Repeatability model and kriging as methods for the estimation of growth curves

Kacper Żukowski, Anna Macierzyńska, Heliodor Wierzbicki

Institute of Genetics and Animal Breeding
Wrocław University of Environmental and Life Sciences
Wrocław, Poland

Aim of the Study

- estimating breeding values from time0 to time530 using repeatability model;
- predicting breeding values for all individuals on time600 using kriging.

Repeatability model

$$y = \mu + Xb + Za + e, \quad (1)$$

where:
y - vector of observations;
b - vector of fixed efects (time effect);
a - vector of random animal efects;
e - vector of random residual efects;
and
X, Z are incidence matrices.

Repeatability model - assumptions

- $a \sim \mathcal{N}(0, A \cdot \sigma_a^2)$
- $e \sim \mathcal{N}(0, I \cdot \sigma_e^2)$;

There is also assumed that A is:

- relationship matrix for all animals **(MA model)** or
- kinship matrix estimated from genetic information for 108 SNPs (distributed at all chromosomes) with $r^2 \geq 0.5$;

Methods used to estimate kinship matrix:

- Loiselle *et al* (1995) **(ML model)**;
- Ritland (1996) **(MR model)**;

Kinship matrix was estimated in *SPAGeDi 1.2g*.

Repeatability model - matrix notation

The mixed model equations for BLUE of $\hat{b}$ and BLUP of $\hat{a}$:

$$\begin{bmatrix} \hat{b} \\ \hat{a} \end{bmatrix} = \begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z + \frac{A^{-1}}{\sigma_a^2} \end{bmatrix}^{-1} \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix}$$

σ_a^2 and σ_e^2 were estimated using *DFREML*:

$$\sigma_a^2 = 2.0215883$$
$$R = I \cdot \sigma_e^2$$

$$\sigma_e^2 = \begin{bmatrix} 0.25 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1.33 & 0 & 0 \\ 0 & 0 & 0 & 10.34 & 0 \\ 0 & 0 & 0 & 0 & 30.94 \end{bmatrix}$$

Kriging, Variogram

Suppose, that for any $s \in D$ and $h \in \mathbb{R}^d$, where h is $s_1 - s_2$

$$\begin{aligned} E\left(Z(s+h) - Z(s)\right) &= 0, && (2) \\ Var\left(Z(s+h) - Z(s)\right) &= 2\gamma(h). && (3) \end{aligned}$$

$2\gamma(\cdot)$ is a structure function and is treated as a parameter of the random process $Z(\cdot)$.

Estimator based on the method-of-moments is:

$$2\hat{\gamma}(h) = \frac{1}{\#N(h)} \sum_{N(h)} \left(Z(s_i) - Z(s_j)\right)^2, \qquad h \in \mathbb{R}^d, \tag{4}$$

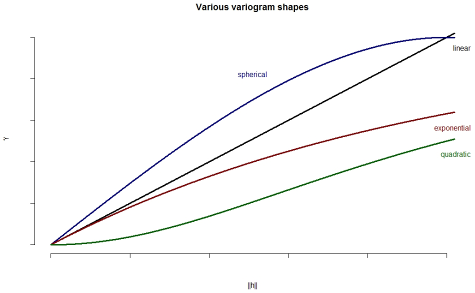
where $|N(h)|$ is the number of distinct pairs in $N(h)$.

Kriging, Variogram models

- Linear
- Spherical
- Exponential
- Quadratic

$$2\gamma_{b,c}(h) = \begin{cases} 0, & h = 0; \\ c + b\left(\frac{\|h\|^2}{1+\frac{\|h\|^2}{a}}\right), & h \neq 0, \end{cases} \quad (5)$$

where $a, b, c \geq 0$.



Method - Universal Kriging

Kriging method is a minimum-mean-squared-error method of spatial prediction that depends on the process $Z(\cdot)$;

$$Z(s) \quad = \quad \mu(s) + \epsilon(s), \tag{6}$$

where:
$Z(s)$ - the weight at time s averaged over all individuals;
$\mu(s)$ - trend of trait;
$\epsilon(s)$ - zero-mean intrinsically stationary random process with variogram $2\gamma(s)$.

