



Česká plemenářská inspekce

Štěpánská 626/63, Praha 1, 11000

v y d á v á

OSVĚDČENÍ

(12641 /2016 - ČPI)

o uznání certifikované metodiky
v souladu s podmínkami „Metodiky hodnocení výsledků výzkumných organizací
a hodnocení výsledků ukončených programů (platné pro léta 2013 až 2016)“

Metodika

GEPH na základě referenčního souboru složeného z domácích TD záznamů a Interbullem přepočtených mezinárodních MACE hodnot

Autoři: J. Příbyl, J. Bauer, A. Kranjčevičová, P. Pešek, J. Příbylová, L. Vostrý, L. Zavadilová
Výzkumný ústav živočišné výroby, v. v. i. Praha Uhřetěves 2016, ISBN: 978-80-7403-157-1

Metodika byla vypracována v rámci výzkumného projektu/podpory na rozvoj výzkumné
organizace č. QJ1510139, Celostátní informační systém genetického hodnocení
hospodářských zvířat.

Projekt využívá „Pravidla pro odvětví zemědělství, lesnictví a rybolovu“ ANO/NE*.
V případě, že projekt využívá „Pravidla pro odvětví zemědělství, lesnictví a rybolovu“, je
výsledek typu N_{met} zdarma k dispozici všem zájemcům na webové stránce www.vuzv.cz.

V Praze, dne 21. 12.2016

Česká plemenářská inspekce
Štěpánská 626/63
110 00 Praha 1
1

Razítko odborného orgánu státní správy:

Jméno zástupce odborného útvaru státní správy: Ing. Zdenka Majzlíková

Funkce zástupce odborného útvaru státní správy: ředitelka

Podpis zástupce odborného útvaru státní správy:

Souhlas Odboru výzkumu, vzdělávání a poradenství MZe:

Datum, jméno a podpis ředitele/ředitelky odboru: 21.12.2016

Razítko

.....
Ing. Pavlína Adam, Ph.D.

* Nehodící se škrtněte.



CERTIKOVANÁ METODIKA

Metodika GEPH na základě referenčního souboru složeného z domácích TD záznamů a Interbullem přepočtených mezinárodních MACE hodnot

Autoři

prof. Ing. Josef Příbyl, DrSc. (52 %)
Ing. Jiří Bauer, Ph.D. (8 %)
Ing. Anita Kranjčevićová (8 %)
Ing. Petr Pešek (8 %)
Ing. Jana Příbylová, CSc. (8 %)
doc. Ing. Luboš Vostrý, Ph.D. (8 %)
Ing. Ludmila Zavadilová, Ph.D. (8 %)

Oponenti

Ing. Zdenka Majzlíková
Česká plemenářská inspekce, Praha
doc. Ing. Karel Mach, CSc.

Metodika byla zpracována za podpory MZe ČR, úkol č. QJ1510139
(Celostátní informační systém genetického hodnocení hospodářských zvířat).

ISBN: 978-80-7403-157-1

Obsah

Seznam zkratk	4
I. Cíl metodiky	5
II. Vlastní popis metodiky	5
1. Úvod	5
2. Postup výpočtu PH / GEPH	6
2.1. Statistický model vyhodnocení	6
2.2. Příprava rodokmenů	7
2.3. Načítání TD údajů	8
2.4. Úprava MACE hodnot a základní odregresování	8
2.5. Úprava SNP souborů	9
2.6. Výběr genotypovaných jedinců	9
2.7. Slučování domácích a MACE hodnot	10
2.8. Parametrové soubory pro řešení BLUP-AM	12
3. Ověření	13
3.1. Výsledky řešení	13
4. Závěr	15
III. Srovnání „novosti postupů“	16
IV. Popis uplatnění Certifikované metodiky	16
V. Ekonomické aspekty	16
VI. Seznam použité literatury	16
VII. Seznam publikací, které předcházely metodice	17
Přílohy	19
1. Příprava rodokmenů	19
2. Načítání TD údajů	33
3. Úprava MACE hodnot a základní odregresování	44
4. Úprava SNP souborů	47
5. Výběr genotypovaných jedinců	55
6. Slučování domácích a MACE hodnot	73
7. Parametrový soubor pro řešení BLUP-AM	82
8. Výsledky řešení	85

Seznam zkratk

- A** – Rodokmenová matice příbuznosti stanovená na základě známých předků.
- A₂₂** - Rodokmenová matice příbuznosti stanovená na základě známých předků pouze u genotypovaných jedinců.
- AM** – Animal model. Metoda stanovení plemenné hodnoty s využitím všech příbuzenských vztahů mezi všemi jedinci (model jedince).
- BLUP** – Nejlepší lineární nestranná předpověď. Metoda vyhodnocení souboru údajů (Best Linear Unbiased Prediction).
- DRP** – Odregresované plemenné hodnoty (Deregressed Proofs), přibližné hodnoty pro DYD, když nejsou dostupné soustavy rovnic, ale pouze PH jedinců.
- DYD** – Opravené odchylky užitkovosti dcer býků, očištěné od systematických vlivů prostředí a plemenných hodnot matek, získané ze soustav rovnic BLUP-AM (Daughter Yield Deviation).
- G** – Genomická matice příbuznosti stanovená na základě podílu mezi jedinci shodných genetických markerů SNP.
- GEPH** – Genomická plemenná hodnota.
- H** – Plemeno holštýn.
- H** – Sloučená matice příbuznosti, která spojuje rodokmenovou (**A**) a genomickou (**G**) příbuznost.
- KU** – Kontrola užitkovosti.
- LP** – Legendův polynom, ortogonální polynom s nezávislými parametry.
- MACE** – Metoda souhrnného celosvětového hodnocení plemenů a přepočítávání celosvětových plemenných hodnot do měřítek a podmínek jednotlivých zemí (Multi Trait Across Country Evaluation).
- PGH** – Přímá genetická hodnota, stanovená na základě regresních koeficientů k jednotlivým genetickým markerům SNP (Direct Genetic Value).
- PH** – Plemenná hodnota.
- RRTDM** – Stanovení plemenné hodnoty na základě jednotlivých kontrol užitkovosti pomocí prokládání laktačních křivek polynomy s náhodnými regresními koeficienty (Random Regression Test Day Model).
- SNP** – Bodová mutace, jednonukleotidový polymorfismus (Single Nucleotide Polymorphism).
- ssGBLUP** – Jednokroková metoda stanovení genomické plemenné hodnoty společně v celé celostátní populaci genotypovaných i negenotypovaných jedinců. Navazuje na postupy BLUP – Animal model a slučuje ve výpočtu rodokmenové a genomické příbuznosti všech jedinců (single step Genomic Best Linear Unbiased Prediction).
- σ^2 – Rozptyl.
- tdr** – Záznam kontroly užitkovosti denního nádoje.
- w** – Efektivní počet případů, váha do vážené analýzy.

I. Cíl

Česká republika je členem mezinárodní chovatelské organizace Interbull. Za členství platí roční poplatky a může využívat služby, které Interbull poskytuje.

Cílem předkládané práce je postup genetického hodnocení chovných zvířat (stanovení plemenné hodnoty (PH) a genomické plemenné hodnoty (GEPH)) mléčné užitkovosti na základě „test day modelů“ s náhodnými regresemi (RRTDM), které pracují s jednotlivými kontrolami užitkovosti a prokládají u každého jedince laktační křivky. Při hodnocení je v navrženém postupu využita referenční populace opřená o sloučenou celou domácí kontrolu užitkovosti (KU) holštýnského skotu a veškeré dostupné údaje z Interbullu od genotypovaných i negenotypovaných jedinců.

Postup je znázorněn u vlastnosti kg mléka.

II. Popis metodiky

1. Úvod

Výsledek šlechtění dojeného skotu je přibližně z 1/3 ovlivněn výběrem plemeníků na stádo a přibližně ze 2/3 výběrem matek a otců býků. Selekcí programy menších populací dojeného skotu jsou silně ovlivněny dlouhodobými dovozy plemeníků a inseminačních dávek ze zahraničí. Domácí zvířata se jako rodiče plemeníků a jako plemeniči prosazují jen obtížně. Dochází k trvalému vyhlazovacímu křížení.

To má dva základní důsledky.

1. Chovatel má na výběr z celosvětové nabídky plemeníků, ovšem šlechtěných do jiných výrobních a ekonomických podmínek. Musí proto počítat pro své stádo s nižším přínosem, než by bylo v zemi původu plemeníka. Snížení přínosu lze odvodit z koeficientů korelací mezi zeměmi, které jsou pravidelně uveřejňovány Interbullem.

2. Dovážení jedinci mají slabou příbuzenskou vazbu na domácí kontrolu užitkovosti, čímž u dovážených jedinců a jejich mladých potomků, je nízká spolehlivost předpokládané domácí plemenné hodnoty (PH). Spolehlivost PH i samotná PH se upřesní (kladným, nebo záporným směrem) teprve, až bude v domácích podmínkách na dovážené jedince dostatečný počet dcer s užitkovostí. K tomu dojde až s časovým zpožděním, kdy již dříve bylo „naslepo“ o zvířatech rozhodnuto. U řady plemeníků, v důsledku omezeného počtu inseminací, zlepšení spolehlivosti nenastane nikdy. Trvale proto v rodokmenu mladých zvířat schází z otcovské strany významnější propojení na domácí užitkovost. Vzhledem k tomu, že šlechtění ovlivňují především plemeniči, jedná se o vážný nedostatek.

Narůstá podíl mladých genotypovaných býků, kteří jsou použiti v plemenitbě. Ti jsou ohodnoceni genomickou plemennou hodnotou (GEPH). V ČR jedнокrokovou metodou ssGBLUP (single step genomic evaluation), která umožňuje společné vyhodnocení a společný žebříček všech jedinců v populaci (Plemdat 2016) a ovlivňuje hodnoty i negenotypovaných jedinců. Jak GEPH, tak její spolehlivost u těchto mladých genotypovaných jedinců přímo závisí na velikosti „referenční populace“ spolehlivě prověřených starých jedinců, kteří byli ohodnoceni na té populaci, kde mají být mladí genotypovaní jedinci použiti. Vztah mladých jedinců k referenční populaci je prostřednictvím rodokmenové příbuznosti a prostřednictvím shody genetických markerů – SNP mladých jedinců s jedinci referenčními. Tyto genetické markéry navyšují spolehlivost předpovědi GEPH nad spolehlivost rodokmenové PH. Vzhledem k tomu, že mladí býci mají zahraniční původ, mají vazbu k domácí referenční populaci slabou a tím i nízkou spolehlivost GEPH. Rovněž dostupní starší genotypovaní býci, ohodnoceni dříve, mají vyšší vztah k užitkovostem zahraničním, než domácími.

Pro částečné překonání těchto těžkostí se nabízí využít v referenční populaci společně spolu s domácími užitkovostmi i údaje zahraniční, které jsou Interbullem pomocí metody MACE

přepočteny na naše domácí podmínky. Interbullový soubor holštýnského skotu zahrnuje přibližně 130.000 prověřených plemeníků, vesměs s vyššími spolehlivostmi, což je významný zdroj informací.

Včlenění MACE (dodatečných) hodnot do vnitrostátního genomického hodnocení ověřovali například Příbyl a kol. (2013, 2014b), Vandenplas a kol. (2014), VanRaden a kol. (2014), Bauer a kol. (2015), Vandenplas, Gengler (2015) a Weller a kol. (2015).

Dodatečné údaje jsou včleňovány prostřednictvím „odchylek užitekosti dcer“ DYD (daughter yield deviation), nebo prostřednictvím odregresovaných plemenných hodnot DRP (deregressed proofs), které vyjadřují „pseudo-fenotypové“ odchylky užitekosti. Vícekrokové metody stanovení GEPH (Meuwissen a kol., 2001, Szyda a kol., 2011) a rovněž „blending“ postupy jednokrokové metody ssGBLUP (Gao a kol., 2012, Su a kol. 2012) používají jako vstupní údaje právě DYD, nebo DRP.

Plemenné hodnoty mléčné užitekosti na Interbullu představují index (průměr) z PH několika, obvykle tří, normovaných laktací. U mléčné užitekosti jsou však v jednotlivých zemích pro stanovení plemenné hodnoty používány postupy „test day modelů“ s náhodnými regresemi RRTDM (Zavadilová a kol. 2005b, Plemdat 2016), které pracují s jednotlivými záznamy kontroly užitekosti a jsou přesnější než „laktační model“. Aby mohly být údaje z Interbullu a domácí kontroly užitekosti porovnávány, musí být převedeny na stejné měřítko. V tom případě je třeba MACE a ostatní dodatečné hodnoty odregresovat do denních záznamů mléčné užitekosti. Následně je možné použít stejný postup vyhodnocení metodou BLUP – animal model, jako u domácích souborů kontroly užitekosti (Liu a kol., 2004, 2014, Vostrý a kol., 2015).

2. Postup výpočtu PH / GEPH

Celý postup se skládá z navazujících kroků, které spočívají v postupné úpravě vstupních údajů a řešení soustavy rovnic. Příprava údajů je podřízena konečné podobě výpočtu PH podle níže uvedeného statistického modelu, který vychází z provozního postupu používaného při pravidelných hodnoceních Plemdatem.

2.1. Statistický model vyhodnocení

Záznamy denních užitekostí (tdr) jsou vyhodnoceny váženou analýzou RRTDM metodami BLUP a ssGBLUP animal model společně pro všechny jedince v populaci, genotypované i negenotypované. Laktační křivka podle záznamů kontroly užitekosti je vyjádřena Legendrovým polynomem (LP) se čtyřmi parametry:

$$\mathbf{f} = \mathbf{p} \cdot \mathbf{b} \quad (1)$$

kde:

\mathbf{b} = vektor regresních koeficientů (\mathbf{b}_{fg} pro pevné skupiny krav s předpokládaným stejným průběhem laktace, \mathbf{b}_{pe} pro náhodné trvalé prostředí uvnitř laktace krávy, \mathbf{b}_{an} pro náhodný genetický efekt jedince);

\mathbf{p} = vektor parametrů polynomu podle dne laktace. Prvky \mathbf{p} jsou stanoveny pro standardizovaný den laktace a proměnlivost každého parametru přibližně 1 (Schaeffer a kol., 2000, Plemdat 2016). Prvý prvek = 1, prvky 2 až 4 se mění podle dne laktace.

První člen polynomu vyjadřuje průměrnou užitekost, druhý až čtvrtý člen polynomu vyjadřují tvar a sklon laktační křivky.

Vyhodnocení je podle 3-laktačního RRTDM (Plemdat 2016):

$$\mathbf{Y}_k = \mathbf{X}_{TD}\mathbf{\beta}_k + \mathbf{X}_{fg}\mathbf{f}_{fg,k} + \mathbf{Z}_{pe}\mathbf{f}_{pe,k} + \mathbf{Z}_{an}\mathbf{f}_{an,k} + \mathbf{e}_k \quad (2)$$

kde:

\mathbf{Y} = vektor denních záznamů užitkovosti krávy na laktaci $k <1,2,3>$,
 \mathbf{X}_{TD} = matice plánu pokusu dnů kontroly užitkovosti,
 $\boldsymbol{\beta}$ = vektor odhadovaných vlivů skupin vrstevnic uvnitř dnů kontroly užitkovosti ve stádě a na dané laktaci (pevný efekt),
 $\mathbf{X}_{fg,k}$ = matice plánu pokusu pro průměrné LP laktačních křivek uvnitř skupin krav s předpokládaným stejným průběhem laktace k (Zavadilová a kol., 2005b),
 $f_{fg,k}$ = pevný průměrný LP uvnitř laktace pro skupinu krav na laktaci k ,
 \mathbf{Z}_{pe} = matice plánu pokusu pro LP trvalého prostředí krávy uvnitř laktace,
 $f_{pe,k}$ = LP trvalého prostředí krávy na laktaci k , náhodný efekt s předpokládaným rozdělením četností

$$\mathbf{b}_{pe} \sim N(0, \mathbf{I} \otimes \mathbf{LE}), \quad (3)$$

\otimes = přímé násobení matic,
 \mathbf{I} = jednotková matice s velikostí podle počtu krav s užitkovostí,
 \mathbf{LE} = (12 x 12) kovarianční matice náhodných regresních koeficientů Legenderova polynomu pro trvalé prostředí jedince přes tři laktace každé krávy (Zavadilová a kol., 2005b),
 \mathbf{Z}_{an} = matice plánu pokusu pro LP genetického efektu jedince uvnitř laktace,
 $f_{an,k}$ = genetický LP jedince na laktaci k , náhodný efekt s předpokládaným rozdělením četností

$$\mathbf{b}_{an} \sim N(0, \mathbf{A} \otimes \mathbf{LG}) \text{ při metodě BLUP – stanovení PH,} \quad (4)$$

$$\text{nebo } \sim N(0, \mathbf{H} \otimes \mathbf{LG}) \text{ při metodě ssGBLUP – stanovení GEPH,}$$
 \mathbf{A} = rodokmenová matice příbuznosti všech jedinců přes pět generací předků,
 \mathbf{H} = sloučená matice příbuznosti (Misztaľ a kol., 2009, Příbyl a kol., 2014b) s váhou $\lambda = 0.80$ pro inverzi \mathbf{G} :

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & \lambda(\mathbf{G}^{-1} - \mathbf{A}_{22}^{-1}) \end{bmatrix}, \quad (5)$$

\mathbf{A}_{22} = rodokmenová matice příbuznosti pouze genotypovaných jedinců,
 \mathbf{G} = genomická matice příbuznosti genotypovaných jedinců založená na odchylkách od průměrných alelových četností hodnocené populace, standardizovaná na diagonálu (Forni a kol., 2011) a na průměr prvků v matici \mathbf{A}_{22} (Vitezica a kol., 2011),
 \mathbf{LG} = (12 x 12) kovarianční matice náhodných regresních koeficientů Legenderova polynomu pro genetický efekt jedince přes tři laktace každé krávy,
 \mathbf{e} = vektor náhodných chyb denních záznamů užitkovosti na laktaci k s předpokládaným rozdělením četností

$$\mathbf{e} \sim N(0, \mathbf{D}\sigma_{rk}^2), \quad (6)$$
 \mathbf{D} = diagonální matice s převrácenými hodnotami vah (w), které zohledňují změnu proměnlivosti během laktace u domácí populace krav, nebo efektivní počet „pseudo-fenotypových“ odchylek užitkovostí u býků z Interbullu na laktaci k ,
 $\sigma_{r,k}^2$ = průměrný chybový rozptyl záznamů užitkovosti uvnitř laktace k .

Dílčí hodnoty PH a GEPH za celé laktace (k) každého jedince jsou stanoveny na základě prvního prvku vypočteného vektoru $\mathbf{b}_{an} * 305$. Celková PH jedince je selekční index, stanovený jako prostý průměr za tři dílčí PH za laktace každého jedince (váhy v indexu 1/3 : 1/3 : 1/3).

2.2. Příprava rodokmenů

Z původů domácích zvířat a původů jedinců na Interbullu je vytvořen společný soubor rodokmenů. Z rodokmenů jsou vyloučeni jedinci s nevěrohodnými údaji a jedinci s „umělými čísly“, která byla doplňována v minulosti. Při slučování rodokmenů nastávají případy, že některý

jedinec je v různých zemích a u různých chovatelských organizací jinak označen. Jsou proto porovnávány zmnožená označení jedinců a převáděna na označení v zemi původu. Je sjednoceno označování země původu, které bývá v rodokmenech uváděno několika různými způsoby.

Jedinci jsou přechíslováni, aby nejstarší měli nejnižší čísla, jak vyžadují návazné postupy sestavování velkých matic příbuznosti. K číslu jedince je proto doplněn rok narození, země původu a podíl holštýna (H). Pokud u některého jedince není rok známý, nebo roky mezi generacemi jsou nevěrohodné, doplní se uměle odstup generací 5 let. Doplnění roků, zemí a podílů H se iterativně opakuje, aby byly údaje postupně doplněny všem zvířatům v mnohogenračním rodokmenu. Vyskytují se „kruhové rodokmeny“, kdy pravděpodobně chybou v minulosti se stejné označení jedince objeví u některého prapředka. To se navíc projeví v neshodě roku narození ve sledu generací. Takový předek je z rodokmenu vyřazen. Nevěrohodní jedinci jsou zapisováni do souborů.

Pro sloučení a úpravu rodokmenů byl vytvořen program v prostředí SAS (příloha 1), ze kterého je zřejmý podrobný pracovní postup a je ho možno použít jako podklad pro tvorbu provozního programu v jiném uživatelském jazyku. Bližší údaje jsou pro zájemce dostupné u autora.

2.3. Načítání TD údajů

Ze souborů kontroly užitkovosti jsou načteny jednotlivé záznamy za první tři laktace krav. Jsou vyřazeny nevěrohodné údaje, které přesahují očekávaná rozmezí užitkovosti. Rovněž jsou vyřazeny krávy s neobvyklým označením a s nevěrohodným údajem o věku krávy. Pokud není známá délka mezibřezosti, nebo její délka přesahuje 190 dnů, dosadí se 190 dnů.

Je porovnán počet záznamů na každou laktaci a krávu, počty záznamů užitkovosti (počty vrstevnic) ve stáji a kalendářním dnu kontroly užitkovosti. Užitkovosti z laktací s méně než třemi záznamy a dny kontroly užitkovosti s méně než osmi záznamy uvnitř téže laktace jsou vyřazeny. Vyřazování probíhá iterativně v několika krocích, aby byly současně splněny obě podmínky.

Na základě věku při otelení, mezidobí, mezibřezosti, ročního období a roku jsou vytvářeny podle jednotlivých laktací skupiny krav pro „pevnou regresi“ ($X_{fg,k}$ v rovnici 2), které by měly mít podobný průběh laktace. Sestavování skupin vychází z postupů používaných při pravidelném vyhodnocování (Plemdat 2016). Pokud počet záznamů kontrol užitkovosti ve skupině je malý (méně než 1000), skupiny jsou slučovány. Slučování proběhlo na základě posouzení tak, aby se slučované skupiny ve většině hledisek shodovaly. Skupiny jsou přechíslovány.

Pro přípravu užitkovosti byl vytvořen program v prostředí SAS (příloha 2), ze kterého je zřejmý podrobný pracovní postup a je ho možno použít jako podklad pro tvorbu provozního programu v jiném uživatelském jazyku.

2.4. Úprava MACE hodnot a základní odregresování

U Interbullového souboru holštýnských býků byla provedena oprava na jednotné označení zvířat (jako v 2.2). Byly vyřazeny býci, kteří pocházejí ze zemí s velmi odlišnými podmínkami chovu od ČR, rokem narození před 1985, podílem holštýna (H+R) pod 85% a spolehlivostí PH pod 0,65.

Pro odregresování lze použít postupy, které rozkládají PH na příspěvky příbuzných jedinců (Jairath a kol., 1998, VanRaden a kol., 2014, Vostrý a kol., 2014), aby bylo vyloučeno opakované započítání stejných užitkovostí. Odregresování velkých souborů s navzájem příbuznými jedinci je počítačově náročné. DRP jsou již očištěny od vlivu ostatních systematických vlivů prostředí, nicméně při odregresování je třeba dbát na správné stanovení vah, které jednotlivým pozorováním náleží, a které závisí na spolehlivostech PH. Správné stanovení vah je důležité především v případě, když se odregresovávají PH s velmi rozdílnými spolehlivostmi (Calus a kol., 2016). Vandenplas a Gengler (2015) však došli u populace skotu v Belgii k závěru, že hodnoty plemeníků z Interbullu jsou dobře použitelné i při jednoduchých,

méně přesných postupech odregresování. Výsledky následné GEPH jsou dostatečně spolehlivé, i když jsou DRP plemeníků, v důsledku příbuznosti, navzájem závislé. Může se proto přistoupit i k pouze přibližným a počítačově méně náročným postupům odregresování.

V našem případě jsme použili pouze plemeníky se spolehlivostí 0,65 a více (po přepočtu na podmínky ČR v průměru 0,72) a proto jsme odregresování provedli „pouze“ vydělením PH její spolehlivostí. Na základě Interbullové spolehlivosti jsme stanovili efektivní počet případů, jakému spolehlivost odpovídá:

$$w = ((1-h^2)/h^2)(rel/(1-rel)) \quad , \quad (7)$$

kde:

w = váha (efektivní počet případů) pro odregresovanou plemennou hodnotu,
 h^2 = koeficient dědivosti indexu ze tří laktací v našich podmínkách. Použili jsme hodnotu 0,39, která pro mléko vychází z kovariančních matic použitých pro RRTDM (Plemdat 2016),
 rel = spolehlivost vstupní plemenné hodnoty.

Z odregresovaných hodnot jsme vyřadili odlehlá pozorování, která přesahovala ± 3 fenotypové směrodatné odchylky ($3 \cdot 794$) kg mléka.

Pro úpravu Interbullových údajů byl vytvořen program v prostředí SAS (příloha 3), ze kterého je zřejmý podrobný pracovní postup a je ho možno použít jako podklad pro tvorbu provozního programu v jiném uživatelském jazyku.

2.5. Úprava SNP souborů

Je načítáno několik souborů SNP údajů. Byla provedena oprava na jednotné označení zvířat (jako v 2.2). Alely v lokusech byly převedeny na jednotná čísla, vyjadřující počet „druhých“ alel $\langle 0,1,2 \rangle$. Lokusy s chybějícími záznamy byly vyřazeny. Jedinci se mohou v různých souborech opakovat, do konečného souboru byl vybrán každý pouze jednou. Byly použity pouze čipy jedinců s podílem holštýna 85% a více.

Na základě četnosti alel v jednotlivých lokusech byly ze souboru vyřazeny údaje, kde bylo nebezpečí, že došlo k chybám buď v laboratořích, nebo v označování vzorků a zvířat. Jmenovitě se jedná o vyřazení jedinců, kteří mají $< 90\%$ použitelných lokusů. Dále byly vyřazeny lokusy, které chybí u $> 5\%$ jedinců a lokusy, které mají malou vypovídací schopnost, to znamená, že jsou téměř homozygotní a četnost jedné alely převyšuje 95%. Jedinci a lokusy byli přečíslováni.

Pro úpravu SNP údajů byly vytvořeny programy v prostředí SAS (příloha 4), ze kterých je zřejmý podrobný pracovní postup a je možno je použít jako podklad pro tvorbu provozního programu v jiném uživatelském jazyku.

2.6. Výběr genotypovaných jedinců

Někteří jedinci při genomickém hodnocení vybočují z obecně platných závislostí (odlehlá pozorování). Důvodů může být několik. Například chybné a nevěrohodné PH u prověřených býků, nízká spolehlivost PH, chybná označení laboratorních vzorků a jedinců, nebo například původ z jiné populace, kdy sledované SNP lokusy se projevují jinak na odlišném genetickém pozadí, které může pocházet od vzdáleného prapředka. Důsledkem je, že u určitého podílu prověřených jedinců nesouhlasí vstupní PH s jejich následnou přímou genetickou hodnotou (PGH), stanovenou pomocí genetických markerů SNP (Pešek a kol., 2015).

U referenčních prověřených genotypovaných jedinců s Interbullovou PH proto víceřadovou metodou ověřujeme regresní koeficienty na jednotlivé SNP lokusy a výslednou PGH. Vyřazujeme jedince, u kterých se vstupní PH a výstupní PGH liší o více než 2 směrodatné odchylky absolutní chyby.

Dále posuzujeme shodu rodokmenové a genomické příbuznosti (matice \mathbf{G} a \mathbf{A}_{22}). Pokud jsou velké rozdíly, je pravděpodobné, že buď rodokmenové, nebo genomické označení není

správné. Pokud jedinec má rozdíl v uvedených způsobech stanovení příbuznosti $> 0,5$ pro více než tři jiné jedince, nebo k jednomu jedinci rozdíl v příbuznosti $> 0,6$, potom ho vyřazujeme.

Pro výběr genotypovaných jedinců byly vytvořeny programy v prostředí SAS (příloha 5), ze kterých je zřejmý podrobný pracovní postup a je možno je použít jako podklad pro tvorbu provozního programu v jiném uživatelském jazyku.

2.7. Slučování domácích a MACE hodnot

Nejprve je vytvořen soubor užitekostí a soubor původů pro samostatné vyhodnocení domácích PH a GEPH podle RRTDM. Soubor zahrnuje užitekostí od roku 2000, což je přibližně ze třech nejmladších generací (Lourenco 2015) a je přidán jeden ročník jalovic.

Dále je vytvořen soubor odregresovaných užitekostí a soubor původů pro samostatné vyhodnocení PH a GEPH pro údaje MACE z Interbullu. Za jedince s užitekostí v tomto případě považujeme přímo býka. Vyhodnocení Interbullových souborů je podle jednoduchého váženého AM s jedním pevným a jedním náhodným efektem. Váha je stanovena podle spolehlivosti Interbullových PH podle (7).

Do příbuznosti u obojích souborů jsou doplněny skupiny neznámých předků podle země původu, roku narození a podílu H. Máločetné skupiny neznámých předků slučujeme, aby dosáhly četností nejméně 100 jedinců. Volbou rozhodujeme, zda budou tyto skupiny při vyhodnocení použity.

Uvedené soubory slučujeme pro vyhodnocení spojených souborů. Pokud má býk v domácích podmínkách více jak 10 dcer s užitekostí, je zároveň zařazen i do Interbullových souborů. Na Interbullu má však potomků mnohem více i z ostatních zemí. Aby nedošlo ke zdvojování stejných podkladových údajů rozhodujeme, zda budou z Interbullu vyloučeni býci, kteří mají domácí dcery. Vzhledem k tomu, že domácích dcer je jen malá část ze všech dcer býka, ponecháváme v tomto případě v Interbullových souborech všechny býky, i když mají domácí dcery.

Z odregresovaných MACE hodnot býků stanovujeme „pseudo-fenotypové“ průměrné denní odchylky užitekostí jedince na jednotlivých laktacích. Údaje odregresovaných denních odchylek užitekostí přidáváme do souboru užitekostí a zařazujeme všechny do samostatné skupiny \mathbf{X}_{TD} dne kontroly užitekostí podle rovnice (2) a do samostatné skupiny $\mathbf{X}_{fg,k}$ pro průměrné LP laktačních křivek podle rovnice (2). Průměrnou denní odchylku užitekostí jedince na dané laktaci stanovíme (Liu a kol., 2014):

$$tdpr_k = DRPI_1 * s_{tdr,k} / s_{DRPI} \quad , \quad (8)$$

kde:

$tdpr_k$ = průměrná odregresovaná odchylky denní užitekostí jedince na laktaci k ,

$DRPI_1$ = odregresovaná hodnota MACE z Interbullu (hodnota indexu ze tří laktací), jak uvedeno v 2.4.,

$s_{tdr,k}$ = směrodatná odchylka denní fenotypové užitekostí na laktaci k , použili jsme 60% ze směrodatné odchylky naměřené užitekostí (Kranjčevićová a kol. 2016),

s_{DRPI} = směrodatná odchylka odregresovaných hodnot MACE z Interbullu.

