

MODELI Z NAKLJUČNO REGRESIJO*

Špela MALOVRH^{a)} in Milena KOVAČ^{b)}

^{a)} Univ. v Ljubljani, Biotehniška fak., Odd. za zootehniko, SI-1230 Domžale, Slovenija, asist.,
e-pošta: spela@mrcina.bfro.uni-lj.si.

^{b)} Isti naslov, izr.prof., dr., mag.

Delo je prispelo 02. oktobra 2000, sprejeto 16. oktobra 2000.

Received October 02, 2000, accepted October 16, 2000.

IZVLEČEK

Proizvodnost plemenskih živali je pogosto merjena v zaporednih intervalih, kot so lastnosti mlečnosti, rasti, zauživanje krme in zunanost. Meritve, zbrane ob različnem času ali na različnem mestu, so med seboj odvisne. Njihova povprečja in standardni odkloni se s časom spreminjajo in korelacije so odvisne od oddaljenosti med meritvami. Modeli z naključno regresijo (MNR), ki opisujejo proizvodne funkcije, v povezavi s kovariančnimi funkcijami, ki zajamejo časovno oz. prostorsko odvisnost komponent (ko)variance, so najprimernejši za analizo take vrste podatkov. MNR smo primerjali z modeli za laktacijske zaključke, modeli z intervali in z modeli s sistematsko regresijo po virih. Za ponazoritev smo uporabili podatke slovenske populacije črno-belih krav. MNR boljše opišejo biološko ozadje proizvodnosti živali v času in omogočajo selekcijo tako na raven kot obliko proizvodnih funkcij. Dovoljujejo gibljivejši postopek preskusa in s tem znižanje stroškov pri zbiranju podatkov. Plemenske vrednosti lahko napovemo pred koncem laktacije, kar skrajšuje generacijski interval, pogostejši izračuni pa so lahko tudi v pomoč pri vodenju reje.

Ključne besede: plemenske živali / proizvodnost / modeli / naključna regresija / MNR

RANDOM REGRESSION MODELS[†]

ABSTRACT

Performance of breeding animals is frequently measured on successive intervals, like milk traits, growth, feed intake, and exterior. The traits collected at different time or different places are not independent. Their average as well as standard deviation changes over time, and the correlation depends on the distance among them. For such kind of data, random regression models (RRM) describing production functions combined with covariance functions accommodating time-(place-)dependent (co)variance components are the method of choice. The models are compared to lactation yield models, to test interval model, and fixed regression models based on literature review, and illustrated using data from Slovenian black-white dairy-cattle population. RRM reflect better biological background of animal performance over time and allow selection on the overall production level as well as the shape of production curve. They are in favor of more flexible recording scheme and thus, cost reduction for data collection. Furthermore, the genetic evaluation can be performed before lactation is finished shortening the generation interval and can be done more frequently giving a useful tool for herd management.

Key words: breeding animals / performance / models / random regression / RRM

* Prispevek je del doktorske disertacije, mentorica izr.prof. dr. Milena Kovač.

[†] This paper is a part of dissertation thesis, supervisor assoc.prof. Milena Kovač, Ph.D.

UVOD

Modeli z naključno regresijo in kovariančnimi funkcijami niso povsem novi. O vključevanju naključne regresije je pisal že Henderson (1982), Laird in Ware (1982) sta jo uporabila za opis rastnih krivulj pri človeku, Kirkpatrick in Heckman (1989) ter Kirkpatrick in sod. (1990) pa so uporabili kovariančne funkcije na področju evolucijske biologije. Sprva so omenjene modele razvijali in uporabljali zgolj za manjše nize podatkov iz preskusov v živinoreji pa postajajo sedaj vse privlačnejši zaradi možnosti praktične uporabe pri selekciji domačih živali. Aktualni so postali šele pred kratkim ne toliko zaradi razvoja teoretičnih osnov kot zaradi razvoja računalniških zmogljivosti. Metoda namreč uvede primernejši statistični model za analizo ponavljajočih meritev na živalih, ki so v časovnem ali prostorskem sosledju, kot so npr. meritve na kontrolni dan za količino, vsebnosti sestavin mleka, število somatskih celic v laktaciji, meritve telesne mase, zaužite krme pri rastočih živalih in podobno.

Vzemimo primer iz govedoreje. Laktacijska krivulja opisuje vpliv faze laktacije na količino mleka. Poznana je splošna oblika laktacijske krivulje, vemo pa tudi, da na potek vplivajo pasma, starost krave oziroma zaporedna laktacija, število molž na dan, sezona telitve, tehnologija reje, čas oploditve itn. Potek je torej povezan z običajnimi sistematskimi vplivi in je opisan z nizom regresijskih koeficientov, na primer pri Wilminkovi funkciji s tremi (Wilmink, 1987), Ali-Schaefferjevi (Ali in Schaeffer, 1987) ter Guo-Swalvejevi (Guo in Swalve, 1995) pa s petimi regresijskimi koeficienti. Zanje imamo navadno veliko število opazovanj, pripadajo večjim skupinam živali, zato so zanesljivo ocenjene in jih uvrščamo med sistematske vplive.

Prav tako smo prepričani, da je za potek laktacijske krivulje pomembna lastnost živali. Žal imamo le malo informacij o poteku laktacijskih krivulj pri živalih - to so le zaključene laktacije pri posameznih kravah. Tako regresijski koeficienti za individualne laktacijske krivulje lahko le napovedujejo razlike med živalmi. V modelu jih torej obravnavamo kot naključne spremenljivke.

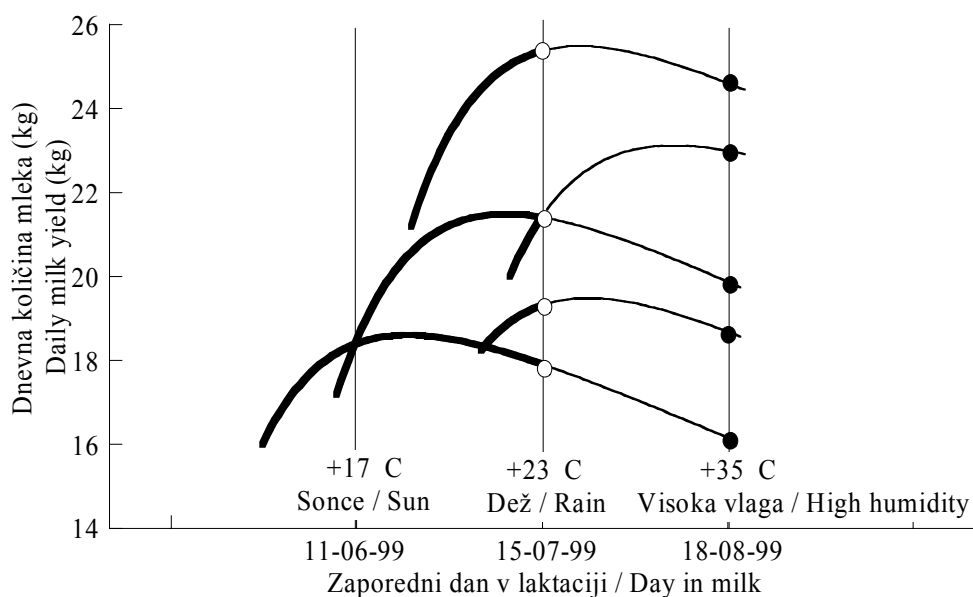
Statistični modeli z naključno regresijo vključujejo tako dva niza regresijskih koeficientov: sistematski in naključni niz. Sistematski niz nadomesti običajne sistematske vplive, naključni pa opisujejo tako naključne okoliške (skupno okolje v hlevu, gnezdu, skupno okolje laktacije itn.) kot genetske komponente (npr. aditivni genetski vpliv).

Namen prispevka je razložiti modele z naključno regresijo, pri čemer bomo kot primer uporabili kontrole mlečnosti pri govedu. Hkrati bomo izkoristili priložnost za primerjavo metod, ki se uporabljajo za analizo meritev na kontrolni dan. Za ilustracijo smo uporabili Potočnikove podatke (1999) za slovensko populacijo črno-belega goveda.

Naključno regresijo se vse več vključuje tudi pri modeliranju rasti in zauživanja krme pri prašičih (Anderson in Pedersen, 1995) in govedu (Koenen in Veerkamp, 1998; Veerkamp in Thompson, 1999; Meyer, 1999; Meyer, 2000), pa tudi za ocene kondicije pri govedu (Koenen in Veerkamp, 1998; Jones in sod., 1999).

