

Field Crops Research

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

--Manuscript Draft--

Manuscript Number:	FIELD-D-23-02059R2
Article Type:	Research Paper
Keywords:	yield map; optimal treatment; spatially varying coefficients; geographically weighted regression; precision agriculture
Corresponding Author:	Zhanglong Cao Curtin University Centre for Crop and Disease Management Bentley, WA AUSTRALIA
First Author:	Zhanglong Cao
Order of Authors:	Zhanglong Cao Jordan Brown Julia Easton Mark Gibberd Suman Rakshit
Abstract:	<p>CONTEXT OR PROBLEM: Randomised designs are often preferred over systematic designs by agronomists and biometrists. 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.</p> <p>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.</p> <p>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.</p> <p>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.</p> <p>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 systematic design should be used as it allows us to obtain better estimates than those obtained from a randomised design in post-experiment statistical modelling.</p> <p>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.</p>
Suggested Reviewers:	Matthew Pringle, doctorate Queensland Department of Environment and Science matthew.pringle@qld.gov.au Our research has been profoundly inspired by the contributions of Matthew. J. Pringle, especially as outlined in his pivotal works, 'Field-Scale Experiments for Site-Specific Crop Management. Part I: Design Considerations' and 'Field-Scale Experiments for Site-Specific Crop Management. Part II: A Geostatistical Analysis.' His comprehensive expertise in spatial experiment analysis directly complements the thematic core of our study. Dr. Pringle's extensive background and insightful explorations in this field

	<p>position him as an ideal candidate to critically assess our research. We kindly recommend him as a potential reviewer, confident that his feedback would greatly enrich the depth and quality of our work.</p>
	<p>Hans-Peter Piepho, doctorate Director, University of Hohenheim piepho@uni-hohenheim.de Prof Hans-Peter Piepho is a distinguished statistician and contributes to the field of agricultural statistics, particularly in relation to large strip trials and advanced statistical methodologies. His expertise is particularly relevant for our study, as the Bayesian model we employed in our simulation draws inspiration from his work, "Statistical aspects of on-farm experimentation". Prof. Piepho's extensive background in statistical analysis and his innovative approaches to on-farm experimentation provide him with the unique insights necessary to critically evaluate and potentially enrich the statistical rigor and application of our findings.</p>
	<p>Fiona Evans, doctorate Research lead, Data Analysis Australia fiona@daa.com.au Fiona has extensive experience and expertise in the field of geospatial analysis, particularly her simulation work using Geographically Weighted Regression (GWR). Her paper, "Assessment of the Use of Geographically Weighted Regression for Analysis of Large On-Farm Experiments and Implications for Practical Application," aligns with the methodological framework of our study. Fiona's understanding of GWR and its application to OFE makes her uniquely positioned to provide insightful and valuable feedback, potentially enriching the depth and applicability of our findings.</p>
	<p>Maria Lie Selle, doctorate Norwegian University of Science and Technology maria.selle@ntnu.no Dr. Maria Lie Selle did some work in the field of Bayesian modelling and statistics, especially in the context of spatial variability in large agricultural trials. Her work, particularly the study titled "Flexible modelling of spatial variation in agricultural field trials with the R package INLA," directly aligns with the methodologies used in our paper for the simulation, and application to a various of spatial scenarios.</p>
	<p>Osval Antonio Montesinos-López, doctorate Universidad de Colima oamontes2@hotmail.com Dr. Osval Antonio Montesinos-López has expertise in OFE trials and advanced statistical analyses. His pioneering work, particularly in "Multivariate Bayesian Analysis of On-Farm Trials with Multiple-Trait and Multiple-Environment Data," showcases his deep understanding of complex statistical methodologies and their application to agricultural research. Dr. Montesinos-López's experience aligns with the analytical approaches we have adopted in our study.</p>
Response to Reviewers:	

Dear Editor,

On behalf of all authors, I am writing to submit our manuscript entitled, "*Optimal design for on-farm strip trials --- systematic or randomised?*" for consideration as a *Field Crops Research* paper.

In recent years, on-farm experimentation (OFE) has garnered substantial attention from researchers. However, the proper design of large-strip trials has received inadequate discussion. Traditionally, agronomists and biometricalians have favoured randomised designs for OFE trials. Our research, employing simulation studies that integrate Bayesian hierarchical models and geographically weighted regression (GWR) to create varying treatment maps, reveals that a systematic design outperforms a randomised design in terms of robustness and smaller mean squared errors (MSE) on coefficients under specific circumstances.

We believe that our work contributes significantly to the ongoing research on OFE topic and hope to raise awareness and stimulate further discussion within the research community. This exploration opens avenues for refining OFE methodologies and encourages a more nuanced consideration of design choices in agricultural experimentation.

We confirm that this manuscript has not been previously published and is not currently under consideration by any other journals, and that all authors have approved the contents of the paper and have agreed to *Field Crops Research*'s submission policies. We are unaware of any conflicts of interest with respect to the named authors.

Yours sincerely,

Zhanglong Cao
Corresponding Author
Research Fellow
Curtin Biometry and Agricultural Data Analytics, Centre for Crop and Disease Management
Curtin University
Bentley, Australia, 6102
zhanglong.cao@curtin.edu.au
+61 (04) 0586 6482

Jordan Brown, jordan.brown@curtin.edu.au
Julia Easton, julia.easton@curtin.edu.au
Mark Gibberd, M.Gibberd@exchange.curtin.edu.au
Suman Rakshit, suman.rakshit@curtin.edu.au

Dear Editor,

On behalf of all authors, I am writing to submit the revised version of our manuscript entitled, "*Optimal design for on-farm strip trials --- systematic or randomised?*" for consideration as a *Field Crops Research* paper.

We have thoroughly addressed all comments provided by the reviewers and have made significant improvements to our manuscript. These enhancements include rewriting certain sections for clarity, reorganising content for better flow and coherence, and strengthening our arguments with additional results. We believe these changes have greatly augmented both the quality and impact of our work.

We are confident that our research contributes significantly to ongoing studies on On-Farm Experimentation (OFE) topics. Our findings aim to raise awareness and stimulate further discussion within the research community. This exploration opens avenues for refining OFE methodologies and encourages a more nuanced consideration of design choices in agricultural experimentation.

We confirm that this manuscript has not been previously published and is not currently under consideration by any other journals. All authors have approved the contents of this paper and have agreed to Field Crops Research's submission policies. We are unaware of any conflicts of interest concerning the named authors.

Yours sincerely,

Zhanglong Cao
Corresponding Author
Research Fellow
Curtin Biometry and Agricultural Data Analytics, Centre for Crop and Disease Management
Curtin University
Bentley, Australia, 6102
zhanglong.cao@curtin.edu.au
+61 (04) 0586 6482

Jordan Brown, jordan.brown@curtin.edu.au
Julia Easton, julia.easton@curtin.edu.au
Mark Gibberd, M.Gibberd@exchange.curtin.edu.au
Suman Rakshit, suman.rakshit@curtin.edu.au

¹ Revision and Response to Reviewers' Comments

² on the manuscript (FIELD-D-23-02059R1)

³ Optimal design for on-farm strip trials — systematic or randomised?

⁴ Zhanglong Cao^{1,*}, Jordan Brown¹, Mark Gibberd¹, Julia Easton¹, and Suman
⁵ Rakshit^{1,2}

⁶ ¹Curtin Biometry and Agricultural Data Analytics, Centre for Crop and Disease
⁷ Management, Curtin University, Perth, Australia

⁸ ²School of Electrical Engineering, Computing and Mathematical Sciences, Curtin
⁹ University, Perth, Australia

¹⁰ *Corresponding author: Zhanglong Cao, zhanglong.cao@curtin.edu.au

¹¹ The authors thank the Associate Editor and Reviewers for their helpful comments and valuable
¹² suggestions. All comments/recommendations are addressed below, where each reviewer's comment
¹³ is presented in **blue text** and each response is given in **black text**. As part of this revision
¹⁴ process, the newly added materials in the revised draft of the paper are presented using **red text**.

¹⁵ **Reviewer 1 writes:**

¹⁶ *This paper compares systematic to randomized designs for on-farm field experimentation. The
17 objective the authors are having in mind is the estimation of spatially varying response functions.
18 The authors find that the systematic design yields more accurate predictions, as anticipated by the
19 authors and some cited references.*

²⁰ *1. My main suggestion is that the authors could also use the exact same simulation to evaluate
21 performance of another objective: Estimation of the marginal regression function, given by the
22 fixed effects b . Randomization theory would be expected to ensure that the randomized design
23 yields unbiased estimates and also yield valid F- and t-tests for treatment effects. It would be very
24 interesting to see how the systematic design fares in comparison when this is the specific purpose
25 of analysis. Of course, as global estimation is the objective in this case, GWR would need to be
26 replaced by classical regression analysis, perhaps just for the model used for simulating the data.
27 Simulating a suitable null hypothesis may also be useful. The linear model could provide a suitable
28 null model for the quadratic as an alternative. Depending on the outcome, such an extension of
29 the simulation would also further accentuate the general conclusion, and potentially enforce the
30 message that a systematic design may be more suitable for local estimation, whereas a randomized
31 design may be preferable for global estimation.*

³² Thank you for your insightful suggestion. We agree that evaluating the performance of the
33 estimation of the marginal regression function, given by the fixed effects b , could provide additional
34 valuable insights. Those are some current research topics that we are working on.

³⁵ The difficulty is how to analyse the global effects and what models are appropriate for the topic.
³⁶ Stefanova et al. (2023) proposed an approach for comparing the treatments/varieties of a large
³⁷ strip OFE trial by partitioning the entire paddock into smaller regions, named pseudo-environment,

38 then used a MET-like approach to predict the best treatments of each pseudo-environment and
39 the entire paddock.

40 We would like to extend our simulation from continuous variables, such as nitrogen rate, to
41 discrete variables, such as varieties, and incorporate Stefanova et al. (2023)'s approach to the
42 data to test the performance of systematic designs and randomised designs. The difficulties are
43 how to generate these simulated data robustly, how to automatically apply the approach to the
44 simulated data in a few iterations, and subsequently to compare the estimated values with the
45 true values.

46 In the meantime, we are investigating the GAM model, the additive model, with incorporating
47 the spline term to accommodate the spatial variation of OFE. Hopefully, we will achieve nice
48 results for publication.

49 *2. My major questions are these (see line 156): For the randomized allocation, was a new
50 randomization generated in each simulation run? Or was the same randomization used in each
51 run. Regarding the systematic design, how was the systematic order determined? Was this the
52 same in each simulation run, and was this equal to that shown in Figure 1? Or was a new
53 systematic order chosen in each run. More importantly, what can be said about the optimal
54 systematic order for the purpose of map estimation? How can this optimum be found? How can
55 optimality be defined? It may well be that the authors do not have an answer to this question.
56 However, the question is relevant, and it is worth at least mentioning in the discussion. At any
57 rate, the choice of systematic order the authors made needs to be explained and justified.*

58 We appreciate the reviewer for pointing out this crucial confusion in this manuscript. It is not
59 clearly explained in the first draft. We have revised the manuscript by the following amendments:

60 We added one more subsection [2.3 Performance evaluation](#) to describe the evaluation of an
61 optimal design:

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

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

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

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

70 At the beginning of section 4, we rewrote the paragraph that:

71 *In this section, we assess the performances of randomised and systematic designs in terms of
72 their utility to accurately estimate the model parameters for both linear and quadratic response
73 models. To this end, we perform 1000 simulations of each of the 24 scenarios, described in the
74 previous section. In each simulation, we first generate the coefficients for all grids and then
75 apply the treatment in each grid to calculate the yield value using the model coefficients. For
76 the treatment order in each iteration, we randomly picked an order of treatments for a single
77 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.*

78 *3. In the Methods section, the subheadings could be reconsidered. Perhaps they can be reworded
79 to clarify the purpose of each model. Which model is used for simulating the data and which
80 model is used for analysing the simulated data?*

81 Thank you for the comment. I have merged the subsections 2.1 and 2.2 to make a clear
82 explanation of the model, and revised the subsection titles to: [2.1 Hierarchical model for generating](#)

83 simulated data and 2.2 Fitting geographically weighted regression to simulated data. Additionally,
84 we have added one more subsection that 2.3 Performance evaluation to explain how to evaluate
85 the performance.

86 4. At the start of Section 3, the authors state an advantage of data simulated from a parametric
87 model. Are there any disadvantages, and what alternatives could be considered? Perhaps this is
88 also a point for the discussion.

89 Thank you for your insightful comment. It is right, while data simulated from a parametric
90 model does have its advantages, it also has certain limitations. One potential disadvantage is
91 that the results are heavily dependent on the assumptions of the parametric model. If these
92 assumptions do not hold, the simulation results may not accurately reflect the real-world scenario.

93 We have put one paragraph in the discussion section that Our study employs simulation to
94 evaluate the effectiveness of GWR in estimating coefficients for both randomised and systematic
95 designs. However, a potential drawback of using simulation is its high dependency on the
96 assumptions underlying the parametric model. If these assumptions are incorrect or fail to
97 accurately represent real-world conditions, the simulation results may not reliably reflect the
98 actual scenario.

99 Additionally, we have revised Section 3, starting with a better description of the data source
100 and why the parameters are chosen at this level. The simulated data aligns with the authentic
101 data from a corn field in Las Rosa. This data has been studied and reported by Rakshit et al.
102 (2020) and Cao et al. (2022). The assumption is based on their findings.

103 5. Also at the start of Section 3, where factors studied in the simulation are listed (from line
104 156), it is not immediately obvious, if the description of details of the simulations refers to the
105 data simulation or the analysis of simulated data. Clarification would be useful. At this point, it
106 is also still not clear which model was used for simulating the data and which for analysing the
107 data. This only becomes clearer gradually downstream in the paper. I feel it could be made very
108 clear much earlier in the paper.

109 We have revised Section 3 for a better and clearer explanation.

110 6. It may help greatly if the statement of the objectives not made in line 190 comes much
111 earlier.

112 We have put a new sentence in the introduction section in line 90: We subsequently evaluate
113 the efficacy of GWR in accurately estimating the spatially varying treatment effects across these
114 scenarios.

115 7. The MSE, used for evaluating predictions, should be described in the section on simulation,
116 as it is part of the methods.

117 Thank you for the comment. We agree and have added a new sub-section 2.3 Performance
118 evaluation to describe MSE..

119 8. Line 168: Is the 75 for the mid N level a typo? Should this be 70? Why are the levels
120 different from those in Figure 1?

