

Evaluación de métodos para el análisis estadístico de ensayos comparativos de rendimiento de girasol (*Helianthus annuus* L.)

MONTIEL, M.G.¹; PERELMAN, S.²; DE LA VEGA, A.J.³

RESUMEN

El rendimiento de aceite por unidad de superficie es el principal criterio de selección de la mayoría de los programas de mejoramiento de girasol. Los ensayos comparativos de rendimiento instalados en red en la región girasolera de la República Argentina contribuyen tanto a la toma de decisión de los productores y asesores como a la selección de genotipos superiores por parte de los fitomejoradores. El objetivo de este trabajo es cuantificar las diferencias en capacidad predictiva de diferentes diseños experimentales y métodos de análisis estadístico en la estimación de los efectos genotípicos para rendimiento de aceite (kg ha^{-1}) y su impacto en la selección de híbridos superiores. Se evaluaron 43 ensayos multiambientales de girasol (261 ensayos locales) en cinco campañas. Se comparó la eficiencia relativa del diseño más frecuente en programas de mejoramiento (Bloques Completos Aleatorizados) con otros dos modelos de diseño y análisis de ensayos: Láttice y Espacial. Se obtuvo un ranking de híbridos para cada modelo y con una presión de selección del 20% se evaluó el grado de superposición en la selección entre pares de modelos mediante un coeficiente de coincidencia. El modelo espacial fue el que mejor ajustó los rendimientos de aceite obtenidos a los esperados y también presentó un aumento en la eficiencia de selección de genotipos superiores en relación con el diseño en bloques a través de un mayor coeficiente de coincidencia con el modelo de mayor eficiencia relativa para cada combinación año-subregión (modelo "ideal"). Según estos resultados se podría mejorar fuertemente la precisión de las estimaciones provenientes de las redes oficiales de ensayos comparativos de rendimiento de cultivos sin aumentar la inversión de recursos en términos de número de parcelas y utilizando programas estadísticos ya disponibles.

Palabras clave: diseño experimental, eficiencia relativa, redes oficiales de ensayos, girasol, modelos lineales mixtos, análisis espacial.

ABSTRACT

The oil yield per unit area is the main criterion for selection of most sunflower breeding programs. Comparative performance testing by a network installed throughout the sunflower region of Argentina contributes to both the decision-making of farmers and their advisers, and to the selection of superior genotypes by plant

¹Dto. Métodos Cuantitativos y Sistemas de Información, Facultad de Agronomía, UBA.

Correo electrónico: mmontiel@agro.uba.ar

²Dto. Métodos Cuantitativos y Sistemas de Información, Facultad de Agronomía, UBA. Ifeva, UBA-CONICET. Correo electrónico: perelman@agro.uba.ar

³DuPont Pioneer, 41309 La Rinconada, Sevilla, España. Correo electrónico: abelardo.delavega@pioneer.com

Recibido 02 de enero de 2016 // Aceptado 10 de abril de 2017 // Publicado online 17 de enero de 2018

breeders. The aim of this study is to quantify the differences in predictive ability of different experimental designs and methods of statistical analysis in the estimation of the genotypic effects oil yield (kg ha^{-1}) and its impact on the selection of superior hybrids. Forty-three multi-environmental sunflower trials (261 local trials) were evaluated for five seasons. The relative efficiency of the most common breeding program design (Randomized Complete Blocks) was compared with two alternative models for design and analysis of trials. A ranking of hybrids was obtained for each model and, with a selection pressure of 20%, the degree of overlap was evaluated in the selection of peer models using a coefficient of coincidence. The spatial model was the best fit between actual and expected oil yields, and also showed an increase in the efficiency of selection of superior genotypes in relation to the block design through a higher coefficient of coincidence with the Ideal Model (that of highest relative efficiency). According to these results, the precision of estimates from official networks of comparative crop yield tests could be strongly improved using available statistical packages without increasing the need for investment of additional resources in terms of the number of plots.

Keywords: experimental design, relative efficiency, official trial networks, sunflower, mixed linear models, spatial analysis.