Z uvedeného vyplývá u jedince na všech třech laktacích stejná předpokládaná PH. Protože máme u jedince na laktaci pouze jednu průměrnou hodnotu, nelze tvarovat laktační křivku. Prvky 2 až 4 vektoru \mathbf{p} v (1) jsou proto nula, nebo hodnoty velmi blízké nule, aby kovarianční matice bylo možno invertovat. V našem případě jsme pro tvorbu těchto prvků použili náhodná čísla vynásobená hodnotou 0.00000001.

Váhy (w) přidělené jednotlivým odchylkám užitekostí $tdpr_k$, které jsou použity v (6), musí odpovídat spolehlivostem vstupních $DRPI_1$ z Interbullu. Váhy (w) představují efektivní počet záznamů užitekostí během laktace. Spolehlivost dílčí PH za laktaci závisí na tomto efektivním počtu a genetických parametrech:

$$rel_k = w_k * h_k^2 / (1 + (w_k - 1) * rep_k) \quad , \quad (9)$$

kde:

rel_k = spolehlivost dílčí PH jedince na laktaci k ,
 w_k = efektivní počet odchylek užítkovosti na laktaci k ,
 h_k^2 = vnitrolaktační koeficient dědivosti na laktaci k , Použili jsme hodnotu 0,20,
 rep_k = vnitrolaktační koeficient opakovatelnosti na laktaci k . Použili jsme hodnotu 0,74.

Z jedné laktace na další přežívá v našich podmínkách přibližně 70% krav. Proto s pořadím laktace efektivní počet klesá, $w_2 = 0.7 * w_1$ a $w_3 = 0.7 * w_2$.

Spolehlivost celkové PH jedince (spolehlivost indexu) (rel_I) na základě dílčích PH ze třech laktací lze přibližně stanovit z rovnice selekčního indexu:

$$\mathbf{Q} * \boldsymbol{\beta} = \mathbf{C} * \boldsymbol{\alpha} \quad , \quad (10)$$

kde:

\mathbf{Q} = kovarianční matice (3 x 3) dílčích PH za prvou, druhou a třetí laktaci, které zahrnují dílčí spolehlivosti podle (9) ,
 $\boldsymbol{\beta}$ = vektor vah laktací v indexu,
 \mathbf{C} = kovarianční matice (3 x 3) mezi plemennými hodnotami a genotypem jedince, které zahrnují dílčí spolehlivosti podle (9) ,
 $\boldsymbol{\alpha}$ = vektor ekonomických hodnot pro prvou, druhou a třetí laktaci (1/3; 1/3; 1/3).

Spolehlivost indexu je:

$$rel_I = (\boldsymbol{\beta}' * \mathbf{C} * \boldsymbol{\alpha}) / (\boldsymbol{\alpha}' * \mathbf{T} * \boldsymbol{\alpha}) \quad , \quad (11)$$

kde:

rel_I = spolehlivost indexu jedince za tři laktace,
 \mathbf{T} = genetická kovarianční matice (3 x 3) mezi laktacemi.

Vážený součet užítkovostí na všech třech laktacích by měl poskytnout stejnou spolehlivost celkové PH, jako vstupní MACE údaje z Interbullu. Jestliže je spolehlivost vstupní Interbullové PH vysoká, rel_I (11) její hodnoty nedosáhne ani při třech laktacích a vysokých w . Jak vyplývá z (3), v RRTDM pracujeme s trvalým prostředím jedince. Jmenovatel ve vzorci spolehlivosti (9) je proto zřetelně větší než číselník, neboť opakovatelnost je vždy vyšší než dědivost. Přistupujeme proto k opakování hodnot indexů za jedince. Celková spolehlivost na základě n opakování (rel_T) je stanovena přibližně z rovnice selekčního indexu:

$$\mathbf{U} * \boldsymbol{\delta} = \mathbf{M} \quad , \quad (12)$$

kde:

\mathbf{U} = kovarianční matice velikosti ($n \times n$) pro (n) opakování hodnot indexů jedince. Prvky matice jsou stanoveny na základě spolehlivostí jednotlivých indexů (rel_I) podle 11.
 $\boldsymbol{\delta}$ = vektor vah jednotlivých indexů,
 \mathbf{M} = kovarianční matice ($n \times 1$) jednotlivých indexů k celkové genetické hodnotě. Prvky matice jsou stanoveny na základě spolehlivostí jednotlivých indexů (rel_I) podle 11. Proměnlivost celkové genetické hodnoty = 1.

Celková spolehlivost jedince je:

$$rel_T = \boldsymbol{\delta}' * \mathbf{M} \quad , \quad (13)$$

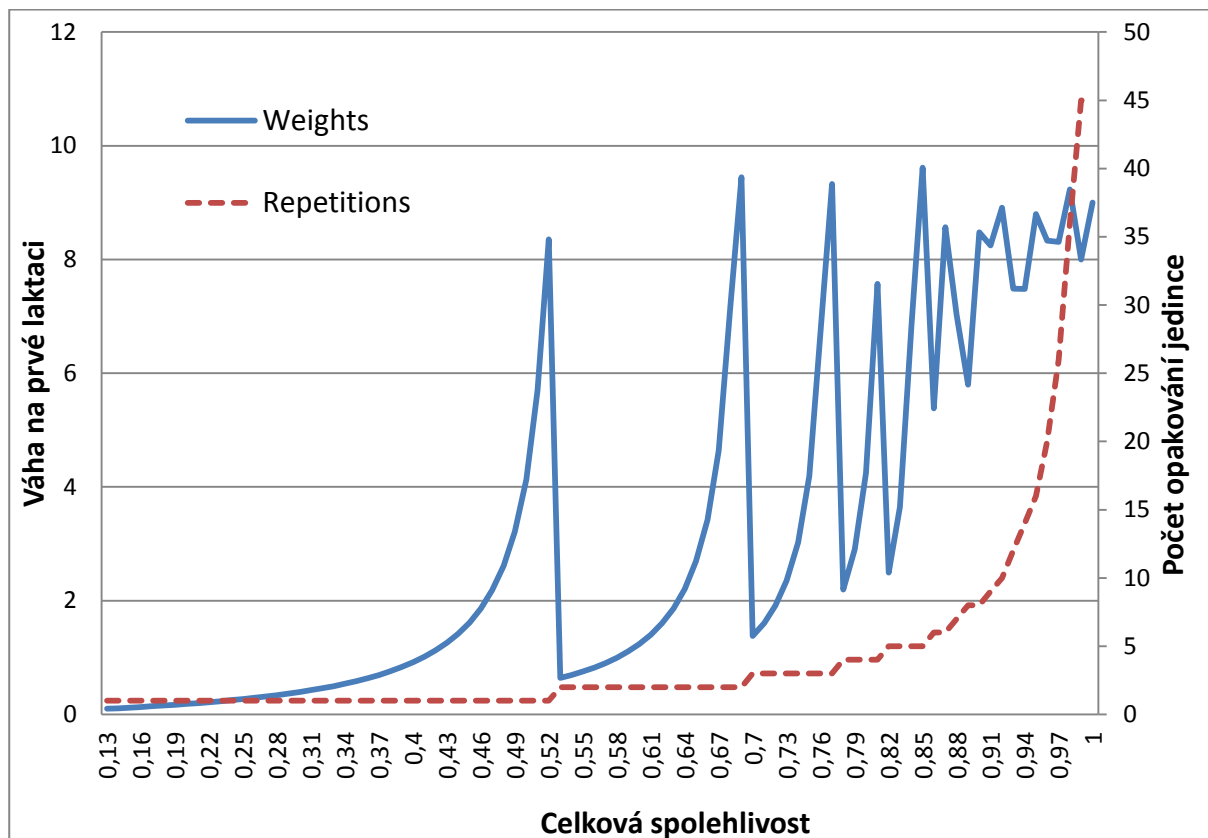
kde:

rel_T = celková spolehlivost sloučených indexů jedince.

Hodnoty vah (w_1, w_2, w_3) a počty opakování (n) nelze na základě spolehlivostí vstupních PH z Interbullu stanovit přímo. Jsou proto na základě (9) až (13) dopředu programem během výpočtu

stanovovány tabulkové hodnoty, které jsou jednotlivým DRP_I přiřazeny. Na základě (n) je každý z Interbullu dodávaný jedinec v souboru opakován. Hodnoty (w_1) a (n) jsou uvedeny na příloženém obrázku 1. Z obrázku vyplývá, že do celkové spolehlivosti 0,50 je jedinec přidán do souboru 1x a postupně narůstá váha přidávané užítkovosti až nad 8, což představuje 8 záznamů užítkovostí během laktace. Od celkové spolehlivosti 0,50 do 0,70 je jedinec přidán do souboru 2x, nad tuto hodnotu spolehlivosti počet opakování jedince narůstá, od spolehlivosti 0,85 poměrně prudce.

Obrázek 1. Váha užítkovosti na první laktaci (w_1) a počet opakování indexu jedince (n) v závislosti na celkové spolehlivosti jedince (rel_T).



Pro přípravu sloučených souborů užítkovostí a rodokmenů byl vytvořen program v prostředí SAS (příloha 6), ze kterého je zřejmý podrobný pracovní postup a lze jej použít jako podklad pro tvorbu provozního programu v jiném uživatelském jazyku.

2.8. Parametrové soubory pro řešení BLUP-AM

Pro stanovení PH a GEPH pomocí programů BLUPF90-family uvádíme v návaznosti na Příbyla a kol. (2014a) parametrový soubor (příloha 7), který byl použit pro stanovení genetických parametrů.

Podle velikostí souborů při genetickém hodnocení je třeba změnit počty úrovní jednotlivých efektů. Parametrový soubor odpovídá stanovení GEPH pomocí postupu ssGBLUP. Do výpočtu načítáme předem připravenou matici genomické příbuznosti „gemat2“. Přestože nenačítáme jednotlivé SNP, je třeba programu přiřadit soubor „genot2“. Do tohoto souboru uměle dosadíme pro každého genotypovaného jedince několik SNP.

Pro stanovení PH je možno použít stejný parametrový soubor, ze kterého je třeba vyřadit (zadat notový křížek) „OPTION“ týkající se „genot2“, „tunedG“, „gemat2“.

3. Ověření

Postup výpočtu je ověřen na rozděleném souboru u mléčné užitkovosti. Údaje celostátních souborů za roky užitkovosti 2000 až 2012 jsou použity pro předpověď PH a GEPH mladých jedinců bez užitkovosti. Jedná se o jalovice a mladé genotypované býky do roku 2012 bez potomstva. Předpovězené hodnoty jsou porovnány s výsledky vyhodnocení celého souboru za roky 2000 až 2016, které již zahrnují vlastní užitkovosti otelených jalovic a u býků výsledky užitkovosti dcer.

Jsou navzájem porovnány spolehlivosti předpovědí v roce 2012 na základě PH a GEPH a při použití pouze domácích, pouze Interbulových, či spojených souborů užitkovosti. Pro každý postup je stanovena „validovaná spolehlivost“ (Gao a kol., 2012):

$$vrel = (r_{PV12, DRP16})^2 / rel16 \quad (14)$$

kde:

vrel = validovaná spolehlivost předpovědi,

$(r_{PV12, DRP16})^2$ = determinační koeficient předpovědi v roce 2012 pro odregresované hodnoty v roce 2016. Determinační koeficient byl stanoven metodou nejmenších čtverců z regresní závislosti DRP16 na PH12, nebo na GEPH12.

rel16 = průměrná spolehlivost odregresovaných hodnot podle PH, nebo GEPH v roce 2016 (Bauer a kol. 2015).

3.1. Výsledky řešení

Pro posouzení byly použity celostátní soubory mléčné užitkovosti a veškeré dostupné údaje z Interbullu, upravené podle výše uvedených postupů. Byly použity veškeré dostupné genotypy H skotu, včetně „red“ holštýna. Nevěrohodné údaje a odlehlé hodnoty, které vybočují ze stanovených závislostí byly vyřazeny.

Rozsahy použitých souborů jsou uvedeny v tabulce 1.

Tabulka 1. Rozsahy souborů pro výpočty PH a GEPH.

Soubor	Domáci 2012	Interbull 2012	D + I 2012	Domáci 2016
Jedinců s užitkovostí	719.545	102.649	822.194	976.283
Jedinců v rodokmenu	1.478.545	283.571	1.704.294	1.725.859
Domácích užitkovostí	11.917.733		11.917.733	16.423.305
Interbull. užitkovostí		102.649	1.088.445	
Součet domácích vah	13.266.069		13.266.069	18.272.116
Součet Interbull. vah		501.851	2.946.914	
Vyřazeno genotypů podle alel	370	370	370	463
odlehlých hodnot	160	160	160	39
příbuznosti	23	23	23	18
Použito genotypů	5.162	5.162	5.162	5.444
z toho prověřených	1.340	3.428	3.428	2.034
Použito lokusů	40.592	40.592	40.592	40.566
Rovnic v modelu	26.726.663	283.572	30.667.502	32.858.308
Použitá paměť GB	17,5	4,5	18,5	23,5
Stroj čas výpočtu PH	26 hod.	15 min.	27 hod.	29 hod.
Počty iterací PH	2.766	398	2.893	2.879
Stroj čas výpo. GEPH	31 hod.	27 min.	32 hod.	36 hod.
Počty iterací GEPH	2.873	400	2.987	2.956

Domácí soubor pro předpověď v roce 2012 zahrnoval 11,9 mil. užitkovostí se součtem vah 13,3 mil. a bylo použito 5.162 genotypovaných jedinců, z nichž však pouze 1.340 mělo v domácím souboru užitkovost dcer. V samostatném souboru z Interbullu bylo použito 102.649 odregresovaných užitkovostí, se součtem vah 501.851 a bylo použito rovněž 5.162 genotypovaných jedinců, z nichž však 3.428 mělo vlastní Interbullovou PH. „Referenční“ populace genotypovaných jedinců na Interbullu je tedy 2,6x větší, než u samostatného domácího souboru. Ve sloučeném souboru (D+I) jsou použiti jak užitkovosti domácí, tak z Interbullu a všichni prověřeni genotypovaní býci, jako v souboru z Interbullu. Celkem je ve sloučeném souboru 13 mil. užitkovostí se součtem vah 16,2 mil.

Soubor domácí a soubor sloučený se liší počtem zahrnutých užitkovostí a počtem referenčních genotypovaných jedinců. Obojí se projevuje ve výsledcích.

Vyhodnocení souboru z Interbullu bylo jednoduchým animal modelem s genetickými parametry, které odpovídají indexu za tři laktace. Soubor domácí a soubor sloučený byly vyhodnoceny RRTDM, kde vstupují kovarianční matice. Proto také odregresování pro samostatný soubor z Interbullu a soubor sloučený vede k jinému počtu vložených Interbullových užitkovostí a jinému součtu vah užitkovostí.

Soubor z roku 2016 obsahuje 16,4 mil. užitkovostí se součtem vah 18,3 mil. a bylo použito 5.444 genotypovaných jedinců, z nichž 2.034 byli býci prověřeni v domácích podmínkách.

Výpočty PH i GEPH byly ukončeny při dosažení konvergenčního hlediska $1 \cdot 10^{-14}$ a vyžadovaly více jak den strojového času počítače.

Aby bylo možno výsledky v roce 2012 a 2016 porovnat byly standardizovány na průměr a směrodatnou odchylku plemenných hodnot podle 2.381 prověřených býků, kteří v obou souborech měli alespoň 60 laktací dcer. V průměru tito býci měli 280 laktací dcer v souboru 2012 a 323 laktací dcer v souboru 2016.

Standardizované hodnoty byly použity pro stanovení regresních závislostí, jak DRP podle skutečné užitkovosti v roce 2016 závisí na předpovědích podle PH a GEPH v roce 2012.

Závislosti byly stanoveny pro 390 mladých genotypovaných býků v roce 2012 bez dcer, kteří měli v roce 2016 nejméně 20 laktací dcer. V průměru tito býci měli v roce 2016 na třetí laktaci 9 dcer, druhých + třetích laktací 32 a v součtu za první, druhé a třetí otelení 107 laktací dcer. Dále byly závislosti stanoveny pro 24.146 jalovic z roku 2012, které nebyly genotypovány, a které dosáhly do roku 2016 tři laktace.

V tabulce 2 (příloha 8) jsou uvedeny determinační koeficienty, parametry regresní závislosti odregresovaných hodnot v roce 2016 na předpovědích mladých zvířat v roce 2012, spolehlivosti v roce 2016 a validované spolehlivosti předpovědí v závislosti na metodě předpovědi (podle PH, nebo GEPH) a na vstupních souborech (domácí KU, Interbullové MACE hodnoty všech býků, spojené soubory). Pro posouzení samostatného vlivu počtu genotypovaných referenčních býků a velikostí souborů s užitkovostmi, byly u Interbullových a sloučených souborů použity buď všechny na Interbullu dostupné hodnoty, nebo pouze hodnoty genotypovaných jedinců..

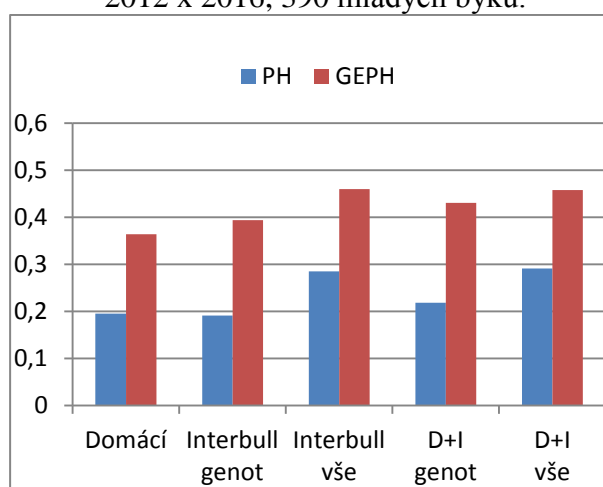
Očekávané hodnoty regresních koeficientů jsou 1. Z výsledků vyplývá, že u mladých býků jsou předpovědi všemi postupy nadhodnoceny (regresní koeficienty menší než 1) a u jalovic jsou předpovědi všemi postupy podhodnoceny (regresní koeficienty větší než 1). Tato skutečnost se projevuje přesto, že všichni jedinci jsou hodnoceni společným výpočtem se vzájemnou provázaností pomocí příbuznosti.

Hlavní výsledky z tabulky 2 jsou uvedeny na obrázcích 2 a 3. Jak vyplývá z výsledků, je validovaná spolehlivost předpovědi budoucí PH mladých plemenů podle rodokmenové PH nízká, 0,195 a podle GEPH se zvýší na 0,364. Při samostatné použití z Interbullu jen genotypovaných býků, je spolehlivost předpovědi mladých býků podle PH 0,192 a podle GEPH 0,394. V důsledku většího počtu prověřených genotypovaných býků na Interbullu je předpověď pomocí GEPH poněkud lepší, než pomocí PH podle domácího souboru. Při použití

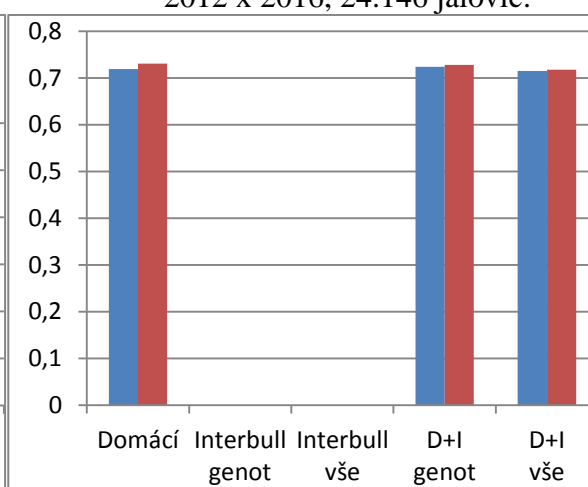
samostatných souborů z Interbullu se všemi dostupnými hodnotami, genotypovaných i negenotypovaných jedinců, mají tyto předpovědi spolehlivost vyšší, a to 0,285 a 0,460. Při použití sloučených domácích a Interbullových souborů mají předpovědi spolehlivost podobnou, jako podle údajů z Interbullu. I zde se projevuje, zda byli k domácímu souboru přidáni pouze genotypovaní, či všichni býci z Interbullu. Zvětšení souborů se více projevuje při předpovědích podle PH, než předpovědích podle GEPH. Při zahrnutí všech jedinců spolehlivost předpovědi dosahuje podle PH 0,291 a podle GEPH 0,458. Předpovědi u jalovic jsou podstatně spolehlivější, kolem 0,72 a nejsou v podstatě rozdíly mezi použitými vstupními soubory užitkovostí a metodami stanovení plemenných hodnot (BLUP nebo ssGBLUP).

Jak vyplývá z tabulek 2 (příloha 8), genomické plemenné hodnoty v roce 2016 jsou předpovídány s podobnými spolehlivostmi, jako plemenné hodnoty.

Obrázek 2. Validovaná spolehlivost 2012 x 2016, 390 mladých býků.



Obrázek 3. Validovaná spolehlivost 2012 x 2016, 24.146 jalovic.



4. Závěr

Cílem je dosáhnout co nejvyšší spolehlivosti předpovědí budoucí hodnoty u mladých chovných zvířat. Je žádoucí pro to využít všechny dostupné zdroje.

Zařazení zahraničních Interbullových údajů, přepočtených pomocí MACE na podmínky ČR, do referenční domácí populace významně přispívá ke zvýšení spolehlivosti hodnocení mladých plemenných býků. Zvýšení spolehlivosti je jak u PH, tak GEPH, u PH je vliv přidaných Interbullových údajů větší. Projevuje se zde jednak zvětšení souborů užitkovostí, ale i zvýšení počtu referenčních prověřených genotypovaných plemenů.

Důležité je především zvýšení spolehlivosti předpovědi PH, neboť GEPH je nadstavba nad PH. Ve spolehlivosti GEPH se mohou projevovat obtíže spojené se správným výběrem genotypovaných jedinců.

Zlepšení se projevuje při použití samostatných Interbullových údajů, nebo při použití spojených domácích a Interbullových souborů. Přednost má použití spojených souborů, které umožňuje hodnotit všechny jedince v domácí populaci.

Genomická plemenná hodnota významně zlepšuje spolehlivost předpovědi PH mladých býků u všech třech použitých souborů. Předpověď pomocí GEPH u sloučených souborů dosahuje přibližně 2,4x vyšší spolehlivost, než předpověď pouze podle domácí PH a 1,3x vyšší spolehlivost, než podle domácí GEPH.

Spolehlivost předpovědí u býků, kteří jsou většinou z dovozů, je podstatně nižší, než spolehlivost předpovědí u jalovic. Jalovice jsou potomky zvířat z domácí populace a mají propojení na užitkovost v obou větvích rodokmenu.

Aby bylo možno výsledky porovnat, pracovali jsme s rozděleným souborem (částečným do roku 2012 a celým do 2016). Předpovědi jsou založeny na údajích do roku 2012, kdy byl počet referenčních genotypovaných prověřených býků značně omezen. S nárůstem počtu genotypovaných jedinců lze v budoucnosti očekávat spolehlivost předpovědí GEPH vyšší.

III. Srovnání „novosti postupů“

Česká republika používá pro genomické hodnocení pokročilý jednokrokový postup ssGBLUP, který překonává metodické nedostatky ostatních vícekových postupů. Umožňuje společné hodnocení všech jedinců v populaci na celostátní úrovni, jak genotypovaných, tak negenotypovaných a vytvoření společného žebříčku podle hodnot všech chovných zvířat.

Včlenění celosvětových Interbulových údajů, přepočtených metodou MACE na podmínky ČR, do RRTDM hodnocení mléčné užitkovosti, zpřesňuje domácího hodnocení.

Tento postup je celosvětově nový a nebyl zatím v ČR, ani v jiných zemích použit.

IV. Popis uplatnění Certifikované metodiky

Základním uživatelem je v ČR společnost Plemdat, s.r.o., která je zodpovědná za pravidelná celostátní vyhodnocování kontroly užitkovosti a hodnocení chovných zvířat.

Navržené metodické postupy mohou využít i další uživatelé z chovatelských organizací a výzkumných pracovišť při vyhodnocování souborů užitkovostí a spojování různých zdrojů informací do společného hodnocení.

V. Ekonomické aspekty

Metodika slouží především pro celostátní hodnocení chovných zvířat, které je podle zákona č. 110/1997 Sb. a podle zákona č. 154/2000 Sb. ČR prováděno pověřenou organizací Českomoravská společnost chovatelů, která má za tím účelem zřízenou organizaci Plemdat, s.r.o. Tato organizace spravuje celostátní chovatelský informační systém a tím je součástí celostátní administrativy. Výsledky pravidelných zpracování jsou uveřejňovány pro potřeby všech chovatelů. Českomoravská společnost chovatelů byla zřízena za tímto účelem podle zákona jako nezisková organizace.

Částku peněžního přínosu nelze přesně určit vzhledem k působení mnoha proměnných, které užitkovost a ekonomické podmínky výroby a zpeněžování ovlivňují. Možné hospodářské zisky z využívání chovatelského informačního systému mohou vyplynout všem chovatelům a podnikům chovatelských služeb z vyššího chovatelského pokroku v důsledku vyšší přesnosti výběru chovných zvířat a tím i lepšího zpeněžování výrobků a snížení nákladů na chov zvířat.

VI. Seznam použité související literatury

- Calus M.P.L., Vandenplas J., ten Napel J., Veerkamp R.F. 2016. Validation of simultaneous deregression of cow and bull breeding values and derivation of appropriate weights. *J. Dairy Sci.*, 99, 6403-6419.
- Forni S., Aguilar I., Misztal I. 2011. Different genomic relationship matrices for single step

- analysis using phenotypic, pedigree and genomic information. *Genet. Sel. Evol.*, 43,1.
- Gao H., Christensen O.F., Madsen P., Nielsen U.S., Zhang Y., Lund M.S., Su G. 2012. Comparison on genomic predictions using three GBLUP methods and two single-step blending methods in the Nordic Holstein population. *Genet. Sel. Evol.*, 44, 8.
- Jairath L., Dekkers J.C.M., Schaeffer L.R., Liu Z., Burnside E.B., Kolstad B. 1998. Genetic evaluation for herd life in Canada. *J. Dairy Sci.*, 81, 550-562.
- Liu Z., Goddard M.E., Reinhardt F., Reents R. 2014. A Single-Step SNP Model Applied to Test-Day Data of Dairy Cows. International Conference XXVI. Genetic Days, Praha, Czech Rep., 3-4 September.
- Liu Z., Reinhardt F., Bünger A., Reents R. 2004. Derivation and calculation of approximate reliabilities and daughter yield deviations of a random regression test-day model for genetic evaluation of dairy cattle. *J. Dairy Sci.*, 87, 1896-1907.
- Lourenco D.A.L. 2015. Osobní sdělení.
- Meuwissen T.H.E., Hayes B.J., Goddard M.E. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, 157, 1819–1829.
- Misztal I., Legarra A., Aguilar I. 2009. Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. *J. Dairy Sci.*, 92, 4648-4655.
- Plemdat 2016. Popis postupu stanovení plemenné hodnoty mléčné užitkovosti. Staženo 18.11. 2016 z www.plemdat.cz.
- Schaeffer L.R., Jamrozik J., Kistemaker G.J., Van Doormaal B.J. 2000. Experience with a test-day model. *J. Dairy Sci.*, 83, 1135–1144.
- Su G., Madsen P., Nielsen U.S., Mäntysaari E.A., Aamand G.P., Christensen O.F., Lund M.S. 2012. Genomic prediction for Nordic Red Cattle using one-step and selection index blending. *J. Dairy Sci.*, 95, 909–917.
- Szyda J., Żarnecki A., Suchocki T., Kaminski S. 2011. Fitting and validating the genomic evaluation model to Polish Holstein–Friesian cattle. *J. Appl. Genet.*, 52, 363–366.
- Vandenplas J., Colinet F.G., Gengler N. 2014. Unified method to integrate and blend several, potentially related, sources of information for genetic evaluation. *Genet. Sel. Evol.*, 46, 59.
- Vandenplas J., Gengler N. 2015. Strategies for comparing and combining different genetic and genomic evaluation: A review. *Livest. Sci.*, 181, 121-130.
- VanRaden P.M., Tooker M.E., Wright J.R., Sun C., Hutchison J.L. 2014. Comparison of single-trait to multi-trait national evaluations for yield, health, and fertility. *J. Dairy Sci.*, 97, 7952-7962.
- Vitezica Z.G., Aguilar I., Misztal I., Legarra A. 2011. Bias in genomic predictions for populations under selection. *Genet. Res. (Camb)*, 93, 357-366.
- Weller J.I., Stoop W.M., Eding H., Schrooten C., Ezra E. 2015. Genomic evaluation of a relatively small dairy cattle population by combination with a larger population. *J. Dairy Sci.*, 98, 4945-4955.

VII. Seznam publikací, které předcházely metodice

- Bauer J., Příbyl J., Vostrý L. 2015. Contribution of domestic production records and Interbull EBV on approximate reliabilities of single-step genomic breeding values in dairy cattle. *Czech J. Anim. Sci.*, 60, 263-267.
- Kranjčevićová A., Brzáková M., Příbyl J., Svitáková A., Fialová Z. 2016. Vliv pohlaví telete na mléčnou užitkovost dojnice. *Náš chov*, LXXVI, 10: 32-33.
- Pešek P., Příbyl J., Vostrý L. 2015. Genetic variances of SNP loci for milk yield in dairy

- cattle. *J. Appl. Genet.*, 56, 339-347.
- Příbyl J., Bauer J., Krupa E., Krupová Z., Milerski M., Novotná A., Pešek P., Příbylová J., Schmidová J., Svitáková A., Veselá Z., Vostrá Vydrová H., Vostrý L., Zavadilová L., Žáková E. 2014a. Genetic evaluation by Linear Models using own algorithms and standard software. Certifikovaná metodika, VÚŽV Uhřetěves. ISBN: 978-80-7403-128-1. Str. 55.
- Příbyl J., Bauer J., Pešek P., Příbylová J., Vostrý L., Zavadilová L. 2014b. Domestic and Interbull information in single step genomic evaluation of Holstein milk production. *Czech J. Anim. Sci.*, 59, 409-415.
- 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. *J. Dairy Sci.*, 96, 1865-1873.
- Vostrý L., Bauer J., Pešek P., Příbyl J., Šplíchal J., Zavadilová L. 2015. Contribution of Interbull MACE values in national Random Regression Test-Day Model in single-step Genomic Evaluation. EAAP / ICAR / Interbull Meeting, 31 August – 4 September 2015, Warsaw, Poland.
- Vostrý L., Příbyl J., Krupa E. 2014. *Deregress*. Software, Praha Uhřetěves: Výzkumný ústav živočišné výroby v.v.i., Dostupný z: <http://www.vuzv.cz/index.php?p=deregress&site=GenetikaSlechteti>
- Zavadilová L., Jamrozik J., Schaeffer L.R. 2005a. Genetic parameters for test-day model with random regressions for production traits of Czech Holstein cattle. *Czech J. Anim. Sci.*, 50, 142-154.
- Zavadilová L., Němcová E., Příbyl J., Wolf J. 2005b. Definition of subgroups for fixed regression in the test-day animal model for milk production of Holstein cattle in the Czech Republic. *Czech J. Anim. Sci.*, 50, 7-13.

