Analysis of Longitudinal Data for Selection and Management

N. Gengler\textsuperscript{1,2} and P. Mayeres\textsuperscript{2}

\textsuperscript{1} National Fund for Scientific Research, Brussels, Belgium
\textsuperscript{2} Gembloux Agricultural University, Belgium
Objectives of this Talk

- Analysis of longitudinal data
  - special respect to dairy test day model
  - development of models
- More than only “genetic” results
  - use of test day model results for herd management purposes
  - development of management tools
- Evolution of genetic evaluation systems
  - towards integrated systems for management and selection of animals
Analysis of Longitudinal Data

- Dynamic biological processes
  - provide longitudinal data (e.g. depending on time)
  - until recently “static” models
  - eliminating influence continuos variable

- Examples
  - test day yields $\Rightarrow$ lactation yields
  - individual weights $\Rightarrow$ standardized weights, ADG

- Selection vs management
  - two clearly different objectives!
Current Use for Management

- Simple management “traits”
  - dynamic aspect nuisance
- Often eliminated using trivial methods
  - computation of a weighted average or sum
  - standardization by using adjustments
- Strictly on a phenotypic level
  - no consideration of genetic differences
- Raw values reported to farmers
Example: Lactation Yields

- Aggregating daily yields over 305 days
- Computed by mostly simple methods
  - test interval (TIM)
  - centering date methods (CDM)
- Recently more advance methods
  - Bayesian (MTP) or Regression (BP)
- Extension of lactation problem
  - strictly on a phenotypic level
  - RIP dip and “Sunny Boy” effect
Use for Selection (until recently)

- Genetic evaluations
  - use of mixed linear models
  - based on aggregated “traits”
  - clearly distinguished from management
  - different organizations?

- Environmental effects ⇔ nuisance
  - generally not used or even reported to farmers
  - lost of potentially interesting information!

- Only EBVs reported to farmers
Example: milk yield (until recently)

- Genetic evaluations
  - based on 305 day yield

- Effects typically included: (reported)
  - contemporary groups (no)
  - age effects (no)
  - permanent environment (nearly never)
  - genetic (yes)

- Few exceptions
Analyses of Longitudinal Data

- Recent advances
- Two central issues
  - describing $E(y)$ and $\text{Var}(y)$
- Description of the mean
  - evolution $E(y)$ over time
- Description of the (co)variances
  - evolution of $\text{Var}(y)$ over time
Modeling $E(y)$ over Time

- Often considered secondary
- **Objective:**
  - allowing correct comparisons among animals
- **Central issue for selection**
  - unbiasedness of genetic solutions
- **Central issues for management**
  - not the same
Modeling $\text{Var}(y)$ over Time

- Central issue for genetic evaluations
- Repeatability models
  $\Rightarrow$ Random coefficient (regression) models
- Multiple trait models
  $\Rightarrow$ (Co)variance functions
- Equivalent
Example: Test Day Models (TDM)

- Direct use of daily milk results
- Most recent TDM directly model
  - variation $E(y)$ over time
  - variation $\text{Var}(y)$ over time
- Numerous advantages
- Feasible due to $\uparrow$ computing power
- Results reported (currently)
  - report of performed yield and EBV
Test Day Models (TDM)

- Interesting for management use
  - strongest argument for TDM?
- Fixed effects
  - herd level, herd lactation curves
  - standard lactation curves
- Random effects
  - individual lactation curves
  - producing abilities persistency, maturity rate
- Prediction
Test Day Models (TDM)

- **Current TDM implementations**
  - focus on genetic effects

- **Some issues partly addressed**
  - standard lactation curves (reported?)
  - persistency (definitions? use?)
  - maturity (definitions? use?)

- **Several unsolved issues**
  - herd/cow specific lactation curves
  - producing abilities
  - “prediction” (herd and individual level)
Cow Specific Lactation Curves

- PE and genetic random regressions
- PE + genetic solutions
  - mostly only EBV considered
- Producing abilities
  - management potential
- Prediction
  - herd specific lactation curves
  - however always only deviations
Herd Specific Lactation Curves

- Herd environmental random regressions
- Large herds
  - herd specific curves
- Small herds
  - regressed towards over population curves
- Now considered in several TDM
- Prediction
  - herd specific lactation curves
  - also deviations
Prediction

- Very important issue for management
  - not only deviations, but also overall level
- Next test and overall production
  - herd level
  - individual level
- Compared with real value measured
  - out of the prediction interval

→ Management decisions!
Prediction with TDM

- Opposition to classical methods
  - TIM, CDM
  - MTP, BP

- They model directly the mean

- Prediction from TDM
  - could be directly obtained from solutions
  - by summing the effects of the model

- Problem: herd test day effect
Herd Test Day (HTD) Fixed Effect

- Results from Mayeres et al. (2002)
  - acknowledge Luxembourgish Herdbook, VIT

- HTD not predictable
  - effect does not model any trend

- Objective:
  - new modeling proposition

- Example how slight changes
  - improve usability of TDM
Study of HTD Fixed Effect

- HTD month’s mean for each year across herds for the 3 traits
- For milk
- Two trends:
  - General upward trend through years
  - Yearly trend with maximum near the pastern release
New Model

- Replacement of HTD fixed effect
- Herd test month fixed effect
  - period of 4 years (5 for newer years)
- Herd test year fixed effect
  - 2 years for current test years
- Herd test day random effect
Comparison of EBV

• Few changes in ranking
  - rank correlation of cows and sires > 0.99 for each trait
  - absolute difference between EBV of cows and sires are low for each trait

• Only few rankings change significantly
Comparison of Herd Effects

- **Model 2:**
  - Herd Effect = HTY + HTMp

- **Similar trend**
  - correlation is > 0.91 for each trait
  - absolute difference is very low for each trait

- **Biggest differences for HTD with few tests**

### Table: Absolute difference

<table>
<thead>
<tr>
<th>Trait</th>
<th>Correlation</th>
<th>Mean</th>
<th>Std</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Milk</td>
<td>0.918</td>
<td>1.00</td>
<td>0.91</td>
<td>17.6</td>
</tr>
<tr>
<td>Fat</td>
<td>0.919</td>
<td>0.046</td>
<td>0.042</td>
<td>0.87</td>
</tr>
<tr>
<td>Protein</td>
<td>0.919</td>
<td>0.037</td>
<td>0.032</td>
<td>0.49</td>
</tr>
</tbody>
</table>

### Graph: Observations by HTD
Comparison of Herd Effect

- Particular herd

- Two special tests (•)
  - 01/12/1994 and 06/04/1992: one animal tested
Some Questions

• Why no provide more results?
  • *we compute them anyway!*

• Why no adapt our models?
  • *we could gain too!*

• What is the real interest in EBVs?
  • *genetic evaluations very much separated from performance recording*
  • *current interest by farmers is decreasing*

• What is need for successful management?
Implications for the Future

• Personal opinions
  • analysis of longitudinal data
  • opportunity to develop advanced management tools
  • large influence evolution of genetic evaluation systems
  • interest in “genetics” only decreasing
  • opportunity to use “optimal” modeling
  • higher integration of selection and management leading eventually to

⇒ Integrated systems!
Integrated Systems for Management and Selection

- Provide optimal useful results for
  - management
  - selection
- Optimal use of computing power
- “Re-conciliate” farmers with EBVs
  - showing link phenotype to genetic values
  - avoiding “black box” syndrome
- Could avoid that genetic evaluations are sidelined