Optimal design for on-farm strip trials — systematic or randomised?

Zhanglong Cao[1,*], Jordan Brown[1], Mark Gibberd[1], Julia Easton[1], and Suman Rakshit[1,2]

[1]Curtin Biometry and Agricultural Data Analytics, Centre for Crop and Disease Management, Curtin University, Perth, Australia

[2]School of Electrical Engineering, Computing and Mathematical Sciences, Curtin University, Perth, Australia

[*]Corresponding author: Zhanglong Cao, zhanglong.cao@curtin.edu.au

Abstract

CONTEXT OR PROBLEM: Randomised designs are often preferred over systematic designs by agronomists and biometricians. For on-farm trials, however, the choice may depend on the objective of the experiments. If the purpose is to create a prescription map of a continuous input for each plot of a grid covering a large strip trial, a systematic design may be a better choice, but it attracts less discussion and attention.

OBJECTIVE OR RESEARCH QUESTION: This study aims to evaluate the performance of systematic designs with geographically weighted regression (GWR) models in addressing spatial variation and estimating continuous treatment effects in large strip trials through numeric simulations.

METHODS: A hierarchical model with spatially correlated random parameters is utilised to generate simulated data for various scenarios of large strip on-farm trials. The study employs GWR models to analyse the simulated data for two assumptions: a linear response and a quadratic response of yield to the treatment effects.

RESULTS: With the assumption of a quadratic response, a systematic design is superior to a randomised design in terms of achieving lower mean squared errors (MSE) with GWR. With the assumption of a linear response, the difference of MSE between a systematic design and a randomised design is not significant, regardless of the presence of spatial variation.

CONCLUSIONS: The findings highlight the superiority of systematic designs in producing smooth spatial maps of optimal input levels for quadratic response models in large strip trials, even when impacted by significant spatial variation. Additionally, we recommend selecting fixed bandwidths in GWR analysis based on the plot configurations used in experimental designs. For a large strip trial, to produce estimates of spatially-varying treatment effects across strips, a systemic design should be used as it allows us to obtain better estimates than those obtained from a randomised design in post-experiment statistical modelling.

IMPLICATIONS OR SIGNIFICANCE: The findings offer practical recommendations for designing large strip trials. By drawing attention to the experiment's main inferential purpose, this research contributes valuable insights for improving the efficacy and planning of large strip trials.

Keywords: yield map, optimal treatment, spatially varying coefficients, geographically weighted regression, precision agriculture.

1 Introduction

The principle of randomisation was first expounded in 1925 by Fisher (1934), who analysed a few systematically arranged experiments and pointed out that randomisation can provide valid tests of significance subject to appropriate restrictions, such as experimental units arranged in blocks or in rows and columns of a Latin square (Verdooren 2020). Traditionally, small-plot trials for agriculture are designed to obtain unbiased estimates of treatment effects using the completely randomised design, where treatments are randomly allocated into plots. More complex designs, such as the randomised complete block design, the split-plot design, the strip-plot design and the Latin square design, are also widely used in agricultural experiments to improve the precision of treatment effect estimates (Petersen 1994). With the primary aim of obtaining unbiased estimates of global treatment effects, randomised designs , which use different layouts of treatments in each replicate, are routinely used for on-farm strip trials, whereas systematic designs, which use the same layout of treatments in all replicates, are rarely used.

On-farm experiment (OFE) enables farmers the flexibility to implement large-scale experiments in order to test management practices on their farms (Evans et al. 2020). The main goal of OFE is to help farmers better understand uncertainties around farm-related decisions and leverage their existing strengths in managing translational and structural uncertainties in decision-making (Cook et al. 2013). In situations where the goal is to compare yield responses between management classes or to select best-performing crop varieties as new market varieties, a randomised design may be superior to a systematic design (Pringle et al. 2004; Selle et al. 2019).

While randomisation is often considered a crucial prerequisite for obtaining valid statistical inferences (Piepho et al. 2013), this is not always the case when the goal of OFE shifts from the conventional analysis. In the application of precision agriculture using variable rate applicators, a prescription map of the experiment is required to optimally apply varying treatments across a field (Pringle et al. 2004). Therefore, in this scenario, the goal of OFE becomes obtaining a smooth spatial map showing the optimal level of a controllable input, such as nitrogen rates, across a grid made of rows and columns covering the whole field. An important point to note here is that only a single treatment level can be directly observed at any one point on the grid, and the response for other levels in the same grid must be interpolated. If a randomised design is conducted, the interpolation distances to locations with treatment levels of interest will vary throughout the field. Such heterogeneous distances increase the uncertainty in the analysis and reduce the efficiency of local prediction. As a result, a systematic design is preferable to a randomised design in this scenario. Unfortunately, this perspective has often been overlooked by researchers, leading to the widespread use of randomised designs.

Analysing a systematic design for the creation of an optimal treatment map is a statistically challenging task. The true responses at each point on the grid corresponding to all the treatment rates are unknown, and the treatment producing the optimum response may vary continuously across the field. Cao et al. (2022) implemented a Bayesian approach with spatially correlated random parameters for analysing large systematic strip trials. These authors considered a quadratic response model with both global and local (spatially-varying) components. However, Bayesian analysis can be computationally expensive and would require at least preliminary knowledge of Bayesian inference to interpret the results, which can be extremely demanding for farmers and agronomists. Alternatively, Rakshit et al. (2020) adopted a local regression approach, called geographically weighted regression (GWR), to obtain spatially-

varying estimates of treatment effects for OFE. Additionally, Evans et al. (2020) concluded through simulation studies that GWR is capable of accurately separating variation in yield response due to treatment from the variation that is not due to the applied treatment. The limitations in their study are the use of a randomised design and the assumption of a linear response model. To compare between the systematic and randomised treatment allocation in the chessboard design, Alesso et al. (2021) simulated corn yield response for four nitrogen levels and estimated the regression coefficients using GWR. They concluded that systematic designs achieved the best results in most cases. However, the use of chessboard design often presents several challenges, particularly during harvesting. Harvesters can produce erroneous data due to the abrupt treatment changes between plots (Pringle et al. 2004). Additionally, the quadratic or plateau feature in a response model was not considered in their simulation study.

Piepho and Edmondson (2018) presented an example where a linear model turns out to be inadequate for analysing sugar beet data (Petersen 1994). Glynn (2007) showed that many curves exist beyond a linear trend for nutrient-response relationships. The response curve often depends on the availability of other macro and micronutrients in the soil (Marschner 2011), which means that a linear relationship is unlikely to be consistent across a large trial. For this reason, it is important to consider models with terms of order higher than unity. For example, a quadratic model can often found to be suitable for modelling nutrient-response relationships (Piepho and Edmondson 2018; Liben et al. 2019).

In this study, we generate simulated data for several scenarios, where each scenario is constructed by choosing one component at a time from the following four categories: (i) randomised and systematic designs; (ii) linear and quadratic responses; (iii) model coefficients with low and high correlations; and (iv) spatial variance-covariance matrix among grids given by identity (no spatial trend), AR1$\otimes$AR1, and Matérn forms. We subsequently evaluate the efficacy of GWR in accurately estimating the spatially varying treatment effects across these scenarios.

The GWR model in this paper is implemented with the `R`-package `GWmodel` (Lu et al. 2014; Gollini et al. 2015).

2 Methods

2.1 Hierarchical model for generating simulated data

In a conventional agricultural study, a field experiment can be considered as a rectangular matrix, representing a regular grid with r rows and c columns, where the total number of plots in the experiment is $n = r \times c$. Let $s_i \in \mathcal{R}^2, i = 1, \ldots, n$, denote the Cartesian coordinate of the i-th plot centroid, located on a regular grid (Zimmerman and Harville 1991). Let $y(s_i)$, $i = 1, \ldots, n$, denote the value of the dependent variable recorded at the i-th plot.

Let $\boldsymbol{Y}$ denote the vector of the plot data ordered as rows nested within columns. The basic model can be written using the matrix notation as follows:

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{b} + \boldsymbol{Z}\boldsymbol{u} + \boldsymbol{e}, \tag{1}$$

where $\boldsymbol{b}$ and $\boldsymbol{u}$ are vectors of fixed and random effects, respectively; $\boldsymbol{X}$ and $\boldsymbol{Z}$ are the associated design matrices; and $\boldsymbol{e}$ is the error vector. We assume that $\boldsymbol{u}$ and $\boldsymbol{e}$ are distributed independently of each other and that their joint

distribution is

$$\begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{e} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Sigma_u & 0 \\ 0 & \Sigma_e \end{bmatrix} \right). \tag{2}$$

Using the notation introduced above and in Cao et al. (2022), the simulation model is given by

$$\begin{aligned} y(s_i) \mid \boldsymbol{u}_i, \theta_u, \sigma_e &\sim \mathcal{N}\left(\sum_{m=1}^{l} b_m x_m(s_i) + \sum_{j=1}^{k} u_j(s_i) z_j(s_i),\ e(s_i) \right), \\ \boldsymbol{u}_i \mid \theta_u &\sim \mathcal{N}(0, V_u(\theta_u)), \\ e(s_i) \mid \sigma_e &\sim \mathcal{N}(0, \sigma_e^2), \end{aligned} \tag{3}$$

where $x_1, \ldots, x_l$ are l fixed effects; $z_1, \ldots, z_k$ are k random effects; b_m and $u_j(s)$ are the coefficients for the fixed and random terms, respectively; $\boldsymbol{u}_i$ is a vector of all random effects at the i-th plot, $i = 1, \ldots, n$; θ_u is a set of parameters of the covariance matrix V_u; and σ_e is a positive latent variable.

