



Výzkumný ústav živočišné výroby, v.v.i., Praha Uhřetěves

Jiří Bauer
Luboš Vostrý
Josef Přibyl

Odhad spolehlivosti jednokrokových genomických plemenných hodnot pro dojený skot





CERTIFIKOVANÁ METODIKA

Odhad spolehlivosti jednokrokových genomických plemenných hodnot pro dojený skot

Autoři

Ing. Jiří Bauer, Ph.D
doc. Ing. Luboš Vostrý, Ph.D.
prof. Ing. Josef Příbyl, DrSc.

Oponenti

prof. Ing. Jindřich Čítek, CSc.
Jihočeská univerzita v Českých Budějovicích

Ing. Zdeňka Majzlíková
Česká státní plemenářská inspekce, Praha

Metodika byla vypracována v rámci řešení výzkumného projektu NAZV QI111A167.

ISBN 978-80-7403-120-5

Obsah

| | | |
|---------|--|----|
| I. | Cíl metodiky..... | 3 |
| II. | Vlastní popis metodiky | 3 |
| II.1. | Úvod | 3 |
| II.2. | Datové soubory a jejich příprava..... | 3 |
| II.2.1. | Struktura datových souborů..... | 4 |
| II.2.2. | Příprava souborů..... | 5 |
| II.3. | Pracovní postup | 5 |
| II.3.1. | Parametry potřebné pro průběh výpočtu | 6 |
| II.3.2. | Vlastní výpočet | 6 |
| II.4. | Ověření pracovních postupů a algoritmy | 12 |
| III. | Srovnání „novosti postupů“ | 12 |
| IV. | Popis uplatnění Certifikované metodiky..... | 12 |
| V. | Ekonomické aspekty | 12 |
| VI. | Seznam použité související literatury | 13 |
| VII. | Seznam publikací, které předcházely metodice | 14 |
| VIII. | Přílohy..... | 15 |

I. Cíl metodiky

Cílem metodiky je shrnout postup odhadu spolehlivosti běžných a jednokrokových genomických (ssGBLUP) plemenných hodnot pro množství nadojeného mléka skotu, který bude nadále využíván v rutinním provozu.

II. Vlastní popis metodiky

II.1. Úvod

Základem pro účinné šlechtění je řízení plemenitby podle genetického hodnocení, které lze vyjádřit pomocí plemenných hodnot. Běžně používanou součástí šlechtitelských programů je předpověď plemenné hodnoty (PH) metodou BLUP – animal model. Díky vědeckému pokroku lze využít k předpovědi plemenných hodnot i genomické informace, jedná se pak o předpověď genomické plemenné hodnoty (GEPH). Existují obecně dva přístupy k předpovědi genomické plemenné hodnoty. Vícekrokový postup (VanRaden, 2008) a jednokrokový postup (Misztal et al., 2009). Zatímco ve vícekových metodách jsou hodnoceni pouze genotypovaní jedinci, v jednokrokovém přístupu je možné hodnotit pomocí předpovědi plemenných hodnot celou populaci.

Důležitou informací pro chovatele je spolehlivost běžné i genomické plemenné hodnoty. Ta je vyjádřena druhou mocninou korelace mezi skutečnou PH a předpovídanou PH (Jakubec et al., 2010), tj. do jaké míry vypovídá předpovídaná plemenná hodnota o opravdovém genetickém založení zvířete. Spolehlivost běžné plemenné hodnoty může být určena pomocí inverze levé strany soustavy rovnic v metodě BLUP (VanRaden, 2008). Tento přístup ovšem není většinou možné uskutečnit z důvodu velkého rozsahu hodnocených populací, což vede k obrovským požadavkům na výpočetní techniku. Spolehlivost PH se tudíž odhaduje pomocí jiných postupů (např. Tier et Meyer, 2004 nebo Misztal et al., 1993). Spolehlivost GEPH se odhaduje metodami uváděnými Szydrou et al., 2011 pro vícekový postup a Misztalem et al., 2013 pro jednokrokový postup.

Postup pro rutinní předpovědi GEPH pomocí jednokrokové metody byl popsán v metodice „Úprava soustav rovnic pro předpověď genomické plemenné hodnoty (GEPH) jednokrokovou metodou“ v roce 2011. Předpověď genomických plemenných hodnot byla dále ověřována na populaci holštýnského skotu chovaného v České republice a výsledky jsou uvedeny v pracích Příbyl et al., 2012 a 2013.

Především pro potřeby celosvětového hodnocení je třeba spolu s plemennými hodnotami uvádět i správné spolehlivosti, aby nedocházelo k znevýhodňování zvířat z České republiky. Ke stanovení plemenných hodnot patří proto i navržení postupu pro odhad spolehlivosti předpovězených běžných a genomických plemenných hodnot, který bude používán pro pravidelné hodnocení dojeného skotu v České republice.

II.2. Datové soubory a jejich příprava

Pro odhad spolehlivosti genomických plemenných hodnot jsou zapotřebí údaje z kontroly užítkovosti, údaje o původech všech hodnocených zvířat z ústřední evidence skotu a předem

připravené inverze rodokmenové („ A_{22}^{-1} “) a genomické („ G^{-1} “) matice příbuznosti, obojí u genotypovaných zvířat. Dále seznam genotypovaných zvířat.

Postup vytvoření matic „ A_{22} “ a „ G “ je popsán v pracích Forni et al., 2011 a Vitezica et al., 2011. Inverze rodokmenové a genomické matice zvířat musí odpovídat použitým maticím v předpovědi genomických plemenných hodnot a celá příprava pro účely pravidelné předpovědi GEPH (a tudíž i spolehlivosti GEPH) byla popsána v metodice Příbyl et al., 2011.

Pro předpovědi GEPH lze využít jako další vstupní informace i plemenné hodnoty býků z údajů organizace Interbull. Jedná se o mezinárodní plemenné hodnoty býků z výpočtu MACE a převedené na poměry České republiky. V tomto případě jsou pak pro odhady spolehlivosti GEPH důležité spolehlivosti interbullových plemenných hodnot.

Pro účel odhadu spolehlivosti GEPH jednkrokovou metodou jsou předávány soubory s následující strukturou:

II.2.1. Struktura datových souborů

II.2.1.1. Soubor se skupinami vrstevníků dle kontroly užítkovosti

- registrační číslo (kód) zvířete
- číslo (kód) skupiny vrstevníků („stáda-roku-období“- zkráceně SRO, nebo dne kontroly - TD) daného záznamu v kontrole užítkovosti, tj. zvířat, která byla měřena ve stejný den na stejném místě a stejného chovu

II.2.1.2. Soubor původů zvířat pocházející z ústřední evidence

- číslo (kód) zvířete
- číslo (kód) otce
- číslo (kód) matky

Soubor zahrnuje záznamy původu zvířat z kontroly užítkovosti a dále předem stanoveného počtu generací jejich předků.

