

Artículo científico

Modelación estadística de la emergencia de tallos de caña de azúcar**Statistical modelling of sugar cane stems emergency**N. del V. Ortiz^{1*}; P.A. Digonzelli^{1,2}; M.B. García^{1,2}; O.E.A. Arce¹; E. Romero^{1,2}

¹Facultad de Agronomía y Zootecnia, Universidad Nacional de Tucumán, Argentina. Avda. Kirchner 1900, (4000), San Miguel de Tucumán, Argentina. *E-mail: ohne265@gmail.com

²Estación Experimental Agroindustrial Obispos Colombres, William Cross 3150, (4101), Las Talitas, Tucumán, Argentina.

Resumen

Se realizó el modelado estadístico de la emergencia de tallos de dos variedades de caña de azúcar (CP 65-357 y LCP 85-384) provenientes de caña semilla saneada por dos metodologías: micro propagación e hidro termoterapia. El objetivo de este trabajo fue la modelación de curvas de crecimiento mediante el ajuste no lineal logístico para evaluar el efecto de las variedades, de las metodologías de saneamiento y de su interacción, sobre la evolución temporal del número de tallos. Se realizaron tres repeticiones por metodología y variedad para cada repetición se usó una parcela que se consideró como unidad experimental. A partir del análisis descriptivo de los datos se seleccionó el ajuste. Inicialmente se aplicó un modelo no lineal de efectos fijos, pero debido a la varianza no constante de los datos se aplicó modelación con efectos aleatorios para datos longitudinales donde la unidad o sujeto fue la parcela. Los parámetros correspondientes a la asíntota horizontal y al valor medio de crecimiento se consideraron como efectos aleatorios con el objeto de disminuir la variabilidad. Los resultados mostraron significación en algunas de las estimaciones de los parámetros de la regresión logística. La asíntota horizontal y el número de días necesarios, desde la plantación, para alcanzar el 50 % de la brotación máxima fueron significativas para variedad, mientras que la metodología y la interacción no resultaron significativas. Se analizaron los supuestos distribucionales de los residuos y de los efectos aleatorios que resultaron adecuados y se verificó la varianza constante de los residuos. Se concluye que el modelo propuesto explica adecuadamente el comportamiento de los datos del porcentaje de emergencia de tallos por metro.

Palabras clave: Datos longitudinales; Regresión logística; Modelos no lineales; Modelos mixtos; R.

Abstract

Growth curve analysis was used to explain the emergence of stems of two varieties (CP 65-357 y LCP 85-384) of healthy sugar cane seed obtained by two methodologies: micropropagation and hydrothermotherapy. The aim of this study was to model the effects of varieties, methodologies used and the interaction between them on the evolution of the number of stems per meter. Three replicates per variety and methodology were carry out and, for each replicate, the plot was considered as the test unit. From the descriptive analysis of the data, a logistic regression fit was selected. Initially, fixed models were applied, but due to the non-constant variance of the data nonlinear modeling with random effects for longitudinal data was applied where the plot was considered to be the unit or subject. Parameters corresponding to the horizontal asymptote and the average value of the horizontal asymptote were considered as random effects in order to reduce variability. The results showed significance in the logistic regression parameters. The horizontal asymptote and the number of days necessary to reach 50 % of maximum emergency were significant for variety, while the methodology and interaction were not significant. The assumptions of normality and constant variance on the residuals were analyzed and turned out to be adequate. As a conclusion, it can be proposed that the logistic regression model adequately explain the behavior of emergency data stems per meter.

Key words: Longitudinal data; Logistic regression; Non-linear models; Mixed models; R.

Recibido 30/10/15; Aceptado 23/11/15.

Los autores declaran no tener conflicto de intereses.

Introducción

La caña semilla que no posea enfermedades ni plagas o presente baja incidencia de ellas brinda la posibilidad de aumentar la productividad del cañaveral. Es por ello que resulta necesario que su producción sea de alta calidad, extremando los cuidados de forma de no ocasionar la difusión de enfermedades sistémicas que se propagan con cada corte y así evitar la necesidad de hacer una frecuente renovación del cañaveral. Para sanear la caña semilla se emplean dos metodologías alternativas que son: micro propagación e hidro termoterapia (Digonzelli *et al.*, 2009, 2010).

