Effective statistical evaluation of grower group on-farm trials

Session 1 ... Concerning trial design

Bev Gogel[1] Dini Ganesalingam[2]

[1]University of Adelaide beverley.gogel@adelaide.edu.au

[2]University of Western Australia aanandini.ganesalingam@uwa.edu.au

Adelaide & Kadina, October 20-21, 2014

Five fundamentals of on farm experimentation

trial design

1. decide on the aim of the research... the theory to be addressed/research question to be answered
2. generate a statistically valid experimental design to address this aim ...

trial management

3. conduct a successful on-farm trial...
   - choose an appropriate trial site
   - establish and manage trial efficiently and effectively
   - collect data as accurately as possible *including* covariates if appropriate
   - prepare the data for analysis

analysis and inference

4. *do a valid statistical analysis...*
5. *make sound statistical inference/ draw some valid conclusions...*

Outline

In this session

- being clear about the aim of the research.. the research question
- importance of *sound statistical design* to allow *valid statistical inference*
- basic design principles...*replication/randomisation/blocking*
  - some well designed trials
  - some not-so-well designed trials/common design flaws
- the role of demonstration trials
- the benefit/advantage/merit of testing over multiple sites

The research question

critical to be clear about the aim of the research and the research question (RQ) *before* the trial is designed:

- *How do two new nitrogen products compare to the standard district practice treatment of Urea?*
- *What is the effect on yield for different levels of nitrogen application for a range of wheat varieties?*
- *What is the effect of various combinations of macro-nutrient level and lime rate on wheat grain yield?*

because

- the RQ will provide a clear focus for the experimentation
- the RQ will determine how the experiment is designed, set up and conducted
- if the RQ is right, it *will* provide a better understanding of the aspect/practice/issue

Trial design

once RQ is clear, design will then involve:

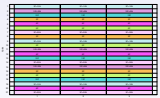
- choosing an appropriate set of treatments to address the research question
- choosing an appropriate trial site and area within that site
- establishing an appropriate experimental layout consisting of *experimental units* to which the treatments are applied:
  - experimental units usually the smallest unit on which a measurement is recorded eg. field plot, pot in glasshouse
- allocating the treatments to the experimental units to
  - allow the data to be analysed using statistical methods
  - allow treatment differences to be estimated in a *statistically valid way* and *as accurately as possible*

Trial design: examples

- *How do two new nitrogen products compare to the standard district practice treatment of Urea?*
  - set of treatments involving the new nitrogen products, treatment(s) involving urea and an untreated control
  - randomised complete block (RCB) design

- *What is the effect on yield for different levels of nitrogen application for a range of wheat varieties?*
  - wheat varieties of interest & different levels of nitrogen application
  - standard split-plot trial

- *What is the effect of various combinations of macro-nutrient level and lime rate on wheat grain yield?*
  - set of lime applications & set of macro-nutrient levels
  - strip-plot trial



Basic design principles

the three basic principles in experimental design include

- replicating the treatments
- randomising the treatments to experimental units (plots, pots,...)
- blocking the trial area into sections that contain similar experimental material

replicating the treatments

- increases the information on each treatment
- allows us to
  - estimate the sources of variation in the data and then
  - make statements about the treatments that have statistical backup
- is a safeguard against losing all of the information on a particular treatment for some unforeseen reason (animal damage, disease, ...)

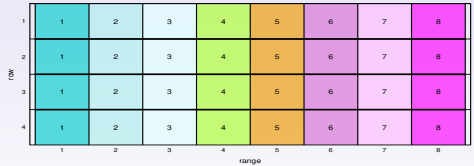
blocking the trial area

- into sections that contain similar experimental material, such that
- the variation *within blocks is minimised* while the variation *between blocks is maximized*
  - blocks might be
    - series of adjacent ranges in a field trial
    - blocks of pots from east to west in a glasshouse
    - morning and afternoon sessions in a malt run
- so we can control for broad trends and extraneous effects (more later)

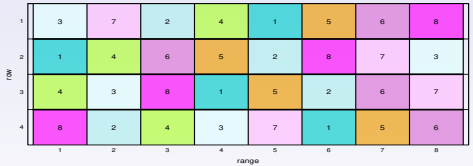
Basic design principles

randomising the treatments to experimental units (plots, pots, ...), that is

- randomising *within* blocks
- re-randomising *between* blocks
  - to avoid biasing and confounding the results
  - so we can make valid comparisons between treatments



completely unrandomised/systematic



randomised within & between reps

A well designed on farm trial

Brome grass control trial

aim:

- to compare a set of grass control treatments

standard RCB trial:

- 4 reps of 9 treatments
- 4 row $\times$ 9 range array of plots
- replicate blocks = rows

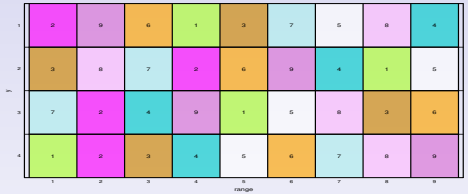


Figure: Allocation of treatments to plots

except ...

