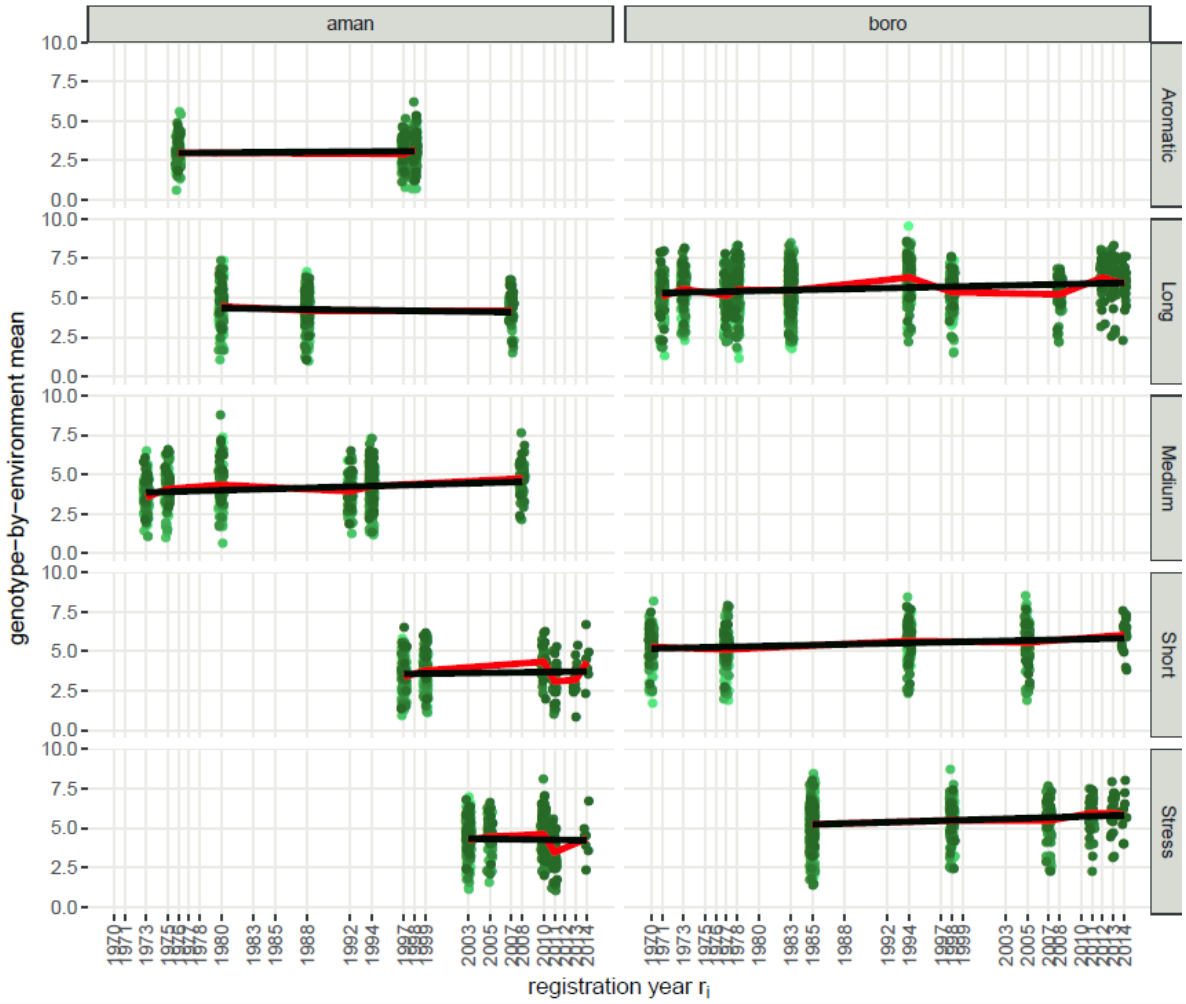


Implementation of genetic gain assessment using the BRRI data

Overview

Two datasets

- Aman & Boro

→ Get mean yield for each genotype at each environment

Genetic gain assessment

1. Incorporating regression terms in basic MET mixed model
    - Genetic trend: βr_i
        - Per group
    - Agronomic trend: γt_k

Data

Two datasets

- Aman & Boro

Data

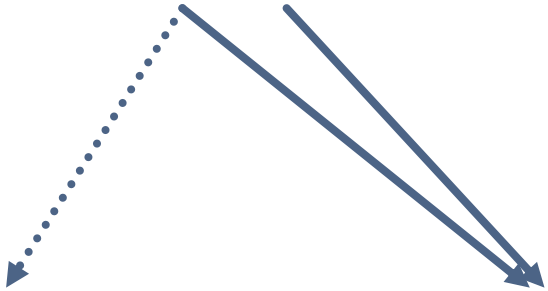
Original

SL	Year	Location	Year of release	Rep	Variety	Group	Yield
1	2001	Barisal	1980	1	BR10	Long	3.5
2	2001	Comilla	1980	1	BR10	Long	5.01104651
3	2001	Gazipur	1980	1	BR10	Long	4.28023256
4	2001	Rajshahi	1980	1	BR10	Long	4.48

