

Validación de modelos estadísticos multivariados en el proceso de selección de variedades de caña de azúcar. Caso: variedad CC 85-92

Fredy Salazar y Jorge I. Victoria K.*



Variedad CC 85-92

Introducción

En los procesos de mejoramiento genético es necesario evaluar nuevas técnicas estadísticas de análisis, adaptarlas e implementarlas para hacer más eficiente la selección de variedades y la predicción del comportamiento de los genotipos en determinados ambientes. En este sentido se han desarrollado diferentes técnicas multivariadas que han sido utilizadas ampliamente en cultivos como maíz, arroz, soya y trigo.

Los modelos multivariados fueron desarrollados a principios de 1900 pero sólo en las últimas décadas han sido aplicados en el campo del mejoramiento genético. Muchos de ellos contribuyen a explicar en detalle algunos fenómenos biológicos como la interacción genotipo por ambiente, lo cual les permite a los fitomejoradores hacer más eficiente la selección de los genotipos, maximizar el potencial genético y predecir el comportamiento y la adaptación en ambientes específicos. Entre los modelos estadísticos multivariados de mayor uso están el modelo de efectos principales aditivos e interacciones multiplicativas (*Additive Main Effects and Multiplicative Interaction: AMMI*) y el modelo de regresión de sitios (SREG).

El modelo SREG se utiliza para el análisis de los datos provenientes de ensayos multiambientales comparativos de rendimiento, en especial cuando el ambiente (E) es la fuente de variación más importante en relación con la contribución del genotipo (G) y la interacción genotipo-ambiente (GE). Este modelo, que incluye el término bilineal G+GE, proporciona un análisis gráfico del comportamiento (rendimiento y estabilidad) de los genotipos denominado biplot GGE. Este gráfico permite identificar el genotipo de mayor potencial en cada ambiente y agrupar genotipos y ambientes con patrones similares de respuesta (Ibáñez *et al.*, 2006). Los modelos AMMI, al permitir una representación biplot de filas (genotipos) y columnas (ambientes), dan la posibilidad de estudiar el grado de estabilidad de los genotipos al ser probados en diferentes ambientes (Varela y Castillo, 2005). Los modelos AMMI suelen denominarse como AMMI1, AMMI2... AMMI n , según el número de componentes principales usadas para modelar la interacción (Balzarini *et al.*, 2005).

El presente trabajo tiene como objetivo validar en caña de azúcar el uso de los dos modelos multivariados mencionados, su capacidad predictiva y la eficiencia para el manejo y uso de la interacción genotipo por ambiente con el enfoque de la agricultura específica por sitio. Los resultados de la validación indican que con el uso de estos modelos de análisis en las pruebas regionales de variedades Cenicaña Colombia (CC) se facilitarán los procesos de selección de genotipos con potencial de multiplicación para la producción a escala comercial.

* Ingeniero Agrónomo, Ph.D., Fitomejorador <fsalazar@cenicana.org>; Ingeniero Agrónomo, Ph.D., Director Programa de Variedades <jivictoria@cenicana.org>. Ambos de Cenicaña

Materiales y métodos

Se usaron los datos de la prueba regional de la serie 1985 (tres cortes, entre marzo de 1990 y diciembre de 1994) que incluyó las variedades de caña de azúcar CC 85-27, CC 85-47, CC 85-53, CC 85-63, CC 85-71, CC 85-92 y CC 85-96 y como testigos la MZC 74-275 y la V 71-51. La prueba fue sembrada en un diseño de bloques completos al azar y cuatro repeticiones, donde el área de la parcela útil varió entre 72–105 m². En el Cuadro 1 se muestran los ingenios azucareros y las zonas agroecológicas de la prueba regional.

Para la validación de los modelos se tuvo en cuenta la información de lotes comerciales del período 1995-2009 reportada por doce ingenios azucareros donde se sembraron las variedades CC 85-92, CC 84-75, CC 85-96, MZC 74-275, PR 61-632 y V 71-51; se analizaron las variables rendimiento (porcentaje de azúcar recuperado por tonelada de caña molida, en peso), toneladas de caña por hectárea (TCH) y toneladas de azúcar por hectárea (TAH), para un total de 289,542 registros.

En el análisis de estabilidad y adaptabilidad de las variedades se usaron el modelo de efectos principales aditivos e interacciones multiplicativas (*Additive Main Effects and Multiplicative Interaction*, AMMI) (Gauch, 2006; Crossa *et al.*, 2004) y el modelo de regresión de sitios (SREG) (Yan *et al.*, 2000; Yan *et al.*, 2001; Yan and Kang, 2003). La información de la prueba regional y los datos comerciales fueron analizados por localidad y mediante análisis combinado a través de ambientes. Se usaron el modelo lineal mixto (Yan, 2007) y el paquete estadístico SAS para Windows® versión 9.2 (SAS, 2008).

