

# **Algunos aspectos estadísticos de interés en los ensayos de fertilización<sup>1</sup>**

C. Ramos  
Instituto Valenciano de Investigaciones Agrarias  
Apartado Oficial  
46113 Moncada (Valencia)  
(jubilado) cramos.ramos@gmail.com

**Palabras clave:** abonado, estadística, diseño experimental, variabilidad espacial

## **Resumen**

**En este trabajo se describen algunos aspectos estadísticos del diseño y análisis de los experimentos que en ocasiones no son considerados por los investigadores o reciben escasa atención. Se presentan las ideas básicas del diseño de experimentos, con especial énfasis en el caso de bloques completos al azar con una y varias observaciones por unidad experimental. Se remarca la importancia de los objetivos del estudio en el diseño experimental y de la relación que existe entre el diseño y el análisis de los resultados. Se exponen algunos de los errores más frecuentes en el diseño y análisis de experimentos y se introduce la geoestadística como una herramienta útil para la caracterización espacial de algunas variables del suelo y de las plantas, así como para la estimación de estas variables en zonas no medidas y se discute el papel de la variabilidad espacial del suelo en el diseño y análisis de los ensayos de campo. Por último se presenta el sistema de dominio público R para el análisis estadístico de los datos experimentales y también para la realización de gráficos de alta calidad.**

## **INTRODUCCIÓN**

La estadística es básica en cualquier investigación, pero la formación de los investigadores en esta materia es, en general, deficiente. Las publicaciones científicas son cada vez más exigentes en los aspectos estadísticos de los ensayos y en algunas existen guías para los autores, como en el European Journal of Soil Science o el Agronomy Journal (ver Referencias). Por otra parte, de vez en cuando se publican artículos en los que se señalan los errores estadísticos más frecuentes en los estudios experimentales (Ioannidis, 2005) y en los de temática agrícola (Bryan-Jones y Finney, 1983; Nelson, 1989; Maindonald, 1992). En este trabajo se hace una revisión de aquellos aspectos estadísticos que se plantean con frecuencia en los ensayos de fertilización en el campo y se intenta remarcar la importancia que tiene para un investigador el dominio de los fundamentos básicos de la estadística y del diseño y análisis de experimentos.

## **Aspectos del diseño experimental**

Una de las primeras cuestiones que se presentan al preparar un experimento en el campo es el diseño experimental, incluyendo el número de tratamientos, repeticiones y tamaño de las unidades experimentales (la unidad experimental es la unidad de material a

---

<sup>1</sup> Este trabajo está dedicado a la memoria de Juan Ramón Castel Sánchez, amigo y compañero de muchos años, y aficionado también a la estadística.

la que se aplica un tratamiento). Sin embargo, un aspecto primordial que a veces se descuida, y que es previo al diseño experimental, es la especificación clara de los objetivos del ensayo, ya que son éstos los que tienen que guiar todo el planteamiento experimental, como señala Cousens (1985) en un artículo muy interesante. El establecimiento claro de los objetivos presupone que previamente se ha hecho una revisión bibliográfica exhaustiva para delimitar mejor los objetivos.

Uno de los primeros aspectos a considerar es el tipo de investigación que queremos hacer. Como señala Pearce (1988) hay dos tipos generales de experimentos: 1) los de testado de hipótesis (en ellos se quiere contestar una pregunta con un sí o un no, por ejemplo: ¿es el abono A mejor que el B?) y, 2) los de estimación (en este caso la pregunta requiere una contestación cuantitativa, por ejemplo: ¿En cuanto aumenta la producción de un cultivo si se aplican al suelo 100 kg/ha de N?).

Cuando el objetivo del experimento es testar una hipótesis, entonces se deben establecer niveles de error del tipo I y la potencia del test. El error tipo I (que viene determinado por lo que tradicionalmente se denomina  $\alpha$ ) es el error que se comete cuando se rechaza la hipótesis nula siendo ésta verdadera (la hipótesis nula establece que no hay diferencia entre las medias). El error tipo II, es el que se comete cuando no se rechaza la hipótesis nula siendo ésta falsa; a este error se le suele denominar  $\beta$ . La potencia de un test es la probabilidad de rechazar la hipótesis nula cuando es falsa; la potencia es igual a  $1 - \beta$  (Steel et al., 1997; Quinn y Keough, 2002). Existe una relación entre los dos tipos de errores, ya que cuando se quiere disminuir el error tipo I, fijando un valor de  $\alpha$  más pequeño, entonces aumenta el error tipo II ( $\beta$ ) y por tanto disminuye la potencia del test ( $1 - \beta$ ) (pp. 118-124 de Steel et al., 1997). La potencia de un test ( $1 - \beta$ ) depende de la diferencia entre la hipótesis nula y la hipótesis alternativa (mayor potencia con mayor diferencia) y también aumenta con el tamaño de las muestras.