Data

Original

SL	Year	Location	Year of release	Rep	Variety	Group	Yield
1	2001	Barisal	1980	1	BR10	Long	3.5
2	2001	Comilla	1980	1	BR10	Long	5.01104651
3	2001	Gazipur	1980	1	BR10	Long	4.28023256
4	2001	Rajshahi	1980	1	BR10	Long	4.48



Formatted

	Y	t.k	r.i	L	Env	G	Rep	Group	Yield
1:	2001	2001	1980	Barisal	2001-Barisal	BR10	1	Long	3.500000
2:	2001	2001	1980	Comilla	2001-Comilla	BR10	1	Long	5.011047
3:	2001	2001	1980	Gazipur	2001-Gazipur	BR10	1	Long	4.280233
4:	2001	2001	1980	Rajshahi	2001-Rajshahi	BR10	1	Long	4.480000
5:	2001	2001	1980	Rangpur	2001-Rangpur	BR10	1	Long	3.176419

7631:	2015	2015	2014	Comilla	2015-Comilla	BRRI dhan66	3	Stress	5.219160
7632:	2015	2015	2014	Gazipur	2015-Gazipur	BRRI dhan66	3	Stress	1.988598
7633:	2015	2015	2014	Rajshahi	2015-Rajshahi	BRRI dhan66	3	Stress	4.926488
7634:	2015	2015	2014	Rangpur	2015-Rangpur	BRRI dhan66	3	Stress	2.200000
7635:	2015	2015	2014	Satkhira	2015-Satkhira	BRRI dhan66	3	Stress	3.820000

Numeric

New columns

Data

Table: Number of unique entries per column

Aman		Boro	
	number		number
Y	15	Y	15
t.k	15	t.k	15
r.i	18	r.i	16
L	9	L	10
Env	116	Env	137
G	29	G	31
Rep	3	Rep	3
Group	5	Group	3

Data

Table: Number of unique entries per column

Aman		Boro	
	number		number
Y	15	Y	15
t.k	15	t.k	15
r.i	18	r.i	16
L	9	L	10
Env	116	Env	137
G	29	G	31
Rep	3	Rep	3
Group	5	Group	3

Basic model for long-term MET data

$$y_{ijk} = \mu + G_i + L_j + Y_k + (LY)_{jk} + (GL)_{ij} + (GY)_{ik} + (GLY)_{ijk} \quad (1)$$

y_{ijk} = mean yield of the i-th genotype in the j-th location and k-th year

Two datasets

- Aman & Boro

→ Get mean yield for each genotype at each environment

GxE means

GxE means

Goal: Obtain yield per genotype and environment averaged over replicates

Option 1: Arithmetic means

```
dat[, ari.mean  := mean(Yield), by=c("Env","G")]
```

GxE means

Goal: Obtain yield per genotype and environment averaged over replicates

Option 1: Arithmetic means

```
dat[, ari.mean  := mean(Yield), by=c("Env","G")]
```

Option 2: Adjusted means

For each environment separately:

1. Fit simple linear model: $y_{ij} = \mu + g_i + rep_j + e_{ij}$
2. Calculate adj. means per genotype
3. Calculate Smith's weights (Smith et al., 2001)

```
mod <- lm(formula = Yield ~ G + Rep,
          data    = subset(dat, Env==dat$Env[sel.env]))
adj.mean <- emmeans(mod, "G")
weight   <- diag(solve(vcov(adj.mean)))
```

GxE means

Aman

	Y	L	Env	t.k	r.i	Group	G	G.ari.mean	G.adj.mean	w.G
1:	2001	Barisal	2001-Barisal	2001	1980	Long	BR10	3.873333	3.873333	58.926346
2:	2001	Barisal	2001-Barisal	2001	1980	Medium	BR11	3.743333	3.743333	58.926346
3:	2001	Barisal	2001-Barisal	2001	1988	Long	BR22	4.236667	4.236667	58.926346
4:	2001	Barisal	2001-Barisal	2001	1988	Long	BR23	4.710000	4.710000	58.926346
5:	2001	Barisal	2001-Barisal	2001	1992	Medium	BR25	4.086667	4.086667	58.926346

2542:	2015	Sonagazi	2015-Sonagazi	2015	2011	Short	BRRI dhan57	2.372395	2.372395	15.605064
2543:	2015	Sonagazi	2015-Sonagazi	2015	2011	Stress	BRRI dhan57	2.372395	2.372395	15.605064
2544:	2015	Sonagazi	2015-Sonagazi	2015	2013	Short	BRRI dhan62	2.784334	2.784334	7.802532
2545:	2015	Sonagazi	2015-Sonagazi	2015	2014	Short	BRRI dhan66	3.523763	3.523763	15.605064
2546:	2015	Sonagazi	2015-Sonagazi	2015	2014	Stress	BRRI dhan66	3.523763	3.523763	15.605064

