Joint Modelling of Competition and Fertility effects in Cocoa Breeding Trials

Frank Owusu-Ansah
University of Reading

This work forms part of my PhD research at the University of Reading

Supervisors

M. Fazil Baksh
and
Eleanor F. Allan

Why breeding trial

To efficiently select best genetic material

Trial constraints

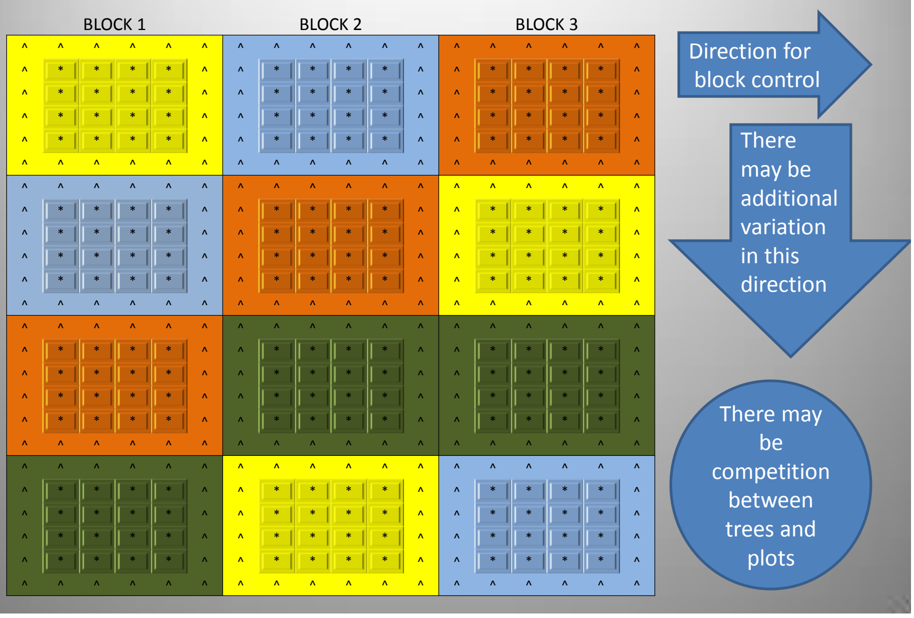
- Resources
- Heterogeneous soil
- Competition effects
- Other confounding environmental factors

Strategy

Design based options complemented with modelling

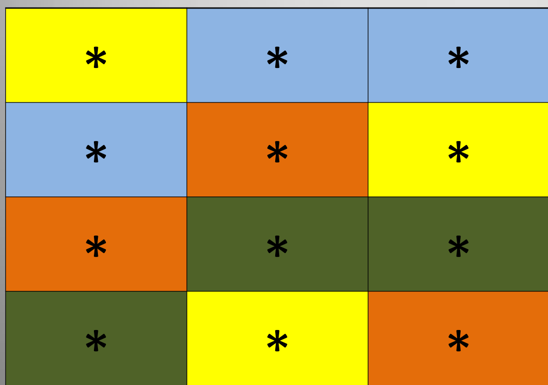
- For spatial
  - Totally Single tree Randomisation designs
  - Multiple tree plot with Complete and Incomplete Block Design
- Competition (options are generally multiple tree plots based)
  - Treatment assignment based on varietal's competitive ability (David et al.)
  - Ignoring the outer or perimeter trees and using the inner core for analysis
- Other
  - Measure to be used as covariate

RCBD with 3 blocks, 4 plots per block and 36 trees per plot



Data usually summarized for plot level analysis

BLOCK 1 BLOCK 2 BLOCK 3



- Analyze plot averages or totals based on core plot or whole plot.

- Account for
  1. Extra spatial variation with eg.Papadakis or Gilmour et al. method
  2. Competition with Kempton method
  3. Both spatial and competition with Lachenaud *et al.* or Durban *et al.*

Warning!

A complex mix of fertility and/or competition effect within a plot

leads to inefficient selection decision

My strategy

Develop a model that will not disappoint when all the basic traditional assumptions are achieved and will lead to a gain when there is a complex mix of fertility and competition effects

I plan to achieve this by extending the Durban *et al.* method from the plot to the tree level

Objectives of study

.

