

## **Evaluación de diseños experimentales y métodos de análisis para ensayos en campos de productores con herramientas de agricultura de precisión**

C. Agustín Alesso<sup>1</sup>, Pablo A. Cipriotti<sup>2,3</sup>, Germán A. Bollero<sup>4</sup>

<sup>1</sup> Facultad de Ciencias Agrarias, Universidad Nacional del Litoral (FCA-UNL). Kreder 2805, S3080HOF, Esperanza. e-mail: [calesso@fca.unl.edu.ar](mailto:calesso@fca.unl.edu.ar).

<sup>2</sup> Facultad de Agronomía, Universidad de Buenos Aires (FAUBA). Av. San Martín 4453, C1417DSE, Buenos Aires. e-mail: [cipriott@agro.uba.ar](mailto:cipriott@agro.uba.ar)

<sup>3</sup> Instituto de Investigaciones Fisiológicas y Ecológicas vinculadas a la Agricultura (IFEVA), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). Av. San Martín 4453, C1417DSE, Buenos Aires.

<sup>4</sup> Department of Crop Sciences, University of Illinois at Urbana-Champaign. 228 Mumford Hall, 61801 IL, United States. e-mail: [gbollero@illinois.edu](mailto:gbollero@illinois.edu)

**Resumen.** Los costos de realizar experimentos en campos de productores se ha reducido con el desarrollo de agricultura de precisión. No obstante, es necesario evaluar los diseños comúnmente utilizados considerando la variabilidad espacial y autocorrelación de los datos. Se evaluaron diseños experimentales sistemáticos simulando condiciones experimentales a campo con distinta estructura de autocorrelación espacial. Los diseños se diferenciaron en la eficiencia de los estimadores como resultado de su sensibilidad al efecto de la estructura espacial. En general, el grado de de estructura espacial (proporción varianza nugget) disminuyó la precisión de los estimadores y aumentó la tasa de error simulada tipo I, siendo mayor el efecto en diseños con parcelas grandes. El diseño en tablero de ajedrez mostró mejor performance general. El método de estimación GLS (con parámetros fijos) mejoró las propiedades de los estimadores obtenidos.

**Palabras Claves:** geoestadística, simulación, estimadores

### **1 Introducción**

La experimentación es una de las formas de ganar conocimiento sobre la respuesta de los diversos sistemas productivos. En la investigación agrícola, los métodos tradicionales de experimentación a campo han jugado un rol importante permitiendo el desarrollo y generalización de recomendaciones agronómicas desde

la elección de cultivares a nivel regional hasta dosis de fertilización sitio específicas [1]. Típicamente, los experimentos agrícolas son conducidos en estaciones experimentales para lograr un mayor control de las condiciones experimentales y mejorar la precisión de las estimaciones. No obstante, los experimentos o ensayos en campo de productores tienen una mayor relevancia práctica aunque menor precisión debido a la mayor heterogeneidad dentro y entre las unidades experimentales [2].

En general, los productores y asesores realizan experimentos en sus campos para evaluar la conveniencia de prácticas agronómicas alternativas o corroborar las recomendaciones obtenidas de los centros de investigación [1]. En la última década, este tipo de experimentos se ha facilitado gracias al desarrollo de herramientas de Agricultura de Precisión (AP). Así la tecnología de dosificación variable permite la modificación de las prácticas de manejo dentro del lote (e.g. densidad de siembra, dosis de fertilizante, dosis de pesticidas, etc.) y los monitores de rendimiento montados en cosechadoras y sensores de cultivo permiten la cuantificación de la respuesta del cultivo [3]. No obstante, si bien estas tecnologías han reducido el costo de la experimentación en campos comerciales, no los han eliminado por completo y la falta de un diseño experimental apropiado en la mayoría de los casos puede ocasionar la toma de decisiones basadas en conclusiones erróneas [4].

Los experimentos en campo de productores pueden orientarse a estimar el efecto global de una determinada práctica de manejo, sin la necesidad de contar con datos georreferenciados, o bien para explorar la variabilidad de la respuesta del cultivo dentro del lote utilizando datos georreferenciados [2]. En este último caso, se puede poner a prueba la hipótesis nula de la agricultura de precisión, i.e. si existe heterogeneidad espacial en la respuesta del cultivo que justifique el manejo diferenciado de insumos dentro del lote [5]; o asumiendo dicha premisa, estimar una función de respuesta del cultivo para obtener mapas de prescripción variable [6]. De este modo, se descarta el supuesto de aditividad de los efectos de los tratamientos y las condiciones ambientales dentro del lote y se explora el efecto sitio-específico de los tratamientos con el objetivo final del manejo sitio-específico del cultivo, continuo o por zonas de manejo [2].

