

Revisión Bibliográfica

LOS MODELOS LINEALES GENERALIZADOS MIXTOS. SU APLICACIÓN EN EL MEJORAMIENTO DE PLANTAS

Review

Generalized linear mixed models. Its application in plant breeding

Evelyn Bandera Fernández✉ y **Leneidy Pérez Pelea**

ABSTRACT. Frequently, in agricultural research, experimental data do not satisfy the assumptions of general linear models, making the recommended transformations very few effective. Generalized Linear Mixed Models (GLMMs) provide a way of analysis for those data that are correlated and that does not necessarily require that the evaluated variable be normally distributed and but to a distribution of the exponential family (Gamma, Poisson, Binomial among others). The objective of this review was to present the applications of generalized linear mixed models in breeding programs. This model has been focused fundamentally in three directions in plant breeding programs: in the prediction of family behavior, in the estimation of variance components and in multi-environment trials. GLMM estimation procedures make it possible to reduce biases when data is incomplete, unbalanced or adjust scattered data and also allow modeling the structure of errors in data from longitudinal measurements. There are several commercial and free programs such as: the GLIMMIX and GENMOD procedure of the SAS package and the lme4 of the R package that enable the use of The Generalized Linear Mixed Models for most current applications in plant genetics.

RESUMEN. Con frecuencia, en las investigaciones agrícolas los datos experimentales no satisfacen las premisas de los modelos lineales generales y las transformaciones recomendadas tienen poca efectividad. Los Modelos Lineales Generalizados Mixtos (GLMMs, Generalized Linear Mixed Models) proporcionan una vía de análisis para aquellos datos que se encuentran correlacionados y no requiere necesariamente que la variable evaluada se distribuya normalmente, sino a una distribución de la familia exponencial (Gamma, Poisson, Binomial, entre otras). El objetivo de esta revisión fue presentar las aplicaciones de los modelos lineales generalizados mixtos en los programas de mejoramiento genético. Este modelo ha estado enfocado fundamentalmente en tres direcciones en los programas de mejoramiento en plantas: en la predicción del comportamiento familiar, en la estimación de los componentes de varianza y en los ensayos multiambientes. Los procedimientos de estimación de los GLMMs posibilitan reducir los sesgos cuando los datos están incompletos, desbalanceados o ajustar datos dispersos y permiten además modelar la estructura de los errores en datos provenientes de mediciones longitudinales. Existen varios programas comerciales y libres tales como: el procedimiento GLIMMIX y GENMOD del paquete SAS y el lme4 del paquete R, que posibilitan el uso de los Modelos Lineales Generalizados Mixtos para la mayoría de las aplicaciones actuales en la genética de plantas.

Key words: plant breeding, quantitative genetics

Palabras clave: fitomejoramiento, genética cuantitativa

INTRODUCCIÓN

Los Modelos lineales tradicionales surgen en el siglo XX y se basan en estimaciones por

el método de los Cuadrados Mínimos como el Análisis de Varianza (ANOVA), el Análisis de Covarianza (ANACOVA), el Análisis de Varianza Multivariado (MANOVA) y la correlación parcial. En general, emplean variables respuesta continuas y una o varias variables independientes, las que pueden ser variables

de clasificación que dividen las observaciones en diferentes grupos o variables continuas. El objetivo de estos modelos puede ser la comparación de grupos o tratamientos (prueba de hipótesis) o la predicción de una respuesta (variable dependiente) a partir de las variables dependientes (1). Las premisas del ANOVA basado

Departamento Biología Vegetal, Facultad de Biología, Universidad de La Habana. Calle 25 # 455 / I y J, Plaza. La Habana. Cuba.

✉ tati@fbio.uh.cu

en un modelo lineal tradicional son asumidas, en ocasiones sin tener un estricto cumplimiento, con el fundamento de que si el número de observaciones es elevado, los datos pueden llegar a ajustarse a la normalidad.

