Use of Hardware and Sensors Towards Phenomics to Deliver Complex Data and Advance Animal Breeding

Juan P. Steibel
Professor of Animal Science
J. Lush Endowed Chair of Animal Breeding and Genetics
Phenome, phenomics, phenotyping

Phenomics = the acquisition of high-dimensional phenotypic data on an organism-wide scale.

“Some of the most scientifically disrupting and industry-relevant challenges relate to ‘phenomics’ as much as to ‘genomics’”
## What is novel in novel phenotyping?

<table>
<thead>
<tr>
<th>Novel / Hard to measure traits</th>
<th>Massive collection of common traits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Welfare/Health related</td>
<td>Real Time</td>
</tr>
<tr>
<td><strong>Behavioral</strong></td>
<td>“All” animals</td>
</tr>
<tr>
<td>Physiological</td>
<td>Across farms world-wide</td>
</tr>
<tr>
<td>Gas emissions</td>
<td>Under typical production conditions</td>
</tr>
<tr>
<td>Feed Conversion</td>
<td></td>
</tr>
</tbody>
</table>
Breeding programs

Production system

↑ pork quantity
↑ pork quality
↓ feed used

Breeding goals

Phenotypes
- Weight gain
- Carcass weight
- Backfat depth
- Marbling
- Feed conversion ratio

Genotypes

Genealogy (Pedigree)

Collect BIG data

Breeding Program

Genetic Progress
- Inbreeding

Evaluate program

Selection and Mating

Estimate genetic merit (breeding values)

\[ y = X\beta + Zu + e \]

\[ u \sim N(0, G\sigma^2) \]
Phenomics in animal breeding

- Raw sensor data
  - Extract relevant measure
- Phenotyping
  - Phenotypic measure
  - Link to perennial animal ID
- Statistical Modeling
  - SNP markers
  - genotype matrix $M$
  - $(\begin{bmatrix} x \\ w \\ y \end{bmatrix}) = f(M)$
- EPD/BV/PTA
The need for phenotyping

Phenotyping

Management
- Group-level indicators
- Individual indicators
- Enable Intervention

Genetic improvement
- Improve relevant traits
- Evaluate in the relevant environment

Defined by socio-economic conditions: The production system and the farmers/community.

It’s hard to improve what we don’t measure (attributed to P. Drucker)
An example from the USA: DHIA

1890s: Babcock’s test: How to measure butterfat.

1905: Helmer Rabild starts DHIA in Michigan

1936 1st proven sire list

- Provided useful data for management
- Built on existing infrastructure
- Fed sire comparisons (genetic evaluation)

Source: Brito et al 2021.
(re) using data collected through precision livestock farming systems

Sensors → Raw Data → Usable, processed data → Action

Unused data → $!

$IOWA STATE UNIVERSITY$
Types of sensors

Remote

Wearable

Proximal

Combination (most phenotyping technologies)

One sensors measures the phenotype, the other sensor IDs the animal

Diosdado et al 2015.
Challenges in using sensors for phenomics
Challenge 1: Extracting (valid) phenotypes from sensor signal

Training data: Images + Annotation

Model

Test data (images):

Performance under cross validation

<table>
<thead>
<tr>
<th>True Positives</th>
<th>False Positives</th>
<th>False Negatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mask accuracy</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Classifying interactions at the feeder

No Contact

Head to Body (direct)

Mounting

Levering
Recall: proportion of videos labeled (ground truth) as $<X>$ that are correctly classified.

Precision: proportion of videos classified as $<X>$ that are actually labeled (ground truth) as such.
Performance under across time validation

Drop in precision for most labels
Performance of validation across feeders
Another example: Animal detection

Goal: draw a box around each animal, evaluate its performance.

Then use the box to extract relevant phenotypes

Performance of pig detection over time

<table>
<thead>
<tr>
<th>Date Range</th>
<th>Test Date</th>
<th>Precision</th>
<th>IOU</th>
</tr>
</thead>
<tbody>
<tr>
<td>10-May</td>
<td>11-May</td>
<td>0.9302</td>
<td>0.71</td>
</tr>
<tr>
<td>10-May</td>
<td>17-May</td>
<td>0.8733</td>
<td>0.725</td>
</tr>
<tr>
<td>10-May</td>
<td>18-May</td>
<td>0.7988</td>
<td>0.687</td>
</tr>
<tr>
<td>10-May</td>
<td>24-May</td>
<td>0.843</td>
<td>0.705</td>
</tr>
<tr>
<td>10-May</td>
<td>25-May</td>
<td>0.8718</td>
<td>0.688</td>
</tr>
</tbody>
</table>
Another example: Animal detection

This model may be useful for predicting feeder use, but not so much for interactions at the feeder.
And another one: Key point detection
Here the goal is to detect previously selected key points in unmarked images of animals

Validated on horses recorded at MSU. It works!!!