Method - Universal Kriging

In this case we could assume that **optimal predictor** can be written as

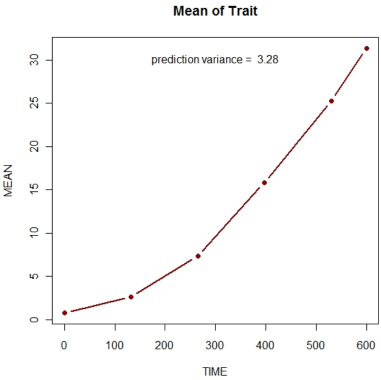
$$p(s, Z) = \sum_{i=1}^{n} \lambda_i Z(s_i). \tag{7}$$

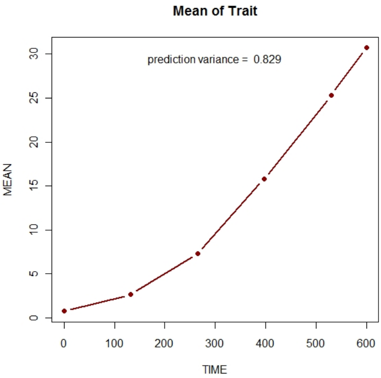
where:
λ_i is optimal weight for each time.

The **optimal prediction variance** *(minimum mean-squared prediction error)* is

$$\sigma^2 = \lambda^T \gamma. \tag{8}$$

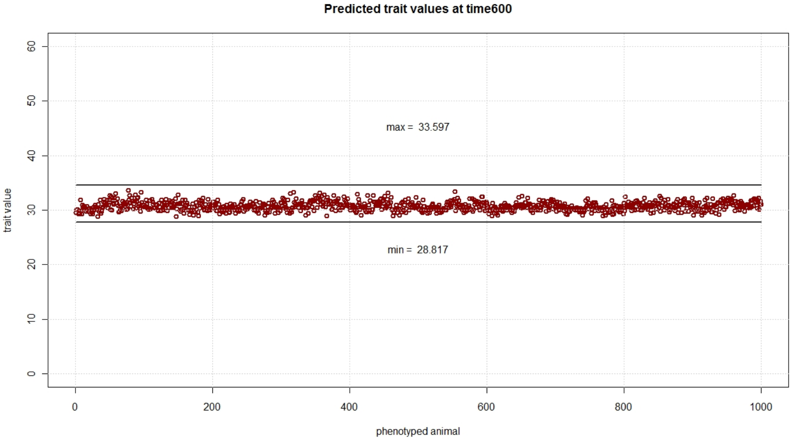
Results, Optimal Prediction using different models of variograms





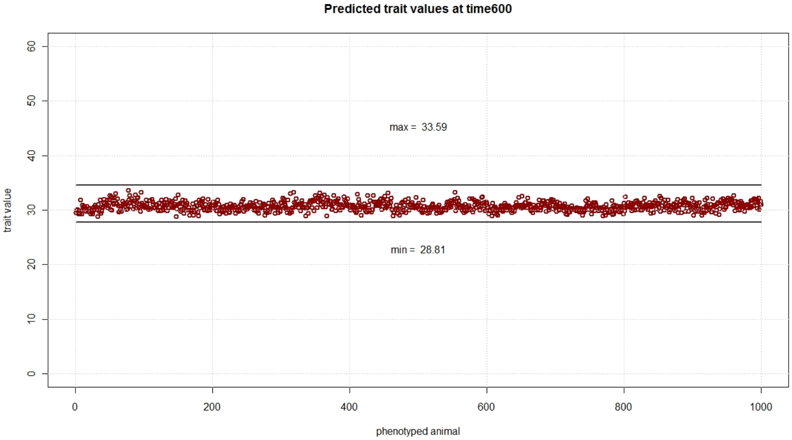
Mean of Trait for different models of variograms (Exponential and Quadratic) at different times for phenotyped animals, ML model.