INTRODUCCIÓN

El girasol (*Helianthus annuus* L.) se cultiva en la Argentina en una región amplia y heterogénea, caracterizada por cambios fuertes y significativos en el comportamiento relativo de los cultivares cuando se evalúan en ambientes diferentes (ej., interacciones genotipo \times ambiente ($G \times A$)) (De la Vega y Chapman, 2010). Estas interacciones a menudo explican una porción de variación superior al efecto genotípico (De la Vega y Chapman, 2001) e implican cambios de ranking complicando el proceso de elección del mejor cultivar para cada sistema de producción. Con el objetivo de facilitar las decisiones, en la Argentina y en otros países se conducen rutinariamente redes públicas de ensayos comparativos de rendimiento (ECR) de diferentes cultivos. Para el caso del girasol, entre otros atributos de interés agronómico, se miden y se analizan el rendimiento de aceite (principal criterio de selección) y sus determinantes inmediatos, el rendimiento de grano y el porcentaje de aceite. La red conducida por INTA-ASAGIR es un ejemplo de este tipo de iniciativas.

Las interacciones $G \times A$ pueden ser predecibles o impredecibles, dependiendo de si el factor ambiental subyacente a estas se conoce con anterioridad a la siembra. El conocimiento actual de las interacciones $G \times A$ del girasol en Argentina sugiere que se podría minimizar el efecto de las interacciones predecibles subdividiendo el área de producción en tres megaambientes (norte, centro y sur), y seleccionando híbridos específicamente adaptados a cada uno (Chapman y De la Vega, 2002; De la Vega y Chapman, 2006; De la Vega y Chapman, 2010). Las interacciones impredecibles, en cambio, deberían tratarse: (1) estimando el número mínimo de años, localidades y repeticiones que permitan alcanzar una determinada repetibilidad (ej., heredabilidad en sentido amplio) en cada megaambiente y (2) mejorando la capacidad predictiva de los ECR. La primera de estas estrategias ha sido abordada por De la Vega y Chapman (2006). La segunda constituye el objetivo de este estudio.

Los ECR presentan variabilidad ambiental entre unidades experimentales igualmente tratadas (ej., error experimental; Petersen, 1994; Qiao *et al.*, 2000), la cual nunca puede ser completamente eliminada. Cuanto más precisa sea la modelización de la respuesta en cada ECR individual y más adecuado el modelo de análisis utilizado para estimar y describir los efectos de interacción $G \times A$, mayor será la correlación entre el valor predicho de un cultivar para un atributo en particular y su valor genético verdadero (Cullis *et al.*, 2000).

El diseño usado con más frecuencia en investigación agrícola es el de bloques completos aleatorizados (DBCA), que ofrece una oportunidad para incrementar la precisión en la medida que los bloques realmente reflejen la variabilidad espacial del lote en que se siembra el ensayo. Una alternativa para experimentos donde se evalúa un número grande de genotipos es el diseño en bloques incompletos o láttice (Yates, 1936; Petersen, 1994; Kempton y Fox, 1996), en el que el tamaño de los bloques que se consideran homogéneos para el análisis se reduce a pocas parcelas. Sin embargo, estos arreglos regulares de bloques no siempre aportan un control efectivo de la variación en ECR (Grondona *et al.*, 1996). Los modelos de análisis espacial, que fueron desarrollados para ajustar las correlaciones espaciales entre parcelas vecinas en una o dos dimensiones, han demostrado ser más eficientes que los análisis convencionales en una gran variedad de situaciones (Grondona *et al.*, 1996; Brownie y Gumpertz, 1997; Gilmour *et al.*, 1997; Qiao *et al.*, 2000), justificándose de esta manera su adopción para el análisis de redes de ECR en la mayoría de los cultivos. Los criterios que hasta ahora se utilizan con mayor frecuencia para la estimación de la eficiencia relativa de estos métodos incluyen: (1) la varianza promedio de la diferencia entre medias (AV) y (2) el error estándar promedio de la diferencia entre medias (SED) (Qiao *et al.*, 2000). Estos parámetros se estiman a través del análisis de modelos lineales mixtos utilizando máxima verosimilitud restringida (REML; Patterson y Thompson, 1971).

