

DISEÑO EXPERIMENTAL DE BLOQUE AUMENTADO: UNA HERRAMIENTA PARA LA VALORACIÓN DE GENOTIPOS DE CAÑA DE AZÚCAR EN EL ESTADO II DE SELECCIÓN DE CENGICAÑA

Fredy Rosales Longo
Fitomejorador, CENGICAÑA

RESUMEN

Debido a la forma en que se plantan los experimentos del estado II, estos pueden ser analizados como un “Diseño de Bloques Aumentado”. Este diseño permite estimar el efecto de los bloques y los genotipos dentro del experimento. Los objetivos de este trabajo fueron establecer la porción de la varianza general de cada uno de los cuatro experimentos que puede ser explicada por los factores incluidos dentro del análisis de variación para cada experimento y la eficiencia relativa del diseño. Se utilizaron los datos de cuatro experimentos del estado II con genotipos de la serie CG16. Se establecieron testigos con lecturas aumentadas en cada experimento. La variable de respuesta fueron lecturas refractométricas (Brix). La información se trató mediante análisis de variación y cálculo de las eficiencias relativas del diseño. Se logró determinar que el diseño experimental aumentado explicó la varianza en cada experimento. No se encontraron diferencias estadísticas entre genotipos. La distribución de los datos, si bien es aceptable, ofrece algunas desviaciones que hay que controlar. Las eficiencias relativas del diseño son altas para todos los experimentos, lo que justifica su uso en términos de la clasificación de los genotipos en evaluación.

ABSTRACT

Due the shape and structure that the trials of the Stage II are planted, the application of a Augmented Block Design is feasible. This design allows to estimate the effects of blocks and genotypes. The new-tested varieties cannot be replicated but the check-varieties can be replicated. The main objectives of this work were to asses the way in what this experimental design estimate the portion of the variance due to the blocks and genotype and to estimate the design relative efficiency. Data of Brix of genotypes from four trials of the Stage II was used to adjust the augmented design model. ANOVA were performed in the agricolae package in R and graphics in other R-packages. Fisher relative design efficiency estimation was done. It was determined that the augmented blocks design explained the variance in six variation sources, included the residuals. No significative differences among genotypes was determined ($\alpha=0.05$). The data dispersion was little inconsistent but in general terms, acceptable. The relative efficiencies showed that the use of the Augmented Block Design is plenty justifiable in the selection Stage II at the CENGICAÑA's sugarcane breeding program.

INTRODUCCIÓN

En las estimaciones de los valores genéticos de la caña de azúcar en los estados tempranos de selección, siempre existe el problema de la escasez de material reproductivo. En el estado II de selección del Programa de Variedades de CENGICAÑA, este problema se presenta todo el tiempo. Se cuenta con solamente una pequeña cantidad de material, suficiente para plantar una pequeña unidad experimental de un surco de tres a cinco metros de largo; no es posible realizar repeticiones para cada material genético (Orozco, Quemé, Ovalle, & Rosales-Longo, 2012). La cantidad de evaluaciones se limitan en muchos casos a la valoración visual por escalas. Solamente el contenido de sólidos solubles (Brix) es determinado por refractometría y se convierte en el único valor susceptible de algún tipo de análisis cuantitativo (Rosales-Longo, *et al.*, 2018). Los análisis que se realizan se han circunscrito a estadística descriptiva y análisis de regresión lineal logística.

Este es un problema generalizado para muchos programas de mejora genética que intentan valorar sus materiales en estados muy tempranos de selección, donde se cuenta con gran cantidad de genotipos para evaluar, pero con poco material reproductivo (Sahagun-Castellanos, 1985) (Cullis, Smith, & Coombes, 2006). En el estado II de selección se cuenta entre 6,000 y 9,500 genotipos diferentes cada año. Una forma de poder hacer inferencias cuantitativas con estos tipos de experimentos es mediante el aumento relativo en r veces el número de observaciones que ocupan, dentro de los ensayos, los tratamientos correspondientes a los testigos. Cada material genético experimental se mantiene sin repeticiones.