Un concepto muy relacionado con el nivel  $\alpha$  es el valor  $p$ . El valor de  $\alpha$  lo establece el investigador y representa el error tipo I que está dispuesto a aceptar, mientras que  $p$  es la probabilidad de obtenerla diferencia observada (o una más grande) por azar, es decir, cuando no hay diferencias entre las medias porque las dos provienen de una misma población. Tradicionalmente se considera que un valor de  $\alpha = 0.05$  (es decir,  $p \leq 0.05$ ) es un nivel de evidencia suficiente para rechazar la hipótesis nula. Sin embargo, como señalan Maindonald y Braun (2010, pag. 108) este valor sería demasiado grande si, por ejemplo, el rechazo de la hipótesis nula fuera a determinar que se cambiara una determinada práctica agrícola en una región; en este caso, sería más razonable emplear un valor  $\alpha = 0.01$  o incluso inferior. En cambio, este valor puede ser demasiado pequeño si el estudio es de tipo exploratorio, por ejemplo, para decidir si un determinado tratamiento puede merecer un estudio específico adicional, en cuyo caso se podría emplear, por ejemplo,  $\alpha = 0.10$ . Actualmente, son muchos los autores que aconsejan expresar los resultados como los valores medios de cada tratamiento y sus intervalos de confianza, ya que estos intervalos dan una idea de la incertidumbre asociada a cada valor (Gardner y Altman, 1986).

Una explicación clara de los tests de hipótesis y de la estimación de intervalos de confianza para las medias se puede ver en una publicación del Statistical Services Centre (2006) y una exposición clara del diseño experimental es la de Maindonald (2000).

En la fig. 1 se resumen las cuatro posibilidades o combinaciones de la realidad de la población sobre la que queremos hacer una inferencia con las decisiones de rechazar o no la hipótesis nula en los test de hipótesis (Quinn y Keough, 2002).

		Conclusión estadística	
		Rechazar $H_0$	No rechazar $H_0$
Realidad	Hay efecto	Decisión correcta	Error tipo II Efecto no detectado
	No hay efecto	Error tipo I Se detecta un efecto que no existe	Decisión correcta

Fig. 1 Decisiones estadísticas y errores al testar una hipótesis nula  $H_0$

Ramsey y Schafer (2002)(pag. 681) dan una lista de ocho tareas necesarias en el diseño de una investigación:

1. Definir el objetivo (¿qué pregunta queremos contestar?)
2. Determinar el ámbito de inferencia (¿será un experimento aleatorizado o un estudio observacional?, ¿qué unidades experimentales o de muestreo se emplearán?, ¿cuál es la población de interés?)
3. Comprender el sistema a estudiar (requiere conocimiento experto sobre el tema de estudio)
4. Decidir cómo se va a medir la respuesta o variable a estudiar
5. Hacer una lista de los factores que pueden afectar la respuesta:
  - a. Factores de diseño:
    - i. Factores a variar (tratamientos y control)
    - ii. Factores a controlar o fijar
  - b. Factores de confusión (“confounding”):
    - i. Factores a controlar por diseño
    - ii. Factores a controlar por análisis (covariables)
    - iii. Factores a controlar por aleatorización.
6. Planificar el cronograma del experimento
7. Bosquejar el análisis estadístico
8. Determinar el tamaño de muestra

Para ayudar a entender la tarea 2, en la fig. 2 se puede ver cómo: 1) la selección aleatoria de la muestra permite inferir los resultados a la población y, 2) la asignación aleatoria de los tratamientos permite concluir que las diferencias observadas en los tratamientos son “causadas” por los mismos (Ramsey y Schafer, 2002).

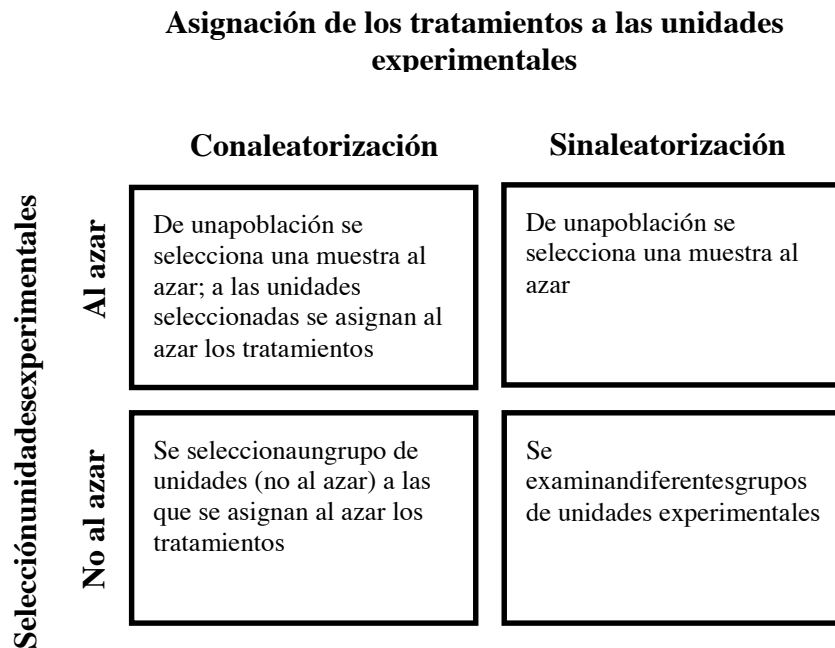


Fig. 2 Inferencias estadísticas permitidas según el diseño del estudio. En la primera columna se puede inferir causalidad y en la primera fila se puede inferir a la población (adaptado de Ramsey y Schafer, 2002).

