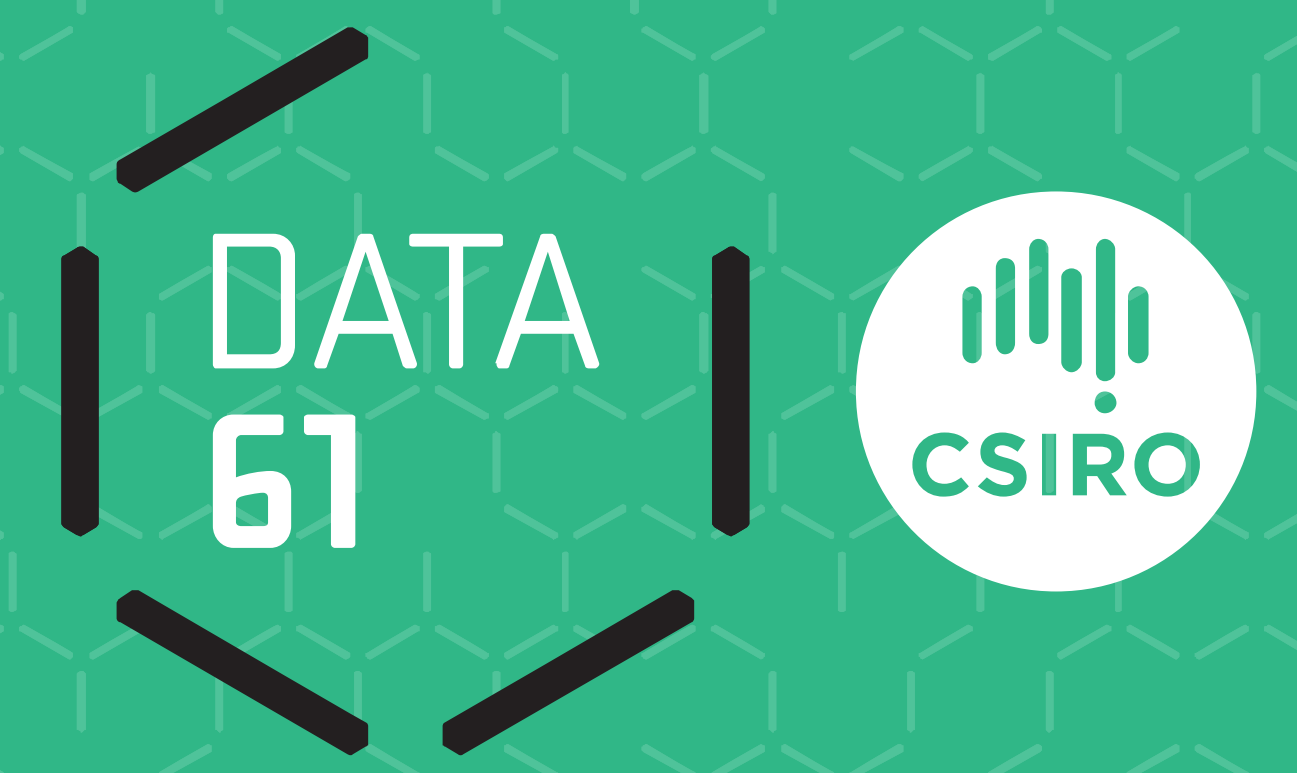


APPLYING THE FACTOR ANALYTIC MIXED MODEL TO META-ANALYSES



ROSS DARNELL, LINDSAY BELL AND JUSTIN FAINGES

www.data61.csiro.au

Introduction

Aim: Given a collection of accession trials performed in Queensland over several years, which legume species are “standouts”, those that

