Genotype by environment interaction

Luis Fernando Delgado Munoz
luis.delgado@cgiar.org
Faisalabad, November 2022

@BiovIntCIAT_eng
@BiovIntCIAT_esp

#Alliance4Science

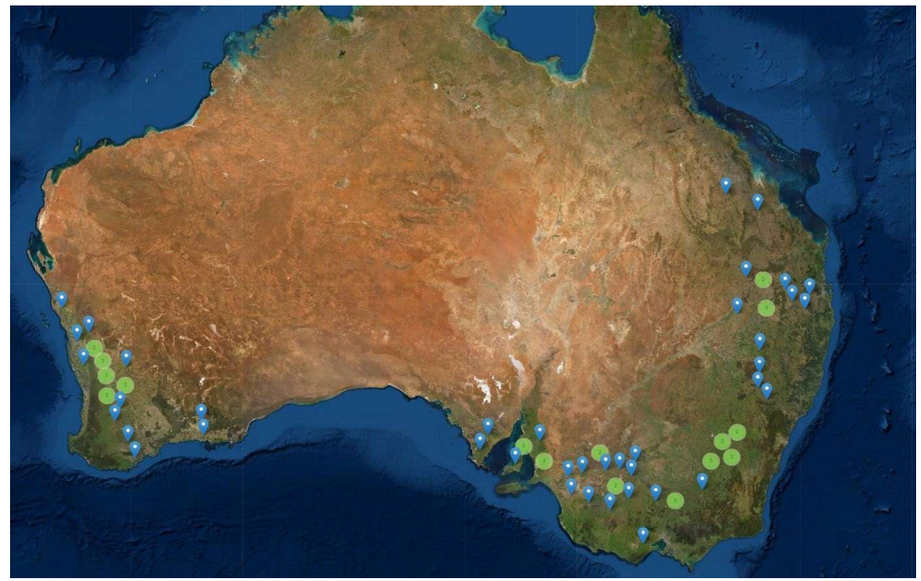
Target population of environments (TPE): what are we breeding for?

- TPE: set of environmental conditions that future varieties are likely to experience when grown by farmers

  (Comstock and Moll 1963)

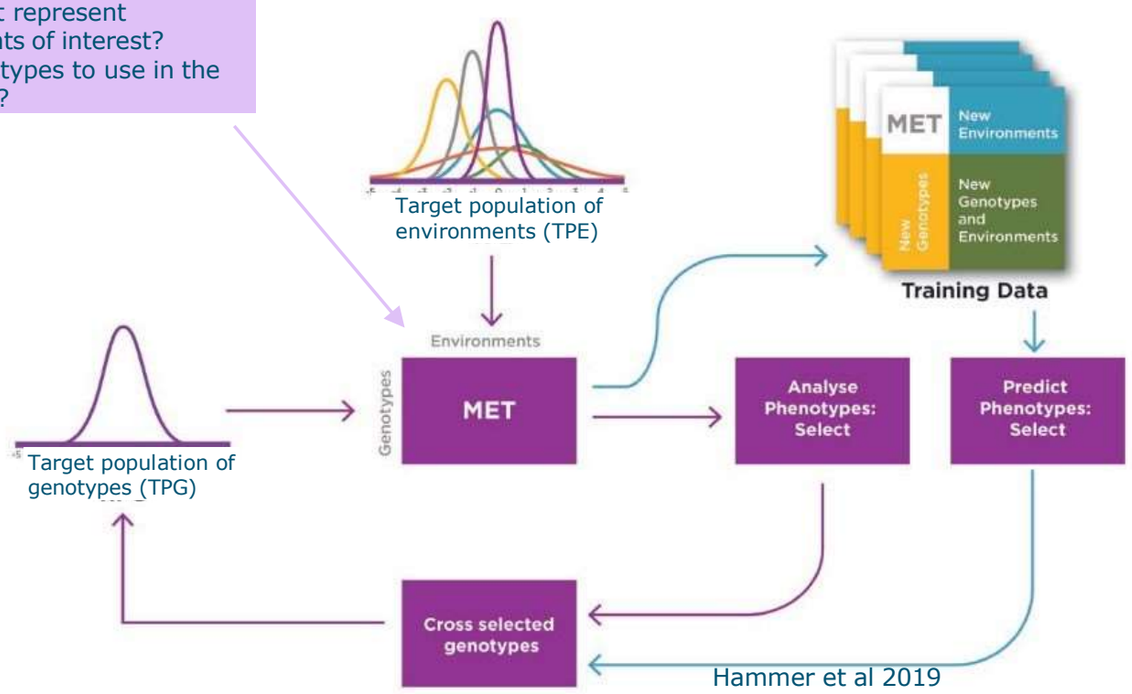
- Multi-environment trials (METs) aim at representing the TPE
  - Locations
  - Years
  - Agronomic management

Example: METs in the Australian TPE for wheat



Predictive technologies have become central to produce better adapted varieties

Design issues
- How to best represent environments of interest?
- Which genotypes to use in the training set?



Hammer et al 2019

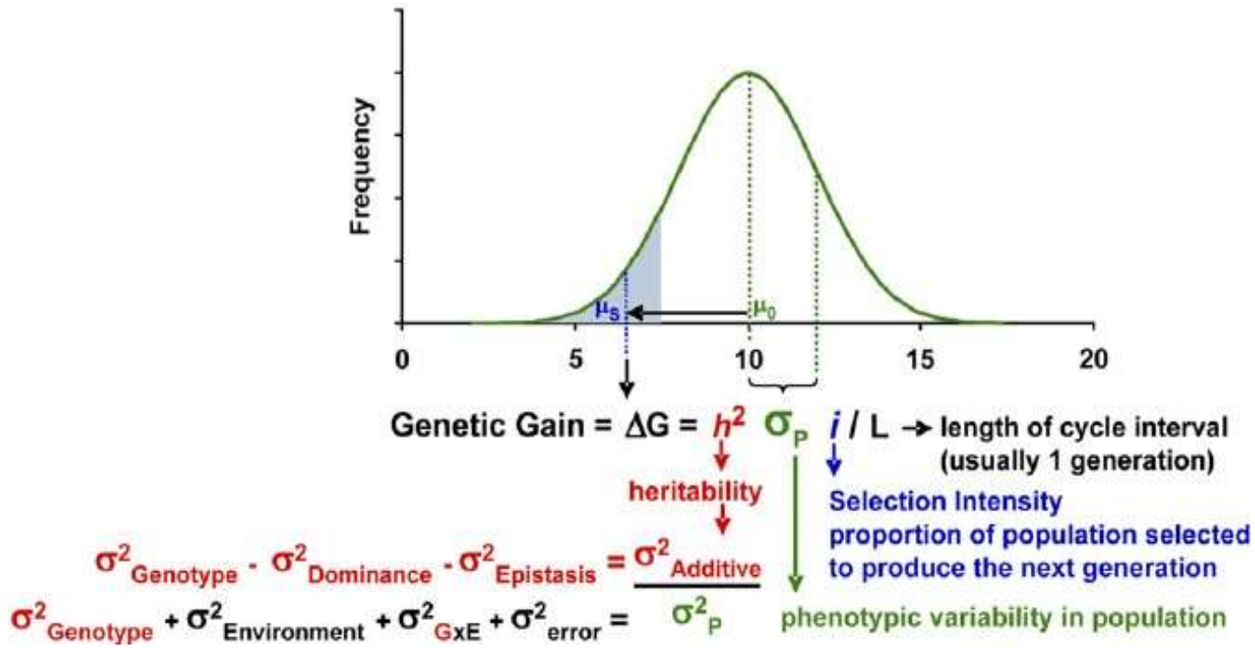
Data to predict new genotypes:
- Genomics (Molecular markers)
- Pedigree
- Phenomics (trait monitoring)

Data to predict new environments:
- Enviromics
  - Weather
  - Soil
  - Agronomic management

Prediction model
- Variable selection
- Structure

Genetic gain

- Genetic gain (phenotype increase) can be accelerated by shortening the breeding cycle and evaluating more genotypes