Results, Prediction at time600 for phenotyped animals



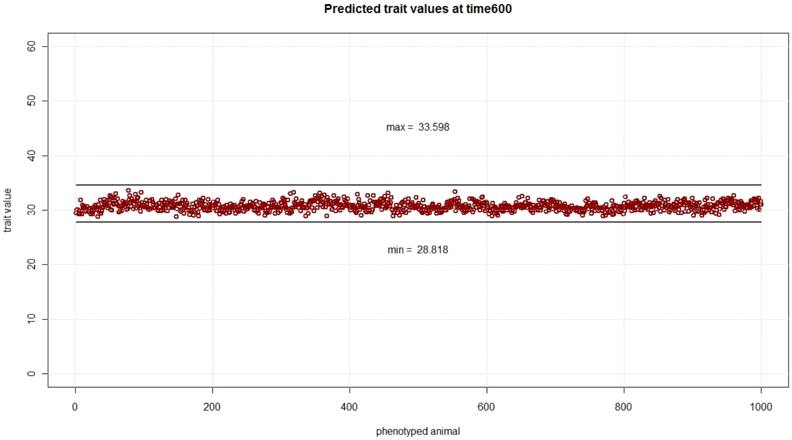
Predicted trait at time600 for phenotyped animals, MA model nad Quadratic variogram.

Results, Prediction at time600 for phenotyped animals



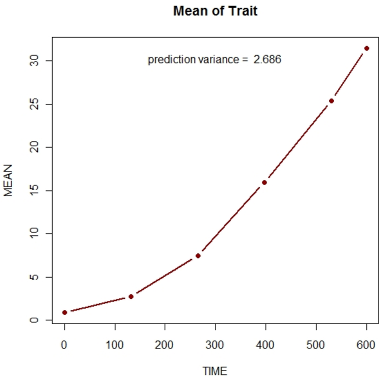
Predicted trait at time600 for phenotyped animals, ML model nad Quadratic variogram.

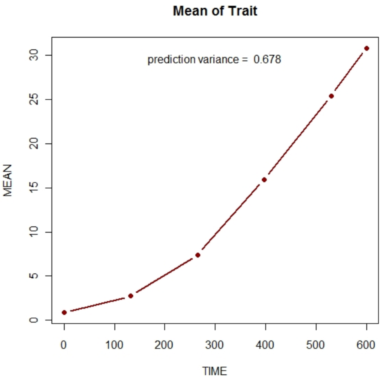
Results, Prediction at time600 for phenotyped animals



Predicted trait at time600 for phenotyped animals, MR model nad Quadratic variogram.

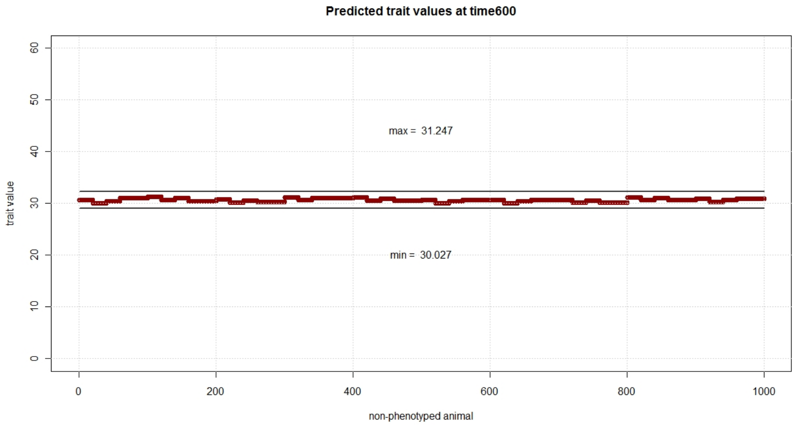
Results, Optimal Prediction using different models of variograms





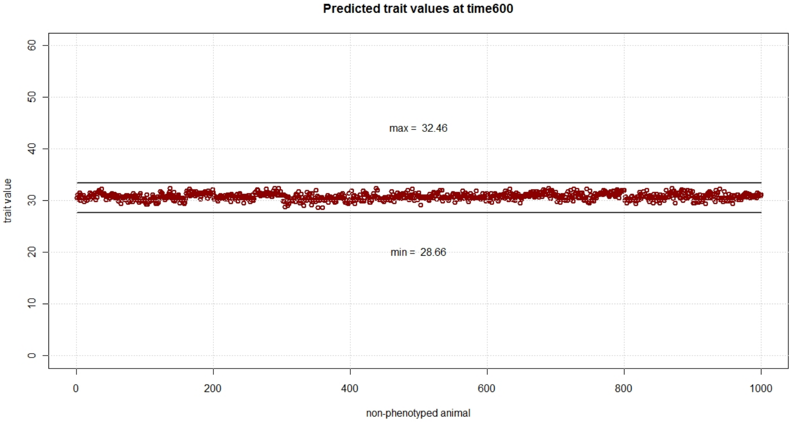
Mean of Trait for different models of variograms (Exponential and Quadratic) at different times for non-phenotyped animals, ML model.