ZAPISI NA KONTROLNI DAN

Zapis oz. meritve na kontrolni dan (angl. test day record) predstavlja zbrano količino mleka ter količino ali vsebnost maščobe in beljakovin v 24-ih urah. Te meritve so odvisne od pasme, zaporedne laktacije, števila molž na dan, faze laktacije, različnih razmer v okolju na kontrolni dan (slika 1) in genetskih vrednosti živali. Prav tako se od ene do druge kontrole spreminja primerjalna skupina (slika 1), saj ene živali laktacijo začenjajo, medtem ko jo druge končujejo in tako zapuščajo skupino molženih krav.



Slika 1. Skupno okolje v čredi na dan kontrole (po Schaefferju, 1998).
Figure 1. Common environment in herd on test day (after Schaeffer, 1998).

Pri klasičnih postopkih vrednotenja plemenskih vrednosti pri kravah zaradi postopka izračunavanja laktacijskih zaključkov (standardna laktacija) zahtevamo, da so kontrole čim bolj enakomerno porazdeljene po laktaciji (ICAR, 1995). Z metodo, kjer v izračun napovedi plemenskih vrednosti vključimo kar meritve na kontrolni dan, so lahko intervali med kontrolami manj enakomerni, meritve lahko izmenično opravljajo različne osebe (kontrolor, rejec), kar tudi poceni kontrolo. Omogoča obdelavo podatkov pred zaključkom laktacije in s tem odločitev, ali bomo žival ponovno pripuščali ali po zaključeni laktaciji izločili, kar je zlasti zanimivo pri drobnici (Brežnik, 1999).

Funkcije za opis laktacijskih krivulj

Za opis laktacijske krivulje na osnovi zapisov dnevnih kontrol lahko uporabimo različne linearne pa tudi nelinearne regresijske modele, ki se razlikujejo v številu regresijskih parametrov (za podrobnejši pregled glej Jamrozik in sod., 1997a) in tem, kako dobro se prilegajo meritvam. Boljše opisujejo krivulje z več parametri, znotraj skupin funkcij z istim številom parametrov so razlike navadno majhne (Guo in Swalve, 1995). Pogosto uporabljene funkcije za opis laktacijskih krivulj so Woodova gama funkcija (1967, funkcija [1]), njena transformacija z naravnim logaritmom (Wood, 1967; funkcija [2]), Wilminkova (Wilmink, 1987; funkcija [3]), Ali-Schaefferjev regresijski model (Ali in Schaeffer, 1987; funkcija [4]) ter Guo-Swalvejeva (Guo in Swalve, 1995; funkcija [5]). Woodova funkcija je med zgoraj naštetimi edina nelinearna, njeno reševanje poteka iterativno in se tudi slabše prilega.

$$y_t = a * t^b * \exp(-c * t) \quad [1]$$

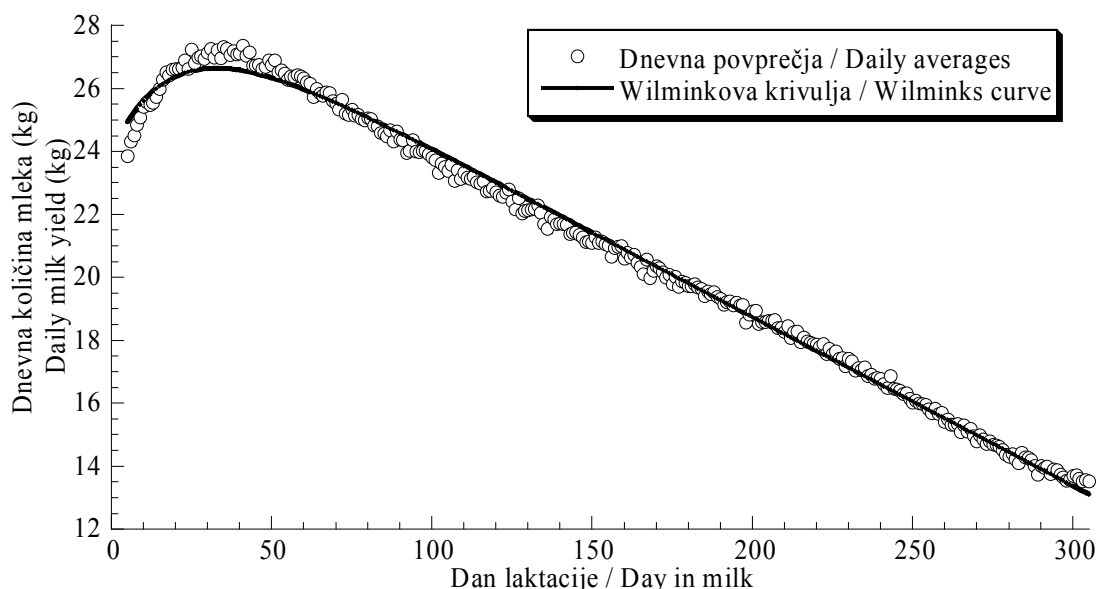
$$\ln(y_t) = \ln(a) + b * \ln(t) - c * t \quad [2]$$

$$y_t = a + b * t + c * \exp(-0.05 * t) \quad [3]$$

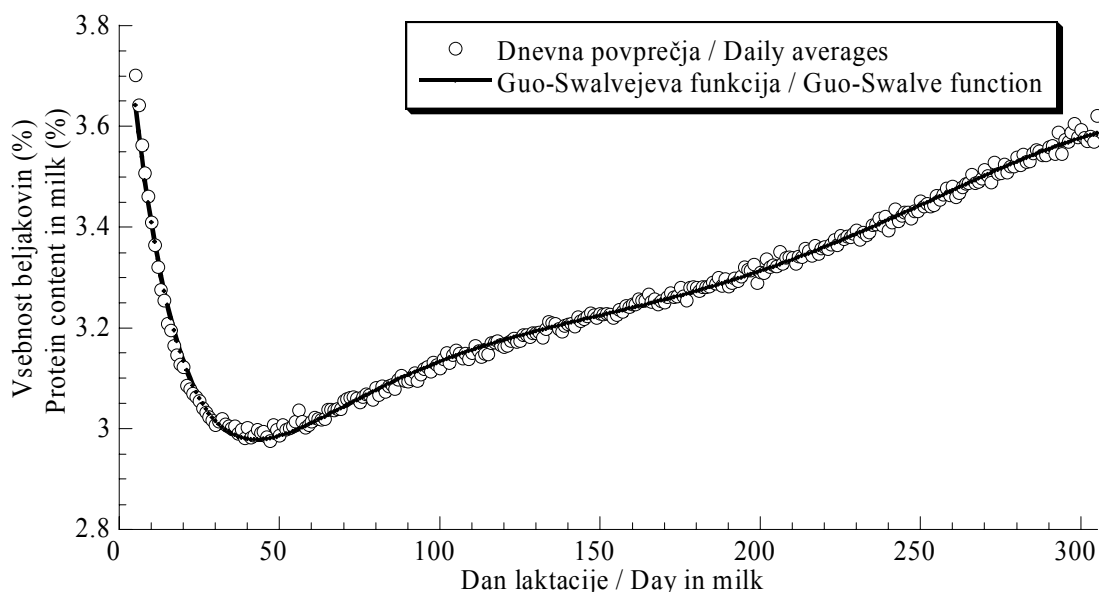
$$y_t = a + b * (t/305) + c * (t/305)^2 + d * \ln(305/t) + f * (\ln(305/t))^2 \quad [4]$$

$$y_t = a + b * (t/305) + c * \sin(t/100) * t^2 + d * \sin(t/100) * t^3 + f * \exp(-0.055 * t) \quad [5]$$

Prileganje uporabljene krivulje za opis spreminjanja dnevne količine in vsebnosti sestavin v mleku s stadijem laktacije je lahko boljše ali slabše. Na primeru dnevne količine mleka slovenskih krav črno-bele pasme (slika 2) vidimo, da Wilminkova funkcija zgreši vrh laktacije, pa tudi kasneje se ne prilega najboljše. Guo-Swalvejeva funkcija na sliki 3 lepše opiše potek laktacijske krivulje za vsebnost beljakovin v mleku. Slednja dobro opiše, če se na krivulji pojavlja „valovanje“.