II.2.1.3. Soubory obsahující příbuzenské matice

Tyto soubory obsahují informace jen o genotypovaných zvířatech.

matice A_{22}^{-1} :

- číslo (kód) prvního zvířete
- číslo (kód) druhého zvířete
- hodnota inverzního prvku matice vzájemné rodokmenové příbuznosti

Jedná se o část matice „ A “, tj. rodokmenové příbuzenské matice pro všechna zvířata v hodnocení.

matice G^{-1} :

- číslo (kód) prvního zvířete
- číslo (kód) druhého zvířete
- hodnota inverzního prvku matice vzájemné genomické příbuznosti

II.2.1.4. Soubor obsahující označení genotypovaných jedinců

-číslo (kód) zvířete

II.2.1.5. Soubory obsahující údaje o býcích z organizace Interbull

-číslo (kód) zvířete

-efektivní počet případů vycházejících z hodnoty spolehlivosti předpovězené Interbullové PH, příklad výpočtu ze spolehlivosti Interbullové PH je uveden ve II.3.3.1.

Pro odhad spolehlivosti PH a GEPH jsou vyžadovány údaje o původu těchto zvířat ve stejném formátu jaký je uveden pro zvířata s užitkovostmi v bodě II.2.1.2.

II.2.2. Příprava souborů

II.2.2.1. Příprava datového souboru užitkovosti

Pro vlastní odhady spolehlivosti plemenných hodnot v prvním kroku metody je nutné datový soubor upravit, přečíslovat jedince a efekt skupiny vrstevníků (SRO). Pro zajištění správného chodu odhadu spolehlivosti PH i GEPH je nutné mít jedince (včetně rodokmenu) přečíslované postupně od 1 až do maximálního počtu zvířat, přičemž jedinci označující genetické skupiny by měli mít nejlépe nejvyšší čísla. Stejným způsobem se přečíslovávají skupiny vrstevníků. V případě výskytu více záznamů u stejného jedince (tj. při opakovaných užitkovostech) je každý záznam uveden v samostatném řádku.

II.2.2.2. Příprava rodokmenového souboru

Rodokmenový soubor ve výše uvedeném formátu je třeba přečíslovat. Jedná se o čísla od 1 do maximálního počtu zvířat, přičemž čísla jednotlivých zvířat musí vzájemně odpovídat ve sloupcích jedince, otce a matky, začlenění do SRO a označení v maticích A_{22} a G . Zároveň musí platit shoda s datovým souborem z kontroly užitkovosti. Každý jedinec uvedený na jakékoliv pozici v rodokmenu musí mít svůj vlastní řádek, neznámý rodič se uvádí jako hodnota 0. I genetické skupiny musí mít svůj vlastní řádek a hodnoty jejich obou rodičů jsou rovny 0. Přečíslování nemusí být závislé na stáří zvířete či příbuzenských vazbách, tj. není třeba, aby rodiči jedince měli nižší číslo než potomek.

II.2.2.3. Příprava souborů matic příbuzností

Čísla všech jedinců v souboru musí odpovídat číslům jedinců v připravených souborech užitkovosti a rodokmenu.

II.2.2.4. Příprava souboru označení genotypovaných jedinců

Čísla všech jedinců v souboru musí odpovídat číslům jedinců v připravených souborech užitkovosti a rodokmenu.

II.2.2.5. Příprava souboru „Interbull“

Identifikace všech jedinců v souboru musí odpovídat číslům jedinců v připravených souborech užitkovosti a rodokmenu.

II.3. Pracovní postup

Vlastní postup odhadu spolehlivosti GEPH je založen na metodách Misztal et al., 1993 a Misztal et al., 2013 a skládá se ze tří hlavních částí:

- I. odhad spolehlivosti běžné PH u celé populace
- II. odhad spolehlivosti GEPH pouze u genotypovaných zvířat
- III. zahrnutí změny spolehlivosti plemenných hodnot genotypovaných zvířat do spolehlivosti běžné PH negenotypovaných příbuzných zvířat v celé populaci

Dále jsou popsány kromě výše uvedených bodů také úpravy pro použití více záznamů na jedno zvíře, využití plemenných hodnot dodaných organizací Interbull do předpovědi GEPH a tudíž i do odhadu spolehlivosti GEPH a úprava pro zohlednění příbuzenských vazeb při odhadu spolehlivosti PH.

II.3.1. Parametry potřebné pro průběh výpočtu

Koeficient „ α “ – podíl reziduálního rozptylu vůči aditivně genetickému rozptylu; lze určit i z dědivosti („ h^2 “) pomocí vzorce $\alpha = (1-h^2)/h^2$. Pokud se při hodnocení zvířat vyskytují opakovaná pozorování, pak $\alpha = (1 - r_{op})/h^2$, kde r_{op} je index opakovatelnosti.

Celkový počet jedinců v rodokmenu (odpovídá nejvyššímu použitému číslu)

Celkový počet řádků v souboru kontroly užitkovosti - počet pozorování

Celkový počet skupin SRO v souboru kontroly užitkovosti (odpovídá nejvyššímu použitému číslu)

Celkový počet genotypovaných zvířat v souboru seznamu genotypovaných zvířat

Celkový počet zvířat v souboru „Interbull“

Koeficient „ τ “ – podíl reziduálního rozptylu vůči rozptylu trvalého prostředí – použito jen pokud mají jedinci více pozorování

II.3.2. Vlastní výpočet

II.3.2.1. Část I. - Odhad spolehlivosti běžné PH u celé populace

Postup vychází z práce Misztal et al., 1993 a odhaduje spolehlivost běžné PH pomocí efektivního počtu případů vyplývajících ze záznamů kontroly užitkovosti a z údajů o příbuzných jedincích v rámci běžné metody BLUP. Jedná se o iterativně založený postup, kde pro jedince i platí, že spolehlivost r_{irp}^2 pocházející z údajů kontroly užitkovosti a z údajů příbuzných jedinců je (Misztal et al., 1993):

$$r_{irp}^2 = \frac{d_{irp}}{d_{irp} + \alpha}$$

kde α je podíl reziduálního rozptylu vůči aditivně genetickému rozptylu a d_{irp} je efektivní počet případů jedince i při započítání údajů z kontroly užitkovosti (d_{ir}) a příbuzenských vztahů (d_{ijp}) k jedincům j , tj.:

$$d_{irp} = d_{ir} + \sum_j d_{ijp}$$

d_{ir} lze vypočítat následujícím způsobem:

$$d_{i_r} = \sum_k \left(1 - \frac{1}{n_{ik}} \right)$$

kde n_{ik} odpovídá počtu záznamů ve stádě, ročníku a sezóně (SRO) dané podtřídy k .

Započítání příspěvků příbuzenských vztahů s jedinci j k efektivnímu počtu případů zvířete i se provádí pomocí iterativního postupu, kdy jsou postupně přičítány důsledky všech vztahů mezi hodnocenými jedinci na spolehlivost plemenných hodnot.

Příbuznost k vrstevníkům

Odhad spolehlivosti PH odpovídá pouze v případě, že se jedinec i setká pouze s malým počtem (polo)sourozenců ve své skupině vrstevníků (SRO). V jiném případě budou hodnoty d_{i_r} jednotlivců nadhodnocené.

To lze ošetřit pomocí snížení příspěvku k efektivnímu počtu případů jedince ze skupiny vrstevníků dle vah odpovídajících podílu sourozenců ve skupině vrstevníků (Tier et Meyer, 2004):

$$w_{ik} = \sqrt{\frac{n_k - q_{ik}}{n_{ik}}}$$

při

w_{ik} odpovídající váze u jedince i pro efektivní počet případů ve skupině k ,

n_k celkovém počtu záznamů ve skupině k ,

q_{ik} počtu sourozenců jedince i ve skupině vrstevníků k .