- treatments *only randomly allocated to 3 of the 4 blocks...* *will come back to this*

Another well designed on farm trial

Herbicide tolerance trial, faba beans

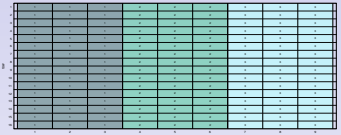
aim:

- to compare the effect of a range of herbicide treatments on the yield of 3 varieties of faba bean

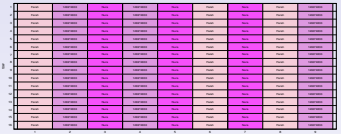
standard strip plot trial:

- 3 varieties & 13 herbicide treatments: 6 at hi/low + 1 UTC
- replicate blocks = 3 ranges
- varieties randomly allocated to main-plots
- herbicide treatments randomly allocated to sub-plots in main-plots

reps in sets of 3 ranges



varieties in mainplots in each replicate



herbicides in rows in reps

A not so well designed on farm trial

Nitrogen products trial

aim:

- to evaluate 3 nitrogen products

trial design:

- 3 rows by 8 ranges
- treatments systematically allocated to all plots in a range
- *... no randomisation*

the consequence:

- treatments confounded with ranges...high risk of a treat *looking good* because of where planted, not due to its true potential to give increased yield...
- no valid statistical analysis/inference ...
- use as a simple demonstration trial only

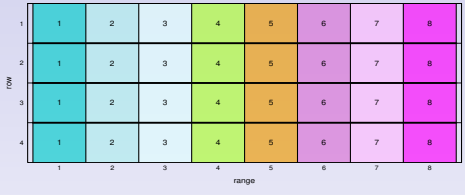


Figure: Systematic allocation of varieties to plots

The treatments were all sown in a direct line meaning that the randomisation for each was greatly compromised. As such, the results here should be treated with some caution and viewed as a comparison against the nearest neighbour.

The Urea, Agrocote and ESN have all shown significant yield responses over the untreated control (UTC) with the Urea and Agrocote both significantly higher than the ESN and Untreated

Another not so well designed trial

Macro-nutrient $\times$ lime trial

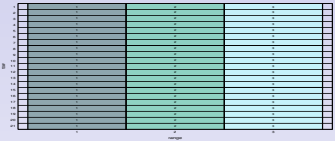
aim:

- to compare the effect of various combinations of macro-nutrient level and lime rate on wheat grain yield

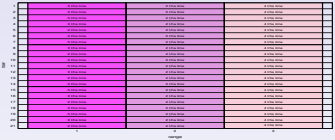
strip plot trial:

- 3 lime applications & 7 treatment levels
- replicate blocks = 3 ranges
- lime rates also aligned with ranges
- treatments
  - randomly allocated to rows 1. . . 7, 8. . . 14, 15. . . 21
  - no re-randomisation across ranges

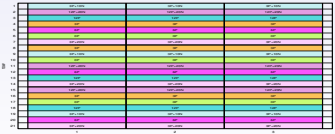
reps in 3 ranges



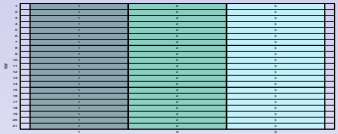
lime rates *also* allocated to 3 ranges



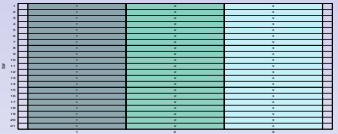
macro-nutrients in rows across ranges



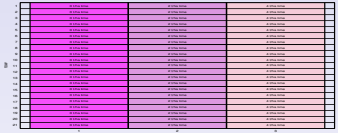
reps in 3 ranges



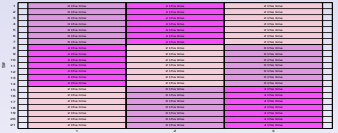
reps in 3 ranges



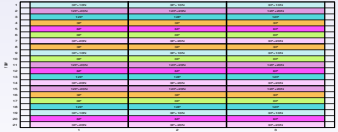
lime rates *completely confounded* with reps



