# MEJOR PREDICTOR LINEAL E INSESGADO PARA APTITUD COMBINATORIA ESPECÍFICA DE LOS DISEÑOS DOS Y CUATRO DE GRIFFING

# BEST LINEAR UNBIASED PREDICTOR FOR SPECIFIC COMBINING ABILITY OF GRIFFING'S DESIGNS TWO AND FOUR

**Osval Antonio Montesinos López[1], Ángel Martínez Garza[1†], Ángel Agustín Mastache Lagunas[2*] y Gilberto Rendón Sánchez[1]**

[1]Programa en Estadística, Instituto de Socioeconomía, Estadística e Informática, Colegio de Postgraduados. Km. 36.5 Carr. México-Texcoco. C.P. 56230. Montecillo, Texcoco, Edo. de México. [2]Centro de Estudios Profesionales, Colegio Superior Agropecuario del Estado de Guerrero. Av. Vicente Guerrero No. 81, 1er. Piso, C.P. 40000 Iguala, Gro. Tel. y Fax: 01(733) 332-4328. Correo electrónico: gama32@starmedia.com

** Autor para correspondencia*

## RESUMEN

**La estimación de parámetros en los experimentos de cruzas dialélicas ayuda a tomar decisiones objetivas en los programas de mejoramiento genético. En estos experimentos la estimación de aptitud combinatoria específica se ha efectuado como si el modelo fuera de efectos fijos, cuando se sabe que su naturaleza es aleatoria. Por ello en el presente trabajo, para los diseños dos y cuatro de Griffing establecidos en un diseño experimental de bloques completos al azar, se derivan los mejores predictores lineales e insesgados de aptitud combinatoria específica, bajo el modelo de efectos mixtos. Además, se presenta un algoritmo computacional en SAS-IML para la obtención de estos estimadores.**

**Palabras clave:** Cruzas dialélicas, aptitud combinatoria específica, efectos fijos, modelo de efectos mixtos.

## SUMMARY

**Parameters estimation in diallel crosses experiments aids to objective decision making in plant breeding programs. In these experiments the estimation of specific combining ability has been done on the basis of a fixed effects model, in spite of their random nature. For this reason, in the present work in Griffing's designs two and four under the correct mixed effects model, the empirical best linear unbiased predictors for specific combining abilities, are derived. Furthermore, a computational algorithm in SAS-IML is presented to obtain such predictors.**

**Index words:** Diallel crosses, specific combining ability, fixed effects, mixed effects model.

## INTRODUCCIÓN

Los experimentos de cruzas dialélicas son utilizados por los genetistas, en investigaciones enfocadas al mejoramiento de plantas y animales. Los diseños dos y cuatro de Griffing (1956a, 1956b) han sido extensamente utilizados para estimar efectos de aptitud combinatoria general (ACG) y específica (ACE) y de sus correspondientes componentes de varianza; todo esto importante para la toma de decisiones en los programas de mejoramiento.

Sin embargo, la metodología existente para el análisis de estos diseños se ha basado en el modelo de efectos fijos con el cual se obtienen estimadores insesgados, pero no de mínima varianza. Aunque los estimadores de aptitud combinatoria general, sobre la base del modelo de efectos mixtos, ya fueron obtenidos por Mastache *et al.* (1999), con la metodología desarrollada por Henderson (1963, 1973), Harville (1976) y Harville y Carriquiry (1992), hasta ahora no se ha hecho una investigación similar sobre los efectos de aptitud combinatoria específica.

En esta investigación se derivó la metodología para estimar los efectos de aptitud combinatoria específica, en los diseños dos y cuatro de Griffing, sobre la base del modelo de efectos mixtos, debido a la naturaleza aleatoria de los efectos de ACE, se ilustra la metodología mediante un ejemplo, y se presenta un algoritmo computacional en comandos SAS-IML que permite realizar la aplicación de la metodología propuesta.

## MARCO TEÓRICO

### Modelo lineal de efectos mixtos

En términos matriciales el modelo lineal de efectos mixtos se representa como

$$Y = X\alpha + Z\theta + \varepsilon \qquad \text{(Ec. 1)}$$

donde $Y$ es un vector $nx1$ de observaciones; $X$ es una matriz diseño $nxf$ conocida; $\alpha$ es un vector $fx1$ de

Recibido: 10 de Noviembre del 2003.
Aceptado: 26 de Julio del 2005.

parámetros desconocidos, de efectos fijos; $Z$ es una matriz diseño $nxp$ conocida; $\theta$ es el vector $px1$ no observable de efectos aleatorios; con $\theta \sim (0, G\sigma_e^2)$, y $\varepsilon$ es un vector $nx1$ no observable de efectos residuales o términos de error tal que $\varepsilon \sim (0, R\sigma_e^2)$.

Además,

$$Var(y) = V\sigma_e^2 = (ZGZ' + R)\sigma_e^2,\ E(y) = X\alpha\ .$$

En la mayoría de las aplicaciones $G$ es una matriz diagonal; es decir, $G = (\sigma_\theta^2 / \sigma_e^2) I_p$, $\sigma_\theta^2$ es la varianza del término aleatorio $\theta$, $\sigma_e^2$ es la varianza de los términos de error, $I_p$ es una matriz idéntica de tamaño $pxp$; y $R$ frecuentemente es una matriz identidad. Con relación al modelo (Ec. 1), Henderson (1963, 1973) desarrolló una técnica para abordar los aspectos aleatorios y derivó los mejores predictores lineales e insesgados (MPLI).

