

Reaction norm models with observed covariates

Exploring the value of genomic prediction, reaction norms and phenotyping strategies

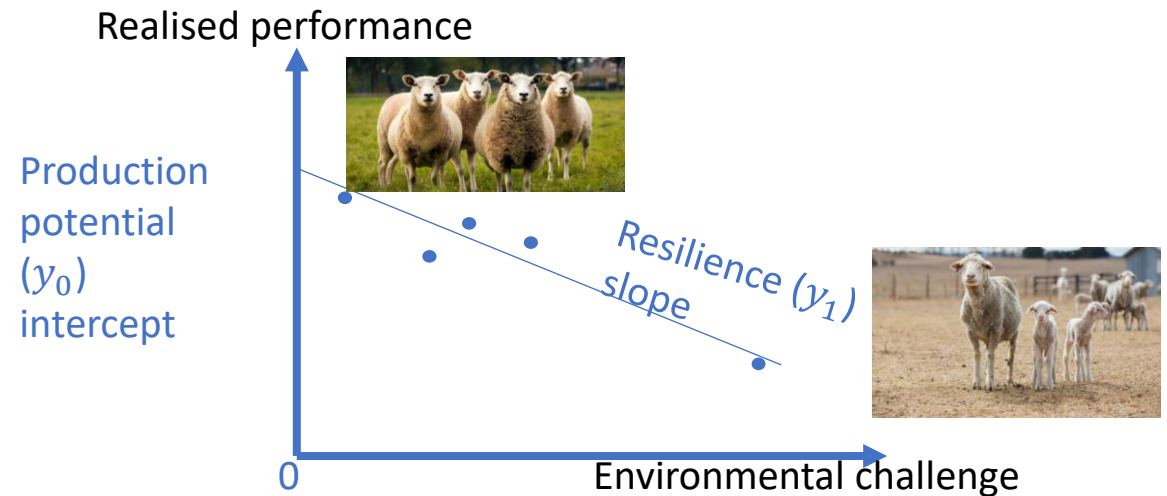
Masoud Ghaderi-Zefreh

Smarter summer-school
March 2023

Animals performance and challenge

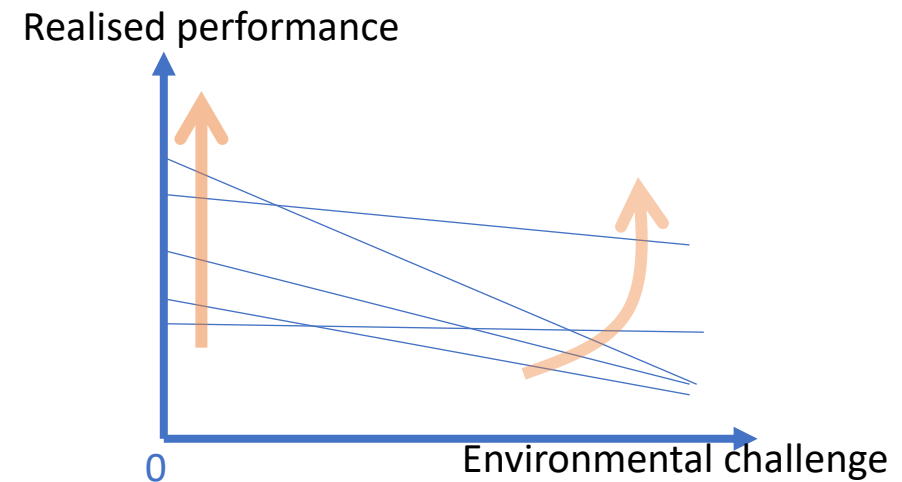
- Challenge: external stimuli provided by the environment that affects animal's phenotype (temperature, humidity, pathogen, quality of food)
- Animals' performance is affected by the level of exposed challenge (on average)
- Resilience: Stay productive under challenging condition
- Production = (Production potential) + (environmental challenge) * resilience

$$y = y_0 + y_1 \cdot t$$



Alternate definitions for resilience

- Plasticity : Resilient animals are non-plastic (don't respond to changes)
 - Sensitivity: Resilient animals are non-sensitive
 - Generalist: Resilient animals are generalist (in contrast with specialist who have high production potential but are not resilient)
 - Robustness*
-
- Favourable: highly resilient animals with high production potential



Random-Regression Reaction-norm model

- Reaction norm: Phenotype affected by environment
- Random regression: dependency to covariate (continuous) is through a random effect
- Random-regression reaction-norm model: effect of environment as a continuous measure is random

