Genomic information in routine evaluation – Genomic prediction, blending and publication criteria NAV

Gert Pedersen Aamand*, Esa Mäntysaari and Jørn Pedersen

*Nordisk Avlsværdivurdering, Agro Food Park 15, DK-8200 Aarhus N, Denmark e-mail: gap@vfl.dk Homepage: www.nordicebv.info
NAV – genomic breeding values

1. The whole set up around genomic prediction
2. Streamlining RDC and Jersey
3. Blending/combining
4. Preliminary results
5. Ongoing work and plans
Bulls with known EBV and markers (SNPs) create ”DNA-dictionary”

Quality of the dictionary depends on the size of the reference population

SNPs

DNA-dictionary

Deregressed proofs DRP)
Abbreviations

- **DGV (SNP effects)**
  - Direct Genomic Value
- **EBV (phenotypic data from farms) – DRP**
  - Estimated breeding value
- **GEBV (SNP effects + phenotypic data)**
  - Genomic Enhanced Breeding value
Flow from phenotypes to EBV
Flow from DNA+phenotypes to GEBV

- SNP data base
- DNA
- Genotyping at Lab
- GS project
- Viking Genetics
- Farmers (females)
- Eurogenomics ref animals
- Extra for GBLUP
- DGV
- GEBV
- DK
- Swe
- Fin

Nordisk Avlsværdi Vurdering • Nordic Cattle Genetic Evaluation
Flow from DNA+phenotypes to GEBV

Requires 3 runs for all traits
1. Traditional EBV estimation
2. Genomic prediction to get DGV
3. Bi-variate blending estimation (EBV&DGV) to get GEBV
Flow from DNA+phenotypes to GEBV

- DGVs have to be estimated often (wish from VG each month) – (>40 traits for all 3 breeds)
- EBV (traditional evaluations) – 4 times a year
- GEBV (DGV & EBV)
  - The official breeding value (printed on certificates etc.)
  - 4 times a year or as often as DGVs – still open
Flow from DNA+phenotypes to GEBV

• A lot of evaluations require things are streamlined

• We need to develop it stepwise: It means we will not be able to estimate DGVs 12 times a year from primo 2011
Prediction model and dependent variable used in unofficial routine runs

<table>
<thead>
<tr>
<th></th>
<th>Holstein</th>
<th>RDC</th>
<th>Jersey</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRP</td>
<td>EBV to DRP</td>
<td>EBV to DRP</td>
<td></td>
</tr>
<tr>
<td>GBLUP</td>
<td>Bayes to GBLUP</td>
<td>Bayes to GBLUP</td>
<td></td>
</tr>
<tr>
<td>Since spring 2010</td>
<td>Change November 2010</td>
<td>Change November 2010</td>
<td></td>
</tr>
</tbody>
</table>

NAV’s goal is to use the same models and dependent variables for all breeds to make things streamlined - given we have no strong scientific reasons to differentiate.
Prediction model used in unofficial routine runs

- **Bayes**
  - Different variance on SNPs (assumes some genes have larger effect)
  - Computationally heavy
- **GBLUP**
  - Same variances (many genes, small effects)
  - Fast computation

No differences in validation correlation
Dependent variable used in unofficial routine runs

DRP

- Equivalent to a daughter group mean corrected for fixed effects – back transformed from EBV
- Do not include pedigree information
Prediction model and dependent variable

- RDC September 2010 and Jersey October 2010 (Gao&Su)
  - Comparison DRP/EBV and GBLUP/Bayes
  - Conclusion DRP is preferable dependent variable and GBLUP and Bayes give similar results

**Nov 2010 changes in DGV:**

- Correlation DGV-old and DGV-new about 95%!
Prediction model and dependent variable used in unofficial routine runs

Status 24.11.2010

- Holstein
  - NAV estimates DGVs
- RDC and Jersey
  - Aarhus University still estimates DGVs
  - NAV will take over during the coming months
Future prediction model and dependent variable

- If future R&D show significantly higher validation reliabilities by changing model for all or a single breed we will also differentiate for the routine run

- Choice for routine evaluation does not influence R&D work, but R&D has to have future possibilities for implementation in mind
Combining DGV and EBV

- Assumptions
  - DGV robust – stabile from evaluation to evaluation
  - We need to know the reliability of EBV and DGV to be able to combine
  - Assumptions are especially critical when comparing 1 year old bull with 5 year old with daughter groups
## Reliabilities DGV, examples Holstein

<table>
<thead>
<tr>
<th>Trait</th>
<th>Validation correlation²</th>
<th>r² (DGV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein</td>
<td>0.55</td>
<td>0.63</td>
</tr>
<tr>
<td>Fertility</td>
<td>0.45</td>
<td>0.55</td>
</tr>
<tr>
<td>Mastitis</td>
<td>0.48</td>
<td>0.62</td>
</tr>
<tr>
<td>Legs</td>
<td>0.24</td>
<td>0.45</td>
</tr>
<tr>
<td>Udder</td>
<td>0.53</td>
<td>0.72</td>
</tr>
</tbody>
</table>

Reliabilities have to be expressed exactly in the same way for all breeds. DGVs will be scaled so there STD match the reliabilities.
Method - blending

NAV decided to use Mäntysaari & Stranden (2010) method

BLUP – bivariate model

DRP – trait 1

DGV – trait 2
Bi-variate model

Bi-variate - “EBV-trait” and “DGV-trait”

Requires 3 runs

1. Traditional EBV estimation
2. Genomic prediction
3. Bi-variate blending estimation

Future maybe one step procedure
Example protein – BLUP
Bi-variate model – one animal

