



Re-Livestock

RESILIENT FARMING SYSTEMS

Estimation of genetic parameters for sniffer methane emissions

Birgit Gredler-Grandl and Coralía Manzanilla-Pech



WAGENINGEN
UNIVERSITY & RESEARCH

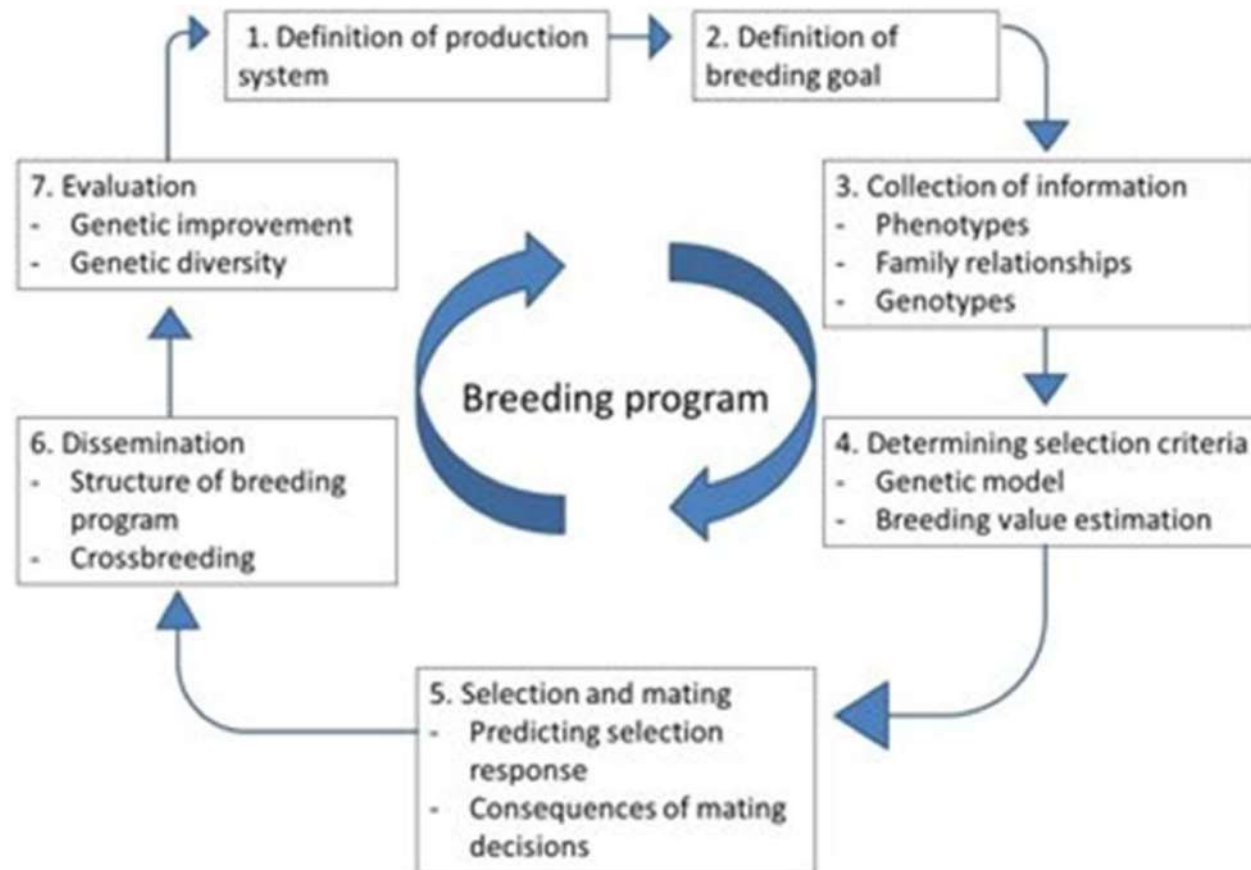


Zaragoza, February 26th, 2025

Overview

- Genetic parameters
- Linear Mixed models
- Random regression models

Seven steps in a breeding program



Heritability

Heritability = proportion of differences in performance for a trait that are attributable to differences in breeding value for the trait

- Ratio between 0 and 1.
- Population-specific.
- No constant value.
- Strongly depends on how variable the environment is and how well it can be recorded

