Investigate the optimal plot length in on-farm trials

Aolin Gong

Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, Urbana, Illinois

A paper from the Proceedings of the
14th International Conference on Precision Agriculture
June 24 – June 27, 2018
Montreal, Quebec, Canada

Abstract. *Agronomic researchers have recently begun running large-scale, on-farm field trials that employ new technologies that enable us to conduct hundreds of farm trials all over the world and, by extension, rigorous quantitative and data-centered analysis. The large-scale, on-farm trials follow traditional small-plot trials where the fields are divided into plots, and different treatments are randomly assigned to each plot. Over the past two years, researchers have been designing trials with plots approximately 90 m in length, following recommendations provided in the agricultural engineering literature in the 1990s. However, in this type of research, smaller plots are preferred for the benefit of more repetitions of the treatments on one field. This is important because advice given to producers is based on the experimental results, and such advice would be of greater value with more repetitions. With the minimum width of experimental plots being fixed due to the size of the farming equipment, the purpose of this research is to investigate optimal plot length. While shorter plot length in the experimental design results in additional plots, there is a tradeoff between the richness of the data from a single plot and the number of plots possible in a field. In order to weigh the tradeoff between the richness of the data from one repetition and the number of repetitions possible in one field with different plot length, Monte-Carlo simulations are conducted to compare the Economic Optimum Rates of fertilization (EOR) derived from the estimated yield on the experiment field with variable plot length and determine the optimal plot length.*

Keywords.
precision agricultural, trial design, plot design, Monte Carlo simulation.

1. Introduction

To improve the efficiency of agricultural practice, new technologies have been introduced to speed adoption of site - specific management practices (Schepers et al. 2008). Specifically, technologies based on computerized geographic information and global positioning systems (GPS) are renovating large-scale commercial agriculture throughout the world. This renovation is often referred to as "precision agriculture". Precision agriculture is a field management concept where information about a field is measured to increase the efficiency of management decisions. New technologies aid in the practice of precision agriculture could be called precision technologies. The realization of this concept gives new life to the old idea of site-specific management by reducing the cost of field experiments and variable rate input application. Producers are also required to process, collect, and analyze data to achieve optimized management of farm operations (Anselin et al. 1994).

However, McBratney et al. points out that as advancing precision agriculture is, it is not as developed as predicted 5 years ago (2005). One major obstacle in the development of precision agriculture is the inefficiency in information acquisition, which greatly depend on the design of field experiment. Precision agriculture is not reducing the cost of field experiments to its full potential if field experiment is not designed most efficiently. In both the field of academic and commercial agriculture, a lot of work have been done in yield monitoring (e.g. Mahasneh et al. 2001), variable rate application of inputs (e.g. Khosla et al. 2002), form of management zone (e.g. Fleming et al. 2000), and etc. However, there is a lack of study in the design of checker-board field experiment with precision technologies such as automatic guidance systems and automatic section control in the past literature.

Perterson made the argument in his book that the fundamental experimental unit in field experiments is the field plot. These plots constitute the piece of field experiment on which treatments are applied, and therefore, serve as the vehicle to evaluate the response of yield to the treatments (1994). In the past literature, the design of very-small plot trials and strips trials are commonly discussed. However, the design of the field plot in both the very-small plot trials and strips trials consider different factors than the checker-board trials. The very-small plot trials are carried manually, and soil variability was the main concern in its design (Peterson, 1994). In the strip trials, the design of treatments is the main focus, and the design of field plot always remain the same. Contrary the strip trials, the design This paper contributes to the literature on efficient checker-board field experiment design. Specifically, I investigate the optimal length of the field plot with precision technologies.

While shorter plot length in the experimental design results in additional plots, and therefore more repetitions, there is an increase in the experimental error in a single plot. The main objective of this paper is to quantify the statistical trade-off between the experimental error per plot and the number of repetitions in one field with variable plot length, and its implications to profit. Monte-Carlo simulations are conducted in this paper to compare the Economic Optimum Rates of fertilization (EOR) derived from the estimated yield on the experiment field with variable plot length and the optimal plot length is determined (Heady et al. 1955).

In particular, I seek to 1) investigate the optimal length of the plot in checker-board trials by comparing the different profits calculated with estimated yield as plot length varies and 2) examine how much better the randomized checker-board trials is compared to the strip-trials in acquiring valuable information.