Boro

	Y	L	Env	t.k	r.i	Group	G	G.ari.mean	G.adj.mean	w.G
1:	2001-2002	Barisal	2001-2002-Barisal	2001	1970	Short	BR1	6.556667	6.556667	16.369190
2:	2001-2002	Barisal	2001-2002-Barisal	2001	1983	Long	BR12	7.556667	7.556667	16.369190
3:	2001-2002	Barisal	2001-2002-Barisal	2001	1983	Long	BR14	6.390000	6.390000	16.369190
4:	2001-2002	Barisal	2001-2002-Barisal	2001	1983	Long	BR15	6.820000	6.820000	16.369190
5:	2001-2002	Barisal	2001-2002-Barisal	2001	1983	Long	BR16	7.073333	7.073333	16.369190

2909:	2015-2016	Sonagazi	2015-2016-Sonagazi	2015	2014	Short	BRRI dhan63	3.930000	3.930000	4.818834
2910:	2015-2016	Sonagazi	2015-2016-Sonagazi	2015	2014	Long	BRRI dhan64	6.000000	6.000000	4.818834
2911:	2015-2016	Sonagazi	2015-2016-Sonagazi	2015	2014	Stress	BRRI dhan67	3.956667	3.956667	4.818834
2912:	2015-2016	Sonagazi	2015-2016-Sonagazi	2015	2014	Short	BRRI dhan68	6.056667	6.056667	4.818834
2913:	2015-2016	Sonagazi	2015-2016-Sonagazi	2015	2014	Long	BRRI dhan69	5.140000	5.140000	4.818834