$$y = y_0 + y_1 \cdot t$$

u : additive genetic

μ : population average

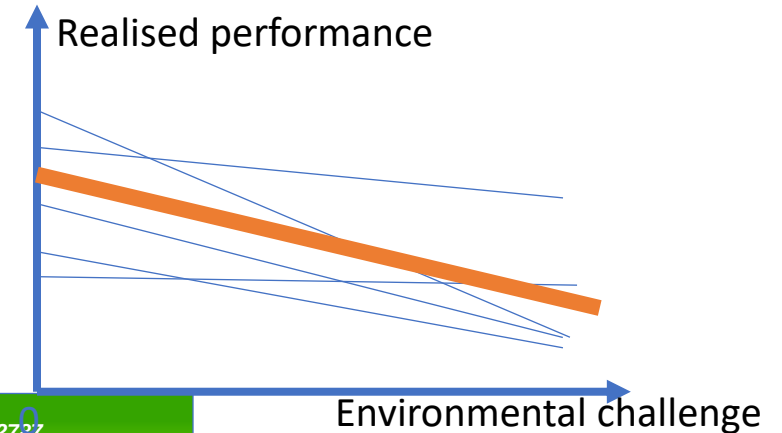
e : environmental deviation

$$y_j = \mu_0 + \mu_1 t + u_{0j} + u_{1j} t + e_{0j} + e_{1j} x$$

Population mean
production

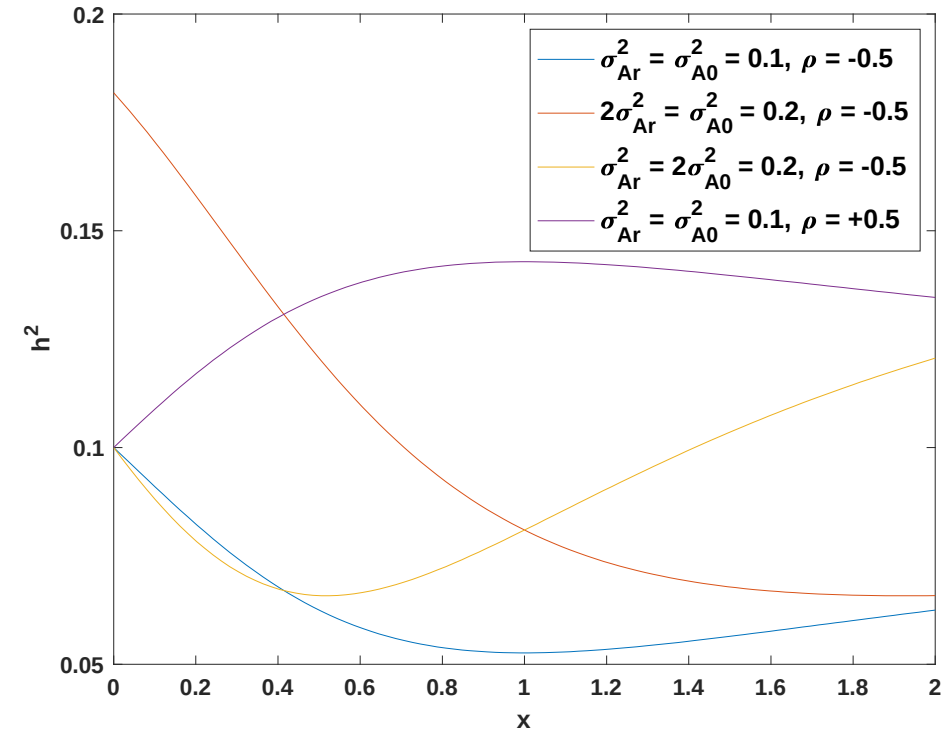
Genetic
deviation

Environmental
deviation



Resilience and production potential as separate traits

- Heritability of production potential $h_0^2 = \frac{\sigma_{u_0}^2}{\sigma_{u_0}^2 + \sigma_{e_0}^2}$
- Heritability of resilience $h_1^2 = \frac{\sigma_{u_1}^2}{\sigma_{u_1}^2 + \sigma_{e_1}^2}$
- Heritability of performance at environment t (Kolmodin et al. 2004)



$$h_y^2(x) = \frac{\sigma_{u_0}^2 + t^2\sigma_{u_1}^2 + 2t\sigma_{u_0u_1}}{\sigma_{u_0}^2 + \sigma_{e_0}^2 + t^2(\sigma_{u_1}^2 + \sigma_{e_1}^2) + 2t(\sigma_{u_0u_1} + \sigma_{e_0e_1})}$$

Linear vs non-linear reaction norm

$$y = y_0 + y_1 \cdot t + y_2 \cdot t^2 + y_3 \cdot t^3 + \dots$$

- Higher flexibility → better fit

However,

- Many traits exhibit linear (or piecewise linear) trend [Sánchez-Molano 2019, 2020; Ramón et al 2021]
- Resilience can be easily inferred from linear traits (slope); what does y_3 refer to?
- More complex selection index in case of higher order



Measuring resilience and production potential using RR

Challenges:

➤ Suitability of Data structure

- Size of phenotypic information (e.g., Number of measurement for individuals (e.g., carcass weight))
- Distribution of related individuals across environments

➤ Unknown environmental challenge (t) [afternoon]

