2012 Interbull Meeting, May 28-31, 2012, Cork, Ireland

A note on using ‘forward prediction’ to assess precision and bias of genomic predictions

Christian Edel, Stefan Neuner, Reiner Emmerling und Kay-Uwe Götz

Motivation

- Investigation focuses on the property of reliabilities as measures of precision of estimates
  - How much will a future, more reliable estimate deviate from the current one?
  - How will selection affect the conditions in a validation sample?

Methods

- calculations based on model-derived reliabilities and related multivariate-normal distributions of TBV and BLUPs

Methods

- calculations bas
  related multivar
  BLUPs

$$\begin{bmatrix} TBV \\ PA \\ GEBV \\ EBV \end{bmatrix} \sim N\{\boldsymbol{\mu}, \mathbf{V}\}$$

$$\begin{bmatrix} TBV \\ PA \\ GEBV \\ EBV \end{bmatrix} \sim N\{\begin{bmatrix} 100 \\ 100 \\ 100 \\ 100 \end{bmatrix}, \begin{bmatrix} R^2_{TBV} & R^2_{PA} & R^2_{GEBV} & R^2_{EBV} \\ R^2_{PA} & R^2_{PA} & R^2_{PA} & R^2_{PA} \\ R^2_{GEBV} & R^2_{PA} & R^2_{GEBV} & R^2_{M1M2} \\ R^2_{EBV} & R^2_{PA} & R^2_{M1M2} & R^2_{EBV} \end{bmatrix} * 144\}$$

$$R^2_{M1M2} = R^2_{PA} + (R^2_{GEBV} - R^2_{PA})(R^2_{EBV} - R^2_{PA})/(1 - R^2_{PA})$$

Methods

- calculations based on model-derived reliabilities and related multivariate-normal distributions of TBV and BLUPs

- effects of selection:
  - explicit calculations in one step selection scenarios via conditional means and (co-)variances

Methods

- calculations based on model-derived reliabilities and related multivariate-normal distributions of TBV and BLUPs
- effects of selection:
  - explicit calculations in one step selection scenarios via conditional means and (co-)variances

$$E(A|B) = \boldsymbol{\mu}_A + \mathbf{V}_{AB}\mathbf{V}_B^{-1}(B - \boldsymbol{\mu}_B)$$

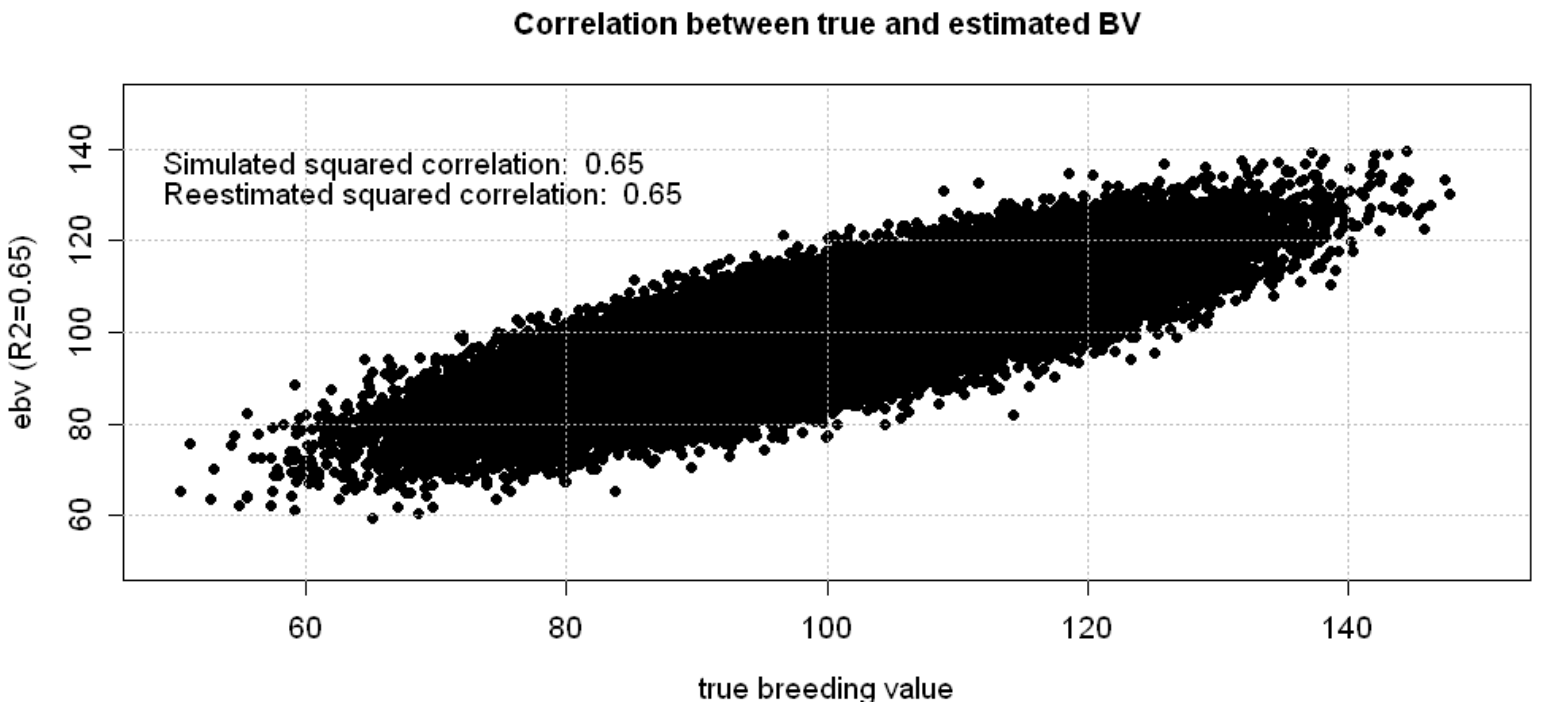
$$Var(A|B) = \begin{bmatrix} \mathbf{V}_A - \mathbf{V}_{AB}\mathbf{V}_0\mathbf{V}_{BA} & \mathbf{V}_{AB}\mathbf{V}_B^{-1}\mathbf{V}_{B_s} \\ \mathbf{V}_{B_s}\mathbf{V}_B^{-1}\mathbf{V}_{BA} & \mathbf{V}_{B_s} \end{bmatrix}$$