En ocasiones no es correcto suponer una distribución normal. Por ejemplo, si se trata de conteos o proporciones que son variables discretas; cuando la media de los datos tiene un rango limitado, como ocurre en datos binarios (0,1) y el predictor lineal de la media no se limita a dicho rango. Además, los modelos lineales clásicos ya sean de ANOVA o de regresión suponen que la varianza del error es constante entre las observaciones; sin embargo, hay datos para los cuales la varianza de las observaciones aumenta con el valor de la media (1).

Un caso muy frecuente en estudios genéticos, es el análisis de datos correlacionados provenientes de mediciones repetidas en los mismos individuos, ya sea en tiempo o en espacio. Existen contextos donde no es posible utilizar modelos lineales clásicos para el análisis de varianzas, porque al realizar mediciones repetidas en el tiempo en las mismas unidades experimentales, se incumplen los supuestos de independencia, normalidad, homocedasticidad y linealidad requeridos para su utilización.

Las transformaciones de escala de los datos, constituye una alternativa para lograr el cumplimiento de las premisas de los análisis paramétricos; sin embargo, aunque mejoran el ajuste de la variable, no siempre logran el cumplimiento de las mismas. Esta situación limita la aplicación de los modelos clásicos, por lo que se trata de determinar qué estrategia analítica resulta más apropiada. Si las medidas repetidas en un individuo pertenecen a diferentes

características correlacionadas, entonces se usan análisis multivariados, como el MANOVA. Si las evaluaciones repetidas son de una misma característica o niveles dentro de una de las fuentes de variación, en ocasiones se asumen de forma incorrecta, con distribución normal y se analizan mediante un ANOVA, identificando los individuos y el momento evaluativo. Sin embargo, lo correcto es no suponer normalidad y obtener una mayor eficiencia mediante ecuaciones de estimación generalizada (2).

La mayoría de los ensayos genéticos que se usan en el fitomejoramiento comprenden factores fijos (controlables por el investigador), así como factores aleatorios que varían en su naturaleza en los ambientes de selección durante las etapas del programa de mejora (3). Estas limitantes han contribuido al surgimiento de los modelos mixtos, como modelos lineales generalizados (GLM, *Generalized Linear Models*), y aportan variaciones al modelo lineal tradicional (LM, *Linear Model*), permitiéndole al mismo ser aplicado a un rango más amplio de datos.

Por otra parte, generalmente los datos que son colectados en las investigaciones agrícolas no satisfacen las premisas de los modelos lineales generales; por lo que los modelos lineales generalizados mixtos proporcionan una vía de análisis que no requiere necesariamente distribución normal de las variables, posibilitando que las mismas se puedan ajustar a una distribución de la familia exponencial (4).

A partir de estos antecedentes y considerando la importancia que poseen estos métodos en la actualidad, el objetivo de esta revisión es presentar las aplicaciones de los modelos lineales generalizados mixtos en los programas de mejoramiento genético.

CLASIFICACIÓN

Los tipos de modelos disponibles en la actualidad pueden ser agrupados por el tipo de la variable respuesta o variable dependiente, la cual puede distribuirse según una distribución normal (Modelo Lineal General) o de otro tipo (Modelos Lineales Generalizados) (Tabla 1).

LOS MODELOS LINEALES MIXTOS

Los análisis de modelos mixtos se aplican particularmente a investigaciones que implican factores con pocos niveles, que pueden ser controlados por los investigadores (fijos); así como, a factores con niveles que están fuera del control del investigador (aleatorios) (3).

Los modelos lineales mixtos son una generalización de modelos lineales generales y se emplean cuando (5):

Los efectos son aleatorios: donde el conjunto de valores de una variable de predictor categórico se ven no como el conjunto completo, sino como una muestra aleatoria de todos los valores

Efectos jerárquicos: donde se miden variables predictoras en más de un nivel.

Medidas repetidas: donde las observaciones se correlacionan de forma independiente.

