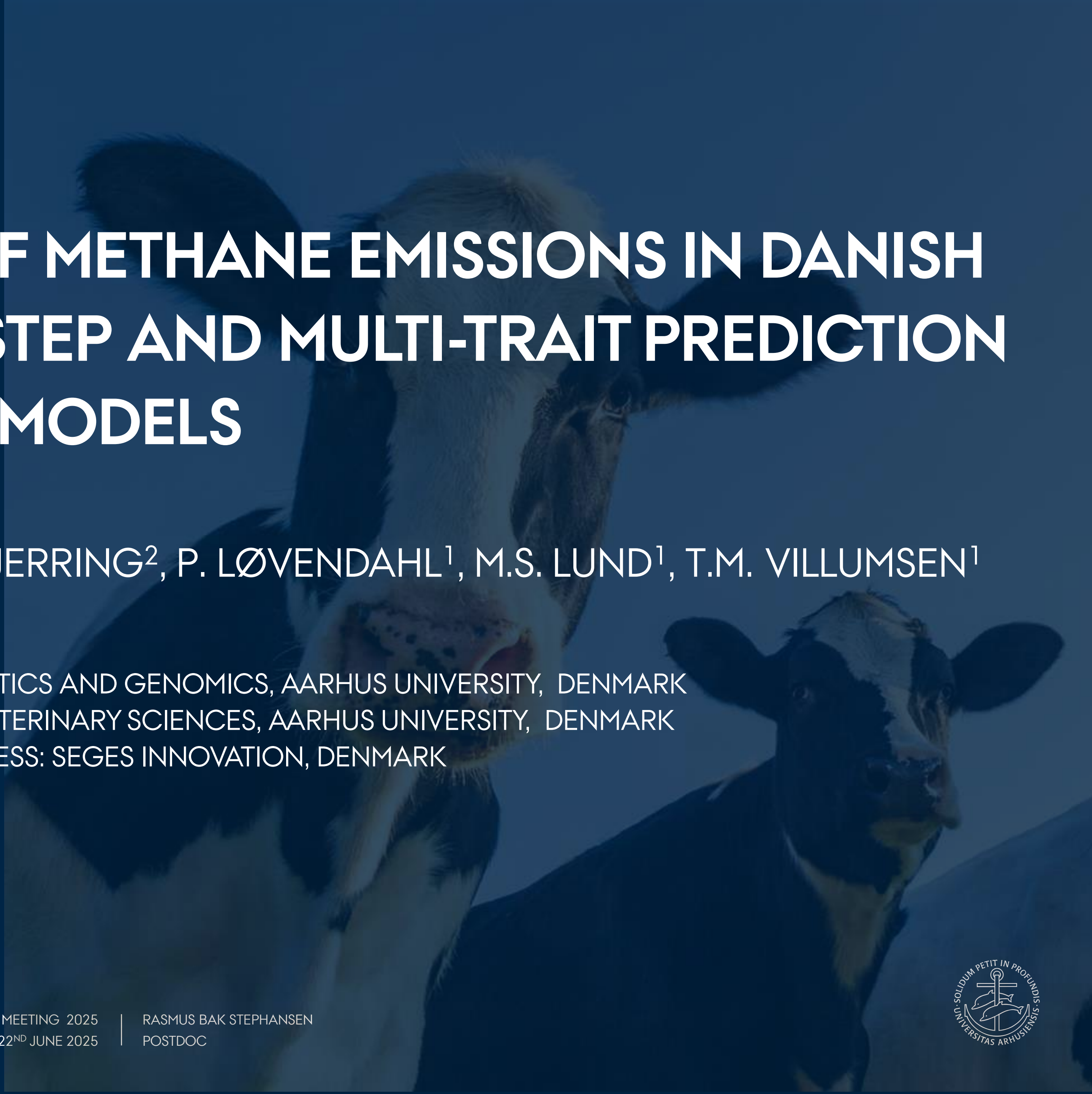


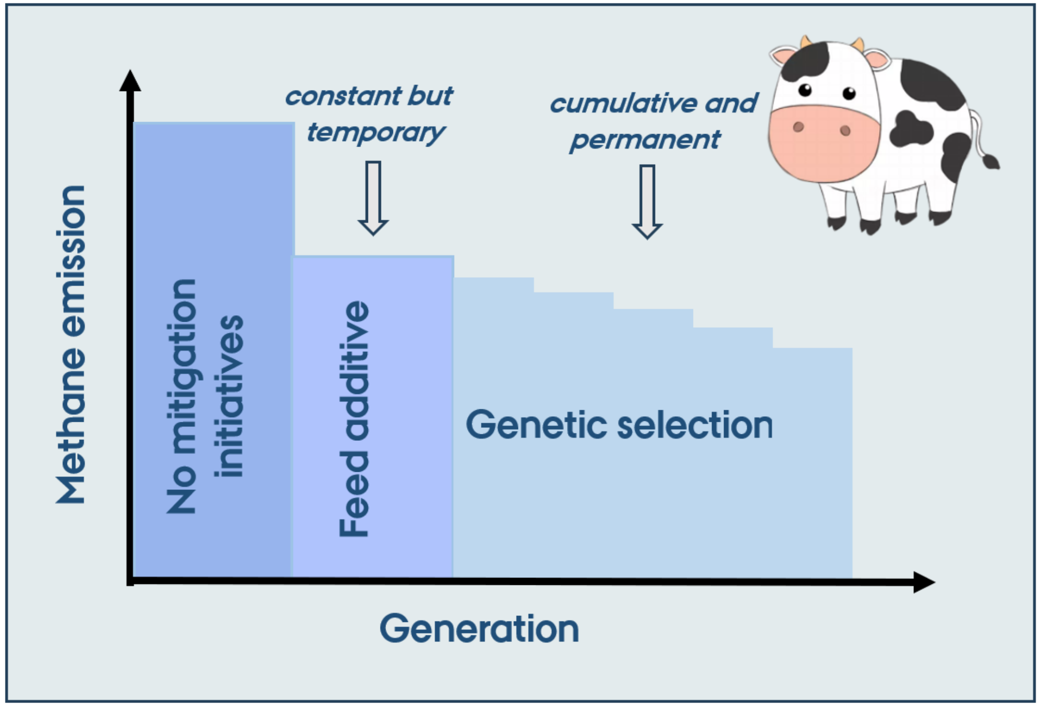
# GENOMIC PREDICTION OF METHANE EMISSIONS IN DANISH HOLSTEIN USING SINGLE STEP AND MULTI-TRAIT PREDICTION MODELS

H. SCHNEIDER[1,3], R.B. STEPHANSEN[1], M. BJERRING[2], P. LØVENDAHL[1], M.S. LUND[1], T.M. VILLUMSEN[1]

[1]CENTER FOR QUANTITATIVE GENETICS AND GENOMICS, AARHUS UNIVERSITY, DENMARK
[2]DEPARTMENT OF ANIMAL AND VETERINARY SCIENCES, AARHUS UNIVERSITY, DENMARK
[3]PRESENT ADDRESS: SEGES INNOVATION, DENMARK

# BACKGROUND

→ Mitigation strategies to **reduce** methane emissions urgently required (e.g., feed additives, genetic selection,…)



Generated with Microsoft Copilot

# BACKGROUND

→ Mitigation strategies to **reduce** methane emissions urgently required (e.g., feed additives, genetic selection,...)

🇩🇰 first breeding values for methane emissions published in **May 2025**

Generated with Microsoft Copilot

# BACKGROUND

**Problem:** Still primarily **small datasets** (limits prediction accuracy & genetic progress)

Prediction accuracies have earlier been shown to benefit from

1. Simultaneous use of genotypic, phenotypic and pedigree information: single step genomic prediction (**ssGBLUP**)
2. Exploiting genetic information from correlated traits: **multi-trait prediction**

# AIM OF THIS STUDY

Evaluate **predictive ability** for methane concentrations (MeC) of

1. pedigree-based BLUP (**pBLUP**) and **ssGBLUP**
2. **univariate** and **multi-trait** models using fat yield (FY) and energy corrected milk yield (ECM) as predictor traits