...Until it doesn’t

Use “canned” software:

DeepLabCut™
A software package for animal pose estimation

Trained on labeled videos of unknown (to us) horses.
And another one: Key point detection

If the key points are well placed, what phenotypes can we extract from them?

<table>
<thead>
<tr>
<th>Frame</th>
<th>Nose x</th>
<th>Nose y</th>
<th>Nose likelihood</th>
<th>Eye x</th>
<th>Eye y</th>
<th>Eye likelihood</th>
<th>Nearknee x</th>
<th>Nearknee y</th>
<th>Nearknee likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>232.87</td>
<td>74.34</td>
<td>1.00</td>
<td>231.50</td>
<td>49.55</td>
<td>1.00</td>
<td>158.60</td>
<td>118.27</td>
<td>0.46</td>
</tr>
<tr>
<td>1</td>
<td>232.90</td>
<td>74.53</td>
<td>1.00</td>
<td>231.51</td>
<td>49.61</td>
<td>1.00</td>
<td>158.49</td>
<td>118.33</td>
<td>0.49</td>
</tr>
<tr>
<td>2</td>
<td>236.78</td>
<td>73.94</td>
<td>1.00</td>
<td>235.15</td>
<td>49.65</td>
<td>1.00</td>
<td>159.46</td>
<td>118.85</td>
<td>1.00</td>
</tr>
<tr>
<td>3</td>
<td>241.27</td>
<td>74.52</td>
<td>0.99</td>
<td>238.44</td>
<td>49.59</td>
<td>1.00</td>
<td>159.35</td>
<td>118.95</td>
<td>1.00</td>
</tr>
<tr>
<td>4</td>
<td>242.60</td>
<td>73.11</td>
<td>1.00</td>
<td>241.56</td>
<td>47.23</td>
<td>1.00</td>
<td>158.83</td>
<td>118.55</td>
<td>1.00</td>
</tr>
<tr>
<td>5</td>
<td>243.56</td>
<td>72.66</td>
<td>0.99</td>
<td>242.64</td>
<td>45.72</td>
<td>1.00</td>
<td>158.11</td>
<td>119.03</td>
<td>1.00</td>
</tr>
<tr>
<td>6</td>
<td>244.86</td>
<td>70.36</td>
<td>0.99</td>
<td>242.40</td>
<td>42.86</td>
<td>1.00</td>
<td>157.53</td>
<td>118.61</td>
<td>1.00</td>
</tr>
<tr>
<td>7</td>
<td>248.45</td>
<td>68.27</td>
<td>0.99</td>
<td>242.95</td>
<td>42.35</td>
<td>1.00</td>
<td>156.34</td>
<td>118.92</td>
<td>1.00</td>
</tr>
<tr>
<td>8</td>
<td>248.28</td>
<td>66.83</td>
<td>0.99</td>
<td>242.89</td>
<td>41.68</td>
<td>1.00</td>
<td>155.15</td>
<td>118.48</td>
<td>1.00</td>
</tr>
<tr>
<td>9</td>
<td>248.75</td>
<td>65.12</td>
<td>0.98</td>
<td>242.10</td>
<td>37.25</td>
<td>1.00</td>
<td>154.70</td>
<td>117.83</td>
<td>1.00</td>
</tr>
<tr>
<td>10</td>
<td>248.56</td>
<td>64.10</td>
<td>0.99</td>
<td>241.65</td>
<td>36.06</td>
<td>1.00</td>
<td>155.12</td>
<td>119.07</td>
<td>1.00</td>
</tr>
<tr>
<td>11</td>
<td>253.06</td>
<td>62.78</td>
<td>1.00</td>
<td>242.50</td>
<td>36.40</td>
<td>1.00</td>
<td>156.83</td>
<td>118.53</td>
<td>0.99</td>
</tr>
<tr>
<td>12</td>
<td>253.92</td>
<td>61.01</td>
<td>1.00</td>
<td>242.20</td>
<td>35.23</td>
<td>1.00</td>
<td>156.62</td>
<td>118.01</td>
<td>0.98</td>
</tr>
<tr>
<td>13</td>
<td>252.46</td>
<td>59.61</td>
<td>1.00</td>
<td>242.01</td>
<td>34.86</td>
<td>1.00</td>
<td>157.04</td>
<td>117.64</td>
<td>0.99</td>
</tr>
<tr>
<td>14</td>
<td>253.91</td>
<td>58.41</td>
<td>1.00</td>
<td>243.56</td>
<td>34.56</td>
<td>1.00</td>
<td>160.37</td>
<td>118.94</td>
<td>0.99</td>
</tr>
<tr>
<td>15</td>
<td>254.22</td>
<td>57.68</td>
<td>1.00</td>
<td>244.50</td>
<td>33.71</td>
<td>1.00</td>
<td>163.35</td>
<td>120.60</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Form follows function?
And another one: Key point detection