Heritability

$$h^2 = \frac{\sigma_A^2}{\sigma_P^2}$$

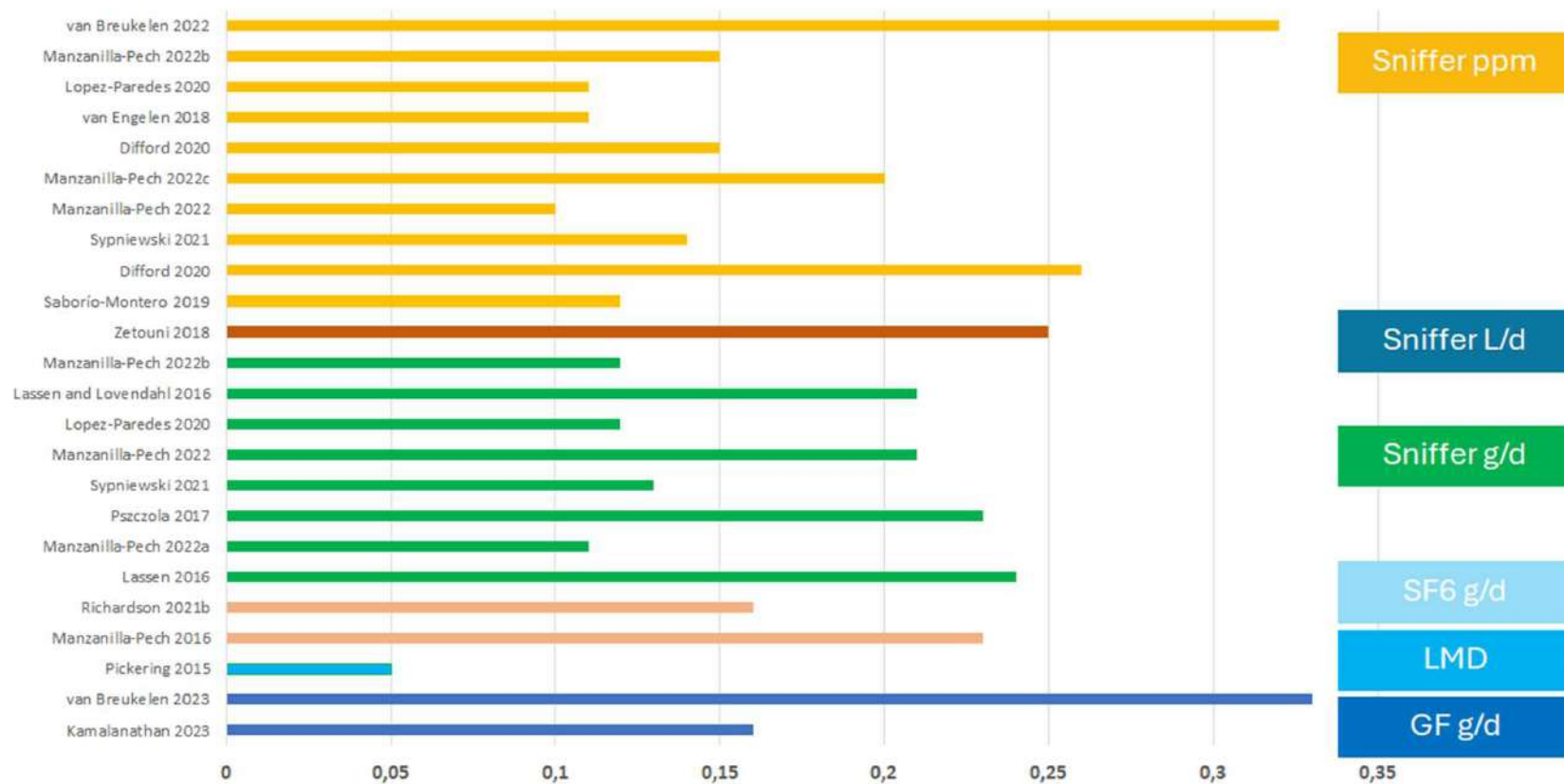
Additive genetic variance

phenotypic Variance (genetic + environment)

Poor environmental recording:
Missing inseminations, survey of
traits, ... reduce heritability!