→ Data splitted into growing (primiparous, PP) and matured (multiparous, MP) cows to account for different correlation structures

# DATA

Danish **Holstein** cows from 15 commercial herds

- 1,744 PP cows with 182,288 daily observations
- 2,989 MP (2nd to 8th parity) cows with 424,888 daily observations

→ between March 2021 and December 2024

# MATERIAL

- Pedigree (47,383 animals) and genotypes (46,342 SNPs) provided by Danish Cattle Database (SEGES Innovation, Skejby, Denmark)

# MATERIAL

- Pedigree (47,383 animals) and genotypes (46,342 SNPs) provided by Danish Cattle Database (SEGES Innovation, Skejby, Denmark)

- MeC recorded using sniffers

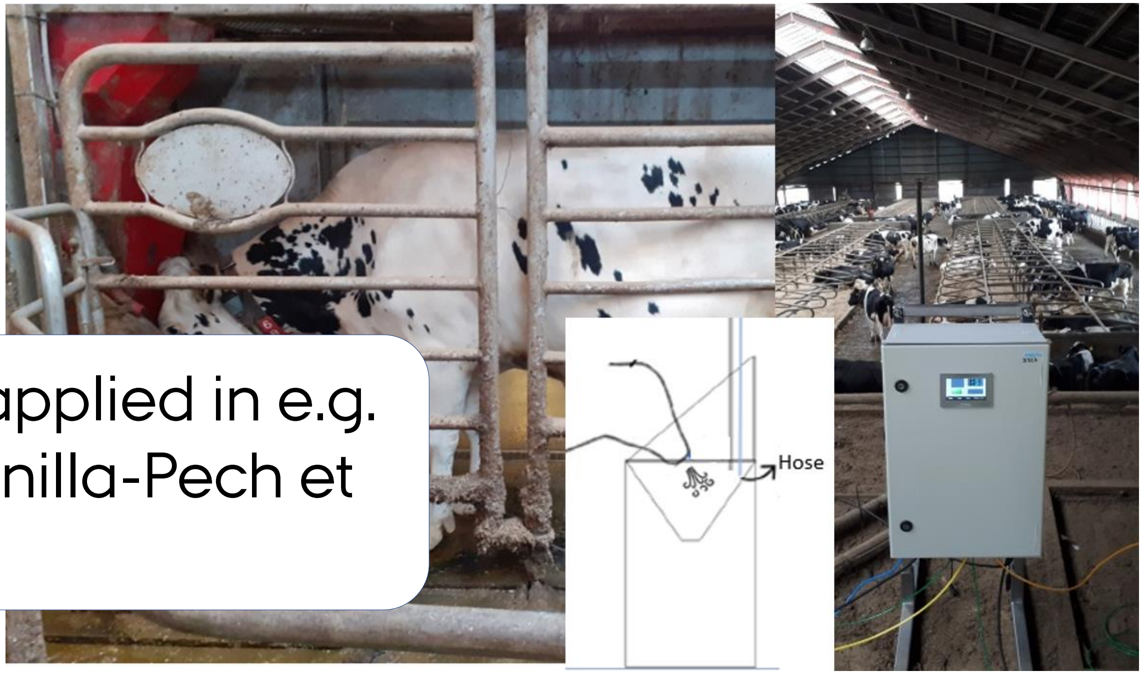
  (Guardian NG, Edinburgh Sensors, Livingston, UK)

# MATERIAL

- Pedigree (47,383 animals) and genotypes (46,342 SNPs) provided by Danish Cattle Database (SEGES Innovation, Skejby, Denmark)

- MeC recorded using sniffers

  (Guardian NG, Edinburgh Sensors, Livingston, UK)



Sniffer technology has already been applied in e.g. Lassen and Løvendahl, (2016), Manzanilla-Pech et al., (2020), Løvendahl et al., (2024)

# MATERIAL

- Pedigree and genotypes provided by Danish Cattle Database (SEGES Innovation, Skejby, Denmark)
- MeC recorded using sniffers (Guardian NG, Edinburgh Sensors, Livingston, UK)
- Milk production data obtained from milking robots (milk yield, MY) and the Danish Cattle Database (milk components). ECM was calculated as

$$ECM\ (kg) = 0.25 * MY\ (kg) + 12.2 * FY\ (kg) + 7.7 * PY\ (kg)$$

(Sjaunja et al., 1991)