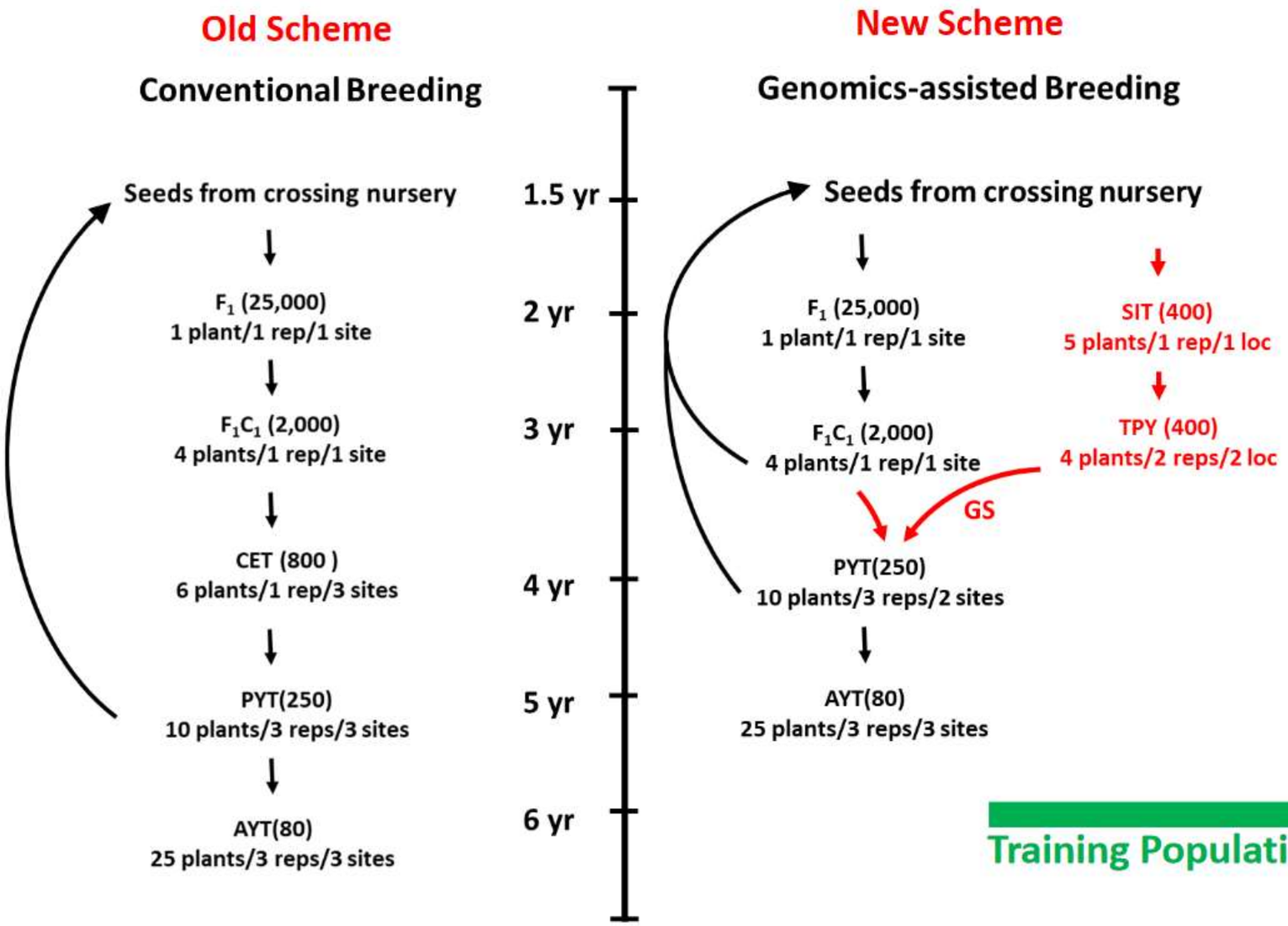
Prediction quality has become central for plant breeding

- Genetic gain (phenotype increase) can be accelerated by shortening the breeding cycle and evaluating more genotypes

$$\Delta G_A = \frac{i * r * \sigma_A}{t}$$

i: measure of selection intensity (do we select the 5 or the 10% best genotypes?)
r: prediction accuracy
σ_A: additive genetic standard deviation of the target trait
t: time to complete one cycle of the breeding program (i.e. to release a new variety). This can be considerably shorter in a breeding programme based on genomic selection, compared to traditional one

Modernization Breeding Scheme: Implement GS



Highlights:

- Reduce the selection cycle by at least 2 years
- Multiple replicates and locations increase phenotype quality
- Evaluate multiple essential traits
- Discard poor clones and reduce population size early
- Genotypic data increases the accuracy of clonal evaluation trials

*SIT, Seed Increase Trial;
TPY, Training Population Yield trial;
F1, seedling nursery;
F1C1, cloned seedling nursery;
CET, single row trial
PYT, Preliminary Yield Trial;
AYT, Advanced Yield Trial

Alliance
Bioversity International
CIAT
CGIAR

Critical issues to achieve higher accuracy

We don’t breed for the past, **we breed for the future**:
which are the future growing conditions that varieties will encounter when planted by farmers?

- How well does our MET data represent the TPE?
- How large is GxL vs GxY?
- Can we partition locations into regions?
- Do we know which is the probability of a stress to occur at a particular location?

Alliance Bioversity & CIAT

CGIAR

What are we selecting for?

- If **GxL** is large and locations can be grouped into **regions**
- If **GxLxY** is large and environments can be grouped into **scenarios**
- If **GxM** is large and there are groups of **management conditions**

Use **group-specific** predictions to select for **specific adaption**

Use predictions for the **general mean** to select for **general adaptation**

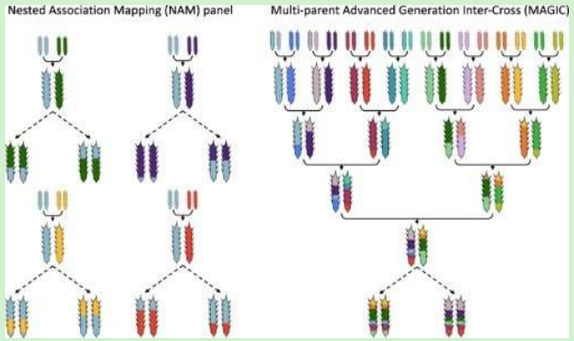
Finlay-Wilkinson/AMMI/GGE models can be used to identify groups of environments.

Mixed models (variance components or variance-covariance models) can be used to make the predictions.

GxE data

Genotypes: how much variation is in the material of interest?

- Association panels / Specific sets of genotypes
- Offspring from biparental and multi-parental crosses
- Genetic relationships of different magnitude.



Environments: which are the environmental conditions in which future varieties will be grown?

- Weather
- Soil
- Agronomic management
- Pest and disease pressure

Multi-environment trials (METs): how do we represent the relevant environmental conditions in our trials?

- Managed stress trials / phenotyping platforms.
- Series of trials over sites and years.

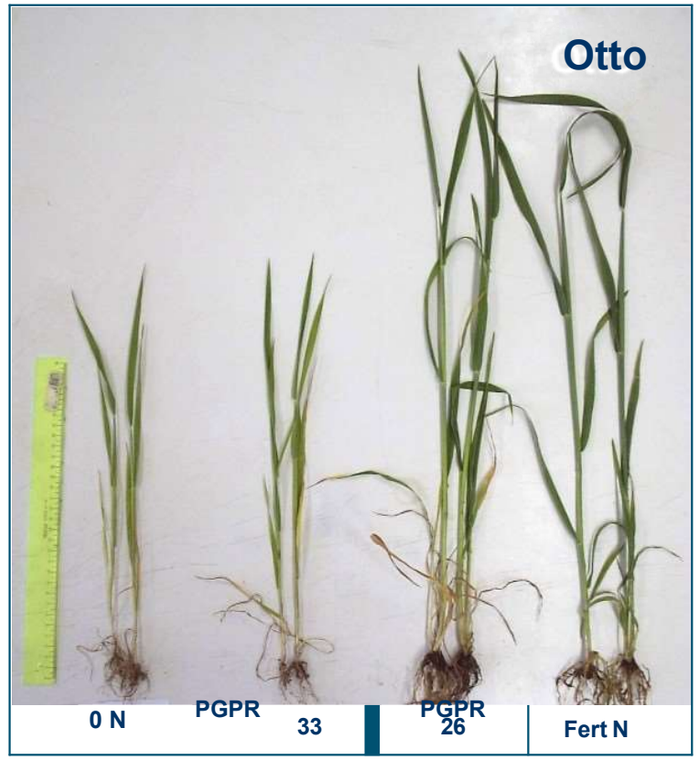
Phenotypic plasticity and sensitivity to the environment

