiche sehr konservativ (Schumacher and Frisch, 1995). Für diesen Fall werden die Beobachtungen innerhalb Block um den Blockmedian korrigiert (bei Unterstellung von additiven Block- und Behandlungseffekten). Anschließend wird mit diesen korrigierten Werten wie beim Permutationstest im vollständig randomisierten Design (Schumacher et al., 1995) verfahren

2.5 Auswertung der Parzellenmittelwerte der Boniturnoten mittels LMM

Eine weitere Möglichkeit der Auswertung von Boniturdaten besteht bei Vorliegen mehrerer Beobachtungen pro Parzelle, indem die arithmetischen Mittel der Boniturnoten je Parzelle gebildet und diese als Realisationen einer kontinuierlichen Zufallsvariablen angesehen werden. Weiter wird angenommen, dass die zufälligen Effekte zumindest approximativ einer Normalverteilung folgen und entsprechend als Modell das lineare gemischte Modell gemäß Abschnitt 2.1 angenommen werden kann. Die Beachtung der räumlichen Kovarianz im Auswertungsmodell kann für die Parzellenmittelwerte ebenfalls analog zu Abschnitt 2.1 erfolgen. Die Einführung eines Parzelleneffektes im Modell ist aber aufgrund der vorherigen Mittelwertbildung über die Boniturnoten innerhalb einer Parzelle nicht möglich. Falls die Beobachtungen jedoch wegen der Existenz einer räumlichen Kovarianz voneinander abhängig sind, weisen auch die Mittelwerte der Parzellen eine räumlich bedingte Kovarianz und Korrelation auf. Diese werden im Folgenden anhand einer normalverteilten stetigen Variablen mit $E(y_{ij}) = \mu$ und $\text{Var}(y_{ij}) = \text{Var}(e_{ij})$ hergeleitet.

Sei $y_{ij} = \mu + e_{ij}$ das Modell für die j -te Beobachtung ($j = 1, \dots, n$) in Parzelle i und es existiere eine bestimmte räumliche Abhängigkeit aller Beobachtungen.

Seien die e_{ij} die Resteffekte der stetigen, normalverteilten Variablen. Dann stehen bei der Betrachtung von zwei benachbarten Parzellen in Parzelle i die e_{ij} und in Parzelle i' die $e_{i'j}$ zur Verfügung, und es existieren Kovarianzen zwischen den Beobachtungen der benachbarten Parzellen $\text{cov}(e_{i1}, e_{i'1}), \text{cov}(e_{i1}, e_{i'2}), \dots, \text{cov}(e_{in}, e_{i'n})$ sowie den Beobachtungen innerhalb einer Parzelle $\text{cov}(e_{i1}, e_{i2}), \dots, \text{cov}(e_{in}, e_{in-1})$.

Die Parzellenmittelwerte ergeben sich nach: $\bar{y}_i = \frac{\sum y_{ij}}{n}$ und $\bar{y}_{i'} = \frac{\sum y_{i'j}}{n}$.

Die Korrelation zwischen den Parzellenmittelwerten ist dann

$$(2.5-1) \quad r_{\bar{y}_i, \bar{y}_{i'}} = \frac{\text{cov}(\bar{y}_i, \bar{y}_{i'})}{\sqrt{\text{var}(\bar{y}_i) \cdot \text{var}(\bar{y}_{i'})}}$$

mit

$$(2.5-2) \quad \text{cov}(\bar{y}_i, \bar{y}_{i'}) = \frac{1}{n^2} [\text{cov}(e_{i1}; e_{i'1}) + \text{cov}(e_{i1}; e_{i'2}) + \dots + \text{cov}(e_{in}; e_{i'n})]$$

und die Varianz der Beobachtungen innerhalb einer Parzelle ist

$$(2.5-3) \quad \text{var}(\bar{y}_i) = \frac{1}{n^2} \left[\sum_{j=1}^n \text{var}(e_{ij}) + 2 \sum_{\substack{i,j=1 \\ j < j'}}^n \text{cov}(e_{ij}; e_{ij'}) \right].$$

Die Tabelle 5.12 in Abschnitt 5 zeigt die Veränderung der Korrelation der Mittelwerte zwischen benachbarten Parzellen bei Existenz unterschiedlich starker räumlicher Kovarianz und unterschiedlicher Anzahl Beobachtungen je Parzelle.

2.6 Auswertung der Boniturnoten als stetiges Merkmal mittels LMM

Ähnlich wie für die Parzellenmittelwerte der Bonituren (vgl. Abschnitt 2.5) ist es auch denkbar, bei entsprechenden Annahmen die einzelnen Bonituren als stetiges Merkmal mittels LMM auszuwerten. Hierfür müssen viele Boniturstufen vorhanden sein, und die zufälligen Effekte sollten approximativ normalverteilt mit gleicher Varianz verteilt sein. Bei typischen Boniturerhebungen, wie sie beispielsweise bei der Beurteilung von Pflanzenbeständen hinsichtlich Lager oder Pilzbefall durchgeführt werden, sind 9 Boniturstufen vorhanden. Die Verteilung der zufälligen Effekte innerhalb der Behandlungen ist jedoch im Allgemeinen nicht mit einer Normalverteilung gleichzusetzen. Die Auswertung der Bonituren als stetiges Merkmal mit dem linearen gemischten Modell wird analog zu Abschnitt 2.1, inklusive aller Möglichkeiten zur Berücksichtigung der räumlichen Kovarianz, durchgeführt.

3 Datensimulation und Maßzahlen zur Auswertung der Simulationen

3.1 Simulierter Versuchsplan und Datenumfänge

Der simulierte Versuchsplan ist in Gesamtgröße sowie der Größe und Form der Teilstücke angelehnt an einen im Jahr 2003 von der Landesanstalt für Landwirtschaft und Fischerei Mecklenburg-Vorpommern in Gülzow angelegten Dauerfeldversuch. Als Versuchsplan wird eine 1-faktorielle randomisierte vollständige Blockanlage mit 20 Prüfgliedern und 4 Wiederholungen (Blocks) mit einer Gesamtgröße von $78 \text{ m} * 99 \text{ m} = 7722 \text{ m}^2$ und einer Parzellengröße von $9 \text{ m} * 9 \text{ m} = 81 \text{ m}^2$ simuliert. Die vier Blocks werden durch zwei Wege von 6 und 9 m Breite getrennt (Abb. 3.1). Die Prüfglieder werden innerhalb der Blocks für jeden Simulationslauf neu randomisiert. Für jede Parzelle werden 9 Beobachtungen simuliert. Jede Beobachtung wird mit Koordinaten in x- und y-Richtung versehen. Die Anordnung der Beobachtungen innerhalb der Parzellen ist in allen Parzellen gleich (Abb. 3.2). Die Anzahl der Beobachtungen je Parzelle musste jedoch aufgrund sehr langer Laufzeiten der Simulationen für 9 Beobachtungen reduziert werden. Dies erfolgte vor der Auswertung der Daten, durch Entfernen bestimmter Beobachtungen aus dem Datensatz. Zur Auswertung wurden 1 (5), 3 ((1),(5),(9)) oder 5 ((1),(3),(5),(7),(9)) Beobachtungen je Parzelle genutzt (Nummerierung gemäß Abbildung 3.2).

3.2 Erzeugung der Zufallszahlen

Die Erzeugung der Zufallszahlen und alle anderen rechentechnischen Umsetzungen zur Datensimulation erfolgten mit der Statistiksoftware SAS, Version 9.2.

Für die vorliegenden Untersuchungen wird ein Vektor Z von Zufallszahlen erzeugt, der dem nachfolgend angeführten Modell genügt:

$$(3.2-1) \quad Z = X\beta + e$$

mit: X = vorgegebene Versuchsplanmatrix der fixen Effekte

β = Vektor der vorgegebenen fixen Effekte

e = Vektor der zufälligen Resteffekte.

Weiter soll gelten:

$$Z \sim N(X\beta, R); e \sim N(0, R).$$

Je Simulationslauf wurde eine Realisation des Vektors e der Dimension N nach der folgenden Vorschrift erzeugt:

$$e = L_R \cdot r$$

Dabei entspricht L_R der Cholesky-Zerlegung von R ($L_R' \cdot L_R = R$) und r ist die Realisation eines Vektors der Dimension N von normalverteilten unabhängigen Zufallsgrößen, also $r \sim N(0, I_N)$.

Für die Erzeugung von r wurde der innerhalb der Prozedur IML bereitgestellte Zufallsgenerator „NORMAL“ zur Erzeugung von Realisationen unabhängiger normalverteilter Zufallsvariablen mit Erwartungswert 0 und Varianz 1 genutzt (SAS Institute Inc., 1988, p. 277).

Zur Erzeugung der VC-Matrix R für vorgegebene Varianzparameter, räumliche Funktionen und euklidische Distanzen wurde die Prozedur MIXED verwendet. Dabei wurde ein Datensatz mit einer Dummy-Variablen und den zugehörigen räumlichen Koordinaten entsprechend des vorgegebenen Planes (vgl. 3.1), durch Nutzung der Datei „Versuchsplan“, eingelesen. Weiter wurden die vorgegebenen Varianzparameter im Statement „parms“ mitgeteilt. Die zu verwendende räumliche Funktion wird im Statement „repeated“ festgelegt. Bei Unterdrückung jeglicher Iterationen wird die den Parametern, euklidischen Distanzen und der gewählten räumlichen Funktion entsprechende VC-Matrix R erzeugt. Nachfolgend ist das zugehörige SAS-Programm aufgeführt:

```
PROC MIXED DATA=Versuchsplan NOPROFILE;  
  MODEL dummy = ;  
  PARMs (sill) (range) (nugget) /NOITER;  
  REPEATED / SUBJECT=intercept LOCAL TYPE=SP(SPH) (x y) R;  
  ODS OUTPUT R = R;  
RUN;
```

Damit wird die im Schwellenwertmodell (siehe Abschnitt 2.2.1) unterstellte normalverteilte latente Zufallsvariable Z mit einer VC-Matrix R entsprechend der vorgegebenen räumlichen Kovarianzstruktur simuliert.

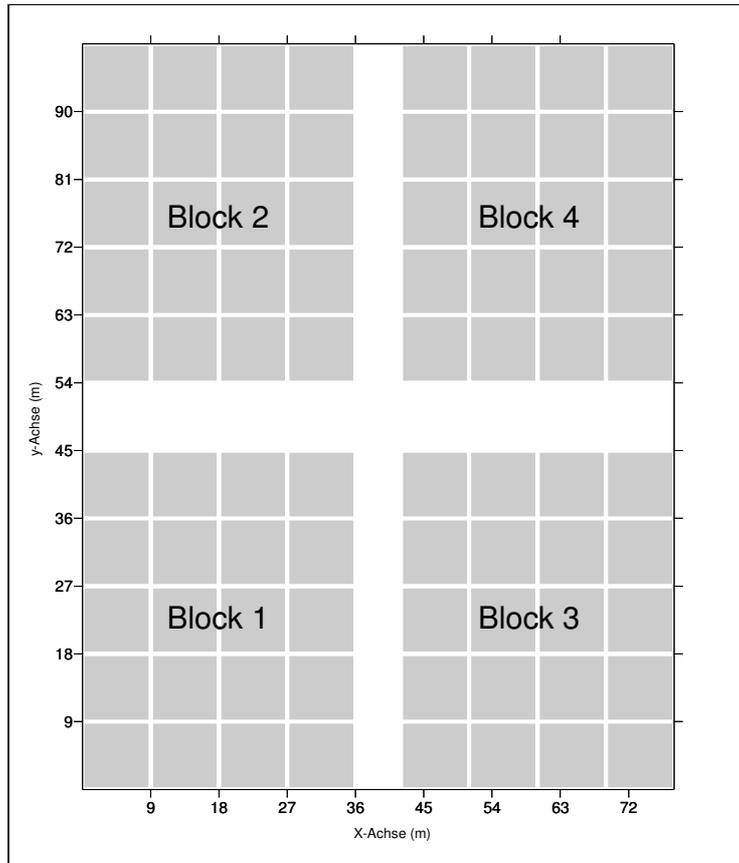


Abb. 3.1: Versuchsplan der simulierten vollständigen Blockanlage mit 4 Wiederholungen und 20 Prüfgliedern (Parzellen) pro Block

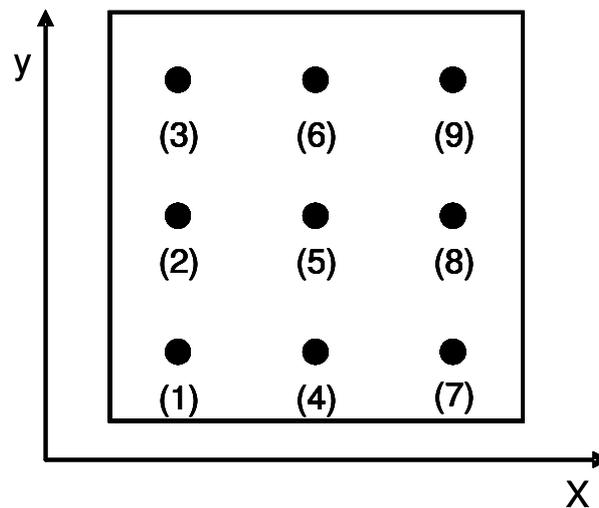


Abb. 3.2: Anordnung der Beobachtungen innerhalb der Parzelle und deren Nummerierung

3.3 Parameter

Die simulierte Versuchsanlage ist eine 1-faktorielle randomisierte vollständige Blockanlage (siehe Abschnitt 3.1), es wird jedoch in den Daten kein Blockeffekt simuliert. Wird eine räumliche Kovarianz simuliert, so entsteht dadurch ein „Pseudoblockeffekt“, da die innerhalb eines Blocks liegenden Beobachtungen stärker korreliert sind als die aus verschiedenen Blocks.

Als räumliches Modell zur Erzeugung der VC-Matrix R der latenten Zufallsvariablen Z wird das sphärische Modell (2.1-8) mit „Nugget“-Effekt gewählt und in den Parametern „Partieller Sill“ und „Range“ variiert. Das sphärische Modell wurde in Analysen von Blindversuchen von Richter und Kroschewski (2012) häufig als zutreffendes räumliches Modell erkannt und wird in dieser Arbeit beispielhaft für räumliche Kovarianzstrukturen genutzt. Zusätzlich zu den räumlichen Ansätzen wird ein Ansatz ohne räumliche Kovarianz simuliert, indem Partieller Sill und Range gleich Null gesetzt werden (Tab. 3.1). Die Abbildung 3.3 zeigt die durch die räumliche Kovarianz resultierende Korrelation zwischen Beobachtungen in Abhängigkeit von der euklidischen Distanz.

Tab. 3.1: Parameter der simulierten Kovarianzstrukturen, sphärisches Modell

Partieller Sill σ_0^2	Range ρ	Nugget-Effekt σ_R^2	Bezeichnung
0	0	1	0_0_1
1	10	1	1_10_1
1	20	1	1_20_1
1	30	1	1_30_1
2	40	1	2_40_1

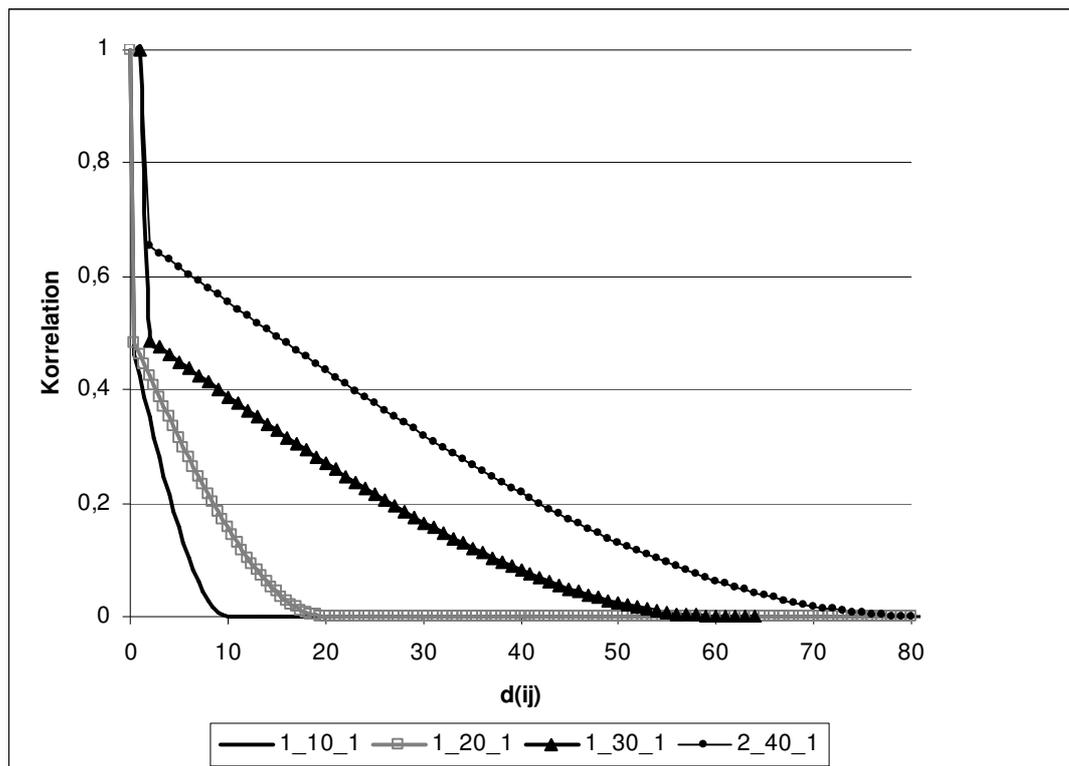


Abb. 3.3: Korrelationsfunktionen der räumlichen Simulationsansätze in Abhängigkeit von der euklidischen Distanz zweier Beobachtungen, sphärisches Modell

Die Realisationen z der simulierten latenten Variablen Z werden bei Nutzung vorgegebener kumulativer Wahrscheinlichkeiten (Tab. 3.2) in Boniturklassen überführt. Die im Schwellenwertmodell unterstellte normalverteilte latente Zufallsvariable Z hat bei Beschränkung auf Prüfglied i die Dichtefunktion

$$(3.3-1) \quad f(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(z-\mu_i)^2}{2\sigma^2}},$$

mit $\mu_i = \mu$ für alle Prüfglieder, und $\sigma^2 = \sigma_R^2 + \sigma_0^2$ aus dem räumlichen Modell.

Um die Aufteilung der latenten Variablen Z in k Klassen mit vorgegebener kumulativer Wahrscheinlichkeit zu gewährleisten, werden die Quantile dieser Wahrscheinlichkeiten bestimmt. Sei q_k die vorgegebene Wahrscheinlichkeit einer Boniturklasse k , dann ist

$$q_k = P(z \leq \theta_k) = \Phi(u_k)$$

mit
$$u_k = \frac{\theta_k - \mu_i}{\sigma} \rightarrow \theta_k = \mu_i + \sigma \cdot u_k = \mu + \sigma \cdot u_k$$

und

$$(3.3-2) \quad \phi(u_k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u_k} e^{-\frac{t^2}{2}} dt$$

Verteilungsfunktion der Standardnormalverteilung (vgl. 2.2.1).

Die gesuchten Quantile für u_k werden bei Gültigkeit der Nullhypothese durch Nutzung der SAS-Funktion „quantile“ mit einem Nichtzentralitätsparameter = 0 bestimmt.

Die θ_k stellen somit die Stützpunkte des Schwellenwertmodells dar (vgl. 2.2.1). Die Realisationen z der latenten Variablen Z können in Boniturnoten überführt werden mit $y = k$ falls $z \in [\theta_{k-1}, \theta_k]$.

Es wurden Datensätze mit $k = 9$ und $k = 4$ Boniturklassen mit homogenen Prüfgliedeffekten erzeugt, und es wurden unterschiedliche Parameter für die kumulativen Wahrscheinlichkeiten vorgegeben. Die kumulativen Wahrscheinlichkeiten für die Parametersätze finden sich in Tabelle 3.2 und die entsprechenden Histogramme für die Wahrscheinlichkeiten der Boniturnoten in Anhang 2. Weiterhin wurden für Parametersatz 1 und 5 auch Datensätze mit heterogenen Prüfgliedeffekten erzeugt. Die Prüfgliedeffekte für Prüfglied 1 und 2 wurden immer mit $\alpha_1 - \alpha_2 = 0$ simuliert, so dass die Nullhypothese für den Vergleich dieser Prüfglieder für jeden Parametersatz gültig ist. Die Varianten der simulierten Prüfgliedeffekte sind Anhang 3 zu entnehmen. Die Tabelle 3.3 zeigt die Kombination aus Parametersatz und Simulation der Prüfgliedeffekte und deren Bezeichnung.

Die Wahrscheinlichkeiten des Parametersatzes 1, mit $k = 9$ Klassen und mit homogenen Prüfgliedeffekten, sind angelehnt an die Auswertung der Bonitur des Lagers von Wintergerste zum Reifezeitpunkt im Jahr 2003, im Dauerfeldversuch der Landesanstalt für Landwirtschaft und Fischerei Mecklenburg-Vorpommern in Gülzow.

Die realisierten Simulationsvarianten pro Parametersatz ergeben sich aus

- 5 Kovarianzstrukturen (Tab. 3.1),
- der Nutzung von 1 (5), 3 ((1),(5),(9)) oder 5 ((1),(3),(5),(7),(9)) Beobachtungen je Parzelle gemäß Abbildung 3.2.

Damit ergeben sich 15 Simulationsvarianten aus der Kombination von Kovarianzstrukturen und Beobachtungen je Parzelle. Diese Varianten wurden für Parametersatz 1 vollständig und mit allen Auswertungsmethoden und Varianten unter H_0 untersucht. Ausgehend von den Ergebnissen für diesen Parametersatz, wurden für die anderen Parametersätze diejenigen Analysen ausgeschlossen, die für Parametersatz 1 unbefriedigende Ergebnisse lieferten. Für die anderen Parametersätze wurden daher nur ausgewählte Analysen der Boniturnoten mittels Schwellenwertmodell und mittels LMM untersucht.

Tab. 3.2: Parametersätze für die kumulativen Wahrscheinlichkeiten der Boniturklassen bei Gültigkeit von H_0 für den Kontrast von Prüfglied 1 zu 2

kum. Wahrscheinlichkeiten k = 9 Klassen				
Boniturklasse	Parametersatz 1	Parametersatz 2	Parametersatz 3	Parametersatz 4
1	0.03	0.09	0.39	0.25
2	0.31	0.26	0.66	0.39
3	0.53	0.32	0.72	0.46
4	0.68	0.35	0.75	0.49
5	0.80	0.48	0.8	0.52
6	0.89	0.71	0.86	0.56
7	0.95	0.95	0.94	0.63
8	0.98	0.98	0.98	0.75
9	1	1	1	1
kum. Wahrscheinlichkeiten k = 4 Klassen				
Boniturklasse	Parametersatz 5	Parametersatz 6	Parametersatz 7	Parametersatz 8
1	0.38	0.68	0.25	0.40
2	0.67	0.88	0.50	0.52
3	0.78	0.96	0.75	0.60
4	1	1	1	1

Tab. 3.3: Bezeichnungen der Kombination von Parametersätzen der kumulativen Wahrscheinlichkeiten und der Prüfgliedeffekte

Anzahl Boniturklassen	Parametersatz (gemäß Tab 3.2)	Prüfgliedeffekte (gemäß Anhang 3)			
		homogen $\alpha_1 = \dots = \alpha_{20} = 0$	heterogen		
			$\alpha_1 = \alpha_2 = 1$ $\alpha_3 = \dots = \alpha_{20} = 0.2$	$\alpha_1 = \alpha_2 = 0.5$ $\alpha_3 = \dots = \alpha_{20} = 0.2$	$\alpha_1 = \alpha_2 = 0.5$ $\alpha_3 \neq \dots \neq \alpha_{20}$
k = 9	1	1	1a	1b	1c
	2	2	-	-	-
	3	3	-	-	-
	4	4	-	-	-
k = 4	5	5	5a	5b	5c
	6	6	-	-	-
	7	7	-	-	-
	8	8	-	-	-

Die Planung der Anzahl Simulationen N je Simulationsvariante erfolgt aufgrund des exakten Konfidenzintervalls für eine Wahrscheinlichkeit nach Sachs (2009, p. 293 ff). Dieser Planungsansatz basiert auf dem Sachverhalt, dass ein wesentliches Ziel der vorliegenden Untersuchungen die Prüfung der Einhaltung eines nominalen Fehlers 1. Art für die verschiedenen Auswertungsmethoden und Modellansätze ist.

Nach Sachs (2009) sind die Grenzen des exakten zweiseitigen Konfidenzintervalls einer Wahrscheinlichkeit p wie folgt gegeben:

$$(3.3-3) \quad p_u = \frac{x}{x + (N - x + 1)F_u}$$

$$p_o = \frac{(x + 1)F_o}{N - x + (x + 1)F_o}$$

mit den F-Quantilen $F_u(1 - \alpha; 2(N - x + 1), 2x)$,

$$F_o(1 - \alpha; 2(x + 1), 2(N - x)),$$

$$x = N \cdot p$$

und mit p = zu schätzender Wahrscheinlichkeitsparameter,
 α = 0.05,
 N = Stichprobenanzahl je Simulationsvariante.

Wird der nominale statistische Fehler 1. Art mit $\alpha = 0.05$ vorgegeben, so führt ein Umfang von 10000 Simulationsläufen zu einer Intervallbreite von 17 % des nominalen Fehlers und den zweiseitige Intervallgrenzen [0.0458; 0.0543], bei einem Konfidenzniveau von $P = 0.95$. Diese Intervallbreite wird für die hier vorliegenden Untersuchungen als ausreichend angesehen.

3.4 Maßzahlen zur Auswertung der Simulationen

Für die Beurteilung der Simulation unterschiedlicher Modellansätze und Auswertungsmethoden kommen folgende Kriterien bzw. Maßzahlen zum Einsatz:

- Anteil konvergenter Simulationsläufe,
- Treffgenauigkeit bzw. Bias
- Differenz des beobachteten und mittleren geschätzten Standardfehlers der Schätzungen
- realisierte Konfidenzniveaus und Fehler 1. Art

Die Kriterien bzw. Maßzahlen sind im Folgenden erläutert.

Der Anteil konvergenter Simulationsläufe ist

$$(3.4-1) \quad \text{Konvergenz}(\%) = \frac{\sum_{i=1}^N K_i}{N},$$

mit $N = \text{Anzahl Simulationsläufe}$ und $K_i = \begin{cases} 0 & \text{falls keine Konvergenz} \\ 1 & \text{falls Konvergenz} \end{cases}$.

Die Schätzung des Bias erfolgt durch:

$$(3.4-2) \quad \text{Bias}(\hat{p}) = \frac{1}{K} \sum_{i=1}^K (\hat{p}_i) - p,$$

$K = \text{Anzahl konvergenter Simulationsläufe}$.

Der Bias erlaubt eine Einschätzung der Treffgenauigkeit.

Die Nutzung der Simulation als Untersuchungsmethode erlaubt einen Vergleich zwischen beobachtetem (STD_{beo}) und mittlerem geschätztem Standardfehler (STD_{ges}).

Der beobachtete Standardfehler eines zu schätzenden Parameters ist:

$$(3.4-3) \quad \text{STD}_{\text{beo}} = \sqrt{\frac{\sum_{i=1}^K (\hat{p}_i - \bar{\hat{p}})^2}{K-1}},$$

$K = \text{Anzahl konvergenter Simulationsläufe}$.

Und der mittlere geschätzte Standardfehler ist:

$$(3.4-4) \quad \text{STD}_{\text{ges}} = \sqrt{\frac{\sum_{i=1}^K \text{Var}(\hat{p}_i)}{K}},$$

K = Anzahl konvergenter Simulationsläufe.

Der Bias der geschätzten zur beobachteten Standardabweichung wird in Prozent der beobachteten Standardabweichung ausgedrückt.

$$(3.4-5) \quad \text{Bias \% (STD}_{\text{ges}}) = \frac{\text{STD}_{\text{ges}} - \text{STD}_{\text{beo}}}{\text{STD}_{\text{beo}}} \cdot 100$$

Für die Schätzungen der Wahrscheinlichkeiten der Boniturklassen im Schwellenwertmodell wird die mittlere Breite der geschätzten 95% Konfidenzintervalle (siehe Abschnitt 2.2.4) und deren empirische Überdeckungswahrscheinlichkeit der Konfidenzintervalle beurteilt. Die Überdeckungswahrscheinlichkeit stellt den Anteil der Simulationsläufe dar, in denen das Konfidenzintervall (siehe Abschnitt 2.2.4) den Parameter p enthält (Burton et al., 2006) und wird pro Boniturnote berechnet. Die empirische Überdeckungswahrscheinlichkeit ist:

$$(3.4-6) \quad \text{ÜDWK \%}_{p \in \text{KI}_k} = \frac{\sum_{i=1}^K p_{i_k}}{K},$$

mit k = Boniturklasse, K = Anzahl konvergenter Simulationsläufe mit Boniturnoten $\in k$

$$\text{und } p_{i_k} = \begin{cases} 0 & \text{falls Parameter nicht innerhalb des Konfidenzintervalls liegt} \\ 1 & \text{falls Parameter innerhalb des Konfidenzintervalls liegt} \end{cases}.$$

Der realisierte Fehler 1. Art ist

$$(3.4-7) \quad \alpha_{\text{realisiert}} = \frac{\sum_{i=1}^K s_i}{K},$$

$$\text{mit } K = \text{Anzahl konvergenter Simulationsläufe und } s_i = \begin{cases} 0 & \text{falls keine Signifikanz} \\ 1 & \text{falls Signifikanz} \end{cases}.$$

Die Beurteilung der empirischen Überdeckungswahrscheinlichkeit und des realisierten Fehlers 1. Art erfolgt mittels des exakten Konfidenzintervalls für eine Wahrscheinlichkeit nach Sachs (2009), wobei in (3.3-3) N durch K ersetzt wird.

4 Rechentechnische Umsetzung der Auswertungsmethoden

Die Möglichkeiten zur Berücksichtigung von räumlichen Kovarianzen sind für die unter Abschnitt 2 vorgestellten Auswertungsmethoden unterschiedlich und teilweise von der Anzahl der Beobachtungen je Parzelle abhängig. Dadurch kommen innerhalb der Auswertungsmethoden verschiedene Modellansätze bzw. Auswertungsvarianten zustande. Die Tabelle 4.1 gibt einen Überblick über alle angewendeten Modellansätze. Im Folgenden werden die Möglichkeiten und Beschränkungen der Auswertungsmethoden sowie deren rechentechnische Umsetzung dargestellt. Alle rechentechnischen Umsetzungen erfolgten mit der Statistiksoftware SAS, Version 9.2.

Die folgenden Abkürzungen werden für die unterschiedlichen Auswertungsvarianten verwendet:

- A-R: Auswertung als vollständig randomisierte Anlage,
- A-BI fix: Auswertung als randomisierte Blockanlage mit fixen (festen) Blocks,
- A-BI rand.: Auswertung als randomisierte Blockanlage mit der Annahme zufälliger Blocks,
- A-BL(fix)-Parz.: Auswertung als randomisierte Blockanlage mit zusätzlichem zufälligen Parzelleneffekt innerhalb der fixen Blocks bei Nutzung von > 1 Beobachtung je Parzelle,
- A-BL(rand.)-Parz.: Auswertung als randomisierte Blockanlage mit zusätzlichem zufälligen Parzelleneffekt innerhalb der zufälligen Blocks bei Nutzung von > 1 Beobachtung je Parzelle,
- A-cov(sph): Auswertung als vollständig randomisierte Anlage mit Modellierung einer räumlichen Kovarianzstruktur nach dem sphärischen Modell.

Simulationsansätze ohne räumliche Kovarianzstrukturen (Ansatz 0_0_1, siehe Abschnitt 3.3) werden nur mit dem Modellansatz A-R ausgewertet.

4.1 Lineares gemischtes Modell

Das lineare gemischte Modell wird für die Auswertung der simulierten Zufallszahlen (vgl. Abschnitt 3) und für die Auswertung der Parzellenmittelwerte der Boniturnoten bzw. der Einzelbonituren als kontinuierliche Variable (vgl. Abschnitt 2.5 und 2.6) eingesetzt. Die simulierten Zufallszahlen stellen die im Schwellenwertmodell unterstellte latente Variable dar (vgl. Abschnitt 2.2.1 und 3.2). Die unterschiedlichen Auswertungsansätze gemäß Abschnitt 4 können innerhalb der SAS Prozedur GLIMMIX formuliert werden (Übersicht 1). Die Prozedur GLIMMIX dient zur Lösung generalisierter linearer gemischter

Modelle (GLMM). Diese Modelle beinhalten bei Annahme von Normalverteilung und bei Verwendung der Identitätsfunktion als Linkfunktion die linearen gemischten Modelle als Spezialfall. Es sei darauf hingewiesen, dass diese Auswertungsvarianten in SAS auch mittels der Prozedur MIXED durchgeführt werden können, die zur Lösung linearer gemischter Modelle dient. Die Prozedur GLIMMIX setzt die räumliche Kovarianz über so genannte „G-seitige“ Kovarianzstrukturen um. Dieses Vorgehen basiert darauf, dass in der Prozedur GLIMMIX generell „R-seitige“ räumliche Kovarianzstrukturen durch „G-seitige“ ersetzt werden (Schabenberger, 2005). Das bedeutet, dass die räumliche Kovarianzstruktur nicht in der VC-Matrix der Resteffekte modelliert wird, sondern in der VC-Matrix der zufälligen Effekte. Deshalb muss jeder Beobachtung ein eigener zufälliger Effekt zugewiesen werden, indem die Beobachtungen durchnummeriert werden und die G-Matrix entsprechend des räumlichen Modells modifiziert wird.

Für den Modellvergleich wird als Schätzmethode die „Restricted Maximum Likelihood“ Methode (REML) gewählt, da diese unverzerrte Schätzer für die Varianzen bzw. Kovarianzen liefert (Patterson and Thompson, 1971; Mielenz et al., 2007). Die Nutzung der REML-Methode und die Einführung zufälliger Effekte, wie beispielsweise Blocks oder räumliche Kovarianzstrukturen im Auswertungsmodell, erfordert die Nutzung einer Freiheitsgradapproximation. Für die Auswertung als vollständig randomisierte Anlage oder als randomisierte Blockanlage mit fixen Blocks ist keine Freiheitsgradapproximation erforderlich, da die Auswertung nach Modell I und damit nur mit fixen Faktoren erfolgt. Der Freiheitsgrad berechnet sich dann nach der Residual-Methode, $FG = N - \text{Rang}(X)$. Für gemischte Modelle (feste und zufällige Faktoren) liefert im balancierten Fall die Freiheitsgradapproximation von Satterthwaite (Satterthwaite, 1941; Giesbrecht and Burns, 1985) und diejenige von Kenward und Roger (1997) das gleiche Ergebnis. Im unbalancierten Fall dagegen ist für gemischte Modelle die Methode von Kenward und Roger (1997) der von Satterthwaite überlegen (Spilke et al., 2005). Allerdings nur, solange in diesen Modellen die Kovarianzstruktur einfach ist, d. h., dass eine sogenannte „lineare Kovarianzstruktur“ vorliegt (Kenward and Roger, 2009). Das gilt beispielsweise für Kovarianzstrukturen, bei Nutzung der Bezeichnungen in SAS, vom Typ = VC („variance components“), CS („compound symmetry“), UN („unstructured“) oder TOEP (Toeplitz). Liegt eine nichtlineare Kovarianzstruktur vor, so wie im Auswertungsmodell mit räumlicher Kovarianz, kann die Nutzung der Approximation von Kenward und Roger (1997) zu Verzerrungen des nominalen Fehlers 1. Art führen (Hu et al., 2006). Für nichtlineare Kovarianzstrukturen sollte die modifizierte Form dieser Methode, die so genannte „Firstorder“ Variante (Kenward and Roger, 2009), genutzt werden. Besonders für Modelle mit räumlichen Kovarianzstrukturen liefert diese Freiheitsgradapproximation die beste Kontrolle der nominalen Fehler 1. Art (Spilke et al., 2010). Im Falle linearer Kovarianzstrukturen unterscheiden sich die Ergebnisse beider Methoden von Kenward und Roger nicht. Da die Nutzung der „Firstorder“ Variante für das räumliche Auswertungsmodell die besten Ergebnisse liefert und gleichzeitig für die anderen Auswertungsvarianten keine Nachteile entstehen, wird sie für alle Auswertungsvarianten eingesetzt.

Übersicht 1: Programmcodes der Auswertungsvarianten für das LMM

```
title 'Nichtbeachtung der räumlichen Kovarianz,  
      Auswertung als vollständig randomisierte Anlage (A-R)';  
PROC GLIMMIX data=daten method=RSPL;  
  /*für stetige, normalverteilte Variablen RSPL=REML*/  
  CLASS trt; /*trt=Behandlung*/  
  MODEL z=trt /CHISQ ddfm=KR(firstorder) /*z=latente Variable*/  
  dist=normal link=identity;  
  ESTIMATE "1-2" trt 1 -1; /* Kontrast von Behandlung 1 zu 2 */ RUN;  
title 'Beachtung der räumlichen Kovarianz durch fixe Blocks (A-Blfix)';  
PROC GLIMMIX data=daten method=RSPL;  
  /*für stetige, normalverteilte Variablen RSPL=REML*/  
  CLASS trt block; /*trt=Behandlung*/  
  MODEL z=trt block/CHISQ ddfm=KR(firstorder)/*z=latente Variable*/  
  dist=normal link=identity;  
  ESTIMATE "1-2" trt 1 -1; /* Kontrast von Behandlung 1 zu 2 */ RUN;
```

```
title 'Beachtung der räumlichen Kovarianz durch fixe Blocks und zu-  
      fällige Parzellen innerhalb der Blocks (A-Bl(fix)-Parz.)';  
PROC GLIMMIX data=daten method=RSPL;  
  /*für stetige, normalverteilte Variablen RSPL=REML*/  
  CLASS trt block Parz; /*trt=Behandlung*/  
  MODEL z=trt block/CHISQ ddfm=KR(firstorder)/*z=latente Variable*/  
  dist=normal link=identity;  
  RANDOM Parz(block);  
  ESTIMATE "1-2" trt 1 -1; /* Kontrast von Behandlung 1 zu 2 */ RUN;
```

```
title 'Beachtung der räumlichen Kovarianz durch zufällige Blocks  
      (A-Bl rand.)';  
PROC GLIMMIX data=daten method=RSPL;  
  /*für stetige, normalverteilte Variablen RSPL=REML*/  
  CLASS trt block; /*trt=Behandlung*/  
  MODEL z=trt /CHISQ ddfm=KR(firstorder)/*z=latente Variable*/  
  dist=normal link=identity;  
  RANDOM block;  
  ESTIMATE "1-2" trt 1 -1; /* Kontrast von Behandlung 1 zu 2 */ RUN;
```

```
title 'Beachtung der räumlichen Kovarianz durch zufällige Blocks  
      und zufällige Parzellen innerhalb der Blocks(  
      A-Bl(rand.)-Parz.)';  
PROC GLIMMIX data=daten method=RSPL;  
  /*für stetige, normalverteilte Variablen RSPL=REML*/  
  CLASS trt block Parz; /*trt=Behandlung*/  
  MODEL z=trt /CHISQ ddfm=KR(firstorder)/*z=latente Variable*/  
  dist=normal link=identity;
```

```

RANDOM block Parz(block);
ESTIMATE "1-2" trt 1 -1; /* Kontrast von Behandlung 1 zu 2 */ RUN;

```

```

title 'Modellierung der räumlichen Kovarianz, sphärisches Modell
      (A-cov(sph))';
/* Zuweisung eines zufälligen Effektes für jede Beobachtung für die
   nachfolgende Random-Anweisung in der Prozedur GLIMMIX*/
DATA daten_random;
  SET daten;
  obs=_N_;
RUN;

PROC GLIMMIX data= daten_random method=RSPL;
/*für stetige, normalverteilte Variablen RSPL=REML*/
  CLASS obs trt ;/*trt=Behandlung*/
  MODEL z=trt /CHISQ ddfm=KR(firstorder)/*z=latente Variable*/
  dist=normal link=identity;
  RANDOM obs / TYPE=SP(sph)(x y);
  ESTIMATE "1-2" trt 1 -1; /* Kontrast von Behandlung 1 zu 2 */ RUN;

```

```

title 'Auswertung der Parzellenmittelwerte der Boniturnoten bei > 1 Be-
      obachtung je Parzelle (mit einer Behandlung), mit Modellierung der
      räumlichen Kovarianz, sphärisches Modell(A-cov(sph))';

/*Berechnung der Parzellenmittelwerte der Boniturnoten und Ausgabe der
   Ergebnisse in einer Datei*/
proc means data=daten ;
CLASS lauf block trt; /*trt=Behandlung*/
VAR Bonitur;
output out=ParzMW ;
run;

/*Weiterbearbeitung der Ergebnisdatei der Mittelwertbildung, dadurch
   Auswahl der Mittelwerte pro Parzelle und Zuordnung der Parzellenmittel-
   werte zur Beobachtung 5, also der Beobachtung in der Parzellenmitte */

DATA ParzMW_1 (keep= lauf Block beob TRT _TYPE_ _STAT_ ParzMW);
  SET ParzMW;
  IF lauf =1 or lauf >1;
  IF _TYPE_ = 7;
  IF _STAT_='MEAN' ;
  RENAME Bonitur=ParzMW;
  beob=5;
RUN;

/*Sortierung der Ursprungsdatei in gleicher Reihenfolge wie die Datei
   der Mittelwerte*/
Proc sort data= daten out=daten_trait_1sort;
by lauf block trt beob; run;

/* Zusammenführung der Ursprungsdatei und der Datei der Mittelwerte und
   Auswahl der Beobachtung 5, dadurch Zuordnung der Parzellenmittelwerte

```

```

zu den in der Ursprungsdatei enthaltenen Koordinaten der Beobachtung
5*/

Data ParzMW_Koord;
merge daten_trait_1sort ParzMW_1;
by lauf block trt beob;
IF beob=5;
run;

Proc sort data =ParzMW_Koord out=ParzMW_Koord_sort;
by lauf x y; run;
/* Zuweisung eines zufälligen Effektes für jede Beobachtung für die
nachfolgende Random-Anweisung in der Prozedur GLIMMIX*/
DATA ParzMW_Koord_sort_1;
SET ParzMW_Koord_sort;
obs=_N_;
RUN;

title 'Glimmix identity';
proc glimmix data=ParzMW_Koord_sort_1;
BY lauf;
CLASS obs trt ;
MODEL ParzMW=trt /CHISQ ddfm=KR(firstorder)
dist=normal link=identity;
/* ParzMW=Parzellenmittelwerte trt=Behandlung */
RANDOM obs / TYPE=SP(sph)(x y);
ESTIMATE "1-2" trt 1 -1;
RUN;

```

Eine weitere Alternative der Auswertung der Bonituren als stetiges Merkmal mittels LMM besteht darin, zusätzlich zur Modellierung der räumlichen Kovarianz im Auswertungsmodell auch heterogene Restvarianzen für die Prüfglieder zuzulassen. Diese Auswertungsvariante ist allerdings nur für die Auswertung der Bonituren als stetiges Merkmal innerhalb der Prozedur MIXED in SAS möglich, da neben der Anweisung des RANDOM Statements zur Modellierung der räumlichen Kovarianz auch ein REPEATED Statement zur Modellierung der heterogenen Restvarianzen angewiesen werden muss. Hierdurch wird innerhalb der VC-Matrix der Resteffekte R ein heterogener *Nugget*-Effekt modelliert. Bei Beschränkung auf 2 Behandlungen wird die VC-Matrix R (siehe Abschnitt 2.1) dadurch wie folgt modifiziert:

$$R = \begin{bmatrix} \sigma_1^2 & \sigma_{12,11} & \cdots & \sigma_{1n,11} & \sigma_{21,11} & \sigma_{22,11} & \cdots & \sigma_{2n,11} \\ \sigma_{11,12} & \sigma_1^2 & & & \sigma_{21,12} & & & \sigma_{2n,12} \\ \vdots & & \ddots & \vdots & \vdots & & & \vdots \\ \sigma_{11,1n} & & & \sigma_{ij,ij} & \sigma_1^2 & \sigma_{ij,ij} & \cdots & \sigma_{2n,1n} \\ \sigma_{11,21} & \sigma_{12,22} & \cdots & \sigma_{ij,ij} & \sigma_2^2 & \sigma_{ij,ij} & \cdots & \sigma_{2n,21} \\ \sigma_{11,22} & & & & \sigma_{ij,ij} & \sigma_2^2 & & \vdots \\ \vdots & & & & \vdots & & \ddots & \vdots \\ \sigma_{11,2n} & \sigma_{12,2n} & \cdots & \sigma_{1n,2n} & \sigma_{21,2n} & \cdots & \sigma_{2n-1,1n} & \sigma_2^2 \end{bmatrix},$$

mit Behandlung $i = 1, 2$ und Beobachtung $j = 1, \dots, n$.

In der Diagonale steht ein prüfgedspezifischer *Sill* $\sigma_i^2 = \sigma_{R_i}^2 + \sigma_0^2$, zusammengesetzt aus $\sigma_0^2 =$ dem so genannten „*Partiellen Sill*“ und den prüfgedspezifischen *Nugget*-Effekten $\sigma_{R_i}^2$, mit $i = 1, \dots, 20$. Die Nichtdiagonalelemente entsprechen den räumlichen Kovarianzen $\sigma_{ij|ij}$ zwischen zwei Beobachtungen (vgl. Abschnitt 2.1). Die Formulierung dieses Auswertungsansatzes innerhalb der SAS Prozedur MIXED ist Übersicht 1a zu entnehmen.

Übersicht 1a : Programmcode der Auswertung der Bonituren als stetiges Merkmal mittels LMM mit Modellierung räumlicher Kovarianz und heterogenen Restvarianzen der Prüfglieder

```

title 'Modellierung der räumlichen Kovarianz, sphärisches Modell
      (A-cov(sph)) mit heterogenen Restvarianzen für die Prüfglieder';

/* Zuweisung eines zufälligen Effektes für jede Beobachtung für die
nachfolgende Random-Anweisung in der Prozedur MIXED*/
DATA daten_random;
  SET daten;
  obs=_N_;
RUN;

proc mixed data=daten_trait_2 ;
  BY lauf;
  CLASS trt obs;
  MODEL boni=trt / ddfm=KR(firstorder) ;

```

```
RANDOM obs /type= sp(sph)(x y) ;  
/*Anweisung heterogener Restvarianzen pro Behandlung (Prüfglied)*/  
REPEATED /group=trt ;  
ESTIMATE "1-2" trt 1 -1; RUN;
```

4.2 Schwellenwertmodell

Generell können alle für das lineare gemischte Modell vorgestellten Auswertungsvarianten (Abschnitt 4.1) auf das Schwellenwertmodell übertragen werden. Da das Schwellenwertmodell zu den generalisierten gemischten Modellen (GLMM) zählt, kann es innerhalb der Prozedur GLIMMIX abgearbeitet werden. Als Schätzmethode wird die „Restricted Pseudo-Likelihood“ Methode (RSPL) eingesetzt. Die Berechnung mittels numerischer Integration (METHOD = QUAD) ist nicht möglich, da diese Methode die erforderliche „G-seitige“ Modellierung der räumlichen Kovarianzen (siehe Abschnitt 4.1) nicht unterstützt.

Durch Linearisierung der Daten unter Verwendung der Taylorreihenentwicklung führt die Pseudo-Likelihood Technik die Maximierung der Likelihoodfunktion auf die Lösung von mehreren linearen gemischten Modellen für so genannte Pseudo-Daten zurück (Wolfinger and O'Connell, 1993). Unter der Annahme von Normalverteilung für die Pseudo-Daten erfolgt in jedem Iterationsschritt die Lösung eines linearen gemischten Modells, beispielsweise mit der REML-Methode. Das beschriebene Schätzverfahren ist doppelt iterativ. Ein lineares Modell für die Pseudo-Daten wird angepasst, und mit seinen Lösungen wird das GLMM erneut linearisiert. Diese Prozedur wird wiederholt, bis die Differenz zwischen aufeinander folgenden Schätzungen der Modellparameter, resultierend aus den Lösungen zweier aufeinander folgender linearer gemischter Modelle, hinreichend gering ist.

Die entsprechenden Modellmodifikationen bezüglich der Berücksichtigung der räumlichen Kovarianz in den verschiedenen Auswertungsvarianten und der Hypothesentest des Behandlungskontrasts beziehen sich im Schwellenwertmodell jeweils auf den linearen Prädiktor.

Bei Vorliegen mehrerer Beobachtungen pro Parzelle kann das Modell für die latente Variable um einen zufälligen Parzelleneffekt erweitert werden (vgl. Abschnitt 2.2.2). Für die Auswertung von Bonituren bietet sich ein solches Modell an, da von einer Schwankung der Multinomialwahrscheinlichkeiten von Parzelle zu Parzelle ausgegangen werden kann (Piepho, 2002), die durch die simulierte räumliche Kovarianzstruktur zusätzlich provoziert wird. Für die Simulationsansätze mit mehreren Beobachtungen pro Parzelle kommen daher als zusätzliche Auswertungsvarianten die um einen zufälligen Parzelleneffekt erweiterten Ansätze mit fixen und zufälligen Blocks zum Einsatz. Weiter werden alle Auswertungsvarianten für das Schwellenwertmodell sowohl mit der Annahme der Nor-

malverteilung (link = cumprobit) als auch der logistischen Verteilung (link = cumlogit) für die latente Variable durchgeführt. Zur rechentechnischen Umsetzung des Schwellenwertmodells müssen die Verteilung der Merkmalswerte und die Linkfunktion in der Modellanweisung angegeben werden. Nachfolgend ist beispielhaft das SAS Programm zur Auswertung von Bonituren mittels Schwellenwertmodell in vollständig randomisierter Anlage mit Modellierung der räumlichen Kovarianz nach dem sphärischen Modell aufgeführt (Übersicht 2). Die Berücksichtigung der räumlichen Kovarianzstrukturen erfolgt für das Schwellenwertmodell innerhalb des RANDOM Statements.

