

Zur Zuchtwertschätzung bei heterogener Populationsstruktur

Engelmann G.¹, Wunsch U.¹, Göpfert R.¹, Schüler L.¹, Bergfeld U.², Krüger W.³

¹ Institut für Tierzucht und Tierhaltung, Agrarwiss. Fakultät der Martin-Luther-Universität Halle-Wittenberg, Adam-Kuckhoff-Straße 35, 06108 Halle

² Fachbereich Tierzucht, Fischerei und Grünland, Sächsische Landesanstalt für Landwirtschaft, Am Park, 04886 Köllitzsch

³ Wissenschaftsbereich Tierzucht und Tierhaltung, Studienprogramm Agrarwissenschaften der Universität Leipzig, Johannisallee 23, 04275 Leipzig

Einleitung

Die Zuchtwertschätzung mittels Tiermodell auf der Grundlage von Mehrmerkmals- oder Wiederholbarkeitsansätzen wird insbesondere in der Milchrindzucht zunehmend zum Standardverfahren. Die auf den von HENDERSON (1972, 1974) abgeleiteten Mixed-Modell-Gleichungen basierenden Verfahren zur Zuchtwertschätzung liefern jedoch nur beste lineare und unverzerrte Zuchtwerte, wenn für die realisierten Leistungen keine ungleichen Erwartungswerte vorliegen. Für die einzelnen, in einer Population auftretenden Umweltklassen werden daher die unterschiedlichen Mittelwerte über die Einbeziehung fixer Effekte in das populationsgenetische Modell korrigiert und somit eine Vergleichbarkeit über die Umweltklassen erreicht. Heterogene genetische und nichtgenetische Varianzen können ebenfalls zu Rangfolgeverschiebungen zwischen den ermittelten Zuchtwerten führen, welche dann keine besten Schätzer mehr darstellen.

Sowohl HENDERSON (1984) als auch GARRICK und VAN VLECK (1987) beschreiben die theoretisch möglichen Situationen der Veränderung der additiv-genetischen und der Restvarianz bzw. des Varianzverhältnisses und leiten die Konsequenzen der Varianzheterogenität für die Zuchtwertschätzung ab. Als Ursachen für das Auftreten von heterogenen Varianzen nennt VISSCHER (1991) sowohl Einflüsse durch einzelne Zuchtprogramme (z.B. ET-Programme, Bullenselektion), Managementeinflüsse (z.B. durch Haltungs- und Fütterungssysteme), Genotyp-Umwelt-Interaktionen, Bullen-Herden-Interaktionen und bevorzugte Behandlung.

Die Selektion von Zuchttieren aus einer Population wird durch das Vorliegen von Varianzheterogenität beeinflusst. Werden Tiere aus Herden mit unterschiedlichen genetischen und nichtgenetischen Varianzen selektiert, so wird der Selektionsentscheid bei gleichen Herdenmitteln wesentlich durch die Differenzen in den Varianzen bestimmt (GIANOLA, 1990). Verhalten sich die Standardabweichungen zwischen zwei Herden mit gleichen Mittelwerten wie 2 : 1 und beträgt die Selektionsintensität 1 %, dann werden 96 % der selektierten Tiere aus der variableren Herde kommen. Auch EVERETT et. al. (1982), HILL (1984) und BROTHERSTONE und HILL (1986) beschreiben, daß über dem Herdenmittelwert liegende Tiere in Herden mit höherer Variabilität überschätzt werden. Das führt zu Verzerrungen in der Bewertung von Bullen, wenn diese auf der Grundlage von Töchterleistungen selektiert werden, welche aus nicht zufällig über die Herden bzw. Umwelten verteilten Anpaarungen stammen. Größere Probleme werden durch heterogene Varianzen bei der Selektion der Bullenmütter hervorgerufen. Da die Rangfolge der Kuhzuchtwerte wesentlich durch den Vergleich innerhalb der Herden bestimmt wird, kann der geschätzte Zuchtwert beträchtlich von der Innerherdenvarianz beeinflusst werden (WINKELMANN, 1988; MEINERT, 1988). Liegen dabei heterogene Umweltvarianzen vor, dann beruht die Selektion mehr auf einer durch die Umwelt

beeinflußten Rangierung der Zuchtwerte. Besteht jedoch Heterogenität der genetischen Varianz zwischen den Herden, dann werden die Kühe aus Herden mit einer höheren Varianz bevorzugt. POWELL (1983) schlußfolgert daraus, daß es zu einer verstärkten Selektion von Tieren aus Herden mit höherer Heritabilität kommen kann.