Přílohy

1. Příprava rodokmenů

```
/*.....prirodTD2.sas.. .....*/
/* priprava rodokmenu pro dalsi programy pro TDM, kdyz nesouhasi
roky rodice, rodic chybejici. Podle oznaceni zeme a vicenasobneho oznaceni
Interbull + seznamy domacich byku a krav
priprava pro navazujici vypocty 24.7.2015...*/
filename byci "/home/pribyl/tdmmace/kgn009.txt";/*ktd*/
filename kravy "/home/pribyl/tdmmace/kgn008.txt";/*ktd*/
filename inter2 "/home/pribyl/tdmmace/pedig_HOL.csv";
filename crosal "/home/pribyl/tdmmace/crossref_ALL.csv";
filename zeme "/home/pribyl/tdmmace/zeme.txt";

/* zapis */
filename rodupr "/home/pribyl/tdmmace/puvo";/*opravene rodokmeny*/
filename zdvjm "/home/pribyl/tdmmace/zdvjm.txt";/*zdvojen jmen*/
filename usbyc '/home/pribyl/tdmmace/usbyc.txt'; /*cisla byku vseh*/
filename uskra '/home/pribyl/tdmmace/uskra.txt';/*cisla krav vseh*/
filename poctchb "/home/pribyl/tdmmace/poctchb";/*chybni byci */
filename poctchk "/home/pribyl/tdmmace/poctchk";/*chybne kravy */
filename alpohl "/home/pribyl/tdmmace/alpohl";/*chyba pohlavi */
filename shodj '/home/pribyl/tdmmace/shodj'; /*shod oznac jedin */
filename shodro '/home/pribyl/tdmmace/shodro';/*shod postav rodic */
filename rodpo '/home/pribyl/tdmmace/rodpo'; /*jedin pred rodic */
filename hrubci '/home/pribyl/tdmmace/hrubci'; /*hruby ciselnik jedince */
/*.....*/
options nodate nonumber ls=89 ps=10000;
proc printto print= "/home/pribyl/tdmmace/vysupravro.lst" ;
proc printto log = "/home/pribyl/tdmmace/logupravro.lst" ;
/*dm output 'clear';
dm log 'clear'; */
/*.....zeme.....*/
data zeme ; title "3x zeme - cislo, jmen2, jmen3 30";
infile zeme firstobs = 2 missover ;
input czm $39-41 ze2 $42-43 ze3 $44-46 ;
if czm = "840" then do ; ze2 = "US" ; ze3 = "USA" ; end ;
if czm = "." then delete; xx = czm; if xx*1 = 0 then delete; if czm = "" then
delete;
proc means ;
proc sort ;by czm ;
data zee ; /* Interbulove zdvojeniny jmen byku */
attrib cdd cii format=$15. length=$15;
attrib cdo ciz format=$16. length=$16;
infile crosal ;
input plin $1-3 zd $4-6 pohd $7 cd $8-19 zc $24-26 pohc $27 cc $28-39 ;
if zd = "840" then zd = "USA"; if zc = "840" then zc = "USA";
cdo = zd||pohd||cd ; ciz = zc||pohc||cc ;
cdd = zd||cd ; cii = zc||cc ;
file alpohl ; /* zapis chybného pohlavi*/
if cdd = cii then do ;
if pohd ne pohc then do ;
put plin $1-3 zd $5-7 pohd $8 cd $9-20 zc $22-24 pohc $25 cc $26-37 ;
end ;
end ;
if cdo = "" then delete ; if ciz = "" then delete ;
if cdo = ciz then delete ;
keep plin cdo ciz ;
/*.....oprava cisla na jmeno zeme.aliasy.....*/
data ze ;
set zee ;
```

```

        czm = substr(ciz,1,3);
    proc sort ; by czm ;                               /* cizi */
data alze ;
    merge ze(in=zem) zeme ; by czm ; if zem ;
data alze ;
    set alze ;
    if ze3 ne "" then ciz = ze3||substr(ciz,4,13);
    drop ze2 ze3 ;
    czm = substr(cdo,1,3);
    proc sort ; by czm ;                               /* domaci */
data ze ;
    merge alze(in=zem) zeme ; by czm ; if zem ;
data ze ;
    set ze ;
    if ze3 ne "" then cdo = ze3||substr(cdo,4,13);
    if cdo = ciz then delete ;
    keep plin cdo ciz;                               /* oboji 16 mistne */
    proc sort ; by ciz ;
data ze ;                                           title "oprava zeme.aliasy 75";
    set ze ; by ciz ; if first.ciz ;
    file zdvjrn ;                                   /* zápis zdvojená jmena na disk */
    put plin $1-3 cdo $5-20 ciz $22-37 ;
    proc means ;
/* .....rodokmeny.....*/
data byk;                                           /* byci 81*/
attrib mat format=$14. length=$14;
attrib pl format=$1. length=$1;
infile byci missover;
input cislob $ 1-6 otec $ 8-13 zm $15-16 cm $17-28
    podilc 30-32 podilh 34-36 podilr 38-40
    Zpuv $42-44 rokna 46-49 /*prizh2 prizkn $*/ pb $59/*-70*/
    zem $72-74 cj $75-86 ;
if zpuv = "840" then zpuv = "USA";          if zem = "840" then zem = "USA";
xx = cj;          if xx*1 = 0 then delete ;
                  if cj = "" then delete ;
                  if zem = "" then delete ;
xx = zm;  if xx*1 = 0 then zm = "" ;
    mat = zm||cm ;
xx = cm;  if xx*1 = 0 then mat = "" ;
          if cm = "" then mat = "" ;
          if zm = "" then mat = "" ;
xx = substr(cm,1,8); if xx*1 = 0 then do ;
xx = substr(cm,9,4); if xx*1 < 1000 and zm = "CZ" then mat = "" ; end ;
if cm="000000970000" or cm="000000981000" or cm="000000983000"
or cm="000000997000" or cm="000000970970" or cm="000000981981"
or cm="000000983983" or cm="000000997997" or cm="000970000000"
or cm="000981000000" or cm="000983000000" or cm="000997000000"
or cm="000970000970" or cm="000981000981" or cm="000983000983"
or cm="000997000997" or cm="000970970000" or cm="000981981000"
or cm="000983983000" or cm="000997997000" or cm="000970970970"
or cm="000981981981" or cm="000983983983" or cm="000997997997"
    then mat= "";
xx = otec ;          if xx*1 = 0 then otec = "" ;
                    if substr(pb,1,1) = "7" then pl = "H" ;
                    if substr(pb,1,1) = "1" then pl = "C" ;
                    czm = zem ;                               /* cislo zem jedinec */
                    if podilc < 0 then podilc = . ;
                    if podilh < 0 then podilh = . ;
                    if podilr < 0 then podilr = . ;
    podcel = podilc + podilh + podilr ;
                    if podcel > 100 then do ;
                        podilc = podilc*100/podcel ;

```

```

                podilh = podilh*100/podcel ;
                podilr = podilr*100/podcel ;
            end ;
        proc sort ; by czm ;
data byk ;                                /*oprava na zem byci 121*/
    merge byk(in=byci) zeme ; by czm ; if byci ;
data byk ;
    set byk ;
    attrib puvc ciz format=$16. length=$16;
    attrib pohb format=$1. length=$1;
    pohb = "M" ;
    if ze3 ne "" then puvc = ze3||pohb||cj;
    else puvc = zem||pohb||cj;          /*nove cislo jedince*/
    ciz = puvc ;
keep cislob otec mat podilC podilH podilR rok nar puvc ciz pl;
    proc sort ; by ciz ;
    proc means ;                          title "byci seznam 133";
    proc freq ; tables pl ;
/*.....*/
    proc sort data=ze ; by ciz ;
data byk3 ;                                /* provereni zdvojenin oznaceni byka 137*/
    merge byk(in = by) ze ; by ciz ; if by ;
data byk3 ;
    set byk3 ;
    if cdo ne "" then puvc = cdo ;
    ciz = puvc ;
    a = 1 ;
keep cislob otec mat podilC podilH podilR rok nar puvc ciz pl a;
    proc sort ; by puvc ;
proc means noprint;var a;by puvc; title "pocty zdvojenych domac byku 146";
    output out=pru mean=;
    data sbc; set pru; pocet=_freq_; keep puvc pocet ;
data pocet ;
    merge byk3 sbc ; by puvc ;
data pocet ;                                /* puvodne jako odlisni, ale jsou stejni*/
    set pocet ;
    if pocet < 2 then delete ;
    file poctchb ;
    put cislob 1-6 puvc $8-23 pocet 25-30;
    keep cislob puvc pocet ;

data pocet ;
    set pocet ; by puvc; if first.puvc ;
    proc means ;
    proc sort data=byk3 ; by cislob ;      /* ciselnik usni cislo byka*/
data uscib ;
    set byk3 ; by cislob ; if first.cislob ;
    file usbyc dlm=";" ;
    put cislob 1-6 puvc $8-23 ;
    keep cislob puvc ;                    /* 6 mistné, 16 místné */
data otcib ;                                title "oprava zdvojeni otci byku 166";
    set byk3 ;
    drop puvc ;
    cislob = otec ;
    proc sort ; by cislob ;
data byk ;
    merge otcib(in = byc) uscib ; by cislob ; if byc ;
    cisj = ciz ;
    o = puvc ;                            /*nove cislo otce*/
    keep cisj o mat podilC podilH podilR rok nar pl ;
    proc means ; title "byci seznam opraveny jedinec + otec 176";
/*.....kravy.....*/
data krava;                                /* kravy 178*/

```

```

        attrib pl pohk format=$1. length=$1;
        attrib cislok mat format=$14. length=$14;
infile kravy missover;
input zj $1-2 cjd $3-14 otec $16-21 zm $23-24 cm $25-36
    podilC 38-40 podilH 42-44 podilR 46-48 Zpuv $50-52 rok nar 54-57 pk $63-74
        zem $76-78 cj $79-90 ;
if zpuv = "840" then zpuv = "USA"; if zem = "840" then zem = "USA";
        pohk = "F" ;
        cislok = zj||cjd ;
/* jedinec*/
xx = cj ; if xx*1 = 0 then delete ; if cj = "" then delete ;
if cj="000000970000" or cj="000000981000" or cj="000000983000"
or cj="000000997000" or cj="000000970970" or cj="000000981981"
or cj="000000983983" or cj="000000997997" or cj="000970000000"
or cj="000981000000" or cj="000983000000" or cj="000997000000"
or cj="000970000970" or cj="000981000981" or cj="000983000983"
or cj="000997000997" or cj="000970970000" or cj="000981981000"
or cj="000983983000" or cj="000997997000" or cj="000970970970"
or cj="000981981981" or cj="000983983983" or cj="000997997997"
then delete;
xx = zem ; if xx*1 = 0 then delete; if zem = "" then delete;
xx = substr(cj,1,8); if xx*1 = 0 then do ;
xx = substr(cj,9,4); if xx*1 < 1000 and zem = "CZE" then delete; end ;
        if substr(pk,1,1) = "7" then pl = "H" ;
        if substr(pk,1,1) = "1" then pl = "C" ;
xx = otec ; if xx*1 = 0 then otec = "" ;
        mat = zm||cm ; /* matka */
xx = cm ; if xx*1 = 0 then mat = "" ;
if cm="000000970000" or cm="000000981000" or cm="000000983000"
or cm="000000997000" or cm="000000970970" or cm="000000981981"
or cm="000000983983" or cm="000000997997" or cm="000970000000"
or cm="000981000000" or cm="000983000000" or cm="000997000000"
or cm="000970000970" or cm="000981000981" or cm="000983000983"
or cm="000997000997" or cm="000970970000" or cm="000981981000"
or cm="000983983000" or cm="000997997000" or cm="000970970970"
or cm="000981981981" or cm="000983983983" or cm="000997997997"
then mat = "" ;
xx = zm ; if xx*1 = 0 then zm = "" ; if zm = "" then mat = "" ;
xx = substr(cm,1,8); if xx*1 = 0 then do ;
xx = substr(cm,9,4); if xx*1 < 1000 and zm = "CZ" then mat = "" ; end ;
        czm = zem ; /* zeme jedince */
        if podilc < 0 then podilc = . ;
        if podilh < 0 then podilh = . ;
        if podilr < 0 then podilr = . ;
        podcel = podilc + podilh + podilr ;
        if podcel > 100 then do ;
                podilc = podilc*100/podcel ;
                podilh = podilh*100/podcel ;
                podilr = podilr*100/podcel ;
        end ;
proc sort ; by czm ;
data krava ; /*oprava na zem kravy 225*/
merge krava(in=byci) zeme ; by czm ; if byci ;
data krava ;
set krava ;
attrib puvc format=$16. length=$16;
if ze3 ne "" then puvc = ze3||pohk||cj; /*krava*/
else puvc = zem||pohk||cj;
        ciz = puvc ;
keep cislok otec mat podilC podilH podilR rok nar puvc ciz pl ;
proc sort ; by ciz ;
data krava ; /* provereni zdvojenin oznaceni kravy 235*/

```

```

merge krava(in = byk) ze ; by ciz ; if byk ;
data krava ;
set krava ;
if cdo ne "" then puvc = cdo ;
ciz = puvc ;
a = 1 ;
keep cislok otec mat podilC podilH podilR rok nar puvc ciz pl a ;
proc sort ; by puvc ; title "pocety zdvojenych domácich krav 243";
proc means noprint; var a ; by puvc ;
output out=pru mean=;
data sbc; set pru; pocet=_freq_; keep puvc pocet ;
data pocet ;
merge krava sbc ; by puvc ;
data pocet ;
set pocet ;
if pocet < 2 then delete ;
file pochtck ;
put cislok $1-14 puvc $16-31 pocet 33-38 ;
keep cislok puvc pocet ;
data pocet ;
set pocet ; by puvc; if first.puvc ;
proc means ;
proc sort data=krava ; by cislok ;
data uscisk ; /* ciselnik usni cislo krava*/
set krava ; by cislok ; if first.cislok ; title "kravy seznam 260";
file uskra dlm=";" ;
put cislok $1-14 puvc $16-31 ;
keep cislok puvc ; /* 14 místné, 16 místné */
proc means ;
data krava ; /* cislo kravy v seznamu jen jednou*/
set krava ; by cislok ; if first.cislok ;
/*.....*/
data otcik ; title "oprava otci krav 268";
attrib cislob format=$6. length=$6;
set krava ; drop puvc ;
cislob = otec ;
proc sort ; by cislob ;
data krava ;
merge otcik(in = by) uscisk ; by cislob ; if by ;
cisj = ciz ;
o = puvc ;
attrib cislok format=$14. length=$14;
cislok = mat ;
keep cisj o cislok mat podilC podilH podilR rok nar pl ;
proc means ; title "kravy opraveny otec 280";
/*.....oprava matek krav.....*/
proc sort data = krava ; by cislok ; title "oprava matky krav 282";
data krava ;
attrib matka format=$16. length=$16;
merge krava(in=byci) uscisk ; by cislok ; if byci ;
matka = puvc ; /* oprava kravy, otci, matky 286*/
keep cisj o matka podilC podilH podilR rok nar pl ;
proc means ;
proc freq ; tables pl ;
/*.....oprava matek byku.....*/
data byk ;
attrib cislok format=$14. length=$14;
set byk ;
cislok =mat ;
proc sort ; by cislok ; title "oprava byci, otci, matky byku 295";
data byk;
attrib matka format=$16. length=$16;

```

```

merge byk(in=byci) uscisk ; by cislok ; if byci ;
      matka = puvc ;
keep cisj o matka podilC podilH podilR rokna pl ;
proc means ;
/*.....interbul rodokmen .....*/
data int2g ; title "interbul 2 generace s predky 303";
  attrib a b c format=$19. length=$19;
  attrib datu format=$8. length=$8;
  attrib czm zmj zmo zmmat pb po plmat format=$3. length=$3;
  attrib pohj poho pohm format=$1. length=$1;
  attrib cj co cmat format=$12. length=$12;
  attrib rokna format=$4. length=$4;
infile inter2 delimiter=';' missover ;
  input a $ b $ c $ datu $ jm $ priz $ ;
xx = a; if xx*1 = 0 then delete ; if a = "" then delete ;
pb = substr(a,1,3) ; zmj = substr(a,4,3) ; pohj= substr(a,7,1) ; cj =
substr(a,8,12) ;
po = substr(b,1,3) ; zmo = substr(b,4,3) ; poho= substr(b,7,1) ; co =
substr(b,8,12) ;
plmat= substr(c,1,3) ; zmmat= substr(c,4,3) ; pohm= substr(c,7,1) ; cmat=
substr(c,8,12) ;
if zmj = "840" then zmj = "USA"; if zmo = "840" then zmo = "USA";if zmmat =
"840" then zmmat = "USA";
xx = b; if xx*1 = 0 then do ; po = "" ; zmo = "" ; co = "" ; end ;
xx = c; if xx*1 = 0 then do ; plmat = "" ; zmmat = "" ; cmat = "" ; end ;
rokna = substr(datu,1,4)*1 ;
      czm = zmj ; /* zeme jedince 320*/
keep pb czm zmj cj po zmo co plmat zmmat cmat pohj poho pohm rokna ;
proc means ;
proc sort ; by czm ;
proc sort data=zeme ; by czm ;
data intjed ; /* oprava zeme jedince 325*/
merge int2g(in=byc) zeme ; by czm ; if byc ;
data intjeda ;
set intjed ;
attrib cisj format=$16. length=$16;
if ze3 ne "" then cisj = ze3||pohj||cj;
else cisj = zmj||pohj||cj;
      czm = zmo ; /* zeme otce 332*/
keep cisj po czm zmo co plmat zmmat cmat pohj poho pohm rokna pb ;
proc sort ; by czm ;
data intot ; /* oprava zeme otec 335*/
merge intjeda(in=bac) zeme ; by czm ; if bac ;
data intot ;
set intot ;
attrib o format=$16. length=$16;
if ze3 ne "" then o = ze3||poho||co;
else o = zmo||poho||co;
      czm = zmmat ; /* zeme matky 342*/
keep cisj o czm plmat zmmat cmat pohj poho pohm rokna pb ;
proc sort ; by czm ;
data intma ; /* oprava zeme matka 345*/
merge intot(in=byc) zeme ; by czm ; if byc ;
data intma ;
set intma ;
attrib matka format=$16. length=$16;
attrib pl format=$1. length=$1;
if ze3 ne "" then matka = ze3||pohm||cmat;
else matka = zmmat||pohm||cmat;
      podilC= 0 ;
      podilH= 0 ;
      podilR= 0 ;

```

```

        if pb="HOL"                then do ; podilH=100; pl = "H" ; end ;
    else if pb="SIM" or pb="MON"  then do ; podilC=100; pl = "C" ; end ;
    else if pb="RED"              then do ; podilR=100; pl = "R" ; end ;

        ciz = cisj ;
        keep cisj o matka podilC podilH podilR roknl pl ciz pb ;
        proc means ;                title "interbul opravy zeme 361";
        proc freq ; tables pl ;
/*.....oprava zdvojeniny Interb.....*/
        proc sort ; by ciz ;        /* oprava jedinec 364*/
data intj ;
    merge intma(in=jedi) ze ; by ciz ; if jedi ;
data intj ;
    set intj ;
    if cdo ne "" then cisj = cdo ;
    ciz = o ;
    if plin ne "" then pb = plin ;
    if pb="HOL"                then do ; podilH=100; pl = "H" ; end ;
    else if pb="SIM" or pb="MON" then do ; podilC=100; pl = "C" ; end ;
    else if pb="RED"              then do ; podilR=100; pl = "R" ; end ;
    podcel = podilc + podilh + podilr ;
        if podcel > 100 then do ;
            podilc = podilc*100/podcel ;
            podilh = podilh*100/podcel ;
            podilr = podilr*100/podcel ;
        end ;
        keep cisj o matka podilC podilH podilR roknl pl ciz;
        proc sort ; by ciz ;        /* oprava otec 382*/
data into ;
    merge intj(in=jedi) ze ; by ciz ; if jedi ;
data into ;
    set into ;
    if cdo ne "" then o = cdo ;
    ciz = matka ;
keep cisj o matka podilC podilH podilR roknl pl ciz;
        proc sort ; by ciz ;        /* oprava matka 390*/
data intm ;
    merge into(in=jedi) ze ; by ciz ; if jedi ;
data intm ;
    set intm ;
    if cdo ne "" then matka = cdo ;
    oo = o ; mm = matka ; podcc = podilc ; podhh = podilh ;
    podrr = podilr ; pll = pl ;
        keep cisj oo mm podcc podhh podrr roknl pll ;
        proc sort ; by cisj ;
data intm ; title "interbul opravy zdvojeniny jedinec, otec, matka 400";
    set intm ; by cisj ; if first.cisj ;
        proc means ;
/*.....cely rodokmen.....*/
        proc sort data=byk; by cisj ; proc sort data=krava; by cisj ;
data rod ; merge byk krava ; by cisj ;
    data rod ; set rod ;
xx = cisj; if xx*1 = 0 then delete ; if cisj = "" then delete ;
xx = roknl; if xx*1 = 0 then roknl= . ;
        proc sort ; by cisj ;
data rod ;
    merge rod intm ; by cisj ;
data rod ;
    set rod ;
xx = cisj; if xx*1 = 0 then delete ; if cisj = "" then delete ;
if o = "" then o = oo ; if matka = "" then matka = mm ;
if podilc = . then podilc = podcc ; if podilh = . then podilh = podhh ;

```

```

if podilr = . then podilr = podrr ;
    podcel = podilc + podilh + podilr ;
    if podcel > 100 then do ;
        podilc = podilc*100/podcel ;
        podilh = podilh*100/podcel ;
        podilr = podilr*100/podcel ;
    end ;
    if rok nar = . then rok nar = rok n ;
    if pl = "" then pl = pll ;
xx = o ; if xx*1 = 0 then o = "" ;
xx = matka ; if xx*1 = 0 then matka = "" ;
    keep cisj o matka podilC podilH podilR rok nar pl ;
data rodupr ; title "rodokmen upraveny 429";
    set rod ; by cisj ; if first.cisj ;
        zj = substr(cisj,1,3);
/*.....uzavreni rodokmenu .....*/
data a ; title "celkovy pocet jedincu v rodokmenu 433";
    set rodupr ;
    keep cisj ; if cisj = "" then delete ;
data b ;
    set rodupr ;
    keep cisj ; cisj = o ; if cisj = "" then delete ;
data c ;
    set rodupr ;
    keep cisj ; cisj = matka ; if cisj = "" then delete ;
data h ;
    set a b c ;
    proc sort ; by cisj ;
data h ;
    set h ; by cisj ; if first.cisj ;
    proc means ;
data rodupr ;
    merge h(in=vse) rodupr ; by cisj ; if vse ;
    proc means ;
/*.....chyby v rodokmenu .....*/
data shodj ; title "chybne shodne oznaceni jedince a jeho rodice 452";
    set rodupr ;
    keep cisj o matka ;
    file shodj;
    if cisj=o or cisj= matka then put cisj $1-16 o $18-33 matka $35-50 ;
    if o ne "" and o = matka then put cisj $1-16 o $18-33 matka $35-50 ;
data tisshj;
    attrib jedin o ma format=$16. length=$16;
    infile shodj missover;
    input jedin $1-16 o $18-33 ma $35-50 ;
    proc means;
data shodo; title "chybne shodni jedinci v postaveni otce a matky 463";
    set shodj;
    if cisj= "" then delete;
    if o = "" then delete;
    rodi= o; jedo=cisj; keep jedo rodi o; /* potomek od otce 467*/
    proc sort ; by rodi;
data shodm;
    set shodj;
    if cisj = "" then delete;
    if matka= "" then delete;
    rodi=matka; jedm=cisj; keep jedm rodi matka; /* tele od matky 473*/
    proc sort ; by rodi;
data shodom;
    merge shodo shodm; by rodi;
data shodom ;
    set shodom ;

```

```

xx = rodi;    if xx*1 = 0 then delete ;
              file shodro;
              if o=matka then do;
                put jedo $1-16 jedm $18-33 rodi $35-50; end;
data tisshom;          /* shoda ruzneho postaveni rodice u jedicu*/
  format jedo jedm rodi $ 16. ;
  infile shodro missover;
  input jedo 1-16 jedm 18-33 rodi 35-50;
  proc means;    run ;
data rodupr ;
  set rodupr ;
if cisj = o then o = "" ; if cisj = matka then matka = "" ;
  if matka = o then matka = "" ;
xx = o;    if xx*1 = 0 then o = "" ;
xx = matka;  if xx*1 = 0 then matka = "" ;
/*.....doplneni roku, zeme a plemen v rodokmenu.....*/
data rro ;          title " zeme z uzitkovosti do rodokmenu 495";
set rodupr ;
  jedin = cisj ; ma = matka ; rokna = roknar ; zem = zj ;
  podc = podilc ; podh = podilh ; podr = podilr ;
              if jedin = "" then delete ;
  if rokna ne . and rokna < 1940 then rokna = 1940 ;
  keep jedin o ma podc podh podr rokna zem ;
  proc means ;
data rje ;          /* jedinci */
set rro;
  keep jedin o ma podcj podhj podrj zemj roknaj ;
  podcj =podc; podhj =podh; podrj =podr;
  zemj =zem; roknaj =rokna;
  /*if jedin = "" then delete ;*/
  proc sort ; by o ;
data rot ;          /* otci 509*/
set rro;
  o = jedin ;
  keep o podco podho podro zemo roknao ;
  podco = podc; podho = podh; podro = podr;
  zemo = zem; roknao = rokna;
  proc sort ; by o ;
data jo ;          /* jedinci + otci*/
merge rje(in=je) rot ; by o ; if je ;
  proc sort ; by ma ;
data rma ;          /* matky */
set rro ;
  ma = jedin ;
  keep ma podcm podhm podrm zemm roknam ;
  podcm = podc; podhm = podh; podrm = podr;
  zemm = zem; roknam = rokna;
  proc sort ; by ma ;
data jolp ;
merge jo(in=je) rma ; by ma ; if je ;
%macro doplrok;          /* doplneni roku */
%do i = 1 %to 8 %by 1;
proc sort data=jolp ; by jedin ;
data jolp ;          title " &i. plemena roky jedinci + otci + matky 531";
set jolp ;
  if roknaj = . then roknaj = roknao + 5 ;
  if roknam = . then roknam = roknao + 4 ;
if roknao = . then roknao = roknaj - 5 ;
if roknam = . then roknam = roknaj - 4 ;
/* .....prepis rodice na jedince.....*/
data rr3 ;          title "&i . prepis matky na jedince 538";
set jolp ;

```

```

/*nejdrive po matce potom prepsat spolehlivejsimi udaji od otce*/
jedin=ma;   rokna=roknam;
    keep jedin rokna ;
        if jedin = "" then delete; if rokna = . then delete ;
    proc sort ; by jedin ;
data rr4;
    set rr3; by jedin; if first.jedin;
data jo3 ;
    merge jolp(in=je) rr4; by jedin; if je ;
data jo3 ;
    set jo3 ;
        if jedin = "" then delete ;
        if rokna ne . then roknaj=rokna;
drop rokna ;
data rr ;           title "&i . prepis otce na jedince 553";
    set jo3 ;
        jedin=o; rokna=roknao;
    keep jedin rokna ;
        if jedin = "" then delete; if rokna = . then delete ;
    proc sort ; by jedin ;
data rr2;
    set rr; by jedin; if first.jedin;
data jo2 ;
    merge jo3(in=je) rr2; by jedin; if je ;
data jo2 ;
    set jo2 ;
        if jedin = "" then delete ;
        if rokna ne . then roknaj=rokna;
drop rokna ;
/*.....prepis jedince na rodice.....*/
proc sort ; by ma ;
data rr3 ;           title "&i . prepis jedince na matky 570";
set jo2 ;
    ma = jedin;   rokna=roknaj;
    keep ma rokna ;
        if ma = "" then delete; if rokna = . then delete ;
    proc sort ; by ma ;
data rr4;
    set rr3; by ma; if first.ma;
data jo3 ;
    merge jo2(in=je) rr4; by ma; if je ;
data jo3 ;
    set jo3 ;
        if jedin = "" then delete ;
        if rokna ne . then roknam=rokna;
drop rokna ;
proc sort ; by o ;
data rr ;           title "&i . dopln roku, prepis jedince na otce 586";
    set jo3 ;
        o = jedin; rokna=roknaj;
    keep o rokna ;
        if jedin = "" then delete; if rokna = . then delete ;
    proc sort ; by o ;
data rr2;
    set rr; by o; if first.o;
data jo2 ;
    merge jo3(in=je) rr2; by o; if je ;
data jolp ;
    set jo2 ;
        if jedin = "" then delete ;
        if rokna ne . then roknao=rokna;
drop rokna ;

```

```

%end;
data jolp ;
    set jolp ;
    /* doplneni, kdyz presto chybi */
if roknoj=. then roknoj=1980;if roknao= . then roknao=1975;if roknam=. then
roknam=1976;
rokkj=roknoj; rokko=roknao; rokkm=roknam; /*rok narozeni jedince zdojene */
proc means ;
%mend doplrok;
%doplrok ;run;
/*.....*/
data dopln ;
    title " plemena roky zeme jedinci + otci + matky 610";
    set jolp ;
if roknoj <= roknao then o = "" ;
if roknoj <= roknam then ma = "" ;
    xx = zemj; if xx*1 = 0 or zemj = " " then zemj = "" ;
    xx = zemo; if xx*1 = 0 or zemo = " " then zemo = "" ;
    xx = zemm; if xx*1 = 0 or zemm = " " then zemm = "" ;
        if zemj = "" then zemj = zemm ;
        if zemj = "" then zemj = zemo ;
if zemo = "" then zemo = zemj ;
if zemm = "" then zemm = zemj ;
    if podcj = . then podcj = (podco +podcm)/2 ;
        if podcm = . then podcm = 2*podcj -podco ;
            if podcm < 0 or podcm >100 then podcm = . ;
        if podco = . then podco = 2*podcj -podcm ;
            if podco < 0 or podco >100 then podco = . ;
if podhj = . then podhj = (podho +podhm)/2 ;
        if podhm = . then podhm = 2*podhj -podho ;
            if podhm < 0 or podhm >100 then podhm = . ;
        if podho = . then podho = 2*podhj -podhm ;
            if podho < 0 or podho >100 then podho = . ;
if podrj = . then podrj = (podro +podrm)/2 ;
        if podrm = . then podrm = 2*podrj -podro ;
            if podrm < 0 or podrm >100 then podrm = . ;
        if podro = . then podro = 2*podrj -podrm ;
            if podro < 0 or podro >100 then podro = . ;
    soucpod = podcj + podhj + podrj ;
        if soucpod > 100 then do ;
            podcj = 100*podcj/soucpod ;
            podhj = 100*podhj/soucpod ;
            podrj = 100*podrj/soucpod ;
        end ;
    soucpod = podcm+ podhm + podrm ;
        if soucpod > 100 then do ;
            podcm = 100*podcm/soucpod ;
            podhm = 100*podhm/soucpod ;
            podrm = 100*podrm/soucpod ;
        end ;
    soucpod = podco + podho + podro ;
        if soucpod > 100 then do ;
            podco = 100*podco/soucpod ;
            podho = 100*podho/soucpod ;
            podro = 100*podro/soucpod ;
        end ;
    drop soucpod ;
proc sort ; by jedin ;
proc means ;
/*.....dopocet roku a plemen v rodokmenu .....*/
%macro roky;
%do i = 1 %to 5 %by 1;
/* .....prepis rodice na jedince.....*/
data rr3 ;
    title "&i . prepis matky na jedince 661";

