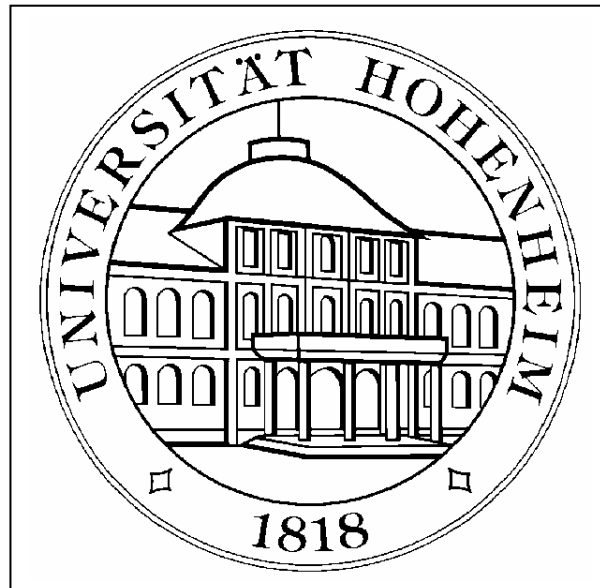


# Single-stage versus stage-wise analysis

Hans-Peter Piepho

Biostatistics Unit  
Universität Hohenheim  
Germany



# 1. Introduction

- Series of experiments over **environments** commonly analysed by mixed model
- Most efficient method: Single-stage analysis based on replicate data
- This can be computationally very demanding:
  - many environments and genotypes
  - different designs in different trials
  - complex mixed models
- Popular alternative: **two-step approach**
  - Step 1: Adjusted genotype means for each trial (environment)
  - Step 2: Fit mixed model to means from Step 1 jointly for series of trials

## 2. Some theory

A mixed model for plot data:

$$y_j = X_{ab(j)}\beta_{ab} + Z_{b(j)}u_{b(j)} + Z_{w(j)}u_{w(j)} + e_{w(j)},$$

$y_j$  = observed data vector of plot yields in  $j$ -th environment

$\beta_{ab}$  = fixed **across**-environment and **between**-environment effects

$u_{b(j)}$  = random **between**-environment effects with  $\text{var}(u_{b(j)}) = G_{b(j)}$

$u_{w(j)}$  = random **within**-environment random design effects for replicates  
blocks, etc. with  $\text{var}(u_{w(j)}) = G_{w(j)}$

$e_{w(j)}$  = plot errors with  $\text{var}(e_{w(j)}) = R_{w(j)}$

## 2. Some theory

Same model in more compact form

$$y_j = X_{ab(j)}\beta_{ab} + Z_{b(j)}u_{b(j)} + f_{w(j)},$$

where  $f_{w(j)} = Z_{w(j)}u_{w(j)} + e_{w(j)}$  and

$$\text{var}(f_{w(j)}) = \Sigma_{w(j)} = Z_{w(j)}G_{w(j)}Z_{w(j)}^T + R_{w(j)}.$$

## 2. Some theory

### Stacking environments

$$y = X_{ab}\beta_{ab} + Z_b u_b + f_w$$

where

$$X_{ab}^T = \left( X_{ab(1)}^T, X_{ab(2)}^T, \dots, X_{ab(q)}^T \right)$$

$$Z_b = \bigoplus_{j=1}^q Z_{b(j)}$$

$$u_b^T = \left( u_{b(1)}^T, u_{b(2)}^T, \dots, u_{b(q)}^T \right)$$

$$f_w^T = \left( f_{w(1)}^T, f_{w(2)}^T, \dots, f_{w(q)}^T \right)$$

Variance-covariance structure is block-diagonal:

$$\text{var}(f_w) = \Sigma_w = \bigoplus_{j=1}^q \Sigma_{w(j)} \quad \text{and} \quad \text{var}(u_b) = G_b = \bigoplus_{j=1}^q G_{b(j)}$$

## 2. Some theory

### Mixed model equations (MME) 1-stage model

$$\begin{pmatrix} X_{ab}^T \Sigma_w^{-1} X_{ab} & X_{ab}^T \Sigma_w^{-1} Z_b \\ Z_b^T \Sigma_w^{-1} X_{ab} & Z_b^T \Sigma_w^{-1} Z_b + G_b^{-1} \end{pmatrix} \begin{pmatrix} \hat{\beta}_{ab} \\ \hat{u}_b \end{pmatrix} = \begin{pmatrix} X_{ab}^T \Sigma_w^{-1} y \\ Z_b^T \Sigma_w^{-1} y \end{pmatrix}$$