Hodnota w_m se pak používá změnou v původním vzorci výpočtu d_{i_r} , a to následujícím způsobem:

$$d_{i_r} = \sum_k \left(1 - \frac{1}{n_{ik} \cdot w_{ik}} \right)$$

Tento vzorec v takovém případě plně nahrazuje vzorec uvedený výše nad částí „Příbuznost k vrstevníkům“.

Vícečetná pozorování

V případě opakovaných pozorování stejného jedince je třeba d_{i_r} před vstupem do iterací snížit. To lze provést pomocí vzorce (Misztal et al., 1991):

$$d_{i_r}^{\#} = \frac{\tau \cdot d_{i_r}}{\tau + d_{i_r}}$$

kde d_{i_r} je efektivní počet případů z vlastních záznamů zvířete i ,

$d_{i_r}^{\#}$ nová hodnota vstupující do první iterace v metodě,

a τ odpovídá:

$$\tau = \frac{\delta_e^2}{\delta_{tp}^2}$$

při δ_e^2 reziduálním rozptylu,

δ_{tp}^2 rozptylu trvalého prostředí.

Údaje „Interbull“

Začlenění případných údajů plemenných hodnot býků dle Interbullu je provedeno následujícím způsobem: Hodnota efektivního počtu případů vyplývající obecně ze všech dalších zdrojů informací je po určení následujícím vzorcem přičtena k původnímu d_i :

$$d_i^b = \sum_{p=1}^N \left(w_{ip} - \frac{w_{ip}^2}{\sum_{j=1}^{n_p} w_{jp}} \right)$$

d_i^b je efektivní počet případů vyplývající z dalších zdrojů informací pro zvíře i ,
 w_{ip} hodnota váhy údajů konkrétního zvířete i ve skupině vrstevníků p ,
 n_p je počet zvířat v dané p skupině vrstevníků,
 N je počet skupin vrstevníků p .

Váhy zvířat i jsou určeny dle spolehlivosti PH býků uváděných organizací Interbull (r_i^2) a to pomocí vzorce:

$$w_{ip} = \alpha \frac{r_i^2}{1 - r_i^2}$$

V provedených odhadech spolehlivosti PH a GEPH jsou brány údaje Interbullu jako jeden rovnocenný zdroj informací, tudíž jako jedna skupina vrstevníků. Pomocí tohoto postupu lze přidávat informace o jedinci z více zdrojů za předpokladu, že informace z těchto zdrojů nemají shodné údaje. Přičtení efektivního počtu případů vyplývajícího z údajů organizace Interbull je provedeno až po případném snížení efektivního počtu případů z důvodu vícečetných pozorování.

Iterační cyklus

Základním vstupem do iterací je vektor efektivního počtu případů vyplývajícího ze záznamů jedinců, tj. d_i . Tyto hodnoty jsou postupně navyšovány o příspěvky d_{ip} následujícím cyklem:

(začátek cyklu)

1. pro každý příbuzenský vztah jedince i s rodiči o_i a m_i je vypočítán přírůstek v množství informací u všech zúčastněných (tj. g_{ii} pro jedince i , g_{io_i} pro otce o_i a g_{im_i} pro matku m_i).

To je provedeno pomocí řešení následující nelineární soustavy rovnic:

$$\left[\mathbf{A}_{3 \times 3}^{-1} + \begin{bmatrix} d_{i_r}^* - g_{ii} & 0 & 0 \\ 0 & d_{o_{i_r}}^* - g_{io_i} & 0 \\ 0 & 0 & d_{m_{i_r}}^* - g_{im_i} \end{bmatrix} \right]^{-1} = \begin{bmatrix} \frac{1}{\alpha + d_{i_r}^*} & \cdot & \cdot \\ \cdot & \frac{1}{\alpha + d_{o_{i_r}}^*} & \cdot \\ \cdot & \cdot & \frac{1}{\alpha + d_{m_{i_r}}^*} \end{bmatrix}$$

kde tečky označují pro výpočet nevýznamné prvky a

* označuje aktuální hodnotu v daném probíhajícím cyklu.

V tomto řešení soustavy rovnic je nutné, aby počáteční množství informací jedince i a jeho rodičů bylo nezáporné, neboli:

pokud $d_{i_r}^* - g_{ii} < 0$ pak je stanoveno $d_{i_r}^* - g_{ii} = 0$,

což platí stejným způsobem i pro $d_{o_{i_r}}^* - g_{io_i}$ a $d_{m_{i_r}}^* - g_{im_i}$.

$\mathbf{A}_{3 \times 3}^{-1}$ je určeno jako 3 x 3 matice příbuznosti mezi jedincem i a jeho rodiči:

$$\mathbf{A}_{3 \times 3}^{-1} = \begin{bmatrix} 2 \cdot \alpha & -\alpha & -\alpha \\ -\alpha & 0,5 \cdot \alpha & 0,5 \cdot \alpha \\ -\alpha & 0,5 \cdot \alpha & 0,5 \cdot \alpha \end{bmatrix}$$

2. přírůstek v množství informací je přičten k původním hodnotám, tj.

$$d_{i_r}^{*+1} = d_{i_r}^* + g_{ii}$$

$$d_{o_{i_r}}^{*+1} = d_{o_{i_r}}^* + g_{io_i}$$

$$d_{m_{i_r}}^{*+1} = d_{m_{i_r}}^* + g_{im_i}$$

3. následuje krok, kdy je přírůstek d_{i_p} daného cyklu snížen na polovinu u jedince i a obou rodičů zprůměrnováním s výsledkem předchozího cyklu, tj. obecně:

$$d_{i_r}^{*+1} = (d_{i_r}^{*+1} + d_{i_r}^*) \cdot 0,5$$

Tento krok je do algoritmu zařazen z důvodu lepší konvergence cyklů k jednotné hodnotě.
(konec cyklu)

Při ověřování metody bylo zjištěno, že pro dojený skot chovaný v České republice je pro ustálení hodnot zvířat 100 iterací dostačující. Výstupní údaj z cyklu $d_{i_r}^{*+1}$ je pak vstupním údajem pro další výpočty jako $d_{i_{rp}}$. Pro určení spolehlivosti běžné plemenné hodnoty jedince jako výsledku z této části postupu, lze $d_{i_{rp}}$ převést na spolehlivost dle vzorce:

$$r_{i_{rp}}^2 = \frac{d_{i_{rp}}}{d_{i_{rp}} + \alpha}$$

II.3.2.2. Část II. - Odhad spolehlivosti GEPH u genotypovaných zvířat

Efektivní počet případů $d_{i_{rp}}$ se využije při odhadu spolehlivosti GEPH u genotypovaných jedinců tak, že jednotlivé hodnoty se umístí na diagonálu nové matice „**D**“ jako prvky d_{ii} . Ostatní prvky matice „**D**“ jsou nulové. Matice „**D**“ je pak zahrnuta do výpočtu matice \mathbf{Q}^{-1} :

$$\mathbf{Q}^{-1} = [\mathbf{D} + (\mathbf{I} + \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1}) \cdot \alpha]^{-1}$$

kde **I** je jednotková matice,

\mathbf{A}_{22}^{-1} inverze rodokmenové matice příbuznosti genotypovaných zvířat,

\mathbf{G}^{-1} inverze genomické matice příbuznosti genotypovaných zvířat,

Vypočítané hodnoty u genotypovaných zvířat se pak převedou na spolehlivost GEPH genotypovaných zvířat pomocí:

$$r_{i_{rpg}}^2 = 1 - \alpha \cdot q^{ii}$$

q^{ii} jsou diagonální hodnoty matice \mathbf{Q}^{-1} .

