



GxE with high-dimensional environmental data: correlated herd effects

Fernando Bussiman, D. Lourenco, C. Chen, J. Holl, I. Misztal, and Z. G. Vitezica



Animal Breeding and Genetics Group

College of Agricultural & Environmental Sciences

UNIVERSITY OF GEORGIA



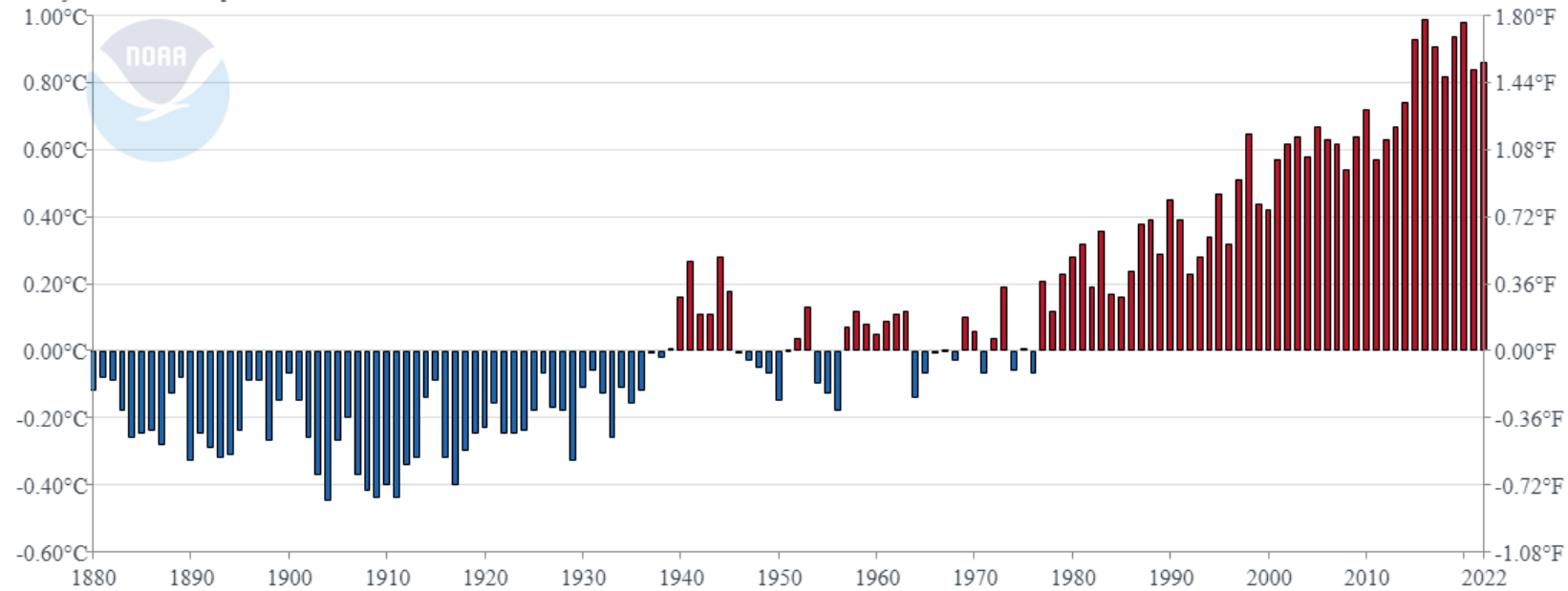
UNIVERSITY OF GEORGIA

fob@uga.edu 1

Introduction

➤ Nearly $+1^{\circ}\text{C}$ since 1980

Global Land and Ocean
January-December Temperature Anomalies

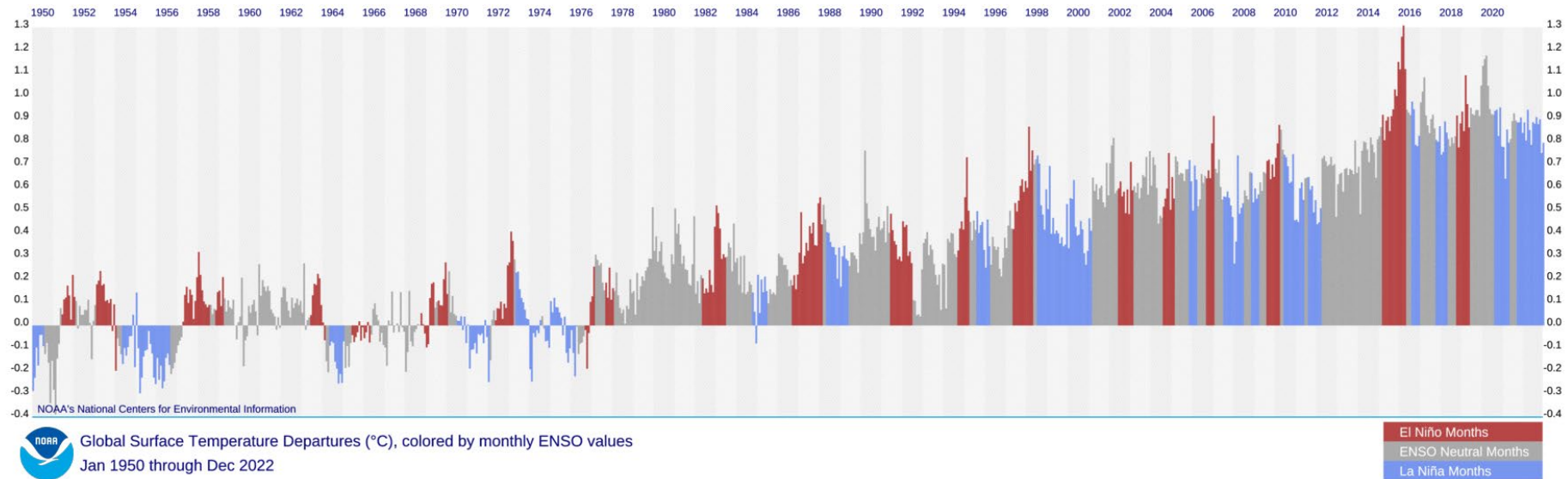


National Centers for Environmental Information:
<https://www.ncei.noaa.gov/access/monitoring/monthly-report/global/202213>



Introduction

➤ Nearly $+1^{\circ}\text{C}$ since 1980

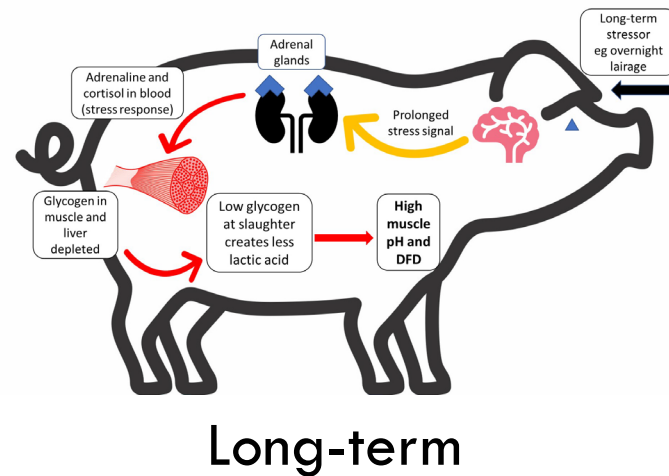