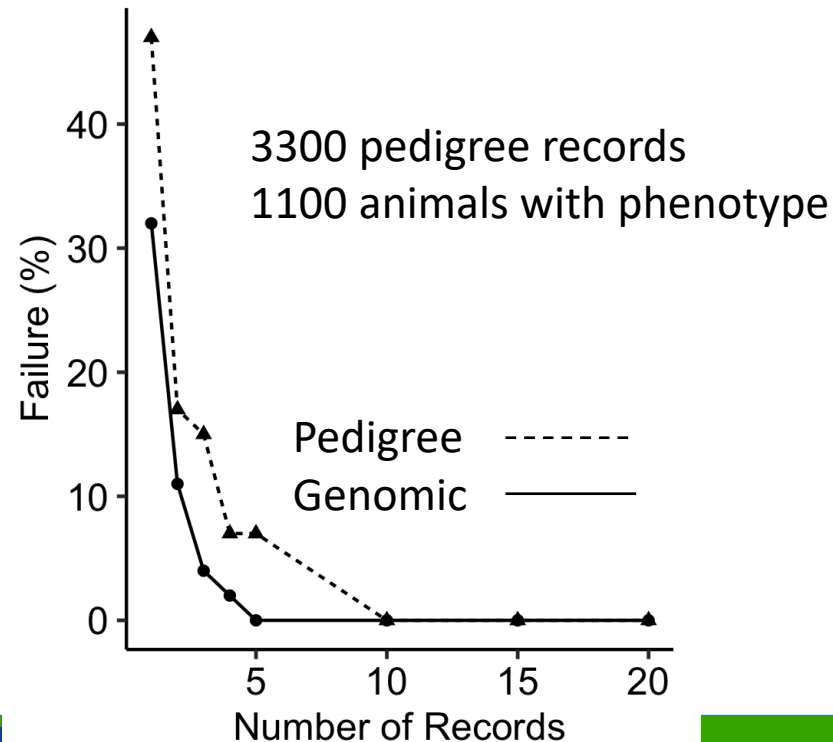
$$y = P_0 + P_R \cdot t$$

✓ Estimation may be poor. Hence, there is scope for genomic prediction



Effect of genomic and number of records on estimation - 1

With low number of records, AIREML may give zero estimate variance component



Example

- We want to estimate 11 effects out of 4 records, but we have relationships and genetic variances and covariances:

$$\text{Var} \begin{pmatrix} \mathbf{u}_0 \\ \mathbf{u}_1 \end{pmatrix} = \mathbf{A} \otimes \mathbf{G}_0 = \begin{pmatrix} 1 & 0 & 0.5 & 0 \\ 0 & 1 & 0 & 0.5 \\ 0.5 & 0 & 1 & 0 \\ 0 & 0.5 & 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 30 & -1 \\ -1 & 3 \end{pmatrix}$$

relationships

trait
covariances

- animals 1,3: family A; animals 2, 4: family B

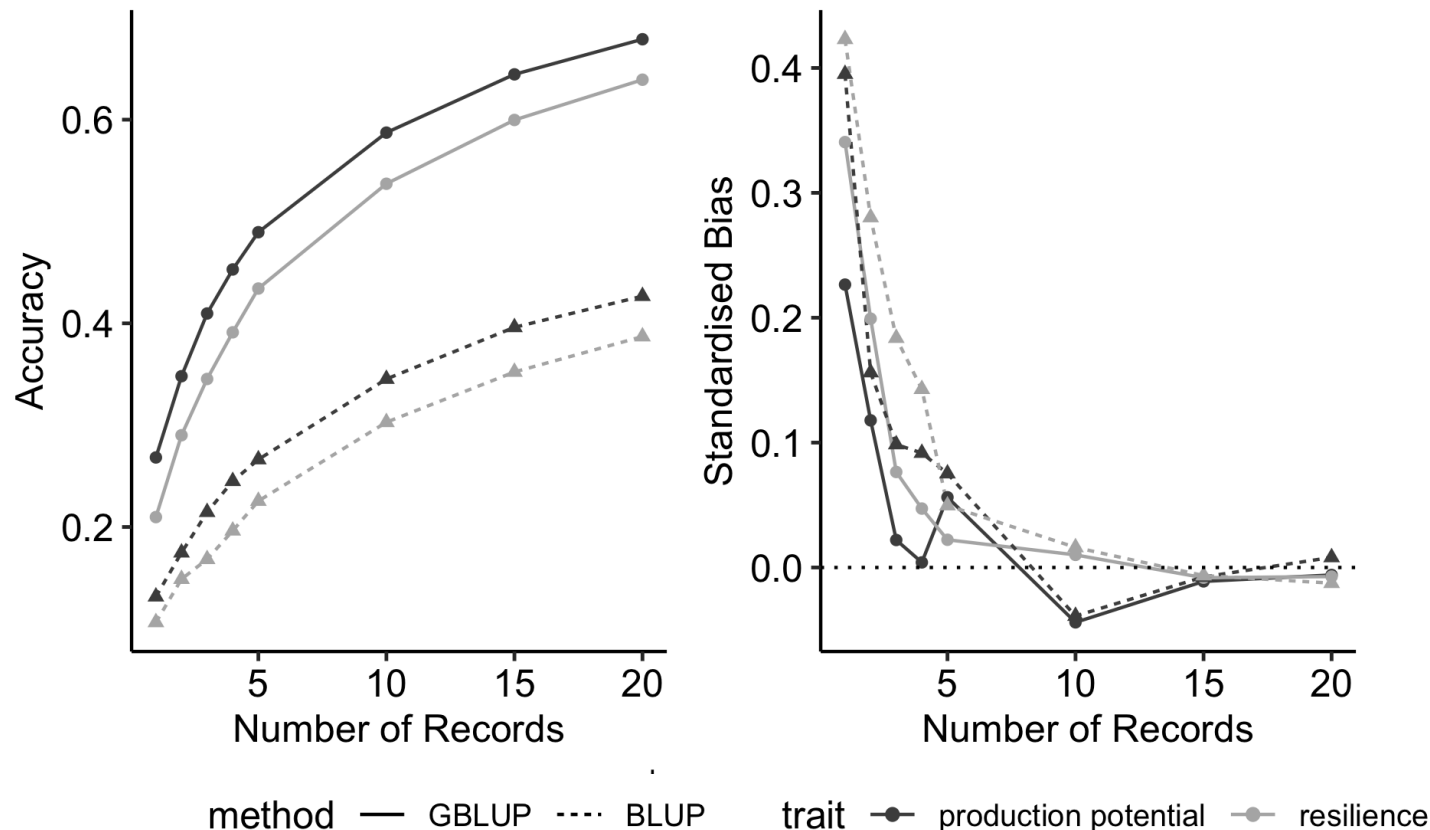
- $\text{Var}(e) = 2$
- $\mathbf{y} = [3000; 2500; 1500; 2000]$