## 2. Some theory

### Two-stage representation of single-stage model

$$\text{Stage 1: } y_j = X_{1(j)}\eta_{1(j)} + f_{w(j)},$$

$\eta_{1(j)}$  = *composite* genotype-environment effects

$$\text{Stage 2: } \eta_{1(j)} = X_{2(j)}\beta_{ab} + Z_{2(j)}u_{b(j)},$$

Plugging last into first equation, it follows that

$$y_j = X_{1(j)}X_{2(j)}\beta_{ab} + X_{1(j)}Z_{2(j)}u_{b(j)} + f_{w(j)}$$

$\Rightarrow$  same as 1-stage model with  $X_{ab(j)} = X_{1(j)}X_{2(j)}$  and  $Z_{b(j)} = X_{1(j)}Z_{2(j)}$ !

## 2. Some theory

Main challenge in devising a two-stage procedure

Must verify that for each environment

$$E(y_j | \eta_{1(j)}) = X_{1(j)}\eta_{1(j)} = X_{1(j)}X_{2(j)}\beta_{ab} + X_{1(j)}Z_{2(j)}u_{b(j)}$$

for some design matrix  $X_{1(j)}$ .

⇒ Estimator of  $\eta_{1(j)}$  contains all information on  
between-environment mean structure



## 2. Some theory

### Estimation in stage 1

$$\hat{\eta}_{1(j)} = \left( X_{1(j)}^T \Sigma_{w(j)}^{-1} X_{1(j)} \right)^{-1} X_{1(j)}^T \Sigma_{w(j)}^{-1} y_j \quad \text{with}$$

$$\text{var}\left(\hat{\eta}_{1(j)} \mid u_{b(j)}\right) = \left( X_{1(j)}^T \Sigma_{w(j)}^{-1} X_{1(j)} \right)^{-1} \quad (\text{composite effect taken as fixed here!})$$

### Stage-2 model for stage-1 estimator

$$\hat{\eta}_{1(j)} = X_{2(j)} \beta_{ab} + Z_{2(j)} u_{b(j)} + e_{b(j)},$$

where  $e_{b(j)} = \left( X_{1(j)}^T \Sigma_{w(j)}^{-1} X_{1(j)} \right)^{-1} X_{1(j)}^T \Sigma_{w(j)}^{-1} f_{w(j)}$  and

$$\text{var}\left(e_{b(j)}\right) = \left( X_{1(j)}^T \Sigma_{w(j)}^{-1} X_{1(j)} \right)^{-1}.$$

## 2. Some theory

### Rotation

Rotated (orthogonalized) composite effects

$$\tilde{\eta}_{1(j)} = W_{1(j)} \hat{\eta}_{1(j)},$$

where

$$W_{1(j)} = S_j \Lambda_j^{-1/2} S_j^T$$

from spectral decomposition

$$\text{var}(e_{b(j)}) = \left( X_{1(j)}^T \Sigma_{w(j)}^{-1} X_{1(j)} \right)^{-1} = S_j \Lambda_j S_j^T$$

## 2. Some theory

### Model for rotated composite effects

$$\tilde{\eta}_{1(j)} = \tilde{X}_{2(j)}\beta_{ab} + \tilde{Z}_{2(j)}u_{b(j)} + \tilde{e}_{b(j)},$$

where

$$\tilde{X}_{2(j)} = W_{1(j)}X_{2(j)}, \quad \tilde{Z}_{2(j)} = W_{1(j)}Z_{2(j)} \quad \text{and} \quad \tilde{e}_{b(j)} = W_{1(j)}e_{b(j)}$$

with

$$\text{var}(\tilde{e}_{b(j)}) = W_{1(j)} \text{var}(e_{b(j)})W_{1(j)}^T = I.$$

## 2. Some theory

### Stacking environments for unrotated model

$$\hat{\eta}_1 = X_2 \beta_{ab} + Z_2 u_b + e_b ,$$

where

$$\hat{\eta}_1^T = \left( \hat{\eta}_{1(1)}^T, \hat{\eta}_{1(2)}^T, \dots, \hat{\eta}_{1(q)}^T \right)$$