La forma general de un modelo lineal mixto es:

$$Y = Xb + Zu + e$$

donde:

Y es el vector de respuesta (datos), **X** y **Z** son matrices de diseño conocidas, **b** es un vector de parámetros fijos, **u** (efectos aleatorios) y **e** (error) son vectores aleatorios no observables. Las esperanzas matemáticas de **u** y **e**, se asumen igual a cero (6).

Tablas I. Tipos de modelos en función del tipo de variable, características y procedimiento a emplear con el programa SAS

Tipo de modelo	Características	PROC-SAS
Modelos Lineales Generales (GLM, <i>General Linear Models</i>)	Efectos fijos Distribución normal. Estimación por Mínimos Cuadrados (LS) Variable dependiente continua Variable independiente continua o discreta	ANOVA, ANACOVA, Análisis de regresión, MANOVA
Modelos Lineales Generales Mixtos (<i>General Linear Mixed Models</i>)	Efectos fijos y aleatorios Distribución normal. Estimación por máxima verosimilitud (MV), REML, MIVQUEO (por sus siglas en inglés) No requiere una función de ligamiento Variable dependiente continua Variable independiente continua o discreta	MIXED-SAS: admite datos correlacionados y variabilidad inconstante. HPMIXED-SAS
Modelos Lineales Generalizados Mixtos (GLMM, <i>Generalized Linear Mixed Models</i>)	Distribuciones de familia exponencial Estimación de mínimos cuadrados de efectos fijos y aleatorios Requiere una función de ligamiento no-lineal Variable dependiente continua o discreta Variable independiente continua o discreta	GENMOD: efectos fijos GLIMMIX: efectos fijos y aleatorios. Admite datos correlacionados, de variabilidad, inconstante, desbalanceados o incompletos

El algoritmo de los modelos mixtos permite calcular el mejor estimador lineal insesgado (BLUE, *Best Linear Unbiased Estimate*) de efectos fijos y el mejor predictor lineal insesgado (BLUP, *Best Linear Unbiased Predictor*) de los efectos aleatorios. El *BLUP* representa la esperanza condicional de los efectos aleatorios dada a los datos observados, y es también un estimador Bayesiano. El *BLUP* de una combinación lineal de efectos fijos y aleatorios, es la combinación lineal de los *BLUE* de efectos fijos y los *BLUP* de efectos aleatorios (7).

Teóricamente, los *BLUPs* tienen los menores cuadrados medios del error de predicción entre todos los predictores lineales insesgados, proveyendo el modelo asumido sostenido y que los parámetros del modelo sean conocidos (8). Para la estimación de parámetros genéticos, los genetistas recurrieron a la máxima verosimilitud (ML – *Maximum Likelihood*) y asumieron normalidad en los datos experimentales. Para dar una solución al sesgo del estimador de máxima verosimilitud de la varianza residual, fue propuesto el método conocido como máxima verosimilitud “restringida” (REML, *Restricted Maximum Likelihood*) (9).

La metodología de modelos mixtos permite analizar correcta y eficientemente los datos de experimentos con medidas repetidas mediante el modelaje de la estructura de covarianzas, que considera las correlaciones entre medidas repetidas y la presencia de varianzas heterogéneas. No considerar la correlación entre sujetos con la utilización de efectos fijos o modelos mixtos con estructuras de covarianzas muy simples, podría originar que aumente la tasa de error tipo I para la prueba de efectos fijos del modelo. Sin embargo, un modelo muy complicado afectaría la potencia y la eficiencia de la prueba para los efectos fijos (10).

APLICACIONES DE LOS MODELOS LINEALES MIXTOS

La aplicación de los modelos mixtos en los programas de mejoramiento en plantas ha estado enfocada en tres direcciones (11):

- ◆ La predicción del comportamiento familiar
- ◆ La estimación de los componentes de varianza
- ◆ Ensayos multiambientes

Predicción del comportamiento familiar

Los modelos mixtos basados en predicción (BLUP) han sido empleados en diferentes especies vegetales como: soya (12), maíz (13), kiwi (14), papa (15), caña (16); con el objetivo de seleccionar individuos y familias que presenten las características deseadas por el investigador, para ser recomendadas en futuras investigaciones.