Effect of genomic and number of records on estimation -2

$$\text{Accuracy} = \text{corr}(\text{EBV}, \text{TBV})$$

$$b = \text{reg.coef}(\text{EBV}, \text{TBV})$$

$$\text{Std bias} = \begin{cases} 1 - b & b < 1 \\ \frac{1}{b} - 1 & b > 1 \end{cases}$$



For detail discussion on data structure, see Calus et al 2004



Possible distribution of families across environments

- Random: Each sire has offspring in different random environments (flock)
- Clustered: A sire has all of its offspring reared in the same flock
- Assortative: Same as clustered with negative correlation between sire TBV and challenge level of the herd: better sires live in better environment



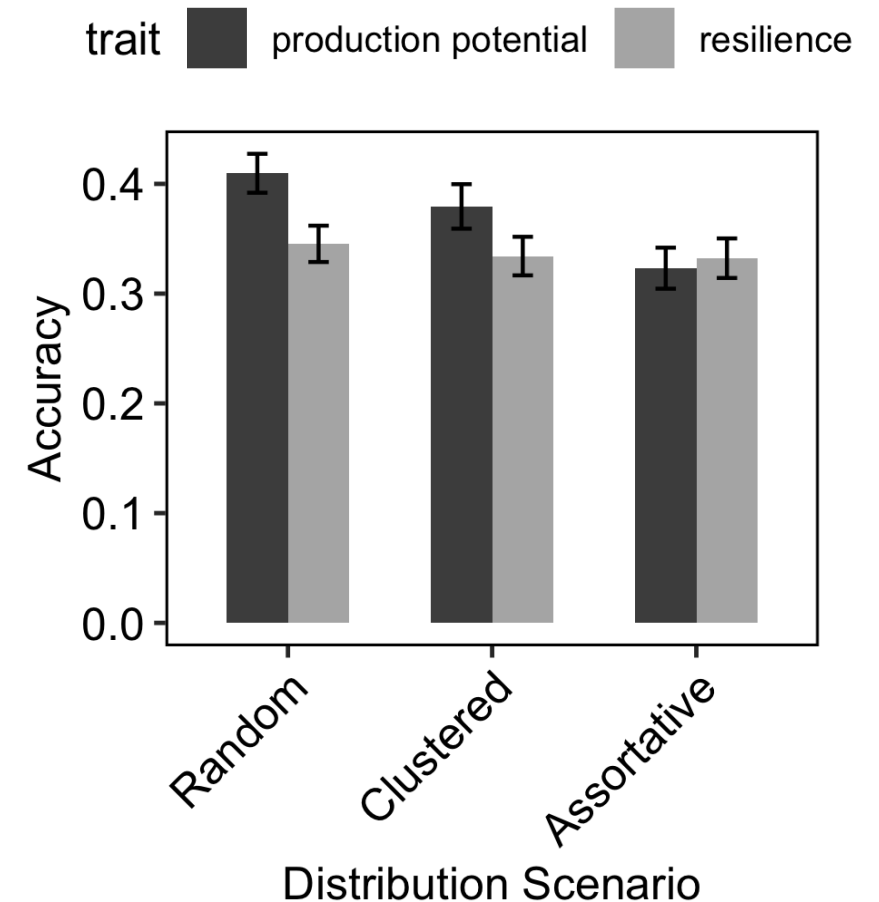
Effect of distribution of families on estimation

- Clustered and assortative distribution have poorest estimation in terms of accuracy of EBV and precision of genetic variance when pedigree information is used (Calus et al 2004)
- This is because genetic connectedness is reduced between herds (Fouilly et al 1990)
- However, using genomic information we obtain a connectedness between unrelated individuals
- Can we avoid this issue with genomic?



Effect of distribution of families on estimates

- No!
- Genomic improve accuracy in general
- However, the detrimental effect of clustering remains for production potential
- This is not a problem for resilience GEBV



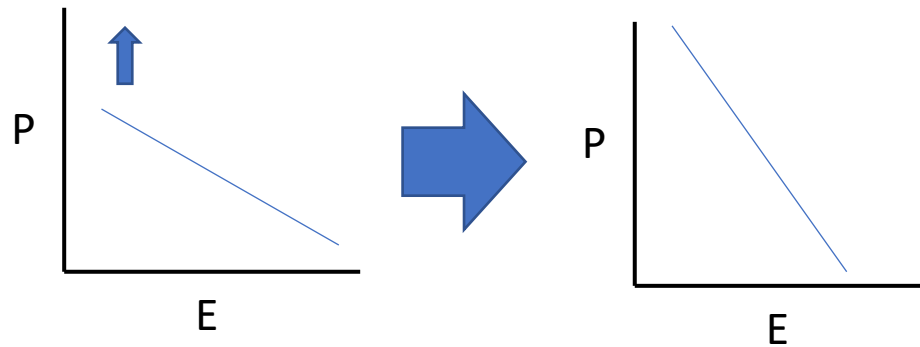
How do we use results from this model

- For instance (the numbers are invented)
 - assume that we increase milk yield by $\Delta G = 100$ liters in 20 years
 - correlated response in “slopes” is
- $\Delta G_{slope} = \Delta G r_{g(0,1)} \frac{\sigma_{g(0)}}{\sigma_{g(1)}} = 100 * (-0.5) \frac{2}{20} = -2$
- now we move to a country with +10 C
- Δ in yield: $\Delta G_0 + \Delta G_1 * 10 = 100 + (-2) * 10 = 80$
- we have lost 4 years of genetic improvement ☹ and our animals are more sensitive



