Integration of single step DGV in conventional genetic evaluations using DGV-PBLUP

H. Eding, L. van Kempen, M. van Pelt and J. Vandenplas

CRV
Coöperatie
BETTER COWS > BETTER LIFE

Reason for integrating

CRV holding is divided into two parts

1. CRV Company
   - Commercial branch (sale of semen and embryo's, breeding program)
2. CRV Herdbook
   - **National genetic evaluation**

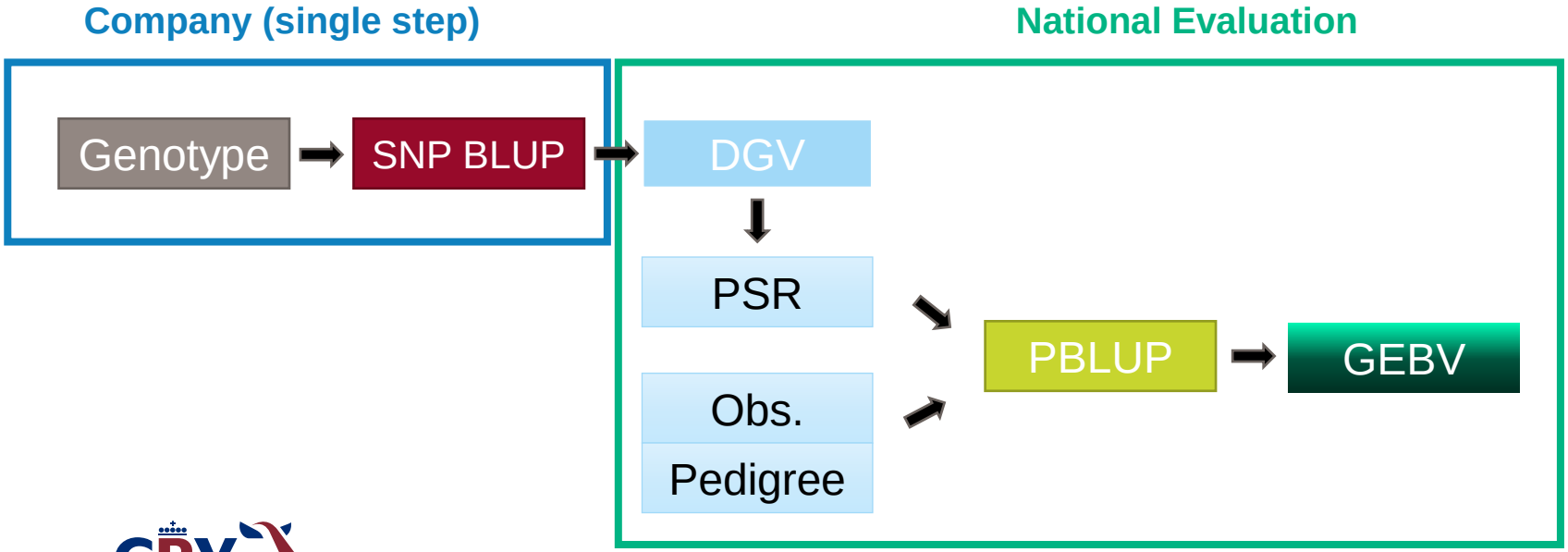
Single step (and earlier genomic evaluations) proprietary to CRV Company

The herdbook cannot use genotypes or SNP effects in the national evaluation.