Para manejar esta situación, se introdujo una clase de diseños llamados “Diseños Aumentados” (Federer & Raghavarao, 1975). Un caso similar se introdujo en el “Experimento Central Compuesto”, donde se aumenta un tratamiento en particular para el ajuste posterior de un modelo cuadrático (Montgomery, 1991). Un grupo de nuevas variedades (v^*), que no pueden repetirse en un experimento y otro grupo (v) de variedades testigos, que sí pueden repetirse en un experimento, son ubicadas apropiadamente dentro del diseño experimental apropiado, de tal forma que pueda controlarse la heterogeneidad del sitio de evaluación (Federer & Raghavarao, 1975). Este es el caso en los experimentos del estado II de selección del Programa de Variedades de CENGICAÑA, en donde las variedades testigos son distribuidos en forma de grilla en todo el experimento (Rosales-Longo, *et al.*, 2018).

Debido a la forma en que se suelen plantar los experimentos del estado II, estos pueden ser analizados como “Diseños de Bloques Aumentado”. Este diseño permite estimar el efecto de los bloques y los genotipos dentro del experimento.

Por definición se acepta que un diseño experimental aumentado es cualquier diseño estándar, aumentado con tratamientos (testigos) en un mismo bloque, un bloque incompleto, una fila, una columna, etcétera (Sahagun-Castellanos, 1985). La aleatorización de los testigos es importante dentro de los bloques (Federer & Raghavarao, 1975), esto permitirá varianzas del error insesgadas. Sin embargo, si los testigos son asignados aleatoriamente en un bloque, su distribución puede ser irregular y en consecuencia pueden no ofrecer un apropiado ajuste a las variaciones en el suelo (Sahagun-Castellanos, 1985). En el estado II, los testigos no son ubicados aleatoriamente dentro del bloque, sino a una distancia equivalente en todos los bloques.

Se presenta una discusión sobre los resultados del uso un diseño experimental aumentado en cuatro experimentos del estado II del Programa de Variedades de CENGICAÑA.

OBJETIVOS

1. Establecer la porción de la varianza general de cada uno de los cuatro experimentos que puede ser explicada por los factores incluidos dentro del análisis de variación para cada experimento.
2. Calcular la eficiencia relativa del diseño en cuatro experimentos del estado II de selección Programa de Variedades de CENGICAÑA.

MATERIALES Y MÉTODOS

Localización: Se utilizaron los resultados de medición de Brix en los cuatro ensayos de selección con variedades de la serie CG16 (dos de variedades con flor y dos de variedades sin flor) en dos estratos de altitud diferentes: Sub-estación experimental del estrato medio (estado II con flor, en la finca San Bonifacio ubicada a 213 msnm a 14°15'14.49" N y 91° 1' 44.63" O. El estado II sin flor se ubicó en la finca El Bálsamo, cañaveral “El Seis” del ingenio Pantaleón a 237 msnm, 14°15'38.89" N y 91° 00' 22.36" O) y la Sub-estación experimental del estrato litoral (finca El Retazo, del ingenio Magdalena a 17 msnm, 14° 01' 14.99" N y 91° 06' 34.42" O).

Material vegetal

Se realizaron una serie de mediones de Brix en cada uno de los experimentos distribuidas según el Cuadro 1.

Cuadro1. Material Vegetal por ensayo, se incluyen los testigos

ensayo	Numero de observaciones	testigos
litoral_con_flor	7	CG02-163
	7	CP72-2086
	7	CP73-1547
	185	CG16's
litoral_sin_flor	8	CP72-2086
	8	SP79-1287
	8	SP83-2847
	217	CG16's
medio_con_flor	8	CP72-2086
	8	CP73-1547
	123	CG16's
medio_sin_flor	10	CP72-2086
	191	CG16's

Se realizaron el número indicado de observaciones por testigo por bloque. Para el caso de los materiales CG16's, se realizó solamente una observación en todo el experimento.