# METHODS

**Pedigree-based repeatability model**

1. **Variance component estimation** for PP and MP cows, separately, using AI-REML algorithm in the DMU software (Madsen and Jensen, 2014)

$$y = X\beta + Za + Wpe + Ie$$

| | |
|---|---|
| $y$ | vector of phenotypic observations for MeC, ECM, and FY |
| $\beta$ | vector of fixed effects (WIM; AFC for PP cows, parity for MP cows; herd-year-season (HYS) × AMS × sniffer for MeC, HYS for ECM and FY) |
| $X$ | corresponding incidence matrix |

# METHODS

**Pedigree-based repeatability model**

1. **Variance component estimation** for PP and MP cows, separately, using AI-REML algorithm in the DMU software (Madsen and Jensen, 2014)

$$y = X\beta + Za + Wpe + Ie$$

$a$ random additive genetic effect with $a \sim N(0, A\sigma_a^2)$ and corresponding incidence matrix $Z$

$pe$ random permanent environmental effect with $pe \sim N\left(0, I\sigma_{pe}^2\right)$ and corresponding matrix $W$

$e$ residual term with $e \sim N(0, I\sigma_e^2)$

# METHODS

**2. GEBV estimation for MeC** using DMU and the same fixed and random effects as in the variance component estimation

→ Different pBLUP and ssGBLUP scenarios

→ Different univariate and multi-trait scenarios

# METHODS

## GEBV estimation

| Traits included in GEBV estimation | Type of analysis | Information included in validation population | Information included in reference population |
|---|---|---|---|
| **MeC** | Univariate | - | MeC |
| **MeC-FY** | Bivariate | FY | MeC, FY |
| | | - | |
| **MeC-ECM** | Bivariate | ECM | MeC, ECM |
| | | - | |
| **MeC-ECM-FY** | Trivariate | ECM, FY | MeC, ECM, FY |

# METHODS

**2. GEBV estimation for MeC -** Calculation of inverse of the H matrix for ssGBLUP (Aguilar et al., 2010; Christensen and Lund, 2010)

$$\boldsymbol{H}^{-1} = \boldsymbol{A}^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & (\omega\boldsymbol{G} + (1-\omega)\boldsymbol{A}_{22})^{-1} - \boldsymbol{A}_{22}^{-1} \end{bmatrix}$$

$\boldsymbol{G}$ genomic relationship matrix (VanRaden, 2008); computed using the invgmatrix software (Su and Madsen, 2011)

$\boldsymbol{A}_{22}$ pedigree relationship matrix for genotyped animals

ω weight assigned to the genomic information (0.8)

# METHODS

**Accuracy calculation** using a 10-fold cross-validation set-up (by sire)

1. Adjusted phenotypes for MeC computed as estimated genetic + permanent environmental effects from the full dataset

# METHODS

Formula adapted from Mrode (2013), accounting for repeated records

**Accuracy calculation** using a 10-fold cross-validation set-up (by sire)

2. Accuracies per cross-validation group calculated as

$$\text{Accuracy} = \frac{r}{\sqrt{\frac{nh^2}{\sqrt{1+(n-1)t}}}}$$

| | |
|---|---|
| $r$ | correlation between the adjusted phenotype and GEBV |
| $n$ | average amount of repeated records for each animal |
| $h^2$ $(t)$ | heritability (repeatability) of MeC |

AARHUS UNIVERSITY
TECHNICAL SCIENCES

# METHODS

**Accuracy calculation** using a 10-fold cross-validation set-up (by sire)

3. Accuracy for each scenario calculated as the average of all cross-validation groups
4. Corresponding standard errors were obtained by dividing the standard deviation of accuracies across cross-validation groups by the square root of the number of validation groups (10)

