

EVALUACIÓN DE GENOTIPOS O VARIEDADES A TRAVÉS DE AÑOS Y LOCALIDADES. FACTORES CRUZADOS vs ANIDADOS

J. Sahagún-Castellanos^{1¶}; J. E. Rodríguez-Pérez²; A. Peña-Lomelí³

Programa Universitario de Investigación y Servicio en Olericultura Departamento de Fitotecnia,
Universidad Autónoma Chapingo. Km 38.5 Carretera México-Texcoco. Chapingo, Texcoco, México. CP 56230.
Correo-e: jsahagun@taurus1.chapingo.mx ([¶]Autor responsable)

RESUMEN

En las series de experimentos que se realizan en varias localidades (L) durante a años (A) para evaluar variedades o genotipos (G) de cultivos anuales en cada uno de los niveles de un factor B asignado a parcelas grandes, en la definición del modelo se debe considerar la verdadera relación entre los factores A y L. Considerando que una elección errónea de modelo afecta la calidad del análisis, el objetivo de este trabajo fue estudiar el efecto que en la estimación de componentes de varianza, heredabilidad y respuesta a la selección tiene la práctica común de considerar que A y L son cruzados (modelo AxL) cuando en realidad A se anida en L (modelo A/L). El estudio incluyó dos casos: a) Todos los factores fueron aleatorios y b) sólo B fue de efectos fijos. En ambos casos los estimadores de los componentes de varianza de G y de GB tuvieron como sesgo el valor negativo de la varianza de GA y de GAB dividida entre a , respectivamente. Además, la varianza de estos estimadores superó a la de los correspondientes estimadores del modelo A/L; en una magnitud igual a la varianza del estimador insesgado de su sesgo, dividida entre a^2 . La subestimación de la varianza de G obtenida en ambos casos a su vez debe producir una subestimación de la heredabilidad y de la respuesta a la selección ya que la varianza fenotípica tendría una misma estimación en ambos modelos.

PALABRAS CLAVE ADICIONALES: estimación de componentes de varianza, cultivos anuales, heredabilidad, respuesta a la selección, interacción ambiental.

EVALUACIÓN DE GENOTIPOS O VARIEDADES A TRAVÉS DE AÑOS Y LOCALIDADES. FACTORES CRUZADOS VS ANIDADOS

ABSTRACT

In series of experiments done in several locations (L) during a years (A) to evaluate genotypes (G) associated with annual crops in several levels of the main plot factor B the definition of the model has to give consideration to the true relationship between A and L. Considering that a wrong model choice affects the quality of the analysis, the objective of this paper was to study the effect on the estimators of variance components, heritability, and response to selection, of the use of a model based on the common consideration that A and L are crossed factors (model AxL) when in reality A is nested in L (model A/L). Two cases were studied: a) All factors were random, and b) only B was a fixed effects factor. In both cases the estimators of the variance components for G and GB from the model AxL showed a bias which was the negative value of the variance of GA and GAB divided by a , respectively; and their variances were larger than the ones of the corresponding estimators generated in model A/L, the excess being the variance of the unbiased estimator of their respective bias divided by a^2 . The underestimation of the variance of G in turn must produce an underestimation of the heritability and response to selection because the phenotypic variance would be equally estimated in both models.

ADDITIONAL KEY WORDS: Variance component estimation, annual crops, heritability, response to selection, environmental interaction

INTRODUCCIÓN

En la definición de un modelo para describir y analizar la información que produce la evaluación de variedades o genotipos (G) de cultivos anuales en series dobles de

experimentos (un experimento en cada combinación de a años con l localidades), el número de niveles del factor años ha sido visualizado como: a) el número (a) de años cronológicos que incluye la serie, y b) el número (al) de combinaciones de años y localidades. Comúnmente, según

lo muestra la literatura, se ha optado por la primera forma, no obstante que en ocasiones la decisión no sea producto de una reflexión profunda por parte del investigador respecto a esta elección. Elegir erróneamente en un contexto como éste, puede producir sesgos en las estimaciones de componentes de varianza y en la de los parámetros que las usen (Nyquist, 1991).