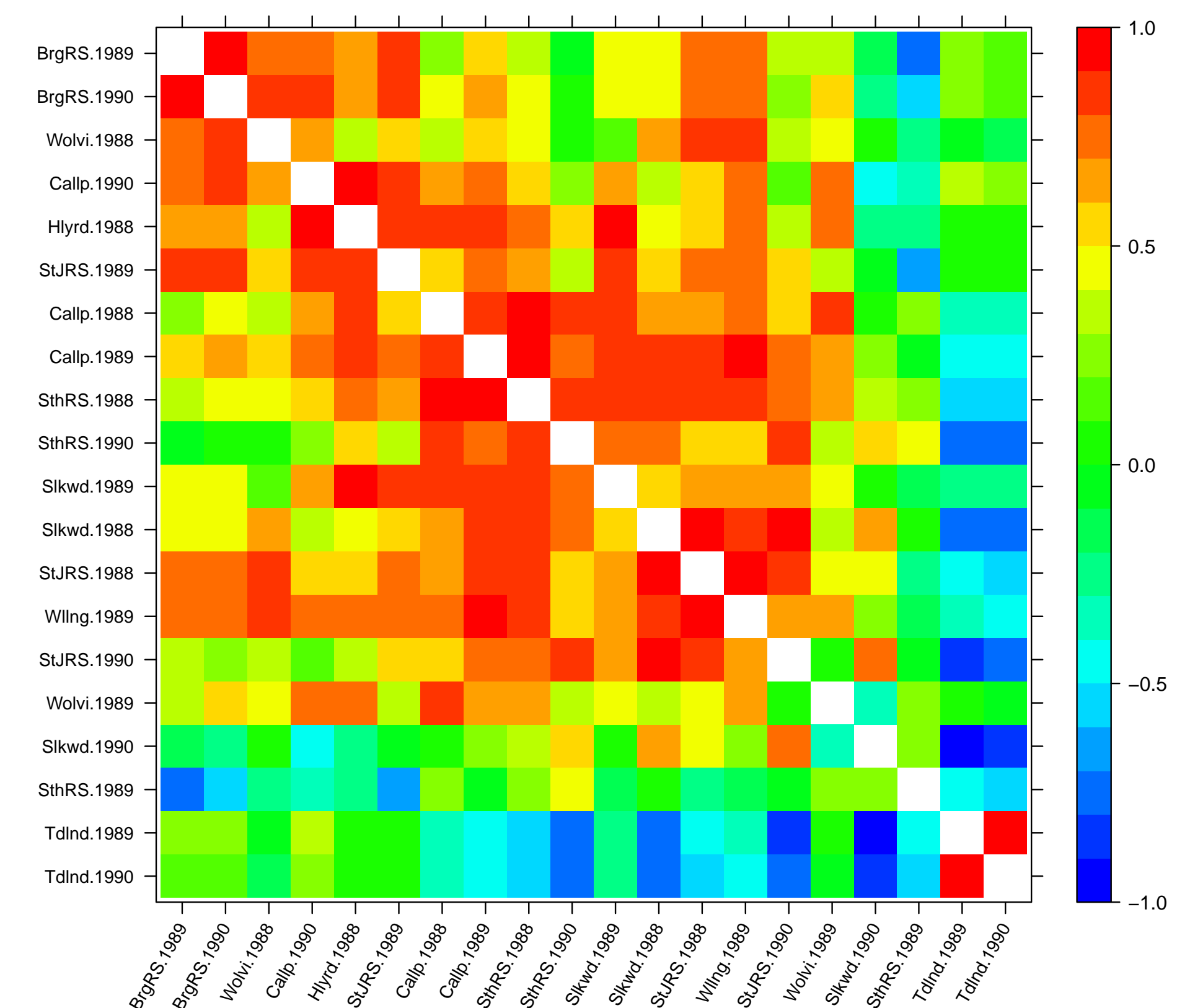
- performed well across a range of locations (robust)
- thrived at particular locations (specific)

Not all accessions were grown at all sites so the analysis is moderately unbalanced.

We follow the analysis described in [1] which is typically used for multi-environment trials when plot information is available. For cases in our legume database for which there was sufficient species data, we used factor analytic mixed models to perform these analyses and explore whether these models are useful when within-field variation information is not available.

The R [2] packages `asreml` [3] and `myf` [4] are used to generate the estimates for the FA model as well as useful plots to display the results.

Genetic correlations



The data

The data included the biomass rating (1-10 scale) of 147 species of legumes grown across 20 environments (years within location) at the first season of growth totalling 4632 records.

Site	Sowing year		
	1988	1989	1990
Brigalow Research Station		✓	✓
Calliope	✓	✓	✓
Holyrood	✓		
Silkwood	✓	✓	✓
South Johnstone Station	✓	✓	✓
Southedge Research Station	✓	✓	✓
Tedlands		✓	✓
Willunga		✓	
Wolvi	✓	✓	