### **El diseño de bloques completos al azar**

Este diseño es uno de los más empleados en los experimentos agrícolas. En ocasiones se aplica sin tener en cuenta cual es el fundamento del mismo. Los tres pilares en que se apoya el diseño experimental son: 1) la asignación de los tratamientos a las unidades experimentales al azar, 2) la replicación de los tratamientos y, 3) la reducción del error experimental. Esta última idea es la que subyace en la creación de bloques cuando en la zona experimental hay una variación del suelo, u otras condiciones que afecten a lo que queremos medir, en una determinada dirección y los bloques se establecen perpendicularmente a esta dirección, lo cual permite tener bloques más homogéneos en relación a la variable cuya variación queremos controlar. En ocasiones no se tiene información sobre la homogeneidad del suelo y, por tanto, la creación de bloques no mejora la homogeneidad de las unidades experimentales en cada bloque. En este caso, el análisis de la varianza (ANOVA) indicará que el efecto de los bloques (un factor aleatorio) no es significativo y por la asignación de grados de libertad a los bloques, los

grados de libertad del error experimental disminuirán en  $b-1$  lo cual aumentará el cuadrado medio residual que es una estimación de la varianza experimental.

A continuación se muestra un ejemplo que ilustra este aspecto (los datos son ficticios). En un experimento se han aplicado tres tratamientos de abonado nitrogenado diferentes a un cultivo: una misma dosis de N pero con tres formas químicas, amoniacal, nítrica y amídica (tratamientos A, B, C, respectivamente). El ensayo se hace en un diseño de bloques completos al azar con 5 repeticiones. La tabla 1 presenta los datos de producción para cada unidad experimental.

Tabla 1. Datos de producción de un ensayo en bloques completos al azar con tres tratamientos y cinco repeticiones

Tratamiento	Bloque	Producción (t/ha)
A	1	29.8
A	2	29.9
A	3	30.6
A	4	22.1
A	5	32.0
B	1	25.8
B	2	20.6
B	3	25.2
B	4	25.9
B	5	19.9
C	1	24.8
C	2	21.5
C	3	26.5
C	4	20.4
C	5	20.1

En la tabla 2 se presenta el ANOVA de estos datos. Se puede ver que hay un efecto estadísticamente significativo de los tratamientos y que el efecto bloques no es significativo ( $p \geq 0.05$ ), ya que  $p = 0.382$ . Ahora, se pueden comparar estos resultados con los que se hubieran obtenido si el experimento hubiera sido sin bloques, es decir, completamente al azar. El ANOVA en este segundo caso se presenta en la tabla 3. Se puede ver que la significación estadística de los tratamientos es algo mayor (el valor de  $p$  pasa de 0.029 a 0.022), debido a un mayor número de grados de libertad para la estimación del error experimental (12 frente a 8 en el primer caso; este cambio afecta a la  $F$  que se emplea para determinar la probabilidad  $p$ ).

Conviene recordar que los bloques no tienen que ser necesariamente unidades de suelo próximas (aunque es lo más frecuente, porque el suelo suele ser más parecido en puntos cercanos que lejanos), también pueden ser bloques las unidades recolectadas en un día (si los días pueden afectar a los resultados), o bien estar asociados a operarios, si hay evidencia o se presume que éstos pueden tener una influencia sobre los resultados.

Tabla 2. ANOVA de los datos de la tabla 1

Fuente de variación	g.l.	Suma cuadrados	Cuadrado medio	F	Pr(>F)
Bloque	4	48.003	12.001	1.1973	0.382
Tratamientos	2	114.201	57.101	5.6969	0.029
Residual	8	80.185	10.023		

Tabla 3. ANOVA de los datos de la tabla 1 ignorando los bloques

Fuente de variación	g.l.	Suma cuadrados	Cuadrado medio	F	Pr(>F)
Tratamientos	2	114.20	57.101	5.34	0.022
Residual	12	128.19	10.682		

Una manera de ver la función del diseño experimental es considerar que para determinar si los tratamientos tienen efectos significativos lo que se hace es comparar la varianza de los datos a partir de la varianza de las medias con la variabilidad experimental (dentro de cada tratamiento). Es decir, se compara la varianza de los datos atribuible a los tratamientos con la varianza entre unidades tratadas igual, que se estima con la varianza del error residual. Así pues, el objetivo del diseño experimental es poder reducir esta variabilidad residual: cuando se establecen bloques, y estos introducen variabilidad porque son diferentes entre sí, entonces el ANOVA determina que parte de la variación es atribuible a los bloques y la descuenta de la variabilidad total, con lo que de esta manera la variabilidad no explicada ni por los tratamientos ni los bloques, es decir, la variabilidad residual, disminuye.

