AARHUS UNIVERSITY

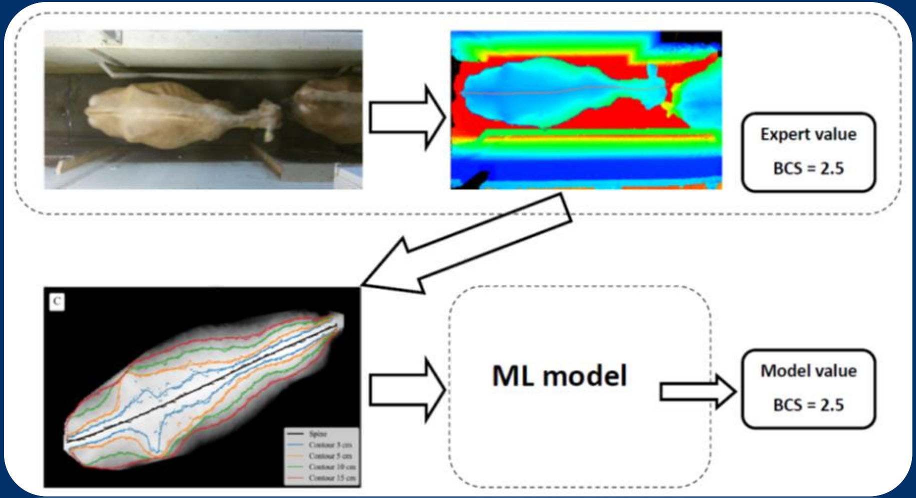
ICAR 2023 Congress
24th May 2023

RASMUS BAK STEPHANSEN
PhD FELLOW

CENTER FOR QUANTITATIVE GENETICS AND GENOMICS

Prediction of body condition in Jersey dairy cattle from 3D-images using Machine Learning techniques

R.B. Stephansen, C.I.V. Manzanilla-Pech, G. Gebreyesus, G. Sahana, J. Lassen



Presentation outline

INTRODUCTION TO THE FIELD

DATA COLLECTION

MODEL DEVELOPMENT

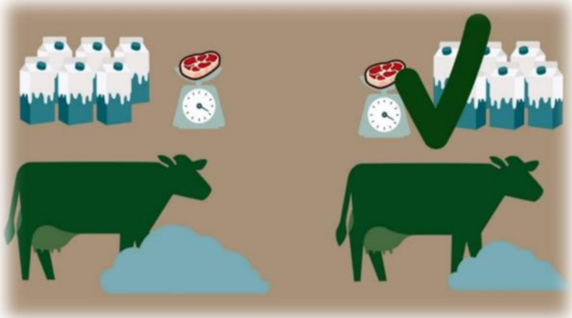
RESULTS

TAKE HOME MESSAGES

Why measure body condition?

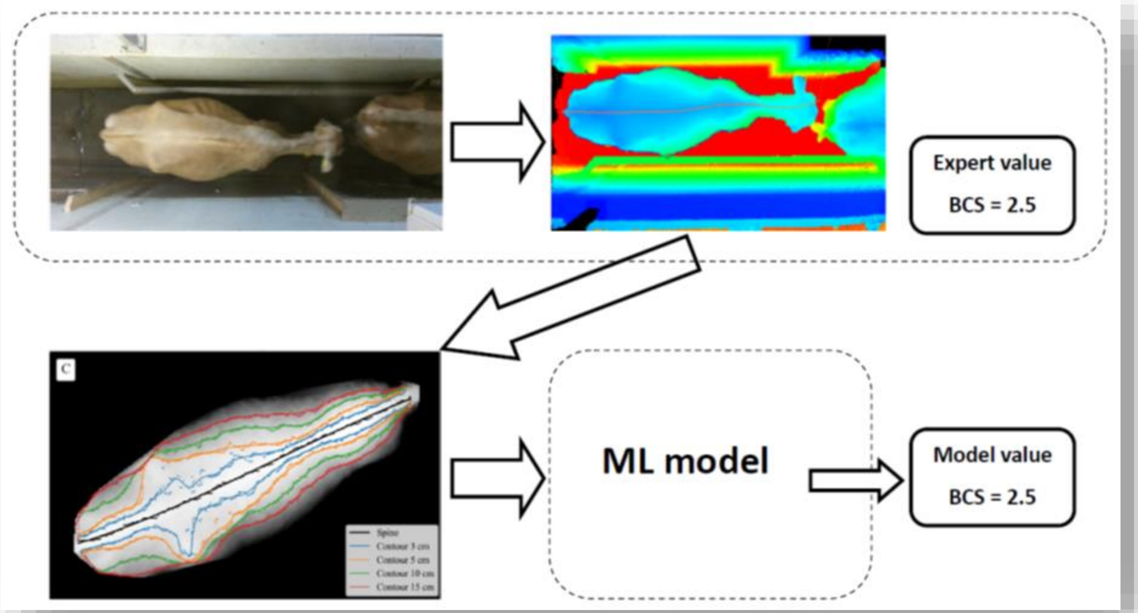
1. **Welfare** indicator
   - High **management** level of body condition gives
     - More **functional** cows (fertility, health, etc.)
     - Higher **production**

