The development of a bivariate mixed model approach for plant survival data

http://agritourismaustralia.com.au

Authors

Aanandini Ganesalingam
School of Plant Biology,
University of Western Australia
Adjunct Assoc. Prof. Alison Smith
School of Plant Biology,
University of Western Australia
Department of Industry & Investment
Dr. Cameron Beeck
The UWA Institute of Agriculture,
University of Western Australia
Winthrop Prof. Wallace Cowling
The UWA Institute of Agriculture,
University of Western Australia
Prof. Brian Cullis
School of Mathematics and Applied Statistics, Faculty of Informatics
University of Wollongong

Outline

* Data set background
* Bivariate model
* Model Results
* Findings
* Conclusion

Photo supplied by Wayne Burton, Victoria DPI

Background: measuring disease resistance

* Plant survival is the main measure of disease resistance
* Often requires multiple measurements before & after infection
* Counts are used to form an index

Project Aim

- To develop a bivariate mixed model approach for plant survival data
- Motivating example
  - 2009 blackleg disease resistance trials for canola

http://www.coggo.net.au

Blackleg disease of Canola

* Caused by *Leptospheria maculans*
* Crown cankers are the main cause of plant death
* In W.A. yield losses in 1998 and 1999 were $20M and $50M

(Khangura & Barbetti 2001)

Photo supplied by Wayne Burton, Victoria DPI

Blackleg disease resistance

- Determined by counting the number of plants at emergence & maturity for each plot
- Historically we convert to:

$$\%survival = \frac{maturity}{emergence} \times 100$$

- The bivariate approach uses both plant counts as two ‘traits’

Disease resistance at maturity

Disease susceptible

Disease resistant

Photo supplied by Canola Breeders Western Australia Pty Ltd

Motivation towards bivariate analysis

Historical

* Single derived variable: %survival
* Assumes counts at emergence are without error
* Confounds errors in emergence and maturity traits

Proposed

* Two traits: emergence & maturity counts
* Model error - spatial field trend for each trait
* Identification trait based outliers
* Examine individual trait genetic effects

Analysis – Single site disease nursery

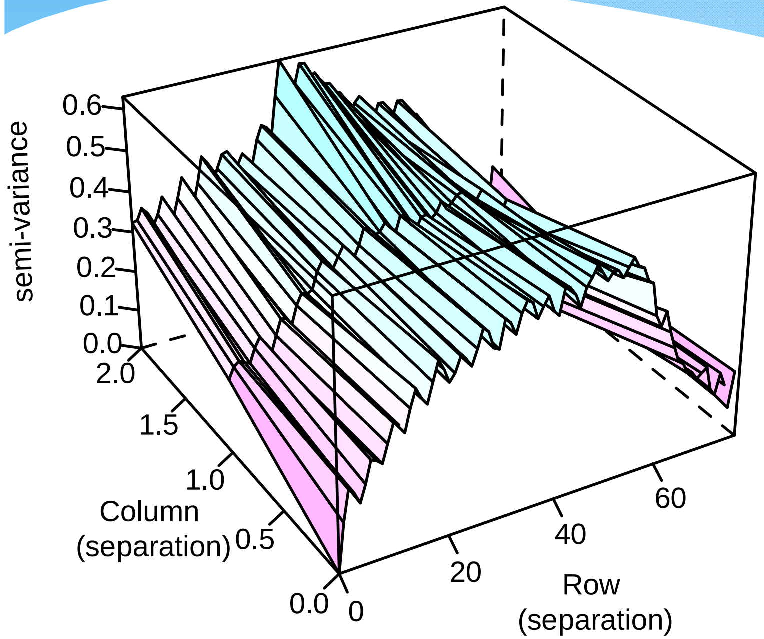
* York, Western Australia - 2009 disease nursery site
  * Plots of approx. 100 plants per variety
  * 3 columns x 79 rows
* Log of emergence & maturity counts
  * Spatial mixed model approach of Gilmour et al. (1997)
* Modelling and analysis undertaken in ASReml – R (Butler et al. 2009)