Diseño Experimental Aumentado

Se analizaron los experimentos según los procedimientos que indica (Sahagun-Castellanos, 1985) y (Federer & Raghavarao, 1975), de acuerdo con este último autor, el análisis se realizó para un Diseño

Experimental de Bloques Aumentados cuando las variedades testigo estuvieron en un diseño de bloques ligados.

Variable de Respuesta: Lecturas Refractométricas en porcentajes (Brix).

Análisis de la Información.

Análisis y gráficas de residuos y de cuantil-cuantil. Análisis de Variación para un modelo fijo de un diseño experimental aumentado, mediante el paquete de R agricolae (De Mendiburu, 2017). Todos los análisis y figuras fueron hechos en el programa para análisis estadístico R (R Core Team, 2019).

El establecimiento de las eficiencias de los experimentos se realizó contra el valor total de la varianza, ya que no hay ningún otro diseño experimental contra el cual compararlo. Para realizar la comparación se realizó una modificación al método propuesto por Fisher, citado por (Steel & Torrie, 1986).

Eficiencia Relativa del Diseño Experimental Aumentado. Como no hay otro tipo de diseños experimentales con los cuales comparar el que aquí se presenta, se realizó una modificación a la fórmula de eficiencia propuesta por Fisher (Steel & Torrie, 1986). Se utilizó el cálculo directo de la varianza para una muestra para cada experimento. Es decir, se comparó el uso de este diseño contra el no uso de ninguno. La ecuación utilizada para el cálculo de la eficiencia relativa se presenta en la Ecuación 1.

$$ER = \frac{(n_1+1)(n_2+3)S_2^1}{(n_2+1)(n_1+3)S_1^1} \quad (1)$$

Donde, n_1 son los grados de libertad del diseño a probar, n_2 son los grados de libertad de la varianza general y S_1^1 y S_2^1 corresponden al cuadrado medio del error experimental del diseño aumentado y la varianza general respectivamente.

RESULTADOS

En la Figura 1 (a, b, c y d), se presentan los gráficos de cuantil-cuantil (QQ). En estos se aprecia que existe una ligera desviación de los datos con respecto

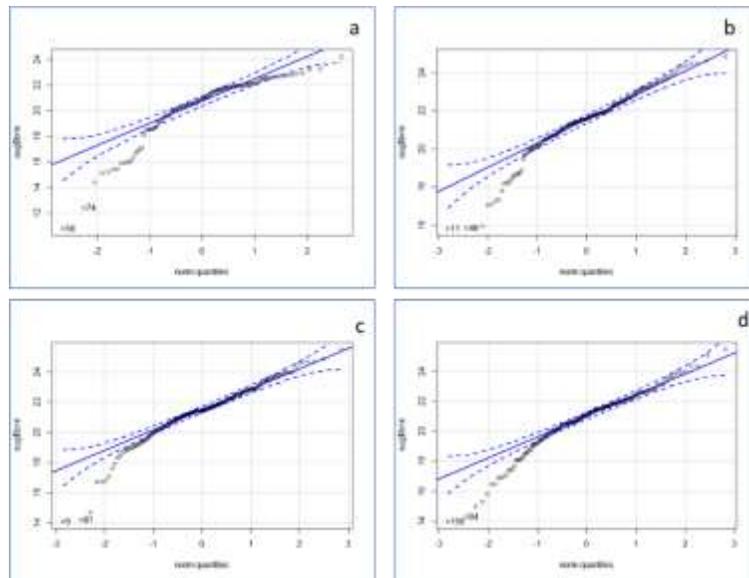


Figura 1. Distribución de los cuantiles normales contra las lecturas de Brix. Experimentos: medio con flor (a), medio sin flor (b), litoral sin flor (c) y litoral con flor (d).