### Método de Henderson para obtener los MPLI

El método propuesto por Henderson (1963, 1973) para obtener los MPLI sobre la base del modelo de efectos mixtos (Ec. 1) supone que $\theta$ y $\varepsilon$ son vectores aleatorios no observables tales que $\theta \sim N_p(0, G\sigma_e^2)$ y $\varepsilon \sim N_n(0, R\sigma_e^2)$, en donde $G$ y $R$ son matrices no singulares. Según Henderson (1963, 1973), la densidad conjunta de $\theta$ y $y$ es $f(y,\theta) = d(y/\theta)c(\theta)$, donde $c(\theta)$ es la densidad marginal de $\theta$ y $d(y/\theta)$ es la distribución condicional de $y$ dado $\theta$. En forma explicita $f(y,\theta)$ es:

$$f(y,\theta) \propto \left\{ e^{\frac{-1}{2\sigma_e^2}(y - X\alpha - Z\theta)' R^{-1} (y - X\alpha - Z\theta)} \right\}$$

$$\left\{ e^{\frac{-1}{2\sigma_e^2}\theta' G^{-1}\theta} \right\} \qquad \text{(Ec. 2)}$$

Al maximizar la densidad conjunta (Ec. 2), con respecto a $\alpha$ y $\theta$, se derivan las ecuaciones normales del modelo mixto:

$$\begin{aligned} XR^{-1}X\hat{\alpha} \quad + \quad XR^{-1}Z\hat{\theta} &= XR^{-1}y \\ Z'R^{-1}X\hat{\alpha} + \left[ Z'R^{-1}Z + G^{-1} \right]\hat{\theta} &= Z'R^{-1}y \end{aligned} \qquad \text{(Ec. 3)}$$

También se demuestra que las soluciones en $\alpha$ de las ecuaciones normales (Ec. 3) son idénticas a las de mínimos cuadrados generalizados (EMCG) obtenidas de la ecuación:

$X'V^{-1}X\hat{\alpha} = X'V^{-1}y$ ; con $V = ZGZ' + R$ y $Var(y) = V\sigma_e^2$

con la ventaja computacional que en la ecuaciones normales (Ec. 3) no se requiere la inversa de $V$ . De las ecuaciones normales (Ec. 3), el mejor predictor lineal e insesgado para el vector aleatorio $\theta$ es:

$$\hat{\theta} = \left[ Z'R^{-1}Z + G^{-1} \right]^{-1} (Z'R^{-1}y - Z'R^{-1}X\hat{\alpha}) \qquad \text{(Ec. 4)}$$

En la mayoría de las aplicaciones los componentes de la varianza involucrados en $G$ y $R$ son desconocidos y deben ser estimados a partir de la información experimental. De acuerdo con Harville y Carriquiry (1992), el procedimiento de estimación se parte en dos etapas; en la primera se estiman los componentes de la varianza y en la segunda se utilizan estos estimadores y se incluyen en las ecuaciones normales (Ec. 3) y se procede a obtener el MPLI, llamado MPLI empírico. En este trabajo los componentes de varianza se obtienen por el método derivado del análisis de varianza, cuyas propiedades son descritas por Robinson (1991). Sin embargo, algunas veces ocurren estimaciones negativas en cuyo caso, de acuerdo con Robinson (1991), se pueden considerar iguales a cero, aunque esto provoca que los estimadores sean sesgados.

### Modelo lineal en experimentos dialélicos

Cuando los efectos maternos no existen, según Martínez (1983, 1988) un modelo lineal apropiado para analizar experimentos dialélicos, establecidos en diseño de bloques completos al azar, es:

$$y_{ijk} = \mu + g_i + g_j + s_{ij} + \delta_k + e_{ijk} \qquad \text{(Ec. 5)}$$

$$1 \leq i, j \leq p, \quad k = 1, 2 \ldots, r$$

donde $y_{ijk}$ es el valor fenotípico observado de la cruza $(i, j)$ en el bloque $k$ ; $\mu$ es el efecto común a todas las observaciones; $g_i$ es el efecto de aptitud combinatoria general del progenitor $i$ ; $s_{ij}$ es el efecto de aptitud combinatoria específica de la cruza $(i, j)$; $\delta_k$ es el efecto del bloque $k$ ; y $e_{ijk}$ es el efecto aleatorio del error correspondiente a la observación $(i, j, k)$. Los términos, $g_i$ , $s_{ij}$ y $e_{ijk}$ se consideran como variables aleatorias normales no correlacionadas entre y dentro de ellas, con media cero y varianzas $\sigma_g^2$ , $\sigma_s^2$ y $\sigma_e^2$ , respectivamente, con $s_{ij} = s_{ji}$ .

Sin pérdida de generalidad, si sólo se consideran las medias de las cruzas y se elimina el efecto de bloques, dado que éste es ortogonal con las cruzas, la representación en (Ec. 5) se reduce a:

$$\bar{y}_{ij.} = \mu + g_i + g_j + s_{ij} + \bar{e}_{ij.} \quad \text{(Ec. 6)}$$

## RESULTADOS

### Estimación de los MPLI de ACE

Cuando sólo se ensayan las cruzas directas, se trata del diseño cuatro de Grifiing, por lo que el número de tratamientos es $t = p(p-1)/2$. Cuando se ensayan tanto las cruzas directas como las autofecundaciones, se trata del diseño dos de Griffing, en donde $t = p(p+1)/2$. La derivación analítica de los MPLI de ACE es similar para ambos diseños. Para facilitar la derivación de los MPLI de ACE, se introduce la variable indicadora $q$, donde $q = 1$ para el diseño dos de Griffing, mientras que si $q = 0$ se tendrá el diseño cuatro de Griffing.

