

## Karjansisäinen lehmien arvostelu

Jarmo Juga  
Kotieläinten jalostustieteen laitos

---

Helsinki 1985

**Julkaisijat:**

Kotieläinten jalostustieteen laitos, Helsingin Yliopisto, Viikki  
Kotieläinjalostuslaitos, Maatalouden Tutkimuskeskus, Jokioinen

KARJANSISÄINEN LEHMIEN ARVOSTELU

Jarmo Juga  
Kotieläinten jalostustieteen  
pro gradu-työ 1985

## TIIVISTELMÄ

Tutkimuksen tarkoituksena oli kehittää laskentarutiini lehmäindeksin laskemiseksi tietokoneella. Lehmäindeksit laskettiin karjan sisäisesti BLUP (Best Linear Unbiased Prediction)-menetelmällä jokaiselle rodulle erikseen siten, että jalostusarvoa laskettaessa otettiin huomioon lehmän omien usean lypsykauden tuotostietojen lisäksi lehmän kaikkien samassa karjassa olevien sukulaisten tuotos-tiedot. Sukulaisten tiedot huomioitiin mallissa sukulaisuusmatriisin välityksellä. Mallissa otettiin huomioon myös karjan lehmien isien jälkeläisarvostelutulokset. Menetelmä testattiin laskemalla indeksi tarkkailulehmille.

Lehmien jalostusarvot laskettiin Maatalouden Tutkimuskeskuksen VAX 11/780 tietokoneella Fortran-kielillä kirjoitetulla ohjelmalla. Arvostelut laskettiin tuhannen vuonna 1982 karjantarkkailuun kuuluneen karjan lehmille. Lehmien maitotuotokset oli esikorjattu lehmien poikimakeran, poikimakuukauden, poikimavälin ja poikimaiän suhteen. Lehmien isille käytettiin karjantarkkailutuloksista laskettuja sonnien jälkeläisarvostelutuloksia.

BLUP-menetelmällä laskettuja lehmien jalostusarvoja verrattiin nykyisin Suomessa käytössä olevalla CHRISTENSENin menetelmällä laskettuihin jalostusarvoihin. Näiden kahdella eri menetelmällä laskettujen jalostusarvojen väliseksi korrelaatioksi saatiin 0.87. Karjojen sisäinen korrelaatio laskettuna kaikkien karjojen yli oli 0.88. Järjestyskorrelaatio näiden jalostusarvojen välillä oli ayrshire rotuisille lehmille 0.85. BLUP-menetelmällä laskettujen jalostusarvojen keskiarvo oli CHRISTENSENin menetelmällä laskettuja alhaisempi ja hajonta suurempi.

Tutkimuksen tuloksista ei voida päätellä kumpi indekseistä on varmempi. Saaduista tuloksista voidaan huomata, että molemmat indeksit asettavat lehmät keskimäärin samaan järjestykseen ja antavat lehmille suurin piirtein samat standardoidut arvot.

# SISÄLLYSLUETTELO

JOHDANTO.....	1
KIRJALLISUUSKATSAUS AIHEESEEN LIITTYVÄÄN TEORIAAN JA TERMINOLOGIAAN.....	2
-VALINTAINDEKSIT.....	2
-Lehmäindeksi Suomessa.....	7
-Esimerkkejä muiden maiden lehmäindekseistä.....	12
-ELÄINTEN GENEETTISEN ARVON ENNUSTAMINEN	
LINEAARIMALLEISTA.....	16
-Yksisuuntainen kiinteiden tekijöiden malli.....	19
-Satunnaistekijöiden malli.....	27
-Sekamalli.....	28
-Sekamallin yhtälöiden muodostaminen.....	30
-Lehmäindeksi laskettuna BLUP-menetelmällä.....	35
-Eläinten arvostelu pelkästään sukulaistietojen perusteella.....	39
-Systemaattisten ympäristötekijöiden ja ei-additiivisten geneettisten tekijöiden absorbointi muihin yhtälöihin.....	40
-Isien jälkeläisarvostelutulosten hyväksikäyttö....	44
-Sukulaismatriisin kääntematriisiin laskeminen....	50
OMAT TUTKIMUKSET.....	57
-ANEISTO JA MENETELMÄT.....	57
-Malli.....	58
-Lehmäindeksin laskeminen.....	65

-TULOKSET.....	71
-TULOSTEN TARKASTELU.....	77
KIRJALLISUUSLUETTELO.....	80
LIITTEET .....	84

#### KUVAT JA TAULUKOT

TAULUKKO 1: Lehmäindeksin varmuus (PHILIPSSON ym. 1978)....	5
TAULUKKO 2: Esimerkkejä indeksin varmuudesta ja painoker- toimista (CHRISTENSEN 1981).....	12
TAULUKKO 3: Indeksien varmuuden kasvu lisättäessä tietoa emästä, emänisästä ja emän puoleisista puo- lisisarista pelkästään lehmän ja isän tie- doista laskettuun indeksiin (POWELL 1977).....	16
KUVA 1: CHRISTENSENin menetelmällä laskettujen lehmäindeksien jakauma.....	72
KUVA 2: BLUP-menetelmällä laskettujen lehmäindeksien jakauma.....	73
KUVA 3: BLUP-indeksien ja CHRISTENSENin indeksien standardoidut arvot satunnaisesti valituilla 250 lehmällä.....	76
KUVA 4: BLUP-indeksien ja CHRISTENSENin indeksien standardoidut arvot laskettuina lehmille, jotka olivat < 10 lehmän karjoissa.....	84

KUVA 5:	BLUP-indeksin ja CHRISTENSENin indeksin standardoidut arvot laskettuina lehmille, jotka olivat 11 - 20 lehmän karjoissa.....	85
KUVA 6:	BLUP-indeksin ja CHRISTENSENin indeksin standardoidut arvot laskettuina lehmille, jotka olivat 21 - 30 lehmän karjoissa.....	86
KUVA 7:	BLUP-indeksin ja CHRISTENSENin indeksin standardoidut arvot laskettuina lehmille, jotka olivat 31 - 40 lehmän karjoissa.....	87
KUVA 8:	BLUP-indeksin ja CHRISTENSENin indeksin standardoidut arvot laskettuina lehmille, jotka olivat 41 - 50 lehmän karjoissa.....	88
KUVA 9 :	Ayrshire rotuisten karjojen fenotyypisten keskiarvojen jakauma.....	89
KUVA 10:	Ayrshire rotuisten karjojen BLUP-menetelmällä laskettujen geneettisten keskiarvojen jakauma....	90
KUVA 11:	Ayrshire rotuisten lehmien geneettinen edistyminen CHRISTENSENin menetelmällä laskettuna.....	91
KUVA 12:	Ayrshire rotuisten lehmien geneettinen edistyminen BLUP-menetelmällä laskettuna.....	92
KUVA 13:	Ayrshire rotuisten lehmien fenotyyppinen edistyminen.....	93

## JOHDANTO

Asetettaessa lehmii geneettiseen paremmuusjärjestykseen täytyy ne pystyä arvostelemaan mahdollisimman suurella varmuudella. Valta-kunnallisesti arvostelun tärkein käyttömuoto on sonninemien va-linta. Karjan sisällä arvostelua tarvitaan valittaessa lehmii, joiden lehmävasikat tullaan käyttämään uusintaan sekä ryhmitel-täessä eläimiä erilaisia jalostusohjelmia varten (esimerkiksi maito-lihaohjelma).

Arvostelua vaikeuttavat erilaiset ympäristötekijät, jotka syste-maattisesti vaikuttavat lehmän tuotokseen. Tällaisten systemaat-tisten tekijöiden, esimerkiksi ruokinnan vaikutus tuotokseen täy-tyy poistaa, jotta esimerkiksi eri karjoissa olevat lehmät tu-lisivat vertailukelpoisiksi. Systemaattisten tekijöiden aiheutta-man virheellisyyden poistamiseksi yksittäisten lehmien iän, poi-kimakerran, poikimakuukauden ja poikimavälin suhteen korjattuja 305 päivän pituisen tuotantojakson tuotoksia verrataan vastaavas-ti korjattuun karjan keskiarvoon.

Lehmien arvostelu perustuu tavallisesti lehmästä ja sen suku-laisista saataviin tietoihin. Sukulaisista saatava tieto lisää arvosteluvarmuutta erityisesti, kun pyritään arvioimaan lehmän ge-neettistä tasoa, siis lehmän arvoa hyvien vasikoiden tuottajana. Keinosiemennysjalostuksessa arviolta 20-30% geneettisestä edisty-misestä johtuu sonninemien valinnasta ja 60-70% huomattavasti te-hokkaammin valittavista sonneista (EVERETT ym. 1977).

Alkionsiirron käyttöön otolla voidaan lisätä lehmien valintainten-siteettiä ja siten tietysti lehmien valinnan merkitystä geneet-



tisen edistymisen kannalta. Tällöin olisi tärkeää kiinnittää huomiota lehmäarvostelujen varmuuteen, jotta uudistamiseen valittaisiin varmasti geneettisesti parhaat lehmät.

Tämän työn tarkoituksena on kehittää tietokoneohjelma, jolla voidaan laskea lehmien jalostusarvo karjansisäisesti HENDERSONin (1975a) kehittämällä menetelmällä. Kirjallisuusosassa luodaan katsaus lehmien arvosteluun, eläinten jalostusarvon ennustamiseen lineaarimallien avulla sekä aiheeseen liittyvään terminologiaan. Omien tutkimusten osassa esitetään aineisto ja menetelmä, jolla laskenta suoritetaan sekä verrataan tällä menetelmällä laskettuja indeksiarvoja Suomessa nykyisin käytössä olevalla menetelmällä laskettuihin indeksiarvoihin.

## KIRJALLISUUSKATSAUS AIHEESEEN LIITTYVÄÄN TEORIAAN JA TERMINOLOGIAAN

### VALINTAINDEKSIT

Lehmien arvostelu perustuu yleensä lehmille laskettavaan indeksiin, joita on periaatteiltaan kahdenlaisia. Tuotantoindeksi mittaa lehmän kykyä tuottaa. Siitä käytetään usein nimitystä todennäköinen tuotantokyky (TTK). Lehmäindeksillä ennustetaan lehmän todennäköistä geneettistä arvoa ja siten lehmän kykyä jättää hyviä vasikoita. Se on siis lehmän todennäköinen jalostusarvo (TJA) (esim. LASLEY 1978).

### Tuotantoindeksi

Tuotantoindeksi voidaan laskea joko absoluuttisilla tai suhteellisilla arvoilla. Indeksillä ilmoitetaan todennäköisen tuotannon tule-

vaisuudessa esim. seuraavalla laktaatiokaudella ja se perustuu edellisten laktaatiokausien tuotoksiin. Indeksiä laskettaessa käytetään periytyvyysasteen ( $h^2$ ) sijasta toistuvuuden ( $r$ ) arvoja. Tuotantoindeksiä laskettaessa voidaan käyttää systemaattisten tekijöiden suhteen korjattuja tuotoksia, jolloin eläinten arvostelu saadaan vertailukelpoisemmaksi kuin käytettäessä korjaamattomia tuotostietoja. Tuotantoindeksiä voidaan käyttää karjansisäiseen arvosteluun järjestämään lehmät arvojärjestykseen niiden tuotantokyvyn mukaan esimerkiksi päätettäessä kuinka kauan yksittäisiä lehmiä kannattaa pitää (PHILIPSSON ym. 1978).

Todennäköinen tuotantokyky lasketaan esimerkiksi Lasleyn (1978) mukaan kaavasta :

$$TTK = \frac{n r}{1+(n-1)r} (y - \bar{y}), \text{ jossa}$$

$n$  = yksilön tuotosten lukumäärä

$r$  = toistuvuuskerroin

$y$  = yksilön omien tuotosten keskiarvo

$\bar{y}$  = karjan samanikäisten eläinten  
tuotosten keskiarvo

Todennäköinen tuotantokyky kuvaa lehmän fenotyyppistä arvoa. Arvosteluvarmuutta voidaan parantaa käyttämällä useiden havaintojen keskiarvoja, mutta toisaalta mitä korkeampi ominaisuuden toistuvuuskerroin on sitä vähemmän arvosteluvarmuus lisääntyy havaintojen lukumäärän lisääntyessä.

Tuotantoindeksiin voidaan liittää myös muita ominaisuuksia kuin pelkkä maitomäärä. Tulos voidaan laskea tulevana nettotuotoksena,

joka ilmoitetaan esim. markkoina ja jossa otetaan huomioon myös muut taloudelliseen tuotokseen vaikuttavat tekijät kuin maitotuos. Lehmien valinnan ja karsinnan karjan sisällä tulisi siten perustua lehmän aikaisemmasta ja nykyisestä fenotyypistä ennustettuun tulevaan taloudelliseen arvoon, johon laskettaisiin mukaan myös vasikan geneettinen arvo (PEARSON ja MILLER 1978).

#### Lehmäindeksin varmuus

Lehmäindeksillä pyritään kuvaamaan parhaalla mahdollisella tavalla lehmän geneettistä arvoa eli jalostusarvoa. Se voidaan määrittää polveutumisen, yksilön omien tuotoksien, puoli- ja täyssisarten, jälkeläisten tai edellä mainittujen eri lähteiden yhdistelmien perusteella. Varmin tulos arvioitaessa jalostusarvoa saadaan jälkeläisarvostelun perusteella, mutta pienen jälkeläismäärän takia lehmäindeksiä laskettaessa päästään varmimpaan arvioon yhdistämällä lehmän omia ja sen sukulaisista saatuja tietoja. Suurin lisäys arvostelun varmuudessa saadaan lisättäessä indeksiin lehmän omien tietojen lisäksi lehmän isän tiedot. Tyttäreiden, emänpuoleisten puolisisarten, emänemän ja emän emänpuoleisten puolisisarten tietojen hyväksikäyttö lisää arvosteluvarmuutta isän tietoja huomattavasti vähemmän (vrt taulukko 1).

Mitä useamman sukulaisen tietoja käytetään hyväksi lehmän omien tuotostietojen ohella lehmäindeksiä laskettaessa sitä varmempi arvio eläimen geneettisestä arvosta saadaan. Nykyisin on käytävissä kaikkien lypsykausien rekisteritiedot, joten haluttaessa lehmälle voidaan laskea jalostusarvo käyttäen hyväksi useampaa laktaatiokautta. Tällöin tulee toisaalta kiinnittää huomiota siihen, ettei sukupolvien väli kasva tarpeettoman suureksi.

Taulukko 1: Lehmäindeksin varmuus eli todellisen jalostusarvon ja indeksin välinen korrelaatio ( $r_{PI}$ ) erilaisista lähteistä laskettuna (PHILIPSSON ym. 1978).

Tietoa	sukutaulu	n = 1	n = 2	n = 3	n = 4	n = 5
	( n = 0 )					
a	-	.50	.60	.65	.67	.69
a + b	.32	.56	.63	.67	.70	.71
a + c	.47	.62	.67	.71	.73	.74
a + b + c	.57	.67	.71	.74	.75	.76

a = n laktaatiokautta lehmältä itseltään

b = 3 laktaatiokautta emältä

c = 1 laktaatiokausi sadalta isänpuoleiselta puolisisarelta

### Lehmäindeksin rakenne

Lehmäindeksiä laskettaessa lehmien arvostelua varten käytetään valtakunnallisessa valinnassa hyväksi useimmissa maissa ns. kaksiportaista valintamenetelmää, jolloin indeksiin lasketaan mukaan vain maidontuotanto-ominaisuus. Se lasketaan samalla tavalla kuin se on laskettu sonnien indeksissä, jotta indeksit olisivat vertailukelpoisia. Tämän jälkeen lehmäindeksin perusteella valitaan yksilöt, jotka tullaan karsimaan vielä muitten haluttujen ominaisuuksien esimerkiksi, rakenteen mukaan. Indeksä muodostuu lehmän omien ja sen sukulaisten tuotoksien tiedoista.

Muiden ominaisuuksien taloudellisen tärkeyden lisääntyessä pitäisi myös lehmäindeksistä kehittää useiden ominaisuuksien lisäksi karjansisäistä arvostelua varten. Laskettaessa lehmäindeksiä

tulee sille asettaa seuraavat yleiset hyvälle indeksille asetettavat tavanomaiset vaatimukset esimerkiksi LASLEYn (1978) mukaan:

- Indeksinkin tulee maksimoida todellisen geneettisen jalostusarvon (T) ja ko. jalostusarvon arvion eli indeksin (I) välinen korrelaatio ( $r$ ), toisin sanoen sen tulee minimoida todellisen ja odotetun jalostusarvon erotuksen neliö.

Sekä edellyttää, että indeksinkin tulee (esimerkiksi CUNNINGHAM 1969)

- johtaa mahdollisimman todelliseen yksilöittäin väliseen geneettiseen arvojärjestykseen.
- johtaa siihen, että valituksi tulee eläin, jolla on mahdollisimman hyvä genotyyppi.
- usean ominaisuuden tapauksessa olla sellainen, että jalostuksellinen kokonaisedistymisen ei tapahdu jonkin yksittäisen ominaisuuden kustannuksella.

Jotta lehmäindeksi kuvaisi mahdollisimman hyvin lehmien välisiä geneettisiä eroja on siinä otettava huomioon kaikki ympäristötekijät, jotka vaikuttavat systemaattisesti maitotuotokseen. Niinpä karjansisäisessä arvostelussa tuotosluvut on korjattava poikimäen, poikimakerran, poikimakuukauden, poikimavuoden, ja laktaatiokauden pituuden suhteen. Korjauksissa on syytä käyttää laajasta materiaalista saatuja valtakunnallisia korjauskertoimia, ettei korjauksista tulisi lisää virhevaihtelua.

Koska lehmien arvostelu samassa karjassa on sen karjan sisäistä yksilöiden keskinäistä vertailua, ei arvosteluun vaikuta ruokintataso. Lehmien isien jälkeläisarvostelutulosten lisääminen leh-

mien arvosteluun varmentaa arvostelua, samalla eri karjojen lehmien geneettiset arvot saadaan vertailukelpoisiksi keskenään.

Lehmäindeksien laskumenetelmissä voidaan ottaa huomioon karjojen väliset erot myös korjaamalla karjojen keskiarvojen erot vastamaan karjojen geneettisiä tasoeroja.

### Lehmäindeksi Suomessa

Suomessa käytössä oleva lehmäindeksi perustuu tanskalaisen CHRISTENSENin (1981) esittämään indeksiin (SYVÄJÄRVI ja HELLMAN 1983). Indeksillä lasketaan vain vähintään 305 päivää lypsäneille lehmille ja se perustuu sekä lehmän omiin eri lypsykausien 305 päivän pituisen tuotosjakson 4-prosenttisiin maitotuotoksiin että sen isän ja emän isän jälkeläisarvostelutuloksiin. Maitotuotokset korjataan poikimaisiin, poikimakerrotaan, poikimavälin ja poikimakuukauden suhteen.

Indeksin laskemisen helpottamiseksi oletetaan, että (CHRISTENSEN 1981):

- Kaikki tuotokset on korjattu isän, poikimisajankohdan, karjan, vuoden, poikimakuukauden ym. kiinteiden tekijöiden suhteen.
- Eri tuotoskerroilla on sama  $h^2$ .
- Saman lehmän kahden tuotoksen välinen fenotyypin ja genotyypin korrelaatio = toistuvuus ( $r$ ),  $r = 1$ .
- Luokansisäinen korrelaatio ( $t$ ) isän puoleisten puolisisarten välillä =  $1/4 h^2$ .
- Valinnan ja geneettisen edistymisen vaikutus tunnetaan.

Jalostusarvo voidaan määrittää sukulaisten tietojen ja eläimen omien tuotostietojen yhdistelmästä eli CHRISTENSENIN (1981) mukaan lehmäindeksi muodostuu polveutumisindeksistä  $I_1$  ja tuotosindeksistä  $I_2$ .

### Polveutumisindeksi $I_1$

Polveutumisindeksi lasketaan isän ja emänisän jälkeläisarvostelujen perusteella. Polveutumisindeksi mahdollistaa lehmän vertaamisen toisten karjojen lehmiin koska vanhemmat yhdistävät lehmän peruspopulaatioon.

$$I_1 = 0.5 (I_F + I_M), \text{ missä}$$

$I_F$  = isän jälkeläisarvostelu 4%-maitotuotoksen perusteella

$$= ((g_i + s_{ij}) - \overline{(g_i + s_{ij})}) \times 2 \\ = (\text{sonni} - \text{rotukeskiarvo}) \times 2$$

$I_M = 0.5 I_{mF} = 0.5 \times (\text{emänisän jälkeläisarvostelu})$   
= emän indeksi

Indeksin varmuus on indeksin ja todellisen jalostusarvon välinen korrelaatio ( $r_{FI}$ ). Varmuuden neliö ( $r_{FI}^2$ ) kuvaa indeksin toistuvuutta. Toistuvuus saadaan CRISTENSENIN (1981) mukaan seuraavasti

$$r_1^2 = 0.25 (r_F^2 + r_M^2), \text{ missä}$$

$$r_F^2 = \frac{n_F \cdot 0.25 \cdot h^2}{1 + (n_F - 1) \cdot t}, \text{ isän jälkeläisarvostelun toistuvuus}$$

$$r_m^2 = \frac{1(n_{m^P} - 0.25h^2)}{4(1+(n_{m^P}-1)t)}, \text{ emänisän jälkeläisarvostelun toistuvuus}$$

$n_{m^P}$  = isän jälkeläisarvostelun tyttärien lukumäärä

$h$  = 0.25 = perityvyysaste

$t$  =  $1/4 h^2 = 0.0625$

$n_{m^P}$  = emänisän jälkeläisarvostelun tyttärien lukumäärä

### Karjakeskiarvojen laskenta

Karjoille lasketaan fenotyyppinen ja genotyyppinen keskiarvo. Fenotyyppistä keskiarvoa tarvitaan verrattaessa lehmää karjan sisällä ja genotyyppistä keskiarvoa verrattaessa lehmää karjojen välillä.

$B_a$  = fenotyyppinen keskiarvo; karjan lehmien esikorjattujen tuotosten keskiarvo

$I_B = b_1 \bar{I}_1$  = genotyyppinen keskiarvo, missä

$$b_1 = \frac{1 + (m - 1) r}{1 + (m - 1) r \bar{r}^2}, \text{ missä}$$

$m$  = karjan lehmäluku

$r$  = 0.15 = lehmien sukulaisuus karjan sisällä

$\bar{r}^2$  = karjan lehmien polveutumisindeksien toistuvuuden keskiarvo

$\bar{I}_1$  = karjan lehmien polveutumisindeksien keskiarvo



Tuotosindeksi I<sub>2</sub>

Lehmän oma vaikutus huomioidaan indeksissä laskemalla lehmälle tuotosindeksi (CHRISTENSEN 1981).

$$I_2 = \frac{n h^2}{1+(n-1) r} Y_i, \text{ missä}$$

$$Y_i = Y_k - \bar{Y} - \frac{m}{m+2} (B_a - \bar{Y}) + I_B, \text{ missä}$$

$Y_k$  = lehmän esikorjattujen tuotosten keskiarvo

$\bar{Y}$  = rodun kaikkien lehmien esikorjattujen tuotosten keskiarvo

$r$  = tuotosten toistuvuus = 0.40

Kaavassa ( $y_k - \bar{Y}$ ) kuvaa lehmän omaa vaikutusta,  $(m/(m+2))(B_a - \bar{Y})$  kuvaa karjan fenotyypistä vaikutusta. Lehmän oman tuotoksen poikkeamasta vähennetään karjan fenotyypinen taso ja siihen lisätään karjan genotyypinen taso, jotta lehmät saataisiin vertailukelpoisiksi indeksinsä välityksellä koko maassa. Geneettistä edistymistä ei ole erikseen otettu huomioon, koska polveutumisindeksit lasketaan viimeksi tehtyjen sonnien jälkeläisarvostelutulosten perusteella. Toisaalta lehmäindeksi on käytössä niin lyhyen ajan, että sinä aikana ei ehdi tapahtua tulokseen oleellisesti vaikuttavaa geneettistä edistymistä.