Se observa en la Figura 1 que las distribuciones alcanzan con cierta dificultad la normalidad. Sin embargo, la cantidad de datos es suficientemente alta para permitir una distribución muy cercana a la normal. Se realizaron las transformaciones rutinarias a la variable *Brix*, sin embargo, los resultados fueron esencialmente los mismos, por lo que se optó por usar los valores originales.

En general un gráfico de dispersión de los residuos (desviaciones), mostraron un nivel aceptable de distribución, si bien es posible que estos puedan ser mejorados mediante un muestreo más amplio de los testigos en el número de observaciones aumentadas por bloque (Figura 2).

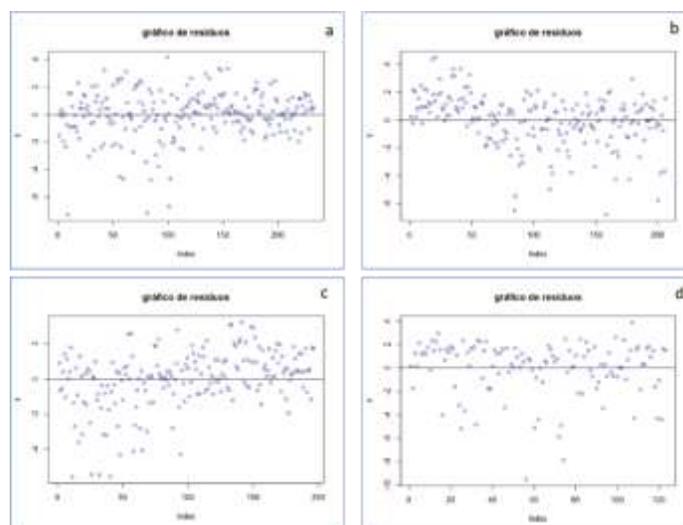


Figura 1. Gráficos de Dispersión de los residuos (desviaciones). Experimentos: litoral sin flor (a), litoral con flor (b), medio sin flor (c) y medio con flor (d)

De los cuatro experimentos el que presenta una distribución menos dispersa es el experimento del estrato medio con variedades con flor. También fue el que contó con el menor número de lecturas de variedades CG y de testigos. En todo caso la dispersión es suficiente para realizar los análisis correspondientes. En general, puede que las distribuciones contengan algunas inconsistencias debido a que emplearon para la medición del Brix dos refractómetros. Si bien ambos son de la misma marca y modelo, es necesario realizar un proceso más exhaustivo de calibración de éstos directamente en el campo a fin de correlacionar lo más alto posible ambas mediciones.

Los valores de las varianzas generales de la variable Brix se presentan en el Cuadro 2 con sus respectivos grados de libertad.

Cuadro 2. Valores de varianza (S²) para cada conjunto de datos (experimentos) calculado a partir de una muestra

experimento	Varianza	Grados de Libertad (n-1)
bajo con flor	3.5057	205
bajo sin flor	2.9633	230
medio con flor	5.4691	122
medio sin flor	2.7954	195

Se aprecia que el experimento de variedades con flor del estrato medio es la que muestra la más alta varianza, lo que es consistente con una dispersión poco uniforme según lo que se observó en la Figura 2. En general este experimento mostró una dispersión muy separada, es decir hubo muchos valores de lecturas en el rango que no fueron encontrados. Esto se debió en parte a que el experimento no mostró un buen desarrollo fenotípico y no fue posible encontrar un buen número de genotipos candidatos a los que se les pudiera medir la concentración de sólidos solubles. En el Cuadro 3 se presenta el resumen de los valores para un análisis de variación para un diseño experimental aumentado con los testigos plantados sistemáticamente, no aleatorizado.

