

Prediction of GEBV in comparison with GMACE values of genotyped young foreign bulls

***J. Příbyl, J. Bauer*, P. Pešek, J. Příbylová, H. Vostrá Vydrová, L. Vostrý,
L. Zavadilová***

(Institute of Animal Science, Praha – Uhřetěves, Czech Rep.)

V. Čermák, J. Šplíchal

(Czech Moravian Breeding Corporation, Hradištko, Czech Rep.)

J. Motyčka

(Holstein Cattle Breeders Association of the Czech Rep.)

Interbull Meeting in Berlin,
May 18-21, 2014

Circumstances

Majority of sires are foreign



Therefore sires in insemination have
weak connection to domestic population

X

But criteria of selection is the production
in domestic condition

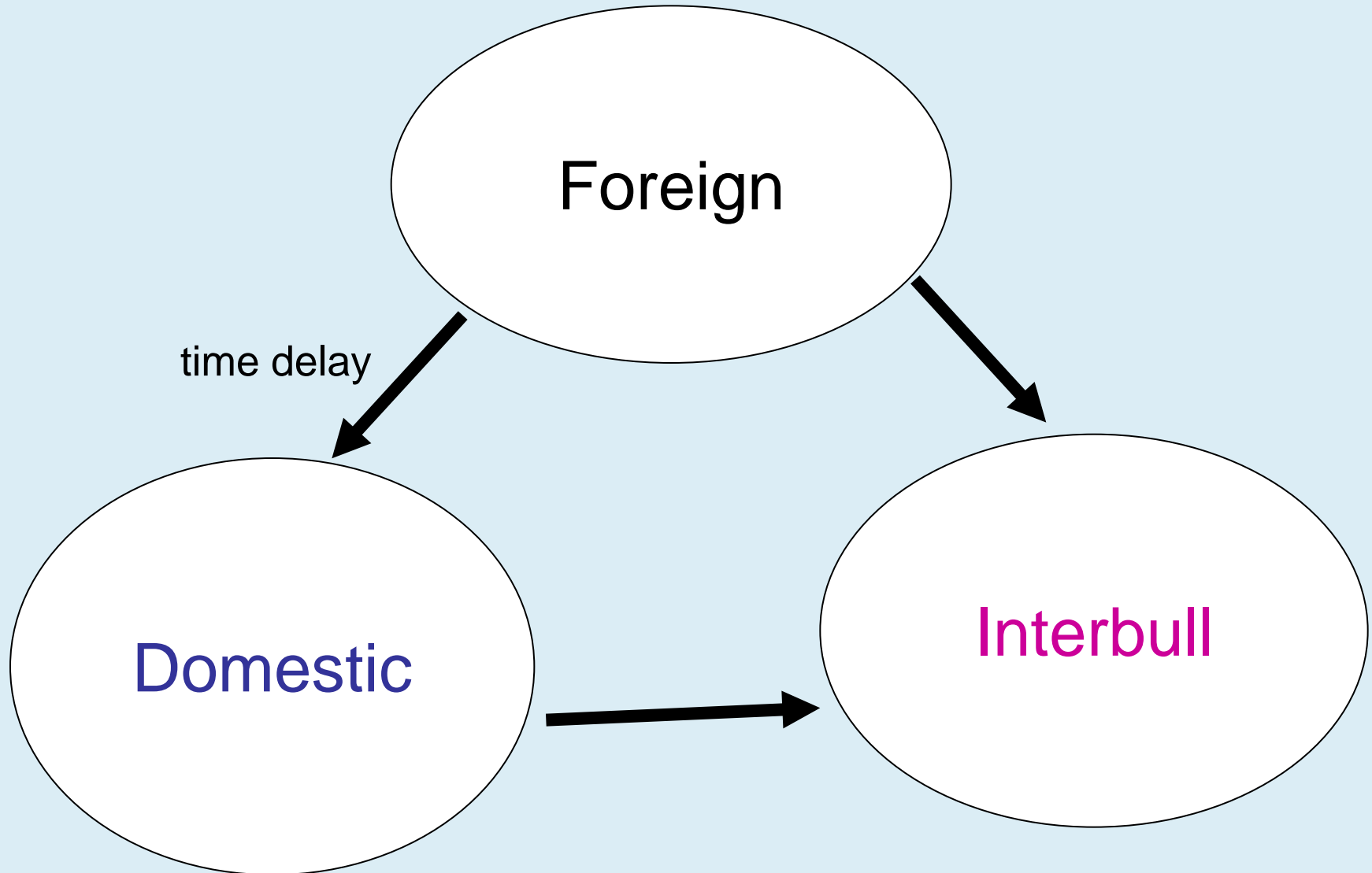
Goal:

best prediction



use all possible information

Prediction of EBV / GEBV



Holstein data for prediction until 2013

Domestic records - 1st lactation
1,257,462 cows

Domestic TD records, 3 lactations
1,086,267 cows,
19,435,367 TD records

Interbull EBVs
112,880 bulls

Genotypes (50k chip)
40,409 loci,
2,627 bulls

training set
2,189

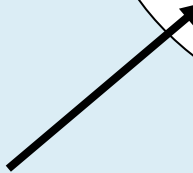


File for verification

GMACE

96 foreign bulls
in a Czech scale

Bulls not used in training set



Methods of prediction

- a) BLUP – Animal Model
- b) RRBLUP
- c) GBLUP
- d) ssGBLUP
- e) Blending ssGBLUP
- f) Combination d + e

Prediction procedures

Method	Value	Domest 1 st l.	Domest TD 3 l.	Interb DRP	D + I 1 st l. + DRP
BLUP	EBV	X		X	X
PA ¹	EBV			X	
RRBLUP	DGV			X	
	GEBV			X²	X³
GBLUP	DGV			X	
	GEBV			X²	X³
ssGBLUP ⁴	GEBV		X	X⁵	X

1) PA (from Interbull EBV) = $0.5 \cdot \text{EBV sire} + 0.25 \cdot \text{EBV maternal grandsire}$;

2) $\text{GEBV} = 0.8 \text{ DGV} + 0.2 \text{ PA}$;

3) $\text{GEBV} = 0.8 \text{ DGV} + 0.2 \text{ Domestic EBV 1}^{\text{st}} \text{ lactation}$;

4) Genomic relationship **G** is 80 %; pedigree relationship **A₂₂** 20 % in **H**

5) One-step blending approach

Correlation of prediction with GMACE

	Method	Value	Correlation
Domest 1st l.	BLUP	EBV	.33
Domest TD 3 l.	ssGBLUP	GEBV	.69
Interbull	PA	EBV	.41
DRP	RRBLUP	DGV	.60
only	RRBLUP	GEBV (EBV)	.67
genotyped	RBLUP	GEBV (PA)	.62
bulls	GBLUP	DGV	.62
+ pedigree	GBLUP	GEBV (EBV)	.68
	GBLUP	GEBV (PA)	.63
Interbull	Blending	EBV	.47
DRP all	Blending	GEBV	.69
D + I	Combination	EBV	.60
all	Combination	GEBV	.75