Estimation of variance
components

Heritability of methane emission



Repeatability

- Measure of strength of the relationship between repeated records for a trait in a population
- Not fixed, population specific

$$r = \frac{\sigma_A^2 + \sigma_{PE}^2}{\sigma_A^2 + \sigma_{PE}^2 + \sigma_E^2}$$

Genetic parameters for repeatedly recorded enteric methane concentrations of dairy cows

A. E. van Breukelen,* M. A. Aldridge, R. F. Veerkamp, and Y. de Haas
Wageningen University & Research Animal Breeding and Genomics, 6700 AH Wageningen, the Netherlands

Table 4. Heritability (h^2) and repeatability (t) for mean methane (CH_4) emissions, median CH_4 emissions, log CH_4 emissions, mean carbon dioxide (CO_2) emissions, median CO_2 emissions, log CO_2 emissions, and CH_4/CO_2 ratio, averaged per visit and per week ($\pm \text{SE}$)

Item	Visit		Week	
	h^2	t	h^2	t
Mean CH_4 (ppm)	0.13 ± 0.01	0.30 ± 0.01	0.32 ± 0.03	0.68 ± 0.01
Median CH_4 (ppm)	0.13 ± 0.01	0.29 ± 0.01	0.32 ± 0.03	0.68 ± 0.01
Log CH_4	0.09 ± 0.01	0.18 ± 0.01	0.23 ± 0.03	0.65 ± 0.01
Mean CO_2 (ppm)	0.16 ± 0.02	0.36 ± 0.01	0.33 ± 0.03	0.71 ± 0.01
Median CO_2 (ppm)	0.16 ± 0.01	0.35 ± 0.01	0.34 ± 0.03	0.71 ± 0.01
Log CO_2	0.07 ± 0.01	0.20 ± 0.01	0.20 ± 0.03	0.57 ± 0.01
CH_4/CO_2	$0.01 \pm <0.01$	$0.08 \pm <0.01$	0.02 ± 0.01	0.15 ± 0.01

Genetic correlation

The correlation coefficient (r) indicates the extent to which two traits are genetically related

- Function of the covariance of the variables and their standard deviations; “standardized” covariance

$$r_g = \frac{\sigma_{A_{12}}}{\sqrt{\sigma_{A_1}^2 \cdot \sigma_{A_2}^2}}$$

Linear models – core of analyses in animal breeding

A model tries to describe reality, e.g. how a cow's milk yield comes about. The model should include all causal effects.

Model of performance:

$$\mathbf{P}_{\text{henotype}} = \mathbf{G}_{\text{enotype}} + \mathbf{E}_{\text{nvironment}}$$

Linear models

A model tries to describe reality, e.g. how a cow's milk yield comes about. The model should include all causal effects.

$$y = X\hat{b} + e$$

Linear model with **fixed** causal effects

$$y = Z\hat{u} + e$$

Linear model with **random** causal effects

$$y = X\hat{b} + Z\hat{u} + e$$

Mixed linear model with **fixed** and **random** causal effects (Mixed Models)

Fixed vs random effects

fixed	random
Number of effect levels low	Number of effect levels high
Interpretation – examined effect levels	Interpretation – population
Sampling – targeted	Sampling – random
Target variable: comparison of means	Target variable: estimation of variance components
Breed, gender, feed ration,	Breeding values (animal effects)

BLUP

BEST: Correlation between the true (a) and the estimated (\hat{a}) breeding value is maximized (minimization of estimation error, solutions have minimal error variance - minimum variance). $E(\hat{a} - a)^2 \Rightarrow$ Minimum!

LINEAR: The estimators for the breeding values are linear functions of the observations

UNBIASED: The expected values of the solutions correspond to the true values ($E(a) = E(\hat{a})$), meaning no systematic overestimation or underestimation (expectation consistency).