In model (3), the structure of the covariance matrix $V_u(\theta_u)$ of $\boldsymbol{u}_i$ can be either diagonal, which implies the random terms at grid i are independent, or in general form, which means a correlation exists. McElreath (2015) suggested that the covariance matrix V_u can be decomposed into $B(\sigma_u)R_uB(\sigma_u)$, where $B(\sigma_u)$ denotes the diagonal matrix with diagonal elements σ_{u_j}, $j = 1, \ldots, k$, and R_u denotes the matrix with correlation coefficients. For the matrix R_u, we specify the Lewandowski-Kurowicka-Joe (LKJ) distribution (Lewandowski et al. 2009), which is given by

$$R_u \sim \text{LKJcorr}(\epsilon), \tag{4}$$

where LKJcorr(ϵ) is a positive definite correlation matrix sampled from the LKJ distribution controlled by a positive parameter ϵ. As ϵ increases, a high correlation among parameters becomes less likely.

Furthermore, by incorporating a spatial correlation structure V_s, the complete form of the covariance matrix of $\boldsymbol{u}$ is presented as

$$\Sigma_u = V_s \otimes V_u. \tag{5}$$

In fact, V_s is the covariance matrix of all grids on the field. For example, if $V_s = I_{n\times n}$ (an identity matrix), the random terms at a grid point are independently distributed from those at other grid points, even though the terms at that grid point are correlated amongst each other. However, the correlation among grids is ubiquitous. Hence, we introduce a simple spatial covariance matrix such as

$$V_s = \text{AR1}(\rho_c) \otimes \text{AR1}(\rho_r), \tag{6}$$

where AR1 $\otimes$ AR1 is the separable first-order auto-regressive model in the column and row directions, controlled by the correlation parameters ρ_c and ρ_r, respectively (Butler et al. 2017).

On the other hand, the Matérn class covariance is given by

$$V_s(d) = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\sqrt{2\nu}\frac{d}{\gamma} \right)^{\nu} K_\nu \left(\sqrt{2\nu}\frac{d}{\gamma} \right), \tag{7}$$

where d is the space lag or distance; γ is a non-negative scaling parameter; $\nu > 0$ is a smoothness parameter; σ^2 is the variance parameter; Γ is the Gamma function; and K_ν is the modified Bessel function of the second kind. The Matérn covariance is commonly used in the analysis of geostatistical data (Cressie and Huang 1999). Moreover, it has also been used in capturing spatial variation in OFE (Selle et al. 2019). If $\nu = \gamma + 1/2$, then the Matérn covariance can be expressed as a product of an exponential and a polynomial of order γ (Pandit and Infield 2019; Abramowitz 1974), which simplifies the model and the computation process. The Matérn models with the values 3/2 and 5/2 for ν are used in most applications.

In each iteration of the simulation process, we used the above formulae and pre-defined parameter values to generate $2 \times n$ coefficients for the linear response and $3 \times n$ coefficients for the quadratic response. Then these coefficients are applied to simulate yield response across strips for both randomised and systematic design layouts.

2.2 Fitting geographically weighted regression to simulated data

Geographically weighted regression (GWR) is a local regression approach and is adapted recently to obtain spatially-varying estimates of treatment effects for OFE (Rakshit et al. 2020). It is a locally weighted regression method that operates by assigning a weight to each observation depending on its distance from the query grid on the field (Páez et al. 2002).

The underlying template model for GWR is given by

$$y(s_i) = \beta_0 + \sum_{j=1}^{k} \beta_j \, x_j(s_i) + \varepsilon_i, \tag{8}$$

where $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_k)^\top$ are model parameters corresponding to the k treatment levels and $\varepsilon_i \sim \mathcal{N}(0, \tau^2)$, $i = 1, \ldots, n$, are independent and identically distributed error terms at n grid points.

For a query location s, the local log-likelihood is given by

$$\log L(s; \boldsymbol{\beta}) = -\frac{1}{2\tau^2} \sum_{i=1}^{n} K(s, s_i) \left(y(s_i) - \beta_0 - \sum_{j=1}^{k} \beta_j x_j(s_i) \right)^2 \tag{9}$$

where $K(\cdot, \cdot)$ is a given kernel function, such as Gaussian, exponential, bi-square or tri-cube (Gollini et al. 2015).

The local-likelihood estimator, obtained by maximising (9), of the regression coefficients $\boldsymbol{\beta}$ at the query location s is given by

$$\hat{\boldsymbol{\beta}}(s) = \left(X^\top W(s) X \right)^{-1} X^\top W(s) Y, \tag{10}$$

where $W(s)$ is an $n \times n$ diagonal matrix of weights with i-th diagonal entry $K(s, s_i)$.

To obtain GWR estimates of model parameters in the simulation study, we have used a Gaussian kernel. In fact, the kernel function is not the crucial factor in the GWR analysis. In contrast, the bandwidth has a higher influence on the estimates.

The optimal bandwidth for a GWR model is usually selected by the lowest AICc, which is given by

$$\text{AICc} = 2n\log(\tau^2) + n\log(2\pi) + \frac{n^2 + n\text{tr}(S)}{n - 2 - \text{tr}(S)}, \tag{11}$$

where S is the matrix with the i-th row given by $X_i \left(X^\top W(s_i) X\right)^{-1} X^\top W(s_i)$ (Evans et al. 2020), and $\text{tr}(\cdot)$ is the trace of a square matrix returning the sum of the elements on the main diagonal. Alternatively, as suggested by Rakshit et al. (2020), it can be chosen according to the experimental design such that the local regressions capture data covering the full range of treatments.

2.3 Performance evaluation

To compare the performances of randomised and systematic designs in terms of accurate estimation of the model coefficients using GWR, we use the mean squared errors (MSE) corresponding to all coefficient estimates. The MSE for a coefficient was computed by first taking the differences between the true coefficient, specified in model (3), and the spatially varying estimates of that coefficient produced by GWR, and then averaging these squared differences across all the grid points, shown in equation (12). The lower the MSE, the better the design's performance.

The MSE corresponding to the estimation of spatially varying β_j is given by

$$\text{MSE}_j = \frac{1}{n}\sum_{i=1}^{n}\left((b_j + u_{ji}) - \hat{\beta}_{ji}\right)^2, \tag{12}$$

where $j = 0, 1$ for a linear response and $j = 0, 1, 2$ for a quadratic response.

3 Simulation study

The simulation study is performed using realistic values for the model parameters, which are selected based on the analysis results of a real-life data recorded from a corn field in Las Rosas, Argentina. This data set was originally provided by Anselin et al. (2004) and can be accessed via the `R`-package `agridat` (White and Evert 2008). In 2001, a systematic design was used, incorporating six rates of nitrogen: 0, 39, 50.6, 75.4, 99.8, and 124.6 kg/ha, in three replicates. Each replicate consists of 93 rows and 6 columns after data preprocessing; see Rakshit et al. (2020) for further details about the preprocessing steps. The unit of yield is quintals per hectare.

Studies by Rakshit et al. (2020) and Cao et al. (2022) suggest that the yield produced by the maximum nitrogen rate of 124.6 kg/ha may be improved by using a higher rate. Thus, we have made some adjustments while selecting the nitrogen rates for our simulation study. We use five evenly-spaced nitrogen rates: 0, 35, 70, 105, and 140 kg/ha. Additionally, we increase the number of replicates from three to four. Consequently, the final layout of the trial used in the simulation consists of 93 rows and 20 columns. Examples of a randomised design, which uses different orders of treatment in each replicate, and a systematic design, which uses the same order of treatment in all replicates, for this layout are presented in Figure 1.