Existen evidencias que muestran que una combinación de selección familiar e individual es más eficaz que una selección de familia sola (17). La selección de individuos dentro de cada familia pasa por un proceso de evaluación visual, lo que conlleva a que sea más difícil la obtención a nivel individual (18,19). El uso de la metodología BLUP individual simulado (BLUPIS, *BLUP individual simulated*) permite indicar un número de individuos a ser seleccionados visualmente por familia, número total de clones o número de familias a contribuir con individuos seleccionados (20).

Con el objetivo de determinar el número de individuos seleccionados por familia en caña de azúcar (16), se estimaron

los mejores predictores lineales insesgados a nivel individual (BLUPIS). Estos autores consideraron en su experimento como efectos fijos, los efectos experimentales, que incluyen la media general y como efectos aleatorios los efectos aditivos del genotipo de los individuos, la capacidad específica de combinación y el efecto de los bloques. A partir del análisis de los resultados, determinaron que las familias que estaban formadas por el grupo dos eran las que debían ser recomendadas en el programa de mejoramiento de la caña en Brasil, por tener variedades óptimamente adaptadas a las condiciones de sabana brasileña.

De igual forma, en otro estudio utilizaron la metodología de los modelos mixtos para predecir los efectos genotípicos de cada familia y los valores genotípicos de cada individuo dentro de cada familia (21). Estos autores consideraron en su modelo mixto, la media general, como de efectos fijos y el efecto genotípico de cada familia, como aleatorio. El número óptimo de genotipos fue seleccionado en las mejores familias, obteniendo una alta eficiencia del método BLUPIS.

Estimación de los componentes de varianza

La cantidad de variación del genotipo se mide y expresa mediante la varianza. La varianza fenotípica se descompone en varianza genética y varianza ambiental. A su vez, la varianza genética se descompone en varianza aditiva, de dominancia y de interacción epistática (22). La partición de la varianza en componentes, permite estimar la importancia relativa de los determinantes del fenotipo, en particular, el papel de la herencia frente al ambiente.

El método REML/BLUP de los modelos mixtos, permite estimar los parámetros genéticos, los BLUP y predecir los valores genéticos

aditivos y genotípicos (23). Estos autores seleccionaron las mejores familias de caña de azúcar para caracteres agroindustriales, a partir de los valores genéticos y aditivos obtenidos por esta metodología. Además obtuvieron valores altos de heredabilidad para la selección familiar y valores bajos y medios para la selección individual.

Otros autores, al emplear la metodología REM/BLUP determinaron las familias que presentaban las características de producción esperadas y los mejores individuos para cada una de ellas (24).

Ensayos multiambientes

La respuesta diferencial de un genotipo o cultivar dado a través de diferentes ambientes se conoce como Interacción Genotipo-Ambiente (IGA) (*GEI, Genotype-Environment Interaction*).

Los caracteres más importantes de los cultivos comerciales están controlados por poligenes con varios tipos de efectos genéticos que son afectados por el ambiente. El uso del operador multiplicativo para modelar la interacción genotipo x ambiente ha sido propuesto por varios investigadores (25-29).

Los modelos lineales-bilineales mixtos son útiles para modelar la interacción genotipo-ambiente y estimar las matrices de varianza-covarianza (30). Un modelo lineal-bilineal de efecto mixto para G genotipos, S sitios y R réplicas es:

$$Y = Xb + Zr + Zg + Zge + e$$

Donde: **X**, **Zr**, **Zg** y **Zge** son las matrices de diseño para efectos fijos, de efectos aleatorios replicados dentro de sitios, genotipos, y la interacción, respectivamente, y **e** es el vector de residuos. El vector **b** denota los efectos fijos de los sitios, y los vectores **r**, **g**, **ge** y **e** contienen efectos aleatorios de repeticiones dentro de los sitios, genotipos, interacción

y residuos, respectivamente, y se supone que son aleatorios y normalmente distribuidos con vectores medios nulos y matrices de varianza-covarianza R, G, GE y E, respectivamente. Estas matrices de varianza-covarianza presentan una estructura simple de componente de varianza (30).