121 Thanks for pointing out the discrepancy. We have revised the mistake by updating the labels
122 in the figures and also in the main context:

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

125 In line 72: We used five evenly spaced nitrogen rates: 0, 35, 70, 105, and 140 kg/ha.

126 Reviewer 2 writes:

127 The paper presents results from a simulation study to compare two experimental designs, a
128 randomized and a systematic. In my opinion, the subject should be of moderate interest to readers

129 of the journal. The paper needs some work to be considered for publication in FCR, assuming that
130 such technical paper is of interest to most readers of the journal. In its current form I believe it
131 is too technical with a lot of statistical jargon. Also the organization and flow of the information
132 needs to be improved. I made some specific comments about my concerns and on how the authors
133 can improve the manuscript.

134 The general suggestion is to revise the text without so much statistical jargon or add explanations
135 that will make it easier for agronomists to follow/understand.

136 We have revised the main text of the manuscript to enhance its readability and comprehension.
137 Additionally, we have streamlined the statistical model sections for simplicity and clarity.

138 *Specific comments: 1) Line 40. Please define OFE here and not in line 46 and 57*

139 We have now defined OFE when it first appears in the main text in line 45. The revised
140 version is **On-farm experiment (OFE)** enables farmers the flexibility

141 *2) Line 71. Constraints in the field.*

142 Thanks for pointing out the error. We have changed it to *constraints in the field*.

143 *3) Lines 73-74. I do not see the value of that sentence.*

144 We have polished this paragraph. The purpose of this sentence is to introduce that a linear
145 model may not be sufficient in capturing the yield curve. That's why we compared linear and
146 quadratic curve in this study.

147 *4) Lines 81-83. These are results and do not fit in the introduction.*

148 We have rewritten this paragraph to better fit the manuscript by: In this study, we generate
149 simulated data for several scenarios, where each scenario is constructed by choosing one component
150 at a time from the following four categories: (i) randomised and systematic designs; (ii) linear
151 and quadratic responses; (iii) model coefficients with low and high correlations; and (iv) spatial
152 variance-covariance matrix 90 among grids given by identity (no spatial trend), AR1 \otimes AR1,
153 and Matérn forms. We subsequently evaluate the efficacy of GWR in accurately estimating the
154 spatially varying treatment effects across these scenarios.

155 *5) Lines 84-89. There is no need to describe the structure of the paper. All papers have
156 introduction, methods, results and discussion. I suggest to use the last paragraph to describe the
157 objectives of the paper and why is important.*

158 We have removed this paragraph. Instead, we put one more sentence here: The GWR model
159 in this paper is implemented with the R-package GWmodel (Lu et al. 2014; Gollini et al. 2015).

160 *6) Lines 91-93. Again there is no need for these three lines.*

161 We have removed these lines and re-structured this section by merging subsections 2.1 and
162 2.2, and inserting a new subsection 2.3. The revised subsection titles are:

163 2.1 Hierarchical model for generating simulated data. 2.2 Fitting geographically weighted
164 regression to simulated data. 2.3 Performance evaluation.

165 *7) 3 Simulation study. More information is needed here. I suggest starting from describing
166 the simulated experiment. For example, I had to read up to line 160 and again after line 166 to
167 understand that the independent variable of interest would be nitrogen fertilizer with 5 levels.*

168 We have rewritten the simulation section by reordering the paragraphs and making a clearer
169 explanation.

170 *8) Line 172. For someone who is not familiar with systematic design, they would have to
171 read up to line 172 and see figure 1 to understand the difference. I suggest describing what a
172 systematic design is early in the introduction so every reader understands what is being compared.*

173 We firstly in line 39 inserted the sentence where treatments are randomly allocated in each
174 replicate for a randomised design.

175 In line 41, we rewrote the sentence as follows: With the primary aim of obtaining unbiased
176 estimates of global treatment effects, randomised designs, which use different layouts of treatments

177 in each replicate, are routinely used for on-farm strip trials, whereas systematic designs, which
178 use the same layout of treatments in all replicates, are rarely used.

179 In line 174, we try to remind the reader by using the following description that Examples of a
180 randomised design, which uses different orders of treatment in each replicate, and a systematic
181 design, which uses the same order of treatment in all replicates, for this layout are presented in
182 Figure 1.

183 9) 4.1 Mean squared error. I recommend to put the figures in order. Now after figure 3 we
184 jump to figure 6 and then figure 4 followed by figure 7.

185 Thank you for your suggestion. We have organised the figures in order and updated them
186 with enhanced visualisation.

187 10) For more practical purposes, apart from comparing MSE and coefficients, it would be
188 valuable to see two maps with optimal nitrogen rate side-by-side. One map generated by the
189 randomized design and one generated by the systematic design. I suspect that for practical purposes
190 maps will be very similar.

191 It's a good point. From the practical point of view, that's the most interests output for
192 growers. So, we added one more sub-section at the end of section 4: 4.3 An example of optimal
193 nitrogen map with the description of the results and two optimal nitrogen maps.

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

199 Figure 1 is the simulated crop yield map with the assumption of a quadratic response curve
200 and Matérn spatial covariance and low within-grid correlation. These two yield maps have the
201 same coefficients but different yields due to different treatment layouts.

202 Figure 2 illustrates an example of the optimal Nitrogen rate (kg/ha) map estimated by
203 GWR with a bandwidth of 9 using the above yield data. The optimal rate at grid i is given by
204 $\hat{N}_i = -\hat{\beta}_{1i}/(2\hat{\beta}_{2i})$ with constraints between 0 and 140, $i = 1, \dots, n$. For the randomised design,
205 GWR underestimated the right part of the paddock. On the contrary, the estimated map from
206 the systematic design is more consistent.

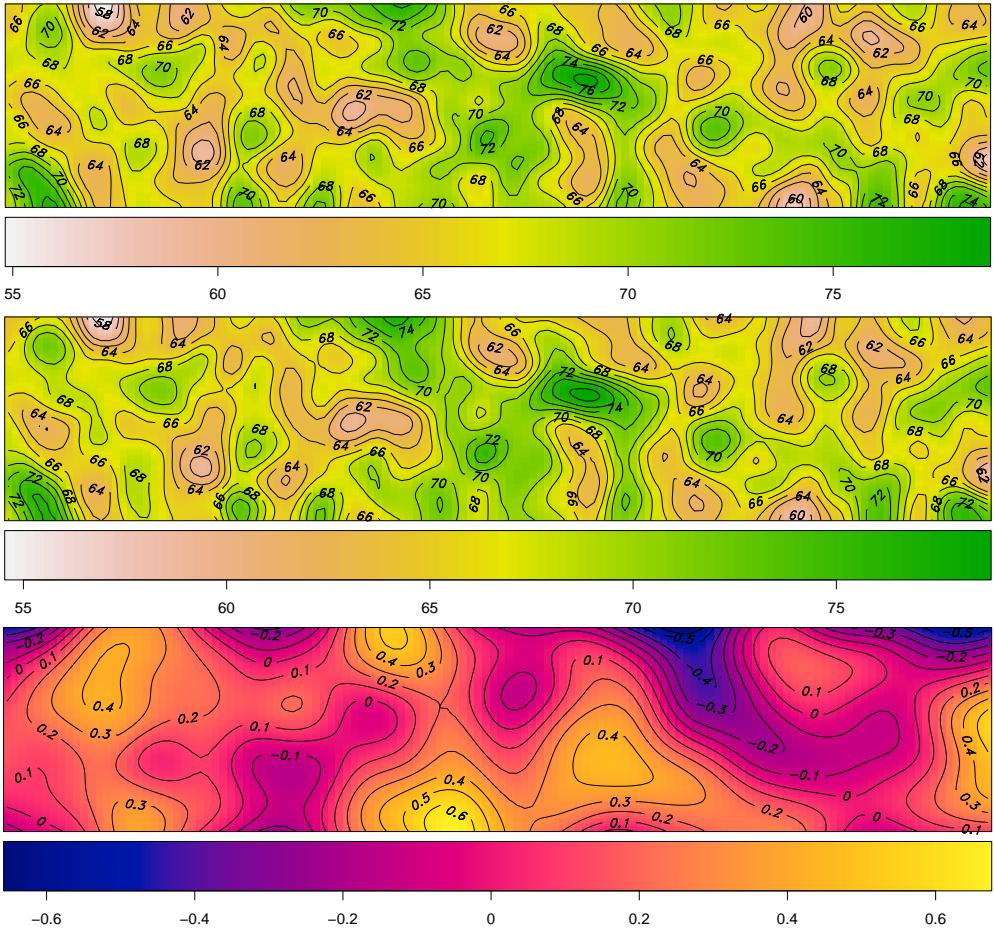


Figure 1: Simulated yield map of a randomised design (top), and a systematic design (middle), and the difference of these two designs (bottom).

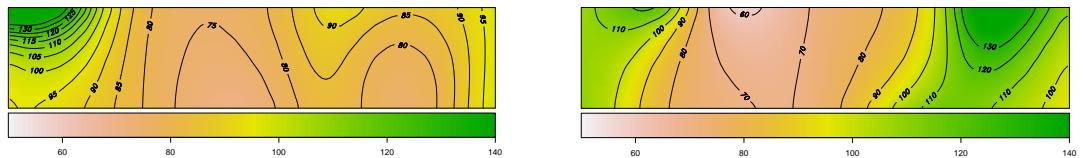


Figure 2: Optimal Nitrogen rate estimated by GWR from a randomised design (left), and a systematic design (right).

Highlights

- Simulation of diverse scenarios generated by Bayesian hierarchical model
- Systematic designs are preferred for producing continuous input maps
- The presence of spatial variation affects the performance of GWR
- A fixed bandwidth based on the experiment is recommended for GWR

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

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

2 Zhanglong Cao^{1,*}, Jordan Brown¹, Mark Gibberd¹, Julia Easton¹, and Suman Rakshit^{1,2}

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

5 ²School of Electrical Engineering, Computing and Mathematical Sciences, Curtin University, Perth,
6 Australia

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

8 Abstract

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

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

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

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

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

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

32 **Keywords:** yield map, optimal treatment, spatially varying coefficients, geographically weighted regression, pre-
33 cision agriculture.

34 1 Introduction

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

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

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

63 Analysing a systematic design for the creation of an optimal treatment map is a statistically challenging task. The
64 true responses at each point on the grid corresponding to all the treatment rates are unknown, and the treatment
65 producing the optimum response may vary continuously across the field. [Cao et al. \(2022\)](#) implemented a Bayesian
66 approach with spatially correlated random parameters for analysing large systematic strip trials. These authors
67 considered a quadratic response model with both global and local (spatially-varying) components. However, Bayesian
68 analysis can be computationally expensive and would require at least preliminary knowledge of Bayesian inference to
69 interpret the results, which can be extremely demanding for farmers and agronomists. Alternatively, [Rakshit et al.
70 \(2020\)](#) adopted a local regression approach, called geographically weighted regression (GWR), to obtain spatially-

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

80 Piepho and Edmondson (2018) presented an example where a linear model turns out to be inadequate for analysing
 81 sugar beet data (Petersen 1994). Glynn (2007) showed that many curves exist beyond a linear trend for nutrient-
 82 response relationships. The response curve often depends on the availability of other macro and micronutrients in the
 83 soil (Marschner 2011), which means that a linear relationship is unlikely to be consistent across a large trial. For this
 84 reason, it is important to consider models with terms of order higher than unity. For example, a quadratic model can
 85 often found to be suitable for modelling nutrient-response relationships (Piepho and Edmondson 2018; Liben et al.
 86 2019).

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

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

93 2 Methods

94 2.1 Hierarchical model for generating simulated data

95 In a conventional agricultural study, a field experiment can be considered as a rectangular matrix, representing
 96 a regular grid with r rows and c columns, where the total number of plots in the experiment is $n = r \times c$. Let
 97 $s_i \in \mathcal{R}^2, i = 1, \dots, n$, denote the Cartesian coordinate of the i -th plot centroid, located on a regular grid (Zimmerman
 98 and Harville 1991). Let $y(s_i), i = 1, \dots, n$, denote the value of the dependent variable recorded at the i -th plot.

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

$$\mathbf{Y} = \mathbf{X}\mathbf{b} + \mathbf{Z}\mathbf{u} + \mathbf{e}, \quad (1)$$

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

103 distribution is

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

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

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

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

108 In model (3), the structure of the covariance matrix $V_u(\theta_u)$ of \mathbf{u}_i can be either diagonal, which implies the random
109 terms at grid i are independent, or in general form, which means a correlation exists. McElreath (2015) suggested
110 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
111 diagonal elements σ_{uj} , $j = 1, \dots, k$, and R_u denotes the matrix with correlation coefficients. For the matrix R_u , we
112 specify the Lewandowski-Kurowicka-Joe (LKJ) distribution (Lewandowski et al. 2009), which is given by

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

113 where $\text{LKJcorr}(\epsilon)$ is a positive definite correlation matrix sampled from the LKJ distribution controlled by a positive
114 parameter ϵ . As ϵ increases, a high correlation among parameters becomes less likely.

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

$$\Sigma_u = V_s \otimes R_u. \quad (5)$$

117 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
118 random terms at a grid point are independently distributed from those at other grid points, even though the terms
119 at that grid point are correlated amongst each other. However, the correlation among grids is ubiquitous. Hence, we
120 introduce a simple spatial covariance matrix such as

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

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

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

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

124 where d is the space lag or distance; γ is a non-negative scaling parameter; $\nu > 0$ is a smoothness parameter; σ^2 is
 125 the variance parameter; Γ is the Gamma function; and K_ν is the modified Bessel function of the second kind. The
 126 Matérn covariance is commonly used in the analysis of geostatistical data (Cressie and Huang 1999). Moreover, it has
 127 also been used in capturing spatial variation in OFE (Selle et al. 2019). If $\nu = \gamma + 1/2$, then the Matérn covariance
 128 can be expressed as a product of an exponential and a polynomial of order γ (Pandit and Infield 2019; Abramowitz
 129 1974), which simplifies the model and the computation process. The Matérn models with the values 3/2 and 5/2 for
 130 ν are used in most applications.

131 In each iteration of the simulation process, we used the above formulae and pre-defined parameter values to generate
 132 $2 \times n$ coefficients for the linear response and $3 \times n$ coefficients for the quadratic response. Then these coefficients are
 133 applied to simulate yield response across strips for both randomised and systematic design layouts.

134 2.2 Fitting geographically weighted regression to simulated data

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

139 The underlying template model for GWR is given by

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

140 where $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_k)^\top$ are model parameters corresponding to the k treatment levels and $\varepsilon_i \sim \mathcal{N}(0, \tau^2)$, $i =$
 141 $1, \dots, n$, are independent and identically distributed error terms at n grid points.