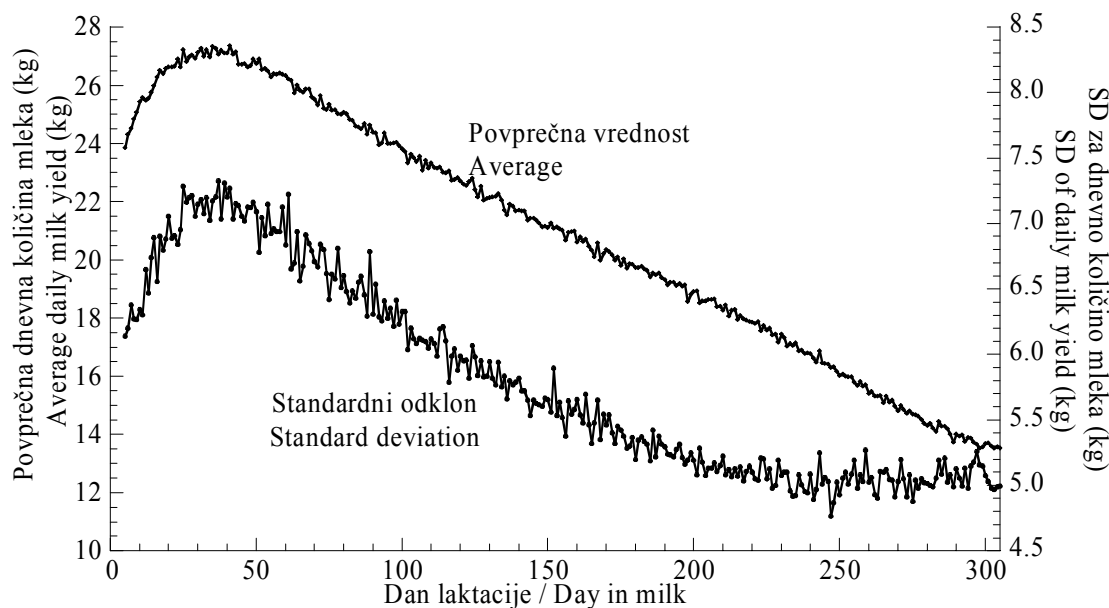
Slika 2. Opis laktacijske krivulje za dnevno količino mleka z Wilminkovo funkcijo.
Figure 2. Fitting lactation curve for daily milk yield with Wilminks function.



Slika 3. Opis laktacijske krivulje za vsebnost beljakovin v mleku s Guo-Swalvejevo funkcijo.
Figure 3. Fitting lactation curve for protein content in milk with function of Guo and Swalve.

Povprečne vrednosti za dnevno količino mleka (slika 4, zgornja krivulja) kažejo tipično obliko laktacijske krivulje. Pri fenotipskem standardnem odklonu (spodnja krivulja) vidimo, da

ni ves čas laktacije enak, narašča do vrha laktacije, čemur sledi padanje nekje do 250. dneva laktacije. To spreminjanje med laktacijo bi bilo potrebno upoštevati pri variancah in kovariancah, s čimer dajemo meritvam različno težo glede na zanesljivost meritev.



Slika 4. Povprečne vrednosti in standardni odklon za dnevno količino mleka v laktaciji.
Figure 4. Mean and standard deviation of daily milk yield during lactation.

PRISTOPI K OBDELAVI MERITEV NA KONTROLNI DAN

Meritve količine namolženega mleka in vsebnosti sestavin mleka so že v osnovi meritve na kontrolni dan. Pri analizi teh meritev v selekcijske namene je možnih več pristopov, ki se razlikujejo po zahtevnosti in porabi računalniških zmogljivosti. Na različne načine (ICAR, 1995) lahko izračunamo oz. ocenimo skupne količine za mleko, mlečno maščobo in beljakovine v standardni laktaciji (305 dni) in v analizi uporabimo laktacijske zaključke za količino mleka in sestavine v enolastnostnem ali večlastnostnem mešanem modelu (model živali). Ocenimo lahko potek laktacijske krivulje za vsako kravo in nato analiziramo parametre krivulj (Schaeffer in Jamrozik, 1996). Če v analizi direktno uporabimo meritve na kontrolni dan, imamo dva pristopa. Prvi je analiza ponavljajočih meritev iste lastnosti (Ptak in Schaeffer, 1993; Reents in sod., 1995), pri čemer predpostavljamo skupno obliko laktacijske krivulje za vse krave. To je model meritev na dan kontrole s sistematsko regresijo (MSR). Tudi za varianco pri tem modelu predpostavljamo, da je konstantna v času laktacije. Nasprotno lahko laktacijo od 5. do 305. dneva razdelimo na intervale oz. stadije (10-14) in meritve na kontrolni dan v različnih stadijih laktacije obravnavamo kot različne lastnosti v večlastnostni analizi (Meyer in sod., 1989; Pander in sod., 1992, 1993). Nadaljnja izboljšava je model z naključno regresijo, ki dopušča (upoštevata) vsaki živali lastno laktacijsko krivuljo (Schaeffer in Dekkers, 1994; Jamrozik in Schaeffer, 1997) in predpostavlja strukturo varianc in kovarianc ponavljajočih meritev (Kirkpatrick, 1994).

Analiza mlečnosti v standardni laktaciji

Za analizo mlečnosti v standardni laktaciji je kot primer model zapisan v skalarni obliki [6], kjer sezona telitve (S_i), starost ob telitvi (A_j) ter čreda (H_k) predstavljajo sistematske vplive, in je aditivni genetski vpliv (a_1) vključen kot naključni vpliv.

$$y_{ijkl} = S_i + A_j + H_k + a_l + e_{ijkl} \quad [6]$$

Pri tem načinu analize je Schaeffer (1998) opozoril na naslednje težave. Pri izračunu praviloma lahko uporabimo samo podatke za zaključene laktacije. Ponekod mlečnost ocenijo (ekstrapolirajo) že pred zaključkom laktacije, s čimer laktacijski zaključek vključuje različno količino informacij za vsako kravo (3 - 10 kontrolnih dni pri različnih zaporednih dneh v laktaciji). Če je laktacijski zaključek ocenjen na prvih treh meritvah (začeta laktacija), je lahko krava s slabšo persistenco ob prvi analizi precenjena (Schaeffer in Jamrozik, 1996). Za vsako kravo se vpliv interakcija čreda-kontrolni dan (HTD) akumulira različno. Nejasno je tudi, kako obravnavati kravo, ki v času laktacije zamenja čredo. Analiza ne vključuje mere za vztrajnost. Kljub naštetim nepopolnostim je metoda dokaj razširjena in je v večini držav v svetu še vedno v uporabi (INTERBULL, 2000).

Model z intervali

Vsako meritev na kontrolni dan lahko obravnavamo kot različno lastnost, tudi genetsko (Meyer in sod., 1989; Pander in sod., 1992, 1993). V primeru 305 dni trajajoče standardne laktacije bi tako imeli v večlastnostnem mešanem modelu kar 305 različnih lastnosti samo za količino mleka na kontrolni dan. Praktično nobena od krav nima vseh meritev, zaradi česar postane matrika varianc in kovarianc zelo prazna (porozna). Genetsko vrednotenje živali s tem modelom omejuje število lastnosti oz. ponovljenih meritev na žival (Thompson in Hill, 1990; Wade in sod., 1990; Wiggans in Goddard, 1997). Model lahko poenostavimo s tem, da zmanjšamo število lastnosti - laktacijo razdelimo na razrede oz. intervale in meritve na kontrolni dan uvrstimo vanje. Kanadčani so laktacijo razdelili v 14 razredov in poskušali uporabiti večlastnostno analizo s 14-imi lastnostmi (Schaeffer, 1998). Mešani model, zapisan v skalarni obliki, je v tem primeru [7]. Pri tem so S_{ti} , A_{tj} in H_{tk} sistematski vplivi za sezono, starost ob telitvi in čredo na intervalu t , a_{tl} je aditivni genetski vpliv. Zaporedni dan laktacije znotraj intervala (t_{ijkl}) je opisan z linearno regresijo.

$$y_{tijkl} = S_{ti} + A_{tj} + H_{tk} + b_t * t_{ijkl} + a_{tl} + e_{ijkl} \quad [7]$$

Matriki varianc in kovarianc za aditivni genetski vpliv in ostanek sta:

$$\text{var}(\mathbf{a}_{14 \times 1}) = \mathbf{G}_{014 \times 14}, \quad \text{var}(\mathbf{e}_{14 \times 1}) = \mathbf{R}_{014 \times 14} \quad [8], [9]$$

Napoved plemenske vrednosti za skupno količino mleka v standardni laktaciji izračunamo nato z enačbo [10], kjer je d_t dolžina t -tega intervala v dnevih. Na posameznih intervalih pomnožimo število dni z napovedjo plemenske vrednosti na tem intervalu ter produkte seštejemo.

