

# Development of a bivariate linear mixed model to describe yield response due to disease

Clayton Forknall   Alison Kelly   Steven Simpfendorfer

December 3, 2015



# Where are we going?

1. Introduction to Yield Response Curves
2. Model Formulation - Bivariate Linear Mixed Model
3. Model Application - Crown Rot Case Study
4. Future Work

# Where are we now?

1. Introduction to Yield Response Curves
2. Model Formulation - Bivariate Linear Mixed Model
3. Model Application - Crown Rot Case Study
4. Future Work

# Introduction to Yield Response Curves

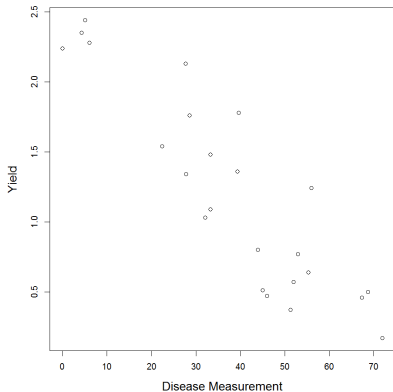
## Quantifying Production Losses due to Disease

- ▶ Production losses due to foliar and root diseases are a major financial constraint to the production of wheat and barley in Australia
- ▶ Annual losses estimated to be \$913 million in wheat alone (Murray and Brennan, 2009)
- ▶ National priority to quantify the potential production losses that could be incurred as a result of these diseases

# Introduction to Yield Response Curves

## Yield Response Curves

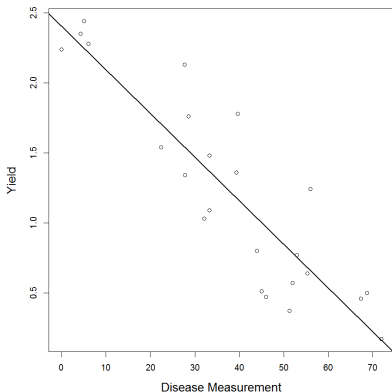
- ▶ In order to quantify production losses, response curves to be constructed
  - ▶ Current Australian wheat and barley varieties
  - ▶ Ranging from moderately resistant to very susceptible
- ▶ Response curves relate a measure of production, namely yield, to a measure of disease



# Introduction to Yield Response Curves

## Yield Potential and Yield Response

- ▶ **Yield Potential** = Intercept
  - ▶ Ability to yield in the absence of disease
- ▶ **Yield Response** = Slope
  - ▶ Response in yield as a result of disease pressure

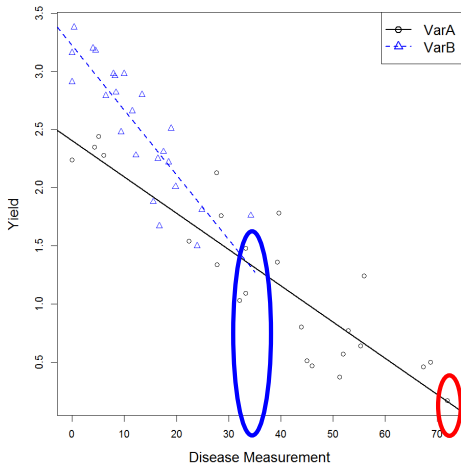


### ► Resistance

- Measure of the compatibility between host and pathogen
- Compare maximum amount of disease observed (VarB more resistant than VarA)

### ► Tolerance

- A variety's ability to maintain yield given an equivalent amount of disease pressure
- Analogous to yield response – rate of change in yield given unit change in disease
- Compare yield response (VarB is less tolerant than VarA)

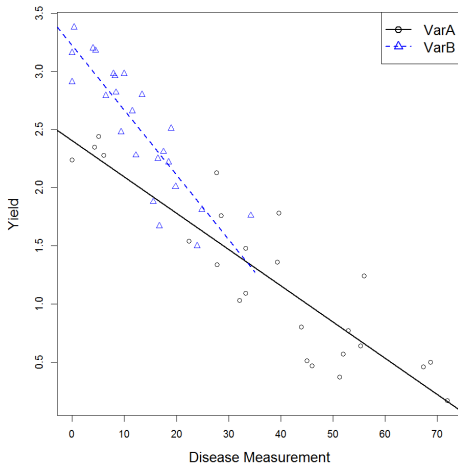


### ▶ Resistance

- ▶ Measure of the compatibility between host and pathogen
- ▶ Compare maximum amount of disease observed (VarB more resistant than VarA)

### ▶ Tolerance

- ▶ A variety's ability to maintain yield given an equivalent amount of disease pressure
- ▶ Analogous to yield response – rate of change in yield given unit change in disease
- ▶ Compare yield response (VarB is less tolerant than VarA)





