



## Monográfico

### Revisiones

Zas R. 2006. Consecuencias de la estructura espacial de los datos en el diseño y análisis de experimentos en campo . *Ecosistemas*. 2006/3

(URL: [http://www.revistaecosistemas.net/articulo.asp?Id=434&Id\\_Categoria=1&tipo=portada](http://www.revistaecosistemas.net/articulo.asp?Id=434&Id_Categoria=1&tipo=portada))

## Consecuencias de la estructura espacial de los datos en el diseño y análisis de experimentos en campo

R. Zas

Dpto. Producción Forestal, Centro de Investigaciones Ambientais de Lourizán, Xunta de Galicia, Apdo. 127, 36080 Pontevedra. ([rzas.cifal@siam-cma.org](mailto:rzas.cifal@siam-cma.org))

*Consecuencias de la estructura espacial de los datos en el diseño y análisis de experimentos en campo. Las variables analizadas en los estudios ecológicos de campo muestran con frecuencia estructuras espaciales no aleatorias que pueden condicionar los resultados de las técnicas de análisis de la estadística convencional. Aunque la aplicación de diseños experimentales más o menos complejos mejora la eficiencia estadística de los análisis, en muchas ocasiones estos diseños no se pueden acoplar suficientemente bien al patrón espacial real de las variables objeto de estudio, siendo necesario la aplicación de técnicas más depuradas de análisis espacial para un correcto análisis de los datos. En este artículo se discuten los efectos que la autocorrelación espacial de los datos tiene sobre su análisis mediante técnicas de estadística paramétrica y se propone un método de ajuste espacial, basado en técnicas de geoestadística, para analizar datos espacialmente autocorrelacionados. A su vez se presenta un caso concreto real donde se puede observar como el análisis convencional de datos espacialmente autocorrelacionados puede generar conclusiones totalmente erróneas.*

*Palabras clave: autocorrelación espacial, análisis espacial, geoestadística, semivariogramas.*

*Consequences of spatial heterogeneity in the analysis and design of experiments. Many traits assessed in field ecological trials show nonrandom spatial structures that may affect the efficiency of standard statistical analyses. Although several more or less sophisticated experimental designs may improve this efficiency by controlling the spatial variation, there are many situations where designs can not be properly arranged to actual spatial patterns. In such cases, spatial analysis techniques become essential to correctly analyse spatial autocorrelated data. In this paper, the effects of spatial autocorrelation on the results of conventional statistical analysis are discussed, and a spatial adjustment procedure, based on geostatistics, is proposed to be used when data are spatially autocorrelated. A case study is presented to show how conventional analysis of spatially autocorrelated data may give completely erroneous conclusions.*

*Key words: Spatial autocorrelation, Spatial analysis, Geostatistics, Semivariograms, Kriging*

### Introducción

La mayoría de las variables analizadas en los experimentos de campo, especialmente en ensayos de ecología forestal, muestran una estructura espacial no aleatoria a mayor o menor escala. De hecho, la heterogeneidad espacial es inherente a todo fenómeno ecológico, y en todos los ecosistemas, tanto los organismos vivos como las variables físicas que describen el medio muestran distribuciones que no son ni uniformes ni aleatorias (Legendre, 1993). En esta falta de homogeneidad están implicados desde fenómenos geomorfológicos y variaciones climáticas a gran escala, hasta variaciones a menor escala en las condiciones del suelo, o en la composición de las comunidades biológicas y la interacción entre ellas a través de fenómenos de competencia, predación, parasitismo, etc.

Esta heterogeneidad ambiental puede reflejarse en las variables evaluadas en los ensayos y condicionar el uso de las técnicas de inferencia de la estadística paramétrica convencional (Dutilleul, 1993). Cuando una variable muestra estructura espacial no aleatoria, los valores de ésta tienden a ser más similares entre puntos cercanos que entre posiciones alejadas.