Tuotosindeksin toistuvuus lasketaan kaavasta (CHRISTENSEN 1981)

$$r_2 = \frac{n h^2}{1+(n-1) r}$$

$n$  = tuotosten lukumäärä

Lehmäindeksi  $I_3$ 

Polveutumisindeksi ja tuotosindeksi painotetaan toistuvuuskertoimista saatavilla kertoimilla ja lasketaan yhteen. Tällöin saadaan lehmälle kokonaisindeksi, joka kuvaa lehmän jalostusarvoa (CHRISTENSEN 1981)

$$I_3 = \frac{1 - r_1^2}{1 - r_1^2 r_2^2} I_1 + \frac{1 - r_2^2}{1 - r_1^2 r_2^2} I_2$$

Lehmäindeksin toistuvuudeksi saadaan

$$r_3^2 = \frac{r_1^2 + r_2^2 - 2 r_1 r_2}{1 - r_1^2 r_2^2}$$

Jos tätä lehmäindeksiä haluttaisiin käyttää pelkästään karjansisäiseen lehmien arvosteluun voitaisiin sitä yksinkertaistaa siten, että tuotosindeksiä laskettaessa ei otettaisi ollenkaan huomioon karjan fenotyyppistä ja genotyyppistä vaikutusta. Tällöin lehmän tuotosindeksi on lehmän edellisten tekijöiden suhteen korjaamaton todennäköinen jalostusarvo:

$$\text{todennäköinen jalostusarvo} = \frac{n h^2}{1+(n-1) r} (Y_k - \bar{Y})$$

Lehmäindeksin standardointi

Jos merkitään rodun lehmäindeksin keskiarvoa  $I_3$ :lla ja keskihajontaa  $\sigma_{I_3}$ :lla saadaan indeksi standardoiduksi (keskiarvo = 100 ja keskihajonta = 10) lausekkeella

$$I = \frac{I_3 - \bar{I}_3}{\sigma_{I_3}} 10 + 100$$

Taulukko 2: Esimerkkejä indeksien toistuvuudesta ja painokertoimista eri tuotostiedoilla. (CHRISTENSEN 1981).

Sukutaulu					Lehmä	Toistuvuus			Painok.	
emänemä	emänisä	emä	isä	itse	$r_1^2$	$r_2^2$	$r_3^2$	$b_1$	$b_2$	
m	n	m	n	m						
0	0	1	0	1	.06	.25	.29	.76	.95	
3	100	3	0	1	.14	.25	.33	.78	.89	
0	0	1	100	1	.28	.25	.42	.81	.77	
3	100	3	100	1	.35	.25	.46	.82	.71	
0	0	1	0	3	.06	.42	.44	.60	.96	
3	100	3	0	3	.14	.42	.47	.62	.92	
0	0	1	100	3	.28	.42	.52	.66	.82	
3	100	3	100	3	.35	.42	.56	.68	.76	
5	200	5	200	5	.38	.48	.61	.64	.76	

$$h^2 = 0.25, r = 0.40$$

m = maitotuotosten lukumäärä

n = tyttärien lukumäärä

Lehmäindeksissä indeksin  $I_1$  kertoimen ( $b_1$ ) suhde indeksin  $I_2$  ker-  
toimeen ( $b_2$ ) ilmoittaa kuinka paljon painoa polveutumisindeksille  
annetaan suhteessa tuotantoindeksiin.

### Esimerkkejä muiden maiden lehmäindekseistä

#### Lehmäindeksi Englannissa

Englannissa käytetään lehmäindeksin laskentaan karjan geneettistä  
tasoa, lehmän isän ja emän ennustettua geneettistä arvoa ja leh-

män omia maitotuotoksia (HILL ja SWANSON 1983). Tuotostiedot korjataan lypsykauden, poikimaiän ja poikimakuukauden suhteen. Lehmät ryhmitellään ensikoihin ja useamman kerran poikineisiin. Tuotostietojen huomioiminen kahtena eri ominaisuutena; ensimmäisen tuotosvuoden tietoina ja myöhempien tuotosvuosien tietoina, tehdään sen vuoksi, että myöhempien laktaatiokausien maitotuotoksien periytyvyysaste on alhaisempi ja varianssi korkeampi kuin ensimmäisen (MAIJALA ja HANNA 1974).

Indeksi lasketaan kaavasta:

$$I = 1/2(\bar{a}_s + \bar{a}_d) + b_s(a_s - \bar{a}_s) + b_d(a_d - \bar{a}_d) + b_1(X_1 - \bar{X}_1) + b_L(X_L - \bar{X}_L), \text{ missä}$$

$a_s$  = isän periyttämiskyky =  $1/2$ (jalostusarvo),  
karjakeskiarvo =  $\bar{a}_s$

$a_d$  = emän lehmäindeksi, karjakeskiarvo =  $\bar{a}_d$

$X_1$  = lehmän 1. tuotos, karjakeskiarvo =  $\bar{X}_1$

$X_L$  = lehmän muiden tuotoksien keskiarvo,  
karjakeskiarvo =  $\bar{X}_L$

$b_s, b_d, b_1, b_L$  = indeksin painokertoimet

Indeksin arvot ilmoitetaan periyttämiskyynä, joka on puolet jalostusarvosta (additiivisesta geneettisestä arvosta). Isien ja emien jalostusarvojen keskiarvo  $1/2(\bar{a}_s + \bar{a}_d)$  on karjan geneettisen tason likiarvo. Isän periyttämiskyky ennustetaan käyttämällä hyväksi tyttärien tuotostiedot kaikista karjoista BLUP tai vastaavaa menetelmää käyttäen. Arvot ovat poikkeamia sovitusta geneettisestä tasosta, jotta voitaisiin poistaa geneettisen edistymisen aiheuttama harha indeksistä. Täten eri-ikäisten sonnien vertailu ei tuota ongelmia. Lehmäindeksit lasketaan sa-

maan tasoon perustuviksi kuin sonneille. Koska sonnien arvostelu nojautuu vain ensikkojen tuotoksiin niin karjan geneettisen tasonkin arvio perustuu vain ensikkojen tuotoksiin.

Kun menetelmä otettiin käyttöön 1982, oli käytettävissä tuotos-tiedot kahdeksalta vuodelta. Jos emälle ei ollut käytettävissä lehmäindeksiä, se arvioitiin karjan geneettistä tasoa laskettaessa isän jälkeläisarvostelutuloksista.

Indeksissä esiintyviä puutteita:

- Karjan geneettinen taso lasketaan sonniarvostelujen keskiarvona ottamatta huomioon luotettavuutta.
- Ei oteta huomioon karjansisäisiä korrelaatioita, jotka aiheutuvat eläinten välisistä sukulaisuussuhteista.

Indeksi on kuitenkin suhteellisen vaivaton laskea suurellekin määrälle eläimiä ja sillä saadaan subjektiivisesti arvioituna järkeviä tuloksia (HILL ja SWANSON 1983).

Jos arvostelu olisi vain karjansisäistä, voitaisiin karjan geneettinen taso jättää indeksissä huomioimatta. Tällöin tosin tuot-taisi vaikeuksia laskea indeksi sellaiselle lehmälle, jonka emä on toisesta karjasta, koska indeksissä otetaan huomioon emän lehmäindeksi, mikä ei ole vertailukelpoinen karjojen välillä, jos karjojen geneettistä tasoa ei oteta huomioon.

### Lehmäindeksin laskeminen USA:ssa

USA:ssa käytetään lehmäindeksin laskemiseen USDA-DHIA-menetelmää (POWELL ym. 1976), jossa otetaan huomioon lehmän omat tuotos-tiedot ja lehmän isän jälkeläisarvostelun tiedot. Indeksi laske-taan kaavasta:

$I = 1/2(w(\text{lehmän } \overline{MCD}) + (1 - w)\text{isän PD})$ , missä  
lehmän  $\overline{MCD}$  = lehmän korjattu tuotoskeskiarvon poikkeama saman ikäryhmän keskiarvosta, lukuun sisältyy ikäryhmän muiden lehmien isien geneettinen arvo

isän PD (predicted difference) = isän ennustettu arvo  
=  $r(\text{isän } \overline{MCD})$ ,

missä  $r$ =isän ennustuksen toistuvuus ja  $\overline{MCD}$ =painotettu keskiarvo tyttärien  $\overline{MCD}$ :stä

$w$  = lehmän tiedoille annettava painokerroin

Indeksi kuvaa lehmän periyttämiskykyä, mikä on puolet lehmän jalostusarvosta. Indeksi lasketaan erikseen maitotuotokselle ja rasvatuotokselle.

POWELL (1978) esitti indeksin laajentamista siten, että indeksissä otetaan huomioon myös emän ja emänisän tiedot. Indeksi saa täl-löin muodon:

$$I = 0.5(w(\text{lehmän } \overline{MCD}) + (1-w)\text{isän PD} + (1-w)\text{emän lehmäindeksi})$$

Tässä kaavassa on emänisän tiedot otettu huomioon jo emän lehmäindeksissä. Tietojen yhdistäminen ei POWELL'n mukaan aiheuta indeksin varmuuden huonontumista (taulukko 3).

Taulukko 3: Indeksin toistuvuuden  $r_{II}$  (indeksin ja lehmän genotyypin välisen korrelaation neliö) kasvu lisättäessä tietoja emästä (E), emänisästä (EI) ja emän puoleisista puolisisarista (EPS) pelkästään lehmän ja isän tiedoista leskettuun indeksiin (POWELL, 1978).

Tietoa	ind.	varmuuden kasvu lisättäessä			
	Toistuvuus	E	EI	E+EI	E+EI+EPS
lehmällä 1 tuotos	.33	.06	.04	.08	.08
lehmällä 5 tuotosta	.43	.04	.03	.06	.06

Kun lehmäindeksissä otettiin huomioon lehmän omien tuotostietojen ja isän jälkeläisarvostelutietojen lisäksi myös emän ja emänisän tiedot, saatiin 8% suurempi korrelaatio lehmäindeksin ja lehmän poikien jälkeläisarvostelutuloksien välille. Vastaavaa hyötyä neljästä lähteestä laskettu indeksi ei antanut ennustettaessa lehmän tyttären arvoa. Tämä johtuu ainakin osittain emän ja tyttären välisestä ympäristökorrelaatiosta. Indeksien hajonta suureni 7.6-11.6% laskettaessa indeksi ko. neljän sukulaisen perusteella. (POWELL ym. 1981).

#### ELÄINTEN GENEETTISEN ARVON ENNUSTAMINEN LINEAARIMALLEISTA

Analysoitaessa esim. biologisia havaintotuloksia tarvitaan ensimmäiseksi tarkka kuvaus käytettävästä tilastollisesta mallista. Malli kuvaa eri muuttujat ja niiden väliset suhteet, tarvittavat parametrit, oletukset ja rajoitukset sekä muuttujiin liittyvien jakaumien ominaisuudet. Kaikkiin tilastollisiin analyyseihin kuuluu malli kirjoitettiinpa se näkyviin tai ei. Mallin kirjoittaminen auttaa ymmärtämään analyysin periaatteet ja helpottaa sen

toistamista ja tarkistamista. Mallit voidaan jakaa esim. seuraavasti (MAO 1982):

1. Todellinen malli on tuntematon täydellinen malli, joka kuvaa tiedon luonnetta.
2. Ideaalinen malli on paras mahdollinen malli, jonka tutkija voi tehdä kuvatakseen tietoa.
3. Työmalli on kompromissi ideaalimallin ja analyysin joustavan suorittamisen välillä.

Analyysiä tehtäessä ollaan kiinnostuneita eri tekijöiden havaintoarvojen vaihtelusta, niiden aiheuttajista ja näiden tekijöiden suhteista toisiinsa. Tekijät voivat olla riippuvia tai riippumattomia sen mukaan, mitä tekijää tutkitaan (riippuva) muista tekijöistä (riippumattomat) selittyvänä. Näiden tekijöiden välisiä suhteita kuvataan yhtälöillä. Yhtälöt voidaan ryhmitellä neljään perustyyppiin sen mukaan selittävätkö ne eksakteja suhteita vai ovatko ne tilastollisia yhtälöitä tai toisaalta, ovatko ne lineaarisia vai ei-lineaarisia.

Kuvattaessa kotieläinten tuotosten havaintoja ja niiden riippuvuutta eri tekijöistä käytetään tilastollista lineaarimallia. Muitakin malleja voitaisiin käyttää, mutta lineaarimallin etuna on sen yleisyys, laajat sovellutusmahdollisuudet ja pitkälle kehitetyt analyysimenetelmät sekä laskentatekniikka.

#### Mallin kuvaus

Minkä tahansa eri tyyppisen mallin kuvaukseen kuuluu kolme osaa (MAO 1982):



1. Yhtälö käsittää kaikki malliin kuuluvat tekijät ja muuttujat, jotka voivat selittää havaintojen vaihtelua sekä näiden tekijöiden ja muuttujien väliset yhdysvaikutukset.
2. Odotusarvot ja varianssi-kovarianssimatriisit kaikista yhtälöön kuuluvista satunnaismuuttujista.
3. Oletukset, rajoitukset ja muut mallin analyysin suorittamista helpottavat tai rajoittavat tekijät.

Lehmien tuotoksia ja niihin vaikuttavia tekijöitä kuvaavista yhtälöistä muodostuvista yhtälöryhmistä ratkaistaan halutut muuttujat. Muuttujan arvo ilmaisee ko. tekijän vaikutuksen suuruuden kokonaistuotoksesta. Tekijät voivat olla joko kiinteitä, esim. vuoden vaikutus, tai satunnaisia, esim. perinnöllisyyden vaikutus. Tällaiset tilastolliset lineaarimallit ovat olleet kotieläinjalostuksen olennaisena osana jo vuosikymmeniä, ja yhä monipuolisempia sekä varmempia malleja pyritään koko ajan kehittämään.

Yksittäisen lehmän maitotuotosta voidaan kuvata lineaarisella yhtälöllä, jossa havainto esitetään vastaavan populaation keskiarvon ja keskiarvosta lasketun poikkeaman summana,

$$y_i = \mu + e_i, \text{ missä}$$

$$y_i = \text{havainto eläimeltä } i$$

$$\mu = \text{keskiarvo}$$

$$e_i = \text{poikkeama keskiarvosta}$$

$e_i$  kuvaa satunnaista 'virhettä' eli poikkeamaa, jota malli ei selitä. Sen odotusarvo  $E(e_i) = 0$  ja varianssi  $\sigma_e^2$ . Tällaisesta yhtälöstä ratkaistaan pienimmän neliösumman (least squares, LS) me-

netelmän mukainen ratkaisu  $\mu$ :lle

$$\mu = (1/n) \sum y_i = (\sum y_i)/n$$

missä  $n$  on havaintojen lukumäärä.

### Yksisuuntainen kiinteiden tekijöiden malli (One-way fixed model)

Lisätään edelliseen yhtälöön yksi selittävä tekijä, joka on kiinteä. Kiinteänä tekijänä pidetään ei-satunnaisia näytteitä tekijän muuntelusta. Tyypillisiä kiinteitä tekijöitä ovat sukupuoli, kausi, ikä, poikimakerta ja alue. Aina ei ole helppoa päättää onko jokin tekijä kiinteä vai satunnainen. Tarkastellaan maitotuotosta ja selittäköön kiinteä tekijä iän vaikutusta maitotuotokseen. Malli kirjoitetaan nyt muotoon

$$y_{ij} = \mu + b_i + e_{ij}$$

$y_{ij}$  = j:n eläimen maitotuotos ikäryhmässä  $i$

$\mu$  = keskiarvo (kiinteä tekijä)

$b_i$  = kiinteä tekijä joka kuvaa iän vaikutusta  $i$ :nessä ikäryhmässä

$e_{ij}$  = satunnainen tekijä eli selvittämättömistä tekijöistä johtuva j:n eläimen poikkeama  $i$ :nnen ikäryhmän sisällä, 'virhe'

Tässä tapauksessa tulokseen  $y_{ij}$  kuuluu aina mukaan geneettinen vaikutus (G) ja ympäristön vaikutus (E) huolimatta siitä onko ne kirjoitettu malliin näkyviin vai ei. Ne vaikutukset joita ei erikseen ole huomioitu mallissa, sisältyvät satunnaiseen virhetermiin ( $e_{ij}$ ).

Esimerkki:

Oletetaan, että on tehty seuraavat havainnot ja kirjoitetaan niistä vastaavat mallit:

$$5670 = y_{11} = \mu + b_1 + e_{11}$$

$$4720 = y_{12} = \mu + b_1 + e_{12}$$

$$4940 = y_{13} = \mu + b_1 + e_{13}$$

$$4200 = y_{21} = \mu + b_2 + e_{21}$$

$$4100 = y_{31} = \mu + b_3 + e_{31}$$

$$5200 = y_{32} = \mu + b_3 + e_{32}$$

Tällöin  $i = 1, 2$  tai  $3$  ja  $n_1 = 3$  (ikäryhmän 1 tuotostietojen lukumäärä),  $n_2 = 1$  sekä  $n_3 = 2$ .

Tämä sama voidaan kirjoittaa myös matriisiyhtälöiden muodossa

$$\begin{bmatrix} 5670 \\ 4720 \\ 4940 \\ 4200 \\ 4100 \\ 5200 \end{bmatrix} = \begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{21} \\ y_{31} \\ y_{32} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} + \begin{bmatrix} e_{11} \\ e_{12} \\ e_{13} \\ e_{21} \\ e_{31} \\ e_{32} \end{bmatrix}$$

tai yleisesti

$$\underline{y} = \underline{X} \underline{b} + \underline{e},$$

missä:

$\underline{y}$  = maitotuotosten vektori

$\underline{X}$  =  $\underline{b}$ :n kerroin- (insidenssi) matriisi, joka on tunnettu ja sisältää ykkösiä ja nollia

$\underline{b}$  = kiinteiden tekijöiden (ikä) vektori, tuntematon

$\underline{e}$  = satunnaisten virhetekijöiden vektori

Kiinteiden tekijöiden  $b_i$  ratkaisut saadaan laskettua pienimmän neliösumman menetelmällä minimoimalla satunnaisen virhetermin neliösumma (esimerkiksi VAN VLECK 1979)

$$SSE = \sum e_{ij}^2 = \sum (y_{ij} - \mu - b_i)^2$$

Ensimmäiseksi lasketaan summa jokaiselle malliin kuuluvalla tekijälle lukuunottamatta virhetermiä. Esimerkin mallissa nämä neljä tekijää ovat  $\mu$ ,  $b_1$ ,  $b_2$  ja  $b_3$ .  $\mu$ :tä vastaava summa sisältää jokaisen tuotoksen, jonka yhtälöön kuuluu  $\mu$ , joka esiintyy luonnollisesti kaikkien yksittäisten tuotosten osana, joten  $y_{..} = 29130$ .  $b_1$ :ta vastaava summa koostuu kaikista tuotoksista joihin kuuluu  $b_1$  eli tuotoksista, joissa  $i = 1$  (yhteensä  $n_1$  kappaletta). Täten  $y_{1.} = 5670 + 4720 + 4940 = 15630$ . Samalla tavalla on  $b_2$ :ta vastaava summa  $y_{2.} = 4200$  ja  $b_3$ :a vastaava summa  $y_{3.} = 9300$ .

Tekijöiden ratkaisemiseksi pienimmän neliösumman menetelmällä kirjoitetaan yhtälöt yhtälöryhmäksi merkitsemällä jokainen summa yhtäsuureksi selittävän mallinsa kanssa. Virhetermejä ei kirjoiteta mukaan yhtälöihin.  $y_{..}$ :tä kuvaava malli koostuu kaikista  $n$ . tuotoksesta

$n_1$  kappaletta tuotoksista kuvaa malli  $\mu + b_1$ ,

$n_2$  kappaletta tuotoksista kuvaa malli  $\mu + b_2$  ja

$n_3$  kappaletta tuotoksista kuvaa malli  $\mu + b_3$ ,

joten  $y_{..}$  saadaan muotoon  $(n_1 + n_2 + n_3)\mu + n_1b_1 + n_2b_2 + n_3b_3$ . Vastaavasti kirjoitetaan mallit tekijöille  $y_1$ ,  $y_2$  ja  $y_3$ . Pienimmän neliösumman yhtälöiksi saadaan siten

$$\begin{aligned}\mu &: n_1 \hat{\mu} + n_1 \hat{b}_1 + n_2 \hat{b}_2 + n_3 \hat{b}_3 = y_{..} \\ b_1 &: n_1 \hat{\mu} + n_1 \hat{b}_1 = y_{1.} \\ b_2 &: n_2 \hat{\mu} + n_2 \hat{b}_2 = y_{2.} \\ b_3 &: n_3 \hat{\mu} + n_3 \hat{b}_3 = y_{3.}\end{aligned}$$

Merkinnällä ' $\hat{\cdot}$ ' kuvataan tässä vaiheessa ratkaisua.

Esimerkkitapauksessa yhtälöt ovat

$$\begin{aligned}6 \hat{\mu} + 3 \hat{b}_1 + 1 \hat{b}_2 + 2 \hat{b}_3 &= 29130 \\ 3 \hat{\mu} + 3 \hat{b}_1 &= 15630 \\ 1 \hat{\mu} + 1 \hat{b}_2 &= 4200 \\ 2 \hat{\mu} + 2 \hat{b}_3 &= 9300\end{aligned}$$

Yhtälöistä huomataan VAN VLECK'n (1979) mukaan, että

- kertoimet ovat symmetrisiä, eli kertoimet ensimmäisellä rivillä ovat samat kuin kertoimet ensimmäisellä sarakkeella
- halkaisijan ulkopuoliset  $b_j$ :n kertoimet saavat arvon nolla

ja ennen kaikkea että

- kolmen  $b_j$ :n yhtälön summa on yhtäsuuri kuin  $\mu$ :n yhtälö

Viimeisestä kohdasta seuraa, että vaikka yhtälöryhmässä on neljä yhtälöä ja neljä tuntematonta, ei yhtälöryhmällä ole yksikäsitteistä ratkaisua, koska yhtälöt eivät ole toisistaan riippumattomia. Ratkaisun löytämiseksi on tehtävä jokin rajoitus. Erilaisia rajoitustapoja on useita (esimerkiksi VAN VLECK 1979 ja MAO 1982):

- tekijän nollaus on laskuteknisesti yksinkertaisin.

Nollaksi voidaan asettaa esimerkiksi  $\hat{\mu}$ ,

jolloin eliminoidaan koko  $\hat{\mu}$ :n

yhtälö sekä  $\hat{\mu}$  jäljellejääneistä yhtälöistä.

Yhtälöryhmä tulee muotoon

$$n_1 \hat{b}_1 = y_1.$$

$$n_2 \hat{b}_2 = y_2.$$

$$n_3 \hat{b}_3 = y_3.$$

Nollaksi voidaan asettaa myös mikä tahansa muu

tekijä esim.  $\hat{b}_3 = 0$ . Tällöin  $\hat{b}_3$ :n yhtälö sekä

$\hat{b}_3$  eliminoidaan yhtälöryhmästä, joka tulee muotoon

$$n \hat{\mu} + n_1 \hat{b}_1 + n_2 \hat{b}_2 = y_{..}$$

$$n_1 \hat{\mu} + n_1 \hat{b}_1 = y_1.$$

$$n_2 \hat{\mu} + n_2 \hat{b}_2 = y_2.$$

- laskuteknisesti monimutkaisin on rajoitus

$$\hat{b}_1 + \hat{b}_2 + \hat{b}_3 = 0.$$

Tämä yhtälö lisätään pienimmän neliösumman

yhtälöihin. Jotta kertoimet saataisiin lisäyksen

jälkeen symmetrisiksi täytyy jokaiseen muuhun

yhtälöön lisätä LaGrange-kerroin.