```

```

set doplň ; /*nejdrive po matce potom prepsat spolehlivejsimi udaji od otce*/
jedin=ma; rokna=roknam; rokk=rokkm; zem=zemm;
podc=podcm; podh=podhm; podr=podrm;
keep jedin rokna rokk zem podc podh podr ;
if jedin = "" then delete; if rokna = . then delete ;
proc sort ; by jedin ;
data rr4;
set rr3; by jedin; if first.jedin;
data jo3 ;
merge doplň(in=je) rr4; by jedin; if je ;
data jo3 ;
set jo3 ;
if jedin = "" then delete ;
if zemj = "" then zemj = zem;
if podc ne . then podcj = 0.5*(podcj +podc) ;
if podc = . then podcj = podc;
if podh ne . then podhj = 0.5*(podhj +podh);
if podh = . then podhj = podh;
if podr ne . then podrj = 0.5*(podrj +podr);
if podr = . then podrj = podr;
drop rokna rokk zem podc podh podr ;
data rr ; title "&i . prepis otce na jedince 685";
set jo3 ;
jedin=o; rokna=roknao; rokk=rokk0; zem=zemo;
podc=podco; podh=podho; podr=podro;
keep jedin rokna rokk zem podc podh podr ;
if jedin = "" then delete; if rokna = . then delete ;
proc sort ; by jedin ;
data rr2;
set rr; by jedin; if first.jedin;
data jo2 ;
merge jo3(in=je) rr2; by jedin; if je ;
data jo2 ;
set jo2 ;
if jedin = "" then delete ;
if zemj = "" then zemj = zem;
if podc ne . then podcj = 0.5*(podcj +podc) ;
if podc = . then podcj = podc;
if podh ne . then podhj = 0.5*(podhj +podh);
if podh = . then podhj = podh;
if podr ne . then podrj = 0.5*(podrj +podr);
if podr = . then podrj = podr;
drop rokna rokk zem podc podh podr ;
/* proc means; */
/*.....prepis jedinci na otce.....710.*/
proc sort data = jo2 ; by o ;
data doplň ;
set jo2 ;
keep o podc podh podr ;
o=jedin; podc=podcj; podh=podhj; podr=podrj ;
proc sort ; by o ;
data jo2 ;
merge doplň jo2 ; by o ;
if jedin = "" then delete ;
if podco = . then podco = podcj ;
if podho = . then podho = podhj ;
if podro = . then podro = podrj ;
drop podc podh podr ;
/*.....prepis jedinci na matky.....724.*/
proc sort data = jo2 ; by ma ;
data doplň ;
set jo2 ;

```

```

keep ma podc podh podr ;
    ma=jedin; podc=podcj; podh=podhj; podr=podrj ;
proc sort ; by ma ;
data jo2 ;
merge dopln jo2 ; by ma ;
    if jedin = "" then delete ;
    if podcm = . then podcm = podcj ;
    if podhm = . then podhm = podhj ;
    if podrm = . then podrm = podrj ;
    drop podc podh podr ;
/*.....*/
proc sort ; by jedin ;
data dopln ;          title "&i . doplneni plemen a roku zeme 740";
set jo2 ;
    if jedin = "" then delete;
        if zemo = "" then zemo = zemj ;
        if zemm = "" then zemm = zemj ;
pod = 0.5*(podcm + (2*podcj -podco));    /* podil matek prumerovani */
    if pod < 0 or pod >100 then pod = . ;
        if pod ne . then podcm = pod ;
pod = 0.5*(podco + (2*podcj -podcm));    /* podil otcu prumerovani*/
    if pod < 0 or pod >100 then pod = . ;
        if pod ne . then podco = pod ;
pod = 0.5*(podhm + (2*podhj -podho));    /* podil matek prumerovani */
    if pod < 0 or pod >100 then pod = . ;
        if pod ne . then podhm = pod ;
pod = 0.5*(podho + (2*podhj -podhm));    /* podil otcu prumerovani*/
    if pod < 0 or pod >100 then pod = . ;
        if pod ne . then podho = pod ;
pod = 0.5*(podrm + (2*podrj -podro));    /* podil matek prumerovani */
    if pod < 0 or pod >100 then pod = . ;
        if pod ne . then podrm = pod ;
pod = 0.5*(podro + (2*podrj -podrm));    /* podil otcu prumerovani*/
    if pod < 0 or pod >100 then pod = . ;
        if pod ne . then podro = pod ;
            if podcj < 0 or podcj > 100 then podcj = . ;
            if podhj < 0 or podhj > 100 then podhj = . ;
            if podrj < 0 or podrj > 100 then podrj = . ;
soucpod = podcj + podhj + podrj ;
    if soucpod > 100 then do ;
        podcj = 100*podcj/soucpod ;
        podhj = 100*podhj/soucpod ;
        podrj = 100*podrj/soucpod ;
    end ;
        if podcm < 0 or podcm > 100 then podcm = . ;
        if podhm < 0 or podhm > 100 then podhm = . ;
        if podrm < 0 or podrm > 100 then podrm = . ;
soucpod = podcm + podhm + podrm ;
    if soucpod > 100 then do ;
        podcm = 100*podcm/soucpod ;
        podhm = 100*podhm/soucpod ;
        podrm = 100*podrm/soucpod ;
    end ;
        if podco < 0 or podco > 100 then podco = . ;
        if podho < 0 or podho > 100 then podho = . ;
        if podro < 0 or podro > 100 then podro = . ;
soucpod = podco + podho + podro ;
    if soucpod > 100 then do ;
        podco = 100*podco/soucpod ;
        podho = 100*podho/soucpod ;
        podro = 100*podro/soucpod ;
    end ;

```

```

        drop soucpod ;
proc means;
        proc sort ; by jedin ;
%end;
%mend roky;
%roky ;run;
/*.....druha opakovani doplneni tam kde zustalo prazdne.....*/
data dopln ; set dopln;          title "druha doplneni kde prazdne 799";
        if podcj = . and podhj = . and podrj = . then podhj = 88 ;
        proc means ;
%roky;run;
/*.....chyby v rodokmenu .....*/
data rodpopot ;          title "chybne jedinci roky pred rodici 804";
        set dopln ;
        file rodpo ;
                if roknao >= roknaj then do ;
put jedin $1-16 roknaj 17-20 o $22-37 roknao 38-41 ma $43-58 roknam 59-62 ;
end ;
                if roknam >= roknaj then do ;
put jedin $1-16 roknaj 17-20 o $22-37 roknao 38-41 ma $43-58 roknam 59-62 ;
end ;
        data rodpo ;
                attrib jedin o ma format=$16. length=$16;
                infile rodpo missover; input
jedin $1-16 roknaj 17-20 o $22-37 roknao 38-41 ma $43-58 roknam 59-62 ;
        proc means ;run ;
/*.....preseknuti kruhoveho rodokmenu.....*/
data rr3 ;          title "preseknuti kruhoveho rod podily plemen 817 ";
        set dopln ;
                if roknaj <= roknao then o = "" ;
                if roknam <= roknam then ma = "" ;
        soucpod = podcj + podhj + podrj ;
                podoj = 100 - soucpod ;
        soucpod = podcm + podhm + podrm ;
                podom = 100 - soucpod ;
        soucpod = podco + podho + podro ;
                podoo = 100 - soucpod ;
        proc means ;
        proc sort ; by jedin ;
/*.....*/
data rodupr ;          title "Zapis rodokmenu 830";
        set rr3 ;
zj = substr(jedin,1,3); zo = substr(o,1,3); zm = substr(ma,1,3) ;
xx = o;          if xx*1 = 0 then o = "" ;
xx = ma;          if xx*1 = 0 then ma = "" ;
        podcel = podcj + podhj + podrj + podoj ;
                if podcel < 10 and zj = "USA" then do ;
                        podhj = 100 - podcj - podrj - podoj ;
                        podcel = 100 ;
                end ;
                if podcel < 10 and zj = "CZE" then do ;
                        podcj = 100 - podhj - podrj - podoj;
                        podcel = 100 ;
                end ;
        data cis1 ;          title "Ciselnik jedincu 844";
                set rodupr ;
                keep jedin ;
        data cis2 ;
                set rodupr ;
                keep jedin ; jedin = o ;
                if jedin = "" then delete ;
        data cis3 ;

```

```

        set rodupr ;
        keep jedin ; jedin = ma ;
        if jedin = "" then delete ;
data cis ;
    set cis1 cis2 cis3 ;
proc sort ; by jedin ;
data cis1 ;
    set cis ; by jedin ; if first.jedin ;
data cis2 ;
    set cis1 ;
    attrib nc format=$9. length=$9;
        nc = _n_ ;
    file hrubci dlm=";" ;
    put jedin $1-16 nc $18-26 ;
proc means ;
proc sort data=rodupr ; by jedin ;
data rod ;
    merge rodupr(in= prvni) cis2 ; by jedin ; if prvni ;
        ncj = nc ;
        drop nc ;
        jedin = o ;
proc sort data=rod ; by jedin ;
data rodo ;
    merge rod(in= druh) cis2 ; by jedin ; if druh ;
        nco = nc ;
        drop nc ;
        jedin = ma ;
proc sort data=rodo ; by jedin ;
data rodm ;
        title "Zapis rodokmen 880";
    merge rodo(in= tret) cis2 ; by jedin ; if tret ;
        ncm = nc ;
        drop nc jedin o ma ;
    file rodupr ;
put ncj $1-16 nco $18-33 ncm $35-50 podCj 52-54 podHj 56-58 podRj 60-62 podOj
64-66 podcel 68-70 roknaj 72-75 rokkj 77-80 zj $82-84 zo $86-88
zm $90-92 ;
proc means;run; /* rokkj= rok nar jedince "puvodni", roknaj= rok nar posun
generaci*/
proc datasets nolist;
delete a alze b byk byk3 c cis cis1 cis2 cis3 dopln h int2g intj intjed
intjeda intm intma into intot jo jolp jo2 jo3 krava otcib otcik pocet rje rma
rod rodm rodo rodpo rodpopot rodupr rot rr rr2 rr3 rr4 rro sbc shodj shodm
shodo shodom tisshj tisshom uscisb uscisk ze zee zeme;
quit; run;
/*.....konec.....*/

```

2. Načítání TD údajů

```

/*.....cteniTD.sas .....
    priprava uzitkovosti pro dalsi programy
    předchazi ctebank.sas, pripadne ctebyc.sas, rodarch.sas
    po druhem kroku navazují .... DMU, rozebr.sas
dosadit: omezeni uzitkovosti radek 55-65,81-104,174,370-372,470-471
        cetnost ve skupinach radek 204-230,254-342,389-412,435-442,489-
513,534-555
        pocty vrstevnic radek 180
        zaznamu na kravu radek 185
/*.....VUZV
10.5.2016 */
options nodate nonumber ls=89 ps=10000;

```

```

proc printto print= "/home/pribyl/tdmmace/vyspruz.lst" ;
proc printto log = "/home/pribyl/tdmmace/logpruz.lst" ;
/*dm output 'clear';
dm log 'clear'; */
filename uzku    "/home/pribyl/tdmmace/kmrp07.csv";/*KU*/
filename uskra   "/home/pribyl/tdmmace/uskra.txt"; /* cisla krav vsech*/
filename hrubci  '/home/pribyl/tdmmace/hrubci'; /*hruby ciselnik jedince */
filename prac   "/home/pribyl/tdmmace/prac"; /* opravena uzitkovot*/
filename kutd   "/home/pribyl/tdmmace/kutd";
/*.....cteni KU .....*/
data uz;          title "nactena uzitkovost ve dnech KU 21";
infile uzku dlm=";" /*obs=1000*/ recfm = v missover lrecl = 2600;
input  cisllok:$14. lakt pl:9. o:$14. m:$14. datnar:$8./*rrrrmdd*/ datotel:$8.
stajukonl:$9. dnylak mllak tuklak bilklak vek sp mezidob stupenku pocettd
/*obdobi KU*/
/*1*/ stajku1:$9. dnyl1  kgml1$  kgtu1$  kgbil1$  somb1  datku1:$8.
pocetdoj1
/*2*/ stajku2:$9. dnyl2  kgml2$  kgtu2$  kgbil2$  somb2  datku2:$8.
pocetdoj2
/*3*/ stajku3:$9. dnyl3  kgml3$  kgtu3$  kgbil3$  somb3  datku3:$8.
pocetdoj3
/*4*/ stajku4:$9. dnyl4  kgml4$  kgtu4$  kgbil4$  somb4  datku4:$8.
pocetdoj4
/*5*/ stajku5:$9. dnyl5  kgml5$  kgtu5$  kgbil5$  somb5  datku5:$8.
pocetdoj5
/*6*/ stajku6:$9. dnyl6  kgml6$  kgtu6$  kgbil6$  somb6  datku6:$8.
pocetdoj6
/*7*/ stajku7:$9. dnyl7  kgml7$  kgtu7$  kgbil7$  somb7  datku7:$8.
pocetdoj7
/*8*/ stajku8:$9. dnyl8  kgml8$  kgtu8$  kgbil8$  somb8  datku8:$8.
pocetdoj8
/*9*/ stajku9:$9. dnyl9  kgml9$  kgtu9$  kgbil9$  somb9  datku9:$8.
pocetdoj9
/*10*/stajku10:$9. dnyl10 kgml10$ kgtu10$ kgbil10$ somb10 datku10:$8.
pocetdoj10
/*11*/stajku11:$9. dnyl11 kgml11$ kgtu11$ kgbil11$ somb11 datku11:$8.
pocetdoj11 ;
    if dnylak = 0 then dnylak = . ;
    if mllak = 0 then mllak = . ;
    if tuklak = 0 then tuklak = . ;
    if bilklak = 0 then bilklak = . ;
    kgm1=translate( kgml1,".", ",")*1; kgt1=translate( kgtu1,".", ",")*1;
    kgb1=translate( kgbil1,".", ",")*1;
    kgm2=translate( kgml2,".", ",")*1; kgt2=translate( kgtu2,".", ",")*1;
    kgb2=translate( kgbil2,".", ",")*1;
    kgm3=translate( kgml3,".", ",")*1; kgt3=translate( kgtu3,".", ",")*1;
    kgb3=translate( kgbil3,".", ",")*1;
    kgm4=translate( kgml4,".", ",")*1; kgt4=translate( kgtu4,".", ",")*1;
    kgb4=translate( kgbil4,".", ",")*1;
    kgm5=translate( kgml5,".", ",")*1; kgt5=translate( kgtu5,".", ",")*1;
    kgb5=translate( kgbil5,".", ",")*1;
    kgm6=translate( kgml6,".", ",")*1; kgt6=translate( kgtu6,".", ",")*1;
    kgb6=translate( kgbil6,".", ",")*1;
    kgm7=translate( kgml7,".", ",")*1; kgt7=translate( kgtu7,".", ",")*1;
    kgb7=translate( kgbil7,".", ",")*1;
    kgm8=translate( kgml8,".", ",")*1; kgt8=translate( kgtu8,".", ",")*1;
    kgb8=translate( kgbil8,".", ",")*1;
    kgm9=translate( kgml9,".", ",")*1; kgt9=translate( kgtu9,".", ",")*1;
    kgb9=translate( kgbil9,".", ",")*1;
    kgm10=translate(kgml10,".", ",")*1;kgt10=translate(kgtu10,".", ",")*1;
    kgb10=translate(kgbil10,".", ",")*1;

```

```

kgm11=translate(kgm111,".",",")*1; kgt11=translate(kgtu11,".",",")*1;
kgb11=translate(kgbil11,".",",")*1;
drop kgm11 kgtu1 kgbil1 kgm12 kgtu2 kgbil2 kgm13 kgtu3 kgbil3 kgm14 kgtu4
kgbil4 kgm15 kgtu5 kgbil5 kgm16 kgtu6 kgbil6 kgm17 kgtu7 kgbil7 kgm18 kgtu8
kgbil8 kgm19 kgtu9 kgbil9 kgm110 kgtu10 kgbil10 kgm111 kgtu11 kgbil11 ;
  if kgm1=0 then kgm1=. ; if kgt1=0 then kgt1=. ; if kgb1=0 then kgb1=. ;
if somb1=0 then somb1=. ;
  if kgm2=0 then kgm2=. ; if kgt2=0 then kgt2=. ; if kgb2=0 then kgb2=. ;
if somb2=0 then somb2=. ;
  if kgm3=0 then kgm3=. ; if kgt3=0 then kgt3=. ; if kgb3=0 then kgb3=. ;
if somb3=0 then somb3=. ;
  if kgm4=0 then kgm4=. ; if kgt4=0 then kgt4=. ; if kgb4=0 then kgb4=. ;
if somb4=0 then somb4=. ;
  if kgm5=0 then kgm5=. ; if kgt5=0 then kgt5=. ; if kgb5=0 then kgb5=. ;
if somb5=0 then somb5=. ;
  if kgm6=0 then kgm6=. ; if kgt6=0 then kgt6=. ; if kgb6=0 then kgb6=. ;
if somb6=0 then somb6=. ;
  if kgm7=0 then kgm7=. ; if kgt7=0 then kgt7=. ; if kgb7=0 then kgb7=. ;
if somb7=0 then somb7=. ;
  if kgm8=0 then kgm8=. ; if kgt8=0 then kgt8=. ; if kgb8=0 then kgb8=. ;
if somb8=0 then somb8=. ;
  if kgm9=0 then kgm9=. ; if kgt9=0 then kgt9=. ; if kgb9=0 then kgb9=. ;
if somb9=0 then somb9=. ;
if kgm10=0 then kgm10=. ;if kgt10=0 then kgt10=. ;if kgb10=0 then kgb10=. ;if
somb10=0 then somb10=. ;
if kgm11=0 then kgm11=. ;if kgt11=0 then kgt11=. ;if kgb11=0 then kgb11=. ;if
somb11=0 then somb11=. ;
  kk = substr(cislok,3,12);
  if kk = "" then delete;
if kk="000000970000" or kk="000000981000" or kk="000000983000"
or kk="000000997000" or kk="000000970970" or kk="000000981981"
or kk="000000983983" or kk="000000997997" or kk="000970000000"
or kk="000981000000" or kk="000983000000" or kk="000997000000"
or kk="000970000970" or kk="000981000981" or kk="000983000983"
or kk="000997000997" or kk="000970970000" or kk="000981981000"
or kk="000983983000" or kk="000997997000" or kk="000970970970"
or kk="000981981981" or kk="000983983983" or kk="000997997997"
or kk="000234988931" or kk*1 < 1000 then delete;
  rokot = substr(datotel,1,4)*1; mesot = substr(datotel,5,2)*1;
proc sort ; by lakt ;
proc means ; by lakt;
data uz ; title "po vyhozeni divnych laktaci 79";
  set uz ;
if mllak >16000 then delete;if tuklak > 800 then delete; if bilklak > 630
then delete;
  if kgm1>70 or kgm1< 5 then kgm1=. ; if kgt1/kgm1 <0.02 or kgt1/kgm1 >0.07
then kgt1=. ;
  if kgm2>80 or kgm2<10 then kgm2=. ; if kgt2/kgm2 <0.02 or kgt2/kgm1 >0.07
then kgt2=. ;
  if kgm3>80 or kgm3<10 then kgm3=. ; if kgt3/kgm3 <0.02 or kgt3/kgm3 >0.07
then kgt3=. ;
  if kgm4>80 or kgm4<10 then kgm4=. ; if kgt4/kgm4 <0.02 or kgt4/kgm4 >0.07
then kgt4=. ;
  if kgm5>80 or kgm5< 8 then kgm5=. ; if kgt5/kgm5 <0.02 or kgt5/kgm5 >0.07
then kgt5=. ;
  if kgm6>70 or kgm6< 8 then kgm6=. ; if kgt6/kgm6 <0.02 or kgt6/kgm6 >0.07
then kgt6=. ;
  if kgm7>70 or kgm7< 8 then kgm7=. ; if kgt7/kgm7 <0.02 or kgt7/kgm7 >0.07
then kgt7=. ;
  if kgm8>60 or kgm8< 5 then kgm8=. ; if kgt8/kgm8 <0.02 or kgt8/kgm8 >0.07
then kgt8=. ;

```

```

if kgm9>50 or kgm9< 5 then kgm9=.; if kgt9/kgm9 <0.02 or kgt9/kgm9 >0.07
then kgt9=.;
if kgm10>40 or kgm10< 2 then kgm10=.; if kgt10/kgm10<0.02 or kgt10/kgm10>0.07
then kgt10=.;
if kgm11>30 or kgm11< 2 then kgm11=.; if kgt11/kgm11<0.02 or kgt11/kgm11>0.07
then kgt11=.;
if kgb1/kgm1 <0.02 or kgb1/kgm1 >0.07 then kgb1=.; if dnyl1 > 305 then
dnyl1 = .;
if kgb2/kgm2 <0.02 or kgb2/kgm2 >0.07 then kgb2=.; if dnyl2 > 305 then
dnyl2 = .;
if kgb3/kgm3 <0.02 or kgb3/kgm3 >0.07 then kgb3=.; if dnyl3 > 305 then
dnyl3 = .;
if kgb4/kgm4 <0.02 or kgb4/kgm4 >0.07 then kgb4=.; if dnyl4 > 305 then
dnyl4 = .;
if kgb5/kgm5 <0.02 or kgb5/kgm5 >0.07 then kgb5=.; if dnyl5 > 305 then
dnyl5 = .;
if kgb6/kgm6 <0.02 or kgb6/kgm6 >0.07 then kgb6=.; if dnyl6 > 305 then
dnyl6 = .;
if kgb7/kgm7 <0.02 or kgb7/kgm7 >0.07 then kgb7=.; if dnyl7 > 305 then
dnyl7 = .;
if kgb8/kgm8 <0.02 or kgb8/kgm8 >0.07 then kgb8=.; if dnyl8 > 305 then
dnyl8 = .;
if kgb9/kgm9 <0.02 or kgb9/kgm9 >0.07 then kgb9=.; if dnyl9 > 305 then
dnyl9 = .;
if kgb10/kgm10<0.02 or kgb10/kgm10>0.07 then kgb10=.; if dnyl10 > 305 then
dnyl10 = .;
if kgb11/kgm11<0.02 or kgb11/kgm11>0.07 then kgb11=.; if dnyl11 > 305 then
dnyl11 = .;
proc means ; by lakt;
proc sort ; by o ; title "dcer na otce 106";
proc means noprint; var mllak ; by o ;
output out=pru mean=;
data sbc; set pru; dcer=_freq_; keep o dcer;
proc means ;
proc sort data=uz ; by rokot ; title "laktaci v roce 111";
proc means noprint; var mllak ; by rokot ;
output out=pru mean=;
data sbc; set pru; laktaci=_freq_; keep rokot laktaci;
proc means ;
proc freq data=uz ; tables rokot ;
proc sort data=uz ; by mesot ; title "laktaci v mesici 117";
proc means noprint; var mllak ; by mesot ;
output out=pru mean=;
data sbc; set pru; laktaci=_freq_; keep mesot laktaci;
proc means ;
proc freq data=uz; tables mesot ;
proc sort data=uz; by cislok ;
data uskra ;
attrib puvc format=$16. length=$16;
attrib cislok format=$14. length=$14;
infile uskra dlm=";" ;
input cislok $1-14 puvc $16-31 ;
proc sort ; by cislok ;
data prvl2; title "16 mistna cisla krav 130";
merge uz(in=uzit) uskra ; by cislok ; if uzit ;
PROC MEANS ;
proc sort data=prvl2; by puvc ; title "ciselnik krav 133";
data cis ;
infile hrubci dlm=";" ;
input puvc $1-16 nc 18-26 ;
proc means ;
data uzkon ;

```

```

merge prvl2(in = krava) cis ; by puvc ; if krava ;
proc means;
data uzkra ;
set uzkon ;
if nc = . then delete ;
PROC MEANS ;
data uz ;
set uzkra ;
array staj stajku1 stajku2 stajku3 stajku4 stajku5 stajku6 stajku7 stajku8
stajku9 stajku10 stajku11;
array ml kgm1 kgm2 kgm3 kgm4 kgm5 kgm6 kgm7 kgm8
kgm9 kgm10 kgm11 ;
array tu kgt1 kgt2 kgt3 kgt4 kgt5 kgt6 kgt7 kgt8
kgt9 kgt10 kgt11 ;
array bi kgb1 kgb2 kgb3 kgb4 kgb5 kgb6 kgb7 kgb8
kgb9 kgb10 kgb11 ;
array sb somb1 somb2 somb3 somb4 somb5 somb6 somb7 somb8
somb9 somb10 somb11 ;
array dl dnyl1 dnyl2 dnyl3 dnyl4 dnyl5 dnyl6 dnyl7 dnyl8
dnyl9 dnyl10 dnyl11 ;
array dku datku1 datku2 datku3 datku4 datku5 datku6 datku7 datku8
datku9 datku10 datku11;
/*.....zapis.pracovni.....*/
file prac dlm=";" ;
do i = 1 to 11 ;
if ml[i] = . or dl[i] = . then do;
i =11 ;
goto konec;
end ;
put nc lakt pl rokot mesot vek sp mezidob staj[i] dku[i] dl[i] ml[i]
tu[i] bi[i] sb[i] ;
konec: end ;
data prover ; title "soubor KU 163";
infile prac dlm=";" ;
input nc lakt pl:9. rokot mesot vek sp mezidob staj:$9. dku:$8. dl ml tu
bi sb ;
attrib htd format=$17. length=$17.;
htd = compress(staj||dku) ;
if sp = 0 or sp > 190 then sp =190 ;
proc means ;
/*.....vrstevnice .....1L.....*/
data zaznam ; title "HTD prvni laktace 171";
set prover ;
if lakt ne 1 then delete ; if vek < 600 or vek > 1100 then delete ;
proc means ; run ;
%MACRO vrstev;
%do i=1 %to 20 %by 1; title " &i zaznamu na kravu a HTD 177";
proc sort data=zaznam; by htd;
proc means noprint; var ml ; by htd ; output out=pru mean=;
data sbc; set pru; n=_freq; if n < 8 then delete; keep htd ;
data zaznam; merge zaznam sbc(in=pocet) ; by htd ; if pocet;
/*... Vyhizeni krav s <3 ku v lakt.*/
proc sort ; by nc ;
proc means noprint; var ml ; by nc ; output out=pru mean=;
data sbc; set pru; zazn=_freq; if zazn < 4 then delete; keep nc ;
data zaznam; merge zaznam sbc(in=pocet) ; by nc ; if pocet;
proc means;
%end;
%mend;
%vrstev;run ;
data zaznam1 ; set zaznam ;proc sort data=zaznam1; by htd;
data cishtd ;

```

```

        set zaznaml ; by htd ; if first.htd ;
        keep htd ;
    data cishtd ;
        set cishtd ;
            nhtd = _n_ ;
    data zaznaml ;                title "pocty v htd uvnitr 1L 200";
        merge zaznaml cishtd ; by htd ;
        proc means noprint; var ml ; by nhtd ; output out=pru mean=;
        data sbc; set pru; poc=_freq_; keep nhtd poc;
    proc means;
    data zaznaml ; set zaznaml ;                title "skupiny 1L 204";
        if vek <631 then skve = 1 ;
    else if vek <661 then skve = 2 ;
    else if vek <691 then skve = 3 ;
    else if vek <721 then skve = 4 ;
    else if vek <751 then skve = 5 ;
    else if vek <781 then skve = 6 ;
    else if vek <811 then skve = 7 ;
    else if vek <841 then skve = 8 ;
    else if vek <871 then skve = 9 ;
    else if vek <901 then skve = 10 ;
    else if vek <931 then skve = 11 ;
    else if vek <961 then skve = 12 ;
    else if vek <991 then skve = 13 ;
    else if vek <1021 then skve= 14;
                                else skve = 15 ;
        if sp < 60 then sksp = 1 ;
    else if sp < 90 then sksp = 2 ;
    else if sp <120 then sksp = 3 ;
    else if sp <150 then sksp = 4 ;
                                else sksp = 5 ;
        rokk = rokot ;
                If mesot < 2 then do ; obd = 4 ; rokk = rokot -1 ; end ;
                else if mesot < 5 then obd = 1 ;
                else if mesot < 8 then obd = 2 ;
                else if mesot <11 then obd = 3 ;
                                else obd = 4 ;
        attrib skup format=$10. length=$10;
        skup = compress(skve||sksp||obd||rokk) ;
        if skup = "" then delete ;
        proc freq ; tables skve ;
        proc freq ; tables sksp ;
        proc freq ; tables obd ;
        proc freq ; tables rokk ;
    proc sort data=zaznaml ; by skup ;                /* ciselnik pevne skupiny */
    data cis ;
        set zaznaml ; by skup ; if first.skup ;
        keep skup ;
    data cis ;
        set cis ;
            ncskup = _n_ ;
    data zaznaml ;
        merge zaznaml cis ; by skup ;                title "pocty ve skupinach 1L 247";
        proc means noprint; var ml ; by ncskup ; output out=pru mean=;
        data sbc; set pru; poc=_freq_; keep ncskup poc;
        proc means;                title "slucovani skupi 1L 251";
    data zaznaml ; set zaznaml ;
        drop ncskup ;
    if skve = 2 then skve = 1;                if skve = 4 then skve = 3;
    if skve = 6 or skve = 7 then skve = 5;    if skve = 9 then skve = 8;
    if skve = 11 then skve = 10;            if skve = 13 then skve = 12 ;
    if skve = 15 then skve = 14;