Die Berechnung der kumulativen Wahrscheinlichkeiten nach (2.2-3) bzw. (2.2-4) wird nicht in der Prozedur GLIMMIX, sondern in der Prozedur IML von SAS durchgeführt. Dies liegt darin begründet, dass ihre Berechnung innerhalb der Prozedur GLIMMIX mittels eines ESTIMATE Statements erfolgen müsste, welches die Vorgabe eines Koeffizientenvektors erfordert, dessen Struktur von der Anzahl beobachteter Kategorien und damit der Stützpunkte sowie der Behandlungseffekte abhängig ist. Für die Auswertung von mehreren Simulationsläufen mit einem vorgegebenen Boniturschlüssel führt dies jedoch zu ungültigen Ergebnissen, wenn in der Stichprobe nicht besetzte Kategorien auftreten. Bei Nutzung der Prozedur IML kann die Berechnung der Wahrscheinlichkeiten der Kategorien unabhängig von eventuell auftretenden leeren Kategorien durchgeführt werden (Hörtl et al., 2010).

Übersicht 2: Programmcode des Schwellenwertmodells als A-cov(sph)

```

title ' Schwellenwertmodell mit Modellierung der räumlichen Kovarianz,
sphärisches Modell(A-cov(sph));

/*      Zuweisung eines zufälligen Effektes für jede Beobachtung für die
nachfolgende Random-Anweisung in der Prozedur GLIMMIX*/

DATA daten_Bonitur; SET daten;  obs=_N_; RUN;

PROC GLIMMIX data=daten_Bonitur method=RSPL;
  CLASS obs trt; /*trt=Behandlung*/
  MODEL Bonitur=trt /CHISQ ddfm=KR(firstorder)
  dist=multinomial link=cumlogit;          /*oder cumprobit*/
  RANDOM obs / TYPE=SP(sph)(x y) ;
  ESTIMATE "1-2" trt 1 -1 / ILINK CL;
      /* Kontrast von Behandlung 1 zu 2 im lin. Prädiktor und
Rücktransformation auf die Boniturskala*/
RUN;

```

4.3 Rangverfahren

In SAS sind die in Abschnitt 2.3 vorgestellten Modelle der Rangverfahren in der Prozedur MIXED unter der Option ANOVAF für Daten nach Rangtransformation umgesetzt. Die erforderlichen SAS Programmcodes sind nachfolgend aufgelistet (Übersicht 3). Innerhalb der Option ANOVAF ist die Formulierung von Kovarianzstrukturen nur mittels REPEATED Statement möglich und auf Kovarianzstrukturen vom Typ=UN(1) („unstructured“) bzw. Typ=VC („variance components“) beschränkt. Die Nutzung des RANDOM Statements ist nicht möglich. Räumliche Kovarianzstrukturen können daher nicht modelliert werden. Die Beachtung der räumlichen Kovarianz über zufällige Blocks ist ebenfalls nicht möglich, da ihre Anweisung innerhalb des REPEATED Statements die Angabe der Kovarianzstruktur Typ= CS („compound symmetry“), zur Erzeugung einer blockdiagonalen Struktur der Kovarianzmatrix, erfordern. Die Auswertungsvarianten für die Rangverfahren reduzieren sich damit auf die Auswertung als vollständig randomisierte Anlage und als randomisierte Blockanlage mit fixen Blocks. Letzteres Modell ist ausschließlich für mehr als eine Beobachtung pro Parzelle definiert, da die Rangverfahren nach Brunner et al. (1997; 2001; 2002) die Formulierung heteroskedastischer Restvarianzen erfordern. Aus Gründen der Vollständigkeit und um die entstehenden Fehlerverzerrungen zu beurteilen, wurde diese Auswertungsvariante aber trotzdem auch für Ansätze mit nur einer Beobachtung pro Parzelle gerechnet. Für die Durchführung der ANOVAF Option in der Prozedur MIXED ist die Schätzmethode mit METHOD = MIVQUE0 (Hartley et al., 1978) vorgeschrieben.

Übersicht 3: Programmcodes der Auswertungsvarianten für das Rangverfahren

```
title 'Rangtransformation';  
PROC RANK data=daten out=daten; var boni; ranks r_boni; RUN;
```

```
title 'Nichtbeachtung der räumlichen Kovarianz,  
Auswertung als vollständig randomisierte Anlage (A-R)';  
PROC MIXED DATA=daten ANOVAF METHOD=MIVQUE0 ;  
CLASS trt ;  
MODEL r_boni = trt / CHISQ ;  
REPEATED /type=UN(1) grp=trt;  
CONTRAST "trt1 -1" trt 1 -1; RUN;
```

```
title 'Beachtung der räumlichen Kovarianz durch fixe Blocks (A-B1 fix)  
für 1 Beobachtung pro Parzelle';  
PROC MIXED DATA=daten ANOVAF METHOD=MIVQUE0 ;  
CLASS trt block;  
MODEL r_boni = trt block / CHISQ;  
REPEATED / type =UN(1) grp=trt*block ;  
CONTRAST "trt1 -1" trt 1 -1 ; RUN;  
title 'Beachtung der räumlichen Kovarianz durch fixe Blocks (A-B1 fix)'
```

```

für >1 Beobachtung pro Parzelle';
PROC MIXED DATA=daten ANOVAF METHOD=MIVQUE0 ;
CLASS trt block;
MODEL r_boni = trt block trt*block / CHISQ;
REPEATED / type =UN(1) grp=trt*block ;
CONTRAST "trt1 -1" trt 1 -1 trt*block 0.25 0.25 0.25 0.25
                                                -0.25 -0.25 -0.25 -0.25;

RUN;

```

4.4 Permutationstests

Die Auswertungsvarianten der Permutationsverfahren beschränken sich auf die Auswertung der Simulationsansätze mit nur einer Beobachtung pro Parzelle in vollständig randomisierter Anlage oder in randomisierten Blockanlagen mit fixen Blocks (siehe Abschnitt 2.4). Die Auswertung als vollständig randomisierte Anlage erfolgte in SAS mit dem Makro CRD_PERM.MAC, die als randomisierte Blockanlage mit dem Makro RIBDPERM.MAC, welches auch für die Analyse von unbalancierten Blockanlagen geeignet wäre, von Schumacher et al. (1995a ; 1995b; 1997). In beiden Auswertungsvarianten wurden die paarweisen Vergleiche mit jeweils 1000 Permutationen durchgeführt. Beide Makros und eine ausführliche Beschreibung ihrer Anwendung sind erhältlich unter:

„<https://www.uni-hohenheim.de/inst110/mitarbeiter/Permutationstests.htm>“
(Stand Oktober 2010).

Voraussetzung für die Anwendung der SAS-Makros ist die Angabe der alphanumerischen Variablen „Versuch“ sowie der drei numerischen Variablen: „Block“ bzw. „Wiederholung“, „Behandlung“ und „Wert“. Die Datei muss hierarchisch sortiert sein nach Versuch (bzw. im Falle von Simulationen nach Simulationslauf), Block und Behandlung. Die Makros bestehen aus einem Eingabeteil, in dem die Anzahl der Permutationen, das Signifikanzniveau sowie die gewünschten Test-Optionen vorgegeben werden können, sowie dem eigentlichen Programmteil. Die in dieser Arbeit gewählten Vorgaben sind der Übersicht 4 zu entnehmen. Die Formulierung von räumlichen Kovarianzstrukturen ist nicht möglich. Bei Existenz räumlicher Kovarianz kann nur versucht werden, diese über die Auswertung mittels randomisierter Blockanlage zu beachten.

Übersicht 4: Eingabeteile der Makros für die Auswertungsvarianten der Permutationstests

```
title 'Nichtbeachtung der räumlichen Kovarianz, Eingabe für den
      Permutationstest für die vollständige Randomisation (A-R),
      Makro CRD_PERM.MAC ';
```

%*crdperm*

```
(g,          /* Approximativer Test der Globalhypothese*/
 pw,         /* Approximative paarweise Vergleiche */
 daten =daten_perm,
 versuch =laeufe,
 rep =block,
 beh =trt,   /* trt = Behandlung */
 wert =boni,
 nperm=1000, /* Anzahl Permutationen*/
 alpha=0.05); /* Signifikanzniveau für paarweise Vergleiche*/
```

```
title 'Beachtung der räumlichen Kovarianz durch fixe Blocks, Eingabe
      für den Permutationstest für die randomisierte Blockanlage
```

```
(A-Blfix), Makro RIBDPERM.MAC ';
```

%*ribdperm*

```
(g,          /* Approximativer Test der Globalhypothese*/
 pw_c,       /* Approximative paarweise Vergleiche, Beobachtungen um
              Blockmedian korrigiert*/

 daten =daten_perm,
 versuch =laeufe,
 block =block,
 beh =trt,   /* trt = Behandlung */
 wert =boni,
 nperm=1000, /* Anzahl Permutationen*/
 alpha=0.05); /* Signifikanzniveau für paarweise Vergleiche*/
```

In Tabelle 4.1 ist zusammenfassend dargestellt, welche Auswertungsvarianten, in Abhängigkeit der Anzahl Beobachtungen je Parzelle und des räumlichen Simulationsansatzes, innerhalb der Auswertungsmethoden angewendet wurden.

Tab. 4.1: Übersicht über die Simulationsansätze und angewendete Auswertungsmethoden

		Auswertungsmethoden															
		SWM			LMM für Bonituren			LMM für Parzellenmittelwerte			Rangverfahren		Permutationstests				
		Auswertungsvarianten															
		Linkfunktionen: cumprobit / cumlogit															
		A-R	A-BI fix	A-BI(fix)-Parz.	A-BI rand.	A-BI(rand.)-Parz.	A-cov(sph)	A-R	A-BI fix	A-BI(fix)-Parz.	A-BI rand.	A-BI(rand.)-Parz.	A-cov(sph)	A-R	A-BI fix	A-R	A-BI fix
Beobachtungszahl / Parzelle		n = 1															
Simulierter räumlicher Kovarianzansatz	0_0_1	x				x								x		x	x
	1_10_1	x	x	x	x	x	x	x	x	x				x	(x)	x	x
	1_20_1	x	x	x	x	x	x	x	x	x				x	(x)	x	x
	1_30_1	x	x	x	x	x	x	x	x	x				x	(x)	x	x
	2_40_1	x	x	x	x	x	x	x	x	x				x	(x)	x	x
Beobachtungszahl / Parzelle		n > 1															
Simulierter räumlicher Kovarianzansatz	0_0_1	x				x								x			
	1_10_1	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
	1_20_1	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
	1_30_1	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
	2_40_1	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	

(x): Rangverfahren für 1 Beobachtung nicht definiert, nur zu Vergleichsgründen gerechnet

5 Ergebnisse

Alle in den Abschnitten 5.1 bis 5.5 dargestellten Ergebnisse für die Auswertung von Bonituren beziehen sich auf den Parametersatz 1 für die Erzeugung von Boniturnoten mit $k = 9$ Klassen (vgl. Abschnitt 3.3). In Abschnitt 5.7 werden ausgewählte Ergebnisse der Auswertung von Bonituren für die alternativen Parametersätze besprochen.

5.1 Konvergenzverhalten

Grundvoraussetzung für den Einsatz einer bestimmten Auswertungsmethode bzw. -variante ist, dass die innerhalb der Schätzmethodik eingesetzten Iterationsverfahren konvergieren. Das Konvergenzverhalten ist primär abhängig von der eingesetzten Auswertungsmethode. Daneben spielen weiterhin die bei der Simulation verwendeten Parameter, das verwendete Auswertungsmodell und der Stichprobenumfang eine Rolle. Das Konvergenzverhalten ist im Rahmen der vorliegenden Untersuchungen nur für die Auswertung der Bonituren mittels Schwellenwertmodell von Bedeutung. Hier kommt es, bei Existenz räumlicher Kovarianzstrukturen, für bestimmte Auswertungsvarianten zu teilweise starken Verlusten an konvergenten Simulationsläufen. Alle anderen Auswertungsmethoden für die Bonituren und für die Auswertung der latenten Variablen mittels LMM weisen keine Konvergenzprobleme auf.

Die Tabelle 5.1 zeigt den Anteil konvergenter Simulationsläufe für die unterschiedlichen Varianten des Schwellenwertmodells zur Auswertung von Bonituren, Parametersatz 1 (vgl. Abschnitt 4.2). Wird mittels Schwellenwertmodell und den Varianten A-R oder A-BI fix ausgewertet, existieren unabhängig von der gewählten Linkfunktion keine Konvergenzprobleme. Dagegen zeigen alle Auswertungsvarianten des Schwellenwertmodells, die einen oder mehrere zufällige Faktoren beinhalten (A-BI rand., A-BI(fix)-Parz. und BI(rand.)-Parz.), leichte Konvergenzverluste, die sich tendenziell mit der Anzahl der zufälligen Faktoren im Modell und mit steigender Beobachtungsanzahl verstärken. Der Verlust ist jedoch gegenüber den räumlichen Modellen (A-cov(sph)), die teilweise starke Konvergenzverluste zeigen, sehr gering. Insgesamt zeigen die Auswertungen mittels Schwellenwertmodell mit der Linkfunktion „cumprobit“ die stärksten Konvergenzverluste. Mit ansteigender Stärke der simulierten räumlichen Kovarianz fällt die Konvergenz für das Schwellenwertmodell mit dieser Linkfunktion unabhängig von der Beobachtungsanzahl stark ab.

Der Verlust an konvergenten Läufen ist für die Auswertungen mit der Linkfunktion „cumlogit“ deutlich geringer, besonders ausgeprägt zeigt sich dies für die Ansätze mit nur einer Beobachtung pro Parzelle. Der Verlust an konvergenten Läufen mit zunehmender räumlicher Kovarianz zeigt für die „cumlogit“ Auswertungsvarianten je nach Beobachtungsanzahl unterschiedliches Verhalten. Für $n = 1$ Beobachtung steigt die Konvergenz mit der räumlichen Kovarianz an. Für $n = 3$ Beobachtungen fällt die Konvergenz bis zum Ansatz 1_30_1 ab, steigt aber für den Ansatz 2_40_1 wieder deutlich an. Für $n = 5$ Beobachtungen sinkt die Konvergenz mit der simulierten Kovarianz.

Mit steigender Beobachtungszahl und bei gleichem räumlichen Ansatz steigt die Konvergenz für die „cumprobit“ Auswertungen tendenziell an, für die „cumlogit“ Auswertungen sinkt sie beim Übergang von $n = 1$ auf $n = 3$ Beobachtungen um bis zur Hälfte ab, steigt aber für $n = 5$ Beobachtungen wieder deutlich an. Der Anteil konvergenter Simulationsläufe für die räumlichen Auswertungsvarianten mit „cumlogit“ Linkfunktion liegt gegenüber denjenigen mit „cumprobit“ Linkfunktion jedoch immer auf deutlich höherem Niveau.

5.2 Schätzung der Varianzkomponenten

Die Schätzung der räumlichen Varianzkomponenten mit der Auswertungsvariante A-cov(sph) ist in dieser Arbeit nicht das Hauptziel, sondern nur Hilfsmittel für die Schätzung der Prüfglieddiffereenz zur Ermittlung der realisierten Fehler 1. Art und im Falle der Auswertung der Boniturnoten mittels SWM auch für die Schätzung der kumulativen Wahrscheinlichkeiten. Der Anhang 4 zeigt den Bias in % vom Parameter für „partieller Sill“ und „Range“ dieser Auswertungen. Da für das Schwellenwertmodell wegen der Standardisierung auf $\sigma_e^2 = 1$ in der Variante A-cov(sph) kein „Nugget-Effekt“ ausgegeben wird, ist dieser auch für die latente Variable nicht angegeben.

Die Schätzung der Varianzkomponenten ist sowohl für die Auswertung der latenten Variablen als auch für diejenige der Boniturnoten extrem verzerrt. Für Beobachtungszahlen > 1 zeigen sich zwar deutlich kleinere Verzerrungen, sie liegen aber trotzdem in einem unakzeptablen Bereich. Damit entsprechen diese Ergebnisse Literaturberichten, wonach es im Allgemeinen schwierig ist, für die Parameter partieller Sill und Range erwartungstreue Schätzungen zu gewinnen, wenn zur Schätzung die REML-Methode genutzt wird. Irvine et al. (2007) fanden in ihren Simulationsuntersuchungen extreme positive Verzerrungen für die Schätzung des Range mittels REML für das exponentielle räumliche Modell, wenn das Verhältnis von Nugget-Effekt zu Sill relativ groß (> 0.5) war. Diese Überschätzung trat umso deutlicher auf, je größer der simulierte Range war. Irvine et al. (2007) führen diese Überschätzungen auf extreme Ausreißer zurück. Die in dieser Arbeit erhaltenen Ergebnisse sind auf extreme Ausreißer und auf die numerische Umsetzung zurückzuführen, deren Einflüsse im Zusammenhang mit der Diskussion zur Einhaltung der Fehler 1. Art in Abschnitt 6.3 besprochen werden.

Eine Interpretation der Schätzung der Varianzkomponenten kann aufgrund der extremen Verzerrungen nicht erfolgen.

Tab. 5.1: Anteil konvergenter Simulationsläufe der Auswertung der Bonituren (Parametersatz 1) mittels Schwellenwertmodell und deren Varianten (10000 Simulationsläufe)

		Auswertungsmethoden											
		Schwellenwertmodell						Schwellenwertmodell					
		Auswertungsvarianten											
		Linkfunktion: „cumprobit“						Linkfunktion: „cumlogit“					
		A-R	A-BI fix	A-BI(fix)- Parz.	A-BI rand.	A-BI(rand.)- Parz.	A-cov(sph)	A-R	A-BI fix	A-BI(fix)- Parz.	A-BI rand.	A-BI(rand.)- Parz.	A-cov(sph)
Beobachtungszahl / Parzelle		n = 1											
Simulierter räumlicher Kovarianzansatz	0_0_1	1.00						1.00					
	1_10_1	1.00	1.00	0.98	0.68	1.00	1.00	0.98	0.91				
	1_20_1	1.00	1.00	0.98	0.41	1.00	1.00	0.98	0.95				
	1_30_1	1.00	1.00	0.98	0.27	1.00	1.00	0.98	0.97				
	2_40_1	1.00	1.00	0.98	0.16	1.00	1.00	0.98	0.98				
Beobachtungszahl / Parzelle		n = 3											
Simulierter räumlicher Kovarianzansatz	0_0_1	1.00						1.00					
	1_10_1	1.00	1.00	0.75	0.91	0.84	0.70	1.00	1.00	0.99	0.97	0.92	0.60
	1_20_1	1.00	1.00	0.97	0.92	0.94	0.38	1.00	1.00	1.00	0.98	0.77	0.49
	1_30_1	1.00	1.00	0.97	0.96	0.96	0.31	1.00	1.00	1.00	0.99	0.98	0.51
	2_40_1	1.00	1.00	0.97	0.98	0.97	0.19	1.00	1.00	1.00	0.99	0.99	0.62
Beobachtungszahl / Parzelle		n = 5											
Simulierter räumlicher Kovarianzansatz	0_0_1	1.00						1.00					
	1_10_1	1.00	1.00	0.94	0.89	0.72	0.78	1.00	1.00	1.00	0.81	0.93	0.78 ^{*)}
	1_20_1	1.00	1.00	0.91	0.90	0.84	0.79	1.00	1.00	0.99	0.91	0.95	0.88
	1_30_1	1.00	1.00	0.93	0.94	0.90	0.57	1.00	1.00	0.99	0.96	0.97	0.74
	2_40_1	1.00	1.00	0.98	0.97	0.97	0.27	1.00	1.00	0.99	0.98	0.99	0.64

^{*)}: Ergebnisse von 1000 Läufen, da Simulationen mit 10000 Läufen aufgrund sehr langer Laufzeiten nicht realisierbar waren

5.3 Kontraste der fixen Effekte

Die Kontraste der fixen Effekte, in dieser Arbeit die Prüfglieddifferenz von PG 1 zu 2, werden über die Erwartungstreue der geschätzten Prüfglieddifferenz, ihrer beobachteten Standardabweichung und des Bias der mittleren geschätzten Standardabweichung zur beobachteten Standardabweichung beurteilt (vgl. Abschnitt 3.4).

Im Falle der Schätzung der Prüfglieddifferenz von PG 1 zu PG 2 unter der Nullhypothese ist der wahre Parameter $p = 0$, wodurch der Bias der mittleren geschätzten Prüfglieddifferenz entspricht. Die Erwartungstreue einer Auswertungsmethode bzw. -variante ist dann gegeben, wenn das 95%ige Konfidenzintervall der geschätzten mittleren Prüfglieddifferenz den wahren Parameter überdeckt (Konfidenzintervalle der geschätzten Prüfglieddifferenzen siehe Anhang 5).

5.3.1 Auswertung der latenten Variablen (LMM)

Die Tabelle 5.2 zeigt den Bias der Schätzung des Prüfgliedkontrasts von PG 1 zu 2, die beobachtete Standardabweichung und den Bias der geschätzten Standardabweichung für die Auswertung der latenten Variablen mittels LMM, für Simulationsansätze mit räumlichen Kovarianzstrukturen. Der Bias des Prüfgliedkontrasts ist für alle Auswertungsvarianten unabhängig von der Anzahl Beobachtungen je Parzelle gering und die geschätzte Prüfglieddifferenz ist für alle Auswertungsvarianten erwartungstreu.

Die beobachtete Standardabweichung sinkt mit steigender Beobachtungsdichte für alle Analysemodelle. Werden, bei Existenz räumlicher Kovarianz in den Daten, Auswertungsvarianten ohne Modellierung einer räumlichen Kovarianz (A-R, A-BL fix/ A-BL random, A-BL(fix)-Parz. / A-BL(rand.)-Parz.) genutzt, steigt die beobachtete Standardabweichung mit der Stärke der simulierten räumlichen Korrelation an. Die Nichtbeachtung der simulierten räumlichen Kovarianz in diesen Auswertungsvarianten führt im Allgemeinen zur Erhöhung der beobachteten Standardabweichung gegenüber der Auswertungsvariante A-cov(sph) (Ausnahmen zeigen sich für den Simulationsansatz 1_10_1, für $n = 1$ sind die Werte fast gleich und für $n = 3$ ist die beobachtete Standardabweichung der Variante A-cov(sph) höher als diejenige der anderen Varianten). Diese Erhöhung nimmt mit steigender Beobachtungsdichte und der Stärke der simulierten räumlichen Kovarianz zu. Die beobachteten Standardabweichungen der nicht räumlichen Auswertungsvarianten sind für $n = 3$ innerhalb der unterschiedlichen räumlichen Simulationsansätze gleich. Für $n = 5$ zeigen die Varianten A-R und A-BL fix und random, sowie die Ansätze mit zufälligem Parzelleneffekt jeweils denselben Wert, wobei die beobachtete Standardabweichung letzterer Ansätze gegenüber der Variante A-cov (sph) noch stärker erhöht ist.

Der Bias der geschätzten Standardabweichung für $n = 1$ ist für die nicht räumlichen Auswertungsvarianten sehr gering, für die Varianten A-cov(sph) zeigt sich mit zunehmender Stärke der räumlichen Korrelation eine deutliche Überschätzung der Standardabweichung, die für den Simulationsansatz 2_40_1 einen extremen Wert annimmt. Diese Verzerrung der geschätzten Standardabweichung entsteht durch vereinzelte Simulationsläufe mit extremen Ausreißern für die Schätzung der Standardabweichung. Werden diese aus der Berechnung des Bias der geschätzten Standardabweichung

chung ausgeschlossen, so zeigen sich nur geringe Abweichungen zwischen beobachteter und geschätzter Standardabweichung. Für $n = 3$ und $n = 5$ zeigen die Auswertungsvarianten A-BI (fix)-Parz. und A-BI(rand.)-Parz. fast keine und die räumliche Variante A-cov(sph) nur geringe Abweichungen von geschätzter und beobachteter Standardabweichung. Für die anderen Auswertungsvarianten dagegen zeigt sich starker negativer Bias, der mit der Stärke der simulierten räumlichen Korrelation ansteigt.

5.3.2 Auswertung der Bonituren mittels Schwellenwertmodell (SWM)

Aufgrund der starken Konvergenzverluste der Auswertungsvariante A-cov(sph) bei Nutzung der Linkfunktion „cumprobit“ (vgl. Abschnitt 5.1) werden im Folgenden nur die Ergebnisse für die Auswertungen der Bonituren bei Nutzung der Linkfunktion „cumlogit“ dargestellt. Die Tabelle 5.3 zeigt den Bias der Schätzung des Prüfgliedkontrasts von PG 1 zu 2, die beobachtete Standardabweichung und den Bias der geschätzten Standardabweichung im linearen Prädiktor für die Auswertung der Boniturnoten (Parametersatz 1) mittels SWM (Linkfunktion „cumlogit“), für Simulationsansätze mit räumlichen Kovarianzstrukturen.

Die Schätzung der Prüfglieddifferenz im linearen Prädiktor ist erwartungstreu, der Bias des Prüfgliedkontrasts ist für alle Simulationsansätze und Auswertungsvarianten klein. Die beobachtete Standardabweichung sinkt mit steigender Beobachtungszahl und steigt mit zunehmender Stärke der simulierten räumlichen Korrelation. Für $n = 1$ ist die beobachtete Standardabweichung der räumlichen Auswertungsvariante A-cov(sph) gegenüber den anderen Varianten deutlich erhöht, dieses Verhalten verliert sich mit steigender Beobachtungszahl und zunehmender räumlicher Korrelation im Simulationsansatz. Die Varianten A-R und A-BI fix zeigen für $n = 1$ einen sehr großen positiven Bias der geschätzten Standardabweichung, der wiederum durch einzelne Simulationsläufe mit extremen Ausreißern für die Schätzung entstanden ist. Die Variante A-BI random zeigt im Gegensatz dazu deutlich geringeren, aber negativen, Bias und die räumliche Auswertungsvariante nur sehr kleinen. Für die Auswertung der Bonituren mittels SWM mit $n = 3$ und 5 Beobachtungen pro Parzelle zeigt der Bias der geschätzten Standardabweichung ein ähnliches Bild wie für die Auswertung der latenten Variablen. Die Auswertungsvarianten A-cov(sph) sowie diejenigen mit zufälligem Parzelleneffekt zeigen die geringste Abweichung von beobachteter und geschätzter Standardabweichung, die anderen Varianten unterschätzen die beobachtete Standardabweichung.

5.3.3 Auswertung der Parzellenmittelwerte der Bonituren (LMM)

Die Tab 5.4 zeigt den Bias der Schätzung des Prüfgliedkontrasts von PG 1 zu 2, die beobachtete Standardabweichung und den Bias der geschätzten Standardabweichung der Auswertung der Parzellenmittelwerte der Boniturnoten (Parametersatz 1) als stetiges Merkmal mittels LMM, für Simulationsansätze mit räumlichen Kovarianzstrukturen. Alle Auswertungsvarianten zeigen für die räumlichen Simulationsansätze Erwartungstreue für den Prüfgliedkontrast. Auch bei der Auswertung der Parzellenmittelwerte der Boniturnoten sinkt die beobachtete Standardabweichung mit der Anzahl der

Beobachtungen je Parzelle und mit der Stärke der simulierten räumlichen Korrelation. Die Auswertungsvariante A-cov(sph) zeigt im Allgemeinen gegenüber den anderen Auswertungsvarianten niedrigere beobachtete Standardabweichungen. Allerdings ist der Bias der geschätzten Standardabweichung dieser Variante für die Simulationsansätze mit stärkerer räumlicher Korrelation extrem verzerrt.

5.3.4 Auswertung der Bonituren als stetiges Merkmal (LMM)

Die Tabelle 5.5 zeigt den Bias der Schätzung des Prüfgliedkontrasts von PG 1 zu 2, die beobachtete Standardabweichung und den Bias der geschätzten Standardabweichung der Auswertung der Boniturnoten (Parametersatz 1) als stetiges Merkmal mittels LMM, für Simulationsansätze mit räumlichen Kovarianzstrukturen.

Für die Auswertungsvariante A-cov(sph) ist die Erwartungstreue der geschätzten Prüfglieddiffereenz nur für Ansätze mit $n = 3$ Beobachtungen und den Simulationsansatz 1_10_1 gegeben. Die Verzerrung für die Ansätze mit stärkerer Kovarianz ist allerdings nur sehr klein (Vergleich Anhang 5, Tabelle A5-4). Die beobachteten und geschätzten Standardabweichungen zeigen für die stetige Auswertung der Boniturnoten ansonsten das gleiche Verhalten wie für die Auswertung der latenten Variablen. Die STD_{beo} sinkt mit der Beobachtungsanzahl und tendenziell auch mit zunehmender Stärke der simulierten räumlichen Kovarianz. Die beobachtete Standardabweichung ist für die Auswertungsvariante A-cov(sph) niedriger als für die anderen Varianten. Die Auswertungsvarianten der Ansätze mit $n = 1$ zeigen außer für die Variante A-R gute Übereinstimmungen von beobachteter und geschätzter Standardabweichung. Für die Ansätze mit $n = 3$ oder 5 Beobachtungen ist diese Übereinstimmung für die Auswertungsvarianten mit zufälligem Parzelleneffekt ebenfalls gegeben, die Variante A-cov(sph) zeigt für die Ansätze 1_10_1, 1_20_1 und 1_30_1 leichte Unterschätzungen, die anderen Auswertungsvarianten zeigen für alle Ansätze stärkere Unterschätzungen.

Vergleicht man die Ergebnisse der Auswertung der Bonituren als stetiges Merkmal mit denen der Auswertung der Parzellenmittelwerte der Bonituren für die Simulationsansätze mit $n = 3$ und 5 Beobachtungen, so zeigen sich für alle Auswertungsvarianten außer der räumlichen Variante A-cov(sph) identische Ergebnisse für den Prüfgliedkontrast und die beobachtete Standardabweichung. Nur der Bias von geschätzter zu beobachteter Standardabweichung unterscheidet sich und ist für die Auswertung der Parzellenmittelwerte der Bonituren kleiner. Auch die Ergebnisse für den Prüfgliedkontrast und die beobachtete Standardabweichung für die Auswertungsvariante A-cov(sph) zeigen nur geringe Unterschiede.

Tab. 5.2: Bias und beobachtete Standardabweichung des Prüfgliedkontrasts PG 1 zu 2 sowie Bias der geschätzten Standardabweichung in Prozent der beobachteten Standardabweichung, Auswertung der latenten Variablen mittels LMM für Simulationsansätze mit räumlichen Kovarianzstrukturen , 10000 Simulationsläufe (fett: Erwartungstreue, d. h. 0.95 %-Konfidenzintervall des geschätzten Prüfgliedkontrasts enthält den wahren Parameter PG1 - PG2 = 0)

Simulierter räumlicher Kovarianzansatz		1_10_1			1_20_1			1_30_1			2_40_1		
Auswertung		Bias PG-Kontrast	STD beo	Bias % STD ges	Bias PG-Kontrast	STD beo	Bias % STD ges	Bias PG-Kontrast	STD beo	Bias % STD ges	Bias PG-Kontrast	STD beo	Bias % STD ges
n = 1	A-R	-0.0010	1.0012	-16.98	-0.0013	0.9837	1.59	-0.0009	0.9562	4.38	0.0006	1.1015	10.49
	A-cov(sph)	-0.0014	1.0026	-0.74	-0.0014	0.9714	10.21	-0.0033	0.9146	18.86	-0.0029	0.9678	212.52
	A-BI fix	-0.0010	1.0012	-0.23	-0.0013	0.9837	-0.56	-0.0009	0.9562	-0.71	0.0006	1.1015	-1.05
	A-BI(fix)-Parz.	---	---	---	---	---	---	---	---	---	---	---	---
	A-BI rand.	-0.0010	1.0012	-1.00	-0.0013	0.9837	-0.96	-0.0009	0.9562	-0.93	0.0006	1.1015	-1.14
	A-BI(rand.)-Parz.	---	---	---	---	---	---	---	---	---	---	---	---
n = 3	A-R	-0.0058	0.6647	-14.22	-0.0057	0.6948	-19.17	-0.0052	0.6896	-19.16	-0.0037	0.8521	-21.52
	A-cov(sph)	-0.0081	0.6464	-10.10	-0.0091	0.6173	-4.60	-0.0090	0.5601	-1.12	-0.0083	0.5947	0.81
	A-BI fix	-0.0058	0.6447	-11.96	-0.0057	0.6948	-21.00	-0.0052	0.6896	-22.91	-0.0037	0.8521	-29.20
	A-BI(fix)-Parz.	-0.0058	0.6447	0.81	-0.0057	0.6948	0.55	-0.0052	0.6896	0.23	-0.0037	0.8521	-0.35
	A-BI rand.	-0.0058	0.6447	-12.07	-0.0057	0.6948	-21.03	-0.0052	0.6896	-22.91	-0.0037	0.8521	-29.20
	A-BI(rand.)-Parz.	-0.0058	0.6447	0.19	-0.0057	0.6948	0.29	-0.0052	0.6896	0.12	-0.0067	0.8521	-0.40
n = 5	A-R	-0.0018	0.5069	-35.55	-0.0023	0.5941	-46.00	-0.0021	0.5932	-46.36	-0.0007	0.7702	-50.68
	A-cov(sph)	-0.0052	0.4885	-7.72	-0.0050	0.4154	-2.89	-0.0050	0.3743	-0.83	-0.0046	0.4025	0.15
	A-BI fix	-0.0018	0.5069	-35.92	-0.0023	0.5941	-47.26	-0.0021	0.5932	-48.85	-0.0007	0.7702	-55.47
	A-BI(fix)-Parz.	-0.0032	0.5368	0.78	-0.0037	0.6198	0.58	-0.0035	0.6218	0.37	-0.0021	0.7912	-0.27
	A-BI rand.	-0.0018	0.5069	-35.94	-0.0023	0.5941	-47.26	-0.0021	0.5932	-48.84	-0.0007	0.7702	-55.82
	A-BI(rand.)-Parz.	-0.0032	0.5368	0.28	-0.0037	0.6198	0.37	-0.0035	0.6218	0.29	-0.0021	0.7912	-0.30

Tab. 5.3: Bias und beobachtete Standardabweichung des Prüfgliedkontrasts PG 1 zu 2 im linearen Prädiktor sowie Bias der geschätzten Standardabweichung in Prozent der beobachteten Standardabweichung, Auswertung der Boniturnoten (Parametersatz 1) mittels SWM (Linkfunktion „cumlogit“) für Simulationsansätze mit räumlichen Kovarianzstrukturen , 10000 Simulationsläufe (fett: Erwartungstreue, d. h. 0.95 %-Konfidenzintervall des geschätzten Prüfgliedkontrasts enthält den wahren Parameter PG1 - PG2 = 0)

Simulierter räumlicher Kovarianzansatz		1_10_1			1_20_1			1_30_1			2_40_1		
Auswertung		Bias PG-Kontrast	STD beo	Bias % STD ges	Bias PG-Kontrast	STD beo	Bias % STD ges	Bias PG-Kontrast	STD beo	Bias % STD ges	Bias PG-Kontrast	STD beo	Bias % STD ges
n = 1	A-R	0.0137	1.6541	812.84	0.0034	1.7026	1026.75	0.0128	1.6363	1013.40	0.0022	1.5324	782.99
	A-cov(sph)	0.0003	2.4982	-7.79	0.0004	2.9933	-5.33	0.0235	3.1509	-5.50	0.0226	3.0378	-3.96
	A-BI fix	0.0122	1.7080	750.40	0.0037	1.7872	923.20	0.0154	1.7643	895.31	0.0019	1.7432	604.68
	A-BI(fix)-Parz.	---	---	---	---	---	---	---	---	---	---	---	---
	A-BI rand.	0.0080	1.5696	-17.62	0.0075	1.5752	-17.78	0.0107	1.5770	-17.67	0.0022	1.5951	-17.94
	A-BI(rand.)-Parz.	---	---	---	---	---	---	---	---	---	---	---	---
n = 3	A-R	0.0113	0.8701	-15.57	0.0120	0.9500	-22.44	0.0103	0.9472	-22.15	0.0060	0.9763	-24.18
	A-cov(sph)	0.0177	0.9503	-11.02	-0.0054	1.2434	-1.86	0.0030	1.1625	0.58	0.0242	1.2187	3.44
	A-BI fix	0.0113	0.8826	-16.58	0.0125	0.9830	-24.71	0.0121	1.0043	-26.06	0.0086	1.0894	-31.16
	A-BI(fix)-Parz.	0.0101	0.9470	-0.38	0.0140	1.1095	0.50	0.0120	1.1524	0.50	0.0102	1.3432	-0.05
	A-BI rand.	0.0093	0.8748	-16.68	0.0133	0.9744	-24.91	0.0130	0.9929	-26.07	0.0089	1.0770	-31.24
	A-BI(rand.)-Parz.	0.0093	0.9261	0.17	0.0148	1.0973	0.45	0.0107	1.1404	0.42	0.0109	1.3254	-0.01
n = 5	A-R	0.0071	0.7013	-19.42	0.0082	0.8192	-30.75	0.0043	0.8282	-31.43	0.0026	0.8811	-35.26
	A-cov(sph)	-0.0038*	0.8071	-8.41	0.0179	0.9041	-0.71	0.0116	0.8436	0.60	0.0197	0.8882	3.20
	A-BI fix	0.0067	0.7094	-20.21	0.0080	0.8445	-32.60	0.0053	0.8737	-34.61	0.0039	0.9768	-40.92
	A-BI(fix)-Parz.	0.0064	0.7433	0.43	0.0095	0.9443	0.56	0.0044	0.9944	0.51	0.0044	1.1958	-0.11
	A-BI rand.	0.0096	0.7067	-20.39	0.0079	0.8416	-32.87	0.0051	0.8700	-34.83	0.0035	0.9683	-40.94
	A-BI(rand.)-Parz.	0.0057	0.7377	0.33	0.0038	0.9376	0.42	0.0058	0.9877	0.51	0.0041	1.1861	0.01

*): Ergebnisse von 1000 Läufen, da Simulationen mit 10000 Läufen aufgrund von Laufzeiten über 5 Monate abgebrochen wurde

Tab. 5.4: Bias und beobachtete Standardabweichung des Prüfgliedkontrasts PG 1 zu 2 sowie Bias der geschätzten Standardabweichung in Prozent der beobachteten Standardabweichung, Auswertung der Parzellenmittelwerte der Boniturnoten (Parametersatz 1) mittels LMM für Simulationsansätze mit räumlichen Kovarianzstrukturen , 10000 Simulationsläufe (fett: Erwartungstreue, d. h. 0.95 %-Konfidenzintervall des geschätzten Prüfgliedkontrasts enthält den wahren Parameter PG1 - PG2 = 0)

Simulierter räumlicher Kovarianzansatz		1_10_1			1_20_1			1_30_1			2_40_1		
Auswertung		Bias PG-Kontrast	STD beo	Bias % STD ges	Bias PG-Kontrast	STD beo	Bias % STD ges	Bias PG-Kontrast	STD beo	Bias % STD ges	Bias PG-Kontrast	STD beo	Bias % STD ges
n = 3	A-R	-0.0089	0.8566	1.02	-0.0104	0.9233	3.94	-0.0099	0.9150	8.73	-0.0068	0.9230	16.86
	A-cov(sph)	-0.0079	0.8591	0.93	-0.0107	0.8800	67.88	-0.0118	0.7945	121.27	-0.0100	0.6956	2253.14
	A-BI fix	-0.0089	0.8566	0.38	-0.0104	0.9233	-0.06	-0.0099	0.9150	0.00	-0.0068	0.9230	-0.22
	A-BI(fix)-Parz.	---	---	---	---	---	---	---	---	---	---	---	---
	A-BI rand.	-0.0089	0.8566	-0.25	-0.0104	0.9233	-0.33	-0.0099	0.9150	-0.13	-0.0068	0.9230	-0.27
	A-BI(rand.)-Parz.	---	---	---	---	---	---	---	---	---	---	---	---
n = 5	A-R	-0.0055	0.7105	1.63	-0.0057	0.8168	5.27	-0.0041	0.8187	10.99	-0.0026	0.8504	19.75
	A-cov(sph)	-0.0058	0.7099	21.01	-0.0064	0.7461	288.00	-0.0084	0.6485	2.64E+06	-0.0071	0.5689	1.69E+04
	A-BI fix	-0.0055	0.7105	0.37	-0.0057	0.8168	0.12	-0.0041	0.8187	0.18	-0.0026	0.8504	-0.04
	A-BI(fix)-Parz.	---	---	---	---	---	---	---	---	---	---	---	---
	A-BI rand.	-0.0055	0.7105	-0.16	-0.0057	0.8168	-0.10	-0.0041	0.8187	0.08	-0.0026	0.8504	-0.09
	A-BI(rand.)-Parz.	---	---	---	---	---	---	---	---	---	---	---	---

Tab. 5.5: Bias und beobachtete Standardabweichung des Prüfgliedkontrasts PG 1 zu 2 sowie Bias der geschätzten Standardabweichung in Prozent der beobachteten Standardabweichung, Auswertung der Boniturnoten (Parametersatz 1) als stetiges Merkmal mittels LMM für Simulationsansätze mit räumlichen Kovarianzstrukturen , 10000 Simulationsläufe (fett: Erwartungstreue, d. h. 0.95 %-Konfidenzintervall des geschätzten Prüfgliedkontrasts enthält den wahren Parameter PG1 - PG2 = 0)

Simulierter räumlicher Kovarianzansatz		1_10_1			1_20_1			1_30_1			2_40_1		
Auswertung		Bias PG-Kontrast	STD beo	Bias % STD ges	Bias PG-Kontrast	STD beo	Bias % STD ges	Bias PG-Kontrast	STD beo	Bias % STD ges	Bias PG-Kontrast	STD beo	Bias % STD ges
n = 1	A-R	-0.0071	1.3333	0.28	-0.0034	1.3123	1.87	-0.0082	1.2797	4.36	-0.0012	1.2093	9.96
	A-cov(sph)	-0.0063	1.3339	-0.06	-0.0061	1.2997	-0.38	-0.0085	1.2347	-0.40	-0.0044	1.0828	0.80
	A-BI fix	-0.0071	1.3333	0.21	-0.0034	1.3123	-0.10	-0.0082	1.2797	-0.27	-0.0012	1.2093	-0.54
	A-BI(fix)-Parz.	---	---	---	---	---	---	---	---	---	---	---	---
	A-BI rand.	-0.0071	1.3333	-0.57	-0.0034	1.3123	-0.53	-0.0082	1.2797	-0.51	-0.0012	1.2093	-0.65
	A-BI(rand.)-Parz.	---	---	---	---	---	---	---	---	---	---	---	---
n = 3	A-R	-0.0089	0.8566	-10.87	-0.0104	0.9233	-18.43	-0.0099	0.9150	-18.26	-0.0068	0.9230	-20.55
	A-cov(sph)	-0.0110	0.8593	-9.72	-0.0169	0.8358	-5.81	-0.0162	0.7637	-1.94	-0.0139	0.6702	0.31
	A-BI fix	-0.0089	0.8566	-11.24	-0.0104	0.9233	-20.11	-0.0099	0.9150	-21.69	-0.0068	0.9230	-27.63
	A-BI(fix)-Parz.	-0.0089	0.8566	0.56	-0.0104	0.9233	-0.05	-0.0099	0.9150	0.00	-0.0068	0.9230	-0.22
	A-BI rand.	-0.0089	0.8566	-11.35	-0.0104	0.9233	-20.14	-0.0099	0.9150	-21.70	-0.0068	0.9230	-27.63
	A-BI(rand.)-Parz.	-0.0089	0.8566	-0.05	-0.0104	0.9233	-0.32	-0.0099	0.9150	-0.13	-0.0068	0.9230	-0.27
n = 5	A-R	-0.0055	0.7105	-16.80	-0.0057	0.8168	-28.82	-0.0041	0.8187	-29.56	-0.0026	0.8504	-33.69
	A-cov(sph)	-0.0102	0.7053	-8.67	-0.0105	0.6600	-3.70	-0.0090	0.6056	-1.04	-0.0090	0.5306	0.46
	A-BI fix	-0.0055	0.7105	-17.19	-0.0057	0.8168	-30.29	-0.0041	0.8187	-32.48	-0.0026	0.8504	-39.50
	A-BI(fix)-Parz.	-0.0055	0.7105	0.41	-0.0057	0.8168	0.12	-0.0041	0.8187	0.18	-0.0026	0.8504	-0.04
	A-BI rand.	-0.0055	0.7105	-17.24	-0.0057	0.8168	-30.30	-0.0041	0.8187	-32.49	-0.0026	0.8504	-39.50
	A-BI(rand.)-Parz.	-0.0055	0.7105	-0.11	-0.0057	0.8168	-0.10	-0.0041	0.8187	0.08	-0.0026	0.8504	-0.09

5.4 Schätzung kumulativer Wahrscheinlichkeiten

Die Schätzung der kumulativen Wahrscheinlichkeiten der Boniturnoten ist ein direktes Ergebnis des Schwellenwertmodells. Aufgrund des sehr niedrigen Anteils konvergenter Simulationsläufe für das Schwellenwertmodell bei Wahl der Linkfunktion „cumprobit“ (siehe Abschnitt 5.1) werden die Ergebnisse für die Schätzung der kumulativen Wahrscheinlichkeiten nur für die Auswertung mittels Schwellenwertmodell mit der Linkfunktion „cumlogit“ dargestellt.

Im Folgenden werden die Schätzungen der kumulativen Wahrscheinlichkeiten der Boniturnoten (Parametersatz 1) anhand der Ergebnisse für das Prüfglied 1 für die Simulationsansätze mit und ohne räumliche Kovarianz und für alle Auswertungsvarianten des SWM besprochen. Beurteilt werden die Erwartungstreue der Schätzung anhand des Bias und die Varianz der Schätzungen anhand der mittleren Breite der 95%-Konfidenzintervalle aller Schätzungen und deren empirischen Überdeckungswahrscheinlichkeiten. Die Konfidenzintervalle wurden gemäß Abschnitt 2.2.4 unterschiedlich berechnet. Direkt für die Wahrscheinlichkeiten (KI) sowie durch Berechnung der Konfidenzintervalle im linearen Prädiktor und Rücktransformation auf die Boniturskala (KI_{\logit}). Da die letzte Boniturnote (9) immer die kumulative Wahrscheinlichkeit von 1 besitzt, wird sie in den nachfolgenden Ausführungen nicht betrachtet. Die Erwartungstreue der Schätzungen ist dann gegeben, wenn der Betrag des Bias der mittleren Schätzung kleiner als die halbe Breite des 95 %-Konfidenzintervall für den Mittelwert der Schätzungen ist. Zur Beurteilung der empirischen Überdeckungswahrscheinlichkeit kann das exakte zweiseitige Konfidenzintervall der Wahrscheinlichkeit p (siehe Abschnitt 3.4) genutzt werden. Da nicht alle Boniturnoten in jedem Lauf auftreten, werden die jeweiligen Intervalle abhängig gemacht von der Anzahl der Läufe, in denen die jeweiligen Boniturnoten realisiert wurden.

In Anhang 6 ist in den Tabellen A6-1 bis A6-6 der Bias für die Schätzungen der kumulativen Wahrscheinlichkeiten von Prüfglied 1 für alle Simulationsansätze und Auswertungsvarianten dargestellt. Erwartungstreue Schätzungen werden, unabhängig vom räumlichen bzw. nichträumlichen Simulationsansatz, nur mit der Auswertungsvariante A-R erhalten (Tabelle A6-1). Der Bias dieser Auswertungsvariante ist für alle Simulationsansätze sehr niedrig und zeigt keine Abhängigkeit von der Stärke der simulierten räumlichen Korrelation. Alle anderen Auswertungsvarianten für Ansätze mit räumlicher Korrelation zeigen nur vereinzelt erwartungstreue Schätzungen, am häufigsten in den Auswertungen der Simulationsansätze mit schwacher räumlicher Korrelation (1_10_1). Die Verzerrungen der Schätzungen sind jedoch insgesamt sehr klein. Mit zunehmender Stärke der simulierten räumlichen Korrelation steigt der Bias. Der größte Bias wird für die Auswertungsvariante A-cov(sph) erhalten (Tabelle A6-2). Die erwartungstreuen Schätzungen dieser Variante für den Ansatz 1_10_1 mit $n = 5$ Beobachtungen pro Parzelle ist darauf zurückzuführen, dass dieser Ansatz aufgrund sehr langer Laufzeiten nicht für 10000 Simulationen, sondern nur für 1000 Simulationen gerechnet wurde. Dies führt zu breiteren Konfidenzintervallen für die Schätzungen der Wahrscheinlichkeiten, so dass der wahre Parameter von diesen Intervallen trotz eines relativ großen Bias der Schätzungen überdeckt wird. Der Vergleich zwischen den Schätzungen von konditionalen und marginalen Wahrscheinlichkeiten für die Auswertungsvarianten A-BI random, A-BI(rand.)-Parz. und

A-BI(fix)-Parz. (Tabellen A6-4a und -4b, A6-5a und -5b, A6-6a und -6b) liefert keine Unterschiede bezüglich der Erwartungstreue der Schätzungen.

Die Abbildung 5.1 zeigt die Parameter (Parametersatz 1, Tab 3.2) und die Konfidenzintervalle sowie die empirischen Überdeckungswahrscheinlichkeiten der Schätzungen der kumulativen Wahrscheinlichkeiten der Boniturnoten (Parametersatz 1) bei Auswertung der Simulationsansätze ohne räumliche Kovarianz mittels SWM (Auswertungsvariante A-R) für Ansätze mit $n = 1, 3$ und 5 Beobachtungen. Die Berechnungsvariante KI zeigt für $n = 3$ und 5 Beobachtungen negative untere Intervallgrenzen für die Boniturote 1. Dies ist bei der Berechnungsvariante KI_{logit} aufgrund der Berechnung der Konfidenzintervalle im linearen Prädiktor und Rücktransformation auf die Boniturskala nicht der Fall, da die Intervallgrenzen auf Werte zwischen 0 und 1 beschränkt sind (vgl. Abschnitt 2.2.4). Durch diese Berechnung der KI_{logit} ist weiterhin begründet, dass die obere und untere halbe Intervallbreite der KI_{logit} nicht gleich ist. Die Konfidenzintervalle sind für die Boniturnoten 2 bis 6 aufgrund der Parameterwahl in der Simulation deutlich breiter (vgl. Abschnitt 3.3). Die Berechnungsvariante KI zeigt, unabhängig von der Beobachtungszahl je Parzelle, für die mittleren Boniturnoten ($2-4$) etwas breitere Intervalle als die Variante KI_{logit} , für die anderen Boniturnoten ist es umgekehrt. Mit steigender Anzahl Beobachtungen je Parzelle verringert sich die Intervallbreite, und die empirische Überdeckungswahrscheinlichkeit erhöht sich. Für die Berechnungsvariante KI zeigt sich die Erhöhung der Überdeckungswahrscheinlichkeit sehr deutlich, für die KI_{logit} nur schwach. Die Abweichung der empirischen Überdeckungswahrscheinlichkeiten vom Konfidenzniveau ist für die KI sehr hoch, für die KI_{logit} deutlich geringer. Die Einhaltung des Konfidenzniveaus ist jedoch auch für die KI_{logit} nur für die Noten 8 des Ansatzes mit $n=3$ und für die Noten $1, 4$ und 8 für $n = 5$ gegeben.