El objetivo general del presente estudio es cuantificar las diferencias en capacidad predictiva de diferentes métodos de diseño y análisis de ECR para girasol en Argentina. El trabajo está dirigido, en primer lugar, a las redes oficiales de ensayos de girasol, que cada año publican sus resultados para contribuir a que el productor elija con mayor certeza el híbrido para sembrar. En segundo lugar, se pretende contribuir al conocimiento aprovechable por los fitomejoradores para los cuales es imprescindible contar con las más avanzadas herramientas analíticas para maximizar la precisión del proceso de selección de genotipos superiores. En este sentido, nuestra aproximación es general y podría tenerse en cuenta para redes de ensayo con otros cultivos.

MATERIALES Y MÉTODOS

Material experimental

Se utilizaron datos provenientes de 43 ECR multiambientales de girasol (16, 13 y 14 para las regiones centro, norte y sur, respectivamente) sembrados en un total de 261 ambientes (ej., combinaciones año-localidad) (121, 62 y 78 para las regiones centro, norte y sur, respectivamente) de las tres subregiones (megaambientes) girasoleras de Argentina a lo largo de cinco campañas (2004/05 a 2009/10). La zona de estudio incluye la porción este del Chaco argentino (región norte) y el noroeste, centro-oeste (región central) y sur (región sur) de la región pampeana argentina. Todos los ensayos fueron conducidos por Advanta Semillas S.A.I.C. como parte de su programa de mejoramiento genético de girasol y se encuentran declarados en el Instituto Nacional de Semillas, dependiente del MAGPyA de la República Argentina. El suelo, clima y tecnologías de estas regiones fueron descriptos por Hall *et al.* (1992), Mercau *et al.* (2001), Chapman y De la Vega (2002) y De la Vega *et al.* (2007a).

Los ensayos fueron sembrados a una densidad de 47000 plantas ha⁻¹. Se utilizaron parcelas de tres o cuatro surcos de seis metros y un espaciamiento entre hileras de 0,70 m. La mayoría de los ensayos fueron establecidos en siembra directa. Todos los ensayos fueron conducidos en seco. Se evitaron deficiencias nutricionales mediante fertilizaciones. Las malezas e insectos fueron controladas químicamente, pero las enfermedades fúngicas no fueron controladas. Los ensayos que sufrieron algún tipo de daño por aves, insectos, granizo o vientos fueron descartados del estudio. Casi todos los ensayos se sembraron durante el periodo óptimo de siembra para el cultivo en cada región: ej., desde fines de julio hasta principios de septiembre en la región norte y desde finales de septiembre hasta principios de diciembre en las regiones central y sur (De la Vega y Chapman, 2010).

Todos los ensayos fueron establecidos como diseños de láttice cuadrado, con tres o cuatro repeticiones y 25 (5 × 5), 36 (6 × 6), o 49 (7 × 7) genotipos por ensayo. Los ensayos analizados corresponden a los niveles A y B (niveles 2 y 3 en las campañas 2004/05 a 2006/07), es decir, fueron híbridos precomerciales y experimentales del 3.º año de

evaluación, respectivamente. La variable respuesta utilizada en el análisis fue el rendimiento en aceite (kg ha⁻¹), calculado a partir de sus componentes inmediatos: rendimiento de grano (kg ha⁻¹) y concentración de aceite (%). El rendimiento de grano se calculó a partir de la cosecha manual de uno o dos surcos centrales de cada parcela. La concentración de aceite se determinó mediante resonancia magnética nuclear (NMR; Robertson y Morrison, 1979). Todos los resultados se presentan al 11% de humedad.

