Trial designs for detecting spatial variability of treatment effects in on-farm precision experiments

Carlos Agustín Alesso[1*], Patricia Melina Acetta[2], Nicolás Federico Martin[3], y Pablo Ariel Cipriotti[4]

[1] ICiAgro Litoral, UNL, CONICET, Fac. de Ciencias Agrarias, Kreder 2805, S3080HOF Esperanza (Argentina)
calesso@fca.unl.edu.ar

[2] Facultad de Ciencias Agrarias, UNL, Kreder 2805, S3080HOF, Esperanza (Argentina)
pacetta@fca.unl.edu.ar

[3] Dept. of Crop Sciences, University of Illinois, 1102 S Goodwin Ave, 61801 Urbana (USA)
nfmartin@illinois.edu

[4] IFEVA, UBA, CONICET, Facultad de Agronomía, Av. San Martín 4453, (C1417DSE) Ciudad de Buenos Aires (Argentina)
cipriott@agro.uba.ar

[*] **Corresponding author**

Abstract. Precision agriculture involves the existence of spatial variability in crop response to input application. Field-scale experiments allow for exploring such variability. However, the interaction between the spatial variability of factors controlling crop response and the applied experimental design conditions the results. It is necessary to identify experimental designs that optimize the acquisition of reliable information on intra-field crop response. Experimental designs at field scale with different spatial resolutions were evaluated to estimate the spatial variability of crop response to input application. Spatial response patterns were simulated as an underlying process to generate yield maps. Geographically weighted regression (GWR) was used to estimate crop response patterns, which were compared with the underlying stochastic field. Designs with high spatial resolution better capture underlying spatial variability patterns across a wide range of considered spatial structures. Furthermore, checkerboard plot designs outperform strip designs as they enable detecting spatial variability in both directions. However, the agreement between GWR-estimated response maps and reference maps is sensitive to kernel selection and bandwidth.

Key words: statistical simulation, non-stationarity, site-specific management

Received May 2024; Accepted June 2024; Published July 2024
https://doi.org/10.24215/15146774e049

ISSN 1514-6774

Diseños experimentales para detectar variabilidad espacial del efecto de tratamientos a escala de lote

Carlos Agustín Alesso[1*], Patricia Melina Acetta[2], Nicolás Federico Martin[3], y Pablo Ariel Cipriotti[4]

[1] ICiAgro Litoral, UNL, CONICET, Fac. de Ciencias Agrarias, Kreder 2805, S3080HOF Esperanza (Argentina)
calesso@fca.unl.edu.ar

[2] Facultad de Ciencias Agrarias, UNL, Kreder 2805, S3080HOF, Esperanza (Argentina)
pacetta@fca.unl.edu.ar

[3] Dept. of Crop Sciences, University of Illinois, 1102 S Goodwin Ave, 61801 Urbana (USA)
nfmartin@illinois.edu

[4] IFEVA, UBA, CONICET, Facultad de Agronomía, Av. San Martín 4453, (C1417DSE) Ciudad de Buenos Aires (Argentina)
cipriott@agro.uba.ar

[*] **Autor de correspondencia**

Resumen. La agricultura de precisión supone la existencia de variabilidad espacial de la respuesta de los cultivos a la aplicación de insumos. Los experimentos a escala de lote permiten explorar dicha variabilidad. No obstante, la interacción entre la variabilidad espacial de los factores que controlan la respuesta del cultivo y el diseño experimental aplicado condicionan los resultados. Es necesario identificar diseños experimentales que optimicen la obtención de información fiable de la respuesta de los cultivos intra-lote. Se evaluaron diseños experimentales a escala de lote con distinta resolución espacial para estimar la variabilidad espacial de la respuesta de un cultivo a la aplicación de insumos. Se simularon patrones espaciales de respuesta como proceso subyacente para generar mapas de rendimiento. Mediante regresión ponderada geográficamente (GWR) se estimaron los patrones de respuesta del cultivo que se compararon con el campo estocástico subyacente. Los diseños con alta resolución espacial permiten capturar mejor los patrones de variabilidad espacial subyacente en un amplio rango de estructuras espaciales consideradas. A su vez, diseños en parcelas tipo tablero de ajedrez superan a los diseños en franjas ya que permiten detectar variabilidad espacial en ambas direcciones. No obstante, la concordancia entre los mapas de respuesta estimados por GWR y los de referencia son sensibles a la selección de *kernel* y ancho de banda.

Palabras clave: simulación estadística, no estacionariedad, manejo sitio-especifico

ISSN 1514-6774

1 Introduction

In agriculture, field trials are useful for collecting information about the crop response to input applications at field scale. Insights from these experiments are useful for supporting better decisions tailored to improve production, efficiency and sustainability considering soil and weather variability [1]. The experimental design is a key aspect for getting reliable information from this kind of experiments [2].

During the last decade, precision agriculture (PA) technologies have helped farmers and crop advisers to conduct field-scale on-farm experiments with little or no disturbance of regular field operations. These experiments are commonly carried out for comparing alternative agronomic practices or for adjusting a given crop management [3]. At the same time, the complexity of data collected by the machinery have challenged the application of classical statistical methods designed for on-station small-plot experiments requiring new approaches [4, 5]. In a simulation study, Alesso et al. [6] have demonstrated that, if not taken into account, spatial autocorrelation can reduce the efficiency of treatment effect estimators and to increase the type I error rates. Among the experimental designs tested, those with smaller experimental units and larger number of replications had the best performances.

The management zones approach (MZ), which is a form of PA, has been widely adopted for managing within field spatial variability [7]. Under this approach, the field is classified into zones having similar soil type and landscape characteristics, and crop response to input applications is expected to be fairly similar within those zones rather than between zones. Thus, on-farm experiments carried out using the MZ approach enable the quantification of such responses by zone.

The increase in the spatial resolution at which the inputs can be controlled today by farmers makes them willing to quantify site-specific crop responses instead of field- or management-zone-level responses [8–10]. To do so, several experimental designs, both randomized and systematic or patterned, and analytical methods have been proposed. Whole-field strip designs are among the most popular designs [11, 12]. Other designs included checkerboard [13] and whole-field block [14] designs. In those approaches, the continuous field spatial variability is divided into management units compatible with the spatial resolution of machinery and a site-specific crop response function is estimated. Due to only one treatment can be assigned at the same time to each management unit, the effect of the remaining treatments needs to be estimated from geostatistical interpolation techniques [3, 15, 16] or by blocks containing another replication of experiment [11, 12]. This methodology increases the spatial resolution of the resulting information compared to the MZ approach.