Se dice entonces que los datos están autocorrelacionados, violándose el requisito de independencia de los datos, necesario en los métodos de estadística paramétrica. La distribución de los estadísticos en las pruebas de inferencia se predice en función de los grados de libertad asociados al factor estudiado y al error del modelo, asumiéndose, bajo el supuesto de independencia, que cada observación suma un grado de libertad. Sin embargo, cuando los datos están autocorrelacionados, conociendo el valor en una posición se puede predecir, en cierta medida, el valor en otras posiciones, de tal manera que cada observación, realmente, no supone un grado de libertad completo (Legendre, 1993). Por tanto, de no tenerse en cuenta esta autocorrelación, la distribución predicha de los estadísticos de inferencia no es correcta y las conclusiones del análisis pueden ser erróneas.

La alternativa más simple para controlar la heterogeneidad espacial dentro de un ensayo en campo es la aplicación de un diseño experimental apropiado y eficiente. En general la heterogeneidad ambiental se controla mediante la subdivisión del área de estudio en bloques *homogéneos*. De esta manera, parte de la variación ambiental es absorbida por el efecto bloque, reduciéndose así la varianza del error experimental y facilitando con ello la comparación entre tratamientos. El diseño más comúnmente empleado es el diseño de bloques completos al azar (BCA), en el que cada bloque se subdivide en tantas unidades experimentales (u.e.) como niveles de tratamiento haya, adjudicando los tratamientos a las u.e. aleatoriamente. Sin embargo, en muchos ensayos en campo, especialmente aquellos con especies arbóreas, los bloques del diseño son excesivamente grandes, debido bien al gran número de niveles de cada tratamiento, bien al elevado tamaño de las unidades experimentales. Si el tamaño resultante de los bloques es mayor que la escala de la estructura espacial de los datos, los bloques no son homogéneos y las conclusiones del análisis podrían, de nuevo, ser erróneas. El número, tamaño y distribución de los bloques son, por tanto, parámetros clave del diseño experimental que deberían ajustarse en función del patrón espacial de los datos, pero esta información no suele conocerse a priori.

Un paso más allá en el diseño experimental son los diseños de bloques incompletos (BI) que, mediante la reducción del tamaño de los bloques, resultan más eficientes en el control de la varianza ambiental (Yates, 1936). En estos diseños, no todos los niveles del factor estudiado se aplican en cada bloque y por ello los bloques resultantes son más pequeños. Para poder estudiar la varianza entre bloques con distintos tratamientos, la distribución de los niveles del factor estudiado en los distintos bloques requiere sofisticados algoritmos y el empleo de programas informáticos específicos (e.g. Whitaker *et al.*, 1999). La mejora de la eficiencia de este tipo de diseños en relación al diseño de BCA ha sido ampliamente demostrada, especialmente en ensayos genéticos forestales donde el número de niveles del tratamiento es muy elevado (e.g., 100-200 familias, clones, procedencias, ...) y el tamaño ocupado por cada árbol es también muy grande (4-25 m<sup>2</sup> en función de la edad de evaluación) (Fu *et al.*, 1998) (**Fig. 1**).



**Figura 1.** Vista parcial de un ensayo genético en campo de *Pinus pinaster*. El gran tamaño del ensayo y la irregularidad del terreno hacen previsible una importante heterogeneidad espacial en las distintas variables de estudio.

La reducción del tamaño de los bloques mejora la situación pero puede no ser suficiente para absorber toda la varianza ambiental. Todo depende de la escala de la heterogeneidad espacial en relación al tamaño de los bloques. Si ésta se presenta en forma de manchas de menor tamaño que los bloques del diseño, los bloques no serán homogéneos y los resultados del análisis no serán fiables. Además, la variación ambiental a pequeña escala tiende a ser continua y suave

(Legendre, 1993) por lo que las fronteras artificiales de los bloques nunca se ajustarán de forma precisa a la variación espacial real. Finalmente, el experimentador no suele conocer a priori la estructura espacial de la zona de ensayo por lo que ajustar la forma y localización de los bloques a ésta puede resultar muy difícil.

El objetivo del presente trabajo es (i) describir el impacto de la heterogeneidad espacial en el análisis de experimentos desde un punto de vista teórico, sencillo e intuitivo, (ii) explicar una posible alternativa para un análisis adecuado ajustando los datos según la autocorrelación espacial observada y (iii) aplicar el método propuesto en un caso concreto real a modo de ejemplo. Este trabajo no pretende ser una revisión exhaustiva de los métodos existentes de análisis espacial, sino simplemente una llamada de atención de la importancia del tema en el análisis de experimentos en campo y de las drásticas consecuencias que el no tenerlo en cuenta puede conllevar.