En general los principios básicos del diseño de experimentos (i.e. aleatorización, replicación y control local) garantizan la obtención de pruebas estadísticas válidas, evitando potenciales fuentes de sesgo y controlando la variación exógena que puede ser confundida con el efecto de los tratamientos [1, 2, 6–8]. No obstante, las características propias de los sistemas de producción así como también la configuración de la maquinaria utilizada imponen ciertas restricciones que deben ser contempladas en el diseño de experimentos en campo de productores [7, 9]. Según Willers et al. [9], en los experimentos a campo con herramientas de AP existen distintos tipos de unidades experimentales producto de la combinación de las características del terreno y el tipo de maquinaria y los manejos evaluados. Más aún, los datos obtenidos con herramientas de AP a menudo resultan de tratamientos asignados de manera sistemática a parcelas grandes con observaciones repetidas que presentan correlación espacial [6]. De este modo, aunque las operaciones de campo

pueden simplificarse, la violación de los principios de experimentación clásica presenta nuevos desafíos para el diseño y análisis de experimentos a fin de obtener conclusiones válidas [2]. El desarrollo de las herramientas de AP y los métodos de estadística espacial han creado nuevas oportunidades para el diseño y análisis de experimentos en campos de productores [10]. Según Whelan et al. [7], los diseños de experimentos deben moverse dentro de las limitantes técnicas, agronómicas y económicas, pero al mismo tiempo ser robustos para inferir estadísticamente distintos atributos del fenómeno bajo estudio.

Las primeras aproximaciones a este tipo de problema han derivado en el uso de diseños altamente complejos con arreglos espaciales tales como tableros de ajedrez [5], patrones sinusoidales bidimensionales [6] o diseños en bloques en todo el campo [1], y la aplicación de técnicas estadísticas avanzadas para modelar los distintos tamaños de unidades experimentales y los efectos aleatorios introducidos por las operaciones de campo [11]. Si bien estos diseños son estadísticamente robustos y optimizados desde el punto de vista de la ganancia de información, el nivel de complejidad requiere un alto nivel de conocimiento agronómico y estadístico para su diseño y análisis; más aún pueden conllevar a una considerable pérdida de producción y costos debido a la inclusión de tratamientos extremos (i.e. controles o dosis extremas) [4, 7].

En contrapartida, recientemente se han propuesto aproximaciones más simplistas tales como diseños en franjas sin réplicas [12] o replicadas en todo el lote [13], o bien pequeñas franjas [7]. Estos diseños son más aceptados por los productores y su implementación no requiere alta tecnología. No obstante, la auto-correlación de las observaciones sobrestima el número de observaciones independientes para evaluar los tratamientos y la orientación de las franjas sin conocimiento previo de los patrones de variación espacial puede resultar en un diseño ineficiente para capturar la variabilidad espacial [7].

En la literatura se han propuesto diversos diseños experimentales y enfoques analíticos para obtener información útil a partir de ensayos a campo con herramientas de AP, no obstante los estudios comparativos de estos diseños son limitados (e.g. Griffin et al. [10]). Es por ello que para avanzar en el diseño y optimización de experimentos en campos de productores es necesario realizar una evaluación comparativa de los mismos en términos de distintas propiedades de los estimadores, tales como la insesgabilidad (i.e. que el valor esperado del estimador coincida con el parámetro), la eficiencia (que la varianza de la distribución muestral del estimador sea mínima), y al mismo tiempo estimar la tasa de error tipo I al declarar diferencias entre tratamientos.

Los objetivos del presente trabajo son: (i) comparar mediante técnicas de simulación estadística, el sesgo y varianza de los estimadores obtenidos por dos métodos de estimación y la tasa error tipo I de diseños experimentales sistemáticos comúnmente empleados en ensayos de campo utilizando herramientas de agricultura de precisión, y (ii) evaluar el impacto de la estructura espacial de la variabilidad

dentro de las unidades experimentales en la precisión de los estimadores y las pruebas de hipótesis asociadas.

## **2 Materiales y Métodos**

Se realizó un estudio de simulación estadística para comparar diseños experimentales utilizados comúnmente en ensayos en campos de productores analizando las propiedades de los estimadores y la tasa de error tipo I .

### **2.1 Experimento y condiciones experimentales hipotéticas**

En el presente ejercicio se simuló la evaluación de dos tratamientos hipotéticos (A y B) en un cultivo de maíz. Los tratamientos pueden ser aplicados en la operación de siembra (e.g. fertilizante) utilizando sembradora de dosificación variable ya que el productor dispone de: (1) tractor con piloto automático; (2) sembradora de grano grueso de 16 surcos a 0,52 m, equipada con sistema de dosificación variable; (3) cosechadora con cabezal maicero de 8 surcos a 0,52 m, equipadas con monitor de rendimiento y piloto automático. De este modo, el ancho mínimo de las parcelas experimentales (franjas o parcelas) queda definido según el ancho del cabezal o medio ancho de la sembradora, i.e. 4,16 m.

Se consideraron dos áreas experimentales rectangulares donde se podría realizar el experimento, ambas de aproximadamente 100 m de ancho (equivalente a 12 anchos de sembradora o 24 de cosechadora) y 200 m de largo. En una de las áreas el relieve es plano o subnormal y en la otra se presenta un gradiente topográfico (loma, media loma y bajo) en dirección N-S (lado más largo). Este gradiente topográfico explicaría parte de la variación de los rendimientos siendo aproximadamente 20% superior al promedio en el bajo y 20% inferior en la loma.