Debe considerarse a niveles del factor *años* sólo cuando los factores meteorológicos de importancia en los cultivos anuales de secano de cada año inciden de igual manera en cada una de las l localidades. Sahagún (2000) ha considerado que la ocurrencia de esta uniformidad meteorológica es prácticamente imposible en regiones como los Valles Altos de México en que en lugares aún poco distantes es frecuente la ocurrencia de diferencias importantes en la forma en que inciden en un mismo año uno o más de estos factores meteorológicos. El efecto de heladas es un ejemplo. Para estos casos este autor considera que es más razonable aceptar que a cada uno de los a años le corresponden l niveles, de manera que si las características que más definen a cada localidad son constantes a través de los años, la relación debe ser de anidamiento de *años* en *localidades*. También es de considerarse que la evaluación de los niveles de un factor en ocasiones se hace en cada uno de los niveles de otro factor en un diseño en parcelas divididas (McIntosh, 1983). Ante el escenario previamente descrito, el objetivo de esta investigación fue determinar el efecto que tiene la consideración de que *años* tiene a niveles (o que tiene una relación de cruzamiento con *localidades*) cuando en realidad *años* tiene al niveles (o que está anidado en *localidades*). Como es común, *años* y *localidades* se consideraron como factores de efectos aleatorios. El efecto se expresa en términos de sesgos y varianzas de los estimadores de los componentes de varianza asociados a los factores G, B y de su interacción, cuando en un diseño con parcelas divididas G y B se asignan a parcela chica y grande, respectivamente.

MARCO TEÓRICO

Para el diseño experimental considerado, bloques al azar en parcelas divididas, variedades o genotipos (G) fue asignado a las parcelas chicas y un factor adicional B a las grandes. Para el caso de cruzamiento entre años (A) y localidades (L), la descripción del valor fenotípico de la parcela del bloque k que recibió los niveles n y m de G y B, respectivamente, en la localidad i y el año j (Y_{ijkmn}) se hizo en la forma:

$$Y_{ijkmn} = \mu + L_i + A_j + (AL)_{ij} + R_{k(ij)} + B_m + (BL)_{im} + (AB)_{jm} + (ABL)_{ijm} + E_{km(ij)} + G_n + (GL)_{in} + (GA)_{jn} + (GB)_{mn} + (GBL)_{imn} + (GAL)_{ijn} + (GAB)_{jmn} + (GABL)_{ijmn} + E_{ijkmn} \quad (1)$$

donde $i=1,2,\dots,l$; $j=1,2,\dots,a$; $k=1,2,\dots,r$; $m=1,2,\dots,b$; $n=1,2,\dots,g$; μ = media general; L_i = Efecto de la localidad i; A_j = Efecto del año j; $R_{k(ij)}$ = Efecto de la repetición k de la localidad i en el año j; B_m = Efecto del nivel m de B; G_n = Efecto del genotipo n; $E_{km(ij)}$ y E_{ijkmn} son los errores de parcela grande y chica, respectivamente. Los efectos restantes son interacciones entre los efectos involucrados.

Para el anidamiento de *años* en *localidades* el modelo utilizado no incluye los términos de interacción que tengan A y L, éstos van incluidos en los términos en que aparece $A_{j(i)}$ (*años* anidado en *localidades*). Así, conservando la notación anterior el modelo utilizado fue (Sahagún, 2000):

$$Y_{ijkmn} = \mu + L_i + A_{j(i)} + R_{k(ij)} + B_m + (BL)_{im} + (AB)_{mj(i)} + E_{mk(ij)} + G_n + (GB)_{mn} + (GL)_{in} + (GBL)_{imn} + (AG)_{nj(i)} + (GAB)_{mnj(i)} + E_{mnkij} \quad (2)$$

Cuando los modelos (1) y (2) fueron considerados aleatorios, todos los términos del lado derecho, excepto μ , fueron tratadas como variables aleatorias no correlacionadas entre ni dentro de sí, con una media igual a cero y una varianza igual para todos los efectos de un factor o de una interacción.

Las esperanzas de cuadrados medios fueron determinadas según las reglas que para ese efecto describe Lindman (1992), y los componentes de varianza se estimaron con el método de momentos.