El ANOVA regular mixto asume que todos los términos de la interacción genotipo-ambiente tienen la misma varianza y son independientes; aunque algunos autores han empleado un modelo mixto que involucra la diagonal de la matriz G con varianzas heterogéneas (por genotipo) para los términos aleatorios de la interacción (31). Por tanto, el modelo asume que todos los términos genotipo-ambiente que involucran un genotipo particular, tienen la misma varianza genotipo-ambiente, y serán tan diferentes los componentes de varianza genotipo-ambiente, como el número de genotipos. Los componentes de varianza REML, asignables a cada genotipo, estiman los mismos parámetros que la varianza de estabilidad de Shukla (32).

La autora comparó los biplots obtenidos cuando colocó en el m-ésimo eje λ_{mj} , como valores genotípicos, en cada término multiplicativo, y el m-ésimo elemento del escalado EBLUP (x_i) como el valor para el i-ésimo ambiente, contra el biplot tradicional obtenido de un modelo fijo (11). Los biplots bajo ambos enfoques fueron obtenidos para varios ficheros completos de ensayos de variedades. Los procedimientos diferentes para obtener los biplots en ambos enfoques, mostraron el mismo patrón de iteración. Datos de experimentos de campo de plantas de caña, de 2007- 2009 (33) fueron empleados para comparar la exactitud de predicción de varios modelos mixtos contra el enfoque de modelos fijos. Los ensayos de campo regulares

involucraron de 20 genotipos por ensayo. Cada año, los experimentos fueron conducidos en varias granjas (6) comerciales. El experimento fue conducido según un diseño de bloques completamente aleatorizados con tres réplicas. El marco teórico de los modelos lineales mixtos proporcionó la posibilidad de un mejor ajuste de los datos con respecto al que se hubiese obtenido bajo los supuestos clásicos de homogeneidad de varianzas a través de ensayos e independencia temporal y espacial de las observaciones, proveyendo un menor error estándar para la comparaciones de medias y, por ende, una mayor eficiencia en el experimento. El cambio de modelo permitió detectar diferencias estadísticamente significativas entre genotipos, no evidenciadas por el modelo clásico, así como también producir alteraciones en la posición relativa de jerarquía de sus desempeños, lo que condujo a cambios en la interpretación de los datos analizados. Cuando se evalúa la ejecución de genotipos en presencia de una pequeña magnitud de variación genotipo-ambiente bajo un enfoque de modelo fijo (con o sin términos genotipo-ambiente), pudiera esperarse una inflación del error tipo I (34).

En otro estudio se realizó una comparación de la eficiencia entre los métodos Eberhart-Russel (ER),

Modelos de Efectos Principales Aditivos e Interacción Multiplicativa (AMMI) y el modelo mixto (REML/BLUP) (35). Estos autores demostraron que el modelo mixto tiene más sensibilidad en la detección de los efectos provocados por la interacción genotipo-ambiente que los otros métodos comparados.

LOS MODELOS LINEALES GENERALIZADOS MIXTOS (GLMM)

Los GLMs, más usados en la actualidad son los Modelos Lineales Generalizados Mixtos (GLMMs, *Generalized Linear Mixed Models*) que permiten que la media de una población dependa de un predictor lineal a través de una función de enlace (*link function*) de tipo no lineal y que la distribución probabilística de la variable respuesta sea cualquiera de la familia exponencial (Tabla II). De este modo, las variables explicativas pueden ser cualquier combinación de variables continuas, variables de clasificación y sus interacciones. La estimación de parámetros en estos modelos se realiza mediante procedimientos de máxima verosimilitud (36).