# RESULTS & DISCUSSION

Variance component estimates

| Trait | Primiparous | | | Multiparous | | |
|---|---|---|---|---|---|---|
| | $h^2$ | $t$ | $r_g$ with MeC | $h^2$ | $t$ | $r_g$ with MeC |
| **MeC** | 0.17 (0.03) | 0.32 (0.01) | | 0.17 (0.02) | 0.32 (0.01) | |
| **ECM** | 0.38 (0.06) | 0.81 (0.01) | 0.15 (0.13) | 0.24 (0.03) | 0.74 (0.01) | 0.41 (0.09) |
| **FY** | 0.31 (0.06) | 0.74 (0.01) | 0.18 (0.13) | 0.20 (0.03) | 0.65 (0.01) | 0.37 (0.09) |

$r_g$ for MeC between PP and MP cows **0.78** (SE: 0.12)

→ Only 580 cows in common between both groups

# RESULTS & DISCUSSION

Variance component estimation

| Trait | Primiparous | | | Multiparous | | |
|---|---|---|---|---|---|---|
| | $h^2$ | $t$ | $r_g$ with MeC | $h^2$ | $t$ | $r_g$ with MeC |
| **MeC** | 0.17 (0.03) | 0.32 (0.01) | | 0.17 (0.02) | 0.32 (0.01) | |
| **ECM** | 0.38 (0.06) | 0.81 (0.01) | 0.15 (0.13) | 0.24 (0.03) | 0.74 (0.01) | 0.41 (0.09) |
| **FY** | 0.31 (0.06) | 0.74 (0.01) | 0.18 (0.13) | 0.20 (0.03) | 0.65 (0.01) | 0.37 (0.09) |

MeC: $h^2$ 0.11 to 0.18 and $t$ 0.27 to 0.41 (van Engelen et al., 2018; van Breukelen et al., 2024)

ECM: $h^2$ 0.26 to 0.37 (Li et al., 2018)