RESULTADOS Y DISCUSIÓN

Modelo aleatorio

De acuerdo con los cuadrados medios (CM) y con sus valores esperados en el modelo con anidamiento de A en L (modelo A/L) (Sahagún, 2000) y en el de cruzamiento entre A y L (modelo AxL) (Cuadro 1), los estimadores de la varianza genotípica (σ_G^2) que se generan en estos modelos, $_{AxL} \hat{\sigma}_G^2$ y $_{A/L} \hat{\sigma}_G^2$, respectivamente, y el de la varianza de la interacción entre G y A ($_{AxL} \hat{\sigma}_{GA}^2$) en el modelo AxL se expresan en la forma:

$$_{A/L} \hat{\sigma}_G^2 = [CM(G) + CM(GBL) - CM(CB) - CM(GL)] / (ablr) \quad (3)$$

CUADRO 1. Cuadrados medios (CM), sus valores esperados [E(CM)] y pruebas de F del análisis de varianza de una serie de experimentos con r repeticiones (R) en a años (A) anidados en l localidades (L) para evaluar g genotipos (G) en b niveles de un factor B asignado a la parcela grande. Modelo Aleatorio.

		Esperanza de cuadrados medios ²																	F	
FV	CM	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	
L	C ₁	p ²	p		p	p			p		p	p		p		p	p		p	$[C_1+C_8+C_{14}+C_{15}]/[C_3+C_6+C_{11}+C_{17}]$
A	C ₂	p	p	p	p			p			p	p	p			p	p	p		$[C_2+C_8+C_{15}+C_{16}]/[C_3+C_7+C_{12}+C_{17}]$
AL	C ₃	p	p		p						p	p				p	p			$[C_3+C_9+C_{17}]/[C_4+C_8+C_{15}]$
R(AL)	C ₄	p									p					p				$[C_4/C_9]$
B	C ₅	p	p	p		p	p				p	p	p	p	p					$[C_5+C_8+C_{14}+C_{16}]/[C_6+C_7+C_{13}+C_{17}]$
BL	C ₆	p	p			p					p	p		p						$[C_6+C_{17}]/[C_8+C_{14}]$
AB	C ₇	p	p	p							p	p	p							$[C_7+C_{17}]/[C_8+C_{16}]$
ABL	C ₈	p	p								p	p								$[C_8+C_{18}]/[C_9+C_{17}]$
Ea	C ₉	p									p									C_9/C_{18}
G	C ₁₀	p	p	p	p	p		p	p	p										$[C_{10}+C_{14}+C_{15}+C_{16}]/[C_{11}+C_{12}+C_{13}+C_{17}]$
GL	C ₁₁	p	p		p	p			p											$[C_{11}+C_{17}]/[C_{14}+C_{15}]$
GA	C ₁₂	p	p	p	p			p												$[C_{12}+C_{17}]/[C_{15}+C_{16}]$
GB	C ₁₃	p	p	p		p	p													$[C_{13}+C_{17}]/[C_{14}+C_{16}]$
GBL	C ₁₄	p	p			p														C_{14}/C_{17}
GAL	C ₁₅	p	p		p															C_{15}/C_{17}
GAB	C ₁₆	p	p	p																C_{16}/C_{17}
GABL	C ₁₇	p	p																	C_{17}/C_{18}
Eb	C ₁₈	p																		

²Cada número representa la varianza del efecto cuyo CM tiene ese número multiplicada por un coeficiente.

Los coeficientes entre (18), (17), (16), (15),..., (5), (4), (3), (2), (1) son 1, r, rl, rb, ra, ral, rbl, abr, g, gr, glr, agr, aglr, bg, brg, bgrl y abgr, respectivamente.

$${}_{AxL}\hat{\sigma}_G^2 = \left[\frac{CM(G) + CM(GBL) + CM(GAL) + CM(GAB)}{-CM(GL) - CM(GA) - CM(GB) - CM(GABL)} \right] / (ablr) \quad (4)$$

$${}_{AxL}\hat{\sigma}_{GA}^2 = [CM(GA) + CM(GABl) - CM(GAB) - CM(GAl)] / (blr) \quad (5)$$

Por lo tanto, según las Ecuaciones 3, 4, y 5,

$${}_{AxL}\hat{\sigma}_G^2 = {}_{A/L}\hat{\sigma}_G^2 - {}_{AxL}\hat{\sigma}_{GA}^2 / a \quad (6)$$