### **ANOVA para bloques completos al azar con varias observaciones por unidad experimental**

En ocasiones, debido a la variabilidad del material o de la técnica de medida es aconsejable tomar varias medidas para tener un valor medio de la variable en cuestión más representativo en cada unidad experimental. Por ejemplo, si se quiere medir el contenido de N mineral en el suelo o la concentración de nitrato en la savia de las plantas normalmente se toman varias muestras de suelo o varias medidas en las plantas en la unidad experimental. El problema que se puede tener en el análisis estadístico en este caso es no calcular correctamente en el ANOVA el valor de F para los tratamientos, ya que el error experimental no es el que tiene  $t \cdot b \cdot (s-1)$  grados de libertad sino  $(t-1)(b-1)$  (tabla 5). En esta situación se pueden emplear dos métodos: 1) el más sencillo es sustituir el conjunto de medidas en cada unidad experimental por su valor medio y operar después como si fuera un diseño sin submuestras (es decir, sin repeticiones dentro de la unidad experimental), y 2) incluir en el ANOVA las submuestras como una fuente adicional de variación, lo cual permite obtener una estimación de la variabilidad debida al submuestreo (tabla 5), que puede ser de interés. Para ilustrar esto se muestra el siguiente

ejemplotomado de: [https://www.ndsu.edu/faculty/horsley/RCBD\\_\(revised\).pdf](https://www.ndsu.edu/faculty/horsley/RCBD_(revised).pdf). En este ejemplo no se especifica ni el cultivo ni los tratamientos y las unidades de producción, pero esto no influye en lo que se quiere mostrar. Se trata de un ensayo con un diseño de bloques completos al azar y tres tratamientos, tres repeticiones y dos determinaciones de producción por unidad experimental (submuestreo). Los datos experimentales se presentan en la tabla 4.

El ANOVA para un experimento de este tipo, empleando el método en el que se tiene en cuenta el submuestreo (y suponiendo que no hay interacción entre los bloques y los tratamientos), es de la forma presentada en la tabla 5. Nótese que el valor de F para los tratamientos se calcula empleando  $CM_T/CM_E$  y no  $CM_T/CM_S$ . Los resultados de este tipo de ANOVA aplicado a los datos de la tabla 4 se presentan en la tabla 6. Se observa que el efecto de los tratamientos es significativo al nivel de significación del 1% (el valor de  $p$  es 0.0003) y que los bloques no son significativamente diferentes al nivel del 5% ( $p = 0.243$ ). Los valores de  $p$  a partir de los valores de F se pueden obtener con cualquier software estadístico (aquí se han calculado con R).

En cambio, si se hace el ANOVA empleando los valores medios de las dos medidas realizadas en cada unidad experimental, es decir, ignorando el submuestreo se obtienen los resultados mostrados en la tabla 7.

En este caso los resultados son similares a los del análisis anterior, aunque hay un pequeño cambio en los valores de significación estadística del efecto de los tratamientos ( $p$  aumenta de 0.0003 a 0.0011). En los dos casos la mínima diferencia significativa (LSD) entre tratamientos es la misma:  $LSD = t_{(t(b-1), \alpha/2)} \cdot \sqrt{2CM_E/b}$  (cuando no se considera submuestreo  $s=1$ ).

La ventaja del ANOVA con submuestreo es que permite obtener información sobre la varianza de las muestras dentro de cada unidad experimental y su participación en la varianza total. Para ver esto más claramente conviene saber que estima cada cuadrado medio (tabla 8).

A partir de los datos de las tablas 6 y 8 tenemos:

- $\sigma_s^2 = 13.11$
- $\sigma_s^2 + s \sigma_e^2 = 8.222$
- $\sigma_s^2 + s \sigma_e^2 + b s \Sigma \tau_i^2 / (t - 1) = 968.222$
- $\sigma_s^2 + s \sigma_e^2 + s t \sigma_b^2 = 16.889$

y resolviendo las ecuaciones anteriores ( $t=3$ ,  $b=3$  y  $s=2$ ) se obtiene:

- $\sigma_s^2 = 13.11$
- $\sigma_e^2 = -2.444$  (como es negativo se toma = 0)
- $\Sigma \tau_i^2 / 2 = 159.19$
- $\sigma_b^2 = 0.63$

Con estas varianzas estimadas de los diferentes componentes se puede estimar la varianza esperable al cambiar el número de bloques y de muestras por unidad experimental.