Resultados y discusión

Análisis de la prueba regional de la serie 85

El análisis combinado de varianza de 45 sitios (15 haciendas por tres cortes) mostró coeficientes de variación bajos para las variables sacarosa % caña (6%), TCH (12%) y toneladas de sacarosa por hectárea TSH (13%) y heredabilidades superiores al 96%, lo cual indicó un adecuado control del error experimental. En los tres caracteres estudiados se encontraron diferencias estadísticas altamente significativas para las fuentes de variación localidad (L), variedad (V) y la interacción localidad por variedad (L*V). En la expresión de las tres variables el efecto más importante fue el de la localidad, seguido por la variedad y la interacción localidad por variedad. Los resultados son explicados por las diferencias genéticas entre las variedades y la respuesta diferente de éstas en los distintos ambientes (agronomía, suelos, agua, fertilización, etcétera). La expresión fenotípica de las variedades dependió directamente de los ambientes. Se encontraron diferencias estadísticas significativas de las variedades entre haciendas y cortes y en las interacciones de haciendas por variedad y corte por variedad, lo cual hace difícil la selección de genotipos de adaptación amplia y lleva a la selección de genotipos de adaptación específica según haciendas y zonas agroecológicas.

En la Figura 1 se presenta el biplot o gráfica de doble entrada de las dos primeras componentes principales (CP1 y CP2) de la interacción localidad por variedad (biplot AMMI2) para el carácter TSH. Las variedades y sitios próximos al centro (0,0) se consideran estables (su TSH

Cuadro 1. Ingenios, haciendas y zonas agroecológicas en donde fue sembrada la prueba regional con variedades Cenicaña Colombia (CC) de la serie 1985 (tres cortes, 1990-1994).

Ingenio	Hacienda	Zona agroecológica ¹
Castilla	Vallecitos	29H1, 30H1, 7H1
	Calanda	10H2, 11H2
Incauca	San Fernando	11H3
	El Naranjo	10H4
La Cabaña	Taula Mejía	6H3
	Orocue	11H3
Manuelita	Florencia	6H1
Mayagüez	La Dominga	11H1
Pichichí	Pichichí	6H1, 6H0, 26H0
Providencia	Providencia	6H1, 11H1
	La Esmeralda	15H1
Riopaila	Normandía	5H3
	Media Luna	14H1
Risaralda	Constancia	10H5
Sancarlos	Ballesteros	6H1

1. Cuarta aproximación de la zonificación agroecológica para el cultivo de la caña de azúcar en el valle del río Cauca (Cenicaña, 2008)

no cambia entre localidades, zonas agroecológicas, años y/o manejo agronómico); en la medida que se alejan de cero, las variedades tienden a ser sensibles a cambios del ambiente, de modo que las localidades permiten observar diferencias entre variedades. Las localidades y las variedades ubicadas en el mismo cuadrante son complementarias, con interacciones positivas, lo cual significa que tales variedades presentan buena adaptación en esas localidades. La ubicación en cuadrantes opuestos expresa interacciones negativas, de forma que las variedades indicadas no se recomiendan para dichas localidades por falta de adaptación. De acuerdo con el análisis representado en el biplot AMMI2, las variedades sensibles a los cambios ambientales fueron MZC 74-275, CC 85-71, CC 85-47 y CC 85-53 y las más estables fueron CC 85-92 y CC 85-96.

La Figura 2 muestra el biplot de la primera componente principal (CP1) de la interacción y los componentes aditivos del modelo (variedades y localidades) para la variable TSH; el gráfico se identifica como biplot AMMI1 por recoger sólo una componente de la interacción. Las variedades estables y agronómicamente deseables presentaron un valor de CP1 bajo o muy cercano a cero y además un valor de TSH por encima de la media general. La variedad MZC 74-275 fue inestable y con tonelaje de sacarosa por hectárea bajo, seguida de la CC 85-47. Las variedades CC 85-96 y CC 85-92 fueron las más estables, y la última fue la de mejor TSH *per se*. Las variedades V 71-51, CC 85-63, CC 85-71 y CC 85-27 fueron estables pero su TSH *per se* fue inferior a la media general.

En la Figura 3 se muestran las gráficas de las componentes principales (CP1 y CP2) del modelo regresión de sitios por variedad (biplot GGE): la Figura 3A para 45 sitios (15 haciendas en cada uno de los tres cortes) y la Figura 3B para 15 haciendas (media de tres cortes). El análisis estableció una regresión positiva entre las variedades y los ambientes de evaluación. El análisis combinado de 45 sitios mostró que las variedades CC 85-96 y CC 85-92 correlacionaron positivamente con la mayoría de los ambientes, excepto en el segundo corte con Normandía (Riopaila, zona 5H3), La Esmeralda (Providencia, 15H1) y San Fernando (Incauca, 11H3) (Figura 3A). Las haciendas correlacionaron positivamente con las variedades CC 85-96 y CC 85-92; Normandía y La Esmeralda correlacionaron con la variedad CC 85-47 y en San Fernando las mejores variedades fueron MZC 74-275 y CC 85-63 (Figura 3B). Las variedades CC 85-71, CC 85-27 y CC 85-83 fueron inestables y no tuvieron sitios específicos de adaptación.