Los datos obtenidos de unidades experimentales, seleccionadas al azar, de una población medida en el tiempo reciben el nombre de datos longitudinales. Los datos longitudinales son mediciones repetidas sobre unidades observacionales sin la creación de condiciones experimentales por el experimentador. Esas unidades son a menudo los llamados sujetos o *clusters*. En la ausencia de la aleatorización los efectos mixtos aparecen porque algunos parámetros del modelo, tales como interceptos y pendientes, se asumen que varían al azar de sujeto a sujeto mientras otros parámetros permanecen constantes a través de los sujetos.

El perfil de las medidas observadas de cualquier unidad puede ser una respuesta a los procesos estocásticos variables en el tiempo que operan dentro de esa unidad y que resulta en una correlación entre los pares de mediciones en la misma unidad dependientes de la separación en el tiempo (Diggle *et al.*, 2002). Esta variabilidad refleja la heterogeneidad debida a factores no medidos, como factores ambientales. Esta heterogeneidad natural a través de los individuos se manifiesta en los coeficientes de la regresión. La correlación entre observaciones para un sujeto o unidad experimental aparece desde variables no observables. En el análisis de datos longitudinales es necesario entender al menos cualitativamente cuáles son las probables fuentes de variación aleatoria (Diggle *et al.*, 2002).

El análisis de datos longitudinales tiene dos enfoques que se expresan como modelos de población promedio y modelos de sujeto específico (Zeger *et al.*, 1988). Los modelos de población promedio se refieren a la expectativa de la variable respuesta como el interés principal de análisis. En modelos de sujeto específico el principal enfoque es la modelación de cambios en una respuesta individual,

lo cual se logra introduciendo efectos aleatorios en el modelo. Estos modelos asumen que “(a) existe un modelo de regresión para cada individuo y (b) los coeficientes de regresión están aleatoriamente distribuidos” (Vonesh y Chinchilli, 1997). Los coeficientes de regresión del modelo sujeto específico describen la curva respuesta de los individuos o unidades experimentales. Se modela la heterogeneidad individual usando efectos aleatorios y esos efectos aleatorios determinan parcialmente la estructura de varianza – covarianza. Este tipo de modelos permite explícitamente tener en cuenta la heterogeneidad unidad a unidad a través del uso de efectos aleatorios del sujeto específico (Zeger *et al.*, 1988). En este trabajo el sujeto específico es una parcela.

Las particularidades mencionadas en este tipo de datos, muy comunes en actividades experimentales agropecuarias, hacen que las metodologías de análisis tradicionales no puedan ser utilizadas. Entre las numerosas propuestas sobre la modelación de estos datos se encuentran los modelos mixtos lineales y no lineales. Los modelos lineales para efectos fijos no alcanzan a modelar la variabilidad proveniente de efectos aleatorios introducidos por la correlación entre mediciones en una misma unidad experimental y la heterocedasticidad de las variables error del modelo. Diversos autores desarrollaron la metodología de modelos mixtos lineales y no lineales aplicados a diversas situaciones experimentales en las que los supuestos básicos de las metodologías tradicionales no se cumplen, entre ellos se mencionan a Laird y Ware (1982), Zeger *et al.* (1988), Bates y Watts (1988), Glasbey (1988), Davidian y Giltinan (1995), Kshirsagar y Smith (1995), Vonesh y Chinchilli (1997), Little *et al.* (2000), Pinheiro y Bates (2000), Venables y Ripley (2002), Diggle *et al.*, (2002), Twisk (2003), Davidian (2005), Faraway (2006), McCulloch *et al.* (2008), Wu (2010).