$$\text{NPV}_{305} = \sum_{t=1}^{14} d_t * \hat{a}_t \quad [10]$$

Pri kravah, ki imajo več kot eno meritev v intervalu, je potrebno te meritve uvrstiti v različne intervale. Vpliv čreda-kontrolni dan še vedno ni vključen. Rezultati izgledajo precej drugače kot tisti iz analize mlečnosti v standardizirani laktaciji, napovedi plemenske vrednosti imajo drugačne vrednosti, rang živali pa se ohrani (Schaeffer, 1998). Do sedaj še ni dokazano, da so rezultati s to metodo tudi zares boljši. Predstavljeni primer modela z intervali pomeni 14 lastnosti zgolj za količino mleka, vključitev vsebnosti maščobe in beljakovin bi pomenila dodatnih 2 x 14 lastnosti. Večlastnostna analiza z modelom z intervali za meritve na kontrolni dan zahteva bistveno večje računalniške zmogljivosti kot analiza mlečnosti v standardni laktaciji. Wiggans in

Goddard (1997) sta zato predlagala kot metodo kanonično transformacijo. V modelu, ki je obsegal količino mleka, vsebnost maščobe in beljakovin na desetih intervalih v dveh laktacijah, sta namesto 60-ih lastnosti (3 lastnosti x 10 intervalov x 2 laktaciji) v analizo vključila samo šest najbolj značilnih kanoničnih spremenljivk. Le-te imajo največje lastne vrednosti in so linearne kombinacije lastnosti, ki prispevajo največ informacij. Bistveno sta zmanjšala rang matrike varianc in kovarianc in s tem razsežnost sistema enačb. Metoda je tudi predvidena kot način genetskega vrednotenja mlečnega goveda v ZDA v prihodnosti (Wiggans in Goddard, 1997), trenutno pa še vedno uporabljajo laktacijske zaključke v enolastnostnem modelu živali s ponovitvami (VanRaden in Wiggans, 1991; INTERBULL, 2000; USDA-DHIA, 2000)

Model meritev na kontrolni dan s sistematsko regresijo

Model živali z meritvami na kontrolni dan sta prva objavila Ptak in Schaeffer (1993), kmalu je temu sledila uporaba meritev na kontrolni dan za lastnosti mlečnosti pri slovenski populaciji koz (Andonov, 1994). Ptak in Schaeffer (1993) sta v svoj model [11] vključila linearno regresijo s petimi členi (Ali in Schaeffer, 1987).

$$y_{ijkln} = \sum_{m=0}^4 b_{mi} * x_m + \text{HTD}_j + a_k + p_{kl} + e_{ijkln} \quad [11]$$

$$x_0 = 1 \quad x_1 = c \quad x_2 = c^2 \quad x_3 = \ln(1/c) \quad x_4 = [\ln(1/c)]^2 \quad [12]$$

Neodvisne spremenljivke x_1 do x_4 [12] so transformacije c , pri čemer je $c=t/305$. Deljenje dneva v laktaciji (t) s 305 zmanjša velikost neodvisne spremenljivke c na vrednosti med 0 in 1. Regresijski koeficienti za laktacijsko krivuljo (b_{mi}) so znotraj interakcije starost-sezona-regija-leto telitve (indeks i v [11]). Uvedla sta tudi interakcijo čreda-kontrolni dan (HTD_j) kot od dneva laktacije neodvisen sistematski vpliv. Le-ta predstavlja skupno okolje vsem meritvam, opravljenim na isti dan kontrole v posamezni čredi. Aditivni genetski vpliv (a_k) predstavlja samo razliko v višini laktacijske krivulje, ne pa tudi v obliki oz. vztrajnosti. Od naključnih vplivov je v modelu še skupno okolje v laktaciji (p_{kl}), e_{ijkln} pa predstavlja ostanek.

Za matriko varianc in kovarianc za ostanek sta predpostavila diagonalno matriko [13], ki ima na diagonali elemente, odvisne od stadija laktacije za vsak kontrolni dan. Funkcijo za izračun $\sigma_{e_t}^2$ sta ocenila posredno. V ta namen sta izvrednotila ostanke (\hat{e}_{ijkln} , [14]) v modelu, v katerem sta predpostavila, da so identični. Nato sta ostankom prilagodila parabolo [15], ki je pri njenih podatkih zadovoljivo opisala spreminjanje variance ostanka skozi laktacijo.

$$\mathbf{R} = \text{var}(\mathbf{e}) = \sum \oplus \sigma_{e_t}^2 \quad [13]$$

$$\begin{aligned} \hat{e}_{ijkln} &= y_{ijkln} - E(y_{ijkln} | a_k, p_{kl}) = \\ &= y_{ijkln} - \sum_{m=0}^4 \hat{b}_{mi} * x_m - \hat{\text{HTD}}_j - \hat{a}_k - \hat{p}_{kl} \end{aligned} \quad [14]$$

$$\hat{\sigma}_{e_t}^2 = \text{var}(\hat{e}_{ijkln}) = 9.722 - 0.072363 * t + 0.000199 * t^2 \quad [15]$$

Matriko varianc za aditivni genetski vpliv predstavlja [16], kjer je \mathbf{A} matrika sorodstva in σ_a^2 aditivna genetska varianca. Matrika varianc za permanentno okolje [17] je diagonalna matrika z varianco za permanentno okolje (σ_p^2) kot diagonalnimi elementi.

$$\text{var}(\mathbf{a}) = \mathbf{A} * \sigma_a^2 \quad [16]$$

$$\text{var}(\mathbf{p}) = \mathbf{I} * \sigma_p^2 \quad [17]$$

Korelacija med napovedmi plemenske vrednosti iz analize mlečnosti v standardni laktaciji in plemenskimi vrednostmi s tem modelom je visoka (0.91, Ptak in Schaeffer, 1993). Oblika laktacijske krivulje je skupna skupini krav znotraj interakcije starost-sezona-regija-leto telitve. Vendar model ne upošteva, da se laktacijske krivulje krav razlikujejo po višini in obliki. Za aditivni genetski vpliv in vpliv stalnega okolja v laktaciji predpostavljamo, da sta konstantna skozi vso laktacijo, prav tako tudi varianca, ki jo pojasnjujeta. Nadalje predpostavljamo, da so med poljubnima kontrolama popolne korelacije, ne glede na oddaljenost med njima. Stalno okolje pa se lahko spreminja v času laktacije. Vpliv čreda-kontrolni dan dovoljuje, da krava zamenja čredo v času laktacije, omogoča pa tudi boljše ovrednotenje vpliva okolja. Kadar so črede majhne, bi morali interakcijo čreda-kontrolni dan uporabiti kot naključni vpliv kot Brežnik (1999). Model rutinsko uporabljajo za genetsko vrednotenje lastnosti mlečnosti v Kanadi, Estoniji, v Nemčiji za črno-belo pasmo, v Švici in na Finskem (INTERBULL, 2000).

MODEL MERITEV NA KONTROLNI DAN Z NAKLJUČNO REGRESIJO

Izpeljava modela

Do sedaj opisani modeli ne dopuščajo individualnosti produkcijskih krivulj za vsako žival, genetska vrednost živali je zgolj odstopanje v eni sami ali nekaj točkah. Smiselno rešitev predstavlja uporaba heterogenih oz. vsaki živali lastnih regresijskih koeficientov. Ker imamo informacij za vsako posamezno žival praviloma malo, vpliva ne moremo zanesljivo oceniti, pač pa lahko le napovemo, kar pomeni naključno spremenljivko v modelu. Model, ki regresijo obravnava kot naključni vpliv (Schaeffer in Dekkers, 1994; Jamrozik in Schaeffer, 1997; Jamrozik in sod., 1997a), predstavlja nadgradnjo modela s sistematsko regresijo. Izgradnjo modela meritev na kontrolni dan z naključno regresijo začnemo z eno od funkcij, ki opisujejo laktacijsko krivuljo npr. Wilminkovo funkcijo [3] (Schaeffer, 1998). V linearnem modelu [18] so A_j , B_j in C_j regresijski koeficienti, neodvisna spremenljivka (t) predstavlja stadij laktacije in ϵ_{jk} ostanek.

$$y_{jk} = A_j + B_j + C_j + \epsilon_{jk} \quad [18]$$

Vsakega od regresijskih koeficientov lahko modeliramo. Vpliv čreda-kontrolni dan (HTD) v modelih [19] do [21] je sestavljen iz vpliva črede (H) in kontrolnega dne (TD), vpliv časovni interval-starost ob telitvi-sezona telitve (TAS) pa sestavljajo vplivi časovni interval (T), starost ob telitvi (A) in sezona telitve (S). V modelih imamo še tri naključne vplive: aditivni genetski vpliv(a), vpliv skupnega (p) in začasnega okolja (e).