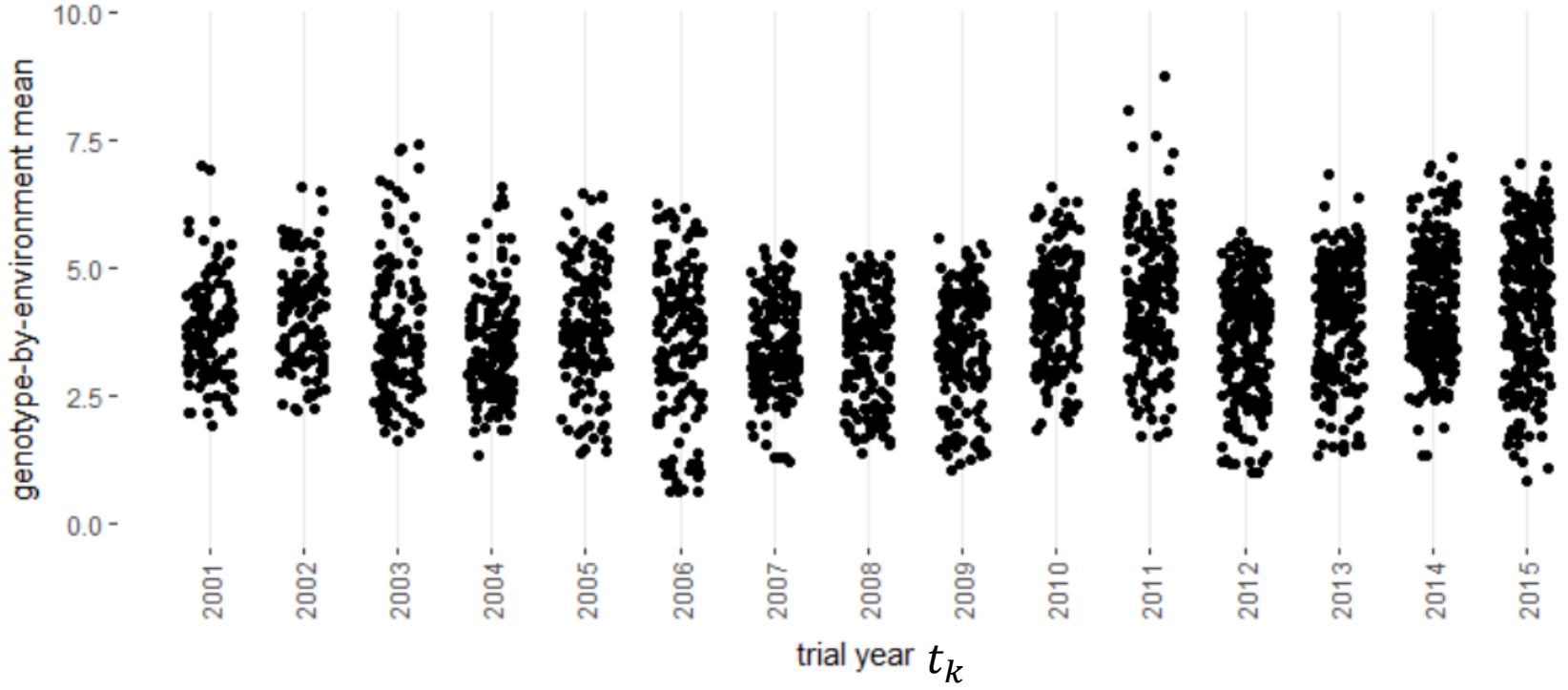
Option 1

Option 2

GxE means

	Y	L	Env	t.k	r.i	Group	G	G.ari.mean	G.adj.mean	w.G
1:	2001	Barisal	2001-Barisal	2001	1980	Long	BR10	3.873333	3.873333	58.926346
2:	2001	Barisal	2001-Barisal	2001	1980	Medium	BR11	3.743333	3.743333	58.926346
3:	2001	Barisal	2001-Barisal	2001	1988	Long	BR22	4.236667	4.236667	58.926346
4:	2001	Barisal	2001-Barisal	2001	1988	Long	BR23	4.710000	4.710000	58.926346
5:	2001	Barisal	2001-Barisal	2001	1992	Medium	BR25	4.086667	4.086667	58.926346

2542:	2015	Sonagazi	2015-Sonagazi	2015	2011	Short	BRRI dhan57	2.372395	2.372395	15.605064
2543:	2015	Sonagazi	2015-Sonagazi	2015	2011	Stress	BRRI dhan57	2.372395	2.372395	15.605064
2544:	2015	Sonagazi	2015-Sonagazi	2015	2013	Short	BRRI dhan62	2.784334	2.784334	7.802532
2545:	2015	Sonagazi	2015-Sonagazi	2015	2014	Short	BRRI dhan66	3.523763	3.523763	15.605064
2546:	2015	Sonagazi	2015-Sonagazi	2015	2014	Stress	BRRI dhan66	3.523763	3.523763	15.605064

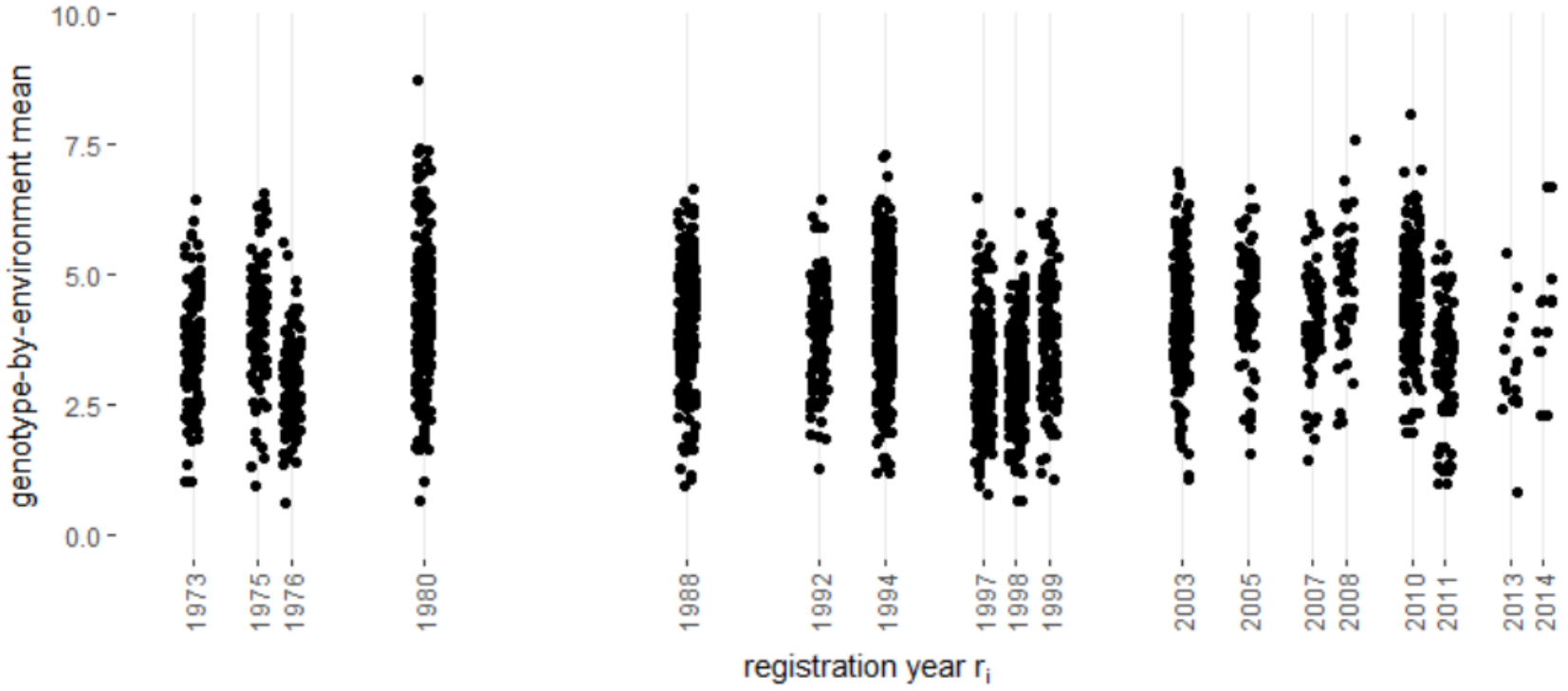


Genotype-by-environment means from aman dataset.