Metodología de análisis

Se analizó el rendimiento de aceite de cada ensayo individual de acuerdo a los tres métodos propuestos (DBCA, láttice y análisis espacial) para ajustar las correlaciones espaciales entre parcelas vecinas en dos dimensiones siguiendo un modelo AR1 × AR1 y (sin considerar las repeticiones) se obtuvieron los valores de eficiencia relativa (ER) de los modelos de láttice y AR1 × AR1 con respecto al tradicional DBCA medido a través del SED (ER_{SED}Láttice; ER_{SED}Espacial). A continuación, se realizaron tres análisis combinados de todas las localidades para cada ensayo dentro de cada campaña y subregión: (1) todos los ensayos individuales analizados según DBCA, (2) todos los ensayos con análisis espacial y (3) aplicando a cada ensayo individual el método de mayor ER, que en el marco de este estudio se denominó modelo preferido (MP, a la escala de ensayo individual), comparando el ranking de materiales obtenido por cada uno de los dos primeros con el del modelo ideal (que usa el MP para el conjunto total de los ensayos).

Para el análisis del rendimiento de aceite (kg ha⁻¹) en los 261 ambientes individuales se utilizaron los siguientes modelos estadísticos: diseño en bloques completos (DBCA), Láttice y Modelo Espacial (AR1 × AR1), usando el paquete REML de GenStat 12.1 (2009) y se estimaron la deviancia (LOG), AV, SED y el criterio de información Akaike (AIC). Estos modelos y los parámetros estimados están descriptos por Montiel (2014). Para la elección del MP siguiendo el enfoque propuesto por Quiao *et al.* (2000), se buscaron los valores más bajos de AIC, AV y SED y el LOG más alto. Es decir, que en el análisis de ensayos individuales el MP para cada ensayo fue el de mayor eficiencia relativa (ER) por presentar los valores más bajos de AIC, AV y SED y el LOG más alto.

Para cada combinación año-subregión, se realizaron tres metaanálisis. En todos los casos, el efecto genotípico se asumió como aleatorio, la localidad como efecto fijo y las varianzas de las diferentes localidades como independientes (Gilmour *et al.*, 1997; De la Vega y Chapman, 2010). La variabilidad en la respuesta por localidad se modeló: (1) todas como DBCA, (2) todas como AR1 × AR1 y (3) cada localidad según el MP descripto anteriormente. En el tercer modelo, denominado ideal, la variabilidad de la respuesta en cada ensayo individual se modeló aplicando el MP en cada caso, con lo que se espera sea el metaanálisis de mayor ER.

Evaluación de las metodologías

Los análisis de láttice y espacial se evaluaron a través de su ER con respecto al modelo convencional (DBCA) y su impacto en la selección de los híbridos de mayor rendimiento en aceite.

La ER se estimó como:

$$ER_{(SED)} = 100 \times (SED_{DBCA} / SED_{AT})$$

donde: AT: modelo alternativo (láttice o espacial) al DBCA.

La intensidad de selección se fijó en el 0,19 o 0,20 superior (de mayor rendimiento) dependiendo del número de híbridos evaluados en cada ensayo. Para la comparación de la lista de genotipos superiores seleccionados por cada uno de los modelos alternativos con respecto al modelo ideal, se utilizó el coeficiente de Czekanowski (C), una medida de similitud entre resultados del proceso de selección (Snijders *et al.*, 1998; Everitt, 1993) estimada como:

$$C = \frac{2a}{2a + b + c} = \frac{a}{(a + c) + (a + c)} = \frac{a}{a + b} = \frac{a}{a + c}$$

Donde a es el número de genotipos seleccionados tanto por la alternativa como por el MP, b es el número de genotipos seleccionados solo por el modelo alternativo y c es el número de genotipos seleccionados solo por el modelo ideal.

RESULTADOS

En el 75% de los casos, el modelo de análisis de ensayo individual de mayor ER (modelo preferido; MP) fue el que incluyó los términos espaciales (AR1 x AR1), mientras que el 25% restante corresponde al análisis de láttice. En solo uno de los 261 ensayos individuales el modelo de mayor ER fue el de DBCA.