```

```

        /*rokk = round(rokk/5 , 1)*5 ;*/
        if rokk < 1999 then rokk = 1995 ;
    else if rokk < 2003 then rokk = 1999 ;
    else if rokk < 2007 then rokk = 2003 ;
    else if rokk < 2011 then rokk = 2007 ;
        else rokk = 2011 ;
    skup = compress(skve||sksp||obd||rokk) ;
if skup = "1111995" then skup = "3111995" ;
if skup = "1111999" then skup = "3111999" ;
if skup = "1112003" then skup = "3112003" ;
if skup = "1112007" then skup = "3112007" ;
if skup = "1112011" then skup = "3112011" ;
if skup = "1121995" then skup = "3121995" ;
if skup = "1121999" then skup = "3121999" ;
if skup = "1122003" then skup = "3122003" ;
if skup = "1122007" then skup = "3122007" ;
if skup = "1122011" then skup = "3122011" ;
if skup = "1131995" then skup = "3131995" ;
if skup = "1131999" then skup = "3131999" ;
if skup = "1132003" then skup = "3132003" ;
if skup = "1132007" then skup = "3132007" ;
if skup = "1132011" then skup = "3132011" ;
if skup = "1141995" then skup = "3141995" ;
if skup = "1141999" then skup = "3141999" ;
if skup = "1142003" then skup = "3142003" ;
if skup = "1142007" then skup = "3142007" ;
if skup = "1142011" then skup = "3142011" ;
if skup = "12112011" then skup = "14112011" ;
if skup = "1211995" then skup = "3211995" ;
if skup = "1211999" then skup = "3211999" ;
if skup = "1212003" then skup = "3212003" ;
if skup = "12122011" then skup = "14122011" ;
if skup = "12132011" then skup = "14132011" ;
if skup = "12142011" then skup = "14142011" ;
if skup = "1221995" then skup = "3221995" ;
if skup = "1221999" then skup = "3221999" ;
if skup = "1222003" then skup = "3222003" ;
if skup = "1231995" then skup = "3231995" ;
if skup = "1231999" then skup = "3231999" ;
if skup = "1232003" then skup = "3232003" ;
if skup = "1241995" then skup = "3241995" ;
if skup = "1241999" then skup = "3241999" ;
if skup = "1242003" then skup = "3242003" ;
if skup = "1311995" then skup = "3311995" ;
if skup = "1311999" then skup = "3311999" ;
if skup = "1312003" then skup = "3312003" ;
if skup = "1321995" then skup = "3321995" ;
if skup = "1321999" then skup = "3321999" ;
if skup = "1322003" then skup = "3322003" ;
if skup = "1331995" then skup = "3331995" ;
if skup = "1331999" then skup = "3331999" ;
if skup = "1332003" then skup = "3332003" ;
if skup = "1341995" then skup = "3341995" ;
if skup = "1341999" then skup = "3341999" ;
if skup = "1342003" then skup = "3342003" ;
if skup = "14112007" then skup = "14122007" ;
if skup = "1411995" then skup = "3411995" ;
if skup = "1411999" then skup = "3411999" ;
if skup = "1412003" then skup = "3412003" ;
if skup = "1412007" then skup = "3412007" ;
if skup = "1412011" then skup = "3412011" ;
if skup = "14132007" then skup = "14122007" ;

```

```

if skup = "14132011" then skup = "14122011" ;
if skup = "14142007" then skup = "14242007" ;
if skup = "14142011" then skup = "14242011" ;
if skup = "1421995" then skup = "3421995" ;
if skup = "1421999" then skup = "3421999" ;
if skup = "1422003" then skup = "3422003" ;
if skup = "1431995" then skup = "3431995" ;
if skup = "1431999" then skup = "3431999" ;
if skup = "1432003" then skup = "3432003" ;
if skup = "1432011" then skup = "3432011" ;
if skup = "14412011" then skup = "14312011" ;
if skup = "1441995" then skup = "3441995" ;
if skup = "1441999" then skup = "3441999" ;
if skup = "1442003" then skup = "3442003" ;
if skup = "1442007" then skup = "3442007" ;
if skup = "1442011" then skup = "3442011" ;
if skup = "14422011" then skup = "14322011" ;
if skup = "14432011" then skup = "14332011" ;
if skup = "14442011" then skup = "14342011" ;
if skup = "1511995" then skup = "3511995" ;
if skup = "1521995" then skup = "3521995" ;
if skup = "1531995" then skup = "3531995" ;
if skup = "1541995" then skup = "3541995" ;
        if skup = "" then delete ;
        proc freq ; tables skve ;
        proc freq ; tables sksp ;
        proc freq ; tables obd ;
        proc freq ; tables rokk ;
proc sort data=zaznam1 ; by skup ;
        title "pocty ve skupinach po slučovani 1L 349";
        proc means noprint; var ml ; by skup ; output out=pru mean=;
        data sbc ; set pru; poc=_freq_ ; keep skup poc;
        proc means ; proc print ;
data skup ;
        set sbc ;
        if poc > 1000 then delete ;
        proc print ;
data cis ;                                /* ciselnik pevne skupiny */
        set zaznam1 ; by skup ; if first.skup ;
        keep skup ;
        data cis ;
        set cis ;
                ncskup = _n_ ;
data zaznam1 ;
        merge zaznam1 cis ; by skup ;
        drop skve sksp rokk obd ;
proc means ;
/*.....vrstevnice .....2L.....*/
data zaznam ;      set prover ;                                title "HTD 2L 368";
        if lakt ne 2 then delete ;
        if vek < 800 or vek > 1650 then delete ;
        if mezidob <210 then delete ;
proc means ; run ;
%vrstev; run;
data zaznam2 ; set zaznam ; proc sort data=zaznam2 ; by htd;
data cishtd ;
        set zaznam2 ; by htd ; if first.htd ;
        keep htd ;
        data cishtd ;
        set cishtd ;
                nhtd = _n_ ;
data zaznam2 ;

```

```

merge zaznam2 cishtd ; by htd ; title "pocety v htd uvnitr 2L 383";
proc means noprint; var ml ; by nhtd ; output out=pru mean=;
data sbc; set pru; poc=_freq_; keep nhtd poc;
proc means;
data zaznam2 ; set zaznam2 ; title "skupiny 2L 387";
if mezidob <301 then skme = 1 ;
else if mezidob <331 then skme = 2;
else if mezidob <361 then skme = 3 ;
else if mezidob <391 then skme = 4 ;
else if mezidob <421 then skme = 5 ;
else if mezidob <451 then skme = 6 ;
else if mezidob <481 then skme = 7 ;
else if mezidob <511 then skme = 8 ;
else if mezidob <541 then skme = 9 ;
else if mezidob <571 then skme = 10 ;
else if mezidob <601 then skme = 11 ;
else skme = 12 ;
if sp < 60 then sksp = 1 ;
else if sp < 90 then sksp = 2 ;
else if sp <120 then sksp = 3 ;
else if sp <150 then sksp = 4 ;
else sksp = 5 ;
rokk = rokot ;
if mesot < 2 then do ; obd = 4 ; rokk = rokot -1 ; end ;
else if mesot < 5 then obd = 1 ;
else if mesot < 8 then obd = 2 ;
else if mesot <11 then obd = 3 ;
else obd = 4 ;
attrib skup format=$10. length=$10;
skup = compress(skme||sksp||obd||rokk) ;
if skup = "" then delete ;
proc freq ; tables skme ;
proc freq ; tables sksp ;
proc freq ; tables obd ;
proc freq ; tables rokk ;
proc sort data=zaznam2 ; by skup ; /* ciselnik pevne skupiny */
data cis ;
set zaznam2 ; by skup ; if first.skup ;
keep skup ;
data cis ;
set cis ;
nckup = _n_ ;
data zaznam2 ;
merge zaznam2 cis ; by skup ; title "pocety ve skupinach 2L 427";
proc means noprint; var ml ; by nckup ; output out=pru mean=;
data sbc; set pru; poc=_freq_; keep nckup poc;
proc means;
data zaznam2 ; set zaznam2 ; title "slucovani skupi 2L 431";
drop nckup ;
if skme = 2 then skme = 1; if skme = 5 then skme = 4;
if skme = 7 then skme = 6; if skme > 7 then skme = 7;
/*rokk = round(rokk/5 , 1)*5 ;*/
if rokk < 1999 then rokk = 1995 ;
else if rokk < 2003 then rokk = 1999 ;
else if rokk < 2007 then rokk = 2003 ;
else if rokk < 2011 then rokk = 2007 ;
else rokk = 2011 ;
skup = compress(skme||sksp||obd||rokk) ;
if skup = "" then delete ;
proc freq ; tables skme ;
proc freq ; tables sksp ;
proc freq ; tables obd ;

```

```

proc freq ; tables rokk ;
proc sort data=zaznam2 ; by skup ;
title "pocety ve skupinach po slučovani 2L 448";
proc means noprint; var ml ; by skup ; output out=pru mean=;
data sbc; set pru; poc=_freq_; keep skup poc;
proc means; proc print ;
data skup ;
set sbc ;
if poc > 1000 then delete ;
proc print ;
data cis ; /* ciselnik pevne skupiny */
set zaznam2 ; by skup ; if first.skup ;
keep skup ;
data cis ;
set cis ;
ncskup = _n_ ;
data zaznam2 ;
merge zaznam2 cis ; by skup ;
drop skme sksp rokk obd ;
proc means ;
/*.....vrstevnice .....3L.....*/
data zaznam ; set prover ; title "HTD 3L 467";
if lakt ne 3 then delete ;
if vek < 1100 or vek > 2200 then delete ;
if mezidob <210 then delete ;
proc means ; run ;
%vrstev; run ;
data zaznam3 ; set zaznam ;proc sort data=zaznam3; by htd;
data cishtd ;
set zaznam3 ; by htd ; if first.htd ;
keep htd ;
data cishtd ;
set cishtd ;
nhtd = _n_ ;
data zaznam3 ;
merge zaznam3 cishtd ; by htd ; title "pocety v htd uvnitr 3L 482";
proc means noprint; var ml ; by nhtd ; output out=pru mean=;
data sbc; set pru; poc=_freq_; keep nhtd poc;
proc means;
data zaznam3 ; set zaznam3 ; title "skupiny 3L 486";
if mezidob <301 then skme = 1 ;
else if mezidob <331 then skme = 2;
else if mezidob <361 then skme = 3 ;
else if mezidob <391 then skme = 4 ;
else if mezidob <421 then skme = 5 ;
else if mezidob <451 then skme = 6 ;
else if mezidob <481 then skme = 7 ;
else if mezidob <511 then skme = 8 ;
else if mezidob <541 then skme = 9 ;
else if mezidob <571 then skme = 10 ;
else if mezidob <601 then skme = 11 ;
else skme = 12 ;
if sp < 60 then sksp = 1 ;
else if sp < 90 then sksp = 2 ;
else if sp <120 then sksp = 3 ;
else if sp <150 then sksp = 4 ;
else sksp = 5 ;
rokk = rokot ;
If mesot < 2 then do ; obd = 4 ; rokk = rokot -1 ; end ;
else if mesot < 5 then obd = 1 ;
else if mesot < 8 then obd = 2 ;
else if mesot <11 then obd = 3 ;

```

```

else obd = 4 ;
attrib skup format=$10. length=$10;
    skup = compress(skme||sksp||obd||rokk) ;
    if skup = "" then delete ;
proc freq ; tables skme ;
proc freq ; tables sksp ;
proc freq ; tables obd ;
proc freq ; tables rokk ;
proc sort data=zaznam3 ; by skup ; /* ciselnik pevne skupiny */
data cis ;
    set zaznam3 ; by skup ; if first.skup ;
    keep skup ;
data cis ;
    set cis ;
        ncskup = _n_ ;
data zaznam3 ;
    merge zaznam3 cis ; by skup ;
        title "pocety ve skupinach 3L 526";
proc means noprint; var ml ; by ncskup ; output out=pru mean=;
    data sbc; set pru; poc=_freq_; keep ncskup poc;
proc means;
data zaznam3 ; set zaznam3 ; title "slucovani skupi 3L 530";
    drop ncskup ;
    if skme = 2 then skme = 1; if skme = 5 then skme = 4;
    if skme > 6 then skme = 6;
        /*rokk = round(rokk/5 , 1)*5 ;*/
        if rokk < 1999 then rokk = 1995 ;
    else if rokk < 2003 then rokk = 1999 ;
    else if rokk < 2007 then rokk = 2003 ;
    else if rokk < 2011 then rokk = 2007 ;
        else rokk = 2011 ;
    skup = compress(skme||sksp||obd||rokk) ;
    if skup = "1111995" then skup = "1121995" ; if skup = "1111999" then
skup = "1121999" ;
    if skup = "1112003" then skup = "1212003" ; if skup = "1131995" then
skup = "1121995" ;
    if skup = "1132003" then skup = "1142003" ; if skup = "1311995" then
skup = "1211995" ;
    if skup = "1411999" then skup = "1311999" ; if skup = "1412007" then
skup = "1312007" ;
    if skup = "1421995" then skup = "1431995" ; if skup = "1421995" then
skup = "1431995" ;
    if skup = "1432007" then skup = "1422007" ; if skup = "1441999" then
skup = "1341999" ;
    if skup = "1112007" then skup = "3112007" ; if skup = "1112011" then
kup = "3112011" ;
    if skup = "1122003" then skup = "3122003" ; if skup = "1122007" then
skup = "3122007" ;
    if skup = "1122011" then skup = "3122011" ; if skup = "1411995" then
skup = "3411995" ;
    if skup = "1412003" then skup = "3412003" ; if skup = "1412011" then
skup = "3412011" ;
    if skup = "7111995" then skup = "6111995" ; if skup = "7112007" then
skup = "6112007" ;
    if skup = "7112011" then skup = "6112011" ; if skup = "7122011" then
skup = "6122011" ;
        if skup = "" then delete ;
proc freq ; tables skme ;
proc freq ; tables sksp ;
proc freq ; tables obd ;
proc freq ; tables rokk ;
proc sort data=zaznam3 ; by skup ;

```

```

                title "pocty ve skupinach po slučovani 3L 559";
proc means noprint; var ml ; by skup ; output out=pru mean=;
  data sbc; set pru; poc=_freq_; keep skup poc;
proc means; proc print ;
data skup ;
  set sbc ;
  if poc > 1000 then delete ;
proc print ;
data cis ; /* ciselnik pevne skupiny */
  set zaznam3 ; by skup ; if first.skup ;
  keep skup ;
data cis ;
  set cis ;
  ncskup = _n_ ;
data zaznam3 ;
  merge zaznam3 cis ; by skup ;
  drop skme sksp rokk obd ;
proc means ;
/*.....ciselnik HTD...vsechny laktace spolecne .....*/
data htd123 ;
  set zaznam1 zaznam2 zaznam3 ;
proc sort data=htd123 ; by htd ;
data cishtd ;
  set htd123 ; by htd ; if first.htd ;
  keep htd ;
data cishtd ;
  set cishtd ;
  nhtdv = _n_ ;
data htd123 ;
  merge htd123 cishtd ; by htd ;
                title "pocty v htd vsechny L 589";
proc means noprint; var ml ; by nhtdv ; output out=pru mean=;
  data sbc; set pru; poc=_freq_; keep nhtdv poc;
proc means ;
data htd123 ; set htd123 ; /*..... zapis 593 */
file kutd dlm=";" ;
  put nc lakt pl rokot mesot ncskup staj dku nhtd
      vek sp mezidob dl ml tu bi sb ;
data prover ; title "soubor htd KU 597";
  infile kutd dlm=";" ;
  input nc:$16. lakt pl:9. rokot mesot ncskup staj:$9. dku:$8. nhtd
      vek sp mezidob dl ml tu bi sb ;
proc means ;
proc datasets nolist;
delete cis cishtd htd123 prover prv12 sbc skup uskra uz uzkon uzkra zaznam1
zaznam2 zaznam3 ;
quit; run;
/*.....konec.....*/

```

3. Úprava MACE hodnot a základní odregresování

```

/*.....dyd9.sas.....*/
/* Interbull hodnoty MACE, priprava pro dalsi programy
   cteni Interbull + oprava meritka radek 43,55
   vstupni spolehlivost radek 46-53,78,89,98
   krajni meze PH radek 68,69
   dosazena dedivost radek 74,85,94
   omezeni rozmezi YD radek 82,83
   omezeni vahy radek
77,88,97

```

```

                vyrazeni zeme                                radek 123-131
                podil H                                    radek 151
                roky narozeni                               radek 152
/*.....VUZV 10.5.2016 */
options nodate nonumber ls=89 ps=10000;
    proc printto print= "/home/pribyl/tdmmace/vysctin9.lst" ;
    proc printto log= "/home/pribyl/tdmmace/logctin9.lst" ;
/*dm output 'clear';
dm log 'clear';*/
filename proh9 '/home/pribyl/tdmmace/hol113.itb';/*Interbull2011*/
filename proh14 '/home/pribyl/tdmmace/prodHOL1404r.itb';/*Interbull2014*/
filename zdvjm "/home/pribyl/tdmmace/zdvjm.txt";/*zdvojen jmen*/
filename zeme "/home/pribyl/tdmmace/zeme.txt";
filename rodupr "/home/pribyl/tdmmace/puvo";/*opravene rodokmeny*/
filename hrubci '/home/pribyl/tdmmace/hrubci'; /*hruby ciselnik jedince */
filename yd9 '/home/pribyl/tdmmace/yda9'; /*zapis dyd9*/
/*.....*/
data zeme ;                                title "3x zeme, cislo, jmen2, jmen3 26";
    infile zeme firstobs = 2 missover ;
    input czm $39-41 ze2 $42-43 ze3 $44-46 ;
        if czm = "840" then do ; ze2 = "US" ; ze3 = "USA" ; end ;
if czm = "." then delete; xx = czm; if xx*1 = 0 then delete; if czm = "" then
delete;
    proc means ;
    proc sort ;by czm ;
data zdvojm ;                                title "cteni zdvojenych jmen 33";
    attrib ciz cdo format=$16. length=$16 ;
    infile zdvjm firstobs = 1 missover ;
    input ple $1-3 cdo $5-20 ciz $22-37;
    proc means ;
    proc sort ; by ciz ;
data intt ;                                title "cteni Interbull 39";
/*infile proh9 recfm = v lrecl = 2900 firstobs = 1 missover ;
input plem $4-6 czm $10-12 id $13-25 spolm 1280-1281 phm 1282-1287
    spolt 1307-1308 pht 1309-1314 spolb 1334-1335 phb 1336-1341 ;
*/
infile proh9 recfm = v lrecl = 2900 firstobs = 1 missover ;
input plem $ 7-9 czm $ 10-12 id $ 13-25 opzem 35-37 opvla 38-40
    zem $ 1778-1780 spolm 1809-1810 phm 1811-1820 ;
    proc means ;
data intt ; set intt ;                                title "spolehlivost > 65 47";
    if spolm > 100 then delete ;
    if spolm < 65 /*20*/ then delete ;
        spolm = spolm*0.01 ;                                /* spolehlivost */
    if spolt > 100 then pht = . ;
    if spolt < 65 /*20*/ then pht = . ;
        spolt = spolt*0.01 ;
    if spolb > 100 then phb = . ;
    if spolb < 65 /*20*/ then phb = . ;
        spolb = spolb*0.01 ;
    phm =phm/1000 ; /*nove od2011*/ /*phm =phm/10; stare*/ /*meritko PH */
    pht = pht/1000 ;phb = phb/1000 ;
    a = 1 ;
    proc means ;
proc means data =intt noprint; var phm; output out=prum mean=;
data prum ;
    set prum ;
        mean=phm; keep mean a;
        a = 1 ;
Data intt ;
    merge intt prum; by a ;
data intt ;                                title "Interbull bez odlehlych 68";

```

```

set intt ;
/*   meze = mean - 3*593 ; if phm < meze then delete ;
   meze = mean + 3*593 ; if phm > meze then delete ;   */
drop meze mean a ;
proc means ;
data intt ; /*.....odregresovani.....*/
set intt ; title "odregresovani 75";
h2m = 0.39 ; /* mleko */
k = (1-h2m)/h2m ;
euz = k*spolm/(1-spolm) ;
If euz > 155 then euz = 155 ; /*max spol 0,99*/
If spolm > 0.99 then spolm = 0.99 ;
euzm = round(euz,0.01) ; /*pocet uzitkovosti*/
/* oprava na prijatelnu smer. odchylku */
yd = /* 0.00046 */ phm/spolm ;
dol = -319 -3*794 ; hor = -319 +3*794 ;
if yd < dol then delete ; if yd > hor then delete ;
ydm = round(yd,1) ; /*odchylka uzitkovosti*/
h2t = 0.38 ; /* tuk */
k = (1-h2t)/h2t ;
euz = k*spolt/(1-spolt) ;
If euz > 155 then euz = 155 ; /*max spol 0,99*/
If spolt > 0.99 then spolt = 0.99 ;
euzt = round(euz,0.01) ; /*pocet uzitkovosti*/
/* oprava na prijatelnu smer. odchylku */
yd = /* 0.00046 */ pht/spolt ;
ydt = round(yd,1) ; /*odchylka uzitkovosti*/
h2b = 0.37 ; /* bilkoviny */
k = (1-h2b)/h2b ;
euz = k*spolb/(1-spolb) ;
If euz > 155 then euz = 155 ; /*max spol 0,99*/
If spolb > 0.99 then spolb = 0.99 ;
euzb = round(euz,0.01) ; /*pocet uzitkovosti*/
/* oprava na prijatelnu smer. odchylku */
yd = /* 0.00046 */ phb/spolb ;
ydb = round(yd,1) ; /*odchylka uzitkovosti*/
proc means ;
proc sort ; by czm ;
data inter ; /* oprava zeme */
merge intt(in=inte) zeme ; by czm ; if inte ;
data inter ;
set inter ;
attrib ciz format=$16. length=$16;
if ze3 = "" then ze3 = czm ;
ciz = ze3||id ;
keep ciz ze3 spolm phm euzm ydm spolt pht euzt ydt spolb phb euzb ydb;
proc univariate data=inter normal plot; title "vsichni jedinci inter 115";
var spolm euzm phm ydm;
proc freq data = inter; tables ze3 ; run ;
proc sort data = inter; by ciz ;
data inter ; /* oprava dvojena jmena */
merge inter(in=intr) zdvojm ; by ciz ; if intr ;
data inter ;
set inter ;
if cdo ne "" then ciz = cdo ;
if ze3 = "ALB" or ze3 = "AND" or ze3 = "ARM" or ze3 = "AUS"
or ze3 = "AZE" or ze3 = "BGR" or ze3 = "BIH" or ze3 = "BLR"
or ze3 = "BRA" or ze3 = "CUB" or ze3 = "EME" or ze3 = "EST"
or ze3 = "FRO" or ze3 = "GEO" or ze3 = "IRL" or ze3 = "ISR"
or ze3 = "JOR" or ze3 = "KAZ" or ze3 = "KGZ" or ze3 = "KWT"
or ze3 = "LBN" or ze3 = "LTU" or ze3 = "LVA" or ze3 = "MDA"
or ze3 = "MKD" or ze3 = "MLT" or ze3 = "NZL" or ze3 = "ROU"

```

```

or ze3 = "SAU" or ze3 = "TJK" or ze3 = "TKM" or ze3 = "TUN"
or ze3 = "TUR" or ze3 = "UZB" or ze3 = "YEM" or ze3 = "ZAF" then delete ;
proc freq ; tables ze3 ; title " po vyrazeni odlehlých zemi 133";
proc means ;
proc sort data = inter; by ciz ;
data cis ; title " ciselnik 136";
infile hrubci dlm=";" ;
input ciz $1-16 nc $18-26 ;
proc means ;
data uzkon ;
merge inter(in = krava) cis ; by ciz ; if krava ;
proc means;
data uzkra ;
set uzkon ;
if nc = . then delete ;
PROC MEANS ;
proc sort ; by nc ;
data rodupr ; title " H v rodokmenu nad 85% 148";
attrib nc format=$9. length=$9;
infile rodupr missover ;
input nc $1-16 podHj 56-58 podrj 60-62 roknej 72-75 ;
pod = podhj + podrj ; if pod < 85 then delete ;
if roknej < 1985 then delete ;
proc means ;
proc sort ; by nc ;
data int ;
merge uzkra(in=mac) rodupr(in=rodk) ; by nc ; if mac ; if rodk ;
proc means ;
proc sort ; by nc spolm ;
data int2 ; title " zapis YD 160";
set int ; by nc ; if last.nc ;
file yd9 dlm=";" ;
put nc roknej spolm phm euzm ydm spolt pht euzt ydt spolb phb euzb ydb;
proc means ;
proc corr ;
var spolm phm euzm ydm spolt pht euzt ydt spolb phb euzb ydb;
proc datasets nolist;
delete cis int int2 inter intt prum rodupr uzkon uzkra zdvojm zeme ;
quit;run;
/*.....konec .....*/

```

4. Úprava SNP souborů

```

/*.....priSNvse.sas.....*/
/* priprava SNP pro dalsi programy podle oznaceni zeme
a vicenasobneho oznaceni nebo podle souboru k dlouhovek radek 54-71
vyber byku se snp VUZV 10.5.2016...*/
/*.....*/
filename zdvjm "/home/pribyl/tdmmace/zdvjm.txt";/*zdvojen jmen*/
filename zeme "/home/pribyl/tdmmace/zeme.txt";
filename hrubci '/home/pribyl/tdmmace/hrubci'; /*hruby ciselnik jedince */
filename cisel '/home/pribyl/tdmmace/ciseln'; /*ciselnik dlouhove*/
filename sezg '/home/pribyl/tdmmace/sezg'; /*seznam g dlouhove*/

filename sn1 "/home/pribyl/tdmmace/Genotypy11-2014/kgnp41u.txt";/*cteni
souboru snp*/
filename sn2 "/home/pribyl/tdmmace/Genotypy11-2014/kgnp42u.txt";
filename sn3 "/home/pribyl/tdmmace/Genotypy11-2014/kgnp43u.txt";
filename sn4 "/home/pribyl/tdmmace/Genotypy11-2014/kgnp44u.txt";
filename sn5 "/home/pribyl/tdmmace/Genotypy11-2014/kgnp45u.txt";
filename sn6 "/home/pribyl/tdmmace/Genotypy11-2014/kgnp46u.txt";

```

```

filename sn7 "/home/pribyl/tdmmace/Genotypy11-2014/kgnp47u.txt";
filename sn8 "/home/pribyl/tdmmace/Genotypy11-2014/kgnp48u.txt";
filename sn9 "/home/pribyl/tdmmace/Genotypy11-2014/kgnp49u.txt";
/* pise snp */
filename snp "/home/pribyl/tdmmace/alsnvse.txt";/*alely SNP */
filename cissn "/home/pribyl/tdmmace/cissnvs";/*ciselnik lokusu*/
filename pomoci "/home/pribyl/tdmmace/gbycvse";/* poradi byka*/
filename souct "/home/pribyl/tdmmace/souct";/*pomocne cetnosti alel*/
/*.....*/
options nodate nonumber ls=89 ps=10000;
proc printto print= "/home/pribyl/tdmmace/vysprsnvs.lst" ;
proc printto log = "/home/pribyl/tdmmace/logprsnvs.lst" ;
/*dm output 'clear';
dm log 'clear'; */
%let maxsn = 165000 ; /* možný počet řádků SN na býka */
/*.....zeme.....*/
data zeme ; title "3x zeme, cislo, jmen2, jmen3 34";
infile zeme firstobs = 2 missover ;
input czm $39-41 ze2 $42-43 ze3 $44-46 ;
if czm = "840" then do ; ze2 = "US" ; ze3 = "USA" ; end ;
if czm = "." then delete; xx = czm; if xx*1 = 0 then delete; if czm = "" then
delete;
proc means ;
proc sort ;by czm ;
/*.....zdvojena jmena.....*/
data ze ; title "zdvojena jmena 42";
attrib cdo ciz format=$16. length=$16;
infile zdvjm ;
input plin $1-3 cdo 5-20 ciz 22-37 ;
proc means ;
proc sort ; by ciz ;
/*.....ciselnik jedincu.....*/
data cis ; title " ciselnik 49";
infile hrubci dlm="," ;
input ciz $1-16 nc 18-26 ;
proc means ;
proc sort ; by ciz ;
/*.....data cisel2 ; title " ciselnik dlouhovekost 54";
infile cisel ;
input ciz $1-16 kod $18-37 ncj 39-46 ; run;
proc means ;
proc sort ; by ncj ;
data sezg ; title " seznam genot dlouhovekost 59";
infile sezg ;
input ncj 1-7 ;
proc means ;
proc sort ; by ncj ;
data dlouho ;
merge cisel2 sezg(in= dloh) ; by ncj ; if dloh ;
keep ciz ;
proc means ;
proc sort ; by ciz ;
data cis ;
merge cis dlouho(in= dlou) ; by ciz ; if dlou ;
proc means ;.....*/
/*.....kontrola snp.....*/
%macro vembykysnp;
%do i=1 %to 9;
data snp; title "lokusy kvalita &i 75";
infile sn&i missover;
input czm $1-3 cj $4-16 snjm $18-57 sncis 59-63 all $65 al2 $67
gcscsco 69-74 gtsco 76-81;

```

```

attrib ciz format=$16. length=$16;
      ciz = czm||cj ;
      obe = compress(all1||al2) ;
proc means ; var sncis gcsco gtsco ;
proc freq ; tables obe ;
data snp ;
  set snp ;
  if all1 ne "A" & all1 NE "B" then all1 = "-" ;
  if al2 ne "A" & al2 NE "B" then al2 = "-" ;
  if all1 = "-" & al2 = "-" then delete ;
proc sort data=snp; by czm ;
  data snp; /*oprava zeme 90*/
    merge snp(in=zesn) zeme ; by czm ; if zesn ;
data snp;
  set snp;
  if ze3 ne "" then ciz = ze3||cj;
  proc sort ; by ciz ;
data snp; /*oprava zdvojeniny 96*/
  merge snp(in=zesn) ze ; by ciz ; if zesn ;
data snp;
  set snp;
  if cdo ne "" then ciz = cdo ;
  drop cj cdo czm plin ;
proc sort ; by ciz ;
data pocet ; title "pocety genot byku ve skupine &i 103";
  set snp ; by ciz ; if first.ciz ;
  keep ciz ;
  proc means ;
data snp11 ; title "pocety alel na byka ve skup &i 107";
  set snp ;
  alela = 1 ;
  keep alela ciz ;
  proc means noprint; var alela ; by ciz ;
  output out=pru mean=;
  data sbc; set pru; pocet=_freq_;
  keep ciz pocet ;
  proc means ;
data snp&i;
  merge snp(in= sni) cis(in=puv) ; by ciz ; if sni ; if puv ;
  sk = &i ; title "skupiny podle skupin vzorku zemi &i 118" ;
  keep nc snjm sncis all1 al2 sk ;
  proc means ;
&end;
  data snp ; /*spojeni souboru 122*/
  set snp1 ;
&do i=2 %to 9 %by 1 ;
  data snp ;
  set snp snp&i ;
&end;
&mend;
%vembykysnp;
proc datasets nolist;
delete
pocet sbc snp1 snp2 snp3 snp4 snp5 snp6 snp7 snp8 snp11 ze zeme ;
quit; run;
/*.....precislovani alel .....*/
proc sort data=snp ; by snjm ; title " prejmenovani genotypy 135";
  proc means ;
data sn ; /* stanoveni snp */
  set snp;
  format sn1 sn2 sn 1. ;
  if nc = . then delete ;