II.3.2.3. Část III. - Odhad spolehlivosti celé populace se započtením GEPH u genotypovaných zvířat

Tato část je shodná s první částí postupu (tj., s II.3.2.1.) s těmito výjimkami:

1. Do iteračního cyklu vstupují u genotypovaných zvířat místo původních d_i hodnoty

$d_{i_{rpg}}$ odvozené z $r_{i_{rpg}}^2$ dle následujícího vzorce:

$$d_{i_{rpg}} = \alpha \frac{r_{i_{rpg}}^2}{1 - r_{i_{rpg}}^2}$$

2. Na konci každé iterace jsou hodnoty $d_{i_{rpg}}$ u genotypovaných zvířat přepsány zpět na hodnotu

d_i vypočítané v části II. - odhadu spolehlivosti GEPH u genotypovaných zvířat (tj., II.3.2.2.)

Spolehlivost GEPH zvířete i je odvozena z výsledného efektivního počtu případů „ d_i “ z iteračního cyklu II.3.2.3. dle vzorce:

$$r_i^2 = \frac{d_i}{d_i + \alpha}$$

Uvedený postup byl ověřován pro využití při odhadu spolehlivosti GEPH jedнокrokovou metodou (ssGBLUP) a první krok samostatně pro odhad spolehlivosti běžných plemenných hodnot z metody BLUP. Dále byly ověřovány úpravy pro zohlednění údajů z Interbullu, opakovaných pozorování u jednotlivců a příbuzenských vazeb mezi jedinci ve skupinách vrstevníků.

Z obecného hlediska by měl být postup vhodný i pro příbuzné metody GBLUP a blending ssGBLUP. V případě potřeby by bylo možné provádět odhady spolehlivosti PH či GEPH popsáním postupem i u tzv. „Test Day“ modelu, pokud by jednotlivé dny pozorování („Test Day“) byly považovány za opakovaná pozorování.

II.4. Ověření pracovních postupů a algoritmy

K přípravě datových souborů, odhadu spolehlivosti běžných i genomických plemenných hodnot a k rozebrání výsledků byl používán program vytvořený v systému SAS (SAS, 2005). Samotný odhad spolehlivosti probíhal v modulu pro maticové výpočty SAS / IML.

Výsledky ověřování jsou uvedeny v pracích Bauer (2013) a Bauer et al. (2014a; 2014b), samotný program vytvořený v SAS pro odhad spolehlivosti je uveden jako příloha této certifikované metodiky.

III. Srovnání „novosti postupů“

Pravidelné genetické hodnocení zvířat je nyní založeno na předpovědi plemenných hodnot metodou BLUP. Díky vědeckému pokroku je nyní možné využít i údaje o zvířatech přímo uložené v jejich DNA. V takovém případě se pak jedná o předpověď genomických plemenných hodnot. Pro správnou šlechtitelskou práci je třeba znát i spolehlivost těchto genomických plemenných hodnot.

Předpověď genomických plemenných hodnot se nyní zavádí do praxe a s těmito hodnotami by měla být šlechtitelům a chovatelům dodána i spolehlivost genomických plemenných hodnot.

Jak předpověď genomických plemenných hodnot tak i odhad jejich spolehlivosti jsou v rámci pravidelných hodnocení zvířat novou záležitostí.

IV. Popis uplatnění Certifikované metodiky

Tato metodika je podkladem pro rutinní odhady spolehlivosti genomických plemenných hodnot pro množství nadojeného mléka skotu. Je jí ovšem možné využít i u dalších vlastností, jejichž plemenné hodnoty jsou předpovídány metodou BLUP či ssGBLUP. Metodika bude uplatněna prostřednictvím ze zákona pověřenou organizací ČMSCH, a.s.. Výsledky této metodiky budou využity příslušnými chovatelskými svazy.

V. Ekonomické aspekty

Podle zákona č. 110/1997 Sb. O potravinách a zákona č. 154/2000 Sb. O šlechtění, plemenitbě a evidenci hospodářských zvířat ve znění pozdějších předpisů je ČMSCH právnická osoba pověřená ministerstvem k výkonu činností podle jednotlivých bodů § 23c. Jmenovitě podle odstavců 1 a 2 a §7 je povinna poskytovat chovatelům a oprávněným osobám údaje, zpracovávat, zveřejňovat a evidovat výsledky, což se týká všech chovatelsky důležitých vlastností, včetně mléčné užitkovosti skotu. V souladu s doporučením Rady vlády pro výzkum uvádíme, že ČMSCH nevytváří těmito činnostmi zisk, ale poskytuje široké chovatelské veřejnosti co nejobektivnější údaje a vyhodnocením celostátních databází vytváří podklady pro prokázání kvality plemenářské práce chovatelů. Nová objektivní hodnocení jsou také nutná pro mezinárodní hodnocení zvířat prostřednictvím organizace Interbull, které je Česká republika členem.

Pokud porovnáme spolehlivost genomické plemenné hodnoty s běžnou plemennou hodnotou zvířat z BLUP metody, může odpovídat přínos genomických údajů 10ti až 20ti dalším dcerám hodnoceného jedince (VanRaden, 2008). Jako důsledek tohoto jevu je, že i mladý neprověřený býk může mít spolehlivost pohybující se okolo 0,6 (oproti běžným 0,4) a tím má předpovídána

plemenná hodnota větší vypovídací schopnost o jeho genetickém založení a o jeho možném využití v plemenitbě, a to dříve než se dostane do testace. Ačkoliv se později tento „náskok“ ve spolehlivosti se vzrůstajícím počtem potomků ostatních býků zmenšuje, umožňuje to šlechtitelům vybírat do plemenitby mladší zvířata se stejnou spolehlivostí, jaká byla dosahována u zvířat starších. Zkrátí se tím generační interval a náklady s testací spojené. Schaeffer (2006) odhaduje u programu mladých býků v USA 12,8 krát nižší náklady na rok, 2,2 krát vyšší roční genetický zisk a 27,8 krát lacinější posun v genetickém založení plemeníků o jednu směrodatnou odchylku.

V podmínkách České Republiky odhadovali přínos využití spolehlivějších (genomických) plemenných hodnot Pešek a Příbyl (2014). Došli k závěru, že použitím genomické selekce dojde díky vyššímu genetickému zisku při vhodném selekčním programu ke zvýšení mléčné užitkovosti o 67,6 kg mléka na rok a dojnici oproti selekci pomocí běžných plemenných hodnot. Ve výpočtech po zahrnutí nákladů na genotypování a navazující výdaje došli k vyššímu zisku u dojnice o 37,1 Kč na laktaci. V případě přepočtu na počet krav ve stávající kontrole užitkovosti (tj. 360 000) se jedná o přínos metody přibližně 11 500 000 Kč ročně.

Tento příklad je ukázkou, jak se výsledek šlechtění projeví v ekonomice chovu. Není možné v tuto chvíli přesně určit, jak vysoký bude ve skutečnosti genetický zisk ani za jak dlouho se populace dostane na tuto předpokládanou úroveň, neboť selekční rozhodnutí závisí na rozhodnutí chovatelů.

Z uvedeného vyplývá, že i když by byl přínos poměrně malý, v rozsahu celé chované populace dosahuje vysokých hodnot. Šlechtění má kumulativní účinek a tudíž ekonomický přínos při účinnějším šlechtění bude každoročně násobně narůstat.