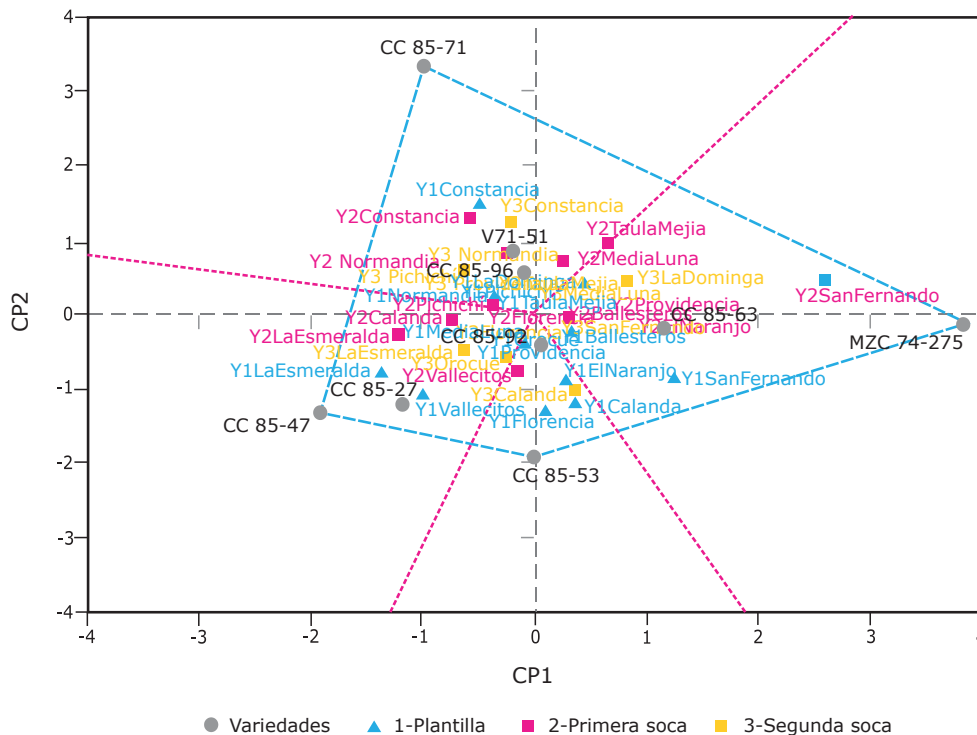


Figura 1. Biplot AMMI2 de la interacción localidad por variedad para la variable TSH (toneladas de sacarosa por hectárea) evaluada en siete variedades CC y dos testigos a través de 15 haciendas y tres cortes. Prueba regional de la serie 85. Las variedades estables son las más cercanas al centro; en los sitios más distantes del centro es donde las variedades dejan ver más diferencias fenotípicas.