PREDICTION: Prediction of random effects

Mixed model equations

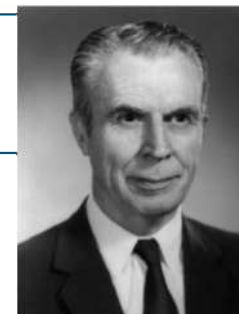
$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}\mathbf{u} + \mathbf{e}$$

Henderson's Mixed Model Equations

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{bmatrix} * \begin{bmatrix} \mathbf{b} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{bmatrix}$$

Often, \mathbf{R}^{-1} is canceled out

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \mathbf{A}^{-1}\lambda \end{bmatrix} * \begin{bmatrix} \mathbf{b} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{bmatrix}$$



Charles Roy Henderson

A Biographical Memoir
by L. Dale Van Vleck

Annahmen:

$$\text{var}(\mathbf{e}) = \mathbf{I}\sigma_e^2 = \mathbf{R}$$

$$\text{var}(\mathbf{u}) = \mathbf{A}\sigma_u^2 = \mathbf{G}$$

$$\text{cov}(\mathbf{u}, \mathbf{e}) = \text{cov}(\mathbf{e}, \mathbf{u}) = 0$$

$$\lambda = \frac{\sigma_e^2}{\sigma_u^2}$$

Repeatability models vs Random regression models

Goal: Estimate genetic merit over time

Applications in Animal Breeding

- Dairy Cattle: Milk yield over lactation
- Pigs: Growth rate modeling
- Poultry: Egg production over time

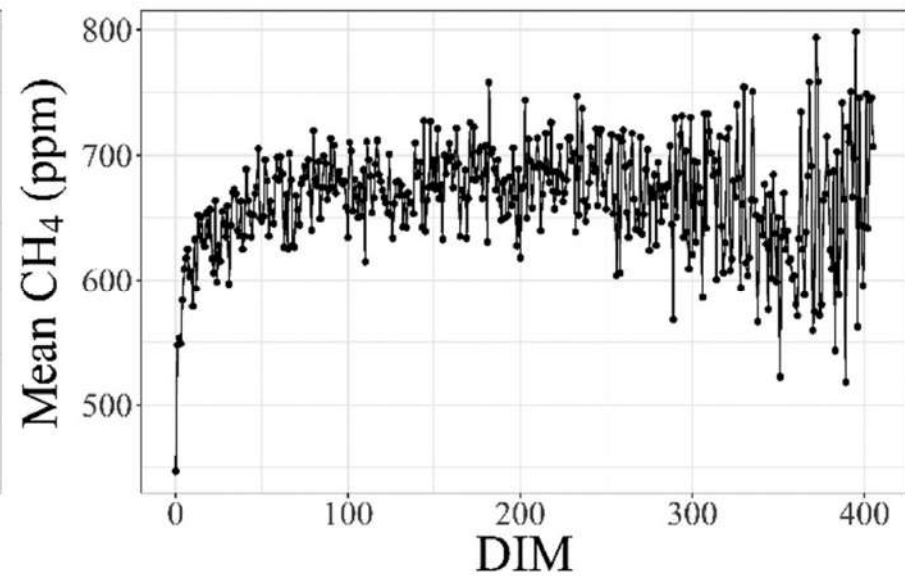
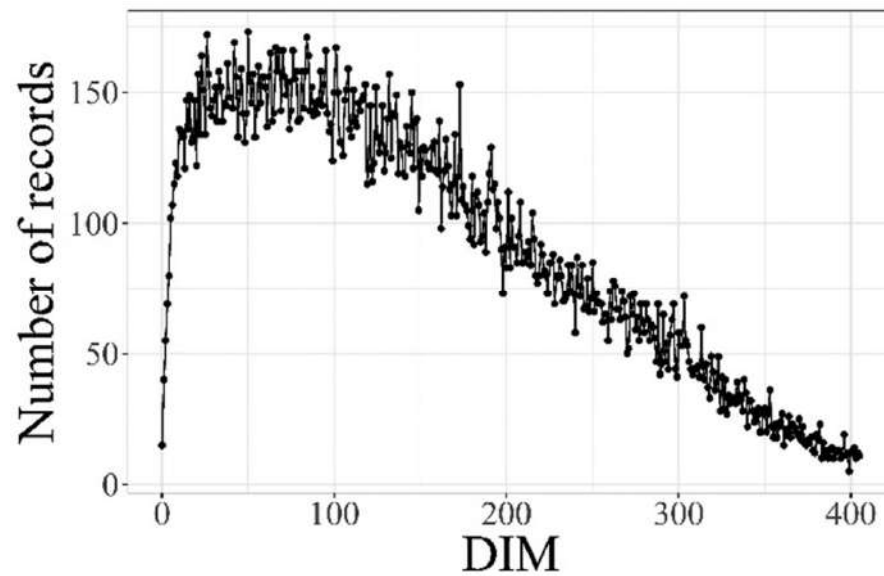
Traditional models vs. Random Regression Models (RRMs)