Correlation of prediction with GMACE

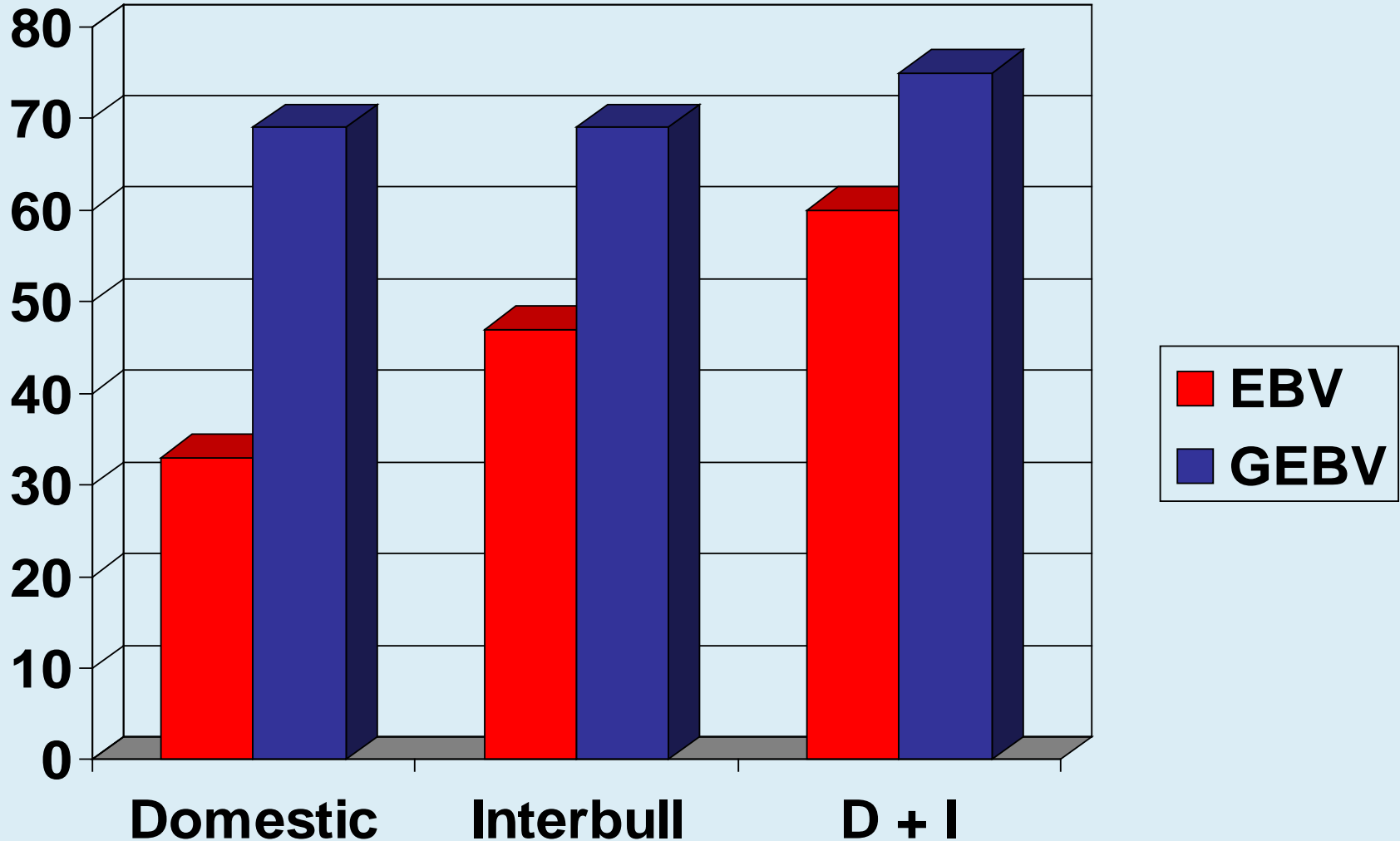
	Method	Value	Correlation
D omest 1st l.	BLUP	EBV	.33
D omest TD 3 l.	ssGBLUP	GEBV	.69
I nterbull DRP only genotyped bulls + pedigree	PA	EBV	.41
	RRBLUP	DGV	.60
	RRBLUP	GEBV (EBV)	.67
	RBLUP	GEBV (PA)	.62
	GBLUP	DGV	.62
	GBLUP	GEBV (EBV)	.68
	GBLUP	GEBV (PA)	.63
I nterbull DRP all	Blending	EBV	.47
	Blending	GEBV	.69
D + I all	Combination	EBV	.60
	Combination	GEBV	.75

Correlation of prediction with GMACE

	Method	Value	Correlation
Domest 1st l.	BLUP	EBV	.33
Domest TD 3 l.	ssGBLUP	GEBV	.69
Interbull DRP only genotyped bulls + pedigree	PA	EBV	.41
	RRBLUP	DGV	.60
	RRBLUP	GEBV (EBV)	.67
	RBLUP	GEBV (PA)	.62
	GBLUP	DGV	.62
	GBLUP	GEBV (EBV)	.68
	GBLUP	GEBV (PA)	.63
Interbull DRP all	Blending	EBV	.47
	Blending	GEBV	.69
D + I all	Combination	EBV	.60
	Combination	GEBV	.75

Summary of correlations of predictions to GMACE

according sources of information



Milk components - Correlations of Domestic TD ssGBLUP with GMACE

(135 bulls)

TD ssGBLUP prediction	GMACE		
	Milk (kg)	Protein (kg)	Fat (kg)
Milk (kg)	0.69	0.43	0.23
Protein (kg)	0.61	0.64	0.36
Fat (kg)	0.37	0.37	0.61
Protein (%)	-0.30	0.21	0.16
Fat (%)	-0.43	-0.15	0.27

Milk components - Correlations of Domestic TD ssGBLUP predictions

(96 bulls)

TD	Protein (kg)	Fat (kg)	Protein (%)	Fat (%)
Milk (kg)	0.81	0.49	-0.53	-0.68
Protein (kg)	-	0.61	0.07	-0.37
Fat (kg)		-	0.04	0.30
Protein (%)			-	0.62

Conclusion

- All genomic predictions > non-genomic predictions
- Single step procedures > Multi-step predictions (according to correlations)
- Addition of pedigree EBV to the DGV in Multi-step procedures increases accuracy
- Interbull + Domestic data > Domestic data
- Interbull + Domestic data help genotyped animals without connection to domestic population
- Our selection for routine is Domestic TD ssGBLUP

**Thank you
for your
attention**

Acknowledgments

We are grateful to Dr. Per Madsen and Dr. Guosheng Su from University of Aarhus, Denmark and Prof. Ignacy Misztal from University of Georgia, USA, for methodical support and provision of computer programs DMU, G-matrix and BLUPF90.