Trevisan et al. [17] have demonstrated the potential of geographically weighted regression (GWR) applied to on-farm field scale experiments to develop prescription maps based on site-specific crop responses rather than sub-field regions or MZ. This technique, developed for tackling non-stationary processes, allows the estimation of

local response functions where model coefficients vary across space, and thus they can be mapped and interpreted as a new spatial variable [18]. Unlike global regression models, the GWR fits local regression models from data within a neighborhood defined by a band width. The coefficients of such models are estimated by weighted least squares procedures. Weights are assigned using a *kernel* function which models the decay of neighbors' influence as they get farther from the target point. Due to the bandwidth and *kernel* choices can result in the detection of misleading non-stationarity patterns [19, 20] several calibration methods were reported [21].

Another important aspect in the application of GWR to the analysis of on-farm precision experiments is the effect of the spatial configuration of experimental units, i.e. the experimental design, and the spatial structure of the underlying spatial process, on the estimation of the spatial variability of the crop response. In a simulation study, Alesso et al. [2] have explored these effects by comparing checkerboard designs with different plot dimensions and randomization under scenarios of spatial structure. These authors concluded that systematic designs with small plots have better performance compared to randomized and large plots alternatives in terms of concordance between the true underlying spatial structure and the patterns approximated by GWR. In a recent study, Li et al. [22] quantified the economic impact of experimental designs used to estimate the economic optimum nitrogen rate (EONR) for corn. They concluded that the choice of the experimental design can significantly influence the economic value of the information obtained from these kind of experiments. However, the effect of spatial structure and GWR parametrization was not addressed.

In this work we proposed an *in silico* study to explore the effect of experimental designs layouts and plot sizes, the underlying spatial structure of true treatment effect, and the GWR parameters, i.e. *kernel* and bandwidth, on the results obtained from an on-farm precision experiment (OFPE) conducted to map the field-scale spatial variability of a single treatment effect.

2 Materials and Methods

2.1 Hypothetical experiment

A simulation study was conducted to evaluate the effect of field-scale experimental designs on capturing the spatial variability of treatment effects under different spatial variability scenarios. The hypothetical experiment aimed to estimate the response of corn to the application of an optimal rate of phosphate fertilizer. An experimental area of 432 m wide by 864 m long (37 ha) was considered. It was assumed that the following machinery configurations were available to conduct the experiment and collect data: (1) a tractor with automatic guidance; (2) a planter with variable rate technology (VRT); and (3) a harvester equipped with yield monitor and automatic guidance. Ac-

cording to the machinery specification, the minimum width of the experimental units was the cutting width, which is equal to 9 m. The field dimension was approximately equivalent to 48 and 96 times the header width, respectively. A total of 4608 data points were simulated for each combination of spatial scenario, experimental design, and GWR parametrization described in the following section. The extent of the experimental field is comparable to that of commercial fields in the Pampas region and allows for a realistic representation of different spatial patterns and allocation of experimental unit sizes while keeping computational time low.

2.2 Experimental designs

The simulated treatments consisted of phosphorus application at planting, in recommended doses based on crop requirements and soil availability, along with a control treatment without application. Treatments were applied with a resolution ranging from 9 to 36 meters, perpendicular to the planting direction, and equivalent to 1 to 4 widths of the harvester header. Along the planting direction, the resolution ranged from 36 to 864 meters. The minimum length was defined to allow sufficient time (approximately 18 seconds) for actuators to adjust the rate on the go, considering delays in the grain threshing system and yield monitor. The maximum resolution was included to represent strip designs. Consequently, the sizes of the experimental units ranged from approximately 324 to 31,104 square meters. The evaluated designs are detailed in the following table, and some examples are visualized in Fig. 1B.

Table 1. Characteristics of the evaluated experimental designs.

Code	Width (m)	Length (m)	Replications	Area (m^2)
W1L4	9	36	1152	324
W1L8	9	72	576	648
W1L16	9	144	288	1296
W1L24	9	216	192	1944
W1L48	9	432	96	3888
W1L96	9	864	48	7776
W2L4	18	36	576	648
W2L8	18	72	288	1296
W2L16	18	144	144	2592
W2L24	18	216	96	3888
W2L48	18	432	48	7776
W2L96	18	864	24	15552
W4L4	36	36	288	1296
W4L8	36	72	144	2592
W4L16	36	144	72	5184
W4L24	36	216	48	7776
W4L48	36	432	24	15552
W4L96	36	864	12	31104

2.3 Yield simulation

The response to the treatment, i.e., phosphorus application, was simulated assuming the following linear model with spatially varying coefficients:

$$y_i(s_i)=\beta_0(s_i)+\beta_1(s_i)x_i+\varepsilon_i \quad (1)$$

where: $y_i(s_i)$ is the yield at site $s_i=(u_i,v_i)$ defined by the spatial coordinates u_i and v_i; $\beta_0(s_i)$ and $\beta_1(s_i)$ are the regression coefficients, where $\beta_0(s_i)$ is the yield from control treatment and $\beta_1(s_i)$ is the treatment effect, i.e. the difference between control and treatment, for the same site s_i; x_i is a dummy variable having 1 for treatment and 0 for control; and ε_i is the residual error which is assumed to be independent and identical normally distributed with 0 mean and constant variance, i.e. $N(0,\sigma_\varepsilon^2)$.

In the first place, the spatial distribution of each coefficient $\beta_j(s_i)$ from Eq (1) was independently simulated by unconditional Gaussian geostatistical simulation [23]. Spatial variability scenarios were simulated assuming a first-order stationary process with the following linear model:

$$\beta_j(s_i)=\mu_j+\varepsilon_i(s_i) \quad (2)$$

where: $\beta_j(s_i)$ is the true regression coefficient j at position $s_i=(u_i,v_i)$; μ_j is the true overall mean of the regression coefficient over the experimental area; and $\varepsilon_i(s_i)$ is the random error term Normal distribution having 0 mean and spatial variance-covariance matrix depending on the separation distances (h) between locations $N(0,\boldsymbol{\Sigma}(h))$. A spherical and isotropic autocorrelation model without nugget effect was assumed, with ranges of 0, 200, and 400 meters (Eq. 2). For the scenario with a range of 0 (no spatial autocorrelation), the variance-covariance matrix simplifies to $\boldsymbol{\Sigma}(h)=\boldsymbol{I}\,\sigma_\varepsilon^2=\sigma_\varepsilon^2$. Alongside the exponential model, the spherical model is one of the most commonly used models in the literature to represent soil and crop variables [23–25].

