Data Stories - Demystifying Agricultural Science

Annie Conway[1], Sam Rogers[1]

[1] Biometry Hub, University of Adelaide
annie.conway@adelaide.edu.au

Introduction to Data Stories

A data story takes an example of a real agricultural data set and tells the story of how we might explore and analyse it. We aim to create a resource that both showcases new and exciting work in agricultural science, and demonstrates statistics and data science in an intuitive and engaging way.

The intended audience ranges from beginners in statistics, such as students and researchers of agriculture, to more experienced users such as junior statisticians and biometricians.

Below is an abridged version of the story “Crop Rotation: Analysis with Linear Mixed Models”. Follow the link in the QR code to see the full story and more.

What is Crop Rotation?

Crop rotation, or sequentially growing different crops on the same plot of land, is an agricultural practice dating back to as early as the Han dynasty of China over 3000 years ago [1], [2]. Increasing crop diversity through crop rotation can help with improving soil nutrition, mitigating weed, insect and pathogen pressure, and improving yield and quality of grain crops. [3], [4].

During the mid-20th century, the use of crop rotations declined due to an increase in the availability of nitrogen fertilizer, herbicides for weed control and pesticides for insect control. However, crop rotations provide a variety of benefits that cannot be entirely replicated by the use of chemicals, and crop rotations still play an important role in farm management today.

Crop Rotation Key Terms

- **Sequence:** a particular ordering of crops to be planted, e.g. Wheat, Lentil, Clover, Wheat, Lentil, Clover etc.
- **Cycle:** a single repetition of a sequence.
- **Phase:** Sequences starting at different points in the same cycle are said to be in different phase.
- **Rotation:** A rotation refers to the completion of a particular cycle. For example, Wheat-Lentil-Clover, Lentil-Clover-Wheat and Clover-Wheat-Lentil are three possible rotations from the Wheat-Lentil-Clover cycle.
- **Break crop:** When considering the effect on a specific crop, for example wheat, a break crop refers to a “break” in the continuous wheat cycle. A **single break** refers to one year of non-wheat planted before wheat, a **double break** refers to two years of non-wheat planted before wheat etc.

Experiment Description

Selected crop rotations were compared over a four year trial in South Australia’s southeast. The experimental design included a continuous wheat rotation (WWWW) to act as a control when comparing the effect of crop rotations on wheat production. Here is a summary of rotations and break crops used in the experiments:

- **Cereal crops:** wheat, barley, oats
- **Annual Pasture Legumes (the gold standard break crops):** balansa clover, sub clover, medic
- **Other crops:** lupin, faba bean, canola.

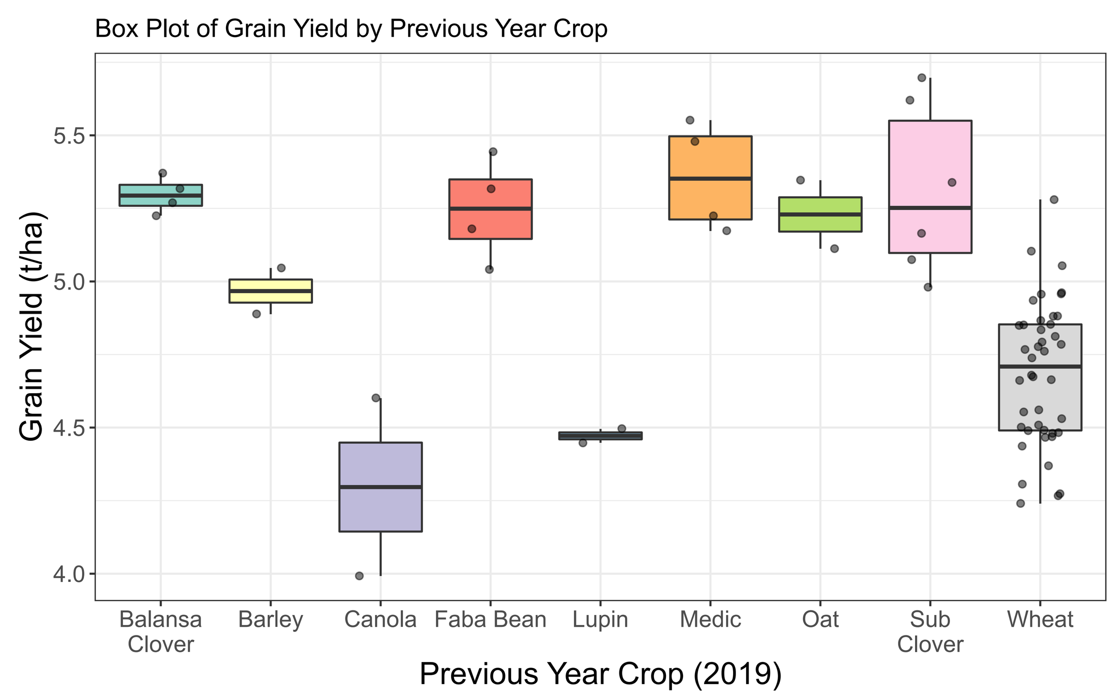
Sixty rotations (covering 18 sequences and a continuous wheat control) were randomised to two blocks, in a randomised complete block design (RCBD).