Para los 121 ensayos locales evaluados en la región central, el MP fue el análisis espacial en el 73% de los casos, el 26% correspondió al diseño láttice y el 1% al DBCA. En el caso de la región norte (n = 62), el análisis espacial ajustó mejor en el 81% de los casos y el de láttice en el 19% restante. Por último, en la región sur, en el 74% de los casos (n=78) el MP fue el análisis espacial y en el 26% restante fue láttice.

En promedio, la ER_(SED) de los modelos láttice, espacial e ideal fue siempre mayor respecto del DBCA en las tres regiones (tabla 1). Se observa que en cada una de las regiones la ER del análisis espacial en términos del SED fue siempre superior con respecto al modelo utilizado regularmente por las redes oficiales de ensayos.

Se observaron diferencias regionales en la distribución de frecuencias de las ER_(SED) del análisis espacial respecto del DBCA en cada ensayo individual (fig.1). Se observa que los tres megaambientes presentan la moda en la clase que contiene al 100%, observándose una mayor asimetría en la región norte, la cual no presenta valores menores

Región	ER _(SED) Láttice	ER _(SED) Espacial	ER _(SED) MP
Central	105,50 (4,52)	110,63 (8,42)	111,62 (8,04)
Norte	107,09 (4,67)	115,14 (6,37)	115,73 (6,35)
Sur	105,61 (3,21)	111,93 (4,43)	111,60 (5,21)

Tabla 1. Promedios de ER_(SED) de los modelos Láttice, Espacial y MP en las tres regiones estudiadas. Valores entre paréntesis muestran el desvío estándar de los datos.

Fuente: Elaboración propia (2014).

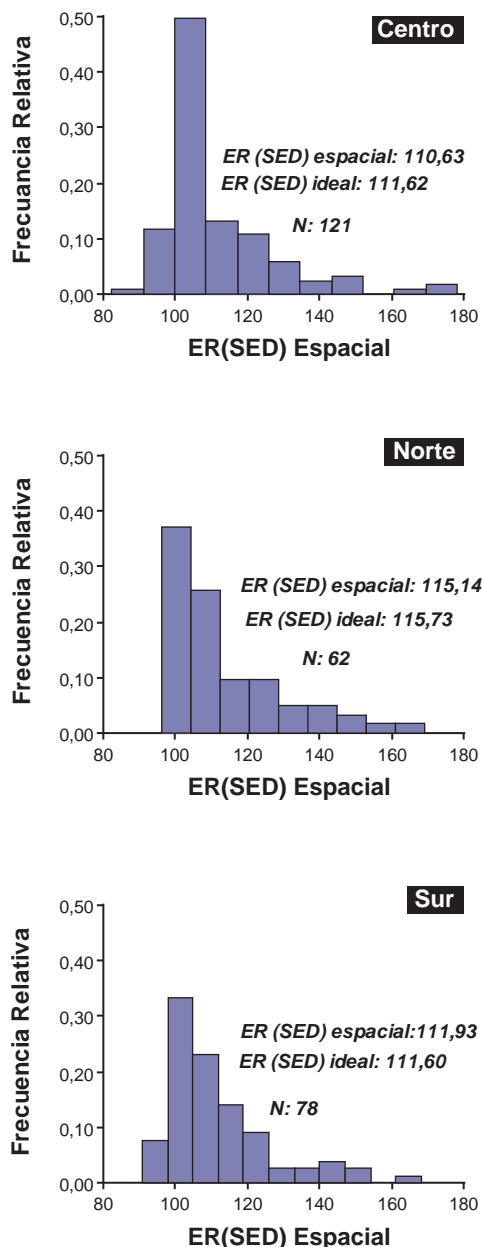


Figura 1. Distribución de frecuencias de la ER_(SED) del análisis espacial en las regiones central, norte y sur. N: número total de ensayos locales en cada región en todos los años. Para cada región se muestra el valor promedio de la ER_(SED) espacial y de la ER_(SED) ideal.

Fuente: elaboración propia (2014).