2. Background

In agricultural production, variable-rate fertilization could reduce total amount of fertilizer applied and increase yields on one field compared to uniform-rate fertilization (Mallarino et al. 1999). Bongiovanni et al. showed an example that variable rate of nitrogen maintains farm profitability even when nitrogen is restricted to less than half of the recommended uniform rate. By using site-specific knowledge, precision technologies such as automatic guidance systems and automatic section control can target rates of fertilizer, seed and chemicals for soil and other conditions (2004). Rigorous quantitative and data-centered analysis are being conducted on field experiments to optimize fertilizer application. Currently, professional managers and consultants are playing the major role in helping producers design field experiments, organize data and make fertilizer application decisions. New management advices are given after statistical analyses confirm its advantages over existing planting plan.

Producers are often faced with highly variable and unpredictable environment and therefore no producer is the same. There is unlikely an overall management strategy adaptive for all producers. Management decisions for site-specific production are achieved through field experiments completed on-farm by producers (McBratney et al. 2005, Whelan et al. 2003), and designing the field experiments efficiently appears increasing important.

As the fundamental piece in field experiment, the design of field plot has been discussed since the 90s. In early stages of experimental trials when field experiments are carried manually, plots were 12.2 meters by 4.6 meters rectangles (Binfod et al. 1990, 1992). These very-small plot experiments phased out as the incorporation of technological advances brought automatization to agricultural production (Zhang et al. 2002). In the past two decades, strip-trials have become the mainstream in field experiments. Ostermeier from Iowa State university described some strip trials conducted at 42 fields in 18 different counties across Iowa, "the treatments were two N fertilizer rates applied alternatively across each strip to accommodate 3 to 17 replications on different site. The width of each treatment strip was the same within each site but varied across site according to the widths of the N applicator and the yield monitors. Strip length was determined by the length of the field, and was at least 500 m" (2007). Two obvious drawbacks of the strip-trials can be easily seen in the description above. The first major weakness is that a strip trial does not provide an adequate number of replications, constraining the reliability of the conclusions or estimates drawn from the data. The second limitation is that the field plot of strip-trials is too large. Field experiments involves N fertilizer rate too high or too low while the selection of the most appropriate rate of N fertilizer affects the profitability of corn production as well as the environment (Cerrato et al. 1990). The opportunity cost of the experiment turns out to be very expensive when inappropriate rate of N fertilizer is applied on some large field plots. One apparent improvement of the strip-trials is to make the field plot shorter, and this improvement is often referred to as checker-board trials.

While it seems apparent that the randomized checker-board designs are superior to strip trials, there exists widespread resistance and doubt about the effectiveness of randomized checker-board design. Part of this resistance is because, although there are obvious advantages in making the plot smaller, there is a tendency toward higher experimental error as plot length decreases (Roger, 1994). This experimental error comes from both the inaccuracy of the machinery. In this context, machinery refers to the yield monitor and the N fertilizer applicator. The applicator cannot adjust to the designed rate immediately when moving from one treatment plot to the next adjacent treatment plot; however, the lag of the applicator remains the same on the whole field and can be easily obtained by looking at the as-applied data. In contrast, the inaccuracy of yield monitor cannot be treated as simply as the change in yield is also related to soil characteristics and other inputs. Mahasneh tested with the Case IH 2166 AFS yield monitor and concluded that the yield monitor accuracy (error percentage per plot) increases as plot length is increased up to about 280 feet (2001). Following this conclusion, researchers have been designing the checker-board field experiment with field plot made 280 feet in length.

In this paper, I continue this work by incorporating both the accuracy model of yield monitor provided by Mahasneh and a spatial error term demonstrating the spatial interaction among yield in Monte Carlo simulations of field experiments with various plot length. As a special case, I will also consider the strip-trials design for comparison.

3. Simulation

3.1 Assumptions

Assumption 1: (Local Control) Elimination of heterogeneity. In practice, field plots that are similar are put together as a group, and by assigning all treatments into each group separately and independently, variation among groups can be removed from the experimental error (Gomez et al. 1984). In this paper, I assume the experimental field in within the same group, specifically, I assume the experimental field is homogenous in yield response to N fertilizer.

Assumption 2: (Error Control) For simplicity, I assume the application of N Fertilizer is accurate since its lag can be observed.

Assumption 3: (Field Control) I assume the field experiment is always designed on the whole field.