El modelo lineal apropiado para derivar los MPLI para estos diseños es el descrito en la Ecuación 6. En notación matricial se tiene:

$$\bar{y} = j\mu + Z_g g + Z_s s + \bar{e} \quad \text{(Ec. 7)}$$

donde $\bar{y} = (\bar{y}_{11.}, \bar{y}_{12.}, \bar{y}_{13.}, \cdots, \bar{y}_{t.})'$ es un vector de $t$ variables aleatorias observables $\bar{y}_{ij.}$; $\mu$ es la media general; $j$ es un vector $tx1$ de unos; $Z_g$ es la matriz diseño correspondiente a los efectos de progenitores de dimensión $txp$ cuya forma es igual a:

$$Z_g = \begin{bmatrix} 2 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ 1 & 0 & 0 & \cdots & 1 \\ 0 & 2 & 0 & \cdots & 0 \\ 0 & 1 & 1 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 1 & 0 & \cdots & 1 \\ & & \vdots & & \\ 0 & 0 & 0 & \cdots & 2 \end{bmatrix}$$

Cuando se trata del diseño dos de Griffing la matriz ***Zg*** contiene las autofecundaciones, por lo que aparecen hileras que tienen un número dos. Mientras que cuando se trata del diseño cuatro estas hileras se eliminan de la matriz ***Zg***, ***Zs*** es la matriz diseño correspondiente a la ACE, la cual es una matriz idéntica de dimensión $txt$; $g = (g_1, g_2, g_3, \cdots, g_p)'$ es un vector de $px1$ variables aleatorias; $s$ = ***(s11, s12, s13,…st)*** es un vector de ***tx1*** variables aleatorias y $\bar{e}$ es el vector de errores de dimensión ***tx1***.

Además,

$$E(\bar{y}) = j\mu,\ E(\bar{e}) = 0,\ E(g) = 0,\ E(s) = 0,\ Var(\bar{e}) = \frac{\sigma_e^2}{r} I_t,$$

$Var(g) = \sigma_g^2 I_p$, $Var(s) = \sigma_s^2 I_t$ y:

$$Var(\bar{y}) = E[(\bar{y} - j\mu)(\bar{y} - j\mu)']$$

$$= E[(Z_g g + Z_s s + \bar{e})(Z_g g + Z_s s + \bar{e})']$$

$$= \left[ Z_g \sigma_g^2 Z'_g + Z_s \sigma_s^2 Z'_s + \frac{\sigma_e^2}{r} I_t \right]$$

$$= \left[ Z_g G_g Z'_g + Z_s G_s Z'_s + I_t \right] \sigma_{e^*}^2 \quad \text{(Ec. 8)}$$

con $\sigma_{e^*}^2 = \frac{\sigma_e^2}{r}$, $G_g = \left( \frac{\sigma_g^2}{\sigma_{e^*}^2} \right) I_p$, $G_s = \left( \frac{\sigma_s^2}{\sigma_{e^*}^2} \right) I_t$ y $R = I_t$.

$I_t$ es una matriz identidad de dimensión $txt$. Con base en los supuestos descritos para el modelo (Ec. 7) y suponiendo adicionalmente que $g$, $s$ y $\bar{e}$ son vectores aleatorios no observables tales que

$$g \sim N_p(0, G_g \sigma_{e^*}^2),\ s \sim N_t(0, G_s \sigma_{e^*}^2),\ \bar{e} \sim N_t(0, I_t \sigma_{e^*}^2)$$

y que $G$ y $R$ son matrices no singulares, ocurre una extensión del método propuesto por Henderson (1963, 1973), lo que implica que la densidad conjunta de $\bar{y}$, $g$ y $s$ puede expresarse como:

$$f(\bar{y}, g, s) = d(\bar{y}|g, s) c_1(g) c_2(s) \quad \text{(Ec. 9)}$$

donde $c_1(g)$ es la densidad marginal de $g$, $c_2(s)$ es la densidad marginal de $s$ y $d(\bar{y}|g,s)$ es la distribución condicional de $\bar{y}$ dados $g$ y $s$.

En forma explícita la Ecuación 9 es igual a:

$$f(\bar{y},g,s) \propto \left\{ e^{\frac{-1}{2\sigma_{e^*}^2}(\bar{y}-j\mu-Z_g g-Z_s s)'(\bar{y}-j\mu-Z_g g-Z_s s)} \right\}$$

$$\left\{ e^{\frac{-1}{2\sigma_{e^*}^2} g'G_g^{-1} g} \right\}\left\{ e^{\frac{-1}{2\sigma_{e^*}^2} s'G_s^{-1} s} \right\} \quad \textit{(Ec. 10)}$$

En la búsqueda de los $\mu, g$ y $s$ que maximizan (Ec. 10) se encuentra el sistema de ecuaciones normales del modelo mixto:

$$\begin{aligned} &j'j\hat{\mu} + j'Z_g\hat{g} + j'Z_s\hat{s} = j'\bar{y} \\ &Z_g' j\hat{\mu} + \left[Z_g'Z_g + G_g^{-1}\right]\hat{g} + Z_g'Z_s\hat{s} = Z_g'\bar{y} \\ &Z_s' j\hat{\mu} + Z_s'Z_g\hat{g} + \left[Z_s'Z_s + G_s^{-1}\right]\hat{s} = Z_s'\bar{y} \end{aligned} \quad \text{(Ec. 11)}$$

Con las restricciones $\sum_{i \le j}^{p}\sum_{i=1}^{p}\hat{s}_{ij} = 0$ y $\sum_{i=1}^{p}\hat{g}_i = 0$ la solución del sistema de ecuaciones simultáneas en (Ec. 11), es:

$$\hat{\mu} = (j'j)^{-1}j'\bar{y} = \frac{2y_{...}}{rp(2q+p-1)} = \bar{y}_{...} \quad \text{(Ec. 12)}$$

y el MPLI de ACE es igual a:

$$\hat{s} = \left[Z_s'Z_s + G_s^{-1}\right]^{-1}(Z_s'\bar{y} - Z_s'j\bar{y}_{...} - Z_s'Z_g\hat{g}) \quad \text{(Ec. 13)}$$