Why use RRM? Capture continuous changes in traits

More flexible than repeatability models

Handles unbalanced data

Longitudinal data



Van Breukelen et al., 2024

Random regression models

Goal: Estimate genetic merit over time

Applications in Animal Breeding

- Dairy Cattle: Milk yield over lactation
- Pigs: Growth rate modeling
- Poultry: Egg production over time

Traditional models vs. Random Regression Models (RRMs)

Why use RRM? Capture genetic continuous changes in traits

More flexible than repeatability models

Handles unbalanced data

Random regression models

- Polynomials
 - **Legendre polynomials**
- Splines
 - Flexible but require knot selection
 - less sensitive to quantity of data (e.g. end of lactation)

Lactation curves - Polynomials

- dependent on lactation day: $y = b_0 + b_1*t + b_2*t^2 + b_3*t^3$
- conventional polynomials, the time-dependent covariates grow very large quickly: e.g., $t = 100 \Rightarrow t^4 = 100,000,000$ (multicollinearity)
- → orthogonal polynomials, e.g. Legendre polynomials
- lactation day transformation -1 to 1

Why Use Legendre Polynomials?

Orthogonal, reducing collinearity

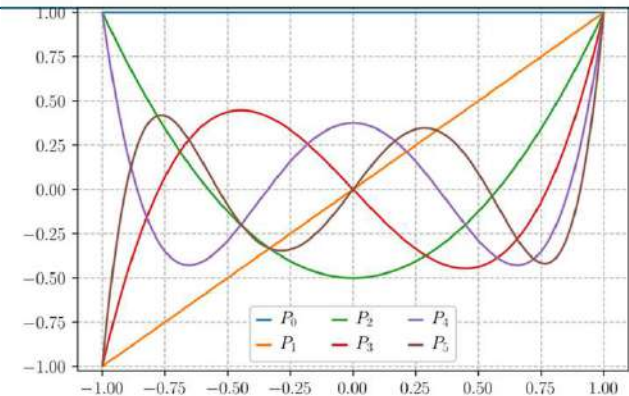
Flexible with minimal parameters

Efficiently captures changes over time

Legendre vs. Standard Polynomials:

Standard polynomials: susceptible to collinearity

Legendre: orthogonality, better numerical stability



$$P_0(x) = 1$$

$$P_1(x) = x$$

$$P_2(x) = \frac{1}{2}(3x^2 - 1)$$

$$P_3(x) = \frac{1}{2}(5x^3 - 3x)$$

$$P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3)$$

$$P_5(x) = \frac{1}{8}(63x^5 - 70x^3 + 15x)$$

$$P_6(x) = \frac{1}{16}(231x^6 - 315x^4 + 105x^2 - 5)$$

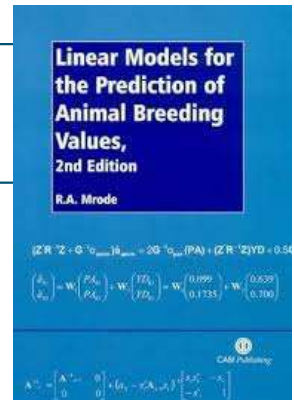
$$P_7(x) = \frac{1}{16}(429x^7 - 693x^5 + 315x^3 - 35x)$$

