

VARIANCE COMPONENTS OF CORN HYBRIDS EVALUATED IN THE USA CORN BELT

COMPONENTES DE VARIANZA DE HÍBRIDOS DE MAÍZ EVALUADOS EN LA FAJA MAICERA DE LOS ESTADOS UNIDOS

Bulmaro Coutiño-Estrada¹, and Victor A. Vidal-Martínez²

¹Campo Experimental Centro de Chiapas. Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias (INIFAP). Km 3, Carretera Ocozocoautla-Cintalapa. Ocozocoautla, Chiapas, México. (coutino.bulmaro@inifap.gob.mx). ²Campo Experimental Santiago Ixcuintla. Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias (INIFAP). Nayarit, México.

ABSTRACT

In the final steps of a corn (*Zea mays* L.) breeding program, the test of the superior crosses is a standard practice involving yield trials in many test environments. The objectives of this research were to compare both the magnitude of the standard error of prediction of three different linear mixed models where location and hybrid effects were considered random or fixed effects and the magnitude of the phenotypic variance. Mixed, GLM, and Varcomp procedures from SAS were applied to grain yield and moisture data obtained from 40 hybrids evaluated at 22 locations within the U.S. Corn Belt. Results showed that when using a linear model which considers both location and hybrid effects as random effects, rather than fixed effects, standard errors of prediction were 33% smaller and differences could be detected with more accuracy. The three SAS procedures computed the same estimates of variance components using balanced data, but with unbalanced data, only Mixed and Varcomp procedures computed the same estimates. Phenotypic variance of grain yield was 52% smaller in Iowa environments, 9% smaller in Indiana environments, and 54% bigger in Illinois locations, compared to Nebraska environments where the hybrids were made.

Key words: *Zea mays* L., corn hybrid, grain moisture, grain yield, variance components.

INTRODUCTION

Single-cross hybrids are the main type of commercial corn hybrids grown in the USA Corn Belt since the 1960's (Hallauer, 1990; Smith *et al.*, 1996) and pedigree selection is the most widely used breeding method for development of inbred lines used in these hybrids (Hallauer *et al.*, 1988). A survey of 130 USA Corn Belt corn breeders indicated that the topcross test has become a standard practice in all corn breeding programs (Bauman, 1981). Usually, commercial corn breeders generate hybrids of different maturities for different environments; in advanced phases of their

RESUMEN

En las etapas finales de un programa de hibridación de maíz (*Zea mays* L.) la evaluación de las cruas superiores es una práctica común que involucra ensayos de rendimiento en muchos ambientes de prueba. Los objetivos de esta investigación fueron comparar la magnitud de los errores estándar de predicción obtenidos de tres modelos estadísticos lineales mixtos en los que los efectos de localidades y de híbridos fueron considerados aleatorios o fijos, y conocer la magnitud de la variación fenotípica. Se aplicaron los procedimientos Mixed, GLM y Varcomp del SAS, a datos de rendimiento y humedad de grano de 40 híbridos evaluados en 22 ambientes de la Faja Maicera de los Estados Unidos. Los resultados mostraron que con el modelo lineal que considera a las localidades y a los híbridos como efectos aleatorios, en lugar de efectos fijos, los errores estándar de predicción fueron 33% más pequeños, lo que permite detectar diferencias con mayor precisión. Los tres procedimientos del SAS produjeron las mismas estimaciones de los componentes de varianza usando datos balanceados, pero con datos desbalanceados sólo Mixed y Varcomp dieron los mismos resultados; la varianza fenotípica del rendimiento de grano fue 52% más pequeña en las localidades de Iowa, 9% menor en las de Indiana y 54% más grande en las de Illinois, que la obtenida en ambientes de Nebraska, donde se formaron los híbridos.

Palabras clave: *Zea mays* L., híbrido de maíz, humedad de grano, rendimiento de grano, componentes de varianza.

INTRODUCCIÓN

Los híbridos de cruza simple son el principal tipo comercial de híbridos de maíz cultivados en la Faja Maicera de los Estados Unidos desde los años sesenta (Hallauer, 1990; Smith *et al.*, 1996) y la selección por pedigree es el método de mejoramiento más utilizado para el desarrollo de líneas endogámicas que se usan en estos híbridos (Hallauer *et al.*, 1988). Una encuesta aplicada a 130 mejoradores de maíz de la Faja Maicera de los Estados Unidos indicó que la prueba de cruas superiores se ha convertido en una práctica estandar en todos los programas de mejoramiento de maíz

breeding programs they evaluate testcrosses in multilocation yield trials in order to identify superior hybrids for each region of their market.

In multilocation testcross yield trials, most breeders assume that fixed effects or variance components are of primary interest (Harville, 1976). When analyzing data collected from yield trials, the main feature is the extent to which different levels of a factor affect the variable of interest. Factor effects (as genotypes, environments, fertilizers, etc.) are either fixed or random, depending on how the different levels are selected.

The GLM procedure of SAS is particularly suited for analysis and interpretation of unbalanced and messy data where effects are fixed, whereas the Mixed procedure allows models that have both random and fixed effects. Sanders (1989) compared six mixed models for analyzing soybean (*Glycine max*) yield data with 32 entries in an incomplete block design with six entries per block, three replications, and two locations; his objectives were to rank entries for each location and over locations. He found that the analysis for his model 6, which considered entries as random and locations as fixed effects, had smaller standard errors in the combined analysis, more flexibility to match predictions with the desired inference space, and predictions for entries over both locations.

Bridges (1989) used data from four cucumber (*Cucurbita moschata*) cultivars grown at two locations, arranged in a Latin Square design to estimate means of cultivar by location combinations and the overall cultivar and location means. Comparison of results from the standard GLM with the Mixed model showed that Henderson's mixed model equations and solutions could easily handle problems with additional random effects, unbalanced data, and heterogeneous variances. A statistical advantage of mixed procedure is that a smaller standard error is found if genotypes are defined as random rather than fixed effects, which permits more accuracy in determining differences among genotypes.