Finalmente, se asume que la resolución espacial de la información de rendimientos es representada mediante píxeles de forma cuadrada con dimensiones iguales al ancho de la cosechadora. Como resultado, el área experimental está representada por una grilla rectangular de 1152 píxeles con resolución de 4,16 m x 4,16 m. Esto último responde a la necesidad de simular distintos tamaños de unidad experimental, número de repeticiones y gradiente topográfico, manteniendo el número de nodos de la grilla relativamente bajo.

### **2.2 Mapas de rendimiento simulados: campos aleatorios**

La variabilidad espacial de los rendimientos del cultivo (maíz) dentro del área de estudio se simuló mediante el procedimiento de simulación geoestadística gaussiana

no condicionada [14]. Los campos aleatorios se generaron a partir del siguiente modelo general:

$$z(\mathbf{s}) = \mu + \varepsilon(\mathbf{s}) \quad (1)$$

donde:  $z(\mathbf{s})$  es el rendimiento en la posición  $s = \{\mathbf{x}, \mathbf{y}\}$ ,  $\mu$  la media general del proceso y  $\varepsilon(\mathbf{s})$  un componente aleatorio con distribución Normal con esperanza 0 y matriz de varianza-covarianza tipo esférica con efecto pepita definida en función de la distancia  $\mathbf{h}$ :

$$C_{\text{sph+nug}}(\mathbf{h}) = \begin{cases} \tau^2 + \sigma^2 \left( 1 - \frac{3\|\mathbf{h}\|}{2a} + \frac{1\|\mathbf{h}\|^3}{2a^3} \right) & 0 \leq \|\mathbf{h}\| \leq a \\ 0 & \|\mathbf{h}\| > a \end{cases} \quad (2)$$

donde:  $a$  es el alcance o distancia a partir de la cual no hay correlación entre rendimientos,  $\tau^2$  es la varianza no estructurada o nugget y  $\sigma^2$  es la varianza estructurada o *partial sill*. La varianza del proceso es  $\tau^2 + \sigma^2 = 64$  qq ha<sup>-1</sup>. Luego, se definieron escenarios combinando los siguientes componentes del modelo espacial:

a. *Estacionariedad de primer orden (modelo de tendencia)*: se consideraron dos escenarios hipotéticos donde los rendimientos son generados por (1) proceso estacionario donde la media poblacional es una constante y (2) proceso no estacionario donde la media poblacional es una función polinómica de tercer orden de la coordenada Y. Estas condiciones representan situaciones comúnmente observadas en la zona: campo con relieve subnormal y campo con relieve normal (toposecuencia de tipo loma-media loma-bajo). Los modelos empleados fueron: (1)  $\mu = 80$ ; y (2)  $\mu = 64,8 - 0,245Y + 0,00453Y^2 - 0,000013Y^3$

b. *Proporción varianza no estructurada (efecto pepita o nugget)*: se simularon procesos con distinto grado de estructuración espacial de la varianza variando la proporción de varianza no estructurada (pepita) sobre la varianza total del proceso. Así un proceso fuertemente estructurado tiene 0% de efecto pepita, por el contrario, si el efecto pepita es 100% el proceso se considera sin estructura espacial. Los niveles de proporción de varianza no estructurada simulados fueron: 0, 20, 40, 60 y 80%. Mapas de rendimiento con 100% de varianza no estructurada son muy poco frecuentes.

c. *Alcance de la autocorrelación espacial (range)*: se consideraron escenarios donde los rendimientos estaban espacialmente autocorrelacionados hasta 20, 40 y 60 m de distancia.

La siguiente tabla resume los 30 escenarios de autocorrelación espacial utilizados para las simulaciones geoestadísticas. Cada escenario se simuló 200 veces obteniéndose 200 realizaciones igualmente probables de la respuesta del cultivo asumiendo que el efecto de los tratamientos fue nulo  $\tau_A = \tau_B = 0$ .

**Tabla 1.** Escenarios de autocorrelación espacial utilizados en la simulación de campos aleatorios mediante el procedimiento de simulación geoestadística gaussiana no condicionada, asumiendo covarianza esférica con varianza 64 qq ha<sup>-1</sup>. En los márgenes los niveles de los atributos modificados y en las celdas el identificador del escenario.

Tendencia	Rango(m)	Efecto pepita (%)				
		0	20	40	60	80
Constante	20	E1	E2	E3	E4	E5
	40	E6	E7	E8	E9	E10
	60	E11	E12	E13	E14	E15
Cúbica	20	E16	E17	E18	E19	E20
	40	E21	E22	E23	E24	E25
	60	E26	E27	E28	E29	E30

### 2.3 Diseños experimentales

Dada la configuración de la maquinaria descrita en el punto 2.1, se asume que los tratamientos pueden asignarse a una resolución espacial mínima igual a medio ancho de sembradora en dirección perpendicular al sentido de avance, y dos veces el ancho de sembradora en el sentido de avance. Esta última restricción responde al tiempo de respuesta para el cambio de dosis del equipo de dosificación variable.