Los Modelos Lineales Generalizados Mixtos (GLMM, *Generalized, Linear Mixed Models*) ofrecen nuevas posibilidades en estos casos, pues hacen posible extender los modelos lineales

clásicos de efectos fijos, al incluir efectos aleatorios y predictores BLUPs, para el análisis de datos con distribuciones de la familia exponencial. Además, permiten procesar datos correlacionados al usar efectos aleatorios y estimar sus componentes de varianza asociados al modelo, adicionales a los del error residual. Los procedimientos de estimación usados posibilitan reducir los sesgos cuando los datos están incompletos, desbalanceados o ajustar datos dispersos y permiten además modelar la estructura de los errores en datos provenientes de mediciones longitudinales (37).

Existen programas comerciales y libres tales como: los procedimientos GLIMMIX y GENMOD del paquete SAS y el lme4 del paquete R que posibilitan el uso de los Modelos Lineales Generalizados Mixtos (GLMMs) para la mayoría de las aplicaciones actuales en la genética de plantas. Sin embargo, la masividad de datos provenientes del nivel molecular y de informaciones fenotípicas históricas hace que se esté trabajando en el desarrollo de herramientas capaces de procesar modelos con miles de niveles de efectos fijos y aleatorios (38, 39).

Se han logrado variantes exitosas que reducen más de un 90 % el tiempo de corrida requerido para datos de distribuciones de Poisson, Binomial y Gamma condicional (40).

Tabla II. Funciones de enlace de los Modelos Lineales Generalizados (GLM) más usados, al emplear el paquete SAS

Modelos	Variable respuesta	Distribución de la variable respuesta	Función de enlace SAS/Stat®9.1.)
Modelo Lineal tradicional (LM)	Continua	Normal	Identity: $g(\mu) = \mu$
Regresión logística	Proporción	Binomial	Logit: $g(\mu) = \log(\mu/1-\mu)$
Regresión de Poisson en modelo Log lineal	Conteo	Poisson	Log: $g(\mu) = \log(\mu)$
Modelo Gamma con enlace Log	Continua positiva	Gamma	Log: $g(\mu) = \log(\mu)$

CONCLUSIONES

Los modelos mixtos han sido ampliamente utilizados en los estudios de genética en plantas, y los modelos mixtos generalizados han permitido ampliar el espectro de trabajo con aquellas variables que por su naturaleza, no se ajustan a una distribución normal y sus datos se encuentran correlacionados. En la actualidad, su utilización requiere de programas computacionales de gran complejidad, pero presentan grandes ventajas con respecto a los métodos clásicos, ya que sus procedimientos de estimación reducen los sesgos cuando los datos están incompletos, desbalanceados o ajustan datos dispersos.