The objectives of this research were: 1) to compare the magnitude of standard errors of prediction from three different mixed linear models considering test location and genotype effects as fixed and as random; 2) to compare three statistical procedures for estimation of variance components; 3) to determine changes in variance component estimates among groups of environments, and in the combined analysis, when all factors are considered as random effects.

MATERIALS AND METHODS

Genetic material

Data for corn grain yield and moisture grain content were obtained from advanced yield trials conducted by the Corn Breeding

(Bauman, 1981). Generalmente, los mejoradores de maíz forman híbridos de diferente maduración para diferentes ambientes; en las fases avanzadas de sus programas de mejoramiento evalúan cruza experimentales en ensayos de rendimiento, en muchas localidades, para identificar los híbridos superiores de cada región de su área de mercado.

En estos ensayos de rendimiento, la mayoría de los mejoradores suponen que el principal interés son los efectos fijos o los componentes de varianza (Harville, 1976). Cuando se analizan los datos recolectados en ensayos de rendimiento, el aspecto más importante es cuánto afectan los niveles de un factor a la variable de interés. Los efectos de un factor (como genotipos, ambientes, fertilizantes, etc.) pueden ser fijos o aleatorios, dependiendo de cómo se seleccionan los diferentes niveles.

El Modelo Lineal General (GLM) del Sistema de Análisis Estadístico (SAS) es un procedimiento muy adecuado para el análisis e interpretación de datos desbalanceados y desordenados donde los efectos son fijos, mientras que el procedimiento mixto admite modelos que tienen ambos efectos, fijos y aleatorios. Sanders (1989) comparó seis modelos mixtos para analizar datos de rendimiento de soya (*Glycine max*), con 32 entradas, en un diseño de bloques incompletos con seis entradas por bloque, tres repeticiones y dos localidades; sus objetivos fueron ordenar la mejor entrada para cada localidad y la mejor para todas ellas. Encontró que el análisis para su modelo seis, que consideró aleatorios los efectos de las entradas y fijos los de las localidades, tuvo errores estándar más pequeños en el análisis combinado, más flexibilidad para coincidir sus predicciones con el espacio de inferencia deseado y las predicciones para las entradas sobre ambas localidades.

Bridges (1989) usó datos de cuatro cultivares de calabaza (*Cucurbita moschata*) cultivadas en dos localidades, arregladas en un diseño cuadro latino para estimar medias de combinaciones de cultivar por localidad y medias de los cultivares y de localidades. La comparación de los resultados del modelo estándar GLM con el modelo Mixto mostraron que las ecuaciones y soluciones del modelo mixto de Henderson resolvieron con facilidad los problemas con efectos aleatorios adicionales, datos desbalanceados y varianzas heterogéneas. Una ventaja estadística del procedimiento mixto es que se obtiene un error estándar más pequeño si los genotipos se definen como efectos aleatorios en lugar de fijos, lo que permite determinar con mayor exactitud las diferencias entre genotipos.

Los objetivos de esta investigación fueron: 1) comparar la magnitud del error estándar de predicción cuando se usan tres diferentes modelos lineales mixtos considerando alternativamente como fijos y aleatorios los efectos de prueba de localidad y los de genotipos; 2)

Program of Mycogen Seeds, at York, NE, USA, In the last stage of its hybrid testing program, testcrosses plus commercial hybrid checks are evaluated at 30 to 40 locations across many environments of the USA Corn Belt in order to identify the future hybrid to be sold in each region. In this study, 39 testcrosses from some hybrid combinations among 32 female inbred lines and 16 male inbred lines (testers) were used.

Field performance trials

The 39 testcrosses, plus one commercial check hybrid, were evaluated during the 1996 growing season at 22 locations of the Corn Belt: York (irrigated), York (dryland), Arlington, Holdrege, NE, and Rossville, KS (grouped as Nebraska locations); Huxley, Earlham, Essex, Fremont, Griswold, and Wyman, IA (grouped as Iowa locations); Lincoln, Tuscola, Macom, Woodhull, IL, and St. Charles, MO (grouped as Illinois locations); Vincennes, Evansville, Franklin, Kentland, IN, Clarksville, OH, and Laurel, DE (grouped as Indiana locations). A randomized complete block design with two replications was used in all field experiments. The hybrids were grown in two-row plots, each row 5 m long and spaced 0.76 m apart, at a plant density of 73 700 plant ha⁻¹. The plots were machine-harvested and grain yield was recorded in bushels per acre, and adjusted to 155 g H₂O kg⁻¹. Grain yield in t ha⁻¹ and grain moisture content in g H₂O kg⁻¹ were analyzed.

Statistical analyses

Based on analyses commonly used in multilocation yield trials, three mixed linear models were used for hybrids: Mixed model 1: both location and hybrid effects fixed; Mixed model 2: location effects as random and hybrid effects as fixed; Mixed model 3: both location and hybrid effects as random. Also, three mixed models were used to analyze female and male parents within hybrid combinations: Mixed model 1: location and male effects as random female effects as fixed; Mixed model 2: location and female effects as random, male effects as fixed; Mixed model 3: all effects considered as random. A small standard error of prediction from these linear models would indicate more statistical efficiency.

As stated earlier, individual locations were grouped by state and analyzed together. The data of Iowa, Illinois, and Indiana groups were combined together to Nebraska group for analysis until all four groups (all 22 locations) were included. The Mixed, GLM, and Varcomp procedures of SAS (version 6.11; Littell *et al.*, 1996) were used to estimate and compare the variance components. From GLM, the random statement and the test option were used to test the hypothesis that the variance components were zero.

RESULTS AND DISCUSSION

Comparison of different mixed linear models

The three mixed models produced different standard errors of prediction (SEP) for each group of locations and for the combined group analyses for the testcross yield

comparar tres procedimientos estadísticos para estimar componentes de varianza; 3) determinar los cambios en las estimaciones de los componentes de varianza entre grupos de ambientes, y en el análisis combinado, cuando todos los factores se consideran aleatorios.