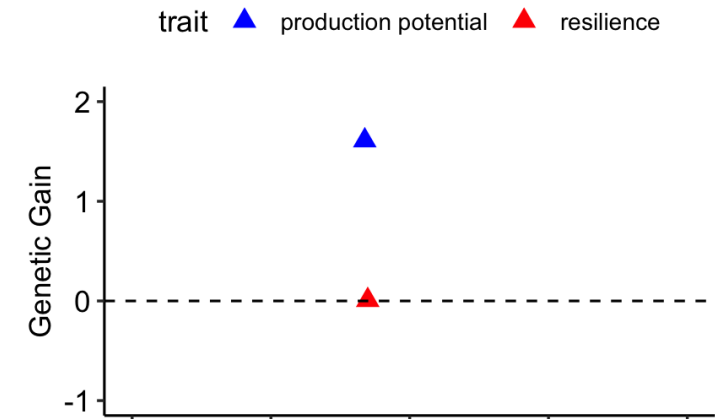
Selection with a model without RR

Conventional model without RR: Response to selection depends on the environment where phenotypes are collected [Falconer 1990]



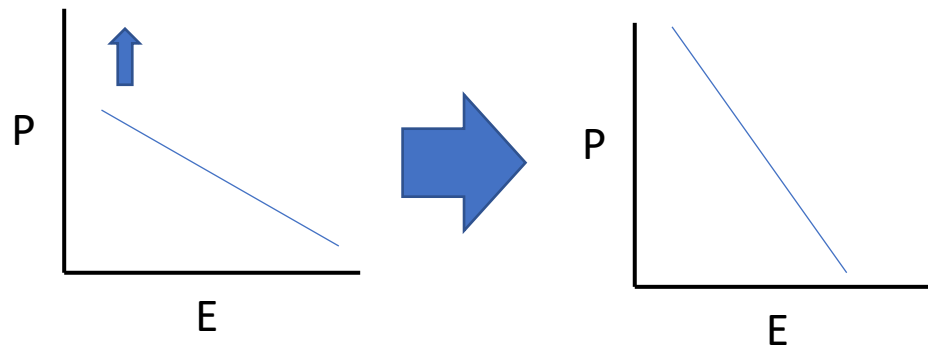
Synergetic selection

Without RR ▲

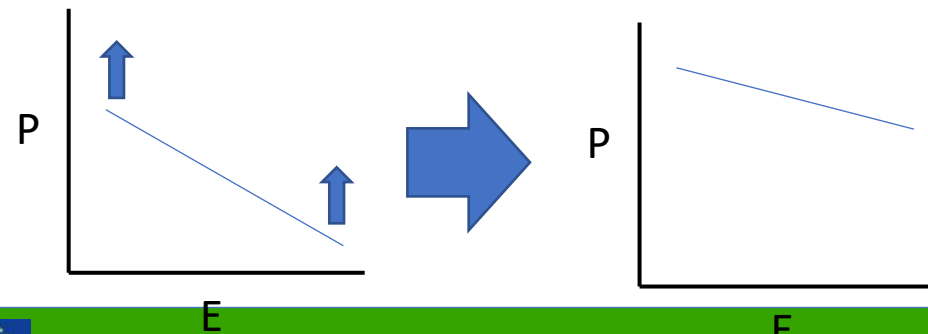


Selection with a model without RR

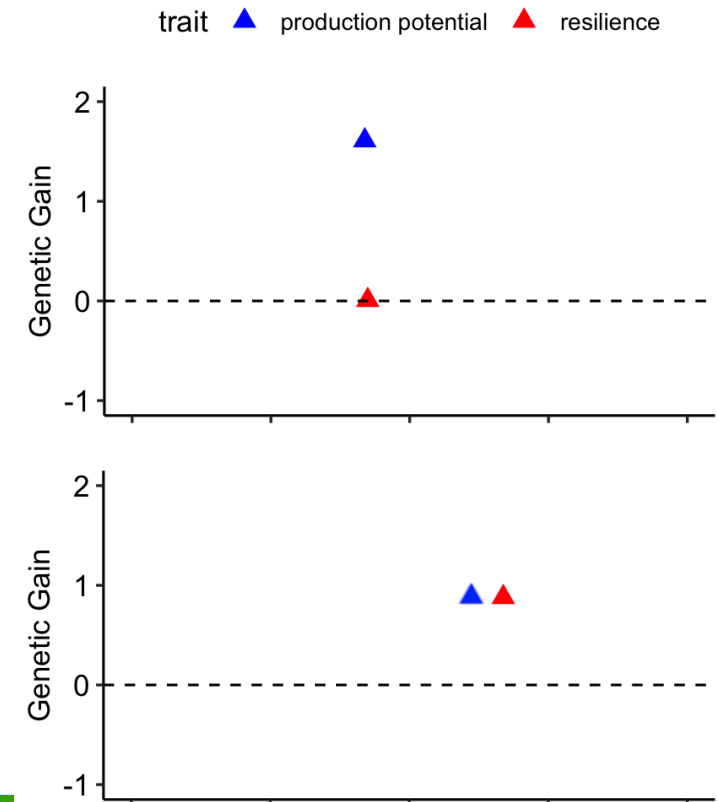
Conventional model without RR: Response to selection depends on the environment where phenotypes are collected [Falconer 1990]