Pure transfer learning using horse model
Challenge 2: Animal identification

Key step in phenotyping is to link measurement to a perennial animal ID that can connect to genotypic and pedigree records.
Challenge 2: Animal identification

Uniform coat colors makes ID without markings difficult…

(in most cases)
Challenge 2: Animal identification

Recognition of feeding behaviour of pigs and determination of feeding time of each pig by a video-based deep learning method

Chen Chen a, b, Weixing Zhu a, b, c, Juan Steibel d, Janice Siegford e, Junjie Han f, Tomas Norton b, a, c

- Labelled feeding time (Day 1)
- Recognised feeding time (Day 1)
- Labelled feeding time (Day 2)
- Recognised feeding time (Day 2)
- Labelled feeding time (Day 3)
- Recognised feeding time (Day 3)
Challenge 2: Animal identification

Morphometrics?

https://www.innovationnewsnetwork.com/

Read ear tags?

Combine it with other wearables
Open set problem: individual records without ID?

Beef breeders have proposed using pooled genotyping. Those ideas can be combined with individual phenotypes without IDs.
Opportunity: better results from “old” models with new data

Social genetic effects

\[ y = X\beta + Z_d u_d + Z_c u_c + e \]

\[ e \sim N(0, I\sigma^2_e) \]
Estimation of indirect social genetic effects for skin lesion count in group-housed pigs by quantifying behavioral interactions

Belcy K Angarita, Rodolfo J C Cantet, Kaitlin E Wurtz, Carly I O’Malley, Janice M Siegford, Catherine W Ernst, Simon P Turner, Juan P Steibel


Published: 03 September 2019

Intensity of interaction

Lesion counts

Response phenotype

Ignoring behavioral interactions and social effects

Modeling social effects accounting for behavior

\[ \sim 50\% > \text{direct } h^2 \]

\[ 1.2\sigma_d^2 < \sigma_c^2 < 2.0\sigma_d^2 \]

Behavioral phenotyping?
More opportunities for improving classic models

GxE using reaction norm models is nothing new, except that… We could have an air temperature sensor mounted on the back of a cow: Measure the environment CHOSEN by the animal 😊.

Maternal effects models are not new, but now we can model the maternal effect as a function of mother-progeny distance: Separate effect of milk production from maternal attention.
Opportunity: New models for new data: Dyadic data

\[
\begin{pmatrix}
    x \\
    w \\
    y
\end{pmatrix} = f(M)
\]

MT-GLMM

Challenge: genomic prediction for traits expressed in pairs of individuals

EPD/BV/PTA

MT-GLMM

Genomic prediction for traits expressed in pairs of individuals

a) Behavioral phenotyping

Animals

b) Genomics and other phenotypes

Commercial trait phenotypes

SNP markers

MT-GLMM

MT-SRM

Dyadic or social relation model (SRM)

Multi-trait model with individual behavior and production trait

Social effect model: production trait and social interaction matrix

Predict dyadic interactions from genomic information

Joint prediction of interactions and production trait

MT-GLMM

SGE
Modeling **directional** dyadic data (probit binary model)

\[ P(y_{ijk} = 1) = \Phi(\mu_{ijk}) \]

Expected interaction

\[ \mu_{ijk} = b_0 + FE_{ijk} + g_i + r_j + d_{ijk} + sg_k \]