Die Auswertung der Simulationsansätze ohne räumliche Kovarianz mit Auswertungsvariante A-R stellt den einfachsten Modellansatz dar. Da in diesen Fällen die simulierten Daten mit dem „wahren“ Modell ausgewertet werden, können ihre Ergebnisse als Vergleichsbasis für die Auswertungsvariante A-cov(sph) der Simulationsansätze mit räumlicher Kovarianz dienen, welche ebenfalls das „wahre“ Auswertungsmodell darstellt. Da die unteren und oberen Intervallbreiten der KI_{logit} nicht gleich sind, wurde, aus Gründen der Vergleichbarkeit mit den Intervallbreiten der Berechnungsvariante KI, die halbe Breite der KI_{logit} als Durchschnitt von unterer und oberer Breite der Konfidenzintervalle berechnet (Abbildungen 5.2 und A7-1 bis A7-3 (Anhang 7)).

Die Auswertungsvariante A-cov(sph) der Simulationsansätze mit räumlicher Kovarianz zeigt im Gegensatz zur Auswertungsvariante A-R der Simulationsansätze ohne räumliche Kovarianz (Abb. 5.2) extrem verbreiterte Konfidenzintervalle für die KI und deutlich verbreiterte für die KI_{logit} . Auch das Verhalten der Intervallbreiten bei Zunahme der Beobachtungen je Parzelle unterscheidet sich. Im Gegensatz zur Auswertung der nichträumlichen Simulationsansätze mit der Variante A-R zeigen die Auswertungen der räumlichen Ansätze mit der Variante A-cov(sph) für die KI je nach Stärke der simulierten Kovarianz für $n = 3$ Beobachtungen eine leichte Verkleinerung der Intervallbreite, für $n = 5$ jedoch wieder eine Vergrößerung, die sogar die Breite der Intervalle für die Ansätze mit $n = 1$ übersteigt. Ausnahme ist der Ansatz 1_{20_1} , der gegenteiliges Verhalten zeigt. Die Breiten der KI_{logit} zeigen deutlich schmalere Intervallbreiten als die KI. Mit steigender Beobachtungsanzahl kommt es bei ihnen unabhängig von der Stärke der simulierten Kovarianz

ebenfalls zu einem Anstieg der Intervallbreiten, der in seiner Höhe zwar nicht mit denjenigen der Berechnungsvariante KI zu vergleichen, aber dennoch deutlich ist. Auch hier zeigt der Ansatz 1_20_1 ein von den anderen Ansätzen abgesetztes Verhalten.

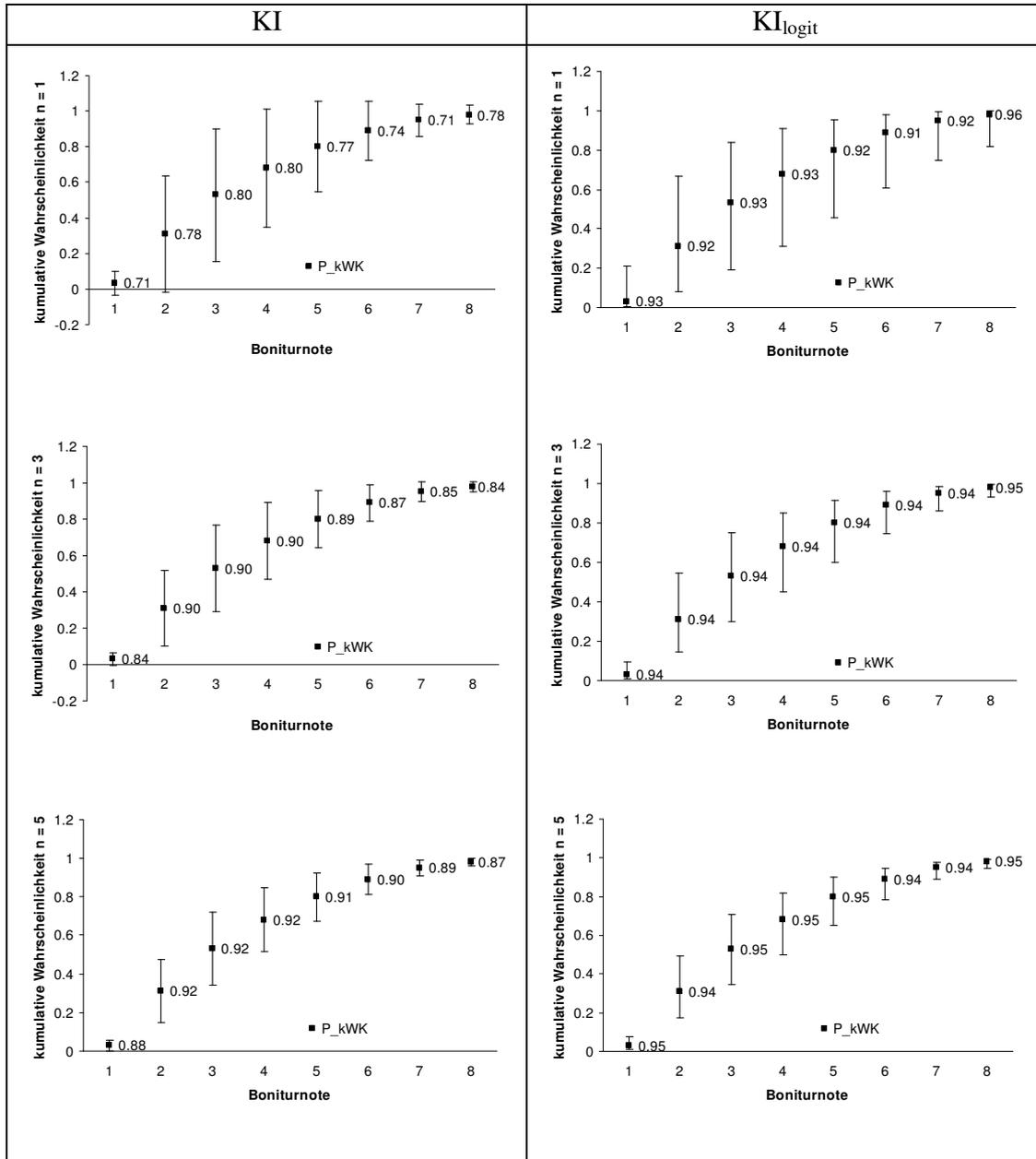


Abb. 5.1: Parameter (P_kWK) gemäß Parametersatz 1 in Tab. 3.2 und Konfidenzintervalle sowie empirische Überdeckungswahrscheinlichkeiten der Schätzungen der kumulativen Wahrscheinlichkeiten der Boniturnoten (Parametersatz 1) bei Auswertung der Simulationsansätze ohne räumliche Kovarianz mittels SWM (Linkfunktion „cumlogit“) Auswertungsvariante A-R, für Ansätze mit n = 1, 3 und 5 Beobachtungen. (KI: Konfidenzintervalle berechnet für die Wahrscheinlichkeiten, KI_logit: Berechnung der Konfidenzintervalle im linearen Prädiktor und Rücktransformation auf die Boniturskala)

Der Vergleich der empirischen Überdeckungswahrscheinlichkeiten für $n = 1$ der Auswertungsvariante A-cov(sph) der Simulationsansätze mit räumlicher Kovarianz mit der Auswertungsvariante A-R der Simulationsansätze ohne räumliche Kovarianz (Abb. 5.3) zeigt für die Berechnungsvariante KI deutliche niedrigere Werte für die Auswertung der räumlichen Ansätze und eine starke Abhängigkeit der Überdeckungswahrscheinlichkeit von der jeweiligen Boniturnote, beides verstärkt sich mit der Intensität der simulierten räumlichen Kovarianz. Mit steigender Beobachtungsanzahl schwächt sich dieses Verhalten für alle Ansätze außer demjenigen mit der stärksten simulierten räumlichen Kovarianz (2_40_1) deutlich ab. Für die Berechnungsvariante KI_{logit} sind die Unterschiede zwischen der Auswertung des nichträumlichen Ansatzes und der räumlichen Ansätze nicht so gravierend. Hier ist für die Ansätze mit $n = 1$ Beobachtung je Parzelle ebenfalls eine Abhängigkeit der empirischen Überdeckungswahrscheinlichkeiten der räumlichen Ansätze von der Boniturnote zu erkennen, in den Ansätzen mit $n = 3$ oder 5 ist dies jedoch nicht mehr sehr deutlich. Die Berechnungsvariante KI_{logit} führt für die Auswertungsvariante A-cov(sph) für $n = 3$ und 5 zu einer Überschreitung des Konfidenzniveaus.

Die Darstellung der Intervallbreiten und der empirischen Überdeckungswahrscheinlichkeiten beider Berechnungsvarianten für die Konfidenzintervalle aller möglichen Auswertungsvarianten der Simulationsansätze mit räumlicher Kovarianz, gemäß Abschnitt 4.2, sowie die Vergleiche dieser Werte für die bedingten und marginalen Wahrscheinlichkeiten (vgl. Abschnitt 2.2.2) finden sich in Anhang 7, Abbildung (A7-1 bis A7-6). Die Abbildung A7-1 zeigt die halbe Breite der Konfidenzintervalle KI. Die Ansätze mit räumlicher Kovarianz und $n = 1$ Beobachtungen je Parzelle zeigen für die Auswertungsvarianten A-R, A-BI fix und random sowie für die Blockansätze mit zusätzlichem zufälligen Parzelleneffekt - unabhängig von der Stärke der simulierten Kovarianz - eine fast identische Intervallbreite der KI für die bedingten Wahrscheinlichkeiten. Die Auswertungsvariante A-cov(sph) zeigt stark erhöhte Intervallbreiten, die mit der Stärke der simulierten Kovarianz zunehmen. Steigt die Stärke der simulierten Kovarianz für die Ansätze mit $n = 3$ oder 5 Beobachtungen je Parzelle, bleibt die Intervallbreite der Varianten A-R und A-BL fix annähernd gleich, die Varianten mit zufälligem Parzelleneffekt zeigen breitere Intervalle, die räumliche Variante A-cov(sph) zeigt unregelmäßiges Verhalten. Bei Erhöhung der Beobachtungszahl pro Parzelle sinken die Intervallbreiten für alle Auswertungsvarianten außer der räumlichen Variante A-cov(sph). Die Ansätze A-BL fix und random zeigen dabei eine deutlichere Reduzierung der Intervallbreiten als die Ansätze mit zusätzlichem, zufälligem Parzelleneffekt. Die zu diesen bedingten KI gehörenden empirischen Überdeckungswahrscheinlichkeiten sind in Abbildung A7-4 des Anhangs 7 dargestellt. Alle Auswertungsvarianten ohne Modellierung einer räumlichen Kovarianz zeigen Überdeckungswahrscheinlichkeiten die deutlich unterhalb des Konfidenzniveaus liegen. Für Simulationsansätze mit geringer Stärke der simulierten Kovarianz (1_10_1) finden sich fast identische Werte für die unterschiedlichen Auswertungsvarianten. Bei Erhöhung der Beobachtungszahl je Parzelle sinkt für diese Varianten die empirische Überdeckungswahrscheinlichkeit des KI für alle räumlichen Ansätze, allerdings differenzieren sich die Auswertungsvarianten mit Blocks und zufälligem Parzelleneffekt von den Varianten A-R und A-BI fix und random. Die räumliche Auswertungsvariante A-cov(sph) zeigt für den Ansatz 1_10_1 bei $n = 1$ und für die Ansätze mit stärkerer räumlicher Kovarianz für alle Beobachtungszahlen die oben beschriebene starke Abhängigkeit der empirischen Ü-

berdeckungswahrscheinlichkeit von der Boniturnote. Mit zunehmender Stärke der räumlichen Kovarianz und mit zunehmender Beobachtungsanzahl zeigen auch die nicht-räumlichen Auswertungsvarianten diese Abhängigkeit, und die Unterschiede der Ergebnisse zwischen den Auswertungsvarianten verstärken sich.

Werden für die räumlichen Simulationsansätze und die Auswertungsvarianten A-BI random und A-BL(fix)-Parz. und A-BL(rand.)-Parz. die marginalen KI für die kumulativen Wahrscheinlichkeiten berechnet (vgl. Abschnitt 2.2.2), so sind die Intervallbreiten schmäler als für die KI der bedingten Wahrscheinlichkeiten und verringern sich bei mehr als $n = 1$ Beobachtungen und zunehmender Stärke der simulierten Kovarianz (Abb. A7-2). Für die Variante A-BI können auch für $n = 1$ marginale Wahrscheinlichkeiten und deren KI berechnet werden, die Intervallbreiten im Vergleich zu denen der bedingten Wahrscheinlichkeiten unterscheiden sich jedoch kaum. Die empirischen Überdeckungswahrscheinlichkeiten der marginalen KI (Abb. A7-5) zeigen insgesamt weniger Reaktion auf die Stärke der räumlichen Kovarianz und die Beobachtungsanzahl je Parzelle. Aber auch die Überdeckungswahrscheinlichkeiten der Konfidenzintervalle der marginalen Wahrscheinlichkeiten schöpfen das Konfidenzniveau nicht aus. Mit zunehmender Stärke der räumlichen Kovarianz verringern sich, unabhängig von der Beobachtungsanzahl, die Überdeckungswahrscheinlichkeiten, und die Ergebnisse der unterschiedlichen Auswertungsvarianten differenzieren sich stärker. Die Variante A-BL random zeigt grundsätzlich niedrigere Werte als die Auswertungsvarianten mit zusätzlichem, zufälligem Parzelleneffekt. Diese Auswertungsvarianten zeigen kaum Reaktion auf die Erhöhung der Beobachtungsanzahl, die Variante A-BL random nur geringe. Im Gegensatz zu den Werten für die bedingten Wahrscheinlichkeiten zeigt sich für alle Varianten kaum Abhängigkeit der empirischen Überdeckungswahrscheinlichkeit von den Boniturnoten.

Die halben Breiten und die Überdeckungswahrscheinlichkeiten der Berechnungsvariante KI_{\logit} für alle Auswertungsvarianten der Simulationsansätze mit räumlicher Kovarianz sind im Anhang in Abb. A7-3 und A7-6 dargestellt. Die Breite der Konfidenzintervalle der Variante A-cov(sph) ist gegenüber den nicht-räumlichen Auswertungsvarianten stark erhöht, jedoch nicht in so extremem Ausmaß wie für die Berechnungsvariante KI der bedingten Wahrscheinlichkeiten. Die Breiten der KI_{\logit} der nicht-räumlichen Auswertungsvarianten sind niedriger als die der bedingten KI. Varianten für die marginale KI berechnet werden können zeigen zu den KI_{\logit} sehr ähnliche Intervallbreiten. Das Verhalten der Intervallbreiten der KI_{\logit} zeigt für die nicht-räumlichen Auswertungsvarianten dasselbe Verhalten wie die Intervallbreiten der KI. Mit zunehmender Stärke der simulierten räumlichen Kovarianz und mit zunehmender Beobachtungsanzahl pro Parzelle zeigt sich eine zunehmende Differenzierung in den Intervallbreiten. Steigt die Beobachtungsanzahl, verringern sich die Intervallbreiten für alle Auswertungsvarianten, außer der räumlichen, am stärksten für die Varianten A-R und A-BL fix. Die räumliche Variante A-cov(sph) dagegen zeigt mit zunehmender Beobachtungsanzahl steigende Intervallbreiten. Die zu den KI_{\logit} gehörigen empirischen Überdeckungswahrscheinlichkeiten (Abb. A7-6) der nicht-räumlichen Auswertungsvarianten zeigen ein ähnliches Muster wie diejenigen der bedingten KI, jedoch auf insgesamt höherem Niveau. Die Überdeckungswahrscheinlichkeiten erreichen jedoch auch für die Berechnungsvariante KI_{\logit} nur für einzelne Boniturnoten und einzelne Auswertungsvarianten das 95 %-Konfidenzniveau. Mit Zunahme der

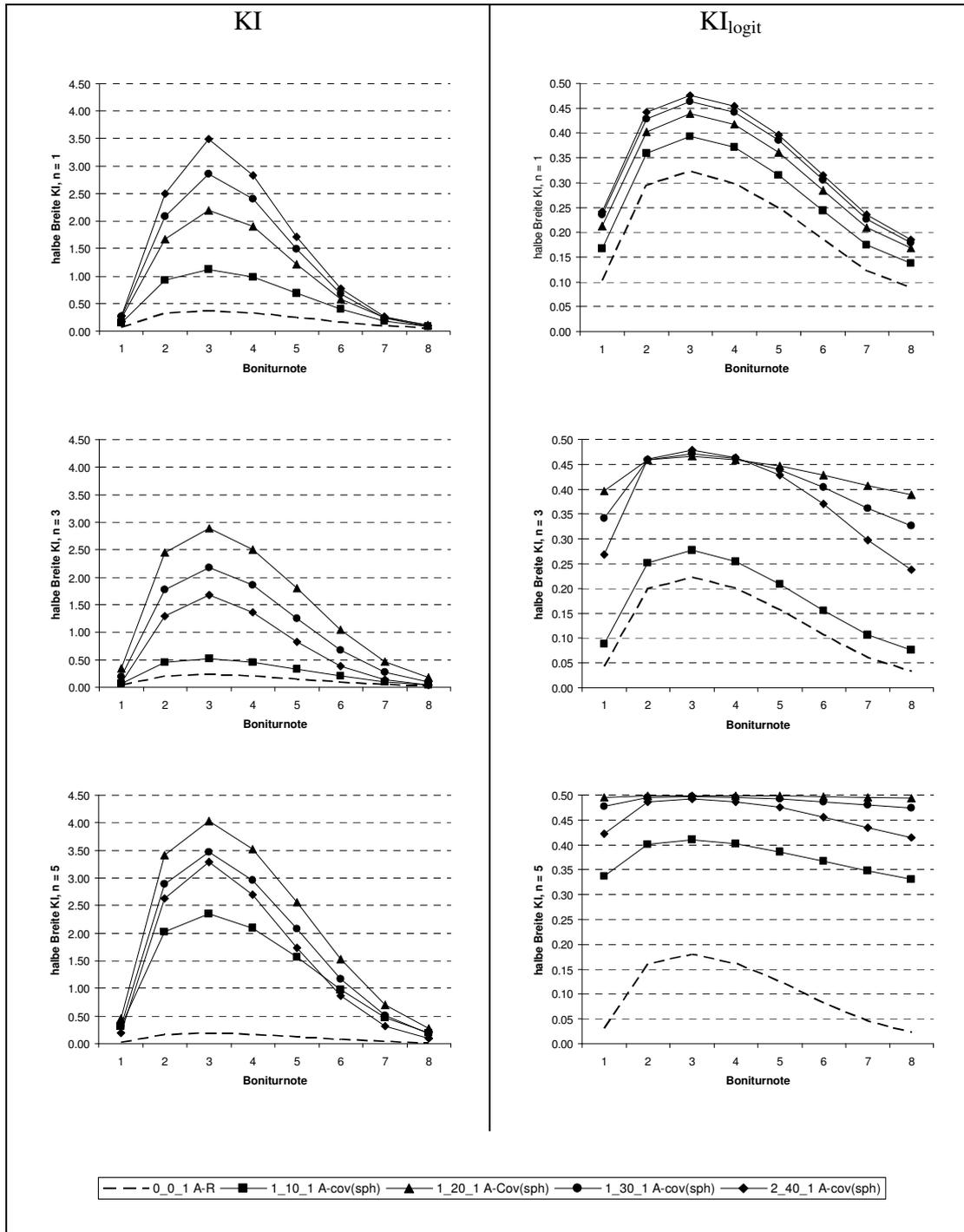


Abb. 5.2: Halbe Breiten der Konfidenzintervalle der kumulativen Wahrscheinlichkeiten für die Boniturnoten (Parametersatz 1) der Auswertung der Simulationsansätze ohne räumliche Kovarianz mittels SWM, Auswertungsvariante A-R und Simulationsansätze mit räumlicher Kovarianz mittels SWM, Auswertungsvariante A-cov(sph), für Ansätze mit $n = 1, 3$ und 5 Beobachtungen.

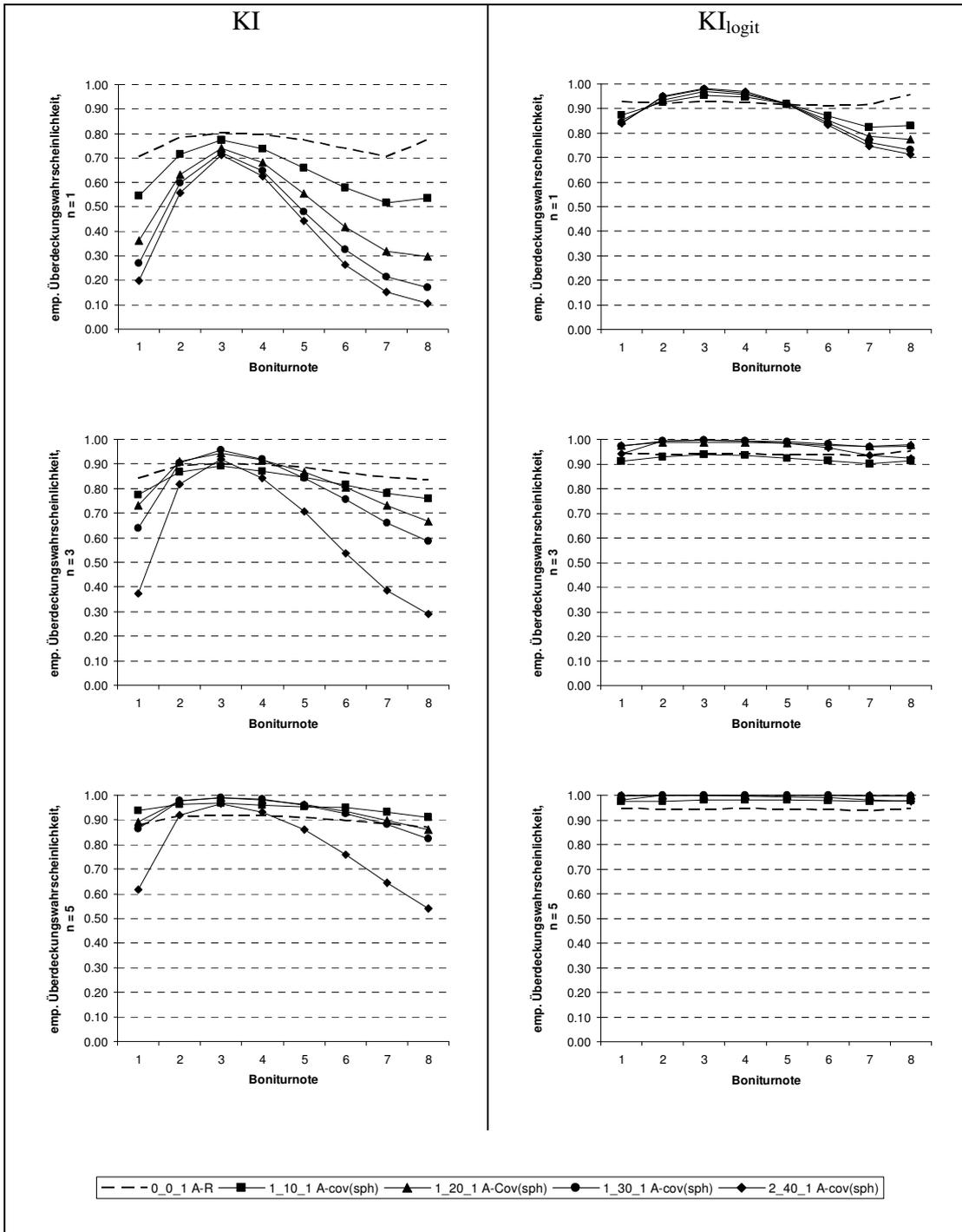


Abb. 5.3: Empirische Überdeckungswahrscheinlichkeiten der Konfidenzintervalle der kumulativen Wahrscheinlichkeiten für die Boniturnoten (Parametersatz 1) der Auswertung der Simulationsansätze ohne räumliche Kovarianz mittels SWM, Auswertungsvariante A-R und Simulationsansätze mit räumlicher Kovarianz mittels SWM, Auswertungsvariante A-cov(sph), für Ansätze mit $n = 1, 3$ und 5 Beobachtungen.

räumlichen Kovarianz und der Beobachtungsanzahl sinkt die Überdeckungswahrscheinlichkeit, die Differenzierung zwischen den Varianten nimmt zu, und die Abhängigkeit von der Boniturnote steigt. Die Variante A-cov(sph) liegt für $n = 1$ auf deutlich höherem Niveau gegenüber den Werten der bedingten KI, und die Abhängigkeit von der Boniturnote ist weniger deutlich. Mit zunehmender räumlicher Kovarianz und Beobachtungsanzahl überschätzt diese Variante jedoch die 95%-ige Überdeckung.

Im Allgemeinen zeigen die Überdeckungswahrscheinlichkeiten der Auswertungen der Boniturnoten mittels Schwellenwertmodell mit der Linkfunktion „cumlogit“ unabhängig vom Simulationsansatz keine Einhaltung des Konfidenzniveaus.

5.5 Einhaltung des Fehlers 1. Art

Ein wichtiges Kriterium für die Beurteilung einer Auswertungsmethode zum Vergleich fester Effekte ist die Einhaltung der Fehler 1. Art (auch statistisches Risiko 1. Art) bei der Hypothesenprüfung. Entsprechend kommt diesem Kriterium auch in den vorliegenden Untersuchungen eine große Bedeutung zu. Der nominale Fehler 1. Art wurde für alle durchgeführten Tests mit $\alpha = 0.05$ vorgegeben. Bewertet wird der realisierte Fehler 1. Art der Auswertung von Simulationen unter der Nullhypothese anhand des exakten Konfidenzintervalls für eine Wahrscheinlichkeit (vgl. Abschnitt 3.4). Für einen realisierten Fehler 1. Art von 0.05 ist das Intervall bei einem Konfidenzniveau von 95 % für 10 000 Simulationsläufe: [0.0458; 0.0543]. Für die Auswertung der latenten Variablen mittels LMM, der Auswertung der Bonituren mittels Schwellenwertmodell und der Auswertung der Bonituren oder ihrer Parzellenmittelwerte mittels LMM (vgl. 2.1, 2.2, 2.5 und 2.6) werden die realisierten Fehler 1. Art des t-Tests für den Kontrast von Prüfglied 1 zu 2 sowie diejenigen des F- und χ^2 - Tests der Globalhypothese bewertet. Die vergleichbaren Teststatistiken für die Rangverfahren sind für den Prüfgliedkontrast die ANOVA-Typ Statistik und für die Globalhypothese die ANOVA-Typ und die Wald-Typ Statistik (vgl. 2.3.3). Die Permutationsverfahren liefern eine mit dem t-Test vergleichbare Statistik für den Prüfgliedkontrast und eine mit dem F-Test vergleichbare für die Globalhypothese (vgl. 2.4).

5.5.1 Auswertung der Simulationsansätze ohne räumliche Kovarianzstrukturen

Die Tabelle 5.6 zeigt die realisierten Fehler 1. Art für die Auswertung von Simulationsansätzen ohne räumliche Kovarianzstrukturen für die latente Variable und für die Boniturnoten (Parametersatz 1). Die Simulationsansätze ohne räumliche Kovarianzstrukturen wurden für alle Auswertungsmethoden nach dem vollständig randomisierten Modell (A-R) simuliert (vgl. 3.2) und ausgewertet. Der Vergleich der realisierten Fehler 1. Art der Auswertung der latenten Variablen und der zugehörigen Boniturnoten gibt Auskunft darüber, wie gut die gewählte Auswertungsmethode die multinomial verteilten Boniturnoten abbilden kann.

Für die Auswertung der latenten Variablen werden die realisierten Fehler 1. Art von t- und F-Test unabhängig von der Beobachtungsanzahl pro Parzelle eingehalten. Mit steigender Anzahl der Beobachtungen sinkt der realisierte Fehler 1. Art für alle Tests. Die realisierten Fehler des χ^2 -Test werden überschätzt. Mit steigender Beobachtungsanzahl verbessern sich die Ergebnisse, dies zeigt deutlich die nur asymptotische Gültigkeit dieses Tests.

Von den unterschiedlichen Auswertungsmethoden zeigen die Auswertung der Bonituren als stetiges Merkmal und diejenige der Parzellenmittelwerte der Bonituren mittels LMM ebenfalls die Einhaltung des nominalen Fehlers 1. Art für den t- und den F-Test und eine Überschätzung für den χ^2 -Test.

Die Schwellenwertmodelle zur Auswertung der Bonituren liefern nur für Ansätze mit $n = 5$ Beobachtungen die Einhaltung des nominalen Fehlers 1. Art für t- und F-Test. Die Fehler des t-Tests werden bei geringerer Beobachtungsanzahl unabhängig von der gewählten Linkfunktion in gleicher Größenordnung überschätzt. Für 1 Beobachtung pro Parzelle und „cumlogit“ Linkfunktion wird der Fehler für den F-Test deutlich unterschätzt, der χ^2 -Test trifft hier zwar den nominalen Fehler, ist aber aufgrund der nur asymptotischen Gültigkeit dieses Tests und der starken Unterschätzung des Fehlers für den F-Test an dieser Stelle ebenfalls nur auf eine Unterschätzung zurückzuführen. Die realisierten Fehler 1. Art des χ^2 -Tests zeigen für die Linkfunktion „cumprobit“ Überschätzungen in der Größenordnung wie diejenigen der latenten Variablen.

Die Rangverfahren halten den nominalen Fehler 1. Art für den Prüfgliedkontrast für alle Beobachtungszahlen ein. Beim Test der Globalhypothese wird der Fehler der ANOVA-Typ Statistik für 1 und 3 Beobachtungen unterschätzt. Der Fehler der Wald-Typ Statistik wird für alle Ansätze deutlich überschätzt.

Die Permutationstests unterschätzen den nominalen Fehler 1. Art für den Prüfgliedkontrast, für die Globalhypothese dagegen wird der nominale Fehler eingehalten.

Mit steigender Anzahl an Beobachtungen zeigen die Schwellenwertmodelle und die Rangverfahren kleinere Abweichungen vom nominalen Fehler 1. Art. Für Ansätze mit $n = 1$ Beobachtung pro Parzelle liefert nur die Auswertung der Bonituren als stetiges Merkmal mittels LMM die Einhaltung des nominalen Fehlers 1. Art für den t-Test des Prüfgliedkontrasts und den F-Test der Globalhypothese.

Tab. 5.6: Realisierte Fehler 1. Art der Auswertungsmethoden für Simulationsansätze ohne räumliche Kovarianzstrukturen (t-Test der Prüfglieddiffenz PG 1 zu 2, F- und χ^2 -Test der Globalhypothese unter H_0), Auswertungsvariante: A-R, 10000 Simulationsläufe (fett: Einhaltung des nominalen Fehlers)

		t-Test	F-Test	χ^2 -Test
Auswertung der latenten Variablen				
LMM	n =1	0.0529	0.0537	0.0910
	n =3	0.0507	0.0518	0.0629
	n =5	0.0466	0.0480	0.0554
Auswertung der Bonituren (Parametersatz 1)				
Schwellenwertmodell (SWM) Linkfunktion: „cumprobit“	n =1	0.0815	0.0434	0.1027
	n =3	0.0593	0.0559	0.0699
	n =5	0.0539	0.0534	0.0599
Schwellenwertmodell (SWM) Linkfunktion: „cumlogit“	n =1	0.0785	0.0134	0.0481
	n =3	0.0595	0.0417	0.0533
	n =5	0.0542	0.0468	0.0532
LMM für Parzellenmittelwerte	n =1	-----	-----	-----
	n =3	0.0486	0.0523	0,0917
	n =5	0.0470	0.0527	0,0902
LMM für Boniturnoten	n =1	0.0479	0.0538	0.0917
	n =3	0.0488	0.0527	0.0628
	n =5	0.0478	0.0507	0.0568
Rangverfahren		ANOVA-Typ Kontrast PG 1 zu 2	ANOVA-Typ Globaltest	Wald- Typ
	n =1	0.0467	0.0283	0.7307
	n =3	0.0497	0.0446	0.2078
	n =5	0.0490	0.0480	0.1241
Permutationstests (1000 Permutationen)		Kontrast PG 1 zu 2	Globaltest	-----
	n =1	0.0169	0.0505	-----

5.5.2 Auswertung der Simulationsansätze mit räumlichen Kovarianzstrukturen

5.5.2.1 Auswertung der latenten Variablen (LMM)

Die Tabelle 5.7 zeigt die realisierten Fehler 1. Art der unterschiedlichen Varianten zur Berücksichtigung der räumlichen Kovarianzen (vgl. Abschnitt 4) für die Auswertung der latenten Variablen mittels LMM.

Für die Auswertungsvarianten A-R, A-BI fix und random steigen für alle räumlichen Ansätze die realisierten Fehler aller Tests mit der Anzahl an Beobachtungen pro Parzelle an.

Mit steigender Stärke der simulierten räumlichen Korrelation (vom Ansatz 1_10_1 mit schwacher räumlicher Korrelation bis zur starken räumlichen Korrelation im Ansatz 2_40_1) zeigt sich ein gegenläufiges Verhalten der Fehler bei den Ansätzen mit 1 Beobachtung gegenüber denjenigen mit 3 oder 5 Beobachtungen, wenn die Auswertung mit den Varianten A-R, A-BI fix oder random durchgeführt wurde.

a) Ansätze mit einer Beobachtung pro Parzelle

Ansätze mit 1 Beobachtung zeigen für die Auswertungsvariante A-R mit Zunahme der simulierten räumlichen Kovarianz eine Verringerung der realisierten Fehler 1. Art für alle Tests. Für den t- und F-Test kommt es zu einer zunehmenden Unterschätzung der nominalen Fehler. Der Fehler des χ^2 -Tests wird im schwachen räumlichen Ansatz 1_10_1 überschätzt, mit zunehmender Stärke der räumlichen Korrelation verringert sich diese Überschätzung, bis hin zu einer Unterschätzung im stärksten räumlichen Ansatz (2_40_1).

Die Auswertungsvarianten A-BI fix oder random sowie A-cov(sph) dagegen zeigen nur geringe Veränderungen der realisierten Fehler. Die nominalen Fehler 1. Art werden für den t- und F-Test bei fast allen räumlichen Konstellationen eingehalten. Nur für den Ansatz mit geringer räumlicher Korrelation (1_10_1) zeigen die Auswertungsvarianten A-BI und A-cov(sph) eine leichte Abweichung vom nominalen Fehler. Der χ^2 -Test zeigt Überschätzungen des nominalen Fehlers 1. Art.

b) Ansätze mit 3 oder 5 Beobachtungen

Die empirischen Fehler der Ansätze mit 3 oder 5 Beobachtungen zeigen ein anderes Verhalten. Sie zeigen für die Auswertungsvarianten A-R und A-BI fix oder random schon für Simulationsansätze mit geringer räumlicher Korrelation (1_10_1) eine extreme Überschätzung des nominalen Fehlers 1. Art für alle Tests, die mit zunehmender Stärke der räumlichen Korrelation noch zunimmt. Die Größenordnung der Überschätzung bei der Auswertung mit Blockanlagen ist unabhängig davon, ob der Block als fixer oder zufälliger Faktor ins Modell eingeht, sie liegt jedoch noch etwas höher als bei der Auswertung nach Modell A-R.

Tab. 5.7: Realisierte Fehler 1. Art bei Auswertung der latenten Variablen mittels LMM, für Simulationsansätze mit räumlichen Kovarianzstrukturen (t-Test der Prüfglieddiffrenz PG 1 zu 2, F- und χ^2 -Test der Globalhypothese unter H_0), 10000 Simulationsläufe (fett: Einhaltung des nominalen Fehlers)

Simulierter räumlicher Kovarianzansatz		1_10_1			1_20_1			1_30_1			2_40_1		
Auswertung		t-Test	F-Test	χ^2 -Test									
n = 1	A-R	0.0468	0.0526	0.0904	0.0438	0.0403	0.0729	0.0383	0.0290	0.0560	0.0327	0.0160	0.0332
	A-cov(sph)	0.0494	0.0590	0.1029	0.0503	0.0537	0.1037	0.0496	0.0490	0.1028	0.0488	0.0573	0.1094
	A-BI fix	0.0479	0.0526	0.0952	0.0483	0.0493	0.0905	0.0501	0.0475	0.0903	0.0498	0.0488	0.0901
	A-BI(fix)-Parz.	----	----	----	----	----	----	----	----	----	----	----	----
	A-BI rand.	0.0496	0.0582	0.1004	0.0493	0.0525	0.0938	0.0508	0.0487	0.0927	0.0500	0.0491	0.0909
	A-BI(rand.)-Parz.	----	----	----	----	----	----	----	----	----	----	----	----
n = 3	A-R	0.0838	0.2158	0.2434	0.1095	0.4022	0.4339	0.1100	0.3952	0.4284	0.1254	0.4546	0.4841
	A-cov(sph)	0.0796	0.1902	0.2224	0.0632	0.1179	0.1406	0.0535	0.0704	0.0858	0.0467	0.0515	0.0666
	A-BI fix	0.0853	0.2221	0.2510	0.1154	0.4546	0.4863	0.1271	0.4987	0.5308	0.1617	0.6585	0.6912
	A-BI(fix)-Parz.	0.0471	0.0499	0.0875	0.0491	0.0492	0.0937	0.0480	0.0501	0.0936	0.0479	0.0505	0.0912
	A-BI rand.	0.0857	0.2246	0.2543	0.1156	0.4555	0.4873	0.1273	0.4992	0.5311	0.1618	0.6588	0.6914
	A-BI(rand.)-Parz.	0.0484	0.0540	0.0928	0.0497	0.0521	0.0956	0.0485	0.0507	0.0946	0.0480	0.0508	0.0916
n = 5	A-R	0.1070	0.3708	0.3896	0.1679	0.7199	0.7339	0.1735	0.7272	0.7424	0.2000	0.9781	0.9783
	A-cov(sph)	0.0730	0.1747	0.1966	0.0579	0.0850	0.0957	0.0538	0.0563	0.0657	0.0512	0.0472	0.0545
	A-BI fix	0.1087	0.3839	0.4031	0.1772	0.7592	0.7724	0.1929	0.8005	0.8138	0.2471	0.9172	0.9238
	A-BI(fix)-Parz.	0.0468	0.0516	0.0900	0.0494	0.0508	0.0935	0.0488	0.0492	0.0925	0.0455	0.0482	0.0862
	A-BI rand.	0.1089	0.3844	0.4040	0.1772	0.7594	0.7727	0.1929	0.8006	0.8140	0.2471	0.9172	0.9238
	A-BI(rand.)-Parz.	0.0475	0.0546	0.0946	0.0498	0.0522	0.0947	0.0490	0.0495	0.0930	0.0457	0.0483	0.0862

Die Auswertungsansätze mit räumlicher Kovarianz zeigen dagegen deutlich kleinere Überschätzungen, welche sich mit zunehmender räumlicher Korrelation (Ansatz 1_30_1 und 2_40_1) verringern. Für diese Ansätze werden die realisierten Fehler 1. Art für den t- und F-Test eingehalten. Auch die Verzerrung der realisierten Fehler des χ^2 -Tests sinkt für die Auswertungsvariante A-cov(sph) für Simulationsansätze mit stärkerer räumlicher Abhängigkeit.

Die Auswertungsvarianten mit zusätzlichem zufälligen Parzelleneffekt innerhalb der Blocks (A-BI(fix)-Parz. und A-BI(rand.)-Parz.) zeigen für $n = 3$ Beobachtungen durchgehend eine gute Realisierung der nominalen Fehler 1. Art für den t- und den F-Test. Für $n = 5$ Beobachtungen gibt es in wenigen Fällen leichte Verzerrungen. Der χ^2 -Test zeigt Verzerrungen in Größenordnungen, die deutlich geringer sind als für die Modellvarianten A-R, A-BI fix und random, aber etwas größer als diejenigen des räumlichen Modells A-cov(sph).

5.5.2.2 Auswertung der Bonituren mittels Schwellenwertmodell (SWM)

Die Tabellen 5.8 und 5.9 zeigen die im linearen Prädiktor realisierten Fehler 1. Art der unterschiedlichen Varianten zur Berücksichtigung der räumlichen Kovarianzen (vgl. Abschnitt 4) für die Auswertung der Boniturnoten (Parametersatz 1) mittels SWM mit den Linkfunktionen „cumprobit“ und „cumlogit“. Die Unterschiede im Verhalten der empirischen Fehler zwischen den Auswertungen mit den unterschiedlichen Linkfunktionen sind sehr gering. Bei der Beurteilung der realisierten Fehler 1. Art für die Auswertung der Bonituren mittels Schwellenwertmodell muss bedacht werden, dass diese empirischen Fehler aus teilweise stark unterschiedlichen Anteilen konvergenter Simulationsläufe entstanden sind (siehe Abschnitt 5.1), wodurch empirische Fehler gleicher Größenordnung - je nach Konvergenzverhalten der jeweiligen Auswertungsvariante - bezüglich der Einhaltung des nominalen Fehlers durchaus unterschiedlich bewertet werden können.

Unabhängig von der gewählten Linkfunktion zeigen die Auswertungsvarianten A-R, A-BI fix und random einen Anstieg der realisierten Fehler 1. Art aller Tests mit der Anzahl an Beobachtungen pro Parzelle. Die realisierten Fehler liegen für die Ansätze mit 3 und 5 Beobachtungen pro Parzelle in der gleichen Größenordnung wie für die Auswertung der latenten Variablen (vgl. Tab. 5.7). Für die Auswertung der Ansätze mit $n = 1$ finden sich leichte Unterschiede zu den Ergebnissen der latenten Variablen, im Schwellenwertmodell wird der nominale Fehler des t-Tests bei den obigen Auswertungsvarianten fast immer überschätzt. Ausnahme ist die Variante A-BI random für den Simulationsansatz 1_10_1. Auch für das Schwellenwertmodell gilt für diese Auswertungsvarianten, dass mit steigender Stärke der simulierten räumlichen Korrelation das Verhalten der realisierten Fehler 1. Art für die Ansätze mit $n = 1$ und diejenigen mit $n > 1$ unterschiedlich ist. Ansätze mit $n = 1$ Beobachtung zeigen für die Auswertungsvariante A-R mit Zunahme der simulierten räumlichen Kovarianz eine Verringerung der realisierten Fehler 1. Art für alle Tests. Für den F-Test kommt es zu einer zunehmenden Unterschätzung der nominalen Fehler. Die Auswertungsvariante A-BI fix dagegen zeigt nur geringe Veränderungen der

realisierten Fehler, ohne eine solche Tendenz zur Verringerung. Die Variante A-BI random zeigt für den t-Test eine zunehmende Überschätzung der Fehler. Für die Auswertung der Bonituren mittels Schwellenwertmodell werden für $n = 1$ die nominalen Fehler 1. Art des t- und F-Tests im Gegensatz zur Auswertung der latenten Variablen (vgl. Tab. 5.7) von den meisten Auswertungsvarianten nicht eingehalten. Die realisierten Fehler der Ansätze mit $n > 1$ zeigen, mit zunehmender Stärke der räumlichen Kovarianz, eine zunehmende Überschätzung der nominalen Fehler 1. Art für die Auswertungsvarianten A-R, A-BI fix und random.

Zur Auswertung von Ansätzen mit $n > 1$ können für das Schwellenwertmodell ebenfalls die um einen zusätzlichen Parzelleneffekt erweiterten Blockmodelle (vgl. Abschnitt 4.2) als Auswertungsvarianten genutzt werden. Auch hier sind die Unterschiede zwischen den Ergebnissen für die Linkfunktionen „cumprobit“ und „cumlogit“ nur gering. Da die empirischen Fehler 1. Art des t-Tests für diese Auswertungsvarianten jedoch in der Größenordnung des nominalen Fehlers liegen, sind sie teilweise entscheidend bezüglich der Einhaltung oder Nichteinhaltung des nominalen Fehlers.

Für die Linkfunktion „cumprobit“ und Ansätze mit $n = 3$ zeigen die Auswertungsvarianten A-BI(fix)-Parz. und A-BI(rand.)-Parz. für den t-Test eine gute Einhaltung der nominalen Fehler 1. Art bei fast allen räumlichen Konstellationen. Für die Simulationsansätze mit $n = 5$ hält die Variante A-BI(fix)-Parz. den nominalen Fehler des t-Tests für fast alle Simulationsansätze mit räumlicher Korrelation ein (Ausnahme Ansatz 2_40_1). Die Variante A-BI(rand.)-Parz. dagegen nur für den 1_10_1er Ansatz, bei stärkeren räumlichen Korrelationen wird der nominale Fehler leicht unterschätzt. Der nominale Fehler des F-Tests wird für $n = 3$ und 5 Beobachtungen je Parzelle von beiden Auswertungsvarianten unterschätzt. Für die Linkfunktion „cumlogit“ zeigen die Auswertungsvarianten A-BI(fix)-Parz. und A-BI(rand.)-Parz. für fast alle räumlichen Ansätze (Ausnahme Ansatz 1_10_1) ein leichte Unterschätzung der nominalen Fehler 1. Art des t-Tests für $n = 3$ und 5 Beobachtungen. Unabhängig von der Linkfunktion werden die Fehler des F-Tests unterschätzt. Die nominalen Fehler des χ^2 -Tests werden mit der Linkfunktion „cumprobit“ überschätzt, mit der Linkfunktion „cumlogit“ werden die nominalen Fehler für $n = 3$ Beobachtungen eingehalten, für $n = 5$ Beobachtungen ebenfalls überschätzt.

Wird mit dem räumlichen Modell (Variante A-cov(sph)) und der Linkfunktion „cumprobit“ ausgewertet, wird nur für einzelne Simulationsansätze der nominale Fehler 1. Art von t-, F- oder χ^2 -Test eingehalten. Für $n = 1$ werden die Fehler von t- und F-Test meist unterschätzt. Diese Unterschätzung nimmt mit zunehmender Stärke der räumlichen Korrelation zu. Für $n > 1$ werden die Fehler in Simulationsansätzen mit schwacher räumlicher Korrelation für den t-Test leicht und für den F- und χ^2 -Test deutlich überschätzt. Mit zunehmender Stärke der räumlichen Korrelation sinken die realisierten Fehler, so dass sich für den räumlichen Ansatz 2_40_1 eine Unterschätzung des nominalen Fehlers von t- und F-Test zeigt und der χ^2 -Test den nominalen Fehler realisiert.

Im Gegensatz dazu zeigt die Auswertungsvariante A-cov(sph) mit der Linkfunktion „cumlogit“ für $n = 1$ die Einhaltung des nominalen Fehlers 1. Art des t-Test für fast alle räumlichen Ansätze. Die Fehler von F- und χ^2 -Test dagegen werden stark unterschätzt.