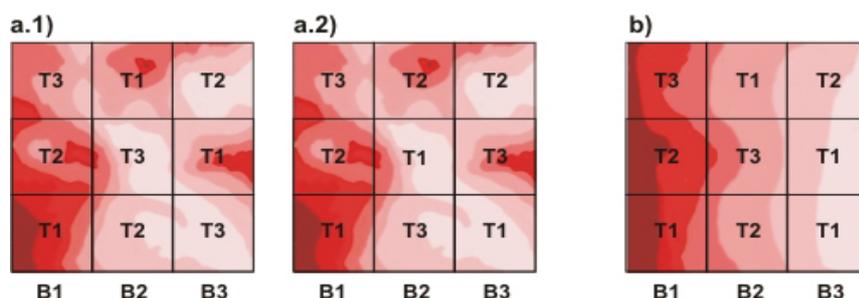
## Repercusiones en el análisis de experimentos

En el análisis de los datos recopilados en experimentos en campo el investigador busca evaluar los siguientes aspectos: i) ¿hay efecto estadísticamente significativo de los factores sobre la variable dependiente?, ii) ¿qué proporción de la variación total observada explica cada factor?, iii) estimación de los efectos del modelo y iv) comparación estadística entre los niveles de cada factor. El análisis convencional de datos autocorrelacionados espacialmente puede afectar de forma diferente (aunque relacionada) a cada uno de estos aspectos del proceso analítico. A continuación se describen estos efectos de forma esquemática centrándonos en el caso más simple de diseños BCA. Las conclusiones son extrapolables a otros diseños experimentales con estructura en bloques.

Significación de los efectos del modelo. En el análisis de varianza, las pruebas de inferencia se basan en estadísticos derivados de la relación entre los cuadrados medios del factor y los del error, o varianza residual. Cuanto mayor son los primeros en relación a los segundos menor es la probabilidad de aceptar la hipótesis nula o, en otras palabras, mayor probabilidad de considerar diferencias *estadísticamente* significativas entre los niveles del factor. La heterogeneidad espacial aumenta la varianza del error experimental y por tanto dificulta la detección de diferencias entre tratamientos. De esta manera, la autocorrelación espacial aumenta la probabilidad de cometer un error de tipo II, aceptar la hipótesis nula ( $H_0$ : sin diferencias entre tratamientos) siendo falsa. Con los diseños en bloques, parte de la varianza ambiental es absorbida por los bloques, reduciéndose la varianza residual. Como ya se ha comentado, la eficacia del diseño depende de lo bien que éstos estén colocados sobre la estructura espacial real y del tamaño de éstos en relación al tamaño de las manchas que definen el patrón espacial de la variable objeto de estudio. En la medida en que la ubicación de los bloques no absorba la variación ambiental, mayor será el error y menor la potencia del análisis convencional (ver Ejemplo más adelante).

Proporción explicada por cada factor. Siguiendo el mismo argumento, el aumento de la varianza del error experimental supone una disminución relativa de la proporción de la variación total explicada por los factores del modelo. Esto es bien conocido en el ámbito de la mejora genética vegetal, donde muchas decisiones se basan en parámetros genéticos derivados de la estimación de los componentes de la varianza en ensayos en campo (e.g. Qiao *et al.* 2000). A modo de ejemplo, Hamman *et al.* (2002) encontraron aumentos de hasta el 40% en la estimación de la heredabilidad en ensayos genéticos de *Alnus rubra* cuando la autocorrelación espacial de los datos se consideraba en los análisis.