Zur Homogenisierung der Varianz und zu Möglichkeiten der Beachtung von Varianzheterogenitäten in Modellen der Zuchtwertschätzung werden verschiedene Ansätze vorgeschlagen. Von HILL (1983), BROTHERSTONE und HILL (1986) werden arithmetische Transformationen, z.B. Logarithmierung, genutzt. Arithmetische Transformationen sind mathematisch einfache Wege zur Varianzhomogenisierung, wenn die Heterogenität auf Skaleneffekten beruht. Logarithmierung führt demzufolge zu homogeneren Varianzen, wenn diese einfache Funktionen des Mittelwertes sind. Die Korrelation zwischen Mittelwert und Varianz ist jedoch häufig zu gering (KACHMAN und EVERETT, 1993). Eine Überkompensation der Herdenvarianz durch log - Transformation wurde durch BERGFELD (1992) beschrieben. Mehrmerkmalsansätze wie von KACHMAN und EVERETT (1993) vorgeschlagen, nutzen multiplikative Skalierungsfaktoren für die einzelnen Umweltklassen bzw. Merkmale. Multiplikative Faktoren liefern bessere Schätzer, wenn sich die Varianz der Effekte über die Umwelten ändert.

Material und Methode

Für die Untersuchungen stehen die Daten von insgesamt 1 831 765 Kühen der ostdeutschen Milchrindpopulation aus den Kalbejahren 1987 bis 1992 zur Verfügung. Dabei liegen für die Merkmale Milchmenge, Fettmenge und Eiweißmenge die erfaßten Leistungen aus den fünf Abschnitten 1. - 100. Tag, 101. - 200. Tag, 201. - 305. Tag der 1. Laktation, 1. - 305 Tag der 2. Laktation und 1. - 305. Tag der 3. Laktation vor, wobei in der ersten Laktation nur bei etwa 40% der Tiere und in der zweiten und dritten Laktation nur bei etwa 20% Leistungen für das Merkmal Eiweiß registriert wurden. Die phänotypischen Leistungen wurden für Kalbealter, sowie für laufende und vorhergehende Zwischenkalbezeit vorkorrigiert. Einen Überblick über das Leistungsniveau der untersuchten Population für die Abschnitte 1. - 100. Tag sowie erste Laktation vermittelt Tabelle 1.

Tabelle 1: Leistungsniveau der bearbeiteten Milchrindpopulation für die Abschnitte 1. - 100. Tag und 1. Laktation

Merkmals	Leistungsabschnitte	N	MW	S	VK
Milch - kg	100 Tage	1358532	1672.18	356.51	21.32
Fett - kg	100 Tage	1359569	68.42	17.28	25.26
Eiweiß - kg	100 Tage	584445	56.43	11.95	21.18
Milch - kg	1. Lakt.	1084641	4078.65	841.30	20.63
Fett - kg	1. Lakt.	1092851	172.59	39.83	23.08
Eiweiß - kg	1. Lakt.	444274	148.88	30.81	20.70

Die im Datenmaterial vorliegende phänotypische Variabilität wurde für die einzelnen Umweltklassen untersucht. Dabei erfolgte die Beschreibung der Varianzheterogenität für das Merkmal Milchleistung 1. - 100. Tag an Hand des nach Herden, Herdenmittel, Region und Subregion klassifizierten Materials. Die Varianz innerhalb der insgesamt 4445 im Material vorhandenen Herden wurde für die Merkmale Milchmenge, Fettmenge und Eiweißmenge 1. - 100. Tag mittels beschreibender Statistik untersucht.

Ergebnisse und Diskussion

Untersucht wurden zunächst die innerhalb der Herden für den Leistungsabschnitt 1. - 100. Tag vorliegenden phänotypischen Variabilitäten der Merkmale Milch-, Fett- und Eiweißmenge. Für die Herden-Standardabweichung und den Herden-Variationskoeffizienten sind die statistischen Parameter in den Tabellen 2.1 und 2.2 dargestellt. Es lassen sich erhebliche Unterschiede in der Variabilität erkennen. Die Korrelationen zwischen dem Leistungsniveau der Herden und der Inner-herdenvarianz liegen dabei im unteren bis mittleren Bereich und sind für alle Merkmale positiv. Zwischen Variationskoeffizienten und Leistungsniveau bestehen geringere und negative Beziehungen.