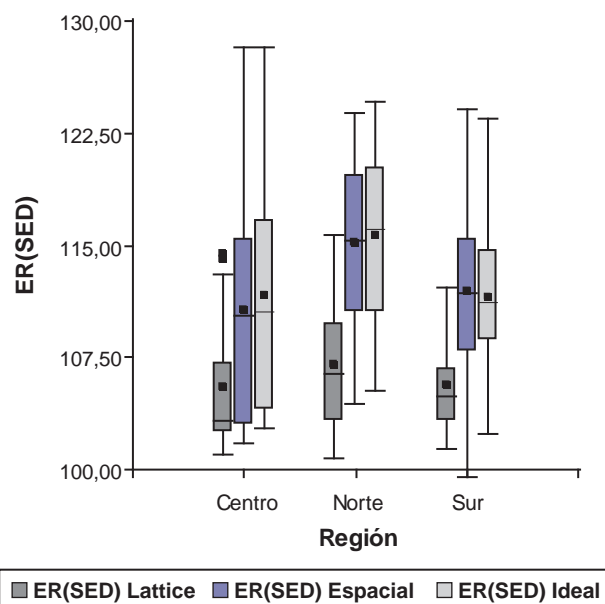


Figura 2. Gráfico Box-plot de la $ER_{(SED)}$ para los modelos Láttice, Espacial e Ideal en cada uno de los megaambientes (región centro, norte y sur).

Fuente: elaboración propia (2014).

al 100% de ER. En promedio, el análisis espacial para la región central fue menos eficiente que en otras regiones, aunque se observan valores extremos de ER muy altos al igual que para la región sur (fig. 1).

En los metaanálisis conducidos por región, la ER del modelo ideal con respecto al DBCA fue siempre mayor en la región norte (fig.2). La región norte mostró una mayor concentración de los valores alrededor de la mediana, con una clara mejora en la eficiencia predictiva de los rendimientos esperados, cuando se utilizó un modelo que contemple las diferencias entre parcelas vecinas.

El porcentaje medio de coincidencia en la selección de los mejores híbridos seleccionados por el modelo espacial y el DBCA con respecto al *MP* mediante el Coeficiente de Czekanowski (C) fue siempre mayor para el análisis espacial que para el DBCA en todas las regiones. En promedio y para las tres regiones en conjunto el coeficiente de coincidencia para el DBCA fue de 0,81 (se estarían rechazando en promedio 8 híbridos por ensayo de los que se hubieran seleccionado con un método de análisis que ajuste mejor los rendimientos esperados a los obtenidos); mientras que para el análisis espacial fueron en promedio 5 los híbridos rechazados cada dos ensayos (coeficiente de coincidencia promedio de 0,94), existiendo diferencias entre subregiones (centro: 0,84-0,95; norte 0,75-0,95; sur: 0,84-0,93).

DISCUSIÓN Y CONCLUSIONES

El *MP* resultó mayormente aquel que incluyó información como el efecto de la repetición, fila, columna, y/o

bloque con muy pocas excepciones. Es decir que tanto la información del diseño del ensayo como los ajustes de las tendencias espaciales resultan relevantes en la elección del *MP* (Qiao *et al.*, 2000; Basford *et al.*, 1996; Grondona *et al.*, 1996; Federer; 1998). Los resultados de esos trabajos coinciden con los del presente estudio, donde la mayor $ER_{(SED)}$ promedio (ca. 113) de 261 ensayos locales coincidió con la adopción de un *MP*, donde el 75% de los casos correspondió al análisis espacial ($AR1 \times AR1$). A igual que lo abordado por Funda *et al.* (2007), la mayoría de estas tendencias espaciales no habrían sido modeladas adecuadamente por el DBCA, el cual siempre presentó menores $ER_{(SED)}$ en la predicción de rindes en comparación con el diseño láttice cuadrado (ca. 106) y espacial (ca. 112). El Criterio de Información Akaike (AIC) también mostró en nuestro trabajo resultados coincidentes con el resto de los criterios considerados y con los presentados por otros autores (Duarte y Vencovsky, 2005; Müller *et al.*, 2010).