Phenotypic plasticity: change in the external characteristics (phenotype)

Sensitivity to the env: phenotypic plasticity in response to environmental conditions

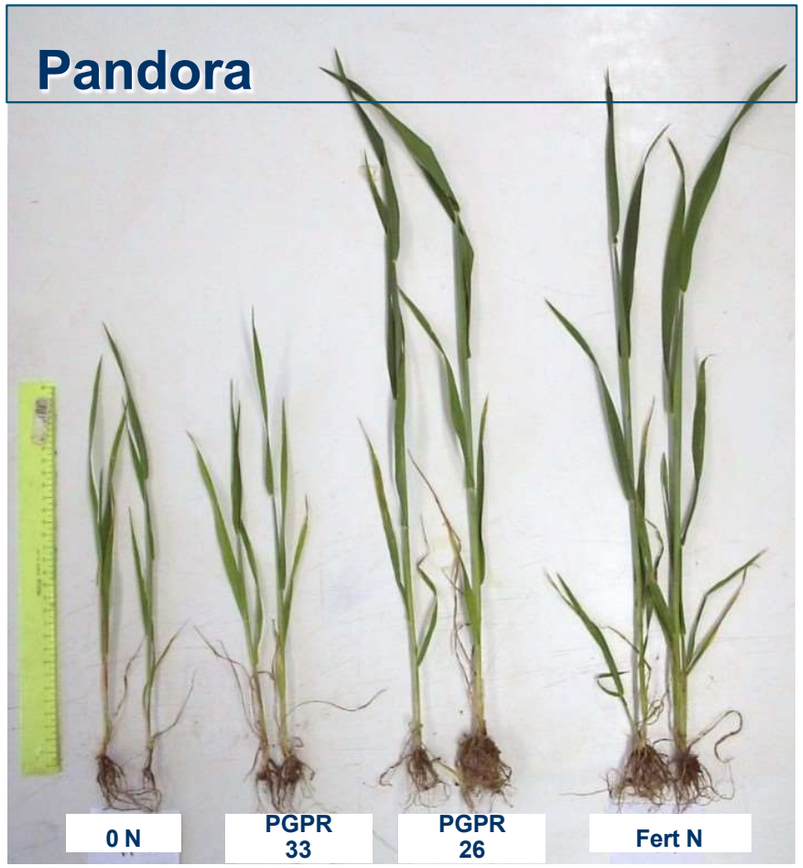
Example: the wheat cultivar Otto produces larger biomass at higher nitrogen availability (Fert N y PGPR26)

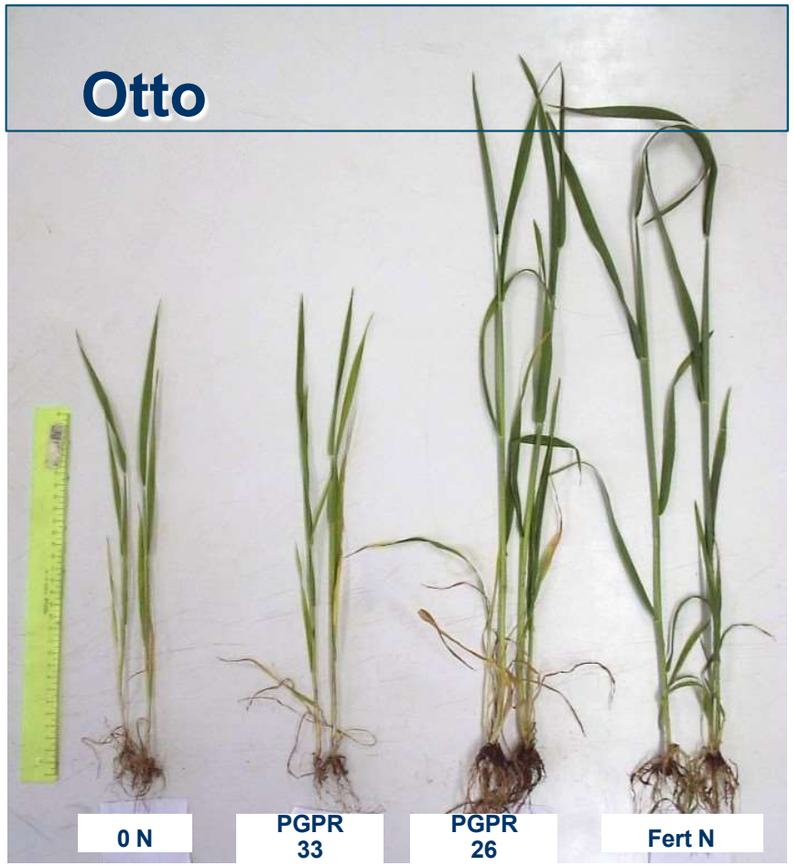
These biomass differences are an expression of **phenotypic plasticity** and reflect **sensitivity to nitrogen**

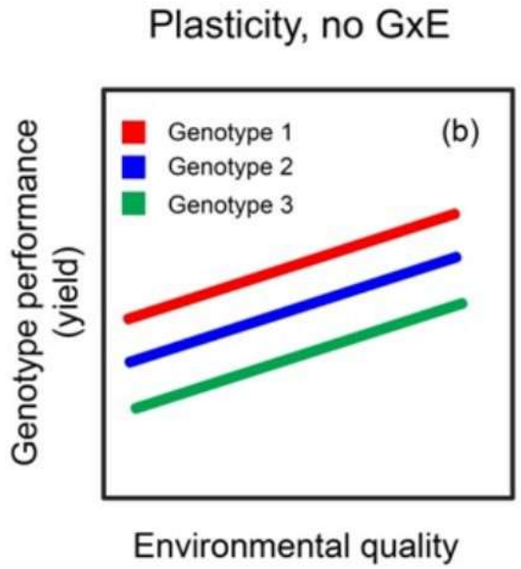


If both genotypes shot the same phenotypic plasticity and sensitivity to the env:

absence of GxE

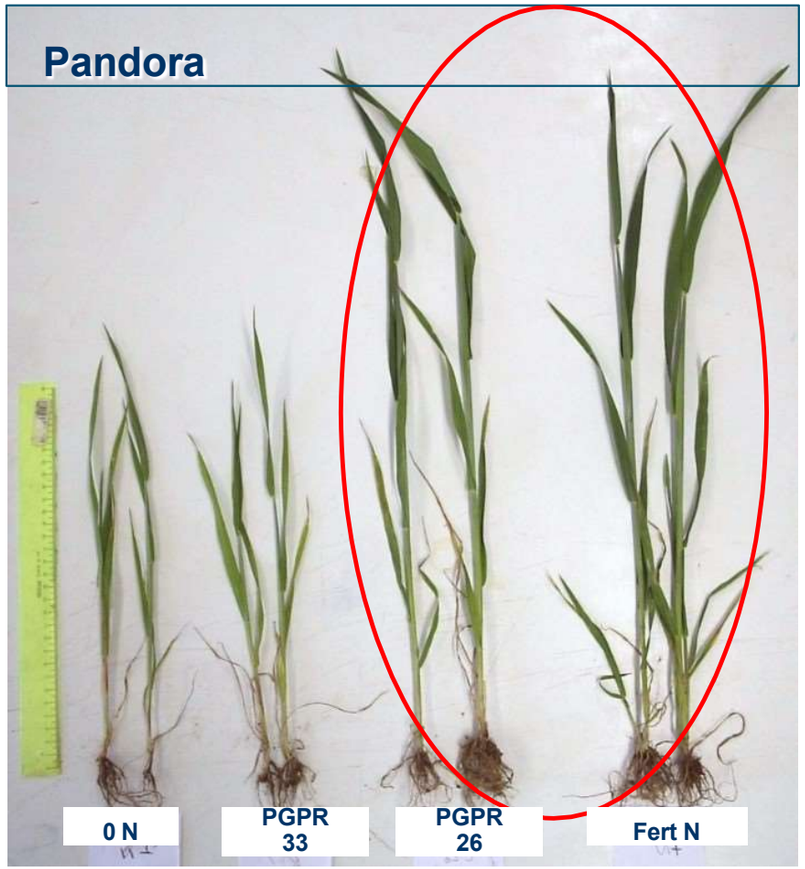


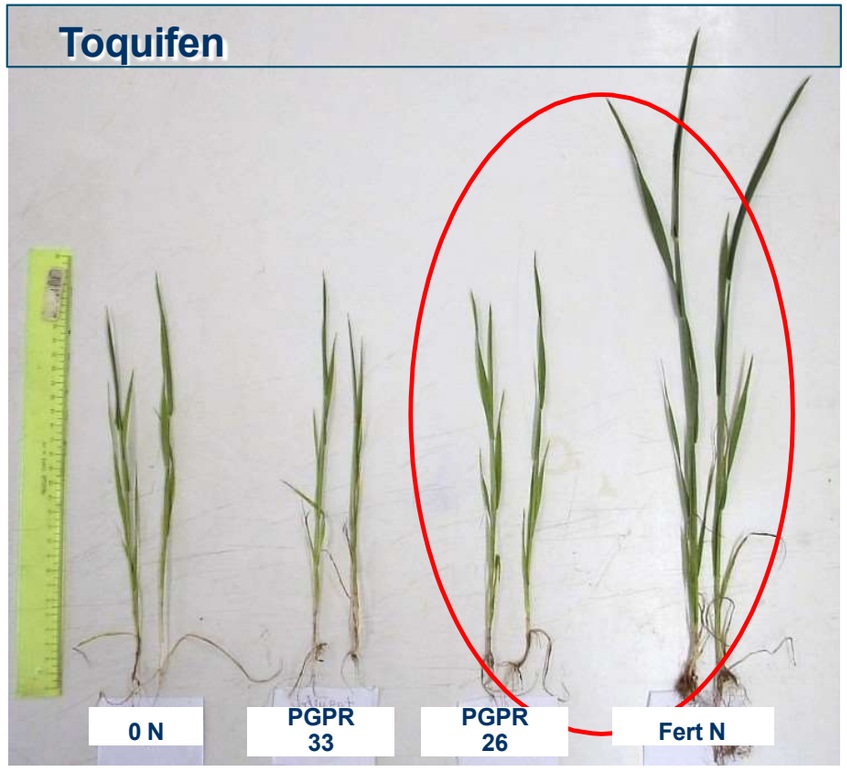


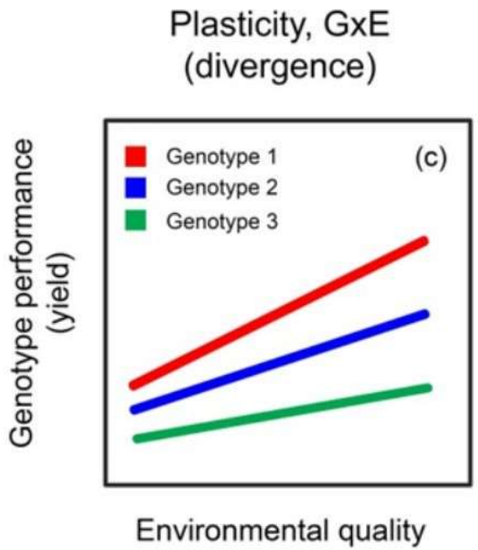


Genotypes differ in their phenotypic plasticity and sensitivity to the env:

presence of GxE







Genotype by env interaction (GxE)

- When modelling GxE interaction (GEI) distinguish:
  - approach by regression
    - non parallelism of individual responses for genotypes (reaction norms)
  - approach by variances and correlations
    - heterogeneity of within environment variance across environments = genetic or genotypic variance
    - heterogeneity of between environment correlations = genetic or genotypic correlation
    - heterogeneity of within genotype variances = stability variances

RESEARCH

What Should Students in Plant Breeding Know About the Statistical Aspects of Genotype × Environment Interactions?

Fred A. van Eeuwijk,* Daniela V. Bustos-Korts, and Marcos Malosetti

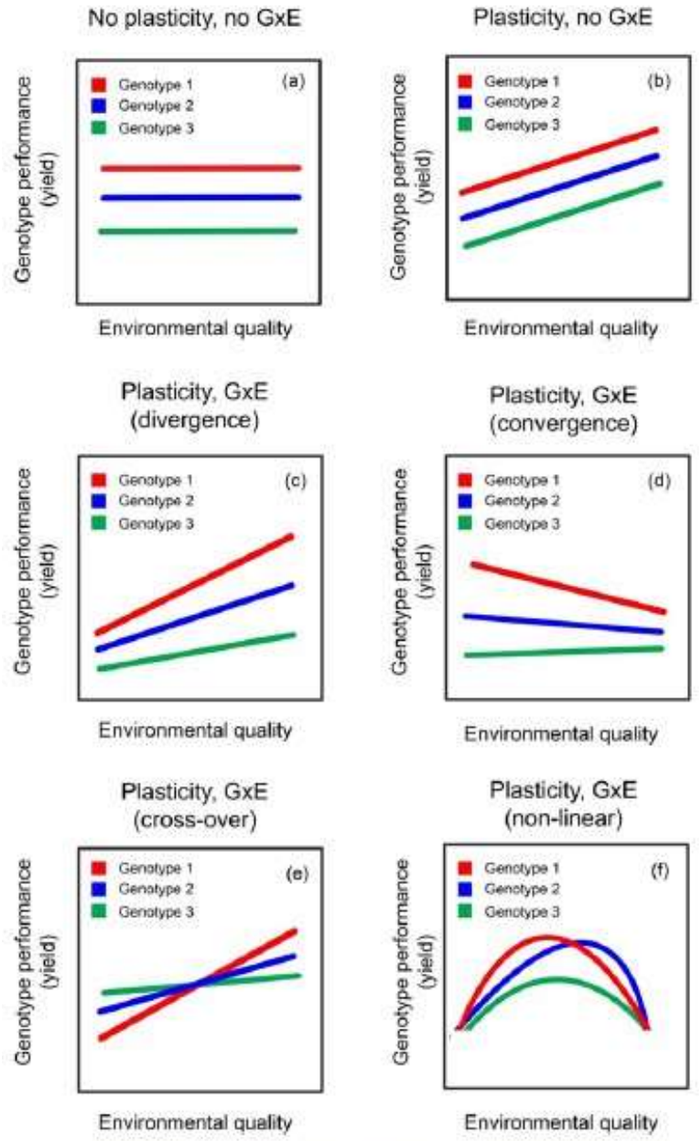


Fig. 1. Reaction norms for three genotypes that illustrate various forms of plasticity and Genotype × Environment interaction (G × E). No plasticity in (a) versus plasticity in (b) to (f), no G × E in (a) and (b) versus various forms of G × E in (c) till (f).

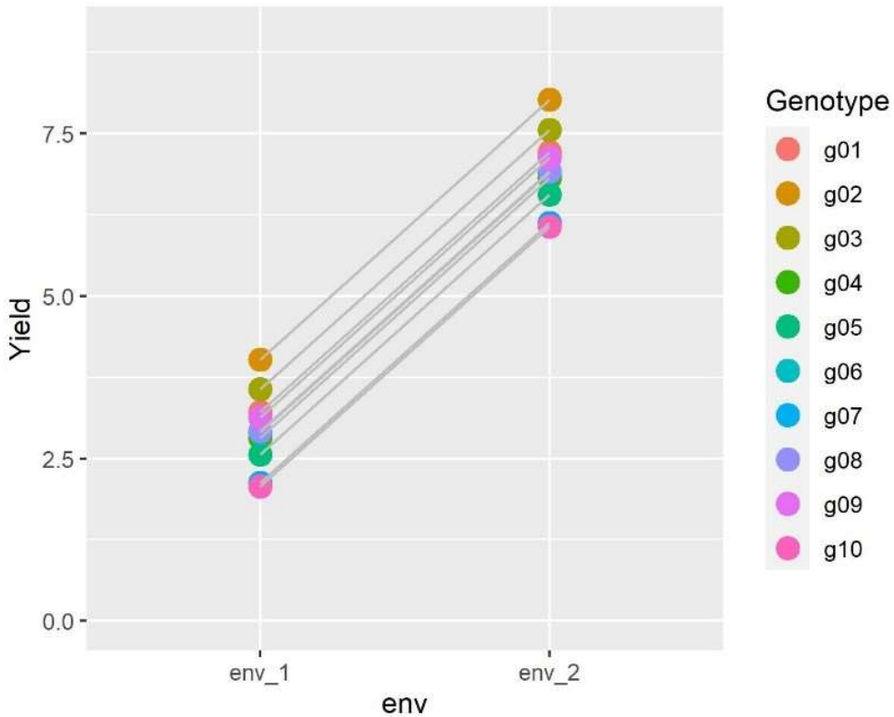
GxE explained as lack of correlation and heterogeneity of variance