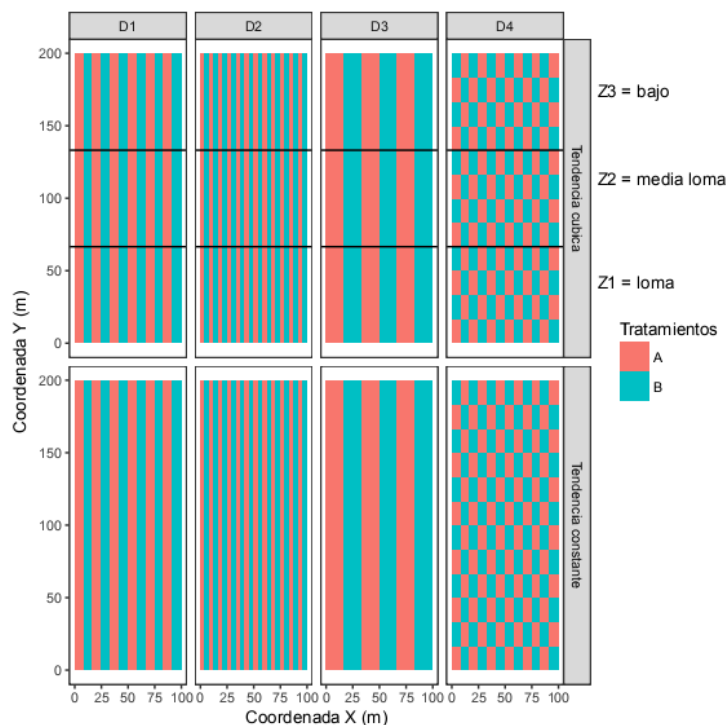
Teniendo en cuenta las restricciones de la maquinaria se consideraron 4 diseños sistemáticos utilizando sembradora entera o media sembradora con dosis uniforme y sembradora entera con dosificación variable. En el área con gradiente topográfico los diseños fueron aplicados agregando zonas delimitadas por topografía (loma, media loma y bajo) como bloques (Tabla 2). En la Fig. 1 se representan gráficamente cada uno de los 8 diseños evaluados..

**Tabla 2.** Resumen de los diseños sistemáticos aplicados en área sin gradiente topográfico, número de réplicas y tamaño de las unidades experimentales (UE). En áreas con gradiente topográfico se agregó un factor correspondiente a las zonas (bloques)

Cód.	Nombre	Réplicas	Sup. UE (m <sup>2</sup> )
D1	Franjas	6	1661
D2	Split-planter	12	831
D3	Franjas dobles (ida y vuelta)	3	3323
D4	Tablero de ajedrez	36	138

Los modelos lineales asociados a cada diseño fueron: (1) diseños sin bloques:  $y_{ij(k)} = \mu + \tau_i + \varepsilon_{ij(k)}$  y (2) diseños con bloques:  $y_{ij(k)} = \mu + \tau_i + \rho_j + \varepsilon_{ij(k)}$  ,

donde:  $y_{ij(k)}$  es el rendimiento correspondiente al  $k$ -ésimo pixel dentro de la  $j$ -ésima unidad experimental sometida al  $i$ -ésimo tratamiento.  $\mu$  representa la media general,  $\tau_i$  el efecto del  $i$ -ésimo tratamiento,  $\rho_j$  el efecto del  $j$ -ésimo bloque (zona) y  $\varepsilon_{ij(k)}$  el error experimental asociado al  $k$ -ésimo pixel dentro de la  $j$ -ésima unidad experimental sometida al  $i$ -ésimo tratamiento.



**Fig. 1.** Esquema de los diseños experimentales utilizados (D1-D4) corresponden a los diseños aplicados en el área sin gradiente topográfico (tendencia constante) y con gradiente topográfico (tendencia cúbica). En este último caso se añadieron las zonas (Z1, Z2, Z3) como bloques.

## 2.4 Métodos de análisis de datos

Los datos de rendimiento simulados con los 30 campos aleatorios definidos en la Tabla 1 fueron analizados ajustando los modelos lineales descritos en el punto 2.3. A cada una de las 200 realizaciones de cada campo aleatorio se le aplicó uno de los 4 diseños correspondientes según sea un campo con o sin gradiente topográfico. Los parámetros de los modelos se estimaron mediante dos procedimientos: mínimos cuadrado ordinarios (OLS, del inglés *ordinary least-squares*) y mínimos cuadrados

generalizados espacial (GLS, del inglés *generalized least-squares*) [10, 11]. En ambos casos se asumió homocedasticidad, i.e. igual varianza entre tratamientos. En total se estimaron 48000 modelos de los cuales se extrajeron los coeficientes y estadísticas asociadas.

El procedimiento OLS se basa en el supuesto clásico de independencia de los errores, lo que equivale a una estructura de covarianza espacial de tipo de efecto pepita puro [15]. Cuando este supuesto se cumple, OLS devuelve el mejor estimador lineal insesgado (BLUE, por sus siglas en inglés) a partir de los datos. En cambio, si las observaciones están autocorrelacionadas, los estimadores OLS siguen siendo insesgados pero no son eficientes, i.e. menor varianza [16]. Por su parte, el procedimiento GLS permite incorporar la estructura de autocorrelación de los errores definida en función de la distancia de separación de las observaciones [11] obteniendo estimadores BLUE mediante máxima verosimilitud restringida (REML, del inglés *restricted maximum likelihood*).