3.2 The Experiment Field

In this paper, I simulate an experiment field that is 151.13-acre in size, and the data generation process mimicked on the experiment field is established from a 2017 checker-board trial field growing corn in Central Illinois with completely randomized nitrogen rates. The treatment on the experiment field is seven N fertilizer rates applied randomly to each plot, namely 110, 130, 150, 180, 200, 220, 240 (pounds/acre). The field plots are made the same size and 60 feet wide determined by the width of the N applicator and the yield monitor. Based on Assumption 2, the accuracy of N applicator is not considered in the simulation, and the yield monitor used on the experiment field is the Case IH 2166 AFS yield monitor tested by Mahasneh et al. 2001.

In the simulation, I consider 11 different length of a field plot, with $70 + 20(n)(n = 0, \dots 10)$ feet being the length of the plot. (Figure I shows an example of the experiment field with 270 feet long and 60 feet wide field plots, and Figure II shows an example of the experiment design map.) Actual yield level is generated at a spatially finer scale than the field plot dimension. Specifically, the experiment field is divided into 30 feet by 30 feet grids, with actual yield spatially correlated among the grids. (Figure III shows the experiment field divided into grids.) Actual yield generated at the grid level will be translated to plot level in regression.

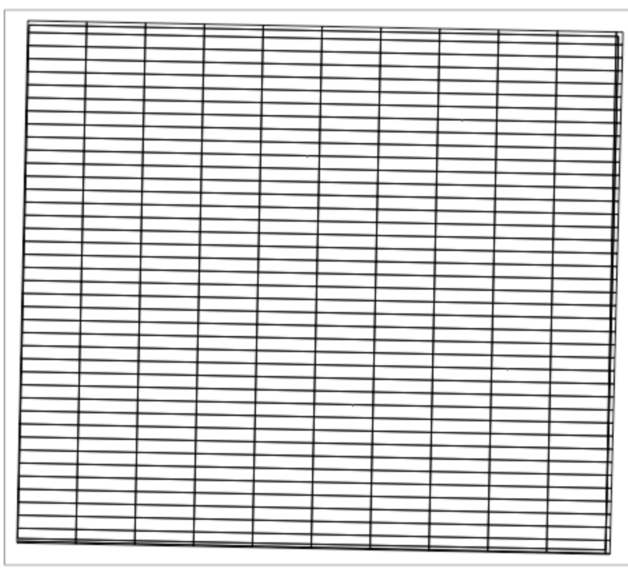

Fig. I An experiment field with 82.3 meters long and 18.3 meters wide plots

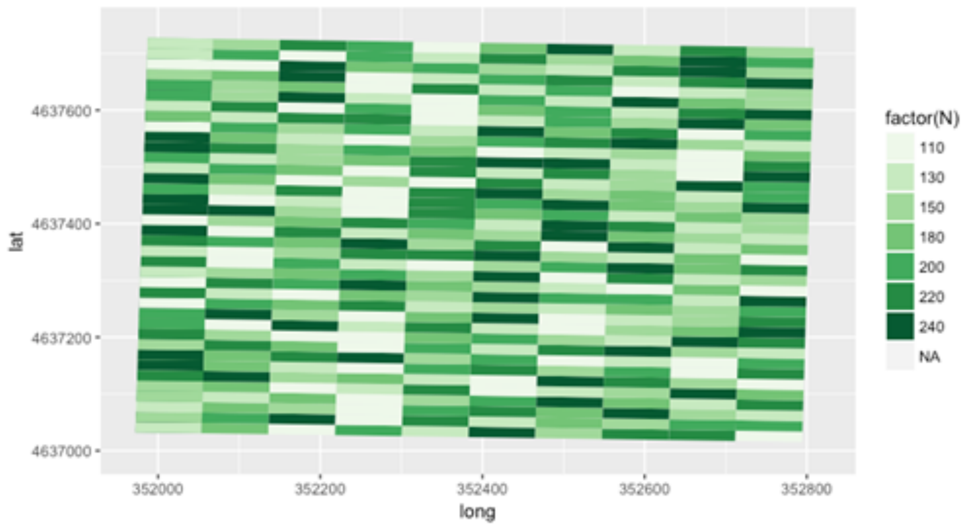

Fig. II The treatment design map with 270 feet by 60 feet field plots