$$C(h)=\begin{cases}\sigma_{sill}^2\left(\dfrac{3h}{2a}+\dfrac{h^3}{2a^3}\right) & 0\le h\le a\\ 0 & 0\le h\le a\end{cases} \quad (3)$$

where: $C(h)$ is the spatial covariance function between pairs separated by h units; a is the range of spatial dependence in meters; σ_{sill}^2 is the total process variance, assumed based on a coefficient of variation of 30% around the mean yield of the control $\beta_0(s_i)$ 9 Mg/ha and the mean treatment effect $\beta_1(s_i)$ (1,3 Mg/ha). Therefore, the values of total variance were 7.29 (Mg/ha)² and 0.1521 (Mg/ha)² respectively. For each scenario, a random realization of both random fields was independently taken, assumed to be the true coefficient maps of the spatial model (Fig. 1A). The pixel-to-pixel correlation between the coefficient maps ($\beta_j(s_i)$) was negligible (Pearson correlation coefficients between -0.3 and 0.12). The combination of these coefficient maps allowed representing a wide range of responses, from situations with low yields with-

out phosphorus and no response, which occurs when the limiting factor is not the treatment, to high yields without phosphorus and high response when the site has high response potential.

The spatial variability scenarios included in this study aim to cover a wide range of response patterns observable in the field, from random patterns (range = 0 m) to well-structured patterns with smoothed variation patterns (range = 200 and 400 m). In the scenarios where coefficients do not have spatial structure, the application of the GWR procedure would not be warranted. Conversely, in scenarios with good spatial structure of the coefficients, the possibility of using linear models with spatially varying coefficients will allow estimating the local response to treatments.

Finally, for each combination of spatial variability scenario (3) and experimental design (18), 200 maps of corn yield were simulated on a grid of 4608 points combining the information of the coefficients $\beta_j(s_i)$, the assigned treatment, and a residual error with mean 0 and constant variance around 30% according to the model of Eq. 1 (Fig. 1C). In total, 10800 yield maps were simulated.

2.4 Data analysis

Each simulated yield map containing 4608 cells (each cell of 81 m²) was analyzed by fitting a GWR model [18] of Eq. 1 (Fig. 2). The regression coefficients at each location were estimated using weighted least squares based on information from nearby observations:

$$\hat{\beta}_i(s_i)=(\boldsymbol{X}'\boldsymbol{W}(s_i)\boldsymbol{X})^{-1}\boldsymbol{W}(s_i)\,\boldsymbol{y} \tag{3}$$

Where: $\boldsymbol{y}$ is the vector containing the yields; $\boldsymbol{X}$ is the predictor matrix composed of a column representing the intercept $\beta_0(s_i)$ and one for the treatment indicator variable $\beta_1(s_i)$; $\hat{\beta}_j(s_i)$ is the vector of estimators $\beta_j(s_i)$ ar locations $s_i=(u_i,v_i)$, and $\boldsymbol{W}(s_i)$ is an n x n diagonal matrix with the weights of each observation relative to point s_i on the diagonal. Three kernels with adaptive bandwidth were tested: exponential (Eq. 4), bisquare (Eq. 5), and Gaussian (Eq. 6):

$$w(s_i,s_j)=\exp\left(-\frac{d(s_i,s_j)}{b(s_i)^{ad}}\right) \tag{4}$$

$$w(s_i,s_j)=\left[1-\left(\frac{d(s_i,s_j)}{b(s_i)^{ad}}\right)^2\right] \text{ if } d(s_i,s_j)<b \text{ and 0 otherwise} \tag{5}$$

$$w(s_i,s_j)=\exp\left[\left(-\frac{1}{2}\frac{d(s_i,s_j)}{b(s_i)^{ad}}\right)^2\right] \tag{6}$$

where $w(s_i, s_j)$ is the weight of the data at location s_i relative to location s_j; $d(s_i, s_j)$ is the Euclidean distance between both locations; and $b(s_i)^{ad}$ is the bandwidth for the site s_i. Due to the impact of the number of neighbors on the coefficient estimation, two neighborhood sizes were evaluated: 3% and 5% of the total observations, i.e., 138 and 230 neighbors. Given the dimensions of the plots and the grid resolution, these numbers of neighbors allowed representing both treatments within a search radius for each target point. As a result, 64800 GWR models were estimated.

2.5 Designs comparison

The agreement between the effect map estimated by the GWR model $\hat{\beta}_1(s_i)$ for each realization of the simulated yield map and the true coefficient map of the effect $\beta_1(s_i)$ was evaluated using the following metrics: correlation coefficient (r), mean absolute error (MAE), coverage probability (CP), and standard deviation ratio (SDR):

$$r = \frac{\sigma_{\beta_1(s_i)\hat{\beta}_1(s_i)}}{\sigma_{\beta_1(s_i)}\sigma_{\hat{\beta}_1(s_i)}} \tag{7}$$

$$MAE = \frac{1}{N}\sum\left|\hat{\beta}_1(s_i) - \beta_1(s_i)\right| \tag{8}$$

$$CP(d) = P\left[\left|\hat{\beta}_1(s_i) - \beta_1(s_i)\right|\right] \leq d \tag{9}$$

$$SDR = \frac{\sigma_{\hat{\beta}_1(s_i)}}{\sigma_{\beta_1(s_i)}} \tag{10}$$

The correlation coefficient indicates the degree of linear association between the estimated coefficients and the reference ones. Absolute values close to unity indicate a nearly perfect relationship, but it does not necessarily imply the absence of bias in the estimates. The MAE represents the average absolute distance between the estimated coefficients and the reference ones. Ideally, this measure should be 0, but the lower it is, the lower the bias, and the higher it is the average concordance between the coefficients. Additionally, for each estimated map, the proportion of locations where the absolute distance between the estimated coefficients and the reference one was less than or equal to 15% relative to the reference one was calculated. In other words, this measure represents the proportion of the plot where the differences between the observed response coefficients and the true parameter value were less than the predefined error margin. Finally, the SDR captures the relationship between the variability of the coefficients estimated by GWR and those used in the simulation. When SDR > 1, the variability of the effect map obtained by GWR is higher than the variability of the true effect.