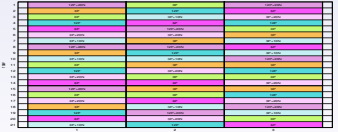
lime rates randomised to main plots in each rep



treatments in rows across reps



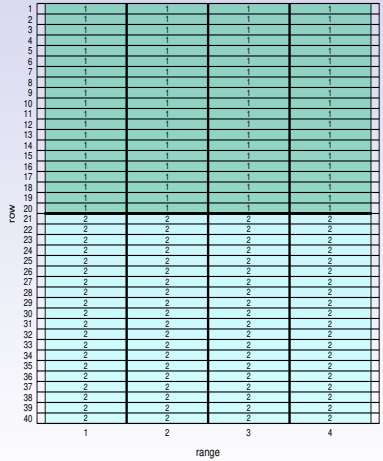
treatments randomised within main plots



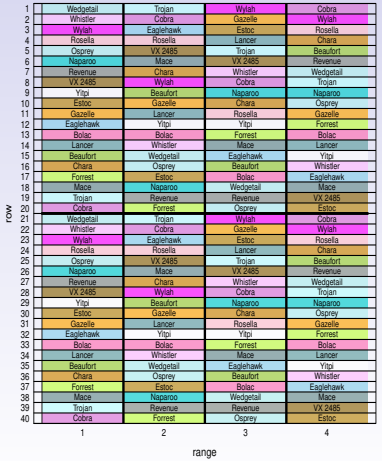
And another not so well designed trial

Time of sowing $\times$ Variety trial '

time of sowing



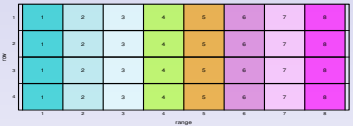
variety allocation the same across TOS



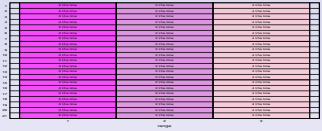
Common design flaws summary

common design flaws in the on farm trial context:

- no true replication ... reps called *pseudo* or *false* replicates
  - not possible to validly compare treatments

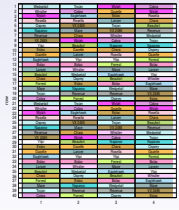


completely unrandomised/systematic



no randomisation to main plots in reps

- no re-randomisation between blocks
  - treatment performance confounded with its location in field
  - for trials with spatial trends (most trials) comparisons between neighboring treatments will be much more reliable



variety allocation the same across TOS

Some statistically significant statements

... poor trial design is a serious issue and can have really bad consequences

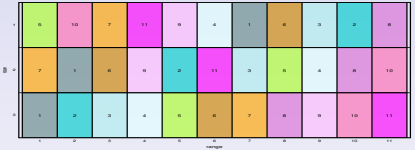
... if in doubt, consult a biometrician/statistician at the design phase of the experimental process...

The place for demonstration trials

demonstration trials are

- common in on farm trialling
- often positioned in a rep on one side of main trial
- treatments ordered systematically 1 . . . no. treatments

fertilizer trial

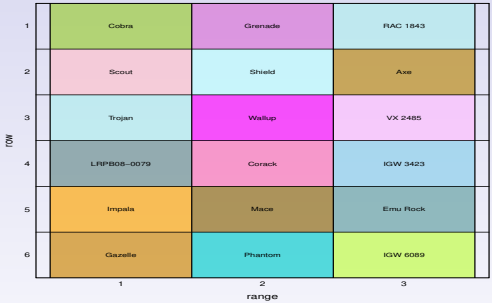


wheat variety trial

- unfortunately always a risk of prejudicial side or neighbour effects
- *by all means* include a demonstration trial
- *do not* include it in the analysis of the main trial!

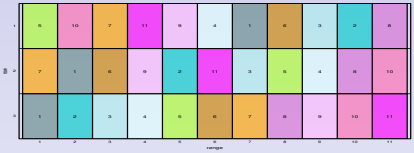
A demonstration trial

Variety demonstration



Binary designs

lots of self adjacencies in ranges: not binary



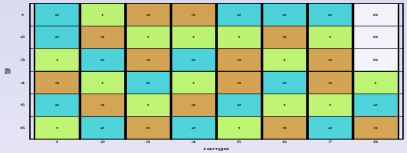
no self adjacencies in ranges: binary design

a design is binary with respect to a given blocking factor (replicate blocks, rows or columns, ...) if each treatment occurs no more that once within each level of the blocking factor...

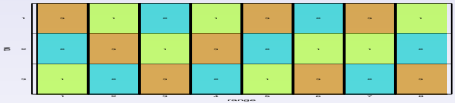
- self adjacencies in ranges mean confounding between treatments and ranges
- self adjacencies also sub-optimal in spatial terms
- *best to utilize binary designs if at all possible!*

Testing over multiple sites

Fertilizer catalyst evaluation in canola



site 1



site 2

the main strength in testing over multiple sites:

- inference over more than 1 set of environmental conditions