En donde

$$\left[Z_s'Z_s + G_s^{-1}\right]^{-1} = \varphi I_t \quad \text{(Ec. 14)}$$

para $\varphi = 1/(1+\frac{\sigma_e^2}{r\sigma_s^2})$ con $\varphi \in (0,1)$ e $I_t$ es una matriz identidad de tamaño $txt$. Además,

$$(Z_s'\bar{y} - Z_s'j\bar{y}_{...}) = (\bar{y} - j\bar{y}_{...}) = \hat{v}_q = (\hat{v}_{1j}, \hat{v}_{2j}, \cdots, \hat{v}_{tj})' \quad \text{(Ec. 15)}$$

$\hat{v}_{ij} = (\bar{y}_{ij.} - \bar{y}_{...})$ para $i \le j$, $1 \le i,j \le p$ es el estimador mínimo cuadrático (EMC) del efecto de cruzas proporcionado por Martínez (1983), que coincide con el EMCG debido a que $R = I_t$. Por otra parte,

$$\hat{g} = k_1\hat{w}_q = k_1(\hat{w}_{q1}, \hat{w}_{q2}, \cdots, \hat{w}_{qp})' \quad \text{(Ec. 16)}$$

con

$$k_1 = \frac{1}{1 + \frac{1}{(4q+p-2)}\left[\frac{r\sigma_s^2 + \sigma_e^2}{\sigma_g^2}\right]} \in (0,1) \text{ y}$$

$$\hat{w}_{qi} = \frac{Q_i}{r(4q+p-2)} - \frac{2y_{...}}{rp(4q+p-2)}$$

donde $Q_i = (\sum_{j \ne 1}^{p} y_{ij.} + 2qy_{ij.})$ con $(y_{ij.} = y_{ji.})$. Aquí $\hat{w}_q$ es el EMC de ACG proporcionado por Martínez (1983), el cual coincide con el EMCG de ACG. De igual manera el subíndice $q$, también es la variable indicadora, donde si $q = 1$ se trata del diseño dos y si $q = 0$ se trata del diseño cuatro.

Finalmente sustituyendo las Ecuaciones 14, 15 y 16 en la Ecuación 13 se obtiene que:

$$\hat{s} = \varphi(\hat{v}_q - Z_g k_1 \hat{w}_q) \quad \text{(Ec. 17)}$$

Si se conocieran los componentes de varianza $\sigma_e^2, \sigma_s^2$ y $\sigma_g^2$, en (Ec. 17) se obtendría el MPLI de ACE, pero si no se conocen, según Harville y Carriquiry (1992), éstas pueden ser sustituidas por sus respectivos estimadores obteniéndose los MPLI empíricos. Martínez (1983) obtuvo los componentes de varianza en los diseños dos $(q = 1)$ y cuatro $(q = 0)$ de Griffing, basado en la técnica del análisis de varianza y en el supuesto de que todos los factores en el modelo (Ec. 5) son fijos:

$$\hat{\sigma}_e^2 = CME \quad \text{(Ec. 18)}$$

$$\hat{\sigma}_s^2 = (CM_{ACE} - CME)/r \quad \text{(Ec. 19)}$$

$$\hat{\sigma}_g^2 = (CM_{ACG} - CM_{ACE})/\left[r(4q+p-2)\right] \quad \text{(Ec. 20)}$$

donde $CME$ es el cuadrado medio del error; $CM_{ACE}$ es el cuadrado medio de la aptitud combinatoria específica y $CM_{ACG}$ es el cuadrado medio de aptitud combinatoria general. Sustituyendo las Ecuaciones 18, 19 y 20 en (Ec. 17 ), se obtienen los MPLI empíricos de ACE en los diseños dos y cuatro de Griffing.

*Cuadro 1. Datos ficticios para un dialélico con cuatro progenitores y cuatro repeticiones.*

| Cruza | Progenitor[¶] i | j | Dialelo | Bloques I | II | III | IV | $y_{ij.}$ | $\bar{y}_{ij.}$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 2 | 2 | 30.780 | 31.200 | 30.450 | 32.040 | 124.470 | 31.118 |
| 2 | 1 | 3 | 3 | 33.720 | 31.200 | 32.400 | 67.200 | 164.520 | 41.130 |
| 3 | 1 | 4 | 4 | 40.260 | 38.040 | 38.520 | 33.030 | 149.850 | 37.463 |
| 4 | 2 | 3 | 6 | 37.620 | 35.700 | 32.640 | 30.390 | 136.350 | 34.089 |
| 5 | 2 | 4 | 7 | 37.290 | 38.640 | 33.780 | 32.250 | 141.960 | 35.490 |
| 6 | 3 | 4 | 9 | 63.360 | 63.540 | 94.020 | 94.440 | 315.360 | 78.840 |

[¶] i y j representan los progenitores masculino y femenino; $y_{ij.}$ es la suma de cada una de las cruzas; $\bar{y}_{ij.}$ es el promedio de cada una de las cruzas.

## Ejemplo ilustrativo

En el Cuadro 1 se muestran datos de un dialélico de 4 progenitores y cuatro repeticiones; es decir, $p = 4$ y $r = 4$. En este caso $t = 4(4-1)/2 = 6$ cruzas directas y están alojadas en un diseño de bloques completos al azar. De estos datos se generará estimadores de los efectos de aptitud combinatoria general y de aptitud combinatoria específica.