Cuadro 3. Resumen de los Análisis de Variación donde se Ajustaron los tratamientos (genotipos). Para cuatro experimentos del Estado II

Response: Brix Litoral con Flor					
	Df	Sum Sq	Mean Sq	F value	Pr(>F)
block.unadj	6	41.16	6.8597		
trt.adj (Genotipos)	180	640.78	3.5599	1.6806	0.09318
Control (testigo)	3	26.18	8.7276	4.1202	0.02076
Control + control.VS.aug.	177	614.6	3.4723	1.6392	0.10431
Residuals (Error)	19	40.25	2.1183		
				CV=7%	Media=20.86
Response: Brix Litoral Sin Flor					
	Df	Sum Sq	Mean Sq	F value	Pr(>F)
block.unadj	4	8.38	2.0946		
trt.adj (Genotipos)	212	654.43	3.0869	1.8753	0.0798795
Control (testigo)	3	52.28	17.4277	10.587	0.0005432
Control + control.VS.aug.	209	602.15	2.8811	1.7502	0.1052865
Residuals (Error)	15	24.69	1.6461		
				CV=6%	Media=21.37

Response: Brix. Medio con flor					
	Df	Sum Sq	Mean Sq	F value	Pr(>F)
block.unadj	3	15.64	5.2134		
trt.adj (Genotipos)	116	641.42	5.5295	1.6306	0.3923
Control (testigo)	1	1	1.0035	0.2959	0.6243
Control + control.VS.aug.	115	640.42	5.5689	1.6422	0.3895
Residuals (Error)	3	10.17	3.3912		
				CV=9%	Media=20.37
Response: brix. Medio con flor					
	Df	Sum Sq	Mean Sq	F value	Pr(>F)
block.unadj	3	10	3.3335		
trt.adj (Genotipos)	189	530.85	2.8088	1.1951	0.5247
Control (testigo)	1	4.71	4.7073	2.003	0.2519
Control + control.VS.aug.	188	526.15	2.7987	1.1908	0.5264
Residuals (Error)	3	7.05	2.3501		
				CV=7.2%	Media=21.44

Se observa en el cuadro 3 que los valores de probabilidad de encontrar un valor absoluto mayor que F son relativamente altos. Solamente en los experimentos del litoral fue posible encontrar diferencias significativas, aunque para un alfa (α) entre 0.05 y 0.10. Las no significancias encontradas en los experimentos del estrato medio pueden deberse a que el número de lecturas para diferentes valores fue menor en los experimentos del estrato medio, si bien los rangos fueron similares a los que se presentaron en el resto de los experimentos.

Los coeficientes de variación fueron relativamente bajos, lo cual es razonable considerando que la medida del error se hace sobre los genotipos testigos, los cuales varían relativamente poco entre ellos. En todo caso los análisis muestran datos valiosos en términos de los efectos en los que la variación es separada. Es posible que debido a que se utilizaron relativamente pocas lecturas sobre los testigos en cada bloque, entonces las diferencias entre los genotipos, no fue del todo claro. Con un mejor número de testigos (controles) y un mayor número de lecturas de éstos en cada bloque puede ser posible ser más preciso con la medición del error experimental. En este sentido es necesario aumentar el número de observaciones por bloque y al menos contar con tres testigos con lecturas aumentadas en cada bloque.

En todo caso es preciso anotar que la utilización de esta herramienta será de amplia utilidad para la clasificación y jerarquización de los genotipos en evaluación en el Estado II, según las lecturas refractométricas. La selección en general debe ser iniciada por este aspecto y luego seguida por el resto de las variables como la acumulación de biomasa y la reacción a enfermedades, seguidas finalmente por características de manejo agronómico.

El uso de esta herramienta debe ser entendida desde el punto de su eficiencia en términos de explicar la proporción de la varianza debida a cada uno de los factores. En este sentido se realizó un cálculo de la eficiencia relativa de los experimentos.