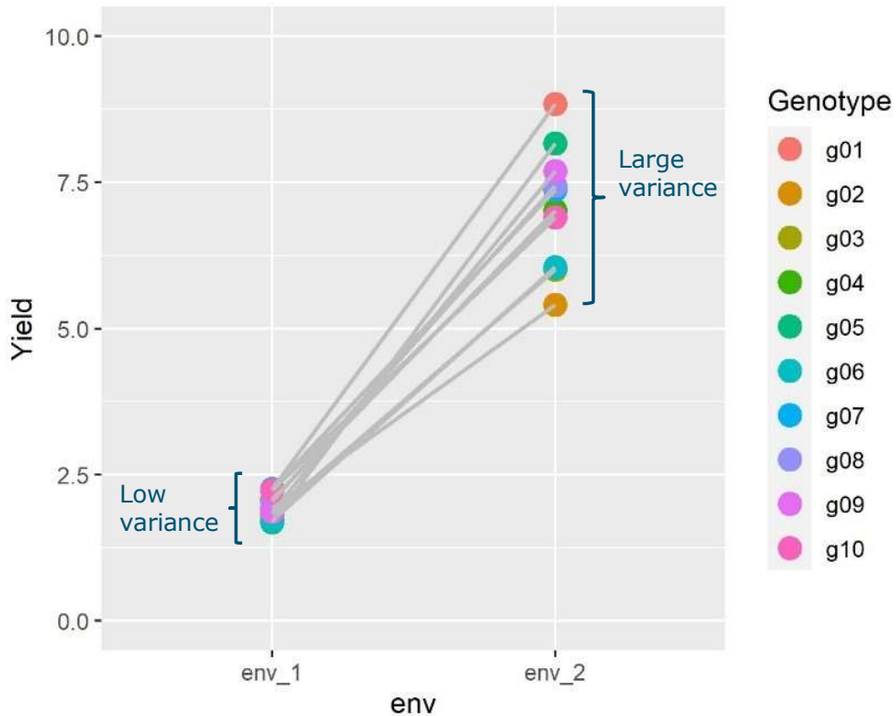
Absence of GxE

- Genotypic differences are preserved across envs
- Correlation between envs= **1**
- envs have the same phenotypic variance

Presence of GxE

- Genotypic differences depend on the env
- Correlation between envs **< 1**
- envs have different phenotypic variance

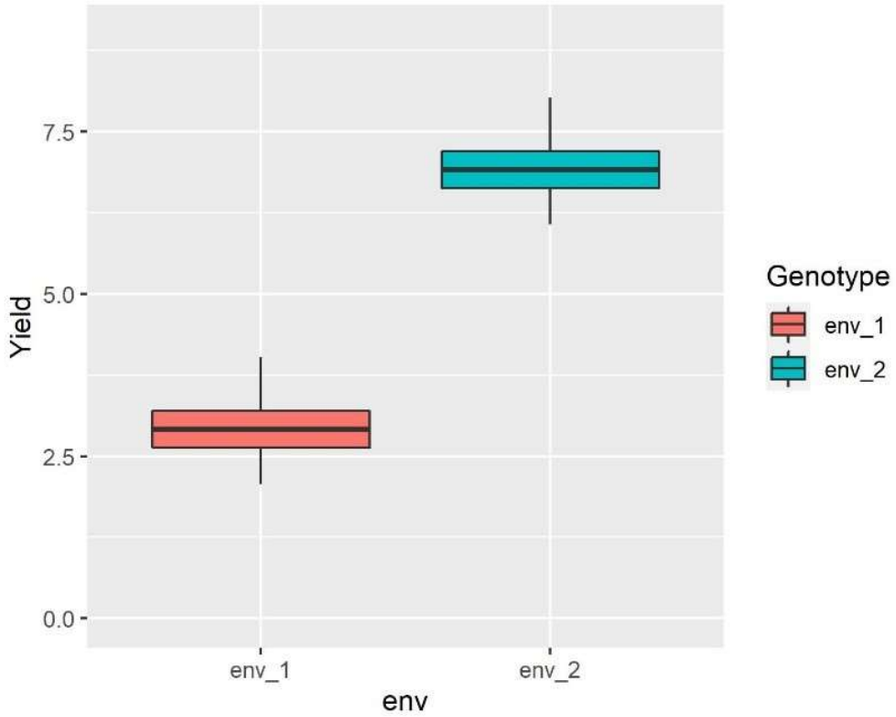




Heterogeneity of variance (boxplot)

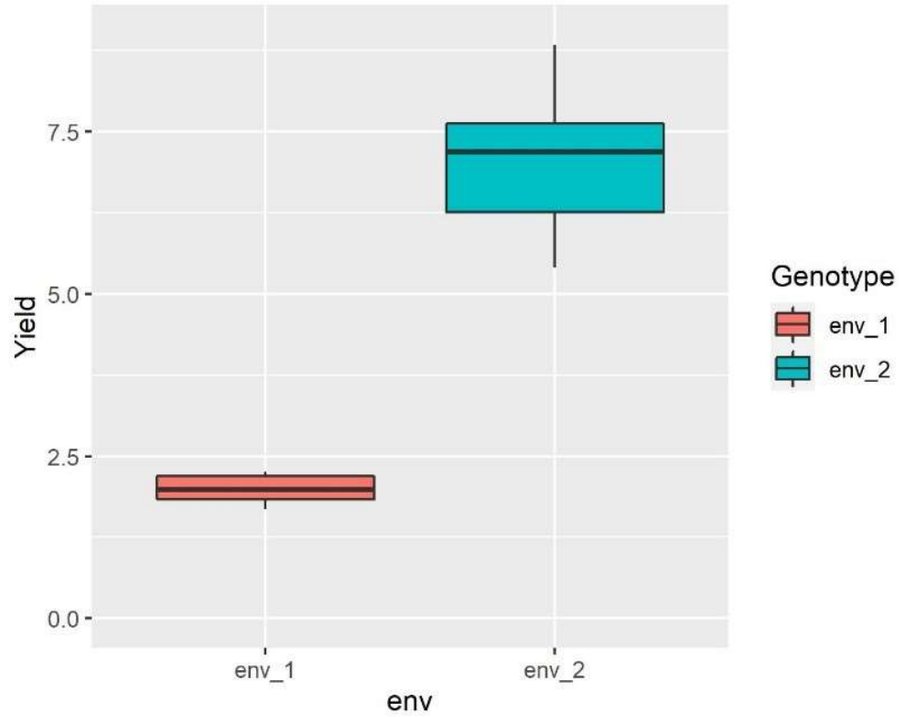
Absence of GxE

- Environments have the same phenotypic variance



Presence of GxE

- Environments have different phenotypic variance



The (very) basics about mixed models

- Mixed models consist of:

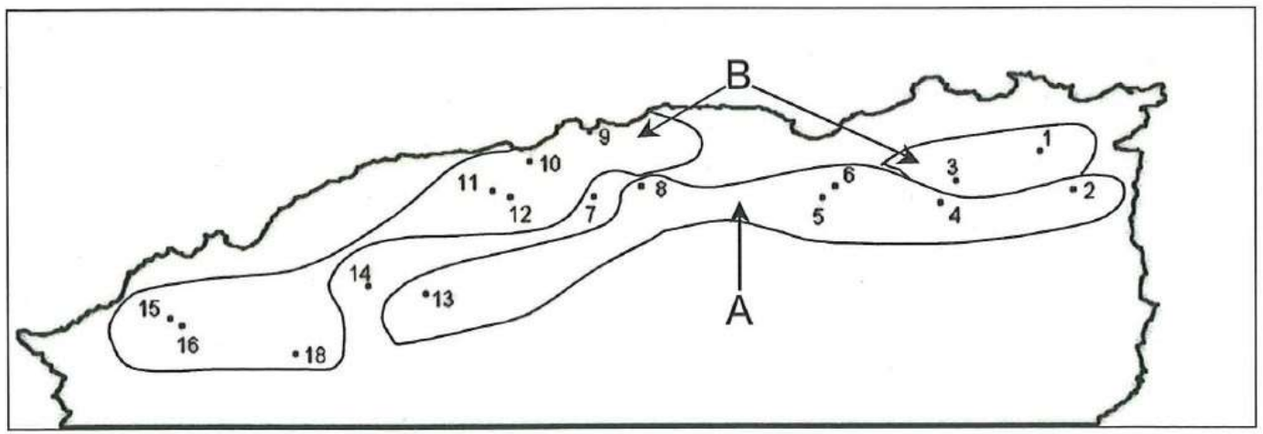
Fixed terms

- Interest in specific levels
- Parameters of interest are individual treatment effects