# RESULTS & DISCUSSION

**Positive $r_g$ is unfavorable!** ⚠

→ Selecting for lower methane emissions leads to reduced milk production

## Variance component estimation

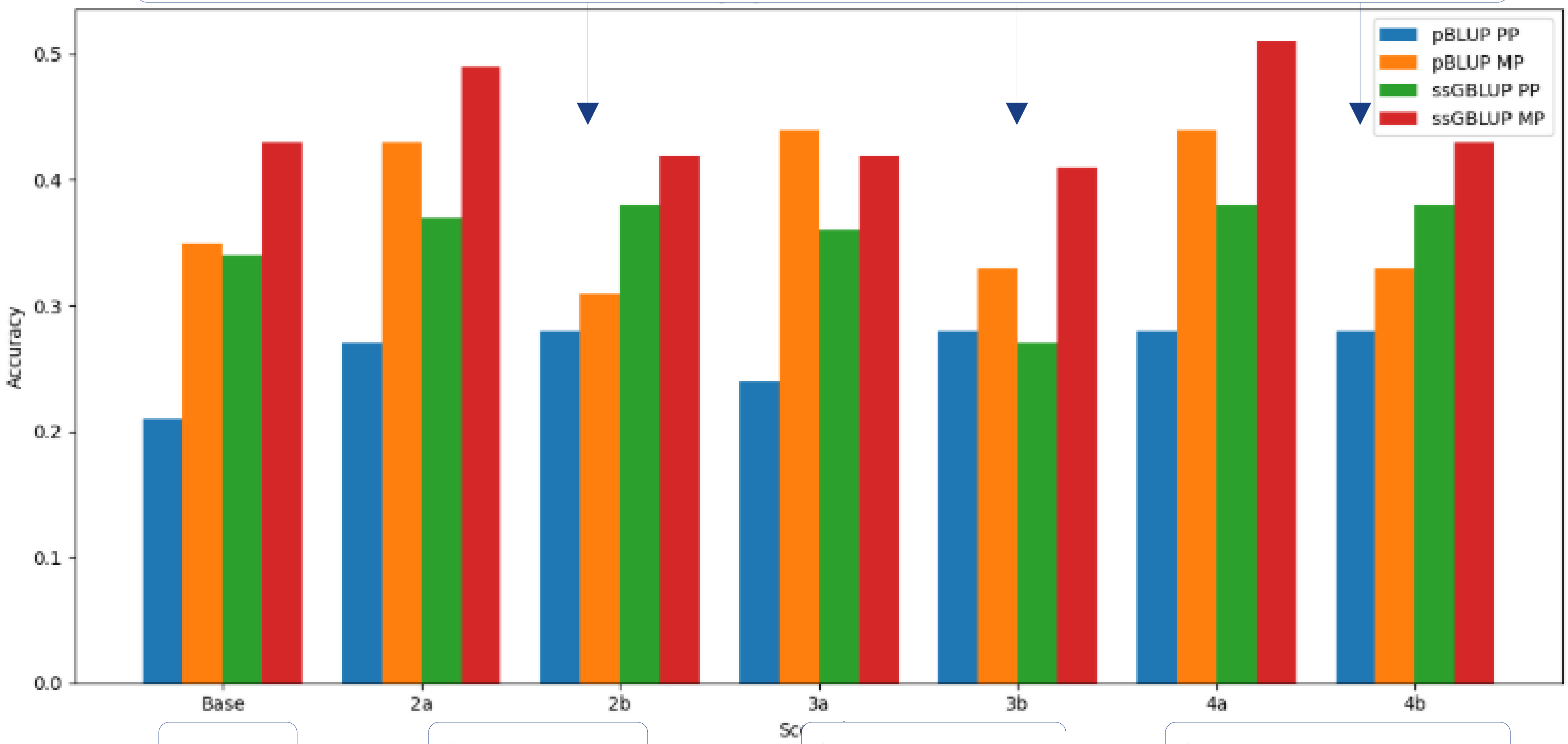
| Trait | Primiparous | | | Multiparous | | |
|---|---|---|---|---|---|---|
| | $h^2$ | $t$ | $r_g$ with MeC | $h^2$ | $t$ | $r_g$ with MeC |
| **MeC** | 0.17 (0.03) | 0.32 (0.01) | | 0.17 (0.02) | 0.32 (0.01) | |
| **ECM** | 0.38 (0.06) | 0.81 (0.01) | **0.15 (0.13)** | 0.24 (0.03) | 0.74 (0.01) | **0.41 (0.09)** |
| **FY** | 0.31 (0.06) | 0.74 (0.01) | **0.18 (0.13)** | 0.20 (0.03) | 0.65 (0.01) | **0.37 (0.09)** |

$r_g$ 0.35 between MeC and ECM (Manzanilla-Pech et al., 2022)

Correlation of 0.27 between GEBV for MeC and FY (Lopez-Paredes et al., 2020)

# RESULTS & DISCUSSION

No information for ECM and FY in reference and validation population



MeC

MeC - FY

MeC - ECM

MeC – ECM - FY

# RESULTS & DISCUSSION

**Accuracies of GEBV** obtained using

- ssGBLUP were mostly higher than those obtained using pBLUP, in line with previous results in dairy cattle (Hayes and Goddard, 2008; VanRaden et al., 2009)

# RESULTS & DISCUSSION

**Accuracies of GEBV** obtained using

- ssGBLUP were mostly higher than those obtained using pBLUP, in line with previous results in dairy cattle (Hayes and Goddard, 2008; VanRaden et al., 2009)
- multi-trait prediction scenarios mostly higher than univariate scenarios, in line with e.g. Tsuruta et al. (2011) for linear type traits.