We investigate all possible combinations of the following factors: (i) types of design with two levels, namely, randomised and systematic; (ii) response relationship with two levels, namely, linear and quadratic; (iii) correlation coefficients corresponding to the random effects within each plot with two levels, namely, low and high; (iv) spatial

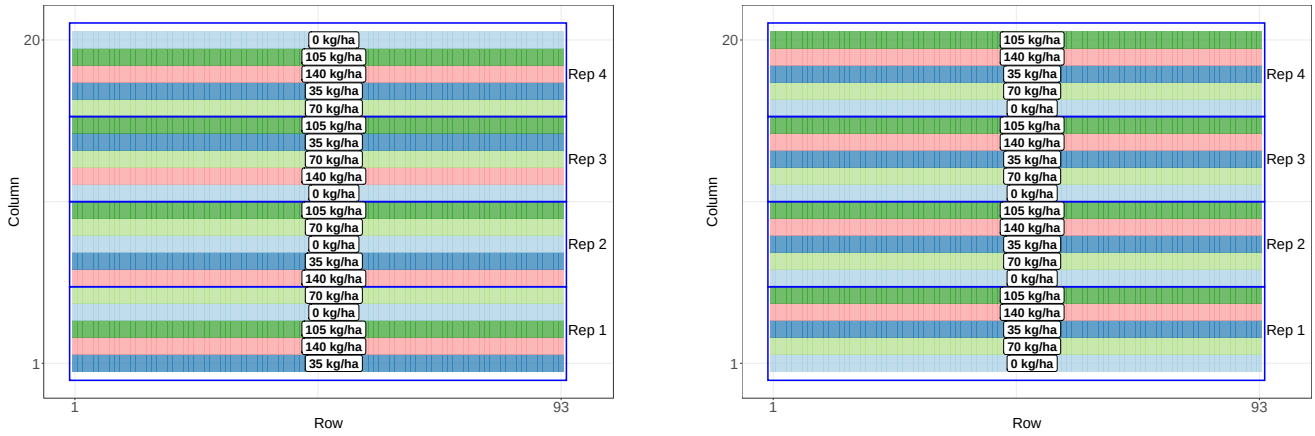

Figure 1: The nitrogen treatments with five levels (0, 35, 70, 105 and 140 kg/ha) randomly (*left*) and systematically (*right*) allocated into strips in each replicate block.

variation between grid points with three levels, namely, identity (no spatial trend), AR1$\otimes$AR1, and Matérn form. This results in 24 unique combinations. For each combination, we simulated yield data and fitted the GWR model for three bandwidth values of 5, 9 and the optimum value selected by AICc. For a systematic design in our simulation study, all the treatment levels (five nitrogen levels) can be covered by the bandwidth of 5, making it adequate for inference based on a quadratic response model. On the other hand, the bandwidth of 9 may be necessary to cover all possible treatment levels in a randomised design. In particular, if identical treatments are positioned at the far left edge of the first replicate block and at the far right edge of the second replicate block, a bandwidth smaller than 9 would result in the GWR model capturing only the treatments between these boundaries, thereby missing the treatment levels at the extremes.

To specify the linear relationship in the simulation study, we consider the values of 65 and 0.05 for the global intercept b_0 and slope b_1 coefficients, respectively, in model (3). The variances of the coefficients $\boldsymbol{u}_i$ are set to 5 for σ_{u_0} and 0.01 for σ_{u_1}. These parameters are chosen according to the estimates reported by Cao et al. (2022). For the AR1 $\otimes$ AR1 covariance matrix in (6), the two correlation parameters ρ_c and ρ_r are set to 0.15 and 0.50, respectively. We assume a higher correlation in the row direction because the crop is traditionally sown and harvested along the column direction, and the correlation is higher in the direction perpendicular to the sowing direction (Marchant et al. 2019). For the Matérn covariance matrix (7), we set the value of the variance parameter σ_d^2 to 1, the value of the parameter r to 1, and the value of the parameter ν to 3/2. After drawing samples of $\boldsymbol{u}$ from $\mathcal{N}(0, \Sigma_u)$, the spatially varying coefficients $\boldsymbol{\beta}_0$ and $\boldsymbol{\beta}_1$ are specified using the relations $\boldsymbol{\beta}_0 = b_0 + \boldsymbol{u}_0$ and $\boldsymbol{\beta}_1 = b_1 + \boldsymbol{u}_1$.

For the quadratic relationship, we consider the values of 65, 0.05 and -0.0003 for the coefficients b_0, b_1, and b_2, respectively. These choices make the response curve concave down. For the variance components, we set to 5 for σ_{u_0}, 0.01 for σ_{u_1}, and 0.0001 for σ_{u_2}. The rest of the parameters left unchanged. Consequently, the true spatially varying coefficients are $\boldsymbol{\beta}_0 = b_0 + \boldsymbol{u}_0$, $\boldsymbol{\beta}_1 = b_1 + \boldsymbol{u}_1$, and $\boldsymbol{\beta}_2 = b_2 + \boldsymbol{u}_2$ for the quadratic model.

Figure 2 illustrates the global yield response to Nitrogen for the linear and quadratic relationships.

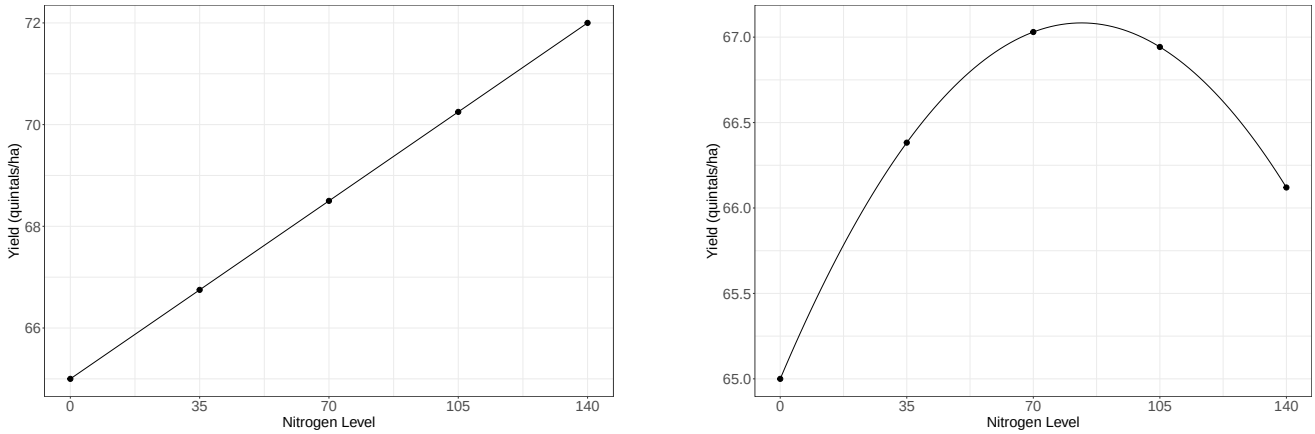

Figure 2: The global linear relationship of yield and nitrogen is $y = 65 + 0.05N$ (*left*), and the global quadratic relationship between yield and nitrogen is $y = 65 + 0.05N - 0.003N^2$ (*right*).

To summarise, the simulated yield response is obtained by

$$\begin{cases} \text{Linear} & y_i = b_0 + u_{0i} + (b_1 + u_{1i})N_i + e_i \\ \text{Quadratic} & y_i = b_0 + u_{0i} + (b_1 + u_{1i})N_i + (b_2 + u_{2i})N_i^2 + e_i \end{cases} \tag{13}$$

where N_i is the nitrogen rate, $e_i \sim \mathcal{N}(0, 1)$ is the error term at grid i, $i = 1, \ldots, n$.

4 Results

In this section, we assess the performances of randomised and systematic designs in terms of their utility to accurately estimate the model parameters for both linear and quadratic response models. To this end, we perform 1000 simulations of each of the 24 scenarios, described in the previous section. In each simulation, we first generate the coefficients for all grids and then apply the treatment in each grid to calculate the yield value using the model coefficients. For the treatment order in each iteration, we randomly picked an order of treatments for a single replicate and repeated this sequence across all other replicates to construct a systematic design. For a randomised design, all replicates have random orders of the treatments.

4.1 Comparison based on mean squared errors

Figures 3 and 4 show the results of linear models for the cases of low ($\epsilon = 1$) and high ($\epsilon = 0.1$) correlations, respectively, while Figures 5 and 6 show the results of quadratic models for the same low and high correlations, respectively. To specify the covariance matrix V_s used in producing the results in these figures, we use the following labels: (i) "NS" for the identity matrix representing no spatial correlation, (ii) "AR1" for AR1(0.15) $\otimes$ AR1(0.5), and (iii) "Matern" for Matérn covariance with $\nu = 3/2$. Note that the model parameters and their corresponding MSEs are small values, and this makes it difficult to compare the MSEs of different scenarios using the original scale of MSE values. Therefore, to enhance clarity in visualisation and comparison, we have multiplied the MSEs of $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ by 10^4 and 10^8, respectively, in the figures presented below.

For the linear response model, both randomised and systematic designs perform similarly, particularly for the case NS. Figure 3 shows that the MSE corresponding to $\hat{\boldsymbol{\beta}}_0$ for all bandwidths are fairly similar for both designs

without spatial correlation. However, when a spatial covariance matrix is incorporated in the model, the MSE results, presented also in Tables 1 and 2, in the figures below show that the MSE medians corresponding to $\hat{\beta}_1$ for AR1$\otimes$AR1 and Matérn cases are lower for the systematic design.