Results, Prediction at time600 for non-phenotyped animals



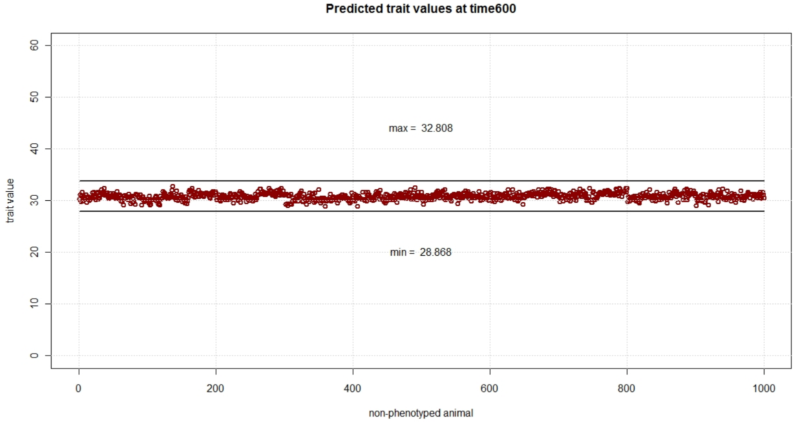
Predicted trait at time600 for non-phenotyped animals, MA model nad Quadratic variogram.

Results, Prediction at time600 for non-phenotyped animals



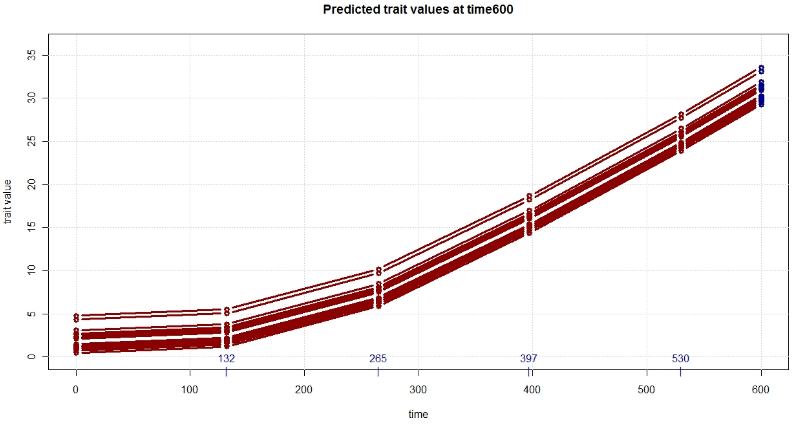
Predicted trait at time600 for non-phenotyped animals, ML model nad Quadratic variogram.

Results, Prediction at time600 for non-phenotyped animals



Predicted trait at time600 for non-phenotyped animals, MR model nad Quadratic variogram.

Results, Predicted trait values at time600



Predicted trait values at time600 for 20 indiwiduals.

Conclusions

- repeatability model is not enough deft instrument for used data set, contrary to anticipation;
- high similarity of all used models in prediction trait for phenotyped animals;
- noticeable difference between standart deviation to not-phenotyped animals for MA and ML/MR models;
- connection repeatability model and kriging - prediction method based on mean is a very conservative approach.

Bibliography

- Arango J.A., Cundiff L.V., Van Vleck L.D. (2004) *Covariance functions and random regression models for cow weight in beef cattle;* Journal of Animal Science, 82:54-67
- Cressie N. (1993) *Statistics for Spatial Data, revised edition;* Wiley, New York
- Hardy, O. J. and X. Vekemans (2002) *SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels;* Molecular Ecology Notes 2: 618-620.
- Interbull (2000) *Proceedings of the 2000 Interbull Meeting,* May 14-15, BLED, SLOVENIA
- Mrode R.A. (2005) *Linear Models for the Prediction of Animal Breeding Values; 2nd Edition,* CABI Publishing

THANK YOU!