$$X_2^T = \left( X_{2(1)}^T, X_{2(2)}^T, \dots, X_{2(q)}^T \right)$$

$$Z_2 = \bigoplus_{j=1}^q Z_{2(j)}$$

$$e_b^T = \left( e_{b(1)}^T, e_{b(2)}^T, \dots, e_{b(q)}^T \right)$$

## 2. Some theory

Stacking environments for rotated model

$$\tilde{\eta}_1 = \tilde{X}_2 \beta_{ab} + \tilde{Z}_2 u_b + \tilde{e}_b ,$$

where

$$\tilde{\eta}_1^T = \left( \tilde{\eta}_{1(1)}^T, \tilde{\eta}_{1(2)}^T, \dots, \tilde{\eta}_{1(q)}^T \right)$$

$$\tilde{X}_2^T = \left( \tilde{X}_{2(1)}^T, \tilde{X}_{2(2)}^T, \dots, \tilde{X}_{2(q)}^T \right)$$

$$\tilde{Z}_2 = \bigoplus_{j=1}^q \tilde{Z}_{2(j)}$$

$$\tilde{e}_b^T = \left( \tilde{e}_{b(1)}^T, \tilde{e}_{b(2)}^T, \dots, \tilde{e}_{b(q)}^T \right)$$

## 2. Some theory

Computation of rotated means can be written compactly as

$$\tilde{\eta}_1 = W_1 \hat{\eta}_1,$$

$$\text{where } W_1 = \bigoplus_{j=1}^q W_{1(j)}$$

**MME for rotated model**

$$\begin{pmatrix} \tilde{X}_2^T \tilde{X}_2 & \tilde{X}_2^T \tilde{Z}_2 \\ \tilde{Z}_2^T \tilde{X}_2 & \tilde{Z}_2^T \tilde{Z}_2 + G_b^{-1} \end{pmatrix} \begin{pmatrix} \beta_{ab} \\ u_b \end{pmatrix} = \begin{pmatrix} \tilde{X}_2^T \tilde{\eta}_1 \\ \tilde{Z}_2^T \tilde{\eta}_1 \end{pmatrix}.$$

## 2. Some theory

Using  $\tilde{X}_2 = W_1 X_2$  and  $\tilde{Z}_2 = W_1 Z_2$ , these can be re-arranged as

$$\begin{pmatrix} X_2^T X_1^T \Sigma_w^{-1} X_1 X_2 & X_2^T X_1^T \Sigma_w^{-1} X_1 Z_2 \\ Z_2^T X_1^T \Sigma_w^{-1} X_1 X_2 & Z_2^T X_1^T \Sigma_w^{-1} X_1 Z_2 + G_b^{-1} \end{pmatrix} \begin{pmatrix} \hat{\beta}_{ab} \\ \hat{u}_b \end{pmatrix} = \begin{pmatrix} X_2^T X_1^T \Sigma_w^{-1} y \\ Z_2^T X_1^T \Sigma_w^{-1} y \end{pmatrix}$$

Observing that  $X_{ab} = X_1 X_2$  and  $Z_b = X_1 Z_2$ , this is seen to be equivalent to the 1-stage MME!

(Piepho, Möhring, Schulz-Streeck & Ogutu, 2012)

## 2. Some theory

### Three-stage analysis

- (1) Compute summaries per environment (adjusted means, regression slopes)
- (2) Average across environments
- (3) Fit model to across-environment summaries
  - mixed model
  - other models / methods, e.g. for QTL mapping, genomic selection (boosting, random forests, SVM, etc)

⇒ apply same approach to three stages



# 3. Examples

## Example 1: Exploiting pedigree information

- Series of trials with **sugar beet** from a hybrid breeding program
  - 831 **inbred lines** were **test-crossed** to testers of the opposite pool
  - A **single tester** for each inbred line
  - Six common **check varieties**
  - The test crosses were **split into groups** tested in separate series of trials
  - Trials performed in a series of **6 × 6 simple lattice designs**
  - **Each series of trials** was replicated at **11 or 12 environments**
  - Total of **21 environments** in 2002
  - Trials were **connected by check varieties**
  - **Imbalance** arises from the facts that
    - not all trials were conducted in all locations
    - checks were replicated more often than were test crosses
    - test crosses were subdivided among the series of lattice experiments
- ⇒ **Pedigree information to be exploited by BLUP of genotype main effect**