Además, como ${}_{A/L}\hat{\sigma}_G^2$ y ${}_{AxL}\hat{\sigma}_G^2$ no involucran cuadrados medios en común, la varianza (Var) de ${}_{AxL}\hat{\sigma}_G^2$ según la Ecuación 6 es

$$\text{Var}({}_{AxL}\hat{\sigma}_G^2) = \text{Var}({}_{A/L}\hat{\sigma}_G^2) + \text{Var}[{}_{AxL}\hat{\sigma}_{GA}^2 / a] \quad (7)$$

En síntesis, el uso del modelo basado en el cruzamiento entre A y L cuando en realidad A está anidado en L genera un estimador de σ_G^2 , ${}_{AxL}\hat{\sigma}_G^2$, que tiene un sesgo igual a $-{}_{AxL}\sigma_{GA}^2/a$ y una mayor varianza (menor precisión) que la del estimador de $\hat{\sigma}_G^2$, que produce el modelo A/L (${}_{A/L}\hat{\sigma}_G^2$). La estimación de la heredabilidad

con base en la estimación que de σ_G^2 hace ${}_{AxL}\hat{\sigma}_G^2$ sería una subestimación y ésta sería transmitida a toda respuesta a la selección que se estime con base esta heredabilidad. Estas conclusiones también se basan en la consideración de que la varianza fenotípica no depende del modelo, es la misma cualquiera que éste sea.

Los estimadores (${}_{AxL}\hat{\sigma}_B^2$ y ${}_{AxL}\hat{\sigma}_{GB}^2$) de los componentes de varianza de B y de GB (σ_B^2 y σ_{BG}^2) que produjo un análisis similar al hecho para σ_G^2 en el modelo de la Ecuación 1 (modelo AxL) cuando el modelo de la Ecuación 2 (modelo A/L) es el que corresponde a la realidad, también resultaron ser sesgados. Sus sesgos (Cuadro 2) siguen el mismo patrón que el seguido por el estimador ${}_{AxL}\hat{\sigma}_G^2$; es decir, son iguales al valor negativo de la varianza de la interacción de B con A y de BG con A dividido entre a , respectivamente. En el ámbito aplicado, la subestimación de σ_B^2 y σ_{BG}^2 puede hacer que el investigador tienda a asignar una menor ponderación que la debida a la necesidad de evaluar los genotipos en varios niveles de B.

Por lo que respecta a la varianza de los estimadores de σ_B^2 y σ_{BG}^2 producidos en el modelo AxL (${}_{AxL}\hat{\sigma}_B^2$ y ${}_{AxL}\hat{\sigma}_{GB}^2$), su comportamiento también es similar a la del estimador ${}_{AxL}\hat{\sigma}_G^2$; es decir, en cada caso supera a la varianza del estimador correspondiente que se genera en modelo adecuado (modelo A/L) en una cantidad igual a la varianza del estimador insesgado de su sesgo dividida entre a^2 (Cuadro 2). Es decir, estos estimadores no sólo son insesgados, también son menos precisos.

CUADRO 2. Sesgo y varianza de los estimadores de σ_G^2 , σ_B^2 y σ_{GB}^2 .

Los estimadores se generaron según el modelo AxL (Ecuación 1) cuando la verdadera relación es de anidamiento de A en L (A/L) (Ecuación 2). Modelo Aleatorio.

Estimador	Sesgo	Varianza (Var)
${}_{AxL}\hat{\sigma}_G^2$	$-{}_{AxL}\sigma_{GA}^2/a$	$\text{Var}[{}_{A/L}\hat{\sigma}_G^2] + \text{Var}[{}_{AxL}\hat{\sigma}_{GA}^2]/a^2$
${}_{AxL}\hat{\sigma}_B^2$	$-{}_{AxL}\sigma_{AB}^2/a$	$\text{Var}[{}_{A/L}\hat{\sigma}_B^2] + \text{Var}[{}_{AxL}\hat{\sigma}_{AB}^2]/a^2$
${}_{AxL}\hat{\sigma}_{GB}^2$	$-{}_{AxL}\sigma_{GAB}^2/a$	$\text{Var}[{}_{A/L}\hat{\sigma}_{GB}^2] + \text{Var}[{}_{AxL}\hat{\sigma}_{GAB}^2]/a^2$