Spatial modelling of individual traits

$$y_j = X\tau_j + Z_v u_{vj} + Z_b u_{bj} + e_j$$

* $\mathbf{y}_j$ is the $n \times 1$ vector of data (j=1 for emergence, j=2 for maturity
* $\boldsymbol{\tau}_j$ is the vector of fixed effects (overall site mean)
* $\mathbf{u}_{vj}$ is the $m{\times}1$ vector of random variety effects
* $\mathbf{u}_{bj}$ is the $b \times 1$ vector of random block effects
* $\mathbf{e}_j$ is the vector of residuals ordered as per the data vector

Spatial effects for the emergence mixed model

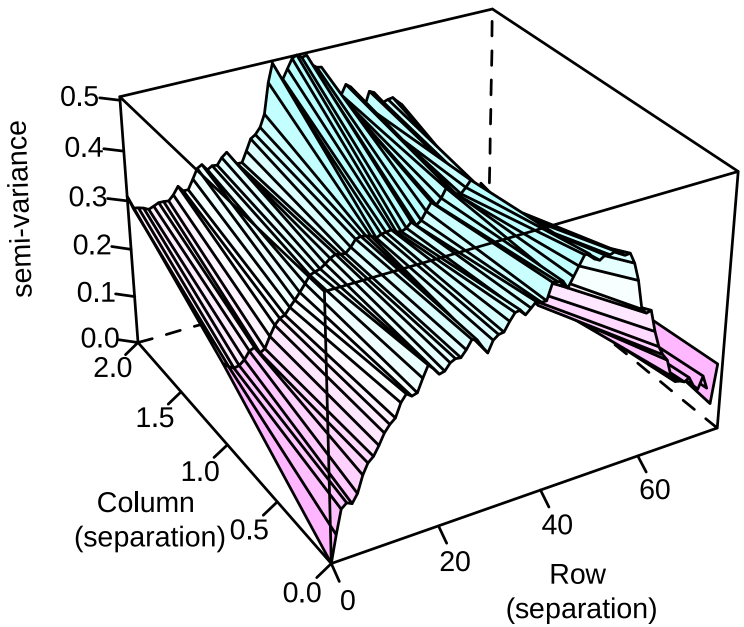


Initial trial variogram

\+

Random Row

Resulting trial variogram



Bivariate analysis

* Combines individual emergence and maturity mixed models
* Retains spatial terms from base modeling

Bivariate analysis

$$y = X^{*}\tau + Z_{v}^{*}u_{v} + Z_{b}^{*}u_{b} + Z_{o}^{*}u_{o} + e$$

- $\boldsymbol{y} = (\boldsymbol{y_1}', \boldsymbol{y_2}')$ is the combined vector of data across sampling times
- $\boldsymbol{\tau}$ is the vector of fixed effects
- $\boldsymbol{u_v} = (\boldsymbol{u_{v1}}', \boldsymbol{u_{v2}}')'$ is the $2m \times 1$ vector of random variety effects
- $\boldsymbol{u_b} = (\boldsymbol{u_{b1}}', \boldsymbol{u_{b2}}')'$ is the $2b \times 1$ vector of random block effect
- $\boldsymbol{u_o}$ includes any random effects determined in the spatial modeling
- $\boldsymbol{e} = (\boldsymbol{e_1}', \boldsymbol{e_2}')'$ is the vector of errors ordered as for the data vector

Bivariate model - assumptions

- Variety effects

$$\begin{pmatrix} u_{v1} \\ u_{v2} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{bmatrix} \sigma_{v11} & \\ \sigma_{v21} & \sigma_{v22} \end{bmatrix} \otimes \mathrm{I_m}\right)$$

↑ Covariance between emergence and maturity

- Errors

$$\begin{pmatrix} e_1 \\ e_2 \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{bmatrix} \sigma_{11} & \\ \sigma_{21} & \sigma_{22} \end{bmatrix} \otimes \Sigma_c \otimes \Sigma_r\right)$$

↑ Covariance between emergence and maturity

Model Comparisons

- Spatial model:
  REML estimates of genetic, error & row autocorrelation variance parameters