```

```

        if snjm = "" then delete ;
        if all = "A" then sn1 = 1 ;
    else if all = "B" then sn1 = 2 ;
    else sn1 = 0 ;
        if al2 = "A" then sn2 = 1 ;
    else if al2 = "B" then sn2 = 2 ;
    else sn2 = 0 ;
        if sn1 = 1 and sn2 = 1 then sn = 0 ;
    else if sn1 = 1 and sn2 = 2 then sn = 1 ;
    else if sn1 = 2 and sn2 = 1 then sn = 1 ;
    else if sn1 = 2 and sn2 = 2 then sn = 2 ;
    else if sn1 = 0 and sn2 = 0 then sn = . ;
    else if sn1 = 0 then sn = . ;
    else if sn2 = 0 then sn = . ;
    keep nc snjm sncis sn1 sn2 sn sk ;
data snp;
    set sn ; by snjm ; if first.snjm ;
data cislsn ; title " ciselnik lokusu vse 158";
    set snp; ; cissn = _n_ ;
    keep snjm sncis cissn ;
    file cissn ;
    put snjm $1-40 sncis 42-47 cissn 49-54 ;
    proc means ;
proc means noprint; var cissn ; title "pocet byku v alelach 164";
    output out=pru mean=;
data sbc; set pru; poceal=_freq_;
    keep poceal ;
    proc means ;
data poceal ;
    set sbc ;
data snp ; title " prejmenovane alely 171";
    merge sn(in=jmen) cislsn ; by snjm ;
    keep nc snjm sncis cissn sn1 sn2 sn sk ;
    proc sort ; by nc ;
    proc means; run;
/*.....pomoc cislo byka .....*/
data cisby ;
    set snp ; by nc ; if first.nc ;
data pomoc ; title " pomocne poradi g byku 179";
    set cisby ;
    keep nc pomcb ;
    pomcb = _n_ ;
    file pomoci ;
    put nc 1-9 pomcb 11-15 ;
    proc means ;
proc means noprint; var pomcb ; title "pocet alel u g byku 186";
    output out=pru mean=;
data sbc; set pru; pocegb=_freq_;
    keep pocegb ;
    proc means ;
data pocegb ;
    set sbc ;
data pruc ;
    merge snp(in= loku) pomoc(in=pcis) ; by nc ; if loku ; if pcis ;
    keep pomcb cissn sn1 sn2 sn sk;
/*.....byk jen jednou .....*/
proc iml ;
    use pocegb ;
        read all into pocgb ;
            pocgb = max(pocgb) ; print pocgb ;
    close pocegb ;
    use poceal ;

```

```

        read all into pocal ;
                poceal = max(pocal) ;                                print poceal ;
close poceal ;
        s1= j(poceal,6,.);
create mark from s1;
use pruc ;
        do i = 1 to pocegb by 1;                                     /* po bycich 209*/
                x = j(&maxsn,6,.) ;
                s1= j(poceal,6,.);
read all var{pomcb cissn sn1 sn2 sn sk} into x where (pomcb=i);
                nro=nrow(x);                                       /* print &maxsn i nro ;*/
                /* při opakovani byka kazde SN jen jednou 214*/
        do k = 1 to nro by 1 ;
                cislo = x[k,2] ;
                if x[k,5] ^= . then do ;
                        s1[cislo,1] = x[k,1];
                        s1[cislo,2] = x[k,2];
                        s1[cislo,3] = x[k,3];
                        s1[cislo,4] = x[k,4];
                        s1[cislo,5] = x[k,5];
                        s1[cislo,6] = x[k,6];
                end ;
                end;
        append from s1 ;
        end ;
close pruc ;
/*.....*/
data pruc ;                                                         title " snp byk 1x 230";
        set mark ;
                keep pomcb cissn sn1 sn2 sn skz;
                pomcb = col1; cissn = col2;
                sn1 = col3; sn2 = col4; sn = col5; skz = col6;
                if pomcb = 0 then delete ; if pomcb = . then delete ;
        proc means ;
proc datasets nolist;
        delete mark ;
        proc sort ; by cissn ;
data pruc ;
        merge pruc cislsn ; by cissn ;
        proc sort ; by pomcb ;
data pruc ;                                                         /* zapis snp soubor 243 */
        merge pruc pomoc ; by pomcb ;
        file snp dlm=";" ;
                put nc pomcb snjm sncis cissn sn1 sn2 sn skz ;
/*.....cetnosti alel .....*/
data pocb ;
        set pruc ; by pomcb ; if first.pomcb ;
proc sort data = pruc ; by cissn ;
data _null_ ;                                                         title "vypocet prumeru alel 251";
        keep cissn souc pocet celpoc ;
        file souct ;
                set pruc ; by cissn ;
                if first.cissn then do ;
                        souc = 0 ; pocet = 0 ; celpoc = 0 ;
                end ;
                        celpoc + 1 ;
                if sn ne . then do ;
                        pocet + 1 ;
                        souc + sn ;
                end ;
if last.cissn then
        put cissn 1-6 pocet 8-13 souc 15-20 celpoc 22-26;

```

```

data prusn ;                               title "cetnosti SN 265";
  infile souct ;
  input cissn 1-6 pocet 8-13 souc 15-20 cel poc 22-26;
  proc means ;
/*.....cetnosti uvnitr byku .....*/
proc sort data = pruc ; by pomcb ;
data _null_ ;                               title "vypocet prumeru byku 271";
  keep pomcb souc pocet cel poc ;
  file souct ;
  set pruc ; by pomcb ;
  if first.pomcb then do ;
    souc = 0 ; pocet = 0 ; cel poc = 0 ;
  end ;
    cel poc + 1 ;
  if sn ne . then do ;
    pocet + 1 ;
    souc + sn ;
  end ;
  if last.pomcb then
  put pomcb 1-6 pocet 8-13 souc 15-20 cel poc 22-26;
data prusn ;                               title "cetnosti byku 285";
  infile souct ;
  input pomcb 1-6 pocet 8-13 souc 15-20 cel poc 22-26;
  proc means ;
proc datasets nolist;
delete cis cisby cislsn mark pocb poceal pocegb pocet pomoc pruc pruc
prusn sbc sn snp;
quit; run;
/*.....konec .....*/

/*.....snvyber.sas.....*/
/*
      rok narozeni                vyber snp
      podil H      + R              radek 48
      pocty lokusu na byka      radek 81
      pocty byku na lokus       radek 120
      MAF                          radek 121
                                  VUZV 10.5.2016...*/
/*.....*/
/* cte */
filename rodupr "/home/pribyl/tdmmace/puvo"; /*opravene rodokmeny*/
filename gbyci  "/home/pribyl/tdmmace/gbycvse";/* ciselnik G byku */
filename cissn  "/home/pribyl/tdmmace/cissnvs";/*ciselnik lokusu*/
filename snpvse  "/home/pribyl/tdmmace/alsnvse.txt";/*alely vyber */
/* pise */
filename snvyhr "/home/pribyl/tdmmace/snvyhr.txt";/*alely vyber H+R */
filename cisnhr "/home/pribyl/tdmmace/cisnhr";/*cisel lokusu vyber H+R*/
filename gbychr "/home/pribyl/tdmmace/gbychr";/* cisel byku vyber H+R*/
filename snvyh  "/home/pribyl/tdmmace/snvyh.txt";/*alely vyber H*/
filename cisnh  "/home/pribyl/tdmmace/cisnh";/*cisel lokusu vyber H*/
filename gbych  "/home/pribyl/tdmmace/gbych";/* cisel byku vyber H*/
filename souct  "/home/pribyl/tdmmace/souct";/*pomocne cetnosti alel*/
/*.....*/
options nodate nonumber ls=89 ps=10000;
  proc printto print= "/home/pribyl/tdmmace/vysvybersn.lst" ;
  proc printto log   = "/home/pribyl/tdmmace/logvybersn.lst" ;
/*dm output 'clear';
dm log 'clear'; */
%let maxsn = 165000 ;                       /* možný počet řádků SN na býka */

```

```

/*.....rodokmen.....*/
data soub ; title " rodokmen 32";
  infile rodupr missover;
  input nc 1-16 nco 18-33 ncm 35-50
         podCj 52-54 podHj 56-58 podRj 60-62 podOj 64-66 podcel 68-70
         roknaj 72-75 roknaj 77-80 zj $82-84 zo $86-88 zm $90-92 ;
  proc means ;

/*.....ciselnik byku .....*/
data cisel ; title " ciselnik g byku vse 39";
  infile gbyci missover;
  input nc 1-9 pomcb 11-15 ;
  proc means ;

/*.....byci podily H + R .....*/
proc sort data = soub ; by nc ;
data podily ; title " vyrazeni roky podil H + R 45";
  merge soub cisel(in=byc) ; by nc ; if byc ;
  podhr = podhj + podrj ;
  if podhr < 85 then delete ; if roknaj > 2012 then delete;
  keep nc pomcb podhj podrj roknaj ;
  proc means ;

/*.....ciselnik lokusy.....*/
data cisel ; title " ciselnik lokusu vse 52";
  attrib snjm format=$40. length=$40;
  infile cissn missover;
  input snjm $1-40 sncis 42-47 cissn 49-54 ;
  proc means ;

/*.....snp.....*/
data soub ; title " soubor snp 58";
  attrib snjm format=$40. length=$40;
  infile snpvse dlm=";" missover;
  input nc pomcb snjm sncis cissn sn1 sn2 sn skz ;
  if sn1 = 0 and sn2 = 0 then delete ;
  if sn1 = . and sn2 = . then delete ;

  proc means ;
  proc sort data = soub ; by nc ;
  data pocb ; title " snp jen H+R 66";
merge soub(in= loku) podily(in= holred); by nc; if holred; if loku;
  proc means ;

/*.....provereni poctu lokusu na byka.....*/
data sn ; title "pocety alel na byka 70";
  set pocb ;
  alela = 1 ;
  keep alela nc ;
  proc means noprint; var alela ; by nc ;
  output out=pru mean=;
  data sbc; set pru; pocet=_freq_;
  keep nc pocet ;

  proc means ;
data sn ;
  merge pocb sbc ; by nc ; /*vyhozeni byka s chybejicimi alel*/
data soub ; title "vyhozeni byku podle poctu alel 81";
  set sn ;
  po = 54609*0.90 ;
  if pocet < po then delete ;
  keep nc pomcb snjm sncis cissn sn1 sn2 sn skz ;
  proc means ;
  proc sort data = soub ; by pomcb ;
data pocb ; title "pocet byku 89";
  set soub ; by pomcb ; if first.pomcb ;
  keep pomcb ;
  proc means noprint; var pomcb ;
  output out=pru mean=;

```

```

data sbc; set pru;
  celpoc = _freq_ ;
  keep celpoc a ;
  a = 1 ;
proc means ;
/*.....cetnosti alel ...MAF.....*/
proc sort data = soub ; by cissn ;
data _null_ ; title "vypocet prumeru alel 101";
  keep cissn souc pocet ;
  file souct ;
  set soub ; by cissn ;
  if first.cissn then do ;
    souc = 0 ; pocet = 0 ;
  end ;
  if sn ne . then do ;
    pocet + 1 ;
    souc + sn ;
  end ;
  if last.cissn then
    put cissn 1-6 pocet 8-13 souc 15-20 ;
data prusn ; title "cetnosti SN 114";
  infile souct ;
  input cissn 1-6 pocet 8-13 souc 15-20 ;
  a = 1 ;
data sn ; /*omezeni na pocet uvnit lokusu */
  merge prusn sbc ; by a ; drop a ;
if pocet = . or pocet < 0.95*celpoc then delete;
  prum = souc/pocet ;
if prum < 0.1 or prum > 1.9 then delete ; /*omezeni na MAF 5%*/
  rozpt = prum*(2-prum)/2 ;
proc means ;
data prusn ; title "alely po vyhozeni SN 125";
  merge soub sn(in=prau) ; by cissn ; if prau ;
  keep nc pomcb snjm sncis cissn sn1 sn2 sn skz ;
proc means ;
/*.....novy cisel byku H+R.....*/
proc sort data=prusn ; by nc ;
data sn ; title "novy ciselnik byku H+R 131";
  set prusn ; by nc ; if first.nc ;
data sbc ;
  set sn ;
  cissb2 = _n_ ;
  file gbychr ;
  put nc 1-16 cissb2 18-24;
proc means ;
data soub ;
  merge prusn sbc ; by nc ;
  keep nc pomcb cissb2 snjm sncis cissn sn1 sn2 sn skz ;
proc means ;
/*.....novy cisel lokusu.....*/
proc sort data=soub ; by cissn ;
data sn ; title "novy ciselnik lokusu H+R 145";
  set soub ; by cissn ; if first.cissn ;
data sbc ;
  set sn ;
  cissn2 = _n_ ;
  file cissnhr ;
  put snjm 1-40 cissn2 42-47 ;
proc means ;
data sn ; title "zapis upravene lokusy H+R 153";
  merge soub sbc ; by cissn ;
  file snvyhr dlm=";" ;

```

```

        put cisb2 cissn2 sn1 sn2 sn skz ;
        keep nc cisb2 snjm sncis cissn2 sn1 sn2 sn skz ;
    proc means ;
/*.....vyber jen H .....*/
        proc sort data = sn ; by nc ;
data podilh ;                                title " vyrazeni podil H 161";
        set podily ;
            if podhj < 85 then delete ;
            keep nc pomcb podhj podrj roknej ;
        proc means ;
data podil ;
        set podilh ;
            keep nc pomcb ;
data soub ;                                title " alely jen H 169";
        merge podil(in=byci) sn(in=alely) ; by nc ; if byci ; if alely ;
        proc means ;
/*.....novy cisel byku H.....*/
data sn ;                                title "novy ciselnik byku H 173";
        set soub ; by nc ; if first.nc ;
        data sbc ;
            set sn ;
                cisb22 = _n_ ;
            file gbych ;
                put nc $1-16 cisb22 18-24 ;
        proc means ;
data prusn ;
        merge soub sbc ; by nc ;
            keep nc pomcb cisb22 snjm sncis cissn2 sn1 sn2 sn skz ;
        proc means ;
/*.....novy cisel lokusu H.....*/
proc sort data=prusn ; by sncis ;
data sn ;                                title "novy ciselnik lokusu H 187";
        set prusn ; by sncis ; if first.sncis ;
        data sbc ;
            set sn ;
                cissn22 = _n_ ;
            file cish ;
                put snjm $1-40 cissn22 42-47 ;
        proc means ;
data soub ;                                title "zapis upravene lokusy 195";
        merge prusn sbc ; by sncis ;
        file snvyh dlm=";" ;
            put cisb22 cissn22 sn1 sn2 sn skz ;
        proc means ;
proc datasets nolist;
delete cisel pocb podil podilh podily prusn sbc sn soub ;
quit; run;
/*.....konec .....*/

```

5. Výběr genotypovaných jedinců

```

/*.....vybPeTD.sas.....2013 Petr Pesek -Diplomova prace.*/
/* vyber byku se snp - vyrazeni odlehlých podle h, nebo hr   radek 15,17,18
   sestaveni soustavy rovnic
   reseni soustavy rovnic pomoci preconditioned conjugate gradient
   vyber byku podle chyby predpovedi PGH
   dosazeni dedivosti                                radek 198
   nastaveni spolehlivosti reseni                    radek      261,267
   nastaveni pripustne chyby                          radek 358
   zapis novych snp, gbyc, dyd                        10.5.2016 VUZV ... */

```

```

/*.....*/
options nodate nonumber ls=89 ps=10000;
    proc printto print= '/home/pribyl/tdmmace/vyschyb2.lst';
    proc printto log = '/home/pribyl/tdmmace/logchyb2.lst';
/*.....vstupy.....*/
Filename snpa      "/home/pribyl/tdmmace/snvyh.txt"; /*vstup alel h, nebo hr*/
filename DYDd      "/home/pribyl/tdmmace/yda9";      /*vstup dyd edc*/
Filename cissnp    "/home/pribyl/tdmmace/cisnh"; /*ciselnik lokusu*/
Filename gbyca     "/home/pribyl/tdmmace/gbych"; /*stary seznam Gbyku*/
/*.....vystupy.....*/
Filename byci_T1   '/home/pribyl/tdmmace/byci_t1';
Filename byci_T2   '/home/pribyl/tdmmace/byci_t2';
Filename bycichyb '/home/pribyl/tdmmace/bycichyb.txt'; /*byci velka chyba
predpovedi PGH*/
Filename efsnp     '/home/pribyl/tdmmace/efsnp.txt'; /*efekty lokusu*/
Filename phbyc     '/home/pribyl/tdmmace/phbyc';     /*PH byku*/
Filename gbyc      '/home/pribyl/tdmmace/gbycb';     /*novy seznam Gbyku*/
Filename snp        '/home/pribyl/tdmmace/alelsnb.txt'; /*vystup alel*/
filename DYD        '/home/pribyl/tdmmace/ydb';      /*vystup dyd edc*/
/*.....*/
Data SNPa;          /*nacita a upravuje pocet snp 29*/
    infile snpa dlm=";" ;
    input cisb22 cissn22 sn1 sn2 sn skz ;
        proc sort; by cisb22;
data snp1 ;
    set snpa ;
        keep cisb22 cissn22 sn ;
data Gbyci;          title "G byci 36";
    infile gbyca;
        input nc 1-16 cisb22 18-24;
        proc means ;
        proc sort; by nc ;
Data DYD1;
    infile dydd dlm=";" ;
    input nc roknaj spolm phm euzm ydm spolt pht euzt ydt spolb phb euzb ydb;
        proc sort; by nc ;
Data celk;          title "Byci s dyd v seznamu 45";
    merge gbyci(in=geno) dyd1(in=int); by nc ; if geno ; if int ;
        proc means;
Data celk;          /*cislo G byku s YD*/
    set celk;
        ncj=_n_;
Data ID;           /*vektor spolehlivosti pro porovnaní pri vystupech 51*/
    set celk;
        keep nc ncj phm spolm;
data pomdyd;       /*vytvareni odmocnin vah a vazenych uzitkovosti 54*/;
    set celk;
        odeuz=sqrt(euzm);
        vazyd=ydm*odeuz;
        keep ncj vazyd ydm odeuz;
data YD;           /*datovy soubor pro porovnaní s predpovedi 59*/
    set pomdyd;
        keep ydm;
Data oded;        /*odmocnene edc 62*/
    set pomdyd;
        keep ncj odeuz;
        proc sort; by ncj;
data oded2;
    set oded;
        keep odeuz;
data pomdyd2;
    Set pomdyd;

```

```

        keep vazyd;
data pomsnp;                               /*ciselnik-stara a nova id 72*/
    set celk;
        keep ncj cisb22;
    proc sort; by cisb22;
Data pomsnp2;                               title "snp byku s dyd 76";
    merge pomsnp snp1(in=jed); by cisb22; if jed;
Data pomsnp2;
    set pomsnp2;
    if ncj="." then delete;
    format pomc $16.;
    pomc=(1000000*ncj)+cissn22 ;
    keep pomc ncj cissn22 sn;
    proc sort; by pomc;
    proc means;
data njed; /*pomocny soubor pro jedincu pro vytvoreni soustavy 86*/
    set pomsnp;
        keep ncj;
data nsnp1;                                title "ciselnik snp 89";
    attrib snjm format=$40. length=$40;
    infile cissnp; input snjm $1-40 cissn22 42-47 ;
    proc means ;
data nsnp ; set nsnp1 ; keep cissn22 ;
    /*vytvareni sestavy snp pro vypocet poctu a souctu lokusu.*/
proc iml;                                  /*vytvareni sestavy snp, pocet jedincu 95*/
    use njed;
        read all into J1;
        close njed ;
        j=max(j1);                               /*max jedincu 99*/
    use nsnp;
        read all into S1;
        close nsnp ;
        s=max(s1);                               /*soucet cely soubor max snp 103*/
    Byk=j(s,2,.);
    create alllok from byk;
do i=1 to j;
    a=i;
    Do j=1 to s;
        b=j;
        Byk[b,1]=a;
        Byk[b,2]=b;
    end;
    append from byk;                               /*pridava dalsi byky 113*/
end;
free byk;
data allsnp;
    set alllok;
    ncj=col1; cissn22=col2;
    drop col1 col2;
    format pomc $16.;
    pomc=(1000000*ncj)+cissn22;
    proc sort; by pomc;
    proc means;                               title"Celkovy vektor jedincu a snp 123";
Data allsnp2;
    merge pomsnp2 allsnp; by pomc;
    keep ncj cissn22 sn;
    proc sort; by ncj;
Data poms;
    set allsnp2;
        keep ncj cissn22 ;
Data allsnp3;
    merge allsnp2 oded; by ncj;

```

```

Data allsnp3;
  set allsnp3;
  if sn="." then sn=4;
  vazgen=sn*odeuz;
  if sn=4 then vazgen=".";
  keep ncj cissn22 vazgen;
  proc means; title "T pro odhad 139";
Data Tmat; /*data pro vytvoreni T matice bez vah 140*/
  merge allsnp2 oded; by ncj;
  keep ncj cissn22 sn ;
  proc sort; by ncj cissn22 ;
  proc means; title "T pro predpoved 144";
/*.....Vytvoreni matice T a doplnovani chybejicich snp prumerem...*/
options fullstimer;
proc iml;
start;
  use allsnp3;
  read all into LHS; /*nacistani jedincu, snp a vazenych genotypu 150*/;
  close allsnp3;
  use pomdyd2;
  read all into VAZYD; /*vazene yield deviations 153*/;
  close pomdyd2;
  use oded2;
  read all into VAZX; /*vazeny vektor 1 156 */;
  close oded2;
  use yd;
  read all into yd;
  close yd;
  use njed;
  read all into J1;
  close njed;
  use nsnp;
  read all into S1;
  close nsnp;
  zaok=1e-8;
  j=max(j1); /*max jedincu 168*/;
  s=max(s1); print j s ; /*max snp 169*/;
T=J(j, s, .);
it=s*j; /*pocet radku v souboru 171*/;
Do w=1 to it;
  m=LHS[w,1];
  n=LHS[w,2];
  T[m,n]=LHS[w,3];
end;
free lhs;
/*doplneni neznamych lokusu prumerem pres byky 178*/;
vek1=j(1, s, 0);
vek2=j(1, s, 0);
Do i=1 to s;
  do h=1 to j;
    if T[h,i]^=. then do;
      vek1[1,i]=vek1[1,i]+T[h,i];
      vek2[1,i]=vek2[1,i]+1;
    end;
  end;
end;end;
Tpom=vek1/vek2;
Tpom=round(tpom, zaok);
do u=1 to j by 1;
  do h=1 to s by 1;
    if T[u,h]=. then do;
      T[u,h]=tpom[1,h];
    end;
  end;
end;

```

```

end; end;
free tpom;
/*.....vytvareno matice G.....*/
h2=0.35;
I=i(s);
  k=(1-h2)/h2;
  f=k*s;
G=f*I;
/*.....vytvareni soustavy rovnic - leva strana .....*/
  XX=VAZX`*VAZX;
  XT=VAZX`*T;
  TX=T`*VAZX;
n=5000; /*velikost bloku po jakych se sestavuje C22 207*/
/* iterativni vytvoreni C22*/
lhs=j(s,s,.);
do i=1 to s by n;
  do h=1 to s by n;
    a1=i+(n-1);
    if s-i<n then do;
      a1=i+(s-i);
    end;
    a2=h+(n-1);
    if s-h<n then do;
      a2=h+(s-h);
    end;
    lhs[i:a1,h:a2]=t[1:j,i:a1]*t[1:j,h:a2];
end;end;
  LHS=lhs + g;
  L1=XX||XT;
  LHS=TX||LHS;
  LHS=L1//LHS;
free l1;
/*.....vytvareni soustavy rovnic- prava strana.....*/
  XY=VAZX`*VAZYD;
  RHS=T`*VAZYD;
  RHS=XY//RHS;
free xx xt tx xy vazyd t;
/*.....reseni soustavy preconditioned conjugate gradient.....*/
  pocn=S+1; /*pocet neznamych 232*/
  v=j(pocn,1,1); /*nastreleny vektor reseni 234*/
  vs=j(pocn,1,1);
  iM=j(pocn,1,.);
do i=1 to pocn;
  p=LHS[i,i];
  Im[i,1]=1/p;
end;
  r=RHS-LHS*v;
  p=im#r;
do i= 1 to 5000 by 50;
  do f=1 to 50;
    cisit=i+f-1;
    w=j(pocn,1,.);
    w=LHS*p;
    lam=p`*w;
    alf=(p`*r)/lam;
    v=v+alf*p;
    r=r-alf*w;
    /*oprava rezidui pro zpresneni pri kazde 50-te iteraci 252*/
    if f=50 then r=rhs-lhs*v;
    beta=(-r`*w)/lam;
    p=r+beta*p;
    w=im#r;
  end;
end;

```

```

vn=v;
rozres=VN-VS;
rozres=max(rozres);
rozres=abs(rozres);
if rozres<1e-15 then goto uk;
vs=vn;
/*pocita rozdil mezi LS*b a PS 263*/

maxr=max(r);
maxr=abs(maxr);
print cisit maxr rozres;
if maxr<10e-6 then goto uk; /*ukoncuje reseni pokud r<10e-8 267*/
end;
end;
uk:
lok=j(pocn,1,0);
do n=2 to pocn;
lok[n,1]=n-1;
end;
celkef=lok||v[1:pocn,1];
free lhs im;
create vys from celkef;
append from celkef;
/*.....vytvareni predpovedi PGH.....*/
/*.....T bez vah.....*/
use tmat;
read all into TT; /*nacistani jedincu, snp a vazenych genotypu 282*/
close tmat ;
T=J(j,s,.);
it=s*j; /*pocet radku v souboru 285*/
do w=1 to it;
m=TT[w,1];
n=TT[w,2];
T[m,n]=TT[w,3];
end;
free lhs;
/*doplneni neznamych lokusu prumerem pres byky 292*/
vek1=j(1,s,0);
vek2=j(1,s,0);
do i=1 to s;
do h=1 to j;
if T[h,i]^=. then do;
vek1[1,i]=vek1[1,i]+T[h,i];
vek2[1,i]=vek2[1,i]+1;
end;
end;end;
tpom=vek1/vek2;
do u=1 to j by 1;
do h=1 to s by 1;
if T[u,h]=. then do;
T[u,h]=tpom[1,h];
end;
end; end;
free tpom;
ef=v[1:pocn,1];
X=j(j,1,1);
T=x||T;
pred=(T*ef);
pred=round(pred,0.1);
chyba=(yd-PRED);
absch=abs(chyba);
pred=J1||pred||yd||chyba||absch;
create pr from pred;

```

```

append from pred;
finish;run; quit;
Data EFSNP;
    set vys;
        cissn22 =col1; efekt=col2;
        keep cissn22 efekt;
Data ciselnik;
    set nsnpl ;
        proc means; var cissn22 ;                               title "Popis ciselniku 327";
        proc sort; by cissn22 ;
Data efsnp2;
    set efsnp;
        efekt=round(efekt, 0.00001);
        efekt2=round(efekt,0.01);
Data vystup;                               title "zapis efektu jednotlivych lokusu 333";
merge efsnp2 ciselnik; by cissn22 ;
file efsnp;
    put snjm $1-40 cissn22 42-47 efekt 49-58 .6 ;
    Proc means;
Data UZPRED;
    set pr;
        ncj=col1; PGH=col2; dyd=col3; chyba=col4; absch=col5;
        keep ncj PGH dyd chyba absch;
        proc means;                               title "Vystup predpovedi 342";
Data UZPRED2;                               title "byci pred odstraneni chybnych 343";
merge id uzpred ; by ncj ;
file phbyc ;
put nc 1-16 PGH 18-22 dyd 24-28 phm 30-34 chyba 36-40 absch 42-46
    spolm 48-55 6;
proc means data =uzpred2 noprint; var absch; output out=prum mean=;
proc means data =uzpred2 noprint; var absch; output out=odch std=;
proc means data =uzpred2; var pgh dyd phm chyba absch spolm ;
Data prumer;
    set prum;
        mean=absch; keep mean;
Data odchylka;
    set odch;
        smrodch=absch; keep smrodch;
Data slouc;                               title "maximalni chyba byka pro ponechani v souboru 356";
merge prumer odchylka;
krit=mean+2*smrodch; /*= 539,krit = .pri velkem nastaveni nic nevyradi..*/
    keep krit;
    proc means ;
proc iml;
    use slouc;
        read all into A;
    close slouc ;
    vek=j(5000,1,.);
do i=1 to 5000;
    vek[i,1]=a;
end;
create kr from vek;
append from vek;
Data slouc2;                               title "byci po odstraneni chybnych 371";
merge uzpred2 kr; if chyba;
    kr=col1;
Data slouc2;
    set slouc2;
        if absch>kr then delete;
        b = 1 ;
    keep nc pgh dyd phm chyba absch spolm kr b;
    proc means ;

```

```

proc sort; by nc;
Data slouc3;
set slouc2;
if dyd="." then delete;
if chyba="." then delete;
title "Popisna statistika vseh byku po odstraneni chybnych 385";
proc corr; var pgh dyd phm chyba absch spolm ;
proc sort data=uzpred2 ; by nc ;
proc sort data=slouc2 ; by nc ;
Data CHyb; title "byci s velkou chybou 389";
merge uzpred2 slouc2; by nc ;
Data CHyb;
set CHyb;
if b = 1 then delete;
a = 1 ;
keep nc a;
file bycichyb;
put nc 1-16 ;
proc means ;
data oprav ; title "YD bez chybnych byku 399";
merge chyb DYD1; by nc ;
data oprav ;
set oprav ;
if a = 1 then delete ;
file dyd;
put nc 1-16 spolm 18-25 6 phm 27-31 euzm 33-40 2 ydm 42-46 ;
proc means ;
data oprav ;
merge chyb gbyci ; by nc ;
data oprav ;
set oprav ;
if a = 1 then delete ;
data gbyci ;
set oprav ; by nc ; if first.nc ;
ncj2 = _n_ ;
file gbyc;
put nc 1-16 ncj2 18-24 ;
proc means ; title "novy seznam G byku 417";
proc sort ; by cisb22 ;
data oprav ; title "novy soubor snp 419";
merge gbyci(in=byc) snpa ; by cisb22 ; if byc ;
file snp dlm="," ;
put ncj2 cissn22 sn1 sn2 sn skz ;
proc sort data=slouc3 ; by nc ;
Data T1;
merge slouc3 gbyci; by nc;
keep nc pgh cisb22; if pgh="." then delete;
file byci_T1;keep nc;
put nc 1-16;
Data T2; title "Byci T2 429";
merge T1 Gbyci ; by nc;
Data T2;
set T2;
if pgh ^= "." then delete; /*odmazani byku s dyd 433*/;
file byci_T2;
put nc 1-16; /*vystup byku, kteri nejsou v t1 435*/;
proc means data=t2; var cisb22;
proc datasets nolist;
delete alllok allsnp allsnp2 allsnp3 celk ciselnik dyd1 efsnp efsnp2 gbyci
chyb id njed nsnp nsnp1 oded oded2 odch odchylka oprav pomdyd pomdyd2 pomsnp
pomsnp2 prum prumer slouc slouc2 slouc3 snp1 snpa t1 t2 tmat uzpred uzpred2
vys vystup yd ;