Estimación de los efectos del modelo. Cuando los bloques no son homogéneos, ciertos tratamientos resultan *favorecidos* en relación a otros. Esto puede conllevar dos efectos opuestos (**Fig. 2**). Si el reparto aleatorio provoca que ciertos tratamientos se sitúen siempre en zonas favorecidas (o desfavorecidas) dentro de los distintos bloques, la estimación de los efectos correspondientes resultarían sobreestimadas (o subestimadas). Sin embargo, esto no es fácil que ocurra si el reparto aleatorio se ejecuta adecuadamente y, especialmente, si el número de bloques es suficientemente grande. Las sobreestimaciones y subestimaciones tienden a compensarse entre bloques y la estimación final de los efectos no se ve demasiado alterada. Por el contrario, lo que más comúnmente ocurre es que estas sobre- y subestimaciones de los efectos de los tratamientos en los distintos bloques generan una importante y significativa interacción tratamiento x bloque. Esta interacción debería interpretarse como una extrema sensibilidad de los tratamientos a la variación ambiental pero, en realidad, se trata de una consecuencia directa de la heterogeneidad dentro de los bloques, que altera las conclusiones derivadas del análisis. Considerando el elevadísimo número de trabajos científicos de diversa índole donde esta interacción resulta significativa y cuantitativamente importante, el efecto de la heterogeneidad espacial es, probablemente, todavía más frecuente de lo en principio esperable. Por ejemplo, en ensayos de progenie de *Pinus pinaster*, la interacción familia x bloque explicó, de media, el 20% de la variación total, y resultó entre 5 y 10 veces superior a la variación entre familias (Zas, 2006).



**Figura 2.** Esquema de un diseño de bloques completos al azar con 3 bloques (columnas, B1-B3) y tres tratamientos (T1-T3) sobre una estructura espacial en manchas (a.1 y a.2) o en gradiente (b). En a.1 y a.2. el patrón espacial es el mismo, con bloques claramente heterogéneos, pero el reparto aleatorio de los tratamientos dentro de cada bloque es diferente. En a.1 los tratamientos adjudicados a las u.e. favorecidas o desfavorecidas de cada bloque se repiten en los tres bloques (el tratamiento T1 se sitúa en las zonas más favorecidas, tono oscuro, mientras que el T3 se sitúa en las desfavorecidas, tono claro). El resultado es una sobreestimación de los efectos principales, una disminución de sus errores estándar y un aumento de probabilidad de cometer error tipo I (rechazar  $H_0$  cuando es cierta). En a.2 el reparto de tratamientos sobre zonas favorecidas-desfavorecidas no se repite entre bloques provocando la aparición de una fuerte interacción tratamiento x bloque y un aumento del error estándar de la estimación de los efectos principales que dificulta la comparación entre tratamientos (error tipo II, aceptar  $H_0$  cuando es falsa). En b) la estructura en bloques absorbe la mayoría de la variación espacial, por lo que el diseño resulta estadísticamente eficiente.

Comparación entre tratamientos. Otro efecto directo de la heterogeneidad dentro de bloques, también relacionado con la interacción tratamiento x bloque, es el aumento de los errores estándar de la estimación de los efectos del modelo. La heterogeneidad dentro de los bloques genera valores más diferentes de un mismo tratamiento en los distintos bloques, disminuyendo la precisión de la estimación del efecto medio del mismo. Esto tiene consecuencias directas en la comparación entre tratamientos ya que, a mayores errores estándar, más difícil resulta detectar diferencias significativas entre tratamientos. A modo de ejemplo, Qiao *et al.* (2000) observaron que la heterogeneidad espacial aumentaba los errores estándar de la estimación de la producción de líneas genéticas de trigo una media del 40%.