Yhtälöryhmä kirjoitetaan tämän jälkeen muotoon

$$n \hat{\mu} + n_1 \hat{b}_1 + n_2 \hat{b}_2 + n_3 \hat{b}_3 + 0 \lambda = y_{..}$$

$$n_1 \hat{\mu} + n_1 \hat{b}_1 + 1 \lambda = y_1.$$

$$n_2 \hat{\mu} + n_2 \hat{b}_2 + 1 \lambda = y_2.$$

$$n_3 \hat{\mu} + n_3 \hat{b}_3 + 1 \lambda = y_3.$$

$$0 + 1 \hat{b}_1 + 1 \hat{b}_2 + 1 \hat{b}_3 + 0 = 0$$

Ratkaisuja jotka saadaan rajoituksella  $\hat{\mu} = 0$ , on yksinkertaisinta tutkia. Tällä tavalla saadaan myös yhtälöryhmä jossa on eniten nollia kertoimina. Ratkaisuiksi saadaan

$$\begin{aligned} \hat{\mu} &= 0 & \hat{b}_2 &= y_2. / n_2 \\ \hat{b}_1 &= y_{1.} / n_1 & \hat{b}_3 &= y_3. / n_3 \end{aligned}$$

On huomattava, että  $\hat{\mu} = 0$  on yksi ratkaisuisista.

Käyttämällä toisia rajoituksia saadaan erilaiset ratkaisut. Siitä huolimatta esim. funktioille  $\mu + b_1$  ja  $b_1 - b_j$  saadaan samat estimaatit riippumatta siitä, mitä rajoitusta käytetään. Tällaisia funktioita sanotaan estimoitavissa oleviksi funktioiksi.

Ratkaisumenetelmä voidaan esittää matriisimuodossa yleisesti. Pienimmän neliösumman ratkaisut  $\underline{b}$ :lle saadaan, kuten edellä, minimoimalla havaittujen  $\underline{y}$ :iden mallin mukaisista odotusarvoista laskettujen poikkeamien neliösumma, ts. minimoimalla vektorin  $\underline{e} = \underline{y} - \underline{X}\underline{b}$  elementtien neliösumma. Merkitsemällä transpoosilla voidaan neliösumma kirjoittaa muotoon

$$\begin{aligned} \underline{e}'\underline{e} &= (\underline{y} - E(\underline{y}))'(\underline{y} - E(\underline{y})) \\ &= (\underline{y} - \underline{X}\underline{b})'(\underline{y} - \underline{X}\underline{b}) \\ &= \underline{y}'\underline{y} - 2\underline{b}'\underline{X}'\underline{y} + \underline{b}'\underline{X}'\underline{X}\underline{b} \end{aligned}$$

Jotta saataisiin ne  $\underline{b}$ :n ratkaisut, jotka minimoivat  $\underline{e}'\underline{e}$ :n derivoidaan  $\underline{e}'\underline{e}$   $\underline{b}$ :n suhteen ja määritetään ne  $\underline{b}$ :n arvot joilla kyseinen derivaatta saa arvon nolla:

$$\frac{\partial(\underline{e}'\underline{e})}{\partial \underline{b}} = -2\underline{X}'\underline{y} + 2\underline{X}'\underline{X}\underline{b} = 0$$

$$\Leftrightarrow \underline{X}'\underline{X}\underline{b} = \underline{X}'\underline{y}$$

Tätä yhtälöä kutsutaan normaaliyhtälöksi ja sen ratkaisut ovat pienimmän neliösumman (least squares, LS) ratkaisuja. Tähän sisältyy oletus, että kaikkien jäännöstermien varianssi on sama, matriisimerkinnällä  $\text{Var}(\underline{e}) = \underline{I}\sigma_e^2$ . Mikäli näin ei voida olettaa, joudutaan yhtälöihin lisäämään varianssi-kovarianssimatriisi  $\underline{V} = \text{Var}(\underline{y}) = \text{Var}(\underline{e})$ . Tällöin saadaan LS-yhtälöryhmien yleinen muoto

$$\underline{X}'\underline{V}^{-1}\underline{X}\underline{b} = \underline{X}'\underline{V}^{-1}\underline{y}$$

Tällaisen yhtälöryhmän ratkaiseminen on  $\underline{V}$ :n suuruuden vuoksi kuitenkin erittäin hankalaa ellei mahdotonta, joten yleensä tyydytään normaaliyhtälöihin.

Ainoastaan silloin kun  $\underline{X}'\underline{X}$ :n yhtälöt ovat toisistaan riippumattomia eli kun matriisin  $\underline{X}'\underline{X}$  aste on täysi (full rank) saadaan  $\underline{b}$ :lle yksikäsitteinen ratkaisu

$$\hat{\underline{b}} = (\underline{X}'\underline{X})^{-1}\underline{X}'\underline{y}$$

Neliömatriisien yhteydessä täysiasteista matriisia kutsutaan säännölliseksi matriisiksi, ts. tällaisen matriisin determinantti on erisuuri kuin nolla. Yleisesti malli, joka sisältää useita kiinteitä tekijöitä, ei voi saada yksikäsitteistä  $\underline{b}$ :n ratkaisua, koska  $\underline{X}'\underline{X}$  ei ole säännöllinen, eikä ole olemassa sen yksikäsitteistä käänteismatriisia  $(\underline{X}'\underline{X})^{-1}$ . Tällaiselle vajaa-asteiselle (non-full rank) matriisille tai singulaarille neliömatriisille voidaan kuitenkin rajoituksen jälkeen laskea yleistetty käänteismatriisi  $(\underline{X}'\underline{X})^{\sim}$ . Tällaisia käänteismatriiseja saadaan äärettömän monta erilaista valittujen rajoitusten mukaan. Ratkaisuina saata-  
vat  $\underline{b}$ :n arvot eivät ole yksikäsitteisiä, vaan riippuvat valitusta käänteismatriisista  $(\underline{X}'\underline{X})^{\sim}$ . Näin saatavia ratkaisuja merkitään



"~":lla, kun taas yksikäsitteisiä ratkaisuja, joista käytetään nimitystä likiarvo (eli estimaatti), merkitään "ˆ":lla.

Edellisen esimerkin pienimmän neliösumman yhtälöt ovat matriisimuodossa

$$\begin{bmatrix} 6 & 3 & 1 & 2 \\ 3 & 3 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} \mu \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 29130 \\ 15630 \\ 4200 \\ 9300 \end{bmatrix}$$

$\underline{X}'\underline{X}$  matriisin rivit ja sarakkeet eivät ole riippumattomia, sillä kolmen viimeisen rivin, vaihtoehtoisesti sarakkeen, summa on yhtäsuuri kuin ensimmäinen rivi (sarake). Matriisi  $\underline{X}'\underline{X}$  on siten epä-säännöllinen eikä sille voida laskea yksikäsitteistä käänteismatriisia. Tällaisessa yhden kiinteän tekijän tapauksessa matriisin asteluku = yhtälöiden lukumäärä - 1. Yhtälöiden ratkaisemiseksi on tehtävä rajoitusehto. Tämä rajoitus tehdään samalla tavalla kuin edellä poistamalla jokin rivi ja sitä vastaava sarake tai käyttämällä LaGrange-kertoimia. Näin saatava käänteismatriisi on siis yleistetty käänteismatriisi.  $\underline{b}$ :n ratkaisu on silloin

$$\underline{\tilde{b}} = (\underline{X}'\underline{X})^{-1}\underline{X}'\underline{y}$$

Vaikka ratkaisut eivät ole yksikäsitteisiä voidaan niitä kuitenkin käyttää hyväksi, koska kuten todettiin, on aina olemassa es-timoitavissa olevia funktioita, joilla on yksikäsitteiset ratkaisut.

Satunnaistekijöiden malli (random model)

Tekijät, joiden luokat jakautuvat satunnaisesti eli joilla on hajonta, ovat satunnaisia tekijöitä. Satunnaisen tekijän vaikutusta kuvaamaan on populaatiosta valittu näyte, joka on satunnainen, eikä siis ole toistettavissa. Tällaisia tekijöitä ovat esimerkiksi eläimessä itsessään olevien perintötekijöiden vaikutus tai isältä tulleiden perintötekijöiden vaikutus. Satunnaisten tekijöiden mallia käytetään yleisesti laskettaessa varianssikomponentteja. Ryhmitellään tulokset esimerkiksi eläimen isän mukaan. Malli on tällöin

$$y_{ij} = \mu + a_i + e_{ij}$$

missä

$\mu$  = keskiarvo

$a_i$  = satunnainen tekijä, yhteinen kaikille eläimille, joiden isä on  $i$ . (Tämä on yhtäkuin puolet isän jalostusarvosta eli additiivisesta geneettisestä arvosta, koska jokaiseen isän jälkeläiseen on siirtynyt satunnaisesti kunkin lokuksen jompikumpi alleeli).  $E(a_i) = 0$  ja  $E(a_i^2) = \sigma_a^2 =$  isänpuoleinen puolisisarkovarianssi  $= (1/4)h^2\sigma_y^2$ , missä  $\sigma_y^2$  on tuotoksen kokonaisvarianssi

$e_{ij}$  = satunnainen virhetermi

Matriisimuodossa satunnaistekijöiden malli kirjoitetaan muotoon

$$\underline{y} = \underline{Za} + \underline{e}$$

missä

$\underline{y}$  = tuotosvektori

$\underline{a}$  = satunnaisten tekijöiden vektori

$\underline{Z}$  = a:n kerroinmatriisi

$\underline{e}$  = satunnaisten virhetermien vektori

Mallissa tuotoksen poikkeama keskiarvosta on yhtäsuuri kuin tuotoksen geneettisen ja ympäristövaikutuksen summa eli  $a_i + e_{ij} = G_{ij} + E_{ij}$  siten, että  $E(e_{ij}) = 0$  ja  $\sigma_e^2 = \sigma_y^2 - \sigma_a^2 = \sigma_y^2 (1 - 1/4h^2)$ .  $n_i$  on isän  $i$  tyttärien lukumäärä.

### Sekamalli (mixed model)

Edellisten mallien matriisiesitysten mukaisesti sekä kiinteitä että satunnaisia tekijöitä sisältävä malli, ns. sekamalli voidaan kirjoittaa yleisessä muodossa MAOn (1982) mukaan seuraavasti

$$\underline{y} = \underline{X}\underline{b} + \underline{Z}\underline{a} + \underline{e},$$

missä:

$\underline{y}$  =  $n \times 1$  vektori havaintoja

$\underline{b}$  =  $p \times 1$  vektori tuntemattomia kiinteitä tekijöitä

$\underline{X}$  =  $n \times p$  tunnettu  $\underline{b}$ :n kerroinmatriisi

$\underline{a}$  =  $q \times 1$  vektori tuntemattomia satunnaisia tekijöitä

$\underline{Z}$  =  $n \times q$  tunnettu  $\underline{a}$ :n kerroinmatriisi

$\underline{e}$  =  $n \times 1$  vektori virhetermejä jokaiselle  $\underline{y}$ :n havainnolle

Jos  $\underline{b}$  sisältää vain vakion  $\mu$ , joka on yhteinen kaikille  $\underline{y}$ :ille, niin  $X = 1$  ja sekamalli on

$$\underline{y} = \mu \underline{1} + \underline{Z}\underline{a} + \underline{e},$$

jota yleisesti kutsutaan satunnaismalliksi vaikka se teknisesti onkin sekamalli.

Satunnaisvektoreiden odotusarvot ovat:

$$E \begin{bmatrix} \underline{y} \\ \underline{a} \\ \underline{e} \end{bmatrix} = \begin{bmatrix} \underline{Xb} \\ \underline{0} \\ \underline{0} \end{bmatrix}$$

Varianssi-kovarianssimatriisi on:

$$\text{Var} \begin{bmatrix} \underline{y} \\ \underline{a} \\ \underline{e} \end{bmatrix} = \begin{bmatrix} \underline{V} & \underline{ZG} & \underline{R} \\ \underline{GZ}' & \underline{G} & \underline{0} \\ \underline{R} & \underline{0} & \underline{R} \end{bmatrix},$$

missä  $\text{Var}(\underline{a}) = E(\underline{a}'\underline{a}) = E \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_q \end{bmatrix} \begin{bmatrix} a_1 & a_2 & \dots & a_q \end{bmatrix} = \underline{G}_{q \times q}$

$\text{Var}(\underline{e}) = E(\underline{e}'\underline{e}) = \underline{R}_{n \times n}$ .  $\text{Cov}(\underline{a}, \underline{e}) = \underline{0}$ , joten  $\underline{e}$ :n ja satunnaistekijöiden välillä ei ole korrelaatiota.

$$\begin{aligned} \text{Cov}(\underline{y}, \underline{e}) &= E(\underline{y} - E(\underline{y}))(\underline{e} - E(\underline{e}))' \\ &= E(\underline{y} - \underline{Xb})(\underline{e} - \underline{0})' = E(\underline{Za} + \underline{e})\underline{e}' \\ &= \underline{Z}E(\underline{a}\underline{e}') + E(\underline{e}\underline{e}') = \underline{R} \end{aligned}$$

$$\begin{aligned} \text{Cov}(\underline{y}, \underline{a}) &= E(\underline{y} - E(\underline{y}))(\underline{a} - E(\underline{a}))' \\ &= E(\underline{y} - \underline{Xb})(\underline{a} - \underline{0})' = E(\underline{Za} + \underline{e})\underline{a}' \\ &= \underline{Z}E(\underline{a}\underline{a}') + E(\underline{e}\underline{a}') = \underline{ZG} \end{aligned}$$

$$\begin{aligned}\text{Var}(\underline{y}) &= \underline{V} = E(\underline{y}-E(\underline{y}))(\underline{y}-E(\underline{y}))' \\ &= E(\underline{y}-\underline{X}\underline{b})(\underline{y}-\underline{X}\underline{b})' \\ &= E(\underline{Z}\underline{y}+\underline{e})(\underline{Z}\underline{a}+\underline{e})' \\ &= \underline{Z}E(\underline{a}\underline{a}')\underline{Z}' + \underline{Z}E(\underline{a}\underline{e}') + E(\underline{e}\underline{a}')\underline{Z}' + E(\underline{e}\underline{e}') \\ &= \underline{Z}\underline{G}\underline{Z}' + \underline{R}\end{aligned}$$

Tarkasti ottaen malliin kuuluu lisäksi oletukset:

- $\underline{y}$  on satunnainen ei-valittu näyte
- $\underline{y}$ :n pituus on ääretön tai ainakin  $n$  on riittävän suuri, jotta se voidaan käsittää äärettömänä
- järjestys jokaisessa luokassa  $\underline{a}_i$  on riippumaton havaintojen määrästä ko. luokassa.

Nämä oletukset eivät kuitenkaan käytännössä ole realistisia esim. jos halutaan geneettistä edistymistä on valintaa pakko suorittaa (MAO 1982).

### Sekamallin yhtälöiden muodostaminen

HENDERSONin (1973) kehittämä BLUP-menetelmä voidaan käsittää sekä pienimmän neliösumman analyysin että tavallisen valintaindeksin laajennettuna muotona tai yhdistelmänä. Menetelmällä korjataan jalostusarvojen ennustamisen ohessa samanaikaisesti ei-geneettisistä lähteistä johtuvat virheet, mikä on suurin ero tavalliseen valintaindeksiin verrattuna. Lisäksi BLUP-menetelmä on valintaindeksiä tehokkaampi, kun eri luokissa on erilaiset määrät havaintoja tai kun halutaan käyttää hyväksi sukulaisten tietoja.

Sekä BLUP- että valintaindeksiä käytettäessä pitää olla tiedossa

genotyyppiset ja fenotyyppiset varianssit. Näiden menetelmien yhteisiä ominaisuuksia ovat:

- Molemmat ovat harhattomia; valintaindeksi on harhaton automaattisesti, kun taas BLUP-yhtälöiden ratkaisuehdoksi asetetaan harhattomuus.
- Ennustusvirheen varianssi minimoidaan.
- Ennusteen ja todellisen jalostusarvon korrelaatiota  $r_{TI}$  maksimoidaan.
- Ennusteet ovat muuten samoja, mutta BLUP-menetelmässä käytetään parhaita harhattomia estimaatteja (BLUE) kiinteistä tekijöistä korjattaessa tuotoksia, kun taas valintaindeksissä oletetaan, että käytetään ko. tekijöiden parametriarvoja. Toisin sanoen ero riippuu siitä, miten hyviä parametrien likiarvoja valintaindeksissä käytetään.

BLUP-menetelmässä yhtälöiden lukumäärä on yhtäsuuri kuin tekijöiden lukumäärä mallissa. Jos kaikki tekijät olisivat kiinteitä olisivat yhtälöt samat kuin pienimmän neliösumman menetelmässä. Ennustuksia laskettaessa joudutaan kuitenkin käyttämään sekamallia, koska mukana on sekä kiinteitä että satunnaisia tekijöitä.

Menemättä yksityiskohtiin voidaan todeta, että yhtälöiden ratkaisu lähtee tarkoituksesta ennustaa funktioryhmää ( $\underline{K}'\underline{b} + \underline{M}'\underline{a}$ ) havaintojen lineaarisen funktion  $\underline{L}'\underline{y}$  avulla siten, että virhevarianssia  $\text{Var}\{\underline{K}'\underline{b} + \underline{M}'\underline{a} - \underline{L}'\underline{y}\}$  minimoidaan jokaiselle ( $\underline{k}'\underline{b} + \underline{m}'\underline{a}$ ) joka kuuluu ( $\underline{K}'\underline{b} + \underline{M}'\underline{a}$ ):n sekä siten, että ennustajan odotusarvo on yhtäkuin ennustettavan odotusarvo  $E(\underline{L}'\underline{y}) = E(\underline{K}'\underline{b} + \underline{M}'\underline{a})$ , siis harhaton. Siten ennustavalla termillä on pienin ennustuksen virhevarienssi (paras, best) kaikkien muiden harhattomien  $\underline{y}$ :n suh-

teen lineaaristen ennustajien joukossa ja sitä kutsutaan parhaaksi suoraviivaiseksi harhattomaksi ennustajaksi (BLUP) (MAO 1982).

Jotta voitaisiin laskea BLUP-ratkaisut eli ennusteet satunnaistekijöille ja BLUE-ratkaisut eli estimaatit kiinteille tekijöille, täytyy ensin muodostaa sekamallin yhtälöt (MME). BLUP-menetelmää käytettäessä ei tarvita tietoja havaintojen jakaumasta, mutta varianssit ja kovarianssit oletetaan tunnetuiksi.

BLUP-menetelmän sekamalliyhtälöiksi saadaan (HENDERSON 1973):

$$\begin{bmatrix} \underline{X}' \underline{R}^{-1} \underline{X} & \underline{X}' \underline{R}^{-1} \underline{Z} \\ \underline{Z}' \underline{R}^{-1} \underline{X} & \underline{Z}' \underline{R}^{-1} \underline{Z} + \underline{G}^{-1} \end{bmatrix} \begin{bmatrix} \underline{b} \\ \underline{a} \end{bmatrix} = \begin{bmatrix} \underline{X}' \underline{R}^{-1} \underline{y} \\ \underline{Z}' \underline{R}^{-1} \underline{y} \end{bmatrix}$$

Näitä yhtälöitä kutsutaan yleisesti HENDERSONin (1973) mukaan sekamallin yhtälöiksi (mixed model equations, MME). Näistä yhtälöistä ratkaistuilla ennusteilla on pienin ennustuksen virhevarianssi kaikkien harhattomien lineaaristen ennusteiden joukosta, eli ennustajan ja ennustettavan välinen korrelaatio maksimoituu. Lisäksi, olettaen että satunnaistekijät jakautuvat normaalisti, maksimoituu eläinten oikean geneettisen arvojärjestyksen ennustamisen todennäköisyys.

Nämä sekamallin yhtälöt ovat yleisiä ja niillä voidaan käsitellä kaikki lineaarisia malleja koskevat ongelmat mukaanlukien kaikki pienimmän neliösumman menetelmät.

Yleisen MME:n kerroinmatriisi on useimmiten vajaa-asteinen, koska rivit ja sarakkeet eivät ole toisistaan riippumattomia kiinteiden tekijöiden vuoksi. Jotta MME-yhtälöt pystytään ratkaisemaan, täytyy kiinteiden tekijöiden osalta tehdä aikaisemmin mainitut ra-

joitukset, jolloin MME:stä saadaan laskettua yleistetty käänteismatriisi. MME-yhtälöihin voidaan tehdä myös haluttuja rajoituksia, jotka eivät vaikuta yhtälöiden ratkaisemiseen, vaan satunnaisten tekijöiden arvoihin. Esimerkiksi jokin tekijä voidaan asettaa vakioksi ym.

### Yleisen sekamallin olettamusten yksinkertaistuksia

Kuten edellä mainittiin, on yleisessä sekamallissa  $\text{Var}(\underline{a}) = \underline{G}$  ja  $\text{Var}(\underline{e}) = \underline{R}$ . Käytännössä joudutaan kuitenkin laskutoimitusten helpottamiseksi tekemään joitakin oletuksia. Yksinkertaisimmassa MME-mallissa oletetaan ensiksikin, että virheet ovat korreloimattomia ja että niillä on sama varianssi, ts. että  $\underline{R} = \underline{I}\sigma_e^2$ , ja toiseksi että s:llä satunnaistekijällä (eläinjalostuksessa yleensä isällä) on sama varianssi eivätkä ne ole korreloituneita keskenään (eivät ole sukulaisia), siis  $\underline{G} = \underline{I}\sigma_a^2$

Kun  $\underline{R}^{-1}$  on halkaisijamatriisi ja halkaisijan elementit ovat samoja eli  $\underline{R}^{-1} = (\underline{I}/\sigma_e^2)$  voidaan yhtälöryhmä sieventää kertomalla  $\sigma_e^2$ :lla. Tällöin yhtälöryhmä saadaan muotoon:

$$\begin{bmatrix} \underline{X}'\underline{X} & \underline{X}'\underline{Z} \\ \underline{Z}'\underline{X} & \underline{Z}'\underline{Z} + \underline{I}\frac{\sigma_a^2}{\sigma_e^2} \end{bmatrix} \begin{bmatrix} \underline{b} \\ \underline{a} \end{bmatrix} = \begin{bmatrix} \underline{X}'\underline{y} \\ \underline{Z}'\underline{y} \end{bmatrix}$$

Näiden varianssista ja kovarianssista tehtyjen oletusten jälkeen voidaan yhtälöryhmä ratkaista kuten normaalit LS-yhtälöt kuitenkin sillä poikkeuksella, että satunnaisten tekijöiden kerroinmatriisin diagonaalielementteihin lisätään varianssisuhde  $(\sigma_e^2/\sigma_a^2)$  (HENDERSON 1973).

Täten MME-yhtälöiden ratkaisuna saatavat kiinteiden tekijöiden BLUE-estimaatit ovat itse asiassa LS-estimaatteja ja satunnaisten



tekijöiden BLUP-estimaatit olisivat valintaindeksin estimaatteja, mikäli valintaindeksissä käytettäisiin  $\mu$  :n likiarvona LS-estimaattia. Yhtäläisyys satunnaisten tekijöiden BLUP-ennusteiden ja valintaindeksin ennusteiden kanssa voidaan todeta kirjoittamalla satunnaisesti tekijäksi  $a_i$  esimerkiksi isän vaikutus, jolloin minkä tahansa isän  $i$  jälkeläisten keskiarvo on  $\bar{y}_i = y_i/n_i$  ja

$$n_i \hat{\mu} + (n_i + \sigma_e^2/\sigma_a^2) \hat{a}_i = y_i.$$

$$(n_i + \sigma_e^2/\sigma_a^2) \hat{a}_i = n_i(\bar{y}_i - \hat{\mu}) \quad \text{joten}$$

$$\hat{a}_i = \frac{n_i}{n_i + \sigma_e^2/\sigma_a^2} (\bar{y}_i - \hat{\mu})$$

mikä on identtinen valintaindeksin kanssa muuten, paitsi että  $\mu$  :n sijasta jälkeläisten keskiarvosta vähennetään BLUE-arvo  $\hat{\mu}$ . (VAN VLECK 1979).