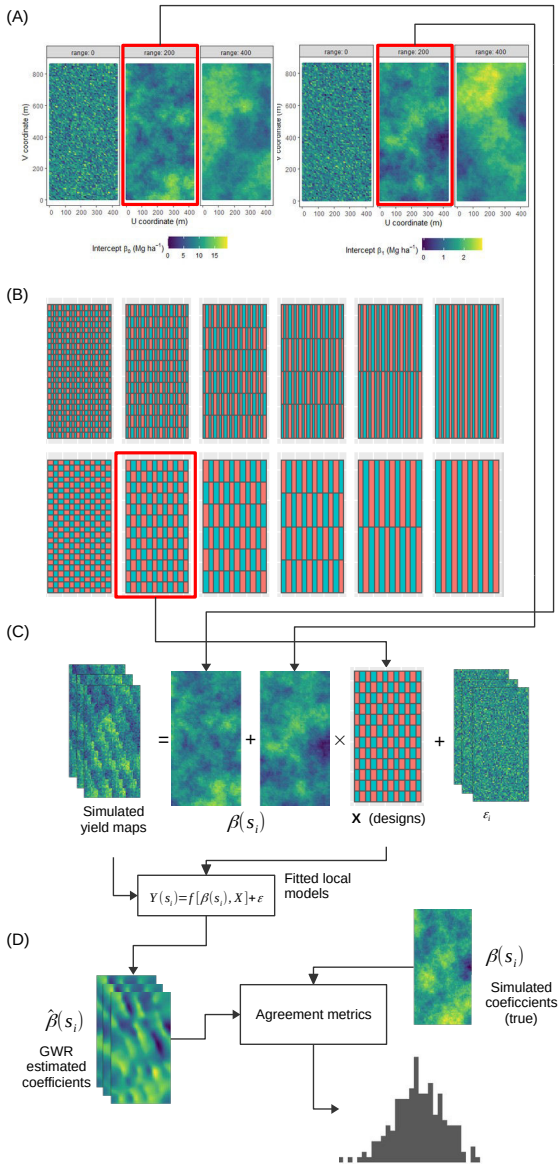

Fig. 1. Stages of the methodological workflow. (A) Generation of spatial variability scenarios of the coefficients $\beta_j(s_i)$ of the response model, (B) design of vector layers associated with each experiment, (C) simulation of yields and estimation of spatially variable coefficient models from the simulated coefficients in (C) and designs in (B), and (D) analysis of agreement between the estimated coefficients in (C) and the simulated coefficients in (A). The example considers design X and variability scenario Y.

As a result, for each combination of spatial structure (ranges) and designs, a sample of 200 values of each metric was obtained (Fig. 1D), and descriptive statistics were calculated on them. The effect of design and underlying spatial structure on the model's ability to capture the spatial structure of the effects was evaluated by fitting a full factorial model: 3 widths × 6 lengths × 3 ranges × 2 bandwidths × 3 kernels. The importance of the factors to explain metrics variability was quantified by calculating the principal and total sensitivity indices from the sums of squares (SS) of the ANOVA [26]. The principal sensitivity index (SP) captures the importance of the factor by discounting interaction effects, while the total sensitivity index (ST) includes the interaction effect. The SS of the factors were subsequently partitioned to evaluate specific contrasts between plot sizes or design types.

Data processing, visualization, and modeling were conducted using the statistical program R [27]. and functions from packages *gstat* [28], *sf* [29], *tidyverse* [30], and *GWmodel* [31, 32].

3 Results and Discussion

The median of the local coefficient of determination (R^2) of the models, grouping all the designs employed, ranged between 5% and 64% depending on the combination of the underlying variability scenario and the GWR parameters used, i.e., bandwidth and kernel (Fig. 2). Regardless of the kernel choice, the moderate to low proportion of variance explained by these models is inversely related to the amount of residual variation. In this study, simulations assumed a residual variation of 30% around the control yield. This residual variation was used to evaluate designs in scenarios of high variability that can cover real situations in experiments of this type at realistic scales in commercial fields of the Pampas region. Since the regression coefficients estimated by GWR depend on the weighting function [18, 19], it was observed that, regardless of the designs and bandwidth used, on average the highest local R^2 values were obtained with the bisquare kernel ($R^2 \approx 48\%$) compared to the exponential and Gaussian kernels (R^2 between 27% and 30%). The differences due to bandwidth were minor, with R^2 values consistently lower for larger bandwidths. On the other hand, it was also observed that the goodness of fit of the local models increased with the spatial structure of the simulated scenario. This is expected since as the spatial structure increases, there is more evidence to consider a model with non-stationary coefficients.

The importance of the spatial structure of the scenarios (depicted by the range), the characteristics of the experimental design (plot dimensions), and the parameters used in the GWR model (kernel and bandwidth) in the comparison between the estimated coefficients and the true ones are shown in Fig. 3. As can be seen, the spatial structure of the simulated scenario, alone or through interaction with other factors, had a significant impact on 3 out of the 4 evaluated metrics.

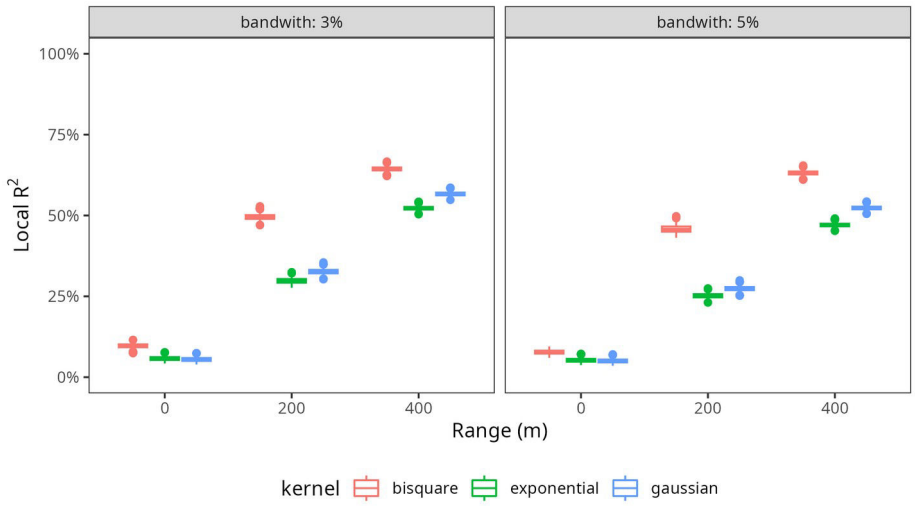

Fig. 2. Distribution of local coefficient of determination obtained by GWR models fitted using different combinations of kernel and bandwidth for spatial dependence ranges.