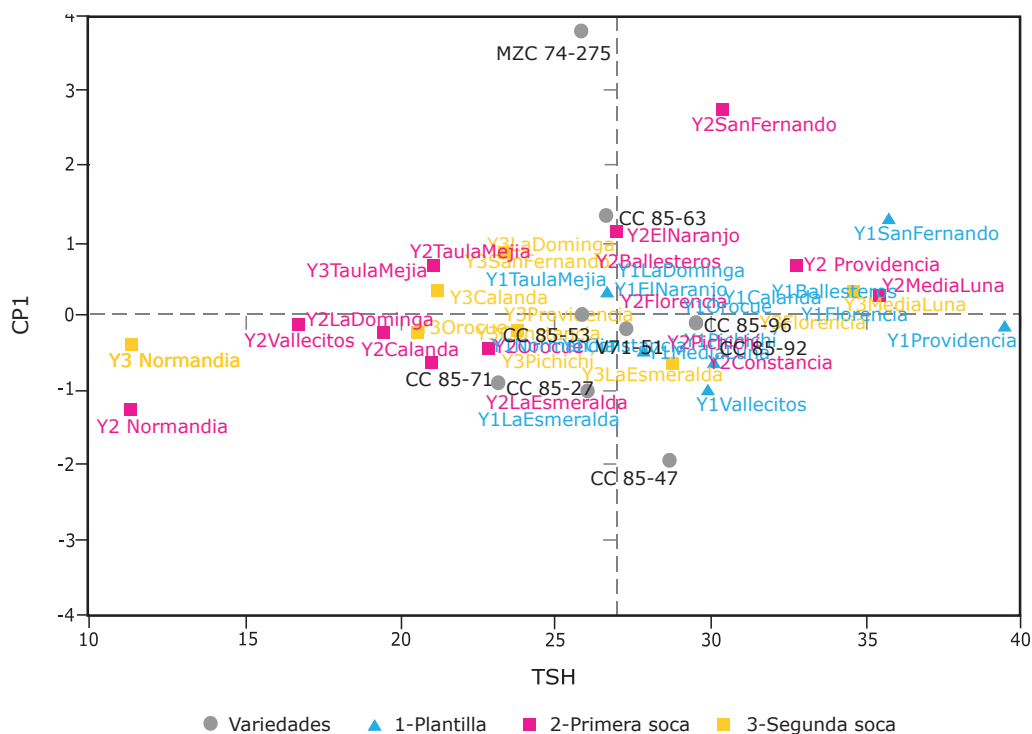


Figura 2. Biplot AMMI1 de la interacción localidad por variedad y la componente aditiva de TSH (toneladas de sacarosa por hectárea) evaluadas en siete variedades CC y dos testigos a través de 15 haciendas y tres cortes. Prueba regional de la serie 85. Las variedades más estables son las más cercanas a cero, con valores de TSH superiores a la media general.

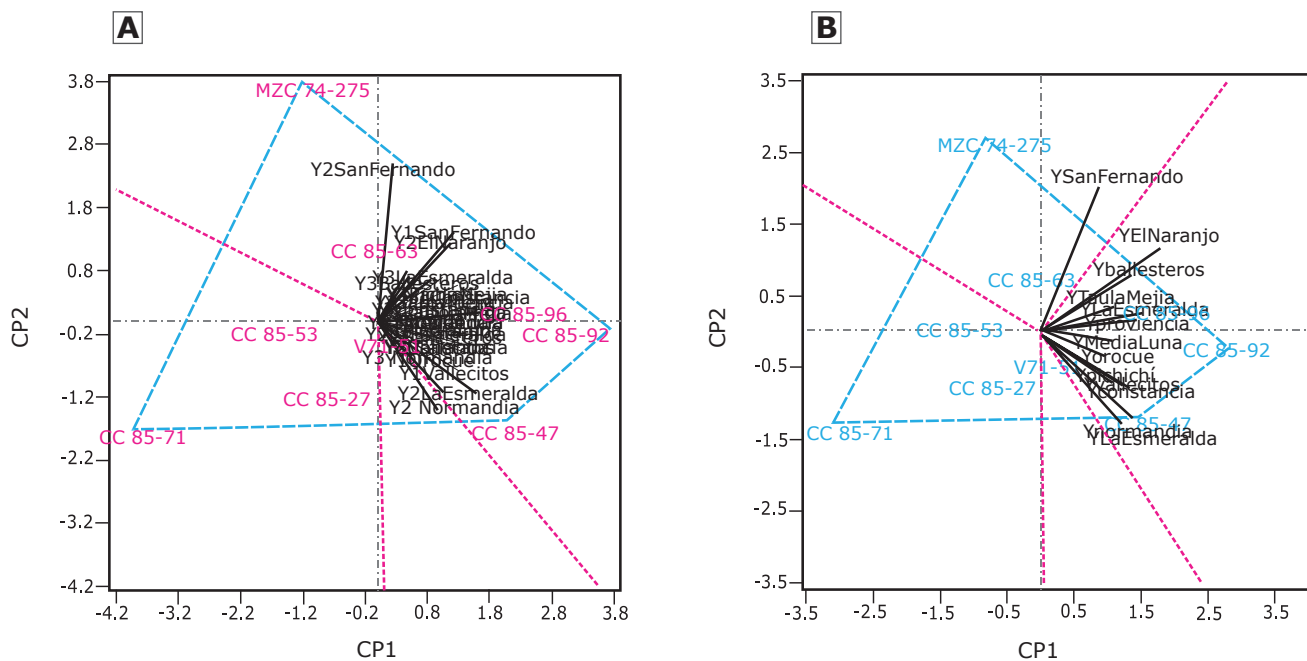


Figura 3. Biplot GGE de las dos primeras componentes principales (CP1 y CP2) de la interacción de localidad por variedad para la variable TSH (toneladas de sacarosa por hectárea) evaluada en siete variedades CC y dos testigos. Prueba regional de la serie 85: (A) 45 sitios (15 haciendas y cada uno de los tres cortes); (B) 15 haciendas (media de los tres cortes).

Análisis de datos comerciales para el carácter rendimiento (1995-2009)

Los datos correspondientes a 33 zonas agroecológicas, 12 ingenios y 15 años para el carácter rendimiento (%) fueron analizados para estimar la interacción localidad por variedad. El coeficiente de variación fue bajo (8%) y la heredabilidad fue del 80%, lo cual señala un adecuado manejo del modelo estadístico en la disminución del error experimental. Se encontraron diferencias estadísticas altamente significativas entre zonas agroecológicas, años, ingenios, variedades y en las interacciones respectivas, lo cual indica la importancia que juega el ambiente en la expresión de las variedades.