## 2.5 Comparación de diseños

La significancia del efecto de los tratamientos fue evaluada en cada combinación de escenario-diseño-método mediante análisis de la varianza (ANOVA). De cada modelo se extrajeron los valores p correspondientes a la prueba de hipótesis sobre el efecto del tratamiento para estimar la tasa de error tipo I simulada, dado que  $H_0 : \tau_i = 0$ .

El sesgo y eficiencia de los estimadores se evaluó examinando las características de la distribución (media y varianza) de los coeficientes estimados aproximada mediante el procedimiento de simulación considerando que el efecto de los tratamientos fue nulo ( $\tau_i = 0$ ). El sesgo se calculó como la diferencia entre el valor esperado del estimador (promedio de la distribución) y el verdadero valor del estimador. Por su parte, la varianza es una medida de la dispersión de la distribución del estimador.

La performance global de los estimadores se estimó mediante la raíz cuadrada del error cuadrático medio (RMSE, por sus siglas en inglés *root mean squared error*) el cual integra la medida de sesgo y varianza. Así, estimadores con bajo RMSE tienen bajos valores de sesgo y varianza a la vez.

## 2.6 Software

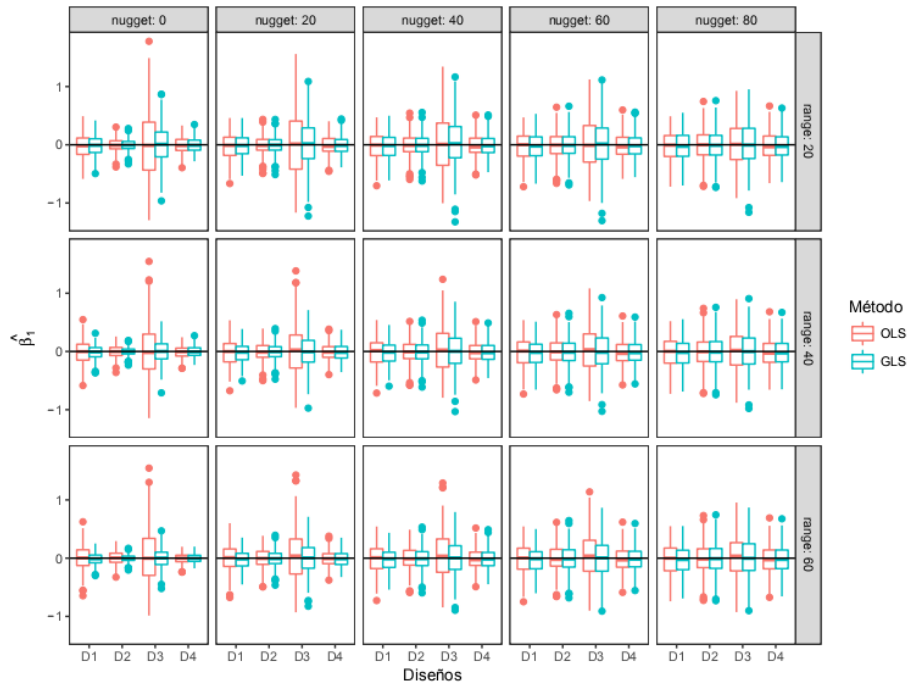
El análisis se desarrolló utilizando el entorno estadístico R [17] y los paquetes *gstat* [18]. La manipulación y visualización de datos a través de los paquetes *dplyr* [19] y *ggplot2* [20]. El ajuste de modelos OLS y GLS con el paquete *nlme* [21]



### 3 Resultados y Discusión

El valor central de las distribuciones de los  $\hat{\beta}_1$  obtenidos por simulación para cada diseño no difiere sustancialmente respecto del valor verdadero ( $\tau = \beta_1 = 0$ ). Tampoco se observan diferencias entre métodos de estimación, y combinaciones de range y nugget (Fig. 2). En cambio, la variabilidad de la distribución si estuvo afectada por el diseño experimental aplicado en combinación con las características de la estructura de variabilidad y el método de estimación. En general, el procedimiento GLS resultó en distribuciones más concentradas en todos los diseños. Las diferencias entre OLS y GLS disminuyen conforme disminuye el grado de estructuración espacial (incrementos de proporción de varianza no estructurada o *nugget*). Para un valor fijo de *nugget*, el efecto del rango es menos importante. En cuanto a las diferencias de los diseños, éstas son más importantes cuando el método de estimación no considera la estructura espacial (i.e. OLS). En los diseños con unidades experimentales más grandes (D3, franjas ida y vuelta), la dispersión de los  $\hat{\beta}_1$  es 2 o 3 veces mayor comparada con diseños tablero de ajedrez (D4) o bien split-planter (D2). Esto podría explicarse por el tamaño de las unidades experimentales, que en el caso de D2 y D4 se aproximan a las dimensiones de la grilla simulada.