# Where are we now?

1. Introduction to Yield Response Curves
2. Model Formulation - Bivariate Linear Mixed Model
3. Model Application - Crown Rot Case Study
4. Future Work

# Formulation of Bivariate Linear Mixed Model

## Limitation of Standard Linear Regression

- ▶ Explanatory variable is measured with error
- ▶ In this case, parameters estimated from standard linear regression are biased (Fuller, 1987)
- ▶ Measurement error models exist, however difficult to pull apart measurement error from signal without making limiting assumptions
- ▶ Seek a flexible model that captures variance and covariance between variables (traits)

# Formulation of Bivariate Linear Mixed Model Model for Individual Trait

$$\mathbf{y}_j = \mathbf{X}\boldsymbol{\tau}_j + \mathbf{Z}\mathbf{u}_j + \mathbf{e}_j \quad (j = Y, D) \quad (1)$$

- ▶  $\mathbf{y}_j$  - vector of length  $n$  containing values for trait  $j$ , where  $j = Y$  for yield and  $j = D$  for disease,
- ▶  $\boldsymbol{\tau}_j = (\mu_j \mathbf{v}_j)'$  - vector of fixed effects corresponding to the mean of trait  $j$  and variety main effects  $\mathbf{v}_j$  for trait  $j$ , where  $\mathbf{v}_j = (v_{1j} \dots v_{vj})'$
- ▶  $\mathbf{u}_j = (\mathbf{b}_j' \mathbf{m}_j)'$  - vector of random effects corresponding to the experimental design structure
  - ▶  $\mathbf{b}_j$  - replicate block effects for trait  $j$ , where  $\mathbf{b}_j = (b_{1j} \dots b_{bj})'$
  - ▶  $\mathbf{m}_j$  - main plot effects for trait  $j$ , where  $\mathbf{m}_j = (m_{1j} \dots m_{mj})'$
- ▶  $\mathbf{X}, \mathbf{Z}$  - designs matrices - same for both traits
- ▶  $\mathbf{e}_j$  - vector of length  $n$  containing the residuals for trait  $j$ .

# Formulation of Bivariate Linear Mixed Model Model for Individual Trait

- ▶ Random and error terms are assumed to be IID Normal random variables, with zero means and variance structures seen below

$$\mathbf{b}_j \sim N(\mathbf{0}, \sigma_{b_j}^2 \mathbf{I}_b), \quad \mathbf{m}_j \sim N(\mathbf{0}, \sigma_{m_j}^2 \mathbf{I}_m), \quad \mathbf{e}_j \sim N(\mathbf{0}, \sigma_j^2 \mathbf{I}_n)$$

By a direct extension of Equation (1), the bivariate model can be formulated as

$$\mathbf{y} = (\mathbf{I}_2 \otimes \mathbf{X})\boldsymbol{\tau} + (\mathbf{I}_2 \otimes \mathbf{Z})\mathbf{u} + \mathbf{e} \quad (2)$$

where,

$$\blacktriangleright \mathbf{y} = (\mathbf{y}'_Y, \mathbf{y}'_D)', \boldsymbol{\tau} = (\boldsymbol{\tau}'_Y, \boldsymbol{\tau}'_D)', \mathbf{u} = (\mathbf{u}'_Y, \mathbf{u}'_D)' \text{ and } \mathbf{e} = (\mathbf{e}'_Y, \mathbf{e}'_D)'$$

$\boldsymbol{\tau}$  - vector of fixed effects extended to include terms for the mean of each trait and variety effects nested within a trait

$$\boldsymbol{\tau}_Y = (\mu_Y, \mathbf{v}'_Y)' \quad \boldsymbol{\tau}_D = (\mu_D, \mathbf{v}'_D)'$$

Using unstructured variance models, the covariance between traits can be captured, leading to the following variance structures for the random and plot level (error) terms

### Random (Replicate Block and Main Plot)

$$\text{var}(\mathbf{b}) = \begin{bmatrix} \sigma_{b_Y}^2 & \sigma_{b_{YD}} \\ \sigma_{b_{YD}} & \sigma_{b_D}^2 \end{bmatrix} \otimes \mathbf{I}_b \quad \text{var}(\mathbf{m}) = \begin{bmatrix} \sigma_{m_Y}^2 & \sigma_{m_{YD}} \\ \sigma_{m_{YD}} & \sigma_{m_D}^2 \end{bmatrix} \otimes \mathbf{I}_m$$

### Plot Level

$$\text{var}(\mathbf{e}) = \begin{bmatrix} \sigma_Y^2 & \sigma_{YD} \\ \sigma_{YD} & \sigma_D^2 \end{bmatrix} \otimes \mathbf{I}_n$$