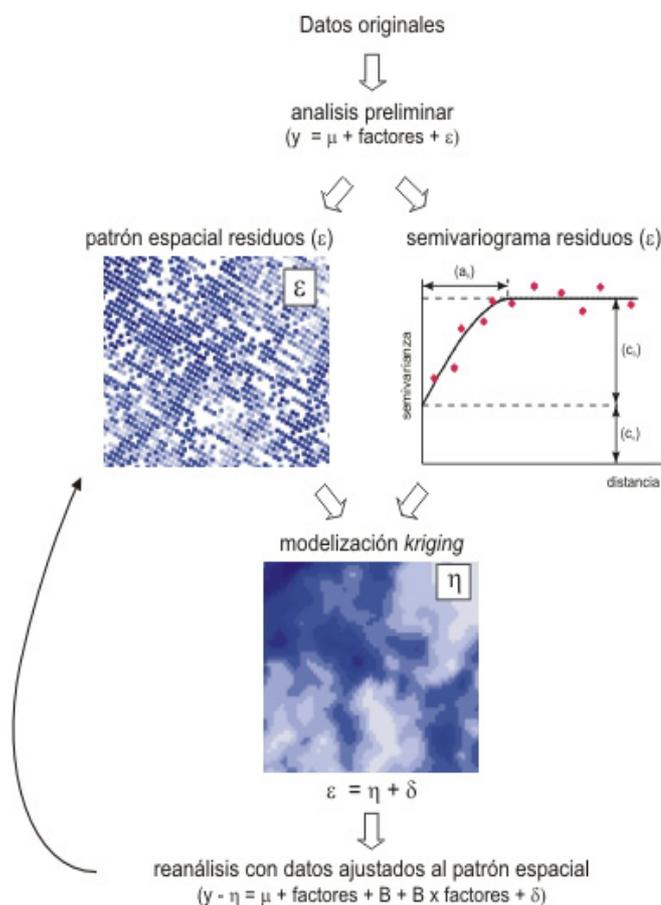
A su vez, cuando hay muchos niveles de tratamiento (como en los ensayos genéticos), aunque en conjunto los efectos de la heterogeneidad dentro de los bloques se diluyen al aumentar el número de bloques, la probabilidad de que algún nivel determinado resulte más o menos favorecido que otro en la mayoría de los bloques aumenta. La comparación estadística entre ellos podría, entonces, resultar completamente inversa a la real. Por ejemplo, Díaz *et al.* (en prensa), seleccionando clones de cerezo resistentes a la cilindrosporiasis, observaron que la heterogeneidad espacial de la infección afecta notablemente al ranking de resistencia del material ensayado.

## Ajustes y correcciones

Existen multitud de métodos propuestos para el análisis de ensayos en campo con datos espacialmente autocorrelacionados (ver Dutkowski *et al.* 2002 para una discusión). En general, todos ellos se basan en la idea de descomponer la variación del error del modelo en un componente espacial y en un error aleatorio espacialmente independiente. Esta descomposición permite *limpiar* (eliminar) de los datos la autocorrelación espacial. El método más intuitivo está basado en la corrección de los datos según los valores de la variable y la variación de ésta en las unidades de muestreo vecinas (*nearest neighbor method*, Wilkinson *et al.*, 1993). Otros métodos modelizan la variación espacial con procedimientos más complejos, como el *trend surface analysis* (e.g. Thompson y El-Kassaby, 1988), *least square smoothing* (Green *et al.*, 1985), o mediante técnicas de geoestadística (e.g. Zaslavski, 2006). Por otra parte, ciertos paquetes estadísticos también permiten la incorporación de la estructura espacial de los residuos en el propio análisis. Así, por ejemplo, en SAS, el procedimiento PROC MIXED permite especificar el modelo de variación espacial esperado de los residuos (lineal, esférico, exponencial,...) (Littell *et al.*, 1996). La descripción detallada de todas estas posibilidades está fuera del alcance de este trabajo y nos limitaremos aquí a exponer brevemente una de ellas.

El procedimiento propuesto para el análisis de datos espacialmente autocorrelacionados (ver **Fig. 3**) se basa en la modelización de la estructura espacial de los residuos del modelo estadístico mediante técnicas geoestadísticas (semivariogramas y *kriging*; Zaslavski, 2006). Cuando el semivariograma de los residuos indica la existencia de un patrón espacial no aleatorio, los parámetros del semivariograma teórico se utilizan, mediante el método *kriging*, para descomponer la variación de los residuos en una componente espacial y un error espacialmente independiente. La componente espacial modeliza la variación espacial de la variable dependiente o, en otras palabras, indica la variación de ésta en función de la ubicación

espacial. Restando esta componente de la variable original, se elimina la autocorrelación espacial de los datos.