– Solution: Integrate Direct Genomic Values into national evaluation

Schematic overview of national genetic evaluation

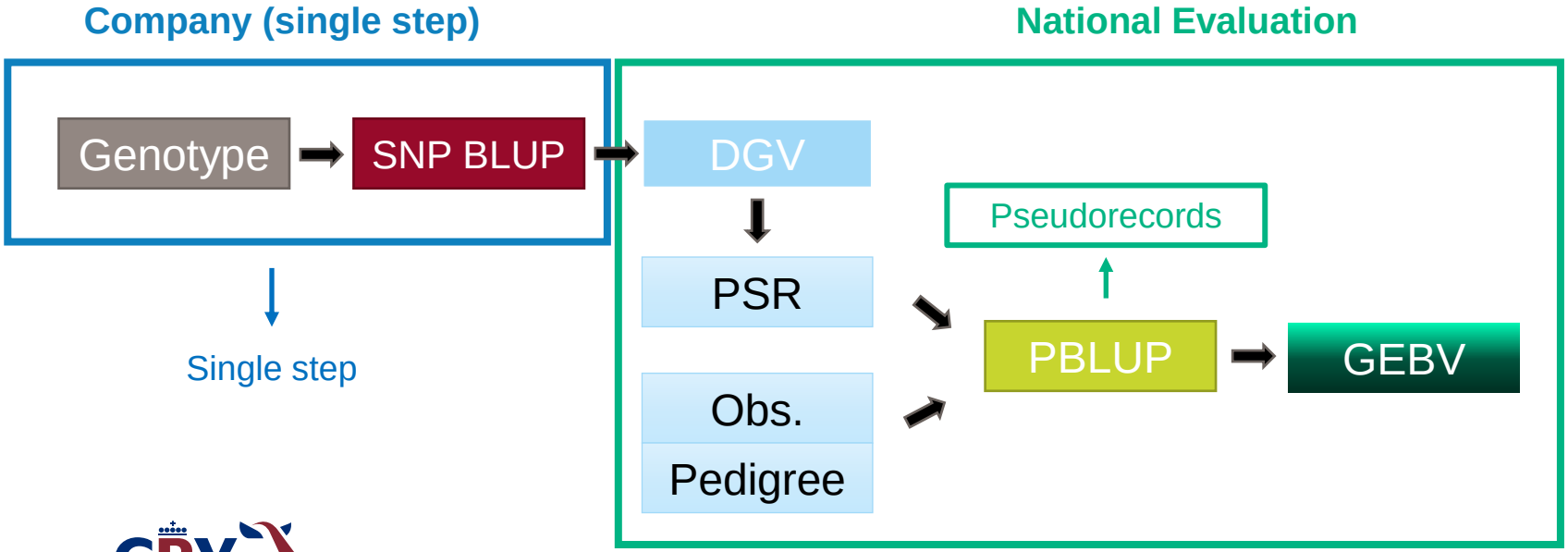
CRV Company supplies DGV for integration into the national genetic evaluations done by the herdbook



CRV Coöperatie
BETTER COWS > BETTER LIFE

Schematic overview of national genetic evaluation

CRV Company supplies DGV for integration into the national genetic evaluations done by the herdbook



CRV Coöperatie
BETTER COWS > BETTER LIFE

Adding genomic information to national evaluation

Current national evaluation uses a pseudorecord system
DGV supplied for national evaluations used as observations on a pseudo-trait

$$\begin{bmatrix} y \\ d \end{bmatrix} = \begin{bmatrix} a_y \\ a_d \end{bmatrix} + \begin{bmatrix} e_y \\ a_e \end{bmatrix} \quad \text{and} \quad \text{var}\begin{bmatrix} a_y \\ a_d \end{bmatrix} = \begin{bmatrix} \sigma_g^2 & r_{yd}^2\sigma_g^2 \\ r_{yd}^2\sigma_g^2 & \sigma_g^2 \end{bmatrix}$$

- The correlation to existing trait integrates genomic data into EBV of existing trait

CRV Coöperatie
BETTER COWS > BETTER LIFE

Adding genomic information to national evaluation

Current national evaluation uses a pseudorecord system
DGV fitted as observations on a pseudo-trait
The correlation to existing trait integrated genomic data in EBV of existing trait

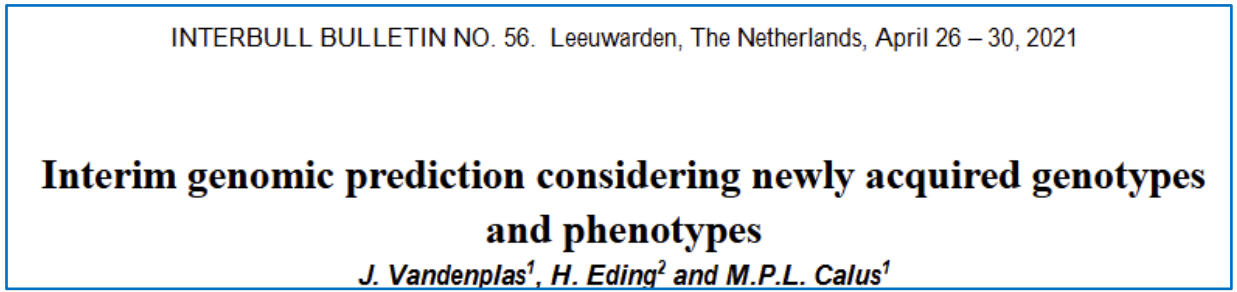
But:
- Many extra traits to fit (runtime and memory requirements)
- Imbalances in variance component matrices cause noise on GEBV/reliability

Ultimate goal: national GEBV == single step GEBV

CRV Coöperatie
BETTER COWS > BETTER LIFE

Theory

Single step SNP BLUP allows for interim genomic predictions
Equations derived assuming prior knowledge of SNP effects ($\hat{\mathbf{g}}$)