## Método de análisis manual

De acuerdo con la Ecuación 17 se tiene que el MPLI de ACE es igual a: $\hat{s} = \varphi(\hat{v}_0 - Z_g k_1 \hat{w}_0)$. Para obtener una estimación del mejor predictor lineal e insesgado de aptitud combinatoria específica, se requiere calcular cada uno de los términos involucrados. Para calcular $\varphi$ se requiere $\sigma_e^2$ y $\sigma_s^2$. Dado que no los conocemos se realizó una estimación de éstos a partir de sus correspondientes estimadores (Ecuaciones 18 y 19), con los datos del Cuadro 1: $\hat{\sigma}_e^2 = 111.7695$ y $\hat{\sigma}_s^2 = 189.7434$, y $\hat{\sigma}_g^2 = 84.5435$. Por tanto,

$$\varphi = \frac{1}{1 + \hat{\sigma}_e^2 / r\hat{\sigma}_s^2} = \frac{1}{1 + 111.7695 /((4)(189.7434))} = \frac{1}{1.1473}$$

Cálculo de $\hat{v}_0$:

$$\hat{v}_0 = \begin{bmatrix} 31.118 & - & 43.0212 \\ 41.130 & - & 43.0212 \\ 37.463 & - & 43.0212 \\ 34.088 & - & 43.0212 \\ 35.490 & - & 43.0212 \\ 78.840 & - & 43.0212 \end{bmatrix} = \begin{bmatrix} -11.9038 \\ -1.8913 \\ -5.5588 \\ -8.9338 \\ -7.5313 \\ 35.8188 \end{bmatrix}$$

Cálculo de $Z_g k_1 \hat{w}_0$: como $k_1 \hat{w}_0 = \hat{g}$. Para los promedios de las cruzas:

$$\begin{bmatrix} \hat{g}_1 \\ \hat{g}_2 \\ \hat{g}_3 \\ \hat{g}_4 \end{bmatrix} = \begin{bmatrix} -4.2305 \\ -6.2010 \\ 5.4633 \\ 4.9682 \end{bmatrix} \text{ y } Z_g = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix},$$

$$\text{así } Z_g \hat{g} = \begin{bmatrix} -10.4315 \\ 1.2328 \\ 0.7377 \\ -0.7377 \\ -1.2328 \\ 10.4315 \end{bmatrix}$$

Finalmente, como $\hat{s} = \varphi(\hat{v}_0 - Z_g k_1 \hat{w}_0) = \varphi(\hat{v}_0 - Z_g \hat{g})$ entonces:

$$\begin{bmatrix} \hat{s}_{12} \\ \hat{s}_{13} \\ \hat{s}_{14} \\ \hat{s}_{23} \\ \hat{s}_{24} \\ \hat{s}_{34} \end{bmatrix} = \frac{1}{1.1472} \begin{bmatrix} -11.9038 & + & 10.4315 \\ -1.8913 & - & 1.2328 \\ -5.5588 & - & 0.7377 \\ -8.9338 & + & 0.7377 \\ -7.5313 & + & 1.2328 \\ 35.8188 & - & 10.4315 \end{bmatrix} = \frac{1}{1.1472} \begin{bmatrix} -1.4722 \\ -3.1241 \\ -6.2965 \\ -8.1960 \\ -6.2984 \\ 25.3872 \end{bmatrix} = \begin{bmatrix} -1.2832 \\ -2.7231 \\ -5.4883 \\ -7.1439 \\ -5.4899 \\ 22.128 \end{bmatrix}$$

## Método de análisis con SAS

Los MPLI empíricos de ACE se pueden obtener con el algoritmo computacional en IML de SAS que se describe a continuación:

```
OPTIONS PS=60 PAGENO=1 NODATE NONUMBER;
DATA MASTACHE;
INPUT CRUZA  I  J  DIALELO REP Y1;
CARDS;
```