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

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

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

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

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

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

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

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

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

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

155 2.3 Performance evaluation

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

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

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

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

163 3 Simulation study

164 The simulation study is performed using realistic values for the model parameters, which are selected based on
165 the analysis results of a real-life data recorded from a corn field in Las Rosas, Argentina. This data set was originally
166 provided by Anselin et al. (2004) and can be accessed via the R-package `agridat` (White and Evert 2008). In 2001,
167 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
168 replicates. Each replicate consists of 93 rows and 6 columns after data preprocessing; see Rakshit et al. (2020) for
169 further details about the preprocessing steps. The unit of yield is quintals per hectare.

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

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

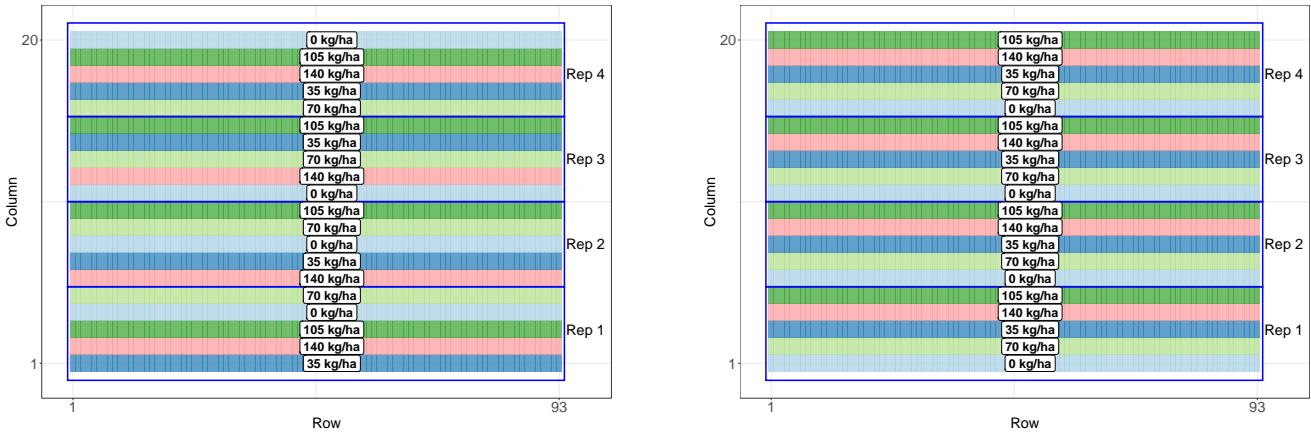


Figure 1: The nitrogen treatments with five levels (0, 35, 70, 105 and 140 kg/ha) randomly (*left*) and systematically (*right*) allocated into strips in each replicate block.

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

189 To specify the linear relationship in the simulation study, we consider the values of 65 and 0.05 for the global
 190 intercept b_0 and slope b_1 coefficients, respectively, in model (3). The variances of the coefficients \mathbf{u}_i are set to 5 for
 191 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
 192 $\text{AR1} \otimes \text{AR1}$ covariance matrix in (6), the two correlation parameters ρ_c and ρ_r are set to 0.15 and 0.50, respectively.
 193 We assume a higher correlation in the row direction because the crop is traditionally sown and harvested along the
 194 column direction, and the correlation is higher in the direction perpendicular to the sowing direction (Marchant et al.
 195 2019). For the Matérn covariance matrix (7), we set the value of the variance parameter σ_d^2 to 1, the value of the
 196 parameter r to 1, and the value of the parameter ν to 3/2. After drawing samples of \mathbf{u} from $\mathcal{N}(0, \Sigma_u)$, the spatially
 197 varying coefficients β_0 and β_1 are specified using the relations $\beta_0 = b_0 + \mathbf{u}_0$ and $\beta_1 = b_1 + \mathbf{u}_1$.

198 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 ,
 199 respectively. These choices make the response curve concave down. For the variance components, we set to 5 for σ_{u_0} ,
 200 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
 201 coefficients are $\beta_0 = b_0 + \mathbf{u}_0$, $\beta_1 = b_1 + \mathbf{u}_1$, and $\beta_2 = b_2 + \mathbf{u}_2$ for the quadratic model.

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

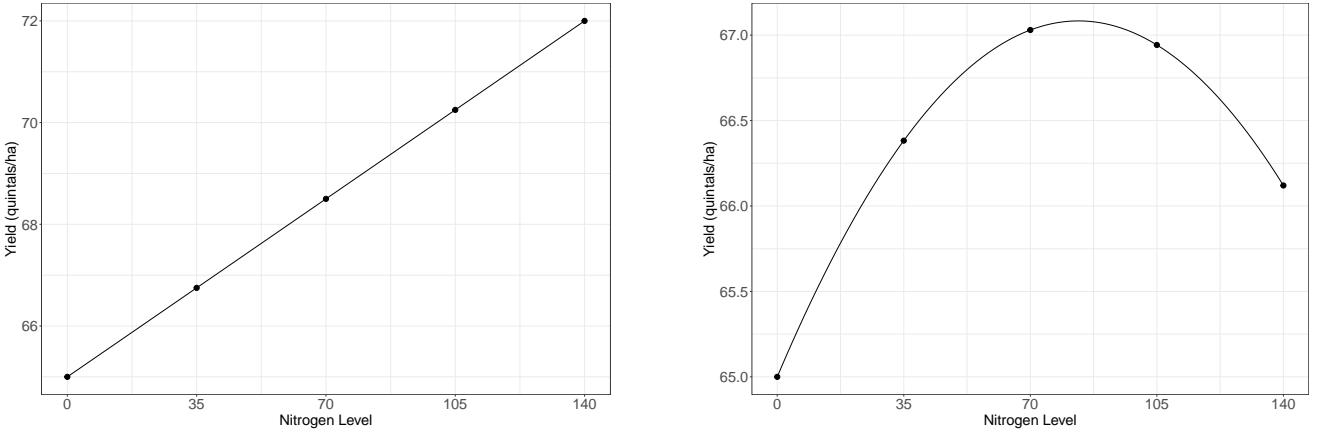


Figure 2: The global linear relationship of yield and nitrogen is $y = 65 + 0.05N$ (left), and the global quadratic relationship between yield and nitrogen is $y = 65 + 0.05N - 0.003N^2$ (right).

203 To summarise, the simulated yield response is obtained by

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

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

205 4 Results

206 In this section, we assess the performances of randomised and systematic designs in terms of their utility to
 207 accurately estimate the model parameters for both linear and quadratic response models. To this end, we perform
 208 1000 simulations of each of the 24 scenarios, described in the previous section. In each simulation, we first generate
 209 the coefficients for all grids and then apply the treatment in each grid to calculate the yield value using the model
 210 coefficients. For the treatment order in each iteration, we randomly picked an order of treatments for a single replicate
 211 and repeated this sequence across all other replicates to construct a systematic design. For a randomised design, all
 212 replicates have random orders of the treatments.

213 4.1 Comparison based on mean squared errors

214 Figures 3 and 4 show the results of linear models for the cases of low ($\epsilon = 1$) and high ($\epsilon = 0.1$) correlations,
 215 respectively, while Figures 5 and 6 show the results of quadratic models for the same low and high correlations,
 216 respectively. To specify the covariance matrix V_s used in producing the results in these figures, we use the following
 217 labels: (i) “NS” for the identity matrix representing no spatial correlation, (ii) “AR1” for $AR1(0.15) \otimes AR1(0.5)$, and
 218 (iii) “Matérn” for Matérn covariance with $\nu = 3/2$. Note that the model parameters and their corresponding MSEs
 219 are small values, and this makes it difficult to compare the MSEs of different scenarios using the original scale of MSE
 220 values. Therefore, to enhance clarity in visualisation and comparison, we have multiplied the MSEs of β_1 and β_2 by
 221 10^4 and 10^8 , respectively, in the figures presented below.

222 For the linear response model, both randomised and systematic designs perform similarly, particularly for the
 223 case NS. Figure 3 shows that the MSE corresponding to $\hat{\beta}_0$ 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 $\text{AR1} \otimes \text{AR1}$ and Matérn cases are lower for the systematic design.

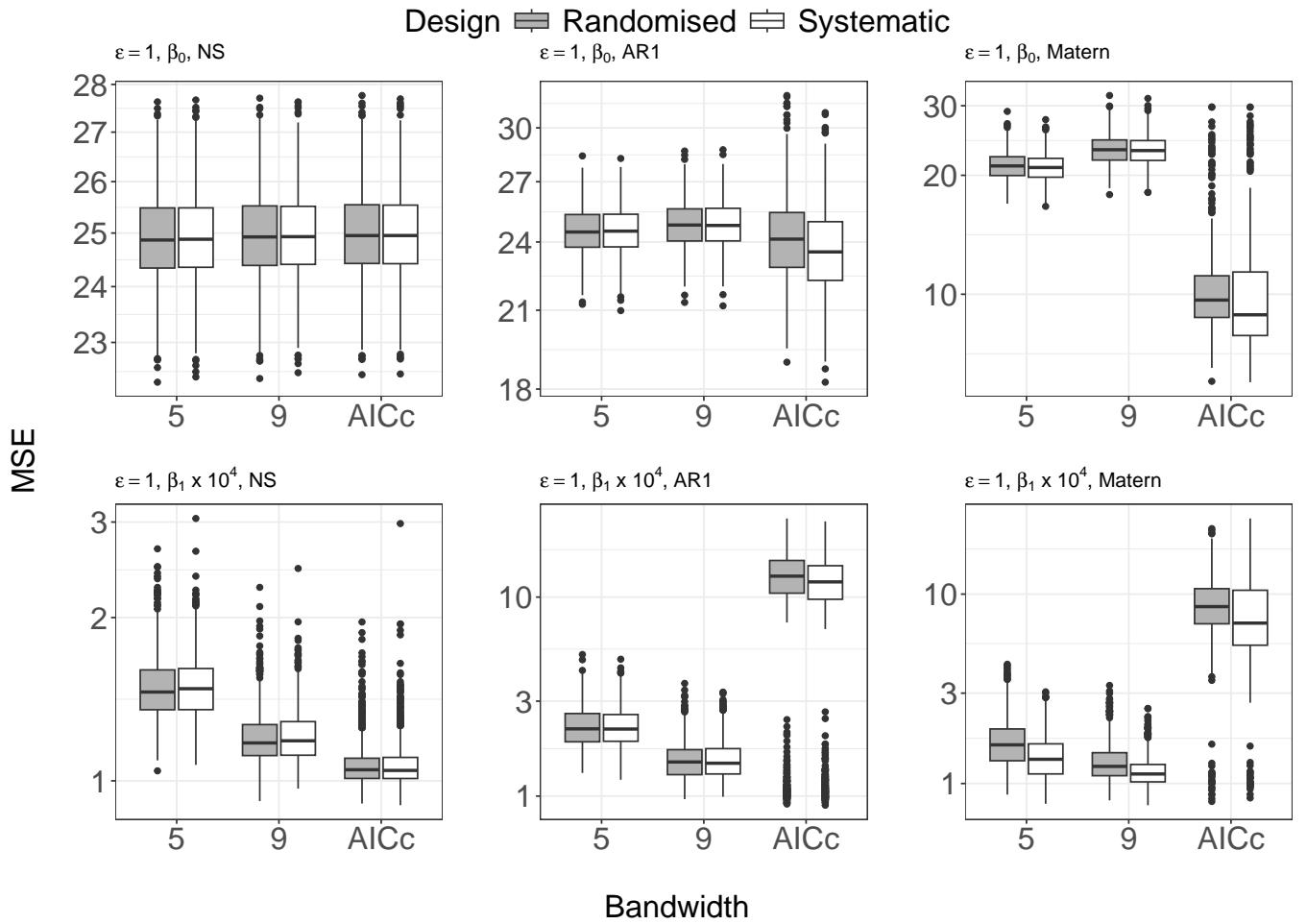


Figure 3: Boxplots of MSE for $\hat{\beta}_0$ and $\hat{\beta}_1$ in GWR models using different bandwidths for the simulated data with a linear response. The simulated data had different spatial covariance matrices (NS, $\text{AR1} \otimes \text{AR1}$ and Matérn) and a low correlation between the parameters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

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 $\text{AR1} \otimes \text{AR1}$ or Matérn covariance structures), the bandwidth of 9 consistently produces the most accurate estimates of β_1 and β_2 ,

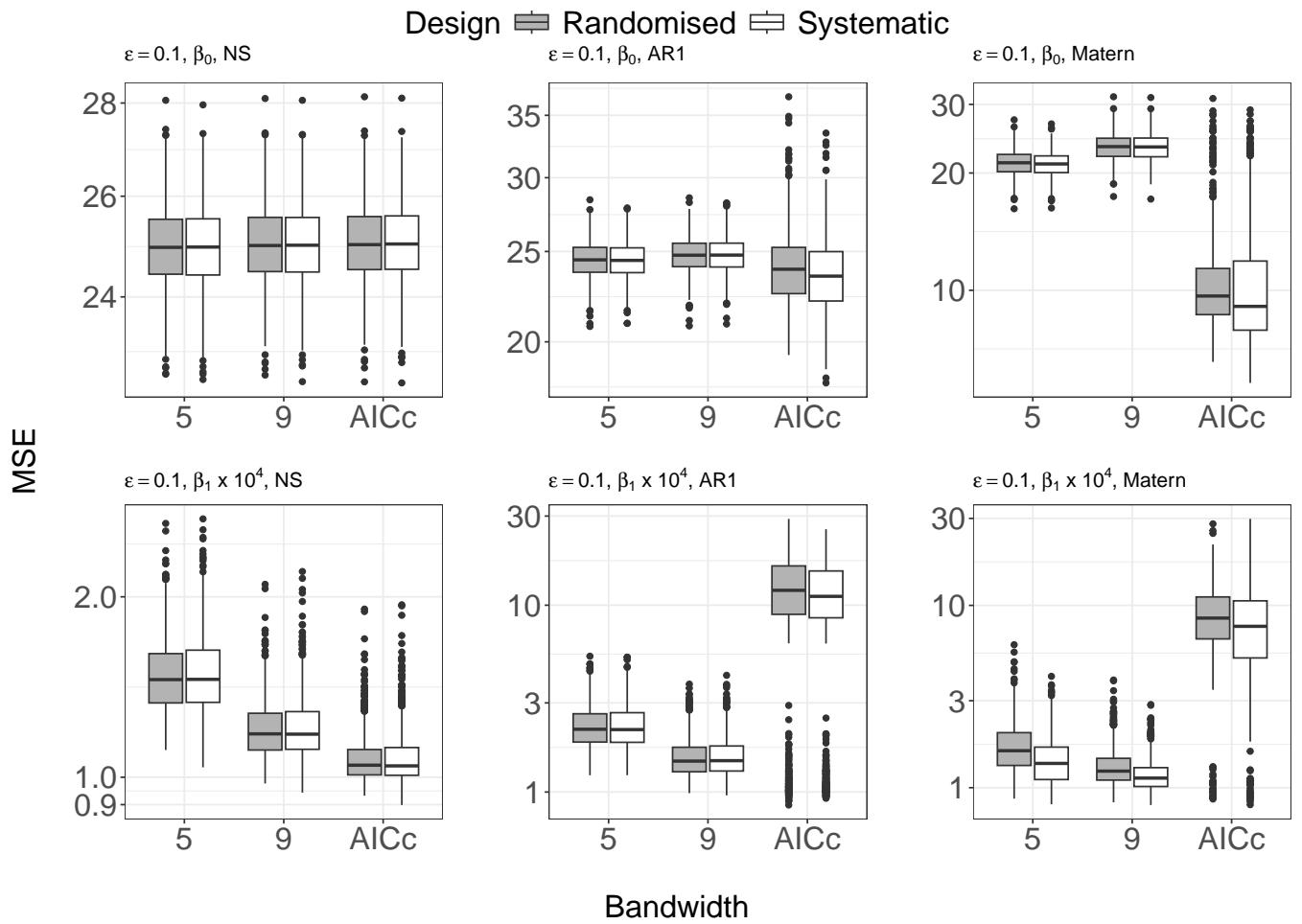


Figure 4: Boxplots of MSE for $\hat{\beta}_0$ and $\hat{\beta}_1$ in GWR models using different bandwidths for the simulated data with a linear response. The simulated data had different spatial covariance matrices (NS, AR1 \otimes AR1 and Matérn) and a high correlation between the parameters ($\epsilon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