Cuadro 4. Eficiencia relativa del diseño Experimental aumentado para los cuatro experimentos

	n ₁	s ₁	n ₂	s ₂	eficiencia
bajo con flor	19	2.1183	205	3.5057	3299.22
bajo sin flor	15	1.6461	230	2.9633	1416.99
medio con flor	3	3.3912	122	5.4691	452.36
medio sin flor	3	2.3501	195	2.7954	159.27

Cualquier valor superior a 1 sería un valor que justifica plenamente el uso de este diseño experimental (Steel y Torrie, 1985). Según lo observado en el Cuadro 4, el uso de este experimento se justifica plenamente y su uso es ampliamente mejor a no usar ninguno. En este sentido, se sigue que la implementación de este tipo de análisis ayudará considerablemente a un mejor entendimiento de la distribución de los genotipos según sus propiedades de sólidos solubles.

CONCLUSIONES

El diseño experimental aumentado fue capaz explicar de la varianza en cada experimento a través de los factores establecidos.

No se encontraron diferencias estadísticas entre genotipos. Sin embargo, se observó una tendencia a la diferenciación en los experimentos del estrato litoral, donde los genotipos se expresaron mejor fenotípicamente.

La distribución de los datos, si bien es aceptable, ofrece algunas desviaciones que hay que controlar, posiblemente debido al empleo de dos refractómetros durante la toma de lecturas.

La eficiencia general del uso de un diseño experimental aumentado es alta para todos los experimentos. Lo que justifica su uso en términos de la clasificación de los genotipos en evaluación.

RECOMENDACIONES

1. Continuar con el uso de este diseño experimental en el estado II.
2. Realizar los ajustes necesarios en términos de aumentar el número de lecturas de los testigos en cada bloque para calcular mejor el error experimental.
3. Calibrar constantemente los refractómetros a fin de generar una función que permita alinear ambas lecturas o corregir las lecturas de uno de ellos.
4. Aumentar el número de genotipos candidatos para ser seleccionados en cada experimento, a fin de contribuir con un mejor ajuste del factor Genotipos en el análisis de variación.

REFERENCIAS

- Cullis, B. R., Smith, A. B., & Coombes, N. E. (2006). On the design of early generation variety. *Journal of Agricultural Biological and Environmental Statistics*, 11:381, 25. doi:10.1198/108571106X154443
- De Mendiburu, F. (2017). agricolae: Statistical Procedures for Agricultural Research. R package version 1.2-8. Retrieved from <https://CRAN.R-project.org/package=agricolae>
- Federer, W. T., & Raghavarao, D. (1975). ON AUGMENTED DESIGNS. *Biometrics*, 31, 29-35.

- Montgomery, D. C. (1991). *Diseño y Análisis de Experimentos*. (N. Grepe, Ed., & J. Delgado-Saldivar, Trad.) México: Iberoamérica, S.A.
- Orozco, H., Quemé, J. L., Ovalle, W., & Rosales-Longo, F. (2012). Mejoramiento Genético de la Caña de Azúcar. In CENGICAÑA, *El Cultivo de la Caña de Azúcar en Guatemala* (pp. 45-77). Guatemala: Artemis Edinter.
- R Core Team. (2019). R: A Language and Environment for Statistical Computing. Vienna, Austria. Retrieved from R Foundation for Statistical Computing: <https://www.R-project.org/>
- Rosales-Longo, F., Ovalle, W., García, B., Catalán, M., Campos, J., Portillo, O., & Paz, V. (2018). EVALUACIÓN Y SELECCIÓN DE GENOTIPOS DE CAÑA DE AZÚCAR EN ELESTADO II DEL PROGRAMA DE VARIEDADES DE CENGICAÑA EN LA ZAFRA 2017/2018. In M. Melgar, A. Meneses, H. Orozco, & O. Pérez (Ed.), *Memoria Presentación de resultados de investigación Zafra 2017 - 2018*, (pp. 151-161). Santa Lucía Cotzumalguapa.
- Sahagun-Castellanos, J. (1985). *EFFICIENCY OF AUGMENTED DESIGNS FOR SELECTION*. Iowa State University.
- Steel, R. G., & Torrie, J. H. (1986). *Bioestadística: Principios y Procedimientos* (2 ed.). (R. Martínez, Trad.) México: McGraw-Hill.