# 3. Examples

## Multiple lattices

### Replicate 1

	Lattice 1					Lattice 2					Lattice 3					Lattice 4					Lattice 5				
Block	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	14	6	C1	C3	11	23	33	39	35	29	C1	C5	42	57	50	C2	71	80	66	68	C1	85	91	C3	82
	1	C4	12	10	C2	C4	34	C5	27	37	58	60	55	C2	43	76	74	67	63	C4	93	C2	86	90	C5
	9	18	2	8	4	24	22	30	C2	C1	53	56	59	41	45	64	C1	C5	77	61	98	84	100	97	87
	13	5	19	15	3	26	21	38	32	25	49	51	46	44	47	70	78	69	C3	75	83	88	C4	94	89
	C5	17	7	16	20	40	C3	28	31	36	52	58	C3	54	C4	62	73	65	79	72	92	99	95	81	96

### Replicate 2

	Lattice 1					Lattice 2					Lattice 3					Lattice 4					Lattice 5				
Block	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
	C3	15	17	13	2	22	C3	27	C4	C2	C2	53	44	C4	58	C1	68	61	73	64	100	87	92	91	C5
	12	C4	7	C1	1	24	37	21	29	34	50	60	42	52	C5	72	69	66	67	75	90	94	96	98	95
	3	4	C5	11	C2	36	35	38	32	23	55	41	45	56	59	76	74	C5	C4	71	89	86	88	82	83
	9	19	16	18	5	C5	26	C1	33	30	51	C3	48	54	57	65	C2	62	70	C3	85	C2	C4	C3	99
	6	14	20	8	10	31	28	40	39	25	49	43	C1	46	47	79	63	78	67	80	C1	93	81	84	97

**Figure:** 5 simple lattices (5 x 5) with block size  $k = 5$  for  $v = 100$  genotypes and **5 checks (C1-C5)**

### 3. Examples

Design factors:

**env** = environment

**lat** = lattice

**rep** = replicate

**blk** = incomplete block

Treatment factors:

**chk** = checks and population of test crosses

**test** = testers

**cross** = test crosses

### 3. Examples

#### Single-stage model

Genotype main effect (**gen**) was modelled by

**gen = test + chk : chk•cross**

Between- and within environment effects were modelled by

**env + env•lat + env•lat•rep + env•lat•rep•blk  
+ test•env + chk•env + chk•cross•env**

Notation: **Patterson (1997)**

⇒ **chk•cross** and **chk•cross•env** have variance-covariance matrix proportional to **numerator relationship matrix** (**A** matrix; computed from pedigree)

### 3. Examples

#### Three-stage analysis

Stage one (within environment effects):

$gen : lat + lat \bullet rep + lat \bullet rep \bullet blk$

Stage two (between environment effects):

$gen : env + test \bullet env + chk \bullet env + chk \bullet cross \bullet env$

Stage three (across environment effects):

$gen = test + chk : chk \bullet cross$

⇒  $chk \bullet cross$  has covariance proportional to **numerator relationship matrix**

# 3. Examples

## Single-stage model

$$y = X_a \beta_a + Z_a u_a + Z_b u_b + Z_w u_w + e_w,$$

where

$\beta_a$  = vector of fixed **across**-environment effects

$u_a$  = vector of random **across**-environment effects with  $\text{var}(u_a) = G_a$