MATERIALES Y MÉTODOS

Material Genético

Se obtuvieron datos de rendimiento y contenido de humedad de grano de maíz de los ensayos avanzados de rendimiento del Programa de Mejoramiento de Maíz de la empresa Mycogen Seeds de York, Nebraska, EE.UU. En la última etapa del programa de evaluación de híbridos, los híbridos experimentales, más híbridos comerciales testigos, son evaluados en 30 a 40 localidades a través de muchos ambientes de la Faja Maicera de los Estados Unidos, para identificar el híbrido a vender en cada región. En este estudio se usaron 39 híbridos experimentales de algunas combinaciones entre 32 líneas endogámicas progenitoras femeninas y 16 líneas endogámicas masculinas (probadoras).

Ensayos de rendimiento

Los 39 híbridos experimentales, más un híbrido comercial testigo, se evaluaron durante 1996 en 22 localidades de la Faja Maicera: York (con riego), York (temporal), Arlington y Holdrege del estado de Nebraska y Rossville en Kansas (agrupados como localidades de Nebraska); Huxley, Earlham, Essex, Fremont, Griswold y Wyman de Iowa (agrupados como localidades de Iowa); Lincoln, Tuscola, Macom y Woodhull de Illinois y St. Charles de Missouri (agrupados como localidades de Illinois); Vincennes, Evansville, Franklin y Kentland de Indiana, Clarksville de Ohio y Laurel de Delaware (agrupados como localidades de Indiana). En todos los experimentos de campo se utilizó un diseño de bloques completos al azar, con dos repeticiones. Los híbridos se sembraron en parcelas de dos surcos, de 5 m de longitud, espaciados 0.76 m, con una densidad de población de 73 700 plantas ha⁻¹. Las parcelas fueron cosechadas con maquinaria y el rendimiento de grano se registró en bushels por acre y ajustados a 155 g H₂O kg⁻¹. Se analizó el rendimiento de grano en t ha⁻¹ y el contenido de humedad del grano en g H₂O kg⁻¹.

Análisis estadístico

Según los análisis comúnmente utilizados en ensayos de rendimiento en muchas localidades, se usaron tres modelos lineales mixtos para los híbridos. Modelo mixto 1: efectos fijos de localidades y de híbridos; Modelo mixto 2: efectos aleatorios de localidades y efectos fijos de híbridos; Modelo mixto 3: efectos aleatorios de localidades y de híbridos. También se utilizaron tres modelos mixtos para analizar líneas progenitoras hembras y machos dentro de las combinaciones híbridas: Modelo mixto 1: efecto aleatorio de localidades y de progenitores machos y efecto fijo de progenitores hembras; Modelo mixto 2: efecto aleatorio de localidades y de progenitores hembras y efecto fijo de progenitores machos; Modelo mixto 3: todos los efectos

data. For grain yield at all individual locations, the SEP for mixed model 3 (0.634) was nearly 33% less than the SEP calculated by mixed model 1 (1.052), and mixed model 2 (0.643) produced a SEP very similar to that of mixed model 3. Within groups of locations, reductions in SEP between mixed models 2 and 3 ranged from 5% in Nebraska to 25% in Indiana locations, because experimental errors were larger in Illinois and Indiana than those in Nebraska and Iowa. For grain moisture the same pattern was found indicating that reduced SEP are obtained when location and hybrid effects are considered random effects.

The three mixed models, when used for parental inbred line effects of the testcrosses, also gave different SEP. Smaller SEP were found when female, male, and location effects were considered as random effects, compared to the mixed model that had one of these effects as fixed. At a specific location, SEP for comparing differences in grain yield between two inbred lines were 0.6 for mixed model 1, whereas for mixed models 2 and 3 the SEP were 0.543 and 0.406. For comparing differences between inbred male lines, at a specific location, the SEP were 0.523, 0.427, and 0.513, for mixed models 1, 2, and 3. When all 22 locations were analyzed together, the SEP for grain moisture followed the same trend. For differences between female inbred lines, SEP of 1.351, 0.763, and 0.705 were found for mixed models 1, 2 and 3, and the SEP for male inbred lines were 1.113, 1.166, and 0.739. Sanders (1989) compared different mixed models for analysis of soybean varieties and found that the SEP for one entry was 2.5 when the entry effect was considered fixed, but when considered random the SEP decreased to 2.2. In plant breeding, the smaller standard errors for the predictable functions indicate that they are more precise estimates than the adjusted means.

Variance components

The Mixed and Varcomp procedures produced the same estimates of variance components for testcross grain yield and moisture for all location groups with both balanced and unbalanced data (Table 1), but small differences in these estimates were obtained when GLM procedure was used, because the Nebraska and Indiana groups had three and two missing observations, respectively. For the Iowa and Illinois groups, which had balanced data, these three methods resulted in the same values for all variance components for both traits. When estimates of variance components were obtained for parental female and male inbred lines, GLM computed different values (compared to Mixed and Varcomp) for each of the variance component estimates within each group of locations (Table 2).

aleatorios. Un error estándar de predicción pequeño obtenido con estos modelos indicaría mayor eficiencia estadística.

Como se indicó previamente, las localidades individuales se agruparon por estado y se analizaron por grupo y en forma combinada. Para el análisis estadístico, los datos de los grupos de localidades de Iowa, Illinois e Indiana se combinaron con los del grupo de localidades de Nebraska, hasta incluir todos los grupos (las 22 localidades) para su análisis combinado. Se utilizaron los procedimientos Mixed, GLM y Varcomp del SAS (versión 6.11; Littell *et al.*, 1996) para estimar y comparar los componentes de varianza. Se le dieron las opciones "random" y "test" al procedimiento GLM para probar la hipótesis de que los componentes de varianza eran cero.

RESULTADOS Y DISCUSIÓN

Comparación de los diferentes modelos mixtos

Los tres modelos mixtos produjeron diferentes errores estándar de predicción (EEP) de los datos de rendimiento de los híbridos para cada grupo de localidades y para el análisis combinado de grupos. Para rendimiento de grano en todas las localidades individuales, el EEP para el modelo mixto 3 (0.634) fue casi 33% menor que el EEP calculado por el modelo mixto 1 (1.052), y el modelo mixto 2 (0.643) produjo un EEP muy similar al del modelo mixto 3. Dentro del grupo de localidades, las reducciones en el EEP entre los modelos mixtos 2 y 3 variaron de 5% en localidades de Nebraska a 25% en las de Indiana, porque los errores experimentales fueron más grandes en Illinois y en Indiana que los de Nebraska e Iowa. El mismo patrón se encontró para humedad de grano, indicando que se pueden obtener menores EEP cuando los efectos de localidades y de híbridos se consideran efectos aleatorios.