Synergetic selection

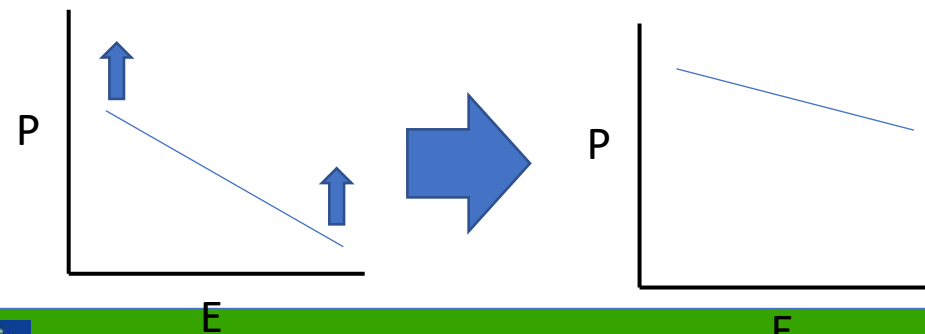
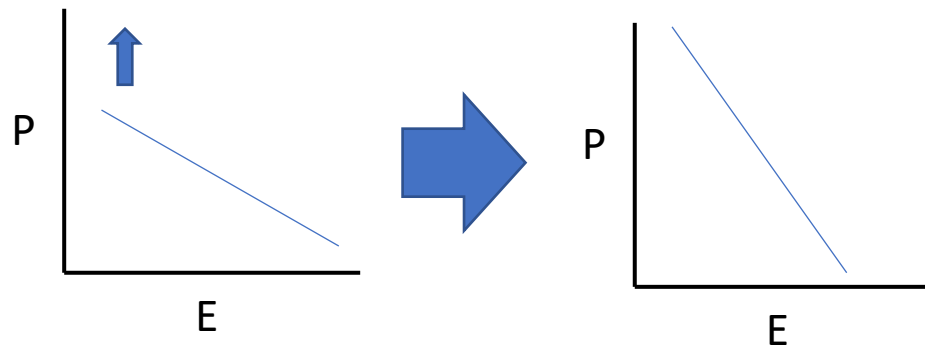


Without RR ▲

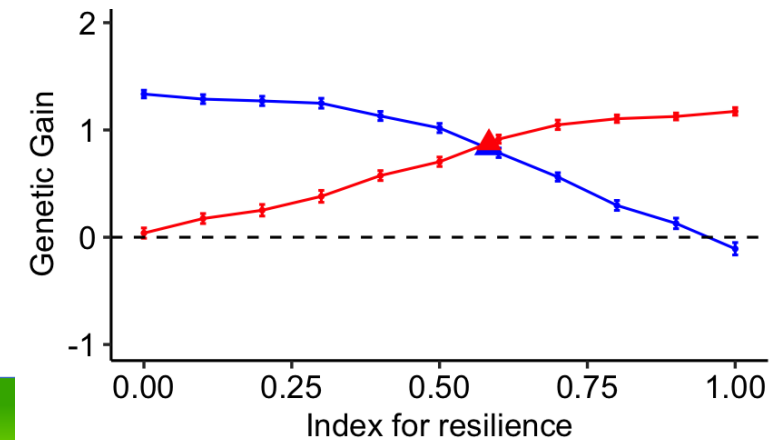
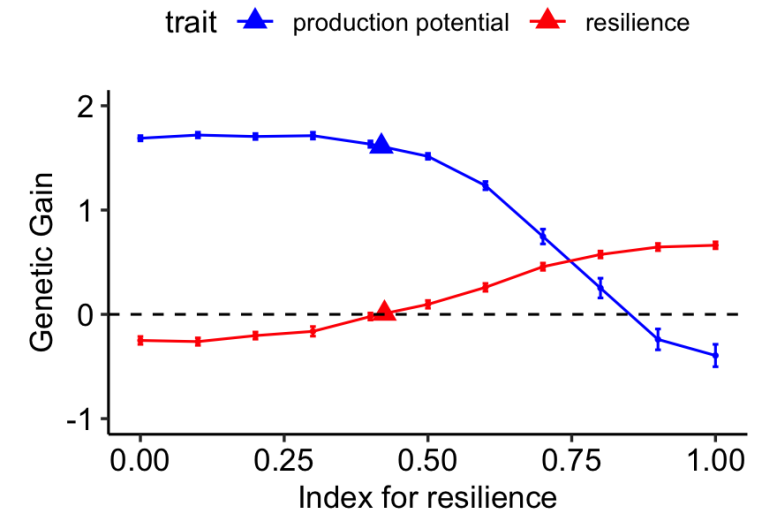