Tabelle 2.1: Mittelwerte, Standardabweichungen und Variationskoeffizienten sowie Korrelationen für die Standardabweichung innerhalb der Herden für den Leistungsabschnitt 1. - 100. Tag

Merkmal	N	MW _S	S _S	VK _S	r _{MW,S}
Milch - kg	3701	320.77	84.90	26.47	0.21
Fett - kg	3695	15.47	4.41	28.49	0.37
Eiweiß - kg	3254	9.87	2.94	29.77	0.25

Tabelle 2.2: Mittelwerte, Standardabweichungen und Variationskoeffizienten sowie Korrelationen für den Variationskoeffizienten innerhalb der Herden für den Leistungsabschnitt 1. - 100. Tag

Merkmal	N	MW _{VK}	S _{VK}	VK _{VK}	r _{MW,VK}
Milch - kg	3701	19.12	5.22	27.30	- 0.26
Fett - kg	3695	22.02	5.89	26.75	- 0.18
Eiweiß - kg	3254	17.76	5.39	30.35	- 0.22

Entsprechend den Mittelwerten für das Merkmal Milchleistung 1. - 100. Tag wurden die Herden, wie aus Tabelle 3 ersichtlich, eingeteilt. Bei linear steigendem Mittelwert bleibt die Variabilität im unteren Leistungsbereich nahezu konstant und steigt erst bei mittlerem und höherem Herdenniveau an. Der Variationskoeffizient sinkt mit steigendem Herdenmittel degressiv. Insgesamt zeigen sich die gleichen Tendenzen, wie sie bereits an Hand der Korrelationskoeffizienten innerhalb der Herden für den Zusammenhang von Mittelwert und Variabilität zu erkennen waren.

Tabelle 3: Mittlere Variabilität der Herden bei Stratifikation nach unterschiedlichem Leistungs-niveau für das Merkmal Milch - kg 1. - 100. Tag

Mittelwert von	Mittelwert bis	N	MW _{MW}	MW _S	MW _{VK}
	1000	3	830.32	456.91	56.73
1000	1200	21	1153.43	317.94	27.48
1200	1400	141	1362.98	318.96	23.40
1400	1600	695	1550.88	301.74	19.46
1600	1800	591	1748.41	325.34	18.61
1800	2000	260	1945.94	343.34	17.65
2000	2200	87	2144.59	366.34	17.07
2200	2400	18	2340.66	378.29	16.16
2400		8	2602.06	424.48	16.29

Eine weitere Klassifizierung des Materials wurde nach Regionen und Subregionen vorgenommen, wobei entsprechend der ehemaligen ostdeutschen Bezirks- und Kreisstruktur unterteilt wurde. Tabelle 4 gibt einen Überblick über die Variabilitätsunterschiede von drei ausgewählten Regionen. Dabei wird deutlich, daß sich die Korrelationen zwischen Mittelwert und phänotypischer Variabilität für die Regionen zum Teil deutlich unterscheiden. Generell liegt die Beziehung zwischen Leistungsniveau und Variabilität innerhalb der Regionen im mittleren bis oberen Bereich. Die statistischen Parameter nach Stratifikation in Subregionen wurden für die Region 1 in Tabelle 5 dargestellt. Es läßt sich eine beträchtliche Variation aller Parameter über die Subregionen erkennen. Am sichtbarsten wird dies an den Variationskoeffizienten, die zwischen 18.23 und 23.31 liegen.

Tabelle 4: Mittelwerte, Standardabweichungen, Variationskoeffizienten und Korrelationen für drei Regionen des gesplitteten Datenmaterials

Region	N	MW _{MW}	S _{MW}	MW _{VK}	S _{VK}	r _{MW.S}
1	117775	1653.18	40.11	20.25	1.71	0.54
2	64309	1631.90	121.73	20.48	0.83	0.86
3	41480	1583.24	79.53	19.58	0.96	0.88

Tabelle 5: Mittelwert, Standardabweichung und Variationskoeffizient für die Subregionen einer ausgewählten Region