Henderson, 1975

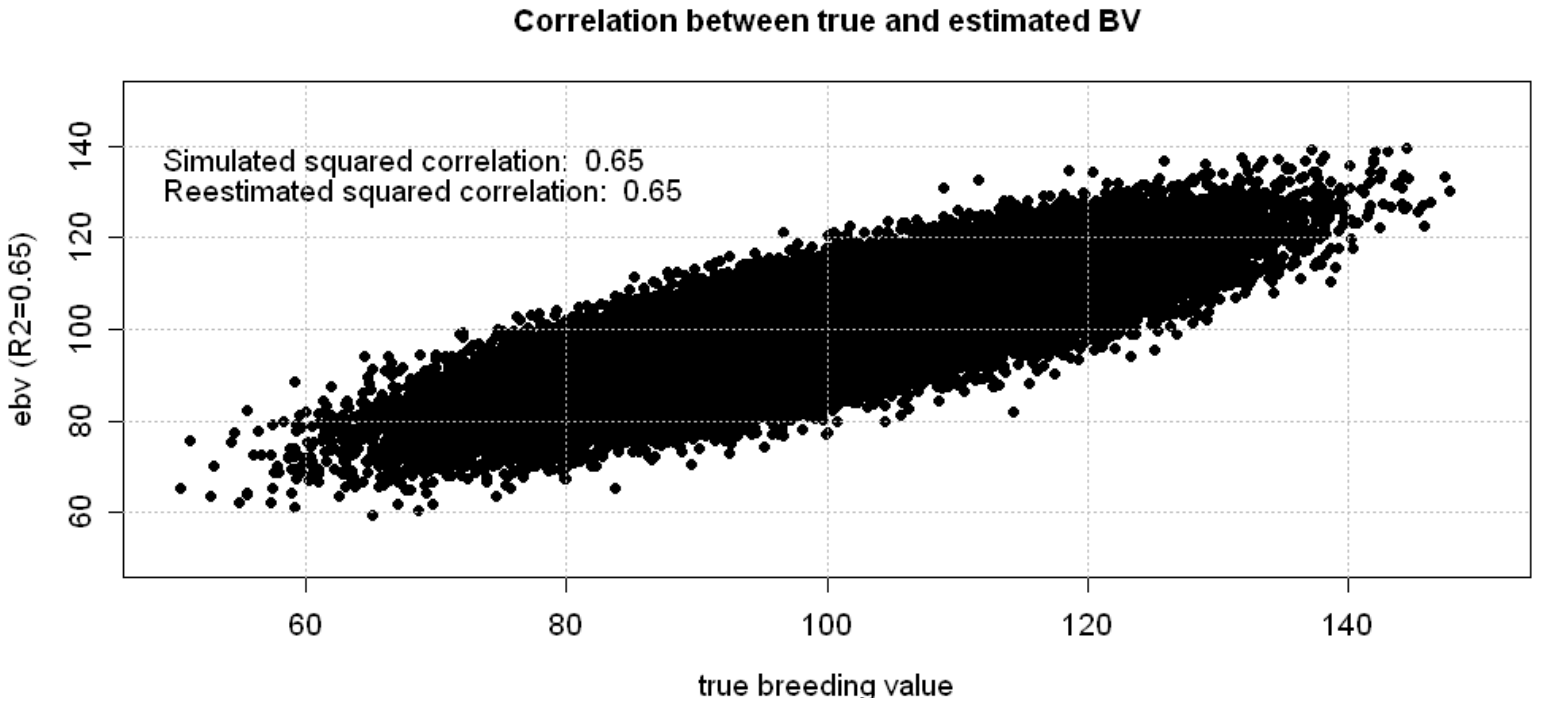
Methods

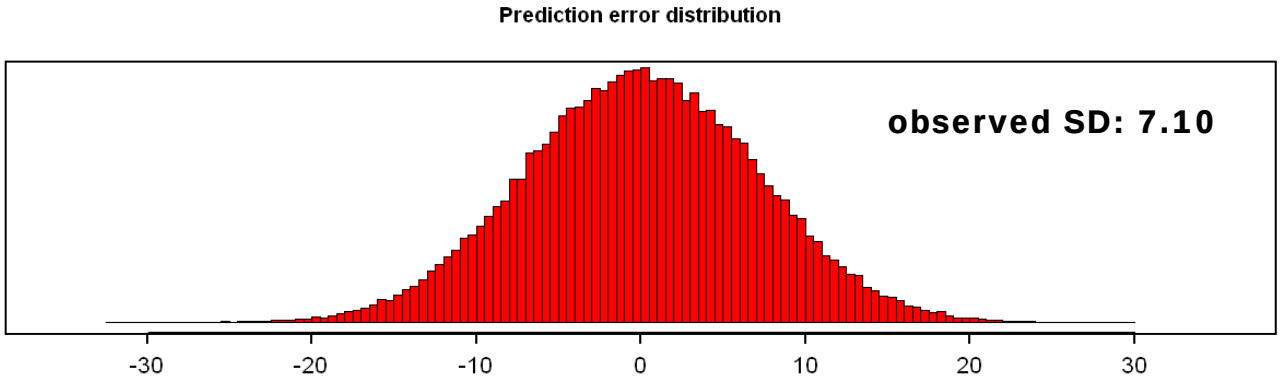
- calculations based on model-derived reliabilities and related multivariate-normal distributions of TBV and BLUPs
- effects of selection:
  - explicit calculations in one step selection scenarios via conditional means and (co-)variances
  - via simulation in multistep scenarios