VI. Seznam použité související literatury

- Forni S., Aguilar, I., Misztal, I. 2011. Different genomic relationship matrices for single step analysis using phenotypic, pedigree and genomic information. *Genetics Selection Evolution* 44:2 43: 1. doi: 10.1186/1297-9686-43-1
- Jakubec, V., Bezdíček, J., Louda, F. 2010. Selekce – inbríding – hybridizace. *Agrovýzkum Rapotín s. r. o. Rapotín*. 382s. ISBN 978-80-87144-22-0.
- Misztal, I., Lawlor, T.J., Short, T.H., Wiggans, G.R., 1991. Continuous genetic evaluation of Holsteins for type. *Journal of Dairy Science*. 74 (6). 2001-2009.
- Misztal, I., Legarra A., Aguilar. I. 2009. Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. *Journal of Dairy Science*. 92. 4648-4655.
- Misztal, I., Legarra, A., Short, T. H. 1993. Implementation of single- and multiple-trait Animal models for genetic evaluation of Holstein type traits. *Journal of Dairy Science*. 76 (5). 1421–1432.
- Misztal, I., Tsuruta, S., Aguilar, I., Legarra, A., Van Raden, P. M., Lawlor, T. J.. 2013. Methods to approximate reliabilities in single-step genomic evaluation. *Journal of Dairy Science*. 96 (1). 647 – 654.
- Pešek, P., Příbyl, J. 2014. Genomika - možnost zvyšování zisku u dojeného skotu. *Náš chov*, 74 (1), 59-60.
- Příbyl, J., Bolečková, J., Haman, J., Kott, T., Příbylová, J., Šimečková, M., Vostrý, L., Zavadilová, L. 2011. Úprava soustav rovnic pro předpověď genomické plemenné hodnoty (GEPH) jednokrokovou metodou. *Metodika. VÚŽV v.v.i., Praha Uhřetěves*. ISBN 978-80-7403-079-6.

- Příbyl, J., Haman, J., Kott, T., Příbylová, J., Šimečková, M., Vostrý, L., Zavadilová, L., Čermák, V., Růžička, Z., Šplíchal, J., Verner, M., Motyčka, J., Vondrášek, L., 2012. Single-step prediction of genomic breeding value in a small dairy cattle population with strong import of foreign genes. *Czech Journal of Animal Science*. 57 (4). 151–159.
- Příbyl, J., Madsen P., Bauer J., Příbylová, J., Šimečková M., Vostrý, L., Zavadilová, L. 2013. Contribution of domestic production records, Interbull estimated breeding values, and single nucleotide polymorphism genetic markers to the single-step genomic evaluation of milk production. *Journal of Dairy Science*. 96 (2). 1865-1873.
- SAS (2005): SAS/STAT® 9.1 User's Guide. Cary, NC: SAS Institute Inc. 5121p.
- Schaeffer, L.R. 2006. Strategy for applying genome-wide selection in dairy cattle. *Journal of Animal Breeding and Genetics*. 123. 218-223.
- Szyda J., Zarnecki, A., Suchocki, T., Kaminski, S. 2011. Fitting and validating the genomic evaluation model to Polish Holstein-Friesian cattle. *Journal of Applied Genetics*. 52. 363-366.
- Tier, B., Meyer, K. 2004. Approximating prediction error covariances among additive genetic effects within animals in multiple-trait and random regression models. *Journal of Animal Breeding and Genetics*. 121. 77–89.
- VanRaden, P.M., 2008. Efficient methods to compute genomic predictions. *Journal of Dairy Science*. 91. 4414-4423.

VII. Seznam publikací, které předcházely metodice

- Bauer, J., 2013. Genomická selekce dojeného skotu - spolehlivost genomických plemenných hodnot. Disertační práce, Česká zemědělská univerzita v Praze, 119s., bez ISBN.
- Bauer J., Příbyl J., Vostrý L., Zavadilová, L. Contribution of domestic production records and Interbull EBV on approximate reliabilities of single-step genomic breeding values for traits in cattle. *Journal of Animal Science (oponentské řízení)*
- Bauer J., Vostrý L., Svitáková A. 2013. Spolehlivost genomických plemenných hodnot. *Náš chov*, 73, (10), 61-62.
- Bauer J., Vostrý L., Příbyl J., Svitáková A., Zavadilová L. Approximation of the reliability of single-step genomic breeding values for dairy cattle in the Czech Republic. *Animal Science Papers and Reports (oponentské řízení)*
- Příbyl, J., Haman, J., Kott, T., Příbylová, J., Šimečková, M., Vostrý, L., Zavadilová, L., Čermák, V., Růžička, Z., Šplíchal, J., Verner, M., Motyčka, J., Vondrášek, L., 2012. Single-step prediction of genomic breeding value in a small dairy cattle population with strong import of foreign genes. *Czech Journal of Animal Science*. 57 (4). 151–159.
- Příbyl, J., Madsen P., Bauer J., Příbylová, J., Šimečková M., Vostrý, L., Zavadilová, L. 2013. Contribution of domestic production records, Interbull estimated breeding values, and single nucleotide polymorphism genetic markers to the single-step genomic evaluation of milk production. *Journal of Dairy Science*. 96 (2). 1865-1873.

VIII. Přílohy

Příloha 1: Program vytvořený v SAS pro odhad spolehlivosti běžných a genomických plemenných hodnot

```
/****** V1 *****/
proc printto print= '/home/pribyl/spolehjb/output.lst' ;
proc printto log= '/home/pribyl/spolehjb/log.lst' ;

data uzit;
infile '/home/pribyl/spolehjb/uzdo5'; input sro j;

data rod;
infile '/home/pribyl/spolehjb/maprd5'; input j o m ; run; /*rodokmen*/

data bull;
infile '/home/pribyl/spolehjb/bullenr' termstr=crlf; input j v d; run; /*hodnoty efekt poctu pripadu z
interbullu*/

proc means data=uzit;run;
proc means data=rod;run;
proc means data=bull;run;

data rod;set rod;
if o<0 then o=0;
if m<0 then m=0;
run;

/*v7 priprava dat*/
proc sql;
create table souroz as
SELECT uzit.sro, rod.j, rod.o, rod.m
FROM uzit, rod
WHERE uzit.j=rod.j;
quit;
/*v7 konec pripravy*/

/*vypocet*/
proc iml; /*odhad spol negenotyp pro vsechny*/
print "nacistani uzitkovosti a rodokmenu";
use uzit; read all into xxx; close uzit;

use rod; read all into yyy; close rod;
print "nacteno";
/*definice konstant*/
/*h2=0.25;
a= (1-h2)/h2;
a=round(a,.01);*/

a=3;
tau=0; /*rezidual variance lomeno variance trvaleho prostredi*/;
```

```

paid=1917416;          *.....pocet v rodokmenu.;
pcg=67459;            *.....pocet sro.;
pozorov = nrow(xxx);  *.....pocet radku s pozorovanim.;
opakovani=0;
bul=1;
pribuz=0;
print "definovany rozmery promennych";
/*zacatek vypoctu odhadu spolehlivosti*/
anm=xxx[,2];          *.....jedinci s uzitkovosti.;
cg=xxx[,1];           *.....vrstevnici.;

aid=yyy[,1];          *.....vektor jedincu.;
sid=yyy[,2];          *.....vektor otcu.;
did=yyy[,3];          *.....vektor matek.;

b=yyy*0;              *.....priprava matic b.;
r=b;                  *.....priprava matice r.;
s=b;
bx=b;
print "definovany hodnoty vektoru";
*.....pocety jedincu v jednotlivych SRO.;
tcg=j(pcg,1,0);      *print tcg;          *..... pocet sro.;

do i=1 to pozorov;    *.....jedincu s pozorovanim.;
  tcg[cg[i]]= tcg[cg[i]]+1;
end; print "vytvoreno tcg";

/**v7-pribuznost v sro**/
if pribuz=1 then do;
  use souroz; read all into sour;close souroz; /*..v7...matice sourozencu.....*/
  print "pribuzenska kontrola sro";
  do sro=1 to pcg;
    rows2keep = loc(sour[,1]=sro);
    sou = sour[rows2keep,];
    do i=1 to nrow(sou);
      x=0;
      do j=1 to nrow(sou); /*sou=sro j o m*/
        if sou[i,2]=sou[j,3] | sou[i,2]=sou[j,4] | sou[i,3]=sou[j,2] | sou[i,4]=sou[j,2] |
          (sou[i,3]=sou[j,3] & sou[j,3]^=0) | (sou[i,4]=sou[j,4]^=0 & sou[j,4]^=0) then
x=x+1;
      end;
      x=x-1;
      zapsat=(sou[i,2]||sro||x);
      qi=qi//zapsat;
    end;

  end;
/*vyhozeni duplikaci*/
call sortndx (sortsez, qi,1:2);
unique_rows = uniqueby(qi, 1:2, sortsez);
qik = qi[sortsez[unique_rows], 1:3];