Trait	Variety	Error	ρ_r
Log eme	0.400	0.382	0.723
Log mat	0.511	0.299	0.218

- Bivariate model:
  REML estimates of genetic, error & row autocorrelation variance parameters
  - Correlation between genetic effects and errors are also shown

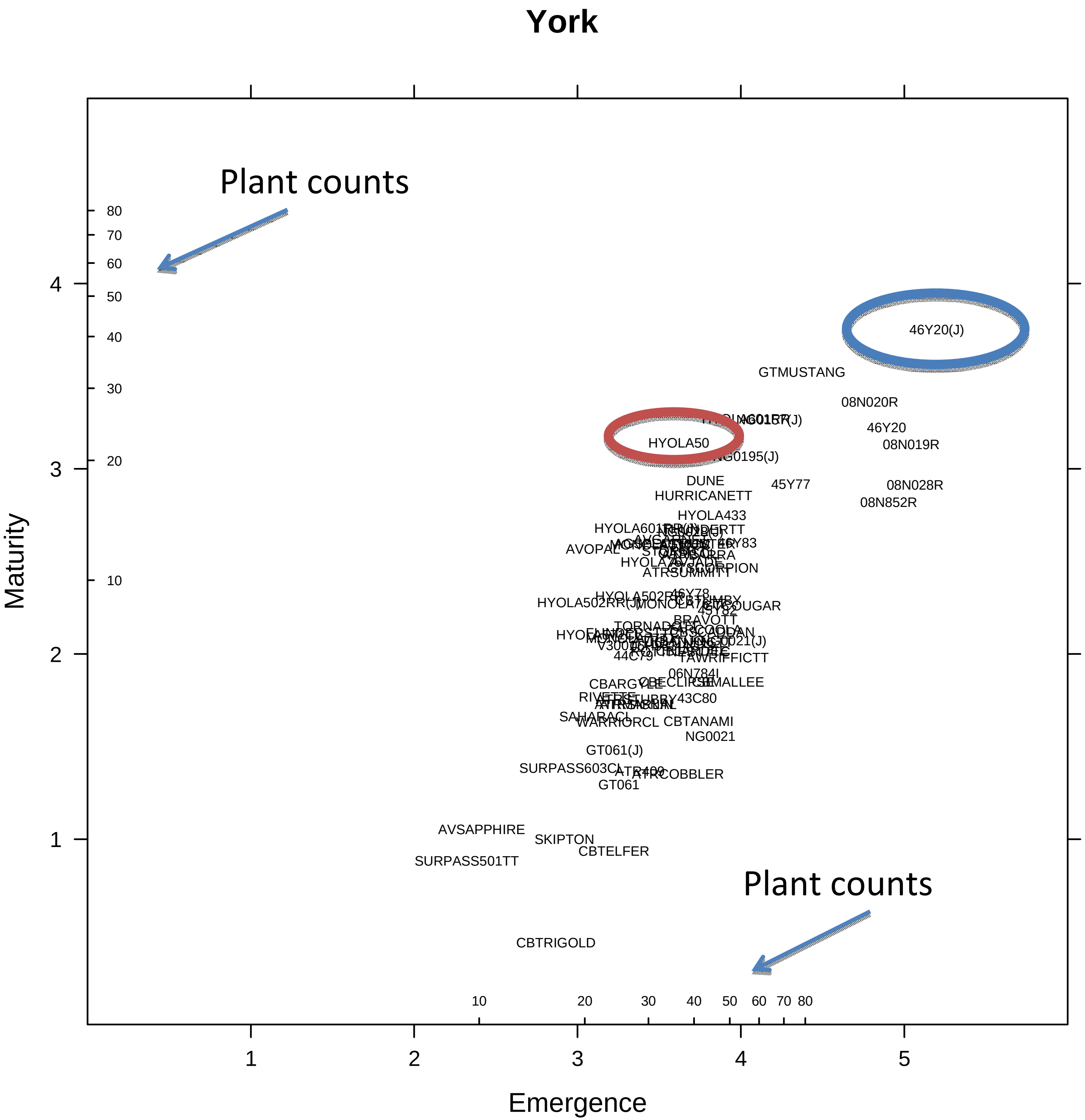
Trait	Variety	Error	ρ_r
Log eme	0.354	0.329	0.362
Log mat	0.493	0.334	
Correlation	**0.71**	**0.59**	