```

```

quit; run;
/*.....konec.....*/

/*.....prv1G.sas.....*/
/*
oprava pribuznosti - PRVYKROK, vsichni G byci,
.....VUZV 10.5.2016 */
options nodate nonumber ls=89 ps=10000;
proc printto print= "/home/pribyl/tdmmace/vysprv1.lst" ;
proc printto log = "/home/pribyl/tdmmace/logprv1.lst" ;
/*dm output 'clear';
dm log 'clear'; */ /* cteni */
filename gbyc "/home/pribyl/tdmmace/gbycb"; /*druhy ciselnik byka*/
filename alelsnb "/home/pribyl/tdmmace/alelsnb.txt";/*vstup alely SNP*/
filename cissn2 "/home/pribyl/tdmmace/cissn";/* ciselnik lokusu*/
/*psani*/
filename marker5 '/home/pribyl/tdmmace/marker.dat'; /* pro DMU G*/
filename map5 '/home/pribyl/tdmmace/map.dat'; /* pro DMU G*/
/*.....genotypovani jedinci .....*/
data genotb; title "G byci 16";
infile gbyc ;
input nc 1-16 ncj2 18-24;
if nc = . then delete ; if ncj2 = . then delete ;
proc means ;
proc means noprint; var ncj2 ; title "pocet g byku 21";
output out=pru mean=;
data sbc; set pru; pocegb=_freq_;
keep pocegb ;
data pocegb ;
set sbc ;
data alely ; title "ciselnik alel 27";
attrib snjm format=$40. length=$40;
infile cissn2 ;
input snjm $1-40 cissn2 42-47;
proc means ;
proc means noprint; var cissn2 ; title "pocet alel 32";
output out=pru mean=;
data sbc; set pru; poceal=_freq_;
keep poceal ;
data poceal ;
set sbc ;
/*.....SN.....*/
data sn ; title " prectene alely 39";
infile alelsnb dlm=";" missover;
input ncj2 cissn2 sn1 sn2 sn sk ;
keep ncj2 cissn2 sn1 sn2 ;
proc means ;
proc sort data=genotb; by nc ;
/*.....priprava pro G .....*/
proc iml ;
use pocegb ;
read all into pocgb ;
pocegb = max(pocgb) ;
close pocegb ;
use poceal ;
read all into pocal ;
poceal = max(pocal) ;
close poceal ;
x = j(poceal,4,0) ;
zt = j(1,poceal+1,"0000") ;
create marker from zt;
use sn ;
do i = 1 to pocegb by 1; /* po bycich */

```

```

read all var{ncj2 cissn22 sn1 sn2} into x where (ncj2=i);
      poca = nrow(x) ;
s1=j(poceal,2,0);
      do k = 1 to poca by 1 ;
          cislo = x[k,2] ;
          if x[k,3] ^= 0 then s1[cislo,1] = x[k,3];
          if x[k,4] ^= 0 then s1[cislo,2] = x[k,4];
      end;
      slch = char(s1) ;
      byk = x[1,1] ;
      bykch=char(byk,5,0); /* spojeni alel do jedne promenne*/
s2 = compress(slch[,1])+ " "+compress(slch[,2]) ;
      zt=compress(bykch)||s2`;
      append from zt ;
      end ;
close sn ;
/*.....zapis pro G .....*/
PROC SQL noprint; /* vypoctena promenna vlozena do globalni promenne */
select poceal
into :pocetal
from work.poceal
;quit;
%let pocet=%eval(&pocetal+1);
data _null_ ;
  set marker ;
  file marker5 RECFM=V LRECL=330000 ;
  put coll-col&pocet;
data map;
  file map5;
  poc = &pocet - 1 ;
  do i = 1 to poc;
    put i i ;
  end ;
proc means ;
  proc datasets nolist;
delete alely genotb map marker poceal pocegb sbc sn ;
quit; run;
/*.....konec .....*/

/*.....druh1.sas.....*/
/*
      uprava G , A22 po castech dosadit:
      zapis vypisu          radek 10
      nastaveni cteni Gmatrix radek 34
      zapis G,A             radek 80,155,472,511
nastavit rozdily pribuznosti radek 508 + prejmenovat v adresari na 1,2,3
/*..... VUZV 18.5.2016 */
options nodate nonumber ls=89 ps=10000;
/*.....celkove promenne .....*/
proc printto print= "/home/pribyl/tdmmace/vysdrul.lst" ;
proc printto log = "/home/pribyl/tdmmace/logdrul.lst" ;
/*dm output 'clear';
dm log 'clear'; */
filename gmatr5 '/home/pribyl/tdmmace/gmatrix.dat';/*matice G Per*/
filename d5 '/home/pribyl/tdmmace/gbycb';
filename rod65 '/home/pribyl/tdmmace/puvo';
filename hrubci '/home/pribyl/tdmmace/hrubci'; /*hruby ciselnik jedince */
/* trojuhelnik G zapis */
filename ga225 '/home/pribyl/tdmmace/ga225';
filename podiv '/home/pribyl/tdmmace/podiv';
filename pocroz '/home/pribyl/tdmmace/pocrozag';/*prejmenovavat 1,2,3*/
filename rozag '/home/pribyl/tdmmace/rozag';
filename gcis5 '/home/pribyl/tdmmace/gcis5';

```

```

/*.....*/
data hrubci ;                                title "hrubcisel 25";
    infile hrubci dlm=";" ;
    input jedin $1-16 nc 18-26 ; proc means ;
data d;                                        title "G byci 28";
    infile d5 ;
    input nc ncj ;
    if nc = . then delete ; if ncj = . then delete ;
    proc means ;
data zapg ;                                    title "cteni gmatrix 33 " ;
    infile gmatr5 missover;
    input ncj1 ncj2 g ;
    g = round(g, 0.0001);
    proc means ;
data dig ;                                    title "diagonala G 38" ;
    set zapg ;
    if ncj1 ne ncj2 then delete ;
    proc means ;
data mimodig ;                                title "mimodiagonalu G 42" ;
    set zapg ;
    if ncj1 = ncj2 then delete ;
    proc means ;
/*.....provereni pribuznosti .....*/
data prover ;                                title "velke pribuznosti g >> 0.5 47";
    set mimodig;
    if g > -0.5 and g < 0.5 then delete ;
    ncj = ncj1 ; drop ncj1 ;
    proc sort ; by ncj ;
    proc sort data=d ; by ncj ;
data pro2 ;
    merge d prover(in=pribu); by ncj ; if pribu ;
    jed1 = nc ;
    ncj = ncj2 ;
    keep jed1 ncj g ;
    proc sort ; by ncj ;
data pro3 ;
    merge d pro2(in=pribu2); by ncj ; if pribu2 ;
    jed2 = nc ;
    keep jed1 jed2 g ;
data pod ;
    set pro3 ;
    nc = jed1;
    proc sort ; by nc ;
data pod ;
    merge pod(in=ge) hrubci ; by nc ; if ge ;
    keep jedin1 nc jed1 jed2 g ;
    jedin1 = jedin ;
    nc = jed2 ;
    proc sort ; by nc ;
data pod ;
    merge pod(in=ge) hrubci ; by nc ; if ge ;
    keep jedin1 jed1 jedin2 jed2 g ;
    jedin2 = jedin ;
    proc sort ; by g ;
data pro4 ;
    set pod ;
    file podiv ;                                /*velke pribuznosti*/
    put jedin1 $1-16 jed1 18-24 jedin2 $26-41 jed2 43-49 g 51-58 6;
    proc means ;
proc datasets nolist;
    delete dig mimodig pro2 pro3 pro4 prover ;
quit; run;

```

```

/*.....A22.....*/
data grouzg ;
  set d ;
      keep nc ;
      proc sort ; by nc ;
data rod6 ;
  infile rod65 missover ;
input nc 1-16 nco 18-33 ncm 35-50 rokkj 77-80 ;
  if nc = . then delete ;
  proc sort ; by nc ;
data ddd ;                                title " kod rok jedinec 96";
  set rod6 ;
  format kodj $20. ;
  kodj = compress(rokkj||nc) ;
  keep nc kodj ;
  proc means ;
  proc sort ; by nc ;
data rod6 ;
  SET ROD6 ;
  KEEP nc nco ncm ;
  proc sort ; by nc ;
/*.....rod na genotypovane.....*/
%macro gpredci;
%do i = 1 %to 7 %by 1;
data grodg ;                                title " g + &i generaci rodicu 110";
merge grouzg(in= uz) rod6; by nc; if uz;
  data grodg ;
    set grodg ;
    keep nc nco ncm ;
  data ge; set grodg ;                                /* do sloupce jedinec + rodice 115*/
    keep nc ;
    if nc = . then delete ;
  data gf; set grodg ;
    keep nc ; nc = nco ;
    if nc = . then delete ;
  data geg; set grodg ;
    keep nc ; nc = ncm ;
    if nc = . then delete ;
  data gh ; set ge gf geg ;
    proc sort ; by nc ;
data grouzg;
  set gh ; by nc ; if first.nc ;
proc means;
%end;
%mend gpredci;
%gpredci ;
/*.....uzavreni rodokmenu pro A22.....*/
data grodgg;                                title " uzavreny rodokmen 133";
  merge grouzg(in=slopec) grodg ; by nc ; if slopec ;
data grodgg;
  set grodgg;
  keep nc nco ncm ;
  proc means ;
  proc sort ; by nc ;
/*.....druhe precislovani na A*22.....*/
data grodl;                                title " druhy ciselnik jedince na A22 141" ;
  merge grodgg(in=roda) ddd; by nc ; if roda ;
  keep nc kodj ;
  proc sort ; by kodj ;
data grodl ;
  set grodl ; by kodj ; if first.kodj ;
data gcisel ;

```

```

set grod1 ;
gncj = _n_ ;
keep nc gncj ;
proc means ;
proc sort ; by nc ;
data sehzap ; title "ciselnik genotypovanych 153" ;
merge gcisel d(in=geno) ; by nc ; if geno ;
file gcis5 ;
put nc 1-16 gncj 18-24 ;
proc means ;
data grod2; title "rodokmen A*22 precislovany jedinec 158";
merge grodgg(in = jj) gcisel ; by nc ; if jj ;
nj = gncj ;
nc = nco ;
keep nj nc ncm ;
proc means ;
proc sort ; by nc ;
data grod4 ; title "rodokmen precislovany otec 165";
merge grod2(in = oo) gcisel ; by nc ; if oo ;
data grod4 ; set grod4 ;
no = gncj ;
if no = . then no = 0 ;
nc = ncm ;
keep nj no nc ;
proc means ;
proc sort ; by nc ;
data grod6 ; title "rodokmen A*22 precislovana matka 174";
merge grod4(in=mm) gcisel ; by nc ; if mm ;
data grod6 ; set grod6 ;
nm = gncj ;
if nm = . then nm = 0 ;
keep nj no nm ;
proc means ;
proc sort data=grod6 ; by nj;
proc datasets nolist;
delete ge geg gf gh grodgg grodgg grod1 grod2 grod4 rod6 ;
quit ;run ;
/*.....matice pribuznosti A*22.....*/
proc iml ;
use grod6;
read all into mb;
close grod6;
n = nrow(mb); /* pocet jedincu pribuznost vctne predku 190*/
*Vytvorenie spodnej triangularnej matice L;
L=i(n);
do i=1 to n;*Zacni diagonalnym prvkom jedinca 1;
o = mB[i,2]; m = mB[i,3];
if o & m = 0 then L[i,i] = 1;
if o & m > 0 then do;
x = L[o,1:o]; x = x#x;
a = (sum(x))*0.25;
y = L[m,1:m]; y = y#y;
c = (sum(y))*0.25;
L[i,i] = sqrt((1 - a - c));
end;
else if o > 0 then do;
x = L[o,1:o]; x = x#x;
a = (sum(x))*0.25;
L[i,i] = sqrt((1-a));
end;
else if m > 0 then do;
y = L[m,1:m]; y = y#y;

```

```

c = (sum(y))*0.25;
L[i,i] = sqrt((1-c));
end;
/*Pokracuj v danom stlpci s dalsim jedincom
a pocitaj mimodiagonalny prvok danho jedinca L[j,i];*/
do j=i+1 to n;
o = mB[j,2]; m = mB[j,3];
if o & m = 0 then L[j,i] = 0;
if o & m > 0 then L[j,i] = 0.5*(L[o,i] + L[m,i]);
else if o > 0 then L[j,i] = 0.5*(L[o,i]);
else if m > 0 then L[j,i] = 0.5*(L[m,i]);
end;
/*... print i ; .....kontrolní tisk */
end; title "konec tvorby L 223";
print n ;
create nn from n ; /*zapis pocet prvku do souboru*/
append from n;
m = int(n/3) ; /* pocy prvku matice A na tretiny m */
if m*3 < n then do ;
a11 = j(1,n,0) ;
mb = a11||0 ;
n = n+1 ;
m = int(n/3) ;
if m*3 < n then do ;
a11 = a11//a11 ;
mb = j(2,2,0) ;
mb = a11||mb ;
n = n+1 ;
m = int(n/3) ;
end ;
L = L//a11 ; /* doplneni nulovych radku a sloupcu */
mb = mb` ;
L = L||mb;
end ; /* m ; velikost opravene tretiny */
create mm from m ; /*zapis pocet prvku do souboru*/
append from m;
*vypocet matice A -matice pribuznosti pre A22 sposob podla Quass (1976);
/*..mG = L` ;...Aa22 = L*mG; ..nize rozepsane..*/ title "matice A*22 247";
/* .....násobeni bloku matic po tretinach - A11*/
A11 = L[1:m,]*L[1:m,]` ;
t = (m*m-m)/2 + m ; /*trojuhelnik*/
mB= j(t, 3, 0) ; /*prevod A11 na slopce*/
k = 0 ;
do i = 1 to m ;
do j = i to m ;
a = j-i+1+k ;
mB[a,1] = i ;
mB[a,2] = j ;
mB[a,3] = A11[i,j] ;
end;
k = k+m+1-i;
/* print i k ; .....kontrolní tisk */
end ;
create maprall from mB ; /*A11 matici do souboru*/
append from mB; title "konec tvorby bloku A11 264";
print m t k ;
/* .....násobeni bloku matic po tretinach - A12*/
z = m + 1 ; u = m+m ;
A11 = L[1:m,]*L[z:u,]` ; /*ctverec*/
mB= j(m*m, 3, 0) ; /*prevod A12 na slopce*/
k = 0 ;
do i = 1 to m ;

```

```

do j = 1 to m ;
    mB[j+k,1] = i ;
    mB[j+k,2] = j ;
    mB[j+k,3] = A11[i,j] ;
end;
k = k+m ;
/* print i k ;                                     ....kontrolní tisk */
end ;
create mapra12 from mB ;                             /*A12 matici do souboru*/
append from mB;                                     title "konec tvorby bloku A12
282";
print m z u k ;
/* .....násobení bloku matic po tretinách - A13*/
z = m + m + 1 ; u = m + m + m ;
A11 = L[1:m,]*L[z:u,]` ;                             /*ctverec*/
mB= j(m*m, 3, 0) ;                                     /*prevod A13 na slopce*/
k = 0 ;
do i = 1 to m ;
    do j = 1 to m ;
        mB[j+k,1] = i ;
        mB[j+k,2] = j ;
        mB[j+k,3] = A11[i,j] ;
    end;
    k = k+m ;
    /* print i k ;                                     ....kontrolní tisk */
end ;
create mapra13 from mB ;                             /*A13 matici do souboru*/
append from mB;                                     title "konec tvorby bloku A13 300";
print m z u k ;
/* .....násobení bloku matic po tretinách - A22*/
z = m + 1 ; u = m+m ;
A11 = L[z:u,]*L[z:u,]` ;                             /*trojuhelnik*/
mB= j(t, 3, 0) ;                                     /*prevod A22 na slopce*/
k = 0 ;
do i = 1 to m ;
    do j = i to m ;
        a = j-i+1+k ;
        mB[a,1] = i ;
        mB[a,2] = j ;
        mB[a,3] = A11[i,j] ;
    end;
    k = k+m+1-i;
    /* print i k ;                                     ....kontrolní tisk */
end ;
create mapra22 from mB ;                             /*A22 matici do souboru*/
append from mB;                                     title "konec tvorby bloku A22 319";
print m t z u k ;
/* .....násobení bloku matic po tretinách -
A23*/
z = m + 1 ; v = m + m ; w = m + m + 1 ; u = m + m + m ;
A11 = L[z:v,]*L[w:u,]` ;                             /*ctverec*/
mB= j(m*m, 3, 0) ;                                     /*prevod A23 na slopce*/
k = 0 ;
do i = 1 to m ;
    do j = 1 to m ;
        mB[j+k,1] = i ;
        mB[j+k,2] = j ;
        mB[j+k,3] = A11[i,j] ;
    end;
    k = k+m ;
    /* print i k ;                                     ....kontrolní tisk */
end ;create mapra23 from mB ;                         /*A23 matici do souboru*/

```

```

        append from mB;                                title "konec tvorby bloku A23 337";
print m z v w u k ;
        /* .....násobeni bloku matic po tretinach - A33*/
z = m + m + 1 ; v = m + m + m ;
A11 = L[z:v,]*L[z:v,]` ;                                /*trojuhelnik*/
mB= j(t, 3, 0) ;                                        /*prevod A33 na slopce*/
        k = 0 ;
        do i = 1 to m ;
            do j = i to m ;
                a = j-i+1+k ;
                mB[a,1] = i ;
                mB[a,2] = j ;
                mB[a,3] = A11[i,j] ;
            end;
            k = k+m+1-i;
            /* print i k ;                                ....kontrolní tisk */
        end ;
create mapra33 from mB ;                                /*A33 matici do souboru*/
        append from mB;                                title "konec tvorby bloku A33 356";
print m t z v k ;
L = i(1) ; A11 = L ; mB = L ;                            /* vynulovani */
/*.....zapis A22 sloupec do souboru .....*/
data nn ; set nn;                                       /* pocet radku v matici A*/
        n = coll ;
        f = 1 ;
data mm ; set mm;                                       /* velikost tretiny matice A*/
        m = coll ;
        f = 1 ;
data mggg11; set mapra11;                                title "trojuhelnik A11 366";
        rad = coll ; slou = col2 ; aa = col3 ;
        keep rad slou aa ;
        aa= round(aa, 0.0001);
        if slou < rad then delete ;
        proc means ;
data mggg12; set mapra12;                                title "ctverec A12 372";
        rad = coll ; slou = col2 ; aa = col3 ;
        f = 1 ;
data mggg12;
        merge mggg12 mm; by f ;
                slou = slou + m ;
        keep rad slou aa ;
                aa= round(aa, 0.0001);
        proc means ;
data mggg13; set mapra13;                                title "ctverecA13 381";
        rad = coll ; slou = col2 ; aa = col3 ;
        f = 1 ;
data mggg13;
        merge mggg13 mm; by f ;
                slou = slou + m + m ;
        keep rad slou aa ;
                aa= round(aa, 0.0001);
        proc means ;
data mggg22; set mapra22;                                title "trojuhelnik A22 390";
        rad = coll ; slou = col2 ; aa = col3 ;
        f = 1 ;
data mggg22;
        merge mggg22 mm; by f ;
                rad = rad + m ; slou = slou + m ;
        keep rad slou aa ;
                aa= round(aa, 0.0001);
        if slou < rad then delete ;
        proc means ;

```

```

data mggg23; set mapra23; title "ctverec A23 400";
    rad = col1 ; slou = col2 ; aa = col3 ;
    f = 1 ;
data mggg23;
    merge mggg23 mm; by f ;
    rad = rad + m ; slou = slou + m + m ;
    keep rad slou aa ;
    aa= round(aa, 0.0001);
proc means ;
data mggg33; set mapra33; title "trojuhelnik A33 409";
    rad = col1 ; slou = col2 ; aa = col3 ;
    f = 1 ;
data mggg33;
    merge mggg33 mm; by f ;
    rad = rad + m + m ; slou = slou + m + m ;
    keep rad slou aa ;
    aa= round(aa, 0.0001);
    if slou < rad then delete ;
proc means ;
data mggg; title "trojuhelnik Avse 419";
    set mggg11 mggg12 mggg13 mggg22 mggg23 mggg33;
    f = 1 ;
proc datasets nolist;
delete mggg11 mapra11 mggg12 mapra12 mggg13 mapra13
    mggg22 mapra22 mggg23 mapra23 mggg33 mapra33 ;
quit;
data mggg;
    merge mggg nn; by f ;
    if rad > n then delete ;
    if slou > n then delete ;
    keep gncj rad slou aa ;
    gncj = rad ;
    aa= round(aa, 0.0001);
    if slou < rad then delete ;
proc means ;
proc sort; by gncj ;
proc sort data=gcisel ; title "ciselnik poradi uvnitr A22 436";
    by nc ;
data genoa ;
    set d ;
    nj = ncj ;
    keep nc ncj nj ;
proc sort data = genoa ;
    by nc ;
data ga22 ; /* spojene ciselniky pro genotypovane 444*/
    merge gcisel genoa(IN= GENO) ; by nc ; if geno ;
    keep nj gncj ;
proc sort ; by gncj ;
proc means ;
data garad ; /* precislovani prveho jedince 449*/
    merge mggg ga22(in=gar) ; by gncj ; if gar ;
    ncj1 = nj ;
    gncj = slou ;
    keep ncj1 gncj slou aa ;
proc sort ; by gncj ;
data gaslou ; title "precislovane A22 455";
    merge garad ga22(in=gas) ; by gncj ; if gas ;
    ncj2 = nj ;
    keep ncj1 ncj2 aa ;
data gaslou ;
    set gaslou ;
    if ncj1 > ncj2 then do ;

```

```

                pomo = ncj1 ;
                ncj1 = ncj2 ;
                ncj2 = pomo ;
            end ;
            keep ncj1 ncj2 aa ;
        proc sort ; by ncj1 ncj2 ;
        proc means ;
data gzapg ;                                title " zapis trojuhelnik A22 na disk 469";
        set gaslou ;
                a22 = round(aa, 0.0001);
        file ga225 ;
                put ncj1 1-7 ncj2 9-15 a22 17-24 6 ;
                keep ncj1 ncj2 a22 ;

        proc means ;
/*.....porovnani G A22 .....*/
data acis1 ;                                title "nacteni G 477" ;
        set zapg ;
                keep ncj ncj1 ncj2 g ;
                ncj = ncj2 ;
        proc sort ; by ncj ;
data acis2 ;
        merge acis1(in=matra) d ; by ncj ; if matra ;
                jedin2 = nc ;
                ncj = ncj1 ;
                keep ncj ncj1 ncj2 jedin2 g ;
        proc sort data=acis2 ; by ncj1 ncj2 ;
data ncj1 ;
        merge acis2(in=mata) d ; by ncj ; if mata ;
                jedin1 = nc ;
                keep ncj1 jedin1 ncj2 jedin2 g ;
data ggaa ;
        merge ncj1 gzapg ; by ncj1 ncj2 ;
                roz = a22 - g ;
                absroz = abs(roz) ;
/*.....G4.....*/
data gdiag ;                                title "slog A - G 497" ;
        set ggaa ;
                if ncj1 ne ncj2 then delete ;
        proc corr ; var a22 g roz absroz;
data gndiag ;                                title "mimodiag A - G 501" ;
        set ggaa ;
                if ncj1 = ncj2 then delete ;
        proc corr ; var a22 g roz absroz;
/*.....*/
data velkroz ;                                title "velke rozdily A G 506" ;
        set ggaa ;
        if absroz < 0.3 then delete ;          /*.....NASTAVIT.....*/
                proc sort ; by descending absroz ;
        data velkroz ; set velkroz ; /*if jedin1 = jedin2 then delete */
                file rozag ;
        put jedin1 1-16 jedin2 18-33 a22 35-42 6 g 44-51 6 roz 53-61 6 absroz 63-70
6 ;
                proc means ;
data velrol ;
        set velkroz ;
                keep nc roz absroz ;
                nc = jedin1 ;
data velro2 ;
        set velkroz ;
                keep nc roz absroz ;
                nc = jedin2 ;
data veroz ;

```

```

set velro1 velro2 ;
proc sort ; by nc ;
proc means noprint; by nc ; output out=poc mean=;
/*..cetnosti a prumery rozdilu podle byka..*/
data b; set poc; n=_freq_; keep n nc roz absroz ;
proc sort ; by descending n descending absroz ;
data b ; set b ;
file pocroz ;
put nc 1-16 n 18-23 roz 25-33 6 absroz 35-42 6 ;
proc means data=b; title "pocety rozdilu A G na byka 532" ;
proc datasets nolist;
delete acis1 acis2 b d ddd dig garad gaslou ga22 gcisel gdiag genoa ggaa ge
geg gf gh gndiag grodg grod6 grodg grouzg gzapg maprall mggg mggg2 mimodig
nn ncj1 poc pro2 pro3 pro4 prover rod6 sehzap velkroz velro1 velro2 veroz
zapg ;
quit;run;
/*.....konec.....*/

/*.....prv2G.sas.....*/
/* oprava pribuznosti - druhy KROK, vsichni G byci,
dosadit: cteni soubory rozdily pribuznosti+ pocety radek 28-50
/*...zapis novych gbyc, snp, dyd.....VUZV 18.5.2016 */
options nodate nonumber ls=89 ps=1000;
proc printto print= "/home/pribyl/tdmmace/vysprv2.lst" ;
proc printto log = "/home/pribyl/tdmmace/logprv2.lst" ;
/*dm output 'clear';
dm log 'clear'; */ /* cteni */
filename gbyc "/home/pribyl/tdmmace/gbycb"; /*druhy ciselnik byka*/
filename alelsnb "/home/pribyl/tdmmace/alelsnb.txt";/*vstup alely SNP*/
filename cissn2 "/home/pribyl/tdmmace/cisnh";/*druhy ciselnik lokusu*/
filename DYDb '/home/pribyl/tdmmace/ydb'; /*vstup dyd edc*/

filename pocroz1 '/home/pribyl/tdmmace/pocrozag'; /* velka odch pribuz*/
filename pocroz2 '/home/pribyl/tdmmace/pocrozag2'; /* velka odch pribuz*/
filename pocroz3 '/home/pribyl/tdmmace/pocrozag3'; /* velka odch pribuz*/

filename gbyc2 "/home/pribyl/tdmmace/gbyc"; /*treti ciselnik G byku*/
filename alelsn "/home/pribyl/tdmmace/alelsn.txt";/*vystup alely SNP*/
filename DYD '/home/pribyl/tdmmace/yd';/*vystup dyd bez chybných byku*/
/*.....genotypování jedinci .....*/
data gbyc ; title "G byci 23";
infile gbyc ;
input nc 1-16 ncj2 18-24;
if nc = . then delete ; if ncj2 = . then delete ;
proc means ;
data rozdil1 ; /* velky rozdil pribuznosti 28 */
infile pocroz1 missover;
input nc 1-16 n 18-23 roz 25-33 absroz 35-42 ;
if n > 3 & absroz > 0.5 then a = 1; if absroz > 0.6 then a = 1; if a = .
then delete ;
keep nc a ;

```

6. Slučování domácích a MACE hodnot

```

/*.....tdydG822.sas.....*/
/* priprava spolecne TD 3 laktace + YD pro GEBV -
navazuje na soubory uzitTD a IntrbYD
dosadit: druh souboru radek 18
dedivost indexu radek 181,264,265
efektivní počet radek 273-340

```

```

    spolehlivosti otcu uzit      radek 183
    směrodat odchylky           radek 348,355,362,437-9
    genet vstupy index spol Inter radek 255
doplneni vah                    radek 328
    opakovani uzit z DYD        radek 248,273-340,344
zapis ciselniku                 radek 208
    zapis rodokmenu            radek 236
zapis uzitkovosti              radek 455
    zapisy souboru G           radek 480,502,504
/*.....VUZV 8.8.2016 */
options nodate nonumber ls=89 ps=10000;
%let soubo = 5 /* 10*/ ; /* druh souboru */
    proc printto print= "/home/pribyl/tdmmace/tdyd22/vysuzdy2.lst" ;
    proc printto log = "/home/pribyl/tdmmace/tdyd22/loguzdy2.lst" ;
/*dm output 'clear';
dm log 'clear'; */
filename slog '/home/pribyl/tdmmace/slo45'; /* G mat */
filename gcis5 '/home/pribyl/tdmmace/gcis52'; /*genotypovanych 5*/
filename spoleh "/home/pribyl/tdmmace/tdyd22/spoleh" ;/*tabulka spoleh podle
w*/
/* cteni YD*/
filename cisyd '/home/pribyl/tdmmace/yd9/cisyd'; /*ciselnik ../
filename maprYD '/home/pribyl/tdmmace/yd9/maprd5'; /*rodokmen DMU do../
filename uziYD '/home/pribyl/tdmmace/yd9/uzdo5'; /*uzit do roku../
filename sezgYD '/home/pribyl/tdmmace/yd9/sezgen5'; /*genotypovanych*/
/* cteni TD */
filename cisTD '/home/pribyl/tdmmace/uztd9/cis5'; /*ciselnik */
filename uziTD '/home/pribyl/tdmmace/uztd9/uzdo5'; /*uzit do roku../
filename maprTD '/home/pribyl/tdmmace/uztd9/maprd5'; /*rodokmen DMU do../
filename sezTD '/home/pribyl/tdmmace/uztd9/sezby5'; /*otci krav do*/
filename sezgTD '/home/pribyl/tdmmace/uztd9/sezgen5'; /*genotypovanych */
/* zapis */
filename cis5 '/home/pribyl/tdmmace/tdyd22/cistddy'; /*ciselnik ../
filename maprd5 '/home/pribyl/tdmmace/tdyd22/maprd5'; /*rodokmen DMU do../
filename prvlDY '/home/pribyl/tdmmace/tdyd22/prvlDY'; /* kazda uzit 30x*/
filename drulDY '/home/pribyl/tdmmace/tdyd22/drulDY';
filename trelDY '/home/pribyl/tdmmace/tdyd22/trelDY';
filename uzit5 '/home/pribyl/tdmmace/tdyd22/uzdo5'; /*uzit do roku../
filename slog45d '/home/pribyl/tdmmace/tdyd22/slog45';
filename aamat '/home/pribyl/tdmmace/tdyd22/a22';
filename sezgen5d '/home/pribyl/tdmmace/tdyd22/sezgen5'; /*genotypovanych*/
/*.....ciselnik YD.....*/
data cisdy ;
    title " ciselnik YD 48";
    infile cisyd ;
    input jedin 1-16 kodj $18-37 ncj 39-45 ;
    keep jedin kodj ncj ;
    if jedin = . then delete ; if ncj = . then delete ;
    proc means ;
    proc sort ; by ncj ;
/*.....rodokmen YD.....*/
data rodDY ; title " rodokmen YD 56";
    infile maprYD recfm = v missover ;
    input ncj 1-7 no 9-15 nm 17-23 nj 25-31;
if ncj = . then delete ; if no = 0 then no = . ; if nm = 0 then nm = . ;
    proc means;
    proc sort ; by ncj ;
data rod1 ;
    merge cisdy rodDY(in=rod) ; by ncj ; if rod ;
    keep jed ncj nm ;
    jed = kodj ;
    ncj = no ;