# Bivariate Linear Mixed Model

## Calculating Yield Response (Slope)

- ▶ Using the regression implicit in the variance structure at the plot level, a yield response (slope) can be calculated
- ▶ This variance structure can be as a result of a regression of the plot level effects for yield on disease:

$$\mathbf{e}_Y = \beta \mathbf{e}_D + \boldsymbol{\epsilon}, \quad \beta = \frac{\sigma_{YD}}{\sigma_D^2}$$

- ▶ As it stands, slope is averaged across all varieties in the trial
- ▶ Revise the variance structure at the plot level to enable the estimation of separate yield responses for each variety

# Bivariate Linear Mixed Model

## Allowing Separate Yield Responses for each Variety

With the aim of developing response curves for each variety, the variance structure at the plot level can be extended further to enable the estimation of heterogeneous variances and covariance between traits for each variety

$$\text{var}(\mathbf{e}_{v_i}) = \begin{bmatrix} \sigma_{v_iY}^2 & \sigma_{v_iYD} \\ \sigma_{v_iYD} & \sigma_{v_iD}^2 \end{bmatrix} \otimes \mathbf{I}_{v_{in}}, \quad (i = 1 \dots v)$$

Using the regression implicit in the variance structure, yield responses for each variety can be calculated

$$\mathbf{e}_{v_iY} = \beta_{v_i} \mathbf{e}_{v_iD} + \boldsymbol{\epsilon}_{v_i}, \quad \beta_{v_i} = \frac{\sigma_{v_iYD}}{\sigma_{v_iD}^2}$$



- ▶ The bivariate linear mixed model was fit using the ASReml-R package (Butler *et al.*, 2009) in the R statistical computing environment (R Core Team, 2014)
- ▶ A general form of the model statement in ASReml-R can be seen below:

```
asr.biv <- asreml(cbind(Yield,Disease) ~ trait + trait:Variety,  
                random= ~ us(trait):Block + us(trait):MainPlot,  
                rcov= ~ at(Variety):Plot:us(trait),  
                data=biv.df)
```

# Where are we now?

1. Introduction to Yield Response Curves
2. Model Formulation - Bivariate Linear Mixed Model
- 3. Model Application - Crown Rot Case Study**
4. Future Work

# Crown Rot Case Study

## Nil Disease vs High Disease Treatment



### ▶ **Treatments**

- ▶ Factorial treatment structure of 5 wheat varieties x 6 rates of crown rot inoculum
- ▶ Varieties ranged in resistance to crown rot from moderately resistant, moderately susceptible to very susceptible
- ▶ Rates of crown rot inoculum serve to establish differing disease epidemics

### ▶ **Design**

- ▶ Split-plot design with four replicates
- ▶ 5 varieties as main plots, 6 rates of crown rot inoculum as subplots

### ▶ **Measurements**

- ▶ Yield (t/ha)
- ▶ Crown Rot Index (%) – industry accepted measure calculated using both the incidence and severity of crown rot on a plant
- ▶ 25 plants per plot randomly sampled for disease measurement

# Crown Rot Case Study

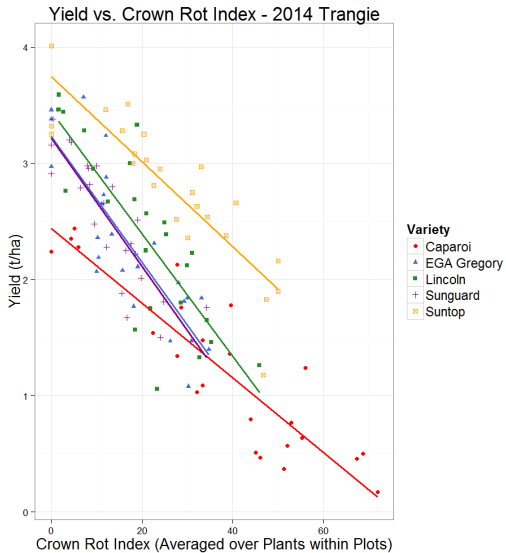
## Calculating Yield Response Curves

- ▶ Applying the bivariate linear mixed model to the 2014 Trangie experiment gives the following estimates of yield response and yield potential
- ▶ Delta method was used to calculate an approximate standard error of the yield response for each variety

Variety	Yield Response ( $\pm$ standard error)	Yield Potential (t/ha)
Caparoi	-0.032 ( $\pm$ 0.0023)	2.44
EGA Gregory	-0.054 ( $\pm$ 0.0071)	3.23
Lincoln	-0.053 ( $\pm$ 0.0081)	3.45
Sunguard	-0.055 ( $\pm$ 0.0072)	3.22
Suntop	-0.037 ( $\pm$ 0.0041)	3.74