$$A_j = \text{HTD}_{A_j} + \text{TAS}_{A_j} + a_{A_j} + p_{A_j} + e_{A_j} \quad [19]$$

$$B_j = \text{HTD}_{B_j} + \text{TAS}_{B_j} + a_{B_j} + p_{B_j} + e_{B_j} \quad [20]$$

$$C_j = \text{HTD}_{C_j} + \text{TAS}_{C_j} + a_{C_j} + p_{C_j} + e_{C_j} \quad [21]$$

Ko vstavimo koeficiente iz (pod)modelov [19] do [21] v model [18], dobimo model [22]. Pri tem ni nujno, da uporabimo isti model za vse koeficiente, npr. HTD ni potreben za koeficienta B_j in C_j .

$$y_{jk} = (\text{HTD}_{Aj} + \text{TAS}_{Aj} + a_{Aj} + p_{Aj} + e_{Aj}) + (\text{HTD}_{Bj} + \text{TAS}_{Bj} + a_{Bj} + p_{Bj} + e_{Bj}) * t + (\text{HTD}_{Cj} + \text{TAS}_{Cj} + a_{Cj} + p_{Cj} + e_{Cj}) * \exp(-0.05 * t) + \varepsilon_{jk} \quad [22]$$

Model [22] nekoliko preuredimo in dobimo model [23], ki je zgolj eden od možnih modelov z naključno regresijo. Prav tako ni smiselno, da ocenjujemo laktacijsko krivuljo za vsak razred vpliva čreda-kontrolni dan (izraz v oklepaju v prvi vrstici v [23]), nadomestimo ga lahko z izrazom [24], ki vključuje (pod)model za laktacijske krivulje znotraj vpliva čreda-letno-sezona telitve (HZS). S tem model sestavljajo tako od časa (stadija laktacije) odvisni kot neodvisni sistematski vplivi.

$$y_{jk} = (\text{HTD}_{Aj} + \text{HTD}_{Bj} * t + \text{HTD}_{Cj} * \exp(-0.05 * t)) + (\text{TAS}_{Aj} + \text{TAS}_{Bj} * t + \text{TAS}_{Cj} * \exp(-0.05 * t)) + (a_{Aj} + a_{Bj} * t + a_{Cj} * \exp(-0.05 * t)) + (p_{Aj} + p_{Bj} * t + p_{Cj} * \exp(-0.05 * t)) + (e_{Aj} + e_{Bj} * t + e_{Cj} * \exp(-0.05 * t)) + \varepsilon_{jk} \quad [23]$$

$$\text{HTD}_{Aj} + (\text{HYS}_{Aj} + \text{HYS}_{Bj} * t + \text{HYS}_{Cj} * \exp(-0.05 * t)) \quad [24]$$

Zapis v matrični obliki za model z naključno regresijo za eno lastnost [25] sestavljajo vektor opazovanj \mathbf{y} , sistematski del $\mathbf{X}\boldsymbol{\beta}$, naključni del ($\mathbf{Z}_1\mathbf{h}$, $\mathbf{Z}_2\mathbf{a}$, $\mathbf{Z}_2\mathbf{p}$ in $\mathbf{Z}_3\mathbf{e}$) ter vektor ostankov $\boldsymbol{\varepsilon}$.

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_1\mathbf{h} + \mathbf{Z}_2\mathbf{a} + \mathbf{Z}_2\mathbf{p} + \mathbf{Z}_3\mathbf{e} + \boldsymbol{\varepsilon} \quad [25]$$

$\boldsymbol{\beta}$	HTD_{Aj} , $\{\text{TAS}_{Aj}, \text{TAS}_{Bj}, \text{TAS}_{Cj}\}$	sistematski vplivi
\mathbf{h}	$\{\text{HYS}_{Aj}, \text{HYS}_{Bj}, \text{HYS}_{Cj}\}$	naključni vplivi
\mathbf{a}	$\{a_{Aj}, a_{Bj}, a_{Cj}\}$	naključni vplivi
\mathbf{p}	$\{p_{Aj}, p_{Bj}, p_{Cj}\}$	naključni vplivi
\mathbf{e}	$\{e_{Aj}, e_{Bj}, e_{Cj}\}$	naključni vplivi

V vektorju neznanih parametrov za sistematski del modela ($\boldsymbol{\beta}$) so vpliv čreda-kontrolni dan (HTD_{Aj}) in sistematski regresijski koeficienti za vpliv časovni interval-starost ob telitvi-sezona telitve (TAS_{Aj} , TAS_{Bj} , TAS_{Cj}), vektor parametrov za naključni vpliv čreda-letno-sezona telitve (\mathbf{h}) sestavljajo naključni regresijski koeficienti HYS_{Aj} , HYS_{Bj} in HYS_{Cj} . Prav tako vektorji parametrov za aditivni genetski vpliv (\mathbf{a}), vpliva stalnega okolja (\mathbf{p}) in začasnega okolja (\mathbf{e}) vsebujejo pripadajoče nize naključnih regresijskih koeficientov ($\{a_{Aj}, a_{Bj}, a_{Cj}\}$, $\{p_{Aj}, p_{Bj}, p_{Cj}\}$, $\{e_{Aj}, e_{Bj}, e_{Cj}\}$).

Matrika dogodkov za sistematski del modela \mathbf{X} je odvisna od vplivov HTD in TAS. Za naključni del modela je struktura matrik dogodkov nekoliko drugačna kot pri mešanem modelu, ki ne vključuje naključne regresije. Vektor \mathbf{f}'_j [26] sestavljajo koeficienti, izpeljani iz neodvisne spremenljivke stadija laktacije. Vrednosti in število koeficientov so odvisni od funkcije, izbrane za opis laktacijske krivulje. Tu smo izbrali Wilminkovo funkcijo [3] in imamo zato tri koeficiente. Matrike dogodkov za naključni del modela \mathbf{Z}_1 , \mathbf{Z}_2 in \mathbf{Z}_3 ([27], [28], [29]) imajo na mestih, kjer bi imeli v navadnem mešanem modelu koeficienti vrednost 1, vektor \mathbf{f}'_j in na mestih s koeficientom z vrednostjo 0, ustrezno velik ničelni vektor ($\mathbf{0}'$).

$$\mathbf{f}'_j = [1 \quad t \quad \exp(-0.05t)] \quad [26]$$

$$\mathbf{Z}_1 = \begin{bmatrix} \mathbf{f}'_1 \\ \mathbf{f}'_2 \\ \mathbf{f}'_3 \\ \mathbf{f}'_4 \end{bmatrix}_{4 \times 3} \quad \mathbf{Z}_2 = \begin{bmatrix} \mathbf{f}'_1 & \mathbf{0}' \\ \mathbf{f}'_2 & \mathbf{0}' \\ \mathbf{0}' & \mathbf{f}'_3 \\ \mathbf{0}' & \mathbf{f}'_4 \end{bmatrix}_{4 \times 6} \quad \mathbf{Z}_3 = \begin{bmatrix} \mathbf{f}'_1 & \mathbf{0}' & \mathbf{0}' & \mathbf{0}' \\ \mathbf{0}' & \mathbf{f}'_2 & \mathbf{0}' & \mathbf{0}' \\ \mathbf{0}' & \mathbf{0}' & \mathbf{f}'_3 & \mathbf{0}' \\ \mathbf{0}' & \mathbf{0}' & \mathbf{0}' & \mathbf{f}'_4 \end{bmatrix}_{4 \times 12} \quad [27], [28], [29]$$

Pričakovana vrednost za opazovanja [30] je enaka sistematskemu delu modela. Strukturo varianc in kovarianc predstavlja enačba [31].

$$E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta} \quad [30]$$

$$\text{var}(\mathbf{y}) = \mathbf{Z}_1(\mathbf{I} \otimes \mathbf{H})\mathbf{Z}'_1 + \mathbf{Z}_2(\mathbf{A} \otimes \mathbf{G})\mathbf{Z}'_2 + \mathbf{Z}_3(\mathbf{I} \otimes \mathbf{P})\mathbf{Z}'_3 + \text{var}(\boldsymbol{\epsilon}) \quad [31]$$