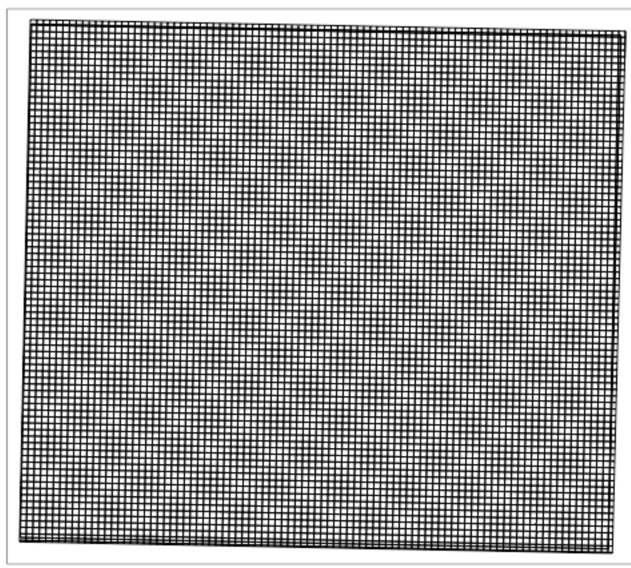

Fig. III The experiment field with 30 feet by 30 feet grids

3.3 Grid-level data generation

Agricultural production process is characterized by input-output relations (Nijland et al. 2008), and these relations are often concluded in functional forms. On the experimental field, I mainly consider the yield response to N fertilizer since it is the most commonly used fertilizer by corn farms in Illinois. There exists multiple functional form describing yield response to N fertilizer, including quadratic; square root; Mitscherlich-Baule; linear von Liebig; nonlinear von Liebig, and etc. (Llewelyn and Featherstone 1997; Frank et. Al 1990) compared different yield response models and showed that the cost of using the Mitscherlich–Baule is lower. Thus, the yield production function of Mitscherlich (also Mitscherlich–Baule) is chosen to generate the deterministic part of actual yield at grid level:

$$f(N) = ymax \times (1 - \exp\,(a + b * N)) \quad (1)$$

where $ymax$ is the maximum yield level attainable on the field, N is the amount of N fertilizer applied, a and b controls the curvature of the production function. As mentioned above, the generation process is established from a 2017 corn field. The value of a and b reported from the field are respectively -0.473343, -0.008973, while the maximum yield level on that field equals 252.40 (bu/acre). (Figure IV shows the yield plateau with the chosen parameters.) The deterministic part of actual yield is generated on the grid level using the Mitscherlich–Baule model with parameters concluded from the 2017 corn field. (Figure V shows the map of the deterministic part of actual yield.)