Solution for trait I is GEBV

Trait I
Protein, kg

Trait II
DGV
Eksample protein – BLUP
Bi variate model – one animal

Solution for trait I is GEBV

Trait I
Protein, kg

DRP
Reliability

Trait II
DGV

Solution for trait I is GEBV

Nordisk Avlsværdi Vurdering • Nordic Cattle Genetic Evaluation
Example protein – BLUP
Bi variate model – one animal

Solution for trait I is GEBV

Trait I
Protein, kg

DRP
Reliability

Trait II
DGV

DGV Heritability = 1.00

Heritability

Example protein – BLUP
Bi variate model – one animal

Solution for trait I is GEBV

Trait I
Protein, kg

Genetic correlation =
Reliability DGV

DRP
Reliability

Trait II
DGV

DGV
Heritability = 1.00
Blending in two steps

1. Early 2011
   • DRP for bulls and DGVs
   • All candidates and bulls get GEBVs

2. Later 2011
   • DRP for cows and DGVs
   • All animals get GEBVs - also female relatives to genotyped animals
DRP

1. We have a method for calculation of DRP for bulls from EBV (Strandén & Mäntysaari, 2010)

2. We investigate if the method works for estimation of DRP from cows from EBVs. (MTT is working on it)
Blending and relationship

- In the traditional estimation of breeding values information goes from an animal to relatives (progeny and parent).
- For genomic information – the same takes place, but if an animal is genomic tested then a genotyped relative does not provide extra genomic information.
How much extra information do we get from pedigree – additional to DGV?

<table>
<thead>
<tr>
<th>Holstein - Protein</th>
<th>Pedigree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>½Sire+½Dam</td>
</tr>
<tr>
<td>DGV+Pedigree</td>
<td>49.8%</td>
</tr>
<tr>
<td>DGV</td>
<td>49.7%</td>
</tr>
<tr>
<td>Pedigree</td>
<td>13.0%</td>
</tr>
</tbody>
</table>

Note pedigree add very little extra information! Bull dam add no extra information!
Preliminary results
GEBV for candidates

Correlation between:

- Sire+MGS pedigree index and EBV about 50%
- Sire+MGS pedigree index and DGV about 70%
- DGV-GEBV about 95%
Combining DGV and pedigree for candidates (young animals)

Results show:
• (Sire +MGS) add very very little extra information
• Bull dam add no extra information, but a potential risk of bias
Reference bulls

• Reference bulls can get more reliable GEBVs by including DGVs, especially for functional traits

• But DGVs for reference bulls are highly correlated to EBV, because we use a very over saturated model – much more SNPs are estimated than number of bulls in the ref. pop.
How to test if blending of DGV and EBV is optimal for reference bulls?

We can estimate GEBVs for the same bulls with daughters in two ways:

1. Treat the bull as a candidate and combine DGV and EBV
2. Treat the bull as a reference bull and combine DGV and EBV

Hopefully we get the same results otherwise we will see fluctuations in GEBV in practice
## Preliminary results - Holstein Protein – correlation EBV-GEBV

<table>
<thead>
<tr>
<th>Ref bulls born 95-04</th>
<th>$R_{ia}^2$, %</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>I Ref bulls born 95-04</td>
<td>&lt;88</td>
<td>0.96</td>
</tr>
<tr>
<td>II Ref bulls born 95-04</td>
<td>88-93</td>
<td>0.996</td>
</tr>
<tr>
<td>III Ref bulls born 95-04</td>
<td>93-97</td>
<td>0.999</td>
</tr>
<tr>
<td>IV Ref bulls born 95-04</td>
<td>&gt;97</td>
<td>0.998</td>
</tr>
</tbody>
</table>

Production traits only small changes if a bull has a full daughter group.
Preliminary results - Holstein Fertility – correlation EBV-GEBV

<table>
<thead>
<tr>
<th>Ref bulls born 95-04</th>
<th>Ria2, %</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>&lt;55</td>
<td>0.70</td>
</tr>
<tr>
<td>II</td>
<td>55-69</td>
<td>0.91</td>
</tr>
<tr>
<td>III</td>
<td>69-83</td>
<td>0.96</td>
</tr>
<tr>
<td>IV</td>
<td>&gt;83</td>
<td>0.98</td>
</tr>
</tbody>
</table>

Functional traits significant changes even though the bull has a full daughter group
1. What will happen in practice – DRP bulls

- Pedigree information adds very little extra information
- EBVs for bulls with milking daughters will change significantly for functional traits, but also for yield, if reliability is moderate

reliability DGV = reliability EBV
2. What will happen in practice—DRP cows

- EBVs for genotyped cows will change significantly – DGV gives a lot of new information – reliability DGV >> reliability EBV
- EBVs for cows related to genotyped animals can also be changed a bit.
How to estimate GEBV for combined traits e.g. mammary system?

- Blending DGV and EBV for linear udder traits or
- Blending DGV for udder and EBV for udder

Need to be analysed!
Ongoing work – genomic selection

- NAV take over unofficial routine for Holstein
- Change of dependent variable and model RDC and Jersey
- Test based on Holstein about blending, combining traits, reference bulls, DRP cows etc.
Key points - publication rules

Under discussion

• Publication of GEBVs from bulls at 18-20 month age (after ”test” inseminations) versus as early as possible

• Extra information DGVs?

• Frequency of evaluations – Final goal DGV 12 times per year and GEBV 4 times or?
Practical implementation

- NAV hope to be able to publish the first official GEBVs early 2011, BUT several things need to be tested and clarified before it is possible.

- GEBV for Holstein might be published before GEBV for RDC and Jersey.
NAV’s plan

1. Blend using DRP for bulls and DGV
2. Blend using DRP cows and DGV – all animals evaluated
3. Test better method

Status:

- Test run ongoing~ routine as soon as possible – hopefully Feb 2011
- DRP-cows under investigation – 2011+
- One step