Varianssisuhde ( $\sigma_e^2/\sigma_a^2$ ) pitäisi teoriassa olla tiedossa, mutta käytännössä on toisin ja sen sijasta joudutaan käyttämään kyseisten varianssien estimaatteja. MME-yhtälöistä voidaan ratkaista uudet varianssikomponentit, ja jos ne poikkeavat huomattavasti oletetuista, sijoittaa MME-yhtälöihin ja ratkaista yhtälöt uudelleen. Tämä prosessi voidaan toistaa eli iteroida niin kauan, että tuloksissa ei enää ole merkittävää eroa. (MAO 1982)

Mallissa näkyviin kirjoitetut odotusarvot ja varianssi-kovarianssien oletukset aiheuttavat mallille lisäksi ns. piilo-oletuksia, jotka täytyy tietää mallia käytettäessä. Tällaisia yksinkertaiseen MME-malliin liittyviä oletuksia ovat (MAO 1982):

- Muut kuin malliin sisällytetyt ympäristötekijät eivät vaikuta ominaisuuteen.
- Vain additiiviset geneettiset vaikutukset huomioidaan.
- Eläimet eivät ole sukusiitettyjä ts. niiden vanhemmat ole sukua keskenään.
- Ei tapahdu valintaa.
- Eläimeltä ei ole kuin yksi tuotostieto (havainto).

Näistä oletuksista täytyy olla tietoinen kerätessä havaintoaineistoa ja tutkittaessa analyysituloksia. Jos jotkut oletukset ovat ilmiselvästi väärinä, täytyy mallia muuttaa niiden osalta. Tämä voi olla laskuteknisesti hankalaa, koska oletukset on tehty yksinkertaistamaan mallia ja siten helpottamaan laskutoimituksia. Oletuksia, jotka on tehty estimoimaan puuttuvaa tietoa, ei voida mallissa muuttaa. Karjansisäisessä lehmien arvostelussa joudutaan mallissa ottamaan huomioon valinta, vanhempien sukulaisuussuhteet ja useamman samalta eläimeltä saadun havainnon olemassaolo. Näiden oletusten salliminen edellyttää monimutkaisempaa menetelmää.

Lehmäindeksi laskettuna BLUP-menetelmällä

HENDERSON (1975a) kehitti BLUP-menetelmään perustuvan lehmien arvostelumenetelmän. Menetelmän etuja ovat:

- Arvostelu voidaan sitoa tiettyyn perustasoon ja poistaa siten geneettisen edistymisen aiheuttama harha
- Arvostelussa otetaan huomioon kaikki karjan sisäiset sukulaisuussuhteet.

- Karjan lehmien isät otetaan huomioon arvosteltaessa karjan geneettistä tasoa.
- Valinnan aiheuttama harha poistuu.

### Malli

Yleinen malli tuotoksille karjan sisällä on:

$$y_{ijk} = \mu + b_i + a_j + p_j + e_{ijk}$$

missä:

$y_{ijk}$  = j:nnen lehmän vuonna i tekemä k:s tuotos

$\mu$  = vakio (keskiarvo)

$b_i$  = kiinteä vuositekijä

$a_j$  = j:nnen lehmän additiivinen geneettinen arvo

$p_j$  = pysyvien ympäristötekijöiden ja ei-additiivisten geneettisten tekijöiden vaikutus j:nnen lehmän kaikille tuotoksille

$e_{ijk}$  = satunnainen ympäristötekijä (virhe)  
vuonna i j:nnen lehmän k:nnele tuotokselle

$E((a_j + p_j)^2) = r\sigma_y^2$ ,  $E(a_j^2) = h^2\sigma_y^2$ ,  $E(p_j^2) = (r - h^2)\sigma_y^2$  ja  $E(e_{ijk}^2) = (1 - r)\sigma_y^2$ . Asetetaan: Vuosia 1:stä H:hon, lehmiä 1:stä C:hen ja j:nnen lehmän tuotoksia 1:stä m:mään. n saa arvon nolla tai yksi riippuen siitä onko lehmällä tuotosta vuonna i vai ei.

Normaaliyhtälöt voidaan kirjoittaa VAN VLECK'n (1979) mukaan sum-  
mamuotoa käyttäen seuraavasti:

yhtälö	$\hat{\mu}$	$\hat{b}_1 \dots \hat{b}_H$	$\hat{a}_1 \dots \hat{a}_C$	$\hat{p}_1 \dots \hat{p}_C$	= summa
$\hat{\mu}$	: $n_{..}$	$n_{1.} \dots n_{H.}$	$n_{.1} \dots n_{.C}$	$n_{.1} \dots n_{.C}$	= $y_{...}$
$\hat{b}_1$	: $n_{1.}$	$n_{1.} \begin{matrix} \leftarrow 0 \\ \downarrow \\ [H*H] \end{matrix}$	$n_{11} \dots n_{1C}$	$n_{11} \dots n_{1C}$	= $y_{1..}$
.	.	.	$[H*A]$	$[H*P]$	.
$\hat{b}_H$	: $n_{H.}$	$\begin{matrix} \uparrow 0 \\ \rightarrow \end{matrix} n_{H.}$	$n_{H1} \dots n_{HC}$	$n_{H1} \dots n_{HC}$	= $y_{H..}$
$\hat{a}_1$	: $n_{.1}$	$n_{11} \dots n_{H1}$	$n_{.1} \begin{matrix} \leftarrow 0 \\ \downarrow \\ [A*A] \end{matrix}$	$n_{.1} \begin{matrix} \leftarrow 0 \\ \downarrow \\ [P*A] \end{matrix}$	= $y_{.1.}$
.	.	$[A*H]$	.	.	.
$\hat{a}_C$	: $n_{.C}$	$n_{1C} \dots n_{HC}$	$\begin{matrix} \uparrow 0 \\ \rightarrow \end{matrix} n_{.C}$	$\begin{matrix} \uparrow 0 \\ \rightarrow \end{matrix} n_{.C}$	= $y_{.C.}$
$\hat{p}_1$	: $n_{.1}$	$n_{11} \dots n_{H1}$	$n_{.1} \begin{matrix} \leftarrow 0 \\ \downarrow \\ [P*A] \end{matrix}$	$n_{.1} \begin{matrix} \leftarrow 0 \\ \downarrow \\ [P*P] \end{matrix}$	= $y_{.1.}$
.	.	$[P*H]$	.	.	.
$\hat{p}_C$	: $n_{.C}$	$n_{1C} \dots n_{HC}$	$\begin{matrix} \uparrow 0 \\ \rightarrow \end{matrix} n_{.C}$	$\begin{matrix} \uparrow 0 \\ \rightarrow \end{matrix} n_{.C}$	= $y_{.C.}$

Pienimmän neliösumman yhtälöryhmissä ei voida erottaa a:ta ja p:tä toisistaan, koska ne molemmat ovat satunnaistekijöitä ja niitä kuvataan siten samanlaisilla yhtälöillä. Näille tekijöille saadaan kuitenkin ratkaistua erilaiset ennusteet tekemällä yhtälöistä sekamallin yhtälöt.

Tällöin lisätään p:n yhtälön halkaisijaan ( $n_{.j}$ )

$$\sigma_e^2 / \sigma_p^2 = (1 - r) \sigma_y^2 / (r - h^2) \sigma_y^2 = (1 - r) / (r - h^2) = t$$

Mikäli lehmien välillä ei ole sukulaisuussuhteita lisätään a:n yhtälöiden halkaisijaan ( $n_{.j}$ )

$$\sigma_e^2/\sigma_a^2 = (1 - r)\sigma_y^2/h^2\sigma_y^2 = (1 - r)/h^2 = k$$

Jos lehmien väliset sukulaisuussuhteet otetaan huomioon niin sukulaisuusmatriisin käänteismatriisin elementit kerrottuna  $\sigma_e^2/\sigma_a^2$  lisätään a:n yhtälöryhmän ( $A^*A$ ) elementteihin, mutta ei p:n yhtälöihin. Mikäli malliin otetaan mukaan lehmien lisäksi myös muita eläimiä kuten isät tai jälkeläiset, joilla ei vielä ole tuotostietoja, niin nämä kaikki eläimet on huomioitava a-vektorissa sekä a:n kerroinmatriisissa. Näille eläimille lasketaan jalostusarvo sukulaistietojen perusteella.

Matriisimuodossa kirjoitetaan yleinen malli tuotoksille karjan sisällä HENDERSONin (1975a) mukaan seuraavasti:

$$\underline{y} = \underline{X}\underline{b} + \underline{Z}\underline{a} + \underline{Z}\underline{p} + \underline{e}$$

missä:

$\underline{y}$  = tuotosvektori

$\underline{b}$  = tuntematon kiinteiden tekijöiden vektori esim.  
vuosi

$\underline{X}$  = tunnettu  $\underline{b}$ :n kerroinmatriisi, joka sisältää  
nollia ja ykkösiä

$\underline{Z}$  = tunnettu  $\underline{a}$ :n ja  $\underline{p}$ :n kerroinmatriisi

$\underline{a}$  = additiivisten geneettisten arvojen vektori kaikille  
mallissa mukana oleville eläimille

$\underline{p}$  = ei-additiivisten geneettisten arvojen ja pysyvien  
ympäristötekijöiden vaikutusten vektori ko.lehmille

$\underline{e}$  = satunnaisten ympäristötekijöiden vektori

Oletukset:

Satunnaistekijöiden  $\underline{a}$ ,  $\underline{p}$ ,  $\underline{e}$  keskiarvo = 0

$\text{Var}(\underline{a}) = \underline{A}h^2\sigma_y^2$ , missä  $\underline{A}$  = karjansisäinen sukulaisuusmatriisi.

$\text{Var}(\underline{p}) = \underline{I}(r - h^2)\sigma_y^2$

$\text{Var}(\underline{e}) = \underline{I}(1 - r)\sigma_y^2$

Mallissa oletetaan, että tuotosten välisiä korrelaatioita eri eläinten välillä aiheutuu vain additiivisesta geneettisestä variaanssista, ts.  $\underline{a}$ ,  $\underline{p}$  ja  $\underline{e}$  tekijöiden välillä ei ole korrelaatiota.

Nyt voidaan MME-yhtälöt kirjoittaa muotoon (HENDERSON 1975a):

$$\begin{bmatrix} \underline{X}'\underline{X} & \underline{X}'\underline{Z} & & \underline{X}'\underline{Z} \\ \underline{Z}'\underline{X} & \underline{Z}'\underline{Z} + t\underline{A}^{-1} & & \underline{Z}'\underline{Z} \\ \underline{Z}'\underline{X} & \underline{Z}'\underline{Z} & & \underline{Z}'\underline{Z} + k\underline{I} \end{bmatrix} \begin{bmatrix} \underline{b} \\ \underline{a} \\ \underline{p} \end{bmatrix} = \begin{bmatrix} \underline{X}'\underline{Y} \\ \underline{Z}'\underline{Y} \\ \underline{Z}'\underline{Y} \end{bmatrix}$$

Ratkaisun tuloksena saadaan parhaat lineaariset harhattomat estimaatit ( $\hat{\underline{b}}$ ) karjan ympäristölle ja parhaat lineaariset harhattomat ennusteet ( $\hat{\underline{a}}$ ) geneettiselle arvolle ja ( $\hat{\underline{p}}$ ) ei-additiivisille ja pysyville ympäristövaikutuksille. Yhtälöstä saadaan laskettua lehmille myös todennäköinen tuotantokyky laskemalla yhteen  $\hat{\underline{a}}$  ja  $\hat{\underline{p}}$ .

### Eläinten arvostelu pelkästään sukulaistietojen perusteella

Eläimet, joilla ei ole tulosta, kuten vanhemmat tai jälkeläiset, voidaan arvostella sukulaisten tuotostietojen perusteella. Tällaiset eläimet otetaan mukaan sukulaisuusmatriisiin ja jokaista

lisättyä eläintä kohti kirjoitetaan uusi yhtälö, jossa oikealla puolella oleva summa on nolla. MME-yhtälöiden kirjoittamisen helpottamiseksi voidaan sukulaismatriisi kirjoittaa HENDERSONIN (1975b) menetelmää sovellettaessa siten, että eläimet, joilla on tuotostietoja kirjoitetaan ennen eläimiä, joilla ei ole tuotostietoja. Tällöin matriisiyhtälö saa muodon:

$$\begin{bmatrix} \underline{X}'\underline{X} & \underline{X}'\underline{Z} & \underline{0} & \underline{X}'\underline{Z} \\ \underline{Z}'\underline{X} & \underline{Z}'\underline{Z} + t\underline{A}^{11} & t\underline{A}^{12} & \underline{Z}'\underline{Z} \\ \underline{0} & t\underline{A}^{21} & t\underline{A}^{22} & \underline{0} \\ \underline{Z}'\underline{X} & \underline{Z}'\underline{Z} & \underline{0} & \underline{Z}'\underline{Z} + k\underline{I} \end{bmatrix} \begin{bmatrix} \underline{b} \\ \underline{a}_1 \\ \underline{a}_2 \\ \underline{p} \end{bmatrix} = \begin{bmatrix} \underline{X}'\underline{y} \\ \underline{Z}'\underline{y} \\ \underline{0} \\ \underline{Z}'\underline{y} \end{bmatrix}$$

Sukulaismatriisin  $\underline{A}$  käänteismatriisi  $\underline{A}^{-1}$  on jaettu neljään osaan:

$$\left[ \begin{array}{c|c} \underline{A}^{11} & \underline{A}^{12} \\ \hline \underline{A}^{21} & \underline{A}^{22} \end{array} \right], \text{ jossa } \underline{A}^{11} \text{ muodostuu eläimistä, joilla}$$

on tuotos ja  $\underline{A}^{22}$  eläimistä, joilla ei

ole tuotostietoja.  $\underline{A}^{12}$  ja  $\underline{A}^{21}$  sisältää

näiden kahden eri ryhmän eläinten vä-

liset sukulaisuudesta johtuvat yhteydet

$\underline{a}_1$  = additiivinen geneettinen arvo eläimille,

joilla on tuotostietoja

$\underline{a}_2$  = additiivinen geneettinen arvo eläimille,

joilla ei ole tuotostietoja.

Systemaattisten ympäristötekijöiden ja ei-additiivisten tekijöiden absorbointi muihin yhtälöihin

Vektori  $\underline{p}$  sisältää ei-additiivisten geneettisten tekijöiden ja systemaattisten ympäristövaikutusten osuuden tulosten vaihtelusta. Lehmien järjestys  $\underline{p}$ :n sisällä on sama kuin  $\underline{a}$ :ssa. Koska  $\underline{Z}'\underline{Z} + \underline{I}k$  on diagonaalimatriisi voidaan matriisiyhtälöiden luku-

määrän pienentämiseksi  $\underline{p}$ :n yhtälöt helposti absorboida  $\underline{b}$ :n ja  $\underline{a}$ :n yhtälöihin.

Olkoon matriisiyhtälöt muotoa

$$\begin{bmatrix} \underline{A} & \underline{B} \\ \underline{B}' & \underline{C} \end{bmatrix} \begin{bmatrix} \underline{b}_1 \\ \underline{b}_2 \end{bmatrix} = \begin{bmatrix} \underline{r}_1 \\ \underline{r}_2 \end{bmatrix}$$

Absorboitaessa jokin yhtälö yhtälöryhmän muihin yhtälöihin, ratkaistaan ensin absorboitava yhtälö sitä vastaavan tuntemattoman suhteen ja sijoitetaan saatu ratkaisu muihin yhtälöihin. Esimerkin tapauksessa ratkaistaan  $\underline{b}_2$  ja sijoitetaan saatu ratkaisu  $\underline{b}_1$ :n yhtälöön. Matriisiyhtälön toiselta riviltä saadaan

$$\underline{B}'\underline{b}_1 + \underline{C}\underline{b}_2 = \underline{r}_2$$

joten

$$\underline{C}\underline{b}_2 = \underline{r}_2 - \underline{B}'\underline{b}_1 \text{ eli}$$

$$\underline{b}_2 = \underline{C}^{-1}\underline{r}_2 - \underline{C}^{-1}\underline{B}'\underline{b}_1, \text{ olettaen että } \underline{C}^{-1} \text{ on olemassa.}$$

Sijoitetaan saatu  $\underline{b}_2$  vastaavaan  $\underline{b}_1$ :n yhtälöön

$$\underline{A}\underline{b}_1 + \underline{B}\underline{b}_2 = \underline{r}_1 \text{ tai}$$

$$\underline{A}\underline{b}_1 + \underline{B}\underline{C}^{-1}\underline{r}_2 - \underline{B}\underline{C}^{-1}\underline{B}'\underline{b}_1 = \underline{r}_1$$

mistä saadaan absorboidut yhtälöt

$$(\underline{A} - \underline{B}\underline{C}^{-1}\underline{B}')\underline{b}_1 = \underline{r}_1 - \underline{B}\underline{C}^{-1}\underline{r}_2$$

Kun  $\underline{C}$ -matriisi on diagonaalinen saadaan

$$\underline{B}\underline{C}^{-1}\underline{B}' = (1/c_{ii})\underline{B}\underline{B}'$$

missä  $c_{ii}$  on  $\underline{C}$ -matriisin diagonaali-elementti

sekä

$$\underline{B}\underline{C}^{-1}\underline{r}_2 = (1/c_{ii})\underline{B}\underline{r}_2$$



MME-yhtälöiden matriisi ositetaan ja suoritetaan absorbointi osamatriiseilla kuten edellisessä esimerkissä.

$$\begin{bmatrix} \underline{X}'\underline{X} & \underline{X}'\underline{Z} & \underline{0} & \underline{X}'\underline{Z} \\ \underline{Z}'\underline{X} & \underline{Z}'\underline{Z} + \underline{A}^{11}t & \underline{A}^{12}t & \underline{Z}'\underline{Z} \\ \underline{0} & \underline{A}^{21}t & \underline{A}^{22}t & \underline{0} \\ \underline{Z}'\underline{X} & \underline{Z}'\underline{Z} & \underline{0} & \underline{Z}'\underline{Z} + \underline{I}k \end{bmatrix} \begin{bmatrix} \underline{b} \\ \underline{a}_1 \\ \underline{a}_2 \\ \underline{p} \end{bmatrix} = \begin{bmatrix} \underline{X}'\underline{Y} \\ \underline{Z}'\underline{Y} \\ \underline{0} \\ \underline{Z}'\underline{Y} \end{bmatrix}$$

Absorboinnin jälkeen matriisiyhtälöt kirjoitetaan muotoon (esimerkiksi BOLGIANO 1981)

$$\begin{bmatrix} \underline{X}'\underline{M}\underline{X} & \underline{X}'\underline{M}\underline{Z} & \underline{0} \\ \underline{Z}'\underline{M}\underline{X} & \underline{Z}'\underline{M}\underline{Z} + \underline{A}^{11}t & \underline{A}^{12}t \\ \underline{0} & \underline{A}^{21}t & \underline{A}^{22}t \end{bmatrix} \begin{bmatrix} \underline{b} \\ \underline{a}_1 \\ \underline{a}_2 \end{bmatrix} = \begin{bmatrix} \underline{X}'\underline{M}\underline{Y} \\ \underline{Z}'\underline{M}\underline{Y} \\ \underline{0} \end{bmatrix}$$

missä  $\underline{M} = \underline{I} - \underline{Z}(\underline{Z}'\underline{Z} + \underline{I}k)^{-1}\underline{Z}'$ .

$\underline{p}$ :n ja  $\underline{a}$ :n yhtälöt, joihin  $\underline{p}$  on absorboitu voidaan muodostaa suoraan tuloslistasta, jossa tulokset on lajiteltu poikimavuoden mukaan. Yhtälöt kirjoitetaan mallin mukaan, jossa malli yhdelle tuotokselle on

$$y_{ijk} = b_i + a_j + p_j + e_{ijk}$$

siten että  $y_{ijk}$  on j:nnen lehmän vuonna i lypsämä maitotuotos k.

Tällöin matriisiyhtälöiden elementit ovat (HENDERSON 1975a):  
(absorboinnin tuloksena saatuja elementtejä merkitään tässä yläindeksi \*:llä)

1)  $\underline{X}'\underline{X}$ -matriisin diagonaalielementit

$$X'X(i,i)^* = \frac{X'Z(i,j) \times (Z'Z(j,j) + k - 1)}{Z'Z(j,j) + k}$$

missä summaus on lehmien yli vuonna i,

$X'Z(i,j) = 1$  tai 0 sen mukaan onko lehmällä j tulos vuonna i vai ei.

2)  $\underline{X}'\underline{X}$ -matriisin ei-diagonaalielementit

$$X'X(i,j)^* = - \frac{X'Z(i,j) \times X'Z(L,j)}{Z'Z(j,j) + k}$$

missä summaus on lehmien yli, joilla on tulos vuosiparilla i,L.

3)  $\underline{X}'\underline{Z}$ -matriisin ei-nollaelementit

$$X'Z(i,j)^* = X'Z(i,j) \times \frac{k}{Z'Z(j,j) + k}$$

saavat nollasta poikkeavan arvon vain mikäli lehmällä j on tulos vuonna i.

4)  $\underline{Z}'\underline{Z}$ -matriisin diagonaalielementit

$$Z'Z(j,j)^* = Z'Z(j,j) \times \frac{k}{Z'Z(j,j) + k}$$

5)  $\underline{Z}'\underline{Z}$ -matriisin ei-diagonaalielementit pysyvät muuttumattomina.

Oikean puolen yhtälöihin saadaan

6)  $\underline{x}'\underline{y}$ -vektori

$$x'y(i)^* = x'y(i) - \frac{x'z(i,j) \times z'y(j)}{z'z(j,j) + k}$$

missä summaus on lehmien yli, joilla on tulos vuonna i.

7)  $\underline{z}'\underline{y}$ -vektori

$$z'y(j)^* = \frac{k z'y(j)}{z'z(j,j) + k}$$

Absorbointi ei vaikuta ratkaisujen lopputulokseen, mutta pienentää oleellisesti käännettävän kerroinmatriisin kokoa. Kun  $\hat{a}$  :t on ratkaistu voidaan  $\hat{p}$ :t ratkaista kaavasta

$$\hat{p} = tA^{-1} \hat{a} / k$$

### Isien jälkeläisarvostelutulosten hyväksikäyttö

HENDERSONIN (1975a) menetelmässä voidaan käyttää isien jälkeläisarvosteluja lehmien arvosteluun. Tällöin oletetaan, että on käytettävissä isien BLUP- tai muulla menetelmällä lasketut jälkeläisarvostelut, jotka ilmoitetaan poikkeamina populaation yhteisestä geneettisestä keskiarvosta. Jälkeläisarvostelussa käytetään tyttären ensikkotuotoksia, jotka oletetaan korjatuiksi kaikkien kiinteiden tekijöiden suhteen. Yhtälöt lehmien geneettisen arvon laskemiseksi tuotosten perusteella karjan sisällä eivät ole suoraan sidottuja vastaaviin toisissa karjoissa olevien lehmien yhtälöihin, koska ainoa yhteys yhtälöillä on sukulaisuusmatriisin

käänteismatriisiin välityksellä ja sen elementit isänpuoleisille puolisisarille saavat arvon nolla. Siten yhteys lehmien välillä on epäsuora isän välityksellä: käänteismatriisin elementit isän ja minkä tahansa tyttären välillä ovat  $-2/3$ .

Isän jälkeläisarvostelun tiedot liitetään lehmäarvosteluun lisäämällä sekamallin yhtälöihin isiä vastaavat yhtälöt siten, että  $\underline{a}$ :n kerroinmatriisiin isän yhtälön halkaisijaan tulee

$$q(1 - r)/(4 - h^2),$$

johon lisätään ko. diagonaalielementti  $\underline{A}^{-1}$ :stä kerrottuna  $t$ :llä.  $q$  on tyttäreiden lukumäärä,  $t = (1 - r)/h^2$ .