BIBLIOGRAFÍA

- Zar J H. Biostatistical Analysis. 5ta ed. New Jersey: Ediciones Pearson Prentice Hall; 2009. 189-449 p.
- Gómez S, Torres V, García Y, Fraga LM, Sarduy L, Savón LL. Comparación de modelos de efectos fijos y mixto en el análisis de un experimento con cepas mutantes de hongos celulolíticos *Trichoderma viride*. Revista Cubana de Ciencia Agrícola. 2012;46(2):127-31.
- Edward E, Gbur E, Walter W, Stroup W, McCarter K, Durham S, et al. Analysis of generalized linear mixed models in the agricultural and natural resources sciences. Madison, Wis: American Society of Agronomy; 2012. 277 p.
- Wang T, He P, Ahn KW, Wang X, Ghosh S, Laud P. A re-formulation of generalized linear mixed models to fit family data in genetic association studies. Frontiers in Genetics [Internet]. 2015 [citado 16 de febrero de 2018];6. doi:10.3389/fgene.2015.00120
- Garson GD. Hierarchical linear modeling: Guide and applications. Sage; 2012.
- Laird NM, Ware JH. Random-effects models for longitudinal data. Biometrics. 1982;38(4):963-74. doi:10.2307/2529876
- Piepho HP, Möhring J, Melchinger AE, Büchse A. BLUP for phenotypic selection in plant breeding and variety testing. Euphytica. 2008;161(1):209-28. doi:10.1007/s10681-007-9449-8
- Searle MP, Waters DJ, Rex DC, Wilson RN. Pressure, temperature and time constraints on Himalayan metamorphism from eastern Kashmir and western Zaskar. Journal of the Geological Society. 1992;149(5):753-73. doi:10.1144/gsjgs.149.5.0753
- Patterson HD, Thompson R. Recovery of inter-block information when block sizes are unequal. Biometrika. 1971;58(3):545-54. doi:10.1093/biomet/58.3.545
- Vallejo G, Fernández MP, Tuero E, Livacic-Rojas PE. Análisis de medidas repetidas usando métodos de remuestreo. Anales de psicología [Internet]. 2010 [citado 16 de febrero de 2018];26(2). Disponible en: <http://www.redalyc.org/resumen.oa?id=16713079025>
- Balzarini M. Applications of mixed models in plant breeding. En: Quantitative genetics, genomics and plant breeding. CAB International; 2002.
- Shu YJ, Yu DS, Wang D, Bai X, Zhu YM, Guo CH. Genomic selection of seed weight based on low-density SCAR markers in soybean. Genetics and Molecular Research. 2013;12(3):2178-88. doi:10.4238/2013.July.3.2
- Massman JM, Gordillo A, Lorenzana RE, Bernardo R. Genomewide predictions from maize single-cross data. Theoretical and applied genetics. 2013;126(1):13-22. doi:10.1007/s00122-012-1955-y
- Silva NHD, Gea L, Lowe R. Genetic analysis of resistance to *Pseudomonas syringae* pv. actinidiae (Psa) in a kiwifruit progeny test: an application of generalised linear mixed models (GLMMs). SpringerPlus. 2014;3(1):547. doi:10.1186/2193-1801-3-547
- Ticona-Benavente CA, Da DSF. Comparison of BLUE and BLUP/REML in the selection of clones and families of potato (*Solanum tuberosum*). Genetics and molecular research : GMR. 2015;14(4):18421-30. doi:10.4238/2015.December.23.30
- Barbosa MHP, Ferreira A, Peixoto LA, Resende MDV, Nascimento M, Silva FF. Selection of sugar cane families by using BLUP and multi-diverse analyses for planting in the Brazilian savannah. Genetics and Molecular Research. 2014;13(1):1619-26. doi:10.4238/2014.March.12.14
- Stringer JK, Cox MC, Atkin FC, Wei X, Hogarth DM. Family selection improves the efficiency and effectiveness of selecting original seedlings and parents. sugar Tech. 2011;13(1):36-41. doi:10.1007/s12355-011-0073-5
- Resende MDV de, Barbosa MHP. Selection via simulated individual BLUP based on family genotypic effects in sugarcane. Pesquisa Agropecuária Brasileira. 2006;41(3):421-9. doi:10.1590/S0100-204X2006000300008
- Pedrozo CÂ, Pereira Barbosa MH, Vilela Resende MD, Peternelli LA, de Almeida Costa PM, Lopes da Silva F. Eficiência da seleção em fases iniciais do melhoramento da cana-de-açúcar. Revista Ceres [Internet]. 2008 [citado 16 de febrero de 2018];55(1). Disponible en: <http://www.redalyc.org/resumen.oa?id=305226830002>
- Oliveira RA, Daros E, Resende MDV, Bessalho-Filho JC, Zambon JLC, De Souza T R. Procedimento Blupis e seleção massal em cana-de-açúcar. Bragantia. 2011;70(4):1-5.
- López S F, Pereira BM, Vilela R M, Paternelli LA, Pedrosa C A. Efficiency of selection within sugarcane families via simulated individual BLUP. Crop Breeding and Applied Biotechnology. 2015;15(1):1-9.
- Falconer DS. Introduction to quantitative genetics. 1989 [citado 16 de febrero de 2018]; Disponible en: <http://agris.fao.org/agris-search/search.do?recordID=XF2016003336>
- Costa K, Carvalho I, Silva JWD, Cunha J, Silva JD, Teixeira JDS. Escolha de famílias em cana-de-açúcar via modelos mistos. Agropecu. Científica no Semi-Árido, Patos. 2013;9(1):7-13.