Tab. 5.8: Realisierte Fehler 1. Art für den linearen Prädiktor bei Auswertung der Boniturnoten (Parametersatz 1) mittels SWM (Linkfunktion „cumprobit“), für Simulationsansätze mit räumlichen Kovarianzstrukturen (t-Test der Prüfglieddiffereenz PG 1 zu 2, F- und χ^2 -Test der Globalhypothese unter H0), 10000 Simulationsläufe (fett: Einhaltung des nominalen Fehlers)

Simulierter räumlicher Kovarianzansatz		1_10_1			1_20_1			1_30_1			2_40_1		
Auswertung		t-Test	F-Test	χ^2 -Test	t-Test	F-Test	χ^2 -Test	t-Test	F-Test	χ^2 -Test	t-Test	F-Test	χ^2 -Test
n = 1	A-R	0.0812	0.0424	0.1012	0.0767	0.0296	0.0827	0.0661	0.0220	0.0648	0.0584	0.0125	0.0405
	A-cov(sph)	0.0437	0.0332	0.0991	0.0428	0.0220	0.0780	0.0420	0.0151	0.0491	0.0329	0.0050	0.0186
	A-BI fix	0.0893	0.0560	0.1316	0.0893	0.0514	0.1343	0.0890	0.0521	0.1317	0.0909	0.0500	0.1255
	A-BI(fix)-Parz.	----	----	----	----	----	----	----	----	----	----	----	----
	A-BI rand.	0.0472	0.0418	0.1203	0.0658	0.0510	0.1183	0.0719	0.0580	0.1182	0.0816	0.0589	0.1100
	A-BI(rand.)-Parz.	----	----	----	----	----	----	----	----	----	----	----	----
n = 3	A-R	0.0944	0.2136	0.2444	0.1211	0.3916	0.4285	0.1211	0.3818	0.4168	0.1333	0.4395	0.4696
	A-cov(sph)	0.0574	0.1223	0.2325	0.0666	0.0947	0.1442	0.0509	0.0487	0.0676	0.0383	0.0313	0.0430
	A-BI fix	0.0979	0.2332	0.2663	0.1309	0.4567	0.4920	0.1403	0.4907	0.5300	0.1744	0.6471	0.6784
	A-BI(fix)-Parz.	0.0409	0.0140	0.0440	0.0470	0.0246	0.0576	0.0474	0.0255	0.0583	0.0466	0.0274	0.0623
	A-BI rand.	0.0583	0.1360	0.2564	0.1158	0.3975	0.4817	0.1348	0.4648	0.5173	0.1681	0.6337	0.6674
	A-BI(rand.)-Parz.	0.0481	0.0247	0.0610	0.0472	0.0278	0.0627	0.0476	0.0271	0.0604	0.0469	0.0273	0.0625
n = 5	A-R	0.1108	0.3623	0.3835	0.1723	0.7006	0.7177	0.1764	0.7045	0.7185	0.2022	0.7771	0.7909
	A-cov(sph)	0.0719	0.1379	0.2151	0.0585	0.0726	0.0831	0.0571	0.0476	0.0557	0.0455	0.0323	0.0401
	A-BI fix	0.1138	0.3836	0.4049	0.1830	0.7464	0.7620	0.1967	0.7870	0.8023	0.2460	0.9013	0.9104
	A-BI(fix)-Parz.	0.0464	0.0365	0.0771	0.0465	0.0367	0.0712	0.0456	0.0325	0.0691	0.0439	0.0311	0.0684
	A-BI rand.	0.0827	0.2562	0.3972	0.1705	0.6849	0.7534	0.1919	0.7619	0.7977	0.2413	0.8922	0.9074
	A-BI(rand.)-Parz.	0.0457	0.0356	0.0737	0.0445	0.0363	0.0701	0.0448	0.0337	0.0711	0.0451	0.0306	0.0680

Tab. 5.9: Realisierte Fehler 1. Art für den linearen Prädiktor bei Auswertung der Boniturnoten (Parametersatz 1) mittels SWM (Linkfunktion „cumlogit“), für Simulationsansätze mit räumlichen Kovarianzstrukturen (t-Test der Prüfglieddifferenz PG 1 zu 2, F- und χ^2 -Test der Globalhypothese unter H0), 10000 Simulationsläufe (fett: Einhaltung des nominalen Fehlers)

Simulierter räumlicher Kovarianzansatz		1_10_1			1_20_1			1_30_1			2_40_1		
Auswertung		t-Test	F-Test	χ^2 -Test	t-Test	F-Test	χ^2 -Test	t-Test	F-Test	χ^2 -Test	t-Test	F-Test	χ^2 -Test
n = 1	A-R	0.0789	0.0126	0.0476	0.0733	0.0105	0.0358	0.0654	0.0057	0.0264	0.0557	0.0029	0.0154
	A-cov(sph)	0.0470	0.0091	0.0494	0.0448	0.0054	0.0343	0.0468	0.0033	0.0241	0.0467	0.0008	0.0120
	A-BI fix	0.0867	0.0189	0.0665	0.0880	0.0196	0.0690	0.0870	0.0180	0.0693	0.0924	0.0198	0.0711
	A-BI(fix)-Parz.	----	----	----	----	----	----	----	----	----	----	----	----
	A-BI rand.	0.0497	0.0186	0.0668	0.0670	0.0229	0.0638	0.0738	0.0254	0.0651	0.0859	0.0263	0.0631
	A-BI(rand.)-Parz.	----	----	----	----	----	----	----	----	----	----	----	----
n = 3	A-R	0.0940	0.1733	0.2040	0.1184	0.3345	0.3738	0.1214	0.3267	0.3635	0.1293	0.3813	0.4178
	A-cov(sph)	0.0493	0.0947	0.1970	0.0476	0.0385	0.0764	0.0487	0.0208	0.0362	0.0402	0.0121	0.0191
	A-BI fix	0.0974	0.1946	0.2254	0.1334	0.4037	0.4448	0.1381	0.4442	0.4854	0.1730	0.6025	0.6370
	A-BI(fix)-Parz.	0.0457	0.0182	0.0490	0.0445	0.0204	0.0480	0.0454	0.0193	0.0466	0.0447	0.0162	0.0465
	A-BI rand.	0.0626	0.1298	0.2296	0.1173	0.3633	0.4401	0.1296	0.4270	0.4791	0.1681	0.5953	0.6283
	A-BI(rand.)-Parz.	0.0467	0.0201	0.0516	0.0455	0.0204	0.0480	0.0458	0.0199	0.0460	0.0440	0.0159	0.0461
n = 5	A-R	0.1100	0.3274	0.3465	0.1698	0.6636	0.6817	0.1717	0.6649	0.6833	0.1966	0.7492	0.7634
	A-cov(sph)	0.0643 ^{*)}	0.1080 ^{*)}	0.1735 ^{*)}	0.0495	0.0381	0.0464	0.0493	0.0255	0.0324	0.0392	0.0195	0.0247
	A-BI fix	0.1138	0.3486	0.3701	0.1833	0.7190	0.7338	0.1956	0.7603	0.7744	0.2419	0.8872	0.8977
	A-BI(fix)-Parz.	0.0473	0.0309	0.0646	0.0455	0.0308	0.0637	0.0454	0.0277	0.0630	0.0424	0.0247	0.0584
	A-BI rand.	0.0824	0.2367	0.3738	0.1713	0.6663	0.7313	0.1901	0.7400	0.7720	0.2403	0.8796	0.8943
	A-BI(rand.)-Parz.	0.0470	0.0338	0.0652	0.0455	0.0309	0.0651	0.0455	0.0269	0.0619	0.0423	0.0248	0.0582

^{*)} Ergebnisse von 1000 Läufen, da Simulationen mit 10000 Läufen aufgrund sehr langer Laufzeiten nicht realisierbar waren

Das Verhalten und die Größenordnung der realisierten Fehler 1. Art für die Ansätze mit $n > 1$ stellen sich für die Linkfunktion „cumlogit“ ähnlich dar wie für die Auswertung mit der Linkfunktion „cumprobit“. Die nominalen Fehler des t-Tests werden jedoch für die Linkfunktion „cumlogit“ deutlich häufiger eingehalten. Man muss bei dem Vergleich mit der Linkfunktion „cumprobit“ jedoch bedenken, dass die Ergebnisse jeweils durch einen sehr unterschiedlichen Anteil konvergenter Simulationsläufe entstanden sind (siehe Tab. 5.1).

Beim Vergleich aller Auswertungsvarianten für das Schwellenwertmodell mit „cumprobit“ Linkfunktion zeigt sich, dass für Ansätze mit $n = 1$ die Variante A-cov(sph) unter allen Auswertungsvarianten die geringsten Abweichungen vom nominalen Fehler 1. Art des t-Tests liefert. Der nominale Fehler des F-Tests dagegen wird für $n = 1$ am ehesten von den Auswertungsvarianten mit Blockeffekt getroffen. Die Fehler des χ^2 -Test werden überschätzt. Für $n > 1$ zeigen die Auswertungsvarianten der Blockmodelle mit zufälligem Parzelleneffekt die häufigste Einhaltung des nominalen Fehlers 1. Art für den t-Test. Die Variante A-cov(sph) zeigt, im Gegensatz zu den Auswertungsvarianten A-R und A-BI fix oder random, nur leichte Verzerrungen. Der nominale Fehler des F-Tests wird von den Auswertungsvarianten mit zufälligem Parzelleneffekt unabhängig von der Stärke der simulierten räumlichen Kovarianz in konstanter Größenordnung unterschätzt. Die räumliche Variante A-cov(sph) dagegen überschätzt den Fehler 1. Art von F- χ^2 - Test bei Ansätzen mit geringer räumlicher Korrelation deutlich. Mit zunehmender Stärke der simulierten räumlichen Kovarianz reduziert sich diese Überschätzung, bis hin zu einer Unterschätzung beim stärksten räumlichen Ansatz (2_40_1).

Unter allen Auswertungsvarianten des Schwellenwertmodells mit „cumlogit“ Linkfunktion zeigt die räumliche Auswertungsvariante A-cov(sph), unabhängig von der Beobachtungsanzahl pro Parzelle, für den t-Test die beste Einhaltung des nominalen Fehlers 1. Art, gefolgt von den Ansätzen mit zufälligem Parzelleneffekt. Auch mit dieser Linkfunktion werden die nominalen Fehler des F-Tests nicht realisiert, diejenigen des χ^2 -Tests nur für die Auswertungsvarianten mit zufälligem Parzelleneffekt für $n = 3$ Beobachtungen pro Parzelle.

Alle bisherigen Analysen der Bonituren mittels Schwellenwertmodell erfolgten anhand des Parametersatzes 1 für Bonituren, der mit homogenen Prüfgliedeffekten und $k = 9$ Boniturklassen simuliert wurde. Für die alternativen Parametersätze (vgl. Abschnitt 3.3) werden ausgewählte Analysen mittels SWM und LMM in Abschnitt 5.7 dargestellt.

5.5.2.3 Auswertung der Bonituren mittels Rangverfahren

Die realisierten Fehler 1. Art der Auswertung der Bonituren (Parametersatz 1) mittels Rangverfahren sind in Tabelle 5.10 wiedergegeben. Die beiden möglichen Auswertungsvarianten der Rangverfahren A-R und A-BI fix sind für Simulationsansätze mit räumlichen Korrelationen nicht in der Lage, die nominalen Fehler 1. Art einzuhalten. Für Ansätze mit $n = 1$ wird mit der Auswertungsvariante A-R der nominale Fehler 1. Art der ANO-

VA-Typ Statistik für den Test der Prüfglied- und Globalhypothese unterschätzt. Die Wald-Typ Statistik liefert eine starke Überschätzung. Mit steigender Stärke der simulierten räumlichen Korrelation verstärken sich diese Verzerrungen des nominalen Fehlers. Die Auswertungsvariante A-BI fix dagegen zeigt Überschätzungen des nominalen Fehlers 1. Art für alle Tests und alle räumlichen Simulationsansätze. Die Verzerrungen sind für diesen Modellansatz bei der Auswertung der Bonituren mittels Rangverfahren stärker als bei der Auswertung mittels Schwellenwertmodell. Die Ergebnisse der Auswertungsvariante A-BI fix für $n = 1$ sind aber unter dem Gesichtspunkt zu betrachten, dass die Methode nur für $n > 1$ definiert ist (siehe Abschnitt 4.3) und sie lediglich aus Vergleichsgründen gerechnet wurde.

Für Ansätze mit $n > 1$ zeigen sich realisierte Fehler 1. Art in der gleichen Größenordnung wie bei der Auswertung der latenten Variablen (vgl. Tab. 5.7) bzw. wie bei der Auswertung der Bonituren mittels Schwellenwertmodell (vgl. Tab 5.8 und 5.9). Auch bei den Rangverfahren verstärken sich die Verzerrungen des nominalen Fehlers 1. Art mit steigender Beobachtungsanzahl und steigender Stärke der räumlichen Korrelation. Für die Auswertungsvariante A-BI fix liefert die Wald-Typ Statistik für Ansätze mit $n = 1$ viel stärkere Verzerrungen als für Ansätze mit $n > 1$.

5.5.2.4 Auswertung der Bonituren mittels Permutationstests

Die Tabelle 5.11 zeigt die realisierten Fehler 1. Art der Varianten zur Auswertung der Boniturnoten (Parametersatz 1) mit nur einer Beobachtung pro Parzelle mittels Permutationstests, für Simulationsansätze mit räumlichen Kovarianzstrukturen. Die Permutationstests zur Auswertung von Bonituren zeigen für beide Auswertungsvarianten eine Unterschätzung des nominalen Fehlers 1. Art für den Test der Differenz von Prüfglied 1 zu 2. Mit steigender räumlicher Korrelation wird die Verzerrung deutlicher. Der nominale Fehler für die Globalhypothese wird in der Variante A-R nur bei schwacher räumlicher Korrelation (Ansatz 1_10_1) eingehalten. Der mit den um den Blockmedian korrigierten Daten durchgeführte Test der Globalhypothese in der Variante A-BI fix (vgl. Abschnitt 2.4.2) dagegen liefert unabhängig von der simulierten räumlichen Kovarianzstruktur eine gute Einhaltung des nominalen Fehlers 1. Art.

Tab. 5.10: Realisierte Fehler 1. Art der Varianten zur Auswertung der Boniturnoten (Parametersatz 1) mittels Rangverfahren, für Simulationsansätze mit räumlichen Kovarianzstrukturen (ANOVA-Typ Statistik der Prüfglieddiffenz PG 1 zu 2, Test der Globalhypothese mittels ANOVA - und Wald-Typ Statistik unter H0), 10000 Simulationsläufe

Simulierter räumlicher Kovarianzansatz		1_10_1			1_20_1			1_30_1			2_40_1		
Auswertung		Anova-Typ Kontrast	Anova-Typ Global	Wald-Type Global	Anova-Typ Kontrast	Anova-Typ Global	Wald-Type Global	Anova-Typ Kontrast	Anova-Typ Global	Wald-Type Global	Anova-Typ Kontrast	Anova-Typ Global	Wald-Type Global
		n = 1	A-R	0.0454	0.0276	0.7278	0.0434	0.0208	0.06988	0.0127	0.0152	0.6629	0.0344
	A-BI fix ^{*)}	0.1075	0.2118	0.9996	0.1266	0.2080	0.9997	0.1027	0.2022	0.9993	0.1002	0.1821	0.9987
n = 3	A-R	0.0799	0.1787	0.4541	0.1065	0.3486	0.6310	0.1069	0.3386	0.6212	0.1168	0.3941	0.6713
	A-BI fix	0.0799	0.2462	0.6245	0.1266	0.5892	0.8602	0.1396	0.6417	0.8850	0.1900	0.8370	0.9590
n = 5	A-R	0.0998	0.3327	0.4944	0.1584	0.6707	0.7961	0.1610	0.6718	0.7997	0.1840	0.7524	0.8586
	A-BI fix	0.1112	0.4289	0.6170	0.2036	0.8508	0.9276	0.2242	0.8853	0.9440	0.2981	0.9701	0.9865

^{*)} Methode für n = 1 nicht definiert, lediglich aus Vergleichsgründen gerechnet

Tab. 5.11: Realisierte Fehler 1. Art der Varianten zur Auswertung der Boniturnoten (Parametersatz 1) mittels Permutationstests, für Simulationsansätze mit räumlichen Kovarianzstrukturen (Test der Prüfglieddiffenz PG 1 zu 2 und Test der Globalhypothese unter H0), 10000 Simulationsläufe, 1000 Permutationen (fett: Einhaltung des nominalen Fehlers)

Simulierter räumlicher Kovarianzansatz		1_10_1		1_20_1		1_30_1		2_40_1	
Auswertung		Kontrast PG 1 zu 2	Global						
		n = 1	A-R	0.0175	0.0494	0.0149	0.0369	0.0123	0.0297
	A-BI fix	0.0237	0.0517^{*)}	0.0220	0.0475^{*)}	0.0212	0.0498^{*)}	0.0195	0.0483^{*)}

^{*)} Globalhypothese mit um den Blockmedian korrigierten Daten

5.5.2.5 Auswertung der Parzellenmittelwerte der Bonituren (LMM)

Existiert in den simulierten Daten eine räumliche Kovarianz, so sind die Resteffekte miteinander korreliert. Werden die Parzellenmittelwerte der Bonituren gebildet, so sind auch die Resteffekte dieser Mittelwerte korreliert (vgl. Abschnitt 2.5). Das ist der Grund, weshalb auch für die Auswertung der Parzellenmittelwerte mittels LMM die unterschiedlichen Varianten zur Berücksichtigung der räumlichen Kovarianz zum Einsatz kommen.

Die Bildung der Parzellenmittelwerte hat jedoch Einfluss auf die Höhe der Korrelation der Resteffekte. Vor der Darstellung der Ergebnisse zur Einhaltung der Fehler 1. Art für die Auswertung der Parzellenmittelwerte der Bonituren wird im Folgenden die Veränderung der Korrelation durch Bildung der Parzellenmittelwerte für die unterschiedlichen räumlichen Simulationsansätze dargestellt. Dies erfolgt anhand der Resteffekte der simulierten latenten Variablen.

Die Tabelle 5.12 zeigt die Korrelationen der Resteffekte benachbarter Parzellen. Mit steigender Stärke der simulierten räumlichen Kovarianz verstärkt sich die Korrelation zwischen den Werten benachbarter Parzellen. Das ist unabhängig davon, ob die Korrelation der Mittelwerte, die Korrelation der Beobachtungen im Parzellenmittelpunkt (Beobachtung 5, vgl. Abb. 3.2) oder die mittlere Korrelation aller Beobachtungen benachbarter Parzellen betrachtet wird. Die Korrelation der Beobachtungen im Parzellenmittelpunkt ist stets etwas höher als die mittlere Korrelation aller Beobachtungen da bei letzterer auch die weiter von einander entfernt liegenden Beobachtungen betrachtet werden, welche aufgrund ihres größeren Abstandes schwächer korreliert sind. Werden jedoch die Mittelwerte der Resteffekte der Parzellen gebildet, so sind diese deutlich stärker korreliert. Gelten die beschriebenen Korrelationsverhältnisse für die latente Variable, so ist dies auch für die Bonituren der Fall.

Die Tabelle 5.13 zeigt die realisierten Fehler 1. Art der Varianten zur Auswertung der Parzellenmittelwerte der Boniturnoten (Parametersatz 1) mittels LMM für Simulationsansätze mit räumlichen Kovarianzstrukturen. Die Größenordnungen der realisierten Fehler aller Auswertungsvarianten sind für $n = 3$ und 5 Beobachtungen je Parzelle vergleichbar mit denen der Auswertung der latenten Variablen mit $n = 1$ (vgl. Tab. 5.7).

Die Auswertungsvariante A-R zeigt unabhängig von der Beobachtungsanzahl je Parzelle eine Unterschätzung des nominalen Fehlers 1. Art für den t- und den F-Test, die sich mit zunehmender simulierter räumlicher Kovarianz verstärkt. Ausnahme ist der Fehler des t-Tests des räumlichen Ansatzes 1_10_1, der nur schwache räumliche Korrelation aufweist. Die nominalen Fehler des χ^2 -Tests werden für schwache räumliche Simulationsansätze überschätzt, mit stärkerer räumlicher Korrelation verringern sich die realisierten Fehler und unterschätzen für die stärkeren räumlichen Ansätze die nominalen Fehler 1. Art.

Mit den räumlichen Auswertungsvarianten A-Cov(sph) werden unabhängig von der Beobachtungsanzahl je Parzelle die nominalen Fehler aller Tests für alle räumlichen Strukturen leicht überschätzt (Ausnahme Fehler des t-Tests für $n = 3$ im Ansatz 1_10_1).

Die Auswertungsvarianten A-BI fix und random dagegen halten für alle Ansätze bei $n = 3$ und 5 die nominalen Fehler der t- und F-Tests ein. Dies ist ein deutlicher Unterschied zur Auswertung der latenten Variablen (vgl. Tab. 5.7) und der Bonituren mittels Schwellenwertmodell (vgl. Tab. 5.8 und 5.9), wo diese Auswertungsvarianten starke Verzerrungen aufweisen. Auch die Variante A-R zeigt für die Auswertung der Parzellenmittelwerte deutlich geringere Verzerrungen.

Tab. 5.12: Korrelationen der Resteffekte benachbarter Parzellen

Simulierter räumlicher Kovarianzansatz		1_10_1	1_20_1	1_30_1	2_40_1
n = 3	Korrelation der Mittelwerte benachbarter Parzellen	0.0623	0.3238	0.4651	0.6263
	Korrelation der Beobachtungen im Parzellenmittelpunkt (Beob 5) benachbarter Parzellen	0.0073	0.1853	0.2818	0.4455
	Mittlere Korrelation aller Beobachtungen benachbarter Parzellen	0.0565	0.1718	0.2669	0.4286
n = 5	Korrelation der Mittelwerte benachbarter Parzellen	0.1378	0.3985	0.5475	0.6909
	Korrelation der Beobachtungen im Parzellenmittelpunkt (Beob 5) benachbarter Parzellen	0.0073	0.1853	0.2818	0.4455
	Mittlere Korrelation aller Beobachtungen benachbarter Parzellen	0.0414	0.1722	0.266	0.4273

Tab. 5.13: Realisierte Fehler 1. Art der Varianten zur Auswertung der Parzellenmittelwerte der Boniturnoten (Parametersatz 1) mittels LMM, für Simulationsansätze mit räumlichen Kovarianzstrukturen (t-Test der Prüfglieddiffereenz PG 1 zu 2, F- und χ^2 -Test der Globalhypothese unter H0), 10000 Simulationsläufe (fett: Einhaltung des nominalen Fehlers)

Simulierter räumlicher Kovarianzansatz		1_10_1			1_20_1			1_30_1			2_40_1		
Auswertung		t-Test	F-Test	χ^2 -Test									
		n = 3	A-R	0.0463	0.0452	0.0833	0.0429	0.0347	0.0653	0.0351	0.0207	0.0433	0.0260
	A-cov(sph)	0.0518	0.0573	0.1039	0.0670	0.0726	0.1267	0.0583	0.0567	0.1130	0.0589	0.0591	0.1075
	A-BI fix	0.0476	0.0491	0.0909	0.0492	0.0506	0.0947	0.0501	0.0486	0.0932	0.0498	0.0488	0.0902
	A-BI rand.	0.0489	0.0529	0.0961	0.0499	0.0525	0.0973	0.0504	0.0492	0.0941	0.0500	0.0494	0.0904
n = 5	A-R	0.0463	0.0431	0.0801	0.0389	0.0313	0.0572	0.0315	0.0175	0.0351	0.0277	0.0075	0.0172
	A-cov(sph)	0.0547	0.0570	0.1026	0.0791	0.0822	0.1295	0.0643	0.0582	0.1038	0.0709	0.0602	0.0987
	A-BI fix	0.0490	0.0520	0.0919	0.0476	0.0497	0.0918	0.0485	0.0494	0.0876	0.0464	0.0469	0.0873
	A-BI rand.	0.0500	0.0541	0.0962	0.0481	0.0512	0.0933	0.0487	0.0497	0.0883	0.0465	0.0470	0.0875

5.5.2.6 Auswertung der Bonituren als stetiges Merkmal (LMM)

Die Auswertung der Bonituren (Parametersatz 1) als stetiges Merkmal mit dem linearen gemischten Modell (Tab. 5.14) zeigt realisierte Fehler 1. Art in derselben Größenordnung und mit dem gleichen Verhalten bei den unterschiedlichen Auswertungsvarianten wie diejenigen der Auswertung der latenten Variablen (vgl. Abschnitt 5.5.2.1 und Tab. 5.7).

Das Verhalten der Fehler mit steigender Stärke der simulierten räumlichen Kovarianz ist für die Auswertungsvarianten A-R, A-BI fix und random unterschiedlich für Ansätze mit $n = 1$ und $n > 1$ Beobachtungen.

a) Ansätze mit einer Beobachtung pro Parzelle

Die Auswertungsvariante A-R unterschätzt die nominalen Fehler 1. Art mit zunehmender Stärke der räumlichen Korrelation. Die Einhaltung des nominalen Fehlers ist nur bei schwacher räumlicher Korrelation gegeben. Wie auch bei der Auswertung der latenten Variablen wird der Fehler des χ^2 -Tests im schwächsten räumlichen Ansatz überschätzt, dann aber sinkt sein realisierter Fehler mit wachsender räumlicher Korrelation.

Die Varianten A-BI fix und random dagegen zeigen nur wenig Reaktion auf die Variation der räumlichen Korrelation, und die nominalen Fehler von t- und F-Test werden für alle räumlichen Ansätze eingehalten. Der Fehler des χ^2 -Tests wird überschätzt.

Die Auswertungsvariante A-cov(sph) hält den nominalen Fehler 1. Art für den t-Test für alle räumlichen Ansätze ein, für den F-Test nur bei stärkerer räumlicher Korrelation. Allerdings sind die Verzerrungen bei den schwachen räumlichen Ansätzen nur gering.

b) Ansätze mit 3 oder 5 Beobachtungen

Die Auswertungsvarianten A-R und A-BI fix oder random zeigen für alle Simulationsansätze mit räumlicher Korrelation eine extreme Überschätzung des nominalen Fehlers 1. Art für alle Tests, die mit zunehmender Stärke der räumlichen Korrelation noch anwächst. Die Unterschiede in der Größenordnung der Verzerrung zwischen diesen Varianten sind nur gering.

Die räumliche Auswertungsvariante zeigt bei den schwachen räumlichen Ansätzen ebenfalls Überschätzungen der Fehler. Bei den Ansätzen mit stärkerer räumlicher Korrelation sinken die realisierten Fehler, so dass für den t- und den F-Test der nominale Fehler 1. Art eingehalten wird.

Die Auswertungsvarianten mit zusätzlichem zufälligen Parzelleneffekt innerhalb der Blocks zeigen eine durchgehend gute Einhaltung der nominalen Fehler 1. Art bei allen räumlichen Ansätzen für den t- und den F-Test, Ausnahme ist der F-Test für $n = 5$ beim schwachen räumlichen Ansatz 1_10_1, der sehr knapp überschätzt wird. Der χ^2 -Tests dagegen wird überschätzt.

Tab. 5.14: Realisierte Fehler 1. Art der Varianten zur Auswertung der Bonituren (Parametersatz 1) als stetige Variable mittels LMM, für Simulationsansätze mit räumlichen Kovarianzstrukturen (t-Test der Prüfglieddiffenz PG 1 zu 2, F- und χ^2 -Test der Globalhypothese unter H_0), 10000 Simulationsläufe (fett: Einhaltung des nominalen Fehlers)

Simulierter räumlicher Kovarianzansatz		1_10_1			1_20_1			1_30_1			2_40_1		
Auswertung		t-Test	F-Test	χ^2 -Test									
		n = 1	A-R	0.0493	0.0508	0.0897	0.0459	0.0368	0.0745	0.0411	0.0308	0.0582	0.0340
A-cov(sph)	0.0515		0.0565	0.1012	0.0515	0.0548	0.1086	0.0506	0.0545	0.1076	0.0484	0.0507	0.1049
A-BI fix	0.0495		0.0514	0.0928	0.0488	0.0465	0.0916	0.0495	0.0507	0.0925	0.0498	0.0480	0.0887
A-BI(fix)-Parz.	----		----	----	----	----	----	----	----	----	----	----	----
A-BI rand.	0.0518		0.0562	0.0984	0.0503	0.0490	0.0961	0.0502	0.0520	0.0942	0.0498	0.0483	0.0891
A-BI(rand.)-Parz.	----		----	----	----	----	----	----	----	----	----	----	----
n = 3	A-R	0.0839	0.1914	0.2211	0.1058	0.3617	0.3976	0.1071	0.3618	0.3948	0.1203	0.4272	0.4599
	A-cov(sph)	0.0784	0.1736	0.2076	0.0659	0.1223	0.1479	0.0540	0.0737	0.0940	0.0529	0.0525	0.0676
	A-BI fix	0.0849	0.1989	0.2300	0.1115	0.4092	0.4446	0.1220	0.4568	0.4901	0.1514	0.6198	0.6497
	A-BI(fix)-Parz.	0.0471	0.0483	0.0845	0.0491	0.0507	0.0944	0.0501	0.0489	0.0931	0.0498	0.0488	0.0902
	A-BI rand.	0.0854	0.2014	0.2323	0.1117	0.4104	0.4457	0.1220	0.4570	0.4903	0.1514	0.6201	0.6500
	A-BI(rand.)-Parz.	0.0489	0.0519	0.0893	0.0498	0.0526	0.0970	0.0504	0.0495	0.0940	0.0500	0.0494	0.0904
n = 5	A-R	0.1010	0.3385	0.3573	0.1613	0.6789	0.6958	0.1637	0.6845	0.7013	0.1947	0.7741	0.7868
	A-cov(sph)	0.0750	0.1779	0.1987	0.0595	0.0919	0.1041	0.0532	0.0633	0.0725	0.0531	0.0521	0.0596
	A-BI fix	0.1034	0.3501	0.3694	0.1693	0.7210	0.7372	0.1830	0.7673	0.7815	0.2323	0.8949	0.9030
	A-BI(fix)-Parz.	0.0488	0.0524	0.0902	0.0476	0.0496	0.0917	0.0484	0.0494	0.0876	0.0464	0.0469	0.0872
	A-BI rand.	0.1036	0.3512	0.3699	0.1693	0.7211	0.7374	0.1830	0.7674	0.7815	0.2324	0.8949	0.9030
	A-BI(rand.)-Parz.	0.0498	0.0546	0.0944	0.0481	0.0511	0.0932	0.0486	0.0497	0.0833	0.0465	0.0470	0.0874

Insgesamt führt die Auswertung der Bonituren als stetiges Merkmal insbesondere für den t-Test zu einer guten Einhaltung des nominalen Fehlers. Dabei ist aber zu beachten, dass bei der Simulation für die Prüfglieder identische Effekte vorgegeben wurden. Auf die daraus resultierenden Konsequenzen wird in Abschnitt 5.7 eingegangen.

5.6 Zusammenfassende Bewertung der Ergebnisse für Parametersatz 1, mit räumlichen Kovarianzstrukturen

Die Ergebnisse der Auswertung der Bonituren, die mittels Parametersatz 1 erstellt und mit räumlichen Kovarianzstrukturen simuliert wurden, zeigen, dass die Wirksamkeit der unterschiedlichen Auswertungsmethoden sehr unterschiedlich ist. Da die Bewertung der Wirksamkeit der Auswertungsmethoden verschiedene Aspekte beinhaltet, die in den Abschnitten 5.1 bis 5.5 detailliert anhand der Ergebnisse dargestellt wurden, wird im Folgenden eine kurze zusammenfassende Bewertung vorgenommen. Die Rangverfahren und die Permutationstests liefern für den Kontrast von Prüfglied 1 zu 2 keine Einhaltung des nominalen Fehlers 1. Art für Simulationsansätze mit räumlichen Kovarianzstrukturen. Für die Auswertungen mittels SWM und LMM wird die Einhaltung der nominalen Fehler nur mit bestimmten Modellvarianten innerhalb der Methoden erreicht. Die Tabelle 5.15 gibt einen Überblick über die nutzbaren Auswertungsverfahren und -varianten. Das Schwellenwertmodell mit Linkfunktion „cumlogit“ ist mit Konvergenzproblemen verbunden, die sich vor allem für die räumliche Auswertungsvariante A-cov(sph) zeigt. Trotz der Konvergenzprobleme zeigt es mit der räumlichen Auswertungsvariante eine gute Einhaltung der nominalen Fehler 1. Art für den t-Test des Prüfgliedkontrasts von Prüfglied 1 zu 2 und nur leichte Verzerrungen für den F-Test. Für Ansätze mit mehr als einer Beobachtung kann als Alternative zur räumlichen Auswertungsvariante ein Modell mit zufälligen Parzelleneffekten innerhalb fixer oder zufälliger Blocks gewählt werden. Das Schwellenwertmodell liefert als einzige Auswertungsmethode Schätzungen für die kumulativen Wahrscheinlichkeiten der Boniturnoten. Diese weisen jedoch leichte Verzerrungen auf. Die Auswertung der Bonituren mittels LMM zeigt für den Parametersatz 1 eine sehr gute Einhaltung der nominalen Fehler von t- und F-Test wenn als Modellvariante bei $n = 1$ Beobachtungen ein Blockansatz und bei $n > 1$ ein Ansatz mit zufälligen Parzelleneffekten innerhalb fixer oder zufälliger Blocks gewählt wird. Letzteres trifft ebenfalls auf die Auswertung der Parzellenmittelwerte der Bonituren zu. Diese Aussagen gelten jedoch nur für den Parametersatz 1 der Bonituren, der mit homogenen Prüfgliedeffekten simuliert wurde.

Tab. 5.15: Zusammenfassende Bewertung der Ergebnisse der nutzbaren Auswertungsmethoden und -varianten der Bonituren (Parametersatz 1, homogene Prüfgliedeffekte) für Simulationsansätze mit räumlichen Kovarianzstrukturen

	Auswertungsmethode	Auswertungsvariante	Konvergenzverhalten	Einhaltung der nominalen Fehler 1. Art		Schätzung kumulativer Wahrscheinlichkeiten für Boniturstufen
				t-Test Prüfgliedkontrast PG 1 zu 2	F-Test Globalhypothese	
n = 1	SWM Linkfunktion „cumlogit“	A-cov(sph)	(-)	(+)	(-)	(-)
	LMM für Bonituren	A-cov(sph) A-BI fix / A-BI rand	(+) (+)	(+) (+)	(-) (+)	
n > 1	SWM Linkfunktion „cumlogit“	A-cov(sph)	(- -)	(+)	(- -)	(-)
		A-BI(fix)-Parz. / A-BI(rand.)-Parz.	(-)	(+)	(- -)	(-)
	LMM für Bonituren	A-BI(fix)-Parz. / A-BI(rand.)-Parz.	(+)	(+)	(+)	
	LMM für Parzellenmittelwerte	A-BI fix / A-BI rand.	(+)	(+)	(+)	

(-) leichte, (- -) mittelstarke und (- - -) starke Verzerrungen bzw. Konvergenzprobleme

(+) Einhaltung der Fehler, bzw. 100 %-ige Konvergenz

5.7 Spezielle Vergleiche der Auswertung der Bonituren mittels LMM und SWM

Für die Auswertung der Bonituren und ihrer Parzellenmittelwerte als stetige Variable mittels linearem gemischtem Modell und mittels Schwellenwertmodell wurden ausgewählte Analysen für die alternativen Parametersätze (vgl. Abschnitt 3.3, Anhang 2 und 3) durchgeführt. Die Analysen der alternativen Parametersätze erlauben den Vergleich von Ergebnissen bei Reduzierung der Klassenanzahl, alternativen Parametern der Wahrscheinlichkeiten der Boniturnoten und der Simulation homogener oder heterogener Prüfgliedeffekte. Die folgenden Analysen wurden durchgeführt.

Für Simulationsansätze ohne räumliche Kovarianz:

- Auswertung der Parametersätze 1 bis 4 und 5 bis 8, mit homogenen Prüfgliedeffekten, für $k = 9$ und $k = 4$ Boniturklassen,
- Auswertung der Parametersätze 1a bis 1c und 5a bis 5c, mit heterogenen Prüfgliedeffekten, im Vergleich zu den Parametersätzen 1 und 5 mit homogenen Prüfgliedeffekten, für $k = 9$ und $k = 4$ Boniturklassen.

Für Simulationsansätze mit räumlicher Kovarianz:

- Auswertung des Parametersatzes 1 a, mit heterogenen Prüfgliedeffekten für $k = 9$ Boniturklassen.

5.7.1 Auswertung der Simulationsansätze ohne räumliche Kovarianz für $k = 4$ und $k = 9$ Boniturklassen

Die Tabellen 5.16 und 5.18 zeigen die realisierten Fehler 1. Art des t-Tests für den Prüfgliedkontrast von Behandlung 1 und 2 für die Auswertung der Bonituren als stetige Variable, für $k = 9$ und $k = 4$ Boniturklassen, mit verschiedenen Parametersätzen für die kumulativen Wahrscheinlichkeiten der Boniturklassen. Die Daten wurden ohne räumliche Kovarianz simuliert (Ansatz 0_0_1) und die Auswertung erfolgte mit der Auswertungsvariante A-R, also mit einem linearen Modell ohne zufällige Effekte (LMM). Die Tabellen 5.17 und 5.19 zeigen für diese Simulationsansätze die entsprechenden Ergebnisse im linearen Prädiktor bei Auswertung der Bonituren mittels Schwellenwertmodell (SWM) mit der Linkfunktion „cumlogit“.

Für $k = 9$ Boniturklassen (Tab. 5.16) werden die nominalen Fehler für die Auswertung der Bonituren als stetiges Merkmal unabhängig von den gewählten Parametern für die kumulativen Wahrscheinlichkeiten für alle Beobachtungszahlen pro Parzelle eingehalten. Die Auswertung mittels Schwellenwertmodell dagegen (Tab. 5.17) zeigt für Ansätze mit $n = 1$ oder 3 Beobachtungen je Parzelle fast immer eine Überschätzung des nominalen Fehlers (Ausnahme Parametersatz 2) Mit zunehmender Beobachtungsanzahl je Parzelle reduziert sich diese Überschätzung so dass für $n = 5$ der nominale Fehler in 2 von 4 Fällen eingehalten wird und die Verzerrung des Fehlers für Parametersatz 2 und 4 nur gering ist.

Für $k = 4$ Boniturklassen (Tab. 5.18) zeigt sich für die Auswertung der Bonituren mittels LMM ein ähnliches Bild wie für $k = 9$ Klassen. Hier werden die nominalen Fehler ebenfalls für alle Beobachtungszahlen und Parametersätze eingehalten. Die realisierten Fehler 1. Art des t-Tests der Auswertung der Bonituren mittels Schwellenwertmodell für $k = 4$ Klassen (Tab. 5.19) unterscheiden sich aber deutlich von denjenigen mit $k = 9$ Klassen. Je nach Parametersatz zeigt sich ein unterschiedliches Verhalten der Fehler. Der Parametersatz 7 zeigt als Einziger eine Überschätzung der Fehler für $n = 1$ und 3. Die anderen Parametersätze zeigen für $n = 1$ Unterschätzungen, der Parametersatz 6 unterschätzt auch die Ansätze mit mehr Beobachtungen, die Parametersätze 5 und 8 dagegen realisieren den nominalen Fehler 1. Art ab $n = 3$ Beobachtungen je Parzelle.

Alle bisherigen Analysen wurden anhand von Parametersätzen durchgeführt, die mit homogenen Prüfgliedeffekten simuliert wurden (Parametersätze 1 bis 8, Tab 3.3). Diese Vorgehensweise induziert homogene Restvarianzen für die Prüfglieder. In praktischen Anwendungen werden jedoch heterogene Prüfgliedeffekte und damit in Folge auch heterogene Restvarianzen die Regel sein. Daher werden im Folgenden die Ergebnisse der Auswertungen der Parametersätze mit heterogenen Prüfgliedeffekten (siehe Abschnitt 3.3). besprochen.

Der gewählte räumliche Simulationsansatz ist 0_0_1, also ohne räumliche Korrelation, die Auswertungsvariante A-R. Die gewählten Parameter für die Wahrscheinlichkeiten sind Tabelle 3.2, Abschnitt 3.3 zu entnehmen.

Die Tabellen 5.20 und 5.21 zeigen die realisierten Fehler 1. Art des t-Tests für den Kontrast von Prüfglied 1 zu 2 der Auswertung der Bonituren mittels LMM sowie der Auswertung der Bonituren mittels Schwellenwertmodell mit „cumlogit“ Linkfunktion (auf der Ebene des linearen Prädiktors) für die Varianten der Simulation der Prüfgliedeffekte für die Parametersätze 1 und 5 für $k = 9$ und $k = 4$ Boniturklassen.

Wenn die Bonituren mit $k = 9$ Klassen (Tab. 5.20) mittels linearem Modell (LMM) ausgewertet werden, wird für den Parametersatz 1 (homogene Prüfgliedeffekte), unabhängig von der Anzahl der Beobachtungen, der nominale Fehler 1. Art eingehalten. Werden die Bonituren in diesem Ansatz mittels Schwellenwertmodell ausgewertet, so wird der Fehler bei 1 und 3 Beobachtungen leicht überschätzt. Für die Parametersätze 1a und 1b (heterogene Prüfgliedeffekte) zeigen sowohl die Auswertung der Bonituren mittels LMM als auch mittels Schwellenwertmodell Überschätzungen des nominalen Fehlers. Für $n = 1$ Beobachtungen liegt die Überschätzung für das Schwellenwertmodell über derjenigen des linearen Modells, für $n = 3$ und 5 dagegen ist das Schwellenwertmodell dem linearen Modell überlegen. Dieser Trend verstärkt sich für 9 Beobachtungen pro Parzelle (Ergebnisse hier nicht gezeigt). Für Parametersatz 1c (heterogene Prüfgliedeffekte) ergibt sich ein Verhalten der Fehler wie für den Parametersatz 1 mit homogenen Prüfgliedeffekten.

Tab. 5.16: Realisierte Fehler 1. Art des t-Tests der Prüfglieddiffrenz PG 1 zu 2, Auswertung der Bonituren (k = 9) als stetige Variable mittels LMM, 10000 Simulationsläufe, für verschiedene Parametersätze der kumulativen Wahrscheinlichkeiten der Boniturklassen (fett: Einhaltung des nominalen Fehlers)

Simulierter räumlicher Kovarianzansatz		Bonituren k = 9			
		Parametersatz 1	Parametersatz 2	Parametersatz 3	Parametersatz 4
		t-Test	t-Test	t-Test	t-Test
n = 1	0_0_1 A-R (LMM)	0.0479	0.0517	0.0470	0.0529
n = 3		0.0488	0.0490	0.0502	0.0504
n = 5		0.0478	0.0488	0.0480	0.0505

Tab. 5.17: Realisierte Fehler 1. Art des t-Tests der Prüfglieddiffrenz PG 1 zu 2 für den linearen Prädiktor, Auswertung der Bonituren (k = 9) mittels Schwellenwertmodell (Linkfunktion „cumlogit“), 10000 Simulationsläufe, für verschiedene Parametersätze der kumulativen Wahrscheinlichkeiten der Boniturklassen (fett: Einhaltung des nominalen Fehlers)

Simulierter räumlicher Kovarianzansatz		Bonituren k = 9			
		Parametersatz 1	Parametersatz 2	Parametersatz 3	Parametersatz 4
		t-Test	t-Test	t-Test	t-Test
n = 1	0_0_1 A-R (SWM)	0.0785	0.0835	0.0479	0.0607
n = 3		0.0595	0.0604	0.0548	0.0590
n = 5		0.0542	0.0544	0.0499	0.0551

Tab. 5.18: Realisierte Fehler 1. Art des t-Tests der Prüfglieddifferenz PG 1 zu 2, Auswertung der Bonituren (k = 4) als stetige Variable mittels LMM, 10000 Simulationsläufe, für verschiedene Parametersätze der kumulativen Wahrscheinlichkeiten der Boniturklassen (fett: Einhaltung des nominalen Fehlers)

Simulierter räumlicher Kovarianzansatz		Bonituren k = 4			
		Parametersatz 5	Parametersatz 6	Parametersatz 7	Parametersatz 8
		t-Test	t-Test	t-Test	t-Test
n = 1	0_0_1 A-R (LMM)	0.0478	0.0480	0.0486	0.0537
n = 3		0.0497	0.0481	0.0522	0.0503
n = 5		0.0456	0.0494	0.0490	0.0534

Tab. 5.19: Realisierte Fehler 1. Art des t-Tests der Prüfglieddifferenz PG 1 zu 2 für den linearen Prädiktor, Auswertung der Bonituren (k = 4) mittels Schwellenwertmodell (Linkfunktion „cumlogit“), 10000 Simulationsläufe, für verschiedene Parametersätze der kumulativen Wahrscheinlichkeiten der Boniturklassen (fett: Einhaltung des nominalen Fehlers)

Simulierter räumlicher Kovarianzansatz		Bonituren k = 4			
		Parametersatz 5	Parametersatz 6	Parametersatz 7	Parametersatz 8
		t-Test	t-Test	t-Test	t-Test
n = 1	0_0_1 A-R (SWM)	0.0403	0.0060	0.0602	0.0120
n = 3		0.0539	0.0296	0.0583	0.0506
n = 5		0.0513	0.0425	0.0531	0.0517

Tab. 5.20: Realisierte Fehler 1. Art t-Test der Prüfglieddiffereenz PG 1 zu 2, Auswertung der Bonituren (9 Noten) als stetige Variable mittels LMM und mit Schwellenwertmodell (Linkfunktion „cumlogit“, Fehler für den linearen Prädiktor), 10000 Simulationsläufe, Parametersätze der Wahrscheinlichkeiten mit unterschiedlicher Simulation der Effekte, entsprechend Abschnitt 3.3 (fett: Einhaltung des nominalen Fehlers)

	Simulierter räumlicher Kovarianzansatz	Bonituren k = 9		Bonituren k = 9		Bonituren k = 9		Bonituren k = 9	
		Parametersatz 1		Parametersatz 1a		Parametersatz 1b		Parametersatz 1c	
	Auswertungsvariante	LMM	SWM	LMM	SWM	LMM	SWM	LMM	SWM
n = 1	0_0_1 A-R	0.0479	0.0785	0.0689	0.0878	0.0616	0.0883	0.0532	0.0861
n = 3		0.0488	0.0595	0.0697	0.0693	0.0623	0.0617	0.0533	0.0587
n = 5		0.0478	0.0542	0.0640	0.0628	0.0595	0.0574	0.0511	0.0482

Tab. 5.21: Realisierte Fehler 1. Art t-Test der Prüfglieddiffereenz PG 1 zu 2, Auswertung der Bonituren (4 Noten) als stetige Variable mittels LMM und mit Schwellenwertmodell (Linkfunktion „cumlogit“, Fehler für den linearen Prädiktor), 10000 Simulationsläufe, Parametersätze der Wahrscheinlichkeiten mit unterschiedlicher Simulation der Effekte, entsprechend Abschnitt 3.3 (fett: Einhaltung des nominalen Fehlers)

	Simulierter räumlicher Kovarianzansatz	Bonituren k = 4		Bonituren k = 4		Bonituren k = 4		Bonituren k = 4	
		Parametersatz 5		Parametersatz 5a		Parametersatz 5b		Parametersatz 5c	
	Auswertungsvariante	LMM	SWM	LMM	SWM	LMM	SWM	LMM	SWM
n = 1	0_0_1 A-R	0.0478	0.0403	0.0337	0.0140	0.0528	0.0396	0.0539	0.0414
n = 3		0.0497	0.0539	0.0292	0.0395	0.0511	0.0551	0.0525	0.0552
n = 5		0.0456	0.0513	0.0314	0.0471	0.0512	0.0520	0.0526	0.0524

Für $k = 4$ Boniturklassen (Tab. 5.21) zeigt die Auswertung der Bonituren mittels LMM für den Parametersatz 5 (homogene Prüfgliedeffekte) die Einhaltung des Fehlers 1. Art bei $n = 1$ und 3 Beobachtungen pro Parzelle und bei $n = 5$ Beobachtungen eine sehr geringe Unterschätzung. Das Schwellenwertmodell zeigt für $n = 1$ mit nur 4 Boniturklassen keine Einhaltung des nominalen Fehlers, bei den höheren Stichprobenumfängen allerdings schon. Für den Parametersatz 5a (heterogene Prüfgliedeffekte) zeigt sich unabhängig von der Anzahl Beobachtungen eine deutliche Unterschätzung des nominalen Fehlers 1. Art, wenn die Bonituren mittels LMM ausgewertet werden. Das Schwellenwertmodell dagegen zeigt ab einer ausreichenden Stichprobengröße ($n = 5$) eine Einhaltung des nominalen Fehlers. Für $n = 1$ ist jedoch die Auswertung der Bonituren mittels LMM besser als die Auswertung mittels Schwellenwertmodell. Für die Parametersätze 5b und 5c (heterogene Prüfgliedeffekte) liefert die Auswertung der Bonituren mittels LMM unabhängig von der Beobachtungsanzahl den nominalen Fehler 1. Art. Die Auswertung der Bonituren mittels Schwellenwertmodell dagegen leistet dies erst bei $n = 5$ Beobachtungen.

5.7.2 Auswertung der Simulationsansätze mit räumlicher Kovarianz und Simulation heterogener Prüfgliedeffekte

Die Tabellen 5.22 bis 5.24 zeigen die Ergebnisse für Parametersatz 1a (heterogene Prüfgliedeffekte) mit ausgewählten Auswertungsmethoden und -varianten bei gleichzeitiger Simulation räumlicher Korrelation. Es wurden jeweils die Auswertungsvarianten gewählt, die für den Parametersatz 1 die besten Ergebnisse zeigten.

Die Auswertung der Bonituren oder ihrer Parzellenmittelwerte mittels LMM zeigen für den Parametersatz 1a unabhängig von der Modellvariante und von der Beobachtungszahl leichte Verzerrungen der nominalen Fehler 1. Art für den t-Test. Die Fehler des F-Tests können nicht mehr beurteilt werden, da durch die Simulation heterogener Prüfgliedeffekte die Nullhypothese für den Globaltest nicht mehr gültig ist.

Werden die Bonituren mit dem Schwellenwertmodell ausgewertet (Tabelle 5.24), so zeigen sich für den Parametersatz 1a mit heterogenen Prüfgliedeffekten teils extreme Konvergenzverluste für die Modelle mit zufälligen Parzelleneffekten innerhalb der Blocks. Diese Modelle zeigten für den Parametersatz 1 mit homogenen Prüfgliedeffekten keine Konvergenzprobleme. Der Vergleich der Einhaltung der nominalen Fehler 1. Art der unterschiedlichen Modellvarianten ist daher nicht möglich. Die räumliche Auswertungsvariante zeigt für $n = 1$ nur geringe Konvergenzverluste und relativ kleine Verzerrungen für den nominalen Fehler des t-Tests.

Werden die Bonituren mittels LMM ausgewertet, kann zusätzlich zur Modellierung der räumlichen Kovarianz die Modellierung heterogener Restvarianzen für die Prüfglieder angewiesen werden (vgl. Abschnitt 4.1), die Ergebnisse dieser Auswertungsvarianten sind in Tabelle 5.25 und 5.26 dargestellt.

Die Anweisung heterogener Restvarianzen im Auswertungsmodell hat starke Auswirkungen auf das Konvergenzverhalten, was wiederum den Vergleich der Fehler erschwert. Für die Auswertung der Bonituren mittels LMM zeigen die Ansätze A-BI(fix)-Parz. und A-BI(rand.)-Parz. gute Konvergenz. Die Fehler des t-Tests werden

jedoch leicht verzerrt geschätzt. Die Größe der Verzerrung der Fehler für den t-Test wird durch die Auswahl heterogener Restvarianzen im Auswertungsmodell im Gegensatz zur Auswahl homogener Restvarianzen reduziert. Die Blockansätze für $n = 1$ Beobachtung weisen starke Konvergenzprobleme auf. Die Auswertungen der Parzellenmittelwerte mit den Blockansätzen zeigen bei Anweisung heterogener Restvarianzen im Auswertungsmodell ebenfalls starke Konvergenzprobleme und Verzerrungen der nominalen Fehler 1. Art für den t-Test.