Additional notes

Weighted analysis

$$\text{ERC} = \delta (\text{rel}/(1-\text{rel}))$$

$$\delta = ((1-h^2)/ h^2)$$

Weighted analysis

Ridge Regression

$$\mathbf{DRP}_j = \mu + \sum \delta_i \cdot \mathbf{g}_{ij} + \mathbf{e}_j$$

Fixed effects

μ – common constant

Random effects

\mathbf{g}_{ij} – genotype of j-th bull in i-th locus

δ_i – regression coefficient

\mathbf{e}_j – residual

Heritability (h^2) = 0.25; Weight = ERC

Weighted analysis

GBLUP

$$\mathbf{DRP}_j = \mu + \mathbf{an}_j + \mathbf{e}_j$$

Fixed effects

μ – common constant

Random effects

\mathbf{an}_j – animal

\mathbf{e}_j – residual

Heritability (h^2) = 0.25; Weight = ERC

Weighted analysis

BLUP / ssGBLUP, lactation model

$$\mathbf{milk}_{ijkl} = \mathbf{HYS}_i + \beta_1 \cdot \mathbf{ca}_k + \beta_2 \cdot \mathbf{ca}_k^2 + \beta_3 \cdot \mathbf{do}_l + \beta_4 \cdot \mathbf{do}_l^2 + \mathbf{an}_j + \mathbf{e}_{ijkl}$$

Fixed effects

HYS – contemporary group

β – regression coefficients

\mathbf{ca}_k ; \mathbf{ca}_k^2 – curvilinear regression
on calving age

\mathbf{do}_l ; \mathbf{do}_l^2 – curvilinear regression
on days open

Random effects

\mathbf{an}_j – animal

\mathbf{e}_{ijkl} – residual

Heritability (h^2) = 0.25; Weight = ERC

ssGBLUP, RR-TDM, 3 lactations

$$\mathbf{y}_{ijn} = \text{HTD}_{in} + \beta_1 \cdot \text{ca}_j + \beta_2 \cdot \text{ca}_j^2 + \beta_3 \cdot \text{do}_{jn} + \beta_4 \cdot \text{do}_{jn}^2 + \beta_5 \cdot \text{ci}_{jn} + \beta_6 \cdot \text{ci}_{jn}^2 + f_{fg,n} + f_{pe,n} + f_{an,n} + e_{ijn} \quad ,$$

y_{ijn} = test-day record of milk yield of cow in lactation $n \in \{1, 2, 3\}$;

HTD_{in} = herd-test-day contemporary group i within a herd in lactation n (fixed effect);

$\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$ and β_6 = fixed regression coefficients;

ci_{jn} and ci_{jn}^2 = parameters for curvilinear regressions on calving interval for second and third lactations (fixed effect);

$f_{fg,n}$ = average LP of lactation curve according to groups of cows within management classes of systematic environment (Zavadilová et al., 2005b) (fixed effect);

$f_{pe,n}$ = permanent environmental within lactation LP of lactation curve of cows, random effect with covariance matrix (Zavadilová et al., 2005a);

$f_{an,n}$ = genetic within lactation LP of lactation curve of animal, random effect with covariance matrix;

e_{ijn} = random residual of test day records within lactation n , reflecting changes of variability along the course of lactation.

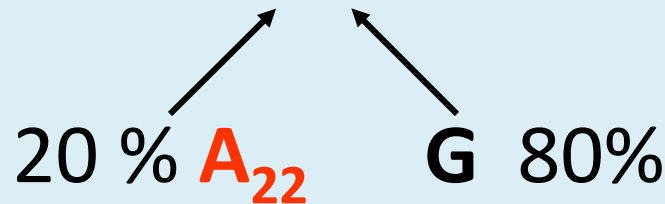
Relationship

Ridge Regression...**I**

BLUP – AM**A**

GBLUP.....**G** (*VanRaden 2008*)

ssGBLUP.....**H** (*Legarra et al., 2009*)



G – normalised (aver. diag. = 1)

(*Forni et al., 2011*)

shifted (aver. **G** = aver. **A₂₂**)

(*Vitezica et al., 2011*)

SNP editing:

- MAF,
- G-score,
- No. of loci per bull,
- No. of bulls per locus,
- Big error of prediction of old bulls
in training set,
- Big discrepancy of relationship $A_{22} \times G$,
- Proportion of H.