2. Important for **Metabolic Efficiency**
   - Gives **ability** to model changes in **energy**
     - **Distinguish** between **fat** & **protein**



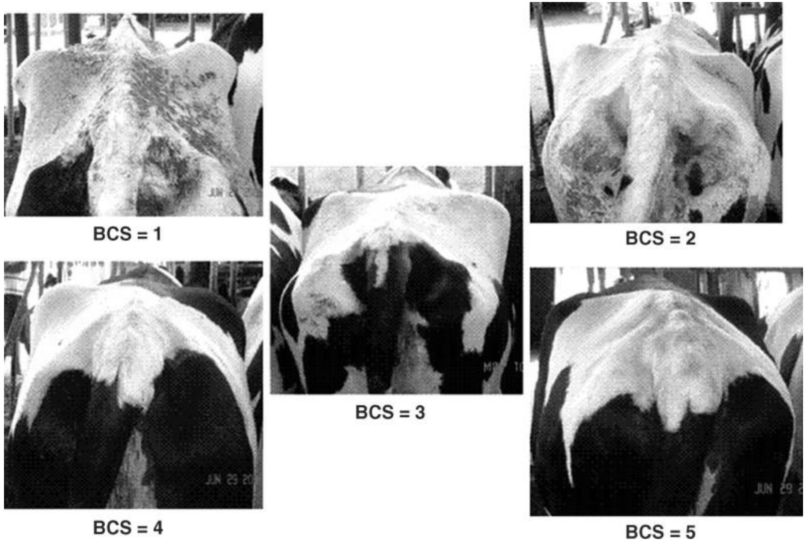
Aim of this project

Establish a **reliable** prediction of **body condition** using **3D-images** and **machine learning** techniques in Danish **Jersey** cows on **commercial farms**.



Trait definition

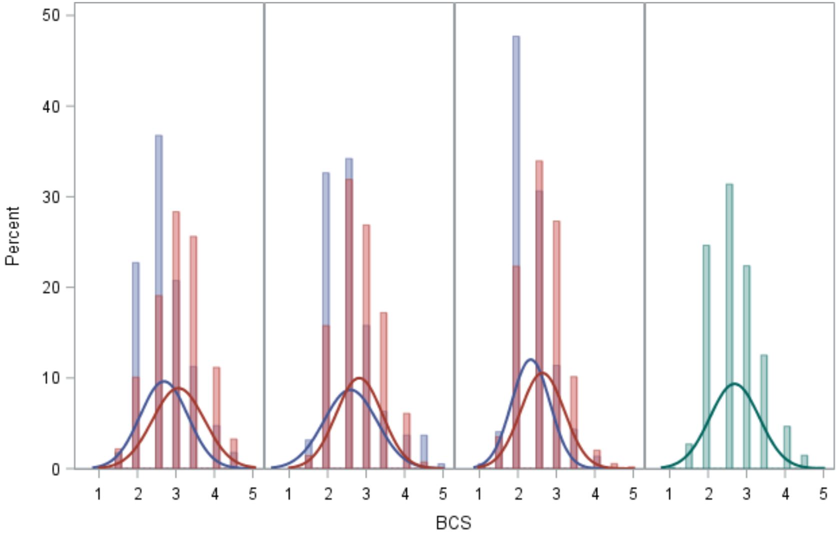
- **Body condition score (BCS)**
  - Typically scored on a **1 to 5 scale** with **0.25-unit**
  - SEGES **classifiers** scored the cows with **0.5-unit** differences

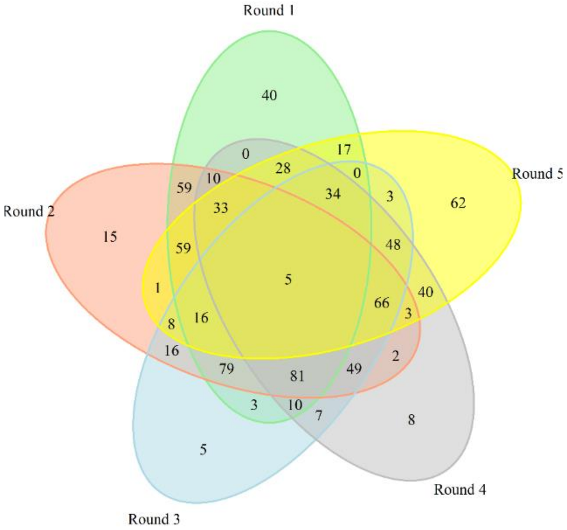


Ferguson et al., (2006)