Tab. 5.22: Anteil konvergenter Simulationsläufe und realisierte Fehler 1. Art des t-Tests der Prüfglieddiffereenz PG 1 zu 2, Auswertung der Bonituren (Parametersatz 1a, heterogene Prüfgliedeffekte) als stetige Variable mittels LMM, mit Auswahl homogener Restvarianzen im Auswertungsmodell, 10000 Simulationsläufe

Simulierter räumlicher Kovarianzansatz		1_10_1		1_20_1		1_30_1		2_40_1	
Auswertung		Konvergenz	t-Test	Konvergenz	t-Test	Konvergenz	t-Test	Konvergenz	t-Test
n = 1	A-BI fix	1.000	0.0681	1.000	0.0700	1.000	0.0719	1.000	0.0713
	A-BI rand.	1.000	0.0702	1.000	0.0718	1.000	0.0726	1.000	0.0715
n = 3	A-BI (fix)-Parz.	1.000	0.0682	1.000	0.0709	1.000	0.0694	1.000	0.0699
	A-BI (rand.)-Parz.	1.000	0.0703	1.000	0.0722	1.000	0.0696	1.000	0.0702
n = 5	A-BI (fix)-Parz.	1.000	0.0661	1.000	0.0707	1.000	0.0670	1.000	0.0688
	A-BI (rand.)-Parz.	1.000	0.0673	1.000	0.0713	1.000	0.0672	1.000	0.0689

Tab. 5.23: Anteil konvergenter Simulationsläufe und realisierte Fehler 1. Art des t-Tests der Prüfglieddiffereenz PG 1 zu 2, Auswertung der Parzellenmittelwerte der Bonituren (Parametersatz 1a, heterogene Prüfgliedeffekte) als stetige Variable mittels LMM, mit Auswahl homogener Restvarianzen im Auswertungsmodell, 10000 Simulationsläufe

Simulierter räumlicher Kovarianzansatz		1_10_1		1_20_1		1_30_1		2_40_1	
Auswertung		Konvergenz	t-Test	Konvergenz	t-Test	Konvergenz	t-Test	Konvergenz	t-Test
n = 3	A-BI fix	1.000	0.0691	1.000	0.0709	1.000	0.0694	1.000	0.0699
	A-BI rand.	1.000	0.0714	1.000	0.0722	1.000	0.0696	1.000	0.0702
n = 5	A-BI fix	1.000	0.0664	1.000	0.0707	1.000	0.0670	1.000	0.0688
	A-BI rand.	1.000	0.0675	1.000	0.0713	1.000	0.0672	1.000	0.0689

Tab. 5.24: Anteil konvergenter Simulationsläufe und realisierte Fehler 1. Art des t-Tests der Prüfglieddifferenz PG 1 zu 2, Auswertung der Bonituren (Parametersatz 1a, heterogene Prüfgliedeffekte) mittels Schwellenwertmodell (Linkfunktion „cumlogit“), 10000 Simulationsläufe für die Auswertungen A-BI (fix)-Parz und A-BI (rand.)-Parz., 1000 Simulationsläufe für die Auswertung A-cov(sph) (fett: Einhaltung des nominalen Fehlers)

Simulierter räumlicher Kovarianzansatz		1_10_1		1_20_1		1_30_1		2_40_1	
Auswertung		Konvergenz	t-Test	Konvergenz	t-Test	Konvergenz	t-Test	Konvergenz	t-Test
n = 1	A-cov(sph) ¹⁾	0.9190	0.0446	0.9430	0.0467	0.9710	0.0566	0.9750	0.0441
n = 3	A-cov(sph) ¹⁾	0.6140	0.0586	0.4560	0.0307	0.5070	0.0572	0.6220	0.0338
	A-BI (fix)-Parz. ²⁾	0.2045	0.0450	0.2441	0.0492	0.2554	0.0415	0.2488	0.0410
	A-BI (rand.)-Parz. ²⁾	0.1561	0.0480	0.1537	0.0527	0.1958	0.0419	0.2225	0.0458
n = 5	A-cov(sph) ¹⁾	0.7700	0.0584	0.8860	0.0474	0.7540	0.0424	0.6490	0.0385
	A-BI (fix)-Parz. ²⁾	0.8278	0.0518	0.6029	0.0496	0.6424	0.0472	0.6883	0.0484
	A-BI (rand.)-Parz. ²⁾	0.4767	0.0539	0.4628	0.0538	0.5114	0.0508	0.6087	0.0465

¹⁾ 1000 Simulationsläufe aufgrund sehr langer Laufzeiten

²⁾ 10000 Simulationsläufe

Tab. 5.25: Anteil konvergenter Simulationsläufe und realisierte Fehler 1. Art des t-Tests der Prüfglieddiffrenz PG 1 zu 2, Auswertung der Bonituren (Parametersatz 1a, heterogene Prüfgliedeffekte) als stetige Variable mittels LMM, mit Auswahl heterogener Restvarianzen im Auswertungsmodell, 1000 Simulationsläufe (fett: Einhaltung des nominalen Fehlers)

Simulierter räumlicher Kovarianzansatz		1_10_1		1_20_1		1_30_1		2_40_1	
Auswertung		Konvergenz	t-Test	Konvergenz	t-Test	Konvergenz	t-Test	Konvergenz	t-Test
n = 1	A-BI fix	0.3349	0.0403	0.3297	0.0394	0.3367	0.0327	0.3206	0.0371
	A-BI rand.	0.6899	0.0430	0.6138	0.0412	0.5328	0.0419	0.4331	0.0379
n = 3	A-BI (fix)-Parz.	0.9997	0.0517	0.9980	0.0568	0.9971	0.0587	0.9946	0.0608
	A-BI (rand.)-Parz.	0.9996	0.0533	0.9976	0.0562	0.9967	0.0585	0.9924	0.0607
n = 5	A-BI (fix)-Parz.	1.000	0.0536	0.9998	0.0605	0.9987	0.0593	0.9920	0.0611
	A-BI (rand.)-Parz.	1.000	0.0543	0.9989	0.0608	0.9961	0.0595	0.9833	0.0613

Tab. 5.26: Anteil konvergenter Simulationsläufe und realisierte Fehler 1. Art des t-Tests der Prüfglieddiffrenz PG 1 zu 2, Auswertung der Parzellenmittelwerte der Bonituren (Parametersatz 1a, heterogene Prüfgliedeffekte) als stetige Variable mittels LMM, mit Auswahl heterogener Restvarianzen im Auswertungsmodell, 1000 Simulationsläufe

Simulierter räumlicher Kovarianzansatz		1_10_1		1_20_1		1_30_1		2_40_1	
Auswertung		Konvergenz	t-Test	Konvergenz	t-Test	Konvergenz	t-Test	Konvergenz	t-Test
n = 3	A-BI fix	0.3728	0.0328	0.3694	0.0363	0.3622	0.0315	0.3595	0.0323
	A-BI rand.	0.7443	0.0410	0.6364	0.0383	0.5328	0.0396	0.4267	0.0401
n = 5	A-BI fix	0.3631	0.0355	0.3665	0.0388	0.3691	0.0339	0.3689	0.0336
	A-BI rand.	0.7286	0.0361	0.6079	0.0421	0.5113	0.0411	0.4230	0.0345

6 Diskussion und Schlussfolgerungen

Die folgenden Ausführungen werden hinsichtlich der Konsequenzen für die biometrische Auswertungspraxis gegliedert. Aus Sicht eines Anwenders muss eine Auswertungsmethode Konvergenz zeigen, erwartungstreue Schätzungen mit minimaler Varianz liefern und bei der Durchführung von Tests für die statistische Inferenz die Einhaltung des Fehlers 1. Art garantieren, das bedeutet, dass das Konfidenzintervall des empirischen Fehlers 1. Art den gegebenen nominalen Fehler überdeckt. Aus Gründen der Übersichtlichkeit wird die Diskussion zur Einhaltung der Fehler 1. Art untergliedert in die Auswertung von Simulationsansätzen ohne und mit räumlicher Kovarianz. Die Diskussion der Simulationsansätze ohne räumliche Kovarianz dient zunächst der grundsätzlichen Einschätzung der unterschiedlichen Methoden zur Auswertung von Bonituren. Darauf aufbauend erfolgt die Besprechung der Simulationsansätze mit räumlicher Kovarianz. Sie dient der Beurteilung der innerhalb der Auswertungsmethoden möglichen Modellvarianten zur Berücksichtigung vorhandener räumlicher Korrelation.

6.1 Konvergenzverhalten

Aus Anwendersicht ist eine Auswertungsmethode nur dann sinnvoll, wenn das eingesetzte Schätzverfahren Konvergenz zeigt. Die vorliegenden Untersuchungen zeigen, dass bei Nutzung des Schwellenwertmodells zur Auswertung der Bonituren in Abhängigkeit von der gewählten Auswertungsvariante teilweise erhebliche Konvergenzprobleme auftreten. Diese sind abhängig von der Komplexität der gewählten Auswertungsvariante, den bei der Simulation verwendeten Parametern und dem Stichprobenumfang. Für die Auswertung von Datensätzen mit homogenen Prüfgliedeffekten zeigt sich der größte Konvergenzverlust für die Auswertungsvarianten des SWM mit räumlicher Kovarianz. Der Verlust an konvergenten Simulationsläufen ist bei Nutzung der Linkfunktion „cumprobit“ deutlich höher als bei Nutzung der Linkfunktion „cumlogit“. Für die Auswertung von Datensätzen mit mehr als einer Beobachtung je Parzelle und heterogenen Prüfgliedeffekten zeigen sich bei Nutzung des Schwellenwertmodells mit der Linkfunktion „cumlogit“ auch für die übrigen Modellvarianten erhebliche Konvergenzprobleme.

Die Konvergenzprobleme sind auf die für das Schwellenwertmodell angewendete „Restricted Pseudo-Likelihood“ Methode (RSPL) zur Schätzung der Modellparameter zurückzuführen. Diese Methode hat insofern Nachteile, als sie „doppelt iterativ“ ist. Durch Taylorreihenentwicklung werden Pseudodaten erzeugt, von denen angenommen wird, dass sie zumindest näherungsweise normalverteilt sind. Dann erfolgt für diese Pseudodaten die Iteration eines LMM, und anschließend wird das SWM, basierend auf der aktuellen Schätzung der Kovarianzparameter, wiederum durch das LMM approximiert. Die Schätzungen der Kovarianzparameter sind jedoch für die räumliche Auswertungsvariante A-cov(sph) extrem verzerrt. Weiterhin ist die Schätzung der G-Matrix aufgrund der zahlreichen zufälligen Effekte (Anzahl zufälliger Effekte = Gesamtanzahl der Beobachtungen) ebenfalls erschwert. Das deutlich schlechtere Konvergenzverhalten des Schwellenwertmodells bei Nutzung der Linkfunktion „cumprobit“ im Gegensatz zur Nutzung der Linkfunktion „cumlogit“ ist darauf zurückzuführen, dass für die Linkfunktion „cumprobit“ das Integral über die Dichte der Normal-

verteilung berechnet werden muss. Dies ist mit der Linkfunktion „cumlogit“ nicht notwendig, wodurch der Lösungsalgorithmus numerisch stabiler wird.

Für Datensätze mit heterogenen Prüfgliedeffekten zeigen die Auswertungen der Parzellenmittelwerte der Bonituren mittels LMM ebenfalls erhebliche Konvergenzprobleme, wenn im Auswertungsmodell heterogene Restvarianzen angewiesen werden. Bei Auswertung der Bonituren mittels LMM ist dies nur für Ansätze mit einer Beobachtung je Parzelle der Fall, sind mehrer Beobachtungen je Parzelle vorhanden, liefert diese Auswertungsmethode fast hundertprozentige Konvergenz. Dies ist auf die zunehmende Modellkomplexität durch die Anweisung heterogener Restvarianzen und die geringe Stichprobengröße zur Schätzung der Restvarianzen zurückzuführen.

Bei der Beurteilung des Konvergenzverhaltens muss beachtet werden, dass dieses zwar primär von der gewählten Auswertungsmethode abhängt, sekundär jedoch auch von den genutzten Iterationsverfahren und Konvergenzkriterien. In dieser Arbeit wurden für alle Auswertungen mittels Schwellenwertmodell die Standardeinstellungen der Prozedur GLIMMIX genutzt. Als Iterationsverfahren wird standardmäßig ein Quasi-Newton Algorithmus mit mehreren Konvergenzkriterien eingesetzt, welche SAS intern simultan berechnet und beurteilt. Die maximale Anzahl der Iterationsschritte wurde auf 150 festgelegt. Für diesen Algorithmus wird das „absolute Gradient“-Kriterium (ABSGCONV), ein „Relativfunktions“-Kriterium (FCONV) und das „relative Gradient“-Kriterium (GCOV) genutzt. Eine Berechnung wird als konvergent definiert, wenn eines dieser Kriterien einen bestimmten Grenzwert unterschreitet (SAS 9.2, online Dokumentation, Shared Concepts and Topics, Nonlinear Optimization: The NLOPTIONS Statement, Choosing an Optimization Algorithm). Das Konvergenzverhalten einer Auswertungsmethode kann durch Veränderung der Standardeinstellungen bezüglich der Optimierungsmethode, des genutzten Konvergenzkriteriums und seines Grenzwertes sowie durch Erhöhung der Iterationsschritte beeinflusst werden. Eine verallgemeinerte Aussage zur Beurteilung des Konvergenzverhaltens kann deshalb nicht getroffen werden. Für den praktischen Anwender wird eine Veränderung der Standardeinstellungen jedoch im Allgemeinen eine zu große Herausforderung darstellen.

Die vorliegenden Untersuchungen zeigen, dass sich aus Sicht des Konvergenzverhaltens bestimmte Auswertungsmethoden bzw. -varianten für die Auswertungspraxis verbieten. Liegt ein Datensatz mit heterogenen Prüfgliedeffekten vor, wie es in der Auswertungspraxis die Regel sein wird und wird für den Anteil konvergenter Läufe beispielsweise ein Grenzwert von 80 % festgelegt, so verbietet sich bei Vorliegen einer räumlichen Kovarianz die Anwendung des SWM mit den Modellvarianten A-cov(sph), A-BI(fix)-Parz. und A-BI(rand.)-Parz., wenn die Anzahl Beobachtungen je Parzelle $n > 1$ ist. Bei Nutzung des LMM mit Anweisung heterogener Restvarianzen im Auswertungsmodell verbieten sich die Modellvarianten A-BI fix und A-BI rand., für die Anwendung des LMM auf die Bonituren bei $n = 1$ und für die Anwendung des LMM auf die Parzellenmittelwerte der Bonituren bei $n > 1$ Beobachtungen.

6.2 Erwartungstreue und Varianz der Parameterschätzungen

Schätzung der Prüfgliedkontraste

Für die Auswertung von Bonituren zählen zu den zu schätzenden Parametern neben den fixen Effekten und ihren Kontrasten auch die kumulativen Wahrscheinlichkeiten der Boniturnoten, wenn zur Auswertung das Schwellenwertmodell (McCullagh, 1980) genutzt wird. Die Erwartungstreue der Schätzungen der fixen Parameter und ihrer Kontraste kann jedoch nicht für alle in dieser Arbeit angewendeten Auswertungsmethoden beurteilt werden. Die nichtparametrischen Verfahren von Brunner et al. (2001; 2002) schätzen relative Effekte für die Prüfglieder, und die Permutationsverfahren von Schumacher et al. (1995 a, b; 1997) liefern keine Parameterschätzungen, sondern berechnen ausschließlich die Teststatistik für die statistischen Tests.

Parameterschätzungen für die fixen Effekte werden daher nur erhalten, wenn die Bonituren mittels Schwellenwertmodell ausgewertet werden oder wenn die Bonituren bzw. ihre Parzellenmittelwerte unter der Annahme der approximativen Normalverteilung als stetige Variable mittels LMM ausgewertet werden. Die Beurteilung der Parameterschätzungen der fixen Effekte erfolgt in dieser Arbeit anhand des Prüfgliedkontrasts von Prüfglied 1 und 2.

Der Bias der Schätzung des Prüfgliedkontrasts für die Auswertungsmethoden der Bonituren für Simulationsansätze mit räumlicher Kovarianz ist im Gegensatz zur Auswertung der latenten Variablen leicht erhöht, liegt jedoch für fast alle Auswertungsmethoden und -varianten noch innerhalb des Bereichs der Erwartungstreue (vgl. Abschnitt 5.3). Nur für die stetige Auswertung der Bonituren mittels LMM wird der Prüfgliedkontrast mit der Auswertungsvariante A-cov(sph) für Simulationsansätze mit stärkerer räumlicher Kovarianz und $n = 3$ Beobachtungen pro Parzelle leicht verzerrt geschätzt. Die Auswertung der Bonituren mittels Schwellenwertmodell liefert eine erwartungstreue Schätzung des Prüfgliedkontrasts auf der Skala des linearen Prädiktors.

Innerhalb der Auswertungsmethoden ist der Bias für die Auswertungsvariante A-cov(sph) meist am höchsten und in der komplexen Kovarianzstruktur dieser Auswertungsvariante begründet. Besonders deutlich zeigt sich dies für die Auswertung mittels Schwellenwertmodell, da hier außer den Prüfgliedeffekten zusätzlich die Stützungs-punkte geschätzt werden müssen.

Unterschiede der Varianz der Parameterschätzungen zwischen den Auswertungsmethoden und -verfahren können wegen der Nutzung der Simulation in dieser Arbeit bei Nutzung der beobachteten Standardabweichung (STD_{beo}) beurteilt werden.

Für alle Auswertungsmethoden und -varianten der Bonituren sinkt erwartungsgemäß die beobachtete Standardabweichung bei Erhöhung der Anzahl der Beobachtungen pro Parzelle. Bei nichträumlichen Auswertungsvarianten steigt die STD_{beo} mit zunehmender Stärke der räumlichen Korrelation für Ansätze mit $n > 1$ Beobachtungen. Diese Aussage gilt auch für die latente Variable.

Die beobachtete Standardabweichung ist in zweifacher Weise abhängig von der Anzahl der Beobachtungen pro Parzelle. Erstens führt eine Erhöhung der Stichproben-

größe grundsätzlich zu einer Verringerung der Varianz einer Parameterschätzung, wodurch begründet ist, dass für alle Auswertungsvarianten die beobachtete Standardabweichung bei zunehmender Beobachtungsanzahl sinkt. Zweitens drückt sie aus, in wie weit die simulierte räumliche Kovarianz bei der Schätzung der fixen Effekte berücksichtigt wird. Wird die Korrelation bei der Schätzung der Effekte berücksichtigt, verringert sich die beobachtete Standardabweichung. Deshalb ist für Ansätze mit $n > 1$ die STD_{beo} der Auswertungsvarianten A-Cov(sph), A-Block(fix)-Parz. und A-Block(rand.)-Parz. gegenüber den anderen Auswertungsvarianten erniedrigt. Dies ist umso deutlicher, je höher die Anzahl der Beobachtungen je Parzelle und je stärker die räumliche Korrelation. Ausnahme ist die Auswertung der Bonituren als stetiges Merkmal mittels LMM.

Die Beurteilung der Differenz von beobachteter Standardabweichung (STD_{beo}) und mittlerer geschätzter Standardabweichung der unterschiedlichen Auswertungsmethoden und -varianten erfolgt in Abschnitt 6.3 innerhalb der Diskussion zur Einhaltung der Fehler 1. Art für die Auswertung der Simulationsansätze mit räumlicher Kovarianz.

Alle genutzten Auswertungsmethoden für Bonituren liefern also unabhängig von der Modellvariante zur Berücksichtigung der räumlichen Kovarianz erwartungstreue Schätzungen für den Prüfgliedkontrast. Die beobachtete Standardabweichung der unterschiedlichen Varianten ist jedoch für die nichträumlichen Auswertungsvarianten groß. Deshalb ist aus dieser Sicht die Variante A-cov(sph) zu bevorzugen. Für das Schwellenwertmodell stellen auch die Varianten A-Block(fix)-Parz. und A-Block(rand.)-Parz. eine Auswertungsalternative dar.

Schätzung der kumulativen Wahrscheinlichkeiten der Boniturnoten

Wird das Schwellenwertmodell zur Auswertung der Bonituren (Parametersatz 1, homogene Prüfgliedeffekte) genutzt, können aus den geschätzten Parametern direkt Schätzungen für die kumulativen Wahrscheinlichkeiten der Boniturnoten abgeleitet werden (vgl. Abschnitt 5.4). Die Ergebnisse der Simulationsansätze ohne und mit räumlicher Kovarianz bei Nutzung der Auswertungsvariante A-R (vgl. Anhang 6, Tabelle A6-1) zeigen, dass für $n = 1, 3$ und 5 Beobachtungen pro Parzelle unabhängig von der simulierten räumlichen Kovarianz erwartungstreue Schätzungen für die Wahrscheinlichkeiten der Boniturnoten erhalten werden. Alle anderen Auswertungsvarianten für die Simulationsansätze mit räumlicher Kovarianz liefern hingegen formal keine erwartungstreuen Schätzungen. Der auftretende Bias ist jedoch klein und ist nur für die unteren Boniturnoten von Bedeutung, da hier sehr kleine Parameter geschätzt wurden, wodurch sich der auftretende Bias relativ stärker auswirkt.

Die Konfidenzintervalle der kumulativen Wahrscheinlichkeiten der Auswertungsvariante A-R für die Simulationsansätze ohne räumliche Kovarianz sind für Ansätze mit $n < 5$ Beobachtungen relativ breit. Das bedeutet, dass für Ansätze mit geringen Beobachtungszahlen hohe Standardabweichungen für die kumulativen Wahrscheinlichkeiten geschätzt werden. Die Überdeckungswahrscheinlichkeiten der kumulativen Wahrscheinlichkeiten erreichen nicht das vorgegebene Konfidenzniveau. Die Schätzung der kumulativen Wahrscheinlichkeiten entspricht zwar im Mittel aller Läufe dem wah-

ren Parameter, für die einzelnen Läufe treten jedoch extreme Unter- oder Überschätzungen auf. Dies liegt wiederum in der unzureichenden Beobachtungsanzahl begründet, was deutlich aus der mit steigender Beobachtungsanzahl auftretenden Annäherung der empirischen Überdeckungswahrscheinlichkeit an das Konfidenzniveau hervorgeht. Für die Konfidenzintervallberechnung im linearen Prädiktor und anschließende Rücktransformation auf die Boniturskala (KI_{logit} Berechnungsvariante, siehe Abschnitt 2.2.4) liegen die Überdeckungswahrscheinlichkeiten über denjenigen der Berechnungsvariante KI. Dies ist darauf zurückzuführen, dass bei dieser Berechnungsvariante gesichert ist, dass die Intervallgrenzen zwischen 0 und 1 liegen (vgl. Abschnitt 2.2.4), wodurch die Auswirkung von extremen Ausreißern abgeschwächt wird. Die vorliegenden Untersuchungen zeigen, dass das Konfidenzniveau nicht unabhängig von n eingehalten wird und für nichträumliche Simulationsansätze mindestens $n = 5$ Beobachtungen je Parzelle vorhanden sein sollten.

Wird eine räumliche Kovarianz simuliert und mit dem räumlichen Modell A-cov(sph) ausgewertet, kommt es zur extremen Verbreiterung der Intervallbreiten, da die geschätzten Standardabweichungen der Boniturnoten extrem überschätzt werden. Für die Berechnungsvariante KI_{logit} ist die Verbreiterung der Intervalle geringer, da die Intervallgrenzen auf Werte zwischen 0 und 1 beschränkt sind. Das nominale Konfidenzniveau wird mit keiner Berechnungsvariante für die Konfidenzintervalle eingehalten.

Aus praktischer Sicht sind die Schätzungen der Wahrscheinlichkeiten der Boniturnoten für den Nutzer zwar anschaulich und gut interpretierbar, es ist jedoch aufgrund der sehr breiten Konfidenzintervalle und der schlechten Einhaltung der Konfidenzniveaus, die schon für den nichträumlichen Simulationsansatz auftreten, davon auszugehen, dass die Schätzungen in praktischen Anwendungen nicht erwartungstreu sind.

Schätzung der Varianzkomponenten

Die Schätzung der räumlichen Varianzkomponenten liefert für die Auswertung der latenten Variablen mittels LMM und für die Auswertung der Bonituren mittels SWM extrem verzerrte Ergebnisse. Obwohl die Schätzung der Varianzkomponenten nicht das Hauptziel der Auswertung ist, soll hier auf einige besondere Ergebnisse eingegangen werden, da sie auf den empirischen Fehler 1. Art teilweise starken Einfluss nehmen.

Die Auswertungsvariante A-cov(sph) der Simulationsansätze mit räumlicher Kovarianz zeigt sowohl für die Auswertung der Bonituren mit dem Schwellenwertmodell als auch für die Auswertung der latenten Variablen mittels LMM, dass nicht in jedem Simulationslauf Ergebnisse für die Schätzungen der Kovarianzparameter und ihrer Varianz erhalten werden oder dass nicht spezifische Werte geschätzt werden (siehe unten). Dieser Fall tritt ein wenn ein Simulationslauf mit den gewählten Standardeinstellungen nicht konvergiert (siehe Abschnitt 6.1) oder wenn die G-Matrix nicht positiv definit ist.

Konvergiert ein Lauf nicht, so leistet er keinen Beitrag zum Fehler 1. Art. Ist die G-Matrix nicht positiv definit, so reduzieren sich die Schätzungen für die Varianzpara-

meter auf die Restvarianz, und es werden für „Partieller Sill“ und „Range“ bestimmte, feste Werte geschätzt.

Hierbei treten bei Auswertung der Bonituren mittels Schwellenwertmodell zwei Fälle auf:

- a) Partieller Sill = 0 oder Partieller Sill ≈ 0 , für n = 1: Range ≈ 160 , für n = 3: Range ≈ 480 , für n = 5: Range ≈ 800 (Erklärung der Werte siehe nächste Seite), **ohne Schätzung der Standardfehler von Partieller Sill und Range**, dann wird der Nenner-Freiheitsgrad des F-Tests bzw. der Freiheitsgrad des t-Tests in der ESTIMATE-Anweisung standardmäßig auf **DF = 1** gesetzt,
- b) Partieller Sill = 0 oder Partieller Sill ≈ 0 , für n = 1: Range ≈ 160 , für n = 3: Range ≈ 480 , für n = 5: Range ≈ 800 , **mit Schätzung des Standardfehlers des Partiellen Sills**, dann wird der Nenner-Freiheitsgrad des F-Tests bzw. der Freiheitsgrad des t-Tests in der ESTIMATE-Anweisung **DF =1 oder > 1** gesetzt.

Bei der Auswertung der latenten Variablen tritt nur der erste Fall auf. Durch die Schätzung von Partieller Sill = 0 erfolgt eine Modellreduktion auf Modell I, so dass die Auswertungsvariante A-R genutzt wird und sich der Nenner-Freiheitsgrad des F-Tests und der Freiheitsgrad des t-Tests im LMM nach **DF = N-Rang(X)** berechnet.

Wird Partieller Sill = 0 geschätzt, so existiert für die Freiheitsgradapproximation nach Kenward und Roger (1997; 2009) keine Approximation für den Erwartungswert der Varianz der Kovarianzparameter, und diese Freiheitsgradapproximation wird nicht genutzt. Die Festsetzung des Freiheitsgrades auf DF = 1 bei der Auswertung der Bonituren mittels Schwellenwertmodell hat Konsequenzen für den Testentscheid und in Folge dessen auf den realisierten Fehler 1. Art. Für den Freiheitsgrad DF=1 wird deutlich häufiger auf Nicht-Signifikanz erkannt als bei großen Freiheitsgraden. Die Einhaltung des Fehlers 1. Art bei den Simulationen wird also durch den Anteil an Läufen mit nicht positiv definiten G-Matrix beeinflusst. Je höher ihr Anteil, desto geringer ist der realisierte Fehler 1. Art von t- und F-Test. Der Anteil an Läufen mit nicht positiv definiten G-Matrix sinkt mit steigender Beobachtungsanzahl und mit der Stärke der simulierten räumlichen Kovarianz. Tabelle 6.1 zeigt den Anteil der konvergenten Simulationsläufe mit nicht positiv definiten G-Matrix für die Auswertung der Bonituren mittels Schwellenwertmodell mit Linkfunktion „cumlogit“.

Tab. 6.1: Anteil der Simulationsläufe mit nicht positiv definiten G-Matrix in Prozent der konvergenten Simulationsläufe (Schwellenwertmodell mit Linkfunktion „cumlogit“ für Simulationsansätze mit räumlicher Kovarianz, Parametersatz 1)

	Simulierte räumliche Kovarianz			
	1_10_1	1_20_1	1_30_1	2_40_1
n = 1	55.05	20.83	8.26	2.01
n = 3	52.2	5.77	0.81	0.00
n = 5	18.51 [*]	0.23	0.00	0.00

*⁾ Ergebnisse von 1000 Läufen, da Simulationen mit 10000 Läufen aufgrund sehr langer Laufzeiten nicht realisierbar waren

Weiter führt die Schätzung von Partieller Sill = 0, dazu, dass ein Modell ohne räumlich strukturierten Varianzanteil geschätzt wird. Das Verhältnis der Schätzung von Range = 160 bei n = 1 Beobachtung und Range = 480 (800) bei n = 3 (5) Beobachtungen in der Parzelle legt nahe, dass es sich bei der Festlegung des Range um eine von der Anzahl der Beobachtungen und derjenigen der Parzellen abhängige Funktion handelt.

Ein Gedanke wäre:

$$160 = 1 \text{ Beobachtung} * 80 \text{ Parzellen} * 2$$

$$480 = 3 \text{ Beobachtungen} * 80 \text{ Parzellen} * 2,$$

$$800 = 5 \text{ Beobachtungen} * 80 \text{ Parzellen} * 2,$$

wobei die Herkunft der 2 fraglich ist, und es werden ja auch nicht immer ganzzahlige Schätzwerte erhalten. Zur Abklärung wurde in einigen dieser Fälle (mit n = 1) getestet, was passiert, wenn für die Schätzung der räumlichen Parameter unterschiedliche Startwerte vorgegeben werden. Dies führte dazu, dass der entsprechende Simulationslauf nun eine positiv definite G-Matrix aufwies, und es wurden, von den Startwerten unabhängig, dieselben Schätzungen für die räumlichen Parameter erhalten. Daraus kann man folgern, dass die ohne Startwertvorgabe erhaltenen Ergebnisse für den Range bei nicht positiv definiter G-Matrix höchstwahrscheinlich Startwerte der Iteration darstellen.

Das dritte Problem bei der Auswertung von Simulationsläufen mit räumlicher Kovarianz ist das Auftreten von Läufen, die einen negativen Eigenwert der Kovarianzmatrix der festen Effekte (CovB) aufweisen. Dies kann zu Problemen bei den Schätzungen der Varianz der Behandlungseffekte führen. Bei der Auswertung der Bonituren mittels Schwellenwertmodell tritt dies meist für Läufe mit Schätzungen Range = 0 auf. Die Eigenwerte der CovB werden offensichtlich im Rahmen der Prozedur GLIMMIX-**nicht** standardmäßig getestet.

Für die Auswertung der latenten Variablen mittels LMM führen Läufe mit negativen Eigenwerten der CovB dazu, dass negative Schätzwerte für die Varianz der Prüfgliedkontraste auftreten können. In diesen Fällen liefert die ESTIMATE-Anweisung keine Ergebnisse für die Schätzungen der Standardfehler der Behandlungskontraste. Diese Simulationsläufe fehlen dann für die Ermittlung der realisierten Fehler 1. Art. Die Schätzungen der räumlichen Kovarianzparameter weisen keine Auffälligkeiten auf. Bei der Auswertung der Bonituren mittels Schwellenwertmodell weisen die Schätzungen der Behandlungseffekte und ihrer Kontraste keine Auffälligkeiten auf und werden mit Standardabweichung erhalten.

Die dargestellten Probleme bei der Nutzung der Auswertungsvariante A-cov(sph) im LMM und im SWM haben Einfluss auf die Simulationsergebnisse wie Bias und Varianz der Schätzung der Varianzkomponenten und den realisierten Fehler 1. Art. Die Abbildung 6.1 stellt die Probleme bei der Ergebnisbildung schematisch dar.

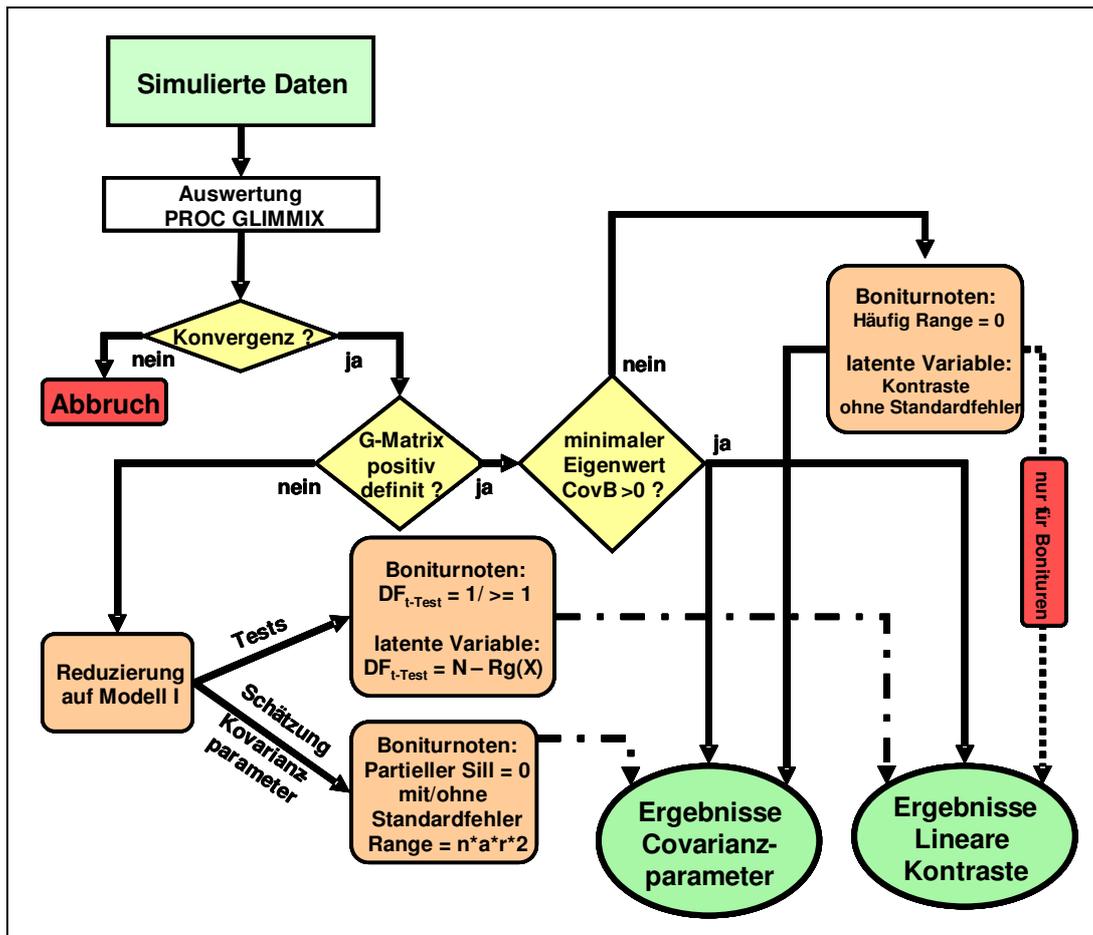


Abb. 6.1: Schematische Darstellung der durch nicht positiv definite G-Matrix oder negative Eigenwerte der Kovarianzmatrix der festen Effekte entstehenden Probleme bei der Ergebnisbildung mit der Auswertungsvariante A-cov(sph) (n = Beobachtungen je Parzelle, a = Behandlungen; r = Wiederholung; $N = n \cdot a \cdot r$)

6.3 Einhaltung des Fehlers 1. Art bei Hypothesentests

Auswertung der Simulationsansätze ohne räumliche Kovarianz

Alle Simulationsansätze ohne räumliche Kovarianz wurden nach dem vollständig randomisierten Modell (Variante A-R) simuliert und ausgewertet (vgl. Abschnitt 5.5.1). Für die Auswertung der latenten Variablen liefert das lineare Modell die erwartete Einhaltung der Fehler 1. Art für den t- und den F-Test. Die Verzerrungen der Fehler des χ^2 -Tests sind aufgrund seiner nur asymptotischen Gültigkeit auf die geringen Stichprobengrößen zurückzuführen. Im Vergleich zur Auswertung der latenten Variablen mittels LM liefern nicht alle Auswertungsmethoden für Bonituren realisierte Fehler 1. Art, die dem nominalen Fehler entsprechen.

Werden die Bonituren mittels Schwellenwertmodell ausgewertet, so werden die nominalen Fehler des t- und F-Test unabhängig von der gewählten Linkfunktion erst ab Ansätzen mit $n = 5$ eingehalten. Das Schwellenwertmodell benötigt für die Schätzung der Parameter (Prüfgliedeffekte und Stützpunkte) eine ausreichend große Stich-

probengröße. Dies bestätigt sich auch für die Ergebnisse der Auswertung der Bonituren mit anderen Parametersätzen für die Wahrscheinlichkeiten oder mit nur $k = 4$ Klassen (vgl. Abschnitt 5.7.1). Die leichten Unterschiede der realisierten Fehler 1. Art zwischen den Linkfunktionen liegen darin begründet, dass die latente Variable mit der Normalverteilung simuliert wurde, wodurch bei der Auswertung mittels Linkfunktion „cumlogit“ die Normalverteilung der latenten Variablen mittels logarithmischer Verteilung approximiert wird.

Werden die Parzellenmittelwerte der Bonituren mittels linearem gemischtem Modell ausgewertet, so wird bei den hier gewählten Parametern der Wahrscheinlichkeiten schon ab $n = 3$ Beobachtungen pro Parzelle eine approximative Normalverteilung der Residuen der Mittelwerte erreicht, und die Fehler von t- und F-Test werden eingehalten. Der Fehler des χ^2 -Tests wird wiederum aufgrund des geringen Stichprobenumfangs (ein Mittelwert pro Parzelle) überschätzt. Ergebnisse von Thöni (1985; 1992) zeigten für die Analyse von Versuchen mit $k = 4$ und $k = 9$ Boniturstufen eine Übereinstimmung der Ergebnisse von Schwellenwertmodell und der varianzanalytischen Auswertung der Parzellenmittelwerte. Allerdings waren in den von Thöni untersuchten Daten deutlich mehr Beobachtungen je Parzelle vorhanden. Der Autor gibt als untere Grenze $n = 25$ Beobachtungen je Parzelle an (Thöni, 1985). Piepho (1997) gibt als Erklärungsansatz für diese Übereinstimmung einerseits die relative Robustheit der Varianzanalyse gegenüber Verletzungen ihrer Voraussetzungen (Varianzhomogenität und Normalverteilung der Residuen) an und zeigt als weiteren möglichen Grund, dass bei Gültigkeit des Schwellenwertmodells der Erwartungswert der Boniturnote in einer Parzelle in annähernd linearer Beziehung zum linearen Prädiktor des Schwellenwertmodells steht. Allerdings weist er darauf hin, dass dieser Zusammenhang abhängig ist von den Schwellenwerten bzw. Stützpunkten und dem Wertebereich des linearen Prädiktors.

Werden die Bonituren unter der Annahme der approximativen Normalverteilung mittels linearem gemischtem Modell ausgewertet, so liefern sie bei $k = 9$ Klassen und dem Parametersatz 1 für ihre Wahrscheinlichkeiten, die Einhaltung der nominalen Fehler 1. Art für den t- und den F-Test. Auch eine Veränderung in den Parametern der Wahrscheinlichkeiten (Parametersätze 2 – 4, Tab. 3.2) oder die Reduktion der Klassenanzahl auf $k = 4$ (Parametersätze 5 - 8, Tab. 3.2) führt für Ansätze ohne räumliche Kovarianzen unabhängig von der Anzahl Beobachtungen je Parzelle zur Einhaltung der Fehler 1. Art für den t-Test des Prüfgliedkontrasts (vgl. Abschnitt 5.7). Werden jedoch Daten mit heterogenen Prüfgliedeffekten simuliert (vgl. Tabelle 3.3), was in Folge zu heterogenen Restvarianzen führt, werden die nominalen Fehler des t-Tests nicht immer eingehalten (Überschätzung für $k = 9$ und Unterschätzung für $k = 4$ Klassen), wenn die Heterogenität der Restvarianzen im Auswertungsmodell nicht beachtet wird. Dies bestätigt die Aussage, dass der t-Test robust ist gegenüber Abweichungen der Normalverteilung, jedoch empfindlich gegenüber Varianzheterogenität. Die Rangverfahren von Brunner et al. (2001; 2002) eignen sich gut für den statistischen Test der Prüfgliedkontraste, der Test der Globalhypothese unterschätzt jedoch den Fehler der ANOVA-Typ Statistik für die Simulationsansätze mit $n < 5$. Die Approximation der F-Verteilung für die ANOVA-Typ Statistik durch eine zentrale F-Verteilung ist nach Brunner und Munzel (2002, p. 148) gut für Stichproben mit Beobachtungszahlen $n_{ij} > 7$ geeignet. Die hier erhaltenen Ergebnisse weisen daraufhin, dass für $k = 9$ Klassen und den hier für die Datensimulation gewählten Parametern für die Wahrscheinlichkeiten auch eine Stichprobengröße von $n = 5$ ausreicht. Die Wald-Typ Statistik ist

nur asymptotisch gültig und ist für kleine Stichproben ungeeignet. Nach Brunner und Munzel (2002, p. 224) liefert sie bei kleinen und mittleren Stichprobengrößen anti-konservative Entscheidungen, dies wird von den hier erhaltenen Ergebnissen bestätigt.

Die Permutationstests sind gut geeignet für den Test der Globalhypothese, aber der Fehler des Prüfgliedkontrasts wird unterschätzt. Dies liegt in der geringen Stichprobengröße begründet. In den simulierten Datensätzen liegen pro Prüfglied je 4 Wiederholungen mit $n = 1$ Beobachtungen vor. Wird der Prüfgliedkontrast getestet, so werden nur die Daten der entsprechenden Prüfglieder permutiert, sind Bindungen in den Beobachtungen vorhanden, so reduziert sich die Anzahl der Permutationen zusätzlich. Dies führt nach Frisch und Schumacher (1998) zu einer Reduzierung der Testgüte bzw. zu konservativen Testergebnissen, wie sie für die vorliegenden Untersuchungen in den Ergebnissen der Prüfgliedkontraste der Auswertung der Simulationen ohne räumliche Kovarianz wiederzufinden sind.

Die Ergebnisse aller Auswertungsmethoden für die Simulationsansätze ohne räumliche Kovarianz zeigen, dass für Parametersatz 1 ab einer ausreichend großen Stichprobengröße ($n \geq 5$ Beobachtungen pro Parzelle) alle Methoden den nominalen Fehler des Tests der Globalhypothese einhalten. Für den Test des Prüfgliedkontrasts führen nur die Permutationsverfahren zu Verzerrungen. Die Untersuchungen der anderen Parametersätze (siehe Abschnitt 5.7.1) der Bonituren mittels Schwellenwertmodell und mittels linearem Modell zeigen, dass eine Reduzierung der Klassenanzahl für das Schwellenwertmodell zu einer größeren benötigten Stichprobe führt, um Verzerrungen im realisierten Fehler des t-Tests zu vermeiden. Die Auswertung der Bonituren als stetiges Merkmal mittels LM ist hier für kleine Beobachtungszahlen vorzuziehen. Weiter zeigen diese Untersuchungen, dass bei Vorliegen heterogener Prüfgliedeffekte auch das LM nicht immer die Einhaltung der nominalen Fehler gewährleistet, seine Wirksamkeit derjenigen des Schwellenwertmodells jedoch meist überlegen ist.

Auswertung der Simulationsansätze mit räumlicher Kovarianz

Die Ergebnisse aller Auswertungsmethoden und -varianten, angewendet auf Simulationsansätze mit räumlicher Kovarianz, können folgendermaßen zusammengefasst werden. Wird eine räumliche Kovarianz in den Daten simuliert und diese im Auswertungsmodell nicht oder unzureichend beachtet, so kommt es mit zunehmender Beobachtungsdichte und zunehmender Stärke der räumlichen Korrelation zu extremen Verzerrungen der realisierten Fehler 1. Art. Durch die steigende Beobachtungsdichte kommt es zur Zunahme der räumlichen Korrelation auf kleinem Raum, was in Folge zu extremen Überschätzungen des Fehlers 1. Art für die Auswertungsvarianten A-R sowie A-BI fix und random führt. Dies gilt gleichermaßen für die Auswertung der latenten Variablen mittels LMM und für die Auswertung der Bonituren mit den unterschiedlichen Auswertungsmethoden. Für die Verzerrung der nominalen Fehler für die unterschiedlichen Auswertungsmethoden der Bonituren sind verschiedene Aspekte zu beachten.

Die Einhaltung des Fehlers 1. Art der statistischen Tests bei Nutzung einer Auswertungsvariante des LMM oder SWM wird für Simulationsansätze mit räumlichen Kovarianzen von mehreren Faktoren beeinflusst. Neben dem Konvergenzverhalten und der Erwartungstreue der Schätzungen der fixen Effekte und ihrer Kontraste kann auf-

grund der vorliegenden Simulation die Analyse erweitert werden auf die beobachtete (STD_{beo}) und geschätzte Standardabweichung (STD_{ges}) der Prüfgliedkontraste (vgl. Abschnitt 3.4). Diese Maßzahlen sind Ausdruck der Beachtung der simulierten Kovarianz bei der Schätzung der Effekte und ihrer Differenz (STD_{beo}) sowie der Schätzung der Varianz der Prüfglieddifferenz (STD_{ges}).

Der Einfluss dieser beiden Aspekte wird im Folgenden anhand der Ergebnisse der Auswertung der latenten Variablen mittels LMM erläutert.

Die nominalen Fehler 1. Art von t- und F-Test werden für die Auswertung der latenten Variablen für Ansätze mit $n = 1$ sowohl von der räumlichen Auswertungsvariante A-cov(sph) als auch von den Varianten A-BI fix und random eingehalten (vgl. Tab. 5.7). Die Variante A-R unterschätzt die Fehler 1. Art von t- und F-Test bei Simulationsansätzen mit stärkerer räumlicher Kovarianz. Für Ansätze mit $n > 1$ Beobachtungen dagegen liefern die Auswertungsvarianten A-BI(fix)-Parz. und A-BI(rand.)-Parz. die besten Ergebnisse. Die räumliche Variante A-cov(sph) hält die Fehler nur bei Simulationsansätzen mit stärkerer räumlicher Kovarianz ein. Alle anderen Auswertungsvarianten führen zu extrem hohen Verzerrungen der empirischen Fehler 1. Art.

Die beobachtete Standardabweichung spiegelt die Beachtung der simulierten Kovarianz bei der Schätzung der Effekte und deren Differenz wieder. Bei Auswertung der latenten Variablen (vgl. Tab. 5.2) werden mit den Auswertungsvarianten A-R, A-BI fix und random identische Prüfglieddifferenzen und beobachtete Standardabweichungen erhalten, obwohl die Blockmodelle einen zusätzlichen Faktor und eine andere Kovarianzstruktur innerhalb R (bzw. bei Modellierung der Kovarianz über die zufälligen Effekte innerhalb G) aufweisen. Da jedoch die Differenz der Effekte im balancierten Fall betrachtet wird, ist die Blockvarianz für die Varianz der Differenz unbedeutend. Die beobachteten Prüfglieddifferenzen und deren beobachtete Standardabweichung sind daher in den Auswertungsvarianten A-R und A-BI fix und random gleich.

Beobachtete und geschätzte Standardabweichungen (vgl. Tab. 5.2) sinken mit steigender Beobachtungsdichte für alle Analysemodelle. Allerdings führt die Nichtbeachtung der räumlichen Kovarianz in den nichträumlichen Auswertungsvarianten zur Erhöhung der beobachteten Standardabweichung gegenüber der Variante A-cov(sph). Diese Erhöhung nimmt mit steigender Beobachtungsdichte und der Stärke der simulierten räumlichen Kovarianz zu.

Für die Einhaltung der nominalen Fehler spielt jedoch auch eine Rolle, inwiefern die beobachtete Standardabweichung von der geschätzten Standardabweichung widergespiegelt wird. Die Auswertungsvarianten, die den nominalen Fehler 1. Art für den t- und F-Test überschätzen, unterschätzen die beobachtete Standardabweichung und umgekehrt. Ausnahme ist hier die Auswertungsvariante A-cov(sph) für Ansätze mit stärkerer simulierter räumlicher Kovarianz und $n = 1$ Beobachtung. Hier wird die Standardabweichung im Mittel deutlich überschätzt, dennoch werden die nominalen Fehler eingehalten. Diese starke Überschätzung der Standardabweichung ist einigen Simulationsläufen mit extremen Ausreißern geschuldet. Durch diese ist zwar die mittlere geschätzte Standardabweichung sehr hoch, der empirische Fehler 1. Art über alle Simulationsläufe wird dadurch aber wenig beeinflusst.

Bei Simulationsansätzen mit $n = 1$ Beobachtung pro Parzelle ist die Erhöhung der beobachteten Standardabweichung der nichträumlichen Auswertungsvarianten gegenüber der Variante A-cov(sph) noch nicht so gravierend. Hier sind die beobachteten Standardfehler der Differenz für die Analysemodelle noch sehr ähnlich. Die beobachteten Standardfehler der nichträumlichen Auswertungsvarianten sind gegenüber der räumlichen Variante nur wenig erhöht. In der Variante A-R wirkt sich aber die Nichtbeachtung der Abhängigkeit beim geschätzten Standardfehler der Differenz vergleichsweise stärker aus und führt zu einer Unterschätzung des nominalen Fehlers 1. Art.

Bei den übrigen Beobachtungszahlen ist für die Varianten A-R und A-BI fix und random die Zunahme der beobachteten Standardabweichungen gegenüber derjenigen der räumlichen Variante A-cov(sph) größer als das im geschätzten Standardfehler widergespiegelt wird, was zu einer Überschätzung des nominalen Fehlers 1. Art führt. In Simulationsanalysen von Hu et al. (2006), in denen als Auswertungsvariante von Daten mit räumlicher Kovarianz ein Blockmodell genutzt wurde, ergab für die Unterschätzung der Standardabweichung ebenfalls eine Überschätzung der nominalen Fehler 1. Art und umgekehrt.

Werden die Effekte mit der räumlichen Auswertungsvariante oder den Varianten mit zufälligem Parzelleneffekt innerhalb der Blocks geschätzt, wird die Information der anderen Daten genutzt, und beobachtete und geschätzte Standardabweichung stimmen überein. Der nominale Fehler kann dann eingehalten werden.

Die Abbildungen 6.2 bis 6.4 veranschaulichen beispielhaft die Auswirkungen dieser Verhältnisse auf den t-Test für die Auswertungsvarianten A-cov(sph) und A-BL fix von Simulationsansätzen mit $n = 5$ Beobachtungen und starker räumlicher Korrelation (2_40_1). Die Häufigkeitsverteilung der für 10000 Simulationen erhaltenen t-Werte der unterschiedlichen Analysemethoden (Abbildung 6.2) zeigt für die Variante A-BI fix einen deutlichen Anstieg der außerhalb der mittleren 95 % der t-Verteilung liegenden Werte, wodurch die Überschätzung der nominalen Fehler 1. Art zustande kommt. Die Erhöhung der beobachteten Standardabweichung der Differenz gegenüber dem räumlichen Modell wird in der Verbreiterung der Häufigkeitsverteilung der geschätzten Prüfglieddiffenzen deutlich (Abbildung 6.3) und die Unterschätzung dieser Erhöhung durch die nichträumlichen Analysevariante A-BI fix zeigt sich in der Linksverschiebung der Häufigkeitsverteilung der geschätzten Standardabweichung der Differenz (Abbildung 6.4).