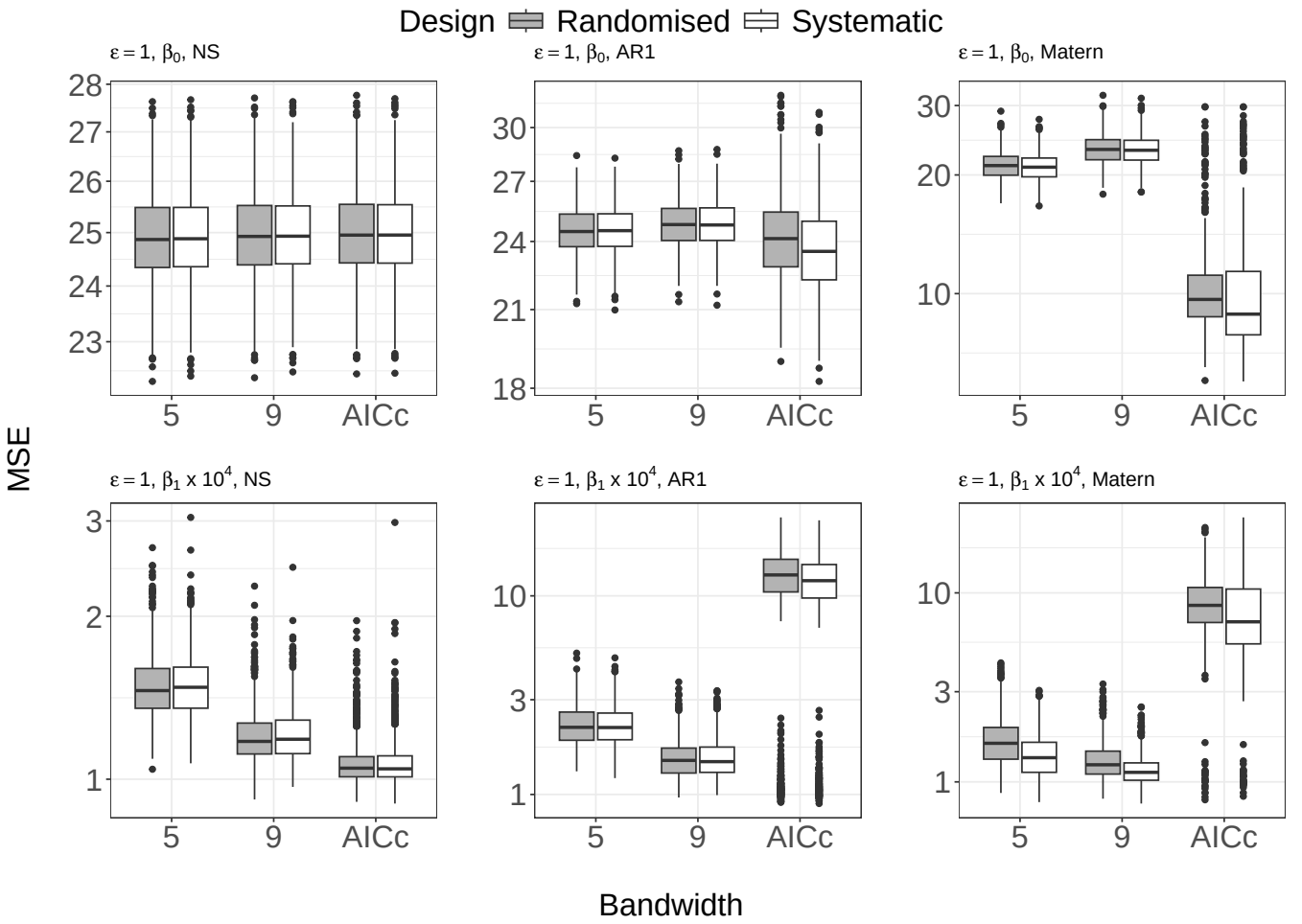

Figure 3: Boxplots of MSE for $\hat{\beta}_0$ and $\hat{\beta}_1$ in GWR models using different bandwidths for the simulated data with a linear response. The simulated data had different spatial covariance matrices (NS, AR1$\otimes$AR1 and Matérn) and a low correlation between the parameters ($\epsilon = 1$).

For the quadratic response model, Figures 5 and 6 show that the GWR estimates of both β_1 and β_2 based on fixed bandwidths of 5 and 9 for systematic designs outperform the estimates obtained for randomised designs when spatial correlation ("AR1" and "Matern") is present in yield data. Using the AICc optimal bandwidth, GWR successfully estimates the intercepts β_0 but fails to accurately estimate linear and quadratic coefficients β_1 and β_2, resulting in MSEs that are relatively larger than those obtained using a fixed bandwidth. Overall, the results of our simulation study indicate that the systematic designs are superior to randomised designs in enabling accurate and precise estimation of spatially varying treatment effects, especially when the response model is a quadratic function of the treatment levels.

Moreover, MSE comparisons reveal that the choice of bandwidth may influence the relative performance of the two designs differently depending on whether the intercept coefficient or the slope coefficients are being estimated. Differences in relative performance are also observed for different forms of spatial covariance matrices considered in the simulation scenarios. In scenarios without spatial variation, when estimating β_0, β_1 and β_2, the AICc-selected bandwidths produce the lowest MSE medians. In contrast, when spatial variation is present (utilising either AR1$\otimes$AR1 or Matérn covariance structures), the bandwidth of 9 consistently produces the most accurate estimates of β_1 and β_2,

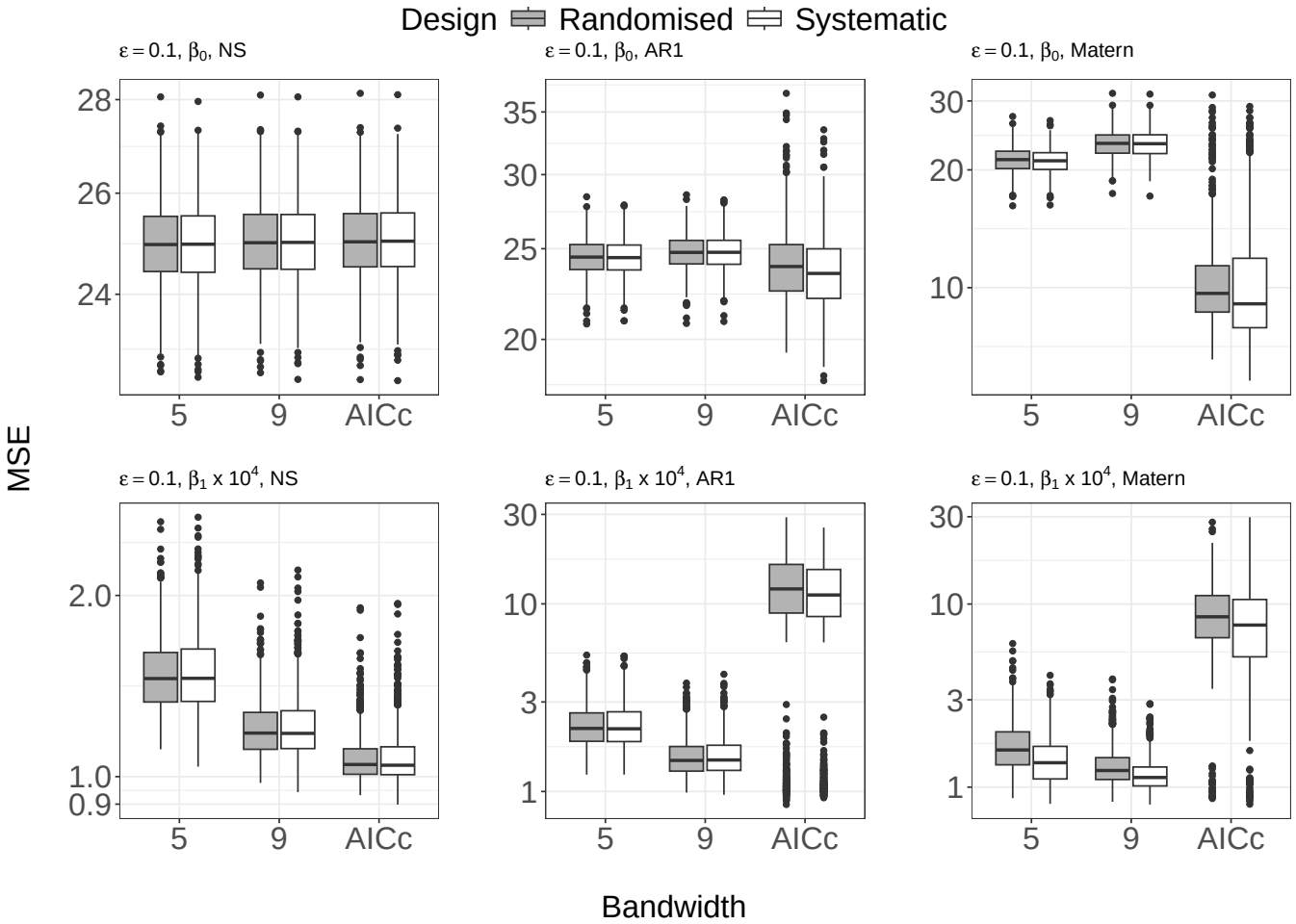

Figure 4: Boxplots of MSE for $\hat{\boldsymbol{\beta}}_0$ and $\hat{\boldsymbol{\beta}}_1$ in GWR models using different bandwidths for the simulated data with a linear response. The simulated data had different spatial covariance matrices (NS, AR1⊗AR1 and Matérn) and a high correlation between the parameters ($\epsilon = 0.1$).