En la Figura 4 se muestran los biplots AMMI2 y GGE de la interacción localidad por variedad para la variable rendimiento (%) evaluada en seis variedades a través de 15 años. Según el biplot AMMI2 (Figura 4A), la variedad CC 85-92 es la más estable, seguida de V 71-51; las condiciones ambientales predominantes en 1996, 1998 y 2007 contribuyeron en la discriminación de genotipos con respecto al carácter rendimiento (%). El biplot GGE (Figura 4B) mostró una fuerte correlación entre la variedad CC 85-92 y los años de evaluación, excepto en 2007 y 2008 cuando la variedad CC 85-96 fue la mejor.

En el análisis por ingenio (Figura 5), las variedades CC 85-92 y CC 84-75 mostraron rendimientos estables en todos ellos, y MZC 74-275, PR 61-932 y CC 85-96 fueron sensibles a los cambios ambientales. En los ingenios Castilla, Mayagüez, Providencia y Risaralda se encontraron más diferencias entre las variedades (Figura 5A), y la variedad CC 85-92 presentó correlación directa y positiva con la mayoría de los sitios de evaluación y la CC 85-96 mostró buena adaptación en Incauca, Manuelita y Providencia (Figura 5B).

Finalmente, en la Figura 6 se presenta la estabilidad de la variedad CC 85-92 en comparación con CC 84-75, CC 85-96, MZC 74-275 y V 71-51 en función del coeficiente de regresión de cada una a través de 15 años (1995-2009) y de las desviaciones de la regresión medidas a través de la falta de ajuste del modelo o coeficiente de determinación (R^2). La variedad con rendimiento estable a través de los sitios y años de evaluación es aquella que muestra un coeficiente de determinación igual o cercano a uno (100%); la respuesta al cambio ambiental fue evaluada con el coeficiente de regresión, de modo que la respuesta es positiva cuando dicho valor es superior a uno. Así, la variedad MZC 74-275 presentó un rendimiento superior al de la CC 85-92 pero su estabilidad a través de sitios, zonas agroecológicas y años de evaluación es errático, lo cual se refleja en el bajo coeficiente de determinación (48%); el coeficiente de determinación de CC 85-92 fue de 96% (Figura 6A). La estabilidad del rendimiento de la variedad CC 84-75 (coeficiente de determinación de 82%) fue inferior a la de CC 85-92 (Figura 6B). Por su parte, la variedad CC 85-96 tuvo un rendimiento muy similar al de CC 85-92 pero su comportamiento fue bastante inestable en el tiempo, lo que se refleja en su bajo coeficiente de determinación (26%) (Figura 6C). La variedad V 71-51 mostró un rendimiento inestable ($R^2=70%$) e inferior al de CC 85-92 (Figura 6D).

