Big Data in Animal Agriculture

December 5th, 2018 – ICBF ICBF & Sheep Ireland Genetics Conference – Athlone, Ireland

Roel Veerkamp, Wageningen Livestock Research
Big Data

1.79 billion monthly active users

317 million monthly active users
Hype cycle
Outline

- What is Big data?
- Big Data in Animal Agriculture?
- Key pointers to make Big Data useful
- Example projects
What is big data field?

**Volume**
- Capability to acquire, and interpret data real-time

**Velocity**
- Forms of data (text, tweets, video, drone images)

**Variety**
- Reliability and quality of data

**Veracity**
- Data whose meaning is constantly changing

**Variability**
- Expectations are huge if analysis of Big Data delivers insights and information
Outline

- What is Big data?
- Big Data in Animal Agriculture?
- Key pointers to make Big Data useful
- Example projects
Sources of Big Data - Machines

- Tractors
- Tillage equipment
- Milking robot / parlour
- Feed boxes
- ....
Sources of Big Data - Fields

- Soil analysis
- Soil type
- Soil temperature
- Ground water level
- Crop history
- .....

WAGENINGEN UNIVERSITY & RESEARCH
Sources of Big Data - Animals

- Genomic data
- Sensors / images
  - ID
  - Behaviour
  - Health
  - Position
  - Smart fencing

.....
Sources of Big Data - Environment

- Gaseous emissions
  - Methane (CH₄)
  - Ammonium (NH₃)
  - Nitrous oxide (N₂O)
- Ground/surface water
- Weather
- .....
Sources of Big Data – production chain

- Slaughter data
- Tracking & tracing
- Farm management program
- Financial accounts
- .....
Outline

- What is Big data?
- Big Data in Animal Agriculture?
- Key pointers to make Big Data useful
- Example projects
Key pointers to make Big Data useful (1)

- Making data available for the benefit of ...
  - farmer
  - consultant
  - legislation
  - technology provider

1. Secured data ownership of dairy farmers
   - Farmer is owner of data
   - Agreements with data suppliers about data availability

2. Sharing data through SDF Datahub
   - Stimulate use of sensor data and statistical data through data exchange
   - Farmer authorizes third parties for data use
   - Datahub managed by SDF Foundation
   - Open to third parties use Q4 2017

3. More efficiency and sustainability through applications
   - Applications to increase efficiency on dairy farms
   - E.g. by increasing nutrition efficiency and reducing environmental impact

van het Land, ICAR, 2017
Key pointers to make Big Data useful (2)

Make sure domain knowledge is present (and leading)

Domain knowledge
- farming
- animal health
- food production

ICT skills
- Cloud computing
- Block chain

Data analytics
- Machine learning
- Data lakes
Key pointers to make Big Data useful (3)

Other ways of working e.g. hackatons

- Multidisciplinary teams
- Combining data, software, hardware and design
- Competition
- 24 - 36 hours
- Pressure cooker setting

Computer Assisted Semen Analysis

Big data analytics & male fertility,
November 2017, Dairy Campus

Hackathon smart farming,
December 2017, Westfort, Nieuwegein
Outline

- What is Big data?
- Big Data in Animal Agriculture?
- Key pointers to make Big Data useful
- Example projects
Predict which animal is going to have best herdlife
Dairy cow’s longevity

- DNA: breeding value for 50 traits
- 72 additional phenotypic records: Pedigree, dam, own birth and calving records, test milk days, movement (transport), inseminations, viability & vitality of calves, survival status at various points, farm...
- Statistical methods: Machine learning
Better management predicting longevity

Combination of genomic breeding values and phenotypic traits important to predict survival, even after first calving.
Resilience and efficiency of animal and farms
Resilience through the theory of critical transitions

A. High resilience
   - Stable state 1
   - Stable state 2
   - Perturbation

B. Low resilience
   - Stable state 1
   - Stable state 2
   - Perturbation

Scheffer et al., 2012
New breeding trait resilience using existing data

Resilient

Not resilient

Variance in deviations

Lag-1 autocorrelation of deviations

Skewness of deviations

Disturbance
Environmental impact

Manure management

Erwin Mollenhorst, Claudia Kamphuis, Gerard Migchels
Current situation:

- Fixed phosphate application norms for crops / grassland
- 3 classes, based on P status of field
- For crops: 50 / 60 / 75 kg P$_2$O$_5$ (app. 22 / 26 / 33 kg P)

Can we predict future maize yields (= P) based on farm data and open source weather data?
Ideas developed at Hackatons

**MestHack October 2017, Dairy Campus**

*(Be)MestWijs* won the incentive prize for most market-ready result
Job de Pater (NMI), Reinier Wieringa (EZ-Dictu), Erwin Mollenhorst (WUR), Justin Steenhuis (VAA ICT), Herbert Meuleman (CRV), Claudia Kamphuis and Gerard Migchels (both WUR). Not on foto: Roel Veerman (Akkerweb)

**MaxiMy-N** won with a data- en IT-implementation to measure and show ecosystem services
Mehrab Marri (MSc), Joost Lahr, Henk Janssen, Yke van Randen, Erwin Mollenhorst (all 4 WUR) and Lucas vd Zee (UvA). In front: Gerard Ros (NMI) and Charon Zondervan (jury)

**BodemHack, May 2018, De Marke**
Norm vs model

Norm (50 kg P₂O₅ = 22 kg P)

Predicted (validation sets)

RMSE = 4.86

RMSE = 4.54

r = 0.4
Most important variables

Cropping scheme
Crop in previous year (grass/maize)

Soil status
Phosphate status field

Weather
Maximum temperature in July

Yield history
Average Pyield maize same field past 7 yrs
Summary

- More and more big data will come available
- Key pointers to success
  - Sharing data (who organises and benefits?)
  - Domain knowledge should not be forgotten
  - Domain experts should adapt
- Technology is not the silver bullet!
Thanks for your attention

Success in Big Data is not about technical tools, but connecting the tools with people and domain expertise

Roel.Veerkamp@wur.nl