| | | | | | |
|---|---|---|---|---|---|
| 1 | 1 | 2 | 1 | 1 | 30.780 |
| 1 | 1 | 2 | 1 | 2 | 31.200 |
| 1 | 1 | 2 | 1 | 3 | 30.450 |
| 1 | 1 | 2 | 1 | 4 | 32.040 |
| 2 | 1 | 3 | 2 | 1 | 33.720 |
| 2 | 1 | 3 | 2 | 2 | 31.200 |
| 2 | 1 | 3 | 2 | 3 | 32.400 |
| 2 | 1 | 3 | 2 | 4 | 67.200 |
| 3 | 1 | 4 | 3 | 1 | 40.260 |
| 3 | 1 | 4 | 3 | 2 | 38.040 |
| 3 | 1 | 4 | 3 | 3 | 38.520 |
| 3 | 1 | 4 | 3 | 4 | 33.030 |
| 4 | 2 | 3 | 4 | 1 | 37.620 |
| 4 | 2 | 3 | 4 | 2 | 35.700 |
| 4 | 2 | 3 | 4 | 3 | 32.640 |
| 4 | 2 | 3 | 4 | 4 | 30.390 |
| 5 | 2 | 4 | 5 | 1 | 37.290 |
| 5 | 2 | 4 | 5 | 2 | 38.640 |
| 5 | 2 | 4 | 5 | 3 | 33.780 |
| 5 | 2 | 4 | 5 | 4 | 32.250 |
| 6 | 3 | 4 | 6 | 1 | 63.360 |
| 6 | 3 | 4 | 6 | 2 | 63.540 |
| 6 | 3 | 4 | 6 | 3 | 94.020 |
| 6 | 3 | 4 | 6 | 4 | 94.440 |

```
TITLE " LOS MPLI EMPÍRICOS EN LOS DISEÑOS 2 Y 4 DE GRIFFING ";
PROC IML;SORT MASTACHE OUT=NUEVO BY CRUZA; USE NUEVO; READ ALL INTO MATRIZ; CRUZA=MATRIZ[,1];
A=MATRIZ[,2]; B=MATRIZ[,3]; REP=MATRIZ[,5];
N=NROW(MATRIZ); UNO=J(N,1,1); CERO=J(N,1,0);
MDIS=DESIGN(CRUZA); X=UNO||MDIS;
XX=X`*X;XXIG=GINV(XX); M=X*XXIG*X`;
BLOQ=DESIGN(REP);
W=X||BLOQ;WW=W`*W;WINV=GINV(WW);WWW=W*WINV*W`;IDEN=I(N);Z=MDIS;ZZ=Z*Z`;FIJO=UNO||BLOQ;
A0=DESIGN(A);B0=DESIGN(B);T=NCOL(MDIS);R=MAX(REP);P=MAX(B); NC=NCOL(MATRIZ); IF ANY (A=B) THEN Q=1;ELSE Q=0; IF Q=1 THEN PRINT "DISEÑO 2 DE GRIFFING";ELSE PRINT "DISEÑO 4 DE GRIFFING";IF Q=1 THEN AB=A0+B0;ELSE
AB=(A0||CERO)+(CERO||B0);X0=UNO||AB;X0X0=X0`*X0;X0IG=GINV(X0X0);M0=X0*X0IG*X0`;
Z=MDIS;ZZ=Z*Z`;FIJO=UNO||BLOQ;Zp=AB; ZpZp=Zp`*Zp;
PRINT  N    T    R    P;
TITLE " ANÁLISIS DE VARIANZA ";
FV=J(6, 5, .); UN=J(P,1,1);PROG=J(P, 4, .);PPP=J(P,1,.);DO LLL=1 TO P BY 1;PPP[LLL,1]=LLL;END;DO F=6 TO NC  BY 1;VARIABLE= F-5; Y=MATRIZ[,F];FC=(UNO`*Y)**2/N; MEDIA=UNO`*Y/N;SCTOT=Y`*Y-FC;
SCE=Y`*(IDEN-WWW)*Y; CME=SCE/((R-1)*(T-1));
CV=(CME**.5)*100/MEDIA; SCB=Y`*(WWW-M)*Y;
CMB=SCB/(R-1); FBLOQ=CMB/CME; SCCRUZA=(Y`*M*Y)-FC;
CMCRUZA=SCCRUZA/(T-1); FCRUZA=CMCRUZA/CME;
SCACG=(Y`*M0*Y)-FC; CMACG=SCACG/(P-1); SCACE=SCCRUZA-SCACG; CMACE=SCACE/(T-P);FACG=CMACG/CMACE;
FACE=CMACE/CME;GLBLOQ=(R-1); GLCRUZA=(T-1);
GLACG=(P-1); GLACE=(T-P);GLE=(R-1)*(T-1);PROBBLOQ= 1-PROBF(FBLOQ,GLBLOQ,GLE);PROBCRUZA= 1-PROBF(FCRUZA,GLCRUZA,GLE);PROBACG= 1-PROBF(FACG,GLACG,GLACE);
PROBACE=1-PROBF(FACE,GLACE,GLE);FV[1,1]=R-1;FV[2,1]=T-1;FV[3,1]=P-1;FV[4,1]=T-P;
FV[5,1]=(R-1)*(T-1);FV[6,1]=T*R-1;
FV[1,2]=SCB;FV[2,2]=SCCRUZA;FV[3,2]=SCACG;
FV[4,2]=SCACE;
FV[5,2]=SCE;FV[6,2]=SCTOT;FV[1,3]=CMB;FV[2,3]=CMCRUZA;FV[3,3]=CMACG;FV[4,3]=CMACE;FV[5,3]=CME;
FV[1,4] = FBLOQ;FV[2,4]=FCRUZA;FV[3,4]=FACG;
FV[4,4]=FACE; FV[1,5]=PROBBLOQ; FV[2,5]=PROBCRUZA;
FV[3,5]=PROBACG;FV[4,5]=PROBACE; CCC={"GL" "SC" "CM" "F" "Pr > F"};
DDD={"BLOQUES" "CRUZAS" "  ACG" "  ACE" "ERROR" "TOTAL"};
TITLE " ESTIMACIÓN DE LOS COMPONENTES DE VARIANZA ";
VARe=CME;VARs=(CMACE-CME)/R;IF VARs>0 THEN VARs=VARs;ELSE VARs=0;
VARg=(CMACG-CMACE)/(R*(4*Q+P-2));IF VARg>0 THEN VARg=VARg;ELSE VARg=0;
TITLE " EMC Y EL MPLI DE Gi ";
MED=MEDIA*UN; TETA=GINV(FIJO`*FIJO)*FIJO`*Y;FIXTETA=FIJO*TETA;ZpY=Zp`*Y;ZpFIXTTA=Zp`*FIXTETA;
EMCg=(ZpY-ZpFIXTTA)/(R*(4*Q+P-2));EMCGg=EMCg;IF VARg > 0 THEN K=(4*Q+P-2)/((4*Q+P-2)+(R*VARs+VARe)/(R*VARg));ELSE
K=1;MPLIg=K*EMCg;MPLIMED=MPLIg+MED;PROG[,1]=EMCg;PROG[,2]=EMCGg;
PROG[,3]=MPLIg;PROG[,4]=MPLIMED;EEE={"EMC" "EMCG" "MPLI" "MPLI+MEDIA"};FFF=CHAR(PPP,3,0);
TITLE " EMC Y MPLIs DE ACE Y DE LOS EFECTO DE CRUZAS";
BLUPSACE=J(T, 4, .); YYY=J(T,1,.); BLUPSCRUZAS=BLUPSACE; DO BBB=1 TO T BY 1; YYY[BBB,1]=BBB;END;
UNN=J(T,1,1);IF VARs>0 THEN IGs=(VARe/VARs)*I(T); ELSE IGs=0*I(T); Zs=Z;Zg=Zp;IF VARg>0 THEN
Gg=(VARg/VARe)*I(P); ELSE Gg=0*(P);Rg=Zg*Gg*Zg`+I(N);
IRg=GINV(Rg);MEDD=MEDIA*UNN;
EMCc=GINV(Z`*Z)*Z`*(Y-UNO*MEDIA); EMCGc=EMCc;IF VARs>0 THEN KK=R*VARs/(VARe+R*VARs); ELSE KK=1;
MPLIs=KK*(EMCc-(Z`*Zg/R)*MPLIg);EMCs=EMCc-(Z`*Zg/R)*EMCg; EMCGs=EMCs;MPLIMEDD=MPLIs+MEDD;
FFF1=CHAR(YYY,3,0); BLUPSACE[,1]=EMCs;BLUPSACE[,2]=EMCGs;BLUPSACE[,3]=MPLIs;
BLUPSACE[,4]=MPLIMEDD;
MPLIc=(Z`*Zg/R)*MPLIg+MPLIs; MPLICMEDD=MPLIc+MEDD;
        BLUPSCRUZAS[,1]=EMCc;BLUPSCRUZAS[,2]=EMCGc;
BLUPSCRUZAS[,3]=MPLIc;BLUPSCRUZAS[,4]=MPLICMEDD;
TITLE "LA MATRIZ DE COEFICIENTES: C DE ACG ";
IF VARg>0 THEN INVGp=(VARe/VARg)*I(P);ELSE INVGp=0*I(P);GRR=GINV((VARs/VARe)*ZZ+I(N));
CC1=(UNO`*GRR*UNO)||(UNO`*GRR*Zp);CC2=(UNO`*GRR*Zp)`||((Zp`*GRR*Zp)+INVGp);CC3=CC1`||CC2`;
CCCACG=GINV(CC3);
TITLE "LA MATRIZ DE COEFICIENTES: Cs DE ACE ";
CCS1=(UNO`*IRg*UNO)||(UNO`*IRg*Zs);CCS2=(UNO`*IRg*Zs)`||((Zs`*IRg*Zs)+IGs);CCS3=CCS1`||CCS2`;
CCCACE=GINV(CCS3);
TITLE " IMPRESIÓN DE RESULTADOS ";
PRINT VARIABLE;PRINT "CUADRO 1. ANALISIS DE VARIANZA";PRINT FV[ROWNAME=DDD  COLNAME=CCC];
PRINT MEDIA[FORMAT=12.5]   CV[FORMAT=12.5];PRINT ,;PRINT "ESTIMACION DE LAS COMPONENTES DE VARIANZA";PRINT VARe[FORMAT=12.5]  VARs[FORMAT=12.5] VARg[FORMAT=12.5];PRINT ,;PRINT "CUADRO 2. ESTIMACION Y PREDICCION DE ACG";PRINT PROG[ROWNAME=FFF COLNAME=EEE FORMAT=12.5];PRINT ,;PRINT
```