Paper 3 – Data collection

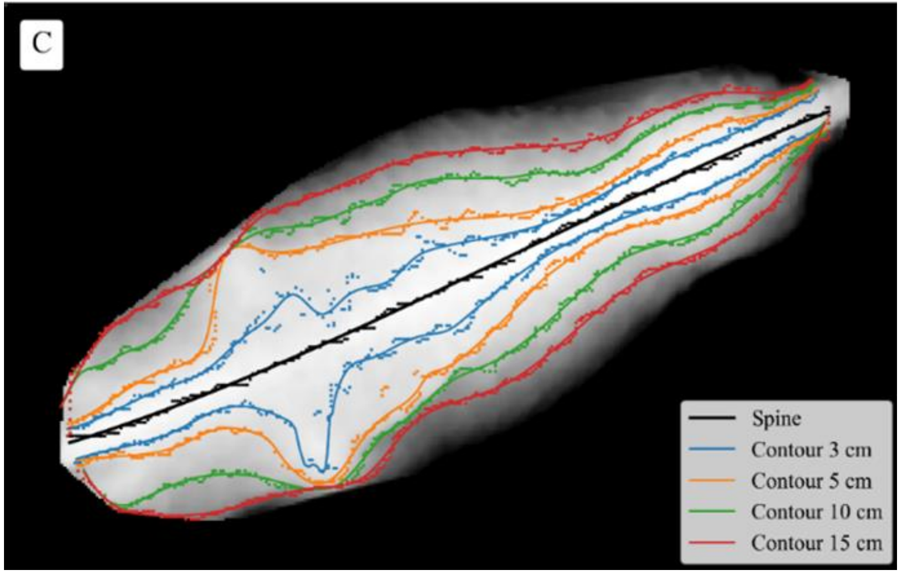
- **3** Jersey herds had **all cows** (808) scored for **BCS**
  - Scored from **December 2021** to **August 2022** (2,253 records)
  - VENN diagram for number of **Jersey** cows per round





Features for prediction

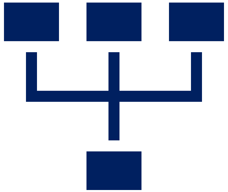
- **Features** used from the **Cattle Feed InTake** system (VikingGenetics)
  - **Mean** feature **calculated** per **round** of classification



Algorithms used

- **Machine Learning algorithms**
  - H_2O ***AutoML*** (R-studio)
    - Tested both **classification** and **regression**

- **Models tested**
  - Deep Learning (**DL**)
    - **Complex** model based on **neural networks**
  - Partial Least Square (**PLS**)
    - **Simpler** model using **dimension reduction**



Defining training and validation

- We **grouped training** and **validation** data as
  - Random **7:3 split**, clustered by cows
    - Replicated 10 times

Evaluation parameters

- Evaluation **parameters**
  - **Accuracy** (proportion of correctly assigned phenotypes) – **higher is better**
  - **F1-Score** (Combined measure of precision and recall) **– higher is better**

  - **R^2** (Coefficient of determination) – **higher is better**
  - **RMSE** (root of mean squared error) – **lower is better**

- Evaluation parameters **assessed** on
  - **Exact** phenotype
  - **0.5-unit** deviation (**DV**)

Results – DL Classification model

- **DeepLearning** results
  - **Accuracy**
    - Exact: **48.1** (45.9-50.7)
    - 0.5-unit DV: **93.5** (92.7-95.3)
  - Rodriguez Alvarez et al. (2019)
    - Exact: **41**
    - 0.5-unit DV: **97**
  - Shi et al., (2023)
    - Exact: **49**
    - 0.5-unit DV: **96**

	F1-Score	
BCS	**Exact**	**0.5-unit DEV**
1.5	3	39
2.0	59	98
2.5	55	96
3.0	36	94
3.5	42	85
4.0	9	81
4.5	4	13
Weighted average	46	91

Results – Regression models

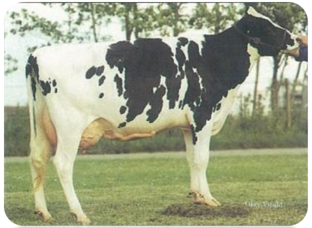
- **Accuracy** based on **rounded** predicted **phenotype** from **regression** models

BCS	PLS		DL	
	Exact	0.5 range	Exact	0.5 range
Accuracy	51.2	96.1	52.0	95.5
R^2	0.67		0.66	
RMSE	0.31		0.29	

Take-home messages

- Accuracy from **tested** models show
  - **Similar** validation **accuracies** as **Holstein** studies

- **PLS** achieve **similar** validation results as **DL**
  - Reduce **computational requirements**
- It is **feasible** to predict **BCS** in **Jersey** from 3D-images
  - **Next** step is to **build** a model for **Holstein**



Acknowledgements

CENTER FOR QUANTITATIVE GENETICS AND GENOMICS QGG



VIKINGGENETICS innovative breeding

SEGES INNOVATION



Innovationsfonden

Thank you for listening