B es de efectos fijos

El caso en que sólo el factor B es de efectos fijos (densidades de siembra, dosis de fertilización, etc) puede ser de alto valor aplicado en la evaluación de genotipos. En este caso, de acuerdo con los cuadrados medios y con sus valores esperados (Cuadro 3), los estimadores de la varianza genética según los modelos A/L y AxL son, respectivamente:

$${}_{A/L}\hat{\sigma}_G^2 = [\text{CM}(G) - \text{CM}(GL)]/(ablr) \quad (8)$$

$${}_{AxL}\hat{\sigma}_G^2 = [\text{CM}(G) + \text{CM}(GAL) - \text{CM}(GA) - \text{CM}(GL)]/(abr) \quad (9)$$

Según estos últimos dos resultados (Ecuaciones 8 y 9) y la expresión del estimador de ${}_{AxL}\sigma_{GA}^2$, también se obtiene que

$${}_{AxL}\hat{\sigma}_G^2 = [\text{CM}(G) + \text{CM}(GL)]/(ablr) - [\text{CM}(GA) - \text{CM}(GAL)]/(ablr) \quad (10)$$

$$= {}_{A/L}\hat{\sigma}_G^2 - {}_{AxL}\hat{\sigma}_{GA}^2/a$$

En virtud de que ${}_{A/L}\hat{\sigma}_G^2$ y ${}_{AxL}\hat{\sigma}_{GA}^2$ son estimadores

insesgados en sus respectivos modelos de σ_G^2 y σ_{GA}^2 , respectivamente, se deduce que el estimador ${}_{AxL}\hat{\sigma}_G^2$ tiene un sesgo igual a $-{}_{AxL}\sigma_{GA}^2/a$ (como en el modelo aleatorio, ${}_{AxL}\sigma_{GA}^2$ es la varianza de la interacción entre G y A, definida sólo en el modelo AxL). Por otra parte, la varianza del estimador de $\hat{\sigma}_G^2$ generado en el modelo AxL cuando sólo B es factor de efectos fijos de acuerdo con la Ecuación 10 es

$$\text{Var}[{}_{AxL}\hat{\sigma}_G^2] = \text{Var}[{}_{A/L}\hat{\sigma}_G^2] + \text{Var}[{}_{AxL}\hat{\sigma}_{GA}^2/a] \quad (11)$$

Con respecto a la varianza de la interacción GB, sus estimadores según los modelos AxL y A/L cuando sólo B es factor de efectos fijos, de acuerdo con los cuadrados medios y sus valores esperados [Cuadro 3, Sahagún (2000), respectivamente] son:

$${}_{AxL}\hat{\sigma}_{GB}^2 = [\text{CM}(GB) + \text{CM}(GABL) - \text{CM}(GBL) - \text{CM}(GAB)]/alr \quad (12)$$

$${}_{A/L}\hat{\sigma}_{GB}^2 = [\text{CM}(GB) - \text{CM}(GBL)]/alr \quad (13)$$

Además, del modelo AxL con sólo B de efectos fijos

$${}_{AxL}\hat{\sigma}_{GBA}^2 = [\text{CM}(GAB) - \text{CM}(GABL)]/(alr) \quad (14)$$

De estos tres resultados, Ecuaciones 12, 13 y 14, se obtiene que

$${}_{AxL}\hat{\sigma}_{GB}^2 = [\text{CM}(GB) - \text{CM}(GBL)]/(alr) - [\text{CM}(GAB) - \text{CM}(GABL)]/(alr)$$

$$= {}_{A/L}\hat{\sigma}_{GB}^2 - {}_{AxL}\hat{\sigma}_{GBA}^2/a \quad (15)$$

Además,

$$\text{Var}[{}_{AxL}\hat{\sigma}_{GB}^2] = \text{Var}[{}_{A/L}\hat{\sigma}_{GB}^2] + \text{Var}[{}_{AxL}\hat{\sigma}_{GBA}^2/a] \quad (16)$$

Los resultados relativos a los estimadores de los componentes de varianza σ_G^2 y σ_{GB}^2 del caso en que sólo el factor B es fijo (Ecuaciones 10, 11, 15 y 16) tienen las mismas expresiones que las relativas a los estimadores de los mismos componentes en el modelo en que todos los factores son aleatorios (Cuadro 1); es decir, el cambio del factor B de aleatorio a fijo no afectó la forma general de los estimadores de σ_G^2 y σ_{GB}^2 , ni de su varianza, aunque sí “simplificó” las expresiones para los estimadores que ahora involucran menos cuadrados medios y son, en consecuencia más precisos.