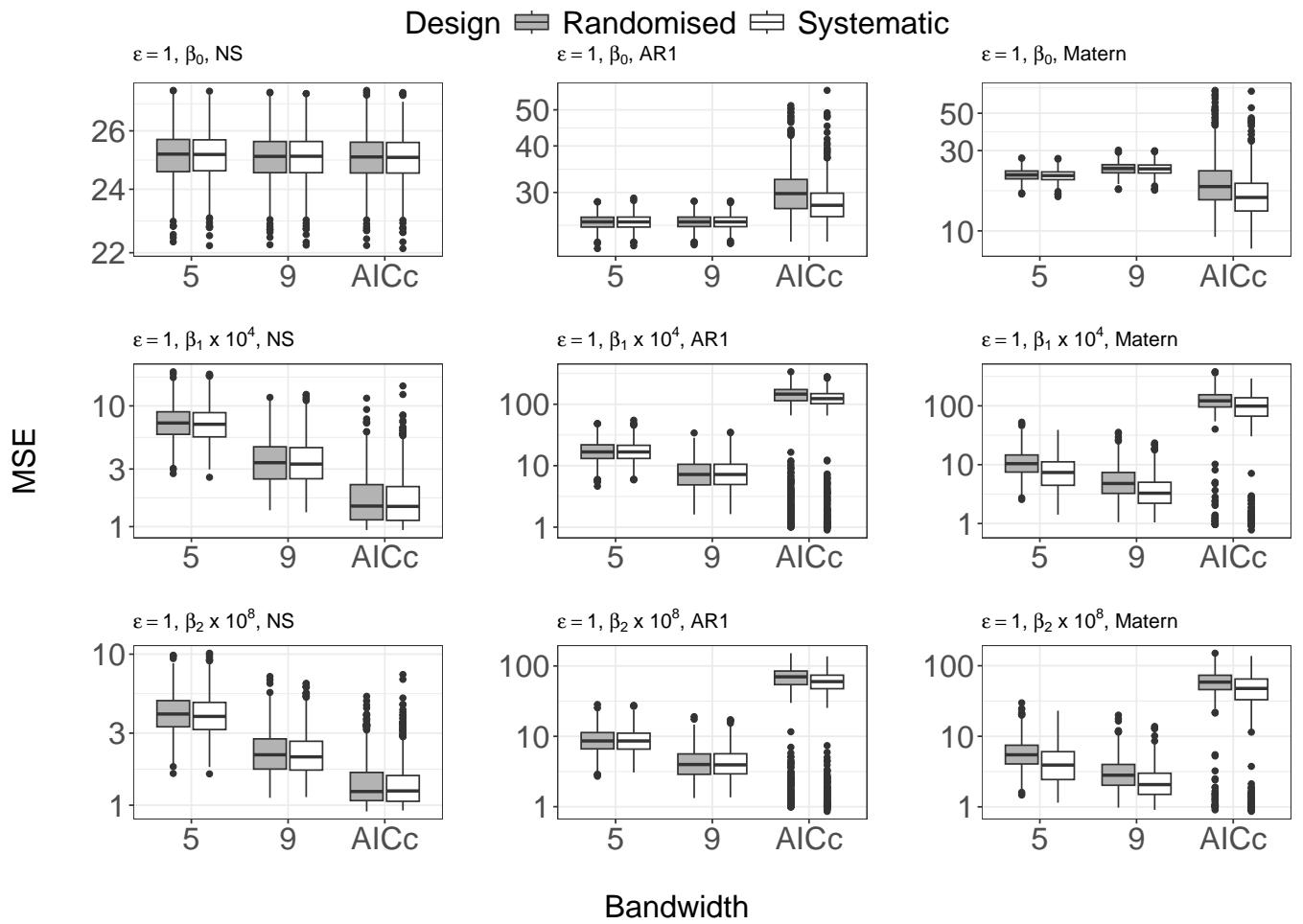


Figure 5: Boxplots of MSE for $\hat{\beta}_0$, $\hat{\beta}_1$ and $\hat{\beta}_2$ in GWR models using different bandwidths for the simulated data with a quadratic response. The simulated data had different spatial covariance matrices (NS, AR1 \otimes AR1 and Matérn) and a low correlation amongst the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

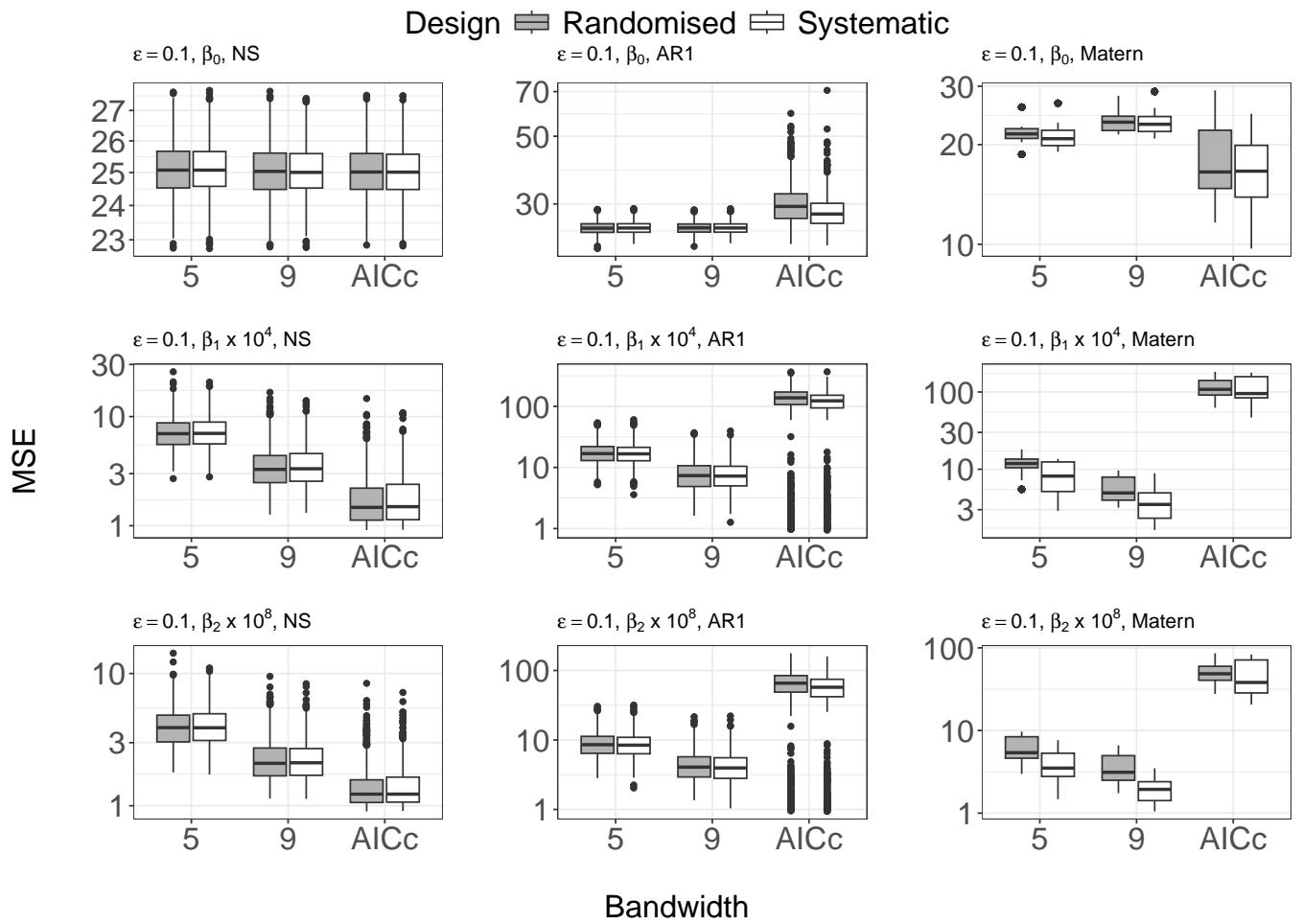


Figure 6: Boxplots of MSE for $\hat{\beta}_0$, $\hat{\beta}_1$ and $\hat{\beta}_2$ in GWR models using different bandwidths for the simulated data with a quadratic response. The simulated data had different spatial covariance matrices (NS, AR1 \otimes AR1 and Matérn) and a high correlation amongst the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

²⁴¹ outperforming the estimates obtained using either the bandwidth of 5 or the one chosen by AICc.
²⁴² Tables 1 and 2 present the median MSEs corresponding to the parameter estimation for the linear response model
²⁴³ 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	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	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 \otimes 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.

²⁴⁴ Tables 3 and 4 present the median MSEs corresponding to the estimation of quadratic response models for the two
²⁴⁵ 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	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 \otimes 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 \otimes 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.

246 4.2 Comparison of density plot

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

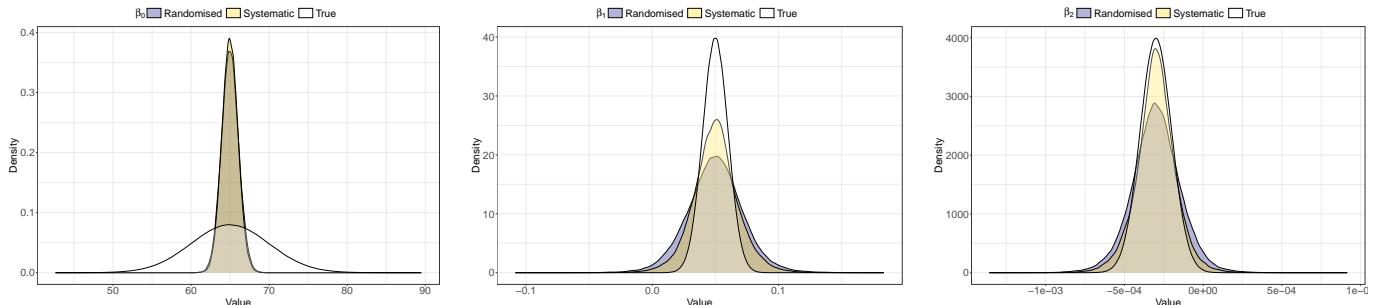


Figure 7: Density plots comparing the true coefficients of β_0 (left), β_1 (middle), and β_2 (right) with their estimates from GWR with bandwidth 9, based on 1000 simulations.

252 4.3 An example of optimal nitrogen map

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

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

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

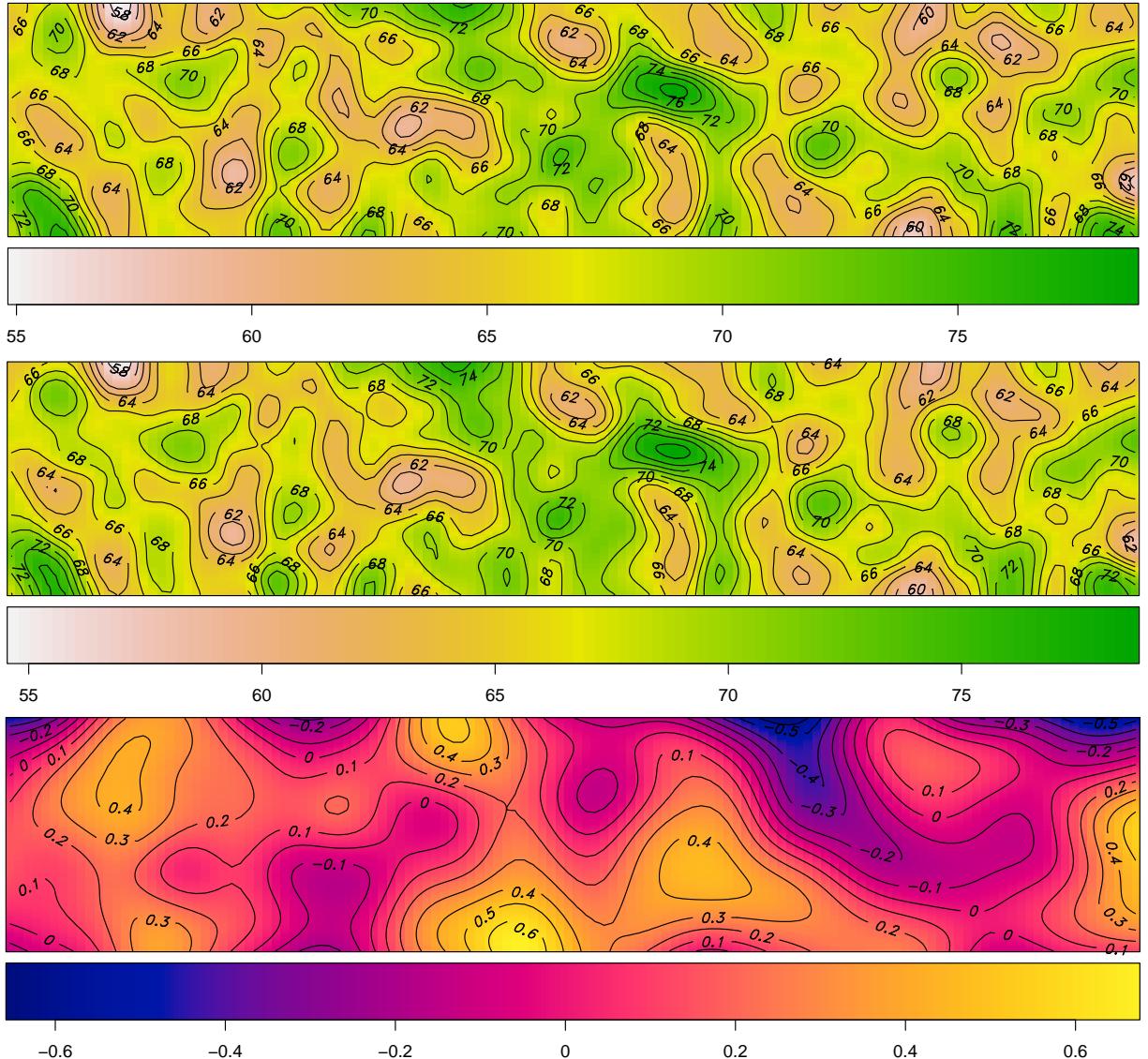


Figure 8: Simulated yield map of a randomised design (*top*), and a systematic design (*middle*), and the difference of these two designs (*bottom*).