Vastaavasti sonni saa yhtälöryhmän oikeallepuolelle:

$$\frac{2(1 - r)(4 + (q - 1)h^2)}{h^2(4 - h^2)} \hat{v}$$

$\hat{v}$  = isän jälkeläisarvostelun perusteella laskettu additiivinen- eli jalostusarvo (BLUP-ennuste)

Kertoimet saadaan Van Vleckin (1982) mukaan seuraavasti:

Olkoon  $\underline{a}_{sp}$  mielivaltaisen lehmän geneettinen arvo karjassa, jossa on yhteensä  $p$  tämän lehmän isän tytärtä ja  $\underline{a}_{sq}$  ( $i = 1, \dots, q$ ) muissa karjoissa olevan  $q$ :n tyttären geneettiset arvot sekä  $\underline{a}_s$  isän geneettinen arvo. MME-mallin yhtälöt kirjoitetaan muotoon (karjan muiden lehmien yhtälöitä sekä kiinteitä yhtälöitä ei tässä kirjoiteta näkyviin):

$$\begin{array}{c}
 \left[ \begin{array}{cc|cc}
 n_{sp} & k/(n+k)+4t/3 & -2t/3 & 0 & \dots & 0 \\
 -2t/3 & & (1+1/3+q/3)t & -2t/3 & \dots & -2t/3 \\
 0 & & -2t/3 & k/(k+1)+4t/3 & 0 & \dots & 0 \\
 \cdot & & \cdot & 0 & \cdot & \cdot & \cdot \\
 \cdot & & \cdot & \cdot & \cdot & \cdot & \cdot \\
 \cdot & & \cdot & \cdot & \cdot & \cdot & \cdot \\
 0 & & -2t/3 & 0 & \dots & k/(k+1)+4t/3 & \dots
 \end{array} \right] \begin{array}{c}
 \left[ \begin{array}{c}
 y_{sp} \\
 y_{sq} \\
 y_{s1} \\
 \cdot \\
 \cdot \\
 \cdot \\
 y_{sq}
 \end{array} \right]
 \end{array}
 \end{array}$$

Yhtälöryhmän oikea puoli kirjoitetaan muotoon

$$\begin{array}{c}
 \left[ \begin{array}{c}
 ky_{sp} / (n+k) \\
 0 \\
 ky_{s1} / (k+1) \\
 \cdot \\
 \cdot \\
 \cdot \\
 ky_{sq} / (k+1)
 \end{array} \right]
 \end{array}$$

t:n kertoimet ovat sukulaisuusmatriisiin käänteismatriisin elementtejä. Isän yhtälön diagonaalinen käännöselementti on jaettu kolmeen osaan: 1 isälle, 1/3 tyttäreille sp ja q/3 muille tyttäreille q. t ja k ovat kuten edellä. Termit, joissa esiintyy k sisältävät absorboidut ei-additiiviset geneettiset ja pysyvien ympäristötekijöiden vaikutukset, missä tyttäreillä sp on n tuotosta joiden summa on y<sub>sp</sub> ja muilla tyttärillä jokaisella yksi tuotosta merkitään y<sub>si</sub>. Pysyvien ympäristötekijöiden absorboinnista johtuen on joihinkin matriisiyhtälöiden termeihin lisätty k/(n+k) tai k/(k+1).

Seuraavaksi  $q$ :n tyttärien yhtälöt absorboidaan isän yhtälöön. Valitaan edellisen esimerkin mukaisesti osamatriisit  $\underline{A}$ ,  $\underline{B}$ ,  $\underline{B}'$  ja  $\underline{C}$  siten, että

$$\underline{A} = (1 + 1/3 + q/3)t$$

$$\underline{B} = -2t/3 \dots -2t/3$$

$$\underline{C}_{ii} = k/(k+1) + 4t/3$$

sekä

$$\underline{b}_1 = \underline{a}_{s1}$$

$$\underline{b}_2 = \begin{bmatrix} \underline{a}_{s1} \\ \cdot \\ \cdot \\ \cdot \\ \underline{a}_{sq} \end{bmatrix}$$

$$\underline{r}_1 = 0$$

$$\underline{r}_2 = \begin{bmatrix} ky_{s1} / (k+1) \\ \cdot \\ \cdot \\ \cdot \\ ky_{sq} / (k+1) \end{bmatrix}$$

Tällöin

$$\underline{A} - (1/\underline{C}_{ii})\underline{B}\underline{B}'$$

$$= (1 + 1/3 + q/3)t - q \frac{(-2t/3)(-2t/3)}{k/(k+1) + 4t/3}$$

$$= (1 + 1/3)t + \frac{qt}{3} - q \frac{4t \cdot 3(k+1)}{9(3k + 4t(k+1))}$$

$$= (1 + 1/3)t + q \frac{(3k + 4t(k+1))t - (4t^2(k+1))}{3(3k + 4t(k+1))}$$

$$= (1 + 1/3)t + q \frac{3kt}{9k + 12t(k+1)}$$

kun sijoitetaan k:n ja t:n paikalle niiden lausekkeet ja suoritetaan supistukset päädytään kaavaan

$$= (1 + 1/3)t + q \frac{(1 - r)}{(4 - h^2)}$$

Oikean puolen yhtälöihin saadaan

$$\underline{r}_1 - (1/c_{ii})\underline{B} \underline{r}_2$$

$$= 0 - \frac{-2t/3 k(k+1)}{k/(k+1) + 4t/3} y_{si}$$

$$= \frac{2tk}{3k + 4t(k+1)} y_{si}$$

Sijoitetaan k:n ja t:n lausekkeet ja suoritetaan supistukset, jolloin saadaan

$$= \frac{2(1 - r)}{(4 - h^2)} y_{si}$$

Absorbointi ei vaikuta muihin yhtälöihin. Absorboinnin jälkeen saadaan:

$$\begin{bmatrix} n_{sp}k/(n_{sp}+k)+4t/3 & -2t/3 \\ -2t/3 & q(1-r)/(4-h^2)+(1+1/3)t \end{bmatrix} \begin{bmatrix} \hat{a}_{sp} \\ \hat{a}_s \end{bmatrix} = \begin{bmatrix} k_{sp}/(n+k) \\ 2(1-r)y_s/(4-h^2) \end{bmatrix}$$

missä  $\underline{y}_s$  on sonniarvostelun tuotosten summa  $q$ :lle tyttärelle. Isän yhtälön halkaisijasta  $qt/3$  sisältyy lauseeseen  $q(1-r)/(4-h^2)$ . Koska  $\underline{y}_s$  ja  $q$  ei ole tiedossa täytyy niiden tilalla käyttää estimaatteja, jotka saadaan isän jälkeläisarvostelusta:  $q$ :lle voidaan käyttää tyttärien lukumäärää. Jos isän arvostelutulos asetetaan yhtäsuureksi kuin ko. valintaindeksin estimaatin  $1/2a_s$  algebrallinen yhtälö, löydetään aproksimaatio  $\underline{y}_s$ :lle:

$$\hat{V} = \underline{y}_s / (q + (4 - h^2)/h^2) ,$$

jolloin  $(q + (4 - h^2)/h^2)\hat{V}$  sijoitetaan  $\underline{y}_s$ :n paikalle.

#### Sukulaismatriisin käänteismatriisin laskeminen

MME-mallissa tarvitaan sukulaismatriisin käänteismatriisia  $\underline{A}^{-1}$ , kun geneettisten arvojen ennustamisessa (BLUP) halutaan ottaa huomioon myös geneettisten vaikutusten kovarianssit. Tavallinen menetelmä olisi laskea sukulaisuusmatriisi  $\underline{A}$  ja kääntää se tietokoneella. Tällöin tulee kuitenkin rajoittavaksi tekijäksi aineiston koko, koneen muistikapasiteetti ja laskutoimituksen hitaus. HENDERSON (1976) kehitti käänteismatriisin laskemiseksi menetelmän, jossa varsinaista sukulaismatriisia ei lasketa ollenkaan. Menetelmässä oletetaan, että eläimet eivät ole sukusiitettyjä, jota ei karjan sisällä keinosiemennyksen yleisyyden takia voida pitää kohtuuttomana. QUAS (1976) kehitti sukulaismatriisin käänteismatriisin laskemiseksi yleisen menetelmän, joka ottaa huomioon myös sukusiitoksen ts. menetelmässä lasketaan eläimelle myös sukusiitoskerroin  $F$ .



### Sukulaismatriisin laskeminen

Johdonmukaisuuden takia tarkastellaan myös sukulaisuusmatriisin laskemista, vaikka sitä ei varsinaisesti käänteismatriisia laskettaessa tarvitakaan. Karjan eläimet merkitään 1-n, joista eläimet 1-b kuuluvat peruspopulaatioon, jonka eläimet eivät ole sukusiitettyjä eivätkä sukua keskenään. Peruspopulaatiota vastaava osamatriisi =  $\underline{I}$  eli vanhempien sukulaisuuskerroin = 1. Yksilön sukusiitosaste lasketaan puolittamalla vanhempien välinen sukulaisuuskerroin. Tällöin, jos yksilö on sukusiitetty, tulee yksilön sukulaisuuskertoimeksi 1 + sukusiitosaste(F) (WRIGHT 1922).

Sukulaisuusmatriisi on siis n x n matriisi, josta peruspopulaation osuus on b x b. Sukulaisuusmatriisin muodostaminen aloitetaan HENDERSONin (1976) mukaan peruspopulaatiosta, jolle merkitään sukulaisuuskertoimet. Tämän jälkeen edetään eläin kerrallaan rivittäin siten, että eläin saa sukulaisuuskertoimekseen seuraavat arvot:

- jos eläimen t molemmat vanhemmat p ja q tunnetaan

$$a_{ti} = a_{it} = .5(a_{ip} + a_{iq}), \text{ jokaiselle } i = 1, \dots, t-1;$$
$$a_{tt} = 1 + .5a_{pq}.$$

- jos eläimeltä t tunnetaan vain toinen vanhemmista, esim p

$$a_{ti} = a_{it} = .5a_{ip}, \text{ jokaiselle } i = 1, \dots, t-1;$$
$$a_{tt} = 1$$

- jos kumpaakaan vanhemmista ei tunneta

$$a_{ti} = a_{it} = 0, \text{ jokaiselle } i = 1, \dots, t-1;$$

$$a_{tt} = 1$$

Kolmiomatriisin laskeminen

Olkoon  $\underline{L}$  sellainen alakolmiomatriisi, että  $\underline{L}\underline{L}' = \underline{A}$ . Eläimet, joi-  
ta on  $n$  kappaletta, järjestetään kuten edellä. Peruspopulaation  
ulkopuolella olevat  $\underline{L}$ -matriisin elementit lasketaan rivittäin  
seuraavasti (HENDERSON 1976):

- jos eläimen  $t$  molemmat vanhemmat  $p$  ja  $q$  tunnetaan  
ja  $p < q$  niin

$$l_{tj} = .5(l_{pj} + l_{qj}), \text{ jokaiselle } j = 1, \dots, p$$

$$= .5l_{qj}, \text{ jokaiselle } j = p+1, \dots, q$$

$$= 0, \text{ jokaiselle } j = q+1, \dots, t-1$$

sekä erikoisesti, jos  $q < t - 1$

$$l_{tt} = \sqrt{1 + .5 l_{pj} l_{qj} - l_{tj}^2}$$

On huomattava, että

$$.5 \sum_{j=1}^p l_{pj} l_{qj} = F_t \quad \text{ja}$$

$$l_{tt} = \sqrt{.5 - .25(F_p + F_q)}$$

- jos eläimeltä  $t$  tunnetaan vain toinen vanhemmista, esim.  $p$

$$l_{tj} = .5l_{pj}, \text{ jokaiselle } j = 1, \dots, p$$

$$= 0, \text{ jokaiselle } j = p+1, \dots, t-1$$

sekä erikoisesti, jos  $p < t - 1$

$$l_{tt} = \sqrt{1 - \sum_{j=1}^p l_{tj}^2} = \sqrt{.75 - .25F_p}$$

- jos kumpaakaan vanhemmista ei tunneta

$$l_{tj} = 0, \text{ jokaiselle } j = 1, \dots, p$$

$$l_{tt} = 1$$

Käänteismatriisin laskemisen yleinen muoto

QUAAS (1976) kehitti käänteismatriisin  $\underline{A}^{-1}$  nopean ratkaisumenetelmän yleisen muodon, jossa sukulaismatriisin  $\underline{A}$  halkaisijaelementit lasketaan ottaen huomioon sukusiitosaste  $F$ . Tällöin ko. elementit saavat arvon  $1 + F$ . Näistä saadaan koko sukulaisuusmatriisin käänteismatriisi  $\underline{A}$  helposti laskettua.

Mikä tahansa  $\underline{L}$ -matriisin elementti esim.  $t$ :s elementti saadaan laskettua  $l_{tt}$ :stä ja listasta, jossa on tiedot isistä ja emistä.

$$l_{tt} = \begin{cases} \sqrt{.5 - .25(F_p + F_q)}, & \text{kun molemmat vanhemmat } p \text{ ja } q \\ & \text{tunnetaan} \\ \sqrt{.75 - .25F_p}, & \text{kun vain toinen vanhemmista tun-} \\ & \text{netaan esim. } p \\ 1, & \text{kun kumpaakaan vanhemmista ei tunneta} \end{cases}$$

$F_p$  ja  $F_q$  ovat vanhempien  $p$  ja  $q$  sukusiitosasteet.

Yhtälöistä:

$$a_{jj} = 1 + F_{j^2}$$

$$\underline{L}\underline{L}' = \underline{A} \quad \text{saadaan}$$

$$F_j = a_{jj} - 1 = \left( \sum_{k \neq j}^i l_{jk}^2 \right) - 1.$$

Sijoittamalla tämä F:n arvo edellisiin  $l_{tt}$ :n yhtälöihin saadaan:

$$l_{tt} = \begin{cases} \sqrt{1 - .25 \sum_{k=1}^p l_{pk}^2 - .25 \sum_{k=1}^q l_{qk}^2}, & \text{kun tunnetaan } p \text{ ja } q \\ \sqrt{1 - .25 \sum_{k=1}^p l_{pk}^2}, & \text{kun vain toinen esim. } p \text{ tunnetaan} \\ 1, & \text{kun kumpaakaan ei tunneta} \end{cases}$$

Siten  $\underline{L}$ :n halkaisijaelementti jollakin rivillä on edellisten rivien elementtien neliöiden summan funktio ja nämä neliöiden summat ovat  $\underline{A}$ :n halkaisijaelementtejä. Laskettaessa  $\underline{L}$ :n ja  $\underline{A}$ :n halkaisijaelementtejä  $n \times n$  kokoisesta  $\underline{A}$ -matriisista tarvitaan  $n$  laskukierrosta, jossa  $k$ :s kierros laskee  $k$ :nnen sarakkeen  $\underline{L}$ :stä.

Olkoon  $\underline{u}$  vektori mihin lisätään jokaisen  $\underline{L}$ -matriisin rivien elementtien neliöiden summa ja olkoon  $\underline{v}$  vektori johon talletetaan  $\underline{L}$ -matriisin halkaisijaelementit sekä väliaikaisesti myös halkaisijan ulkopuoliset elementit.  $\underline{u}$  ja  $\underline{v}$  ovat molemmat  $n \times 1$  vektoreita. Eläimet oletetaan järjestetyiksi iän mukaan ja numeroidun  $1, \dots, n$ . Tuntemattomat vanhemmat merkitään nolllalla.

Seuraavat laskutoimitukset suoritetaan  $k$ :nnella kierroksella (QUAAS 1976):

1) Lasketaan

$$a) \quad v_k = l_{kk} = \begin{cases} \sqrt{1 - .25(u_p + u_q)}, & \text{jos } 0 < p < q \\ \sqrt{1 - .25 u_q}, & \text{jos } p = 0 < q \\ 1, & \text{jos } p = q = 0 \end{cases}$$

missä  $p$  ja  $q$  ovat  $k$ :nnen eläimen vanhemmat.

$$b) \quad v_i = \begin{cases} l_{ik} & , i = k + 1, \dots, n \\ .5 v_r + .5 v_s & , \text{jos } k < r < s \\ .5 v_s & , \text{jos } r < k < s \\ 0 & , \text{jos } r < s < k \end{cases}$$

missä r ja s ovat i:nnen eläimen vanhemmat.

c) Neliöidään  $\underline{v}$  ja lisätään  $\underline{u}$  :hin kun  
 $i = k, \dots, n.$

2) Jos halutaan  $\underline{A}^{-1}$  niin lasketaan

$$d = v_k^{-2} = l_{kk}^{-2}, \text{ jonka jälkeen}$$

a) jos  $0 < p < q$

$$\begin{array}{ll} \text{lisätään } \underline{A}^{-1} : \text{n elementtiin } (k, k) & d \\ & (p, k) \quad (-1/2)d \\ & (p, p) \quad (1/4)d \\ & (p, q) \quad (1/4)d \\ & (q, q) \quad (1/4)d \end{array}$$

b) jos  $p = 0 < q$

$$\begin{array}{ll} \text{lisätään } \underline{A}^{-1} : \text{n elementtiin } (k, k) & d \\ & (q, k) \quad (-1/2)d \\ & (q, q) \quad (1/4)d \end{array}$$

c) jos  $p = q = 0$

$$\text{lisätään } \underline{A}^{-1} : \text{ elementtiin } (k, k) \quad d$$

missä k:nnen eläimen vanhemmat ovat p ja q. Jos  $\underline{A}^{-1}$  lasketaan pitää tehdä lajittelu sarakkeet rivin sisällä ja laskea yhteensä nelioidyt termit. Nämä summat ovat ei-nollaelementtejä puoliksi talletetussa  $\underline{A}^{-1}$ -matriisissa. n:nnen kierroksen jälkeen  $\underline{L}$ :n hal-

kaisijaelementit talletetaan  $\underline{y}$ -vektoriin ja  $\underline{A}$ :n halkaisijaelementit  $\underline{u}$ -vektoriin.

### Käänteismatriisin laskemisen erikoistapaus

HENDERSONin (1976) menetelmä on erikoistapaus yleisestä mallista, koska siinä oletetaan, että populaatiossa ei ole tapahtunut sukusiitosta. Sukulaismatriisin käänteismatriisi  $\underline{A}^{-1}$  voidaan laskea käyttäen hyväksi kolmiomatriisia  $\underline{L}$ , joka voidaan esittää muodossa

$\underline{L} = \underline{T}\underline{D}$ , missä  $\underline{D}$  on halkaisijamatriisi, jonka halkaisijaelementit ovat samat kuin  $\underline{L}$ :n halkaisijaelementit.  $\underline{T}$ -matriisi lasketaan kuten  $\underline{L}$  paitsi, että  $\underline{T}$ :n halkaisijaelementit ovat kaikki ykkösiä.

$$\underline{A}^{-1} = (\underline{L}\underline{L}')^{-1} = (\underline{T}\underline{D}\underline{D}\underline{T}')^{-1} = (\underline{T}^{-1})'(\underline{D}^{-1})^2\underline{T}^{-1}$$

Siten  $\underline{T}$  on alakolmiomatriisi, jonka halkaisijaelementit ovat kaikki ykkösiä ja ainoat nollasta poikkeavat elementit ovat  $-.5$  jokaisella  $i$ :nnellä rivillä, tunnettujen vanhempien sarakkeella.

Tärkein laskutoimitus  $\underline{A}^{-1}$ :n laskemiseksi on  $\underline{L}$ -matriisin halkaisijaelementtien, siis  $\underline{D}$ :n, laskeminen.  $\underline{A}$ :ta ja  $\underline{T}$ :tä ei tarvita. Ei-sukusiitetyssä populaatiossa  $\underline{D}$ -matriisi voidaan laskea muodostamatta itse matriisia  $\underline{L}$ , koska  $\underline{D}$ -matriisin elementit voivat saada vain kolme mahdollista arvoa:

$\sqrt{.5}$ , jos molemmat vanhemmat tunnetaan

$\sqrt{.75}$ , jos vain toinen vanhemmista tunnetaan

1, jos kumpaakaan vanhemmista ei tunneta.

Täten  $\underline{D}^{-1}$ :n elementit saavat vain arvoja 2, 4/3 ja 1.  $\underline{A}^{-1}$  voidaan laskea tekemällä lista kaikista n:stä eläimestä ja lisäämällä seuraavat luvut  $\underline{A}^{-1}$ -matriisin elementteihin, jotka alunperin on asetettu nolliksi (HENDERSON 1976):

- jos eläimen i molemmat vanhemmat p ja q tunnetaan lisätään

2 elementtiin (i,i)

-1 elementteihin (p,i),(i,p),(q,i),(i,q)

.5 elementteihin (p,p),(p,q),(q,p),(q,q)

- jos vain toinen vanhemmista esim. p tunnetaan lisätään

4/3 elementtiin (i,i)

-2/3 elementteihin (p,i),(i,p)

1/3 elementtiin (p,p)

- jos kumpaakaan vanhemmista ei tunneta lisätään

1 elementtiin (i,i).

Eläimet voidaan esittää missä järjestyksessä tahansa. Myös peruspopulaation eläimet, jotka ovat sukua muille populaation eläimille, mutta eivät keskenään, täytyy ottaa mukaan vaikka niillä kaikilla ei olisi edes tulosta. Sijoittamalla peruspopulaation eläimet sekä muut tuloksettomat eläimet listan loppuun voidaan MME-yhtälöiden määrää pienentää.

## OMAT TUTKIMUKSET

### AINEISTO JA MENETELMÄT

Lähdeaineistona työssä käytettiin karjantarkkailutietoja, jotka oli kerätty vuonna 1982. Karjantarkkailun piiriin kuuluu noin 40 % Suomen lehmistä eli noin 300000 lehmää. Karjantarkkailun lehmärekisteriin sisältyy tuotoshistoria, jossa on kaikkien tarkkailuun kuuluvien lehmien tuotostiedot kaikilta vuosilta. Vuosittain tästä rekisteristä poimitaan lehmäindeksien laskemista varten kaikkien elossa olevien lehmien tiedot. Lehmillä täytyy olla tiedossa syntymäaika sekä vähintään yksi tuotosjakso 305-päivän ajalta. Lehmäindeksi lasketaan CHRISTENSENIN (1981) kehittämällä menetelmällä.

Tästä vuosirekisteristä saatiin BLUP-menetelmällä laskettavaa karjansisäistä lehmäindeksiä varten maatalouskeskuksittain ja karjoittain lajitellut lehmien poikimakuukauden, poikimavälin, poikimäen ja poikimakerran suhteen korjatut tuotostiedot poikimavuosittain, tuotosten lukumäärä, lehmän rotu, emän tiedot, isän ja emänisän jälkeläisarvosteluindeksi ja tytärien lukumäärä sekä vertailua varten CHRISTENSENIN (1981) menetelmällä lasketut lehmäindeksit vuodelle 1982.

Rekisteristä saatavat tiedot järjestettiin SPSSX-lajitteluohjelmalla tiedostoksi, jossa tietue sisälsi yhden lehmän tiedot. Tietueet lajiteltiin juoksevaan järjestykseen maatalouskeskuksen ja karjan mukaan. Lehmäindeksin laskemiseksi kirjoitettu Fortran-kielinen ohjelma luki indeksin laskemiseen tarvittavat tiedot suoraan tästä SPSSX-ohjelmalla lajitellusta tiedostosta.



### Karjojen valinta

Lehmien jalostusarvon laskemista varten valittiin satunnaisesti 1000 karjaa, jolloin lehmien lukumääräksi saatiin 10781. Karjoissa oli keskimäärin 10.8 lehmää. Karjaluvun suuruudesta johtuen karjoihin sisältyi kooltaan ja tuotostasoltaan hyvin erilaisia karjoja. Lisäksi erirotuisia eläimiä sisältyi otokseen riittävästi tilastolistien analyysien suorittamiseksi roduittain. Jalostusarvot laskettiin vain sellaisten karjojen lehmille, missä lehmien lukumäärä oli alle viisikymmentä. Rajoitus tehtiin tietokoneen muistikapasiteetin vuoksi. Sitä suurempien karjojen MME-yhtälöiden ratkaisuun vaadittu yleistetyn käänteismatriisin muodostaminen on tehtävä iteroimalla.