1. Evaluate Durban Methodology through simulation
2. Extend model to tree level
3. Application on a cocoa trial at CRIG

Materials and Methods

1. Simulation study
   - Type I error
   - Power
2. Application (Data from 8th PTA CRIG)
   - Incomplete Block design
   - Planted in 1952
   - Spaced at 2.44 meter square
   - 14 progenies
   - 4 incomplete blocks

Materials and methods cont'd

Models at tree level

$$y = X\beta + P\tau + e$$

$$y = X\beta + \rho Wy + P\tau + e$$

$$y = X\beta + X_s\beta_s + Z_s u_s + P\tau + e$$

$$y = X\beta + X_s\beta_s + Z_s u_s + \rho Wy + P\tau + e$$

The plot level versions come without the component $P\tau$

Materials and methods cont'd

$$y = X\beta + X_s\beta_s + Z_s u_s + \rho W y + P\tau + e$$

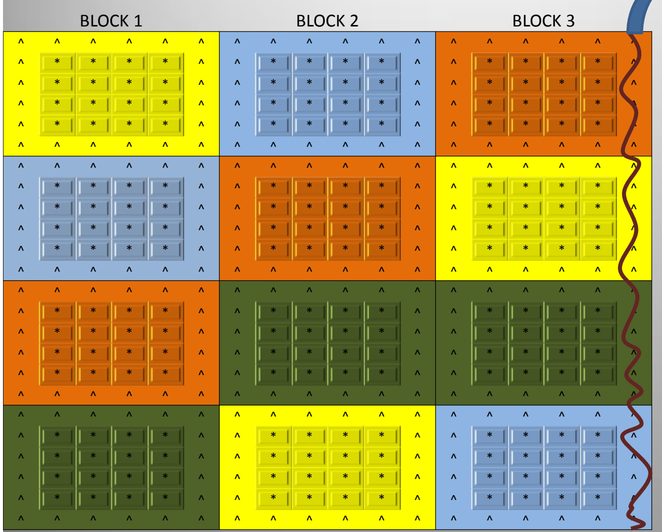
Component for fixed effect

$$X\beta$$

Here variety and blocks are treated as fixed

Materials and methods cont'd

$$y = X\beta + X_s\beta_s + Z_s u_s + \rho W y + P\tau + e$$



Component for fertility trend

$$X_s\beta_s + Z_s u_s$$

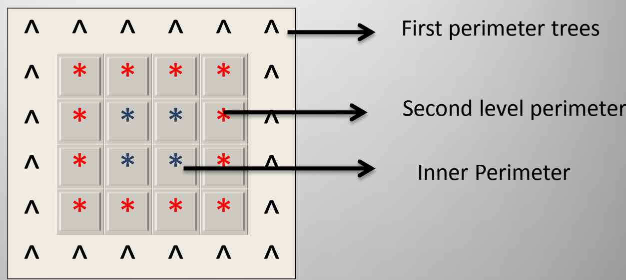
Can be formulated to have several options.
•Common trend pattern within block
To
•Different trend pattern of all columns across trial

Materials and methods cont’d

$$y = X\beta + X_s\beta_s + Z_s u_s + \rho Wy + P\tau + e$$

Component for interplot competition

$$\rho Wy$$



Choice of W can be constructed to allow for different levels of perimeter trees to be included in estimating interplot competition.

A special case is when all trees are included for estimating competition with equal weights. It is identical to plot level analysis

Materials and methods cont’d

$$y = X\beta + X_s\beta_s + Z_s u_s + \rho Wy + P\tau + e$$

The component that models the variation between plots and within plots

$$P\tau + e$$

It can also be formulated to give estimates identical to plot level analysis

Materials and methods cont'd

Analysis at the tree level here have all the model components conforming to plot level options

Results (simulation on Type I error)

Single parameter

Multiple parameter

nblocks	σ_b	T1rcbd	T1simJs	T1simJm
3	0	0.053	0.091	0.143
4	0	0.053	0.077	0.126
5	0	0.047	0.071	0.117
3	1	0.054	0.090	0.133
6	1	0.049	0.064	0.116
4	2	0.055	0.082	0.133
3	3	0.052	0.094	0.139
3	4	0.051	0.101	0.138

Results (simulation on Type I error cont'd)

Nested models for model selection using LRT

$$CH1 = RCBD \rightarrow Sm \rightarrow Jm$$
$$CH2 = RCBD \rightarrow Ss \rightarrow Js$$
$$CH3 = RCBD \rightarrow Ss \rightarrow Sm \rightarrow Jm$$
$$CH4 = RCBD \rightarrow Ss \rightarrow Js \rightarrow Jm$$