Auch die nominalen Fehler von F- und χ^2 -Test werden in ähnlicher Weise von den oben beschriebenen Verhältnissen beeinflusst. Die im Vergleich zum t-Test jedoch sehr großen realisierten Fehler 1. Art des F- und χ^2 -Tests liegen darin begründet, dass diese Tests auf Verteilungsabweichungen empfindlicher reagieren.

Die hier für die Auswertung der latenten Variablen mittels LMM gezeigten Zusammenhänge können auf die Auswertung der Bonituren übertragen werden. Im Falle der Auswertung mittels Schwellenwertmodell gilt der Zusammenhang jedoch auf der Ebene des linearen Prädiktors, und die Verhältnisse stellen sich für Simulationsansätze mit $n = 1$ etwas anders dar. Sie werden im Folgenden für die Auswertung des Parametersatz 1 der Bonituren erläutert.

Die Ergebnisse der Auswertung der Bonituren (Parametersatz 1, homogene Prüfgliedeffekte) mittels Schwellenwertmodell liefern unabhängig vom Simulationsansatz, von der Linkfunktion und der Auswertungsvariante verzerrte Fehler für den F- und den χ^2 -Test. Weiter treten bei der Nutzung des Schwellenwertmodells abhängig von der Linkfunktion und dem Simulationsansatz teilweise extreme Konvergenzprobleme auf. Für Ansätze mit $n = 1$ Beobachtungen pro Parzelle liefert die Nutzung der Linkfunktion „cumlogit“ die beste Einhaltung der nominalen Fehler 1. Art des t-Tests für die Auswertungsvariante A-cov(sph), mit der Linkfunktion „cumprobit“ sind die Fehler verzerrt. Für Ansätze mit $n > 1$ führt die Linkfunktion „cumlogit“ mit der Auswertungsvariante A-cov(sph) zur besten Einhaltung der nominalen Fehler des t-Tests, gefolgt von den Varianten A-BI(fix)-Parz. und A-BI(rand.)-Parz., die nur leichte Verzerrungen zeigen. Mit der Linkfunktion „cumprobit“ ist es umgekehrt. Die Variante A-cov(sph) zeigt mit der Linkfunktion „cumprobit“ jedoch auch die stärksten Konvergenzprobleme.

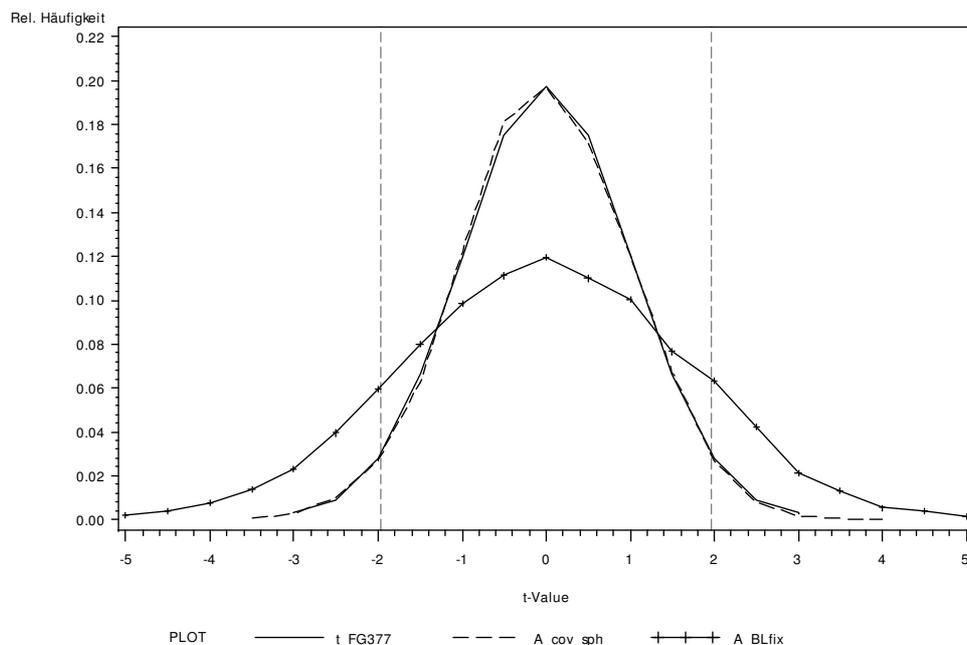


Abb. 6.2: Verteilung der t-Werte der Auswertung der latenten Variablen mittels LMM (Auswertungsvarianten A-cov(sph) und A-BL fix, 5 Beobachtungen je Parzelle, räumliche Korrelation (2_40_1), gestrichelte Grenzen zeigen die mittleren 95% der t-Verteilung bei dem in der Auswertungsvariante A-BL fix für die Tests verwendeten Freiheitsgrad von FG=377)

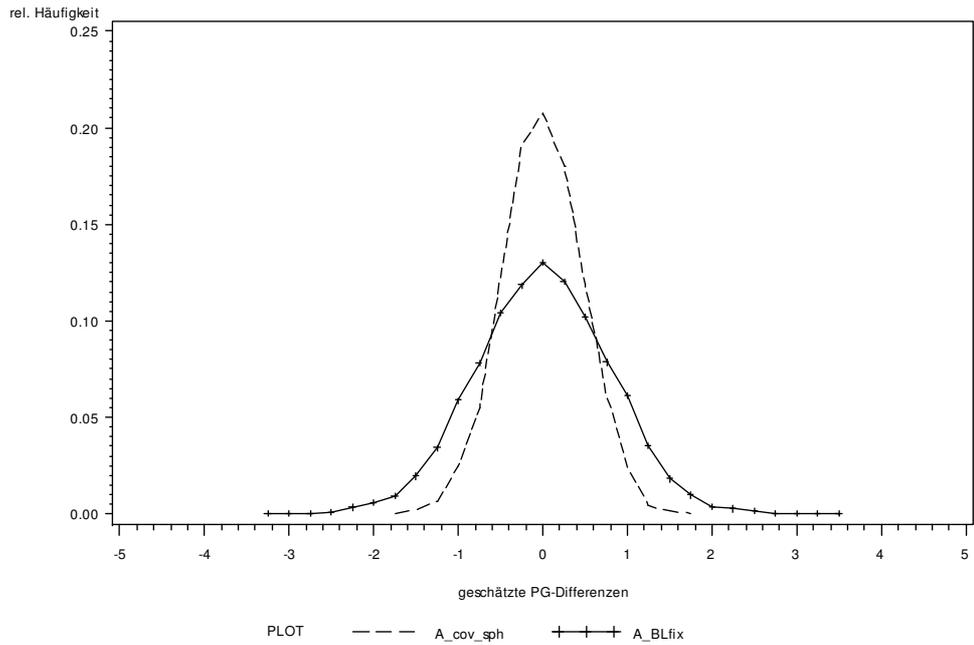


Abb. 6.3: Verteilung der geschätzten Prüfglieddiffenzen der Auswertung der latenten Variablen mittels LMM (Auswertungsvarianten A-cov(sph) und A-BL fix, 5 Beobachtungen je Parzelle, und räumliche Korrelation (2_40_1))

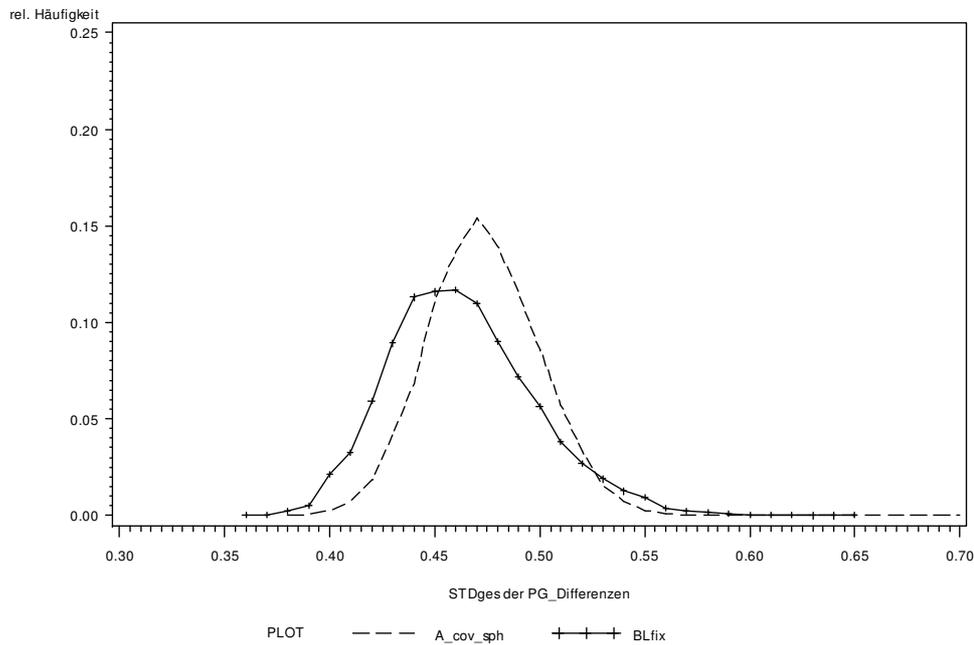


Abb. 6.4: Verteilung der geschätzten Standardabweichungen der Prüfglieddiffenzen der Auswertung der latenten Variablen mittels LMM (Auswertungsvarianten A-cov(sph) und A-BL fix, 5 Beobachtungen je Parzelle, und räumliche Korrelation (2_40_1))

Betrachtet man für die Auswertung der Bonituren (Parametersatz 1, homogene Prüfgliedeffekte) mittels Schwellenwertmodell mit der Linkfunktion „cumlogit“ die Zusammenhänge von beobachteter und geschätzter Standardabweichung, so finden sich im Vergleich mit den Verhältnissen bei der Auswertung der latenten Variablen Unterschiede. Auch für das Schwellenwertmodell sinken die beobachteten Standardabweichungen aller Auswertungsvarianten mit steigender Beobachtungsanzahl. Es ist jedoch die beobachtete Standardabweichung der räumlichen Variante A-cov(sph) gegenüber den nicht-räumlichen Auswertungsvarianten erhöht. Besonders deutlich zeigt sich dies für Ansätze mit $n = 1$ Beobachtung pro Parzelle (vgl. Tab. 5.3). Mit steigender Beobachtungsanzahl schwächt sich diese Erhöhung ab. Dies ist darauf zurückzuführen, dass im Schwellenwertmodell mit der gegebenen Stichprobengröße neben den fixen Parametern auch die Stützpunkte geschätzt werden müssen. Bei der geringen Stichprobengröße der Simulationsansätze mit $n = 1$ Beobachtungen pro Parzelle kommt dies besonders deutlich zum Tragen. Der Einfluss einer Unterschätzung der beobachteten Standardabweichung hat aber auch für die Auswertung der Bonituren mittels Schwellenwertmodell dieselben Auswirkungen wie für Auswertung der latenten Variablen, eine Unterschätzung führt zu einer Überschätzung des nominalen Fehlers 1. Art.

Werden für Simulationsansätze mit räumlicher Kovarianz die Parzellenmittelwerte der Bonituren (Parametersatz 1, homogene Prüfgliedeffekte) mittels LMM ausgewertet, werden die nominalen Fehler 1. Art von t- und F-Test eingehalten, wenn die Auswertungsvarianten A-BI fix oder random genutzt werden. Die räumliche Auswertungsvariante A-cov(sph) liefert hier verzerrte Fehler. Durch die Bildung der Parzellenmittelwerte der Bonituren reduzieren sich die Beobachtungen auf eine Beobachtung pro Parzelle. Durch die Mittelwertbildung verstärkt sich jedoch die räumliche Korrelation (vgl. Tab. 5.12), was zu einem hohen Anteil an Simulationsläufen mit deutlich erhöhten geschätzten Standardabweichungen führt. Der Anteil dieser Simulationsläufe ist so groß, dass die nominalen Fehler nicht eingehalten werden können. Die stärkere Korrelation kann jedoch in den Blockmodellen gut abgebildet werden. Hier ist die Schätzung der Prüfgliedkontraste erwartungstreu, geschätzte und beobachtete Standardabweichung sind etwa gleich.

Die stetige Auswertung der Bonituren (Parametersatz 1, homogene Prüfgliedeffekte) mittels LMM führt für Simulationsansätze mit räumlicher Kovarianz zu ähnlichen Ergebnissen wie die Auswertung mittels Schwellenwertmodell. Mit dem Unterschied, dass bei denjenigen Auswertungsvarianten, die den nominalen Fehler 1. Art des t-Tests einhalten, dies in den meisten Fällen auch für den F-Test zutrifft bzw. dieser nur leicht verzerrt ist. Für Ansätze mit $n = 1$ Beobachtungen je Parzelle liefern die räumliche Auswertungsvariante und die Blockmodelle die Einhaltung der Fehler 1. Art. Für Ansätze mit $n = 3$ oder 5 Beobachtungen dagegen werden die Fehler am besten von den Auswertungsvarianten mit zufälligem Parzelleneffekt innerhalb der Blocks (fix oder random) eingehalten. Die Variante A-cov(sph) dagegen liefert nur bei Simulationsansätzen mit stärkerer räumlicher Kovarianz die nominalen Fehler, denn nur dort stimmen geschätzte und beobachtete Standardabweichung überein. Die Ergebnisse für die Prüfglieddiffereenz und die beobachtete Standardabweichung sind für die Auswertung der Bonituren und ihrer Parzellenmittelwerte mittels LMM gleich. Die geschätzte Standardabweichung unterscheidet sich und ist für die Auswertung der Parzellenmittelwerte aufgrund der Mittelwertbildung über die Beobachtungen je Parzelle niedriger.

Die Permutationsverfahren von Schumacher et al. (1995 a, b; 1997) sind für Bonituren (Parametersatz 1, homogene Prüfgliedeffekte) und Simulationsansätze mit räumlicher Kovarianz nicht geeignet, den Prüfgliedkontrast unter Einhaltung des nominalen Fehlers 1. Art zu testen. Nur die Globalhypothese kann unabhängig von der simulierten räumlichen Kovarianz mit der Auswertungsvariante A-BI fix unter Einhaltung des nominalen Fehlers getestet werden. In dieser Arbeit wurden aufgrund der geringen Blockzahlen die um den Blockmedian korrigierten Daten vollständig randomisiert ausgewertet. Diese Korrektur fängt bei den hier gewählten räumlichen Simulationsparametern die räumliche Korrelation hinreichend ab. Für den Prüfgliedkontrast ist aufgrund der vorhandenen Bindungen die Stichprobengröße jedoch zu gering. Werden deutlich größere räumliche Parameter gewählt (Sill = 4, Range = 80, Nugget = 1, Ergebnisse nicht gezeigt) kann die Korrektur mittels Blockmedian die räumliche Korrelation jedoch nicht mehr abfangen.

Die nichtparametrischen Rangverfahren von Brunner et al. (2001; 2002) halten für die Bonituren (Parametersatz 1, homogene Prüfgliedeffekte) mit keiner Auswertungsvariante die nominalen Fehler 1. Art ein. Mit der Vergabe von Rängen ist immer ein Informationsverlust verbunden, gleichzeitig wird in den simulierten Daten die Anzahl der Bindungen mit der simulierten räumlichen Kovarianz verstärkt, dies wirkt sich negativ auf die Güte der Approximation der F-Verteilung für die ANOVA-Statistik aus (Brunner and Munzel, 2002, p. 224).

Zunächst wurden alle Untersuchungen in dieser Arbeit an dem Parametersatz 1 der Bonituren durchgeführt, der mit homogenen Prüfgliedeffekten simuliert wurde. Eine realistischere Annahme sind jedoch heterogene Prüfgliedeffekte. Da die parameterfreien Auswertungsmethoden (Rang- und Permutationsverfahren) für den Parametersatz 1 zu unbefriedigenden Ergebnissen führten, wurden sie aus der Analyse der weiteren Parametersätze ausgeschlossen. Daher wurden für Parametersätze mit heterogenen Prüfgliedeffekten (vgl. Abschnitt 5.7) nur die Auswertungen der Bonituren mittels Schwellenwertmodell und mittels LMM durchgeführt.

Die Analysen der räumlichen Simulationsansätze für Parametersatz 1a mit heterogenen Prüfgliedeffekten sind in den Tabellen 6.2 und 6.3 dargestellt und zeigen eine gemeinsame Bewertung der Auswertungsmethoden und Varianten mit der Auswahl homogener bzw. heterogener Restvarianzen im Auswertungsmodell des LMM, sie führen zu den im folgenden Kapitel aufgeführten Konsequenzen für die Auswertungspraxis.

6.4 Konsequenzen für die Auswertungspraxis

Die Simulationsuntersuchungen in dieser Arbeit zeigen, dass die Auswertung von Bonituren bei Existenz räumlicher Kovarianz und heterogener Prüfgliedeffekte viele Schwierigkeiten aufweist. Die Nichtbeachtung der räumlichen Kovarianz im Auswertungsmodell führt zu Verzerrungen der Fehler 1. Art. Die Permutations- und Rangverfahren sind im momentanen Stand der Umsetzung dieser Methoden innerhalb SAS nicht geeignet die räumliche Kovarianz durch einen entsprechenden Modellansatz abzubilden und liefern daher keine Einhaltung des Fehlers 1. Art. Sie scheiden damit als

Auswertungsmethoden aus. Als mögliche Auswertungsmethoden kommen nur die Nutzung des SWM oder LMM in Frage.

Die vorliegenden Untersuchungen zeigen, dass die Auswertungsmethode und der Modellansatz zur Beachtung der räumlichen Kovarianz in Abhängigkeit von der Anzahl Beobachtungen je Parzelle auszuwählen sind. Weiter muss ein Kompromiss gefunden werden zwischen dem Konvergenzverhalten einer Methode und der zu erwartenden Verzerrung der Fehler 1. Art. Ausgehend von der Analyse des Parametersatzes 1a mit heterogenen Prüfgliedeffekten (Tab. 6.2 und 6.3) können die folgenden Aussagen getroffen werden.

Für Datensätze mit $n = 1$ Beobachtung pro Parzelle wird empfohlen zur Auswertung das Schwellenwertmodell (Linkfunktion „cumlogit“) mit Formulierung einer räumlichen Kovarianz, Auswertungsvariante A-cov(sph) zu nutzen. Mit dieser Modellvariante liefert das Schwellenwertmodell nur leicht verzerrte Fehler 1. Art für den t-Test des Prüfgliedkontrasts und den F-Test. Die Erwartungstreue für die Schätzung der Wahrscheinlichkeiten der Bonituren ist jedoch nicht gegeben. Da die Hypothesentests innerhalb des Schwellenwertmodells auf der Ebene des linearen Prädiktors durchgeführt werden, ist die Interpretation erschwert. Treten Konvergenzprobleme auf, so stellt die Beachtung der räumlichen Kovarianz über einen Blockeffekt eine Alternative dar, die jedoch zu etwas stärker verzerrten Fehlern 1. Art führt.

Für Ansätze mit $n > 1$ Beobachtungen treten beim Schwellenwertmodell starke Konvergenzprobleme auf, wenn eine räumliche Kovarianz existiert. Für diese Ansätze wird empfohlen, die Bonituren mittels LMM als stetige Variable auszuwerten und als Modellansatz einen Ansatz mit zufälligen Parzelleneffekten innerhalb fixer oder zufälliger Blocks zu nutzen. Bei Vorliegen heterogener Restvarianzen der Prüfglieder sollten diese im Auswertungsmodell durch eine entsprechende Anweisung berücksichtigt werden. Diese Auswertung liefert dann nur sehr leicht verzerrte Fehler für den t-Test bei fast hundertprozentiger Konvergenz.

Tab. 6.2 Zusammenfassende Bewertung der Ergebnisse der nutzbaren Auswertungsmethoden und -varianten der Bonituren (Parametersatz 1a, heterogene Prüf- gliedeffekte) für Simulationsansätze mit räumlichen Kovarianzstrukturen, mit Auswahl homogener Restvarianzen im LMM

	Auswertungsmethode	Auswertungsvariante	Konvergenz- verhalten	Einhaltung der nominalen Fehler 1. Art des t-Tests des Prüfgliedkontrasts PG 1 zu 2
n = 1	SWM Linkfunktion „cumlogit“	A-cov(sph)	(-)	(-)
	LMM für Bonituren	A-BI fix / A-BI rand.	(+)	(-)
n > 1	SWM Linkfunktion „cumlogit“	A-cov(sph)	(- -)	(-)
		A-BI(fix)-Parz. / A-BI(rand.)-Parz.	(- - -)	(+)
	LMM für Bonituren	A-BI(fix)-Parz. / A-BI(rand.)-Parz.	(+)	(-)
	LMM für Parzellenmittelwerte	A-BI fix / A-BI rand.	(+)	(-)

(-) leichte, (- -) mittelstarke und (- - -) starke Verzerrungen, bzw. Konvergenzprobleme
 (+) Einhaltung der Fehler, bzw. 100 %-ige Konvergenz

Tab. 6.3: Zusammenfassende Bewertung der Ergebnisse der nutzbaren Auswertungsmethoden und -varianten der Bonituren (Parametersatz 1a, heterogene Prüf-gliedeffekte) für Simulationsansätze mit räumlichen Kovarianzstrukturen, mit Auswahl heterogener Restvarianzen im LMM

	Auswertungsmethode	Auswertungsvariante	Konvergenz-verhalten	Einhaltung der nominalen Fehler 1. Art des t-Tests des Prüfgliedkon-trasts PG 1 zu 2
n = 1	LMM für Bonituren	A-BI fix	(- - -)	(-)
		A-BI rand.	(- -)	(-)
n > 1	LMM für Bonituren	A-BI(fix)-Parz. / A-BI(rand.)-Parz.	(+)	(-)
	LMM für Parzellenmittelwerte	A-BI fix	(- - -)	(-)
		A-BI rand.	(- -)	(-)

(-) leichte, (- -) mittelstarke und (- - -) starke Verzerrungen, bzw. Konvergenzprobleme

(+) Einhaltung der Fehler, bzw. 100 %-ige Konvergenz

7 Zusammenfassung

Im landwirtschaftlichen Versuchswesen hat die Bewertung nicht zählbarer oder messbarer Merkmale wie des Pilzbefalls bei Pflanzen oder der Kondition bei Tieren durch Bonituren eine große Bedeutung. Die Bonituren werden in diesen Analysen als geordnete kategoriale Merkmale angesehen. Eine aktuelle Fragestellung aus Sicht der Datenauswertung besteht darin, mit welchen Methoden Bonituren zu analysieren sind, und inwiefern zufällige Effekte mit komplexen Kovarianzstrukturen im Auswertungsmodell berücksichtigt werden können. Komplexe Kovarianzstrukturen sind beispielsweise räumlich bedingte Kovarianzen im Feldversuchswesen. Daraus folgt die Frage nach den Konsequenzen bezüglich der Schätzung fester Effekte sowie Inferenzaussagen über die festen Effekte, wenn diese Kovarianzen im Auswertungsmodell vernachlässigt werden. Aus diesem Problemkreis resultierend ergab sich als Zielstellung dieser Arbeit ein umfassender Methodenvergleich bei Existenz einer räumlichen Kovarianzstruktur. Der Methodenvergleich erfolgte mittels stochastischer Simulation und wurde anhand eines im Feldversuchswesen typischen Versuchsplanes mit Beachtung einer durch räumliche Abhängigkeiten bedingten Kovarianzstruktur durchgeführt. In dieser Arbeit wurden die folgenden methodischen Ansätze zur Auswertung von Bonituren untersucht:

- das Schwellenwertmodell (SWM) (McCullagh, 1980),
- die Anwendung von Rangverfahren (Brunner et al., 2001; 2002),
- die Anwendung von Permutationsverfahren (Schumacher et al., 1995 a, b; 1997),
- das lineare gemischte Modell (LMM) (Henderson, 1990).

Das simulierte Versuchsdesign entspricht einer einfaktoriellen vollständigen Blockanlage mit $r = 4$ Blocks und $a = 20$ Stufen des Behandlungsfaktors und 1, 3 oder 5 Beobachtungen je Parzelle. Es wurden geordnete kategoriale Daten (Boniturnoten) mit 9 und 4 Boniturklassen und unterschiedlichen Parametersätzen bezüglich der vorgegebenen Wahrscheinlichkeiten der Boniturklassen erzeugt. Die Datensätze wurden mit homogenen und verschiedenen Varianten für heterogene Prüfgliedeffekte simuliert. Auch bei Simulation heterogener Prüfgliedeffekte wurde für den Kontrast von Prüfglied 1 zu 2 die Nullhypothese beibehalten. Es wurden 5 Kovarianzstrukturen simuliert: ein nichträumlicher Ansatz und 4 sphärische Ansätze. Der Hauptteil der Analysen erfolgte anhand eines Datensatzes mit 9 Boniturklassen. Je Simulationsvariante wurden 10000 Datensätze erzeugt (mit einer Ausnahme von 1000 Datensätzen aus Gründen der Rechenzeit für einen Modellansatz des SWM). Jedem Datensatz lag eine andere Randomisationsstruktur zugrunde. Sowohl die Erzeugung der simulierten Daten als auch die Auswertungen erfolgten mit der Statistiksoftware SAS, Version 9.2.

Die Möglichkeiten zur Berücksichtigung einer räumlichen Kovarianz sind innerhalb der oben genannten Auswertungsmethoden unterschiedlich, entsprechend ergeben sich innerhalb der Auswertungsmethoden differenzierte Modellansätze bzw. Auswertungsvarianten.

n = 1 Beobachtungen je Parzelle

- SWM, LMM: vollständige Randomisation (A-R); vollständig randomisierte Blockanlage mit Block fix oder zufällig (A-BI fix / A-BI rand.), Modellierung einer sphärischen Kovarianzstruktur (A-cov(sph)).

n > 1 Beobachtungen je Parzelle

- SWM, LMM: zusätzlich mit zufälligen Parzelleneffekt innerhalb der fixen oder zufälligen Blocks (A-BL(fix)-Parz. / A-BL(rand.)-Parz.).

Für das LMM zusätzlich die Beachtung heterogener Restvarianzen in den Auswertungsvarianten A-BI fix / A-BI rand. und A-BL(fix)-Parz. / A-BL(rand.)-Parz.

- Permutations- und Rangverfahren: Auswertungsvarianten A-R und A-BI fix (die Beschränkung ergibt sich, da die Einbeziehung von speziellen Kovarianzstrukturen für diese Modelle noch nicht ausgearbeitet und entsprechend auch nicht rechentechnisch umgesetzt ist).

Die verschiedenen Auswertungsmethoden und -varianten zur Berücksichtigung der räumlichen Kovarianz wurden hinsichtlich der folgenden Kriterien untersucht: Konvergenzverhalten, Treff- und Wiederholungsgenauigkeit der Schätzungen der Varianzkomponenten sowie des Prüfgliedkontrasts von Prüfglied 1 zu Prüfglied 2 und der Wahrscheinlichkeiten für Boniturnoten, Einhaltung des nominalen Fehlers 1. Art für den F- und t-Test. Die nachfolgende Darstellung der Ergebnisse ist nach der Anzahl Beobachtungen je Parzelle und, wenn bedeutsam, bei Beachtung homogener und heterogener Prüfgliedeffekte gegliedert.

n = 1 Beobachtungen je Parzelle

Bei Existenz räumlicher Kovarianzstrukturen führt die Nutzung des SWM mit der Probitfunktion als Linkfunktion für die räumliche Auswertungsvariante (A-cov(sph)) zu extremen Konvergenzproblemen. Bei Nutzung der Logitfunktion zeigen sich nur geringe Konvergenzprobleme, weshalb die weiteren Untersuchungen nur für diese Linkfunktion dargestellt werden.

Die räumlichen Varianzkomponenten können nur innerhalb des SWM oder des LMM mit dem Modellansatz A-cov(sph) geschätzt werden, sie werden jedoch mit beiden Methoden extrem überschätzt.

Bei Existenz räumlicher Kovarianzstrukturen wird der Prüfgliedkontrast bei Nutzung des SWM und des LMM unabhängig vom gewählten Modellansatz erwartungstreu geschätzt. Seine Standardabweichung jedoch wird verzerrt geschätzt. Die Verzerrung ist für das SWM und das LMM mit der räumlichen Auswertungsvariante am geringsten. Bei Existenz räumlicher Kovarianz sind die Schätzungen der Wahrscheinlichkeiten der Boniturnoten unabhängig vom Modellansatz leicht verzerrt (Über- und Unterschätzung) und das Konfidenzniveau wird im Allgemeinen unterschritten.

Das SWM zeigt mit der räumlichen Auswertungsvariante A-cov(sph) für homogene Prüfgliedeffekte eine gute Einhaltung der nominalen Fehler 1. Art für den t-Test des Prüfgliedkontrasts. Für heterogene Prüfgliedeffekte zeigt es leichte Verzerrungen in einem tolerierbaren Bereich (< 1 Prozentpunkt). Die Fehler des F-Tests sind verzerrt. Die Auswertung der Boniturnoten als stetiges Merkmal mittels LMM mit der räumlichen Auswertungsvariante und den Blockmodellen zeigt für homogene Prüfgliedeffekte unverzerrte Fehler für den t- und F-Test, für heterogene Prüfgliedeffekte liefern die Blockmodelle für beide Tests etwas stärker verzerrte Fehler (ca. 2 Prozentpunkte). Das LMM mit der räumlichen Auswertungsvariante dagegen zeigt stärker verzerrte Fehler für den F-Test.

Die Permutations- und Rangverfahren unterschätzen die Fehler 1. Art für den Prüfgliedkontrast deutlich.

n > 1 Beobachtungen je Parzelle

Für heterogene Prüfgliedeffekte zeigen sich für das SWM Konvergenzprobleme, wenn die räumliche Auswertungsvariante oder Modelle mit zufälligem Parzelleneffekt genutzt werden. Auch die Anwendung des LMM auf die Parzellenmittelwerte der Bonituren zeigt Konvergenzprobleme, wenn ein Blockansatz mit Modellierung heterogener Restvarianzen gewählt wird.

Die Schätzung des Prüfgliedkontrasts mit dem SWM und dem LMM ist erwartungstreu, jedoch mit verzerrt geschätzter Standardabweichung. Die Verzerrung ist für das SWM und das LMM am geringsten mit der räumlichen Auswertungsvariante und den Varianten mit zufälligem Parzelleneffekt. Ausnahme ist die Auswertung der Parzellenmittelwerte mit dem LMM, hier führt der räumliche Modellansatz zu extremen Verzerrungen der Schätzung der Standardabweichung des Prüfgliedkontrasts.

Das SWM und das LMM liefern extreme Überschätzungen für die räumlichen Parameter. Die Schätzungen der Wahrscheinlichkeiten der Boniturnoten sind bei Existenz räumlicher Kovarianz leicht verzerrt und das Konfidenzniveau wird mit der räumlichen Auswertungsvariante A-cov(sph) meist über-, mit den anderen Auswertungsvarianten unterschritten.

Bei Existenz räumlicher Kovarianz und für homogene Prüfgliedeffekte liefern die folgenden Methoden zur Auswertung der Bonituren die Einhaltung der nominalen Fehler 1. Art für den t-Test: Das SWM mit der räumlichen Auswertungsvariante oder den Modellen mit Parzelleneffekt und das LMM, angewendet auf die Parzellenmittelwerte der Bonituren, mit den Blockmodellen. Das LMM mit den Modellen mit Parzelleneffekt liefert dagegen zusätzlich die Einhaltung der Fehler 1. Art für den F-Test.

Für heterogene Prüfgliedeffekte liefert das SWM (räumliche Auswertungsvariante oder Modelle mit Parzelleneffekt) und das LMM, angewendet auf die Parzellenmittelwerte der Bonituren (Blockmodelle), leicht verzerrte Fehler 1. Art (Überschätzung von ca. 2 Prozentpunkten). Werden im LMM, angewendet auf die Parzellenmittelwerte der Bonituren, für die Blockmodelle zusätzlich heterogene Restvarianzen angewiesen, kommt es zu extremen Konvergenzverlusten. Die Auswertung der Bonituren mittels LMM dagegen zeigt fast hundertprozentige Konvergenz, und die Überschätzung der nominalen Fehler 1. Art für den Prüfgliedkontrast verringert sich auf ca. 1 Prozentpunkt. Das Rangverfahren überschätzt die Fehler 1. Art des Prüfgliedkontrasts. Die Permutationsverfahren sind für $n > 1$ nicht definiert.

Für die Auswertungspraxis ergeben sich die folgenden Empfehlungen:

Die Existenz räumlich bedingter Kovarianzen führt bei deren Nichtbeachtung im Auswertungsmodell zu Verzerrungen insbesondere der Fehler 1. Art, deren Ausmaß mit Zunahme der räumlichen Kovarianz ansteigt. Bei $n = 1$ stellt die Nutzung des SWM mit räumlicher Kovarianz das Vorzugmodell dar, wobei sich aber aufgrund von Konvergenzproblemen Einschränkungen ergeben. Die Beachtung der räumlichen Kovarianz über einen Blockeffekt stellt in diesem Fall eine Auswertungsalternative dar, die jedoch mit einer Überschätzung des Fehlers 1. Art verbunden ist. Bei $n > 1$ ergibt die Anwendung des SWM erhöhte Konvergenzprobleme. Hier führt das LMM mit heterogenen Restvarianzen zu einer nur geringen Verschätzung des Fehlers 1. Art und stellt daher eine tragfähige Auswertungsalternative dar. Im vorliegenden Sachzusammenhang sollte die Situation $n > 1$ angestrebt werden, um die dann mögliche Alternative eines LMM mit heterogenen Restvarianzen nutzen zu können.

Permutations- und Rangverfahren stellen im Fall räumlicher Kovarianzen keine Auswertungsalternative dar.

8 Summary

The analysis of non-countable or measurable traits – such as the infection of plants with fungi or the body condition of animals – with scores plays an important role in agricultural science. In this analysis, the scores were considered as ordered categorical data. From the perspective of data evaluation, there is the current question of which methods should be used to analyze scores and how random effects with complex covariance structures can be taken into account. One example of complex covariance is spatial covariance in field trials. Hence, there is the question of what the consequences for the estimations and inference statements of the fixed effects are if such covariance is disregarded in the analyzing model. Resulting from this issue, the goal of this work is a comprehensive comparison of methods with the presence of spatial covariance. The comparison of methods was carried out by means of stochastic simulation on the basis of a typical field trial design with regard to covariance structures caused by spatial dependencies. Within this work, the following methodical approaches to analyze scores were evaluated:

- the threshold model (THM) (McCullagh, 1980),
- the rank method (Brunner et al., 2001; 2002),
- the permutation method (Schumacher et al., 1995 a, b; 1997),
- the linear mixed model (LMM) (Henderson, 1990).

The simulated field trial design equates a complete randomized block design with $r = 4$ blocks and $a = 20$ levels for the treatment factor and 1, 3 or 5 observations per plot. The ordered categorical data (scores) were developed with 9 and 4 score classes and with different parameter sets for the cumulative probabilities of the score classes. The datasets were simulated with homogenous and heterogeneous variants of treatment effects. For the simulation of heterogeneous treatment effects the validity of the null hypothesis for the contrast of treatment 1 to treatment 2 was also retained. Five structures were simulated for the covariance: one non-spatial approach and 4 spherical approaches. The main analyses were carried out for a dataset with 9 score classes. For every simulation variant, 10000 datasets were generated (with one exception of 1000 datasets for one model structure of the THM for calculating time reasons). Each dataset was based on a different randomization structure. The simulation of the data as well as the analyses were performed with the statistical software SAS, version 9.2.

The opportunities to address existing spatial covariance in the model structure vary amongst the different methods for analyzing scores mentioned above. Hence, differentiated model structures or analyzing variants result for the different analyzing methods.

n = 1 observations per plot

- THM, LMM: completely randomized design (A-CRD); randomized complete block design with fixed or random blocks (A-RCBD fix / A-RCBD rand.), modeling of a spherical covariance structure (A-cov(sph)).

n > 1 observations per plot

- THM, LMM: with an additional random plot effect within the fixed or random blocks (A-RCBD-plot fix / A-RCBD-plot rand.).

For the LMM, additionally with regard to heterogeneous rest variances in the analyzing variants A-RCBD fix / A-RCBD rand. and A-RCBD-plot fix / A-RCBD-plot rand.

- Permutation and rank methods: analyzing variants A-CRD and A-RCBD fix (the constraint results since the factoring-in of special covariance structures for these models has not yet been worked out and also not implemented in software calculations).

The different analyzing methods and their variants for taking spatial covariance into account were evaluated with regard to the following criteria: convergence properties, unbiasedness and precision of the estimations of the variance components as well as the treatment contrast of treatment 1 to treatment 2 and the probabilities of the scores, and the maintenance of the nominal type one error of the F- and t-test.

The subsequent description of the results is structured according to the observations per plot and, if necessary, by the consideration of homogeneous and heterogeneous treatment effects.

n = 1 observations per plot

If spatial covariance structures exist, the use of the THM with the probitfunction as linkfunction in the case of the spherical covariance structure (A-cov(sph)) analyzing variant leads to extreme convergence problems. The utilization of the logitfunction exhibits only slight convergence problems, for which reason, all further evaluations are described for this linkfunction alone.

The spatial variance components can only be estimated by using the THM or the LMM with the spatial model structure A-cov(sph), however, both methods lead to severe overestimations.

If spatial covariance structures exist, the estimation of the treatment contrast is unbiased for the THM and the LMM independent of the model structure that is chosen. The standard derivation of the contrast is, however, estimated with bias. The bias is smallest if the THM or the LMM with the spatial analyzing variant is used. For existing spatial covariance, the estimations of the probabilities of scores are slightly biased (over- and underestimation) and the confidence level is generally undercut.

The THM with the spatial model structure A-cov(sph) shows a good maintenance of the nominal type one errors for the t-test of the treatment contrast in the case of homogenous treatment effects. In the case of heterogeneous treatment effects, it shows slight bias within an acceptable range (< 1 percentage point).

The evaluation of scores as a continuous trait by usage of the LMM with the spatial model structure or the block designs shows unbiased errors for the t-test and for the F-test in the case of homogeneous treatment effects. In the case of heterogeneous treatment effects, the LMM with the block designs shows only a little bit more bias (ca. 2 percentage points). The LMM with the spatial model structure in contrast shows more heavily biased errors for the F-test.

The permutation method and the rank method considerably underestimate the type one error of the treatment contrast.

n > 1 observations per plot

For heterogeneous treatment effects, the THM shows convergence problems if the analyzing variants with the spatial model structure or the model structures with an additional plot effect are used. The utilization of the LMM, used for the analysis of the

means of the scores, also shows convergence problems if a block design with modeling of heterogeneous rest variances is used.

The estimation of the treatment contrast with the THM or the LMM is unbiased, but does show biased estimations for the standard derivation. The bias is lowest if the THM and the LMM is used with the spatial model structure or the model structures with an additional plot effect. One exception is the analysis of the mean values of the plots with the LMM. This leads to extremely biased estimations of the standard derivation of the treatment contrast if the spatial model structure is used.

The THM and the LMM yield extreme overestimations for the spatial covariance parameters. If spatial covariance exists, the estimations of the probabilities of the scores are slightly biased and the confidence level is usually overestimated when the spatial model structure A-cov(sph) is used, while the other model structures usually underestimate it.

With existing spatial covariance and heterogeneous treatment effects, the following methods for the analysis of scores yield the maintenance of the nominal type one errors for the t-test: the THM with the spatial model structure or with the model structures with an additional plot effect and the LMM, used for the mean values of the plots, with the block designs. In contrast, the LMM with the model structures with an additional plot effect also yields the maintenance of the type one errors for the F-test.

For heterogeneous treatment effects, the THM (spatial model structure or model structures with an additional plot effect) and the usage of the LMM for the mean values of the plots (block designs) yield slightly biased type one errors (overestimation of ca. 2 percentage points).

If the LMM, used for the mean values of the plots, is used with a block design, the assignment of heterogeneous rest variances leads to extreme losses of convergence. The analysis of scores with the LMM in contrast shows nearly hundred percent convergence, and the overestimation of the nominal type one errors for the treatment contrast decreases to ca. 1 percent point.

The rank method overestimates the type one errors for the treatment contrasts. The permutation method is not defined for $n > 1$.

The following suggestions result for the evaluation practice:

The existence of spatially determined covariance leads to bias, especially for the type one errors, if the covariance is not taken into account within the evaluation model. The size of the bias increases with increasing spatial covariance.

For $n = 1$ observation per plot, the THM with the spatial model structure is the preferable model, but does have constraints resulting from convergence problems.

In this case, consideration of the spatial covariance with the help of a block effect is an alternative analysis variant; this is, however, connected with an overestimation of the type one error.

For $n > 1$, the utilization of the THM leads to increasing convergence problems. In this case, the LMM with the assignment of heterogeneous rest variances only leads to a small bias for the type one errors and is, therefore, a good analysis alternative.

Within the given context, the $n > 1$ situation should be aimed for in order to then be able to utilize the alternative of the LMM with heterogeneous rest variances.

In the case of existent spatial covariance, the rank method and the permutation method do not represent possible analysis methods.

Literaturverzeichnis

- Agresti, A., 2002, *Categorical Data Analysis* John Wiley & Sons, Inc., New York.
- Akritas, M. G., and E. Brunner, 1997, A unified approach to rank tests for mixed models: *Journal of Statistical Planning and Inference*, v. 61, no. 2, p. 249-277.
- Bortz, J., G. A. Lienert, and K. Boehnke, 2000, *Verteilungsfreie Methoden in der Biostatistik* Springer-Verlag Berlin Heidelberg.
- Brunner, E., H. Dette, and A. Munk, 1997, Box-type approximations in nonparametric factorial designs: *Journal of the American Statistical Association*, v. 92, no. 440, p. 1494-1502.
- Brunner, E., and U. Munzel, 2002, *Nichtparametrische Datenanalyse* Springer-Verlag Berlin Heidelberg.
- Brunner, E., and M. L. Puri, 2001, Nonparametric methods in factorial designs: *Statistical Papers*, v. 42, no. 1, p. 1-52.
- Burton, A., D. G. Altman, P. Royston, and R. L. Holder, 2006, The design of simulation studies in medical statistics: *Statistics in Medicine*, v. 25, no. 24, p. 4279-4292.
- Conover, W. J., and R. L. Iman, 1981, Rank transformations as a bridge between parametric and nonparametric statistics (with discussion): *American Statistician*, v. 35, no. 3, p. 124-129.
- Cressie, N. A. C., 1993, *Statistics for Spatial Data* New York, Wiley & Sons.
- Edgington, S., 1995, *Randomizations Tests*. Marcel Dekker, New York.
- Frisch, M., E. Schumacher. SAS - Macros zur Durchführung von Permutationstests. 2. Konferenz für SAS-Benutzer in Forschung und Entwicklung (KSFE) , Friedrich-Schiller-Universität Jena. Proceedings der 2. Konferenz der SAS-Anwender in Forschung und Entwicklung , 42-50. 1998. Shaker V.
- Giesbrecht, F. G., and J. C. Burns, 1985, Two-Stage Analysis Based on A Mixed Model - Large-Sample Asymptotic Theory and Small-Sample Simulation Results: *Biometrics*, v. 41, no. 2, p. 477-486.
- Good, P., 1995, *Permutation Tests* Springer, New York.
- Greene, W. H., 2008, *Econometric Analysis* Pearson International Edition.
- Hartley, H. O., J. N. K. Rao, and L. R. Lamotte, 1978, Simple Synthesis-Based Method of Variance Component Estimation: *Biometrics*, v. 34, no. 2, p. 233-242.

Henderson, C. R., 1990, Statistical methods in animal improvement: historical overview, *Advances in Statistical Methods for Genetic Improvement of Livestock*: Springer-Verlag, New York, p. 1-14.

Hörtl, K., K. Thamm, N. Mielenz, J. Spilke. Schätzung kumulativer Wahrscheinlichkeiten geordneter kategorialer Daten bei festgelegter Kategorienanzahl und unterschiedlich leeren Klassen. 105-113. 2010. Shaker Verlag. Proceedings der 14. Konferenz der SAS-Anwender in Forschung und Entwicklung (KSFE).

Hu, X., J. Spilke, and C. Richter, 2006, The influence of spatial covariances on the Type I error and the power for different evaluation methods: *Biometrical Letters*, v. 43, no. 1, p. 19-37.

Irvine, K. M., A. I. Gitelman, and J. A. Hoeting, 2007, Spatial Designs and Properties of Spatial Correlation: Effects on Covariance Estimation: *Journal of Agricultural, Biological, and Environmental Statistics*, v. 12, no. Number 4, p. 450-469.

Kenward, M. G., and J. H. Roger, 1997, Small sample inference for fixed effects from restricted maximum likelihood: *Biometrics*, v. 53, no. 3, p. 983-997.

Kenward, M. G., and J. H. Roger, 2009, An improved approximation to the precision of fixed effects from restricted maximum likelihood: *Computational Statistics & Data Analysis*, v. 53, no. 7, p. 2583-2595.

Khuri, A. I., T. Mathew, and Sinha B.K., 1998, *Statistical Tests for Mixed Linear Models* Wiley, New York.

McCullagh, P., 1980, Regression-Models for Ordinal Data: *Journal of the Royal Statistical Society Series B-Methodological*, v. 42, no. 2, p. 109-142.

McCulloch, C. E., and S. R. Searle, 2001, *Generalized, Linear and Mixed Models* New York, NY, John Wiley & Sons, Inc.

Mielenz, N., H. Krejcová, J. Pribyl, and Schüler.L., 2007, Anpassung eines Fixed Regression Modells für die tägliche Zunahme von Fleckviehbullen mit Hilfe von Informationskriterien: *Archives of Animal Breeding*, v. 50-1, p. 47-58.

Patterson, H. D., and R. Thompson, 1971, Recovery of Inter-Block Information when Block Sizes are Unequal: *Biometrika*, v. 58, no. 3, p. 545-&.

Petrondas, D. A., and K. R. Gabriel, 1983, Multiple Comparisons by Rerandomization Tests: *Journal of the American Statistical Association*, v. 78, no. 384, p. 949-957.

Piepho, H.-P., 1997, Schwellenwertmodelle mit festen und zufälligen Effekten für Boniturdaten aus landwirtschaftlichen Versuchen: *Informatik, Biometrie und Epidemiologie in Medizin und Biologie*, v. 28, no. 3, p. 183-195.

Piepho, H.-P., 2002, Auswertung von Bonituren mit der SAS Prozedur NLMIXED: *Zeitschrift für Agrarinformatik*, v. 10/3, p. 30-41.

Reiczigel, J. Resampling Methoden - Dortmund. 2005.

Richter, C., and B. Kroschewski, 2012, Geostatistical Models in Agricultural Field Experiments: Investigations Based on Uniformity Trials: Agronomy Journal, V. 104, no. 1, p. 91-105.

Ruymgaart, F. H., 1980, A unified approach to the asymptotic distribution theory of certain midrank statistics (with discussion), in Raoult ed., Lecture Notes on Mathematics, No. 821, Springer, Berlin, p. 1-8.

SAS Version 9.2, SAS Institute Inc., Cary, NC, USA, 2009.

SAS Institute Inc., 1988, SAS / IML User's Guide.

Satterthwaite, F., 1941, Synthesis of variance: Psychometrika, v. 6, no. 5, p. 309-316.

Schabenberger, O. Introducing the GLIMMIX Procedure for Generalized Linear Mixed Models. Sugi 30 Proceedings; Philadelphia. 2005.

Schumacher, E., H Bleiholder, H Thöni. Methodische Untersuchungen zur biometrischen Analyse von Boniturwerten aus Freilandversuchen mit Herbiziden. Proc. Vol. 1, 283-290. 1995a. 9th EWRS Symposium, Budapest.

Schumacher, E., and M. Frisch, 1995b, Permutationstests zur Analyse von Boniturwerten in einfachen Versuchsanlagen: Zeitschrift für Agrar informatik, no. 5, p. 107-113.

Schumacher, E., and M. Frisch, 1997, Ein SAS-Macro zur Durchführung von Permutationstests in vollständigen und unvollständigen Blockanlagen: Zeitschrift für Agrar informatik, v. 6, p. 125-130.

Schumacher, E, 2010: persönliche Mitteilung

Searle, S. R., 1971, Linear Models New York, John Wiley & Sons, Inc.

Searle, S. R., G. Casella, and C. E. McCulloch, 1992, Variance Components Wiley, New York.

Shah, D. A., and L. V. Madden, 2004, Nonparametric analysis of ordinal data in designed factorial experiments: Phytopathology, v. 94, no. 1, p. 33-43.

Spilke, J., H. P. Piepho, and X. Y. Hu, 2005, A simulation study on tests of hypotheses and confidence intervals for fixed effects in mixed models for blocked experiments with missing data: Journal of Agricultural, Biological, and Environmental Statistics, v. 10, no. 3, p. 374-389.

Spilke, J., C. Richter, and H.-P. Piepho, 2010, Model selection and its consequences for different split-plot designs with spatial covariance and trend: Plant Breeding, v. 129, p. 590-598.

Stroup, W. W., 2002, Power analysis based on spatial effects mixed models: A tool for comparing design and analysis strategies in the presence of spatial variability: *Journal of Agricultural Biological and Environmental Statistics*, v. 7, no. 4, p. 491-511.

Thöni, H., 1985, Auswertung von Bonituren: ein empirischer Methodenvergleich: *EDV in Medizin und Biologie*, v. 16, no. 3, p. 108-114.

Thöni, H., 1992, Auswertung von Bonituren: Ein empirischer Methodenvergleich. II. Signifikanzprüfung von Prüfglied-Effekten: *Biometrie und Informatik in Medizin und Biologie*, v. 23, no. 3, p. 144-156.

Webster, R., and M. A. Oliver, 2001, *Geostatistics for Environmental Scientists* Wiley & Sons, Ltd. England.

Wolfinger, R., and M. O'Connell, 1993, Generalized Linear Mixed Models - A Pseudo-Likelihood Approach: *Journal of Statistical Computation and Simulation*, v. 48, no. 3-4, p. 233-243.