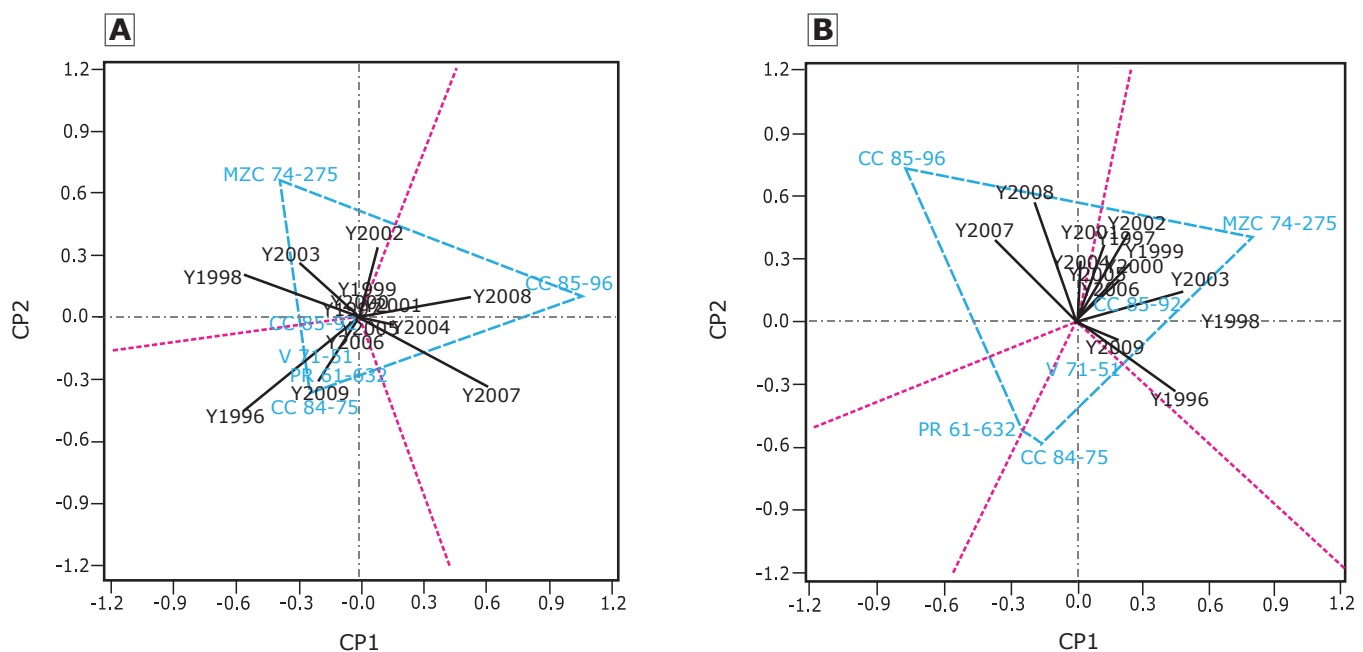


Figura 4. (A) Biplot AMMI2 y (B) Biplot GGE de la interacción localidad por variedad para la variable rendimiento (%) evaluada en seis variedades a través de 15 años (1995-2009). Datos comerciales reportados por doce ingenios azucareros.

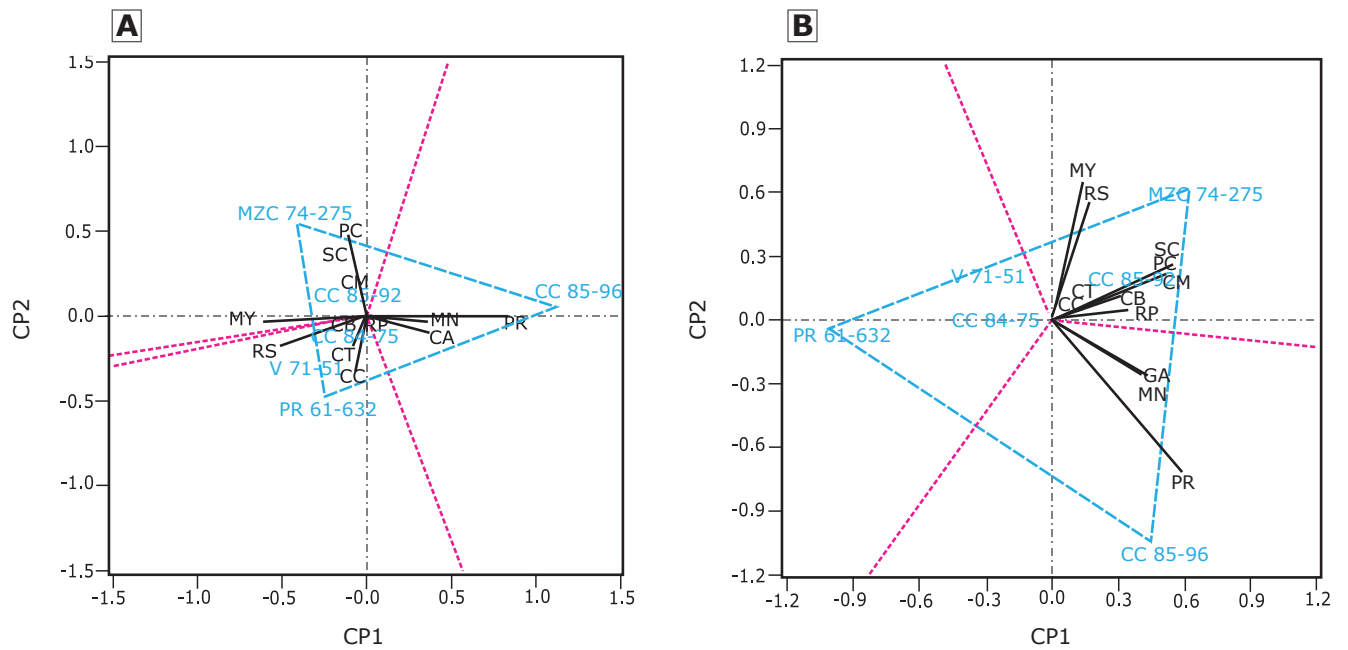


Figura 5. (A) Biplot AMMI2 y (B) Biplot GGE de la interacción localidad por variedad para la variable rendimiento (%) evaluada en seis variedades y en doce ingenios azucareros a través de 15 años (1995-2009). Datos comerciales.