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
 261 140, $i = 1, \dots, n$. For the randomised design, GWR underestimated the right part of the paddock. On the contrary,
 262 the estimated map from the systematic design is more consistent.
 263

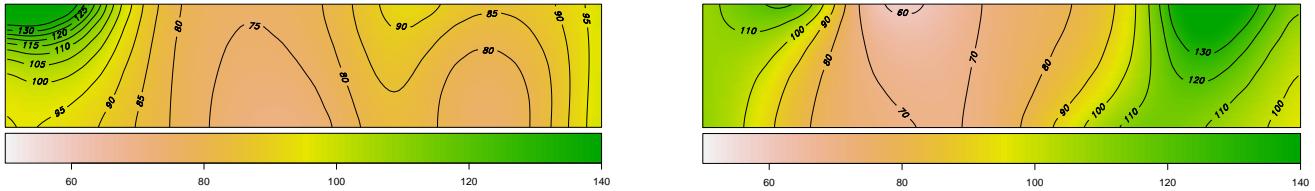


Figure 9: Optimal Nitrogen rate estimated by GWR from a randomised design (*left*), and a systematic design (*right*).

5 Discussion

264 Agronomists and biometricalians generally prefer randomised designs for OFE trials. This is likely due to their
 265 experience with small-plot field experiments, where randomised designs are employed. Our simulation study shows a
 266 systematic design performs either preferably or similarly to a randomised design for the purposes of creating a varying
 267 treatment map when using the performance metrics stated. The differentiating factors primarily included the response
 268 type and the spatial covariance model, while the correlation amongst the treatment coefficients is not found to be
 269 important. These are factors that can be assessed by the farmer and their agronomists beforehand, and this should
 270 dictate which design should be used. Given that a systematic design is easier to implement in the field, and shows
 271 little downside to creating a varying treatment map, the use of systematic designs is recommended.
 272

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

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

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

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

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

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

306 6 Conclusion

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

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

313 Acknowledgements

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

316 CRedit authorship contribution statement

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

323 References

- 324 Abramowitz, M. (1974). *Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables*.
325 USA: Dover Publications, Inc. (cit. on p. 5).
- 326 Alesso, C. A., Cipriotti, P. A., Bollero, G. A., and Martin, N. F. (2021). “Design of on-farm precision experiments to
327 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).
- 328 Anselin, L., Bongiovanni, R., and Lowenberg-DeBoer, J. (2004). “A Spatial Econometric Approach to the Economics
329 of Site-Specific Nitrogen Management in Corn Production”. In: *American Journal of Agricultural Economics* 86.3,
330 pp. 675–687. DOI: <https://doi.org/10.1111/j.0002-9092.2004.00610.x> (cit. on p. 6).
- 331 Bramley, R., Cook, S., Adams, M., and Corner, R. (1999). “Designing your own on-farm experiments: How precision
332 agriculture can help”. In: (cit. on p. 17).
- 333 Butler, D., Cullis, B., Gilmour, A., Gogel, B., and Thompson, R. (2017). *ASReml-R Reference Manual Version 4*
334 (cit. on p. 4).
- 335 Cao, Z., Stefanova, K., Gibberd, M., and Rakshit, S. (2022). “Bayesian inference of spatially correlated random
336 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).
- 337 Cook, S., Cock, J., Oberthür, T., and Fisher, M. (2013). “On-Farm Experimentation”. In: *Better Crop. Plant Food.*
338 4th ser. 97, pp. 17–20 (cit. on p. 2).
- 339 Cressie, N. and Huang, H.-C. (Dec. 1999). “Classes of Nonseparable, Spatio-Temporal Stationary Covariance Func-
340 tions”. In: *Journal of the American Statistical Association* 94.448, pp. 1330–1339. DOI: [10.1080/01621459.1999.10473885](https://doi.org/10.1080/01621459.1999.10473885) (cit. on p. 5).
- 341 Evans, F. H., Recalde Salas, A., Rakshit, S., Scanlan, C. A., and Cook, S. E. (Nov. 2020). “Assessment of the Use
342 of Geographically Weighted Regression for Analysis of Large On-Farm Experiments and Implications for Practical
343 Application”. en. In: *Agronomy* 10.11, p. 1720. DOI: [10.3390/agronomy10111720](https://doi.org/10.3390/agronomy10111720) (cit. on pp. 2, 3, 6).
- 344 Fisher, R. A. (1934). *Statistical Methods for Research Workers*. English. fifth. Edinburgh: Oliver and Boyd (cit. on
345 p. 2).
- 346 Glynn, C. (2007). “Testing the Growth-Differentiation Balance Hypothesis: Dynamic Responses of Willows to Nutrient
347 Availability”. In: *New Phytologist* 176.3, pp. 623–634. DOI: [10.1111/j.1469-8137.2007.02203.x](https://doi.org/10.1111/j.1469-8137.2007.02203.x) (cit. on p. 3).
- 348 Gollini, I., Lu, B., Charlton, M., Brunsdon, C., and Harris, P. (Feb. 2015). “GWmodel: An R Package for Exploring
349 Spatial Heterogeneity Using Geographically Weighted Models”. en. In: *Journal of Statistical Software* 63.1, pp. 1–
350 50. DOI: [10.18637/jss.v063.i17](https://doi.org/10.18637/jss.v063.i17) (cit. on pp. 3, 5).
- 351 Lewandowski, D., Kurowicka, D., and Joe, H. (2009). “Generating random correlation matrices based on vines and
352 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).
- 353 Liben, F. M., Midega, T., Tufa, T., and Wortmann, C. S. (2019). “Soil Fertility & Crop Nutrition Barley and wheat
354 nutrient responses for Shewa, Ethiopia”. In: *Agronomy Journal*. DOI: [10.1002/agj2.20020](https://doi.org/10.1002/agj2.20020) (cit. on p. 3).

- 359 Lu, B., Harris, P., Charlton, M., and Brunsdon, C. (2014). “The GWmodel R package: further topics for exploring
360 spatial heterogeneity using geographically weighted models”. In: *Geo-spatial Information Science* 17.2, pp. 85–101.
361 DOI: [10.1080/10095020.2014.917453](https://doi.org/10.1080/10095020.2014.917453) (cit. on p. 3).
- 362 Marchant, B., Rudolph, S., Roques, S., Kindred, D., Gillingham, V., Welham, S., Coleman, C., and Sylvester-Bradley,
363 R. (Jan. 2019). “Establishing the Precision and Robustness of Farmers’ Crop Experiments”. en. In: *Field Crops
364 Research* 230, pp. 31–45. DOI: [10.1016/j.fcr.2018.10.006](https://doi.org/10.1016/j.fcr.2018.10.006) (cit. on p. 7).
- 365 Marschner, H. (2011). *Marschner’s Mineral Nutrition of Higher Plants*. San Diego: San Diego: Elsevier Science &
366 Technology. DOI: [10.1016/C2009-0-63043-9](https://doi.org/10.1016/C2009-0-63043-9) (cit. on p. 3).
- 367 McElreath, R. (Dec. 2015). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. English. First.
368 Vol. 122. Chapman and Hall/CRC Texts in Statistical Science Ser. CRC Press LLC (cit. on p. 4).
- 369 Páez, A., Uchida, T., and Miyamoto, K. (Apr. 2002). “A General Framework for Estimation and Inference of Geo-
370 graphically Weighted Regression Models: 1. Location-Specific Kernel Bandwidths and a Test for Locational Hetero-
371 geneity”. In: *Environment and Planning A: Economy and Space* 34.4, pp. 733–754. DOI: [10.1068/a34110](https://doi.org/10.1068/a34110) (cit. on
372 p. 5).
- 373 Pandit, R. K. and Infield, D. (Sept. 2019). “Comparative Analysis of Gaussian Process Power Curve Models Based
374 on Different Stationary Covariance Functions for the Purpose of Improving Model Accuracy”. en. In: *Renewable
375 Energy* 140, pp. 190–202. DOI: [10.1016/j.renene.2019.03.047](https://doi.org/10.1016/j.renene.2019.03.047) (cit. on p. 5).
- 376 Petersen, R. G. (1994). *Agricultural Field Experiments: Design and Analysis*. en. 1st. Boca Raton: CRC Press (cit. on
377 pp. 2, 3).
- 378 Piepho, H. P. and Edmondson, R. N. (2018). “A tutorial on the statistical analysis of factorial experiments with
379 qualitative and quantitative treatment factor levels”. In: *Journal of Agronomy and Crop Science* 204.5, pp. 429–
380 455. DOI: <https://doi.org/10.1111/jac.12267> (cit. on p. 3).
- 381 Piepho, H. P., Möhring, J., and Williams, E. R. (2013). “Why Randomize Agricultural Experiments?” In: *Journal of
382 Agronomy and Crop Science* 199.5, pp. 374–383. DOI: [10.1111/jac.12026](https://doi.org/10.1111/jac.12026) (cit. on p. 2).
- 383 Pringle, M. J., Cook, S. E., and McBratney, A. B. (Dec. 2004). “Field-Scale Experiments for Site-Specific Crop
384 Management. Part I: Design Considerations”. en. In: *Precision Agriculture* 5.6, pp. 617–624. DOI: [10.1007/s11119-004-6346-1](https://doi.org/10.1007/s11119-004-6346-1) (cit. on pp. 2, 3).
- 386 Rakshit, S., Baddeley, A., Stefanova, K., Reeves, K., Chen, K., Cao, Z., Evans, F., and Gibberd, M. (2020). “Novel Ap-
387 proach to the Analysis of Spatially-Varying Treatment Effects in on-Farm Experiments”. In: *Field Crops Research*
388 255.October 2019, p. 107783. DOI: [10/gg2vv7](https://doi.org/10/gg2vv7) (cit. on pp. 2, 5, 6, 16).
- 389 Selle, M. L., Steinsland, I., Hickey, J. M., and Gorjanc, G. (2019). “Flexible Modelling of Spatial Variation in Agri-
390 cultural Field Trials with the R Package INLA”. In: *Theoretical and Applied Genetics* 132.12, pp. 3277–3293. DOI:
391 DOI: [10.1007/s00122-019-03424-y](https://doi.org/10.1007/s00122-019-03424-y) (cit. on pp. 2, 5).
- 392 Verdooren, L. R. (Dec. 2020). “History of the Statistical Design of Agricultural Experiments”. en. In: *Journal of
393 Agricultural, Biological and Environmental Statistics* 25.4, pp. 457–486. DOI: [10.1007/s13253-020-00394-3](https://doi.org/10.1007/s13253-020-00394-3)
394 (cit. on p. 2).
- 395 White, J. W. and Evert, F. K. van (2008). “Publishing Agronomic Data”. In: *Agronomy Journal* 100.5, pp. 1396–1400.
396 DOI: <https://doi.org/10.2134/agronj2008.0080F> (cit. on p. 6).

³⁹⁷ Zimmerman, D. L. and Harville, D. A. (1991). "A Random Field Approach to the Analysis of Field-Plot Experiments
³⁹⁸ and Other Spatial Experiments". In: *Biometrics* 47.1, pp. 223–239. doi: [10.2307/2532508](https://doi.org/10.2307/2532508) (cit. on p. 3).

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

2 Zhanglong Cao^{1,*}, Jordan Brown¹, Mark Gibberd¹, Julia Easton¹, and Suman Rakshit^{1,2}

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

5 ²School of Electrical Engineering, Computing and Mathematical Sciences, Curtin University, Perth,
6 Australia

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

8 Abstract

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

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

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

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

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

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

32 **Keywords:** yield map, optimal treatment, spatially varying coefficients, geographically weighted regression, pre-
33 cision agriculture.

34 1 Introduction

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

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

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

63 Analysing a systematic design for the creation of an optimal treatment map is a statistically challenging task. The
64 true responses at each point on the grid corresponding to all the treatment rates are unknown, and the treatment
65 producing the optimum response may vary continuously across the field. Cao et al. (2022) implemented a Bayesian
66 approach with spatially correlated random parameters for analysing large systematic strip trials. These authors
67 considered a quadratic response model with both global and local (spatially-varying) components. However, Bayesian
68 analysis can be computationally expensive and would require at least preliminary knowledge of Bayesian inference to
69 interpret the results, which can be extremely demanding for farmers and agronomists. Alternatively, Rakshit et al.
70 (2020) adopted a local regression approach, called geographically weighted regression (GWR), to obtain spatially-

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

80 Piepho and Edmondson (2018) presented an example where a linear model turns out to be inadequate for analysing
 81 sugar beet data (Petersen 1994). Glynn (2007) showed that many curves exist beyond a linear trend for nutrient-
 82 response relationships. The response curve often depends on the availability of other macro and micronutrients in the
 83 soil (Marschner 2011), which means that a linear relationship is unlikely to be consistent across a large trial. For this
 84 reason, it is important to consider models with terms of order higher than unity. For example, a quadratic model can
 85 often found to be suitable for modelling nutrient-response relationships (Piepho and Edmondson 2018; Liben et al.
 86 2019).