Random terms

- Treatment levels are chosen to represent a population of possible treatments
- Parameters of interest variance parameters

Example durum wheat Algeria

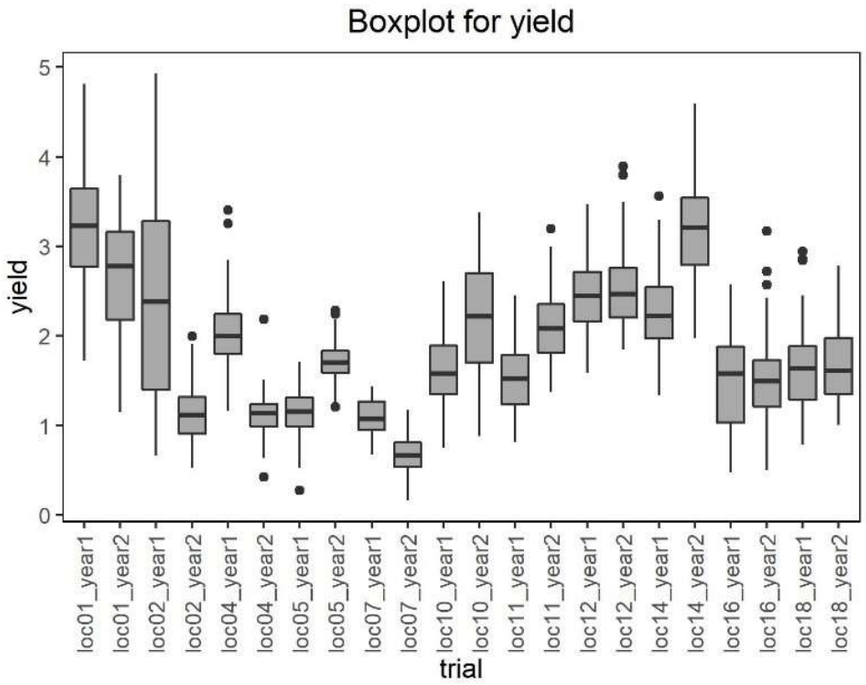


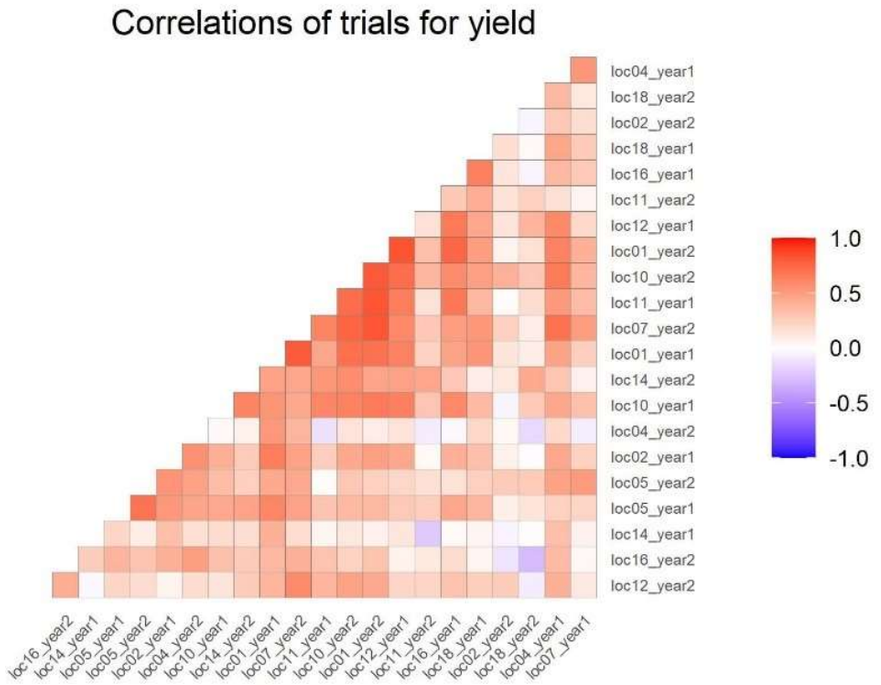
Source: Annicchiarico, 2002c.

Multi-environment trial (MET) data:

- 24 genotypes in 22 environments
  - 11 sites
  - 2 years
- Experiments designed as RCBD in each site (4 replicates).

Visualizing the raw data (without correcting for blocks)





- Large heterogeneity of variance across trials
- Correlations lower than 1 (and very heterogeneous)

} Indicators of GxE

GxE analysis: One vs two-stages

- **One-stage analysis (one-hit models):**
  - Simultaneous modelling of the within and between trial variation.
  - Uses all the information (plot data).
  - Can quickly become complicated (to account individual trials specifics).

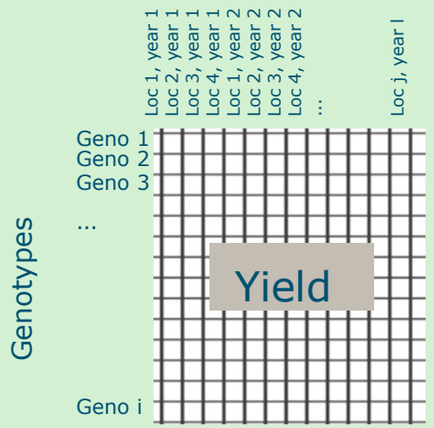
- **Two-stage analysis:** Simpler / pragmatic approach

Stage 1: analysis per trial (focus is on the modelling of variation in a particular trial).

Adjusted means per trial

Weights

Stage 2: focus on GxE.



MET data: GxLxY

	A	B	C	D	E	F	G
1	Year	Location	Environment	Region	Block	Genotype	yield
2	year1	loc01	loc01_year1	B	Bl1	ARDENTE	3.502
3	year1	loc01	loc01_year1	B	Bl1	BIDI17	2.491
4	year1	loc01	loc01_year1	B	Bl1	MBBACHIR	2.519
5	year1	loc01	loc01_year1	B	Bl1	HEBD/GDO	3.652
6	year1	loc01	loc01_year1	B	Bl1	HEBDA03	2.47
7	year1	loc01	loc01_year1	B	Bl1	BID/WAHA	3.21
8	year1	loc01	loc01_year1	B	Bl1	SIMETO	2.481
9	year1	loc01	loc01_year1	B	Bl1	GTADUR	3.68
10	year1	loc01	loc01_year1	B	Bl1	POLONICU	2.586
11	year1	loc01	loc01_year1	B	Bl1	B.DUR194	3.664
12	year1	loc01	loc01_year1	B	Bl1	DUILIO	3.533
13	year1	loc01	loc01_year1	B	Bl1	EIDER	4.71

- 24 genotypes (Genotype)
- 22 environments (Environment)
    - 11 locations (Location)
    - 2 years (Year)
- Each trial a randomized complete block design (Block) with 4 replicates.
- A one-stage model for this data (using Environments as factor)

Treatment design: Crossing and nesting

Crossing:

All combinations of factors A and B potentially present

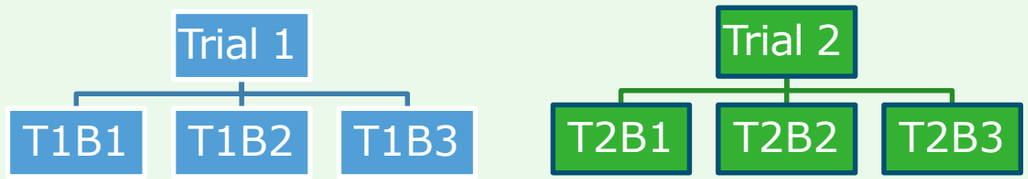
	G1	G2	G3
N1	N1G1	N2G1	N3G1
N2	N1G2	N2G2	N3G2
N3	N1G3	N2G3	N3G3

Example:

- 3 nitrogen levels and 3 genotypes combined in a factorial way

Nesting:

Treatment levels of factor A ONLY occur within specific levels of factor B



Example:

- Blocks are nested within each experiment

A basic two-way ANOVA model: one-stage analysis

$$y_{ijk} = \mu + E_j + b_{k(j)} + G_i + GE_{ij} + \epsilon_{ijk} \qquad \epsilon_{ijk} \sim N(0, \sigma^2_\epsilon)$$

- Grain yield =
  - Intercept +
  - Effect of environment j (E_j) +
  - Block nested within environment ($b_{k(j)}$)
  - Effect of genotype i (G_i) +
  - Effect of genotype x environment interaction (GE_{ij}) +
  - error (ϵ_{ijk})

- Linear model (fixed effects model):
  - One parameter per genotype, environment, and combination of genotype and environment.
  - Assumes constant residual variance.
  - Requires all G-E combinations to be present at least once.