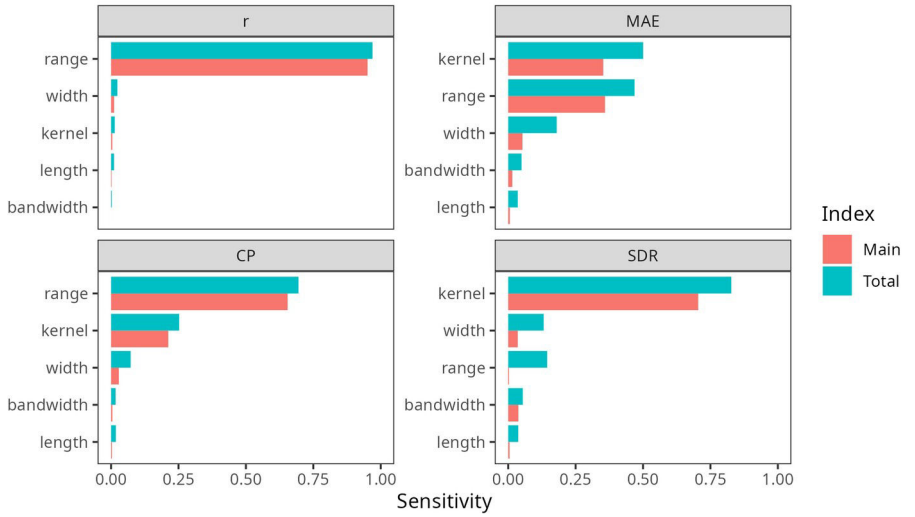

Fig. 3. Importance of the underlying spatial structure (range), the design (width and length), and parameters of the GWR model (kernel and bandwidth) in the variability of metrics of agreement between the coefficients estimated by GWR and the reference realization of the random field (Eq. 2): correlation coefficient (r), mean absolute error (MAE), coverage probability (CP), and standard deviation ratio (SDR). The main sensitivity index represents the proportion of explained variation accounted for by each factor alone. The total sensitivity index includes not only the main factor but all the interactions involving it.

The presence of spatial structure had a significant impact on the correlation between the maps of coefficients estimated by GWR and the reference map (Fig. 3). Nearly 95% of the total variation was associated with this factor alone. This effect can be primarily explained by the contrast between the scenario of no spatial structure and those with structured variability patterns. When the spatial structure was absent, the correlation coefficients obtained were nearly zero (r = 0.008), indicating that the use of GWR in such scenarios is unnecessary (Fig. 4). In scenarios with spatial structure, the average correlation coefficient was around 0.66. These differences accounted for 93% of the variation associated with this factor.

When analyzing the effect of plot size, differences were observed depending on the width. In general, narrower plots had better correlation values. In designs with wide plots (36 m), variable effects of plot length were observed, where shorter plots maintained high correlation values. When comparing strip designs with checkerboard designs, the checkerboard resulted in slightly higher correlation values than the strip designs. These differences were larger as the spatial structure had a greater range.

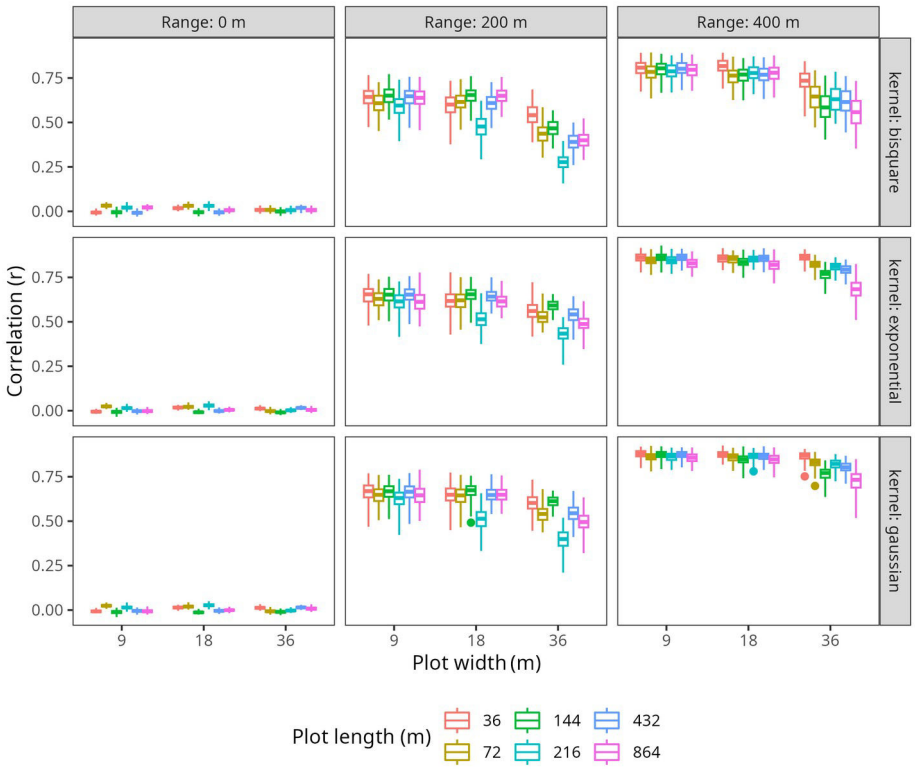

Fig. 4. Effect of spatial structure, kernel, and plot dimensions on the correlation of maps estimated by GWR with the reference random field realization.

Regarding the MAE, spatial structure and kernel were the factors with the greatest impact. Of the total variation associated with these factors and interactions, 78% and 70% were explained by the main effect, respectively (Fig. 3). With greater spatial structure, the error or bias in the estimation of the treatment effect was lower, with some differences observed among designs (Fig. 5). As in the previous case, the effect of spatial structure corresponds to the difference between the results of scenarios without structure (where GWR is not justified) and the average of scenarios with spatial structure. The effect of the kernel is due to the greater bias produced in the models where bisquare was used. The differences between exponential and Gaussian kernels are minimal. Unlike these kernels, where all combinations of plot length and width had comparable results, the bisquare kernel consistently produced higher errors in designs with wider and longer plots. The comparison between checkerboard and strip designs also indicated differences in favor of the former, due to the shorter length of the plots.