Cuando se emplea submuestreo, si el diseño es completamente al azar, para decidir si es mejor aumentar el número de repeticiones o el número de muestras por unidad experimental, se puede emplear la relación (McPherson, 2001, pag. 576):

$$\text{var}(\bar{y}) = \frac{\sigma_e^2}{b} + \frac{\sigma_s^2}{bs}$$

donde la varianza de la media de cualquier tratamiento ( $\bar{y}$ ) se expresa en función de la varianza de las unidades experimentales ( $\sigma_e^2$ ) y de la varianza de las muestras dentro de cada unidad experimental ( $\sigma_s^2$ ), así como del número de repeticiones ( $b$ ) y de muestras por cada unidad experimental ( $s$ ). Así pues, para un número total de muestras ( $bs$ ) interesará aumentar  $b$  ó  $s$  dependiendo de los valores de  $\sigma_e^2$  y  $\sigma_s^2$  para minimizar la varianza de la media ( $\bar{y}$ ). En esta selección también pueden intervenir criterios como el coste de cada opción.

Tabla4. Datos de producción de un cultivo en un ensayo de bloques completos al azar con tres tratamientos, tres bloques y dos determinaciones (muestras) por unidad experimental

Bloque	Muestra	Tratamiento		
		A	B	C
1	1	78	68	89
	2	82	64	87
2	1	74	62	88
	2	78	66	92
3	1	80	70	90
	2	84	60	96

Tabla 5. ANOVA para un experimento de bloques completos al azar con  $b$  bloques,  $t$  tratamientos y  $s$  muestras por unidad experimental (con efecto bloques aleatorio y tratamientos fijo) (Gates, 1991)

O también: (<http://www.tfrec.wsu.edu/ANOVA/RCBsub.html>)

Fuente de variación	g.l.	Suma cuadrad.	Cuadrado medio	F
Bloques (B)	$b - 1$	$SS_B$	$CM_B = SS_B / (b - 1)$	$CM_B / CM_E$
Tratamientos (T)	$t - 1$	$SS_T$	$CM_T = SS_T / (t - 1)$	$CM_T / CM_E$
Error experimental (E)	$(b-1)(t-1)$	$SS_E$	$CM_E = SS_E / ((b - 1)(t - 1))$	$CM_E / CM_S$
Error de muestreo (S)	$t \cdot b \cdot (s-1)$	$SS_S$	$CM_S = SS_S / (t \cdot b \cdot (s-1))$	
Total	$t \cdot b \cdot s - 1$	$SS_{Total}$		

Tabla 6. ANOVA para los datos de la tabla4 utilizando el modelo de la tabla 5



Fuente de variación	g.l.	Suma cuadrad.	Cuadrado medio	F
Bloques (B)	2	33.78	16.889	2.054 (p=0.243)
Tratamientos (T)	2	1936.44	968.222	117.76 (p=0.0003)
Error experimental (E)	4	32.89	8.222	
Error de muestreo (S)	9	118	13.111	
Total	17	2121.111		

Tabla 7. ANOVA para los datos de la tabla 4 tomando los valores medios de las dos medidas de cada unidad experimental (i.e., sin considerar el submuestreo).

Fuente de variación	g.l.	Suma cuadrad.	Cuadrado medio	F	p
Bloques	2	16.89	4.22	1.027	0.44
Tratamientos	2	968.22	242.06	58.88	0.0011
Error residual	4	16.44	4.111		
Total	8				

Tabla 8. Valores esperados de los cuadrados medios en el ANOVA del diseño de bloques completos al azar con submuestreo (tomando los bloques como efecto aleatorio y los tratamientos como efectos fijos y considerando que no hay interacción bloque x tratamiento) (pag. 226 de Steel et al., 1997)<sup>(1)</sup>

Fuente de variación	Grados de libertad	Cuadrado medio esperado
Bloques (b)	b- 1	$\sigma_s^2 + s \sigma_e^2 + s t \sigma_b^2$
Tratamientos (t)	t- 1	$\sigma_s^2 + s \sigma_e^2 + b s \sum \tau_i^2 / (t - 1)$
Error experimental (e)	(b - 1)(t - 1)	$\sigma_s^2 + s \sigma_e^2$
Error de muestreo (s)	bt(s - 1)	$\sigma_s^2$

<sup>(1)</sup> Si hay interacción entre bloques y tratamientos entonces hay una fuente adicional de variación y esta tabla cambia (Gates, 1995)

### Errores frecuentes en el diseño y análisis de resultados

En todas las áreas de investigación, incluida la agronómica, ocasionalmente se publican artículos en los que se señalan deficiencias y errores estadísticos que invalidan o disminuyen el valor de las conclusiones de los mismos, o bien se dan recomendaciones para ayudar a eliminar estos problemas. Algunas de estas publicaciones en el área agronómica son las de Preece (1982), Bryan-Jones and Finney (1983), Maindonald y Cox (1984), Gates, (1991) y Onofri et al. (2010), y en ecología la de Zuur et al. (2009). Aunque estas publicaciones son de ayuda, para entender algunos de los puntos que los autores señalan hace falta tener una cierta base estadística. Entre los errores más frecuentes citados en estas publicaciones y otras similares están:

- La confusión entre la verdadera replicación (medidas en diferentes unidades experimentales) y la pseudoreplicación (en donde se consideran como repeticiones aquellas medidas repetidas que se toman en una misma unidad experimental)
- La utilización de las comparaciones múltiples cuando la variable es cuantitativa (en este caso la regresión es la técnica apropiada) (Perry (1988) y Tukey (1991))
- La comparación de medias sugerida por los datos utilizando tests que están pensados para comparaciones planeadas de antemano.
- Utilizar de manera incorrecta los términos de error en los ANOVA para determinar el valor F correspondiente a los tratamientos.
- No contrastar si los requerimientos de los tests y otros procedimientos estadísticos se cumplen (por ejemplo, normalidad de los datos, independencia de los errores, homocedasticidad (igualdad de varianzas), etc.)

Parte de estos errores se deben al empleo de la estadística como un recetario de procedimientos para el análisis de los resultados de los experimentos siguiendo el patrón de estudios similares publicados. Esta situación es bien descrita por Clewer y Scarisbrick (2001) en el prefacio de su excelente libro, en el que señalan que la existencia del software estadístico facilita la aplicación inapropiada de algunos programas y enfatizan la necesidad de considerar de manera conjunta el diseño y el análisis dentro de cada experimento, ya que el diseño determina en gran parte el análisis adecuado. Una buena referencia sobre la aplicación de la estadística en la investigación agrícola es Mead et al. (2003).

### **La geoestadística y su aplicación a los ensayos de fertilización**

En los apartados anteriores se ha visto como la variabilidad del suelo y de las plantas influye en la variabilidad de los datos y ello determina el “error experimental” que actúa como una referencia sobre la que se comparan los efectos observados de los tratamientos, y cómo algunos diseños experimentales pueden disminuir estos efectos y con ello aumentar la capacidad de detectar diferencias entre tratamientos (es decir, aumentar la potencia de los tests estadísticos). Desde hace ya bastantes años se ha investigado el papel de la variabilidad espacial en los experimentos de campo y en su aplicación a la agricultura de precisión. Los temas estadísticos en esta área están dominados por la dependencia espacial de los datos, mientras que en la teoría en que se basa el ANOVA de los diseños experimentales presentados en este trabajo la independencia de los errores es una hipótesis básica. La geoestadística permite determinar la estructura de esta variabilidad espacial y el rango de la dependencia espacial mediante el cálculo y representación de los variogramas (McBratney y Pringle, 1999; Oliver y Webster, 2015). Así mismo, la técnica de kriging permite interpolar o estimar valores de

las variables en puntos no medidos a partir de los valores medidos en puntos más o menos cercanos, mediante medias ponderadas según diferentes modelos.

Desde hace bastantes años, la existencia de estos patrones de variabilidad espacial habían sugerido la posibilidad de ajustar los valores experimentales de cada unidad experimental en estudios de campo, en función de los valores obtenidos en unidades experimentales próximas, mediante la técnica denominada del “nearest neighbour”(NN) (Besag y Kempton, 1986). Yang y Juskiw (2011) muestran una comparación de un análisis de resultados en un ensayo experimental con diseño de bloques completos al azar con y sin ajuste de los datos con la técnica del NN.

Algunas referencias básicas en este campo son: De Gruijter et al. (2006), Plant (2012) y Webster y Lark (2013). Un ejemplo de aplicación de estos métodos a la determinación espacial del N mineral en el suelo es Baxter et al. (2003).

Los mapas de variabilidad espacial permiten identificar zonas diferenciales del suelo y contribuyen a optimizar el muestreo del mismo (de Gruijter et al., 2006). Esto puede ser muy útil en trabajos de campo con parcelas de grandes dimensiones o heterogéneas respecto a una determinada variable, para identificar zonas representativas de la parcela. Un ejemplo de aplicación en ensayos de riego para localizar zonas de diferente textura y capacidad de retención de agua del suelo se puede ver en Fortes et al. (2015).

Una dificultad para la aplicación de la geoestadística a los experimentos de campo en agronomía es, como apuntó Hengl (2009), la variabilidad temporal de algunas variables (por ejemplo, el contenido de N mineral del suelo) que hace que el período de validez de los mapas obtenidos en un momento determinado pueda ser corto.

### **El sistema R para el análisis y representación gráfica de datos**

En los primeros años 90 apareció el software libre R para el análisis estadístico de datos y la creación de gráficos (Ihaka y Gentleman, 1996). Uno de los principales beneficios de R es que funciona en cualquier plataforma o sistema operativo estándar: Mac, Windows, Linux, etc. Este sistema ha revolucionado el uso de la estadística, no sólo por ser libre y gratuito sino porque permite hacer gráficos de la calidad exigida en las publicaciones científicas y, además, porque gracias a internet ha permitido que expertos de todo el mundo hayan contribuido a crear paquetes (“packages”) para multitud de tareas específicas (en junio de 2015 había cerca de 7000). Esta difusión de R se ha manifestado en que hay muchos foros en internet sobre R y, debido a que R no es fácil al principio y tiene muchas posibilidades, esta información accesible en la red es de gran ayuda. Para ejecutar los análisis de datos, R permite la escritura de los comandos directamente en la consola o bien se pueden emplear “scripts” o programas preparados previamente por el usuario que contienen las diferentes líneas de comandos. Estos “scripts” pueden ser empleados, con pocos cambios, para otros datos que requieran el mismo tipo de análisis o bien por otros usuarios.