Matrika $\mathbf{Z}_3(\mathbf{I} \otimes \mathbf{E})\mathbf{Z}'_3$ [32] je diagonalna in vsebuje 301 različno vrednost ($5 \leq t \leq 305$). Vsak od elementov $\mathbf{f}'_j\mathbf{E}\mathbf{f}_j$ v matriki je odvisen od dneva laktacije. Za varianco ostanka ($\text{var}(\boldsymbol{\epsilon})$) v [31] in [33] lahko uporabimo različno strukturo, matrika je lahko diagonalna ali avtokorelirana (Wade in sod., 1990). Kljub vsemu je matrika varianc in kovarianc za ostanek \mathbf{R} [33] problematična zaradi svoje strukture (Schaeffer, 1998). Kettunen in sod. (1998) ugotavljajo, da bi bilo lastnosti modela z naključno regresijo možno izboljšati tudi z upoštevanjem kovarianc za ostanek med kontrolnimi dnevi. Do enakega zaključka so prišli pri študiju strukture varianc in kovarianc za ostanek tudi Olori in sod. (1999a, 1999b).

$$\begin{aligned} \mathbf{Z}_3(\mathbf{I} \otimes \mathbf{E})\mathbf{Z}'_3 &= \begin{bmatrix} \mathbf{f}'_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{f}'_2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{f}'_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{f}'_4 \end{bmatrix} \begin{bmatrix} \mathbf{E} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{E} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{E} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{E} \end{bmatrix} \begin{bmatrix} \mathbf{f}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{f}_2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{f}_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{f}_4 \end{bmatrix} = \\ &= \begin{bmatrix} \mathbf{f}'_1\mathbf{E}\mathbf{f}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{f}'_2\mathbf{E}\mathbf{f}_2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{f}'_3\mathbf{E}\mathbf{f}_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{f}'_4\mathbf{E}\mathbf{f}_4 \end{bmatrix}_{4 \times 4} = \\ &= \sum_{j=1}^4 \oplus \mathbf{f}'_j\mathbf{E}\mathbf{f}_j \end{aligned} \quad [32]$$

$$\mathbf{R} = \mathbf{Z}_3(\mathbf{I} \otimes \mathbf{E})\mathbf{Z}'_3 + \text{var}(\boldsymbol{\epsilon}) \quad [33]$$

Sistem enačb mešanega modela [34] je pri tem trikrat večji kot pri navadni enolastnostni analizi, kar pomeni bistveno več računanja in večjo porabo pomnilniških zmogljivosti. Velikost sistema enačb (N v [35]) je namreč vsota enkrat število razredov vpliva HTD, trikrat število razredov vpliva TAS, trikrat število razredov vpliva HZS, trikrat število živali in trikrat število krav z meritvami na kontrolni dan, če izberemo za opis laktacijske krivulje funkcijo s tremi parametri. Schaeffer (1998) opozarja tudi na počasnejše konvergirane.

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}_1 & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}_2 & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}_2 \\ \mathbf{Z}'_1\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'_1\mathbf{R}^{-1}\mathbf{Z}_1 + \mathbf{I} \otimes \mathbf{H}^{-1} & \mathbf{Z}'_1\mathbf{R}^{-1}\mathbf{Z}_2 & \mathbf{Z}'_1\mathbf{R}^{-1}\mathbf{Z}_2 \\ \mathbf{Z}'_2\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'_2\mathbf{R}^{-1}\mathbf{Z}_1 & \mathbf{Z}'_2\mathbf{R}^{-1}\mathbf{Z}_2 + \mathbf{A} \otimes \mathbf{G}^{-1} & \mathbf{Z}'_2\mathbf{R}^{-1}\mathbf{Z}_2 \\ \mathbf{Z}'_2\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'_2\mathbf{R}^{-1}\mathbf{Z}_1 & \mathbf{Z}'_2\mathbf{R}^{-1}\mathbf{Z}_2 & \mathbf{Z}'_2\mathbf{R}^{-1}\mathbf{Z}_2 + \mathbf{I} \otimes \mathbf{P}^{-1} \end{bmatrix} * \begin{bmatrix} \tilde{\boldsymbol{\beta}} \\ \tilde{\mathbf{h}} \\ \tilde{\mathbf{a}} \\ \tilde{\mathbf{p}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'_1\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'_2\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'_2\mathbf{R}^{-1}\mathbf{y} \end{bmatrix} \quad [34]$$

$$N = 1 * n_{HTD} + 3 * n_{TAS} + 3 * n_{HYS} + 3 * n_{TAS} + 3 * n_a + 3 * n_p \quad [35]$$

Kot kažejo analize v Kanadi (Jamrozik in sod., 1997b), model meritev na kontrolni dan z naključno regresijo zmanjša vpliv začetih, nezaključenih laktacij na to, da so mladi biki precenjeni ali podcenjeni, vendar tega vpliva ne odstrani v celoti. V primerjavi med modelom s sistematsko regresijo in modelom z naključno regresijo se je izkazalo, da ima model z naključno regresijo vzdolž celotne laktacije manjšo varianco za ocene ostanka. Model s sistematsko regresijo ima zelo veliko varianco ocen ostanka na začetku in koncu laktacije, iz česar Jamrozik in sod. (1997a) sklepajo, da model ni najprimernejši.

Tako kot so meritve lahko funkcija neke neodvisne spremenljivke (npr. časa), so lahko tudi variance in kovariance. Modeli, ki omogočajo, da se kovariance med opazovanji postopoma spreminjajo s časom in so pri tem razlike odvisne od časovne oddaljenosti med meritvami, so modeli s kovariančnimi funkcijami. Meyer in Hill (1997) sta prikazala, da so modeli z naključno regresijo in modeli s kovariančnimi funkcijami v bistvu isto.

Napoved plemenske vrednosti za nivo in vztrajnost

Za vsako kravo z meritvami na kontrolni dan izračunamo po tri napovedi za aditivni genetski vpliv in za stalno okolje. Same po sebi te napovedi niso preveč uporabne za razvrščanje živali, ker so naključni regresijski koeficienti, vendar njihova funkcija nudi uporabno informacijo za odločitve v selekciji (Jamrozik in sod., 1997b). Napoved plemenske vrednosti lahko potem računamo za določen kontrolni dan ali pa za skupno količino v standardni laktaciji [36].

$$\begin{aligned} NPV_{305} &= \sum_{t=1}^{305} (\hat{A} + \hat{B} * t + \hat{C} * \exp(-0.05 * t)) = \\ &= \hat{A} * \left(\sum_{t=1}^{305} 1 \right) + \hat{B} * \left(\sum_{t=1}^{305} t \right) + \hat{C} * \left(\sum_{t=1}^{305} \exp(-0.05 * t) \right) = \\ &= 305 * \hat{A} + 46665 * \hat{B} + 19.504162 * \hat{C} \end{aligned} \quad [36]$$

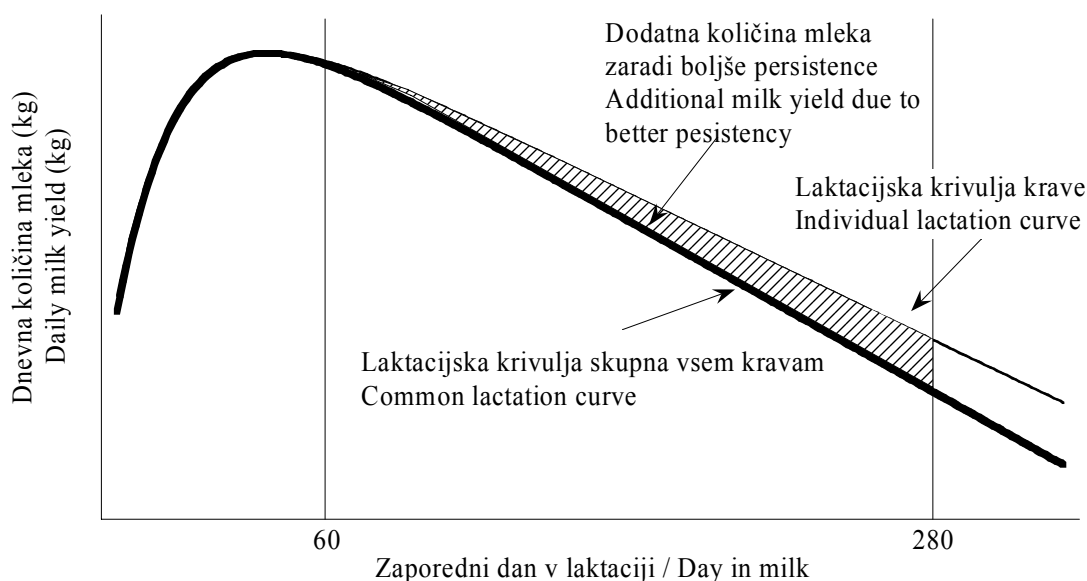
Genetska varianca za količino na vsak kontrolni dan je v modelu z naključno regresijo funkcija neodvisnih spremenljivk (Jamrozik in Schaeffer, 1997). Oceno genetske variance za nivo na kontrolni dan t izračunamo po [37], \mathbf{f}'_t je vektor s koeficienti od neodvisnih spremenljivk [38] in $\hat{\mathbf{G}}$ je matrika ocen genetskih parametrov za naključne regresijske koeficiente. Oceno genetske kovariance med količinama na kontrolna dneva t_1 in t_2 podaja enačba [39].

