The design and analysis of experiments using yield monitoring technology

Sonia Knight and Tony Pettitt

School of Mathematical Sciences, Queensland University of Technology, GPO Box 2434, Brisbane QLD, 4001. www.fsc.qut.edu.au

Email s.knight@qut.edu.au

Abstract

Precision agriculture technology provides on-the-go information in the harvesting process on spatial location of grain yield, water and protein content. Such spatial measurements are subject to various errors. For example, the spatial resolution of yield data is mainly affected by mixing of grain within the harvester, which smooths yield and leads to highly correlated data. Some errors can be removed or corrected by appropriate filtering and aggregation techniques. Care has to be taken in interpreting the resulting data. The design and analysis of experiments need to consider variation and bias in all aspects of the farming process including the potential of contolled traffic, management zones, and variable rate technology.

Key words

Convolution, design of experiments, GPS, Spatial data

Introduction

Yield monitoring technology and other precision agricultural (PA) practices can be useful tools in the conduct of on-farm experiments, however there are errors and biases associated with the technologies that must be considered. The accuracy of the yield sensor relies heavily on proper calibration of the sensor and is also sensitive to environmental conditions. Accounting for the convolution of grain within the harvester as it is cut and threshed is another of the major challenges of yield monitoring technology. Consequently, yield being measured at a particular second is actually a weighted-average of yield over about 30 m. Furthermore, inaccuracies occur when the harvester starts and stops. If double harvesting occurs, where the harvester partially overlaps a previously harvested area, more error is introduced, a source of bias that is difficult to detect. On non-controlled traffic fields, GPS may be used to locate the position of not only the harvester, but also other equipment used to prepare and treat the paddock (such as variable rate equipment). The accuracy of the GPS system will dictate the accuracy of plot boundaries.

This poster describes the error process inherent in these measurements with the aim of designing trials that can use PA technology rather than small plot technology and be adopted by producers with minimal impact on their normal management practices.

Design of Experiments

Harvesting practices

Careful calibration of yield and water and protein content sensors is essential. Maintaining a constant speed during harvesting, including avoiding unnecessary stopping and starting is advised. Controlled traffic increases positioning accuracy and therefore minimises plot dimensions.

Plot location and homogeneity

As in any field experiment, differences between plot yields should reflect differences in treatments only. Crop and management history will be useful in determining appropriate blocks or areas of the field to be used for experimentation. Yield maps and remote sensed data could be useful for this task. In the absence of blocking, each plot should cover the same or similar fertility conditions. Buffer zones should specified in the planning of the experiment.

Plot dimensions

Assuming plots are laid out in strips, the farming practices employed will determine the minimum plot length and width. A minimum plot length of 100 m will ensure the middle 50 m will be representative of the treatment. The plot width will be a multiple of the harvester width. The minimum plot width of one harvester width will be easily achievable when controlled traffic is in place, where harvesting is carried out in the same direction, and the ground is level. When yield readings are affected by direction bias (eg. on a sloping field) and the harvester follows an up-and-back pattern, the minimum plot width will need to increase to two harvester widths so the observations can be averaged. A plot width of three harvester widths will be necessary when either treatment interference is expected between adjacent plots, or when GPS is the sole predictor of position and accuracy is unknown (ie. non-controlled traffic applications). The implications of using variable rate technology in strip-plot experiments has not been dealt with in the literature and needs further investigation.

Replication and Randomisation

Replication and randomisation are an important part of field experimentation. Replication within vs. across paddocks depends on the intended scope of the study. We are currently in the process of preparing a procedure for determining the amount of replication that will be appropriate for different situations and types of treatment allocation designs. With regards to variable rate technology, the trade-off between an increased number of replicates possible from small plots and any loss of data from plot boundaries needs investigating.

Treatment Allocation

Traditional experimental designs such as randomised block, latin squares and split-plot designs are all suitable along with adequate replication. Spatial designs such as nearest neighbour designs and geostatistical designs should be superior over non-spatial designs in theory, however, a small number of replicates may diminish any gain in precision.

Analysis of Experiments

Data Management

The most important recommendations for experimental design were proper calibration and designs to take account of direction and ensure a full, or at least fixed, harvester cutting width, because they can have substantial and uneven effects across the field and cannot be remedied once the data is collected. The resulting yield monitor data should then be processed using basic and advanced filtering techniques, with some information supplied by the user. Finally, aggregating data over sections of the field will create a unit of measurement which is more robust to the problems associated with an approximate delay time and correlation along the harvester path. We have compiled a step-by-step Excel based procedure for preparing yield monitor data for the analysis of experiments.

Statistical Models

Spatial models are more efficient than traditional experimental block analyses because of substantial spatial correlation. However, unlike small-plot experiments whose correlation structure captures the variation in fertility, high correlations in yield monitor data exist along the direction of travel due to mixing of grain within the harvester and trend. Mixed models including trend and correlated errors are very appropriate for yield monitor data particularly if the data has been aggregated over 25-100 m intervals. Competing models of adequate fit may exist, but as long as inferences for treatment differences are similar the particular model chosen should not matter.

Conclusion

We have developed a protocol for the conduct of on-farm experiments using yield-monitoring technology that minimises systematic errors so as to make scientifically substantiated comparisons between treatments. We aim to provide

alternative designs for a given level of input, so it will be possible to calculate whether a treatment difference with adequate precision can be detected.

Acknowledgments

The GRDC and Queensland University of Technology funded this work. We thank the farmers and QDPI for suppling yield data and expertise.