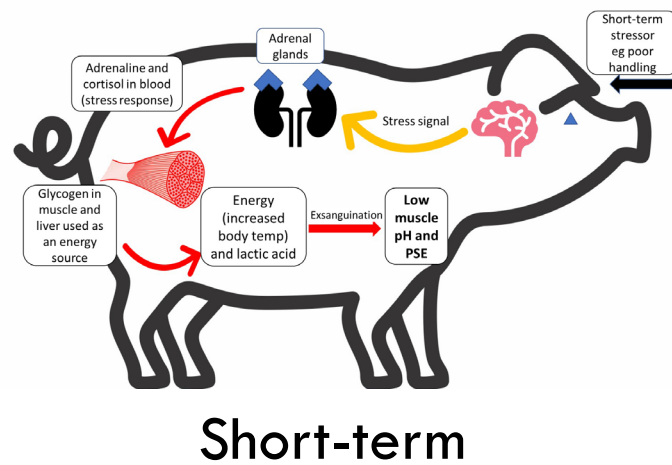


National Centers for Environmental Information:
<https://www.ncei.noaa.gov/access/monitoring/monthly-report/global/202213>



Introduction

- Environmental effect:
 - Seasonal productivity fluctuations
 - Heat Stress



➤ \$299 million loss
(St-Pierre et al., 2003)

➤ Welfare

➤ Diseases

The Pig Site:
<https://www.thepigsite.com/articles/what-are-the-impacts-of-stress-on-pork-quality>

Introduction

➤ Environmental effect:

➤ Seasonal productivity fluctuations

➤ Heat Stress

- Sweat glands are not stimulated (Ingram, 1967)
- Max 50 % heat production dissipated by respiratory evaporation (Renaudeau et al., 2012)
- Evidences of worsening with selection (Brown-Brandl et al., 2001)



Introduction

➤ Environmental effect:

- Seasonal productivity fluctuations

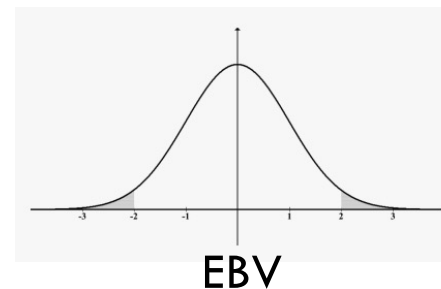
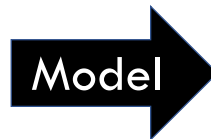
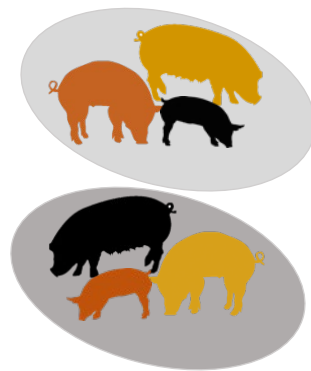
➤ Heat Stress

- Sweat glands are not stimulated (Ingram, 1967)

- Max 50 % heat production dissipated by respiratory evaporation (Renaudeau et al., 2012)

- Evidences of worsening with selection (Brwon-Brandl et al., 2001)

➤ Statistically



➤ Fixed

➤ Random

➤ Uncorrelated

Introduction

➤ What if:

- Herds are geographically close?
- Climate conditions are similar?
- Management/manager is the same?
- ...
- Correlated random herd effects

(Tiezzi et al., 2017; Selle et al., 2020; Cuyabano et al., 2021; Makanjuola et al., 2022)



Introduction

➤ Heat stress and climate effect:

- Temperature and Humidity → THI
- Heat Load → $f(\text{THI})$
 - Same THI for all locations – need to adapt (Bohmanova et al., 2007)
 - Temperature as good as THI (Dikmen and Hansen, 2009; Dado-Senn et al., 2023)

➤ Jarquín et al., 2014

- To accommodate environmental covariates (EC) by (co)variance structures
 - Reduces number of parameters
 - Better characterization of the environment



Objective

- To investigate GxE by using the (co)variances approach to model correlated herd effects and their impact on the prediction accuracy of genomic evaluation in pigs



Datasets Provided

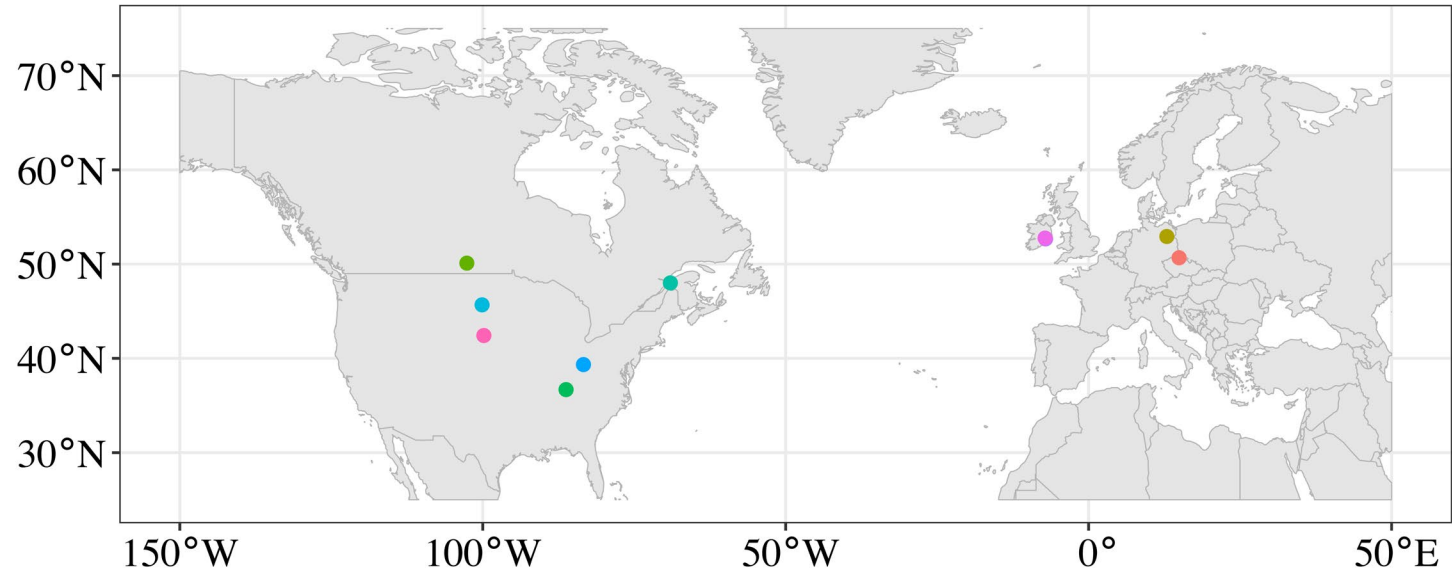
➤ Growth Purebred Pigs:

➤ Average Daily gain (ADG)

- 35,597 records
- All genotyped
- 11 farms

➤ Backfat Thickness (BFT)