Variety predictions

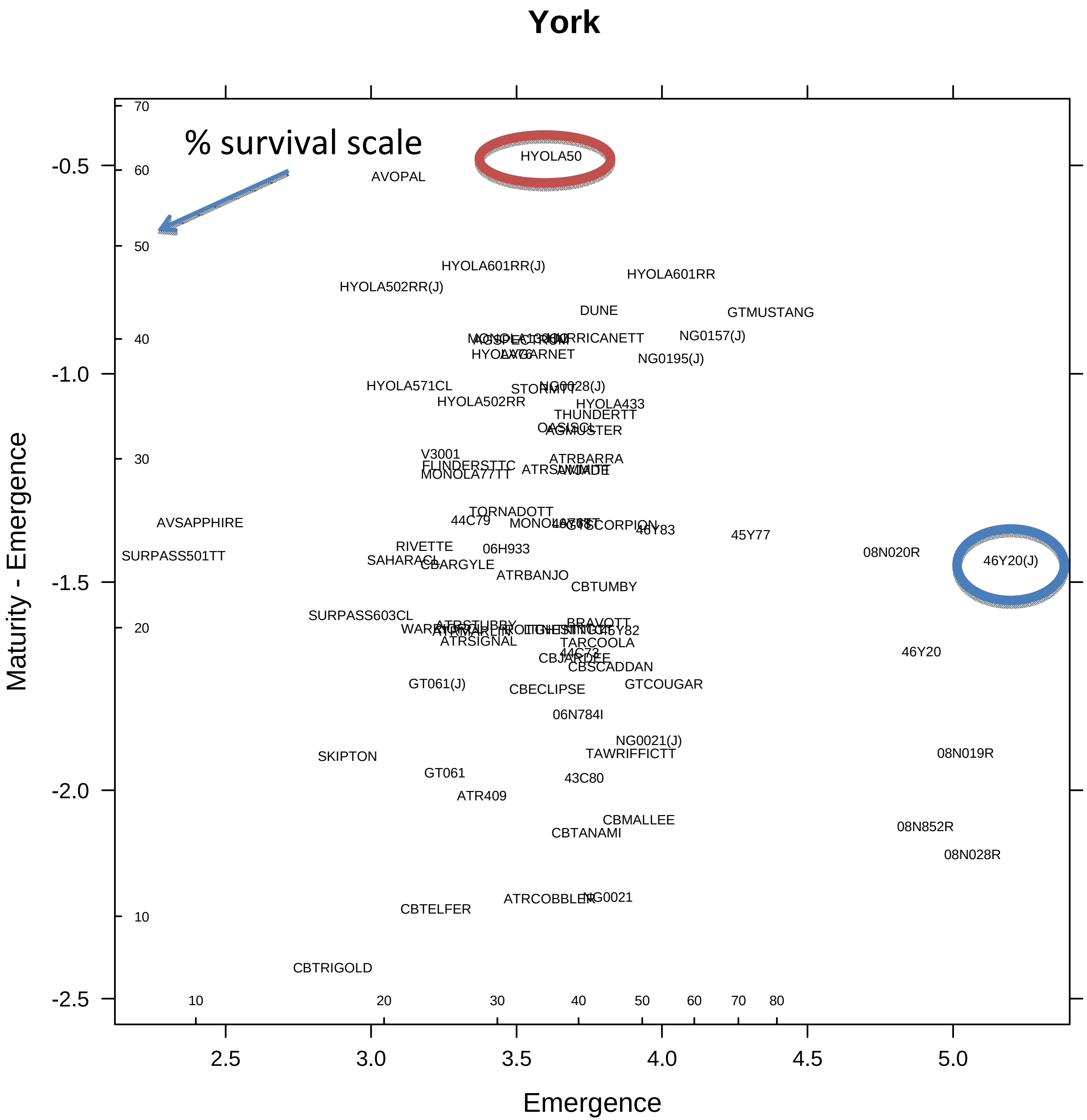
* Best linear unbiased predictors (BLUPS) of variety means
  * At emergence
  * At maturity
* Difference between predicted maturity & emergence means for variety k on the analyzed scale can be back transformed to the % survival scale

$$\exp(\hat{\pi}_{2k} - \hat{\pi}_{1k}) = \frac{\exp(\hat{\pi}_{2k})}{\exp(\hat{\pi}_{1k})}$$