Dentro de modelos mixtos lineales y no lineales que modelan datos longitudinales, se encuentra el desarrollo de curvas de crecimiento. En este estudio, cada parcela se considera como un grupo de medidas correlacionadas entre sí y, además, las observaciones entre grupos son independientes (Wu, 2010). En esta situación las metodologías tradicionales subestiman o sobrestiman los errores estándares de las medias (Cochran, 1947). Cuando se elige un modelo para describir cómo una variable respuesta es explicada por covariables existe la opción de seleccionar modelos mixtos lineales

o no lineales. Cuando se elige un ajuste no lineal se asume un modelo para el mecanismo que genera la respuesta y, en consecuencia, los parámetros tienen una interpretación natural física. En cambio, aun cuando se obtenga un buen ajuste para los datos, los coeficientes estimados por modelos lineales mixtos carecen de interpretación física (Pinheiro y Bates, 2000).

Para la aplicación de esta metodología se desarrollaron paquetes estadísticos, entre otros, R, S-plus, SAS, STATA. Para modelos no lineales Pinheiro y Bates (2000) desarrollaron el programa *nlme* (Pinheiro *et al.*, 2013) en el paquete R (R Core Team, 2013). La bibliografía que proporciona descripciones detallada está dada por Pinheiro y Bates (2000) y Venables y Ripley (2002). Otros autores que tratan sobre modelos no lineales con R, son Maindonald (2003), Crawley (2007), Sarkar (2008), West *et al.* (2015), entre otros.

En este trabajo se plantea la modelación de curvas de crecimiento mediante el ajuste no lineal de efectos mixtos, para evaluar el efecto de las variedades, de las metodologías usadas para obtener semilla saneada y de su interacción, sobre la evolución temporal del número de tallos.

Materiales y métodos

La información se obtuvo de un ensayo realizado en el marco del Subprograma Agronomía de Caña de Azúcar de la Estación Experimental Agroindustrial Obispo Colombes de Tucumán (EEAOC). Dicha información consistió en el número de tallos medido en intervalos diferentes de tiempo. La parcela experimental estuvo constituida por 5 surcos convencionales de 3 metros de longitud. Las evaluaciones se realizaron en los tres surcos centrales de cada parcela, dejando una bordura de 0,50 m en cada extremo de la parcela, es decir, se evaluaron 6 metros lineales de surco por parcela.

Se realizó una plantación invernal en agosto, con una densidad de plantación de 15 yemas por metro lineal de surco. Con intervalos de 3 a 5 días se contó el número de brotes emergidos definiendo como tal aquel que tenía al momento de evaluación por lo menos una hoja verde con lígula visible. Los conteos se realizaron en 12 parcelas desde los 31 hasta los 60 días desde la plantación, momento en que finalizó la fase de emergencia, y se lograron 9 observaciones para cada parcela. Los factores evaluados incluyeron dos variedades: LCP 85-384 y CP 65-357 y dos orígenes de la caña semilla saneada: micro propagada y termo tratada libre de RSD o raquitismo de la caña soca (*Leifsonia xyli* subsp. *xyli*) en ambos casos.

El arreglo de las unidades experimentales correspondió a un experimento factorial 2 x 2 completamente aleatorizado con tres repeticiones. La variable evaluada fue el porcentaje de tallos.

En la Tabla 1 se detallan las variedades y los tratamientos por parcela con el objeto de brindar específica explicación del ensayo y facilitar la interpretación de las Figuras 1 y 2.

La metodología usada fue modelos mixtos no lineales para datos longitudinales. Los datos son porcentajes por lo que admiten la modelación de una regresión logística (Crawley, 2007). La modelación no lineal incluye el análisis de curvas de crecimiento para el que se requiere modelar la variable respuesta como una función no lineal en los parámetros. Los modelos mixtos no lineales asumen errores con distribución normal para describir el perfil de la respuesta individual y tener en cuenta la heterogeneidad dentro y entre sujetos. En general, para modelos de medidas repetidas cada sujeto, la parcela en este caso, puede expresarse de la siguiente forma:

$$y_i = f(\mathbf{X}_i, \mathbf{a}_i, \beta_i) + \varepsilon_i \quad \text{con } i = 1, \dots, n$$