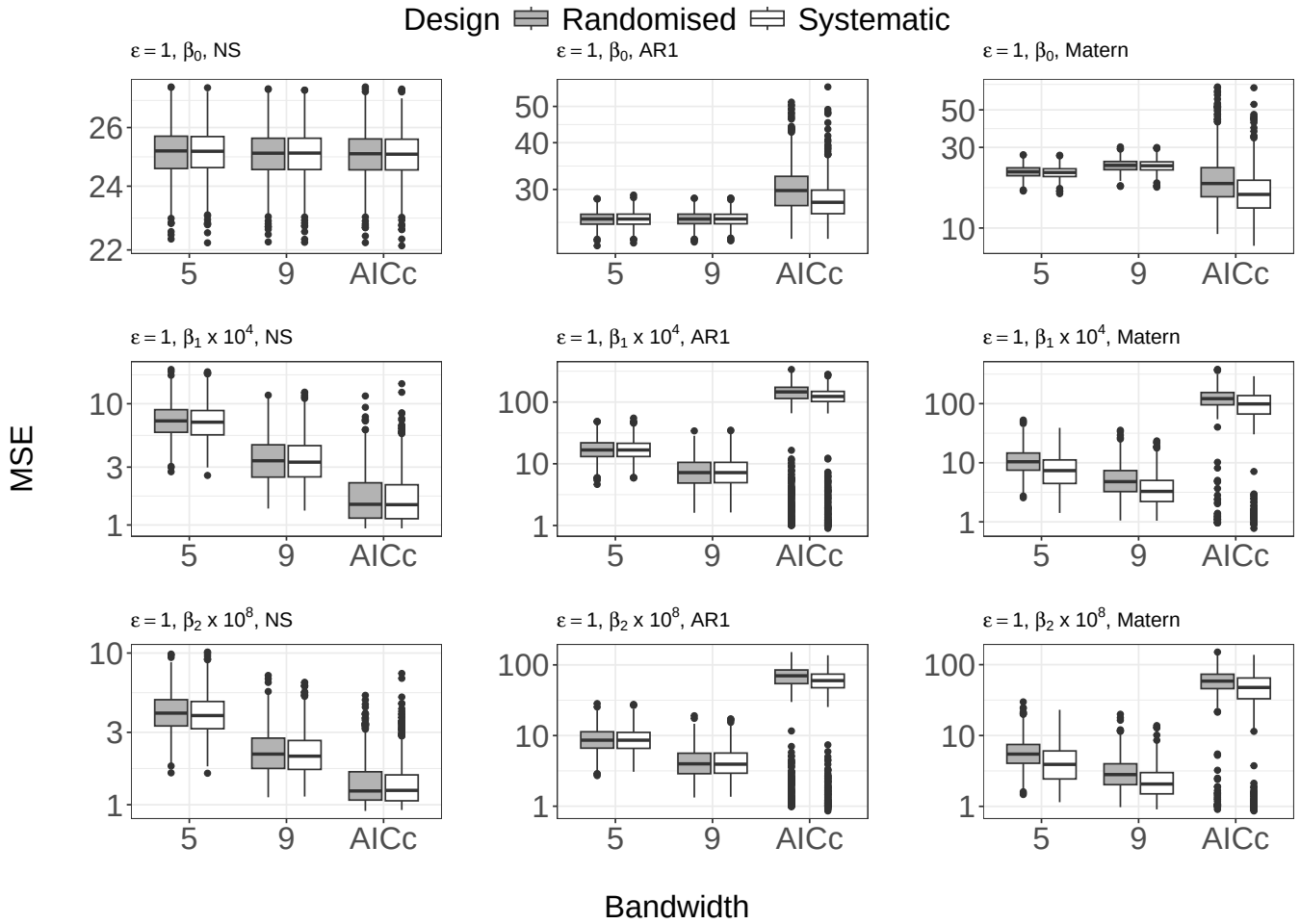

Figure 5: Boxplots of MSE for $\hat{\boldsymbol{\beta}}_0$, $\hat{\boldsymbol{\beta}}_1$ and $\hat{\boldsymbol{\beta}}_2$ in GWR models using different bandwidths for the simulated data with a quadratic response. The simulated data had different spatial covariance matrices (NS, AR1$\otimes$AR1 and Matérn) and a low correlation amongst the parameters ($\epsilon = 1$).

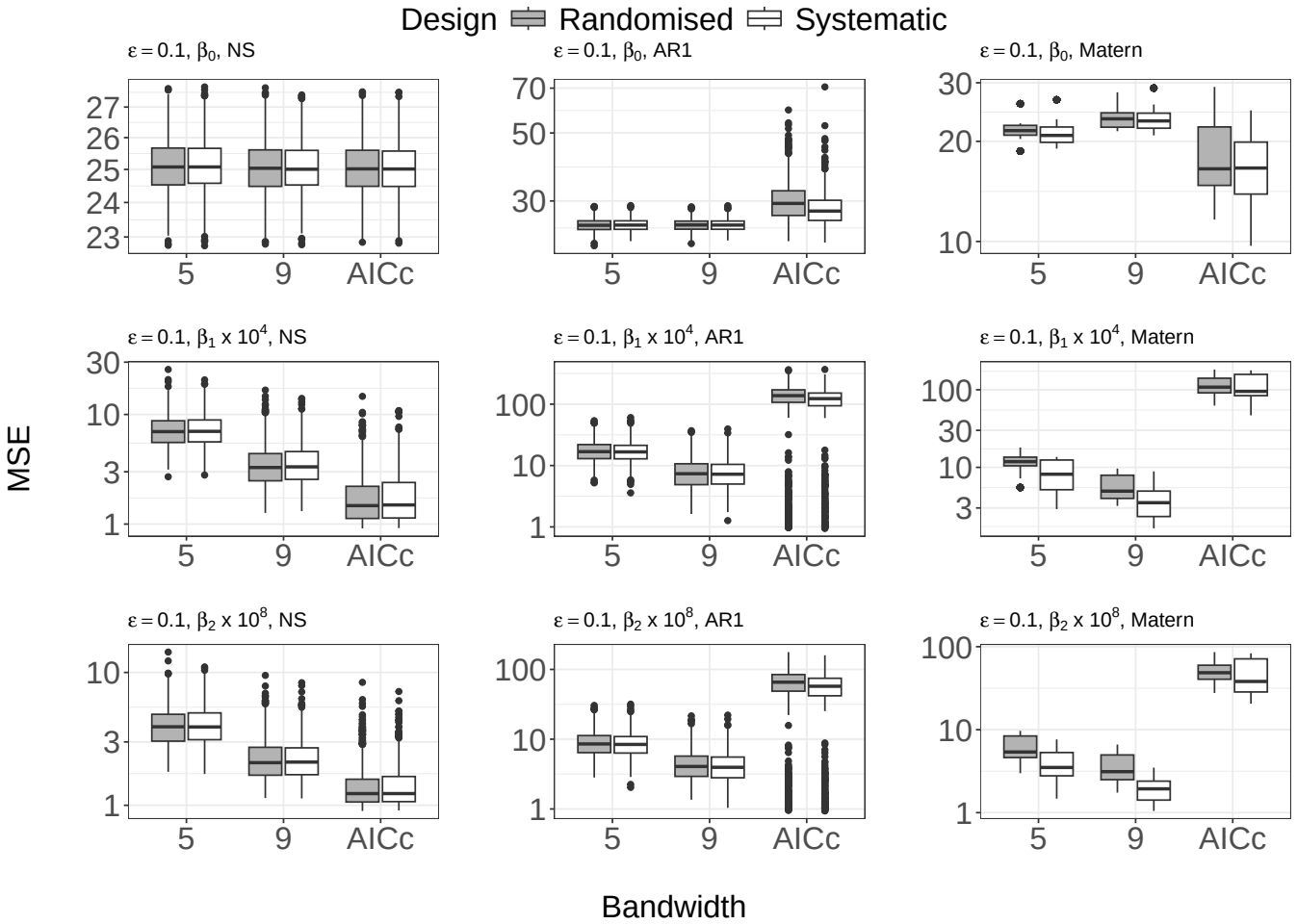

Figure 6: Boxplots of MSE for $\hat{\boldsymbol{\beta}}_0$, $\hat{\boldsymbol{\beta}}_1$ and $\hat{\boldsymbol{\beta}}_2$ in GWR models using different bandwidths for the simulated data with a quadratic response. The simulated data had different spatial covariance matrices (NS, AR1$\otimes$AR1 and Matérn) and a high correlation amongst the parameters ($\epsilon = 0.1$).

outperforming the estimates obtained using either the bandwidth of 5 or the one chosen by AICc.