Verzeichnis der Tabellen und Abbildungen:

Tabellen:

Tab. 3.1: Parameter der simulierten Kovarianzstrukturen, sphärisches Modell	29
Tab. 3.2: Parametersätze für die kumulativen Wahrscheinlichkeiten der Boniturklassen bei Gültigkeit von H_0 für den Kontrast von Prüfglied 1 zu 2	32
Tab. 3.3: Bezeichnungen der Kombination von Parametersätzen der kumulativen Wahrscheinlichkeiten und der Prüfgliedeffekte	32
Tab. 4.1: Übersicht über die Simulationsansätze und angewendete Auswertungsmethoden	47
Tab. 5.1: Anteil konvergenter Simulationsläufe der Auswertung der Bonituren (Parametersatz 1) mittels Schwellenwertmodell und deren Varianten (10000 Simulationsläufe)	50
Tab. 5.2: Bias und beobachtete Standardabweichung des Prüfgliedkontrasts PG 1 zu 2, sowie Bias der geschätzten Standardabweichung in Prozent der beobachteten Standardabweichung, Auswertung der latenten Variablen mittels LMM für Simulationsansätze mit räumlichen Kovarianzstrukturen , 10000 Simulationsläufe ..	54
Tab. 5.3: Bias und beobachtete Standardabweichung des Prüfgliedkontrasts PG 1 zu 2 im linearen Prädiktor, sowie Bias der geschätzten Standardabweichung in Prozent der beobachteten Standardabweichung, Auswertung der Boniturnoten (Parametersatz 1) mittels SWM (Linkfunktion „cumlogit“) für Simulationsansätze mit räumlichen Kovarianzstrukturen , 10000 Simulationsläufe.....	55
Tab. 5.4: Bias und beobachtete Standardabweichung des Prüfgliedkontrasts PG 1 zu 2, sowie Bias der geschätzten Standardabweichung in Prozent der beobachteten Standardabweichung, Auswertung der Parzellenmittelwerte der Boniturnoten (Parametersatz 1) mittels LMM für Simulationsansätze mit räumlichen Kovarianzstrukturen , 10000 Simulationsläufe.....	56
Tab. 5.5: Bias und beobachtete Standardabweichung des Prüfgliedkontrasts PG 1 zu 2, sowie Bias der geschätzten Standardabweichung in Prozent der beobachteten Standardabweichung, Auswertung der Boniturnoten (Parametersatz 1) als stetiges Merkmal mittels LMM für Simulationsansätze mit räumlichen Kovarianzstrukturen , 10000 Simulationsläufe	57
Tab. 5.6: Realisierte Fehler 1. Art der Auswertungsmethoden für Simulationsansätze ohne räumliche Kovarianzstrukturen (t-Test der Prüfglieddifferenz PG 1 zu 2, F- und χ^2 -Test der Globalhypothese unter H_0), Auswertungsvariante: A-R, 10000 Simulationsläufe	67
Tab. 5.7: Realisierte Fehler 1. Art bei Auswertung der latenten Variablen mittels LMM, für Simulationsansätze mit räumlichen Kovarianzstrukturen (t-Test der	

Prüfgliedddifferenz PG 1 zu 2, F- und χ^2 -Test der Globalhypothese unter H0), 10000 Simulationsläufe	69
Tab. 5.8: Realisierte Fehler 1. Art für den linearen Prädiktor bei Auswertung der Boniturnoten (Parametersatz 1) mittels SWM (Linkfunktion „cumprobit“), für Simulationsansätze mit räumlichen Kovarianzstrukturen (t-Test der Prüfgliedddifferenz PG 1 zu 2, F- und χ^2 -Test der Globalhypothese unter H0), 10000 Simulationsläufe..	72
Tab. 5.9: Realisierte Fehler 1. Art für den linearen Prädiktor bei Auswertung der Boniturnoten (Parametersatz 1) mittels SWM (Linkfunktion „cumlogit“), für Simulationsansätze mit räumlichen Kovarianzstrukturen (t-Test der Prüfgliedddifferenz PG 1 zu 2, F- und χ^2 -Test der Globalhypothese unter H0), 10000 Simulationsläufe..	73
Tab. 5.10: Realisierte Fehler 1. Art der Varianten zur Auswertung der Boniturnoten (Parametersatz 1) mittels Rangverfahren, für Simulationsansätze mit räumlichen Kovarianzstrukturen (ANOVA-Typ Statistik der Prüfgliedddifferenz PG 1 zu 2, Test der Globalhypothese mittels ANOVA - und Wald-Typ Statistik unter H0), 10000 Simulationsläufe	76
Tab. 5.11: Realisierte Fehler 1. Art der Varianten zur Auswertung der Boniturnoten (Parametersatz 1) mittels Permutationstests, für Simulationsansätze mit räumlichen Kovarianzstrukturen (Test der Prüfgliedddifferenz PG 1 zu 2 und Test der Globalhypothese unter H0), 10000 Simulationsläufe, 1000 Permutationen.....	76
Tab. 5.12: Korrelationen der Resteffekte benachbarter Parzellen	78
Tab. 5.13: Realisierte Fehler 1. Art der Varianten zur Auswertung der Parzellenmittelwerte der Boniturnoten (Parametersatz 1) mittels LMM, für Simulationsansätze mit räumlichen Kovarianzstrukturen (t-Test der Prüfgliedddifferenz PG 1 zu 2, F- und χ^2 -Test der Globalhypothese unter H0),.....	79
Tab. 5.14: Realisierte Fehler 1. Art der Varianten zur Auswertung der Boniturnoten (Parametersatz 1) als stetige Variable mittels LMM, für Simulationsansätze mit räumlichen Kovarianzstrukturen (t-Test der Prüfgliedddifferenz PG 1 zu 2, F- und χ^2 -Test der Globalhypothese unter H0), 10000 Simulationsläufe	81
Tab. 5.15: Zusammenfassende Bewertung der Ergebnisse der nutzbaren Auswertungsmethoden und -varianten der Boniturnoten (Parametersatz 1, homogene Prüfgliedeffekte) für Simulationsansätze mit räumlichen Kovarianzstrukturen.....	83
Tab. 5.16: Realisierte Fehler 1. Art des t-Tests der Prüfgliedddifferenz PG 1 zu 2, Auswertung der Boniturnoten (k = 9) als stetige Variable mittels LM, 10000 Simulationsläufe, für verschiedene Parametersätze der kumulativen Wahrscheinlichkeiten der Boniturklassen.....	86
Tab. 5.17: Realisierte Fehler 1. Art des t-Tests der Prüfgliedddifferenz PG 1 zu 2 für den linearen Prädiktor, Auswertung der Boniturnoten (k = 9) mittels Schwellenwertmodell (Linkfunktion „cumlogit“), 10000 Simulationsläufe, für verschiedene Parametersätze der kumulativen Wahrscheinlichkeiten der Boniturklassen.....	86

Tab. 5.18: Realisierte Fehler 1. Art des t-Tests der Prüfglieddiffere ⁿ z PG 1 zu 2, Auswertung der Bonituren (k = 4) als stetige Variable mittels LM, 10000 Simulationsläufe, für verschiedene Parametersätze der kumulativen Wahrscheinlichkeiten der Boniturklassen.....	87
Tab. 5.19: Realisierte Fehler 1. Art des t-Tests der Prüfglieddiffere ⁿ z PG 1 zu 2 für den linearen Prädiktor, Auswertung der Bonituren (k = 4) mittels Schwellenwertmodell (Linkfunktion „cumlogit“), 10000 Simulationsläufe, für verschiedene Parametersätze der kumulativen Wahrscheinlichkeiten der Boniturklassen.....	87
Tab. 5.20: Realisierte Fehler 1. Art t-Test der Prüfglieddiffere ⁿ z PG 1 zu 2, Auswertung der Bonituren (9 Noten) als stetige Variable mittels LM und mit Schwellenwertmodell (Linkfunktion „cumlogit“, Fehler für den linearen Prädiktor), 10000 Simulationsläufe, Parametersätze der Wahrscheinlichkeiten mit unterschiedlicher Simulation der Effekte, entsprechend Abschnitt 3.3	88
Tab. 5.21: Realisierte Fehler 1. Art t-Test der Prüfglieddiffere ⁿ z PG 1 zu 2, Auswertung der Bonituren (4 Noten) als stetige Variable mittels LM und mit Schwellenwertmodell (Linkfunktion „cumlogit“, Fehler für den linearen Prädiktor), 10000 Simulationsläufe, Parametersätze der Wahrscheinlichkeiten mit unterschiedlicher Simulation der Effekte, entsprechend Abschnitt 3.3	88
Tab. 5.22: Anteil konvergenter Simulationsläufe und realisierte Fehler 1. Art des t-Tests der Prüfglieddiffere ⁿ z PG 1 zu 2, Auswertung der Bonituren (Parametersatz 1a, heterogene Prüfgliedeffekte) als stetige Variable mittels LMM, mit Auswahl homogener Restvarianzen im Auswertungsmodell, 10000 Simulationsläufe.....	91
Tab. 5.23: Anteil konvergenter Simulationsläufe und realisierte Fehler 1. Art des t-Tests der Prüfglieddiffere ⁿ z PG 1 zu 2, Auswertung der Parzellenmittelwerte der Bonituren (Parametersatz 1a, heterogene Prüfgliedeffekte) als stetige Variable mittels LMM, mit Auswahl homogener Restvarianzen im Auswertungsmodell, 10000 Simulationsläufe	91
Tab. 5.24: Anteil konvergenter Simulationsläufe und realisierte Fehler 1. Art des t-Tests der Prüfglieddiffere ⁿ z PG 1 zu 2, Auswertung der Bonituren (Parametersatz 1a, heterogene Prüfgliedeffekte) mittels Schwellenwertmodell (Linkfunktion „cumlogit“), 10000 Simulationsläufe für die Auswertungen A-BI (fix)-Parz und A-BI (rand.)-Parz. , 1000 Simulationsläufe für die Auswertung A-cov(sph).....	92
Tab. 5.25: Anteil konvergenter Simulationsläufe und realisierte Fehler 1. Art des t-Tests der Prüfglieddiffere ⁿ z PG 1 zu 2, Auswertung der Bonituren (Parametersatz 1a, heterogene Prüfgliedeffekte) als stetige Variable mittels LMM, mit Auswahl heterogener Restvarianzen im Auswertungsmodell, 1000 Simulationsläufe.....	93
Tab. 5.26: Anteil konvergenter Simulationsläufe und realisierte Fehler 1. Art des t-Tests der Prüfglieddiffere ⁿ z PG 1 zu 2, Auswertung der Parzellenmittelwerte der Bonituren (Parametersatz 1a, heterogene Prüfgliedeffekte) als stetige Variable mittels LMM, mit Auswahl heterogener Restvarianzen im Auswertungsmodell, 1000 Simulationsläufe	93

Tab. 6.1: Anteil der Simulationsläufe mit nicht positiv definiten G-Matrix in Prozent der konvergenten Simulationsläufe (Schwellenwertmodell mit Linkfunktion „cumlogit“ für Simulationsansätze mit räumlicher Kovarianz, Parametersatz 1) 99

Tab. 6.2 Zusammenfassende Bewertung der Ergebnisse der nutzbaren Auswertungsmethoden und -varianten der Bonituren (Parametersatz 1a, heterogene Prüfgliedeffekte) für Simulationsansätze mit räumlichen Kovarianzstrukturen, mit Auswahl homogener Restvarianzen im LMM 111

Tab. 6.3: Zusammenfassende Bewertung der Ergebnisse der nutzbaren Auswertungsmethoden und -varianten der Bonituren (Parametersatz 1a, heterogene Prüfgliedeffekte) für Simulationsansätze mit räumlichen Kovarianzstrukturen, mit Auswahl heterogener Restvarianzen im LMM 112

Abbildungen:

Abb. 3.1: Versuchsplan der simulierten vollständigen Blockanlage mit 4 Wiederholungen und 20 Prüfgliedern (Parzellen) pro Block 28

Abb. 3.2: Anordnung der Beobachtungen innerhalb der Parzelle und deren Nummerierung 28

Abb. 3.3: Korrelationsfunktionen der räumlichen Simulationsansätze in Abhängigkeit von der euklidischen Distanz zweier Beobachtungen, sphärisches Modell 30

Abb. 5.1: Parameter (P_{kWK}), Konfidenzintervalle sowie empirische Überdeckungswahrscheinlichkeiten der kumulativen Wahrscheinlichkeiten der Boniturnoten (Parametersatz 1) für die Auswertung der Simulationsansätze ohne räumliche Kovarianz mittels SWM (Linkfunktion „cumlogit“) Auswertungsvariante A-R, für Ansätze mit $n = 1, 3$ und 5 Beobachtungen. (KI: Konfidenzintervalle berechnet für die Wahrscheinlichkeiten, KI_{logit} : Berechnung der Konfidenzintervalle im linearen Prädiktor und Rücktransformation auf die Boniturskala) 60

Abb. 5.2: Halbe Breiten der Konfidenzintervalle der kumulativen Wahrscheinlichkeiten für die Boniturnoten (Parametersatz 1) der Auswertung der Simulationsansätze ohne räumliche Kovarianz mittels SWM, Auswertungsvariante A-R und Simulationsansätze mit räumlicher Kovarianz mittels SWM, Auswertungsvariante A-cov(sph), für Ansätze mit $n = 1, 3$ und 5 Beobachtungen 63

Abb. 5.3: Empirische Überdeckungswahrscheinlichkeiten der Konfidenzintervalle der kumulativen Wahrscheinlichkeiten für die Boniturnoten (Parametersatz 1) der Auswertung der Simulationsansätze ohne räumliche Kovarianz mittels SWM, Auswertungsvariante A-R und Simulationsansätze mit räumlicher Kovarianz mittels SWM, Auswertungsvariante A-cov(sph), für Ansätze mit $n = 1, 3$ und 5 Beobachtungen 64

Abb. 6.1: Schematische Darstellung der durch nicht positiv definite G-Matrix oder negative Eigenwerte der Kovarianzmatrix der festen Effekte entstehenden Probleme

bei der Ergebnisbildung mit der Auswertungsvariante A-cov(sph) (n = Beobachtungen je Parzelle, a = Behandlungen; r = Wiederholung; $N = n*a*r$ 101

Abb. 6.2: Verteilung der t-Werte der Auswertung der latenten Variablen mittels LMM (Auswertungsvarianten A-cov(sph) und A-BL fix, 5 Beobachtungen je Parzelle, räumliche Korrelation (2_40_1), gestrichelte Grenzen zeigen die mittleren 95% der t-Verteilung bei dem in der Auswertungsvariante A-BL fix für die Tests verwendeten Freiheitsgrad von $FG=377$)..... 106

Abb. 6.3: Verteilung der geschätzten Prüfglieddiffenzen der Auswertung der latenten Variablen mittels LMM (Auswertungsvarianten A-cov(sph) und A-BL fix, 5 Beobachtungen je Parzelle, und räumliche Korrelation (2_40_1)) 107

Abb. 6.4: Verteilung der geschätzten Standardabweichungen der Prüfglieddiffenzen der Auswertung der latenten Variablen mittels LMM (Auswertungsvarianten A-cov(sph) und A-BL fix, 5 Beobachtungen je Parzelle und räumliche Korrelation (2_40_1))..... 107

Anhang 1: Berechnung der Standardfehler für marginale Wahrscheinlichkeiten nach der Delta-Methode (Greene, 2008; persönliche Mitteilung Mielenz, 2011)

A1-1 a) 1 zufälliger Faktor (z.B. Block zufällig), logit Linkfunktion

Gegeben:

$$q^* = \frac{\exp(a \cdot \eta)}{1 + \exp(a \cdot \eta)} = q^*(\sigma_{\text{Bl}}^2, \eta) \text{ mit } \eta = x' \beta \text{ und } a = (1 + \tilde{c}^2 \cdot \sigma_{\text{Bl}}^2)^{-\frac{1}{2}}, \tilde{c} = \frac{16 \sqrt{3}}{15 \pi}$$

Für die gesuchte $\text{Var}(q^*)$ gilt nach der Delta-Methode:

$$\text{Var}(\hat{q}^*) = \left(\frac{\partial q^*}{\partial \eta} \right)^2 \cdot \text{Var}(\hat{\eta}) + \left(\frac{\partial q^*}{\partial \sigma_{\text{Bl}}^2} \right)^2 \cdot \text{Var}(\hat{\sigma}_{\text{Bl}}^2)$$

$$\frac{\partial q^*}{\partial \eta} = \frac{\partial q^*}{\partial u} \cdot \frac{\partial u}{\partial \eta} \text{ mit } u = a \cdot \eta$$

$$\frac{\partial q^*}{\partial u} = e^u (1 + e^u)^{-1} + e^u (-1) (1 + e^u)^{-2} e^u = e^u (1 + e^u)^{-1} (1 - e^u (1 + e^u)^{-1}) = q^* (1 - q^*)$$

$$\frac{\partial q^*}{\partial \eta} = q^* (1 - q^*) a$$

$$\frac{\partial q^*}{\partial \sigma_{\text{Bl}}^2} = \frac{\partial q^*}{\partial u} \cdot \frac{\partial u}{\partial \sigma_{\text{Bl}}^2} \text{ mit } u = a \cdot \eta$$

$$\frac{\partial u}{\partial \sigma_{\text{Bl}}^2} = \eta \cdot \frac{\partial a}{\partial \sigma_{\text{Bl}}^2}$$

$$\frac{\partial a}{\partial \sigma_{\text{Bl}}^2} = -\frac{1}{2} (1 + \tilde{c}^2 \cdot \sigma_{\text{Bl}}^2)^{-\frac{3}{2}} \cdot \tilde{c}^2 = -\frac{1}{2} \cdot \tilde{c}^2 \left[(1 + \tilde{c}^2 \cdot \sigma_{\text{Bl}}^2)^{-\frac{1}{2}} \right]^3 = (-1) \frac{\tilde{c}^2}{2} \cdot a^3$$

$$\frac{\partial q^*}{\partial \sigma_{\text{Bl}}^2} = q^* (1 - q^*) \cdot \eta \cdot (-1) \frac{\tilde{c}^2}{2} \cdot a^3$$

$$\text{Var}(\hat{q}^*) = \left[\hat{q}^* \cdot (1 - \hat{q}^*) \cdot \hat{a} \right]^2 \cdot \text{Var}(\hat{\eta}) + \left[\hat{q}^* \cdot (1 - \hat{q}^*) \cdot \hat{\eta} \cdot (-1) \cdot \frac{\tilde{c}^2}{2} \hat{a}^3 \right]^2 \cdot \text{Var}(\hat{\sigma}_{\text{Bl}}^2)$$

A1-1 b) 2 zufällige Faktoren (z.B. Block zufällig, Parzelle zufällig),

logit Linkfunktion

Gegeben: $q^* = \frac{\exp(a \cdot \eta)}{1 + \exp(a \cdot \eta)} = q^*(\sigma_{\text{Parz}}^2, \sigma_{\text{Bl}}^2, \eta)$ mit $a = (1 + \tilde{c}^2 \cdot (\sigma_{\text{Parz}}^2 + \sigma_{\text{Bl}}^2))^{-\frac{1}{2}}$

$$\text{Var}(\hat{q}^*) = \left(\frac{\partial q^*}{\partial \eta} \right)^2 \cdot \text{Var}(\hat{\eta}) + \left(\frac{\partial q^*}{\partial a} \right)^2 \cdot \text{Var}(\hat{a})$$

$$\frac{\partial q^*}{\partial \eta} = q^* (1 - q^*) a$$

$$\frac{\partial q^*}{\partial a} = q^* (1 - q^*) \eta$$

$$\text{Var}(\hat{a}) = \left(\frac{\partial a}{\partial \sigma_{\text{Parz}}^2} \right)^2 \cdot \text{Var}(\hat{\sigma}_{\text{Parz}}^2) + 2 \cdot \frac{\partial a}{\partial \sigma_{\text{Parz}}^2} \cdot \frac{\partial a}{\partial \sigma_{\text{Bl}}^2} \text{Cov}(\hat{\sigma}_{\text{Parz}}^2, \hat{\sigma}_{\text{Bl}}^2) + \left(\frac{\partial a}{\partial \sigma_{\text{Bl}}^2} \right)^2 \cdot \text{Var}(\hat{\sigma}_{\text{Bl}}^2)$$

$$\frac{\partial a}{\partial \sigma_{\text{Parz}}^2} = -\frac{1}{2} \cdot (1 + \tilde{c}^2 (\sigma_{\text{Parz}}^2 + \sigma_{\text{Bl}}^2))^{-\frac{3}{2}} = -\frac{1}{2} a^3 \cdot \tilde{c}^2$$

$$\frac{\partial a}{\partial \sigma_{\text{Bl}}^2} = -\frac{1}{2} \cdot (1 + \tilde{c}^2 (\sigma_{\text{Parz}}^2 + \sigma_{\text{Bl}}^2))^{-\frac{3}{2}} = -\frac{1}{2} a^3 \cdot \tilde{c}^2$$

Einsetzen liefert:

$$\text{Var}(\hat{a}) = \left(-\frac{1}{2} a^3 \cdot \tilde{c}^2 \right)^2 \cdot [\text{Var}(\hat{\sigma}_{\text{Parz}}^2) + 2 \cdot \text{Cov}(\hat{\sigma}_{\text{Parz}}^2, \hat{\sigma}_{\text{Bl}}^2) + \text{Var}(\hat{\sigma}_{\text{Bl}}^2)]$$

Also:

$$\text{Var}(\hat{q}^*) = [\hat{q}^* \cdot (1 - \hat{q}^*) \cdot \hat{a}]^2 \cdot \text{Var}(\hat{\eta}) + [\hat{q}^* \cdot (1 - \hat{q}^*) \cdot \hat{\eta}]^2 \cdot \left(-\frac{1}{2} \hat{a}^3 \cdot \tilde{c}^2 \right)^2 \cdot [\text{Var}(\hat{\sigma}_{\text{Parz}}^2) + 2 \cdot \text{Cov}(\hat{\sigma}_{\text{Parz}}^2, \hat{\sigma}_{\text{Bl}}^2) + \text{Var}(\hat{\sigma}_{\text{Bl}}^2)]$$

A1-2 a) 1 zufälliger Faktor (z.B. Block zufällig), probit Linkfunktion

gegeben: $q^* = \Phi(\hat{c} \cdot \hat{\eta})$ mit $c = (1 + \sigma_{\text{Bl}}^2)^{-\frac{1}{2}}$ gesucht $\text{Var}(\hat{q}^*) = \Phi(\hat{c} \cdot \hat{\eta})$

nach der Delta-Methode gilt: $\text{Var}(\hat{q}^*) = \left(\frac{\partial \Phi}{\partial \eta}\right)^2 \cdot \text{Var}(\hat{\eta}) + \left(\frac{\partial \Phi}{\partial \sigma_{\text{Bl}}^2}\right)^2 \cdot \text{Var}(\sigma_{\text{Bl}}^2)$

$$\frac{\partial \Phi}{\partial \eta} = \frac{\partial \Phi}{\partial u} \cdot \frac{\partial u}{\partial \eta} \quad \text{mit } u = c \cdot \eta$$

$$\Phi(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^u e^{-\frac{x^2}{2}} dx \quad \frac{\partial \Phi}{\partial u} = \varphi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$$

$$\frac{\partial \Phi}{\partial \eta} = \varphi(c \cdot \eta) \cdot c$$

$$\frac{\partial \Phi}{\partial \sigma_{\text{Bl}}^2} = \frac{\partial \Phi}{\partial u} \cdot \frac{\partial u}{\partial \sigma_{\text{Bl}}^2} = \varphi(c \cdot \eta) \cdot \eta \cdot \frac{\partial u}{\partial \sigma_{\text{Bl}}^2}; \quad \frac{\partial u}{\partial \sigma_{\text{Bl}}^2} = \frac{\partial u}{\partial c} \cdot \frac{\partial c}{\partial \sigma_{\text{Bl}}^2} = \eta \cdot \frac{\partial c}{\partial \sigma_{\text{Bl}}^2}$$

$$\frac{\partial c}{\partial \sigma_{\text{Bl}}^2} = -\frac{1}{2} (1 + \sigma_{\text{Bl}}^2)^{-\frac{1}{2}-1} = -\frac{1}{2} (1 + \sigma_{\text{Bl}}^2)^{-\frac{3}{2}} = -\frac{1}{2} \left[(1 + \sigma_{\text{Bl}}^2)^{-\frac{1}{2}} \right]^3 = -\frac{1}{2} c^3$$

Also: $\frac{\partial \Phi}{\partial \sigma_{\text{Bl}}^2} = \varphi(c \cdot \eta) \cdot \eta \cdot \left(-\frac{1}{2}\right) \cdot c^3$ und $\frac{\partial \Phi}{\partial \eta} = \varphi(c \cdot \eta) \cdot c$

$$\text{Var}(\hat{q}^*) = [\varphi(c \cdot \hat{\eta}) \cdot c]^2 \cdot \text{Var}(\hat{\eta}) + \left[\varphi(c \cdot \hat{\eta}) \cdot \hat{\eta} \cdot \left(-\frac{1}{2}\right) \cdot c^3 \right]^2 \cdot \text{Var}(\hat{\sigma}_{\text{Bl}}^2)$$

**A1-2 b) 2 zufällige Faktoren (z.B. Block zufällig, Parzelle zufällig),
probit Linkfunktion**

gegeben: $q^* = \Phi(a \cdot \eta)$ mit $a = (1 + (\sigma_{\text{Parz}}^2 + \sigma_{\text{Bl}}^2))^{-\frac{1}{2}}$ gesucht $\text{Var}(\hat{q}^*) = \Phi(\hat{a} \cdot \hat{\eta})$

nach der Delta-Methode gilt: $\text{Var}(\hat{q}^*) = \left(\frac{\partial \Phi}{\partial \eta}\right)^2 \cdot \text{Var}(\hat{\eta}) + \left(\frac{\partial \Phi}{\partial a}\right)^2 \cdot \text{Var}(\hat{a})$

a) $\frac{\partial \Phi}{\partial \eta} = \frac{\partial \Phi}{\partial u} \cdot \frac{\partial u}{\partial \eta} = \varphi(a \cdot \eta) \cdot a$ mit $u = a \cdot \eta$

$$\frac{\partial \Phi}{\partial a} = \frac{\partial \Phi}{\partial u} \cdot \frac{\partial u}{\partial a} = \varphi(a \cdot \eta) \cdot \eta$$

b)

$$\text{Var}(\hat{a}) = \left(\frac{\partial a}{\partial \sigma_{\text{Parz}}^2}\right)^2 \cdot \text{Var}(\hat{\sigma}_{\text{Parz}}^2) + 2 \cdot \frac{\partial a}{\partial \sigma_{\text{Parz}}^2} \cdot \frac{\partial a}{\partial \sigma_{\text{Bl}}^2} \text{Cov}(\sigma_{\text{Parz}}^2, \sigma_{\text{Bl}}^2) + \left(\frac{\partial a}{\partial \sigma_{\text{Bl}}^2}\right)^2 \cdot \text{Var}(\hat{\sigma}_{\text{Bl}}^2)$$

$$\frac{\partial a}{\partial \sigma_{\text{Parz}}^2} = -\frac{1}{2} \cdot (1 + (\sigma_{\text{Parz}}^2 + \sigma_{\text{Bl}}^2))^{-\frac{3}{2}} = -\frac{1}{2} a^3$$

$$\frac{\partial a}{\partial \sigma_{\text{Bl}}^2} = -\frac{1}{2} \cdot (1 + (\sigma_{\text{Parz}}^2 + \sigma_{\text{Bl}}^2))^{-\frac{3}{2}} = -\frac{1}{2} a^3$$

beide Ableitungen identisch, da innere Ableitung jeweils = 1

Einsetzen liefert:

$$\text{Var}(\hat{a}) = \left(-\frac{1}{2} a^3\right)^2 \cdot [\text{Var}(\sigma_{\text{Parz}}^2) + 2 \cdot \text{Cov}(\sigma_{\text{Parz}}^2, \sigma_{\text{Bl}}^2) + \text{Var}(\sigma_{\text{Bl}}^2)]$$

Also:

$$\text{Var}(\hat{q}^*) = (\varphi(\hat{a} \cdot \hat{\eta}) \cdot \hat{a})^2 \cdot \text{Var}(\hat{\eta}) + (\varphi(\hat{a} \cdot \hat{\eta}) \cdot \hat{\eta})^2 \left(\left(-\frac{1}{2} \hat{a}^3\right)^2 \cdot [\text{Var}(\hat{\sigma}_{\text{Parz}}^2) + 2 \cdot \text{Cov}(\hat{\sigma}_{\text{Parz}}^2, \hat{\sigma}_{\text{Bl}}^2) + \text{Var}(\hat{\sigma}_{\text{Bl}}^2)] \right)$$

A1-3 a) 1 fixer (Blocks) und 1 zufälliger Faktor (Parzellen), logit-Linkfunktion (Mielenz)

Die Berechnung der marginalen Wahrscheinlichkeiten erfolgt durch Akkumulation der Schätzungen der marginalen Wahrscheinlichkeiten je Block. Für eine beliebige Boniturnote sei die marginale Wahrscheinlichkeit einer Behandlung i im l-ten Block

$$q_{li}^* = \frac{\exp(a \cdot \eta_{li})}{1 + \exp(a \cdot \eta_{li})} \quad \text{mit} \quad \eta_{li} = \mu + \alpha_1 + \beta_1 \text{ als marginalem Erwartungswert}$$

$$\text{und} \quad a = (1 + \tilde{c}^2 \cdot \sigma_{\text{Parz}}^2)^{-\frac{1}{2}}, \quad \tilde{c} = \frac{16 \sqrt{3}}{15 \pi}$$

Es wird also für jede Kombination von Block und Prüfglied (und für jede beliebige Boniturnote) ein linearer Prädiktor geschätzt. Daraus werden die marginalen Wahrscheinlichkeiten geschätzt und diese werden über die 4 Blocks gemittelt. Die mittlere marginale Wahrscheinlichkeit erhält man für r = 4 Blocks wie folgt:

$$\bar{q}_i^* = \frac{1}{r} \sum_{l=1}^r q_{li}^* = x' q^* \quad \text{mit} \quad x = \begin{bmatrix} 0.25 \\ 0.25 \\ 0.25 \\ 0.25 \end{bmatrix} \quad \text{und} \quad q^* = \begin{bmatrix} q_{1i}^* \\ q_{2i}^* \\ q_{3i}^* \\ q_{4i}^* \end{bmatrix}$$

Die Varianz der Schätzung \bar{q}_i^* ergibt sich wiederum mittels der Delta-Methode:

$$\text{Var}(\bar{q}_i^*) = x' \text{Cov}(q^*) x \quad \text{mit} \quad \text{Cov}(q^*) = \text{Cov}(q_{li}^*; q_{l'i}^*) \quad l, l' = 1, \dots, r$$

$$\text{Cov}(q_{li}^*; q_{l'i}^*) = \frac{\partial q_{li}^*}{\partial \eta_{li}} \cdot \frac{\partial q_{l'i}^*}{\partial \eta_{l'i}} \cdot \text{Cov}(\hat{\eta}_{li}, \hat{\eta}_{l'i}) + \frac{\partial q_{li}^*}{\partial \sigma_{\text{Parz}}^2} \cdot \frac{\partial q_{l'i}^*}{\partial \sigma_{\text{Parz}}^2} \cdot \text{Var}(\sigma_{\text{Parz}}^2)$$

$$\frac{\partial q_{li}^*}{\partial \eta_{li}} = \hat{q}_{li}^* (1 - \hat{q}_{li}^*) \cdot \hat{a}$$

$$\frac{\partial q_{l'i}^*}{\partial \eta_{l'i}} = \hat{q}_{l'i}^* (1 - \hat{q}_{l'i}^*) \cdot \hat{a}$$

$$\frac{\partial q_{li}^*}{\partial \sigma_{\text{Parz}}^2} = \hat{q}_{li}^* (1 - \hat{q}_{li}^*) \cdot \hat{\eta}_{li} (-1) \cdot \frac{\tilde{c}^2}{2} \hat{a}^3$$

$$\frac{\partial q_{l'i}^*}{\partial \sigma_{\text{Parz}}^2} = \hat{q}_{l'i}^* (1 - \hat{q}_{l'i}^*) \cdot \hat{\eta}_{l'i} (-1) \cdot \frac{\tilde{c}^2}{2} \hat{a}^3$$

$$\text{Cov}(\hat{\eta}_{li}, \hat{\eta}_{l'i}) = k' \cdot c_{ll} \cdot k \quad \text{mit} \quad c_{ll} = \text{Kovarianzmatrix der festen Effekte im Modell}$$

$k_{s,r}$ = Matrix aus 0 und 1 mit s = Anzahl fester Effekte, r = Anzahl Blocks

Anhang 2: Histogramme der Parametersätze für die Wahrscheinlichkeiten der Boniturklassen gemäß Tabelle 3.2

a) $k = 9$ Boniturklassen

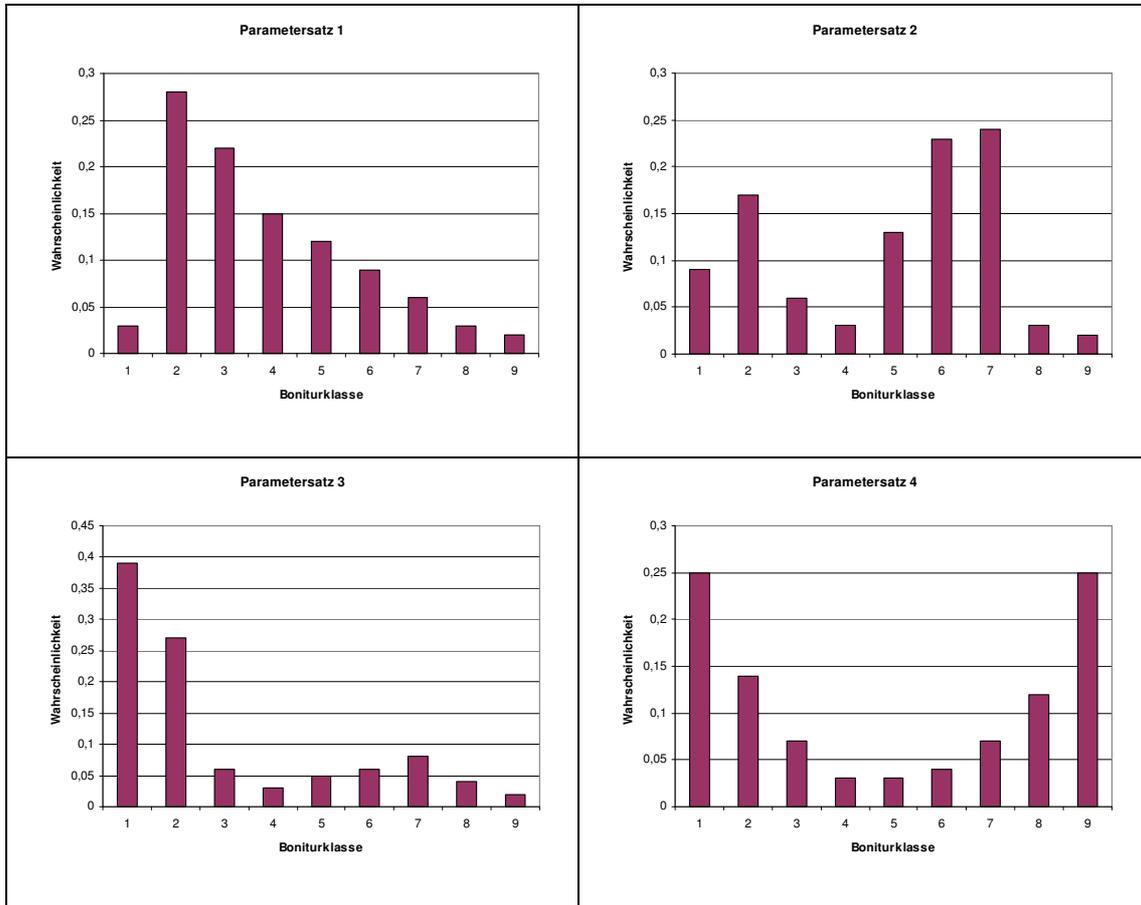


Abbildung A2a: Parametersätze für $k = 9$ Boniturklassen bei Gültigkeit von H_0

b) $k=4$ Boniturklassen

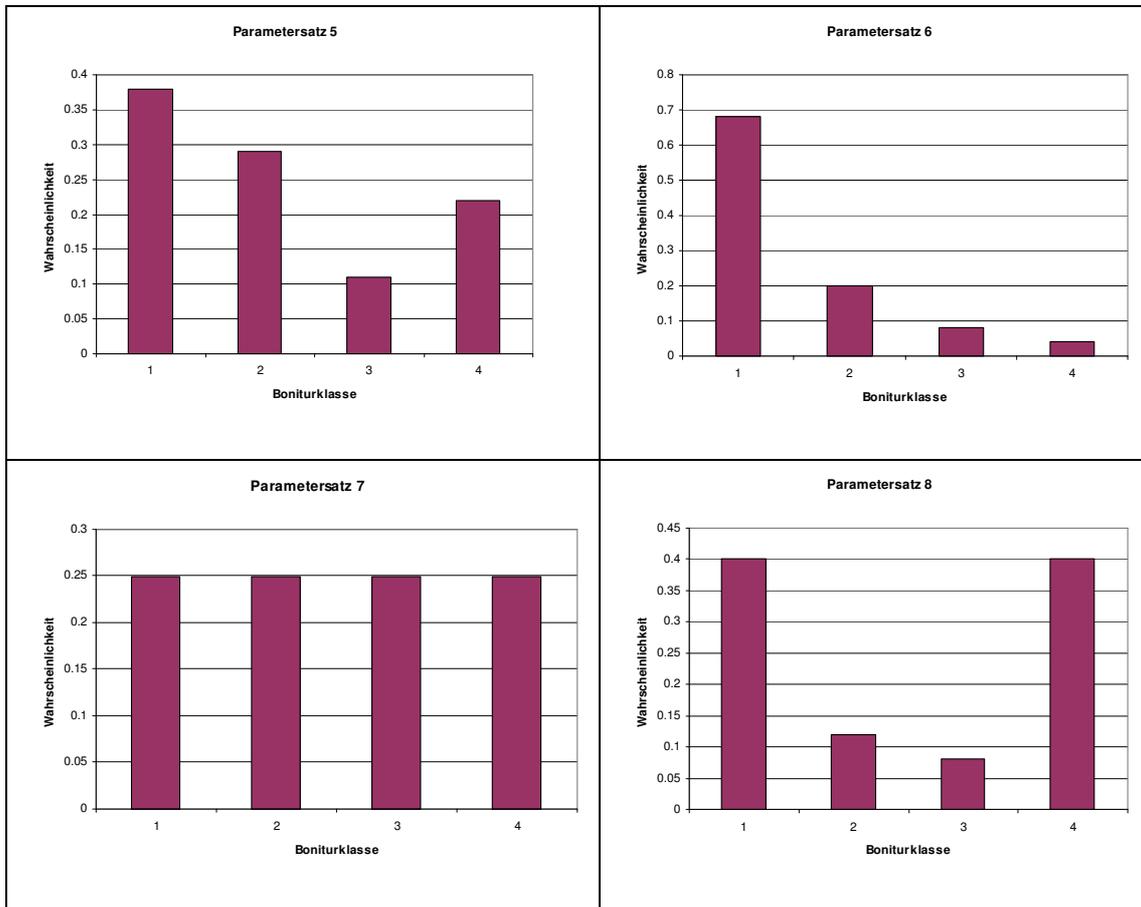


Abbildung A2b: Parametersätze für $k = 4$ Boniturklassen bei Gültigkeit von H_0

Anhang 3: Varianten der Simulation der Prüfgliedeffekte

Die Prüfgliedeffekte werden jeweils unter Gültigkeit der Nullhypothese für den Kontrast von Prüfglied 1 zu 2 simuliert.

Mit homogenen Prüfgliedeffekten, also $\alpha_1 - \alpha_{20} = 0$

oder mit heterogenen Prüfgliedeffekten gemäß der folgenden Varianten:

- a) $\alpha_1 = \alpha_2 = 1$ und alle anderen $\alpha_3 = \alpha_4 = \dots = \alpha_{20} = 0.2$
- b) $\alpha_1 = \alpha_2 = 0.5$ und alle anderen $\alpha_3 = \alpha_4 = \dots = \alpha_{20} = 0.2$
- c) $\alpha_1 = \alpha_2 = 0.5$,
- | | |
|------------------------|------------------------|
| $\alpha_3 = 0.2$, | $\alpha_4 = 0.3$, |
| $\alpha_5 = 0.5$, | $\alpha_6 = 0.6$, |
| $\alpha_7 = 0.25$, | $\alpha_8 = 0.15$, |
| $\alpha_9 = 0.1$, | $\alpha_{10} = 0.7$, |
| $\alpha_{11} = 0.28$, | $\alpha_{12} = 0.39$, |
| $\alpha_{13} = 0.4$, | $\alpha_{15} = 0.5$, |
| $\alpha_{16} = 0.32$, | $\alpha_{17} = 0.27$, |
| $\alpha_{18} = 0.9$, | $\alpha_{19} = 0.6$, |
| $\alpha_{20} = 0.8$ | |

Anhang 4: Bias der räumlichen Kovarianzparameter für die Auswertung der Simulationsansätze mit räumlicher Kovarianz für die Auswertung der latenten Variablen mittels LMM und der Bonituren mittels SWM (Linkfunktion „cumlogit“), (A-cov(sph))

Tab. A4-1: Bias (in % vom Parameter) von partiellem Sill und Range der Auswertung der latenten Variablen mittels LMM

Simulierter räumlicher Kovarianzansatz	1_10_1		1_20_1		1_30_1		2_40_1	
	Bias % part. Sill	Bias % Range						
n = 1	374	22020	2072	19294	4209	13550	4782	9231
n = 3	89	4699	1596	2298	1411	1499	1142	1099
n = 5	1643	7899	3622	3898	2630	2566	1992	1899

Tab. A4-2: Bias (in % vom Parameter) von partiellem Sill und Range der Auswertung der Bonituren mittels SWM (Linkfunktion „cumlogit“)

Simulierter räumlicher Kovarianzansatz	1_10_1		1_20_1		1_30_1		2_40_1	
	Bias % part. Sill	Bias % Range						
n = 1	21300	10270	65623	5189	93707	3785	63701	4033
n = 3	1090	4035	14366	2219	8274	998	3229	456
n = 5	7710 *)	8069 *)	16712	3868	11570	2413	6885	1523

*) : Ergebnisse von 1000 Läufen, da Simulationen mit 10000 Läufen aufgrund sehr langer Laufzeiten nicht realisierbar waren

Anhang 5: Geschätzter Prüfgliedkontrast von PG 1 zu 2, sowie seine 95 %-Konfidenzintervalle für die in Abschnitt 5.3 besprochenen Auswertungsmethoden, in Abhängigkeit der konvergenten Simulationsläufe

Tab. A5-1: Prüfgliedkontrasts PG 1 zu 2, sowie untere und obere Grenze seines 95 % Konfidenzintervalls der Auswertung der latenten Variablen mittels LMM

Simulierter räumlicher Kovarianzansatz		1_10_1			1_20_1			1_30_1			2_40_1		
Auswertung		PG Kontrast	KI unten	KI oben	PG Kontrast	KI unten	KI oben	PG Kontrast	KI unten	KI oben	PG Kontrast	KI unten	KI oben
n = 1	A-R	-0.0010	-0.0206	0.0186	-0.0013	-0.0206	0.0180	-0.0009	-0.0196	0.0178	0.0006	-0.0210	0.0222
	A-cov(sph)	-0.0014	-0.0211	0.0183	-0.0014	-0.0204	0.0176	-0.0033	-0.0212	0.0146	-0.0029	-0.0219	0.0161
	A-BI fix	-0.0010	-0.0206	0.0186	-0.0013	-0.0206	0.0180	-0.0009	-0.0196	0.0178	0.0006	-0.0210	0.0222
	A-BI (fix)-Parz	---			---			---			---		
	A-BI random	-0.0010	-0.0206	0.0186	-0.0013	-0.0206	0.0180	-0.0009	-0.0196	0.0178	0.0006	-0.0210	0.0222
	A-BI (rand.)-Parz	---			---			---			---		
n = 3	A-R	-0.0058	-0.0188	0.0072	-0.0057	-0.0193	0.0079	-0.0052	-0.0187	0.0083	-0.0037	-0.0204	0.0130
	A-cov(sph)	-0.0081	-0.0208	0.0046	-0.0091	-0.0212	0.0030	-0.0090	-0.0200	0.0020	-0.0083	-0.0200	0.0034
	A-BI fix	-0.0058	-0.0184	0.0068	-0.0057	-0.0193	0.0079	-0.0052	-0.0187	0.0083	-0.0037	-0.0204	0.0130
	A-BI (fix)-Parz.	-0.0058	-0.0184	0.0068	-0.0057	-0.0193	0.0079	-0.0052	-0.0187	0.0083	-0.0037	-0.0204	0.0130
	A-BI random	-0.0058	-0.0184	0.0068	-0.0057	-0.0193	0.0079	-0.0052	-0.0187	0.0083	-0.0037	-0.0204	0.0130
	A-BI (rand.)-Parz.	-0.0058	-0.0184	0.0068	-0.0057	-0.0193	0.0079	-0.0052	-0.0187	0.0083	-0.0067	-0.0234	0.0100
n = 5	A-R	-0.0018	-0.0117	0.0081	-0.0023	-0.0139	0.0093	-0.0021	-0.0137	0.0095	-0.0007	-0.0158	0.0144
	A-cov(sph)	-0.0052	-0.0148	0.0044	-0.0050	-0.0131	0.0031	-0.0050	-0.0123	0.0023	-0.0046	-0.0125	0.0033
	A-BI fix	-0.0018	-0.0117	0.0081	-0.0023	-0.0139	0.0093	-0.0021	-0.0137	0.0095	-0.0007	-0.0158	0.0144
	A-BI (fix)-Parz.	-0.0032	-0.0137	0.0073	-0.0037	-0.0158	0.0084	-0.0035	-0.0157	0.0087	-0.0021	-0.0176	0.0134
	A-BI random	-0.0018	-0.0117	0.0081	-0.0023	-0.0139	0.0093	-0.0021	-0.0137	0.0095	-0.0007	-0.0158	0.0144
	A-BI (rand.)-Parz.	-0.0032	-0.0137	0.0073	-0.0037	-0.0158	0.0084	-0.0035	-0.0157	0.0087	-0.0021	-0.0176	0.0134

Tab. A5-2: Prüfgliedkontrasts PG 1 zu 2, sowie untere und obere Grenze seines 95 % Konfidenzintervalls der Auswertung der Boniturnoten (Parametersatz 1) mittels SWM (Linkfunktion „cumlogit“)

Simulierter räumlicher Kovarianzansatz		1_10_1			1_20_1			1_30_1			2_40_1		
Auswertung		PG Kontrast	KI unten	KI oben	PG Kontrast	KI unten	KI oben	PG Kontrast	KI unten	KI oben	PG Kontrast	KI unten	KI oben
n = 1	A-R	0.0137	-0.0187	0.0461	0.0034	-0.0300	0.0368	0.0128	-0.0193	0.0449	0.0022	-0.0278	0.0322
	A-cov(sph)	0.0003	-0.0509	0.0515	0.0004	-0.0600	0.0608	0.0235	-0.0393	0.0863	0.0226	-0.0377	0.0829
	A-BI fix	0.0122	-0.0213	0.0457	0.0037	-0.0313	0.0387	0.0154	-0.0192	0.0500	0.0019	-0.0323	0.0361
	A-BI (fix)-Parz	----	----	----	----	----	----	----	----	----	----	----	----
	A-BI random	0.0080	-0.0231	0.0391	0.0075	-0.0238	0.0388	0.0107	-0.0206	0.0420	0.0022	-0.0294	0.0338
	A-BI (rand.)-Parz	----	----	----	----	----	----	----	----	----	----	----	----
n = 3	A-R	0.0113	-0.0058	0.0284	0.0120	-0.0066	0.0306	0.0103	-0.0083	0.0289	0.0060	-0.0131	0.0251
	A-cov(sph)	0.0177	-0.0064	0.0418	-0.0054	-0.0404	0.0296	0.0030	-0.0291	0.0351	0.0242	-0.0061	0.0545
	A-BI fix	0.0113	-0.0060	0.0286	0.0125	-0.0068	0.0318	0.0121	-0.0076	0.0318	0.0086	-0.0128	0.0300
	A-BI (fix)-Parz.	0.0101	-0.0086	0.0288	0.0140	-0.0078	0.0358	0.0120	-0.0106	0.0346	0.0102	-0.0162	0.0366
	A-BI random	0.0093	-0.0081	0.0267	0.0133	-0.0060	0.0326	0.0130	-0.0066	0.0326	0.0089	-0.0123	0.0301
	A-BI (rand.)-Parz.	0.0093	-0.0097	0.0283	0.0148	-0.0071	0.0367	0.0107	-0.0119	0.0333	0.0109	-0.0157	0.0375
n = 5	A-R	0.0071	-0.0066	0.0208	0.0082	-0.0079	0.0243	0.0043	-0.0119	0.0205	0.0026	-0.0147	0.0199
	A-cov(sph)	-0.0038*	-0.0606	0.0530	0.0179	-0.0010	0.0368	0.0116	-0.0076	0.0308	0.0197	-0.0021	0.0415
	A-BI fix	0.0067	-0.0072	0.0206	0.0080	-0.0086	0.0246	0.0053	-0.0118	0.0224	0.0039	-0.0152	0.0230
	A-BI (fix)-Parz.	0.0064	-0.0082	0.0210	0.0095	-0.0091	0.0281	0.0044	-0.0152	0.0240	0.0044	-0.0191	0.0279
	A-BI random	0.0096	-0.0058	0.0250	0.0079	-0.0094	0.0252	0.0051	-0.0123	0.0225	0.0035	-0.0157	0.0227
	A-BI (rand.)-Parz.	0.0057	-0.0093	0.0207	0.0038	-0.0151	0.0227	0.0058	-0.0138	0.0254	0.0041	-0.0192	0.0274

*) : Ergebnisse von 1000 Läufen, da Simulationen mit 10000 Läufen aufgrund sehr langer Laufzeiten nicht realisierbar waren

Tab. A5-3: Prüfgliedkontrasts PG 1 zu 2, sowie untere und obere Grenze seines 95 % Konfidenzintervalls der Auswertung der Parzellenmittelwerte der Boniturnoten (Parametersatz 1) mittels LMM