También se obtuvieron diferentes EEP en los tres modelos mixtos que se utilizaron para los efectos de las líneas progenitoras de los híbridos experimentales. Se encontraron EEP más pequeños cuando los efectos de las líneas hembras, las líneas machos y las localidades se consideraron aleatorios, comparado con el modelo mixto que tuvo fijo uno de estos efectos. Para una localidad específica, el EEP para comparar diferencias en rendimiento de grano entre líneas endogámicas fue 0.600 para el modelo mixto 1, mientras que para los modelos mixtos 2 y 3, los EEP fueron 0.543 y 0.406. Para comparar diferencias entre líneas endogámicas machos, en una localidad específica, los EEP fueron 0.523, 0.427 y 0.513, para los modelos mixtos 1, 2 y 3. Cuando las 22 localidades fueron analizadas en conjunto, el EEP para humedad de grano siguió la misma tendencia. Para diferencias entre las líneas endogámicas hembras, los EEP encontrados fueron de 1.351, 0.763 y 0.705 para los modelos mixtos 1, 2 y 3, y los EEP para las líneas endogámicas macho fueron 1.113, 1.166

Table 1. Estimates of variance components from corn testcrosses using three statistical procedures with mixed model 3: both location and hybrid effects as random.**Cuadro 1. Estimaciones de componentes de varianza de híbridos de maíz usando tres procedimientos con el modelo mixto 3: efectos aleatorios de localidad y de híbrido.**

Group of locations	Variance	Mixed	GLM	Varcomp	Mixed	GLM	Varcomp
		Grain yield (t ha ⁻¹)			Moisture (g H ₂ O kg ⁻¹)		
Nebraska [†]	σ^2_{loc}	1.919	1.920	1.919	9.093	9.131	9.131
	σ^2_{hyb}	0.124	0.125	0.124	0.399	0.404	0.403
	$\sigma^2_{loc*hyb}$	0.352	0.355	0.352	0.171	0.169	0.169
	$\sigma^2_{rep(loc)}$	0.009	0.009	0.009	0	0	0
	σ^2_{error}	0.765	0.763	0.765	0.495	0.496	0.496
Iowa	σ^2_{loc}	0.508	0.507	0.508	2.927	2.927	2.927
	σ^2_{hyb}	0.154	0.154	0.154	1.738	1.738	1.738
	$\sigma^2_{loc*hyb}$	0.116	0.115	0.116	0.664	0.664	0.664
	$\sigma^2_{rep(loc)}$	0.058	0.058	0.058	0.039	0.039	0.039
	σ^2_{error}	0.690	0.690	0.690	1.133	1.133	1.133
Illinois	σ^2_{loc}	2.653	2.653	2.653	29.153	29.153	29.153
	σ^2_{hyb}	0.093	0.093	0.093	1.794	1.795	1.795
	$\sigma^2_{loc*hyb}$	0.453	0.452	0.453	1.871	1.871	1.871
	$\sigma^2_{rep(loc)}$	0.294	0.294	0.294	0.351	0.351	0.351
	σ^2_{error}	1.392	1.392	1.392	1.273	1.273	1.273
Indiana [‡]	σ^2_{loc}	1.580	1.591	1.580	9.266	9.286	9.286
	σ^2_{hyb}	0.004	0.003	0.004	0.915	0.911	0.911
	$\sigma^2_{loc*hyb}$	0.216	0.217	0.216	0.655	0.656	0.656
	$\sigma^2_{rep(loc)}$	0.092	0.087	0.092	0.057	0.057	0.057
	σ^2_{error}	1.575	1.574	1.575	0.979	0.977	0.977

[†] Within this group of locations there were three missing observations.[‡] Within this group of locations there were two missing observations.

The Varcomp of SAS is a procedure used to obtain estimates of variance components in a mixed model. Varcomp calculates analysis of variance estimators (based on Henderson's Method III) from the Type I sum of squares. The GLM procedure of SAS was the principal procedure used for analyzing mixed models data prior to the advent of Mixed procedure. Even though the basic computations used in GLM were developed for fixed effects, they give appropriate estimates of variance components for balanced data (Searle *et al.*, 1992).

The advantages of the Mixed procedure are that even though the basic computations are based on likelihood principles, many of the statistical computations are the same as those used in analysis of variance methods for balanced data (Littell *et al.*, 1996), and estimates of random effects can be obtained in the same analysis. Since all procedures in SAS use Restricted Maximum Likelihood method to estimate variance components, the results are very similar.

Estimates of variance components

All estimates of variance components for grain yield and grain moisture content were significantly different from zero, thus indicating that all the variance components

y 0.739. Sanders (1989) comparó diferentes modelos mixtos para el análisis de variedades de soya y encontró que el EEP para variedad fue 2.5 cuando el efecto de variedad se consideró fijo, pero cuando se consideró aleatorio, el EEP disminuyó a 2.2. En mejoramiento de plantas, los errores estandar más pequeños para las funciones predecibles indican que son estimaciones más precisas que las medias ajustadas.