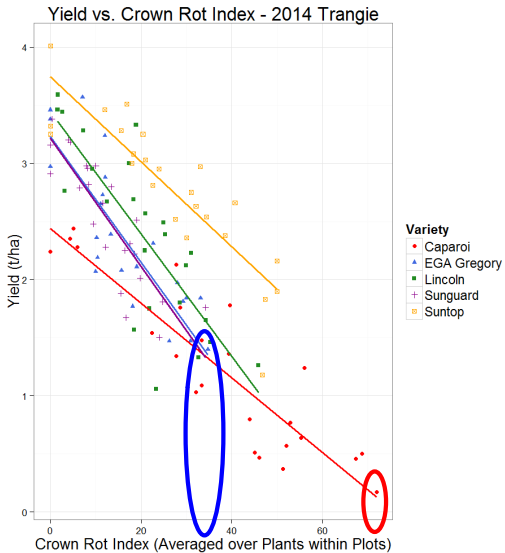
# Crown Rot Case Study

## Yield Response Curves



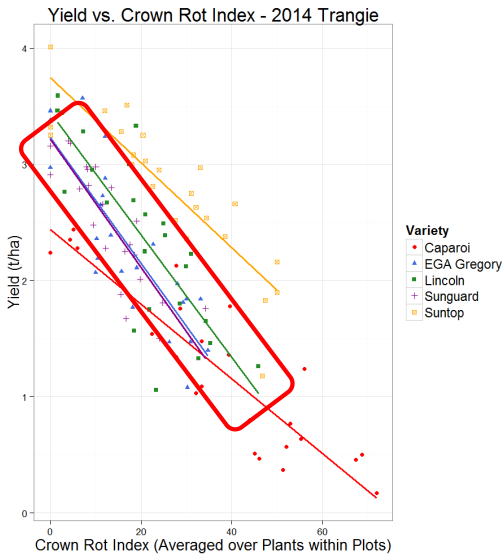
# Crown Rot Case Study

## Interpreting Yield Response Curves - Resistance



# Crown Rot Case Study

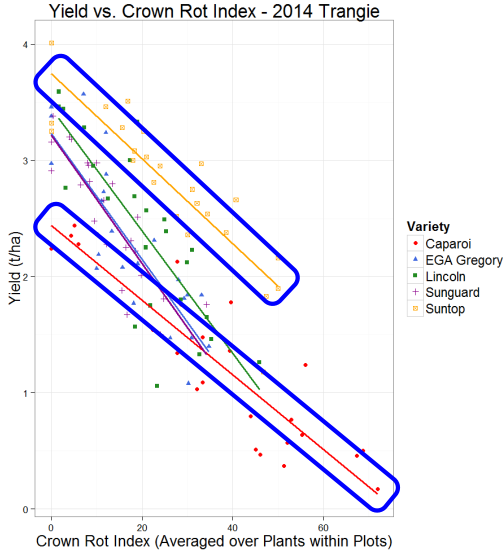
## Interpreting Yield Response Curves - Tolerance





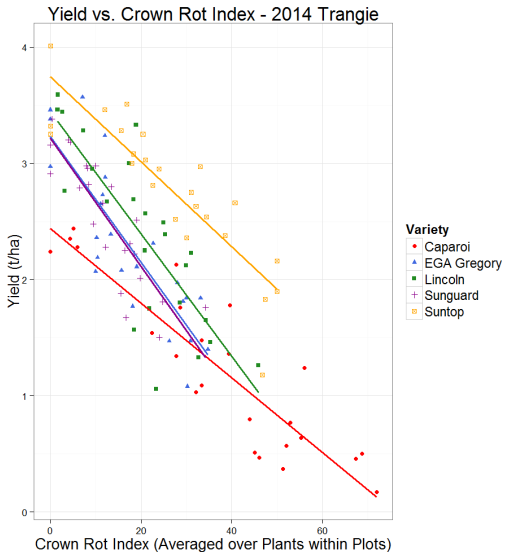
# Crown Rot Case Study

## Interpreting Yield Response Curves - Tolerance



# Crown Rot Case Study

## Interpreting Yield Response Curves - Conclusions



# Where are we now?

1. Introduction to Yield Response Curves
2. Model Formulation - Bivariate Linear Mixed Model
3. Model Application - Crown Rot Case Study
4. Future Work

# Future Work

## Extension to Across Environments

- ▶ Case study only presented for one experiment – have many experiments, conducted across different sites and years
- ▶ Seeking a method of extending the current model to investigate consistency of yield response across environments -  $G \times D \times E$ ?

# Future Work

## Testing for Equivalence of Yield Response

- ▶ If considered across multiple environments, require a method for testing the equivalence of yield responses
- ▶ This can also apply to comparing yield responses within an environment
- ▶ Currently considering the use of a REML likelihood ratio test (REMLRT)