Al igual que lo observado para  $\hat{\beta}_1$ , el valor central de las distribuciones de  $\hat{\beta}_0$ , i.e. la media general, obtenidos por simulación para cada diseño no difiere sustancialmente respecto del valor verdadero ( $\mu = 80$ ). En escenarios sin tendencia se observa ligera disminución de la dispersión de los estimadores estimados por GLS. En escenarios con tendencia, la variabilidad de la distribución de los estimadores son similares entre métodos pero con un ligero sesgo positivo para GLS y negativo para OLS. En cuanto al efecto de las combinaciones de varianza no estructurada y alcance de dependencia espacial, se verificó que a mayor proporción de varianza no estructurada (más *nugget*) o menos distancia de autocorrelación (menor *range*), la variabilidad de los estimadores es menor. Es decir, ante una estructura más débil o de corta distancia la precisión de los estimadores sería mayor.



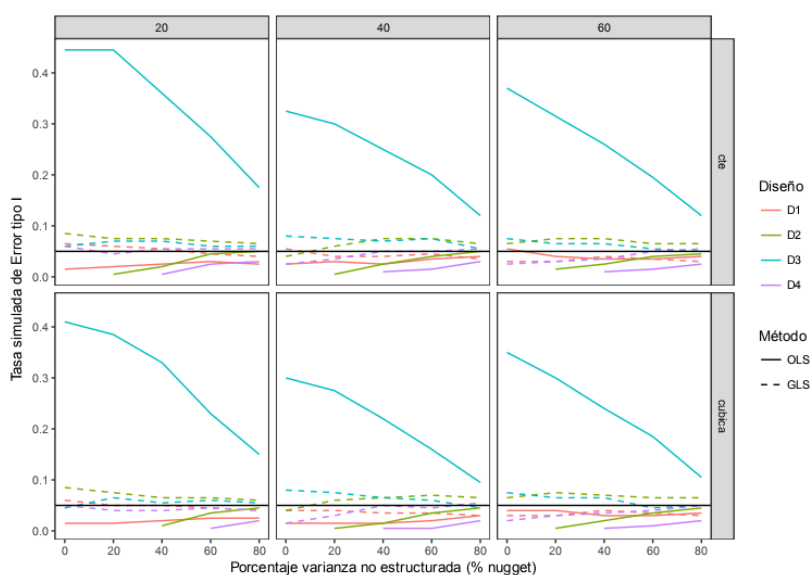
**Fig. 2.** Relación entre la distribución de los estimadores  $\hat{\beta}_1$  (efecto de tratamiento) y los diseños (D1-D4), métodos de ajuste (OLS y GLS) y escenarios de variabilidad espacial (*nugget* y *range*). La línea continua indica el valor verdadero del parámetro  $\tau = \beta_1 = 0$ .

En la Fig. 3 se presentan las tasas de error simulada tipo I, i.e. rechazo de  $H_0$ ;  $\tau_i = 0$ , dado que  $H_0$  es verdadera) en función de los escenarios de variabilidad y diseños. En general, cuando los datos se modelan utilizando GLS (con parámetros conocidos), los niveles de rechazo de  $H_0$  se mantienen cerca o ligeramente sobre del nivel declarado ( $\alpha = 5\%$ ) independientemente de las características de la variabilidad espacial). Esto es así porque el procedimiento en vez de asumir la independencia de los errores, incluye la información de la matriz de covarianza espacial en la estimación del error experimental y por lo tanto en las pruebas de hipótesis. En el caso de OLS, en el diseño con unidades experimentales grandes (franja ida y vuelta), independientemente si es aplicado con o sin bloques, el grado de autocorrelación afecta la probabilidad de rechazar  $H_0$  inflando así las tasas de error tipo I. De este modo a menor autocorrelación (más *nugget*) las tasas de error tipo I disminuyen por igual en escenarios con o sin tendencia, asumiendo que en el primer caso el diseño incorpora un factor de bloqueo. En cuanto al efecto del alcance de dependencia espacial o *range*, éste modula el efecto del *nugget* ya que a medida que aumenta el *range* el efecto del *nugget* disminuye.

Al igual que para las distribuciones de  $\hat{\beta}_0$ , la distribución del error estándar residual no se ve sustancialmente afectado por los diseños. En general la

variabilidad observada para  $\hat{\sigma}$  está relacionada con el método de estimación, siendo menor para GLS, y las características de la estructura espacial. A mayor proporción de varianza no estructurada (más *nugget*) la distribuciones se hacen más concentradas. Lo mismo ocurre con el efecto del *range*: a mayor *range*, menor dispersión de  $\hat{\sigma}$ .

La Fig. 4 muestra la relación del sesgo y varianza de  $\hat{\beta}_1$  para los distintos diseños, escenarios de variabilidad espacial y métodos de estimación. En general se observa que el D3 fue el que mostró mayores niveles de variabilidad, los cuales difirieron según el procedimiento de estimación utilizado. Dichas diferencias se redujeron conforme la proporción de varianza no estructurada aumenta. En cuanto al sesgo, este diseño resulto en sesgo ligeramente positivo mientras el resto de los diseños el sesgo fue casi nulo (D2 y D4) o negativos (D1). Asimismo, las diferencias entre ambos métodos de estimación disminuye en función del *range*.