Selection with a model with RR

Comparison



Without RR ▲
With RR —



Selection with a model with RR

Comparison

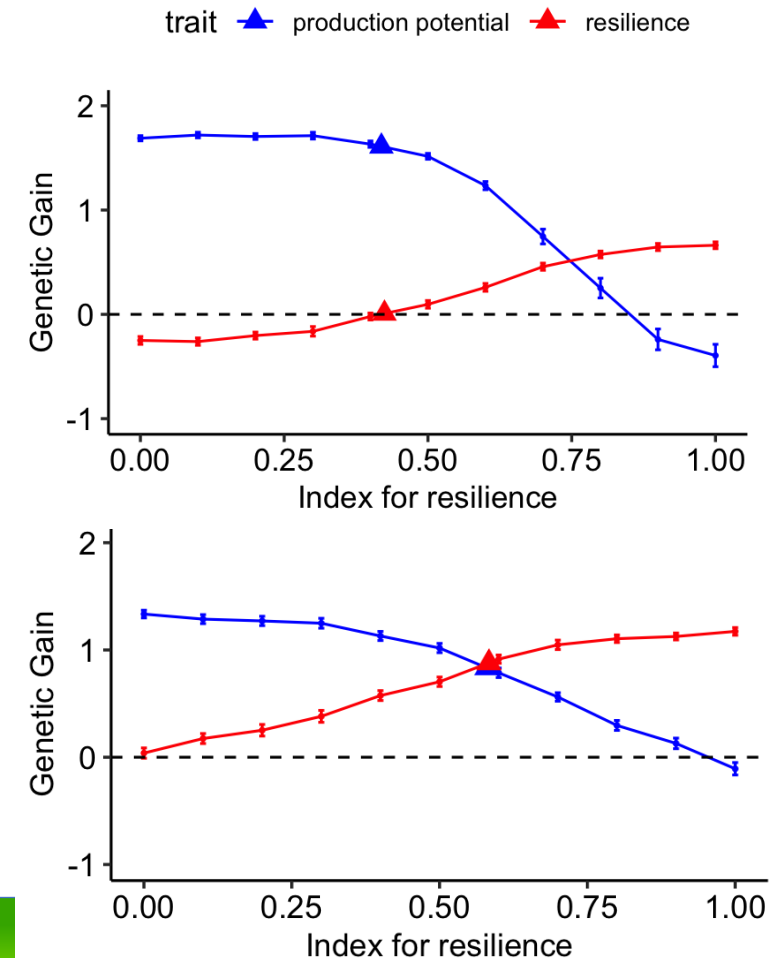
Using model without RR.

- Low variability of environments:
Selection for production may lead to loss of resilience – vice versa
- High variability of environment:
Selection for production also increases resilience