INTERBULL BULLETIN NO. 56. Leeuwarden, The Netherlands, April 26 – 30, 2021

Interim genomic prediction considering newly acquired genotypes and phenotypes

J. Vandenplas[1], H. Eding[2] and M.P.L. Calus[1]

Allows prediction of GEBV given a genotype using a conventional PBLUP

CRV Coöperatie
BETTER COWS > BETTER LIFE

Theory

Interim prediction equation

Equivalent to a full single step genomic evaluation

$$\begin{bmatrix} \mathbf{X'RX} & \mathbf{X'RW} \\ \mathbf{W'RX} & \mathbf{W'RW} + \frac{1}{\sigma^2}\mathbf{A}^{*-1} \end{bmatrix} \begin{bmatrix} \mathbf{b} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{X'Ry} \\ \mathbf{W'Ry} + \frac{1}{\sigma^2}\mathbf{A}^{*-1}\boldsymbol{\mu} \end{bmatrix}$$

Where:

$$\boldsymbol{\mu} = \begin{bmatrix} \mathbf{I} \\ \mathbf{A}_{ng}\mathbf{A}_{gg}^{-1} \end{bmatrix} \mathbf{Z}\hat{\mathbf{g}}$$

Theory

Intermediate prediction equation
Equivalent to a full single step genomic evaluation

$$\begin{bmatrix} \mathbf{X'RX} & \mathbf{X'RW} \\ \mathbf{W'RX} & \mathbf{W'RW} + \frac{1}{\sigma^2}\mathbf{A}^{*-1} \end{bmatrix} \begin{bmatrix} \mathbf{b} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{X'Ry} \\ \mathbf{W'Ry} + \frac{1}{\sigma^2}\mathbf{A}^{*-1}\boldsymbol{\mu} \end{bmatrix}$$

Where:

$$\boldsymbol{\mu} = \begin{bmatrix} \mathbf{I} \\ \mathbf{A}_{ng}\mathbf{A}_{gg}^{-1} \end{bmatrix} \mathbf{Z}\hat{\mathbf{g}}$$

$\mathbf{Z}\hat{\mathbf{g}}$ → allele substitution effects

$\mathbf{Z}\hat{\mathbf{g}}$ → genotypes

$\mathbf{A}_{ng}\mathbf{A}_{gg}^{-1}$ → ‘imputes’ from genotyped to non-genotyped relatives

CRV Coöperatie
BETTER COWS > BETTER LIFE

Theory

Intermediate prediction equation
Equivalent to a full single step genomic evaluation

$$\begin{bmatrix} \mathbf{X'RX} & \mathbf{X'RW} \\ \mathbf{W'RX} & \mathbf{W'RW} + \frac{1}{\sigma^2}\mathbf{A}^{*-1} \end{bmatrix} \begin{bmatrix} \mathbf{b} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{X'Ry} \\ \mathbf{W'Ry} + \frac{1}{\sigma^2}\mathbf{A}^{*-1}\boldsymbol{\mu} \end{bmatrix}$$

Where:

$$\boldsymbol{\mu} = \begin{bmatrix} \mathbf{I} \\ \mathbf{A}_{ng}\mathbf{A}_{gg}^{-1} \end{bmatrix} \mathbf{Z}\hat{\mathbf{g}}$$

→ genotype x ASE = DGV

Theory

DGV-PBLUP

Replace $\mathbf{Zg}$ with a vector of DGV $\mathbf{d}$:

$$\begin{bmatrix} \mathbf{X'RX} & \mathbf{X'RW} \\ \mathbf{W'RX} & \mathbf{W'RW} + \frac{1}{\sigma^2}\mathbf{A}^{*-1} \end{bmatrix} \begin{bmatrix} \mathbf{b} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{X'Ry} \\ \mathbf{W'Ry} + \frac{1}{\sigma^2}\mathbf{A}^{*-1}\boldsymbol{\mu} \end{bmatrix}$$

Where:

$$\boldsymbol{\mu} = \begin{bmatrix} \mathbf{I} \\ \mathbf{A}_{ng}\mathbf{A}_{gg}^{-1} \end{bmatrix} \mathbf{d}$$

- Possible replacement for pseudorecord system?

Theory

DGV-PBLUP: Conventional pedigree BLUP evaluation, augmented with DGV of genotyped animals.

DGV from genomic prediction fitted as a prior mean.