```

```

/*konec vyhozeni duplikaci*/
end;
print "pribuzenska kontrola dokoncena";
/**konec v7**/

dz= j(paid,1,0);
d= j(paid,1,0);          *...priprava vektoru - jedincu.;

if pribuz=1 then
do;
  do i=1 to pozorov;
    row= loc(qik[,1]=anm[i] & qik[,2]=cg[i]);
    w=sqrt ((tcg[cg[i]]-qik[row,3])/tcg[cg[i]]);
    dz[anm[i]]= dz[anm[i]]+((1-1/(tcg[cg[i]]*w))); /*nacteni v pripade pribuznych-v7*/
  end;
end;
else
do;
  do i=1 to pozorov;          *.....jedincu s pozorovnanim.....z.....;
    dz[anm[i]]= dz[anm[i]]+((1-1/tcg[cg[i]])); *... vytvoreni vektoru .....;
  end;
end;

if opakovani=1 then do; /*zapocitani opakovani*/
  d=tau*dz/(tau+dz);
  end;
else do;
d=dz;
  end;

if bul=1 then do;
  use bull; read all into zzz; close bull;
  pozorovbul = nrow(zzz);
  dbulj=zzz[,1];          *identifikace jedincu interbull;
  dbuld=zzz[,3];          *hodnota efekt poctu jedincu z interbull;
  dbulrod=j(paid,1,0); *vektor pro vsechny jedince v rodok;

  do i=1 to pozorovbul;
    dbulrod[dbulj[i]]=dbuld[i];
  end;
  d=d+dbulrod; /*nacisti efekt poctu pripadu z vice zdroju*/
end;

print "vytvoren vektor pro d";

d0=d;

deltad= j(paid,1,10);

do m=1 to 100;
  do i=1 to paid;

```

```

ba=d[aid[i]];
if sid[i]=0 then bs=0; else bs=d[sid[i]];
if did[i]=0 then bd=0; else bd=d[did[i]];

b[i,]=ba||bs||bd;
end;
q=b;

do j=1 to 10;
  do i=1 to paid;
qa=q[i,1];
qs=q[i,2];
qd=q[i,3];

bxs=0.5*a+qs-(0.25*(a**2)*(2*a+qa)-(a**3)+(a**2)*(1.5*a+qd))/((1.5*a+qd)*(2*a+qa)-(a**2));
bxd=0.5*a+qd-(0.25*(a**2)*(2*a+qa)-a**3+(a**2)*(1.5*a+qs))/((1.5*a+qs)*(2*a+qa)-a**2);
bxa=a+qa-((a**2)*(1.5*a+qd)-a**3+(a**2)*(1.5*a+qs))/((1.5*a+qs)*(1.5*a+qd)-0.25*(a**2));

bx[i,]=bxa||bxs||bxd;

  end;

do k=1 to paid;
do l=1 to 3;

if abs(b[k,l])<0.00001 then s[k,l]=0; else s[k,l]=bx[k,l]-b[k,l];
end;
end;

q=q-s;

do k=1 to paid;
  do l=1 to 3;
    if q[k,l]<0 then q[k,l]=0;/*else q[k,l]=q1[k,l]; */
  end;
end;

r=bx-q;
*print r;

end;
d1=d0;

do k=1 to paid;
tem=aid[k];
d1[tem]=d1[tem]+r[k,1];
end;

do k=1 to paid;
tem=sid[k];
if tem^=0 then d1[tem]=d1[tem]+r[k,2];      *.....cislo otce - 2. sloupec v r patri otci.....;

```

```

end;

do k=1 to paid;
  tem=did[k];
  if tem^=0 then d1[tem]=d1[tem]+r[k,3];          *.....cislo matky - 3. sloupec v r patri
matce.....;
end;

      d=0.5*(d+d1);

  print "cyklus prvnych iteraci c." m; /*kontrolni na cyklus stovkovy*/

  *.....;
end; /*stovkova end*/
print "konec prvnych iteraci";

c= j(paid,1,0);
di=j(paid,1,0);

do i=1 to paid;
  c[i]=1/(a+d[i]); /*# aproximace diagonalnich prvnu matice C.....*/

  spol=1-(a*c);

  di[i]=a*(spol[i]/(1-spol[i])); /*efektivni pocet pripadu*/

end;

*print c spol di;
print "vypocteno c, spol, di";

vysl1=c||spol||di;
create vysledky1 from vysl1;
  append from vysl1;
quit;
data vysledky1; set vysledky1;
filename out '/home/pribyl/spolehjb/vysledky1';
file out; put col1-col3;
run;

/***** V2 *****/
proc printto print= '/home/pribyl/spolehjb/output2.lst' ;
proc printto log= '/home/pribyl/spolehjb/log2.lst' ;

data maticeg;
infile '/home/pribyl/spolehjb/slog5'; input zvirex zvirey hodnota ; run; /*nacistani g matice*/

data matice22;
infile '/home/pribyl/spolehjb/a224'; input zvirex zvirey hodnota ; run; /*nacistani a22 matice*/

data sezge;