Donde $\mathbf{y}_i = [y_{i1} \dots y_{ip}]$ es el vector de medidas

Tabla 1. Detalle de variedades y tratamientos

Niveles por tratamiento y parcela de cada variedad	
Variedad CP65-357	Variedad LCP85-384
P1: CP65-357.Micropropagada	P3: LCP85-384.Micropropagada
P2: CP65-357.Termotrada	P4: LCP85-384.Termotrada
P5: CP65-357.Micropropagada	P7: LCP85-384.Micropropagada
P6: CP65-357.Termotrada	P8: LCP85-384.Termotrada
P9: CP65-357.Micropropagada	P11: LCP85-384.Micropropagada
P10: CP65-357.Termotrada	P12: LCP85-384.Termotrada
Pi es la parcela i-ésima, con $i=1, \dots, 12$	

repetidas para cada parcela que por tratarse de un ensayo balanceado es de dimensión $p \times 1$, es decir que son n unidades con p observaciones.

X_i es la matriz de diseño, de dimensión $p \times t$ con $t =$ intervalos de tiempo, de variables dentro de la unidad

a_i es un vector $q \times 1$ de covariables entre unidades.

β_i es un vector de parámetros de dimensión $r \times 1$ para la i -ésima unidad.

f es una función específica de (X_i, a_i, β_i) .

ε_i es vector aleatorio de dimensión $p \times 1$.

Cada y_{ij} puede expresarse como

$$y_{ij} = f(x_{ij}, \beta_i) + \varepsilon_{ij}$$

Donde $i = 1, \dots, n_i$ indexa las parcelas y $j = 1, \dots, n_i$ son las ocasiones dentro de parcelas

De esta forma, según lo propuesto por Diggle *et al.* (2002), se tiene:

a) Estructuras de correlación para el error: las desviaciones ε_{ij} con $j = 1, \dots, n_i$ dentro del i -ésimo sujeto pueden estar correlacionadas.

b) Efectos aleatorios no lineales: uno o más parámetros β de la regresión pueden ser modelados como perturbaciones estocásticas del sujeto específico de un valor de la población promediada. Sea el vector B_i realizado independientemente para cada sujeto con distribución normal multivariada, con media y matriz de varianza V_β .

En modelos lineales esos dos casos se pueden tratar como uno solo, porque en ese contexto el supuesto de parámetros sujeto específico deja la respuesta media poblacional sin cambio y afecta solamente la forma de la estructura de la covarian-

za de los Y_{ij} . En cambio, para modelos no lineales los dos casos tienen diferentes implicaciones para el análisis estadístico y para la interpretación de los parámetros del modelo.

Para la selección de la función que permite la modelación de los datos se realizó un análisis descriptivo. La Figura 1 muestra el comportamiento de los datos por parcela y la Figura 2 por metodología empleada de saneamiento.

En consecuencia, la curva seleccionada fue una regresión logística cuya forma, de acuerdo a Pinheiro y Bates (2000), es

$$y(x) = \frac{\beta_1}{\{1 + \exp[(\beta_2 - x)/\beta_3]\}}$$

Donde:

β_1 : asíntota horizontal para $x \rightarrow \infty$, siendo nula para $x \rightarrow -\infty$ (en adelante Asym). Físicamente se interpreta como el valor de porcentaje máximo de emergencia de tallos alcanzado.

β_2 : valor para el cual $y = \beta_1 / 2$ (en adelante x_{mid}). Se interpreta como el tiempo necesario para alcanza la mitad del porcentaje máximo de emergencia. Es el punto de inflexión de la curva

β_3 : parámetro de escala que representa la distancia sobre el eje x entre el punto de inflexión y el punto donde la respuesta es $\beta_1 / (1+e^{-1}) \cong 0,73 \beta_1$ (en adelante $scal$).