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

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

93 2 Methods

94 2.1 Hierarchical model for generating simulated data

95 In a conventional agricultural study, a field experiment can be considered as a rectangular matrix, representing
 96 a regular grid with r rows and c columns, where the total number of plots in the experiment is $n = r \times c$. Let
 97 $s_i \in \mathbb{R}^2, i = 1, \dots, n$, denote the Cartesian coordinate of the i -th plot centroid, located on a regular grid (Zimmerman
 98 and Harville 1991). Let $y(s_i), i = 1, \dots, n$, denote the value of the dependent variable recorded at the i -th plot.

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

$$\mathbf{Y} = \mathbf{X}\mathbf{b} + \mathbf{Z}\mathbf{u} + \mathbf{e}, \quad (1)$$

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

103 distribution is

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

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

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

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

108 In model (3), the structure of the covariance matrix $V_u(\theta_u)$ of \mathbf{u}_i can be either diagonal, which implies the **random**
109 **terms** at grid i are independent, or in general form, which means a correlation exists. McElreath (2015) suggested
110 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
111 diagonal elements σ_{uj} , $j = 1, \dots, k$, and R_u denotes the matrix with correlation coefficients. For the matrix R_u , we
112 specify the Lewandowski-Kurowicka-Joe (LKJ) distribution (Lewandowski et al. 2009), which is given by

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

113 where $\text{LKJcorr}(\epsilon)$ is a positive definite correlation matrix sampled from the LKJ distribution controlled by a positive
114 parameter ϵ . As ϵ increases, a high correlation among parameters becomes less likely.

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

$$\Sigma_u = V_s \otimes V_u. \quad (5)$$

117 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**
118 **random terms at a grid point are independently distributed from those at other grid points**, even though the **terms**
119 **at that grid point are correlated amongst each other**. However, the correlation among grids is ubiquitous. Hence, we
120 introduce a simple spatial covariance matrix such as

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

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

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

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

124 where d is the space lag or distance; γ is a non-negative scaling parameter; $\nu > 0$ is a smoothness parameter; σ^2 is
 125 the variance parameter; Γ is the Gamma function; and K_ν is the modified Bessel function of the second kind. The
 126 Matérn covariance is commonly used in the analysis of geostatistical data (Cressie and Huang 1999). Moreover, it has
 127 also been used in capturing spatial variation in OFE (Selle et al. 2019). If $\nu = \gamma + 1/2$, then the Matérn covariance
 128 can be expressed as a product of an exponential and a polynomial of order γ (Pandit and Infield 2019; Abramowitz
 129 1974), which simplifies the model and the computation process. The Matérn models with the values 3/2 and 5/2 for
 130 ν are used in most applications.

131 In each iteration of the simulation process, we used the above formulae and pre-defined parameter values to generate
 132 $2 \times n$ coefficients for the linear response and $3 \times n$ coefficients for the quadratic response. Then these coefficients are
 133 applied to simulate yield response across strips for both randomised and systematic design layouts.

134 2.2 Fitting geographically weighted regression to simulated data

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

139 The underlying template model for GWR is given by

$$y(s_i) = \beta_0 + \sum_{j=1}^k \beta_j \mathbf{x}_j(s_i) + \varepsilon_i, \quad (8)$$

140 where $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_k)^\top$ are model parameters corresponding to the k treatment levels and $\varepsilon_i \sim \mathcal{N}(0, \tau^2)$, $i =$
 141 $1, \dots, n$, are independent and identically distributed error terms at n grid points.

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

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

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

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

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

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

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

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

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

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

155 2.3 Performance evaluation

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

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

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

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

163 3 Simulation study

164 The simulation study is performed using realistic values for the model parameters, which are selected based on
165 the analysis results of a real-life data recorded from a corn field in Las Rosas, Argentina. This data set was originally
166 provided by Anselin et al. (2004) and can be accessed via the R-package `agridat` (White and Evert 2008). In 2001,
167 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
168 replicates. Each replicate consists of 93 rows and 6 columns after data preprocessing; see Rakshit et al. (2020) for
169 further details about the preprocessing steps. The unit of yield is quintals per hectare.

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

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

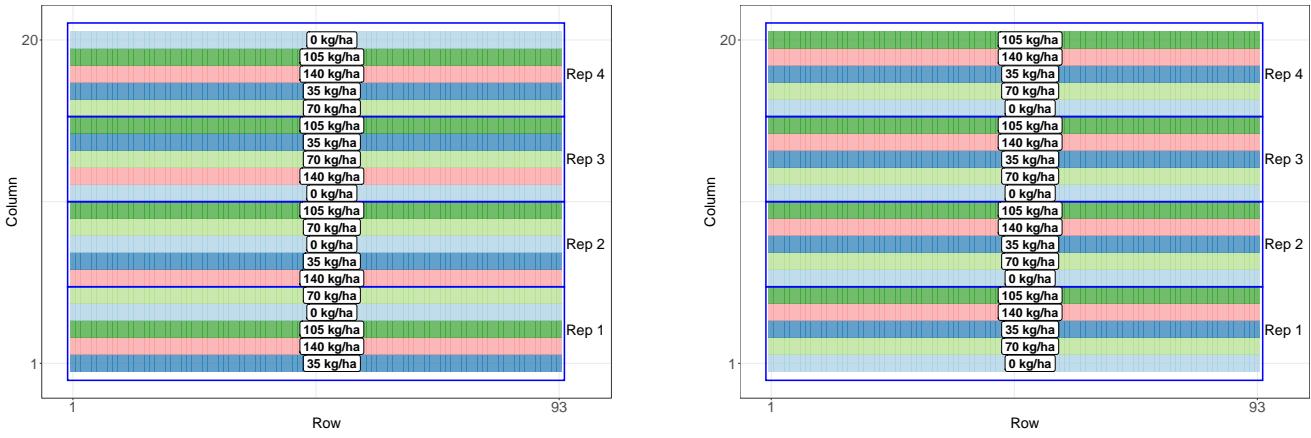


Figure 1: The nitrogen treatments with five levels (0, 35, 70, 105 and 140 kg/ha) randomly (*left*) and systematically (*right*) allocated into strips in each replicate block.

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

189 To specify the linear relationship in the simulation study, we consider the values of 65 and 0.05 for the global
 190 intercept b_0 and slope b_1 coefficients, respectively, in model (3). The variances of the coefficients \mathbf{u}_i are set to 5 for
 191 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
 192 $\text{AR1} \otimes \text{AR1}$ covariance matrix in (6), the two correlation parameters ρ_c and ρ_r are set to 0.15 and 0.50, respectively.
 193 We assume a higher correlation in the row direction because the crop is traditionally sown and harvested along the
 194 column direction, and the correlation is higher in the direction perpendicular to the sowing direction (Marchant et al.
 195 2019). For the Matérn covariance matrix (7), we set the value of the variance parameter σ_d^2 to 1, the value of the
 196 parameter r to 1, and the value of the parameter ν to 3/2. After drawing samples of \mathbf{u} from $\mathcal{N}(0, \Sigma_u)$, the spatially
 197 varying coefficients β_0 and β_1 are specified using the relations $\beta_0 = b_0 + \mathbf{u}_0$ and $\beta_1 = b_1 + \mathbf{u}_1$.

198 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 ,
 199 respectively. These choices make the response curve concave down. For the variance components, we set to 5 for σ_{u_0} ,
 200 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
 201 coefficients are $\beta_0 = b_0 + \mathbf{u}_0$, $\beta_1 = b_1 + \mathbf{u}_1$, and $\beta_2 = b_2 + \mathbf{u}_2$ for the quadratic model.

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

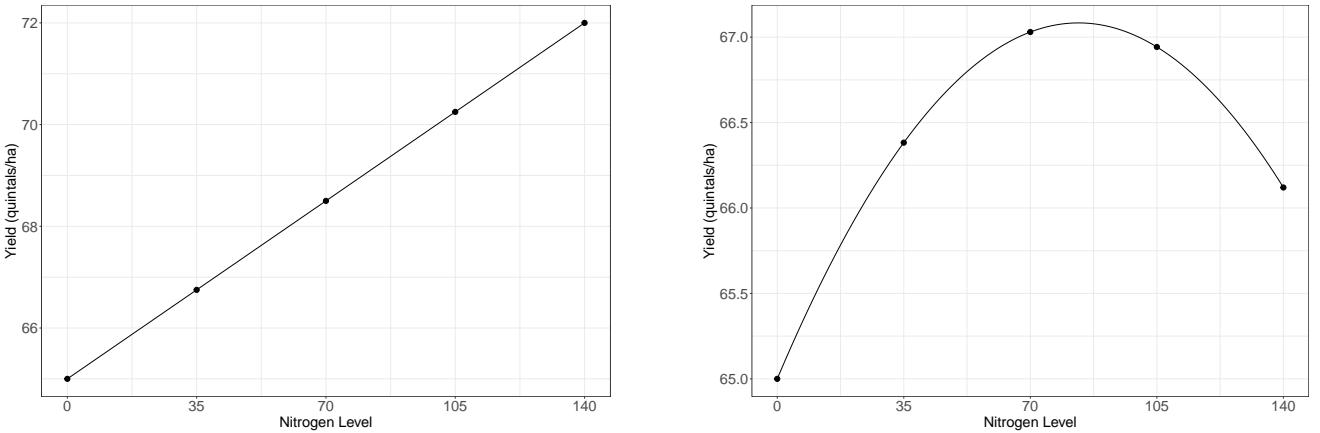


Figure 2: The global linear relationship of yield and nitrogen is $y = 65 + 0.05N$ (left), and the global quadratic relationship between yield and nitrogen is $y = 65 + 0.05N - 0.003N^2$ (right).

203 To summarise, the simulated yield response is obtained by

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

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

205 4 Results

206 In this section, we assess the performances of randomised and systematic designs in terms of their utility to
 207 accurately estimate the model parameters for both linear and quadratic response models. To this end, we perform
 208 1000 simulations of each of the 24 scenarios, described in the previous section. In each simulation, we first generate
 209 the coefficients for all grids and then apply the treatment in each grid to calculate the yield value using the model
 210 coefficients. For the treatment order in each iteration, we randomly picked an order of treatments for a single replicate
 211 and repeated this sequence across all other replicates to construct a systematic design. For a randomised design, all
 212 replicates have random orders of the treatments.

213 4.1 Comparison based on mean squared errors

214 Figures 3 and 4 show the results of linear models for the cases of low ($\epsilon = 1$) and high ($\epsilon = 0.1$) correlations,
 215 respectively, while Figures 5 and 6 show the results of quadratic models for the same low and high correlations,
 216 respectively. To specify the covariance matrix V_s used in producing the results in these figures, we use the following
 217 labels: (i) “NS” for the identity matrix representing no spatial correlation, (ii) “AR1” for $AR1(0.15) \otimes AR1(0.5)$, and
 218 (iii) “Matérn” for Matérn covariance with $\nu = 3/2$. Note that the model parameters and their corresponding MSEs
 219 are small values, and this makes it difficult to compare the MSEs of different scenarios using the original scale of MSE
 220 values. Therefore, to enhance clarity in visualisation and comparison, we have multiplied the MSEs of β_1 and β_2 by
 221 10^4 and 10^8 , respectively, in the figures presented below.

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

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

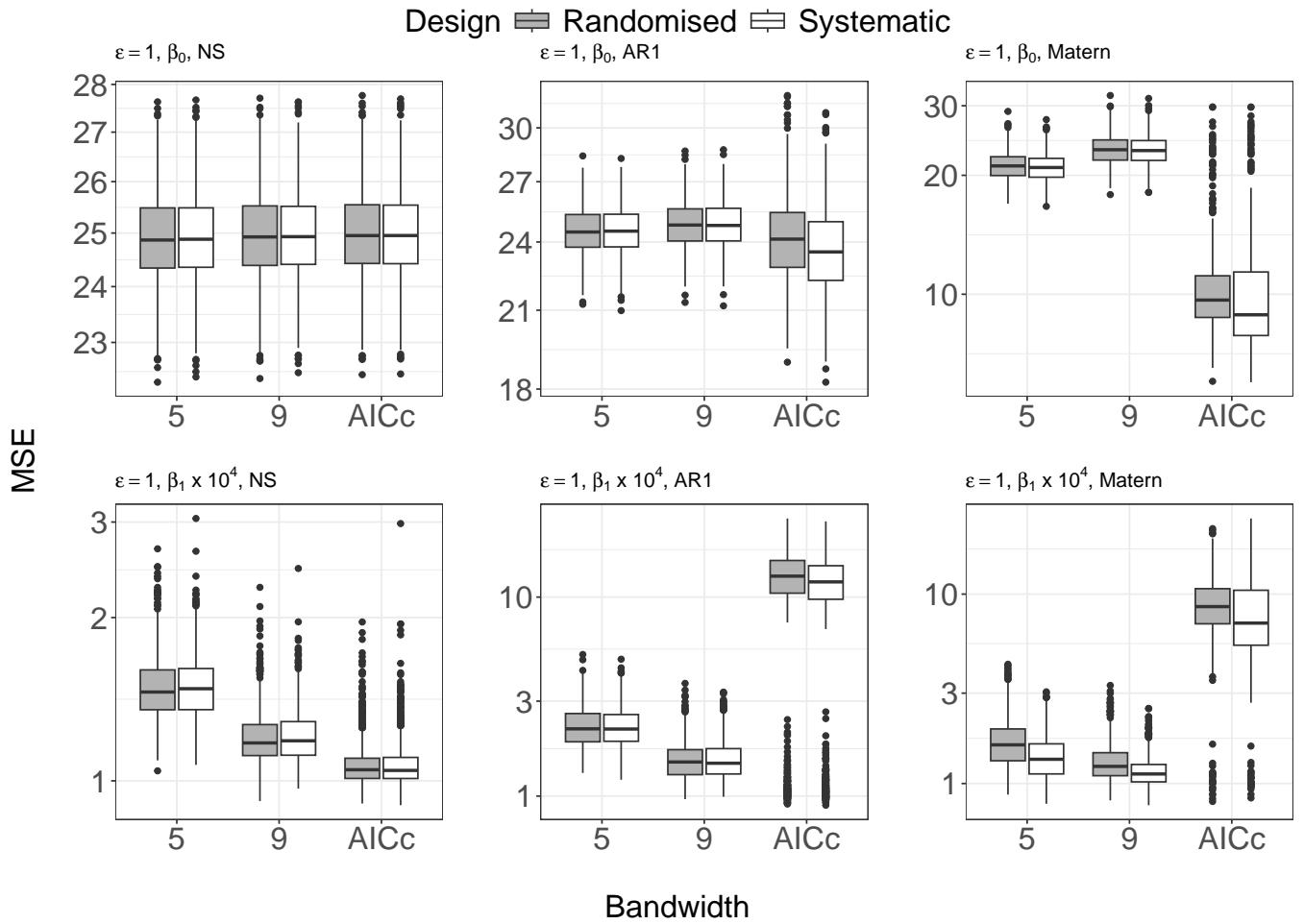


Figure 3: Boxplots of MSE for $\hat{\beta}_0$ and $\hat{\beta}_1$ in GWR models using different bandwidths for the simulated data with a linear response. The simulated data had different spatial covariance matrices (NS, $\text{AR1} \otimes \text{AR1}$ and Matérn) and a low correlation between the parameters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