$u_b$ ,  $u_w$  and  $e_w$  are defined as in two-stage analysis

## Hierarchical three-stage representation

$$y = X_1 \eta_1 + f_w \quad \text{with} \quad f_w = Z_w u_w + e_w,$$

$$\eta_1 = X_2 \eta_2 + Z_2 u_b, \quad \text{and}$$

$$\eta_2 = X_3 \beta_a + Z_3 u_a.$$

### 3. Examples

**Table 1:** Key elements of computations in stages 1 and 2 of a 3-stage analysis.

Element	Stage	
	1	2
Linear model for computed means	$\hat{\eta}_1 = X_2 \eta_2 + Z_2 u_b + e_b$	$\hat{\eta}_2 = X_3 \beta_a + Z_3 u_a + e_a$
Error variance for computed means	$\text{var}(e_b) = (X_1^T \Sigma_w^{-1} X_1)^{-1}$	$\text{var}(e_a) = (\tilde{X}_2^T \Sigma_b^{-1} \tilde{X}_2)^{-1}$
Weighting matrix	$\Sigma_w = Z_w G_w Z_w^T + R_w$	$\Sigma_b = \tilde{Z}_b G_b \tilde{Z}_b^T + I$
Decomposition for rotation	$X_1^T \Sigma_w^{-1} X_1 = W_1^2$	$\tilde{X}_2^T \Sigma_b^{-1} \tilde{X}_2 = W_2^2$
Rotation	$\tilde{\eta}_1 = W_1 \hat{\eta}_1$	$\tilde{\eta}_2 = W_2 \hat{\eta}_2$
Rotated model	$\tilde{\eta}_1 = \tilde{X}_2 \mu_2 + \tilde{Z}_2 u_b + \tilde{e}_b$	$\tilde{\eta}_2 = \tilde{X}_3 \beta_a + \tilde{Z}_3 u_a + \tilde{e}_a$

### 3. Examples

**Table 2:** Results of 1-stage, 2-stage and 3-stage analyses for sugar beet data. Objective: BLUPs for **random main test cross effects**.

Comparison of test cross means	2-stage using diagonal weights	3-stage using diagonal weights	3-stage using rotation
Correlations with estimates from 1-stage analysis	0.9864	0.9862	0.9954
Mean squared error compared to estimates from 1-stage analysis	22.41	30.53	8.26



# 3. Examples

## Example 2: Genomic selection

- Series of trials with **maize** from a hybrid breeding program
- 177 DH-lines were **test-crossed** to one testers of the opposite pool and genotyped with 768 **SNP markers**, where 275 markers are polymorphic
- Two common **check varieties**
- An **augmented design** with three to five incomplete blocks was used
- Incomplete blocks were connected by check varieties
- The test-crosses were replicated at **six locations**
- Imbalance arises from the facts that
  - not all test crosses were tested in all locations
  - checks were replicated more often than test crosses
  - test crosses tested within the same or different incomplete block

### 3. Examples

Ridge-regression BLUP:

$$u_a = Mv$$

where

$u_a$  = genotype main effects

$M$  = marker matrix

$v$  = SNP effects with  $\text{var}(v) = I\sigma_a^2$

$$\Rightarrow \text{var}(u_a) = \sigma_a^2 MM^T$$

### 3. Examples

**Ridge-regression for plot data (single-stage analysis):**

$$y = 1\beta_a + Z_a u_a + Z_{b(2)} u_{b(1)} + Z_{b(2)} u_{b(2)} + Z_w u_w + e,$$

$y$  = observed data vector of plot yields

$1$  = vector with ones

$\beta_a$  = intercept

$u_a$  = random genotype main effect with  $\text{var}(u_a) = \sigma_a^2 M M^T$

$u_{b(1)}$  = random environment main effect with  $\text{var}(u_b) = I \sigma_b^2$

$u_{b(2)}$  = random genotype by environment effect with  $\text{var}(u_c) = I \sigma_c^2$