K[FORMAT=**12.5**];PRINT /;PRINT "CUADRO 3. ESTIMACION Y PREDICCION DE ACE";PRINT BLUPSACE[ROWNAME=FFF1 COLNAME=EEE FORMAT=**12.5**];PRINT ,;PRINT "CUADRO 4. ESTIMACION Y PREDICCION DEL EFECTO DE CRUZAS";PRINT BLUPSCRUZAS[ROWNAME=FFF1 COLNAME=EEE FORMAT=**12.5**];PRINT /;PRINT " LA MATRIZ DE COEFICIENTES: C DE ACG ";PRINT CCCACG[FORMAT=**7.3**];PRINT ,;PRINT " LA MATRIZ DE COEFICIENTES: C DE ACE ";PRINT CCCACE[FORMAT=**5.3**];PRINT /;END;**QUIT**;

*Cuadro 2. Análisis de varianza y estimación de componentes de varianza generados con el algoritmo en IML de SAS en el diseño 4 de Griffing.*

| FV | GL | SC | CM | F | Pr>F |
|---|---|---|---|---|---|
| Bloques | 3 | 268.11 | 89.37 | 0.80 | 0.51 |
| Cruzas | 5 | 6382.76 | 1276.55 | 11.42 | 0.0001 |
| ACG | 3 | 4641.27 | 1547.09 | 1.77 | 0.38 |
| ACE | 2 | 1741.48 | 870.74 | 7.79 | 0.005 |
| Error | 15 | 1676.54 | 111.77 | . | . |
| Total | 23 | 8327.41 | | | |

La información completa proveniente de un experimento en que se evaluó las cruzas dialélicas según un diseño dos o cuatro, establecido en bloques completos al azar, debe organizarse como se muestra para el diseño cuatro en el algoritmo IML de SAS, para los datos del Cuadro 1. Además, debe respetarse el nombre del archivo y el orden de las variables especificadas en el comando INPUT, ya que el programa hace uso de ellas. En general, se tendrán $t$ cruzas, donde **I** se refiere al progenitor femenino y **J** al masculino con **I=i, J=j,** $1 \le i, j \le p,$ en DIALELO **ij=ji**, REP se refiere a las repeticiones $r$; $\mathbf{Y_1}$ es la variable respuesta. Finalmente, el programa SAS-IML, produce los resultados que se presentan en los Cuadros 2 y 3, con lo cual se corrobora la metodología descrita anteriormente.