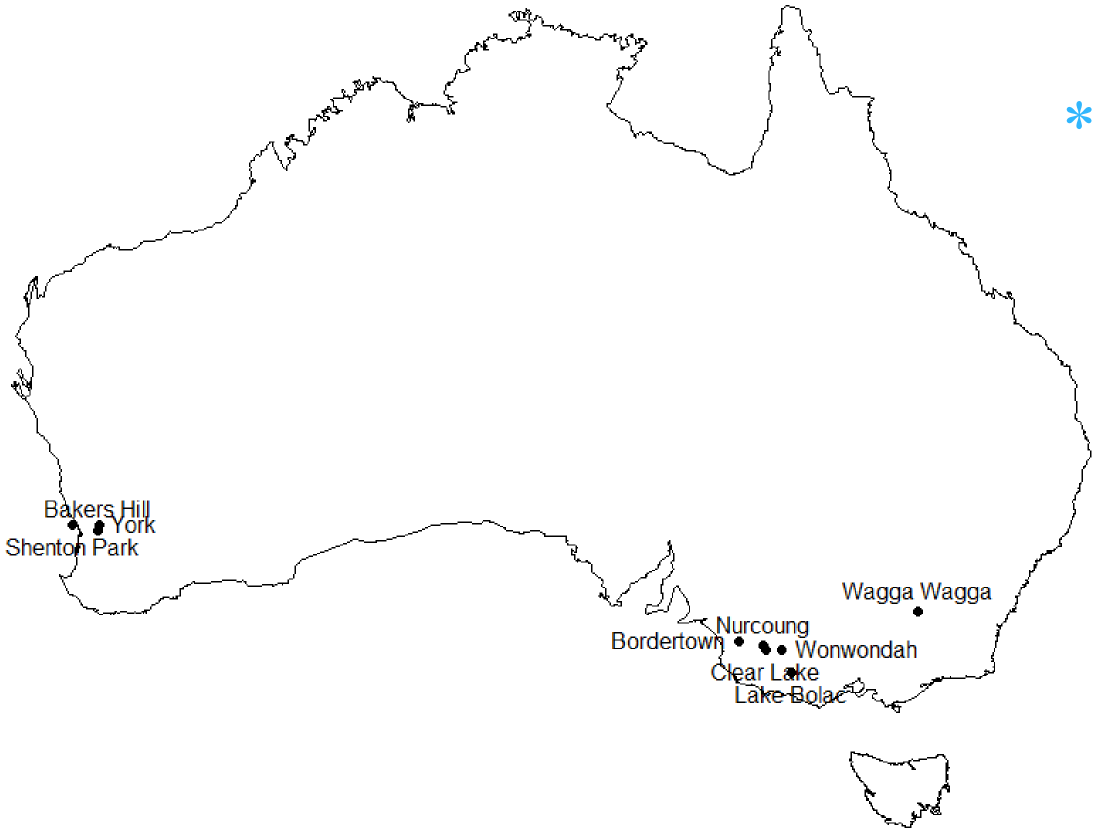
Results


York
Plant counts
Plant counts
Maturity
Emergence
1
2
3
4
5
10
20
30
40
50
60
70
80
46Y20(J)
GTMUSTANG
08N020R
46Y20
08N019R
HYOLA50
DUNE
45Y77
08N028R
HURRICANETT
08N852R
HYOLA433
AVOPAL
ATRSUMMITT
COUGAR
BRAVOTT
TAWRIFFICTT
44C79
06N784I
CBARGYLE
43C80
WARRIORCL
CBTANAMI
NG0021
GT061(J)
SURPASS603CL
ATRCOBBLER
GT061
AVSAPPHIRE
SKIPTON
CBTELFER
SURPASS501TT
CBTRIGOLD