Random regression testday model

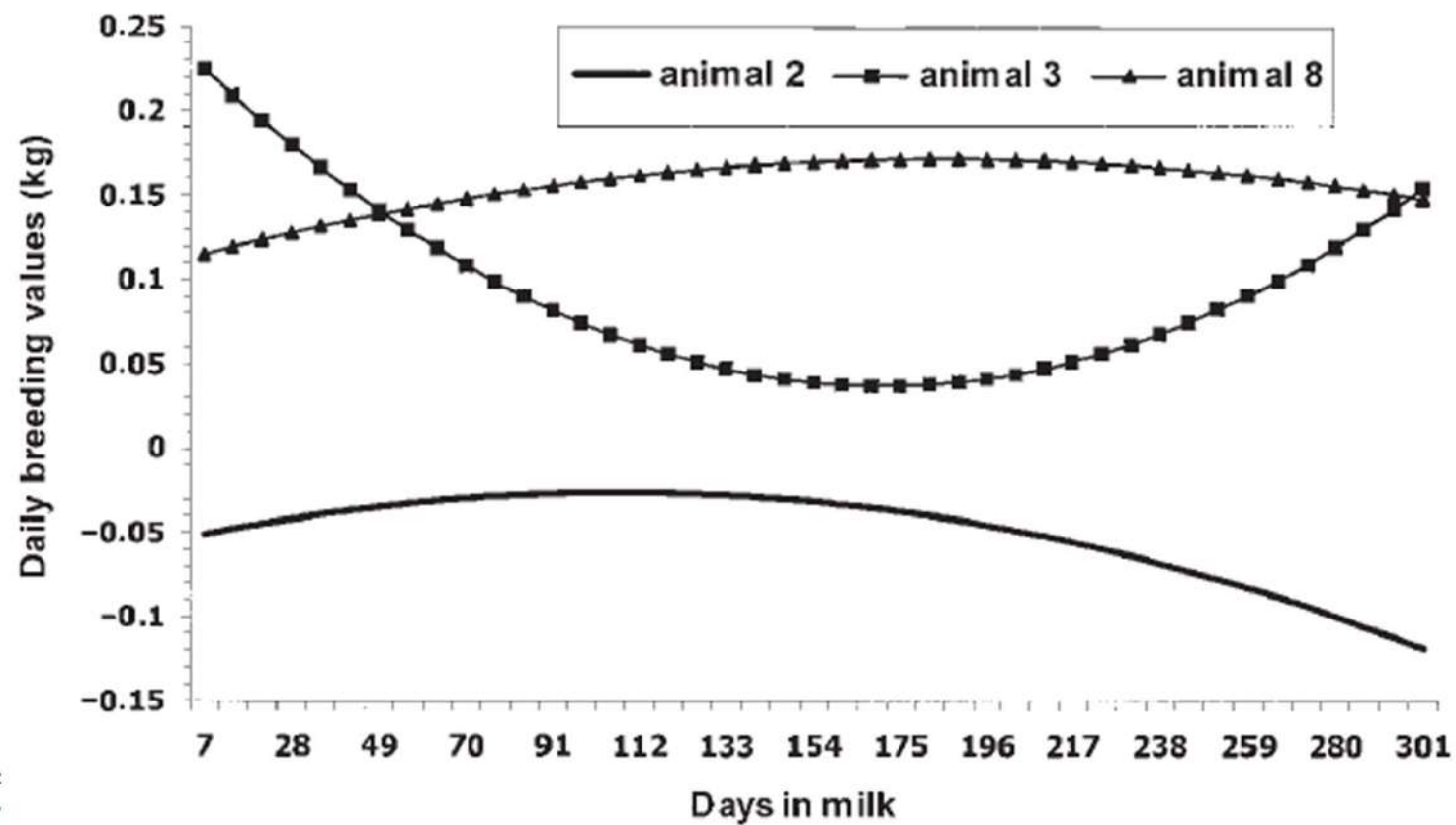
$$y_{tijk} = htd_i + \sum_{k=0}^{nf} \phi_{jtk} \beta_k + \sum_{k=0}^{nr} \phi_{jtk} u_{jk} + \sum_{k=0}^{nr} \phi_{jtk} pe_{jk} + e_{tijk}$$

where y_{tijk} is the test day record of cow j made on day t within htd subclass i ; β_k are fixed regression coefficients; u_{jk} and pe_{jk} are the k th random regression for animal and permanent environmental effects, respectively, for animal j ; ϕ_{jtk} is the k th Legendre polynomial for the test day record of cow j made on day t ; nf is the order of polynomials fitted as fixed regressions; nr is the order of polynomials for animal and pe effects; and e_{tijk} is the random residual. The model in matrix notation is:

Mrode, 2005, p 143



Daily breeding values



Order of polynomials

Figure 1

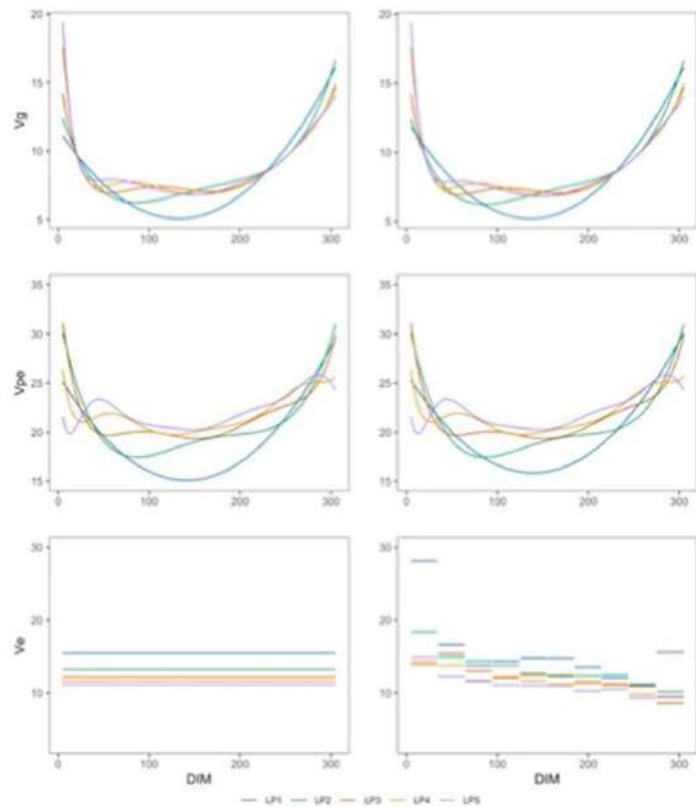


Figure 1. Genetic variances (V_g), permanent environmental variances (V_{pe}), and residual variances (V_e) at each test day along the lactation from models with different orders of LP based on assumption of homogeneous (left column) or heterogeneous residual variance (right column).

<https://www.frontiersin.org/journals/genetics/articles/10.3389/fgene.2020.586155/full>

Li et al. 2020

Order of polynomials

Figure 2

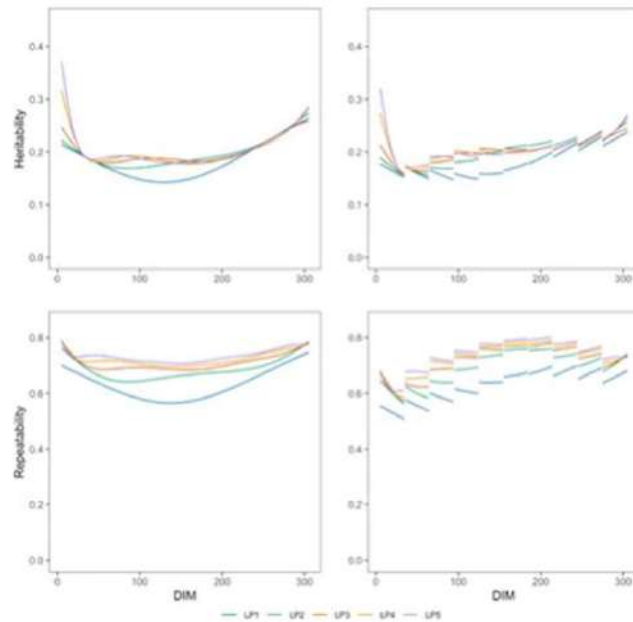


Figure 2. Heritabilities and repeatabilities at each test day along the lactation from models with different orders of LP based on assumption of homogeneous (left column) or heterogeneous residual variance (right column).

<https://www.frontiersin.org/journals/genetics/articles/10.3389/fgene.2020.586155/full>

Li et al. 2020

Choosing the right order

Linear (first order), quadratic (second order), cubic (third order)

- **Low Order:**
 - Simpler, less flexible
 - May not capture complex changes
- **Higher Order:**
 - More flexibility
 - Risk of overfitting
- **Practical Choice:**
 - Often cubic (third-order) is a balance between flexibility and stability