Illustration

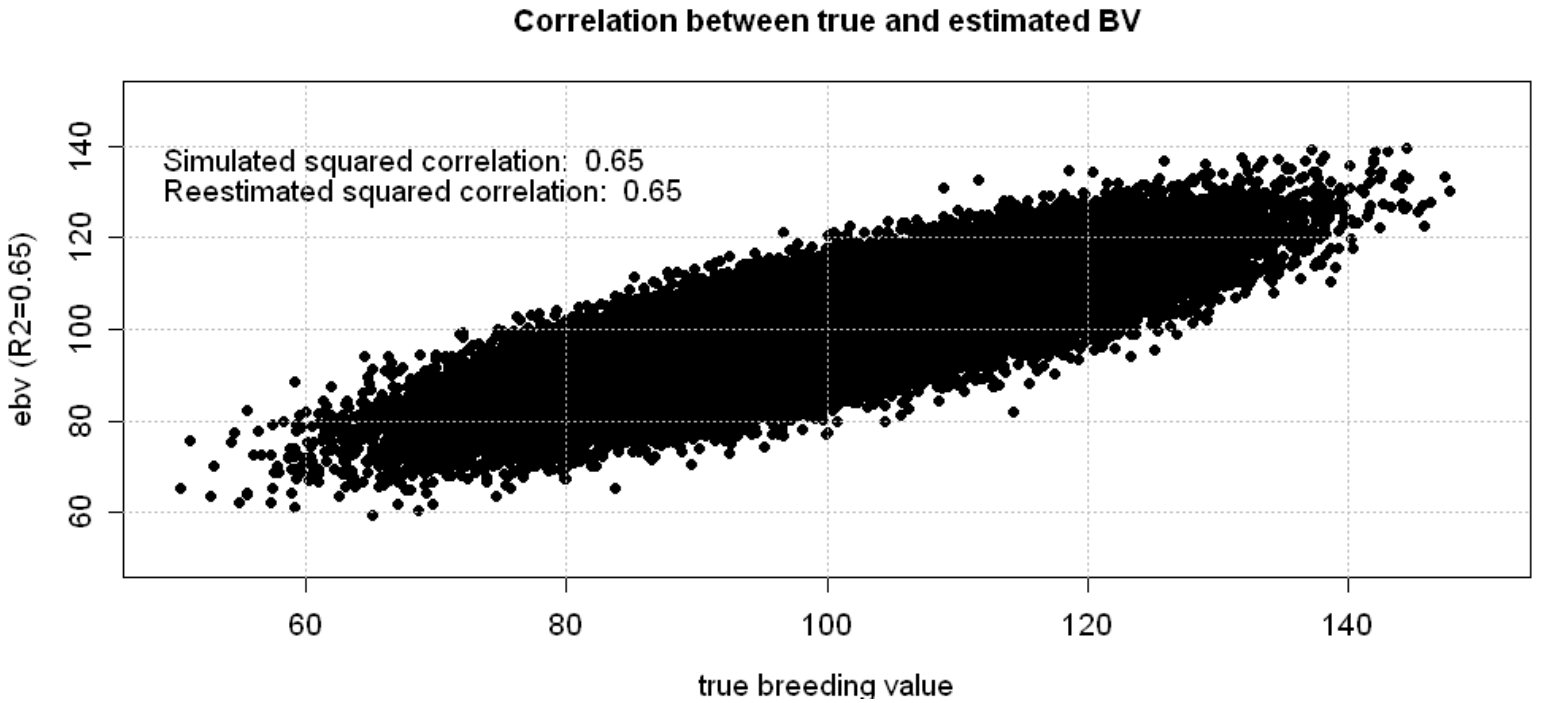


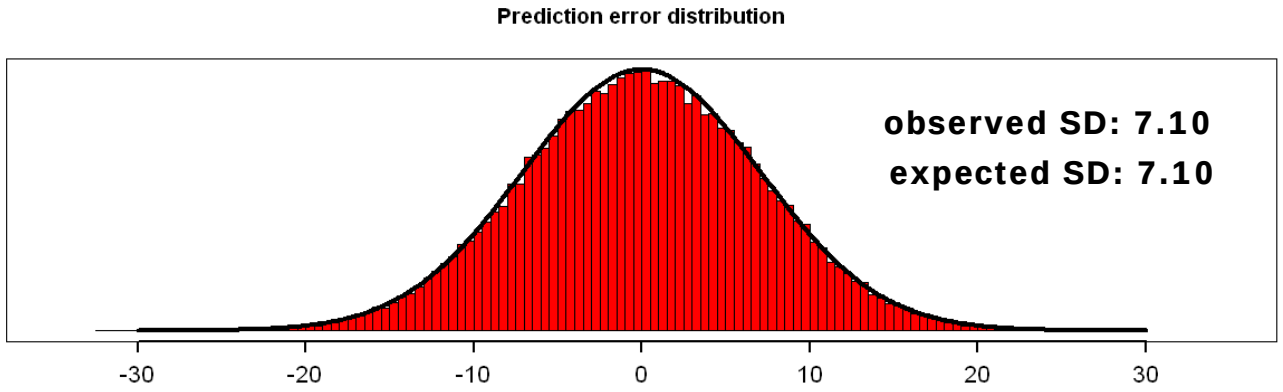
Illustration



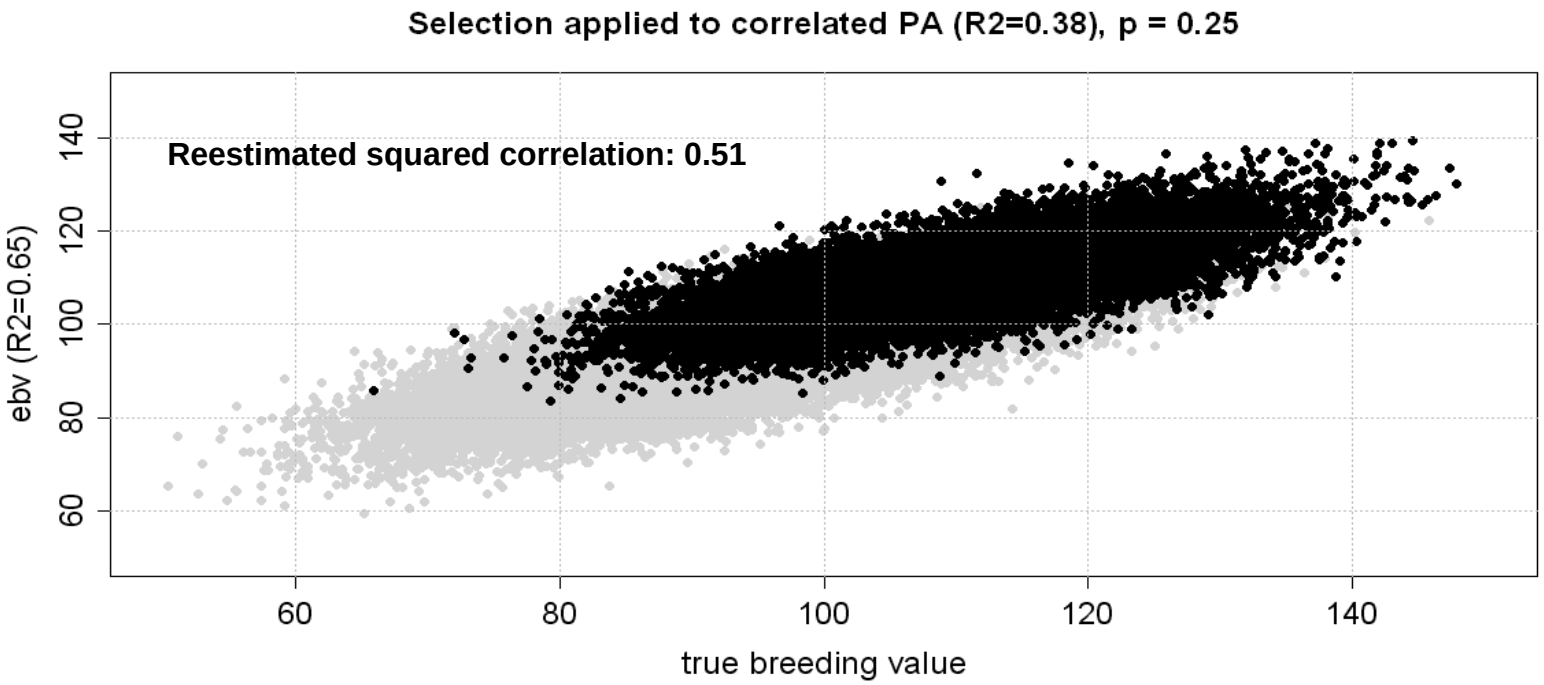