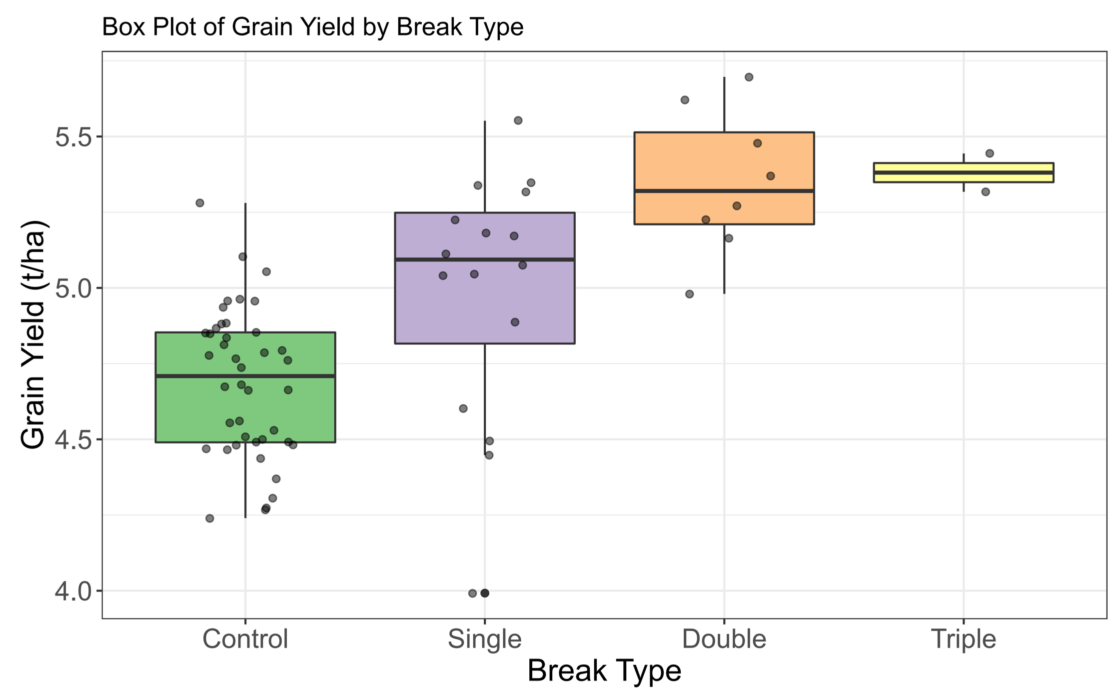
Research Questions

Consider only the wheat harvest in 2020. As this was the last year of the trial, we can consider the effect of the whole four year rotation on the performance of wheat in that year. The following questions relate to the performance of wheat yield.

- How is the current season’s wheat yield affected by the previous crops planted in the same soil?
- What is the effect of the break type? For example, is a double break better than a single break at maximising wheat yield in the current year?
- How well do annual pasture legumes (medic, balansa clover, sub clover) perform as break crops, relative to the other break crops (lupin, faba bean, canola, cereal crops) and continuous wheat?
- Is the whole 4 year rotation important for improving wheat yield, or is only the previous year (or 2, 3 years) important?