```

```

proc sort ; by ncj ;
data rod2 ;
merge cisdy rod1(in=rodo) ; by ncj ; if rodo ;
keep jed o ncj ;
o = kodj ;
ncj = nm ;
proc sort ; by ncj ;
data rod3 ; title "precislovaný rodokmen DY 74";
merge cisdy rod2(in=rodm) ; by ncj ; if rodm ;
keep jed o m ;
m = kodj ;
proc means ;
proc sort ; by jed ;
/*.....genotypovani jedinci YD.....*/
data genotDY ; title "genotypovani byci YD 81";
infile sezgYD ;
input ncj 1-7 ; if ncj = . then delete ;
proc means ;
proc sort ; by ncj ;
data genbDY ;
merge cisdy genotDY(in= geno) ; by ncj ; if geno ;
keep jed ;
jed = kodj ;
proc means ;
/*.....uzitkovost YD.....*/
data dy ; title "uzitkovost YD 92";
infile uziYD ;
input ntd ncj ydm euzm ydt euzt ydb euzb spolm spolt spolb;
if ydm = -9999.999 then ydm = .;if euzm = -9999.999 then euzm = .;
if ydt = -9999.999 then ydt = .;
if euzt = -9999.999 then euzt = .;if ydb = -9999.999 then ydb = .;
if euzb = -9999.999 then euzb = .;
if spolm= -9999.999 then spolm= .;if spolt= -9999.999 then spolt = .;
if spolb= -9999.999 then spolb= .;
if ncj = . then delete ;
proc means ;
proc sort ; by ncj ;
data dyd ;
merge cisdy DY(in= uzi) ; by ncj ; if uzi ;
keep ntd jed ydm euzm ydt euzt ydb euzb spolm spolt spolb;
jed = kodj ;
proc means ;
/*.....ciselnik TD.....*/
data cistd ; title "
ciselnik TD 107";
/*attrib jedin format=$16. length=$16;*/
infile cistd ;
input jedin 1-16 kodj $18-37 ncj 39-45 ;
keep jedin kodj ncj ;
if jedin = . then delete ; if ncj = . then delete ;
proc means ;
proc sort ; by ncj ;
/*.....rodokmen TD.....*/
data rodTD ; title " rodokmen TD 116";
infile maprTD recfm = v missover ;
input ncj 1-7 no 9-15 nm 17-23 nj 25-31;
if ncj = . then delete ; if no = 0 then no = . ; if nm = 0 then nm = . ;
proc means;
proc sort ; by ncj ;
data rod1 ;
merge cistd rodtd(in=rod) ; by ncj ; if rod ;
keep jed ncj nm ;

```

```

        jed = kodj ;
        ncj = no ;
    proc sort ; by ncj ;
data rod2 ;
    merge cistd rod1(in=rodo) ; by ncj ; if rodo ;
    keep jed o ncj ;
    o = kodj ;
    ncj = nm ;
    proc sort ; by ncj ;
data rotd ; title "precislovaný rodokmen TD 134";
    merge cistd rod2(in=rodm) ; by ncj ; if rodm ;
    keep jed o m ;
    m = kodj ;
    proc means ;
    proc sort ; by jed ;
/*.....genotypovani jedinci TD.....*/
data genottd ; title "genotypovani byci TD 141";
    infile sezgtd ;
    input ncj 1-7 ; if ncj = . then delete ;
    proc means ;
    proc sort ; by ncj ;
data genbTD ;
    merge cisTD genottd(in= geno) ; by ncj ; if geno ;
    keep jed ;
    jed = kodj ;
    proc means ;
/*.....uzitkovost TD.....*/
data td ; title "uzitkovost TD 152";
    infile uzitd ;
input ntd 1-6 ncsr 8-13 lakt 15-16 rokot 18-21 ncj 23-29 /*ctp 31-37*/
    dl 39-41 sv 43-48 4 b1p 50-56 4 b2p 58-64 4 b3p 66-72 4
    ml1 74-82 3 tu1 84-92 3 bi1 94-102 3 sb1 104-112 3
    ml2 114-122 3 tu2 124-132 3 bi2 134-142 3 sb2 144-152 3
    ml3 154-162 3 tu3 164-172 3 bi3 174-182 3 sb3 184-192 3
    vaha 194-203 5;
if dl= -9999.999 then dl= .; if sv= -9999.999 then sv= .;
if b1p= -9999.999 then b1p= .; if b2p= -9999.999 then b2p= .;
if b3p= -9999.999 then b3p= .;
if ml1= -9999.999 then ml1= .; if tu1= -9999.999 then tu1= .;
if bi1= -9999.999 then bi1= .; if sb1= -9999.999 then sb1= .;
if ml2= -9999.999 then ml2= .; if tu2= -9999.999 then tu2= .;
if bi2= -9999.999 then bi2= .; if sb2= -9999.999 then sb2= .;
if ml3= -9999.999 then ml3= .; if tu3= -9999.999 then tu3= .;
if bi3= -9999.999 then bi3= .; if sb3= -9999.999 then sb3= .;
    if ncj = . then delete ;
    proc means ;
    proc sort ; by ncj ;
data uztd ;
    merge cistd td(in= krav) ; by ncj ; if krav ;
keep ntd ncsr lakt rokot jed ctp dl sv b1p b2p b3p
    ml1 tu1 bi1 sb1 ml2 tu2 bi2 sb2 ml3 tu3 bi3 sb3 vaha ;
    jed = kodj ; ctp = kodj ;
    proc means ;
/*..... byci DY + TD.....*/
data byctd ; title "opakujici se byci v souborech DY + TD 178";
    infile sezbttd;
    input ncj 1-7 pocetd 9-13 ; if ncj = . then delete ;
    h2 = 0.39;
    k = (4-h2)/h2 ; r2 = pocetd*0.83*0.2/(pocetd*0.83*0.2 + k) ;
    if r2 > 0.47 then a = 1 ; /* odpovida 10 dcer*/
    if a = . then delete ;
    proc means ;

```

```

proc sort ; by ncj ;
data bycopa ;
merge byctd(in=dcer) cistd ; by ncj ; if dcer ;
keep jed a ;
jed = kodj ;
/*.....spojeni DY + TD .....*/
data rod ; title "spojeny rodokmen DY + TD 192";
set rodtd rodty ;
if jed = "" then delete ;
keep jed o m ;
proc means ;
proc sort ; by jed ;
data ro ;
set rod ; by jed ; if first.jed ;
proc means ;
/* .....precislovani rodokmenu .....*/
data ciselj ; title "spojeny ciselnik jedinec 202";
set ro ;
attrib jedin format=$16. length=$16;
nccjj = _n_ ;
jedin = substr(jed,5,16);
keep jedin jed nccjj ;
file cis&soubo ;
put jedin 1-16 jed $18-37 nccjj 39-45 ;
proc means ;
data rod1; title "rodokmen precislovany jedinec 211";
merge ro(in = jj) ciselj ; by jed ; if jj ;
nj = nccjj ;
jed = o ;
keep nj jed m ;
proc means ;
proc sort ; by jed ;
data rod2 ; title "rodokmen precislovany otec 218";
merge rod1(in = oo) ciselj ; by jed ; if oo ;
no = nccjj ;
jed = m ;
keep nj no jed ;
proc means ;
proc sort ; by jed ;
data rod ; title "rodokmen precislovana matka 225";
merge rod2(in=mm) ciselj ; by jed ; if mm ;
data rod6 ;
set rod ;
nm = nccjj ;
keep nj no nm ;
proc means ;
proc sort ; by nj ;
data roddmu ; title "rodokmen pro DMU 233";
set rod6 ;
if no = . then no = 0 ; if nm = . then nm = 0 ;
file maprd&soubo ;
put nj 1-7 no 9-15 nm 17-23 nj 25-31;
keep nj no nm ;
proc means ;
/* .....precislovani uzitkovosti .....*/
proc sort data=dyd ; by jed ;
proc sort data=bycopa ; by jed ;
data dypo ; title "ponizeny DY o opakujici byky z TD 243";
merge bycopa dyd(in=mac) ; by jed ; if mac ;
proc means ;
data bycdy ;
set dypo ;

```

```

/* if a = 1 then delete ; nebo...ponechani vseh Inter byku....*/
keep ntd jed ydm euzm ydt euzt ydb euzb spolm spolt spolb;
proc means ;
/*.....vahy pro den KU pro odregresovany Interbull.....*/
proc iml; /*reset print ;*/
/*..... cele laktace .....*/
/* genetické korelace laktaci 254*/
rg0 = { 1          0.8924722 0.8143218 ,
        0.8924722 1          0.9038041 ,
        0.8143218 0.9038041 1          };
sglak = { 629.3577 , 809.084 , 871.68 };
/* vahy v indexu pro 3 laktace 259*/
EH = { 0.3333 , 0.3333 , 0.3333 };
/*.....matice pro soustavu.....261*/
g = sglak`#rg0#sglak ; /* genetická kov celych laktaci 262*/
s2g = eh`*g*eh ;
h21 = /*0.07*//*0.15*/0.20/*0.30*//*0.37*/; h22=h21; h23=h21; /* dne KU*/
op1 = 0.7354 ; op2 = 0.7539 ; op3 = 0.7540 ;
/*.....spolehlivosti dilcich hodnot.....266*/
k = 0 ;
spoleh = 0.05 ;
zaps = j(195000,3,0);
spol = j(3,1,0) ;
/*.....násobné užitkovosti celeho jedince.....271*/
do pocop = 1 to 45 by 1;
do w1 = 0.1 to 10 by 0.001 ; /* krok*/
w2 = w1*0.7 ; w3 = w2*0.7 ;
spol[1,1] = w1*h21/(1+(w1-1)*op1) ; /*spol za laktaci*/
spol[2,1] = w2*h22/(1+(w2-1)*op2) ;
spol[3,1] = w3*h23/(1+(w3-1)*op3) ;
cn = spol#g ; /*index za zvire*/
pn = spol`#g#spol ;
pn[1,1] = cn[1,1] ;
pn[2,2] = cn[2,2] ;
pn[3,3] = cn[3,3] ;
bn = inv(pn)*cn*eh ;
r2in = (bn`*cn*eh)/s2g ;
cc = j(pocop,1,1) ; /* index opakovane zvire*/
cc = cc*r2in ;
pc = cc*cc` ;
do j = 1 to pocop ;
pc[j,j] = cc[j,1] ;
end ;
bc = inv(pc)*cc ;
r2 = bc`*cc ;
r2 = round(r2,0.01) ;
if r2 >spoleh then do ;
k = k+1 ;
zaps[k,1]=pocop ;
zaps[k,2]=w1;
zaps[k,3]=r2;
end ;
end ;
spoleh = r2 ;
end ;
create vah from zaps ; /* vahy do souboru 303*/
append from zaps ;
quit;
data vaha ;
set vah ;
if coll=0 then delete; if col2=0 then delete; if col3=0 then delete ;
opak = round(coll,1.) ;

```

```

        w1 = col2 ;
        spolm = round(col3, .01) ;
        keep opak w1 spolm ;
    proc sort ; by spolm w1 ;
data spod ;
    set vaha ; by spolm ; if first.spolm ;
data hor ;
    set vaha ; by spolm ; if last.spolm ;
        wh = w1 ;
        oph = opak ;
        drop w1 opak;
data vahy ;
        merge spod hor ; by spolm ;
        w = (w1+wh)/2 ; w = round(w, .001) ;
        opak = (opak+oph)/2 ; opak = round(opak, 1.) ;
        keep opak w spolm ;
proc means ;
data vahy2 ;
        infile spoleh missover termstr=crlf;
        input w2 opak2 spolm ;
        proc sort ; by spolm ;
proc means ;
data vahyy ;
        merge vahy vahy2 ; by spolm ;
data vahyy ;
        set vahyy ;
        if w = . then do ;
            w = w2 ;
            opak = opak2 ;
        end ;
        keep opak w spolm ;
/*.....odregresovani z Interbullu na den KU .....*/
proc sort data= bycdy ; by spolm ;
data odrdyd ;
        merge bycdy(in=inter) vahyy ; by spolm ; if inter ;
proc means ;
data prvni ;
        set odrdyd ;
        lakt = 1 ; yd1 = ydm ; vaha = w ;
        file prvlDY ;
        do i = 1 to opak by 1 ;
            put jed yd1 vaha lakt ;
        end ;
data druha ;
        set odrdyd ;
        lakt = 2 ; yd2 = ydm ; vaha = w*0.7 ;
        file druldY ;
        do i = 1 to opak by 1 ;
            put jed yd2 vaha lakt ;
        end ;
data treti ;
        set odrdyd ;
        lakt = 3 ; yd3 = ydm ; vaha = w*0.7*0.7 ;
        file trelDY ;
        do i = 1 to opak by 1 ;
            put jed yd3 vaha lakt ;
        end ;
data prvni ;
        attrib jed format=$20. length=$20;
        infile prvlDY ;
        input jed yd1 vaha lakt ;
data druha ;

```

```

        attrib jed format=$20. length=$20;
        infile drulDY ;
            input jed yd2 vaha lakt ;
data treti ;
        attrib jed format=$20. length=$20;
        infile trelDY ;
            input jed yd3 vaha lakt ;
data interb ;
        title "DYD pro dny podle laktaci 379";
        set prvni druha treti ;kk=0.00000001 ;
/*b1p=0; b2p=0;b3p=0;*/b1p=kk*rannor(0);b2p=kk*rannor(0);b3p=kk*rannor(0) ;
        ctp = jed ; ncsr = 999000 ; ntd = 999000 ;
        proc means ;
data uzit ;
        title "oboje uzitkovosti DY +TD 384";
        set interb uztd ;
        proc means ;
            proc sort ; by ntd ;
data pordsro ;
        title "precislovani spolecne TD 388";
        set uzit; by ntd ; if first.ntd ;
data ciselsro ;
        set pordsro ;
            ntd2 = _n_ ;
        keep ntd ntd2 ;
        proc means ;
data uzit2 ;
        merge uzit ciselsro ; by ntd ;
            drop ntd ;
        proc sort ; by ncsr ;
data pordsro ;
        title "precislovani spolecne skupina na pev reg 399";
        set uzit2; by ncsr ; if first.ncsr ;
data ciselsro ;
        set pordsro ;
            ncsr2 = _n_ ;
        keep ncsr ncsr2 ;
        proc means ;
data uzit3 ;
        merge uzit2 ciselsro ; by ncsr ;
            drop ncsr ;
        proc sort ; by ctp ;
data pordsro ;
        title "precislovani spolecne trvale prostredi 410";
        set uzit3; by ctp ; if first.ctp ;
data ciselsro ;
        set pordsro ;
            ctp2 = _n_ ;
        keep ctp ctp2 ;
        proc means ;
data uzit4 ;
        title "prumery na odregresovani 417";
        merge uzit3 ciselsro ; by ctp ;
keep ntd2 ncsr2 lakt rokot jed ctp2 dl sv b1p b2p b3p yd1 yd2 yd3
        m11 tu1 bi1 sb1 m12 tu2 bi2 sb2 m13 tu3 bi3 sb3 vaha a ;
            a = 1 ;
        proc means ;
        proc sort ; by jed ;
proc means noprint ;
        var yd1 yd2 yd3 m11 m12 m13 ;
        output out = pru ;
data smod ;
        title "smer odch 427";
        set pru ; if _stat_ = "STD";
            syd1 = yd1 ; sml1 = m11 ;
            syd2 = yd2 ; sml2 = m12 ;
            syd3 = yd3 ; sml3 = m13 ;
            a= 1 ;
        keep syd1 syd2 syd3 sml1 sml2 sml3 a ;

```

```

proc means ;
data uzit4a ; title "odregresovani do dnu uzitkovosti 435";
merge uzit4 smod ; by a ;
if yd1 ne . then do ; yd1 =yd1*0.60*sml1/syd1 ; m11 = yd1 ; end ;
if yd2 ne . then do ; yd2 =yd2*0.60*sml2/syd2 ; m12 = yd2 ; end ;
if yd3 ne . then do ; yd3 =yd3*0.60*sml3/syd3 ; m13 = yd3 ; end ;
keep ntd2 ncsr2 lakt rokot jed ctp2 dl sv b1p b2p b3p yd1 yd2 yd3
m11 tu1 bi1 sb1 m12 tu2 bi2 sb2 m13 tu3 bi3 sb3 vaha ;
proc means ;
data uzi ;
merge uzit4a(in= zvire) ciselj ; by jed ; if zvire ;
if ncsr2 = . then ncsr2= -9999.999 ;
if lakt= . then lakt= -9999.999; if rokot= . then rokot= -9999.999;
if ctp2 = . then ctp2 = -9999.999;
if dl = . then dl = -9999.999; if sv = . then sv = -9999.999;
if b1p = . then b1p = -9999.999; if b2p = . then b2p = -9999.999;
if b3p = . then b3p = -9999.999;

if m11 = . then m11 = -9999.999; if m12 = . then m12 = -9999.999;
if m13 = . then m13 = -9999.999;
if tu1 = . then tu1 = -9999.999; if tu2 = . then tu2 = -9999.999;
if tu3 = . then tu3 = -9999.999;
if bi1 = . then bi1 = -9999.999; if bi2 = . then bi2 = -9999.999;
if bi3 = . then bi3 = -9999.999;
if sb1 = . then sb1 = -9999.999; if sb2 = . then sb2 = -9999.999;
if sb3 = . then sb3 = -9999.999;
file uzit&soubo ;
put ntd2 ncsr2 lakt rokot nccjj ctp2 dl sv b1p b2p b3p
m11 tu1 bi1 sb1 m12 tu2 bi2 sb2 m13 tu3 bi3 sb3 vaha ;
proc means ;
/*.....genotypovani .....*/
data genot ; title "genotypovani DY +TD 460";
set genbTD genbDY ;
if jed = "" then delete ;
proc means ;
proc sort ; by jed ;
data geno ;
set genot ; by jed ; if first.jed ;
proc means ;
data gvmat ; title "genotypovani v matici 468";
attrib jedin format=$16. length=$16;
infile gcis5 ;
input jedin 1-16 gncj 18-24 ;
if jedin = "" then delete ; if gncj = . then delete ;
proc means ;
proc sort ; by jedin ;
proc sort data =ciselj ; by jedin ;
data gvmat ; title "precislovani sezgen5 476" ;
merge gvmat(in=gg) ciselj ; by jedin ; if gg ;
keep nccjj gncj ;
if nccjj = . then delete ;
file sezgen5d ;
put nccjj 1-7 ;
proc means ;
proc sort ; by gncj ;
data gmat ; title
"G matice 484";
infile slog ;
input gncj 1-7 gncj2 9-15 g 17-24 aa 26-33 ;
if gncj = . then delete ; if gncj2 = . then delete ;
proc means ;
proc sort ; by gncj ;

```

```

data slog ;                                title "precislovani G matice 490";
merge gmat gvmat(in=slo) ; by gncj ; if slo ;
      jel = nccjj ;
      gncj = gncj2 ;
      if jel = . then delete ;
      keep jel gncj g aa ;
proc means ;
proc sort ; by gncj ;
data slog2;
merge slog gvmat(in=slo2) ; by gncj ; if slo2 ;
      keep jel nccjj g aa ;
      if nccjj = . then delete ;
file slog45d ;
      put jel 1-7 nccjj 9-15 g 17-24 6 ;
file aamat ;
      put jel 1-7 nccjj 9-15 aa 17-24 6 ;
proc means ;
run;
/*.....konec .....*/

```

7. Parametrový soubor pro řešení BLUP-AM

```

#      paraRR3l
# Random Regression TD model for MT 3 lactations
# 25 columns: htd classes, group of lact classes, lact, yearcalv, anim, cow,
# dim, sv, lp1, lp2, lp3,
# test-day milk1, fat1, prot1, sbs1, milk2, fat2, prot2, sbs2, milk3, fat3, prot3, sbs3,
# weights groupRez. Missing values = 0 ;
#
# milk = HTD + fix reg (within classes of fixed effects)
#      + random reg PE (within cow)
#      + random genetic reg (within animal) + e
DATAFILE
uzdomu
NUMBER_OF_TRAITS
3
NUMBER_OF_EFFECTS
13
OBSERVATION(S)
12 16 20
WEIGHT(S)
24
EFFECTS: POSITIONS_IN_DATAFILE NUMBER_OF_LEVELS TYPE_OF_EFFECT
[EFFECT NESTED]
1 1 1 39123 cross
2 2 2 163 cross
9 9 9 163 cov 2 2 2
10 10 10 163 cov 2 2 2
11 11 11 163 cov 2 2 2
6 6 6 140691 cross
9 9 9 140691 cov 6 6 6
10 10 10 140691 cov 6 6 6
11 11 11 140691 cov 6 6 6

```

5 5 5 544287 cross
9 9 9 544287 cov 5 5 5
10 10 10 544287 cov 5 5 5
11 11 11 544287 cov 5 5 5

RANDOM_RESIDUAL VALUES

7.29638 0 0
0 10.8101 0
0 0 14.0196

RANDOM_GROUP

6 7 8 9

RANDOM_TYPE

diagonal

FILE

(CO)VARIANCES

8.96632 4.72962 4.03267 -0.003617 -0.295957 -0.027363 -0.510039 -0.224115 -0.204849 -
0.043919 -0.06277 0.024566
4.72962 14.8695 6.96711 0.597946 -0.522916 -0.683812 -0.330395 -0.976687 -0.260635
0.068563 -0.136729 -0.091258
4.03267 6.96711 20.0066 0.30522 0.410954 -0.328394 -0.198196 -0.369932 -1.64072 0.035612
-0.071148 -0.260597
-0.003617 0.597946 0.30522 1.34258 -0.066505 -0.155126 -0.113189 -0.103244 -0.077379 -
0.09675 -0.029119 -0.024141
-0.295957 -0.522916 0.410954 -0.066505 2.40916 0.182615 -0.010096 -0.133765 -0.079578 -
0.030618 -0.236216 -0.09712
-0.027363 -0.683812 -0.328394 -0.155126 0.182615 3.62742 -0.085846 0.022972 -0.433617 -
0.044352 -0.078243 -0.573036
-0.510039 -0.330395 -0.198196 -0.113189 -0.010096 -0.085846 0.714045 0.07795 0.016117 -
0.100648 -0.031868 -0.007493
-0.224115 -0.976687 -0.369932 -0.103244 -0.133765 0.022972 0.07795 1.25262 0.23999
0.01139 -0.213644 -0.076836
-0.204849 -0.260635 -1.64072 -0.077379 -0.079578 -0.433617 0.016117 0.23999 1.80933
0.019211 0.003822 -0.259404
-0.043919 0.068563 0.035612 -0.09675 -0.030618 -0.044352 -0.100648 0.01139 0.019211
0.328473 0.030383 0.014558
-0.06277 -0.136729 -0.071148 -0.029119 -0.236216 -0.078243 -0.031868 -0.213644 0.003822
0.030383 0.519975 0.110337
0.024566 -0.091258 -0.260597 -0.024141 -0.09712 -0.573036 -0.007493 -0.076836 -0.259404
0.014558 0.110337 0.709867

RANDOM_GROUP

10 11 12 13

RANDOM_TYPE

add_animal

FILE

maprm5g

(CO)VARIANCES

5.38463 5.01312 4.46822 0.471885 0.110301 -0.200918 -0.783269 -0.751558 -0.845773
0.013446 0.038964 0.089572
5.01312 6.00567 5.56752 0.980788 0.938762 0.563052 -0.72677 -0.756457 -0.799156 -
0.005003 -0.04284 -0.019388

4.46822 5.56752 5.54193 0.973211 0.989551 0.622439 -0.66338 -0.748437 -0.802825 -
0.017084 -0.058028 -0.055586
0.471885 0.980788 0.973211 1.05353 1.32061 1.28323 -0.056905 -0.025642 -0.056773
0.038797 0.019269 0.014763
0.110301 0.938762 0.989551 1.32061 2.68015 2.66918 0.311827 0.340501 0.419846 0.015116 -
0.013635 -0.145802
-0.200918 0.563052 0.622439 1.28323 2.66918 2.7954 0.331903 0.354187 0.4179 0.05249
0.033616 -0.103598
-0.783269 -0.72677 -0.66338 -0.056905 0.311827 0.331903 0.433888 0.45426 0.491596 -
0.036461 -0.061443 -0.104934
-0.751558 -0.756457 -0.748437 -0.025642 0.340501 0.354187 0.45426 0.582598 0.629542 -
0.012665 -0.062225 -0.105752
-0.845773 -0.799156 -0.802825 -0.056773 0.419846 0.4179 0.491596 0.629542 0.729495 -
0.03744 -0.081725 -0.147627
0.013446 -0.005003 -0.017084 0.038797 0.015116 0.05249 -0.036461 -0.012665 -0.03744
0.062203 0.065567 0.065384
0.038964 -0.04284 -0.058028 0.019269 -0.013635 0.033616 -0.061443 -0.062225 -0.081725
0.065567 0.092882 0.085039
0.089572 -0.019388 -0.055586 0.014763 -0.145802 -0.103598 -0.104934 -0.105752 -0.147627
0.065384 0.085039 0.112655
OPTION SNP_file genot2
OPTION saveAscii
OPTION tunedG 0
#OPTION AlphaBeta 0.8 0.2
OPTION readG gemat2
#OPTION prior 6 5 10 5 -1 5
#OPTION hetres_int 25 10
#OPTION conv_crit 1e-17
#OPTION maxrounds 20000

8. Výsledky řešení

Tabulka 2. Spolehlivost v 2016, determinace, validovaná spolehlivost předpovědi a regresní předpověď v závislosti na metodě předpovědi (PH, nebo GEPH) a vstupních souborech (domácí KU, Interbullové MACE hodnoty všech býků, spojené soubory).

Soubor Předpověď	Domácí		Interbull		D+I	
	PH	GEPH	PH	GEPH	PH	GEPH

Býků	390, 1.+2.+3.laktací dcer = 107, 2.+3.laktací = 32, 3.laktací = 9					
Rel16 = 0,85	Ověření podle PH 2016					
Determinace	0,166	0,309	0,242	0,391	0,247	0,389
Valid.spolehliv.	0,195	0,364	0,285	0,460	0,291	0,458
Společ. konst.	6,174	3,864	7,360	6,120	5,182	3,646
Regresní koef.	0,857	0,955	0,857	0,891	0,806	0,836
Determinace	Z Interbullu jen genotypování býci		0,163	0,335	0,185	0,366
Valid.spolehliv.			0,192	0,394	0,218	0,431
Společ. konst.			9,138	7,550	6,300	3,938
Regresní koef.			0,781	0,857	0,761	0,846

Rel16 = 0,90	Ověření podle GEPH 2016					
Determinace	0,166	0,334	0,230	0,418	0,236	0,405
Valid.spolehliv.	0,184	0,371	0,256	0,464	0,262	0,450
Společ. konst.	5,655	3,226	6,903	5,357	4,991	3,204
Regresní koef.	0,765	0,886	0,746	0,823	0,703	0,763
Determinace	Z Interbullu jen genotypování býci		0,162	0,374	0,182	0,394
Valid.spolehliv.			0,180	0,416	0,202	0,438
Společ. konst.			8,336	6,522	5,849	0,305
Regresní koef.			0,696	0,809	0,674	0,784

Soubor Předpověď	Domácí		Interbull		D+I	
	PH	GEPH	PH	GEPH	PH	GEPH

Jalovic	24.146, laktací = 3					
Rel16 = 0,72	Ověření podle PH 2016					
Determinace	0,518	0,526			0,515	0,517
Valid.spolehliv.	0,719	0,731			0,715	0,718
Společ. konst.	7,925	7,844			7,640	7,654
Regresní koef.	1,210	1,217			1,184	1,171
Determinace	Z Interbullu jen genotypování býci				0,521	0,524
Valid.spolehliv.			0,724	0,728		
Společ. konst.			7,785	7,794		
Regresní koef.			1,185	1,175		

Rel16 = 0,73	Ověření podle GEPH 2016					
Determinace	0,518	0,527			0,515	0,518
Valid.spolehliv.	0,710	0,722			0,705	0,710
Společ. konst.	7,961	7,873			7,675	7,684
Regresní koef.	1,212	1,220			1,185	1,174
Determinace	Z Interbullu jen genotypování býci				0,522	0,525
Valid.spolehliv.			0,715	0,719		
Společ. konst.			7,820	7,824		
Regresní koef.			1,189	1,178		

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

Název METODIKA GEPH NA ZÁKLADĚ REFERENČNÍHO SOUBORU SLOŽENÉHO
Z DOMÁCÍCH TD ZÁZNAMŮ A INTERBULLEM PŘEPOČTENÝCH
MEZINÁRODNÍCH MACE HODNOT

Autoři prof. Ing. Josef Příbyl, DrSc. (52 %)
Ing. Jiří Bauer, Ph.D. (8 %)
Ing. Anita Kranjčevičová (8 %)
Ing. Petr Pešek (8 %)
Ing. Jana Příbylová, CSc. (8 %)
doc. Ing. Luboš Vostrý, Ph.D. (8 %)
Ing. Ludmila Zavadilová, Ph.D. (8 %)

Oponenti: Ing. Zdeňka Majzlíková
Česká plemenářská inspekce, Praha
doc. Ing. Karel Mach, CSc.

ISBN 978-80-7403-157-1

Dedikace: Metodika byla zpracována za podpory MZe ČR, úkol č. QJ1510139 (Celostátní
informační systém genetického hodnocení hospodářských zvířat).

Vydáno bez jazykové úpravy