Simulierter räumlicher Kovarianzansatz		1_10_1			1_20_1			1_30_1			2_40_1		
		PG Kontrast	KI unten	KI oben	PG Kontrast	KI unten	KI oben	PG Kontrast	KI unten	KI oben	PG Kontrast	KI unten	KI oben
Auswertung													
n = 3	A-R	-0.0089	-0.0257	0.0078	-0.0104	-0.0285	0.0077	-0.0099	-0.0278	0.0081	-0.0068	-0.0249	0.0113
	A-cov(sph)	-0.0079	-0.0247	0.0090	-0.0107	-0.0279	0.0066	-0.0118	-0.0273	0.0038	-0.0100	-0.0237	0.0036
	A-BI fix	-0.0089	-0.0257	0.0078	-0.0104	-0.0285	0.0077	-0.0099	-0.0278	0.0081	-0.0068	-0.0249	0.0113
	A-BI (fix)-Parz.	---	---	---	---	---	---	---	---	---	---	---	---
	A-BI random	-0.0089	-0.0257	0.0078	-0.0104	-0.0285	0.0077	-0.0099	-0.0278	0.0081	-0.0068	-0.0249	0.0113
	A-BI (rand.)-Parz.	---	---	---	---	---	---	---	---	---	---	---	---
n = 5	A-R	-0.0055	-0.0194	0.0084	-0.0057	-0.0217	0.0103	-0.0041	-0.0201	0.0120	-0.0026	-0.0192	0.0141
	A-cov(sph)	-0.0058	-0.0197	0.0081	-0.0064	-0.0210	0.0082	-0.0084	-0.0211	0.0043	-0.0071	-0.0182	0.0041
	A-BI fix	-0.0089	-0.0257	0.0078	-0.0104	-0.0285	0.0077	-0.0099	-0.0278	0.0081	-0.0068	-0.0249	0.0113
	A-BI (fix)-Parz.	---	---	---	---	---	---	---	---	---	---	---	---
	A-BI random	-0.0055	-0.0194	0.0084	-0.0057	-0.0217	0.0103	-0.0041	-0.0201	0.0120	-0.0026	-0.0192	0.0141
	A-BI (rand.)-Parz.	---	---	---	---	---	---	---	---	---	---	---	---

Tab. A5-4: Prüfgliedkontrasts PG 1 zu 2, sowie untere und obere Grenze seines 95 % Konfidenzintervalls der Auswertung der Boniturnoten (Parametersatz 1) als stetiges Merkmal mittels LMM

Simulierter räumlicher Kovarianzansatz		1_10_1			1_20_1			1_30_1			2_40_1		
Auswertung		PG Kontrast	KI unten	KI oben	PG Kontrast	KI unten	KI oben	PG Kontrast	KI unten	KI oben	PG Kontrast	KI unten	KI oben
n = 1	A-R	-0.0071	-0.0332	0.0191	-0.0034	-0.0291	0.0223	-0.0082	-0.0333	0.0169	-0.0012	-0.0249	0.0226
	A-cov(sph)	-0.0063	-0.0325	0.0198	-0.0061	-0.0316	0.0194	-0.0085	-0.0327	0.0157	-0.0044	-0.0256	0.0168
	A-Bl fix	-0.0071	-0.0332	0.0191	-0.0034	-0.0291	0.0223	-0.0082	-0.0333	0.0169	-0.0012	-0.0249	0.0226
	A-Bl (fix)-Parz	---	---	---	---	---	---	---	---	---	---	---	---
	A-Bl random	-0.0071	-0.0332	0.0191	-0.0034	-0.0291	0.0223	-0.0082	-0.0333	0.0169	-0.0012	-0.0249	0.0226
	A-Bl (rand.)-Parz	---	---	---	---	---	---	---	---	---	---	---	---
n = 3	A-R	-0.0089	-0.0257	0.0078	-0.0104	-0.0285	0.0077	-0.0099	-0.0278	0.0081	-0.0068	-0.0249	0.0113
	A-cov(sph)	-0.0110	-0.0278	0.0059	-0.0169	-0.0333	-0.0005	-0.0162	-0.0311	-0.0012	-0.0139	-0.0270	-0.0007
	A-Bl fix	-0.0089	-0.0257	0.0078	-0.0104	-0.0285	0.0077	-0.0099	-0.0278	0.0081	-0.0068	-0.0249	0.0113
	A-Bl (fix)-Parz.	-0.0089	-0.0257	0.0078	-0.0104	-0.0285	0.0077	-0.0099	-0.0278	0.0081	-0.0068	-0.0249	0.0113
	A-Bl random	-0.0089	-0.0257	0.0078	-0.0104	-0.0285	0.0077	-0.0099	-0.0278	0.0081	-0.0068	-0.0249	0.0113
	A-Bl (rand.)-Parz.	-0.0089	-0.0257	0.0078	-0.0104	-0.0285	0.0077	-0.0099	-0.0278	0.0081	-0.0068	-0.0249	0.0113
n = 5	A-R	-0.0055	-0.0194	0.0084	-0.0057	-0.0217	0.0103	-0.0041	-0.0201	0.0120	-0.0026	-0.0192	0.0141
	A-cov(sph)	-0.0102	-0.0240	0.0036	-0.0105	-0.0234	0.0024	-0.0090	-0.0209	0.0029	-0.0090	-0.0194	0.0014
	A-Bl fix	-0.0055	-0.0194	0.0084	-0.0057	-0.0217	0.0103	-0.0041	-0.0201	0.0120	-0.0026	-0.0192	0.0141
	A-Bl (fix)-Parz.	-0.0055	-0.0194	0.0084	-0.0057	-0.0217	0.0103	-0.0041	-0.0201	0.0120	-0.0026	-0.0192	0.0141
	A-Bl random	-0.0055	-0.0194	0.0084	-0.0057	-0.0217	0.0103	-0.0041	-0.0201	0.0120	-0.0026	-0.0192	0.0141
	A-Bl (rand.)-Parz.	-0.0055	-0.0194	0.0084	-0.0057	-0.0217	0.0103	-0.0041	-0.0201	0.0120	-0.0026	-0.0192	0.0141

Anhang 6: Bias der geschätzten kumulativen Wahrscheinlichkeiten der Boniturnoten (Parametersatz 1) von Prüfglied 1 der Auswertung mittels Schwellenwertmodell (Linkfunktion „cumlogit“)

Tab. A6-1: Bias der geschätzten kumulativen Wahrscheinlichkeiten der Boniturnoten (Parametersatz 1) von Prüfglied 1 für den nichträumlichen Simulationsansatz und alle Simulationsansätze mit räumlicher Kovarianz, Auswertungsvariante A-R (fett: Erwartungstreue Schätzung der kumulativen Wahrscheinlichkeiten), 10000 Simulationsläufe

Auswertungsvariante A-R						
Simulierter räumlicher Kovarianzansatz		0_0_1	1_10_1	1_20_1	1_30_1	2_40_1
Boniturnote						
n = 1	1	0.0004	0.0002	0.0004	0.0004	0.0003
	2	-0.0011	0.0004	-0.0009	-0.0016	-0.0017
	3	-0.0015	0.0003	-0.0018	-0.0014	-0.0018
	4	-0.0015	0.0002	-0.0015	-0.0008	-0.0010
	5	-0.0001	0.0008	-0.0006	0.0006	-0.0001
	6	-0.0003	0.0005	0.0000	0.0002	-0.0004
	7	-0.0005	0.0001	-0.0004	-0.0001	-0.0004
	8	-0.0005	-0.0004	-0.0004	-0.0003	-0.0001
n = 3	1	0.0002	0.0003	0.0004	0.0003	0.0003
	2	0.0015	0.0020	0.0012	0.0011	0.0004
	3	0.0015	0.0018	0.0010	0.0011	0.0003
	4	0.0011	0.0016	0.0005	0.0007	0.0000
	5	0.0009	0.0013	0.0004	0.0006	0.0002
	6	0.0004	0.0005	0.0000	0.0001	0.0000
	7	0.0001	0.0001	-0.0001	-0.0002	-0.0002
	8	-0.0001	0.0001	-0.0001	-0.0001	-0.0001
n = 5	1	0.0001	0.0003	0.0003	0.0003	0.0003
	2	0.0010	0.0014	0.0009	0.0006	0.0001
	3	0.0011	0.0013	0.0008	0.0004	-0.0002
	4	0.0009	0.0013	0.0003	0.0002	-0.0005
	5	0.0007	0.0009	0.0002	0.0001	-0.0004
	6	0.0005	0.0004	-0.0001	-0.0002	-0.0004
	7	0.0000	0.0001	-0.0001	-0.0003	-0.0003
	8	0.0000	0.0001	-0.0001	-0.0001	-0.0001

Tab. A6-2: Bias der geschätzten kumulativen Wahrscheinlichkeiten der Boniturnoten (Parametersatz 1) von Prüfglied 1 für alle Simulationsansätze mit räumlicher Kovarianz, Auswertungsvariante A-cov(sph) (fett: Erwartungstreue Schätzung der kumulativen Wahrscheinlichkeiten), 10000 Simulationsläufe

Auswertungsvariante A-cov(sph)					
Simulierter räumlicher Kovarianzansatz		1_10_1	1_20_1	1_30_1	2_40_1
Boniturnote					
n = 1	1	-0.0082	-0.0166	-0.0218	-0.0249
	2	-0.0426	-0.0827	-0.1052	-0.1232
	3	-0.0085	-0.0184	-0.0181	-0.0212
	4	0.0235	0.0459	0.0630	0.0765
	5	0.0352	0.0701	0.0906	0.1060
	6	0.0289	0.0572	0.0719	0.0811
	7	0.0159	0.0311	0.0388	0.0415
	8	0.0066	0.0134	0.0165	0.0160
n = 3	1	-0.0022	-0.0087	-0.0156	-0.0227
	2	-0.0077	-0.0275	-0.0490	-0.0835
	3	0.0014	0.0024	0.0010	0.0042
	4	0.0083	0.0230	0.0374	0.0669
	5	0.0102	0.0318	0.0515	0.0848
	6	0.0075	0.0285	0.0441	0.0674
	7	0.0039	0.0182	0.0266	0.0375
	8	0.0013	0.0091	0.0124	0.0168
n = 5	1	-0.0019*	-0.0082	-0.0126	-0.0202
	2	-0.0072*	-0.0243	-0.0337	-0.0604
	3	-0.0005*	-0.0014	0.0024	0.0068
	4	0.0046*	0.0152	0.0270	0.0523
	5	0.0069*	0.0231	0.0368	0.0673
	6	0.0060*	0.0216	0.0325	0.0562
	7	0.0038*	0.0145	0.0206	0.0330
	8	0.0023*	0.0075	0.0100	0.0152

* Ergebnisse von 1000 Läufen, da Simulationen mit 10000 Läufen aufgrund sehr langer Laufzeiten nicht realisierbar waren

Tab. A6-3: Bias der geschätzten kumulativen Wahrscheinlichkeiten der Boniturnoten (Parametersatz 1) von Prüfglied 1 für alle Simulationsansätze mit räumlicher Kovarianz, Auswertungsvariante A-BI fix (fett: Erwartungstreue Schätzung der kumulativen Wahrscheinlichkeiten), 10000 Simulationsläufe

Auswertungsvariante A-BI fix					
Simulierter räumlicher Kovarianzansatz		1_10_1	1_20_1	1_30_1	2_40_1
Boniturnote					
n = 1	1	-0.0018	-0.0029	-0.0048	-0.0086
	2	-0.0039	-0.0080	-0.0127	-0.0210
	3	0.0011	-0.0004	0.0008	0.0015
	4	0.0045	0.0056	0.0103	0.0180
	5	0.0063	0.0087	0.0152	0.0247
	6	0.0053	0.0081	0.0130	0.0209
	7	0.0031	0.0047	0.0078	0.0127
	8	0.0011	0.0021	0.0036	0.0061
n = 3	1	-0.0006	-0.0019	-0.0037	-0.0074
	2	-0.0003	-0.0043	-0.0084	-0.0175
	3	0.0021	0.0019	0.0028	0.0034
	4	0.0036	0.0055	0.0097	0.0172
	5	0.0040	0.0070	0.0122	0.0222
	6	0.0029	0.0057	0.0100	0.0187
	7	0.0015	0.0033	0.0059	0.0113
	8	0.0007	0.0015	0.0027	0.0054
n = 5	1	-0.0005	-0.0018	-0.0035	-0.0072
	2	-0.0005	-0.0041	-0.0085	-0.0176
	3	0.0016	0.0017	0.0021	0.0029
	4	0.0030	0.0050	0.0089	0.0163
	5	0.0032	0.0063	0.0113	0.0212
	6	0.0023	0.0051	0.0093	0.0180
	7	0.0012	0.0030	0.0055	0.0108
	8	0.0006	0.0014	0.0026	0.0052

Tab. A6-4a: Bias der geschätzten konditionalen kumulativen Wahrscheinlichkeiten der Boniturnoten (Parametersatz 1) von Prüfglied 1 für alle Simulationsansätze mit räumlicher Kovarianz, Auswertungsvariante A-BI random (fett: Erwartungstreue Schätzung der kumulativen Wahrscheinlichkeiten), 10000 Simulationsläufe

Auswertungsvariante A-BI random					
Simulierter räumlicher Kovarianzansatz		1_10_1	1_20_1	1_30_1	2_40_1
Boniturnote					
n = 1	1	-0.0001	-0.0011	-0.0029	-0.0067
	2	-0.0018	-0.0057	-0.0102	-0.0182
	3	0.0002	-0.0018	-0.0007	-0.0001
	4	0.0015	0.0020	0.0063	0.0137
	5	0.0026	0.0044	0.0104	0.0198
	6	0.0022	0.0045	0.0090	0.0171
	7	0.0011	0.0025	0.0054	0.0106
	8	0.0002	0.0010	0.0024	0.0053
n = 3	1	-0.0001	-0.0013	-0.0031	-0.0069
	2	0.0007	-0.0032	-0.0068	-0.0162
	3	0.0018	0.0013	0.0026	0.0031
	4	0.0024	0.0039	0.0084	0.0157
	5	0.0025	0.0051	0.0105	0.0204
	6	0.0016	0.0042	0.0085	0.0173
	7	0.0008	0.0024	0.0050	0.0104
	8	0.0004	0.0011	0.0024	0.0051
n = 5	1	-0.0002	-0.0013	-0.0031	-0.0068
	2	0.0008	-0.0034	-0.0075	-0.0169
	3	0.0021	0.0013	0.0021	0.0024
	4	0.0028	0.0039	0.0082	0.0152
	5	0.0027	0.0050	0.0103	0.0200
	6	0.0017	0.0040	0.0083	0.0170
	7	0.0007	0.0024	0.0049	0.0103
	8	0.0004	0.0010	0.0023	0.0050

Tab. A6-4b: Bias der geschätzten marginalen kumulativen Wahrscheinlichkeiten der Boniturnoten (Parametersatz 1) von Prüfglied 1 für alle Simulationsansätze mit räumlicher Kovarianz, Auswertungsvariante A-BI random (fett: Erwartungstreue Schätzung der kumulativen Wahrscheinlichkeiten), 10000 Simulationsläufe

Auswertungsvariante A-BI random					
Simulierter räumlicher Kovarianzansatz		1_10_1	1_20_1	1_30_1	2_40_1
Boniturnote					
n = 1	1	0.0013	0.0023	0.0034	0.0053
	2	0.0003	-0.0006	-0.0008	-0.0003
	3	-0.0002	-0.0025	-0.0022	-0.0029
	4	-0.0006	-0.0028	-0.0026	-0.0035
	5	-0.0002	-0.0023	-0.0020	-0.0041
	6	-0.0005	-0.0020	-0.0030	-0.0056
	7	-0.0008	-0.0022	-0.0032	-0.0056
	8	-0.0009	-0.0016	-0.0024	-0.0038
n = 3	1	0.0007	0.0017	0.0029	0.0048
	2	0.0020	0.0014	0.0020	0.0013
	3	0.0015	0.0005	0.0010	-0.0002
	4	0.0012	-0.0005	-0.0001	-0.0014
	5	0.0008	-0.0010	-0.0012	-0.0027
	6	0.0001	-0.0016	-0.0025	-0.0044
	7	-0.0003	-0.0016	-0.0029	-0.0050
	8	-0.0002	-0.0012	-0.0021	-0.0038
n = 5	1	0.0007	0.0018	0.0030	0.0049
	2	0.0022	0.0015	0.0016	0.0009
	3	0.0019	0.0004	0.0004	-0.0008
	4	0.0015	-0.0008	-0.0006	-0.0020
	5	0.0008	-0.0015	-0.0018	-0.0034
	6	0.0000	-0.0021	-0.0030	-0.0049
	7	-0.0005	-0.0019	-0.0032	-0.0053
	8	-0.0003	-0.0014	-0.0023	-0.0041

Tab. A6-5a: Bias der geschätzten konditionalen kumulativen Wahrscheinlichkeiten der Boniturnoten (Parametersatz 1) von Prüfglied 1 für alle Simulationsansätze mit räumlicher Kovarianz, Auswertungsvariante A-BI (rand.)-Parz. (fett: Erwartungstreue Schätzung der kumulativen Wahrscheinlichkeiten), 10000 Simulationsläufe

Auswertungsvariante A-BI(rand.)-Parz.					
Simulierter räumlicher Kovarianzansatz		1_10_1	1_20_1	1_30_1	2_40_1
Boniturnote					
n = 1	1	---	---	---	---
	2	---	---	---	---
	3	---	---	---	---
	4	---	---	---	---
	5	---	---	---	---
	6	---	---	---	---
	7	---	---	---	---
	8	---	---	---	---
n = 3	1	-0.0042	-0.0090	-0.0113	-0.0169
	2	-0.0115	-0.0256	-0.0320	-0.0494
	3	0.0025	0.0031	0.0037	0.0053
	4	0.0121	0.0224	0.0279	0.0428
	5	0.0153	0.0295	0.0364	0.0553
	6	0.0126	0.0247	0.0304	0.0458
	7	0.0073	0.0147	0.0181	0.0270
	8	0.0034	0.0068	0.0085	0.0126
n = 5	1	-0.0034	-0.0087	-0.0110	-0.0170
	2	-0.0092	-0.0245	-0.0303	-0.0485
	3	0.0020	0.0028	0.0045	0.0061
	4	0.0097	0.0210	0.0278	0.0428
	5	0.0120	0.0274	0.0355	0.0547
	6	0.0097	0.0229	0.0294	0.0451
	7	0.0055	0.0136	0.0173	0.0264
	8	0.0026	0.0063	0.0081	0.0123

Tab. A6-5b: Bias der geschätzten marginalen kumulativen Wahrscheinlichkeiten der Boniturnoten (Parametersatz 1) von Prüfglied 1 für alle Simulationsansätze mit räumlicher Kovarianz, Auswertungsvariante A-BI (rand.)-Parz (fett: Erwartungstreue Schätzung der kumulativen Wahrscheinlichkeiten), 10000 Simulationsläufe

Auswertungsvariante A-BI (rand.)-Parz.					
Simulierter räumlicher Kovarianzansatz		1_10_1	1_20_1	1_30_1	2_40_1
Boniturnote					
n = 1	1	---	---	---	---
	2	---	---	---	---
	3	---	---	---	---
	4	---	---	---	---
	5	---	---	---	---
	6	---	---	---	---
	7	---	---	---	---
	8	---	---	---	---
n = 3	1	0.0057	0.0100	0.0112	0.0142
	2	0.0042	0.0049	0.0047	0.0045
	3	-0.0002	-0.0019	-0.0023	-0.0032
	4	-0.0029	-0.0064	-0.0068	-0.0080
	5	-0.0050	-0.0096	-0.0103	-0.0120
	6	-0.0064	-0.0115	-0.0126	-0.0148
	7	-0.0060	-0.0106	-0.0118	-0.0144
	8	-0.0041	-0.0076	-0.0086	-0.0106
n = 5	1	0.0042	0.0085	0.0100	0.0129
	2	0.0030	0.0036	0.0044	0.0041
	3	-0.0001	-0.0018	-0.0014	-0.0027
	4	-0.0020	-0.0057	-0.0053	-0.0071
	5	-0.0038	-0.0085	-0.0087	-0.0110
	6	-0.0050	-0.0103	-0.0111	-0.0138
	7	-0.0047	-0.0095	-0.0107	-0.0137
	8	-0.0032	-0.0069	-0.0079	-0.0105

Tab. A6-6a: Bias der geschätzten konditionalen kumulativen Wahrscheinlichkeiten der Boniturnoten (Parametersatz 1) von Prüfglied 1 für alle Simulationsansätze mit räumlicher Kovarianz, Auswertungsvariante A-BI (fix.)-Parz (fett: Erwartungstreue Schätzung der kumulativen Wahrscheinlichkeiten), 10000 Simulationsläufe

Auswertungsvariante A-BI(fix.)-Parz.					
Simulierter räumlicher Kovarianzansatz		1_10_1	1_20_1	1_30_1	2_40_1
Boniturnote					
n = 1	1	---	---	---	---
	2	---	---	---	---
	3	---	---	---	---
	4	---	---	---	---
	5	---	---	---	---
	6	---	---	---	---
	7	---	---	---	---
	8	---	---	---	---
n = 3	1	-0.0049	-0.0097	-0.0120	-0.0175
	2	-0.0133	-0.0272	-0.0334	-0.0512
	3	0.0028	0.0037	0.0045	0.0057
	4	0.0138	0.0244	0.0300	0.0445
	5	0.0175	0.0318	0.0388	0.0574
	6	0.0143	0.0265	0.0322	0.0473
	7	0.0083	0.0158	0.0191	0.0277
	8	0.0038	0.0073	0.0089	0.0129
n = 5	1	-0.0037	-0.0090	-0.0114	-0.0173
	2	-0.0100	-0.0248	-0.0314	-0.0493
	3	0.0023	0.0038	0.0045	0.0065
	4	0.0108	0.0226	0.0286	0.0439
	5	0.0133	0.0291	0.0365	0.0559
	6	0.0108	0.0242	0.0302	0.0460
	7	0.0062	0.0143	0.0178	0.0269
	8	0.0028	0.0066	0.0083	0.0125

Tab. A6-6b: Bias der geschätzten marginalen kumulativen Wahrscheinlichkeiten der Boniturnoten (Parametersatz 1) von Prüfglied 1 für alle Simulationsansätze mit räumlicher Kovarianz, Auswertungsvariante A-BI (fix.)-Parz (fett: Erwartungstreue Schätzung der kumulativen Wahrscheinlichkeiten), 10000 Simulationsläufe

Auswertungsvariante A-BI(fix.)-Parz.					
Simulierter räumlicher Kovarianzansatz		1_10_1	1_20_1	1_30_1	2_40_1
Boniturnote					
n = 1	1	---	---	---	---
	2	---	---	---	---
	3	---	---	---	---
	4	---	---	---	---
	5	---	---	---	---
	6	---	---	---	---
	7	---	---	---	---
	8	---	---	---	---
n = 3	1	0.0059	0.0095	0.0097	0.0109
	2	0.0048	0.0058	0.0058	0.0054
	3	-0.0004	-0.0018	-0.0019	-0.0033
	4	-0.0036	-0.0069	-0.0071	-0.0089
	5	-0.0058	-0.0101	-0.0103	-0.0121
	6	-0.0071	-0.0115	-0.0118	-0.0132
	7	-0.0064	-0.0101	-0.0104	-0.0116
	8	-0.0043	-0.0070	-0.0072	-0.0078
n = 5	1	0.0043	0.0080	0.0084	0.0097
	2	0.0036	0.0051	0.0050	0.0051
	3	-0.0001	-0.0013	-0.0017	-0.0027
	4	-0.0023	-0.0058	-0.0062	-0.0079
	5	-0.0042	-0.0087	-0.0092	-0.0111
	6	-0.0052	-0.0100	-0.0106	-0.0122
	7	-0.0047	-0.0088	-0.0094	-0.0107
	8	-0.0031	-0.0062	-0.0065	-0.0075

Anhang 7: Die folgenden Abbildungen zeigen die halbe Breite der Konfidenzintervalle und die empirische Überdeckungswahrscheinlichkeiten für die Varianten der Auswertung der Bonituren (Parametersatz 1) mittels SWM (Linkfunktion „cumlogit“) für alle Simulationsansätze mit räumlicher Kovarianz

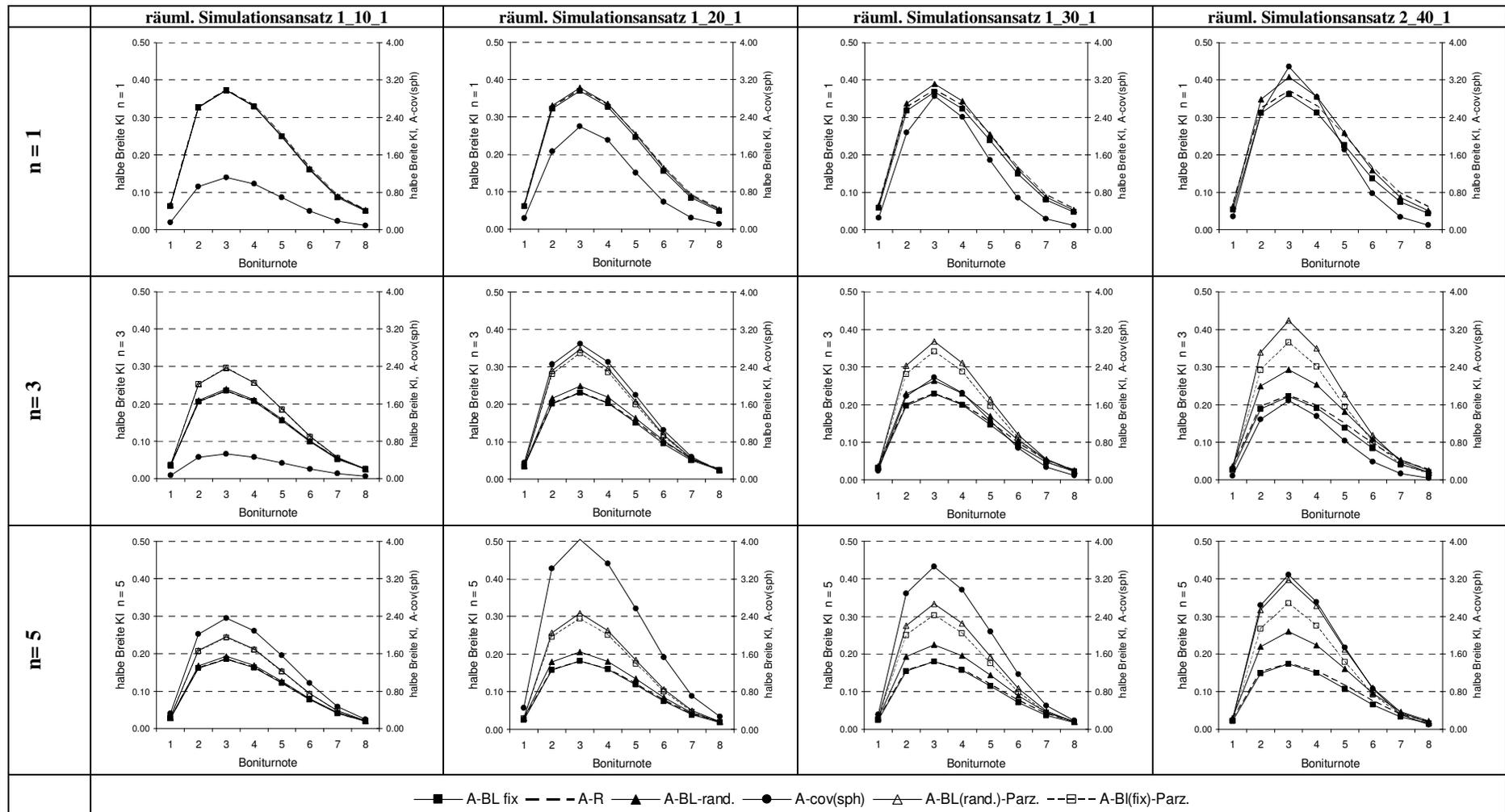


Abb. A7-1: Halbe Breite der Konfidenzintervalle KI für die bedingten Wahrscheinlichkeiten der Boniturnoten (Parametersatz 1) alle Varianten der Auswertung der Boniturnen mittels SWM (Linkfunktion „cumlogit“) für Simulationsansätze mit räumlicher Kovarianz (ACHTUNG: Ansatz A-cov(sph) rechte Skala)

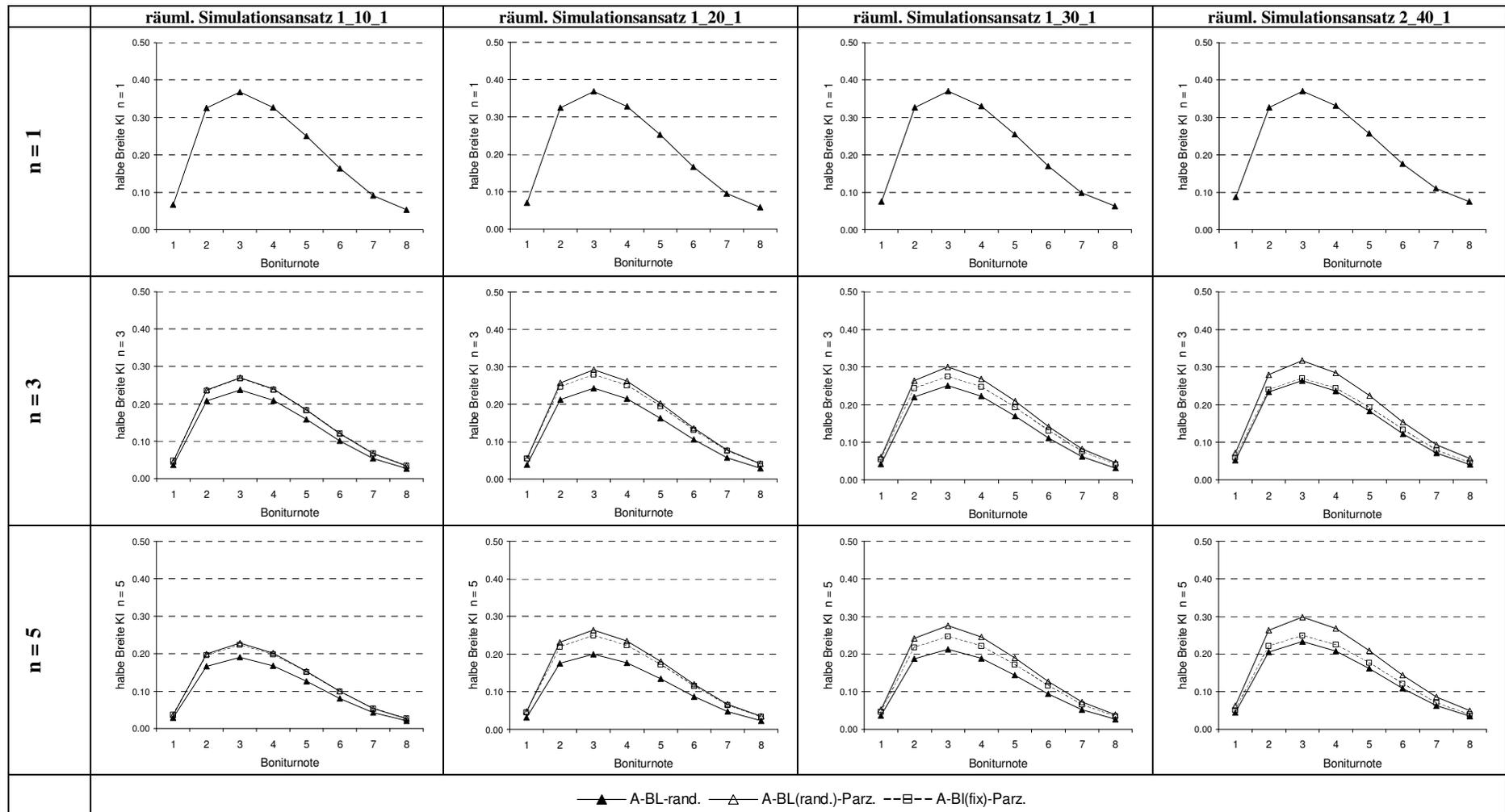


Abb. A7-2: Halbe Breite der Konfidenzintervalle KI für die marginalen Wahrscheinlichkeiten der Boniturnoten (Parametersatz 1) alle Varianten der Auswertung der Boniturnoten mittels SWM (Linkfunktion „cumlogit“) für Simulationsansätze mit räumlicher Kovarianz

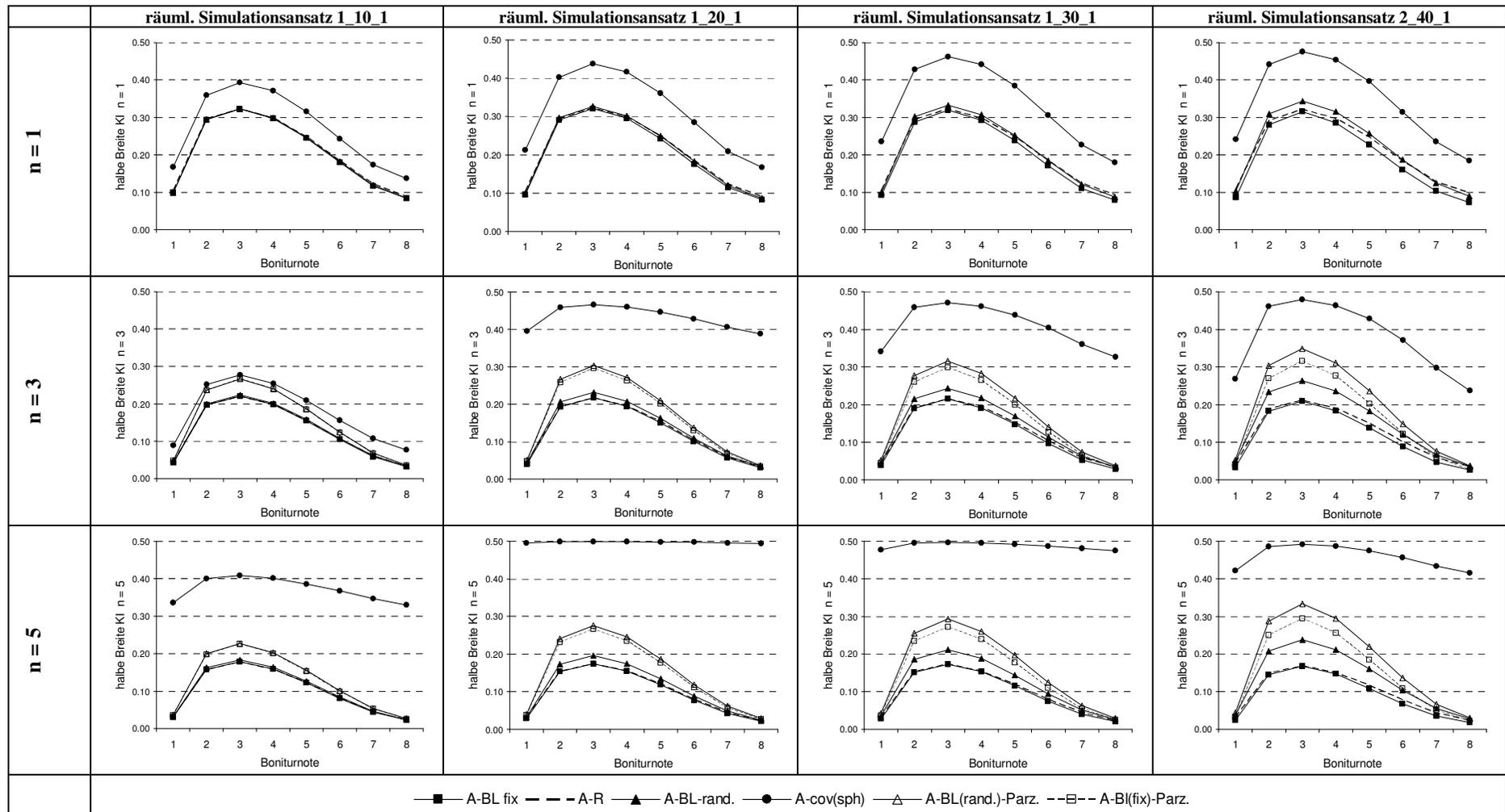


Abb. A7-3: Halbe Breite der Konfidenzintervalle KI_{\logit} für alle Varianten der Auswertung der Bonituren (Parametersatz 1) mittels SWM (Linkfunktion „cumlogit“) für Simulationsansätze mit räumlicher Kovarianz

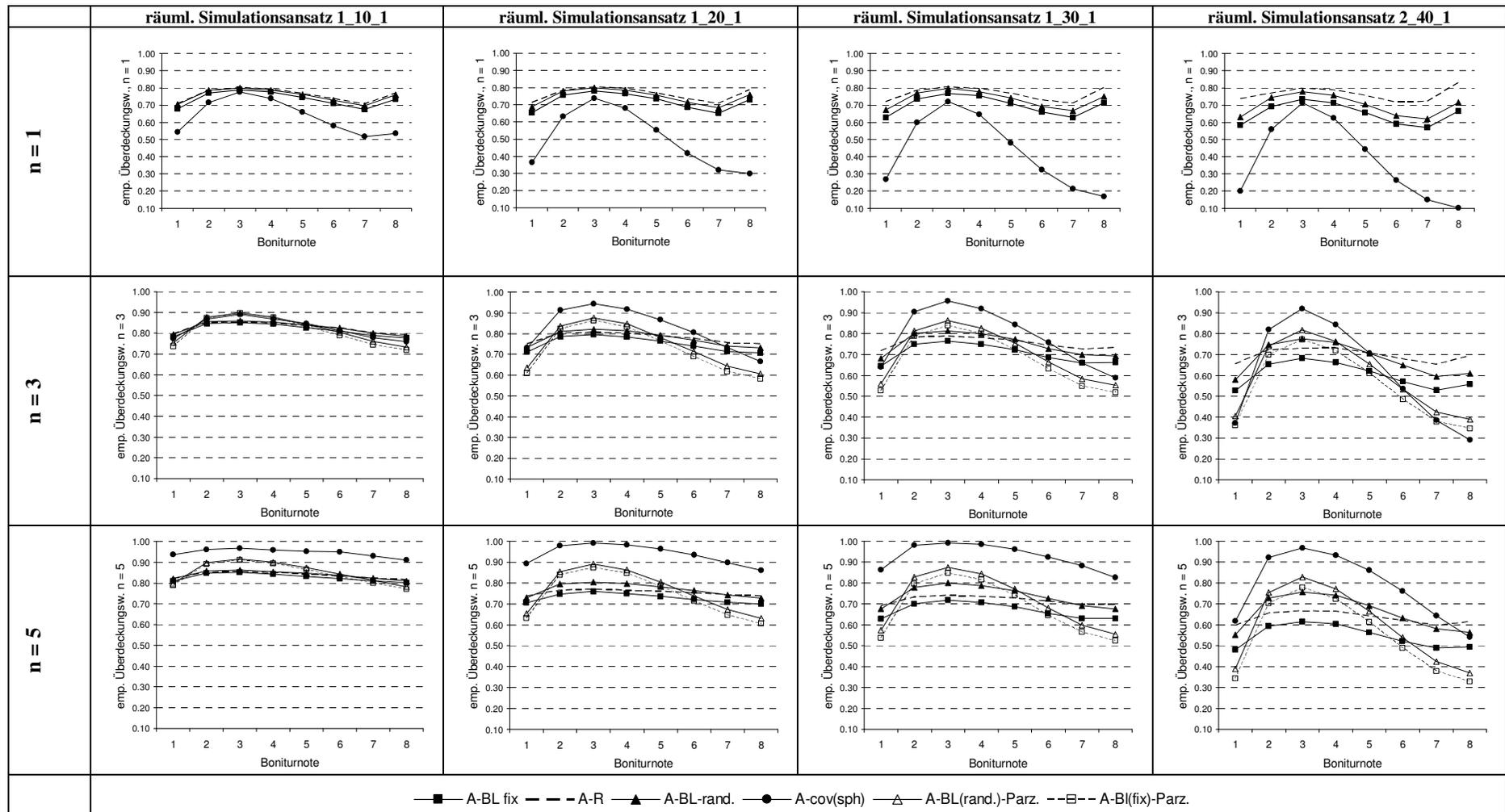


Abb. A7-4: Überdeckungswahrscheinlichkeiten der Konfidenzintervalle KI für die bedingten Wahrscheinlichkeiten der Boniturnoten (Parametersatz 1) alle Varianten der Auswertung der Bonituren mittels SWM (Linkfunktion „cumlogit“) für Simulationsansätze mit räumlicher Kovarianz

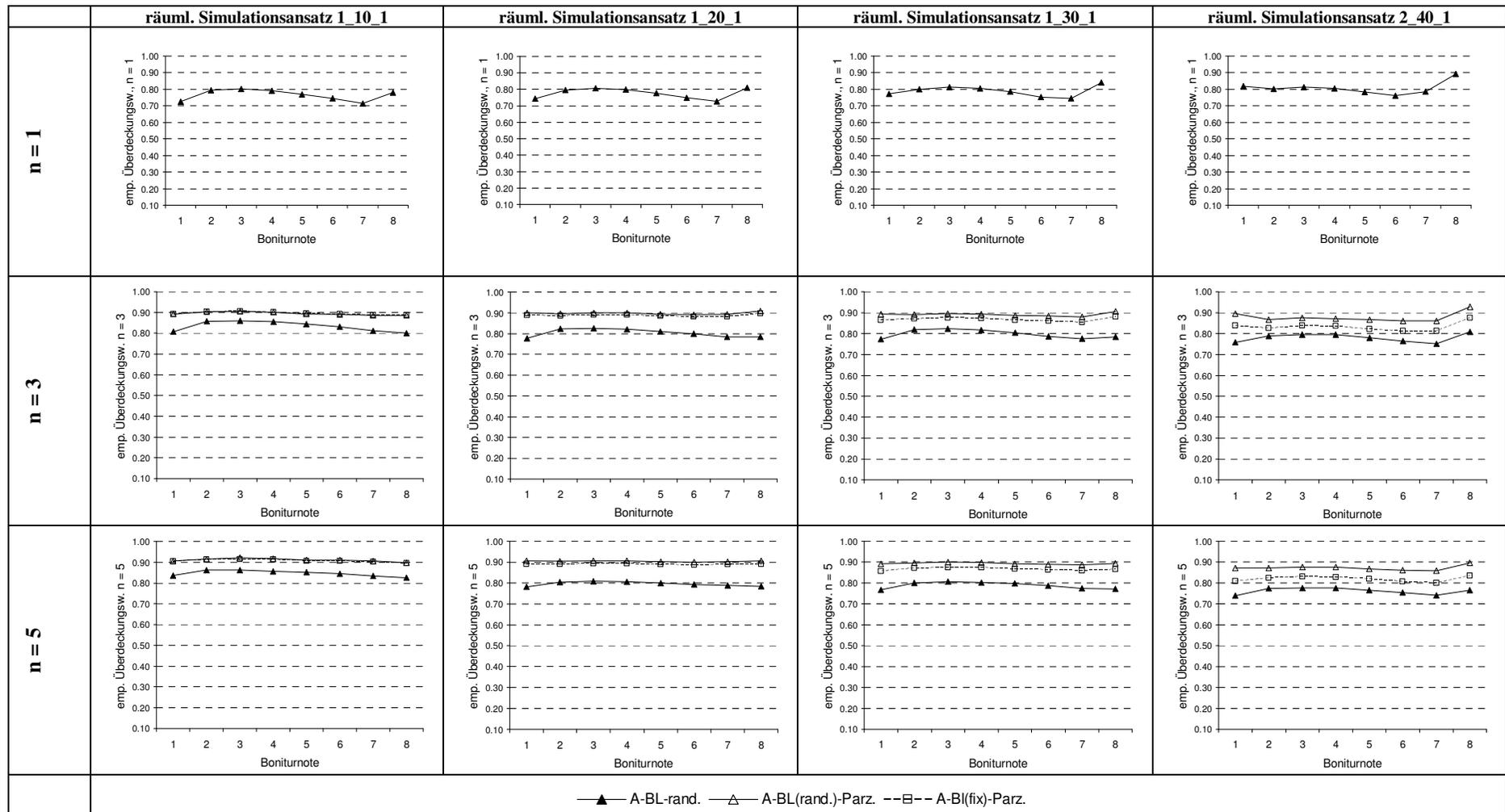


Abb. A7-5: Überdeckungswahrscheinlichkeiten der Konfidenzintervalle KI für die marginalen Wahrscheinlichkeiten der Boniturnoten (Parametersatz 1) alle Varianten der Auswertung der Bonituren mittels SWM (Linkfunktion „cumlogit“) für Simulationsansätze mit räumlicher Kovarianz

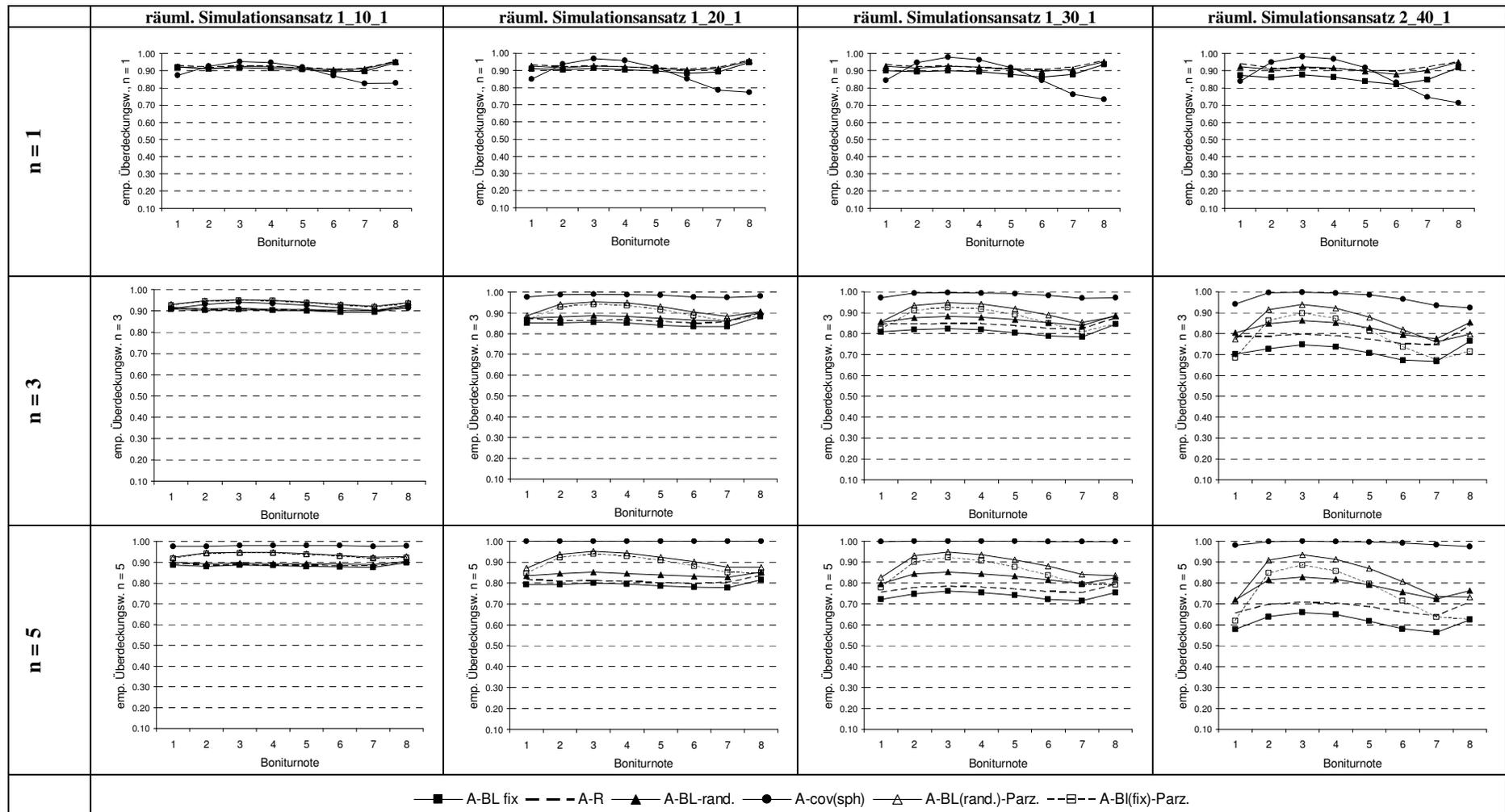


Abb. A7-6: Überdeckungswahrscheinlichkeiten der Konfidenzintervalle KI_{\logit} für alle Varianten der Auswertung der Bonituren (Parametersatz 1) mittels SWM (Linkfunktion „cumlogit“) für Simulationsansätze mit räumlicher Kovarianz

Verzeichnis der verwendeten Abkürzungen

CRD	Completely Randomized Design
FG	Freiheitsgrad
GLM	Generalisiertes lineares Modell
GLMM	Generalisiertes lineares gemischtes Modell
KI	Konfidenzintervall
LMM	Lineares gemischtes Modell
ML	Maximum Likelihood
RCBD	Randomized Complete Block Design
REML	Restricted Maximum Likelihood
RSPL	Restricted Pseudo-Likelihood
STD	Standardabweichung
SWM	Schwellenwertmodell
THM	Threshold Model
ÜDWK	Überdeckungswahrscheinlichkeit
VC-Matrix	Varianz-Kovarianz-Matrix

Erklärung

Hiermit erkläre ich, Karen Höttl, diese wissenschaftliche Arbeit selbständig
angefertigt und keine anderen als die angegebenen Hilfsmittel benutzt zu haben.

Berlin, 07.02.2012

.....

Karen Höttl

Danksagung

Diese Arbeit entstand im Rahmen eines durch die Deutsche Forschungsgemeinschaft (DFG) geförderten Projektes.

Meinen speziellen Dank möchte ich Herrn Prof. Dr. J. Spilke für die Überlassung des Themas und die unermüdliche Betreuung aussprechen. Ebenfalls danke ich Herrn Dr. Mielenz für die Beratung in allen mathematischen Fragen und Frau Prof. C. Richter für die zusätzlichen Anregungen.

Weiterhin danke ich allen ehemaligen Kollegen der Arbeitsgruppe Biometrie und Agrarinformatik für das angenehme Miteinander.

Ausdrücklicher Dank gebührt meinem Vater Helmut Höttl, für die moralische und finanzielle Unterstützung in den letzten drei Jahren, die mir das Berufspendlerdasein sehr erleichtert hat. Meinem Lebenspartner Kai Hellmann danke ich von Herzen für die Engelsgeduld, mit der er diese anstrengende Lebensphase mit mir durchgestanden hat.

Danke.