R se puede descargar de la red: <https://cran.r-project.org/>. En internet hay mucha información disponible para aprender R, desde tutoriales escritos hasta videos (en varios idiomas). En la propia página principal de R mencionada se puede ver (y descargar) mucha información (ver bajo Documentation (Contributed)).

Algunos de los paquetes desarrollados para uso en la investigación agrícola, son: “agricolae” para el diseño de experimentos, y “soiltexture” que permite representar triángulos de textura del suelo de la mayoría de los sistemas existentes y determinar la

denominación textural de los suelos en los diferentes sistemas en función de su granulometría.

En el IVIA en 2015 y 2016 se dio un pequeño curso sobre “Métodos estadísticos para la investigación agronómica” cuya grabación está disponible en la red en: <http://www.ivia.gva.es/curso-de-metodos-estadisticos-para-la-investigacion-agronomica> y, en esta misma página, al final hay unos videos sobre un curso de introducción a R. La bibliografía recomendada sobre R y que está accesible en la red aparece en la misma página, al final, en: Bibliografía → Otro Material (Referencias adicionales).

### **Reflexiones finales**

La estadística es una disciplina que requiere interés y dedicación pero que cada investigador debería conocer en sus principios más básicos para poder hacer mejor su trabajo. Como han señalado diferentes autores, el conocimiento del tema de trabajo debe dirigir la investigación y la estadística debe emplearse para ayudar a conseguir de manera más eficiente los objetivos científicos planteados. Es función del investigador obtener datos de calidad (no sesgados y de la mayor precisión posible) sobre los que aplicar los métodos estadísticos, ya que no hay ningún procedimiento estadístico que pueda arreglar datos obtenidos de manera errónea o sesgada. Un ejemplo muy ilustrativo de la interacción entre el conocimiento sobre la materia y las cuestiones estadísticas en un experimento sencillo lo presenta Easterling (2004).

Debido a la dificultad de algunos temas de la estadística es necesaria en muchos casos la ayuda de los expertos en esta disciplina. Sin embargo, para que estos expertos puedan contribuir mejor al diseño y análisis de experimentos necesitan conocer también cuanto más mejor la materia de investigación. Así pues, un mayor conocimiento de estadística por parte del investigador y de la materia objeto de la investigación por parte del estadístico facilita el entendimiento mutuo y con ello la mejor colaboración para definir y conseguir los objetivos.

Para concluir, se reproduce una cita de Bryan-Jones y Finney (1983) con la que es fácil identificarse:

*“In interpreting and in presenting experimental results there is no adequate substitute for thought – thought about the questions to be asked, thought about the nature and weight of evidence the data provide on these questions, and thought about how the story can be told with clarity and full honesty to a reader. Statistical techniques must be chosen and used to aid but not to replace, relevant thought.”*

### **Agradecimientos**

Quiero agradecer la cuidadosa revisión de una primera versión de este trabajo realizada por: Carlos Campillo, Claudia Ximena Jaramillo, Antonio Lidón y Henar Prieto.

### **Referencias**

Agronomy Journal. Instructions to authors:

<https://dl.sciencesocieties.org/publications/aj/instructions-to-authors>

Baxter, S.J., Oliver, M.A., Gaunt, J. 2003. A geostatistical analysis of the spatial variation of soil mineral nitrogen and potentially available nitrogen within an arable field. *Prec. Agr.* 4: 213–226

Besag, J. and Kempton R. 1986. Statistical analysis of field experiments using neighbouring plots. *Biometrics* 42: 231-251