Illustration

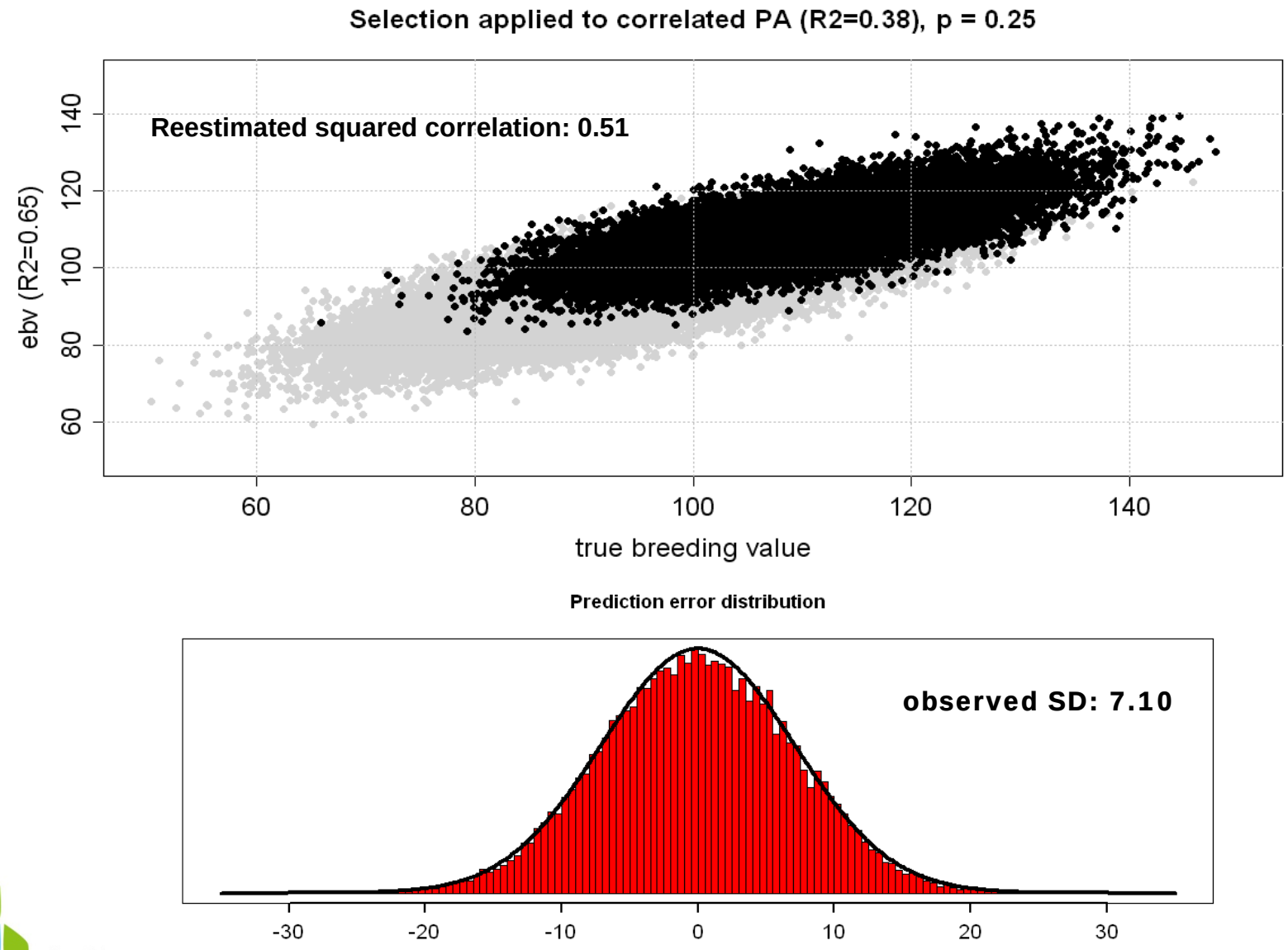




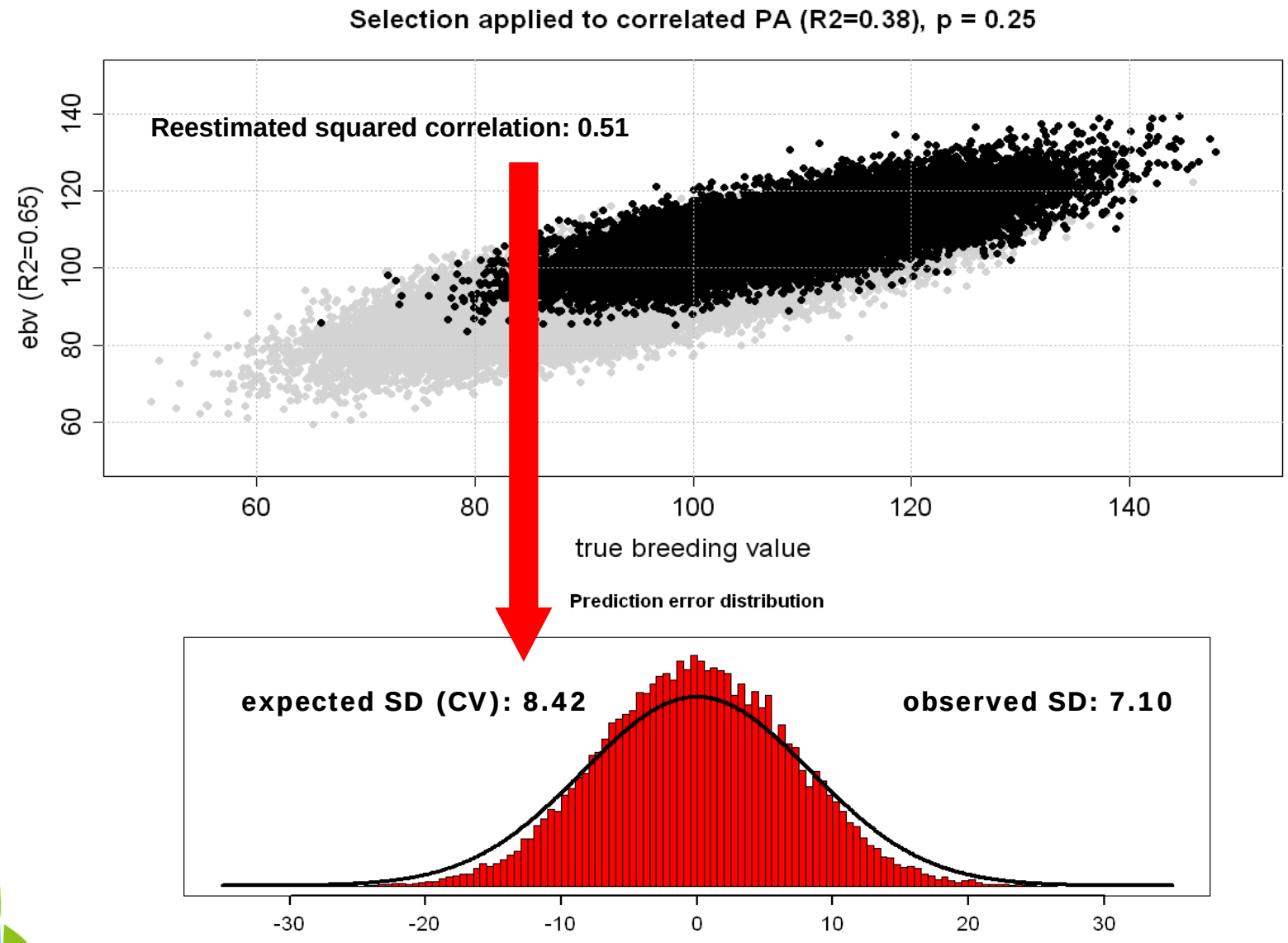
Illustration



Illustration

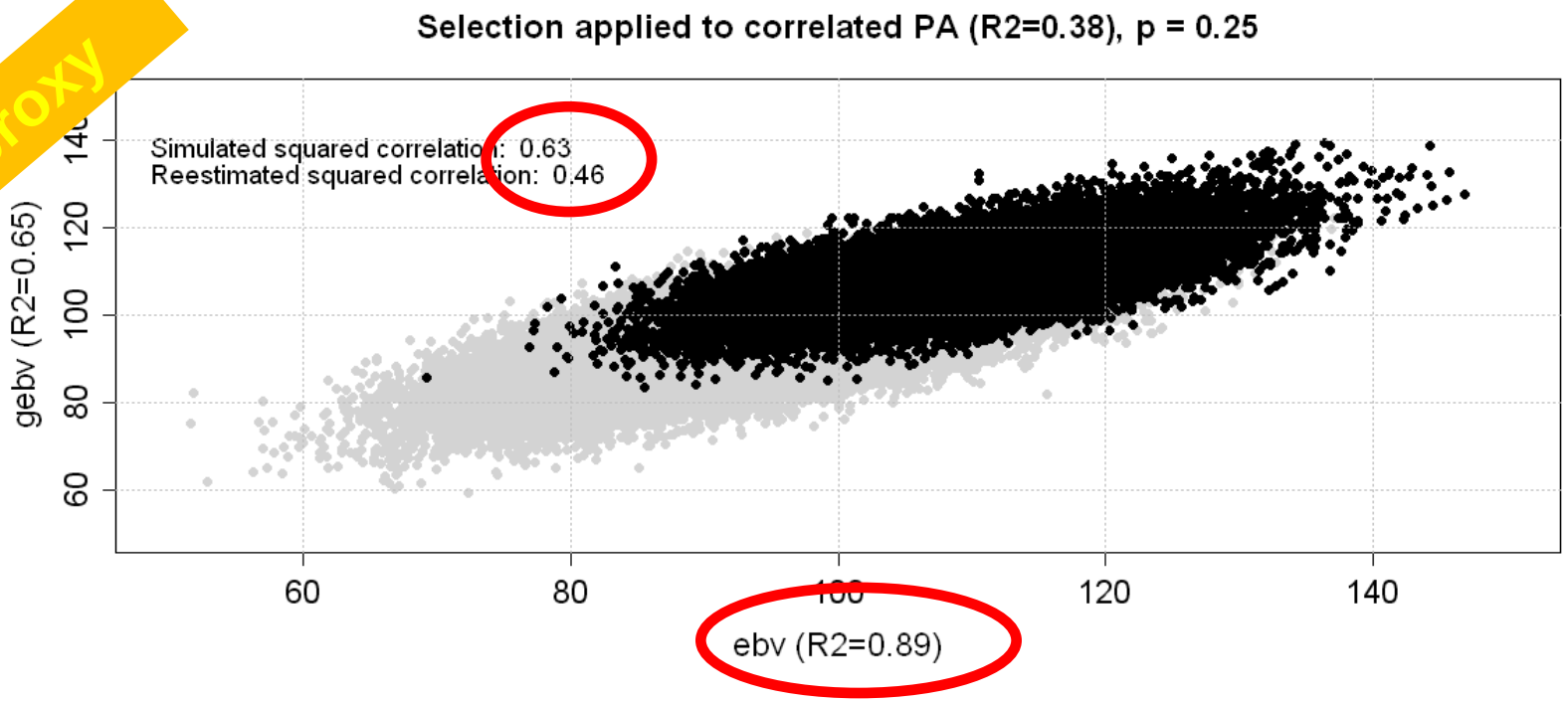