GxE means

	Y	L	Env	t.k	r.i	Group	G	G.ari.mean	G.adj.mean	w.G
1:	2001	Barisal	2001-Barisal	2001	1980	Long	BR10	3.873333	3.873333	58.926346
2:	2001	Barisal	2001-Barisal	2001	1980	Medium	BR11	3.743333	3.743333	58.926346
3:	2001	Barisal	2001-Barisal	2001	1988	Long	BR22	4.236667	4.236667	58.926346
4:	2001	Barisal	2001-Barisal	2001	1988	Long	BR23	4.710000	4.710000	58.926346
5:	2001	Barisal	2001-Barisal	2001	1992	Medium	BR25	4.086667	4.086667	58.926346

2542:	2015	Sonagazi	2015-Sonagazi	2015	2011	Short	BRRI dhan57	2.372395	2.372395	15.605064
2543:	2015	Sonagazi	2015-Sonagazi	2015	2011	Stress	BRRI dhan57	2.372395	2.372395	15.605064
2544:	2015	Sonagazi	2015-Sonagazi	2015	2013	Short	BRRI dhan62	2.784334	2.784334	7.802532
2545:	2015	Sonagazi	2015-Sonagazi	2015	2014	Short	BRRI dhan66	3.523763	3.523763	15.605064
2546:	2015	Sonagazi	2015-Sonagazi	2015	2014	Stress	BRRI dhan66	3.523763	3.523763	15.605064