- Animal effect estimated as a deviation from the mean
  - In regular eval. BV = **0** +/- deviation
  - In DGV-PBLUP BV = **DGV** +/- deviation

Run as conventional pedigree BLUP evaluation
  - No SNP effects, no genotypes

CRV Coöperatie
BETTER COWS > BETTER LIFE

Proof of concept: Longevity

Routine single-step evaluation:

- #Phenotypes: 472,767,427
- #Pedigree : 16,660,929
- #Genotypes : 803,898

PSR:

- #Phenotypes: 472,767,427
- #Pedigree : 16,660,929
- #DGVs : 773,340

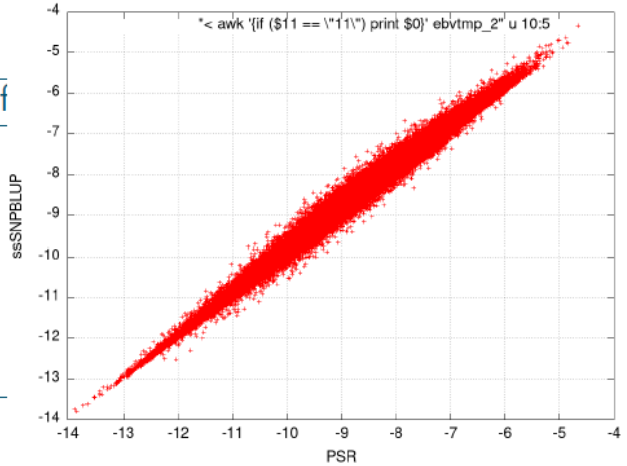
Longevity model is a 5th order random regression model:
6 polynomials (Effect 1…6)

CRV Coöperatie
BETTER COWS > BETTER LIFE

Proof of concept: Longevity

Genotyped animals with DGV in system

Effect	#animals	Correlation	Intercept	Reg. Coef
1	773340	0.9965	0.2771	1.01782
2	773340	0.9975	0.1108	1.01561
3	773340	0.9968	-0.013	1.01229
4	773340	0.9969	-0.006	1.00339
5	773340	0.9970	-0.001	0.98558
6	773340	0.9988	-0.000	0.98822



CRV Coöperatie
BETTER COWS > BETTER LIFE

Proof of concept: Longevity

DGV-PBLUP seems a promising replacement of our current pseudorecord system

- More accurate integration of DGV with phenotypic data
  - Expectation: Higher correlations and less GEBV changes compared to single step

- No need for extra evaluated pseudo-traits
  - Further simplifies evaluations (same model for all types of flows)
  - Easy to fit all traits with DGV from single step.
  - Less computer resources

Application: Milk production test day model

Five lactation random regression model with 4th order Legendre polynomials

- 5 polynomials per lactation: 25 animal solutions

Traits using this test day model

- Milk, fat, protein and lactose production
- Somatic cell score
- Urea content

Single-step evaluation:

- #Phenotyped: 13,662,463
- #Pedigree : 16,313,781
- #Genotyped : 828,590

PSR:

- #Phenotyped: 13,662,463
- #Pedigree : 16,382,568
- #DGVs : 851,704

CRV Coöperatie
BETTER COWS > BETTER LIFE

Run time iteration phase evaluation

Trait group	Single step	DGV-PBLUP	Conventional
Milk	51:28	18:17	17:53
Fat	47:26	18:39	17:37
Protein	49:05	19:20	18:08
Somatic Cell	48:47	19:02	19:30
Urea	56:02	18:22	18:31
Lactose	63:00	18:49	19:45

Evaluations were run on a test machine using 5 CPU for parallel computing

Intel(R) Xeon(R) Gold 6448H 64bit X86 4000MHz

CRV Coöperatie
BETTER COWS > BETTER LIFE

Results: Young bulls with genotype and DGV

Bulls without offspring born after 2020

	corr		< 1/4 s.d.	
Trait	**dgv**	**psr**	**dgv**	**psr**
Milk	0,999	0,928	100,0%	72,2%
Fat	0,999	0,964	100,0%	80,1%
Protein	0,998	0,934	100,0%	74,8%
Somatic Cell	0,995	0,972	99,8%	85,3%
Urea	0,999	0,994	100,0%	98,3%
Lactose	0,999	0,918	100,0%	70,5%

- **corr** **:** correlation GEBV with those from single step
- **< ¼s.d** : fraction of GEBV changes less than 25% genetic s.d.

CRV Coöperatie
BETTER COWS > BETTER LIFE

Discussion

Relative to current PSR system DGV-BLUP improves integration of genomic data

- Higher correlations
- Smaller differences in GEBV
- Improved agreement between national and single step evaluations

Runtime acceptable for operational use

- Comparable to conventional evaluation, faster than single step

Next steps

1. Implement and test in all other flows
2. Provide stakeholders with test file and await go/no go
3. Roll out as replacement of current pseudorecord system

CRV Coöperatie
BETTER COWS > BETTER LIFE

CRV
Coöperatie
BETTER COWS > BETTER LIFE