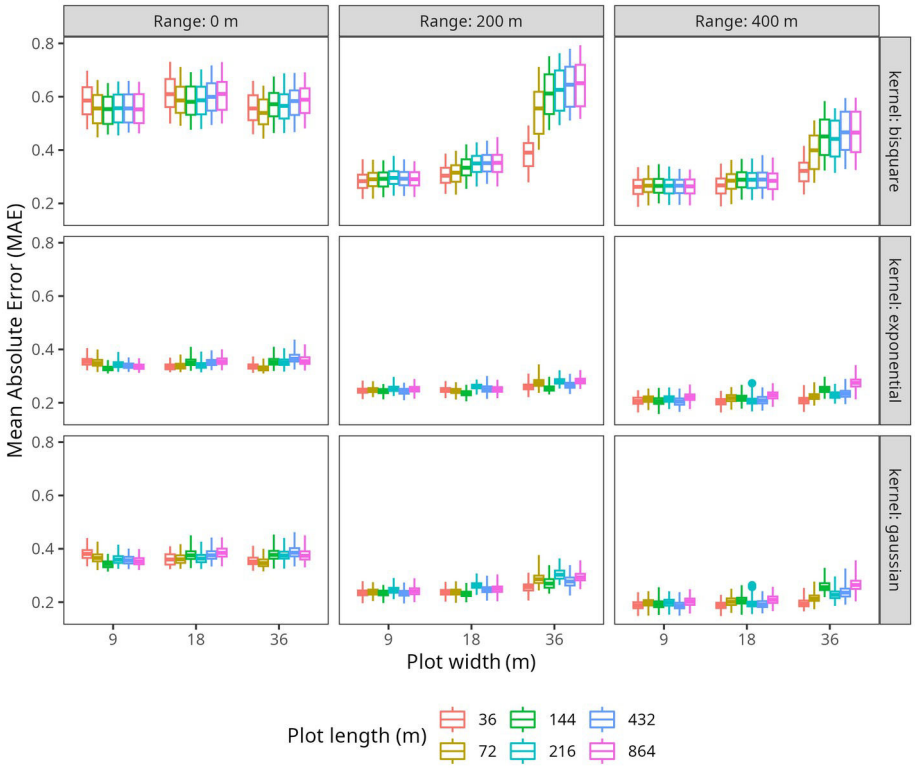

Fig. 5. Effect of spatial structure, kernel, and plot dimensions on the mean absolute error (MAE) and the correlation of the maps estimated by GWR with the reference random field realization.

The CP was mostly affected by the magnitude of the spatial structure (70% of total variation, 65.5% main effect), and to a lesser extent by the kernel (25% of total varia-

tion, 21% main effect) (Fig. 3). Conversely to the results for MAE, the proportion of the plot area where the difference between the estimated and true effect was less than or equal to 15% increased with the range of spatial structure. In the case of the bisquare kernel, in designs with wider plots, smaller increases in CP were observed compared to the exponential and Gaussian kernels, where designs were not differentiated and CP levels were higher due to the general effect of the kernel (Fig. 6). Similar to the previous metrics, the comparison between checkerboard and strip designs indicated differences in favor of the former, due to the shorter length of the plots.

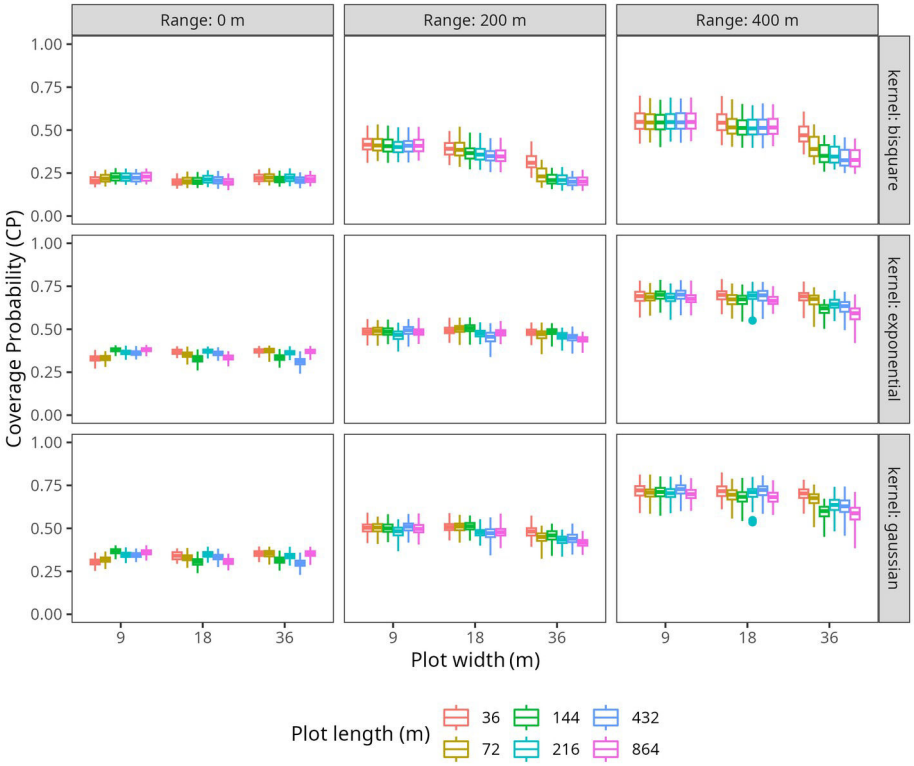

Fig. 6. Effect of spatial structure, kernel, and plot dimensions on the coverage probability (CP) of the maps estimated by GWR and the reference random field realization.

Finally, a high impact of the kernel on the SDR, i.e., the relationship between the variability estimated by the GWR coefficients and the underlying simulated variability, is observed (Fig. 3). The choice of kernel alone explained 70% of the total variation, while an additional 10% was explained by its interactions, mainly with plot dimensions. In this regard, the bisquare kernel tends to overestimate variability, with SDR values around 1.5, while the exponential, and to a lesser extent, the Gaussian underestimate it (Fig. 7). In both cases, over and underestimation decreases as the spatial structure becomes stronger. On the other hand, in the exponential and Gaussian ker-

nels, although they underestimate the underlying variability, the effect of plot dimensions, both length and width, is less pronounced. In contrast, with the bisquare kernel, designs with wider plots tend to overestimate more, even in scenarios of structured underlying variability. The comparison between checkerboard and strip designs resulted in differences in favor of checkerboard designs only in scenarios with long spatial dependence ranges.