Average biomass ratings at first season



Factor model

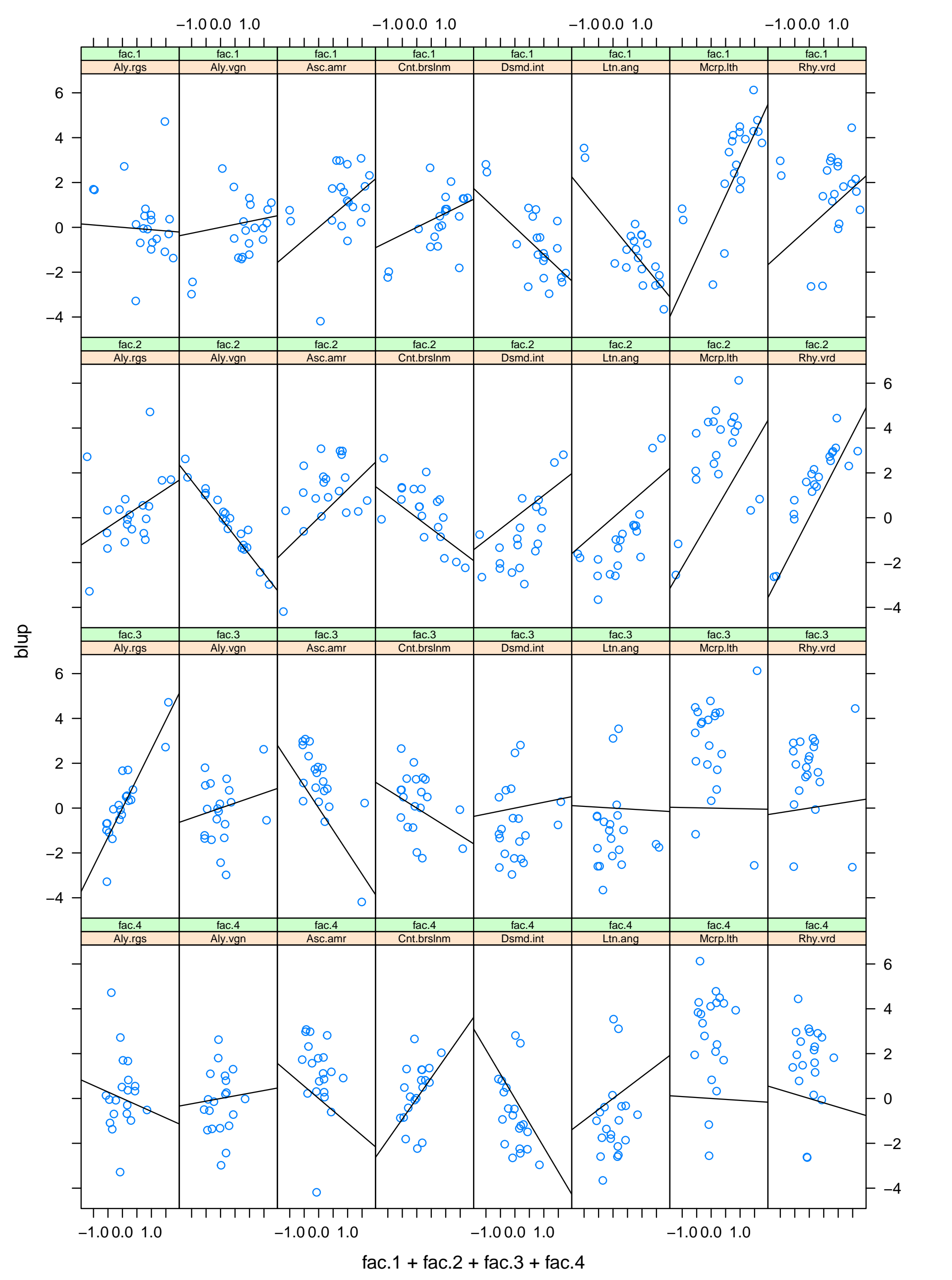
The key outputs for an FA analysis using the `asreml` and `fa.summary` functions are

- the predicted $G \times E$ interaction effects
- the estimated rotated environmental loadings for the factors and
- their factor scores (slopes) indicate responsiveness of the genotypes
- the percentage of $G \times E$ variance explained by each factor
- the estimated genetic correlation matrix
- a set of latent regression plots

Asreml model

```
asreml(fixed = value ~ Env - 1,
       random = ~fa(Env, 4):abgs +
       diag(Env):units,
       G.param = met.asr4$G.param,
       R.param = met.asr4$R.param,
       data = met1,
       family = asreml.gaussian(dispersion = 1e-04),
       na.method.X = "omit")
```

FA regression plots



Conclusion

The **genetic correlation plots** show generally good agreement among species across the sites except between Tedlands and the other sites. A possibility is that the Tedlands scorer reversed the biomass scale. If not, the performance of species at Tedlands was very different to that at other locations.

The **FA regression plots** show the response to environment for a sample of species for the first four factors. These factors explained 69, 17, 10 and 4% of the genotype \times environment interaction. The loadings for the first factor are all positive weights except for the Tedlands sowing years which agrees with the genetic correlations result above.

Looking at the plots associated with first factor (top row), the species labelled *Ltn.ang* and *Dsmd.int* do not thrive in better environments while *Mcrp.lth* responds with improving conditions.

CONTACT INFORMATION

Email Ross.Darnell@csiro.au
Phone +61 7 3833 5534

REFERENCES

- [1] Alison B. Smith, Aanandini Ganesalingam, Haydn Kuchel, and Brian R. Cullis. Factor analytic mixed models for the provision of grower information from national crop variety testing programs. *Theoretical and Applied Genetics*, 128(1):55–72, January 2015.
- [2] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2015.
- [3] David Butler. *asreml: asreml() fits the linear mixed model*, 2009. R package version 3.0.
- [4] David Butler. *myf: Utility functions for asreml objects.*, 2009. R package version 1.0.