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

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

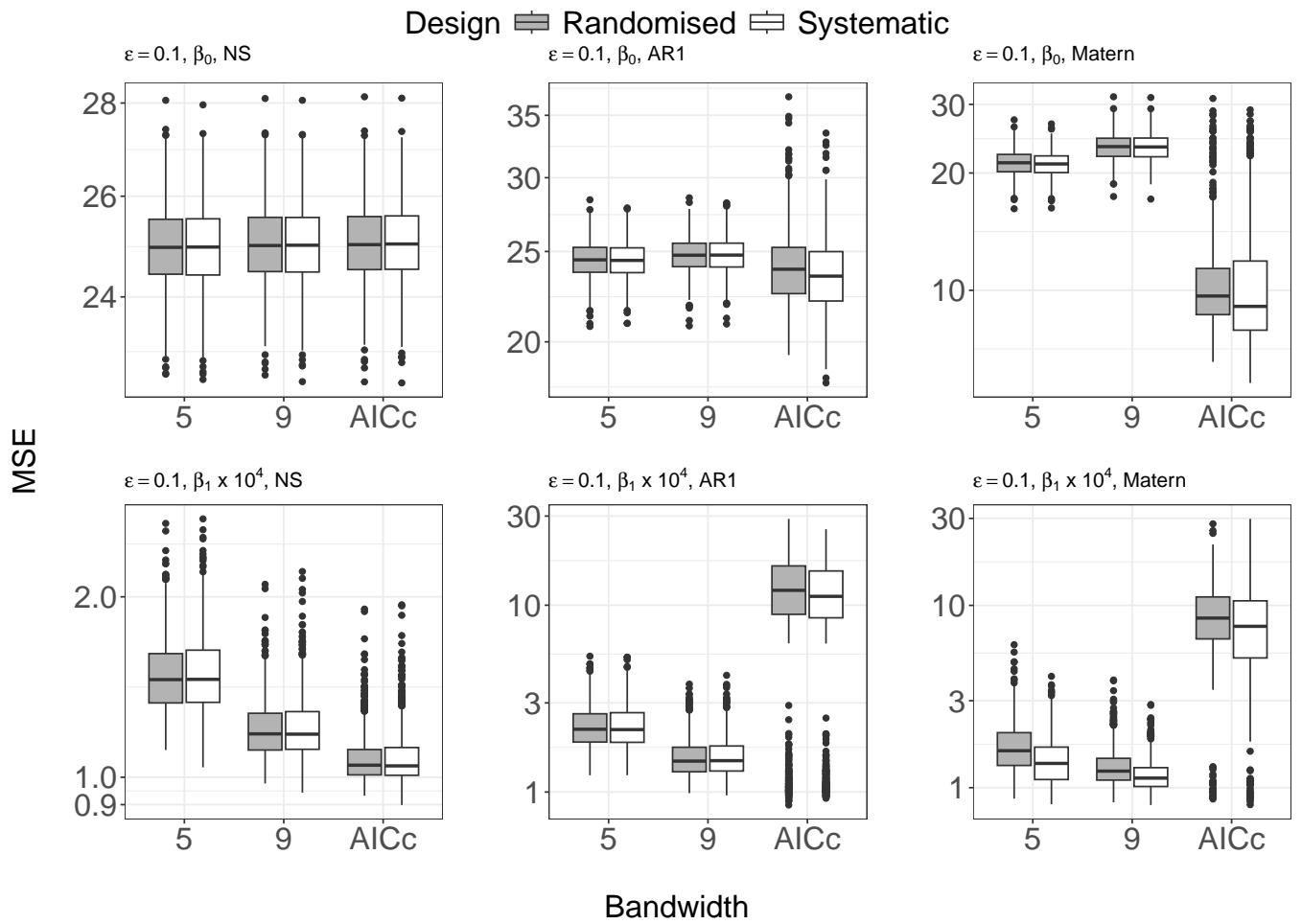


Figure 4: Boxplots of MSE for $\hat{\beta}_0$ and $\hat{\beta}_1$ in GWR models using different bandwidths for the simulated data with a linear response. The simulated data had different spatial covariance matrices (NS, AR1 \otimes AR1 and Matérn) and a high correlation between the parameters ($\epsilon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

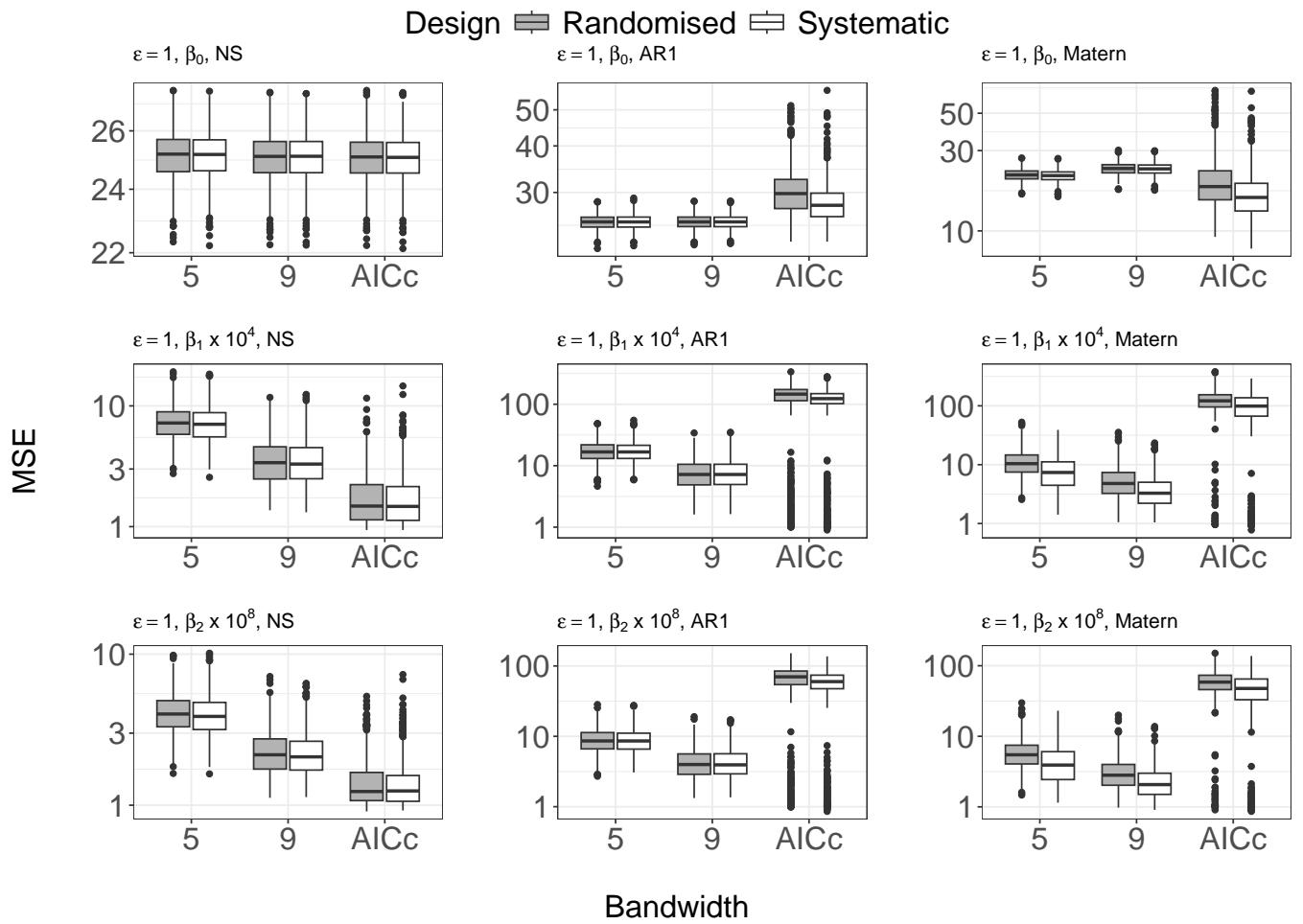


Figure 5: Boxplots of MSE for $\hat{\beta}_0$, $\hat{\beta}_1$ and $\hat{\beta}_2$ in GWR models using different bandwidths for the simulated data with a quadratic response. The simulated data had different spatial covariance matrices (NS, AR1 \otimes AR1 and Matérn) and a low correlation amongst the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

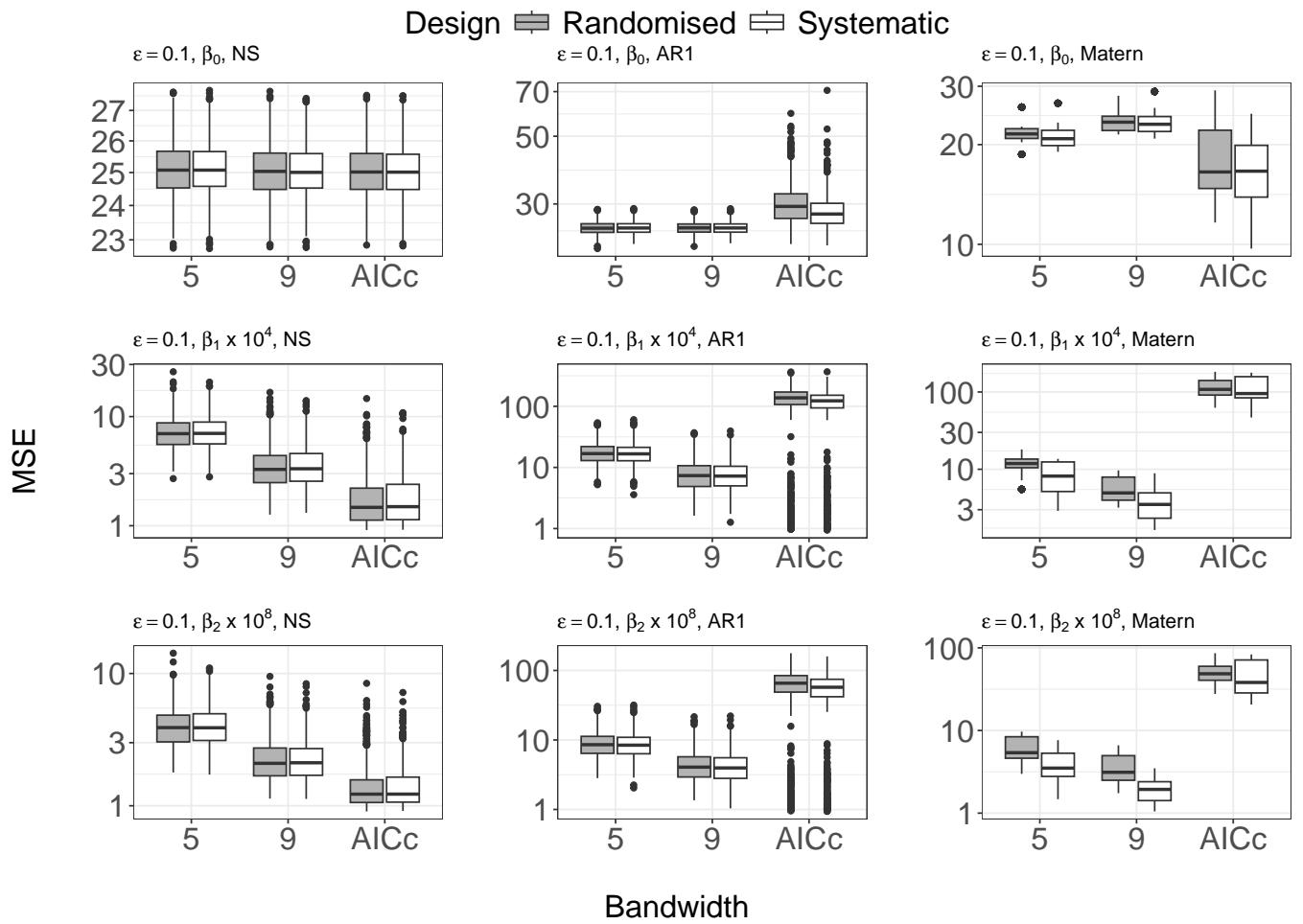


Figure 6: Boxplots of MSE for $\hat{\beta}_0$, $\hat{\beta}_1$ and $\hat{\beta}_2$ in GWR models using different bandwidths for the simulated data with a quadratic response. The simulated data had different spatial covariance matrices (NS, AR1 \otimes AR1 and Matérn) and a high correlation amongst the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

246 4.2 Comparison of density plot

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

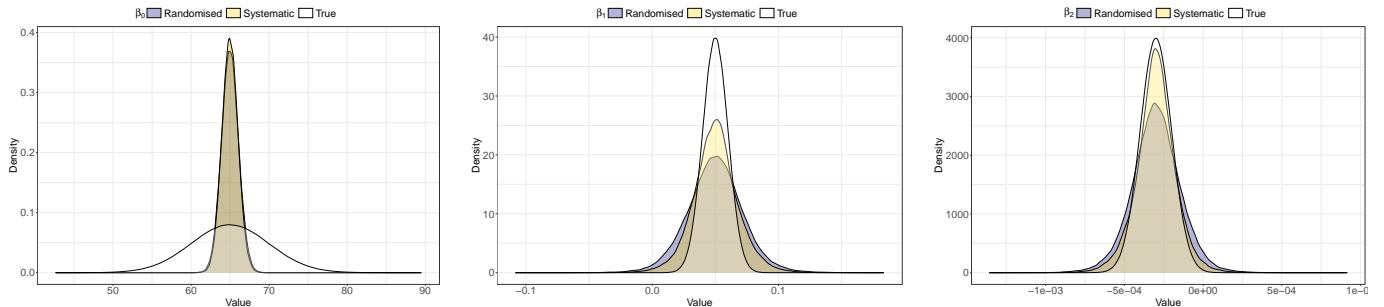


Figure 7: Density plots comparing the true coefficients of β_0 (left), β_1 (middle), and β_2 (right) with their estimates from GWR with bandwidth 9, based on 1000 simulations.