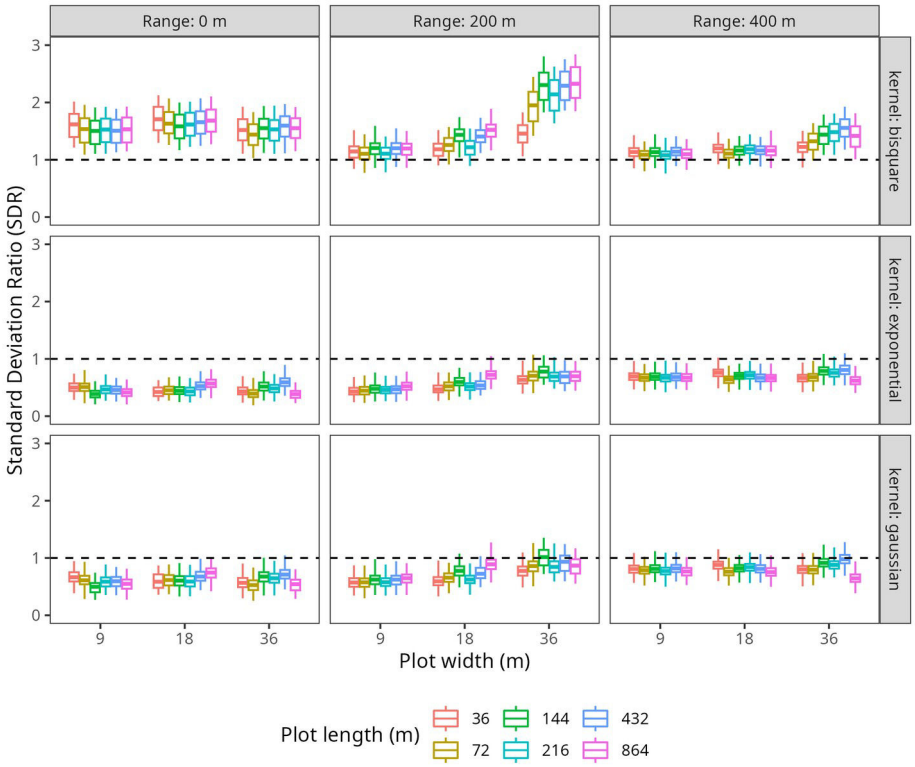

Fig. 7. Effect of spatial structure, kernel, and plot dimensions on the relationship between the variability estimated by GWR and that present in the reference random field realization.

The results presented highlight the impact of the experimental design used and the underlying spatial variability of the response on the ability of the GWR model to capture the spatial variability of crop response to variable input application, as well as the impact of alternative parametrization of this model. Since the spatial structure of the response is unknown and not under the experimenter's control when designing the field experiment, simulation studies like this allows for some conclusions to be drawn for designing field experiments at the plot scale.

In this regard, among the designs evaluated here, which were systematic, the checkerboard design with small plots represents an option that ensures good results

for the range of spatial variability conditions studied and options for parameterization of the GWR model. Similar results were reported by Alesso et al. [2]. In a similar study, Li et al. [22] observed that grid or checkerboard designs did not always show the best performance in terms of information return. The authors concluded that systematic designs with some restriction on spatial distribution, or even strip designs with high spatial heterogeneity, had comparable results.

From the perspective of field application, complex designs such as checkerboard with spatial balance are less adopted because they represent a greater challenge for the prescription design and field implementation, as they requires more precise and expensive variable application systems (e.g., electronic rate controllers, high precision GPS receivers, etc.) to execute dose changes over short distances. In contrast, strip designs are simpler to execute, even without VRT technology. However, it is advisable to balance simplicity and the rigor of the results.

Finally, while the general recommendation regarding plot dimensions is: the smaller dimension, the greater the chance of capturing the underlying spatial variability, the minimum size of plots is limited by machinery footprint, e.g., header or planter width, the smoothing process along the harvesting direction [33, 34], and the accuracy of positioning systems.

4 Conclusions

This study confirmed the effects of underlying spatial structure on the ability of the GWR model to estimate the spatial variability of treatment effects, and its interaction with the field designs. The significant effect of model parameters on the obtained response maps was also verified. In summary, greater spatial structure led to better performance of the GWR model in capturing the non-stationarity of coefficients. Among the evaluated designs, the width and length of the plots had an impact on the evaluated metrics. Larger plot sizes resulted in lower agreement between the coefficients estimated by GWR and the simulated random field. The greatest impact was observed in the measure of mean absolute error (MAE) and the overestimation of variability when the bisquare kernel was employed. Finally, the use of designs with smaller plots, especially checkerboards, improves the estimation of underlying response variability compared to designs with larger plot dimensions.

5 Acknowledgments

This work was supported by the following grants FONCyT-PICT-2021-GRF-TI-00795 and ASACTEI-PEICI+D-2021-180.