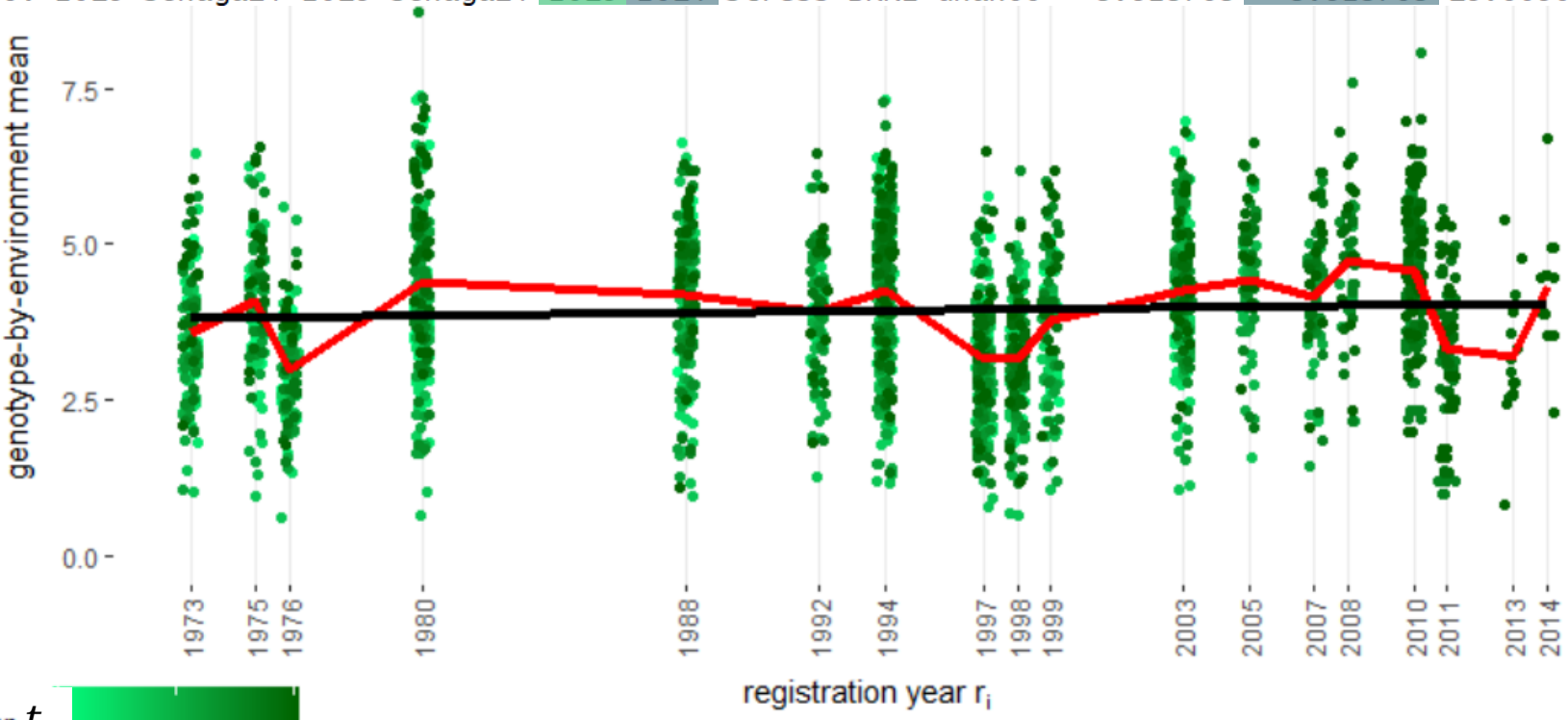


Genotype-by-environment means from aman dataset.