252 4.3 An example of optimal nitrogen map

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

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

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

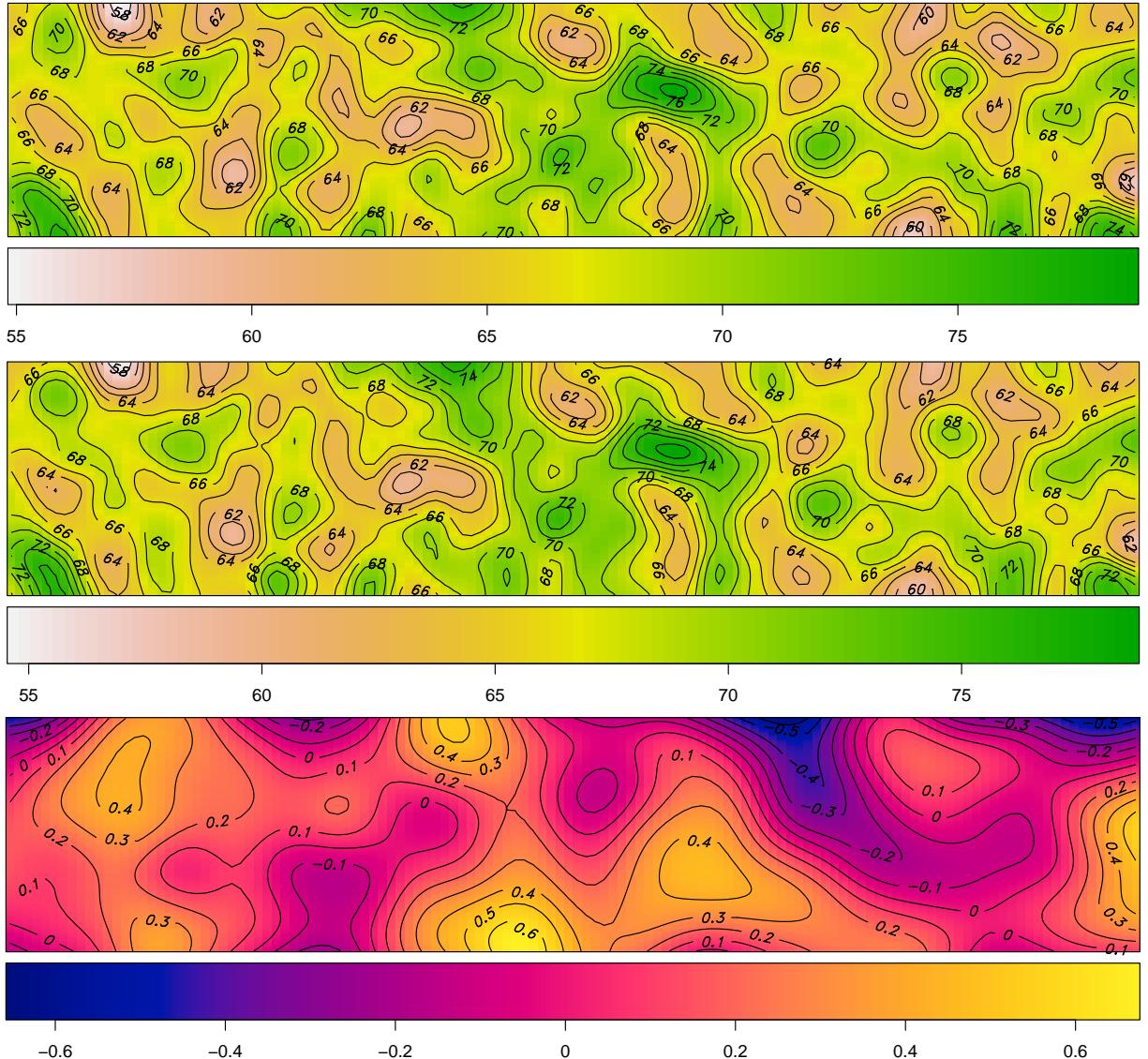


Figure 8: Simulated yield map of a randomised design (*top*), and a systematic design (*middle*), and the difference of these two designs (*bottom*).

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
 262 140, $i = 1, \dots, n$. For the randomised design, GWR underestimated the right part of the paddock. On the contrary,
 263 the estimated map from the systematic design is more consistent.

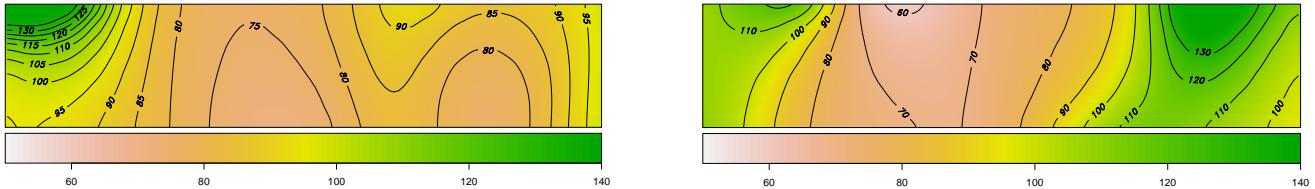


Figure 9: Optimal Nitrogen rate estimated by GWR from a randomised design (*left*), and a systematic design (*right*).

5 Discussion

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

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

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

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

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

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

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

306 6 Conclusion

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

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

313 Acknowledgements

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

316 CRedit authorship contribution statement

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

323 References

- 324 Abramowitz, M. (1974). *Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables*.
325 USA: Dover Publications, Inc. (cit. on p. 5).
- 326 Alesso, C. A., Cipriotti, P. A., Bollero, G. A., and Martin, N. F. (2021). “Design of on-farm precision experiments to
327 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).
- 328 Anselin, L., Bongiovanni, R., and Lowenberg-DeBoer, J. (2004). “A Spatial Econometric Approach to the Economics
329 of Site-Specific Nitrogen Management in Corn Production”. In: *American Journal of Agricultural Economics* 86.3,
330 pp. 675–687. DOI: <https://doi.org/10.1111/j.0002-9092.2004.00610.x> (cit. on p. 6).
- 331 Bramley, R., Cook, S., Adams, M., and Corner, R. (1999). “Designing your own on-farm experiments: How precision
332 agriculture can help”. In: (cit. on p. 17).
- 333 Butler, D., Cullis, B., Gilmour, A., Gogel, B., and Thompson, R. (2017). *ASReml-R Reference Manual Version 4*
334 (cit. on p. 4).
- 335 Cao, Z., Stefanova, K., Gibberd, M., and Rakshit, S. (2022). “Bayesian inference of spatially correlated random
336 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).
- 337 Cook, S., Cock, J., Oberthür, T., and Fisher, M. (2013). “On-Farm Experimentation”. In: *Better Crop. Plant Food.*
338 4th ser. 97, pp. 17–20 (cit. on p. 2).
- 339 Cressie, N. and Huang, H.-C. (Dec. 1999). “Classes of Nonseparable, Spatio-Temporal Stationary Covariance Func-
340 tions”. In: *Journal of the American Statistical Association* 94.448, pp. 1330–1339. DOI: [10.1080/01621459.1999.10473885](https://doi.org/10.1080/01621459.1999.10473885) (cit. on p. 5).
- 341 Evans, F. H., Recalde Salas, A., Rakshit, S., Scanlan, C. A., and Cook, S. E. (Nov. 2020). “Assessment of the Use
342 of Geographically Weighted Regression for Analysis of Large On-Farm Experiments and Implications for Practical
343 Application”. en. In: *Agronomy* 10.11, p. 1720. DOI: [10.3390/agronomy10111720](https://doi.org/10.3390/agronomy10111720) (cit. on pp. 2, 3, 6).
- 344 Fisher, R. A. (1934). *Statistical Methods for Research Workers*. English. fifth. Edinburgh: Oliver and Boyd (cit. on
345 p. 2).
- 346 Glynn, C. (2007). “Testing the Growth-Differentiation Balance Hypothesis: Dynamic Responses of Willows to Nutrient
347 Availability”. In: *New Phytologist* 176.3, pp. 623–634. DOI: [10.1111/j.1469-8137.2007.02203.x](https://doi.org/10.1111/j.1469-8137.2007.02203.x) (cit. on p. 3).
- 348 Gollini, I., Lu, B., Charlton, M., Brunsdon, C., and Harris, P. (Feb. 2015). “GWmodel: An R Package for Exploring
349 Spatial Heterogeneity Using Geographically Weighted Models”. en. In: *Journal of Statistical Software* 63.1, pp. 1–
350 50. DOI: [10.18637/jss.v063.i17](https://doi.org/10.18637/jss.v063.i17) (cit. on pp. 3, 5).
- 351 Lewandowski, D., Kurowicka, D., and Joe, H. (2009). “Generating random correlation matrices based on vines and
352 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).
- 353 Liben, F. M., Midega, T., Tufa, T., and Wortmann, C. S. (2019). “Soil Fertility & Crop Nutrition Barley and wheat
354 nutrient responses for Shewa, Ethiopia”. In: *Agronomy Journal*. DOI: [10.1002/agj2.20020](https://doi.org/10.1002/agj2.20020) (cit. on p. 3).

- 359 Lu, B., Harris, P., Charlton, M., and Brunsdon, C. (2014). “The GWmodel R package: further topics for exploring
360 spatial heterogeneity using geographically weighted models”. In: *Geo-spatial Information Science* 17.2, pp. 85–101.
361 DOI: [10.1080/10095020.2014.917453](https://doi.org/10.1080/10095020.2014.917453) (cit. on p. 3).
- 362 Marchant, B., Rudolph, S., Roques, S., Kindred, D., Gillingham, V., Welham, S., Coleman, C., and Sylvester-Bradley,
363 R. (Jan. 2019). “Establishing the Precision and Robustness of Farmers’ Crop Experiments”. en. In: *Field Crops
364 Research* 230, pp. 31–45. DOI: [10.1016/j.fcr.2018.10.006](https://doi.org/10.1016/j.fcr.2018.10.006) (cit. on p. 7).
- 365 Marschner, H. (2011). *Marschner’s Mineral Nutrition of Higher Plants*. San Diego: San Diego: Elsevier Science &
366 Technology. DOI: [10.1016/C2009-0-63043-9](https://doi.org/10.1016/C2009-0-63043-9) (cit. on p. 3).
- 367 McElreath, R. (Dec. 2015). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. English. First.
368 Vol. 122. Chapman and Hall/CRC Texts in Statistical Science Ser. CRC Press LLC (cit. on p. 4).
- 369 Páez, A., Uchida, T., and Miyamoto, K. (Apr. 2002). “A General Framework for Estimation and Inference of Geo-
370 graphically Weighted Regression Models: 1. Location-Specific Kernel Bandwidths and a Test for Locational Hetero-
371 geneity”. In: *Environment and Planning A: Economy and Space* 34.4, pp. 733–754. DOI: [10.1068/a34110](https://doi.org/10.1068/a34110) (cit. on
372 p. 5).
- 373 Pandit, R. K. and Infield, D. (Sept. 2019). “Comparative Analysis of Gaussian Process Power Curve Models Based
374 on Different Stationary Covariance Functions for the Purpose of Improving Model Accuracy”. en. In: *Renewable
375 Energy* 140, pp. 190–202. DOI: [10.1016/j.renene.2019.03.047](https://doi.org/10.1016/j.renene.2019.03.047) (cit. on p. 5).
- 376 Petersen, R. G. (1994). *Agricultural Field Experiments: Design and Analysis*. en. 1st. Boca Raton: CRC Press (cit. on
377 pp. 2, 3).
- 378 Piepho, H. P. and Edmondson, R. N. (2018). “A tutorial on the statistical analysis of factorial experiments with
379 qualitative and quantitative treatment factor levels”. In: *Journal of Agronomy and Crop Science* 204.5, pp. 429–
380 455. DOI: <https://doi.org/10.1111/jac.12267> (cit. on p. 3).
- 381 Piepho, H. P., Möhring, J., and Williams, E. R. (2013). “Why Randomize Agricultural Experiments?” In: *Journal of
382 Agronomy and Crop Science* 199.5, pp. 374–383. DOI: [10.1111/jac.12026](https://doi.org/10.1111/jac.12026) (cit. on p. 2).
- 383 Pringle, M. J., Cook, S. E., and McBratney, A. B. (Dec. 2004). “Field-Scale Experiments for Site-Specific Crop
384 Management. Part I: Design Considerations”. en. In: *Precision Agriculture* 5.6, pp. 617–624. DOI: [10.1007/s11119-004-6346-1](https://doi.org/10.1007/s11119-004-6346-1) (cit. on pp. 2, 3).
- 386 Rakshit, S., Baddeley, A., Stefanova, K., Reeves, K., Chen, K., Cao, Z., Evans, F., and Gibberd, M. (2020). “Novel Ap-
387 proach to the Analysis of Spatially-Varying Treatment Effects in on-Farm Experiments”. In: *Field Crops Research*
388 255.October 2019, p. 107783. DOI: [10/gg2vv7](https://doi.org/10/gg2vv7) (cit. on pp. 2, 5, 6, 16).
- 389 Selle, M. L., Steinsland, I., Hickey, J. M., and Gorjanc, G. (2019). “Flexible Modelling of Spatial Variation in Agri-
390 cultural Field Trials with the R Package INLA”. In: *Theoretical and Applied Genetics* 132.12, pp. 3277–3293. DOI:
391 DOI: [10.1007/s00122-019-03424-y](https://doi.org/10.1007/s00122-019-03424-y) (cit. on pp. 2, 5).
- 392 Verdooren, L. R. (Dec. 2020). “History of the Statistical Design of Agricultural Experiments”. en. In: *Journal of
393 Agricultural, Biological and Environmental Statistics* 25.4, pp. 457–486. DOI: [10.1007/s13253-020-00394-3](https://doi.org/10.1007/s13253-020-00394-3)
394 (cit. on p. 2).
- 395 White, J. W. and Evert, F. K. van (2008). “Publishing Agronomic Data”. In: *Agronomy Journal* 100.5, pp. 1396–1400.
396 DOI: <https://doi.org/10.2134/agronj2008.0080F> (cit. on p. 6).

³⁹⁷ Zimmerman, D. L. and Harville, D. A. (1991). "A Random Field Approach to the Analysis of Field-Plot Experiments
³⁹⁸ and Other Spatial Experiments". In: *Biometrics* 47.1, pp. 223–239. doi: [10.2307/2532508](https://doi.org/10.2307/2532508) (cit. on p. 3).

[Click here to access/download Figure Dif-BJR qua 1/b9.pdf](#)