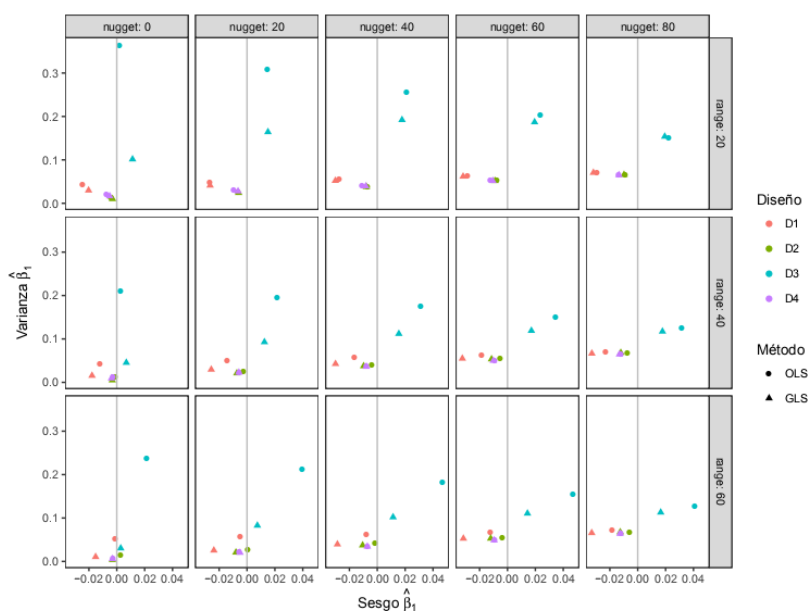


**Fig. 3.** Relación entre la tasa de rechazo de  $H_0 : \tau_i = 0$ , en función de la proporción de varianza no estructurada (*nugget*), método de estimación, rango (*range*) y diseños (D1-D4). La línea continua indica el nivel de significancia declarado  $\alpha = 5\%$ .

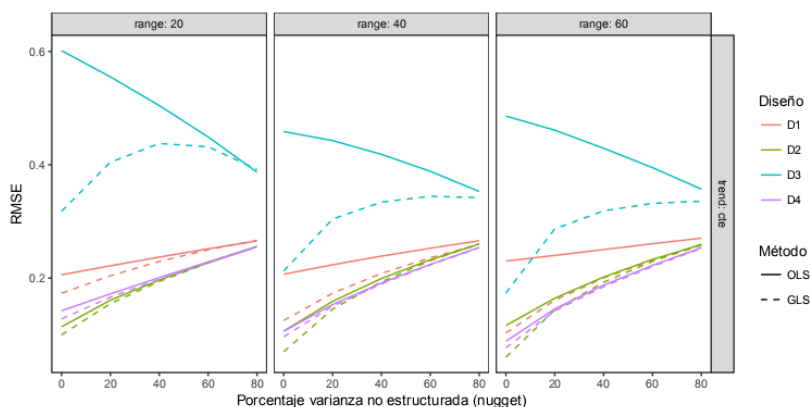
El RMSE resume la bondad del estimador combinando la información del sesgo y eficiencia obtenida a partir de la distribución simulada. En el caso de  $\hat{\beta}_1$  los valores de RMSE difirieron según el diseño aplicado, en combinación con los métodos de estimación y características de la estructura espacial (Fig. 5). En general los estimadores obtenidos por GLS tienen menor RMSE que OLS, y esta diferencia se hace menor a medida que disminuye la autocorrelación. Al igual que lo reportado por Griffin et al. [10] los valores de RMSE aumentaron conforme disminuyó el

porcentaje de varianza estructurada. En cuanto a la diferencia entre diseños, el diseño donde se verifican los mayores efectos de método y combinaciones de *nugget* y *range* es en D3, que utiliza franjas grandes (ida y vuelta). En el otro extremo, i.e. menos afectado por las variaciones de estructura espacial, se encuentra checker board (D4). esto sugiere que a medida que el número de réplicas se incrementa, el RMSE decrece. Por otro lado, esto indicaría que el D4 sería aplicable para un mayor rango de escenarios de variabilidad espacial dado que mantiene la calidad de los estimadores.

Al contrario que lo descrito para  $\hat{\beta}_1$ , todos los diseños tuvieron performance similar para  $\hat{\beta}_0$ , detectándose una relación con las combinaciones de proporción de varianza no estructurada y rango que depende del método de estimación. El patrón general indica que a medida que aumenta la proporción de varianza no estructurada el RMSE decrece, indicando una degradación de la calidad del estimador. Este decrecimiento es magnificado cuando el rango de autocorrelación aumenta. En resumen, una mayor estructura espacial (menos *nugget* y más *range*) la bondad de los estimadores es mejor. Este patrón es similar para ambos estimadores, siendo menor el RMSE para GLS.



**Fig. 4.** Relación entre la sesgo y varianza de la distribución de  $\hat{\beta}_1$  en función de la proporción de varianza no estructurada (*nugget*), método de estimación, rango (*range*) y diseños (D1-D4). La línea continua indica sesgo nulo.