$$\hat{\sigma}_{g_t}^2 = \mathbf{f}'_t \hat{\mathbf{G}}_{3 \times 3} \mathbf{f}_t, \quad [37]$$

$$\mathbf{f}'_t = [1 \quad t \quad \exp(-0.05 * t)] \quad [38]$$

$$\hat{\sigma}_{g_{t_1 t_2}}^2 = \mathbf{f}'_{t_1} \hat{\mathbf{G}} \mathbf{f}_{t_2} \quad [39]$$

Model z naključno regresijo omogoča, da napovemo plemensko vrednost živali tudi za mlečno vztrajnost. Za vztrajnost uporabljajo v svetu različne načine ovrednotenja (za pregled glej npr. Gengler, 1996). Kanadčani (Jamrozik in sod., 1997b) jo računajo po formuli [40] kot količino mleka (pridobljeno ali izgubljeno) med 60-im in 280-im dnevom glede na povprečno žival, ki ima na 60. dan enako dnevno količino mleka ter predstavlja aproksimacijo površine trikotnika (slika 5). Če v formulo [40] vstavimo izraza (izračuna) [41] in [42], dobimo poenostavljeno obliko formule za vztrajnost [43] z napovedmi regresijskih koeficientov \hat{A} , \hat{B} in \hat{C} .



Slika 5. Vztrajnost, ocenjena s površino trikotnika (osenčeno) med skupno in individualno laktacijsko krivuljo (po Jamrozik in sod., 1997b).

Figure 5. Persistency estimated as the area of triangle (shaded) between common and individual lactation curve (after Jamrozik *et al.*, 1997b).

$$NPV_p = \sum_{t=60}^{280} (NPV_t - NPV_{60}) = \sum_{t=60}^{280} NPV_t - 221 * NPV_{60} \quad [40]$$

$$\sum_{t=60}^{280} NPV_t = \hat{A} * \left(\sum_{t=60}^{280} 1 \right) + \hat{B} * \left(\sum_{t=60}^{280} t \right) + \hat{C} * \left(\sum_{t=60}^{280} \exp(-0.05 * t) \right) =$$

$$= 221 * \hat{A} + 37570 * \hat{B} + 1.0208 * \hat{C} \quad [41]$$

$$NPV_{60} = \hat{A} * 1 + \hat{B} * t + \hat{C} * \exp(-0.05 * t) = 1 * \hat{A} + 60 * \hat{B} + 0.0498 * \hat{C} \quad [42]$$

$$NPV_p = 0 * \hat{A} + 24310 * \hat{B} - 9.9821 * \hat{C} \quad [43]$$

Genetska varianca za vztrajnost [44] je odvisna od varianc za parametre laktacijske krivulje in kovarianc med njimi.

$$\begin{aligned} var(NPV_p) &= var(0 * \hat{A} + 24310 * \hat{B} - 9.9821 * \hat{C}) = \\ &= 24310^2 * var(\hat{B}) + 9.9821^2 * var(\hat{C}) + 2 * 24310 * 9.9821 * cov(\hat{B}, \hat{C}) \end{aligned} \quad [44]$$

ZAKLJUČKI

Meritve lastnosti mlečnosti na kontrolni dan sodijo v razred ponavljajočih meritev. Izmerjene so v časovnem sosledju in so med seboj odvisne. Med lastnosti s ponavljajočimi meritvami z vključeno časovno komponento sodijo tudi rast, zauživanje in izkoriščanje krme itn.

Model bolje opiše biološko plat proizvodnih funkcij živali. Naključna regresija omogoča genetsko ovrednotenje živali tako za nivo v celotni laktaciji, delni laktaciji ali samo na določen dan kot za vztrajnost pri lastnostih mlečnosti. Pri rasti dobimo napovedi plemenskih vrednosti za

maso ob izbranih starostih kot hitrost rasti pri pitanju. S tem je možna tudi selekcija na te lastnosti. Omogoča tudi pogostejše izračune napovedi plemenskih vrednosti, npr. mesečne obračune. Prek rednih mesečnih poročil lahko pomaga rejcem pri uravnavanju reje.

Prav tako model z naključno regresijo dovoljuje različno število opazovanj na žival, neenakomerno porazdelitev meritev, različno zanesljivost izmerjenih podatkov. V laktaciji nekatere kontrole mlečnosti opravi kontrolor, druge pa lahko opravi rejec sam.

Modeli z naključno regresijo zahtevajo v primerjavi z modeli s sistematsko regresijo zmogljivejše računalnike, ker namesto ene plemenske vrednosti napovemo plemenske vrednosti za več regresijskih koeficientov na žival. V primerjavi z analizo laktacijskih zaključkov pa so zahteve večje še za mnogokratnik števila kontrolnih dni na laktacijo.

Prednost naključne regresije v primerjavi z večlastnostnim modelom z intervali je v boljšem načinu ocenjevanja parametrov, saj izkoristi prednosti, ki jih v modelu nudi uporaba regresije.

Modeli z naključno regresijo imajo kompleksno strukturo varianc in kovarianc za ostanek, ki je tudi odvisna od časa. Nadaljnje raziskave bodo potrebne pri izboru najprimernejših funkcij za regresijo in kovariančne funkcije ter iskanju primerne interpretacije.

Za genetsko vrednotenje živali z opisanimi modeli bodo potrebni še hitrejši numerični algoritmi, pa tudi hiter napredek v razvoju računalnikov bo prispeval pomemben delež.

Metoda se bo vse več uporabljala pri modeliranju laktacijskih krivulj pri mlečnem govedu in drobnici, rastnih krivulj pri drugih vrstah, pri katerih je glavni cilj prireja mesa, ter preživitvenih krivulj pri procesih, povezanih z mikroorganizmi (prebava, mlekarstvo itn.).

SUMMARY

Random regression (RR) and covariance function (CF) models are not new. At first, they were used in analyses of small experimental data sets, nowadays they become more and more attractive both from theoretical point of view in statistics and application in animal breeding. Their feasibility is a consequence of hardware and software development rather than development in theoretical background.

The aim of this paper is to explain random regression models where test-day records in dairy-cattle serve as example, together with discussion about existent methods in genetic evaluation for dairy traits.

The RR models deal with traits measured sequentially over time or place, called also longitudinal data or test day measurements. Repeated measurements may have different expected value (mean) as well as variance. Because they are not independent, changes can be described by some functions with time as an independent variable. Such traits in animal production are body weight and feed intake in growing animals, test-day records for milk yield, milk components and somatic cell count. Common examples of functions that describe changes in trait over time are growth curves and lactation curves, so called production functions.

The simplest method for prediction of breeding values is analysis of 305-day lactation yield and it is still widely used. The method has many deficiencies. Test interval method accounts for residual covariance structure among measurements within lactation, however this approach has extensive computational demands. Fixed regression models adjust for differences caused by stage of lactation. Different regression functions like Wilmink, Ali - Schaeffer, and Guo - Swalve are used for modelling lactation curves. Common shape is caused by fixed effects like breed, parity, and season. Furthermore, each animal has its own - specific curve influenced by additive genetic effect, but it is not accounted for in model with fixed regression only.

Adding random regression in the model covers variation in shape due to common or permanent environment as well as genetic effects. There are many advantages of use RR and CF models over conventional repeatability models or multitrait models. They describe trait at all

rather than at a finite number of points and are effective in studying changes in trait over time. Such longitudinal studies increase statistical power. Moreover, models allow the covariance between measurements to change gradually over time. Individual curves estimated by RR allow more biological approach. Thus, integrating the area below lactation curve gives total milk yield in lactation, while the first derivative predicts persistency. Animals may be evaluated before lactation is finished allowing selection of breeding animals before mating season, which is important especially in small ruminants. In dairy farming, more frequent evaluation runs of test-day models could offer management aid to farmers.