Fit the model with R (option A)

Use this R statement if blocks are named 1 to 4 in each environment

```
# Fit 2-way anova
m1 <- lm(yield~  Environment/Block +  Genotype +
                 Genotype:Environment, data = dat)

# Display anova table
anova(m1)
```

- lm() function: to fit any linear model, anova or regression (or ancova).
- Geno, Env and Block are factors (block could have been fitted as random instead)
- The plus sign (+) is the additive operator
- The colon (:) is the operator for interaction, so Geno:Env stands for Genotype by Environment interaction
- The forward slash (/) indicates nesting, so Env/Block indicates that blocks should be nested within environments

Fit the model with R (option B)

```
# Fit 2-way anova
m1 <- lm(yield~  Environment + Block +  Genotype +
                 Genotype:Environment, data = dat)

# Display anova table
anova(m1)
```

- Use this option if blocks have labels nested within environment
  - Blocks 1 to 4 in environment 1
  - Blocks 5 to 8 in environment 2
  - ...
- Then, they can be fitted as a main effect (as in the model above)

Classical ANOVA results

```
> anova(m1)
Analysis of Variance Table

Response: yield
                     Df Sum Sq Mean Sq  F value    Pr(>F)
Environment          21 935.95  44.569 537.1651 < 2.2e-16 ***
Block                66 117.69   1.783  21.4909 < 2.2e-16 ***
Genotype             23  80.84   3.515  42.3627 < 2.2e-16 ***
Environment:Genotype 483 151.66   0.314   3.7845 < 2.2e-16 ***
Residuals           1518 125.95   0.083
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

- Environment largest effect.
- Both Genotype and Genotype by Environment interaction significant

Quantifying importance of GxE

```
> # Fit 2-way anova
> m1 <- lm(yield~  Environment + Block +  Genotype +
+               Genotype:Environment, data = dat)
> # Display anova table
> anova(m1)
Analysis of Variance Table

Response: yield
                     Df Sum Sq Mean Sq  F value    Pr(>F)
Environment          21 935.95  44.569 537.1651 < 2.2e-16 ***
Block                66 117.69   1.783  21.4909 < 2.2e-16 ***
Genotype             23  80.84   3.515  42.3627 < 2.2e-16 ***
Environment:Genotype 483 151.66   0.314   3.7845 < 2.2e-16 ***
Residuals           1518 125.95   0.083
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

- Compare variation (sums of squares) due to genotype main effects versus genotype by environment interaction.
- From a plant breeding point of view more insightful to think in terms of variation, in particular genetic variation
    - More useful to compare variance components → switch to random genotypic effects, mixed model.

Classical ANOVA (fixed effect) models…

- Not suitable to answer questions related to variation in relation to genotypes/environments.
    - genetic variances (and covariances) in relation to total phenotypic variance?
- Poorly equipped for unbalanced data (which is usually the case with GxE data).
- When many genotypes/environments tested, not parsimonious (high number of parameters to estimate).
- In many situations it is best to consider genotypes and/or environments as random terms in the model
    - We switch from fixed to mixed models

A basic mixed model: one-hit model

$$\underline{y}_{ijk} = \mu + E_j + b_{k(j)} + G_i + GE_{ij} + \underline{\epsilon}_{ijk}$$

$$\underline{G}_{ijk} \sim N(0, \sigma_g^2)$$

$$\underline{GE}_{ijk} \sim N(0, \sigma_{ge}^2)$$

$$\underline{\epsilon}_{ijk} \sim N(0, \sigma_\epsilon^2)$$

- Grain yield =
- Intercept +
- Effect of environment j (E_j) +
- Block within environment (b_{kj}) +
- Effect of genotype i (G_i) +
- Effect of genotype x environment interaction (GE_{ij}) +
- error $(\underline{\epsilon}_{ijk})$

- We regard genotypes as a sample from a larger population.
- Parameters of interest → variance components.
  - Genotype main effect variation: σ^2_G
  - Genotype by environment interaction variation: σ^2_{GE}
- No problem with unbalanced data.

Testing in mixed models

- Two types of tests:
  - Wald (or F-test): fixed effects.
    - $Ho: G_1 = G_2 = \ldots = G_I = 0$
    - $Ha: \text{at least one } G_i \text{ is different from } 0$
  - Likelihood Ratio Test (LRT) for variance components.
    - $Ho: \sigma_g^2 = 0$
    - $Ha: \sigma_g^2 > 0$

When testing for variance components, the fixed model part must be kept the same between the full and the reduced model (=nested models)

Testing of fixed effect parameters

- In our example we can test whether there are differences between the environments.
  - $H_0 : E_j = 0$ for all j (or that there are no differences)
  - $H_a : E_j \neq 0$ for at least one value of j (or at least one environment is different).
- Test statistics: Wald test
  - can be interpreted as squared t-test for individual (1 df) parameters/ contrasts
  - Under the null follows a chi-square with df corresponding to number of independent parameters (eg: number of environments our example).
- Most packages show an approximate F test statistics (more familiar than the Wald test). The F test statistics is ~ Wald / df
  - Some complications with the estimation of the denominator degree of freedom (these have to be estimated in some form, R uses the Satterthwaite's approximation)

Testing random effects in a mixed model

- Likelihood Ratio Test (LRT) or deviance test implies comparing two nested models (FULL versus REDUCED model):
  - FULL: $y = \mu + E + \underline{G} + \underline{GE} + \underline{\varepsilon}$
  - REDUCED: $y = \mu + E + \underline{G} + \underline{\varepsilon}$

- $H_0: \sigma^2_{GE} = 0$ vs $H_a: \sigma^2_{GE} > 0$ (note that the alternative is one-sided).

- $LRT = Deviance_{REDUCED} - Deviance_{FULL} \sim \chi^2_{df}$
  - Difference in deviance (-2*log-likelihood) is approximately Chi-square distributed with 1 df (for one variance parameter)

- We need to fit two models! Although R has functions to produce this test automatically.

Fit a mixed model with lme4

```
# Fit a mixed model
m2 <- lmer(yield~  Environment/Block +  (1|Genotype) +
             (1|Genotype:Environment), data = dat)

#Display summary (variance components and fixed effects)
summary(m2, ddf = 'lme4')
```

- If using the r package **lme4**: The expression (1| "term") is used to define a random term in the model
  - (1|Geno) for a random genotype main effect
  - (1|Geno:Env) for a random genotype by environment interaction effect.
- Here, blocks have been fitted as fixed, but they could also have been fitted as random

Variance components and deviance

```
> summary(m2, ddf = 'lme4')
Linear mixed model fit by REML ['lmerMod']
Formula: yield ~ Environment/Block + (1 | Genotype) + (1 | Genotype:Environment)
   Data: dat

REML criterion at convergence: 1714.2

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.3201 -0.4712 -0.0083  0.4552  5.0896

Random effects:
 Groups                Name        Variance Std.Dev.
 Genotype:Environment (Intercept) 0.05776  0.2403
 Genotype             (Intercept) 0.03637  0.1907
 Residual                         0.08297  0.2880
Number of obs: 2112, groups:  Genotype:Environment, 528; Genotype, 24

Fixed effects:
                              Estimate Std. Error t value
(Intercept)                   3.378625   0.085902  39.331
Environmentloc01_year2       -0.630292   0.108293  -5.820
Environmentloc02_year1       -2.188750   0.108293 -20.211
Environmentloc02_year2       -2.173042   0.108293 -20.066
Environmentloc04_year1       -1.634000   0.108293 -15.089
     ...                        ...        ...      ...
```

deviance

Variance components

Fixed effects
(here we show only the first part)