Componentes de varianza

Los procedimientos Mixed y Varcomp produjeron las mismas estimaciones de los componentes de varianza del rendimiento y de la humedad de grano de los híbridos experimentales, para todos los grupos de localidades, tanto con datos balanceados como con desbalanceados (Cuadro 1), pero se obtuvieron pequeñas diferencias en esas estimaciones cuando se utilizó el procedimiento GLM, debido a que los grupos de Nebraska e Indiana tuvieron tres y dos observaciones perdidas. Para los grupos Iowa e Illinois, con datos balanceados, los tres procedimientos proporcionaron los mismos valores para todos los componentes de varianza y para ambas variables. Cuando se obtuvieron las estimaciones de componentes de varianza de las líneas

Table 2. Estimates of variance components of parental inbred lines of testcrosses on mixed model 3: location, female, and male effects as random, using three procedures.**Cuadro 2. Estimaciones de componentes de varianza de líneas progenitoras de híbridos usando tres procedimientos con el modelo mixto 3: efectos aleatorios de localidad y de líneas.**

Group of locations	Variance	Mixed	GLM	Varcomp	Mixed	GLM	Varcomp
		Grain yield (t ha ⁻¹)			Moisture (g H ₂ O kg ⁻¹)		
Nebraska	σ^2_{loc}	2.050	2.263	2.050	8.838	10.140	8.839
	σ^2_{fem}	0.045	0.056	0.045	0.187	0.218	0.187
	σ^2_{male}	0.076	0.170	0.076	0.405	0.606	0.405
	$\sigma^2_{loc*fem}$	0.048	0.023	0.048	0.021	0.062	0.021
	$\sigma^2_{loc*male}$	0.412	0.344	0.412	0.161	0.413	0.161
Iowa	σ^2_{loc}	0.522	0.519	0.522	2.527	2.478	2.527
	σ^2_{fem}	0.166	0.147	0.166	0.866	0.967	0.886
	σ^2_{male}	0.005	0	0.005	1.709	2.527	1.709
	$\sigma^2_{loc*fem}$	0.056	0.070	0.056	0.093	0.103	0.093
	$\sigma^2_{loc*male}$	0.071	0.001	0.071	0.809	1.076	0.809
Illinois	σ^2_{loc}	2.796	2.227	2.796	34.933	29.115	31.990
	σ^2_{fem}	0.115	0.107	0.115	0.799	0.882	0.724
	σ^2_{male}	0	0	0	1.432	1.983	1.611
	$\sigma^2_{loc*fem}$	0.138	0.034	0.138	0	0.041	0.463
	$\sigma^2_{loc*male}$	0.486	0.367	0.486	2.703	2.099	1.887
Indiana	σ^2_{loc}	1.624	1.558	1.624	9.082	8.692	9.082
	σ^2_{fem}	0	0	0	0.500	0.585	0.500
	σ^2_{male}	0.021	0.044	0.021	0.728	0.952	0.728
	$\sigma^2_{loc*fem}$	0.070	0	0.070	0.327	0.416	0.327
	$\sigma^2_{loc*male}$	0.291	0.538	0.291	0.486	0.996	0.486

in the model were required to adequately describe the variation in the data, except for replications with the Nebraska group for both traits and for testcross grain yield in Indiana and Illinois groups (Table 3). Phenotypic variation for grain yield in Iowa was -52%, in Illinois was +54%, and nearly the same in Indiana (-9%) compared to Nebraska, where the testcrosses were made. The main component of the phenotypic variation was attributable to differences among locations: 33% for Iowa, 45% for Indiana, 54% for Illinois, and 60% for Nebraska. The small phenotypic variation observed in the Iowa group (1.526) could indicate a result of homogeneity of six locations that comprised that group. Conversely, the greater variation found in Illinois (4.885) was due to the heterogeneity among the locations within that group.

The second largest component of phenotypic variation of grain yield was that unexplained part due to experimental errors, which were 24 and 28% in Nebraska and Illinois, and 45% in Iowa and Indiana. The third largest component was variation due to genotypic by environmental interaction, which was somewhat larger in Nebraska (11%) and Illinois (9%) than in Iowa (7%) and Indiana (6%). Thus, testcrosses responded differentially according with the proportion of variability attributable to the several components among groups of environments. Groups with greater heterogeneity also had greater interaction terms.

endogámicas progenitoras hembras y machos, GLM calculó diferentes valores (comparados con Mixed y Varcomp) para cada una de las estimaciones de los componentes de varianza dentro de cada grupo de localidades (Cuadro 2).

El procedimiento Varcomp del SAS es utilizado para obtener estimaciones de componentes de varianza en un modelo mixto. Varcomp calcula estimadores de análisis de varianza (basado en el Método III de Henderson) de la suma de cuadrados Tipo I. El procedimiento GLM de SAS fue el principal método usado para modelos mixtos, antes del advenimiento del procedimiento Mixed. Aun cuando los cálculos computacionales básicos usados en GLM se desarrollaron para efectos fijos, proporcionan estimaciones apropiadas de componentes de varianza para datos balanceados (Searle *et al.*, 1992).

Aun cuando los cálculos computacionales básicos del procedimiento Mixed están basados en principios de verosimilitud, las ventajas son que muchos de los cálculos estadísticos computacionales son los mismos los que se usan en los métodos de análisis de varianza de datos balanceados (Littell *et al.*, 1996) y pueden obtenerse los estimadores de efectos aleatorios en el mismo análisis. Desde que todos los procedimientos del SAS usan el método de máxima verosimilitud restringida para estimar componentes de varianza, los resultados obtenidos son muy similares.

Table 3. Combined analysis for estimation of phenotypic variance components of corn testcrosses evaluated in 22 environments of the USA Corn Belt.**Cuadro 3. Análisis combinado para estimar componentes de varianza fenotípica de híbridos de maíz evaluados en 22 ambientes de la Faja Maicera de Estados Unidos.**