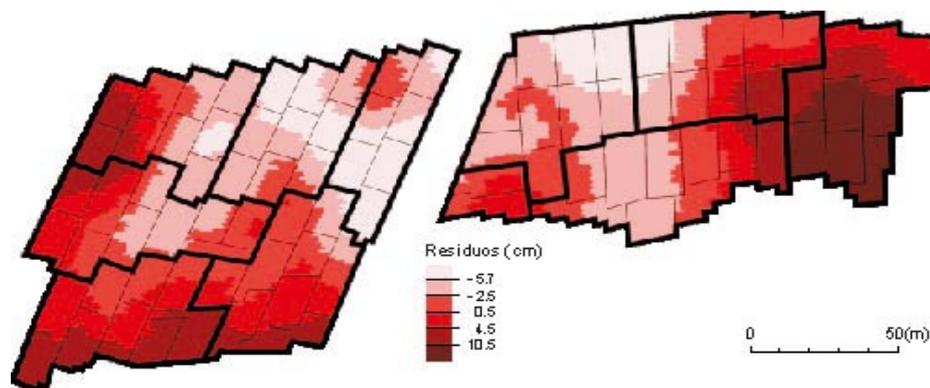
**Figura 3.** Esquema de un procedimiento para analizar ensayos espacialmente heterogéneos basado en técnicas de geoestadística (Zas, 2006). Primero se analiza la variable dependiente sin considerar la estructura en bloques del diseño experimental, obteniéndose los residuos ( $\varepsilon$ ) ajustados según los efectos de los factores del modelo. Estos residuos se analizan espacialmente mediante la construcción del semivariograma observado, el ajuste del semivariograma teórico y la modelización espacial mediante kriging. El método divide la variación espacial de los residuos en una componente espacialmente autocorrelacionada ( $\eta$ ) y un error aleatorio espacialmente independiente ( $\delta$ ). El componente espacial  $\eta$  puede interpretarse como la cantidad de la variable dependiente originada por la posición de la unidad de muestreo y puede utilizarse para corregir la variable original eliminando la estructura espacial de ésta ( $y - \eta$ ). Esta variable corregida puede ahora re-analizarse sin los problemas derivados de la autocorrelación de los datos. El re-análisis genera una nueva estimación de los efectos de los factores del modelo a partir de la cual pueden obtenerse nuevos residuos y repetir el proceso de ajuste iterativamente hasta la convergencia de la estimación de los efectos. Ver Zas (2006) para una descripción detallada del método.

## Ejemplo

Para mostrar los efectos de la heterogeneidad espacial y el ajuste propuesto se ha escogido un ensayo familia x fertilización de *Pinus pinaster* instalado en 2003 en Cotobade (Pontevedra). El ensayo forma parte de una colección de ensayos destinados a evaluar la variación genética en la plasticidad fenotípica respecto a la disponibilidad de nutrientes dentro de la población de mejora de la zona costera de Galicia de ésta y otras especies de pino (ver Zas *et al.*, 2006a, b). El ensayo sigue un diseño *split-plot* replicado en 10 bloques, con 9 tratamientos de fertilización actuando como factor principal y 32 unidades genéticas como factor secundario (28 familias de polinización abierta de árboles superiores seleccionados en la zona costera, 3 lotes comerciales y 1 planta extra de *P. radiata* no incluida en los análisis). Los tratamientos de fertilización se formaron a partir de combinaciones de fertilizantes comerciales mono o bielementales, y se aplicaron en el momento de la plantación. El número total de plantas del ensayo es 2880 que, a marco de 3 x 2 m, ocupan una superficie superior a 2 ha. La variable analizada en este ejemplo es el incremento de altura durante el tercer período vegetativo, buscando determinar si se mantiene

el efecto de la fertilización tres años después de su aplicación.

Como se puede observar en la **Figura 4**, los bloques del diseño (de aproximadamente 0.2 ha cada uno) son claramente heterogéneos. Los residuos  $\varepsilon$  del modelo del análisis de varianza preliminar ( $y = \mu + \text{Fertilización} + \text{Familia} + \text{Fertilización} \times \text{Familia} + \varepsilon$ ) muestran un patrón espacial no aleatorio, con una semivarianza que disminuye claramente a distancias cortas. El modelo esférico se ajusta muy bien al semivariograma observado ( $r^2 = 0.92$ ,  $p < 0.0001$ ), con un rango de ~85 m y una intensidad de la estructura espacial de  $c_o/(c_o+c_n) = 20\%$ , siendo  $c_n$  el valor de la varianza *nugget* y  $(c_o+c_n)$  el valor de la asíntota *sill* del semivariograma teórico (ver artículo de A. Gallardo en este número para información detallada sobre los parámetros del semivariograma).