- 32,105
- All genotyped
- 11 farms

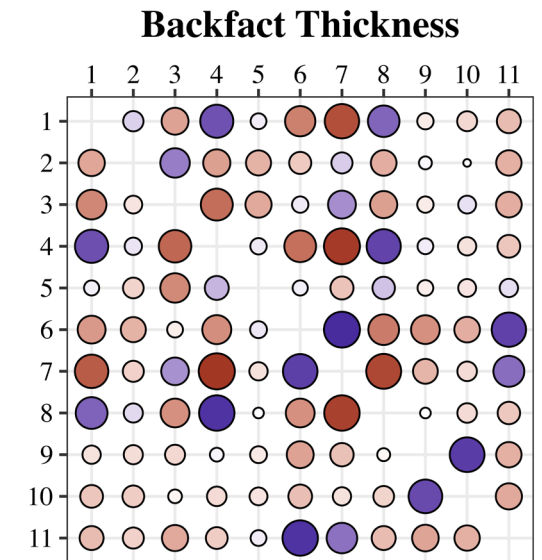
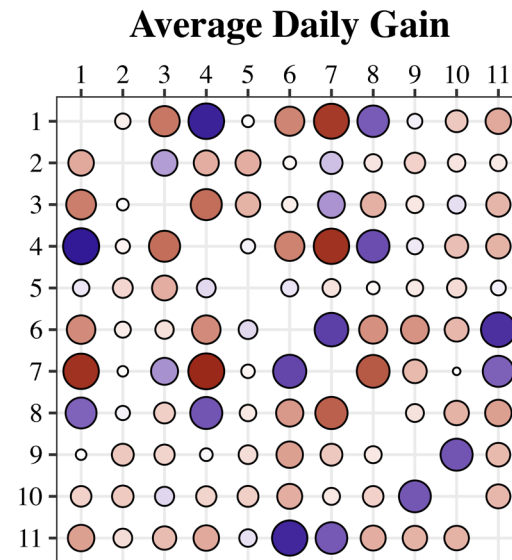
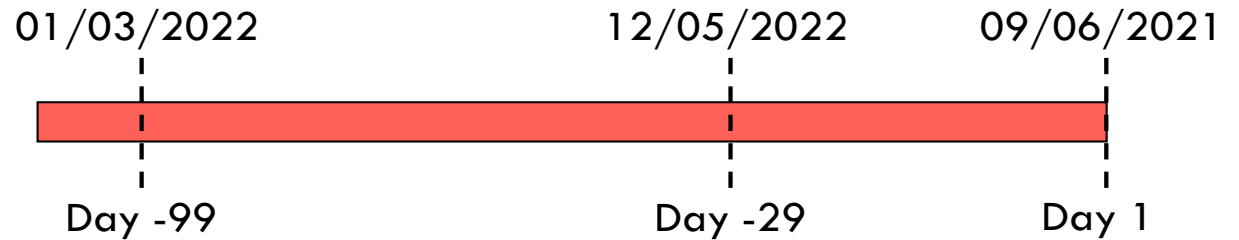


Environmental Data

➤ NASA POWER (<https://power.larc.nasa.gov/data-access-viewer/>) - EC

- T (Temperature - °C)
- Td (Dew/Frost temperature - °C)
- Tw (Wet-bulb temperature - °C)
- Ts (Earth surface temperature - °C)
- H (Relative humidity - %)
- R (Rainfall – mm/day)
- Ws (Wind speed – m/s)
- Md (Wind direction - °)

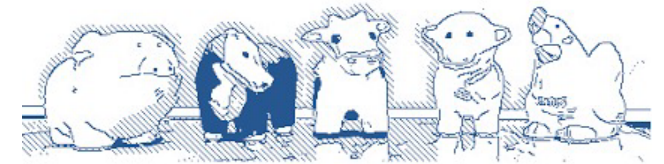
➤ Linear Regression every 10 days



Model of Analysis

➤ $ADG_{ijkl} = \mu + CG_i + l_j + g_k + e_l + \epsilon_{ijkl}$

➤ $BFT_{ijkl} = \mu + \beta_1 EW_k + CG_i + l_j + g_k + e_l + \epsilon_{ijkl}$



BLUPF90 Programs

➤ $y = X\beta + Z_1l + Z_2g + Z_3e + \epsilon$

➤ $l \sim MVN(0, I\sigma_l^2)$

➤ $g \sim MVN(0, G\sigma_g^2)$

➤ $e \sim MVN(0, I\sigma_e^2)$

➤ $\epsilon \sim MVN(0, I\sigma_\epsilon^2)$



$e \sim MVN(0, E_{30}\sigma_e^2)$



$e \sim MVN(0, E_{100}\sigma_e^2)$

MG

ME₃₀

ME₁₀₀



Validation

- Focal animals - born in 2020

- One whole (w, from 2009 to 2020) and one partial (p, from 2009 to 2019) datasets