6 References

1. Maat, H.: The history and future of agricultural experiments. NJAS - Wageningen Journal of Life Sciences. 57, 187–195 (2011). https://doi.org/10.1016/j.njas.2010.11.001.
2. Alesso, C.A., Cipriotti, P.A., Bollero, G.A., Martin, N.F.: Design of on-farm precision experiments to estimate site-specific crop responses. Agronomy Journal. 113, 1366–1380 (2021). https://doi.org/10.1002/agj2.20572.
3. Panten, K., Bramley, R.G.V., Lark, R.M., Bishop, T.F.A.: Enhancing the value of field experimentation through whole-of-block designs. Precision Agriculture. 11, 198–213 (2010). https://doi.org/10.1007/s11119-009-9128-y.
4. Hicks, D., Vanden Heuvel, R., Fore, Z.: Analysis and practical use of information from on-farm strip trials. Better Crops. 81, 18–21 (1997).
5. Plant, R.E.: Comparison of means of spatial data in unreplicated field trials. Agronomy Journal. 99, 481–488 (2007). https://doi.org/10.2134/agronj2005.0150.
6. Alesso, C.A., Cipriotti, P.A., Bollero, G.A., Martin, N.F.: Experimental Designs and Estimation Methods for On-farm Research: A Simulation Study of Corn Yields at Field Scale. Agronomy Journal. 111, 2724–2735 (2019). https://doi.org/10.2134/agronj2019.03.0142.
7. Whelan, B.M., Taylor, J.A.: Precision agriculture for grain production systems. CSIRO Publ, Collingwood (2013).
8. Bullock, D.S., Bullock, D.G.: From agronomic research to farm management guidelines: A primer on the economics of information and precision technology. Precision Agriculture. 2, 71–101 (2000).
9. Pringle, M.J., Cook, S.E., McBratney, A.B.: Field-Scale Experiments for Site-Specific Crop Management. Part I: Design Considerations. Precision Agriculture. 5, 617–624 (2004). https://doi.org/10.1007/s11119-004-6346-1.
10. Piepho, H.P., Richter, C., Spilke, J., Hartung, K., Kunick, A., Thöle, H.: Statistical aspects of on-farm experimentation. Crop and Pasture Science. 62, 721 (2011). https://doi.org/10.1071/CP11175.
11. Scharf, P.C., Kitchen, N.R., Sudduth, K.A., Davis, J.G., Hubbard, V.C., Lory, J.A.: Field-Scale Variability in Optimal Nitrogen Fertilizer Rate for Corn. Agronomy Journal. 97, 452–461 (2005). https://doi.org/10.2134/agronj2005.0452.
12. Kyveryga, P.M., Blackmer, A.M., Zhang, J.: Characterizing and Classifying Variability in Corn Yield Response to Nitrogen Fertilization on Subfield and Field Scales. Agronomy Journal. 101, 269–277 (2009). https://doi.org/10.2134/agronj2008.0168.
13. Kindred, D.R., Milne, A.E., Webster, R., Marchant, B.P., Sylvester-Bradley, R.: Exploring the spatial variation in the fertilizer-nitrogen requirement of wheat within fields. The Journal of Agricultural Science. 153, 25–41 (2015). https://doi.org/10.1017/S0021859613000919.
14. Panten, K., Bramley, R. g. v.: Whole-of-block experimentation for evaluating a change to canopy management intended to enhance wine quality. Australian

Journal of Grape and Wine Research. 18, 147–157 (2012). https://doi.org/10.1111/j.1755-0238.2012.00183.x.
15. Pringle, M.J., McBratney, A.B., Cook, S.E.: Field-Scale Experiments for Site-Specific Crop Management. Part II: A Geostatistical Analysis. Precision Agriculture. 5, 625–645 (2004). https://doi.org/10.1007/s11119-004-6347-0.
16. Bishop, T.F.A., Lark, R.M.: The geostatistical analysis of experiments at the landscape-scale. Geoderma. 133, 87–106 (2006). https://doi.org/10.1016/j.geoderma.2006.03.039.
17. Trevisan, R.G., Bullock, D.S., Martin, N.F.: Spatial variability of crop responses to agronomic inputs in on-farm precision experimentation. Precision Agric. 22, 342–363 (2020). https://doi.org/10.1007/s11119-020-09720-8.
18. Fotheringham, A.S., Brunsdon, C., Charlton, M.: Geographically weighted Regression: the analysis of spatially varying relationships. Wiley, Chichester, England (2002).
19. Páez, A., Farber, S., Wheeler, D.: A Simulation-Based Study of Geographically Weighted Regression as a Method for Investigating Spatially Varying Relationships. Environ Plan A. 43, 2992–3010 (2011). https://doi.org/10.1068/a44111.
20. Bivand, R.S., Pebesma, E., Gómez-Rubio, V.: Applied Spatial Data Analysis with R. Springer New York, New York, NY (2013). https://doi.org/10.1007/978-1-4614-7618-4.
21. Farber, S., Páez, A.: A systematic investigation of cross-validation in GWR model estimation: empirical analysis and Monte Carlo simulations. J Geograph Syst. 9, 371–396 (2007). https://doi.org/10.1007/s10109-007-0051-3.
22. Li, X., Mieno, T., Bullock, D.S.: The economic performances of different trial designs in on-farm precision experimentation: a Monte Carlo evaluation. Precision Agric. 24, 2500–2521 (2023). https://doi.org/10.1007/s11119-023-10050-8.
23. Webster, R., Oliver, M.A.: Geostatistics for enviromental sicentists. Wiley, Chichester, England (2007).
24. Richter, C., Kroschewski, B., Piepho, H.-P., Spilke, J.: Treatment comparisons in agricultural field trials accounting for spatial correlation. Journal of Agricultural Science. 153, 1187–1207 (2015). https://doi.org/10.1017/S0021859614000823.
25. Thöle, H., Richter, C., Ehlert, D.: Strategy of statistical model selection for precision farming on-farm experiments. Precision Agric. 14, 434–449 (2013). https://doi.org/10.1007/s11119-013-9306-9.
26. Wallach, D., Makowski, D., Jones, J.W.: Working with Dynamic Crop Models: Evaluation, Analysis, Parameterization, and Applications. Elsevier Science (2006).
27. R Core Team: R: A Language and Environment for Statistical Computing, https://www.R-project.org/, (2023).
28. Pebesma, E.: Spatio-temporal geostatistics using gstat. (2011).
29. Pebesma, E.: Simple Features for R: Standardized Support for Spatial Vector Data. The R Journal. 10, 439–446 (2018). https://doi.org/10.32614/RJ-2018-009.
30. Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T.L., Miller, E., Bache, S.M., Müller, K., Ooms, J., Robinson, D., Seidel, D.P., Spinu,

V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., Yutani, H.: Welcome to the tidyverse. Journal of Open Source Software. 4, 1686 (2019). https://doi.org/10.21105/joss.01686.

31. Gollini, I., Lu, B., Charlton, M., Brunsdon, C., Harris, P.: GWmodel : An R Package for Exploring Spatial Heterogeneity Using Geographically Weighted Models. J. Stat. Soft. 63, (2015). https://doi.org/10.18637/jss.v063.i17.
32. Lu, B., Harris, P., Charlton, M., Brunsdon, C.: The GWmodel R package: further topics for exploring spatial heterogeneity using geographically weighted models. Geo-spatial Information Science. 17, 85–101 (2014). https://doi.org/10.1080/10095020.2014.917453.
33. Lark, R.M., Stafford, J.V., Bolam, H.C.: Limitations on the spatial resolution of yield mapping for combinable crops. J. agric. Engng Res. 66, 183–193 (1997).
34. Marchant, B., Rudolph, S., Roques, S., Kindred, D., Gillingham, V., Welham, S., Coleman, C., Sylvester-Bradley, R.: Establishing the precision and robustness of farmers' crop experiments. Field Crops Research. 230, 31–45 (2019). https://doi.org/10.1016/j.fcr.2018.10.006.