La utilización de un *MP* de análisis combinado de ensayos en el que la variabilidad en las respuestas de cada ensayo individual se modela de acuerdo al método de mayor ER (DBCA, láttice o $AR1 \times AR1$) permitió aumentar la certeza de selección en cada una de las tres subregiones girasoleras de la Argentina. El DBCA fue el que presentó menor coincidencia en la selección de híbridos superiores en un ranking de materiales con respecto al modelo ideal, observándose diferencias entre subregiones al igual que para la $ER_{(SED)}$. Esto podría estar afectado por la intensidad de selección utilizada, por lo que los resultados obtenidos de haberse usado una intensidad menor, podrían diferir de los encontrados en este trabajo (Qiao *et al.*, 2000).

En este estudio se muestra una interacción G x A diferencial a lo largo de las distintas subregiones evaluadas. La subregión norte siempre presentó mayores ER del análisis espacial, lo que resalta la importancia de este tipo de análisis en regiones donde las tendencias espaciales no son bien conocidas a la hora de diseñar los ensayos o que presentan una mayor heterogeneidad espacial (Vollmann *et al.*, 2000).

Los resultados de este trabajo muestran que sería posible mejorar la capacidad predictiva actual de las redes oficiales de ensayos de girasol (y probablemente también de otros cultivos) sin aumentar la inversión de recursos en términos de número de parcelas y usando software disponible (Smith *et al.*, 2005; Federer y Crossa, 2005). Esto permitiría la evaluación de mayor cantidad de material genético de una manera eficiente, y a los usuarios de los resultados de estas redes les aportaría una mayor certeza en sus decisiones de elección de genotipos. El análisis espacial demostró ser una herramienta muy efectiva para mejorar la capacidad predictiva de los ECR con respecto a los modelos de análisis tradicionales en casi todas las condiciones ambientales evaluadas, siendo su ER mayor en situaciones de alta heterogeneidad ambiental.

AGRADECIMIENTOS

A Advanta Semillas por su valioso aporte. A Renato González y Aldo Martínez por su colaboración en este trabajo.