Illustration



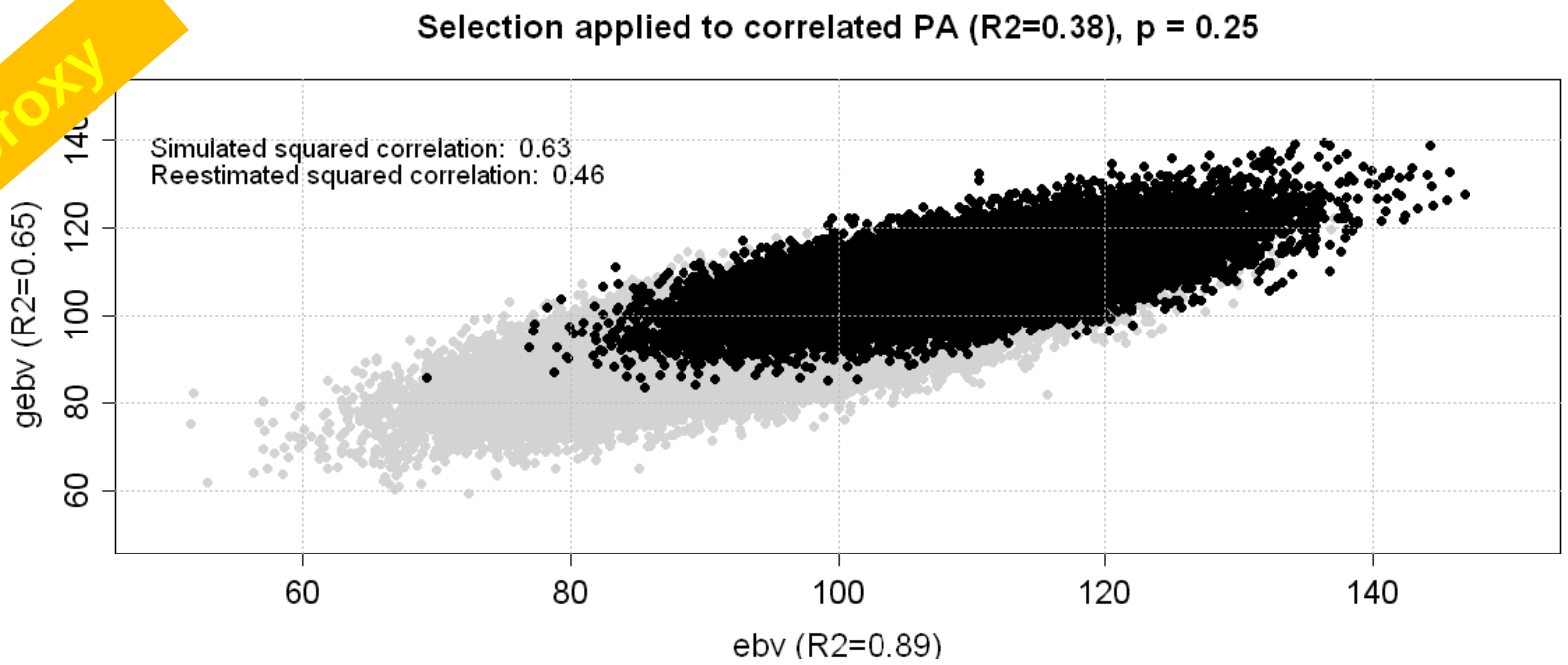
Illustration

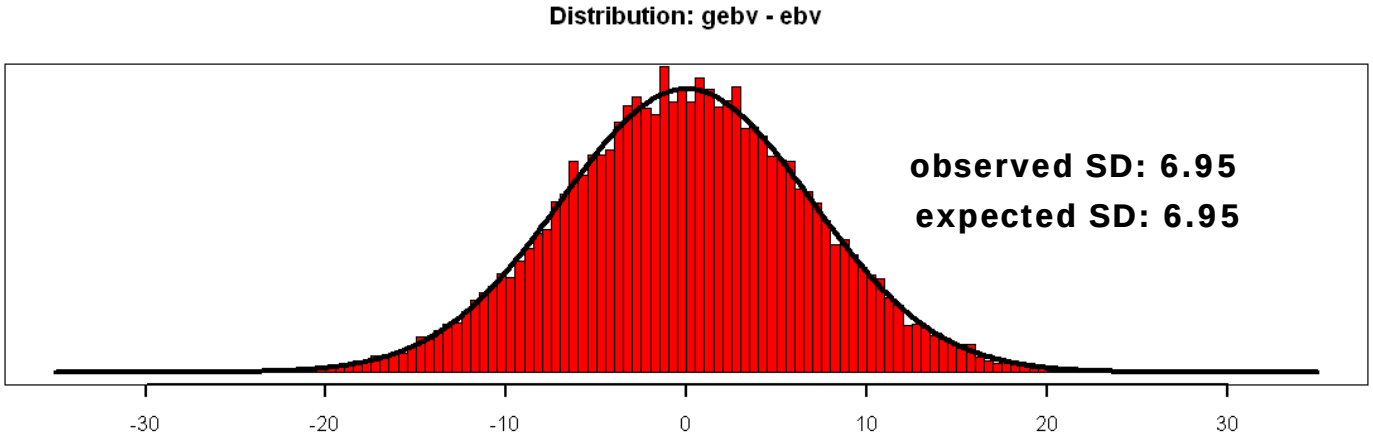
proxy



Illustration

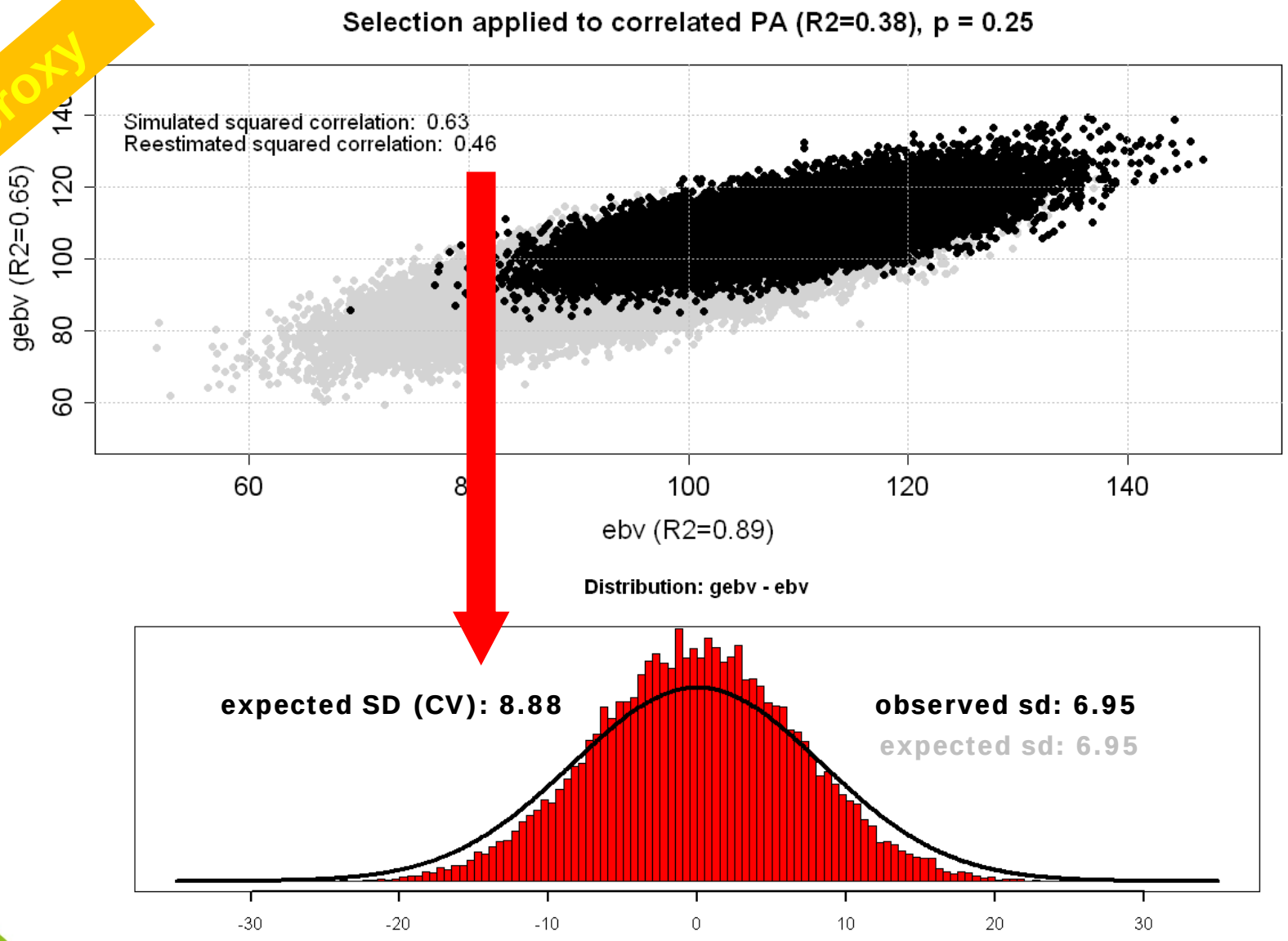
proxy





Illustration

proxy



Results

Table of expected values accounting for selection:

squared correlations and deviations

selection applied to	proportion selected (%)	$\rho^2_{EBV\_GEBV}$	peSD proxy (GEBV-EBV) *"observed"*	peSD CV *„expected" from R^2 in CV*
	100			
PA	25			
EBV	75			
PA/EBV	25/75			