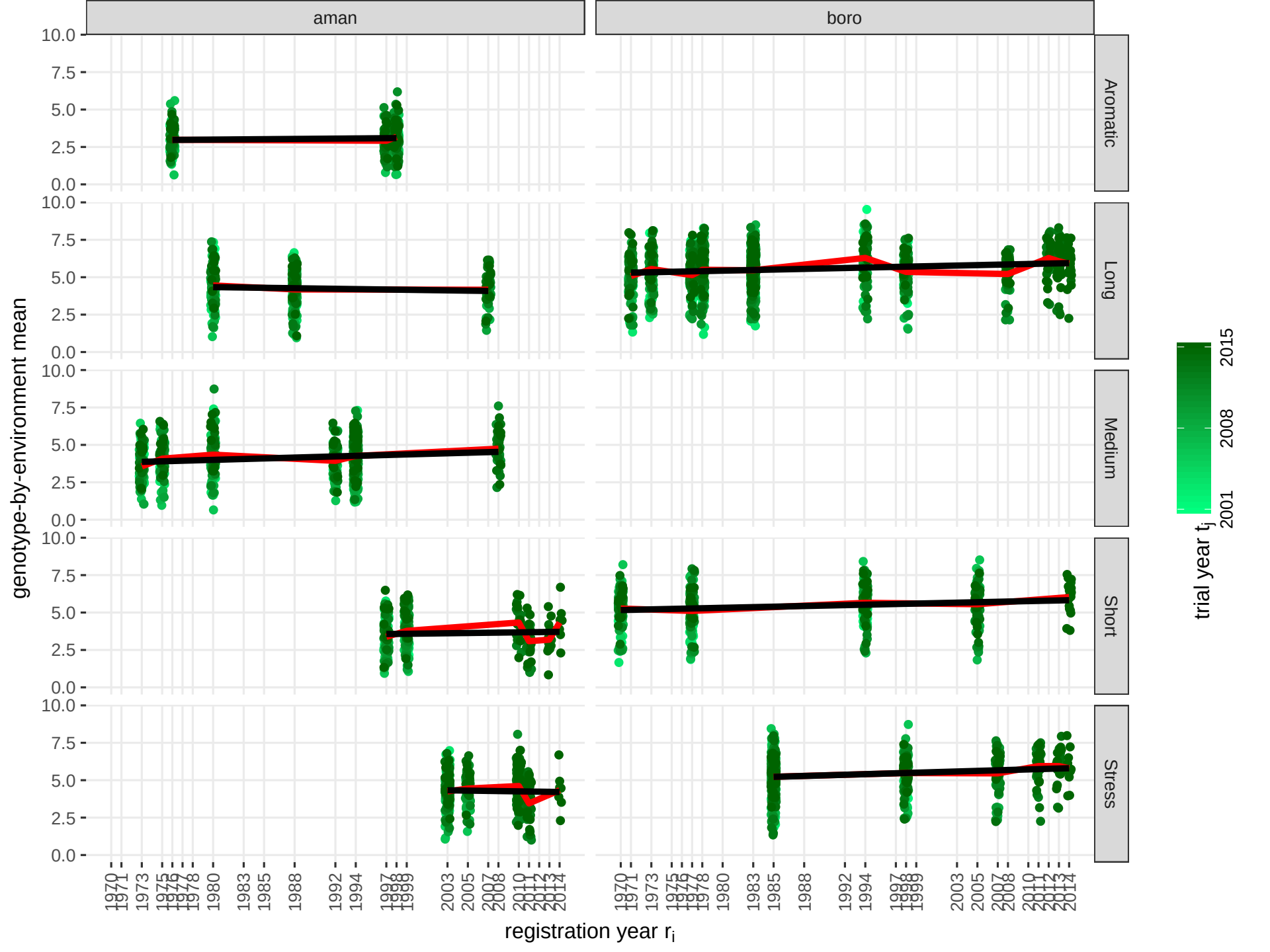
GxE means

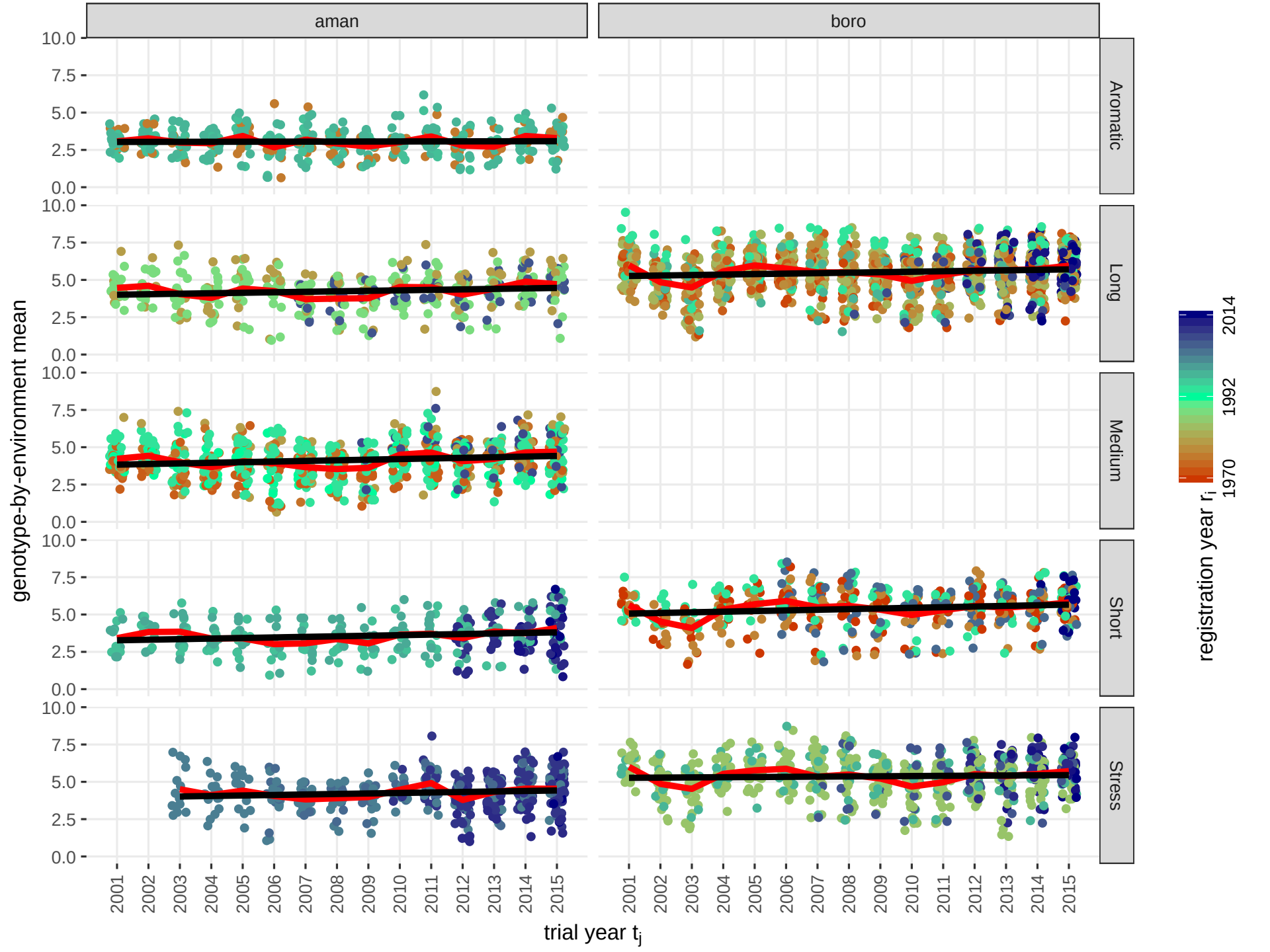
	Y	L	Env	t.k	r.i	Group	G	G.ari.mean	G.adj.mean	w.G
1:	2001	Barisal	2001-Barisal	2001	1980	Long	BR10	3.873333	3.873333	58.926346
2:	2001	Barisal	2001-Barisal	2001	1980	Medium	BR11	3.743333	3.743333	58.926346
3:	2001	Barisal	2001-Barisal	2001	1988	Long	BR22	4.236667	4.236667	58.926346
4:	2001	Barisal	2001-Barisal	2001	1988	Long	BR23	4.710000	4.710000	58.926346
5:	2001	Barisal	2001-Barisal	2001	1992	Medium	BR25	4.086667	4.086667	58.926346