**Figura 4.** Modelización de la distribución espacial de los residuos del modelo estadístico mediante *kriging* para el incremento de altura en el tercer período vegetativo en un ensayo familia x fertilización de *P. pinaster*. Las líneas gruesas indican las fronteras de los bloques y las líneas finas la subdivisión de éstos en 9 *whole plots* sobre los que se reparten al azar los 9 tratamientos de fertilización. Nótese la alta heterogeneidad ambiental dentro de los bloques en relación a la existente dentro de los *whole plots*.

Los efectos de esta estructura espacial sobre el análisis estadístico se revelan al comparar los resultados del análisis convencional y los obtenidos tras el ajuste iterativo descrito en el apartado anterior (**Tabla 1**). Cuando la autocorrelación de los datos no se considera en el análisis, no se detectan diferencias significativas entre los tratamientos de fertilización. Por el contrario, se observan diferencias significativas entre bloques y una importante interacción fertilización x bloque, lo que ya está sugiriendo, como se vio anteriormente, una importante interferencia de la heterogeneidad espacial de los datos en el análisis.

**Tabla 1.** Análisis de varianza del incremento de altura en el tercer período vegetativo en un ensayo familia x fertilización de *P. pinaster* mediante análisis convencional y ajustando los datos según la autocorrelación espacial que presentan. Tras el ajuste espacial se observa una drástica reducción de los cuadrados medios (CM) asociados al factor bloque y a la interacción fertilización x bloque, manteniéndose el resto de los CM sin cambios significativos. Como consecuencia, ambos factores dejan de ser significativos, mientras que las diferencias entre los tratamientos de fertilización se vuelven altamente significativas para la variable ajustada.

GL: grados de libertad, CM: cuadrados medios

| Fuente de variación | GL   | Término de error | Análisis estándar |      |        | Análisis con ajuste espacial |      |        |
|---------------------|------|------------------|-------------------|------|--------|------------------------------|------|--------|
|                     |      |                  | CM                | F    | p<F    | CM                           | F    | p<F    |
| Bloque (B)          | 9    | F x B            | 5771.4            | 7.59 | 0.0000 | 50.6                         | 0.28 | 0.9783 |
| Fertilización (F)   | 8    | F x B            | 987.2             | 1.30 | 0.2584 | 922.6                        | 5.10 | 0.0000 |
| Familia (G)         | 30   | error            | 999.0             | 4.33 | 0.0000 | 990.5                        | 4.53 | 0.0000 |
| G x F               | 240  | error            | 229.5             | 0.99 | 0.5123 | 223.7                        | 1.02 | 0.3962 |
| F x B               | 72   | error            | 760.6             | 3.30 | 0.0000 | 180.8                        | 0.83 | 0.8492 |
| G x B               | 270  | error            | 241.1             | 1.04 | 0.3087 | 230.5                        | 1.05 | 0.2757 |
| Error               | 1755 |                  | 230.8             |      |        | 218.6                        |      |        |

La escena se vuelve completamente distinta cuando se ajustan los datos en función de la autocorrelación espacial que presentan. Las diferencias entre bloques y la interacción fertilización x bloque dejan de ser significativas y, por el contrario, las diferencias entre los tratamientos de fertilización resultan altamente significativas. Al tratarse de un diseño *split-plot*, el factor principal se testa frente al error de los *whole plots* (la interacción tratamiento x bloque). La autocorrelación espacial infla los cuadrados medios de éstos impidiendo el correcto análisis del factor principal. Por el contrario, el efecto del patrón espacial sobre el factor secundario, las familias, y sobre la interacción familia x fertilización (ambos testados frente al error) es prácticamente insignificante. Esto es debido a que el rango, que indica el tamaño medio de las manchas, es elevado (85 m), resultando mucho más patente la heterogeneidad espacial entre las unidades experimentales del factor principal (*whole plots*) que entre las del factor secundario (*split plots*). De hecho, los cuadrados medios del error no disminuyen tras el ajuste espacial, mientras que los correspondientes a la interacción fertilización x bloque disminuyen notablemente.