```

```
infile '/home/pribyl/spolehjb/sezgen5';input poradi ; run;/*nacistani kteri jsou genotypovani*/
```

```
data vysledky1;  
infile '/home/pribyl/spolehjb/vysledky1'; input c spol di;run;
```

```
/*precislovani genotyp*/  
data sezge;set sezge;  
kod=_n_;  
run;
```

```
proc sql noprint;  
create table matice22 as  
SELECT matice22.hodnota, matice22.zvirey, sezge.kod as kodx  
FROM matice22, sezge  
WHERE matice22.zvirex=poradi;
```

```
create table matice22 as  
SELECT matice22.hodnota, matice22.kodx, sezge.kod as kody  
FROM matice22, sezge  
WHERE matice22.zvirey=poradi;
```

```
create table maticeg as  
SELECT maticeg.hodnota, maticeg.zvirey, sezge.kod as kodx  
FROM maticeg, sezge  
WHERE maticeg.zvirex=poradi;
```

```
create table maticeg as  
SELECT maticeg.hodnota, maticeg.kodx, sezge.kod as kody  
FROM maticeg, sezge  
WHERE maticeg.zvirey=poradi;
```

```
quit;
```

```
proc iml;/*odhad pro genotyp*/  
print "priprava na inverzi matic g,a22";  
a=3;  
/*nacteni matic*/  
print "nacistani uzitkovosti a rodokmenu";  
use maticeg; read all into matg; close maticeg;  
use matice22; read all into mat22; close matice22;  
print "nacteno";  
/*konec*/  
print "nacistani vysledku1";  
/*nacteni seznamu genotypovanych*/  
use sezge;read all into sezgenot;close sezge;  
/*nacteni vysledku negenot*/  
use vysledky1;read all into vysl1;close vysledky1;  
print "nacteno";  
nogenotyp=nrow(sezgenot);  
nomatice=nogenotyp;  
*.....pocet genotypovanych.;  
*.....rozmer matic g,a22.;
```

```
noudmatice=((nomatice*nomatice)/2)+(nomatice/2);  
trojuhelniku matic g, a22.;
```

*.....pocet udaju v

```
/*inverze matic g,a22*/
```

```
mg=j(nomatice,nomatice,0);  
ma22=j(nomatice,nomatice,0);
```

```
do i=1 to noudmatice;  
    tempa=matg[i,2];tempb=matg[i,3];  
    mg[tempa,tempb]=matg[i,1];  
    mg[tempb,tempa]=matg[i,1];  
end;
```

```
do i=1 to noudmatice;  
    tempa=mat22[i,2];tempb=mat22[i,3];  
    ma22[tempa,tempb]=mat22[i,1];  
    ma22[tempb,tempa]=mat22[i,1];  
end;
```

```
/*do i=1 to noudmatice;  
    if mg[i,i]<1 then do;  
        gdiag=mg[i,i];  
        print i gdiag;  
        mg[i,i]=1;  
        end;  
    if ma22[i,i]<1 then do;  
        m22diag=ma22[i,i];  
        print i m22diag;  
        ma22[i,i]=1;  
        end;  
end;*/
```

```
ginv=inv(mg);  
a22inv=inv(ma22);
```

```
*print ginv;  
*print a22inv;
```

```
/*vypocet inverze (bod 3) dle Misztal Stavanger 2011 -matice Q*/  
dd=j(nogenotyp,nogenotyp,0);
```

```
do i=1 to nogetotyp; /*nogenotyp je jich genotypovano*/  
    j=sezgenot[i,1];  
    dd[i,i]=vysl1[j,3];  
end;  
*print dd;
```

```
/*****/
```

```
jedn=i(nogenotyp);
```

```
qq=dd+((jedm+ginv-a22inv)*a);
```

```

qqinv=inv(qq);

*print dd invi qq qqinv ginv a22inv;
print "vypocteno qq dle misztal (3)";
/*prevod na r (bod 4)*/
rg=j(nogenotyp,1,0);

do i=1 to nogenotyp;
    rg[i]=1-(a*qqinv[i,i]);/*spolehlivost genotypovanych*/
end;

*print rg;/*spolehlivost genotypova*/

do i=1 to nogenotyp;/*nulovani a vypis zapornych rg*/
    if rg[i]<0 then do;
        genomr=rg[i];
        print i genomr;
        rg[i]=0;
    end;
end;

/*rozdil stejne zvire genr-negenr*/
delta=j(nogenotyp,3,0);

do i=1 to nogenotyp;
    j=sezgenot[i,1];
    delta[i,1]=rg[i];/*genot*/
    delta[i,2]=vysl1[j,2];/*negenot*/
end;

delta[,3]=delta[,1]-delta[,2];/*rozdil*/

*print delta;

create vysledky2 from delta;
    append from delta;
quit;
data vysledky2;set vysledky2;
filename out '/home/pribyl/spolehjb/vysledky2';
file out;put col1-col3;
run;

/***** V3 *****/
proc printto print= '/home/pribyl/spolehjb/output3.lst' ;
proc printto log= '/home/pribyl/spolehjb/log3.lst' ;

data uzit;
infile '/home/pribyl/spolehjb/uzdo5'; input sro j /*ml vek vek2 sp sp2*/ ;

data rod;
infile '/home/pribyl/spolehjb/maprd5'; input j o m ; run;/*rodokmen*/

```



```

data bull;
infile '/home/pribyl/spolehjb/bullenr' termstr=crlf; input j v d; run;/*hodnoty efekt poctu pripadu z
interbullu*/

data vysledky2;
infile '/home/pribyl/spolehjb/vysledky2'; input rg spol rozdil;run;

data sezge;
infile '/home/pribyl/spolehjb/sezgen5';input poradi ; run;/*nacistani kteri jsou genotypovani*/

proc means data=uzit;run;
proc means data=rod;run;
proc means data=bullenr;run;

data rod;set rod;
if o<0 then o=0;
if m<0 then m=0;
run;

/*v7 priprava dat*/
proc sql;
create table souroz as
SELECT uzit.sro, rod.j, rod.o, rod.m
FROM uzit, rod
WHERE uzit.j=rod.j;
quit;
/*v7 konec pripravy*/

/*precislovani genotyp*/
data sezge;set sezge;
kod=_n_;
run;

proc iml;/*bod 5 - iteracni uprava spolehlivosti negenotypovanych zvirat pres pribuznost*/
print "nacistani uzitkovosti a rodokmenu";
use uzit; read all into xxx; close uzit;

use rod; read all into yyy; close rod;
print "nacteno";

/*seznam hodnot genotypovanych*/
print "nacistani vysledku2";
use vysledky2;read all into vysl2;close vysledky2;
/*nacteni seznamu genotypovanych*/
use sezge;read all into sezgenot;close sezge;
print "nacteno";
a=3;
tau=0;/*rezidual variance lomeno variance trvaleho prostredi*/;
paid=1917416; * .....pocet v rodokmenu.;

```

```

pcg=67459;          *.....pocet sro.;
pozorov = nrow(xxx);      *.....pocet radku s pozorovanim.;
opakovani=0;
bul=1;
pribuz=0;
nogenotyp=nrow(sezgenot);

print "definovany rozmery promennych";
/*zacatek vypoctu odhadu spolehlivosti*/
anm=xxx[,2];          *.....jedninci s uzitkovosti.;
cg=xxx[,1];          *.....vrstevnici.;

aid=yyy[,1];          *.....vektor jedincu.;
sid=yyy[,2];          *.....vektor otcu.;
did=yyy[,3];          *.....vektor matek.;

b=yyy*0;             *.....priprava matic b.;
r=b;                 *.....priprava matice r.;
s=b;
bx=b;
print "definovany hodnoty vektoru";
*.....pocety jedincu v jednotlivych SRO.;
tcg=j(pcg,1,0);      *print tcg;          *.....pocet sro.;

do i=1 to pozorov;      *.....jedincu s pozorovnanim.;
  tcg[cg[i]]= tcg[cg[i]]+1;
end; print "vytvoreno tcg pro doplneni negenotypovanych";
*print tcg;

/**v7-pribuznost v sro**/
if pribuz=1 then do;
  use souroz; read all into sour;close souroz; /*..v7...matice sourozencu.....*/
  print "pribuzenska kontrola sro";
  do sro=1 to pcg;
    rows2keep = loc(sour[,1]=sro);
    sou = sour[rows2keep,];
    do i=1 to nrow(sou);
      x=0;
      do j=1 to nrow(sou); /*sou=sro j o m*/
        if sou[i,2]=sou[j,3] | sou[i,2]=sou[j,4] | sou[i,3]=sou[j,2] | sou[i,4]=sou[j,2] |
          (sou[i,3]=sou[j,3] & sou[j,3]^=0) | (sou[i,4]=sou[j,4]^=0 & sou[j,4]^=0) then
        x=x+1;
      end;
      x=x-1;
      zapsat=(sou[i,2]||sro||x);
      qi=qi//zapsat;
    end;
  end;

  end;
/*vyhozeni duplikaci*/
call sortndx (sortsez, qi,1:2);