### Malli

Jalostusarvot laskettiin lehmille jokaisesta karjasta erikseen. Lehmien arvotelussa käytettiin lehmien tuotosta kuvaamaan mallia:

$$\underline{y} = \underline{X}\underline{b} + (\underline{z} \ 0) \begin{bmatrix} \underline{a}_1 \\ \underline{a}_2 \end{bmatrix} + \underline{z}_p + \underline{e}$$

missä:

$\underline{y}$  = tuotosvektori

$\underline{X}$  = tunnettu kiinteiden tekijöiden kerroinmatriisi

$\underline{b}$  = tuntematon kiinteiden tekijöiden (vuosi) vektori

$\underline{z}$  = tunnettu a:n ja p:n kerroinmatriisi

0-matriisi on tuloksettomien eläinten kerroinmatriisi

$\underline{a}$  = additiivisten geneettisten arvojen vektori, jossa

$a_1$  eläimille, joilla on tuotostietoja ja

$a_2$  sekä isille että eläimille, joilla ei ole tuotostietoja

$\underline{p}$  = ei-additiivisten geneettisten arvojen ja pysyvien ympäristötekijöiden vektori

$\underline{e}$  = satunnaisten ympäristötekijöiden (virheiden) vektori

Oletukset:

$E(\underline{y}) = \underline{X}\underline{D}$ , joka pätee vain niin kauan kun lehmien isien arvoja ei ole lisätty malliin.

$$\text{Var} \begin{bmatrix} \underline{a} \\ \underline{p} \\ \underline{e} \end{bmatrix} = \begin{bmatrix} A\sigma_a^2 & 0 & 0 \\ 0 & I\sigma_p^2 & 0 \\ 0 & 0 & I\sigma_e^2 \end{bmatrix}$$

### Sukulaismatriisin käänteismatriisin lisääminen

Sukulaismatriisin käänteismatriisi  $\underline{A}^{-1}$  laskettiin HENDERSONIN (1976) menetelmällä, missä oletetaan, että populaatio ei ole sukusiitetty. Menetelmän etuna on, ettei se vaadi mitään erityistä järjestystä eläinten listaamisessa. Tällöin eläimet voidaan ryhmitellä kahteen ryhmään käännettävän matriisin koon pienentämiseksi. Ensiksi tulevat eläimet, joilla on tuotostietoja (arvosteltavat lehmät) ja toiseksi eläimet, joilla ei ole tuotostietoja (nuoret eläimet ja lehmien isät).

$\underline{A}^{-1}$ -matriisi jaettiin alamatriiseiksi siten, että eläimet eriteltiin sen mukaan, onko niillä tuotostietoja vai ei.

$$\underline{A}^{-1} = \begin{bmatrix} \underline{A}^{11} & \underline{A}^{12} \\ \underline{A}^{21} & \underline{A}^{22} \end{bmatrix}, \text{ missä}$$

$\underline{A}^{11}$  ; eläimet, joilla on tuotostietoja

$\underline{A}^{22}$  ; eläimet, joilla ei ole tuotostietoja

$\underline{A}^{12}$  ; edellisten kahden ryhmän eri eläinten väliset yhteydet

Näin saatu sukulaismatriisin käänteismatriisi  $\underline{A}^{-1}$  lisättiin  $\underline{a}$ :n kerroinmatriisiin ( $\underline{z} \ 0$ ).

#### Isien jälkeläisarvostelutulosten lisääminen

Isien BLUP-jälkeläisarvostelutulokset huomioitiin lisäämällä sekamallin yhtälöihin isiä vastaavat yhtälöt. Yhtälöryhmän vasemmalle puolelle isän yhtälön halkaisijaelementtiin lisättiin  $q(1-r)/(4-h^2)$  ja yhtälöryhmän oikealle puolelle kirjoitettiin sonnin jalostusarvo (VAN VLECK 1982):

$$\frac{2(1-r)(4+(q-1)h^2)}{h^2(4-h^2)} \hat{v}$$

Lisäämällä malliin isien jälkeläisarvostelut, saadaan lehmien arvostelu sidottua samaan geneettiseen perustasoon ja täten välteetään karjojen geneettisten tasoerojen ja geneettisen edistymisen aiheuttama harha lehmien arvostelussa.

MME-yhtälöt

Ratkaistaviksi MME-yhtälöiksi saatiin:

$$\begin{bmatrix} \underline{X}'\underline{M}\underline{X} & \underline{X}'\underline{M}\underline{Z} & \underline{0} \\ \underline{Z}'\underline{M}\underline{X} & \underline{Z}'\underline{M}\underline{Z} + \underline{A}^{11} t & \underline{A}^{12} t \\ \underline{0} & \underline{A}^{21} t & \underline{D} + \underline{A}^{22} t \end{bmatrix} \begin{bmatrix} \underline{b} \\ \underline{a}_1 \\ \underline{a}_2 \end{bmatrix} = \begin{bmatrix} \underline{X}'\underline{M}\underline{y} \\ \underline{Z}'\underline{M}\underline{y} \\ \underline{Q}'\underline{y} \end{bmatrix}$$

missä  $\underline{M} = \underline{I} - \underline{Z}(\underline{Z}'\underline{Z} + \underline{I}_k)^{-1}\underline{Z}'$  ja  $\underline{D}$  sisältää isien halkaisijaelementteihin lisätyn halkaisijamatriisin eli  $q_1 (1 - r)/(4 - h^2)$  ja  $\underline{Q}$  on halkaisijamatriisi, joka sisältää painokertoimet isien jalostusarvoille sekä  $t = (1 - r)/h^2 = \sigma_e^2/\sigma_a^2$ .

Kiinteiden tekijöiden esikorjaus

Malliin otettiin mukaan kiinteänä tekijänä vain vuosivaikutus. Kaikille lehmillä, jotka olivat poikineet ennen huhtikuun loppua merkittiin poikimisvuodeksi edellinen vuosi, koska tuotoksen katsottiin kuuluvan edelliseen satovuoteen. Muut kiinteät tekijät esikorjattiin multiplikatiivisesti samalla menetelmällä kuin Suomessa korjataan karjantarkkailusta saatavat lehmien tuotostiedot laskettaessa sonnien jälkeläisarvostelua ja lehmäindeksejä (SYVÄ-JÄRVI 1984. Suullinen tiedonanto). Kiinteiden tekijöiden korjaus edeltäkäsini tehtiin varmuuden parantamiseksi, koska kärjasta ei saada luotettavia estimaatteja kiinteille tekijöille. Lisäksi tietokoneen rajoitetun muistikapasiteetin vuoksi ei mallia voida laajentaa miten suureksi tahansa. Korjaamalla tuotostiedot edeltäkäsini vähennetään ratkaistavien yhtälöiden määrää sekä helpotetaan yhtälöiden muodostamista.

Tuotostiedot korjattiin kahdessa vaiheessa, ensin poikimakerran sisällä poikimaiän, poikimakuukauden ja poikimavälin suhteen. Sen jälkeen näin korjatut tulokset korjattiin poikimakerran suhteen.

Tuotoskauden pituus oli kaikilla lehmillä 305 päivää. Lehmille, jotka lypsivät lyhyemmän ajanjakson, merkittiin koko tuotos 305-päivän tuotokseksi ja lehmiltä, jotka lypsivät yli 305 päivää otettiin huomioon vain 305 päivän aikana tuotettu maitomäärä.

Korjauskertoimien laskemiseksi lehmät luokitellaan poikimakerran, poikimakuukauden, poikimavälin ja poikimaiän suhteen seuraavasti:

poikimakerta	1	=	1
	2	=	2
	>3	=	3
tyhjä tai			
virhe		=	4
poikimakuukausi	1-2	=	1
	3-4	=	2
	5-6	=	3
	7-8	=	4
	9-10	=	5
	11-12	=	6
poikimaväli	<315	=	1
	316-345	=	2
	346-375	=	3
	376-405	=	4
	406-435	=	5
	436-465	=	6
	466-540	=	7

Jos poikimaväliä ei voida laskea (esim. päivämäärät ovat virheellisiä), lasketaan tyhjäkausi. Tällöin laskennallinen poikimaväli = tyhjäkausi + 270 pv .  
Jos kumpaakaan ei voida laskea oletetaan luokka 4.

Ikä poikiessa eri poikimakerroilla:

Poikimakerta	1	2	3	luokka
	<690	<1055	1420	= 1
	691-720	1056-1085	1421-1450	= 2
	721-750	1086-1115	1451-1480	= 3
	751-780	1116-1145	1481-1510	= 4
	781-840	1146-1205	1511-1570	= 5
	841-900	1205-1265	1571-1630	= 6
	>900	>1265	>1630	= 7

Jos ikää poikiessa ei voida laskea, niin otetaan luokka 3. Jos poikimakerta on tuntematon tai >3 ei ikää poikiessa lasketa.

Jokaiselle yksittäiselle luokalle lasketaan keskiarvo ja korjaustermi k siten, että luokan keskiarvoa verrataan ryhmän keskiarvoon. Esimerkiksi, jos poikimakerroksen tuotosten keskiarvo on 5500 kg ja keskiarvo jossakin ryhmässä on 5000 kg niin tällöin korjauskertoimeksi saadaan:

$$k = \frac{5500}{5000} = 1.1$$

Lehmän jokainen yksittäinen tuotos korjataan vastaavalla k-ker-  
toimella. Korjatuista tuotoksista lasketaan poikimakerroittain  
keskiarvot sekä korjauskerroin, joka korjaa poikimakerran vaiku-  
tuksen. Keskiarvot ja korjauskertoimet lasketaan seuraavissa ryh-  
missä:

poikimakerta	luokka
1	= 1
2	= 2
3	= 3
4	= 4
5-6	= 5
7-8	= 6
9-10	= 7
>10	= 8

Tuotokset korjataan poikimakertakorjauskertoimilla, jolloin saa-  
daan lehmälle korjattu 305-päivän tuotos.

Tässä työssä käytettiin korjauskertoimina valtakunnallisesta kar-  
jantarkkailusta saatujen tuotostietojen perusteella laskettuja  
korjauskertoimia vertailukelpoisuuden varmentamiseksi.

Lehmäindeksi laskettiin rotujen sisällä karjoittain. Rotujen vai-  
kutusta korjattiin poistamalla sonnin jälkeläisarvostelun perus-  
teella saadusta arvosta rodun vaikutus. Tämä tehtiin vähentämällä  
sonnin arvosta, joka on roturyhmän geneettisen arvon ja sonnin  
periyyttämiskyvyn (predicted difference PD) summa, kaikkien  
vähintään kahdenkymmenen tyttären perusteella arvosteltujen rodun  
sonnien keskiarvo. Kertomalla tämä kahdella saatiin sonnin addi-  
tiivinen jalostusarvo rodun sisällä.

Työssä käytettyjen perinnöllisten tunnuslukujen arvot

Mallissa tarvittiin toistuvuuden ( $r$ ) ja periytyvyysasteen ( $h^2$ ) parametreja. Näitä parametreja ei kuitenkaan ole käytettävissä, vaan sopivat likiarvot löytyvät esimerkiksi kirjallisuudesta.

Maidontuotannossa on periytyvyysasteelle ja toistuvuudelle kirjallisuudesta löydettävissä useita toisistaan poikkeavia likiarvoja. MAIJALAN ja HANNAN (1974) kirjallisuuskatsauksen keskiarvo periytyvyysasteelle oli 0.27 ja toistuvuudelle 0.49. Samassa katsauksessa todettiin, että eri laktaatiokausien väliset toistuvuuskerroimet poikkeavat selvästi toisistaan ja että ensimmäisen laktaatiokauden periytyvyysaste on korkeampi kuin seuraavien laktaatiokausien.

Tässä työssä käytettiin toistuvuudelle ( $r$ ) arvoa 0.40 ja periytyvyysasteelle ( $h^2$ ) arvoa 0.25, jotka ovat samoja, mitä Suomessa käytetään sonnien jälkeläisarvostelussa ja lehmäindeksin laske-  
misessa (SYVÄJÄRVI 1984. Suullinen tiedonanto). Lisäksi oletettiin, että ko. tunnusluvuille voidaan käyttää em. arvoja riippumatta laktaatiokaudesta.

#### Lehmien jalostusarvojen laskeminen

Jalostusarvojen laskemiseksi luettiin havaintoaineistoa tietueittain. Tietueet sisälsivät maatalouskeskuksen, karjan numeron, lehmän identiteetin, joka koostui korvanumerosta ja syntymävuodesta sekä rodun, tuotosten lukumäärän, CHRISTENSENIN (1981) menetelmällä lasketun lehmäindeksin ja tuotostiedot poikimavuositain. Tietueeseen sisältyi lisäksi isän kantakirjanumero, NASTA:ssa käytettyjen tytärien lukumäärä sekä NASTA-indeksi, emän maatalouskes-



kus, karjan numero, identiteetti, sekä rotu, emänisän kantakirjanumero, emänisän NASTA-arvostelussa käytettyjen tytärten lukumäärä sekä NASTA-arvo.

Laskentamenetelmän havainnollistamiseksi voidaan esittää numeerinen esimerkki. Olkoon havaintoaineisto seuraavan kaltainen, missä lehmien tuotokset on esitetty vuosittain. Taulukossa on lisäksi eläimet, joilla ei ole tuotostietoja.

vuosi	lehmä		emä		isä			summa
	1	2	3	4	5	6	7	
81	-	3696	-	-	-	-	-	3696
82	-	4704	-	-	-	-	-	4704
83	8103	4108	-	-	-	-	-	12211

Tiedoista muodostettiin suoraan karjan sisällä  $\underline{X}'\underline{Z}$  ja  $\underline{Z}'\underline{Z}$  matriisit sekä  $\underline{X}'\underline{y}$  ja  $\underline{Z}'\underline{y}$  vektorit.

Esimerkin aineistosta saadaan

$\underline{X}'\underline{Z}$  ennen absorbointia

$\underline{Z}'\underline{Z}$  ennen absorbointia

$$\begin{bmatrix} 0 & 1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$$

Tämän jälkeen kirjoitettiin lista lehmistä, joilla oli tuotostiedot, sekä emistä ja lehmien isistä ja emänisistä. Listaan kirjoitettiin ensin lehmät, sitten emät, joille ei ollut saatavissa tuotostietoja ja jotka siten eivät vielä olleet listassa lehminä, sekä viimeiseksi isät ja emänisät. Tässä yhteydessä varastoitiin myös isien ja emänisien NASTA-indeksit sekä niiden tytärten lukumäärä. Listaa kirjoitettaessa eläimet numeroitiin uudelleen.

Kirjoitetaan lista esimerkin eläimistä:

yksilö	isä	emä	sonnin maitopoik- keama, kg	tyttärien lkm
1	5	3		
2	6	4		
3	6	0		
4	7	0		
5	0	0	166.40	160
6	0	0	76.40	5428
7	0	0	-258.00	0

Sonnille numero seitsemän on ennustettu maitopoikkeaman arvo sukulaisten tietojen perusteella.

Tämän karjan sisäisen eläinlistan tiedoista laskettiin suoraan sukulaismatriisin käänteismatriisi  $\underline{A}^{-1}$ , käyttäen HENDERSONin (1976) kehittämää suoraa menetelmää.  $\underline{A}^{-1}$ -matriisi varastoitiin yläkolmiomatriisina vektorimuodossa.

Esimerkissä sukulaismatriisin käänteismatriisiksi  $\underline{A}^{-1}$  saadaan

$$1/6 \begin{bmatrix} 12 & 0 & -6 & 0 & -6 & 0 & 0 \\ 0 & 12 & 0 & -6 & 0 & -6 & 0 \\ -6 & 0 & 11 & 0 & 3 & -4 & 0 \\ 0 & -6 & 0 & 11 & 0 & 3 & -4 \\ -6 & 0 & 3 & 0 & 9 & 0 & 0 \\ 0 & -6 & -4 & 3 & 0 & 11 & 0 \\ 0 & 0 & 0 & -4 & 0 & 0 & 8 \end{bmatrix}$$

Ei-additiivisen geneettisen vaihtelun ja satunnaisen ympäristöstä johtuvan vaihtelun selittävä  $p$  absorboitiin  $\underline{X}'\underline{X}$ -matriisiin samanaikaisesti kun  $\underline{X}'\underline{X}$ -matriisi muodostettiin  $\underline{Z}'\underline{X}$ :sta ja varastoitettiin matriisina vektorimuodossa.

Sekamallin yhtälöryhmä muodostettiin  $\underline{X}'\underline{X}$ ,  $\underline{X}'\underline{Z}$  ja  $\underline{Z}'\underline{Z}$  matriiseista lisäämällä  $\underline{Z}'\underline{Z}$  matriisien diagonaaliin varianssisuhteet  $t$  ja  $k$  sekä sukulaismatriisin käänteismatriisi  $\underline{A}^{-1}$ . Samalla absorboitiin  $p$   $\underline{a}$ :ta vastaavaan kerroinmatriisiin  $\underline{Z}'\underline{Z}$  ja  $\underline{X}'\underline{Z}$ -matriisiin, sekä lisättiin  $\underline{Z}'\underline{Z}$ -matriisiin halkaisijaan sonnien yhtälöihin tytärten lukumäärä. Sekamallin yhtälöt talletettiin yläkolmiomatriisina vektorimuodossa.

Esimerkin MME-yhtälöiksi, joihin samalla absorboidaan  $p$ :n yhtälöt, saadaan

0.857	-0.143	-0.143	0.0	0.571	0.0	0.0	0.0	0.0	0.0
-0.143	0.857	-0.143	0.0	0.571	0.0	0.0	0.0	0.0	0.0
-0.143	-0.143	1.657	0.8	0.571	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.800	5.6	0.0	-2.4	0.0	-2.4	0.0	0.0
0.571	0.571	0.571	0.0	6.514	0.0	-2.4	0.0	-2.4	0.0
0.0	0.0	0.0	-2.4	0.0	4.4	0.0	1.2	-1.6	0.0
0.0	0.0	0.0	0.0	-2.4	0.0	4.4	0.0	1.2	-1.6
0.0	0.0	0.0	-2.4	0.0	1.2	0.0	29.2	0.0	0.0
0.0	0.0	0.0	0.0	-2.4	-1.6	1.2	0.0	872.88	0.0
0.0	0.0	0.0	0.0	0.0	0.0	-1.6	0.0	0.0	3.2

Yhtälöryhmän oikealle puolelle muodostettiin  $\underline{X}'\underline{y}$  ja  $\underline{Z}'\underline{y}$  yhtälöt, joihin absorboitiin  $p$ :n vaikutus. Samalla lisättiin oikealle puolelle sonnien jälkeläisarvostelutulokset.

Esimerkissä yhtälöryhmän oikealle puolelle saadaan

1909.1
2917.1
8803.5
6482.4
7147.4
0.0
0.0
9318.4
133070.4
-1238.4

MME-yhtälöiden käänteismatriisi saatiin HENDERSONIN kirjoittamalla DJNVHF-aliohjelmalla. Ratkaisut saatiin kertomalla yhtälöryhmän oikea puoli MME-yhtälöiden käänteismatriisilla. Samalla ratkaistiin myös  $\underline{p}$ :n estimaatit. Ratkaisut tulostettiin listana: karjan numero, lehmän rotu, poikimakertojen lukumäärä, keskituotos, CHRISTENSENIN (1981) menetelmällä laskettu lehmäindeksi, BLUP-lehmäindeksi ja sen standardoitu ennustusvirheen hajonta sekä ei-additiivisten -ja pysyvien ympäristötekijöiden vaikutus  $\hat{\underline{p}}$ . Ennusteen luotettavuutta kuvaava standardoitu ennustusvirheen hajonta ( $SPE/\sigma$ ) saadaan  $\underline{p}$ :n absorboinnin vuoksi helposti laskettua ottamalla  $\underline{g}$ :n kerroinmatriisin yleistetyn käänteismatriisin diagonaalista neliöjuuri. Ennustusvirheen hajonta kuvaa ennustuksen varmuutta hajonnan pienetessä havaintojen lisääntyessä.

Esimerkin ratkaisuksi saadaan

Vuosi	1981	]	=	[	4425.266
	1982				5433.266
	1983				6039.920
Lehmät	1				559.762
	2				-428.602
	3				264.142
	4				-508.553
Sonnit	5	354.276			
	6	152.455			
	7	-641.276			

Verrattaessa sonnien 5-7 saamia jalostusarvoja lähtötietojen arvoin huomataan, että ne vastaavat hyvin toisiaan, sillä saatu jalostusarvo on likimain 2 \* sonnien maitopoikkeama.

Tulokset voidaan kirjoittaa esimerkiksi seuraavaan tietokoneella luettavaan muotoon

Lehmä	rotu	$\hat{a}$	SPE/ $\sigma$	$\hat{p}$
1	1	559.76	0.52	300.66
2	1	-428.60	0.53	-300.66

missä  $\hat{a}$  = additiivinen jalostusarvo

SPE/ $\sigma$  = standardoitu jalostusarvon ennustusvirheen hajonta

$\hat{p}$  = pysyvien ympäristötekijöiden arvo

## TULOKSET

Lehmien rodun ja karjan sisällä lasketut BLUP-indeksit ilmoittavat suoraan lehmän jalostusarvon. Indeksejä voidaan verrata myös karjojen välillä arvosteluvarmuuden huonontumatta, koska indeksiin lisätyt sonnien jälkeläisarvostelutulokset poistavat karjojen välisten geneettisten tasoerojen vaikutukset (BOLGIANO 1981). Tuloslistaan tulostettiin jokaiselle otoksessa mukana olleelle lehmälle lehmän additiivinen jalostusarvo laskettuna sekä BLUP-menetelmällä että CHRISTENSENin (1981) menetelmällä. Näin indeksejä pystyttiin vertaamaan keskenään.

Koko aineistosta laskettuna indekseille saatiin seuraavat tunnusluvut keskiarvoksi, hajonnaksi, minimiksi ja maksimiksi (kg 4%-maitoa)

Menetelmä	Keskiarvo	Hajonta	Minimi	Maksimi	Lukumäärä
CHRISTENSEN	123.768	270.084	-992.180	1383.540	10781
BLUP	-113.117	404.730	-1654.150	1712.400	10781

Siihen miksi BLUP-menetelmällä lasketun indeksin keskiarvo on CHRISTENSENin menetelmällä lasketun indeksin keskiarvoa huomattavasti pienempi, on varmaan monta syytä. Esimerkiksi laskentamenetelmien erilaisuus sukulaisten tietojen hyväksikäytössä, saattaa vaikuttaa siihen, että indeksit saavat erilaiset keskiarvot. BLUP-indeksin suurempi hajonta johtunee BLUP-indeksin pienemmästä virhevaihtelusta.

Indeksien jakauma on kuvissa 1 ja 2. Indeksien jakaumat ovat keskenään samanlaisia ja erittäin lähellä normaalijakaumaa.







BLUP-indeksin ja Christensenin indeksin välille laskettiin sekä Pearsonin korrelaatio että Spearmanin järjestyskorrelaatio. Pearsonin korrelaatio, laskettaessa koko aineistosta, oli tilastollisesti erittäin merkitsevä 0.865. Karjojen sisäinen Pearsonin korrelaatio karjojen yli laskettuna oli 0.880 \*\*\* , n = 8783. Laskettaessa karjan sisäinen Pearsonin korrelaatio karjojen yli rotujen sisällä saatiin BLUP-indeksin ja Christensenin indeksin väliseksi korrelaatioksi

Ayrshire	0.883 ***	n = 7501
Suomen karja	0.881 ***	n = 41
Friisiläinen	0.881 ***	n = 505

Järjestyskorrelaatio kuvaa lehmien välisen järjestyksen välistä yhteyttä, eli sillä voidaan tutkia asettaako indeksit lehmät erillaiseen järjestykseen. Suoritettaessa valintaa täytyy indeksin luotettavasti asettaa eläimet geneettiseen arvojärjestykseen. Spearmanin järjestyskorrelaatiokerroin on järjestyslukuista laskettu korrelaatiokerroin. Karjojen sisäinen Spearmanin korrelaatio karjojen yli laskettuna ayrshire rotuisille lehmille oli erittäin merkitsevä 0.854. Otoksessa oli mukana 889 karjaa, joiden keskimääräinen karjakoko oli 10.5 lehmää. Täten otos oli edustava, sillä karjakoko oli lähellä maan keskiarvoa.