Region	N	MW	S	VK
1. 1	6085	1620.47	295.35	18.23
1. 2	5900	1617.90	342.82	21.19
1. 3	17524	1681.13	341.92	20.34
1. 4	24226	1694.58	362.75	21.41
1. 5	15222	1593.16	313.25	19.66
1. 6	7464	1697.31	309.57	18.43
1. 7	9503	1703.36	375.50	22.04
1. 8	16799	1663.19	387.71	23.31
1. 9	9725	1673.95	310.72	18.56
1.10	5327	1604.80	310.74	19.36

Zusammenfassung

Das Datenmaterial wurde nach Herden, Herdenniveau, Region und Subregion stratifiziert und die Heterogenität der phänotypischen Varianz untersucht. Innerhalb der Herden zeigt sich eine positive Beziehung zwischen Leistungsniveau und Variabilität. Bei der Klassifizierung entsprechend des Herdenniveaus wird sichtbar, daß der Anstieg der Varianz langsamer erfolgt, als der Anstieg des Herdenmittels. Die Aufspaltung in Regionen und Subregionen läßt ebenfalls Heterogenität der Varianzen erkennen. So differieren die Korrelationen zwischen Leistungsniveau und Standardabweichung zwischen den Regionen. Die Unterschiede zwischen den Subregionen werden an den Varianzkoeffizienten, welche für die einzelnen Subregionen deutlich schwanken, erkennbar. Für die Beachtung heterogener Varianzen in einem Modell der Zuchtwertschätzung sind zunächst weitere Untersuchungen zur Quantifizierung sowohl der genetischen als auch der Restvarianzen notwendig. Bei Vorhandensein von Varianzheterogenitäten müssen geeignete Verfahren zur Varianzhomogenisierung innerhalb eines Zuchtwertschätzmodelles abgeleitet werden.

LITERATUR:

- Bergfeld, U. (1992): Heterogenität der Varianz in der Nachkommenschaftsprüfung beim Ostdeutschen Schwarzbunten Rind.- 65. Sitzung des genetisch-statistischen Ausschusses der DGfZ, Neumühle/Pfalz, 4.-6.3.1992
- Brotherstone, S.; Hill, W.G. (1986): Heterogeneity of variance amongst herds for milk production.- Anim. Prod. 42, 297-303
- Everett, W.; Koewn, J.F.; Taylor J.F. (1982): The problem of heterogeneous within herd error variances when identifying elite cows.- J. Dairy Sci. 65, 100 (Supp. 1)
- Garrick, D.J.; Van Vleck, L.D. (1987): Aspects of selection for performance in several environments with heterogeneous variances.- J. Anim. Sci. 65, 409-421
- Gianola, D.; Foulley, J.L.; Fernando, R.L.; Henderson, C.R. (1990): Estimation of heterogeneous variances using empirical bayes methods: Theoretical considerations.- Mimeo, Dep. Anim. Sci., Univ. Illinois
- Henderson, C.R. (1972): Sire evaluation and genetic trends. Proceedings of the Animal Breeding and Genetics Symposium in Honor of Dr. J. L. Lush, 10-41. Illinois
- Henderson, C.R. (1974): General flexibility of linear model techniques for sire evaluation. J. Dairy Sci. 58, 963-972
- Henderson, C.R. (1984): Application of linear models in animal breeding.- University of Guelph, 1984
- Hill, W.G. (1984): On selection among groups with heterogeneous variance.- Anim. Prod. 39, 473-477
- Kachman, S.; Everett, R. (1993): A multiplicative mixed model when the variances are heterogeneous.- J. Dairy Sci. 76, 859-867
- Meinert, T.R.; Pearson, R.E.; Vinson, W.E.; Cassel, B.G. (1988): Prediction of daughters performance from dam's cow index adjusted for within-herd variance.- J. Dairy Sci. 71, 2220-2231
- Powell, R.L.; Norman, H.D.; Weinland, B.T. (1983): Cow evaluations at different milk yields of herds.- J. Dairy Sci. 66, 148
- Visscher, P.M. (1991): On the estimation of variances within herd-mean production groups.- J. Dairy Sci. 74, 1987
- Wiggans, G.R.; Van Raden, P.M. (1991): Method and effect of adjustment for heterogeneous variance.- J. Dairy Sci. 74, 4350-4357
- Winkelman, A.; Schaeffer, L. R. (1988): Effect of heterogeneity of variance on dairy sire evaluation. - J. Dairy Sci. 71, 3033-3039