Tables 1 and 2 present the median MSEs corresponding to the parameter estimation for the linear response model in the two scenarios of low ($\epsilon = 1$) and high ($\epsilon = 0.1$) correlations, respectively.

Table 1: Median MSE of GWR coefficient estimates for a linear response when the correlation between the parameters is low ($\epsilon = 1$).

Linear Response with $\epsilon = 1$		Randomised Design			Systematic Design		
V_s	Coefficients	Bandwidth 5	9	AICc	5	9	AICc
NS	$\hat{\beta}_0$	**24.868**†	24.927	24.952	24.883	24.931	24.953
	$\hat{\beta}_1(\times 10^4)$	1.458	1.175	1.048	1.479	1.185	**1.045**†
AR1	$\hat{\beta}_0$	24.471	24.807	24.134	24.513	24.788	**23.538**†
	$\hat{\beta}_1(\times 10^4)$	2.178	1.484	12.744	2.173	**1.463**†	11.922
Matern	$\hat{\beta}_0$	21.125	23.215	9.667	20.928	23.100	**8.871**†
	$\hat{\beta}_1(\times 10^4)$	1.600	1.232	8.591	1.343	**1.123**†	7.030

† Indicates the smallest MSE for the row.

Table 2: Median MSE of GWR coefficient estimates of linear response when the correlation between the parameters is high ($\epsilon = 0.1$).

Linear Response with $\epsilon = 0.1$		Randomised Design			Systematic Design		
V_s	Coefficients	Bandwidth 5	9	AICc	5	9	AICc
NS	$\hat{\beta}_0$	**24.964**†	25.000	25.018	24.970	25.006	25.031
	$\hat{\beta}_1(\times 10^4)$	1.455	1.181	1.047	1.456	1.180	**1.044**†
AR1⊗AR1	$\hat{\beta}_0$	24.490	24.773	23.928	24.457	24.780	**23.524**†
	$\hat{\beta}_1(\times 10^4)$	2.168	**1.464**†	12.014	2.159	1.472	11.153
Matérn	$\hat{\beta}_0$	21.247	23.376	9.668	21.100	23.330	**9.093**†
	$\hat{\beta}_1(\times 10^4)$	1.596	1.233	8.518	1.359	**1.129**†	7.676

† Indicates the smallest MSE for the row.

Tables 3 and 4 present the median MSEs corresponding to the estimation of quadratic response models for the two scenarios of low ($\epsilon = 1$) and high ($\epsilon = 0.1$) correlations, respectively.

Table 3: Median MSE of GWR coefficient estimates of quadratic response when the correlation amongst the parameters is low ($\epsilon = 1$).

Quadratic Response with $\epsilon = 1$		Randomised Design			Systematic Design		
V_s	Coefficients	Bandwidth 5	9	AICc	5	9	AICc
NS	$\hat{\beta}_0$	25.184	25.106	25.086	25.172	25.109	**25.072**†
	$\hat{\beta}_1(\times 10^4)$	7.215	3.385	1.480	7.045	3.297	**1.471**†
	$\hat{\beta}_2(\times 10^8)$	4.016	2.157	**1.232**†	3.871	2.090	1.243
AR1⊗AR1	$\hat{\beta}_0$	25.012	**25.005**†	29.797	25.008	25.013	27.712
	$\hat{\beta}_1(\times 10^4)$	16.741	7.1907	146.097	16.730	**7.1906**†	123.256
	$\hat{\beta}_2(\times 10^8)$	8.601	3.979	70.112	8.595	**3.933**†	59.913
Matérn	$\hat{\beta}_0$	21.503	23.470	18.331	21.305	23.359	**15.800**†
	$\hat{\beta}_1(\times 10^4)$	10.397	4.790	121.474	7.368	**3.276**†	98.902
	$\hat{\beta}_2(\times 10^8)$	5.470	2.808	58.626	3.912	**2.068**†	47.653

† Indicates the smallest MSE for the row.

Table 4: Median MSE of GWR coefficient estimates of quadratic response when the correlation amongst the parameters is high ($\epsilon = 0.1$).

Quadratic Response with $\epsilon = 0.1$		Randomised Design			Systematic Design		
V_s	Coefficients	Bandwidth					
		5	9	AICc	5	9	AICc
NS	$\hat{\beta}_0$	25.076	25.036	25.017	25.076	**25.006**†	25.013
	$\hat{\beta}_1(\times 10^4)$	6.974	3.281	**1.473**†	7.011	3.326	1.494
	$\hat{\beta}_2(\times 10^8)$	3.900	2.093	**1.224**†	3.888	2.114	1.225
AR1⊗AR1	$\hat{\beta}_0$	**24.993**†	25.051	29.454	25.027	25.032	27.809
	$\hat{\beta}_1(\times 10^4)$	16.835	7.350	137.643	16.678	**7.220**†	123.024
	$\hat{\beta}_2(\times 10^8)$	8.547	4.065	65.609	8.413	**3.955**†	57.375
Matérn	$\hat{\beta}_0$	21.542	23.376	**16.560**†	20.837	23.039	16.630
	$\hat{\beta}_1(\times 10^4)$	11.864	4.953	108.094	8.248	**3.532**†	95.428
	$\hat{\beta}_2(\times 10^8)$	5.389	3.121	48.640	3.506	**1.940**†	38.378

† Indicates the smallest MSE for the row.

4.2 Comparison of density plot

Figure 7 illustrate the density plots comparing the true coefficients of β_0, β_1, and β_2 with their estimates derived from both randomised and systematic designs with Matérn covariance and low within-grid correlation and fitted by GWR with bandwidth 9. The plots reveal that the true coefficients are well-represented by the GWR estimates, with the systematic design showing slightly tighter distributions compared to the randomised design of $\hat{\beta}_1$ and $\hat{\beta}_2$. This suggests the systematic design provides more precise estimates under the given conditions.

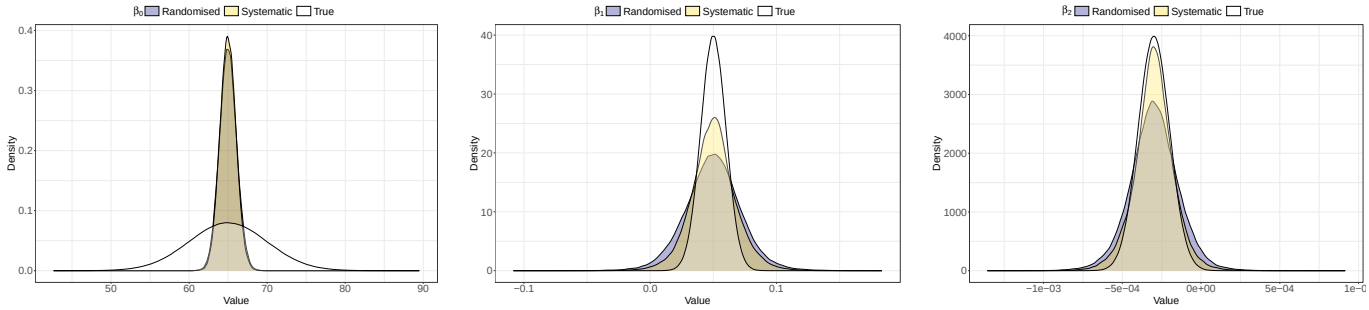

Figure 7: Density plots comparing the true coefficients of β_0 (left), β_1 (middle), and β_2 (right) with their estimates from GWR with bandwidth 9, based on 1000 simulations.

4.3 An example of optimal nitrogen map

In practice, growers are more interested in the prescription map that tells them where the appropriate nitrogen should be applied on the paddock. With the application of GWR, we can find the local variations in crop needs, allowing for more precise and efficient nitrogen application. Each grid of the paddock receives the optimal amount of fertiliser. Consequently, this leads to improved crop yields, reduced investment cost and high profit.

Figure 8 is the simulated crop yield map with the assumption of a quadratic response curve and Matérn spatial covariance and low within-grid correlation. These two yield maps have the same coefficients but different yields due to different treatment layouts.