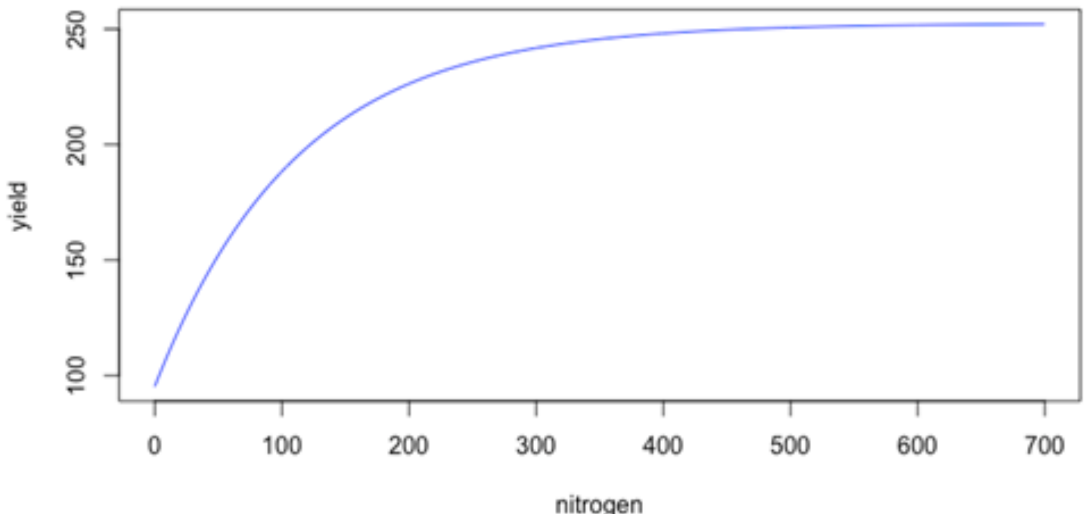

Fig. IV The yield plateau

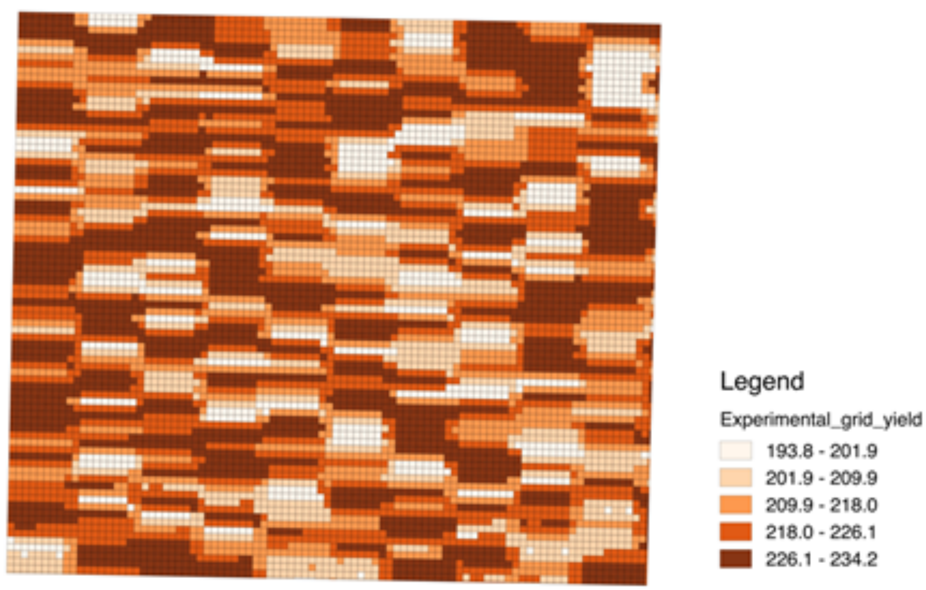

Fig. V The deterministic part of actual yield

The variance of residuals of the Mitscherlich–Baule model is calculated and tested for spatial correlation. To estimate the spatial interaction between yield, the spatial error is introduced, and the spatial lag is also calculated for comparison. Spatial lag, by definition, is the weighted average of neighboring values defined by spatial weights. The spatial lag model I used in this project is:

$$y = \rho W y + X\beta + \varepsilon \tag{2}$$

and I used a spatial simultaneous autoregressive error model to estimate the spatial error:

$$y = X\beta + u \tag{3}$$

$$u = \lambda W u + e \quad (4)$$

Spatial error appears to be a better estimate of the spatial interaction as it can demonstrate the error term due to omitted random factors in the Mitscherlich–Baule model imposed above. Table I summarizes the results from both the spatial error model and the spatial lag model, and the value of ρ and λ are very high, indicating that a spatial correlation term is needed. In this project, I use λ to generate the spatial error on grid level as the spatial correlation term, and add it to the deterministic part of yield as actual yield. Actual yield map at the grid level is then translated to plot level for regression. Figure VI shows the actual yield on the plot level (the incomplete plots were taken out).

Table I Summary of the spatial error model and the spatial lag model

	Spatial Error Model	Spatial Lag Model
Intercept	-0.49837074	1.10553956
	0.17986836(SE)	0.12341101
nitrogen	-0.00862938	-0.00754783
	0.00073617(SE)	0.00066638
λ	0.89806	-
ρ	-	0.88864
Wald test	778.19	833.4
AIC	921.92	926.42