Using model with RR

- Can improve both or based on an index

Without RR ▲
With RR —



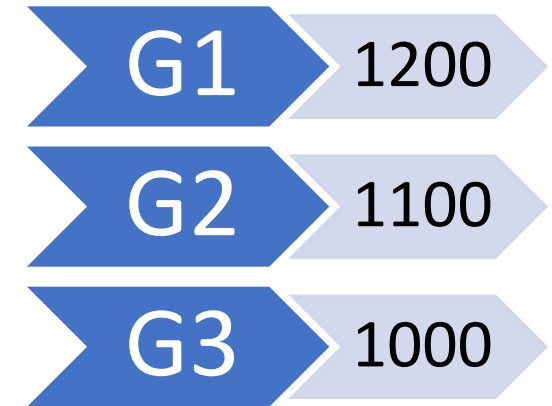
SMARTER PARTNERS



Thank you for your attention

www.smarterproject.eu

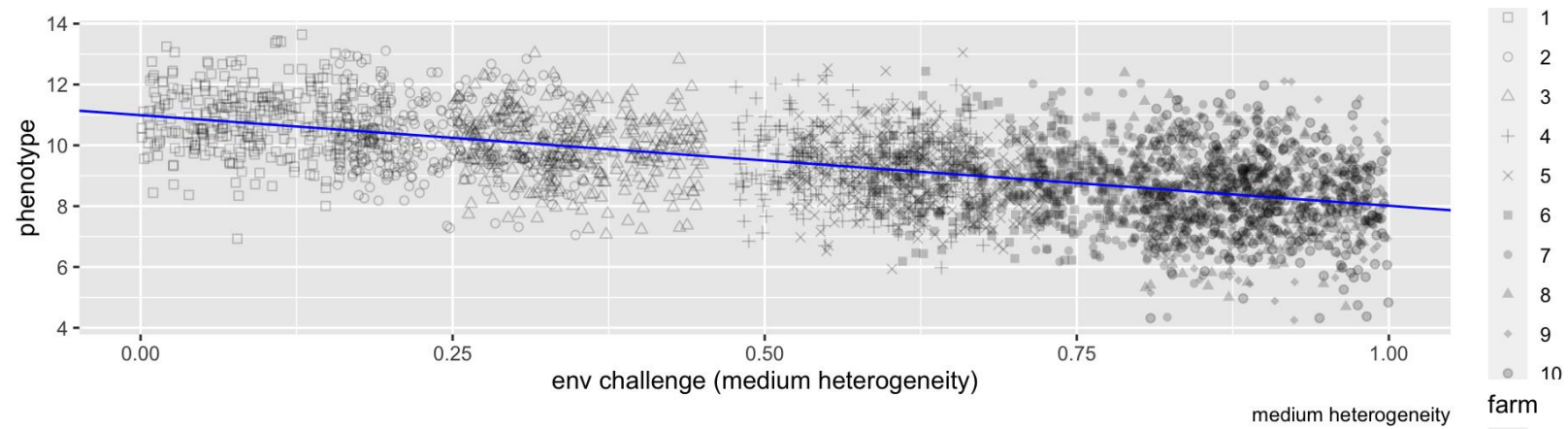
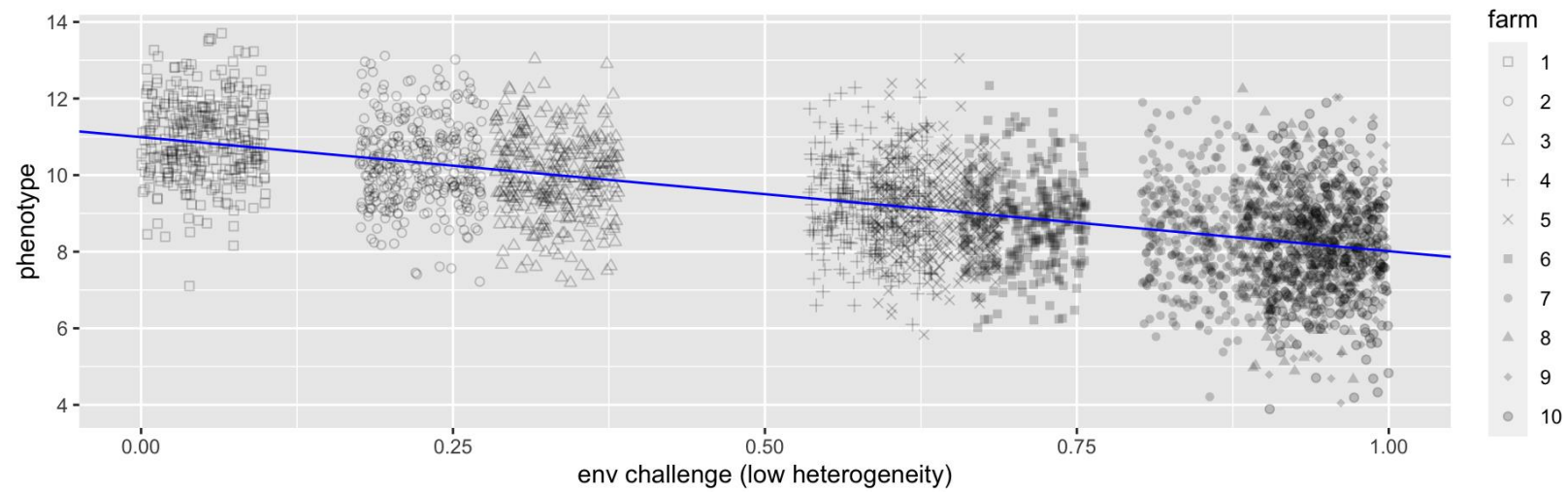
- 3300 individuals in 3 generations
- Id, sire, dam, farm, TBV
- 3 scenarios for challenge level from low, medium and high heterogeneity farm



```
> head(data)
```

	id	sire	dam	farm	x_lowHet	y_lowHet	x_midHet	y_midHet	x_highHet	y_highHet	tbv_r	tbv_0
1	1	0	0	4	0.56571469	10.229946	0.53642721	10.294395	0.4778523	10.423295	0.3010956	-0.0540811
2	2	0	0	8	0.91634493	5.128051	0.85521197	5.369157	0.7329461	5.851367	-0.1163798	-0.3482756
3	3	0	0	1	0.02677978	12.816753	0.05355957	12.754122	0.1071191	12.628859	-1.0049537	0.4142338
4	4	0	0	3	0.31620874	9.343520	0.31706821	9.340000	0.3187871	9.332961	0.1497635	0.3070735
5	5	0	0	3	0.36273375	10.959990	0.41011824	10.809085	0.5048872	10.507276	-0.0551632	0.1980272
6	6	0	0	9	0.92363147	7.151264	0.84820410	7.475782	0.6973493	8.124818	0.0173676	-0.0141196

```
> |
```



```

DATAFILE
phenotype_ex01.txt
# 1      2      3      4      5
# mean id x_value phenotype number_of_records
NUMBER_OF_TRAITS
1
NUMBER_OF_EFFECTS
6
OBSERVATION(S)
4
WEIGHT(S)

```

```

EFFECTS: POSITIONS_IN_DATAFILE NUMBER_OF_LEVELS TYPE_OF_EFFECT[EFFECT NESTED]
1 1 cross      # overall mean
3 1 cov        # covariate X
2 3300 cross   # random animal
3 3300 cov 2   # slope for individual
5 3300 cross   # pseudo-residual
3 3300 cov 5   # covariate nested to pseudo-residual

```

```
> head(df)
```

```

one id      env      phen n
1  1  1 0.56571469 10.229946 1
2  1  2 0.91634493  5.128051 2
3  1  3 0.02677978 12.816753 3
4  1  4 0.31620874  9.343520 4
5  1  5 0.36273375 10.959990 5
6  1  6 0.92363147  7.151264 6

```

```
> |
```

$$y_j = \mu_0 + \mu_1 t + u_{0j} + u_{1j} t + e_{0j} + e_{1j} x$$



	Accuracy of GEBV	Accuracy of EBV
resilice:	0.2132054	0.1730527
production potential:	0.3417898	0.259263