$u_w$  = random incomplete block effect with  $\text{var}(u_d) = I \sigma_d^2$

$e$  = plot errors with  $\text{var}(e) = I \sigma_e^2$

# 3. Examples

## Two-stage analysis

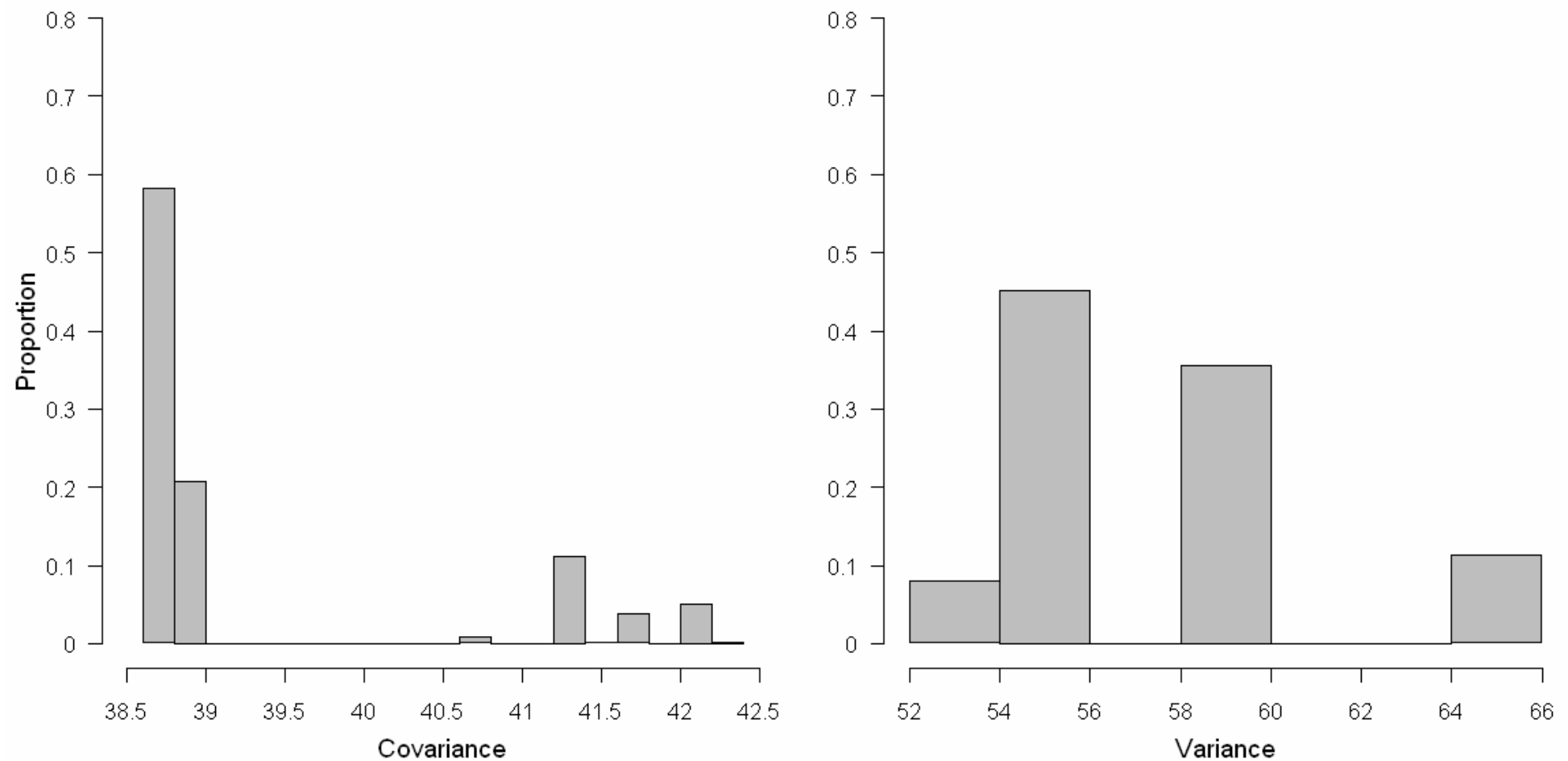
### First stage:

- Compute genotype means across locations
- Genotype main effects modelled as fixed