El ajuste del modelo se realizó en base al cumplimiento de los supuestos referidos a los efectos aleatorios (normalidad e independencia) y a los residuos (normalidad, independencia, media nula y varianza del error constante). Se seleccionó el

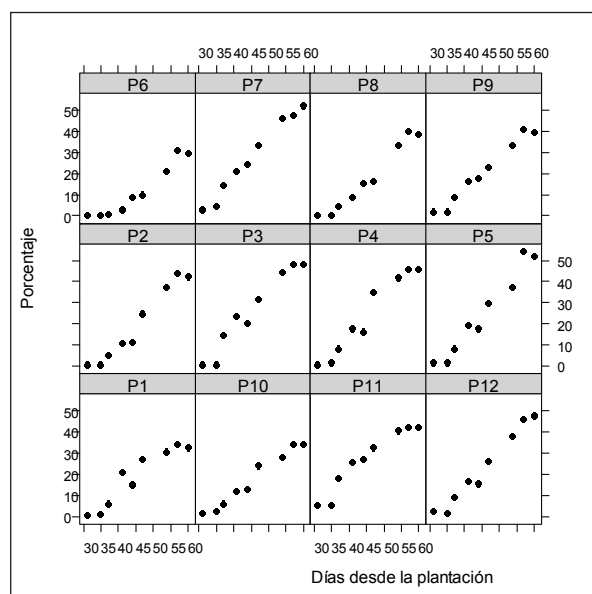


Figura 1. Porcentaje de emergencia de tallos por parcela.

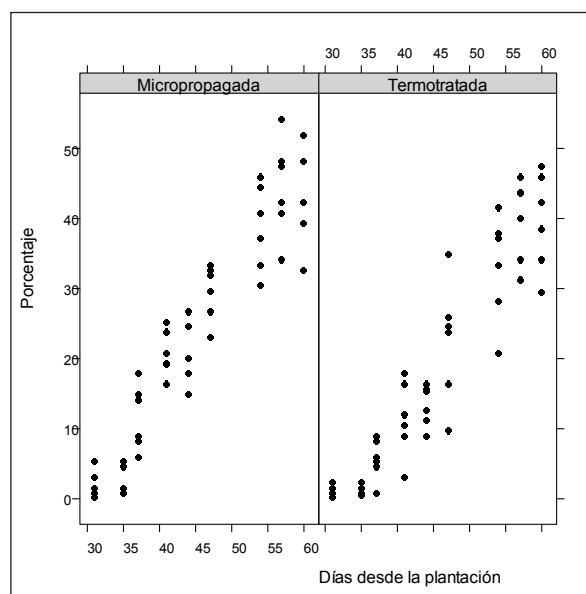


Figura 2. Porcentaje de emergencia por metodología de saneamiento.

mejor modelo, es decir, el del menor número de parámetros estimados que cumpla con los supuestos. La comparación de modelos alternativos se resolvió mediante el uso del test de razón de verosimilitud, como muestra Pinheiro y Bates (2000).

En este trabajo la variable aleatoria fue días desde la plantación dado que las observaciones se realizaron en intervalos diferentes debido a condiciones climáticas, por lo que el ajuste fue

$$y = \frac{\text{Asym}}{1 + \exp[(x_{\text{mid}} - \text{día}) / \text{scal}]}$$

con el que se obtuvieron las estimaciones de los parámetros Asym, xmid y scal.

El modelo de efectos fijos se realizó actualizando el modelo logístico mediante el ajuste de la variable respuesta (Asym + xmid) y como covariable a la interacción entre variedad y metodología de saneamiento.

Para el análisis de los datos se usó el paquete *nlme* (Pinheiro *et al.*, 2013) del software R (R Core Team, 2013). Se trabajó con un nivel de significancia del 5%.

Resultados y discusión

El ajuste mediante una regresión logística de efectos fijos, por parcela, permitió obtener los valores iniciales de los parámetros para el modelo (Tabla 2). Del gráfico de diagnóstico (no se muestra) de los residuos entre parcelas no fue posible asumir varianza constante por lo que se modeló usando una función de varianza que es una potencia del valor absoluto de la covariable de la varianza (Pinheiro y Bates, 2000).