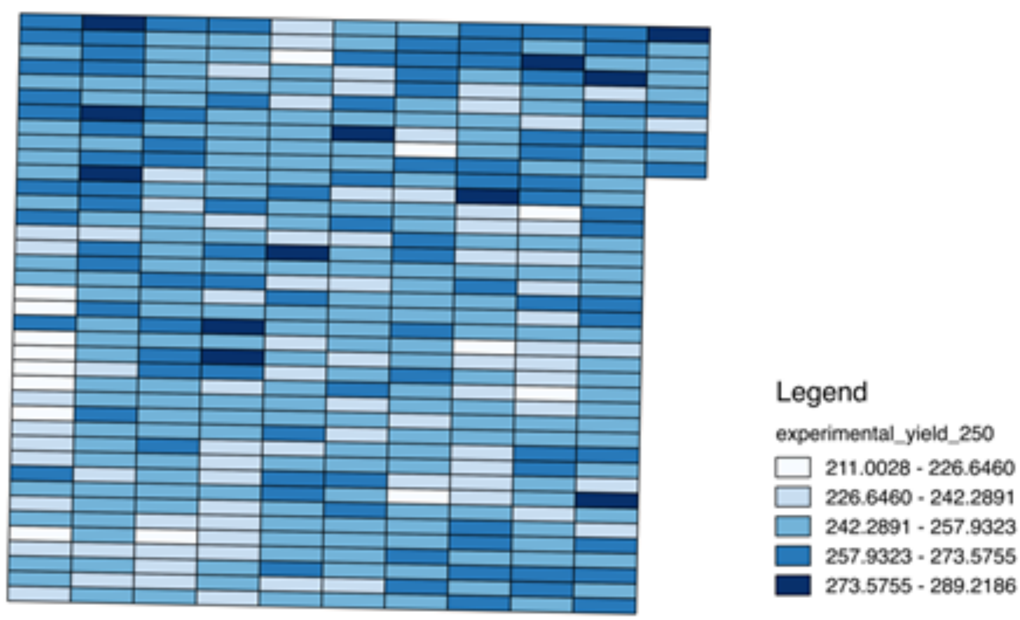

Fig. VI The actual yield at plot level

3.4 Estimation of measurement errors

Based on Assumption 2, yield monitor is the only inaccurate machinery in this project. Let y and e denote the actual yield at plot-level and the measurement error introduced by the yield monitor. The yield level reported by the yield monitor is therefore $y + e$, denoted by $\tilde{y}$. The correlation coefficient of ρ between y and $\tilde{y}$ follows:

$$\rho = \frac{\sigma_{y,\tilde{y}}}{\sqrt{\sigma_y^2 \times \sigma_{\tilde{y}}^2}} = \frac{\sigma_{y,y+e}}{\sqrt{\sigma_y^2 \times \sigma_{y+e}^2}} \quad (5)$$

where $\sigma_{y,\tilde{y}}$ denotes the covariance between y and $\tilde{y}$, and σ_y^2 denotes the variance of y. The measurement error term generated by the yield monitor is assumed to only depend on the plot length and is statistically independent of the actual yield. Under this assumption, covariance between y and e is zero. Thus,

$$\rho = \frac{\sigma_y^2}{\sqrt{\sigma_y^2 \times (\sigma_y^2 + \sigma_e^2)}} \quad (6)$$

Rearranging the equation above, I obtain:

$$\sigma_e^2 = \sigma_y^2 \left(\frac{1-\rho^2}{\rho^2}\right) \quad (7)$$

The variance of actual yield is calculated on plot level, and the value of the correlation coefficient ρ is estimated by Mahasneh et al. dependent on the plot length (2001). The graph of the relation between ρ and the plot length is shown in Figure VII and is summarized in Table II. The measurement error e is assumed to follow a normal distribution $N(0, \sigma_e^2)$ so that e can be added to the actual yield y to obtain $\tilde{y}$.