- Bryan-Jones, J. and D.J. Finney. 1983. On an error in "Instructions to Authors". *HortScience* 18:279-282
- Clewer, A.G and Scarisbrick, D. H. 2001. *Practical statistics and experimental design for plant and crop science*. John Wiley & Sons Inc
- Cousens, R. 1985. Underlying principles in the design and interpretation of experiments. *Aspects of Applied Biology* 10. Field trials methods and data handling, pp. 1-12.
- De Grijter, J. J., Brus, D. J., Bierkens, M. F. P., & Knotters, M. 2006. *Sampling for natural resources*. Springer-Verlag.
- Easterling R.G. 2004. Teaching experimental design. *The American Statistician* 58:244-252.
- European Journal of Soil Science (guide to authors):  
[http://onlinelibrary.wiley.com/journal/10.1111/\(ISSN\)1365-2389/homepage/ForAuthors.html](http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1365-2389/homepage/ForAuthors.html) (consultado el 5-10-2016).
- Fortes, R., Millán, S., Prieto, M. H., & Campillo, C. 2015. A methodology based on apparent electrical conductivity and guided soil samples to improve irrigation zoning. *Precision Agriculture*, 16:441–454
- Gardner, M.J. & Altman, D.G. 1986. Confidence intervals rather than P values: estimation rather than hypothesis testing. *British Medical J.* 292(6522): 746-750.
- Gates, C. E. 1991. A user's guide to misanalyzing planned experiments. *HortScience* 26: 1261-1265.
- Gates, C. E. 1995. What really is experimental error in block designs? *American Statistician*, 49: 362-363.
- Hengl, T. 2009. *A Practical Guide to Geostatistical Mapping*. University of Amsterdam  
<http://spatial-analyst.net/book/> (consultado el 5-10-2016)
- Ihaka, R. and Gentleman, R. 1996. R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics* 5:299--314.
- Ioannidis, J.P.A. 2005. Why most published research findings are false. *PLoS Med*,  
<http://medicine.plosjournals.org/> 2(8): e124, 696–701, consultado el 6-9-2016.
- Maindonald, J. 1992. Statistical design, analysis, and presentation issues. *New Zealand J. Agricultural Res.* 35:121-141.
- Maindonald J. 2000. *The Design of Research Studies – A Statistical Perspective*. Centre for Bioinformation Science, John Curtin School of Medical Research and School of Mathematical Sciences Australian National University. <http://maths-people.anu.edu.au/~johnm/Docs/planning/protgs2002.pdf> (consultado el 5-10-2016)
- Maindonald, J., Braun, J. 2010. *Data Analysis and Graphics Using R. An Example-based Approach*. 3<sup>rd</sup> ed., Cambridge University Press, Cambridge
- Maindonald, J.H. & Cox, N.R. 1984. Use of statistical evidence on some recent issues of DSIR agricultural journals. *New Zealand J. Agricultural Res.* 27: 597-610.
- McBratney A.B. and M. J. Pringle. 1999. Estimating Average and Proportional Variograms of Soil Properties and Their Potential Use in Precision Agriculture. *Precision Agriculture* 1:219-236
- McPherson, G. 2001. *Applying and Interpreting Statistics. A comprehensive Guide*. 2<sup>nd</sup> ed, Springer
- Mead, R., Curnow, R.N. and Hasted, A.M. 2003. *Statistical Methods in Agriculture and Experimental Biology* (3rd edition). Chapman & Hall, London. 472 pp.
- Nelson, L A 1989 A statistical editor's viewpoint of statistical usage in horticultural science publications. *HortScience* 24:53–57
- Oliver M.A. and Webster R. 2015. *Basic Steps in Geostatistics: The Variogram and Kriging*. Springer

- Onofri, A., Carbonell, E.A., Piepho, H.P., Mortimer, A.M. & Cousens, R.D. 2010. Current statistical issues in weed research. *Weed Res* 50:5–24.
- Pearce, S.C. 1988. Analysis of data from agricultural experiments. *Trop. Agric. (Trinidad)* 65:2-5.
- Perry, J. 1986. Multiple-comparison procedures: a dissenting view. *Journal of Economic Entomology* 79:1149-1155.
- Plant, R. E. 2012. *Spatial Data Analysis in Ecology and Agriculture Using R*. CRC Press/Taylor & Francis Group
- Preece, D.A. 1982. The design and analysis of experiments: what has gone wrong?. *Util. Math. A*, 21, 201-244.
- Quinn, G.P. & Keough, M.J. 2002. *Experimental Design and Data Analysis for Biologists*. Cambridge University Press, Cambridge, UK.
- Ramsey, F.L. and Schafer, D.W. 2002. *The Statistical Sleuth: A Course in Methods of Data Analysis*, 2nd Edition, Duxbury Press.
- Statistical Services Centre. 2006. *Confidence & Significance: Key Concepts of Inferential Statistics*. University of Reading.  
[www.reading.ac.uk/ssc/resource-packs/RUFORUM\\_DVD\\_2009-Aug/documents/Stats\\_Guides/inf.pdf](http://www.reading.ac.uk/ssc/resource-packs/RUFORUM_DVD_2009-Aug/documents/Stats_Guides/inf.pdf) (consultado el 5-10-2016)
- Steel, R. G. D., Torrie, J. H., and Dickey, D. 1997. *Principles and Procedures of Statistics: A Biometrical Approach*, Third Edition, New York: McGraw-Hill, Inc.
- Tukey, J. W. 1991. The philosophy of multiple comparisons. *Statistical Science* 6: 100–16.
- Webster, R., & Lark, R. M. 2013. *Field sampling for environmental science and management*. London: Routledge.
- Yang, R.-C. and Juskiw, P. 2011. Analysis of covariance in agronomy and crop research. *Can. J. Plant Sci.* 91: 621-641
- Zuur, A.F., Ieno, E.N., Elphick, C.S. 2009. A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution* 1: 3–14