Random effects

\[ FE_{ijk} = sex_k + \alpha w_{ijk} + \beta w_{ik} + \delta_1 z'_{ijk} + \delta_2 z''_{ijk} + \delta_3 s'_{ijk} \]

Measure (genetic) of similarity

\[ g \sim N(0, G\sigma_g^2) \]
\[ r \sim N(0, G\sigma_r^2) \]
\[ d \sim N(0, I\sigma_d^2) \]
\[ sg \sim N(0, I\sigma_{sg}^2) \]
### Estimation of genetic parameters

**Only binary data: modeling probability of attacks**

#### Quantiles of the posterior distribution of fixed effects

<table>
<thead>
<tr>
<th>Effect</th>
<th>2.5%</th>
<th>50%</th>
<th>97.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>-0.193</td>
<td>-0.015</td>
<td>0.170</td>
</tr>
<tr>
<td>Common Nursery</td>
<td>-0.391</td>
<td>-0.310</td>
<td>-0.237</td>
</tr>
<tr>
<td>Common Litter</td>
<td>-0.199</td>
<td>0.001</td>
<td>0.212</td>
</tr>
<tr>
<td>Weight receiver</td>
<td>-0.007</td>
<td>0.000</td>
<td>0.007</td>
</tr>
<tr>
<td>Weight giver</td>
<td>0.002</td>
<td>0.013</td>
<td>0.023</td>
</tr>
<tr>
<td>Similarity</td>
<td>-0.167</td>
<td>0.198</td>
<td>0.582</td>
</tr>
</tbody>
</table>

#### Quantiles of the posterior distribution of variance components

<table>
<thead>
<tr>
<th>Component</th>
<th>2.5%</th>
<th>50%</th>
<th>97.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>0.108</td>
<td>0.168</td>
<td>0.269</td>
</tr>
<tr>
<td>Receiver</td>
<td>0.030</td>
<td>0.047</td>
<td>0.072</td>
</tr>
<tr>
<td>Giver</td>
<td>0.543</td>
<td>0.670</td>
<td>0.823</td>
</tr>
<tr>
<td>Dyad</td>
<td>0.097</td>
<td>0.165</td>
<td>0.242</td>
</tr>
<tr>
<td>% giver</td>
<td>0.281</td>
<td>0.327</td>
<td>0.372</td>
</tr>
<tr>
<td>% receiver</td>
<td>0.015</td>
<td>0.023</td>
<td>0.035</td>
</tr>
<tr>
<td>% dyad</td>
<td>0.049</td>
<td>0.08</td>
<td>0.111</td>
</tr>
</tbody>
</table>

Posterior correlation between giver and receiver was not significant.
Another example of dyadic data: Co-occurrence at the feeder

Estimation of direct and social effects of feeding duration in growing pigs using records from automatic feeding stations

Belcy K Angaria, Junjie Han, Rodolfo J C Cantet, Sarah K Chewning, Kaitlin E Wurtz, Janice M Siegfried, Catherine W Ernst, Juan Pedro Steibel


Published: 03 May 2021  Article history

*While previous research has assumed a constant social interaction value for all animals in the group, this study examined individual feeding events - using the timing and sequence in which pigs visited the feeder as a proxy for social effects:

1. Direct effects
2. Less competitive pig replaces pig at feeder
3. Pigs spend more time at the feeder
4. Social effects
   - More competitive pig replaces pig at feeder
   - More social interaction
   - More total interaction

In cases where a pig at the feeder is replaced immediately, it was assumed that a shorter meal duration of the feeding pig is a result of the social effect of the pig that follows it.

These effects should be further explored in housing with multi-space feeders, as well as expanded to incorporate genetic data and direct behavioral observations. The modeling approach used in this study can be easily applied to other systems with automatic feeding records.*
Conclusions: benefits of sensor-based phenomics for breeders

- New Traits
  - Behavior
  - Activity
  - Feed Intake
  - Gas emissions
  - Physiological

- Classic traits and environmental variables measured in more relevant contexts

- Improve genomic predictions

- Better modeling
Conclusions: Challenges in livestock phenomics

- Validate phenotyping algorithms in broad contexts
- Link phenotypes through interoperable ID
- Validate genomic predictions of novel traits

Work across disciplines, but remember we (breeders, animal scientists) understand better the sources of phenotypic variation.

Integrate data streams from multiple sensors, keep working across disciplines.

This is where our realm, let's shine 😊