### Second stage:

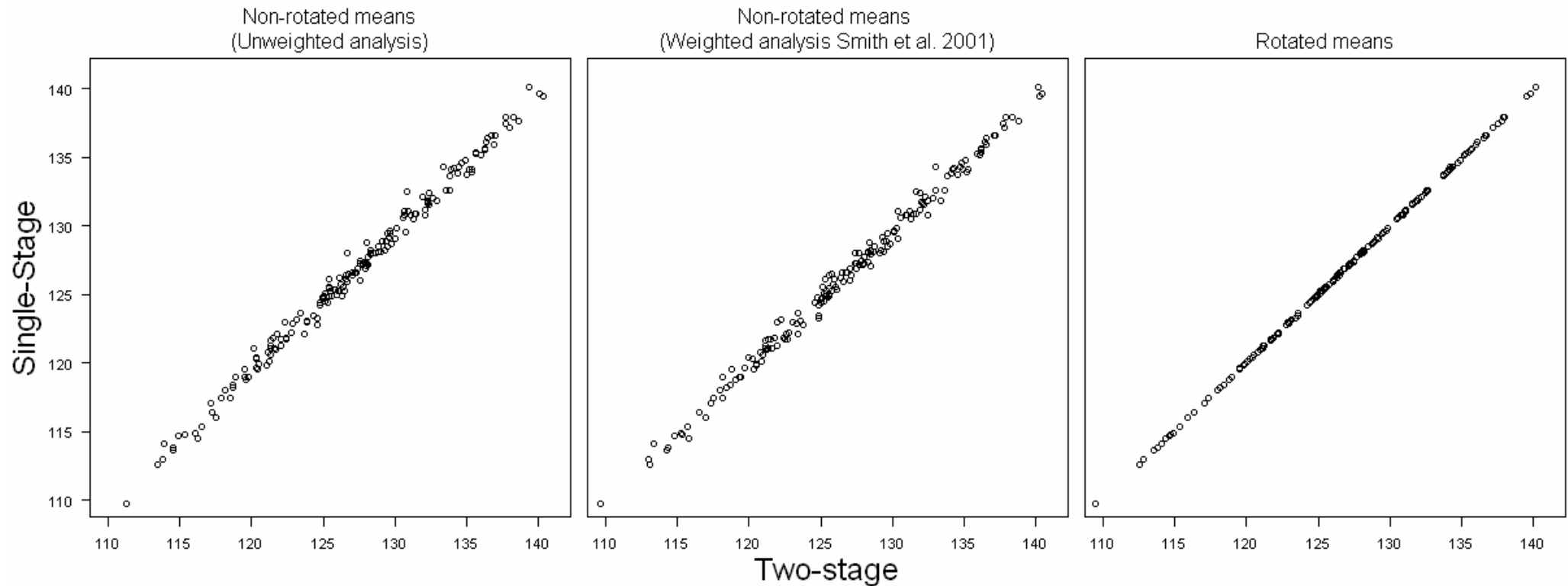
- Compute RR-BLUP for genotype main effects
- Genotype main effects modelled as random

### 3. Examples



**Figure:** Covariances and variances of adjusted means (maize data).

### 3. Examples



**Figure:** Comparisons of the predictions of the single-stage approach with predictions of the three two-stage approaches (maize data).  
(Schulz-Streeck, Ogutu & Piepho, 2013)

## 4. Summary

- Can make stage-wise and single-stage analysis agree very closely
- Stage-wise analysis computationally less demanding
- Easier to account for specifics of trial designs in different environments
- Similar to use of DYD and deregressed proofs in animal breeding

## 5. References

- Piepho, H.P., Möhring, J., Schulz-Streeck, T., Ogutu, J.O. (2012): A stage-wise approach for analysis of multi-environment trials. *Biometrical Journal* **54**, 844-860.
- Schulz-Streeck, T., Ogutu, J.O., Piepho, H.P. (2013): Comparisons of single-stage and two-stage approaches to genomic selection using ridge regression BLUP and componentwise boosting. *Theoretical and Applied Genetics* **126**, 69-82.



## 2. Some shorter theory

A mixed model for plot data:

$$y = X\beta_b + Z_b u_b + Z_w u_w + e_w, \quad (1)$$

$y$  = observed data vector of plot yields

$\beta_b$  = fixed between-environment and across-environment effects

$u_b$  = random between-environment effects with  $\text{var}(u_b) = G_b$ ,

$u_w$  = random within-environment random design effects for replicates, blocks, etc. with  $\text{var}(u_w) = G_w$ , and

$e_w$  = plot errors with  $\text{var}(e_w) = R_w$ .