Otra forma de ver el efecto de la estructura espacial sobre el análisis es comparando los errores estándar de la estimación de los efectos del modelo. Concordando con lo anterior, la precisión de estas estimaciones tras el ajuste espacial mejora notablemente para el factor principal, reduciéndose los errores estándar alrededor de un 50%. Por el contrario los errores estándar de la estimación de los efectos familiares sólo se reducen un 7%.

Este ejemplo refleja como la heterogeneidad espacial puede afectar de forma drástica a las conclusiones derivadas de los análisis estadísticos: mientras que con el análisis convencional no se detectan diferencias entre los tratamientos de fertilización, éstas resultan altamente significativas cuando se analizan correctamente. A su vez, este ejemplo evidencia la importancia de la escala del patrón espacial en relación al diseño experimental, con impactos claramente distintos en el análisis del factor principal y del factor secundario.

## Agradecimientos

Este trabajo ha sido financiado por los proyectos INIA-RTA2-109 e INIA-RTA5-173. Se agradecen las sugerencias aportadas por L. Sampredo, E. Merlo y F. Maestre, y la ayuda en la edición de las figuras de María y Pablo Zas.

## Referencias

- Díaz, R., Zas, R., Fernández-López, J. *en prensa*. Genetic variation of *Prunus avium* in susceptibility to cherry leaf spot (*Blumeriella jaapii*) in spatially heterogeneous infected seed orchards. *Ann. For. Sci.* *en prensa*.
- Dutilleul, P. 1993. Spatial heterogeneity and the design of ecological field experiments. *Ecology* 74: 1646-1658.
- Dutkowski, G.W., Costa-Silva, J., Gilmour, A.R., Lopez, G.A. 2002. Spatial analysis methods for forest genetic trials. *Can. J. For. Res.* 32: 2201-2214.
- Fu, Y., Clarke, G.P.Y., Namkoong, G. Yanchuk, A.D. 1998. Incomplete block designs for genetic testing: statistical efficiencies of estimating family means. *Can. J. For. Res.* 28: 977-986.
- Green, P.L., Jennison, C., Seheult, A. H. 1985. Analysis of field experiments by least squares smoothing. *J. R. Statis. Soc. B* 47: 299-315.
- Hamann, A., Namkoong, G., Koshy, M.P. 2002. Improving precision of breeding values by removing spatially autocorrelated variation in forestry field experiments. *Silvae Genet.* 51: 210-215.
- Legendre, P. 1993. Spatial autocorrelation: a trouble or new paradigm? *Ecology* 74: 1659-1673.
- Littell, R.C., Milliken, G.A., Stroup, W.W., Wolfinger, R.D. 1996. SAS System for mixed models. SAS Institute, Cary, NC.
- Qiao, C.G., Basford, K.E., DeLacy, I.H., Cooper, M. 2000. Evaluation of experimental designs and spatial analyses in wheat breeding trials. *Theor. App. Gen.* 100: 9-16.
- Thompson, A.J., El-Kassaby, Y.A.. 1988. Trend surface analysis of a Douglas-fir provenance-progeny transfer test. *Can. J. For. Res.* 18: 515-520.
- Whitaker, D., Williams, E.R., John, J.A. 1999. CycDesign. A package for the computer generation of experimental designs. CSIRO Forestry and Forest products, Camberra, Australia.



Wilkinson, G.N., Eckert, S.R., Hancock, T.W., Mayo, O. 1983. Nearest neighbour (NN) analysis of field experiments. *J. R. Statis. Soc. B* 45: 151-211.

Yates, F. 1936. A new method of arranging variety trials involving a large number of varieties. *J. Agr. Sci.* 26: 424-455.

Zas, R. 2006. Iterative kriging for removing spatial autocorrelation in analysis of forest genetic trials. *Tree Genet. & Genomics* DOI 10.1007/s11295-006-0042-4.

Zas, R., Pichel, F., Martins, P., Fernández-López, J.. 2006a. Fertilization x genotype interaction in *Pinus radiata* open pollinated seedlings planted in three locations in Galicia (NW Spain). *New For.* DOI 10.1007/s11056-006-9000-0.

Zas, R., Sampedro, L., Prada, E., Lombardero, M. J., Fernández-López, J. 2006b. Fertilization increases *Hylobius abietis* L. damage in *Pinus pinaster* Ait. seedlings. *For. Ecol. Manage.* 222: 137-144.