Figure 9 illustrates an example of the optimal Nitrogen rate (kg/ha) map estimated by GWR with a bandwidth of

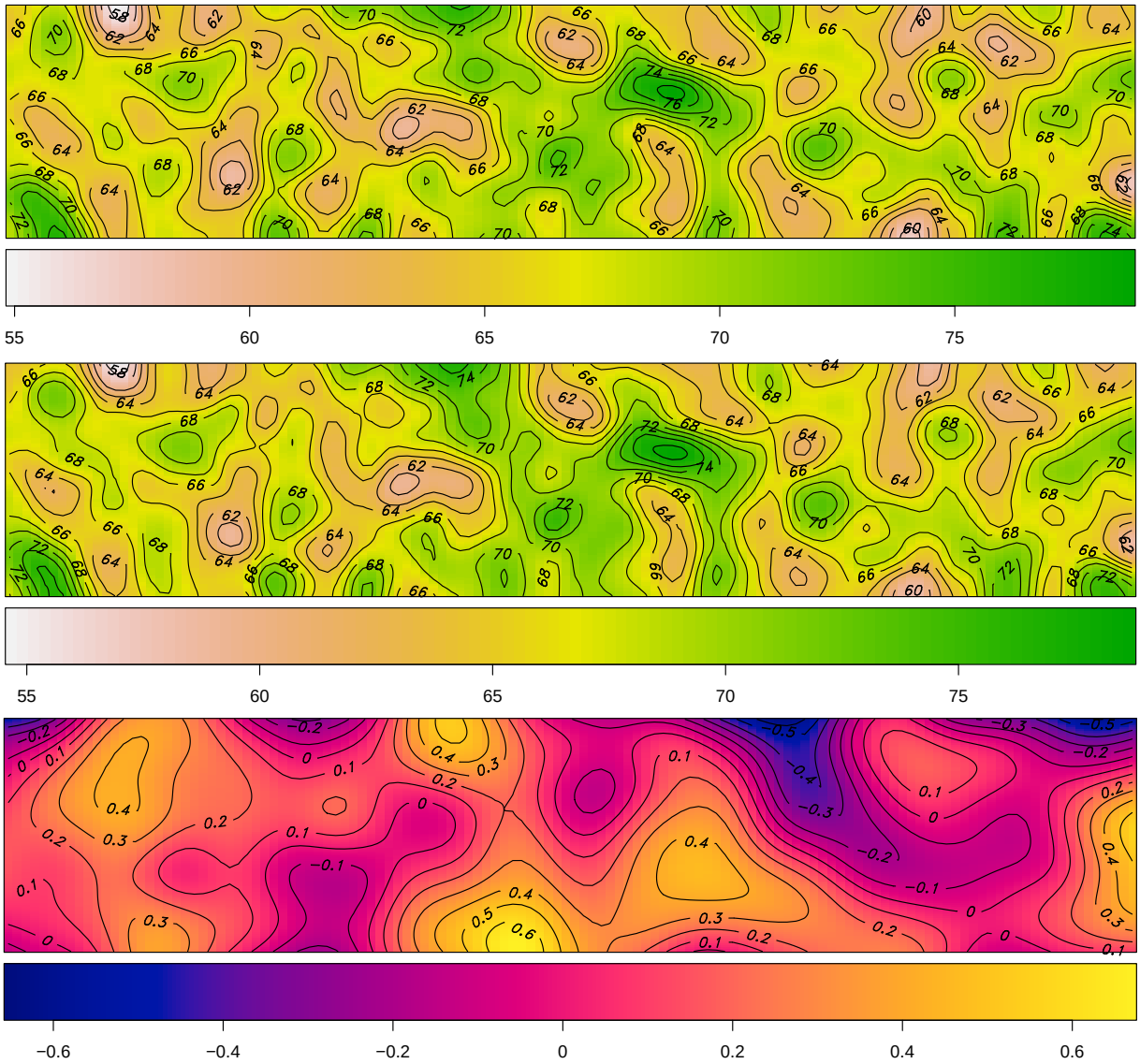

Figure 8: Simulated yield map of a randomised design (*top*), and a systematic design (*middle*), and the difference of these two designs (*bottom*).

9 using the above yield data. The optimal rate at grid i is given by $\hat{N}_i = -\hat{\beta}_{1i}/(2\hat{\beta}_{2i})$ with constraints between 0 and 140, $i = 1, \ldots, n$. For the randomised design, GWR underestimated the right part of the paddock. On the contrary, the estimated map from the systematic design is more consistent.

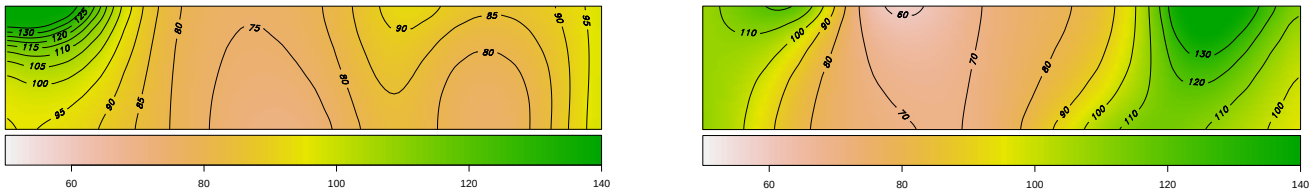

Figure 9: Optimal Nitrogen rate estimated by GWR from a randomised design (*left*), and a systematic design (*right*).

5 Discussion

Agronomists and biometricians generally prefer randomised designs for OFE trials. This is likely due to their experience with small-plot field experiments, where randomised designs are employed. Our simulation study shows a systematic design performs either preferably or similarly to a randomised design for the purposes of creating a varying treatment map when using the performance metrics stated. The differentiating factors primarily included the response type and the spatial covariance model, while the correlation amongst the treatment coefficients is not found to be important. These are factors that can be assessed by the farmer and their agronomists beforehand, and this should dictate which design should be used. Given that a systematic design is easier to implement in the field, and shows little downside to creating a varying treatment map, the use of systematic designs is recommended.

The response type is the main differentiating factor between randomised and systematic designs. When the response is quadratic, the systematic design performed favourably, contrasting with the result of the linear design. Given this, if an approximately linear response is expected in the field, then the selection of the design may not be important. However, given the variable nature of the relationship between response and treatment over a large field (see Rakshit et al. (2020)), it may be wise to implement a systematic design for the potential outcome of a quadratic relationship.

Another consideration for agronomists and biometricians when selecting which design to use is the expected spatial covariance structure in the field. When no spatial structure was simulated, the differences between the prediction from the systematic and random design are minimal. This result should be expected given that if there are no spatial autocorrelations, then the individual query grids are independent observations, and therefore the design is not important. However, when a first-order auto-regressive structure was simulated, the differences are noticeable when a quadratic response was used, showing systematic designs to be preferential. The largest difference between the two designs occurred when considering the Matérn spatial covariance structure, which showed a clear preference for systematic designs when a quadratic response was considered, and also a small preference for systematic designs for a linear response. Therefore, only if spatial variability is predicted to be negligible in the field would using a randomised design be reasonable given a quadratic response. This assumption of negligible spatial variability would be difficult to reason with given the large fields used in on-farm experimentation, meaning that in the application a systematic design should be used.

There are significant deficiencies found in using AICc for bandwidth selection. The AICc-selected bandwidths skewed to 1 and, in a few cases, ended in 93 (number of rows). Even though the bandwidth is optimal according to

AICc, the MSE is higher than when using a fixed bandwidth. Therefore, a fixed bandwidth based on the experimental design (5 or 9 in this case) is recommended, rather than AICc-selected bandwidth. Selecting the bandwidth based on the experimental design is also theoretically better since when only a single measurement is observed in each grid, all levels of the treatment factor should be included in a GWR window at the same time to interpolate the relationship. If more than one level is missing, then the interpolation is incomplete.

Our study employs simulation to evaluate the effectiveness of GWR in estimating coefficients for both randomised and systematic designs. However, a potential drawback of using simulation is its high dependency on the assumptions underlying the parametric model. If these assumptions are incorrect or fail to accurately represent real-world conditions, the simulation results may not reliably reflect the actual scenario.

Given the scope of the paper, some designs and factors were not considered. Designs such as chequerboard or wave designs have been suggested for on-farm experiments (Bramley et al. 1999), however, were not considered here. Topographical factors (spatial zones) were also not entertained in our study. Since GWR estimates a global template model and then adjusts it at a local scale across the study region, the variation between zones is "flushed out" by the spatial covariance.

6 Conclusion

This research offers a number of recommendations for agronomists and biometricians for designing OFE trials.

- Systematic designs are suitable for OFE trials, particularly when the results will be used to develop a variable rate map for the treatment, such as fertiliser.
- Incorporation of a quadratic function for GWR is preferred for systematic designs when spatial variability is present.
- Using a fixed bandwidth for GWR analysis based on experimental design, such as the number of treatments.

Acknowledgements

Curtin Biometry and Agricultural Data Analytics (CBADA) gratefully acknowledges the support from the Grains Research and Development Corporation of Australia (GRDC).