2542:	2015	Sonagazi	2015-Sonagazi	2015	2011	Short	BRRI dhan57	2.372395	2.372395	15.605064
2543:	2015	Sonagazi	2015-Sonagazi	2015	2011	Stress	BRRI dhan57	2.372395	2.372395	15.605064
2544:	2015	Sonagazi	2015-Sonagazi	2015	2013	Short	BRRI dhan62	2.784334	2.784334	7.802532
2545:	2015	Sonagazi	2015-Sonagazi	2015	2014	Short	BRRI dhan66	3.523763	3.523763	15.605064
2546:	2015	Sonagazi	2015-Sonagazi	2015	2014	Stress	BRRI dhan66	3.523763	3.523763	15.605064



Genotype-by-environment means from aman dataset.
Red line: arithmetic mean per year. Black line: linear Regression y = a + bx.


aman
boro
Aromatic
Long
Medium
Short
Stress
genotype-by-environment mean
registration year r_i
trial year t_j
2001
2008
2015
0.0
2.5
5.0
7.5
10.0
1970
1971
1973
1975
1976
1977
1978
1980
1983
1985
1988
1992
1994
1997
1998
1999
2003
2005
2007
2008
2010
2011
2012
2013
2014


aman
boro
Aromatic
Long
Medium
Short
Stress
genotype-by-environment mean
trial year t_j
registration year r_i
1970
1992
2014
0.0
2.5
5.0
7.5
10.0
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015

Genetic gain assessment

1. Incorporating regression terms in basic MET mixed model

Genetic gain assessment

Genetic trend: βr_i
Agronomic trend: γt_k

```
lmer(formula = G.ari.mean ~ r.i + t.k +
       (1|G) + (1|L) + (1|Y) + (1|Y:L) + (1|L:G) + (1|Y:G),
     data    = dat)
```

Basic model for long-term MET data

$$y_{ijk} = \mu + G_i + L_j + Y_k + (LY)_{jk} + (GL)_{ij} + (GY)_{ik} + (GLY)_{ijk} \quad (1)$$

y_{ijk} = mean yield of the *i*-th genotype in the *j*-th location and *k*-th year
μ = overall mean
G_i = main effect of the *i*-th genotype
L_j = main effect of the *j*-th location
Y_k = main effect of the *k*-th year
$(LY)_{jk}$ = *jk*-th location × year interaction
$(GL)_{ij}$ = *ij*-th genotype × location interaction
$(GY)_{ik}$ = *ik*-th genotype × year interaction
$(GLY)_{ijk}$ = residual comprising both genotype × location × year interaction as well as the error of a mean

2. Incorporating regression terms

Genetic trend

$$G_i = \beta r_i + H_i \quad (2)$$

β = fixed regression coefficient for genetic trend
r_i = year of first trial for *i*-th variety
$H_i \sim N(0, \sigma_H^2)$

Non-genetic trend

$$Y_k = \gamma t_k + Z_k \quad (3)$$

γ = fixed regression coefficient for agronomic trend
t_k = calendar year
$Z_k \sim N(0, \sigma_Z^2)$

Genetic gain assessment

Genetic trend per group: βr_{il}
Agronomic trend: γt_k

```
lmer(formula = G.ari.mean ~ r.i:Group + t.k +
       (1|G) + (1|L) + (1|Y) + (1|Y:L) + (1|L:G) + (1|Y:G),
     data    = dat)
```

Basic model for long-term MET data

$$y_{ijk} = \mu + G_i + L_j + Y_k + (LY)_{jk} + (GL)_{ij} + (GY)_{ik} + (GLY)_{ijk} \quad (1)$$

y_{ijk} = mean yield of the *i*-th genotype in the *j*-th location and *k*-th year
μ = overall mean
G_i = main effect of the *i*-th genotype
L_j = main effect of the *j*-th location
Y_k = main effect of the *k*-th year
$(LY)_{jk}$ = *jk*-th location × year interaction
$(GL)_{ij}$ = *ij*-th genotype × location interaction
$(GY)_{ik}$ = *ik*-th genotype × year interaction
$(GLY)_{ijk}$ = residual comprising both genotype × location × year interaction as well as the error of a mean

2. Incorporating regression terms

Genetic trend

$$G_i = \beta r_i + H_i \quad (2)$$

β = fixed regression coefficient for genetic trend
r_i = year of first trial for *i*-th variety
$H_i \sim N(0, \sigma_H^2)$

Non-genetic trend

$$Y_k = \gamma t_k + Z_k \quad (3)$$

γ = fixed regression coefficient for agronomic trend
t_k = calendar year
$Z_k \sim N(0, \sigma_Z^2)$

Summary: Genetic gain assessment

Dataset	Group	Genetic Trend	Agronomic Trend
aman	Aromatic	0.0053 (p=0.562)	
	Long	0.0059 (p=0.520)	
	Medium	0.0059 (p=0.523)	0.0192 (p=0.353)
	Short	0.0057 (p=0.527)	
	Stress	0.0057 (p=0.529)	
boro	Long	0.0124 (p<0.001)	
	Short	0.0123 (p<0.001)	0.0189 (p=0.779)
	Stress	0.0123 (p<0.001)	