BIBLIOGRAFÍA

- BASFORD, K.E.; WILLIAMS, E.R.; CULLIS, B.R.; GILMOUR, A. 1996. Experimental design and analysis for variety trials. United Kingdom: Oxford CAB International.
- BROWNIE, C.; GUMPERTZ, M. 1997. Validity of spatial analyses for large field trials. *J. Agric. Biol. Environ. Stat.* 2:1–23.
- CHAPMAN, S.C.; DE LA VEGA, A.J. 2002. Spatial and seasonal effects confounding interpretation of sunflower yields in Argentina. *Field Crop Res.* 73:107–120.
- CULLIS, B.R.; SMITH, A.; HUNT, C.; GILMOUR, A. 2000. An examination of the efficiency of Australian crop variety evaluation programmes. *J. Agric. Sci. Cambridge*, 135:213–222.
- DE LA VEGA, A.J.; CHAPMAN, S.C. 2001. Genotype by environment interaction and indirect selection for yield in sunflower: II. Three-mode principal component analysis of oil and biomass yield across environments in Argentina. *Field Crop Res.* 72:39–50.
- DE LA VEGA, A.J.; CHAPMAN, S.C. 2006. Defining sunflower selection strategies for highly heterogeneous target population of environments. *Crop Sci.* 46:136–144.
- DE LA VEGA, A.J.; CHAPMAN, S.C. 2010. Mega-environment differences affecting genetic progress for yield and relative value of component traits. *Crop Sci.* 50:574–583.
- DE LA VEGA, A.J.; DELACY, I.H.; CHAPMAN, S.C. 2007a. Progress over 20 years of sunflower breeding in central Argentina. *Field Crop Res.* 100:61–72.
- DUARTE, J.B.; VENCOVSKY, R. 2005. Spatial statistical analysis and selection of genotypes in plant breeding. *Pesq. agropec. Bras.*, vol. 40, n.º 2. 107–114.
- EVERITT, B.S. 1993. Cluster analysis. 3rd ed. Halsted Press, Nueva York.
- FEDERER, W.T. 1998. Recovery for interblock, intergradient, and intervariety information in incomplete block and Láticeice rect-angle design experiments. *Biometrics* 54:471–481.
- FEDERER, W.T.; CROSSA, J. 2005. Designing for and analysing results from field experiments. *Journal of Crop Improvement* 14:29–50.
- FUNDA, T.; LSTIBŮREK, M.; KLAPŠTĚ, J.; PERMEDLOVA, I.; KOBLIHA, J. 2007. Addressing spatial variability in provenance experiments exemplified in two trials with black spruce. *Journal of forest science* 53, 2007 (2): 47–56.
- GENSTAT 12.1. 2009. GenStat for Windows: Introduction. 12nd ed. VSN Int., Oxford, Reino Unido.
- GILMOUR, A.R.; CULLIS, B.R.; VERBYLA, A.P. 1997. Accounting for natural and extraneous variation in the analysis of field experiments. *J. Agric Biol Environ Stat.* 2:269–293.
- GRONDONA, M.R.; CROSSA, J.; FOX, P.N.; PFEIFFER, W.H. 1996. Analysis of variety yield trials using two-dimensional separable ARIMA processes. *Bometrics.* 52:763–770.
- HALL, A.J.; REBELLA, C.M.; GHERSA, C.M.; CULOT, J.P. 1992. Field-crop systems of the Pampas. En: PEARSON, C.J. (ed.) *Field crop ecosystems*. Elsevier, Amsterdam, pp. 413–450.
- KEMPTON, R.A.; FOX, P.N. 1996. *Statistical Methods for Plant Variety Evaluation*.
- MERCAU, J.L.; SADRAS, V.O.; SATORRE, E.H.; MESSINA, C.; BALBI, C.; URIBELARREA, M.; HALL, A.J. 2001. On-farm assessment of regional and seasonal variation in sunflower yield in Argentina. *Agric. Syst.* 37:83–103.
- MÜLLER, B.U.; KLEINKNECHT, K.; MÖHRING, J.; PIEPHO, H.P.; 2010. Comparison of spatial models for sugar beet and barley trials. *Crop Sci.* 50:794–802.
- MONTIEL, M.G. 2014. Evaluación de métodos para el análisis estadístico de ensayos comparativos de rendimiento. Trabajo de intensificación para obtener el grado de Ingeniero Agrónomo otorgado por la Facultad de Agronomía de la Universidad de Buenos Aires.
- PATTERSON, H.D.; THOMPSON, R. 1971. Recovery of inter block information when block sizes are unequal. *Biometrika* 31:100–109.
- PETERSEN, R.G. 1994. *Agricultural Field Experiments: Design and Analysis*. Oregon State University. Corvallis, Oregon.
- QIAO, C.G.; BASFORD, K.E.; DELACY, I.H.; COOPER, M. 2000. Evaluation of experimental designs and spatial analyses in wheat breeding trials. *Theor Appl Genet.* 100:9–16.
- ROBERTSON, J.A.; MORRISON, W.H. 1979. Analysis of oil content of sunflower seed by wide-line NMR. *J. Am. Oil Chem. Soc.*, 56: 961–964.
- SMITH, A.B.; CULLIS, B.E.; THOMPSON, R. 2005. The analysis of crop cultivar breeding and evaluation trials: An overview of current mixed model approaches. *Journal of Agricultural Science*, 143 (6), 449–462.
- SNIJDERS, T.A.B.; KERSHAW, C.D.; ELLIS, R.P. 1998. An investigation of two-dimensional yield variability in breeders' small plot barley trials. *J. Agric. Sci.* 111:419–426.
- VOLLMANN, J.; WINKLER, J.; FRITZ, C.N.; GRAUSGRUBER, H.; RUCKENBAUER, P. 2000. Spatial field variations in soybean (*Glycine max* [L.] Merr.) performance trials affect agronomic characters and seed composition. *Eur J. Agron.* 12:13–22.
- YATES, F. 1936. A new method of arranging variety trials involving a large number of varieties. *J. Agric. Sci.* 23:108–145.