VIRI

- Ali, T.E./ Schaeffer, L.R. Accounting for covariances among test day milk yields in dairy cows. *Can. J. Anim. Sci.*, 67(1987), 637-644.
- Andersen, S./ Pedersen, B. Growth and food intake curves for group-housed gilts and castrated male pigs. *Anim. Sci.*, 63(1996), 457-464.
- Andonov, S. Estimation of genetic parameters for milk production in dairy goats using the first records of lactation. *Mag. delo, Domžale, Univ. v Ljubljani*, 1994, 68 s.
- Brežnik, S. Ocenjevanje parametrov disperzije za lastnosti mlečnosti pri ovcah. *Mag. delo, Domžale, Univ. v Ljubljani*, 1999, 73 s.
- Gengler, N. Persistency of lactation yields: a review. V: *Proceedings of the International Workshop on Genetic Improvement of Functional Traits in Cattle, Gembloux, 1996-01-21/23. INTERBULL Bul. No. 12. Uppsala, Interbull Bull Evaluation Service, 1996, 87-96.*
- Guo, Z./ Swalve, H.H. Modelling of the lactation curve as a sub-model in the evaluation of test day records. V: *Interbull Meeting, Praga, 1995-09-7/8. INTERBULL Bul. No. 11. Uppsala, Interbull Bull Evaluation Service, 1995, 4 s.*
- Henderson, C.R. jr. Analysis of the covariance in the mixed model: Higher-level, nonhomogeneous, and random regressions. *Biometrics*, 38(1982), 623-640.
- ICAR. Constitution - Rules for granting and for the use of Special Stamp International Agreement of recording practices, Section 2.1, Rim, International Committee for Animal Recording (ICAR), 1995, Appendix A in C.
- INTERBULL, 2000., URL: www-interbull.slu.se/eval/table1-003.html (2000-09-28).
- Jamrozik, J./ Kistemaker, G.J./ Dekkers, J.C.M./ Schaeffer, L.R. Comparison of possible covariates for use in a random regression model for analyses of test day yields. *J. Dairy Sci.*, 80(1997a), 2550-2556.
- Jamrozik, J./ Schaeffer, L.R. Estimates of genetic parameters for a test day model with random regressions for yield traits of first lactation Holsteins. *J. Dairy Sci.*, 80(1997), 762-770.
- Jamrozik, J./ Schaeffer, L.R./ Dekkers, J.C.M. Genetic evaluation of dairy cattle using test day yields and random regression model. *J. Dairy Sci.*, 80(1997b), 1217-1226.
- Jones, H.E./ White, I.M.S./ Brotherstone, S. Genetic evaluation of Holstein Friesian sires for daughter condition-score changes using random regression model. *Anim. Sci.*, 68(1999), 467-475.
- Kettunen, A./ Mäntysaari, E.A./ Strandén, I./ Pösö, J./ Lindauer, M. Estimation of genetic parameters for first lactation test day milk production using random regression models. V: *Proceedings of the 6th World Congress on Genetics Applied to Livestock Production, Armidale, 1998-01-11/16. Armidale, University of New England, 23(1998), 307-310.*
- Kirkpatrick, M./ Heckman, N. A quantitative genetic model for growth, shape, reaction norms, and other infinite-dimensional characters. *J. Math. Biol.*, 27(1989), 429-450.
- Kirkpatrick, M./ Lofsvold, D./ Bulmer, M. Analysis of the inheritance, selection and evolution of growth trajectories. *Genetics*, 124(1990), 979-993.
- Kirkpatrick, M./ Hill, W.G./ Thompson, R. Estimating the covariance structure of traits during growth and ageing, illustrated with lactation in dairy cattle. *Genet. Res.*, 64(1994), 57-69.
- Koenen, E.C.P./ Veerkamp, R.F. Genetic covariance functions for live weight, condition score, and dry-matter intake measured at different lactation stages of Holstein Friesian heifers. *Livest. Prod. Sci.*, 57(1998), 67-77.
- Laird, N.M./ Ware, J.H. Random-effects models for longitudinal data. *Biometrics* 38(1982), 963-974.
- Meyer, K. Estimates of genetic and phenotypic covariance functions for postweaning growth and mature weight of beef cows. *J. Anim. Breed. Genet.*, 116(1999), 181-205.
- Meyer, K. Random regressions to model phenotypic variation in monthly weights of Australian beef cows. *Livest. Prod. Sci.*, 65(2000), 19-38.
- Meyer, K./ Graser, H.U./ Hammond, K. Estimation of genetic parameters for first lactation test day production of Australian black and white cows. *Livest. Prod. Sci.*, 21(1989), 177-199.

- Meyer, K./ Hill, W.G. Estimation of genetic and phenotypic covariance functions for longitudinal or 'repeated' records by restricted maximum likelihood. *Livest. Prod. Sci.*, 47(1997), 185-200.
- Olori, V.E./ Hill, W.G./ Brotherstone, S. The structure of the residual error variance of test day milk yield in random regression models. V: Proceedings of the Computational Cattle Breeding '99 Workshop, Tuusula 1999-03-18/20. INTERBULL Bul. No. 20. Uppsala, Interbull Bull Evaluation Service, 1999a, 103-108.
- Olori, V.E./ Hill, W.G./ McGuirk, B.J./ Brotherstone, S. Estimating variance components for test day milk records by maximum likelihood with random regression animal model. *Livest. Prod. Sci.*, 61(1999b), 53-63.
- Pander, B.L./ Hill, W.G./ Thompson, R. Genetic parameters for test day records of British Holstein-Friesian heifers. *Anim. Prod.*, 55(1992), 11-21.
- Pander, B.L./ Thompson, R./ Hill, W.G. The effect of increasing interval between recordings on genetic parameters for test day yields of British Holstein-Friesian heifers. *Anim. Prod.*, 56(1993), 159-164.
- Potočnik, K. Napovedi plemenskih vrednosti za lastnosti mlečnosti pri govedu na podlagi dnevne kontrole. Mag. delo, Domžale, Univ. v Ljubljani, 1999, 90 s.
- Ptak, E./ Schaeffer, L.R. Use of test day yields for genetic evaluation of dairy sires and cows. *Livest. Prod. Sci.*, 34(1993), 23-34.
- Reents, R./ Dekkers, J.C.M./ Schaeffer, L.R. Genetic evaluation for somatic cell score with a test day model for multiple lactations. *J. Dairy Sci.*, 78(1995), 2858-2870.
- Schaeffer, L.R., Dekkers, J.C.M. Random regressions in animal models for test-day production in dairy cattle. V: 5th World Congress on Genetics Applied to Livestock Production, Guelph, 1994-08-7/12. Guelph, University of Guelph, 18(1994), 443-446.
- Schaeffer, L.R./ Jamrozik, J. Multiple-trait prediction of lactation yields for dairy cows. *J. Dairy Sci.*, 79(1996), 2044-2055.
- Schaeffer, L.R. Random Regression Models. PhD Course. Wageningen, 1998-07-13/17. Wageningen, Wageningen Institute of Animal Sciences, 1998, 147 s.
- Thompson, R./ Hill, W.G. Univariate REML analysis for multivariate data with animal model. V: 4th World Congress on Genetics Applied to Livestock Production, Edinburgh, 1990-07-23/27. Edinburgh, The East of Scotland College of Agriculture, 12(1990), 484-487.
- USDA-DHIA, 2000, URL: aipl.arsusda.gov/main/ (2000-09-28).
- VanRaden, P.M./ Wiggans, G.R. Derivation, calculation, and use of national animal model information. *J. Dairy Sci.* 74(1991), 2737-2742.
- Veerkamp, R.F./ Thompson, R. A covariance function for feed intake, live weight, and milk yield estimated using a random regression model. *J. Dairy Sci.*, 82(1999), 1565-1573.
- Wade, K.M./ Quass, R.L./ Van Vleck, L.D. Mixed linear models with autoregressive error structure. V: 4th World Congress on Genetics Applied to Livestock Production, Edinburgh, 1990-07-23/27. Edinburgh, The East of Scotland College of Agriculture, 13(1990), 508-511.
- Werf, van der J.H.J./ Goddard, M.E./ Meyer, K. The use of covariance functions and random regressions for genetic evaluation of milk production based on test day records. *J. Dairy Sci.*, 81(1998), 3300-3308.
- Wilmink, J.B.M. Adjustment for test-day milk, fat and protein yields for age, season and stage of lactation. *Livest. Prod. Sci.*, 16(1987), 335-348.
- Wood, P.D.P. Algebraic model of the lactation curve in cattle. *Nature*, 216(1967), 164-165.