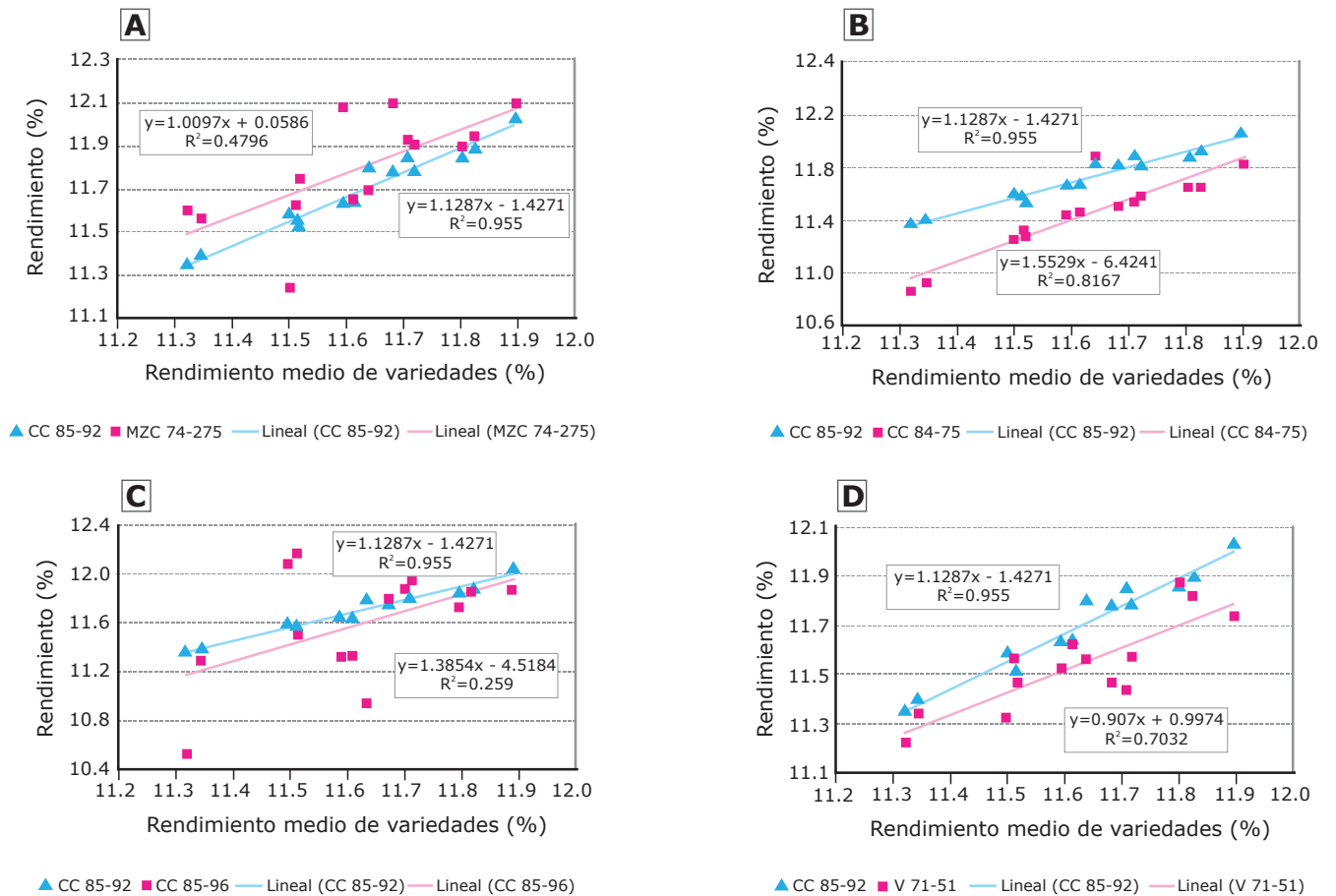


Figura 6. Estabilidad de CC 85-92 versus MZC 74-275, CC 84-75, CC 85-96 y V 71-51 en función del coeficiente de regresión y de los desvíos de la regresión estimados en la falta de ajuste del modelo (R^2), evaluadas a través de 15 años (1995-2009). Datos comerciales de doce ingenios.

Conclusiones

Los análisis de la prueba regional mostraron que la variedad CC 85-92 fue la más estable a través de los ambientes de evaluación, lo que la perfiló como una variedad promisoría. Los resultados fueron validados con el análisis de la información comercial que muestran que la variedad CC 85-92 es realmente una de las variedades más estables a través de ambientes.

Con estos resultados se concluye que los modelos multivariados AMMI y SRGE pueden ser utilizados en ensayos multiambientales de las pruebas regionales de caña de azúcar, dado que son apropiados para estudiar la interacción genotipos por ambiente y para predecir el comportamiento de las variedades a través de sitios. Los modelos multivariados son un complemento estadístico que suma eficiencia en la selección de variedades de adaptación específica, enfoque adoptado por Cenicaña para contribuir al desarrollo de la agricultura específica por sitio en el cultivo de la caña de azúcar en el valle del río Cauca.