```

```

        unique_rows = uniqueby(qi, 1:2, sortsez);
        qik = qi[sortsez[unique_rows], 1:3];
/*konec vyhozeni duplikaci*/
end;
print "pribuzenska kontrola dokoncena";
/**konec v7**/

rg=vysl2[,1];

genotdi=j(13,2,0);

genotdi=j(nogenotyp,2,0);
do i=1 to nogenotyp;
    genotdi[i,1]=sezgenot[i,1];/*cisla genotypovanych zvirat*/
    genotdi[i,2]=a*(rg[i]/(1-rg[i]));/*prevod na efektivni pocet pripadu*/
end;

*print rg genotdi;

/*vytvoreni matice d*/
dz= j(paid,1,0);
d= j(paid,1,0);          *...priprava vektoru - jedincu.;

if pribuz=1 then
do;
    do i=1 to pozorov;
        row= loc(qik[,1]=anm[i] & qik[,2]=cg[i]);
        w=sqrt ((tcg[cg[i]]-qik[row,3])/tcg[cg[i]]);
        dz[anm[i]]= dz[anm[i]]+((1-1/(tcg[cg[i]]*w))); /*nacteni v pripade pribuznych-v7*/
    end;
end;
else
do;
    do i=1 to pozorov;          *.....jedincu s pozorovnanim.....z.....;
        dz[anm[i]]= dz[anm[i]]+((1-1/tcg[cg[i]])); *... vytvoreni vektoru .....;
    end;
end;

if opakovani=1 then do; /*zapocitani opakovani*/
    d=tau*dz/(tau+dz);
    end;
else do;
d=dz;
    end;

if bul=1 then do;
    use bull; read all into zzz; close bull;
    pozorovbul = nrow(zzz);
    dbulj=zzz[,1];          *identifikace jedincu interbull;
    dbuld=zzz[,3];          *hodnota efekt poctu jedincu z interbull;
    dbulrod=j(paid,1,0); *vektor pro vsechny jedince v rodok;

```

```

do i=1 to pozorovbul;
    dbulrod[dbulj[i]]=dbuld[i];
end;
d=d+dbulrod;
end;

print "vytvoren vektor pro d pro doplneni negenotypovanych";

*print d;
do i=1 to nogenotyp; /*prepis hodnot u genotypovanych*/
tem=genotdi[i,1];
d[tem]=genotdi[i,2];
end;
*print d;

d0=d;

deltad= j(paid,1,10);

do m=1 to 100;
do i=1 to paid;
ba=d[aid[i]];
if sid[i]=0 then bs=0; else bs=d[sid[i]];
if did[i]=0 then bd=0; else bd=d[did[i]];

b[i,]=ba||bs||bd;
end;
q=b;

do j=1 to 10;
do i=1 to paid;
qa=q[i,1];
qs=q[i,2];
qd=q[i,3];

bxs=0.5*a+qs-(0.25*(a**2)*(2*a+qa)-(a**3)+(a**2)*(1.5*a+qd))/((1.5*a+qd)*(2*a+qa)-(a**2));
bxd=0.5*a+qd-(0.25*(a**2)*(2*a+qa)-a**3+(a**2)*(1.5*a+qs))/((1.5*a+qs)*(2*a+qa)-a**2);
bxa=a+qa-((a**2)*(1.5*a+qd)-a**3+(a**2)*(1.5*a+qs))/((1.5*a+qs)*(1.5*a+qd)-0.25*(a**2));

bx[i,]=bxa||bxs||bxd;

end;

do k=1 to paid;
do l=1 to 3;
if abs(b[k,l])<0.00001 then s[k,l]=0; else s[k,l]=bx[k,l]-b[k,l];
end;
end;

q=q-s;

```

```

do k=1 to paid;
    do l=1 to 3;
        if q[k,l]<0 then q[k,l]=0;
    end;
end;

r=bx-q;
*print r;

end;
d1=d0;

do k=1 to paid;
    tem=aid[k];
    d1[tem]=d1[tem]+r[k,1];
end;

do k=1 to paid;
    tem=sid[k];
    if tem^=0 then d1[tem]=d1[tem]+r[k,2];    *.....cislo otce - 2. sloupec v r patri otci.;
end;

do k=1 to paid;
    tem=did[k];
    if tem^=0 then d1[tem]=d1[tem]+r[k,3];    *.....cislo matky - 3. sloupec v r patri matce.;
end;

    d=0.5*(d+d1);
print "cyklus druhych iteraci c." m; /*kontrolni na cyklus stovkovy*/

*.....;
/*opraveni genotypovanych pro novy cyklus*/

do i=1 to nogenotyp; /*prepis hodnot u genotypovanych*/
    tem=genotdi[i,1];
    d[tem]=genotdi[i,2];
end;
/*konec uprav dle genotypovanych*/
*.....;
end; /*stovkova end*/
print "konec druhych iteraci";
c= j(paid,1,0);
di=j(paid,1,0);

do i=1 to paid;
    c[i]=1/(a+d[i]); /*# aproksimace diagonalnich prvku matice C.....*/

    spol=1-(a*c);

    di[i]=a*(spol[i]/(1-spol[i])); /*efektivni pocet pripadu*/

```

```
end;

*print c spol di;
print "vypocteno c,spol,di";
vysl3=c||spol||di;

create vysledky3 from vysl3;
  append from vysl3;
quit;
data vysledky3;set vysledky3;
filename out '/home/pribyl/spolehjb/vysledky3';
file out;put col1-col3;
run;
```

Vydal: Výzkumný ústav živočišné výroby, v.v.i.
Přátelství 815, 104 00 Praha Uhřetěves

Název: **Odhad spolehlivosti jednokrokových genomických plemenných hodnot pro dojený skot**

Autoři: Ing. Jiří Bauer, Ph.D. (podíl práce 70 %)
doc. Ing. Luboš Vostrý, Ph.D. (podíl práce 20 %)
prof. Ing. Josef Příbyl, DrSc. (podíl práce 10 %)

Oponenti: prof. Ing. Jindřich Čítek, CSc.
Jihočeská univerzita v Českých Budějovicích
Ing. Zdeňka Majzlíková
Česká státní plemenářská inspekce, Praha

ISBN 978-80-7403-120-5

Metodika byla vypracována v rámci řešení výzkumného projektu NAZV QI111A167.

Vydáno bez jazykové úpravy.