La matriz de varianzas y covarianzas mostró alta correlación entre las estimaciones de la asíntota y el parámetro de escala lo que indicó sobreparametrización en la estructura de efectos aleatorios (Pinheiro y Bates, 2000). En el ajuste la correlación estimada entre Asym y scal resultó ser 0,994. El gráfico de diagnóstico se muestra en la Figura 3 en el cual es posible apreciar el alineamiento entre los efectos aleatorios de la asíntota horizontal, Asym, y los efectos aleatorios de scal por lo que el modelo resultó sobreparametrizado.

Se graficaron los intervalos con el 95% de confianza para los estimadores de Asym, xmid y scal, en donde se observó la variación de las estimaciones por parcela (Figura 4). Si los intervalos están traslapados indica poca variación de lo contrario se debe modelar la variación usando efectos alea-

torios en las estimaciones de los parámetros (Pinheiro y Bates, 2000).

Tabla 2: Valores iniciales de los parámetros estimados. E.e.: error estándar, t: valor del estadístico.

	Asym			
	Estimación	E. e.	t	Pr(> t)
P6	35,47759	8,591938	4,129172	9,056177e-04
P1	32,95917	2,265248	14,549915	2,500077e-05
P10	35,18335	3,604600	9,760679	7,088027e-06
P8	43,48897	5,723730	7,598012	3,451959e-05
P9	42,85780	4,798904	8,930747	3,995157e-05
P11	42,98004	2,626383	16,364724	2,447466e-06
P2	45,06090	3,926462	11,476210	3,673885e-06
P4	46,85014	2,754111	17,010985	5,555943e-06
P12	53,61680	7,141122	7,508176	1,264082e-04
P3	50,80781	3,661594	13,875873	8,250533e-05
P7	53,25820	3,970546	13,413321	1,766098e-06
P5	59,74705	7,575311	7,887076	1,695998e-03

	xmid			
	Estimación	E. e.	t	Pr(> t)
P6	51,12979	3,3789996	15,13164	5,386060e-07
P1	42,23430	1,1073271	38,14076	8,498375e-08
P10	45,10051	1,7271921	26,11204	2,062414e-08
P8	48,17462	2,1029090	22,90856	5,149397e-08
P9	45,64428	1,9854341	22,98957	1,538951e-07
P11	40,60193	1,1090385	36,61002	2,036056e-08
P2	46,97339	1,3743834	34,17780	5,554274e-09
P4	44,34295	0,9411802	47,11420	1,307134e-08
P12	47,88082	2,3212882	20,62683	3,494026e-07
P3	44,21886	1,2676750	34,88186	3,778012e-07
P7	44,25030	1,3382831	33,06497	8,186192e-09
P5	48,08782	2,2206650	21,65469	6,052831e-06

	scal			
	Estimación	E.e.	t	Pr(> t)
P6	4,644946	1,6392577	2,833567	0,0058977430
P1	3,712039	0,9532046	3,894273	0,0211609065
P10	4,638361	1,3249052	3,500900	0,0021010057
P8	4,770077	1,2726307	3,748202	0,0016643030
P9	5,450689	1,3666618	3,988324	0,0031090281
P11	4,849602	0,9967162	4,865580	0,0021719216
P2	4,054457	0,9927767	4,083957	0,0012342473
P4	3,942028	0,7984066	4,937368	0,0048212850
P12	5,839479	1,3144290	4,442598	0,0021051572
P3	5,060372	0,9825994	5,149985	0,0130229818
P7	5,350944	1,0141448	5,276311	0,0003736353
P5	5,969173	1,2263963	4,867247	0,0160052336

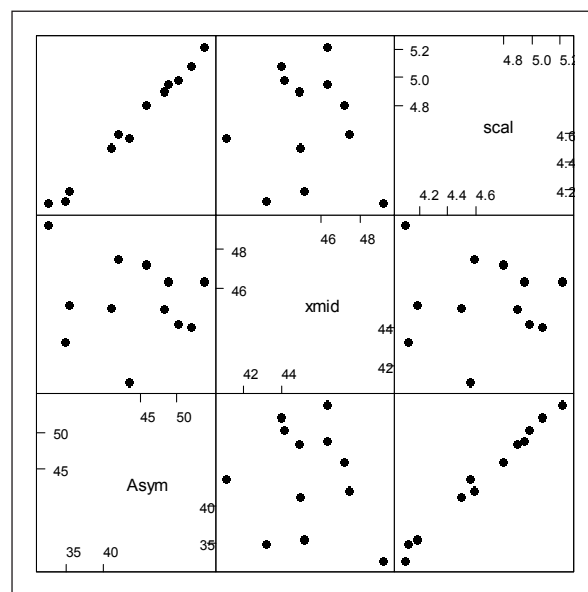


Figura 3. Matriz de estimaciones de los parámetros.