CRediT authorship contribution statement

Zhanglong Cao: Conceptualization; Formal analysis; Methodology; Software; Validation; Visualization; Roles/Writing - original draft; and Writing - review & editing. **Jordan Brown**: Investigation; Methodology; Visualization; Roles/Writing - original draft; and Writing - review & editing. **Mark Gibberd**: Project administration; Roles/Writing - original draft; and Writing - review & editing. **Julia Easton**: Roles/Writing - original draft; and Writing - review & editing. **Suman Rakshit**: Conceptualization; Methodology; Supervision; Roles/Writing - original draft; and Writing - review & editing.

References

Abramowitz, M. (1974). *Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables*. USA: Dover Publications, Inc. (cit. on p. 5).

Alesso, C. A., Cipriotti, P. A., Bollero, G. A., and Martin, N. F. (2021). "Design of on-farm precision experiments to estimate site-specific crop responses". In: *Agronomy Journal* 113.2, pp. 1366–1380. DOI: https://doi.org/10.1002/agj2.20572 (cit. on p. 3).

Anselin, L., Bongiovanni, R., and Lowenberg-DeBoer, J. (2004). "A Spatial Econometric Approach to the Economics of Site-Specific Nitrogen Management in Corn Production". In: *American Journal of Agricultural Economics* 86.3, pp. 675–687. DOI: https://doi.org/10.1111/j.0002-9092.2004.00610.x (cit. on p. 6).

Bramley, R., Cook, S., Adams, M., and Corner, R. (1999). "Designing your own on-farm experiments: How precision agriculture can help". In: (cit. on p. 17).

Butler, D., Cullis, B., Gilmour, A., Gogel, B., and Thompson, R. (2017). *ASReml-R Reference Manual Version 4* (cit. on p. 4).

Cao, Z., Stefanova, K., Gibberd, M., and Rakshit, S. (2022). "Bayesian inference of spatially correlated random parameters for on-farm experiment". In: *Field Crops Research* 281, p. 108477. DOI: https://doi.org/10.1016/j.fcr.2022.108477 (cit. on pp. 2, 4, 6, 7).

Cook, S., Cock, J., Oberthür, T., and Fisher, M. (2013). "On-Farm Experimentation". In: *Better Crop. Plant Food.* 4th ser. 97, pp. 17–20 (cit. on p. 2).

Cressie, N. and Huang, H.-C. (Dec. 1999). "Classes of Nonseparable, Spatio-Temporal Stationary Covariance Functions". In: *Journal of the American Statistical Association* 94.448, pp. 1330–1339. DOI: 10.1080/01621459.1999.10473885 (cit. on p. 5).

Evans, F. H., Recalde Salas, A., Rakshit, S., Scanlan, C. A., and Cook, S. E. (Nov. 2020). "Assessment of the Use of Geographically Weighted Regression for Analysis of Large On-Farm Experiments and Implications for Practical Application". en. In: *Agronomy* 10.11, p. 1720. DOI: 10.3390/agronomy10111720 (cit. on pp. 2, 3, 6).

Fisher, R. A. (1934). *Statistical Methods for Research Workers*. English. fifth. Edinburgh: Oliver and Boyd (cit. on p. 2).

Glynn, C. (2007). "Testing the Growth-Differentiation Balance Hypothesis: Dynamic Responses of Willows to Nutrient Availability". In: *New Phytologist* 176.3, pp. 623–634. DOI: 10.1111/j.1469-8137.2007.02203.x (cit. on p. 3).

Gollini, I., Lu, B., Charlton, M., Brunsdon, C., and Harris, P. (Feb. 2015). "GWmodel: An R Package for Exploring Spatial Heterogeneity Using Geographically Weighted Models". en. In: *Journal of Statistical Software* 63.1, pp. 1–50. DOI: 10.18637/jss.v063.i17 (cit. on pp. 3, 5).

Lewandowski, D., Kurowicka, D., and Joe, H. (2009). "Generating random correlation matrices based on vines and extended onion method". In: *Journal of Multivariate Analysis* 100.9, pp. 1989–2001. DOI: https://doi.org/10.1016/j.jmva.2009.04.008 (cit. on p. 4).

Liben, F. M., Midega, T., Tufa, T., and Wortmann, C. S. (2019). "Soil Fertility & Crop Nutrition Barley and wheat nutrient responses for Shewa, Ethiopia". In: *Agronomy Journal*. DOI: 10.1002/agj2.20020 (cit. on p. 3).

Lu, B., Harris, P., Charlton, M., and Brunsdon, C. (2014). "The GWmodel R package: further topics for exploring spatial heterogeneity using geographically weighted models". In: *Geo-spatial Information Science* 17.2, pp. 85–101. DOI: 10.1080/10095020.2014.917453 (cit. on p. 3).

Marchant, B., Rudolph, S., Roques, S., Kindred, D., Gillingham, V., Welham, S., Coleman, C., and Sylvester-Bradley, R. (Jan. 2019). "Establishing the Precision and Robustness of Farmers' Crop Experiments". en. In: *Field Crops Research* 230, pp. 31–45. DOI: 10.1016/j.fcr.2018.10.006 (cit. on p. 7).

Marschner, H. (2011). *Marschner's Mineral Nutrition of Higher Plants*. San Diego: San Diego: Elsevier Science & Technology. DOI: 10.1016/C2009-0-63043-9 (cit. on p. 3).

McElreath, R. (Dec. 2015). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. English. First. Vol. 122. Chapman and Hall/CRC Texts in Statistical Science Ser. CRC Press LLC (cit. on p. 4).

Páez, A., Uchida, T., and Miyamoto, K. (Apr. 2002). "A General Framework for Estimation and Inference of Geographically Weighted Regression Models: 1. Location-Specific Kernel Bandwidths and a Test for Locational Heterogeneity". In: *Environment and Planning A: Economy and Space* 34.4, pp. 733–754. DOI: 10.1068/a34110 (cit. on p. 5).

Pandit, R. K. and Infield, D. (Sept. 2019). "Comparative Analysis of Gaussian Process Power Curve Models Based on Different Stationary Covariance Functions for the Purpose of Improving Model Accuracy". en. In: *Renewable Energy* 140, pp. 190–202. DOI: 10.1016/j.renene.2019.03.047 (cit. on p. 5).

Petersen, R. G. (1994). *Agricultural Field Experiments: Design and Analysis*. en. 1st. Boca Raton: CRC Press (cit. on pp. 2, 3).

Piepho, H. P. and Edmondson, R. N. (2018). "A tutorial on the statistical analysis of factorial experiments with qualitative and quantitative treatment factor levels". In: *Journal of Agronomy and Crop Science* 204.5, pp. 429–455. DOI: https://doi.org/10.1111/jac.12267 (cit. on p. 3).

Piepho, H. P., Möhring, J., and Williams, E. R. (2013). "Why Randomize Agricultural Experiments?" In: *Journal of Agronomy and Crop Science* 199.5, pp. 374–383. DOI: 10.1111/jac.12026 (cit. on p. 2).

Pringle, M. J., Cook, S. E., and McBratney, A. B. (Dec. 2004). "Field-Scale Experiments for Site-Specific Crop Management. Part I: Design Considerations". en. In: *Precision Agriculture* 5.6, pp. 617–624. DOI: 10.1007/s11119-004-6346-1 (cit. on pp. 2, 3).

Rakshit, S., Baddeley, A., Stefanova, K., Reeves, K., Chen, K., Cao, Z., Evans, F., and Gibberd, M. (2020). "Novel Approach to the Analysis of Spatially-Varying Treatment Effects in on-Farm Experiments". In: *Field Crops Research* 255.October 2019, p. 107783. DOI: 10/gg2vv7 (cit. on pp. 2, 5, 6, 16).

Selle, M. L., Steinsland, I., Hickey, J. M., and Gorjanc, G. (2019). "Flexible Modelling of Spatial Variation in Agricultural Field Trials with the R Package INLA". In: *Theoretical and Applied Genetics* 132.12, pp. 3277–3293. DOI: 10.1007/s00122-019-03424-y (cit. on pp. 2, 5).

Verdooren, L. R. (Dec. 2020). "History of the Statistical Design of Agricultural Experiments". en. In: *Journal of Agricultural, Biological and Environmental Statistics* 25.4, pp. 457–486. DOI: 10.1007/s13253-020-00394-3 (cit. on p. 2).

White, J. W. and Evert, F. K. van (2008). "Publishing Agronomic Data". In: *Agronomy Journal* 100.5, pp. 1396–1400. DOI: https://doi.org/10.2134/agronj2008.0080F (cit. on p. 6).

Zimmerman, D. L. and Harville, D. A. (1991). "A Random Field Approach to the Analysis of Field-Plot Experiments and Other Spatial Experiments". In: *Biometrics* 47.1, pp. 223–239. DOI: 10.2307/2532508 (cit. on p. 3).