Group of locations	Variances					
	σ^2_{loc}	σ^2_{hyb}	$\sigma^2_{loc*hyb}$	$\sigma^2_{rep(loc)}$	σ^2_{error}	$\sigma^2_{phenotypic}$
Grain yield (t ha ⁻¹)						
Nebraska (NE)	1.919**	0.124**	0.352**	0.009 ^{ns}	0.765	3.169
Iowa (IA)	0.508**	0.154**	0.116**	0.058**	0.690	1.526
Illinois (IL)	2.653**	0.093 ^{ns}	0.453**	0.294**	1.392	4.885
Indiana (IN)	1.580**	0.004 ^{ns}	0.216*	0.092**	1.575	3.467
NE+IA	1.219**	0.118**	0.245**	0.036 ^{ns}	0.724	2.342
NE+IA+IL	1.540**	0.116**	0.304**	0.117**	0.933	3.010
NE+IA+IL+IN	1.770**	0.077**	0.289**	0.109**	1.109	3.354
Moisture (g H ₂ O kg ⁻¹)						
Nebraska	9.131**	0.404**	0.169**	0	0.496	10.200
Iowa	2.927**	1.738**	0.664**	0.039*	1.133	6.501
Illinois	29.153**	1.795**	1.871**	0.351**	1.273	34.443
Indiana	9.286**	0.911**	0.656**	0.057**	0.977	11.887
NE+IA	6.363**	0.999**	0.572**	0.021*	0.845	8.800
NE+IA+IL	12.192**	1.253**	0.972**	0.124**	0.979	15.520
NE+IA+IL+IN	11.344**	1.143**	0.903**	0.106**	0.979	14.475

*, ** Significantly different ($p \leq 0.05$) and ($p \leq 0.01$).^{ns} Not significantly different ($p > 0.05$).

The smallest component of phenotypic variation was replications within locations. In Nebraska, experimental sites were very homogeneous, indicating that a completely randomized design could have been better than a randomized complete block design. In contrast, this source of variation was large in Illinois and a completely randomized design would have been inappropriate. Similar amount of variation due to hybrids was found at Nebraska and Iowa, but at Illinois and Indiana groups it was not different from zero with large experimental errors (1.3 and 1.5); also, the group of hybrids were made for Nebraska environments, but not for far away locations, such as Laurel, Delaware, which was included into the Indiana group. For grain moisture, the main component of phenotypic variation was the environmental effect, which proportionally ranged from 45 in Iowa to 89% in Illinois. The environment effects had a larger influence in the variation for both traits than the influence of genotypes and their interaction effects. The effect of location by hybrid interaction was large when the location effect was large too, so the magnitude of the genotype by environment interaction was related to the magnitude of those environment effects. The larger the relative size of the interaction component the more complex the problem of identifying broadly adapted hybrids and, therefore, large series of experiments are necessary to get good estimates for genotypic means.

Estimaciones de componentes de varianza

Todas las estimaciones de componentes de varianza para rendimiento de grano y contenido de humedad de grano fueron significativamente diferentes de cero, indicando que se requirieron todos los componentes de varianza en el modelo para describir adecuadamente la variación en los datos, excepto por las repeticiones en el grupo Nebraska para ambos caracteres y para rendimiento de grano de los híbridos en los grupos Indiana e Illinois (Cuadro 3). La variación fenotípica para rendimiento de grano en Iowa fue 52%, en Illinois fue + 54% y casi igual en Indiana (−9%), comparada con la de Nebraska, donde se formaron los híbridos experimentales. El principal componente de la variación fenotípica fue atribuible a diferencias entre localidades: 33% para Iowa, 45% para Indiana, 54% para Illinois y 60% para Nebraska. La pequeña variación fenotípica observada en el grupo Iowa (1.526) puede indicar el resultado de la homogeneidad de las seis localidades que formaron ese grupo. Contrariamente, la mayor variación encontrada en el grupo Illinois (4.885) fue debida a la heterogeneidad entre las localidades de ese grupo.

El segundo mayor componente de la variación fenotípica del rendimiento de grano fue la parte no explicada atribuible a los errores experimentales, que fueron 24 y 28% en Nebraska e Illinois, y 45% en Iowa e Indiana. La tercer componente más grande fue la variación

The consequences of genotypic by environment interaction, from the genotype standpoint, involves genetic diversity (multiple genotypes) as an adaptive response to different environments, and different genotypes for different environmental niches are required. A more homogeneous environment allows fewer genotypes to prevail, whereas a completely homogeneous environment allows a single genotype to always be superior. Hebert *et al.* (1995) concluded that the genotypic by environmental interaction for root traits in corn, could essentially be due to differences in precipitation and temperature (either in the soil or in the air), which occur during the period preceding the emergence of primary adventitious roots.

The combined analysis carried out to estimate variances for grain yield in female and male inbred lines, and their interactions with environments, showed significance only for female lines and location by male interaction effects. For grain moisture, all estimates of variances were significantly different from zero, except for that due to location by female interaction in Nebraska and Illinois. The main component of testcross variation was due to the female line effects. Smaller and null effects of the male lines were found, which could be related to their similar genetic origin.

When groups of locations were added for combined analysis, the variance components changed in accordance to the variability of the groups added. Combined analysis of locations of Nebraska and Iowa gave smaller estimates of variance components compared with those obtained from only Nebraska locations, which was due to small variances from Iowa. When locations of Illinois were added, the variances increased because of the greater variances from Illinois. When all 22 test locations were combined, the analysis gave almost the average of the variance components obtained from each group of locations.

Testing sites for hybrid development and recommendations are usually defined by State boundaries, but these political boundaries seldom coincide with major changes in biological and environmental parameters (Jensen, 1997). Data from nine traditional testing sites were examined over a five-year period by Horney and Frey (1957) to see if a more logical subdivision would result in more meaningful oat (*Avena sativa* L.) variety recommendations. Trying to find combinations of locations whose variety by location mean squares would be lower than those for the whole region, they found that similar climatic conditions of southern Iowa lowered the variety by location interaction by 30%. This regional division is still in use for the oat breeding program. Liang *et al.* (1966) examined the environmental conditions for Kansas wheat (*Triticum aestivum* L.), oat and barley (*Hordum vulgare* L.) testing

debida a la interacción genotipo-ambiente, que fue un poco mayor en Nebraska (11%) e Illinois (9%) que en Iowa (7%) e Indiana (6%). Así, los híbridos experimentales respondieron diferencialmente de acuerdo a la proporción de variabilidad atribuible a varios componentes entre los grupos de ambientes. Los grupos con mayor heterogeneidad también tuvieron mayor interacción.