The following box-plots illustrate the raw data, and help us visualise the research questions.





Testing Hypotheses

We can turn our research questions from above into hypotheses to be tested.

Firstly, we can test whether having a break crop (versus continuous wheat) makes any difference.

- H_0: A break crop leads to no difference in mean wheat yield.
- H_A: There is a difference in mean yield where there was a break crop versus no break.

We would also like to consider the effect of the type of break.

- H_0: There is no difference between single, double or triple breaks when it comes to wheat yield.
- H_A: At least one type of break is different to the others.

Lastly, we also need to test whether the specific combinations of crops have an effect. It may be that only the previous year crop is important, or it may be that the previous two or three years have an impact, so we will need to consider all these.

- H_0: Previous year crop mean yields are all equal.
- H_0: Combined two year crop mean yields are all equal.
- H_0: Combined three year crop mean yields are all equal.

We will model the crop rotation data with a linear mixed model, fitted with REML. REML is robust to unbalanced models (where the allocation of treatments to plots is not even). The model equation is as follows:

Wheat Yield in 2020 ~ Control + BreakType + 2019 Crop + 2018-2019 Crop + 2017-2018-2019 Crop + *block* + *row* + *bay.*

The terms in italics are modelled as random effects.

The model is fit in R, with the package ASReml-R [5]. The hypotheses are tested with the Wald Test, where adjustments according to Kenward and Roger [6] are carried out, giving us approximate F-values in the ANOVA table.

Table 1: ANOVA Table

	Df	denDF	F.inc	F.con	Pr
Control	1	22.2	93.800	93.800	0.000
Break Type	2	30.4	10.680	1.671	0.205
2019 Crop	7	26.6	9.221	9.221	0.000
2018-2019 Combined	11	25.4	4.171	4.171	0.001
2017-2018-2019 Combined	12	26.9	1.669	1.669	0.131

The effect of control is tested first. This tests whether there is a difference in the mean grain yield of plots where there was no break (control) versus the mean of all plots where there was a break.

Secondly, we check the effect of BreakType. Since the Control has already been tested, this tests whether there is a difference in the single, double or triple break. The associated p-value = 0.2049, and this is greater than 0.05, so we accept the null hypothesis: there is insufficient evidence to say that there is a difference in the three types of break.

The two year combination has associated p-value = 0.0014, which is less than 0.05 and therefore significant. Since there are significant differences in the means of 2 year combinations of crops, we will not consider the effect of the 2019 crop only, as this would ignore the significant effect of variation due to the crop from the year before.

From our statistical modelling, we are able to conclude that crop rotations have an impact on the yield of wheat crops.

To see more of this data story, including R-code to implement the analysis, follow the link in the QR code. We would like to acknowledge Amanda Pearce (SARDI) for providing the data for this story, and thank the Yitpi Foundation and GRDC for the funding of the data stories project.

References

1. R. J. MacRae & G. R. Mehuys, The Effect of Green Manuring on the Physical Properties of Temperate-Area Soils. In B.A. Stewart,ed., *Advances in Soil Science* (New York, NY: Springer, 1985), pp. 71–94. https://doi.org/10.1007/978-1-4612-5090-6_2.
2. D. L. Karlen, G. E. Varvel, D. G. Bullock, & R. M. Cruse, Origin of Crop Rotations. *Advances in agronomy*, **53** (1994) 2–.
3. D. G. Bullock, Crop rotation. *Critical Reviews in Plant Sciences*, **11** (1992) 309–326. https://doi.org/10.1080/07352689209382349.
4. M. D. McDaniel, L. K. Tiemann, & A. S. Grandy, Does agricultural crop diversity enhance soil microbial biomass and organic matter dynamics? A meta-analysis. *Ecological Applications*, **24** (2014) 560–570. https://doi.org/10.1890/13-0616.1.
5. D. Butler, *Asreml: Asreml() fits the linear mixed model* (2009).
6. M. G. Kenward & J. H. Roger, Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*, **53** (1997) 983–997.