Ayrshire-rotuisille karjoille laskettiin Pearsonin korrelaatio karjojen BLUP-menetelmällä lasketun geneettisen ja fenotyypin keskiarvon välille yli karjojen. Korrelaatioksi saatiin 0.352 \*\*, ts. karjojen välisestä vaihtelusta  $100(0.352)^2$  % = 12 % johtuu geneettisistä tekijöistä. Karjojen yli lasketuksi fenotyypiksi ja additiivisten arvojen rotukeskiarvoksi saatiin

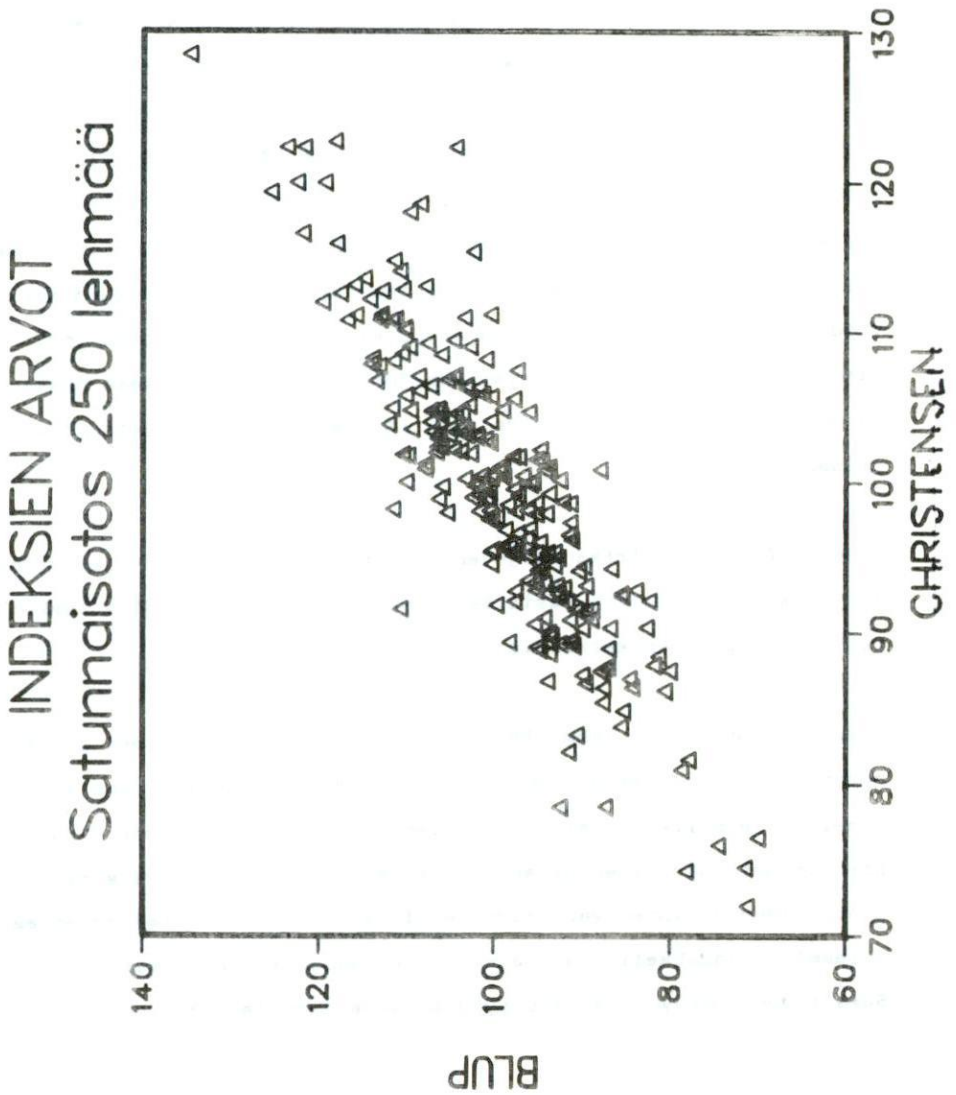
	Keskiarvo	Hajonta	Minimi	Maksimi	N
FENOT.ARVO	5943.282	726.705	3892.000	8869.622	954
JALOSTUSARVO	-120.791	191.683	-822.190	690.450	954

Kuva 3 esittää näiden kahden indeksin standardoidut arvot satunnaisotokselle. Kuvasta nähdään, että indeksit asettavat lehmät melkein samaan järjestykseen. Indeksien arvot samalla lehmällä poikkeavat hieman toisistaan. Kuvat 4, 5, 6, 7 ja 8 (Liite) esittävät indeksien standardoituja arvoja erikokoisissa karjoissa. Kuvista huomataan, että alle kymmenen lehmän karjoissa on indeksien välillä huomattavasti enemmän hajontaa kuin suuremmissa karjoissa. Voidaan olettaa, että eläinten arvostelu pienissä karjoissa on aina epävarmempaa kuin suurissa karjoissa. Lisäksi voidaan huomata hajonnan indeksien arvojen välillä lisääntyvän indeksien ääripäissä.

Kuva 9 (Liite) esittää karjojen fenotyypisten keskiarvojen jakauman ja kuva 10 (Liite) esittää karjojen BLUP-menetelmällä laskettujen genotyypisten keskiarvojen jakauman.

Ayrshire rotuisille lehmille laskettiin myös fenotyypinen ja genotyypinen edistyminen rodun sisällä laskemalla eri vuonna syntyneille lehmille geneettiset ja fenotyypiset keskiarvot. Sonnien arvostelutulosten hyväksikäyttö mahdollistaa eri vuosina syntyneiden eläinten vertailun. Kuvista 11 ja 12 (Liite) huomataan molemmilla indekseillä selvä samansuuntainen kasvava trendi. Sama trendi näkyy myös fenotyypisessä edistymisessä, kuva 13 (Liite).

Kuva 3: BLUP-indeksin ja CHRISTIENSENin indeksin standardoidut arvot satunnaisesti valituille 250 lehmälle.



## TULOSTEN TARKASTELU

Kotieläinjalostajalle on tärkeintä, että eläimet pystytään asettamaan geneettiseen arvojärjestykseen. Siinä suhteessa Christensenin indeksi ja BLUP-indeksi olivat hyvin lähellä toisiaan sillä järjestyskorrelaatio niiden välillä oli 0.85 Eläinten järjestys ei siis merkittävässä määrin muuttunut laskettaessa additiivinen jalostusarvo kummalla tahansa menetelmällä. Pearsonin korrelaatio näiden kahden indeksin välillä oli tilastollisesti erittäin merkittävä. Korrelaatiokertoimien perusteella indeksit ovat siis erittäin yhdenmukaiset.

Hajonta oli BLUP-menetelmällä lasketussa indeksissä huomattavasti suurempi. Tämä oli toisaalta odotettavissa, sillä mitä tehokkaammin jalostusarvon ennustuksen virhevaihtelua saadaan pienemmäksi eli mitä varmempi indeksi on, sitä suurempi on jalostusarvojen hajonta. BLUP-indeksin varmuutta lisää CHRISTENSENIN indeksiin verrattuna suurempi sukulaistietojen varsinkin emän tietojen hyväksikäyttö, sekä vuositekijän mukaanotto indeksiin. Samansuuntaisen tuloksen sai myös POWELL (1978).

Kumpi indekseistä olisi parempi, ei edellä olevilla mittauskeinoilla pysty varsinaisesti osoittamaan. Lisätietoa jommankumman indeksin paremmuudesta, esimerkiksi sonninemien valinnassa, saisi vasta laskemalla korrelaation lehmäindeksin ja lehmän poikien jälkeläisarvostelujen välillä ja vertaamalla näin saatuja korrelaatioita, jotka on laskettu kummallekin lehmäindeksille erikseen.

Geneettistä edistymistä oli selvästi havaittavissa kummankin indeksin yhteydessä. Eläinten valinnassa on tärkeää, että malli huomioi geneettisen edistymisen populaatiossa. Lehmien geneettinen edistyminen seuraa sonnien geneettistä edistymistä jonkin verran jäljessä mikä on luonnollista, koska sonnit voidaan arvostella paljon varmemmin ja sonneilla voidaan karsinta pitää paljon tiukempuna kuin lehmillä. Tulevaisuudessa mahdollisesti käytettävä alkionsiirto mahdollistaa lehmien ankaramman valinnan, jolloin lehmäindeksin varmuuden ja erottelukyvyn tarve lisääntyy.

HENDERSONin (1975a) kehittämä BLUP-menetelmään perustuva lehmäindeksi mahdollistaa yhtäaikaisten arvostelujen laskemisen eläimille, joilla jo on tuotostietoja sekä sellaisille eläimille, joilla ei ole tuotostietoja. Näin voidaan hiehoille ja nuorsonneille laskea sukulaisten tietojen perusteella odotusarvoindeksi. Eläimelle voidaan laskea indeksi karjan sisällä heti kun eläin saavuttaa täysimittaisen 305 päivän tuotosjakson.

Tietokoneaikaa ja muistitilaa BLUP-menetelmä vaatii suhteellisen paljon. Laskettaessa tälle aineistolle lehmäindeksit Maatalouden Tutkimuskeskuksen tietokoneella (Vax-11/780) tarvittiin 55 minuuttia keskustietokoneaikaa. Yhdelle karjalle lehmäindeksejä laskettaessa laskuaika on lyhyt, mutta laskettaessa koko tarkkailuun kuuluvalla populaatiolla indeksejä on ajan tarve suuri. Toisaalta MME-yhtälöiden kääntäminen yli viidenkymmenen lehmän karjassa ei onnistu suoraan, vaan tällaisille suurille karjoille indeksejä laskettaessa on kääntäminen suoritettava iteroimalla.

Lehmäindeksikin voidaan tulevaisuudessa kehittää useiden ominaisuuksien analyysiksi (vrt. HENDERSON ja QUAAS 1976), jolloin menetelmällä lasketaan samanaikaisesti kaikkien ominaisuuksien BLUP-arvot huomioiden eläinten ja ominaisuuden väliset korrelaatiot. Edellytyksenä on tietysti, että ominaisuuksien väliset korrelaatiot ovat tiedossa, mikä ei aina ole varmaa.

KIRJALLISUUSLUETTELO

- BOLGIANO, D.C. 1981. An examination of the accuracy of cow evaluations for intraherd versus intraherd comparison. A thesis. Presented to the Faculty of the Graduate School of Cornell University. Ithaca.
- CHRISTENSEN, L.G. 1981. Basic principles in the estimation of breeding values by direct updating. Z.Tierzuchtung und Zuchtgsbiol. 98:132-150.
- CUNNINGHAM, E.P. 1969. Animal breeding theory. Institute of Animal Genetics and Breeding Agricultural College of Norway. Universitetsforlaget, Oslo. Moniste 272 s.
- DANELL, B. 1981. Evaluation of sires on first-lactation yield of Swedish dairy cattle. Report 51. Swedish University of Agricultural Sciences, Department of Animal Breeding and Genetics. Uppsala.
- EVERETT, R.W., HENDERSON, C.R. & HANSON, C.F. 1977. The Northeast cow ETA report. Animal Sci. Mimeo Series. Department of Animal Science, Cornell University. Ithaca.
- HENDERSON, C.R. 1973. Sire evaluation and genetic trends. Page 10 in Proc. of the animal breeding and genetics symposium in honor of Dr. Jay L. Lush. ASAS and ADSA, 1973 ASAS, 4251 Illinois Building, 113 North Neilstreet, Champaign, Illinois 61820. 32 s.
- HENDERSON, C.R. 1975a. Use of all relatives in intraherd prediction of breeding values and producing abilities. J.Dairy Sci. 58:1910-1916.

- HENDERSON, C.R. 1975b. Inverse of a matrix of relationships due to sires and maternal grandsires. J.Dairy Sci. 58:1917-1921
- HENDERSON, C.R. 1975c. Best linear unbiased estimation and prediction under selection model. Biometrics 31:423.
- HENDERSON, C.R. 1976. A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. Biometrics 32:69-83.
- HENDERSON, C.R. 1977. Best linear unbiased prediction of breeding values not in the model for records. J.Dairy Sci. 60:783-787.
- HENDERSON, C.R. & QUAAS, R.L. 1976. Multiple trait evaluation using relatives' records. J.Anim.Sci. 43:1188-1197.
- HILL, W.G. & SWANSON, G.J.T. 1983. A selection index for dairy cows. Anim.Prod. 37:313-319.
- LASLEY, J.F. 1978. Genetics of livestock improvement 3. painos. Prentice-Hall, Inc. Englewood Cliffs, New Jersey.
- LINDSTRÖM, U.B. Jalostusoppi 2. Helsingin Yliopisto, Kotieläinten Jalostustieteen laitos. Luentomoniste, 147s.
- MAIJALA, K. & HANNA, M. 1974. Reliable phenotypic and genetic parameters in dairy cattle. Proc.1st Wld Congr. Genett. Appl. Livest.Prod. Madrid 1974. 1:541-563.



- MAO, I.L. 1982. Modeling and data analysis in animal breeding. Notes for an internordic post-graduate course Uppsala, Sweden. Department of Animal Breeding and Genetics Swedish University of Agricultural Sciences. Uppsala, Sweden.
- PEARSON, R.E. & MILLER, R.H. 1978. Cow evaluation in North America. *Livest.Prod.Sci.* 5:19-28.
- PHILIPSSON, J., DOMMERHOLT, J., FIMLAND, E., GAILLARD, C., GJOL-CHRISTENSEN, L., LEDERER, J., McCLINTOCK, A.E. & MOCQUOT, J.C. 1978. Problems in cow evaluation and current use of cow index. Report of a Working Group on Cow Evaluation. *Livest.Prod.Sci.* 5:3-18.
- POWELL, R.L. 1978. A Procedure for including the dam and maternal grandsire in USDA-DHIA cow indexes. *J.Dairy Sci.* 61:794-800.
- POWELL, R.L., NORMAN, H.D. & DICKINSON, F.N. 1976. The USDA-DHIA modified contemporary cow index. The USDA-DHIA modified contemporary comparison. Sire summary and cow index procedures. *Agric.Res.Serv.United States Department of Agriculture. Prod.Res.Rep.* 165:35-40.
- POWELL, R.L., NORMAN, H.D. & ELLIOT, R.M. 1981. Accuracy of genetic indexes of cows from adding relatives. *J.Dairy Sci.* 64:838-843.
- QUAAS, R.L. 1976. Computing the diagonal elements and inverse of a large numerator relationship matrix. *Biometrics.* 32:949-953.
- SEEGER, P., LUNDSTRÖM, K. & DANELL, Ö. 1978. Statistisk introduktion till Harvey's program. Rapport 27. Institutionen för husdjurförädling och sjukdomsgenetik. Sveriges Lantbruksuniversitet. Uppsala.

SLANGER, W.D., JENSEN, E.L., EVERET, R.W. & HENDERSON, C.R. 1976.

Prgramming cow evaluation. J.Dairy Sci. 59:1589-1594.

SYVÄJÄRVI, J. & Hellman, T. 1983. Lehmäkortin indeksit ja tulostiedot.

Nautakarja 3/1983:26-28.

VAN VLECK, L.D. 1979. Notes on the theory and application of  
selection principles for the genetic improvement of animals.

2. painos. Dept. of Animal Science, Cornell University.

Ithaca, N.Y. s.188-212.

VAN VLECK, L.D. 1982. Derivation of Henderson's method of

incorporating artificial insemination sire evaluations into

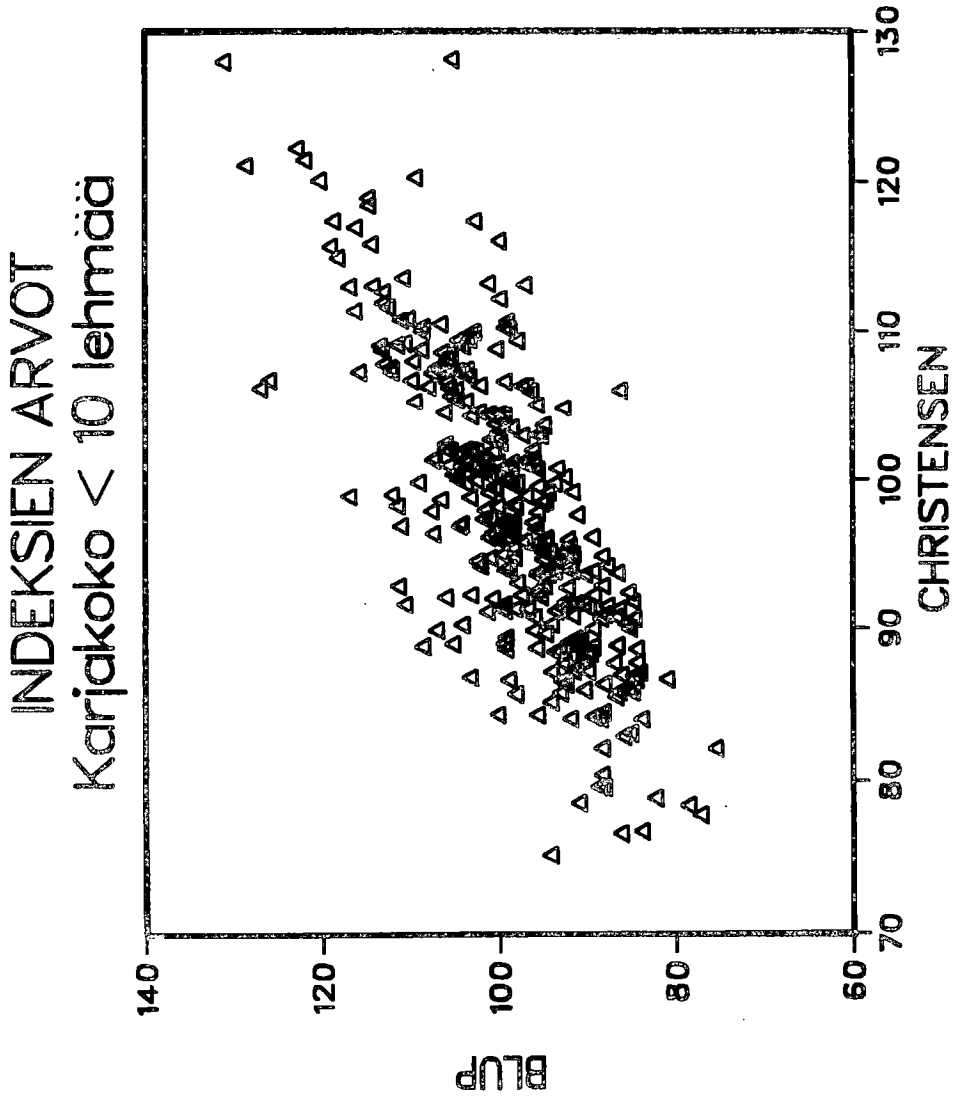
intraherd prediction of breeding values. J.Dairy Sci. 65:284-286.

WRIGHT, S. 1922. Coefficients of inbreeding and relationship.

Amer.Nat. 56:330-338.

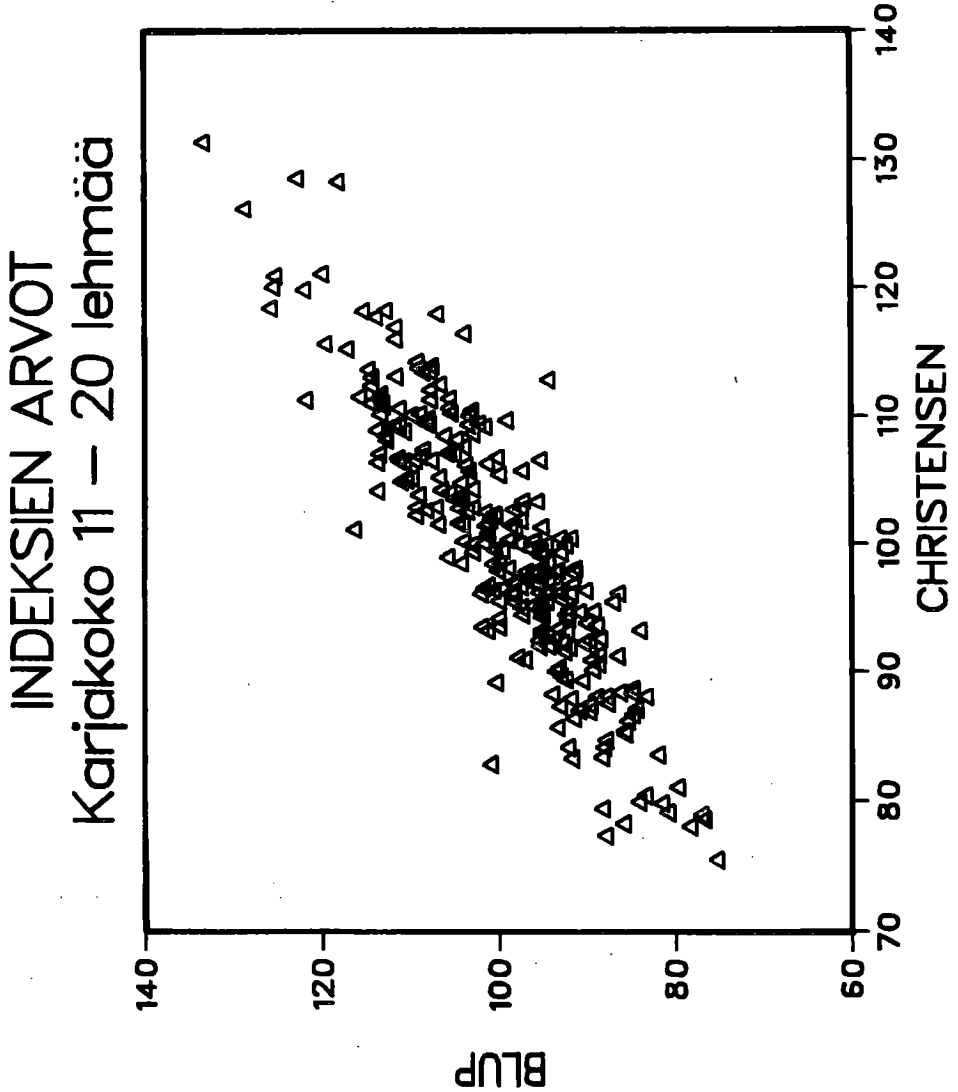
LIITE

Kuva 4: BLUP-indeksin ja CHRISTENSENin indeksin standardoidut arvot laskettuina lehmillä, jotka olivat <10 lehmän karjoissa.



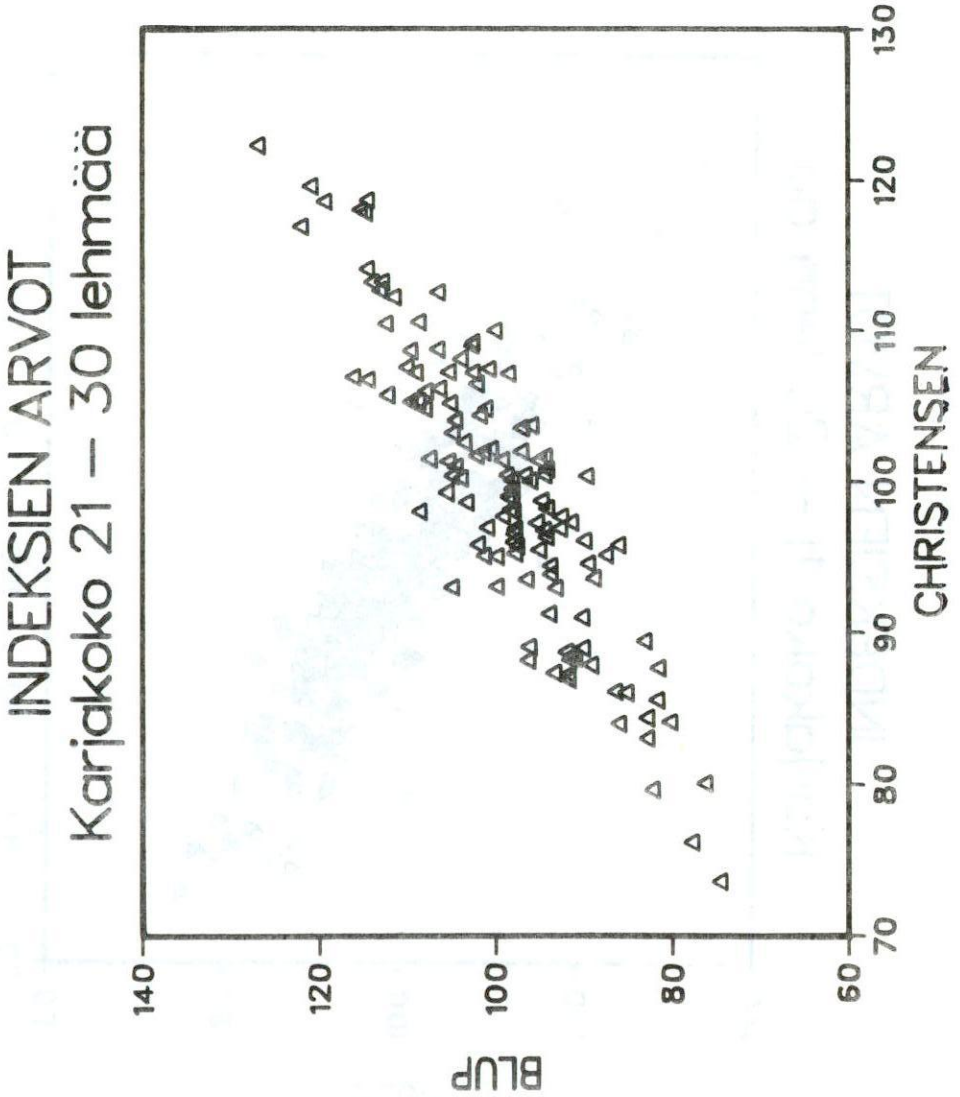
LIITE

Kuva 5: BLUP-indeksin ja CHRISTENSENin indeksin standardoidut arvot laskettuina lehmillä, jotka olivat 15 - 20 lehmän karjoissa.