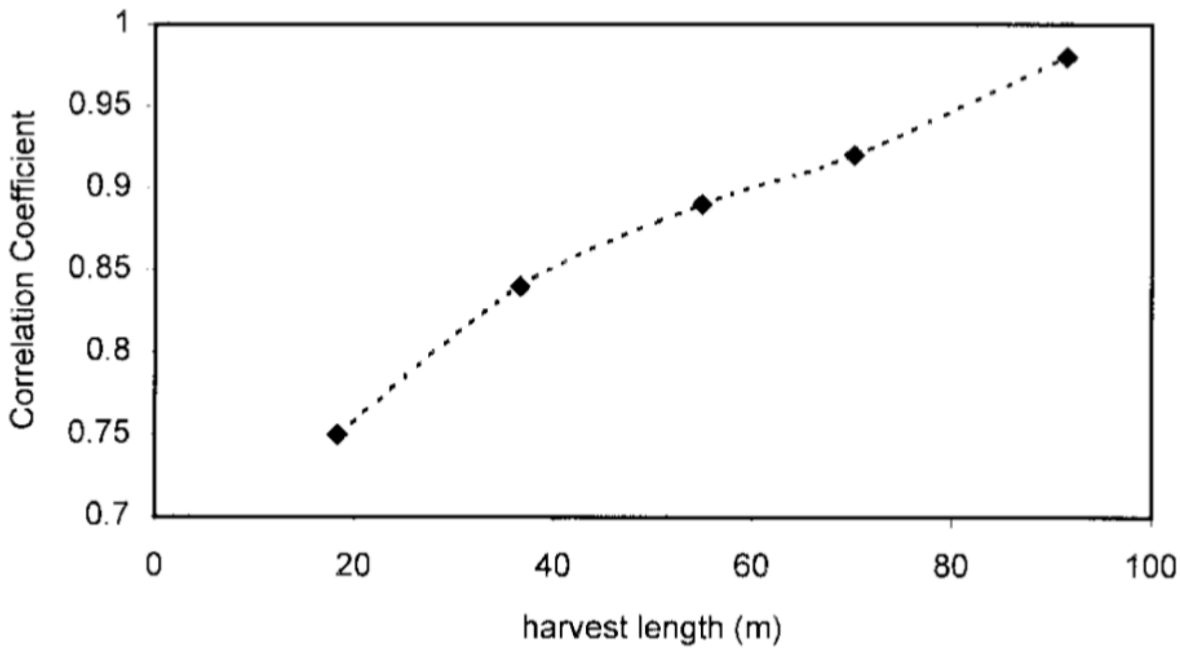

Fig. VII The relation of the correlation coefficient and the plot length by Mahasneh et al. (2001)

Table II The correlation coefficient of different plot length

Plot length (in meter)	ρ
21.336	0.762
27.432	0.803
33.528	0.823
39.624	0.844
45.720	0.861
51.816	0.870
57.912	0.881
64.008	0.915
70.104	0.926
76.200	0.930
82.296	0.944

Note:ρ is the correlation coefficient between y and $\tilde{y}$

3.5 Production function estimation

In the estimation, the functional form with respect to a subset of repressors is unknown. In order to capture the curvature of the production function and the existence of the yield plateau, I adopt a generalized additive model that contains a parametric form for some component of the data with weak nonparametric restrictions on the remainder of the model. Let ϕ_k denote the kth cubic spline, then the parametric form of the structural function is $\sum_{k=1}^{K} \beta_k \phi_k(N_i)$, and β_k are parameters to be estimated. Let ε denote the remainder of the model, following the conditional mean restriction on ε: $E(\varepsilon|N) = 0$. Thus, the estimating equation is:

$$y_i = \sum_{k=1}^{K} \beta_k \phi_k(N_i) + \varepsilon_i \tag{8}$$

3.6 Comparison of Economic optimum rates (EOR)

Kyveryga et al. (2007) imposed a definition of Economic optimum rates (EOR) as “Economic optimum rates explicitly denote rates of N fertilization expected to maximize net returns from fertilizer applied per unit of land area”. In this project, I seek to compare the Economic optimum rates derived from estimated yield on the experiment field with variable plot length by substituting the Economic optimum rates derived from experiment field designed differently into a profit function and compare the estimated profits:

$$\pi = P_c \times f(N) - P_N \times N \tag{9}$$

where P_c denoted the price of corn (\$/pounds), and P_N denotes the price of N fertilizer (\$/pounds). By solving the maximization of Equation 2, I can obtain the Economic optimum rate of the experiment field. Let $P_c = 3.5$, and $P_N = 0.32$, the Economic optimum rate is:

$$N^o = \frac{1}{-0.008973} \log\left(\frac{0.32}{3.5 \times 252.4 \times 0.008973}\right) + \frac{0.473343}{0.008973} = 106.773$$

The maximum profit is therefore $\pi^o = 3.5 \times 192 - 0.32 \times 106.773 = \$\ 637.8326$ per acre. The estimated profits on the experiment field with variable plot length are compared to π^o in order to decide the optimal plot length. The optimal plot length generates the smallest difference between the estimated profit and π^o.