J. Dairy Sci. 107:5853–5868

<https://doi.org/10.3168/jds.2024-24285>

© 2024, The Authors. Published by Elsevier Inc. on behalf of the American Dairy Science Association®.
This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Genetic parameter estimates for methane emission from breath during lactation and potential inaccuracies in reliabilities assuming a repeatability versus random regression model

$$y_{ijkl} = \mu + HYW_i + Par_j \cdot \sum_{k=0}^3 \phi(t)_{lk} \beta_k + Breed \cdot \sum_{k=0}^2 \phi(u) \beta_k \\ + \sum_{k=0}^n \phi(t)_{lk} a_{lk} + \sum_{k=0}^n \phi(t)_{lk} pepar_{lk} + pe_l + e_{ijl},$$



J. Dairy Sci. 107:5853–5868

<https://doi.org/10.3168/jds.2024-24285>

© 2024, The Authors. Published by Elsevier Inc. on behalf of the American Dairy Science Association®.
This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

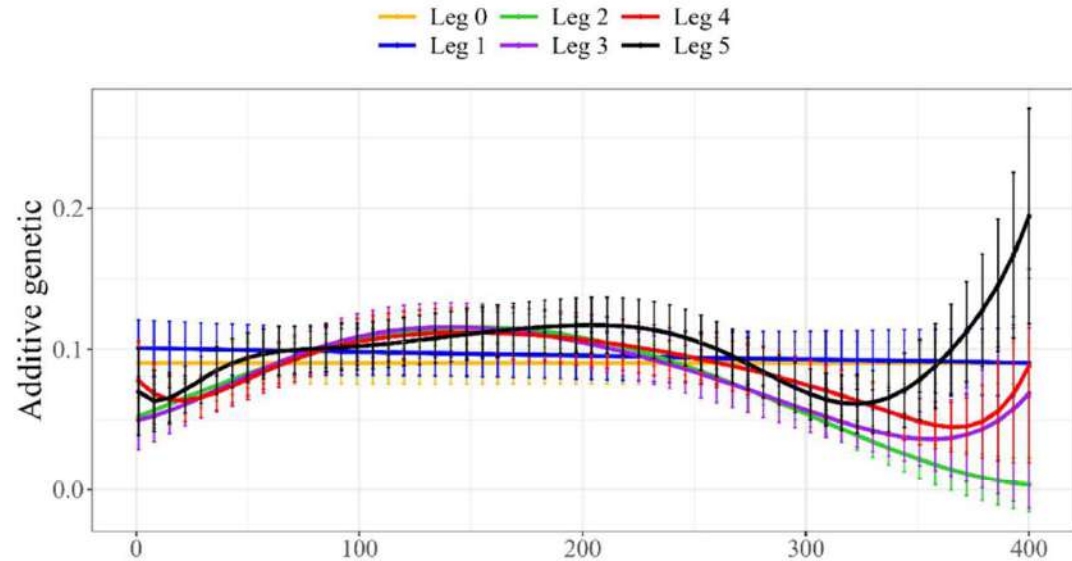
Genetic parameter estimates for methane emission from breath during lactation and potential inaccuracies in reliabilities assuming a repeatability versus random regression model

$$\mathbf{G} = \Phi \mathbf{K} \Phi',$$

$$\mathbf{PE} = \Phi \mathbf{KPE} \Phi', \text{ and}$$

$$\mathbf{P} = \mathbf{G} + \mathbf{PE} + \sigma_e^2,$$

where \mathbf{G} is the genetic (co)variance matrix; \mathbf{PE} is the permanent environmental (co)variance matrix, to which the across parity permanent environmental variance was added to all (co)variance elements in the matrix; \mathbf{P} is the phenotypic (co)variance matrix per DIM ($n \times n$, where n is the level of DIM, consisting of 58 classes of 7 DIM, up to 400 DIM); Φ is a matrix of order $t \times n$, where t is equal to the number of orthogonal polynomial coefficients; \mathbf{K} and \mathbf{KPE} are matrices of order $t \times t$, which contain the estimated covariance functions that describe the genetic (co)variance components and permanent environmental



WAGENINGEN
UNIVERSITY & RESEARCH

Have fun with ASReml!



<https://asreml.kb.vsnr.co.uk/wp-content/uploads/sites/3/ASReml-R-Reference-Manual-4.2.pdf>

<https://asreml.kb.vsnr.co.uk/wp-content/uploads/sites/3/ASReml-4.2-Functional-Specification.pdf>