## DISCUSIÓN

En el Cuadro 3 se presenta la estimación y predicción de ACE por mínimos cuadrados, mínimos cuadrados generalizados y bajo el modelo de efectos mixtos (MPLI). Aunque estas estimaciones son muy similares las estimaciones basadas en el modelo de efectos mixtos, tienen menor varianza que las estimaciones basadas en el modelo de efectos fijos. Lo anterior obedece a la manera en que se obtiene el MPLI de ACE, equivalente a la diferencia entre una constante, que toma valores en el intervalo (0,1), multiplicada por el estimador de mínimos cuadrados generalizados de efectos de cruzas ($\varphi\hat{v}_q$) menos otra constante que también toma valores en el intervalo (0,1), multiplicada por el estimador de mínimos cuadrados generalizados de ACG ($\varphi k_1\hat{w}_q$). De acuerdo con Henderson (1963, 1973), el MPLI de ACE difiere de los estimadores basados en el modelo de efectos fijos por la presencia de la matriz $G_S^{-1}$ en la Ecuación 13, la cual involucra los componentes de varianza y afecta a la matriz $Z'_S Z_S$, que implica que el MPLI de ACE tenga menor varianza que los obtenidos del modelo de efectos fijos. Por otra parte, Martínez (1983) demostró que los estimadores de mínimos cuadrados generalizados de los efectos de cruzas y de ACG, $\hat{v}_q$ y $\hat{w}_q$, respectivamente, son insesgados. Propiedad también presentada por los MPLI de ACE $(\hat{s})$, puesto que:

$$E(\hat{s}) = E(\varphi\hat{v}_q - Z_g\varphi k_1\hat{w}_q) = \varphi E(\hat{v}_q) - Z_g\varphi k_1 E(\hat{w}_q) = 0$$

Por último, esta técnica de estimación desarrollada por Henderson (1963, 1973) para obtener los MPLI tiene mayor precisión que la de mínimos cuadrados generalizados, aún cuando se obtienen los MPLI empíricos; es decir, cuando se sustituye a los componentes de varianza involucrados por sus respectivos estimadores.

## CONCLUSIONES

En este trabajo se derivan los MPLI de ACE, sobre la base del modelo de efectos mixtos, para los diseños dos $(q=1)$ y cuatro $(q=0)$ de Griffing, los cuales se obtienen a partir de la siguiente expresión:

$$\hat{s} = \varphi(\hat{v}_q - Z_g k_1\hat{w}_q)$$

En donde $\hat{v}_q$ y $\hat{w}_q$ son vectores de los estimadores de mínimos cuadrados generalizados de los efectos de cruzas y de ACG, de dimensiones $tx1$ y $px1$, respectivamente, $Z_g$ es la matriz diseño correspondiente a los efectos de progenitores de dimensión $txp$ y con $\varphi$ y $k_1$ según las Ecuaciones 14 y 16, respectivamente.

*Cuadro 3. Estimación y predicción de aptitud combinatoria específica para los datos del Cuadro 1. El análisis se realizó con el algoritmo en IML de SAS†.*

| Cruza | EMC | EMCG | MPLI | MPLI+MEDIA |
|---|---|---|---|---|
| 1 | 11.95750 | 11.95750 | -1.28325 | 41.73800 |
| 2 | -4.71125 | -4.71125 | -2.72307 | 40.29818 |
| 3 | -7.24625 | -7.24625 | -5.48826 | 37.53299 |
| 4 | -7.24625 | -7.24625 | -7.14397 | 35.87728 |
| 5 | -4.71125 | -4.71125 | -5.48995 | 37.53130 |
| 6 | 11.95750 | 11.95750 | 22.12850 | 65.14975 |

†.SAS son las siglas del software Statistical Analysis System; EMC es el estimador de mínimos cuadrados de ACE; EMCG es el estimador de mínimos cuadrados generalizados de ACE; MPLI es el mejor predictor lineal e insesgado de ACE; MPLI+MEDIA es el mejor predictor lineal e insesgado de ACE más el promedio de todas las cruzas.

Cuando no se conocen los componentes de varianza involucrados en $\phi$ y $k_1$, éstos se sustituyen por sus respectivos estimadores para obtener los MPLI empíricos de ACE, que también tienen mejor precisión que los estimadores de mínimos cuadrados generalizados. Por otro lado, el algoritmo computacional proporciona de una forma rápida y compacta un análisis completo de estimación y predicción en los diseños dos y cuatro de Griffing, que considera aleatorios a los efectos de aptitud combinatoria general y específica, con lo que se obtiene un panorama completo de estimación y predicción.

## BIBLIOGRAFÍA

**Griffing B (1956a)** A generalized treatment of the use of dialell crosses in quantitative inheritance. Heredity 10:31-50.

**Griffing B (1956b)** Concept of general and specific combining ability in relation to dialell crossing systems. Aust. J. Biol. Sci. 9:463-491.

**Harville D A (1976)** Extension of the Gauss-Markov theorem to include the estimation of random effects. Ann. Statist. 4:384-395.

**Harville D A, A L Carriquiry (1992)** Classical and Bayesian prediction as applied to an unbalanced mixed linear model. Biometrics 48:987-1003.

**Henderson C R (1963)** Selection index and expected genetic advance. *In*: Statistical Genetics and Plant Breeding. W D Hanson, H F Robinson (eds). Nat. Acad. Sci., Nat. Res. Council, Publication 982, Washington, DC. pp:141-163.

**Henderson C R (1973)** Sire evaluation and genetics trends. Proc. Anim. Breed. Genet. Symp. in Honor of Dr. Jay L. Lush. Am. Soc. Anim. Sci., Champaign, ILL. pp:10-41.

**Martínez G A (1983)** Diseño y Análisis de los Experimentos de Cruzas Dialélicas. Centro de Estadística y Cálculo, Colegio de Postgraduados, Chapingo, México. 252 p.

**Martínez G A (1988)** Diseños Experimentales. Editorial Trillas. 1a. Edición. México. 756 p.

**Mastache L A A, A Martínez, A Castillo (1999)** Los mejores predictores lineales e insesgados (MPLI) en los diseños dos y cuatro de Griffing. Agrociencia 33:81-90.

**Robinson G K (1991)** That BLUP is a good thing: The estimation of random effects. Stat. Sci. 6(1):15-51.

**SAS Institute Inc. (1989)** SAS/IML Software: Usage and Reference, Version 6, First Edition. Cary, N. C. 501 p.