Results

Table of expected values accounting for selection:

squared correlations and deviations

selection applied to	proportion selected (%)	$\rho^2_{EBV\_GEBV}$	peSD proxy (GEBV-EBV) *"observed"*	peSD CV *„expected" from R^2 in CV*
	100	0.63	6.95	6.95
PA	25	0.45	6.95	8.88
EBV	75	0.47	6.49	8.71
PA/EBV	25/75	0.33	6.24	9.86

Results

Table of expected values accounting for selection:

means

selection applied to	proportion selected (%)	Ø TBV	Ø PA	Ø GEBV	Ø EBV
	100	100	100	100	100
PA	25	109	109	109	109
EBV	75	105	102 ⟷	103 ⟷	105
PA/EBV	25/75	113	110 ⟷	111 ⟷	113

Results

Table of expected values accounting for selection:

intercepts and slopes

selection applied to	proportion selected (%)	$b0_{TBV\_GEBV}$	$b1_{TBV\_GEBV}$	$b0_{EBV\_GEBV}$	$b1_{EBV\_GEBV}$
	100	0.00	1.00	7.37	0.93
PA	25	0.00	1.00	14.48	0.87
EBV	75	21.87	0.80	32.61	0.70
PA/EBV	25/75	28.30	0.76	46.15	0.60

General Conclusions

- squared correlations derived from CV can be heavily influenced by effects of selection on validation sample
  - selected animals (preselection on PA, selective genotyping, etc.)
  - data selection (best animals, most reliable animals, etc.)
- effects of selection might be hard to characterize both in scope and nature
- conclusions from CV-correlations about true precision and bias of estimates might be limited

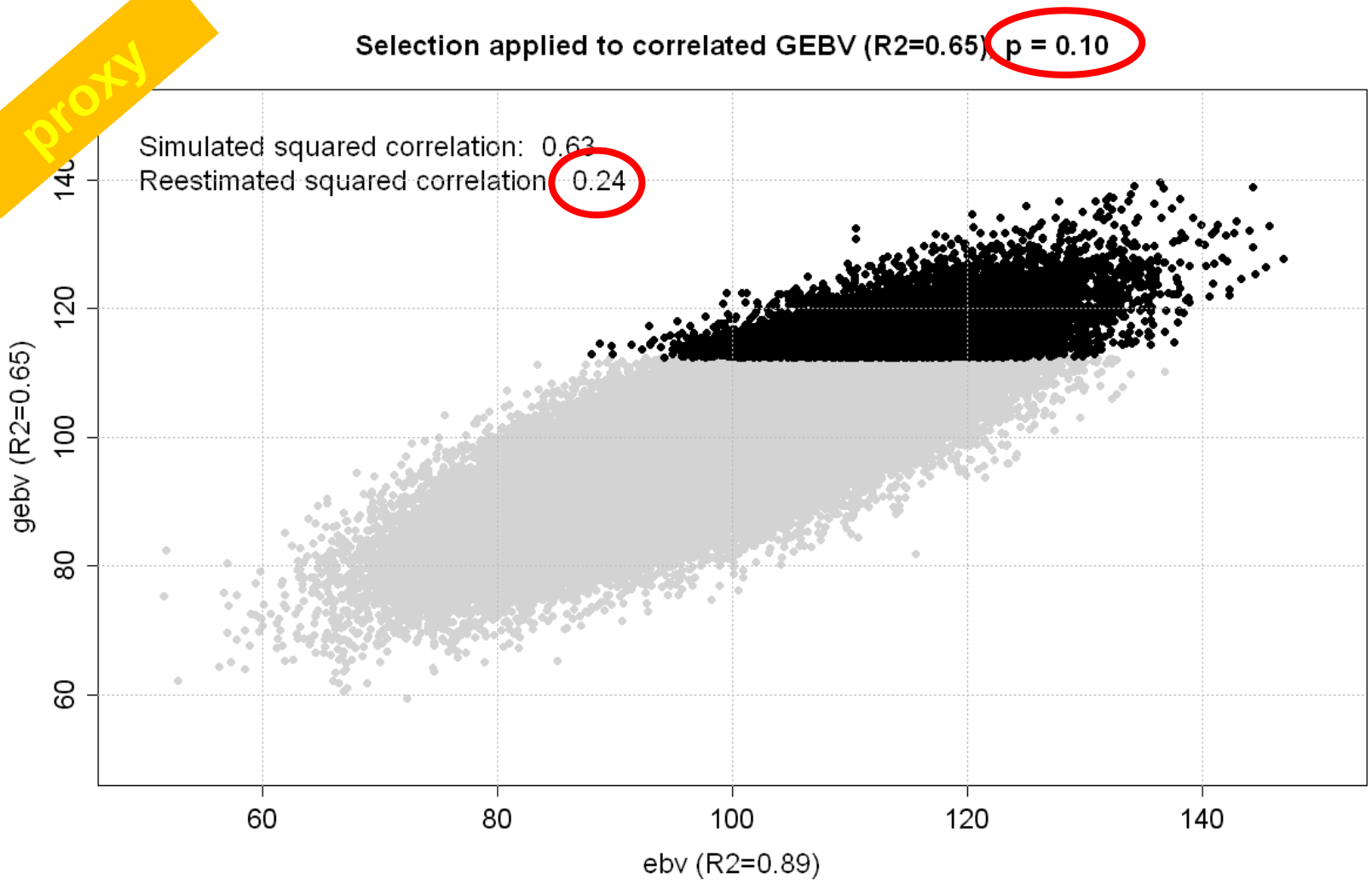
General Conclusions

- distribution of differences (“peSD proxy”)
  - relatively robust: effects of selection are weak
  - could be a helpful extension to GEBV-test
- approach capable for calculation of “expected b1” under selection
  - **using $b1_{TBV\_GEBV}$:** everything depends on deregression
  - **using $b1_{EBV\_GEBV}$:** tests simultaneously for coherence of estimates and assigned reliabilities

In the near future….

Perspectives

proxy



Thank you for your attention