CUADRO 3. Cuadrados medios (CM), sus valores esperados y Fs del análisis de varianza de una serie de experimentos con *r* repeticiones (R) en *a* años (A) anidados en *l* localidades (L) para evaluar *g* genotipos (G) en *b* niveles de un factor B asignado a la parcela grande. Sólo B es de efectos fijos.

FV	Esperanza de cuadrados medios ^z																		F	
	CM	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2		1
L	C ₁	P ^z			p				p							p	p		p	[C ₁ +C ₅]/[C ₃ +C ₁₁]
A	C ₂	p			p			p								p	p	p		[C ₂ +C ₁₅]/[C ₃ +C ₁₂]
AL	C ₃	p			p											p	p			[C ₃ +C ₁₈]/[C ₁₄ +C ₁₅]
R(AL)	C ₄	p														p				C ₄ /C ₁₈
B	C ₅	p	p	p		p	p				p	p	p	p	p					[C ₅ +C ₈ +C ₁₄ +C ₁₆]/[C ₆ +C ₇ +C ₁₃ +C ₁₇]
BL	C ₆	p	p			p					p	p		p						[C ₆ +C ₁₇]/[C ₈ +C ₁₄]
AB	C ₇	p	p	p							p	p	p							[C ₇ +C ₁₇]/[C ₈ +C ₁₆]
ABL	C ₈	p	p								p	p								[C ₈ +C ₁₈]/[C ₉ +C ₁₇]
Ea	C ₉	p									p									C ₉ /C ₁₈
G	C ₁₀	p			p			p	p	p										[C ₁₀ +C ₁₅]/[C ₁₁ +C ₁₂]
GL	C ₁₁	p			p				p											C ₁₁ /C ₁₅
GA	C ₁₂	p			p			p												C ₁₂ /C ₁₅
GB	C ₁₃	p	p	p		p	p													[C ₁₃ +C ₁₇]/[C ₁₄ +C ₁₆]
GBL	C ₁₄	p	p			p														C ₁₄ /C ₁₇
GAL	C ₁₅	p			p															C ₁₅ /C ₁₈
GAB	C ₁₆	p	p	p																C ₁₆ /C ₁₇
GABL	C ₁₇	p	p																	C ₁₇ /C ₁₈
Eb	C ₁₈	p																		

²Los números corresponden a la varianza del efecto cuyo CM tiene ese número multiplicada por un coeficiente.

Los coeficientes en (18), (17), (16), (15),..., (5), (4), (3), (2), (1) son 1, r, rl, rb, ra, ral, rbl, abr, abrl, g, gr, glr, agr, aglr, bg, brg, bgrl y abgr, respectivamente.

CONCLUSIONES

El uso inapropiado del modelo aleatorio que da cuenta del cruzamiento de *a* años (A) y *l* localidades (L) cuando en realidad A se anida en L produjo estimadores de los componentes de varianza de genotipos (G), de un factor adicional B y de la interacción GB que tienen un sesgo negativo igual a la varianza de la interacción GA, AB y GAB dividida entre *a*, respectivamente. Además, la varianza de dichos estimadores siempre superó a la de los correspondientes estimadores que se generaron en el modelo con anidamiento. En el caso en que sólo B fue de efectos fijos, los estimadores de los componentes de

varianza de G y de GB tuvieron, en lo general, las mismas propiedades que las anotadas para el modelo aleatorio.

LITERATURA CITADA

- LINDMAN, H. R. 1992. Analysis of Variance in Experimental Designs. Springer Verlag. New York, USA. 531 p.
- McINTOSH, M. S. 1983. Analysis of combined experiments. Agron. J. 75: 153-155.
- NYQUIST, W. E. 1991. Estimation of heritability and prediction of selection response in plant populations. Critical Reviews in Plant Sciences 10: 235-322.
- SAHAGÚN C., J. 2000. Evaluación de genotipos con varios niveles de un factor en series de experimentos. Agrociencia 34: 193-206.