Results (simulation on Type I error cont'd)

nblocks	σ_b	CH1	CH2	CH3	CH4
3	0	0.054	0.053	0.054	0.054
4	0	0.053	0.053	0.053	0.053
5	0	0.047	0.047	0.047	0.047
3	1	0.054	0.054	0.054	0.054
6	1	0.049	0.049	0.049	0.050
4	2	0.055	0.055	0.055	0.055
3	3	0.052	0.052	0.052	0.052
3	4	0.051	0.051	0.051	0.051

Results (simulation on Power)

b	Nrcbd	CH1	CH2	CH3	CH4
3	0.791	0.791	0.791	0.791	0.791
4	0.934	0.934	0.934	0.934	0.934
5	0.987	0.987	0.987	0.987	0.987
6	0.998	0.997	0.998	0.998	0.998
7	1.000	0.999	1.000	1.000	1.000

Results (Plot level analysis from 1960 to 1966)

Yield	σ_f^2	σ_{ss}^2	σ_{js}^2	ρ_s	σ_{sm}^2	σ_{jm}^2	ρ_m
Y60	15.74	5.56	2.45	-0.26	5.27	2.47	-0.27
Y61	6.59	5.45	22.31	0.65	3.82	4.91	0.11
Y62	17.16	16.21	12.90	-0.11	16.16	13.25	-0.09
Y63	8.92	4.85	29.26	0.65	3.88	2.53	-0.14
Y64	8.43	7.48	0.01	-0.33	1.73	0.29	-0.29
Y65	28.81	32.46	8.02	-0.45	29.58	8.41	-0.43
Y66	10.03	5.94	4.60	-0.12	6.23	3.69	-0.08

Results (Tree level analyses from 1952 to 1960)

Variable	Model 1 σ_p^2	Model 1 σ_t^2	Model 2 σ_p^2	Model 2 σ_t^2	Model 2 ρ
H52JUN	15.87	86.82	15.42	86.82	0.07
H52DEC	32.11	116.54	29.22	116.54	0.14
H53MAR	92.17	305.59	67.15	305.59	0.23
D52DEC	0.39	45.34	0.39	45.34	0.01
D53MAR	0.61	46.12	0.58	46.12	0.06
D53NOV	2.43	50.39	2.19	50.39	0.10
D54MAR	3.87	52.77	3.28	52.77	0.13
D54DEC	9.22	70.21	7.39	70.21	0.15
D55NOV	27.02	105.65	18.29	105.65	0.22
D56DEC	31.70	154.09	26.93	154.09	0.14
C60JAN	383.01	3653.65	308.67	3653.65	0.12

Results (Tree level analyses from 1952 to 1960)

Variable	Model 3 σ_{pss}^2	Model 3 σ_{tss}^2	Model 4 σ_{pjs}^2	Model 4 σ_{tss}^2	Model 4 ρ
H52JUN	16.02	87.14	14.48	87.13	0.11
H52DEC	36.96	115.06	32.26	115.05	0.16
H53MAR	105.53	296.48	81.98	296.54	0.21
D52DEC	0.18	45.79	0.09	45.79	0.08
D53MAR	0.49	46.58	0.18	46.58	0.14
D53NOV	2.34	50.95	1.62	50.97	0.14
D54MAR	3.79	53.35	2.43	53.38	0.17
D54DEC	9.05	70.57	6.82	70.60	0.14
D55NOV	29.81	105.72	22.32	105.76	0.19
D56DEC	30.54	153.72	25.55	153.75	0.13
C60JAN	472.73	3590.11	361.54	3591.07	0.13

We have shown that competition effect is reflected on yield of the cocoa breeding trial

Summary

1. Proposed modelling approach gives precise estimates
2. It is important that model selection is done to ensure correct type I error
3. Absence of fertility trend within blocks or interplot competition, the joint model is equally efficient as the traditional models. When these two factors are present, the joint model could have much power to detect effects

Future work

1. Extend model evaluation to assess the implication of high missing values, and complicated pattern of dependence at both plot level and tree level analysis

2. Extend model to cope with non-normal responses

Acknowledgement

CRUK

CRIG


Thanks for your attention
End