## 2. Some shorter theory

Same model in more compact form

$$y = X\beta_b + Z_b u_b + f_w, \quad (2)$$

where  $f_w = Z_w u_w + e_w$  and

$$\text{var}(f_w) = \Sigma_w = Z_w G_w Z_w^T + R_w. \quad (3)$$

**Mixed model equations (MME) for model (2)**

$$\begin{pmatrix} X^T \Sigma_w^{-1} X & X^T \Sigma_w^{-1} Z_b \\ Z_b^T \Sigma_w^{-1} X & Z_b^T \Sigma_w^{-1} Z_b + G_b^{-1} \end{pmatrix} \begin{pmatrix} \beta_b \\ u_b \end{pmatrix} = \begin{pmatrix} X^T \Sigma_w^{-1} y \\ Z_b^T \Sigma_w^{-1} y \end{pmatrix}. \quad (4)$$

## 2. Some shorter theory

### Two-stage representation of single-stage model (2)

$$\text{Stage 1: } y = X_1\mu_1 + f_w, \quad (5)$$

$\mu_1$  = genotype-environment means

$$\text{Stage 2: } \mu_1 = X_2\beta_b + Z_2u_b, \quad (6)$$

Plugging (6) into (5), it follows that

$$y = X_1X_2\beta_b + X_1Z_2u_b + f_w, \quad (7)$$

which is equivalent to 1-stage model (2) with  $X = X_1X_2$  and  $Z_b = X_1Z_2$ .

## 2. Some shorter theory

### Estimation in stage (1)

$$\hat{\mu}_1 = \left( X_1^T \Sigma_w^{-1} X_1 \right)^{-1} X_1^T \Sigma_w^{-1} y \quad \text{with} \quad (8)$$

$$\text{var}(\hat{\mu}_1 | u_b) = \left( X_1^T \Sigma_w^{-1} X_1 \right)^{-1} . \quad (9)$$

### Stage (2) model for stage (1) estimator from (8)

$$\hat{\mu}_1 = X_2 \beta_b + Z_2 u_b + e_b , \quad (10)$$

where  $e_b = \left( X_1^T \Sigma_w^{-1} X_1 \right)^{-1} X_1^T \Sigma_w^{-1} f_w$  and  $\text{var}(e_b) = \left( X_1^T \Sigma_w^{-1} X_1 \right)^{-1}$ .

## 2. Some shorter theory

### Rotation

Rotated (orthogonalized) means  $\tilde{\mu}_1 = W_1 \hat{\mu}_1$ ,

where  $X_1^T \Sigma_w^{-1} X_1 = W_1^2$

(a spectral decomposition)

## 2. Some shorter theory

Model for rotated means

$$\tilde{\mu}_1 = \tilde{X}_2 \beta_b + \tilde{Z}_2 u_b + \tilde{e}_b, \quad (11)$$

where  $\tilde{X}_2 = W_1 X_2$ ,  $\tilde{Z}_2 = W_1 Z_2$  and  $\tilde{e}_b = W_1 e_b$  with  $\text{var}(\tilde{e}_b) = I$ .

The MME for (11) are

$$\begin{pmatrix} \tilde{X}_2^T \tilde{X}_2 & \tilde{X}_2^T \tilde{Z}_2 \\ \tilde{Z}_2^T \tilde{X}_2 & \tilde{Z}_2^T \tilde{Z}_2 + G_b^{-1} \end{pmatrix} \begin{pmatrix} \beta_b \\ u_b \end{pmatrix} = \begin{pmatrix} \tilde{X}_2^T \hat{\mu}_1 \\ \tilde{Z}_2^T \hat{\mu}_1 \end{pmatrix}. \quad (12)$$

## 2. Some shorter theory

Using  $\tilde{X}_2 = W_1 X_2$  and  $\tilde{Z}_2 = W_1 Z_2$ , these can be re-arranged as

$$\begin{pmatrix} X_2^T X_1^T \Sigma_w^{-1} X_1 X_2 & X_2^T X_1^T \Sigma_w^{-1} X_1 Z_2 \\ Z_2^T X_1^T \Sigma_w^{-1} X_1 X_2 & Z_2^T X_1^T \Sigma_w^{-1} X_1 Z_2 + G_b^{-1} \end{pmatrix} \begin{pmatrix} \beta_b \\ u_b \end{pmatrix} = \begin{pmatrix} X_2^T X_1^T \Sigma_w^{-1} y \\ Z_2^T X_1^T \Sigma_w^{-1} y \end{pmatrix}. \quad (13)$$

Observing that  $X = X_1 X_2$  and  $Z_b = X_1 Z_2$ , this is seen to be equivalent to the single-stage MME!

(Piepho, Schulz-Streeck & Ogutu, 2012)