Results


York
% survival scale
Maturity - Emergence
Emergence
70
60
50
40
30
20
10
-0.5
-1.0
-1.5
-2.0
-2.5
2.5
3.0
3.5
4.0
4.5
5.0
10
20
30
40
50
60
70
80
HYOLA50
AVOPAL
HYOLA601RR(J)
HYOLA601RR
HYOLA502RR(J)
DUNE
GTMUSTANG
NG0157(J)
HYOLA76
NG0195(J)
HYOLA571CL
NG0028(J)
HYOLA502RR
HYOLA433
THUNDERTT
V3001
ATRBARRA
FLINDERSTTC
MONOLA77TT
TORNADOTT
44C79
46Y83
AVSAPPHIRE
45Y77
RIVETTE
06H933
SURPASS501TT
08N020R
46Y20(J)
SAHARACL
CBARGYLE
ATRBANJO
CBTUMBY
SURPASS603CL
BRAVOTT
ATRSIGNAL
TARCOOLA
46Y20
CBSCADDAN
GT061(J)
CBECLIPSE
GTCOUGAR
06N784I
NG0021(J)
TAWRIFFICTT
SKIPTON
08N019R
GT061
43C80
ATR409
CBMALLEE
CBTANAMI
08N852R
08N028R
NG0021
CBTELFER
CBTRIGOLD

Findings



* Individual analysis of 9 blackleg disease nurseries from the 2009 growing season

1. Trait spatial modeling

- Spatial components Gilmour et al. (1997) differed for each trait
  - Local trend, global trend (extraneous variation)
- The number of outliers also differed between traits
  - Emergence trait having the largest number
- Such differences not obvious from the historical approach

2. Variation for emergence

* Demonstrated there is large variation between varieties for emergence
  * Variation in seed source
    * Seed lot factors: age of seed, storage (Ellis and Roberts 1980)
  * ‘Juvenile’ blackleg (Li et al., 2007)
* Emergence & maturity have different causes of variation but are strongly correlated at most sites
    * averaged 0.57 (range 0.22 to 0.94)

3. Selection for disease resistance

* Emergence & maturity BLUPs provide 3 indices for selection
  * Emergence counts
  * Maturity counts
  * % Survival values (analogous to historical method)
* Method of selection is up to the breeder's discretion

Conclusion

* The bivariate analysis is statistically more accurate than the percentage survival approach
  * Model error individually for each trait
  * Identification trait based outliers
  * Examine individual trait genetic effects

Future work…….

* A bivariate approach within a Multi Environment Trial (MET) framework

References

* Butler D, Cullis BR, Gilmour AR, Gogel BJ (2009) Analysis of Mixed Models for S Language Environments: ASReml-R Reference Manual. (Brisbane: DPI&F Publications).
* Ellis R., Roberts E. (1980) Improved equations for the prediction of seed longevity. Annals of Botany 45:13.
* Gilmour A., Cullis B., Verbyla A. (1997) Accounting for natural and extraneous variation in the analysis of field experiments. Journal of Agricultural, Biological, and Environmental Statistics:269-293.
* Khangura, R. K. and Barbetti, M. J. (2001). Prevalence of blackleg (*Leptosphaeria maculans*) on canola (*Brassica napus*) in Western Australia. Australian Journal of Experimental Agriculture, 41(1):71-80.
* Li H., Sivasithamparam K., Barbetti M.J. (2007) Soilborne ascospores and pycnidiospores of Leptosphaeria maculans can contribute significantly to blackleg disease epidemiology in oilseed rape (*Brassica napus*) in Western Australia. Australasian Plant Pathology 36:439-444.

Acknowledgements

- National Blackleg Group for the use of the 2009 blackleg disease resistance data set
- Financial support from the Grains Research and Development Corporation of Australia (GRDC)
- Bayer Crop Science PhD Research Scholarship