El componente de variación fenotípica más pequeño fue repeticiones dentro de localidades. En Nebraska, los sitios experimentales fueron muy homogéneos, indicando que un diseño completamente al azar pudo haber sido mejor que uno en bloques completos al azar. En contraste, esta fuente de variación fue grande en Illinois y un diseño completamente al azar hubiera sido inapropiado. Una cantidad similar de variación debida a los híbridos fue encontrada en Nebraska e Iowa, pero en los grupos de Illinois e Indiana esa variación no fue diferente de cero, con errores experimentales grandes (1.3 y 1.5); el grupo de híbridos utilizados fueron formados para ambientes de Nebraska, pero no para localidades tan lejanas como Laurel, en el estado de Delaware, la cual se agrupó con las localidades de Indiana. Para humedad de grano, el principal componente de la variación fenotípica fue el efecto ambiental, el cual varió proporcionalmente de 45 en Iowa a 89% en Illinois. Los efectos ambientales tuvieron más influencia en la variación de ambos caracteres que los efectos de genotipos y de interacción. El efecto de interacción genotipo por híbrido fue grande cuando el efecto ambiental también lo fue, así que la magnitud de la interacción genotipo-ambiente estuvo relacionada con esos efectos ambientales. Si el tamaño relativo del componente de interacción es grande, también es más complejo el problema de identificar híbridos de amplia adaptación, y por tanto, se necesita un número grande de experimentos para obtener buenos estimadores de las medias genotípicas.

Desde el punto de vista genotípico, las consecuencias de la interacción genotipo-ambiente involucran diversidad genética (varios genotipos) como una respuesta adaptativa a diferentes ambientes y se requieren diferentes genotipos para diferentes nichos ambientales. Un ambiente más homogéneo permite recomendar pocos genotipos, mientras que uno completamente homogéneo permite que un solo genotipo siempre sea el mejor. Herbert *et al.* (1995) concluyeron que este efecto de interacción pudo deberse a diferencias en precipitación y temperatura (tanto en el suelo como en el aire), que ocurren durante el periodo que precede a la emergencia de las raíces primarias adventicias.

El análisis combinado hecho para estimar varianzas del rendimiento de grano de las líneas hembras y machos y sus interacciones con ambientes mostró significancia solamente para efectos de líneas hembras y

sites and came to a similar conclusion, resulting in reduction of the number of site tests.

Johnson (1977) discussed clustering of corn hybrids in terms of its relationship to comparative stability as derived from regression coefficient analysis and genotypic similarity, and suggested that clustering of environments might be done. He stated that strong demand for a hybrid depends more upon minimizing of performance problems in stress or low productivity environments than upon high average yields over all environments.

The purpose of this research was to examine procedures for analyzing testcross data according to the linear statistical models that can be chosen by the breeder. The mixed model of choice will depend on the type of inferences to be made. If inferences are going to be drawn about just the genotypes that occur in the data, then the genotypes may be considered as fixed effects. If not, then inferences will be made about some population of genotypes, from which those in the data set are presumed to be a random sample, in this case, the genotype effects may be considered as random. In this study, smaller standard errors were obtained when genotype and environment were considered random rather than fixed effects. Therefore, differences among genotypes are determined with more accuracy and confidence if both, genotype and environmental effects, are considered random.

Estimates of variance components can be obtained with balanced or unbalanced data from multilocation yield trials using the Mixed procedure of SAS. Partitioning observed phenotypic variation into their components is the basis of a quantitative genetics study and allows estimation of relative importance of different effects over the phenotype and identification of components associated with heritability and environmental factors. Commercial breeding programs may use these statistical methods to allocate their test locations or environments according to their objectives and marketing areas for corn production.

CONCLUSIONS

When location and genotype effects were considered as random in multilocation corn yield trials, standard errors of prediction were 33% smaller compared to those obtained from location and genotype fixed effects, and differences among genotypes were determined with more precision. GLM, Mixed and Varcomp procedures computed the same estimates of variance components using balanced data, but with unbalanced data only Mixed and Varcomp gave the same estimates. All estimates of variance components for both grain yield and grain moisture content were significantly different from zero. Compared to Nebraska environments, where the hybrids

de interacción líneas macho por localidades. Para humedad de grano, todos los estimadores de varianzas fueron significativamente diferentes de cero, excepto para aquel debido a la interacción líneas hembra por localidad, en localidades de Nebraska e Illinois. El principal componente de variación de los híbridos experimentales fue debido a los efectos de las líneas hembras. Los efectos de las líneas macho fueron más pequeños y nulos, lo cual pudo estar relacionado con su origen genético similar.

Cuando se agregaron todos los grupos de localidades para el análisis combinado, los componentes de varianza cambiaron de acuerdo con la variabilidad de los grupos agregados. Los análisis combinados de las localidades de Nebraska e Iowa dieron estimaciones de varianza más pequeñas comparadas con aquellas obtenidas solamente de las localidades de Nebraska, debido a las pequeñas varianzas de Iowa. Cuando las localidades de Illinois fueron agregadas las varianzas se incrementaron por las varianzas más grandes de Illinois. Cuando se combinaron las 22 localidades de prueba, el análisis arrojó casi el promedio de componentes de varianza obtenido de cada grupo de localidades.

Los sitios de prueba para el desarrollo y recomendaciones de híbridos generalmente están definidos por límites estatales, pero estos límites políticos raramente coinciden con los cambios en parámetros biológicos y ambientales (Jensen, 1997). Datos de nueve sitios de prueba tradicionales fueron examinados por un periodo de cinco años por Homey y Frey (1957) para analizar si una subdivisión más lógica podía dar recomendaciones más adecuadas de variedades de avena (*Avena sativa* L.). Tratando de encontrar combinaciones de localidades cuyos cuadrados medios de variedad por localidad pudieran ser más bajos que los de toda la región, esos autores encontraron que en condiciones climáticas similares a las del sur de Iowa disminuyeron 30% la interacción variedad por localidad. Esta división regional aún se usa en el programa de mejoramiento de avena. Liang *et al.* (1966) examinaron las condiciones ambientales de los sitios de prueba de trigo (*Triticum aestivum* L.), avena y cebada (*Hordeum vulgare* L.) y llegaron a una conclusión similar de reducción del número de sitios de prueba.