3.7 Monte-Carlo simulation

For a given length of plots, $70 + 20(n)$, one iteration of an experiment follows the following process:

i. One of the treatment N fertilizer level (110, 130, 150, 180, 200, 220, 240) is randomly assigned to each plot;

ii. Translate the treatment at the plot level to grid level. If a grid is completely contained within a plot, the same nitrogen rate assigned to the plot is also assigned to the grid. If a grid spans across two plots, then the area weighted average N rate for the corresponding plots will be assigned to the grid;

iii. For each grid, generates the actual yield by adding the deterministic part produced by the production function and the spatial error;

iv. Translate the actual yield at the grid level to plot level;

v. Generates the measurement error e from the normal distribution $N(0, \sigma_e^2)$, and obtain the estimated yield $\tilde{y}$ by adding the actual yield with the measurement error e;

vi. Estimate the production function y_i;

vii. Estimate the Economic optimum rate (EOR) N^*based on the estimated function;

viii. Calculate the maximum profit using the EOR, $\hat{\pi} = P_c \times f(N^*) - P_N \times N^*$, then calculate the difference between π^o and the maximum profit estimated with the estimated yield, $\Delta_{70+20(n)} = \pi^o - \hat{\pi}$.

This process is repeated for 1000 times for each plot length, and $\Delta_{70+20(n)}$ are compared among different n..

References

Ancev, T., Whelan, B. M., & McBratney, A. B. (2004). On the economics of precision agriculture: technical, informational and environmental aspects. 2004 Annual Conference of the Australian Agricultural and Resource Economics Society, Melbourne, Victoria, Australia, February 11–13, 2004.

Anselin, L., Bongiovanni, R., & Lowenberg-DeBoer, J. (2004). A spatial econometric approach to the economics of site-specific nitrogen management in corn production. American Journal of Agricultural Economics, 86(3), 675-687.

Bongiovanni, R., & Lowenberg-DeBoer, J. (2004). Precision agriculture and sustainability. Precision Agriculture, 5(4), 359-387.

Cerrato, M. E., & Blackmer, A. M. (1990). Comparison of models for describing; corn yield response to nitrogen fertilizer. Agronomy Journal, 82(1).

Fleming, K. L., & Westfall, D. G. (2000). Evaluating farmer defined management zone maps for variable rate fertilizer application. Precision Agriculture, 2(2), 201–215.

Frank, M. D., Beattie, B. R., & Embleton, M. E. (1990). A Comparison of Alternative Crop Response Models. American Journal of Agricultural Economics. 72(3), 597-603.

Khosla, R., Fleming, K., Delgado, J. A., Shaver, T. M., & Westfall, D. G. (2002). Use of site-specific management zones to improve nitrogen management for precision agriculture. Journal of Soil and Water Conservation, 57(6):513-518.

Kyveryga, P. M., Blackmer, A. M., & Morris, T. F. (2007). Alternative Benchmarks for Economically Optimal Rates of Nitrogen Fertilization for Corn. Agronomy journal. 99 (4), 1057-1065.

Llewelyn, R. V., & Featherstone, A, M. (1997). Comparison of crop production functions using simulated data for irrigated corn in western Kansas. Agricultural Systems. 54(4), 521-538.

Mahasneh, M., & Colvin, T. S. (2000). Verification of yield monitor performance for on-the-go measurement of yield with an in-board electronic scale. Transactions of the ASAE. 43(4), 801-807.

Mallarino, A. P., Wittry, D. J., & Hinz, P. N. (1999). Variable-Rate Phosphorus Fertilization: On-Farm Research Methods and Evaluation for Corn and Soybean. Precision Agriculture. Madison, WI: American Society of Agronomy, Crop Science Society of America, Soil Science Society of America. 687–696.

McBratney, A., Whelan, B., & Ancev, T., & Bouma, J. (2005). Future directions of precision agriculture. Precision Agriculture, 6(1), 7-23.

Ostermeier, G. C. G. (2007). Using on-farm strip-trials to improve nitrogen management for corn. Retrospective Theses and Dissertations. 15952.

Pertersen, R. G. (1994). Agricultural field experiments: design and analysis. New York: Marcel Dekker, Inc.

Schepers, J. S., & Francis, D. D. (2008). Precision agriculture—What is in our future. Communications in Soil Science and Plant Analysis. 1463-1469.

Whelan, B. M., McBratney, A. B., & Stein, A. (2003). On-farm field experiments for precision agriculture. Precision Agriculture, Proceedings of the 4th European Conference on Precision Agriculture Wageningen Academic Publishers. The Netherlands731-737.