**Fig. 5.** Relación RMSE y la proporción de varianza no estructurada (*nugget*), método de estimación, alcance de dependencia espacial (*range*) y diseño para  $\beta_1$ .

## 4 Conclusiones

La comparación de las distribuciones simuladas de los estimadores del efecto de tratamiento obtenidas aplicando diseños sistemáticos indica que: (i) si bien el sesgo no es afectado por el de diseño ni por las características de variabilidad espacial o método de estimación, la dispersión de las distribuciones (indicador de la eficiencia del estimador) varió en función del diseño, la estructura de autocorrelación espacial y el método de estimación. En general el RMSE se incrementó a medida que disminuyó la autocorrelación espacial, principalmente por el efecto sobre la varianza de la distribución.

Entre los diseños evaluados, D4 (tablejo ajedrez) tuvo mejor performance en la mayoría de las situaciones evaluadas, independientemente del método de estimación. Por el contrario, los peores resultados se obtuvieron con el diseño con mayor tamaño de unidad experimental (D3) y el método OLS. Si bien la inclusión de la estructura espacial mediante el procedimiento GLS mejoró la bondad del estimador (menor varianza, menor RMSE), las estimaciones obtenidas con este diseño (D3) fueron menos precisas que el resto de los diseños. Así mismo, el efecto de sobre la distribución del estimador del efecto de tratamiento resultó en niveles variables de tasas de error tipo I.

Entre los diseños considerados, a excepción del split-planter, la precisión lograda con los diseños en franja fue pobre y requirió de la utilización de métodos de estimación que incorporen la autocorrelación espacial para lograr estimaciones confiables. Por su parte, el diseño en tablero de ajedrez mostró mejores propiedades estadísticas independientemente del método de análisis. A diferencia de éste último,

que por su mayor complejidad sólo es aplicable si se dispone de tecnología de dosificación variable, los diseños en franjas muy utilizados por los productores (i.e. franjas de distinto tamaño), los cuales podrían ser aplicados si necesidad de tecnología de dosificación variable.

## 5 Referencias

1. 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).
2. 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).
3. Griffin, T.: Making the most of yield monitor data for field-scale on-farm experiments. Presented at the American Society of Agricultural and Biological Engineers Annual International Meeting 2009, ASABE 2009 (2009).
4. Griffin, T.W., Mark, T.B., Dobbins, C.L., Lowenberg-DeBoer, J.: Estimating whole farm costs of conducting on-farm research on midwestern US corn and soybean farms: A linear programming approach. *International Journal of Agricultura Management*. 4, 21–27 (2014).
5. Whelan, B.M., McBratney, A.B.: The “Null Hypothesis” of Precision Agriculture Management. *Precis. Agric*. 2, 265–279 (2000).
6. 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).
7. Whelan, B.M., Taylor, J.A., McBratney, A.B.: A ‘small strip’ approach to empirically determining management class yield response functions and calculating the potential financial ‘net wastage’ associated with whole-field uniform-rate fertiliser application. *Field Crops Research*. 139, 47–56 (2012).
8. Piepho, H.P., Möhring, J., Williams, E.R.: Why Randomize Agricultural Experiments? *J Agro Crop Sci*. 199, 374–383 (2013).
9. Willers, J.L., Milliken, G.A., Jenkins, J.N., O’Hara, C.G., Gerard, P.D., Reynolds, D.B., Boykin, D.L., Good, P.V., Hood, K.B.: Defining the experimental unit for the design and analysis of site-specific experiments in commercial cotton fields. *Agricultural Systems*. 96, 237–249 (2008).
10. Griffin, T.W., Lambert, D.M., Lowenberg-DeBoer, J.M.: Testing appropriate on-farm trial designs and statistical methods for precision farming: a simulation approach. Presented at the (2004).
11. Milliken, G., Willers, J., McCarter, K., Jenkins, J.: Designing experiments to evaluate the effectiveness of precision agricultural practices on research fields: part I concepts for their formulation. *Oper Res Int J*. 10, 329–348 (2010).
12. Lawes, R.A., Bramley, R.G.V.: A Simple Method for the Analysis of On-Farm Strip Trials. *Agronomy Journal*. 104, 371 (2012).

13. 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).
14. Webster, R., Oliver, M.A.: *Geostatistics for enviromental sicientists*. Wiley (2007).
15. Pinheiro, J.C., Bates, D.M.: *Mixed Effects Models in S and S-Plus*. Springer, New York (2000).
16. Cressie, N.: *Statistics for Spatial Data*. Wiley-Interscience, New York, USA (1993).
17. R Core Team: *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria (2018).
18. Pebesma, E.J.: Multivariable geostatistics in S: the gstat package. *Comput. Geosci.* 30, 683–691 (2004).
19. Wickham, H., Francois, R., Henry, L., Müller, K.: *dplyr: A Grammar of Data Manipulation*. (2017).
20. Wickham, H.: *ggplot2 elegant graphics for data analysis*. Springer, Dordrecht; New York (2009).
21. Pinheiro, J.C., Bates, D.M., DebRoy, S., Sarkar, D., R Core Team: *nlme: Linear and Nonlinear Mixed Effects Models*. R Foundation for Statistical Computing (2017).