Test for fixed effects

```
> # Display anova table
> anova(m2, type = 1)
Type I Analysis of Variance Table with Satterthwaite's method
                  Sum Sq Mean Sq NumDF DenDF F value    Pr(>F)
Environment       247.31 11.7768    21   483 141.939 < 2.2e-16 ***
Environment:Block 117.69  1.7831    66  1518  21.491 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

- If the package ‘lmerTest’ is loaded, ‘lme4’ gives us the approximate F-test statistics (instead of the Wald test).
- The argument “type = 1” asks the sequential test (following the order of terms in the model).

Variance components estimates

```
> summary(m2, ddf = 'lme4')
Linear mixed model fit by REML ['lmerMod']
Formula: yield ~ Environment/Block + (1 | Genotype) + (1 | Genotype:Envi
   Data: dat

REML criterion at convergence: 1714.2

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.3201 -0.4712 -0.0083  0.4552  5.0896

Random effects:
 Groups               Name        Variance Std.Dev.
 Genotype:Environment (Intercept) 0.05776  0.2403
 Genotype             (Intercept) 0.03637  0.1907
 Residual                         0.08297  0.2880
Number of obs: 2112, groups:  Genotype:Environment, 528; Genotype, 24
```

- $\hat{\sigma}^2_g = 0.036; \quad \hat{\sigma}^2_{ge} = 0.058$
- $\hat{\sigma}^2_{ge}$ larger than $\hat{\sigma}^2_g$

- Variance components due to GxE much larger than the G one.
- GxE seems to be very important.
    - Relative high with respect to main effect

Comparison with other examples (from literature)

Crop	Region	Vg	Vgxe	Ve	Vg/Vge
Spring Barley	Canada	62	110	174	0.56
Spring Oat	Canada	122	132	178	0.92
Wheat	Australia	23	70	87	0.33
Winter wheat	UK	99	142	128	0.70
Potatoes	UK	9780	20570	18790	0.48
Lowland rice	Thailand	198	299	178	0.66
Lowland rice	Thailand	60	311	440	0.19
durum wheat	Algeria	0.0364	0.0578	0.0830	0.63

- In general Vgxe > Vg (ratio lower than 1), pointing to high GxE
  - Spring Oat possibly an example of relatively low GxE

Testing variance components (LRT)

```
# Fit a mixed model reduced to test GxE variance
m2red <- lmer(yield~  Environment/Block + (1|Genotype),
              data = dat)
```

- First fit a "reduced" model (ie. a model without the (1|Geno:Env) term
- Then compare full with reduced model

```
> anova(m2, m2red)
refitting model(s) with ML (instead of REML)
Data: dat
Models:
m2red: yield ~ Environment/Block + (1 | Genotype)
m2: yield ~ Environment/Block + (1 | Genotype) + (1 | Genotype:Environment)
      npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
m2red   90 1989.7 2498.7 -904.85   1809.7
m2      91 1589.1 2103.7 -703.53   1407.1 402.64  1  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

GxE analysis: One vs two-stages

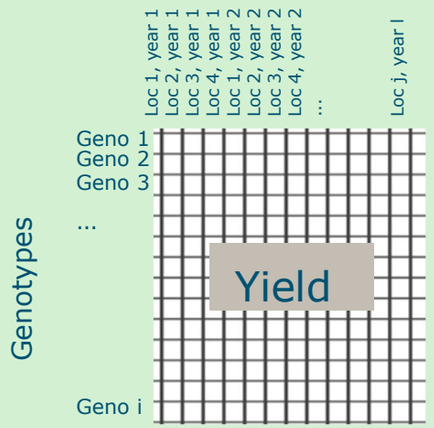
- **One-stage analysis (one-hit models):**
  - Simultaneous modelling of the within and between trial variation.
  - Uses all the information (plot data).
  - Can quickly become complicated (to account individual trials specifics).

- **Two-stage analysis:** Simpler / pragmatic approach

Stage 1: analysis per trial (focus is on the modelling of variation in a particular trial).

Adjusted means per trial

Stage 2: focus on GxE.



2-stage analysis (unweighted)

Stage 1: within-trial calculation of adjusted means to produce tables of means

Option A: One by one ('by hand')

```
# stage  1
# For each trial, fit a model for a RCBD
tr1 <- subset(dat, Environment == "loc01_year1")
m0 <- lm(yield ~  Block +  Genotype, data = tr1)
anova(m0)
```

Option B: In an automated way using statgenSTA

```
# Calculate the adjusted means considering a RCBD
mrcbd <- fitTD(TD = TD, traits = "yield",
               design = "rcbd",
               what = c("fixed", "random"),
               spatial = FALSE,
               engine = "lme4")

# Extract the BLUEs to a TD object
BLUEs <- STAtoTD(mrcbd,
                 what = c("BLUEs", "seBLUEs"))

# Print the first rows for trial 1
head(BLUEs[[1]])
```

Stage 2: Use the tables of means to fit GxE model

```
> m2 <- lm(BLUEs_yield ~  Environment +  genotype, data = bl)
> anova(m2)
Analysis of Variance Table

Response: BLUEs_yield
             Df  Sum Sq Mean Sq F value    Pr(>F)
Environment  21 233.987 11.1422 141.938 < 2.2e-16 ***
genotype     23  20.210  0.8787  11.194 < 2.2e-16 ***
Residuals   483  37.916  0.0785
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

One and two-stage analysis

- If trials are similar in precision (heritability), results of the 1 and 2 stage analysis should be very similar
- If trials are heterogeneous in their precisions, this can be taken care by the use of weights.
- Weights give more importance in the analysis to the more precise trials.
- Some literature about it:

Comparison of Weighting in Two-Stage Analysis of Plant Breeding Trials

J. Möhring and H.-P. Piepho*

Biometrical Journal **54** (2012) 6, 844–860 DOI: 10.1002/bimj.201100219

A stage-wise approach for the analysis of multi-environment trials

Hans-Peter Piepho*, **Jens Möhring, Torben Schulz-Streeck,** and **Joseph O. Ogutu**
Bioinformatics Unit, Institute of Crop Science, University of Hohenheim, 70593 Stuttgart, Germany

Received 26 October 2011; revised 22 June 2012; accepted 25 June 2012

Simple weighting methods (there are more)

- When genotype means can be assumed to be uncorrelated (for example, in a RCBD and experiments with small spatial effects)
  - Method 1: weight means by 1/(average seBLUE)^2 in designs like RCBD, where all means have the same s.e.m
  - Method 2: weight means by 1/(seBLUE)^2 in designs where genotype means have different s.e.m

- When genotype means could be correlated (for example, in designs with incomplete blocks and in experiments with strong spatial effects).
  - Methods 3 to 5: use the full variance-covariance matrix (see Piepho and Möhring 2009)

2-stage analysis (weighted)

Stage 1: within-trial calculation of adjusted means to produce tables of means

Option A: One by one ('by hand')

```
# stage  1
# For each trial, fit a model for a RCBD
tr1 <- subset(dat, Environment == "loc01_year1")
m0 <- lm(yield ~  Block +  Genotype, data = tr1)
anova(m0)
```

Option B: In an automated way using statgenSTA

```
# Calculate the adjusted means considering a RCBD
mrcbd <- fitTD(TD = TD, traits = "yield",
               design = "rcbd",
               what = c("fixed", "random"),
               spatial = FALSE,
               engine = "lme4")

# Extract the BLUEs to a TD object
BLUEs <- STAtoTD(mrcbd,
                 what = c("BLUEs", "seBLUEs"))

# Print the first rows for trial 1
head(BLUEs[[1]])
```

Stage 2: Use the tables of means and s.e.m to fit GxE model

```
# Calculate weights
BLUEs$wt <- 1/BLUEs$seBLUEs_yield^2
```

```
# Fit GxE model with weights
m3w <- asreml(fixed = BLUEs_yield ~ trial,
              random = ~genotype + genotype:trial,
              weights = wt,
              family = asr_gaussian(dispersion = 1),
              data = BLUEs)
```

```
# Predict BLUPs for the genotype main effect from the model with weights
predm3 <- predict(m3w, classify = 'genotype')
```

Summary

- Discussed and motivated the application of mixed models when analysing GxE data.
- Described models and corresponding parameters.
- Mentioned main differences between one- and two-hit models

Thanks