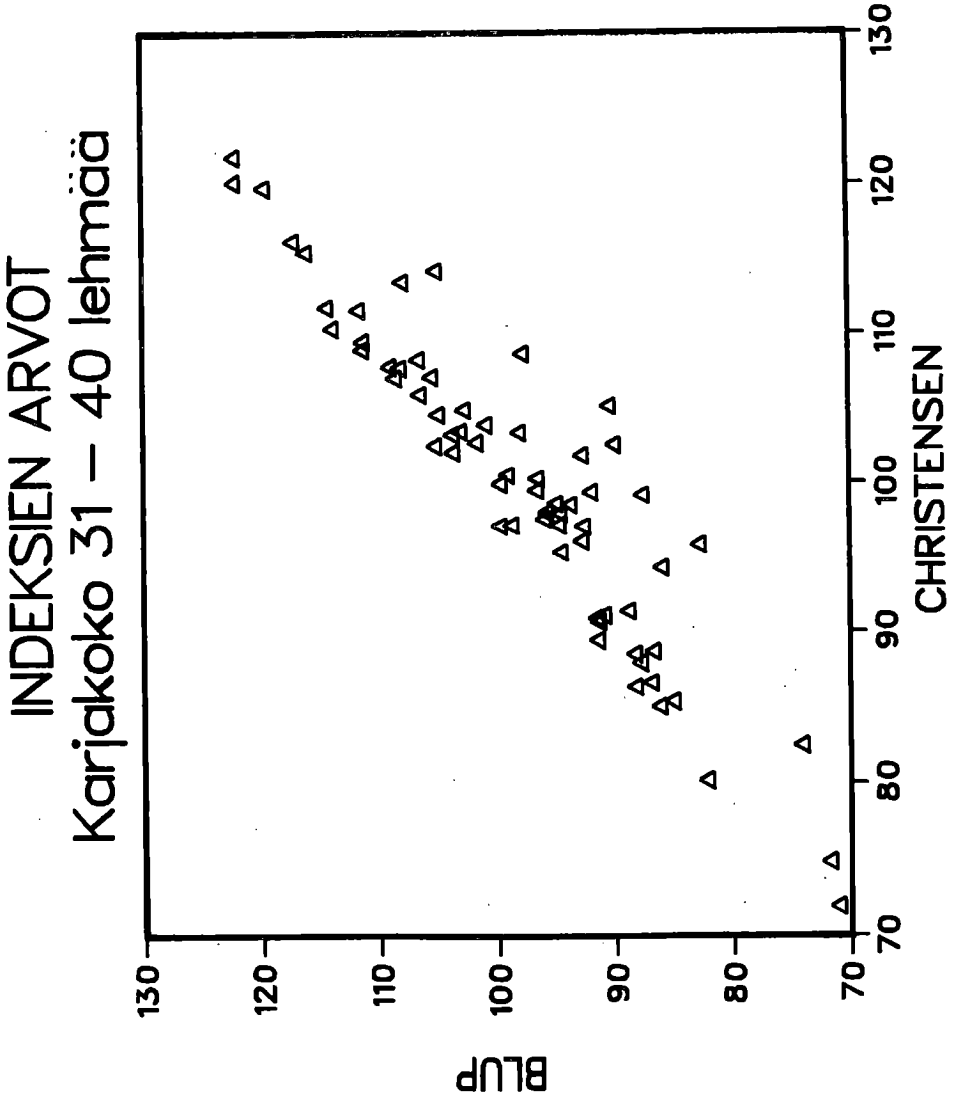
LIITE

Kuva 6: BLUP-indeksin ja CHRISTENSENin indeksin standardoidut arvot laskettuina lehmillä, jotka olivat 21 - 30 lehmän karjoissa.



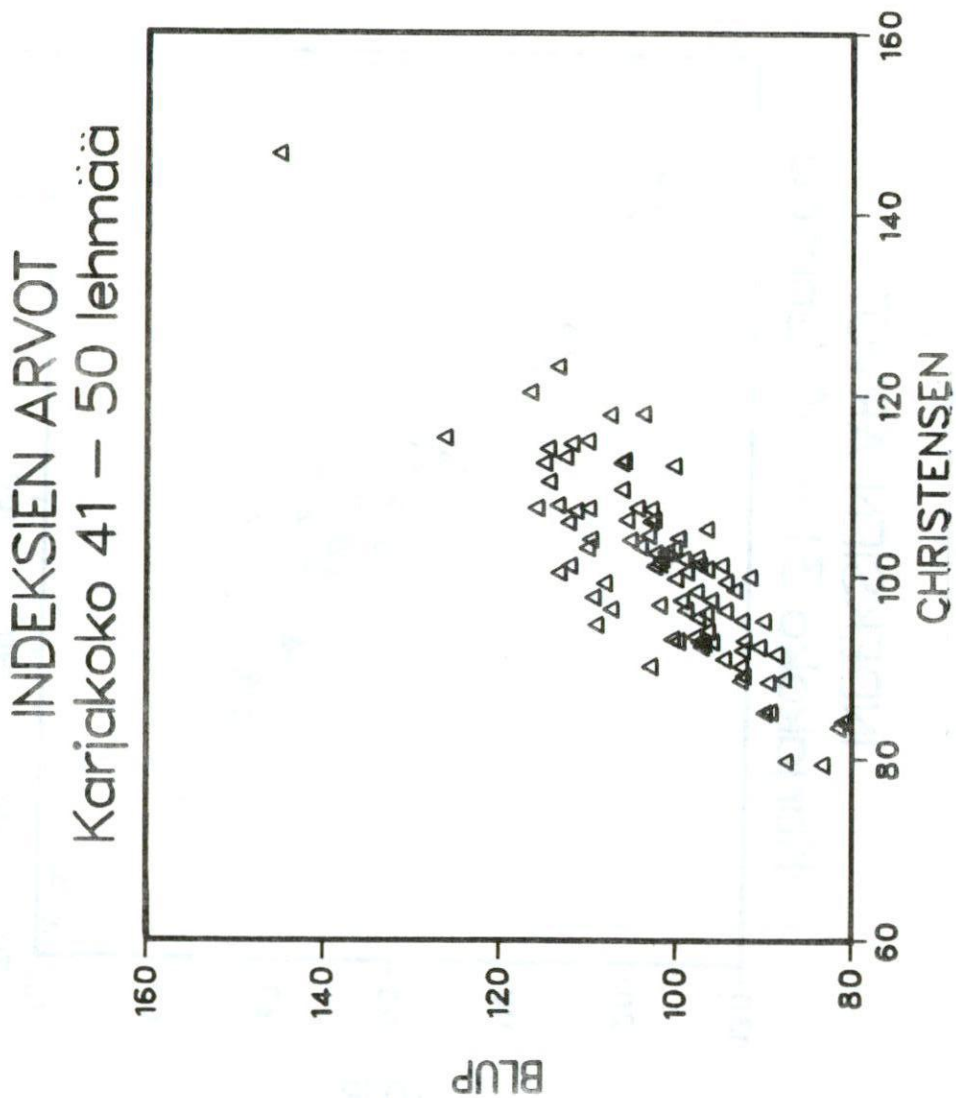
LIITE

Kuva 7: BLUP-indeksin ja CHRISTENSENin indeksin standardoidut arvot laskettuina lehmillä, jotka olivat 31 - 40 lehmän karjoissa.



LIITE

Kuva 8: BLUP-indeksin ja CHRISTENSENin indeksin standardoidut arvot laskettuina lehmillä, jotka olivat 41 - 50 lehmän karjoissa.





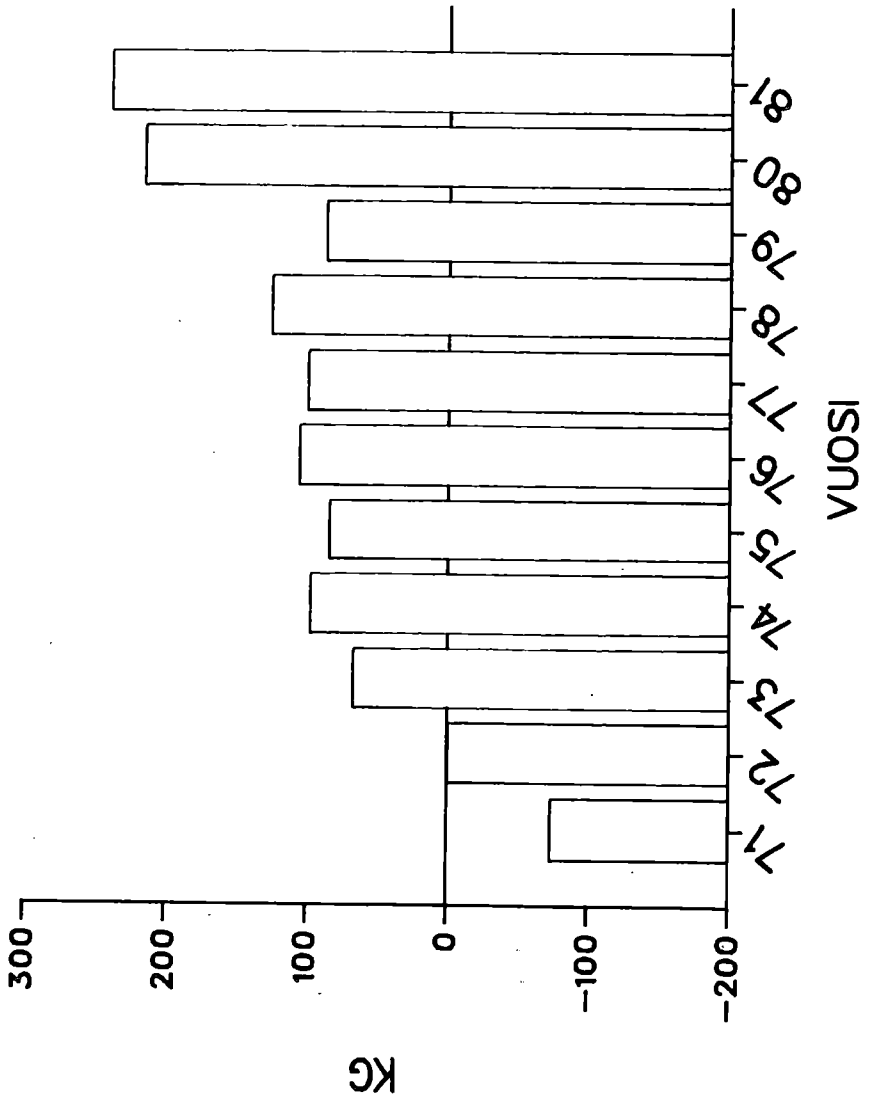




LIITE

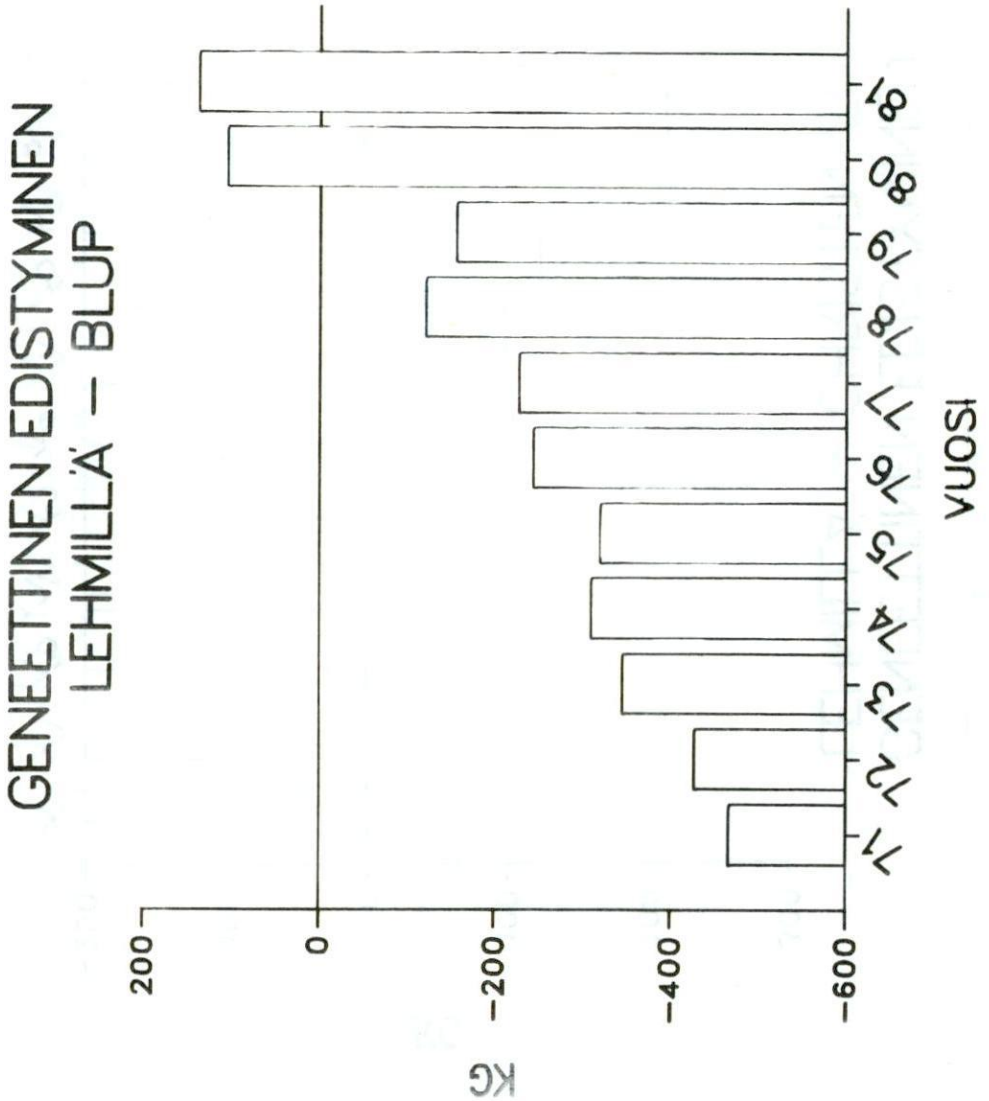
Kuva 11: Ayrshire rotuisten lehmien geneettinen edistymisen CHRISTENSENin menetelmällä laskettuna.

GENEETTINEN EDISTYMINEN  
LEHMILLÄ - CHRISTENSEN



LIITE

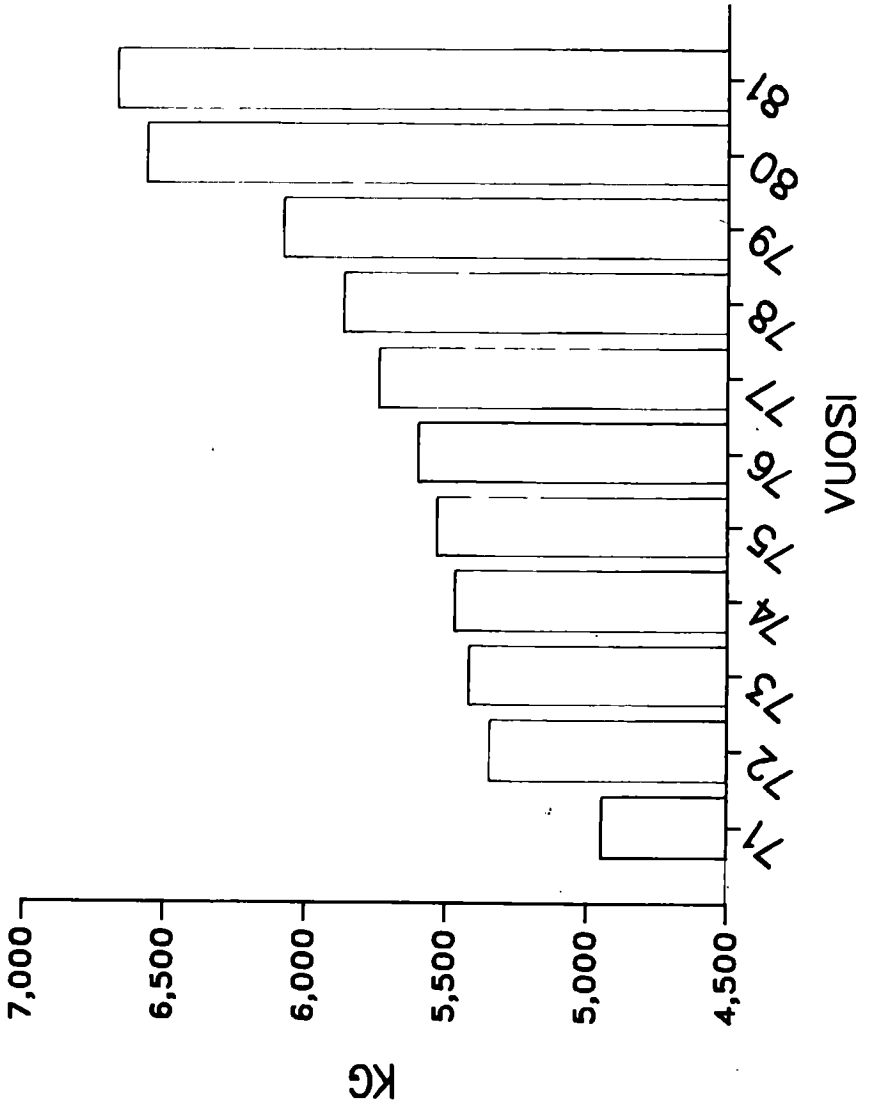
Kuva 12: Ayrshire rotuisten lehmien geneettinen edistyminen BLUP-menetelmällä laskettuna.



LIITE

Kuva 13: Ayrshire rotuisten lehmien fenotyypin edistyminen.

# FENOTYYPIN EDISTYMINEN LEHMILLÄ



## SARJASSA ILMESTYNYT VUODESTA 1980 LÄHTIEN:

40. RUOHOMÄKI, HILKKA, 1980. Lihakarjakokeiden tuloksia IV. 29 s.
41. JALOSTUSPÄIVÄ 9.4.1980. 43 s.
42. LAMMASPÄIVÄ 24.4.1980. 33 s.
43. SIRKKOMAA, S., 1980. Simulointitutkimus sukusiitoksen ja voimakkaan valinnan käytöstä munijakanojen jalostuksessa. Progradu-työ, 90 s.
44. RUOHOMÄKI, HILKKA, 1980. Eri rotuisten lihanautojen elopainot ja iät 160, 180, 210 ja 250 kilon teuraspainossa. 13 s.
45. MAIJALA, K., 1981. Kotieläinten perinnöllisen muuntelun säilyttäminen. 52 s.
46. RUOHOMÄKI, HILKKA, 1981. Lihakarjakokeet vuosina 1960—1980. 30 s.
47. JÄLKEÄISARVOSEMINAARI 12.5.1981. 44 s.
48. MAIJALA, K., 1981. Jalostus ja lisääntyminen vaikuttavina tekijöinä lihanaudan tuotannossa. 20 s.
49. SYRJÄLÄ-QVIST, LIISA, BOMAN, MARJATTA & MOISIO, S., 1981. Lammastalouden rakenne ja merkitys elinkeinona Suomessa, 25 s.
50. LEUKKUNEN, ANU, 1982. Keinosiemennyskarjujen jälkeläisarvostelu tyttärien porsimistulosten perusteella. Liseniaattityö, 88 s.
51. LAURILA, TERHI, 1982. Kilpailutulosten käyttö ratsuhevosten suorituskyvyn mittaamisessa. Pro gradu-työ, 84 s.
52. LINDSTRÖM, U., 1982. Merkkigeenien ja -aineiden käyttöarvosta kotieläinjalostuksessa, 13 s.
53. LEUKKUNEN, ANU, 1982. Heikkolaatuisen rehun hyväksikäytön geneettinen edistäminen, 24 s.
54. OJALA, M., 1982. Eri kudoslajien kasvurytmi naudoilla, 22 s.
55. OJALA, M., 1982. Vanhempien tuotantotietojen ja eräiden ympäristötekijöiden yhteys sonnien kasvukoetuloksiin. Laudaturtyö, 54 s.
56. OJALA, M., 1982. Kilpailutulosten käyttöarvosta ravihevosten jalostuksessa. Liseniaattityö, 16 s.
57. KENTTÄMIES, HILKKA, 1982. Naudanlihantuotantoon vaikuttavista geneettisistä tekijöistä ja ympäristötekijöistä sekä kasvun mittaamisesta kenttäkokeissa. Liseniaattityö, 104 s.
58. HUHTANEN, P., 1982. Suomenkarjan kokonaistaloudellisuus muihin rotuihin verrattuna. Laudaturtyö, 82 s.
59. KUOSMANEN, S., 1983. 305 pv:n maitotuotoksen ennustaminen osatuotostietojen perusteella. Pro gradu-työ, 100 s.
60. HEISKANEN, MINNA-LIISA, 1983. Hevoson keinosiemennys tuore- ja pakastespermalla. Pro gradu-työ, 63 s.
61. MARKKULA, MERJA, 1984. Kanojen yleiseen sairaudenvastustuskykyyn liittyviä tekijöitä, 24 s.

62. MÄNTYSAARI, E., 1984. Valintaindeksi jälkeläisarvosteltujen keinosiemennyssonnien kokonaisjalostusarvon kuvaajana. Pro gradu-työ, 86 s.
63. LAUKKANEN, HANNELE, 1984. Maidon sähkönjohtokykyyn vaikuttavat tekijät ja johtokyvyn käyttömahdollisuuksista utaretulehduksen vastustamisessa. Pro gradu-työ, 68 s.
64. SYVÄJÄRVI, J., 1984. Tutkimuksia maitorotuisten sonnien jälkeläisarvostelun varmistamiseksi ja monipuolistamiseksi. Lisensiaattityö, 14 s. LIITE: Tarkkailulehmien maidon solupitoisuuden vaihtelu ja yhteys maitotuotokseen. 78 s.
65. MAIJALA, K., 1984. Ulkomaisia kokemuksia suomenlampaasta ja sen risteytyksistä. 27 s.
66. ARONEN, PIRJO, 1985. Liharotuisten nautojen painoihin vaikuttavista tekijöistä ja painojen korjaamisesta. Pro gradu-työ, 80 s.
67. JUGA, J., 1985. Karjansisäinen lehmien arvostelu. Pro gradu-työ, 93 s.

**ISBN 951-45-3695-9**  
**ISSN 0356-1429**  
Helsinki 1985  
Yliopistopaino