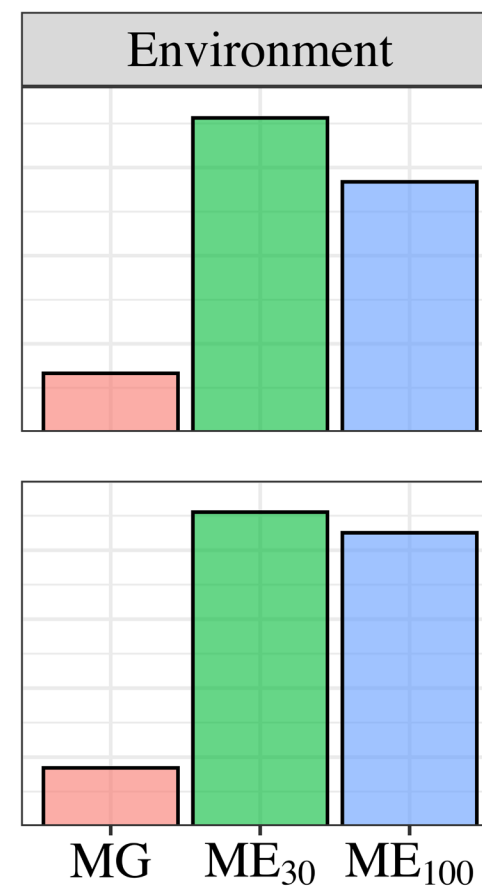
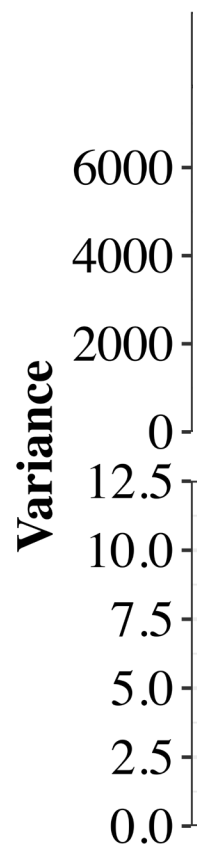
- LR (Legarra and Reverter, 2018)

- $\text{acc} = \sqrt{\text{cov}(\hat{\mathbf{g}}_p, \hat{\mathbf{g}}_w) / (1 - \bar{F}) \sigma_g^2}$

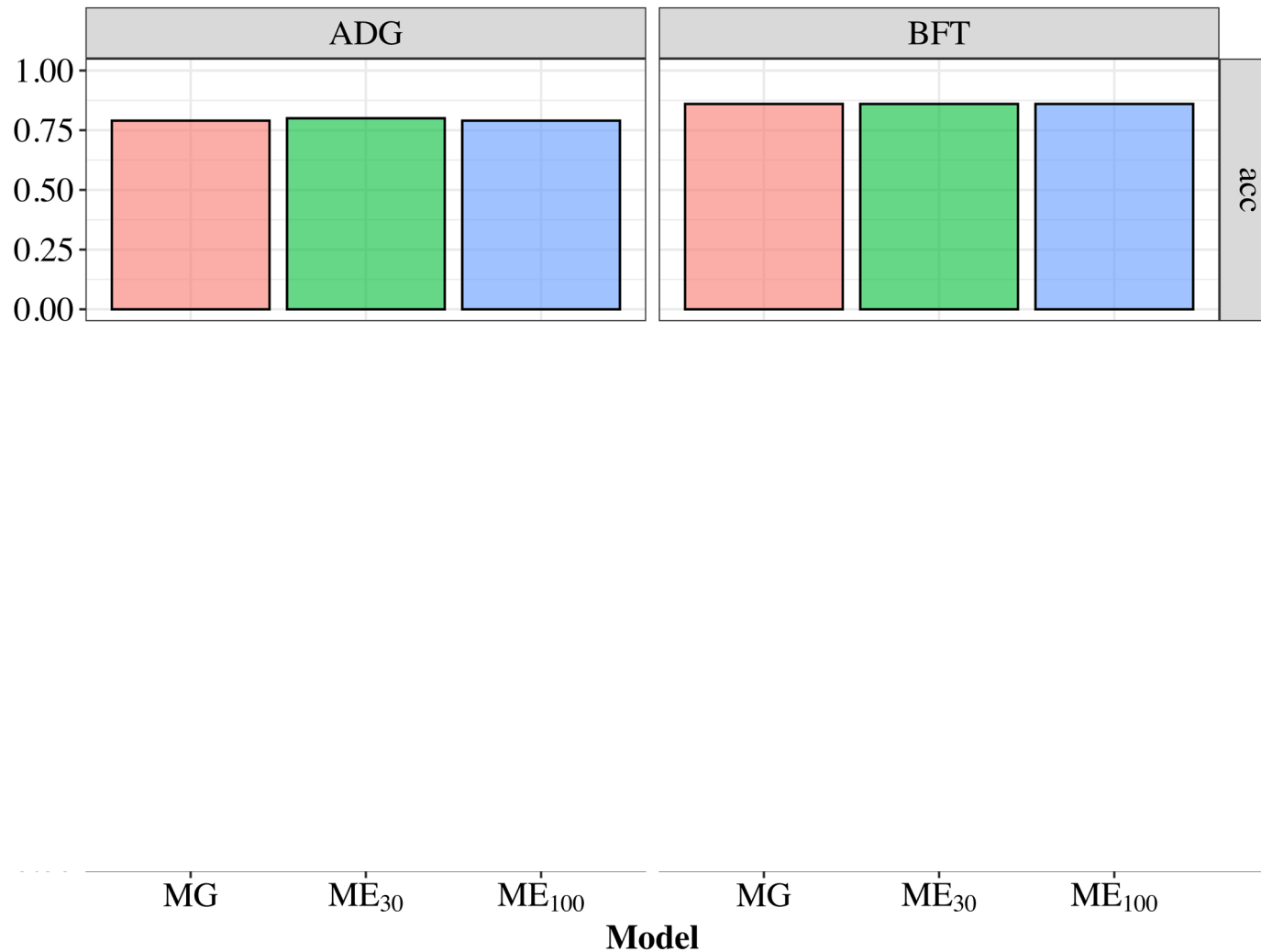
- $\delta = (\bar{\hat{\mathbf{g}}}_p - \bar{\hat{\mathbf{g}}}_w) / \sigma_g$