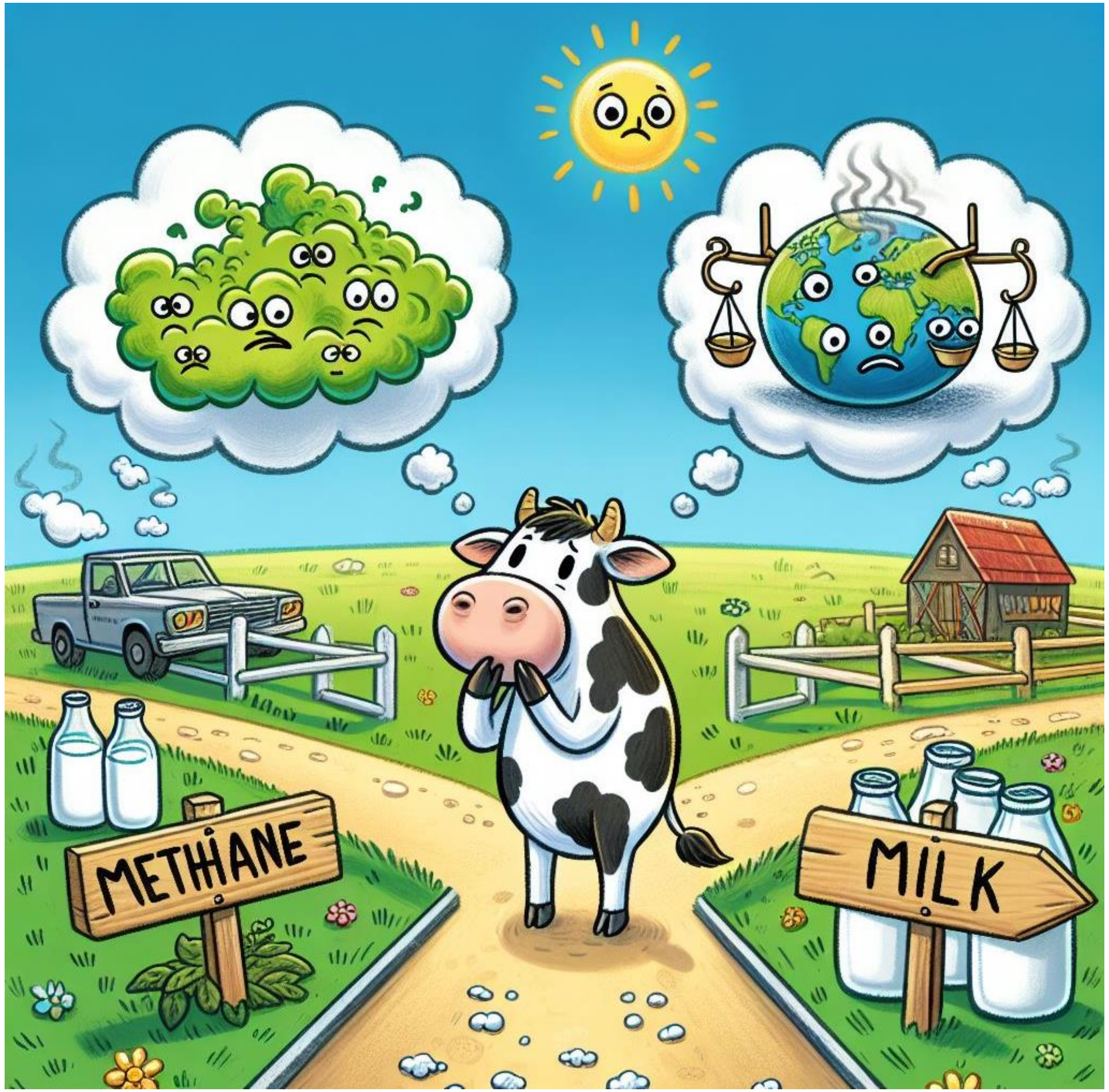
→ Most improvement when phenotypic information for ECM and FY available in reference and validation populations, has also earlier been shown (Pszczola et al., 2013)

# CONCLUSION

☺ ssGBLUP and multi-trait models with ECM and FY records can improve accuracy of MeC breeding values

- However: genetic correlations between FY, ECM and MeC is unfavorable

→ Genetic gain in MeC based on these models might be at the expense of milk production



Generated with Microsoft Copilot

# PERSPECTIVES

Further efforts are needed to

- Record methane emissions in more animals

→ Supported by (international) projects like **GMG**

- Develop a methane emission trait that is genetically independent from economically important, correlated traits like FY or ECM

- Design a multi-trait selection index including all economically important traits

# ACKNOWLEDGEMENTS

This study is part of GUDP-funded project "ONIMIT: On-farm monitoring of methane from dairy cows is a prerequisite for the implementation of mitigation strategies on animal, farm and national level"

Ministeriet for Fødevarer,
Landbrug og Fiskeri
gudp

AARHUS
UNIVERSITY