Los intervalos para la estimación de scal se mostraron todos traslapados (Figura 4) y además sus errores estándar por parcela resultaron pequeños (Tabla 2), en consecuencia, se puede considerar como efecto fijo, como lo establece Pinheiro y Bates (2000).

Los intervalos para Asym mostraron tendencia y en el caso de la estimación de xmid se observó que algunos de los intervalos no resultaron traslapados (Figura 4). Ambas estimaciones presentaron mayores errores estándar respecto a la estimación del parámetro scal, destacándose Asym (Tabla 2). En consecuencia, Asym y xmid se modelaron con efectos aleatorios según lo establecido por Pinheiro y Bates (2000).

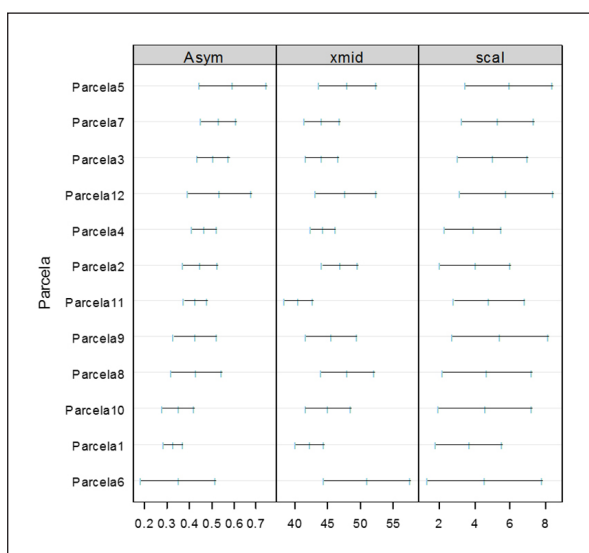


Figura 4. Intervalos de los parámetros por parcela.

En la comparación con el ajuste inicial, no se observaron diferencias significativas, entonces se ajustaron modelos con efectos aleatorios para Asym o xmid. Se compararon con el modelo inicial y se observaron diferencias altamente significativas en ambos modelos. Se seleccionó al modelo con efectos aleatorios para la estimación de xmid porque es el que presentó mayor cantidad de intervalos no traslapados (Figura 4). Este procedimiento es sugerido por Pinheiro y Bates (2000).

Se procedió a continuación al ajuste del modelo de efectos fijos. Los resultados obtenidos mostraron interacción no significativa entre efectos de metodología y variedad. Los valores máximos de brotación (Asym) presentaron diferencias significativas y las estimaciones obtenidas fueron 45,80 % para LCP 85-384 y 40,92 % para cp-65-357.

Para la estimación del número de días desde la plantación hasta el 50 % de la emergencia máxima

(xmid) también se obtuvieron diferencias significativas y los valores estimados correspondientes fueron 44,19 días para la variedad CP 65-357 y 42,22 días para LCP 85-384.

El ajuste de la regresión logística se muestra en la Figura 5.