- $b_1 = \text{cov}(\hat{\mathbf{g}}_p, \hat{\mathbf{g}}_w) / \text{var}(\hat{\mathbf{g}}_p)$

Variance Components



Validation Statistics



- No improvement in acc, bias, and dispersion
- What if within environment?
 - No improvement also

Model of Analysis

$$\text{Var} \begin{bmatrix} \mathbf{l} \\ \mathbf{g} \\ \mathbf{e} \\ \boldsymbol{\epsilon} \end{bmatrix} = \begin{bmatrix} \mathbf{I}\sigma_1^2 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}\sigma_g^2 & \text{Cov}(\mathbf{g}, \mathbf{e}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{E}_i\sigma_e^2 & \mathbf{0} \\ \text{sym.} & \mathbf{0} & \mathbf{0} & \mathbf{I}\sigma_\epsilon^2 \end{bmatrix}$$

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_1\mathbf{l} + \mathbf{Z}_2\mathbf{g} + \mathbf{Z}_3\mathbf{e} + \mathbf{Z}_4\mathbf{ge} + \boldsymbol{\epsilon}$$

$$\mathbf{l} \sim \text{MVN}(\mathbf{0}, \mathbf{I}\sigma_1^2)$$

$$\mathbf{g} \sim \text{MVN}(\mathbf{0}, \mathbf{G}\sigma_g^2)$$

$$\mathbf{e} \sim \text{MVN}(\mathbf{0}, \mathbf{E}_i\sigma_e^2)$$

$$\mathbf{ge} \sim \text{MVN}[\mathbf{0}, (\mathbf{G} \odot \mathbf{Z}_4 \mathbf{E}_i \mathbf{Z}_4') \sigma_{ge}^2]$$

$$\boldsymbol{\epsilon} \sim \text{MVN}(\mathbf{0}, \mathbf{I}\sigma_\epsilon^2)$$



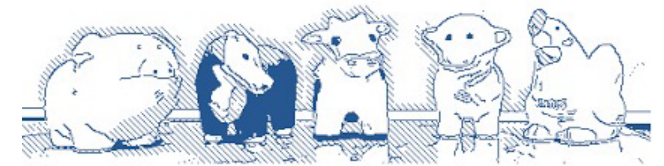
$$\mathbf{e} \sim \text{MVN}(\mathbf{0}, \mathbf{E}_{30}\sigma_e^2)$$

$$\mathbf{ge} \sim \text{MVN}[\mathbf{0}, (\mathbf{G} \odot \mathbf{Z}_4 \mathbf{E}_{30} \mathbf{Z}_4') \sigma_{ge}^2] \quad \text{MGE}_{30}$$



$$\mathbf{e} \sim \text{MVN}(\mathbf{0}, \mathbf{E}_{100}\sigma_e^2)$$

$$\mathbf{ge} \sim \text{MVN}[\mathbf{0}, (\mathbf{G} \odot \mathbf{Z}_4 \mathbf{E}_{100} \mathbf{Z}_4') \sigma_{ge}^2] \quad \text{MGE}_{100}$$



BLUPF90 Programs

$$= (\mathbf{G} \odot \mathbf{Z}_4 \mathbf{E}_i \mathbf{Z}_4') \sigma_{ge}^2$$

Two Questions



Two Questions

➤ Is there GxE?