Referencias bibliográficas

- Balzarini, M.; Bruno, C.; Arroyo, A. 2005. Análisis de Ensayos Agrícolas Multiambientales. Info-Gen, Córdoba, Argentina. 141 pp.
- Cenicaña (Centro de Investigación de la Caña de Azúcar de Colombia). 2008. Informe Anual 2007. Cenicaña, Colombia. 108 p.
- Crossa, J.; Yang, R.-C.; Cornelius, P.L. 2004. Studying crossover genotype \times environment interaction using linear–bilinear models and mixed models. *J. Agric. Biol. Environ. Stat.* 9:362–380.
- Gauch, H.G. 2006. Statistical analysis of yield trials by AMMI and GGE. *Crop Sci.* 46:1488–1500
- Ibañez, M.A.; Cavanagh, M.M.; Bonamico, N.C.; Di Renzo, M.A. 2006. Análisis gráfico mediante biplot del comportamiento de híbridos de maíz. *Revista de Investigaciones Agropecuarias.* 35, 3 (diciembre): 83-93. Disponible en <www.inta.gov.ar/ediciones/ria/35_3/art6.htm> consultado el 30/3/2010.
- Varela, M. y Castillo, J.G. 2005. Modelos con término multiplicativo. Aplicación en el análisis de la interacción genotipo ambiente. *Cultivos Tropicales.* 26, 3: 71-75. Disponible en <www.inca.edu.cu/otras_web/revista/pdf/2005/3/CT26310.pdf> consultado el 30/3/2010.
- Yan, R. 2007. Mixed-Model Analysis of Crossover Genotype-Environment Interactions. *Crop. Sc.* 47:1051-1062.
- Yan, W.; Hunt, L.A.; Sheng, Q.; Szlavics, Z. 2000. Cultivar evaluation and mega-environment investigation based on the GGE biplot. *Crop Sci.* 40:597–605.
- Yan, W.; Cornelius, P.L.; Crossa, J.; Hunt, L.A. 2001. Two types of GGE biplots for analyzing multi-environment trial data. *Crop Sci.* 41:656–663.
- Yan, W. and Kang, M.S. 2003. GGE biplot analysis: A graphical tool for breeders, geneticists, and agronomists. CRC Press, Boca Raton, FL.



Centro de Investigación de la Caña de Azúcar de Colombia - Cenicaña

Agroindustria unida en la investigación y el desarrollo

Cenicaña es una institución privada de carácter científico y tecnológico, sin ánimo de lucro, fundada en 1977 por iniciativa de la agroindustria azucarera localizada en el valle del río Cauca. Su misión es contribuir por medio de la investigación, evaluación y divulgación de tecnología y el suministro de servicios especializados al desarrollo de un sector eficiente y competitivo, de manera que éste juegue un papel importante en el mejoramiento socioeconómico y en la conservación de un ambiente productivo, agradable y sano en las zonas azucareras.

Las actividades de investigación y desarrollo son financiadas por los ingenios azucareros y los cultivadores de caña a través de donaciones directas definidas cada año como un porcentaje del valor de la producción de azúcar.

Las áreas de investigación se enmarcan en tres programas: Variedades, Agronomía y Procesos de Fábrica. Los servicios de apoyo son: Información y documentación, Economía y Estadística, Cooperación Técnica y Transferencia de Tecnología y Tecnología Informática. El Centro Experimental está ubicado a 3°13' latitud norte, a 1024 metros de altura sobre el nivel del mar. En este sitio la temperatura media anual es de 23.5 °C, la precipitación de 1160 mm y la humedad relativa de 77%.

La **Carta Trimestral** es una publicación periódica, editada por Cenicaña con el propósito de difundir información y conocimientos científicos y tecnológicos relacionados con el desarrollo de la agroindustria azucarera colombiana. Ofrece documentación resumida sobre los resultados generados por el centro de investigación y las experiencias de ingenios y cañicultores con las nuevas tecnologías, al tiempo que provee las referencias bibliográficas complementarias sobre cada tema. El primer volumen fue editado en 1978, y los cambios más significativos de diseño y concepto editorial se dieron en 1997 cuando la versión impresa comenzó a publicarse también en Internet.

Título: Validación de modelos estadísticos multivariados en el proceso de selección de variedades de caña de azúcar. Caso: variedad CC 85-92

Autores: Fredy Salazar y Jorge I. Victoria K.

Publicado en: Carta Trimestral. Cenicaña, 2010. v.32, nos. 1 y 2. p.40-46

© Centro de Investigación de la Caña de Azúcar de Colombia, 2010.

Centro Experimental: vía Cali-Florida, km 26
Tel: (57) (2) 6876611 – Fax: (57) (2) 2607853w
Oficina de enlace: Calle 58 norte no.3BN-110
Apartado aéreo: 9138
Cali, Valle del Cauca – Colombia

www.cenicana.org
buzon@cenicana.org