24. Mauro AX, Perecin D, Rodrigues Toledo K, Guimaraes de Andrade LM, Cese AF. Seleção de famílias e progênies de irmãos completos de cana-de-açúcar para atributos tecnológicos e de produção pelo método de REML/BLUP. *Bragantia*. 2014;73(3):253-62. doi:10.1590/1678-4499.0193
25. Piepho HP, Möhring J. Best linear unbiased prediction of cultivar effects for subdivided target regions. *Crop Science*. 2005;45(3):1151-9. doi:10.2135/cropsci2004.0398
26. Crossa J, Burgueño J, Cornelius PL, McLaren G, Trethowan R, Krishnamachari A. Modeling genotype \times environment interaction using additive genetic covariances of relatives for predicting breeding values of wheat genotypes. *Crop Science*. 2006;46(4):1722-33. doi:10.2135/cropsci2005.11-0427
27. Crossa J, Yang R-C, Cornelius PL. Studying crossover genotype \times environment interaction using linear-bilinear models and mixed models. *Journal of Agricultural, Biological, and Environmental Statistics*. 2004;9(3):362-80. doi:10.1198/108571104X4423
28. Burgueño J, Crossa J, Cornelius PL, Yang R-C. Using FactorAnalytic models for joining environments and genotypes without crossover genotype \times environment interaction. *Crop Science*. 2008;48(4):1291-305. doi:10.2135/cropsci2007.11.0632
29. Burgueño J, Crossa J, Cotes JM, Vicente FS, Das B. Prediction assessment of linear mixed models for multi-environment trials. *Crop Science*. 2011;51(3):944-54. doi:10.2135/cropsci2010.07.0403
30. Crossa J, Vargas M, Cossani M, Alvarado G, Burgueño J, Mathew K. Evaluation and interpretation of interactions. *Agronomy Journal*. 2013;105:1-12. doi:10.2134/agronj2012.0491
31. Magari R, Kang MS. SAS-STABLE: Stability analyses of balanced and unbalanced data. *Agronomy Journal*. 1997;89(6):929-32. doi:10.2134/agronj1997.00021962008900060013x
32. Shukla GK. Some statistical aspects of partitioning genotype environmental components of variability. *Heredity*. 1972;29(2):237-45.
33. Ostengo S, Cuenya MI, Balzarini M. Modelos mixtos en el análisis de ensayos comparativos de rendimiento multiambientales de una especie perenne. *Revista industrial y agrícola de Tucumán*. 2013;90(2):53-7.
34. Stroup WW. Mixed model issues in the analysis of multi-location trials. En *University of California at Berkley*; 2000. p. 135-42.
35. Moraes FG, Dilemardo PD. Mixed Model, AMMI and Eberhart-Russel Comparison via simulation on genotype \times environment interaction study in sugarcane. *Applied Mathematics*. 2014;05(14):2107-19. doi:10.4236/am.2014.514205
36. Institute SAS. SAS/IML 9.3 user's guide. Sas Institute; 2011.
37. Stroup W. Generalized linear mixed models: modern concepts, methods and applications. En: *Texts in Statistical Science*. 2012. (Chapman & Hall/CRC Texts in Statistical Science).
38. SAS Institute Inc. 100 SAS Campus Drive Cary [Internet]. USA; 2015. Disponible en: http://www.sas.com/en_us/home.html
39. Bates D, Maechler M, Bolker B, Walker S. lme4: Linear mixed-effects models using Eigen and S4. R package version. 2014;1(7):1-23.
40. Xie L, Madden LV. %HPGLIMMIX&58; A High-Performance SAS Macro for GLMM Estimation. *Journal of Statistical Software*. 2014;58(1):1-25

Recibido: 10 de marzo de 2017

Aceptado: 14 de diciembre de 2017