➤ MTM

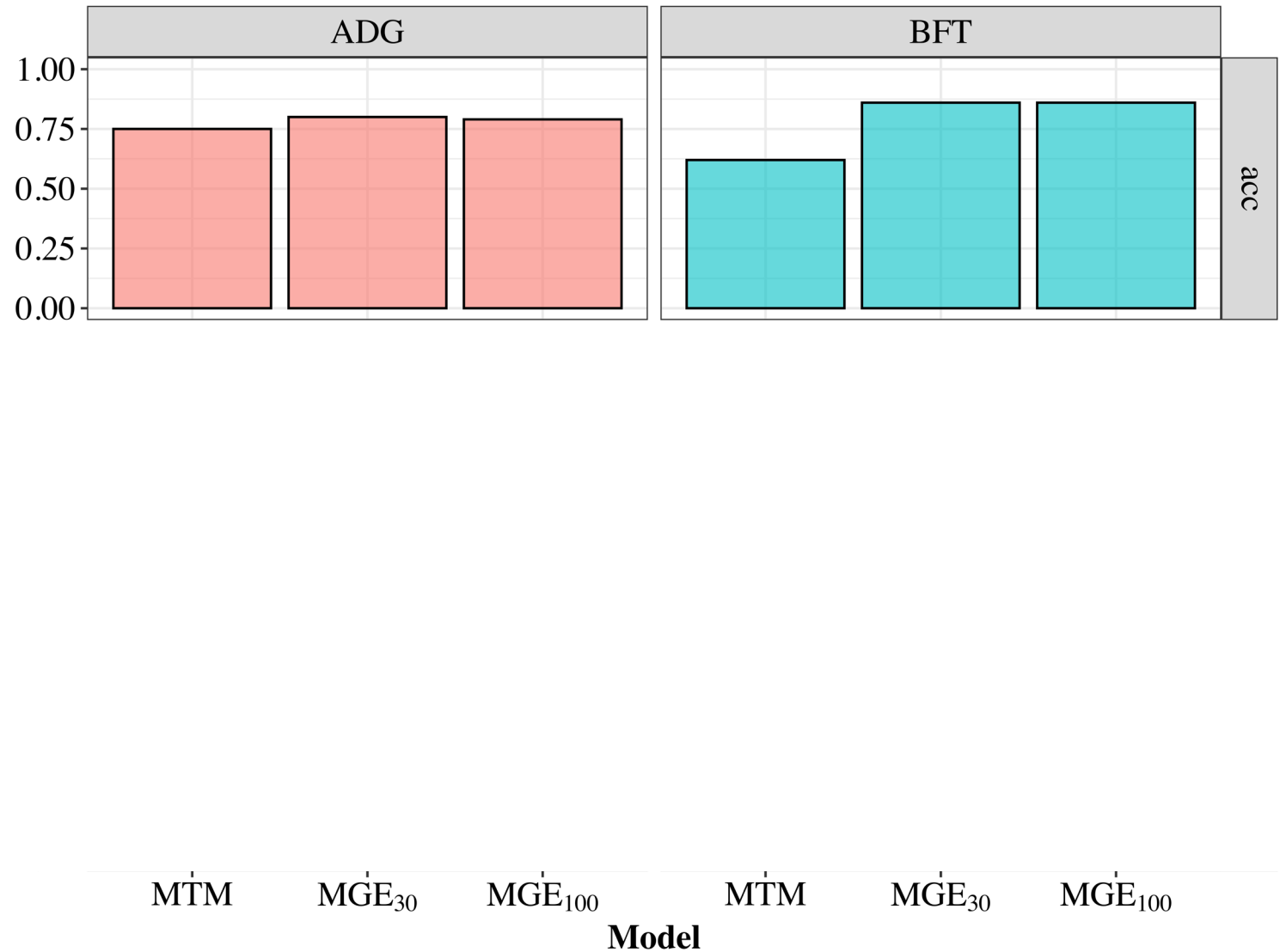
➤ $r_{geiej} < 0.80$

➤ Bending

➤ Number of records/envi

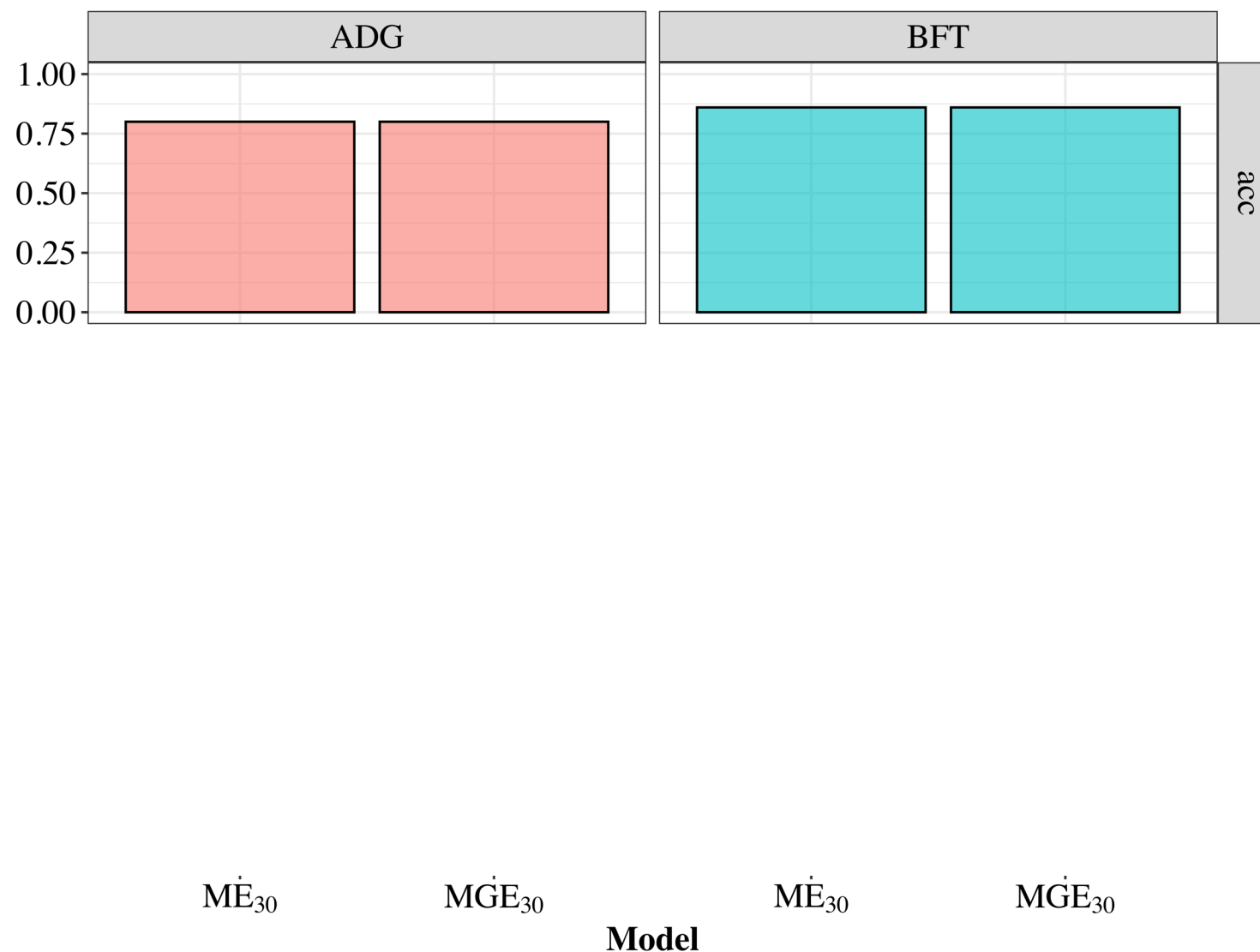
➤ Feeding and measureme

➤ Can we improve accuracy by including GxE?



ME_{30} vs MGE_{30}

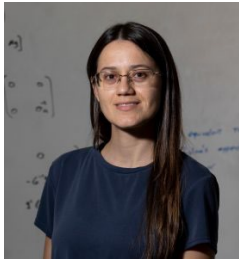
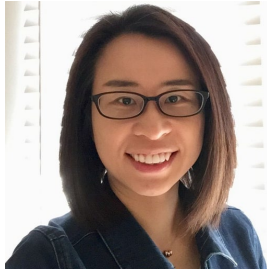
- No improvement in acc, bias, and dispersion
- $MGE = ME$?
 - Possibly no reranking



Remarks

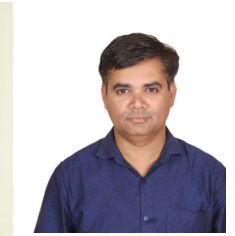
- The (co)variances approach could increase accuracy when environment accounts for a lower proportion of phenotypic variance
 - fixed herd effect is usually enough
- Considering GxE where E is a correlated effect does not improve accuracy
- (Co)variances approach had higher accuracy than MTM for BFT
- This model provides a “ge” breeding value – specific environmental change on the genotype (recalling MGE_i are analogous to a reaction norm, “g” acts as a b_0 and “ge” as a b_1)
- Overall...
 - Using outdoor EC to correlate environments has little benefit for genomic predictions
 - MGE would be better if “g” was observed in each different “e”
 - MTM performs as good if there is borrowing of information

Acknowledgements



Animal Breeding and Genetics Group
College of Agricultural & Environmental Sciences
UNIVERSITY OF GEORGIA

UGA – ABG Group
<http://nce.ads.uga.edu>



UNIVERSITY OF GEORGIA

fob@uga.edu 21