Johnson (1977) discutió el agrupamiento de híbridos en términos de su relación con la estabilidad comparativa derivada del análisis del coeficiente de regresión y la similitud genotípica, y sugirió que puede hacerse el agrupamiento de ambientes. Este autor estableció que la gran demanda de un híbrido depende más de la reducción de los problemas de comportamiento en ambientes estresados o de baja productividad que de los altos rendimientos promedio en todos los ambientes.

El propósito de esta investigación fue examinar procedimientos para analizar datos de híbridos experimentales de acuerdo con los modelos lineales estadísticos

were made, phenotypic variance of grain yield was 52% smaller in Iowa locations, 9% smaller in Indiana locations and 54% bigger in Illinois locations. This information could be useful in determining more appropriate testing sites over the USA Corn Belt.

CITED LITERATURE

- Bauman, L. F. 1981. Review of methods used by breeders to develop superior hybrids. *In: Proc. 36th Annual Corn and Sorghum Res. Conf. ASTA. Chicago, IL.* pp: 199-208.
- Bridges, Jr., W. C. 1989. Analysis of a plant breeding experiment with heterogeneous variances using mixed model equations. *In: Applications of Mixed Models in Agriculture and Related Disciplines. Southern Cooperative Series. Bulletin No. 343. Louisiana Agric. Exp. Stn., Baton Rouge.* pp: 145-154.
- Hallauer, A. R. 1990. Methods used in developing maize inbreds. *Maydica.* 35: 1-16.
- Hallauer, A. R., W. A. Russell, and K. R. Lamkey. 1988. Corn breeding. *In: Sprague, G. (ed). Corn and Corn Improvement. Agronomy Monograph no. 18. 3rd Ed. Wiley and Sons. Madison, WI.* pp: 463-564.
- Harville, D. A. 1976. Confidence intervals and sets for linear combinations of fixed and random effects. *Biometrics* 32: 403-407.
- Hebert, Y., C. Plomion, and N. Harzic. 1995. Genotype by environment interaction for root traits in maize, as analyzed with factorial regression models. *Euphytica* 81: 85-92.
- Horney, T. W., and K. J. Frey. 1957. Methods for determining natural areas for oat varietal recommendations. *Agron. J.* 49: 313-315.
- Jensen, N. F. 1997. *Plant Breeding Methodology.* J. Wiley and Sons, Inc. N.Y., USA. 676 p.
- Johnson, G. R. 1977. Analysis of genotypic similarity in terms of mean yield and stability of environmental response in a set of maize hybrids. *Crop Sci.* 17: 837-842.
- Liang, G. H. L., E. G. Heyne, and T. L. Walter. 1966. Estimates of varietyxenvironment interactions in yield tests of three small grains and their significance on the breeding programs. *Crop Sci.* 6: 135-139.
- Littell, R. C., G. A. Milliken, W. W. Stroup, and R. D. Wolfinger. 1996. *SAS System for Mixed Models.* SAS Institute Inc., Cary, NC, USA. 633 p.
- Sanders, W. L. 1989. Choice of Models. *In: Applications of Mixed Models in Agriculture and Related Disciplines. Southern Cooperative Series Bulletin No. 343. Louisiana Agric. Exp. Stn., Baton Rouge.* pp: 49-56.
- Searle, S. R., Casella, G., and McCulloch, C. E. 1992. *Variance Components.* J. Wiley and Sons, Inc. N.Y., USA. 501 p.
- Smith, O. S., K. Hoard, F. Shaw, and R. Shaw. 1996. Prediction of single-cross performance. *In: Proc. 51th Annual Corn and Sorghum Res. Conf. ASTA. Chicago, IL.* pp: 67-74.

que pueden ser elegidos por el mejorador. El modelo mixto a elegir dependerá del tipo de inferencias que se van a hacer. Si las inferencias van a extenderse solamente a los genotipos que están en los datos, entonces los genotipos pueden ser considerados como efectos fijos. De otra forma, las inferencias se harán sobre alguna población de genotipos, de la cual se supone que estos genotipos son una muestra aleatoria y, en este caso, sus efectos pueden ser considerados como aleatorios. En este estudio, se obtuvieron errores estándar más pequeños cuando genotipo y ambiente fueron considerados efectos aleatorios en lugar de fijos. Por esa razón, las diferencias entre genotipos se determinan con más precisión y confianza si ambos efectos se consideran aleatorios.

Se pueden obtener estimaciones de componentes de varianza con datos balanceados o desbalanceados de ensayos de rendimiento de muchas localidades utilizando el procedimiento Mixed de SAS. La descomposición de la variación fenotípica en sus componentes es básica en un estudio de genética cuantitativa y permite una estimación de cierta importancia de los diferentes efectos sobre el fenotipo e identificación de componentes asociados con la heredabilidad y los factores ambientales. Programas de mejoramiento comerciales pueden utilizar estos métodos estadísticos para definir sus localidades o ambientes de prueba, de acuerdo a sus objetivos y áreas de mercado para la producción de maíz.

CONCLUSIONES

Cuando los efectos localidad y genotipo se consideraron aleatorios, en ensayos de rendimiento de muchas localidades, los errores estándar de predicción fueron 33% más pequeños, comparados con los obtenidos cuando los efectos de localidad y genotipo fueron fijos, y las diferencias entre genotipos se determinaron con mayor precisión. Los procedimientos GLM, Mixed y Varcomp calcularon las mismas estimaciones de componentes de varianza utilizando datos balanceados; pero con datos desbalanceados, solamente Mixed y Varcomp proporcionaron las mismas estimaciones. Todas las estimaciones de componentes de varianza, tanto para rendimiento como para humedad de grano, fueron significativamente diferentes de cero. En comparación con los ambientes de Nebraska, donde se formaron los híbridos, la varianza fenotípica de rendimiento de grano fue 52% más pequeña en las localidades de Iowa, 9% más pequeña en localidades de Indiana y 54% más grande en las de Illinois. Esta información podría ser útil para determinar más apropiadamente los sitios de prueba en la Faja Maicera de los Estados Unidos.

—Fin de la versión en Español—