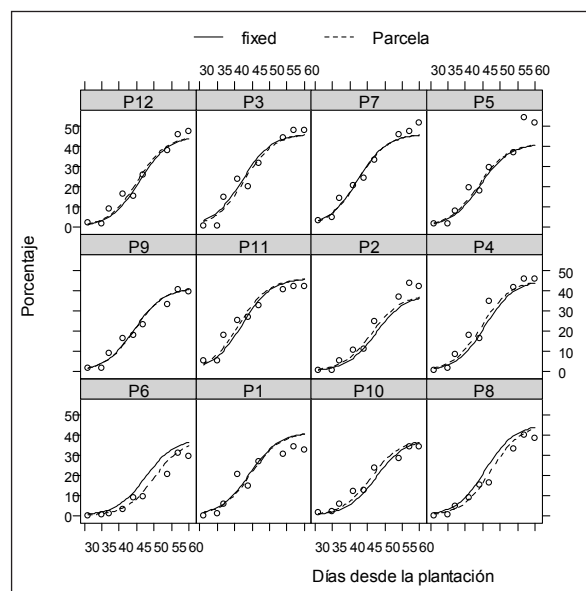


Figura 5. Predicciones poblacionales (fixed), predicciones por parcela, emergencia de tallos por metro versus días (puntos).

Se cumplieron los supuestos por lo que fue posible asumir normalidad y varianza constante de los residuos como se observa en las Figuras 6, 7, 8 y 9.

Conclusión

El modelo propuesto resultó adecuado para evaluar el efecto de metodologías y variedades. Brindó la estimación del máximo porcentaje de brotación en el período estudiado y del número de días desde la plantación en el que se produjo el 50 % de la emergencia máxima.

Permitió, además, establecer las diferencias significativas entre los tratamientos mediante la comparación de las estimaciones de los parámetros de las curvas ajustadas dentro de la estructura de un experimento factorial con efectos principales e interacciones.

Agradecimientos

Este estudio fue financiado por la Secretaría de Ciencia Arte e Innovación Tecnológica (SCAIT) de la Universidad Nacional de Tucumán en el marco del Proyecto PIUNT Código 26/A519.

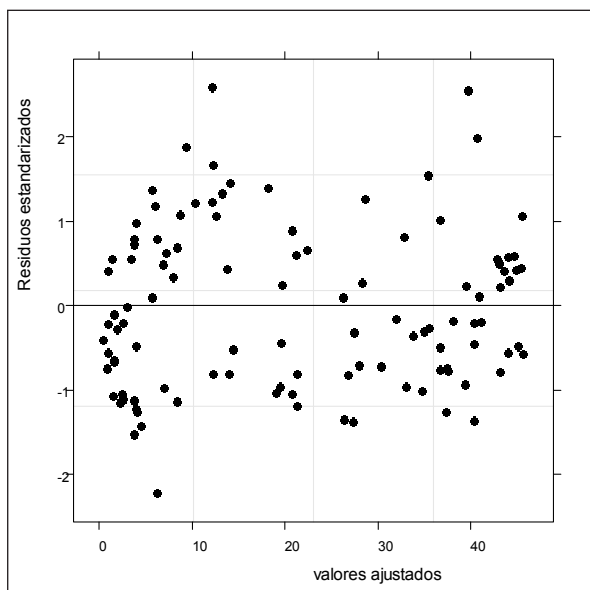


Figura 6. Valores ajustados y residuos estandarizados.

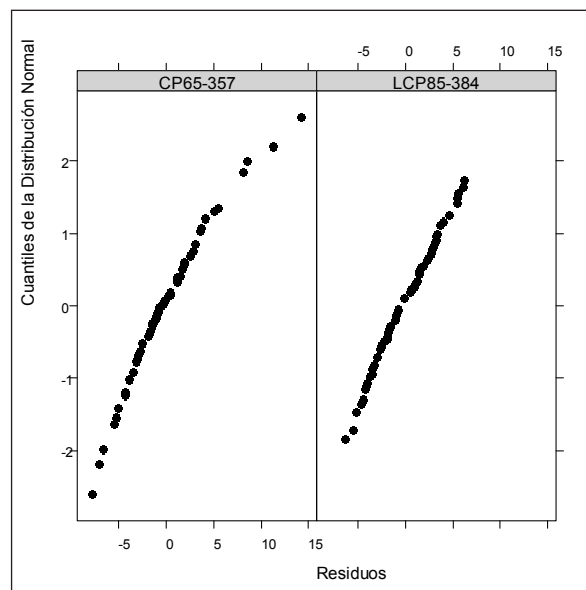


Figura 8. Normalidad de los residuos por variedad.

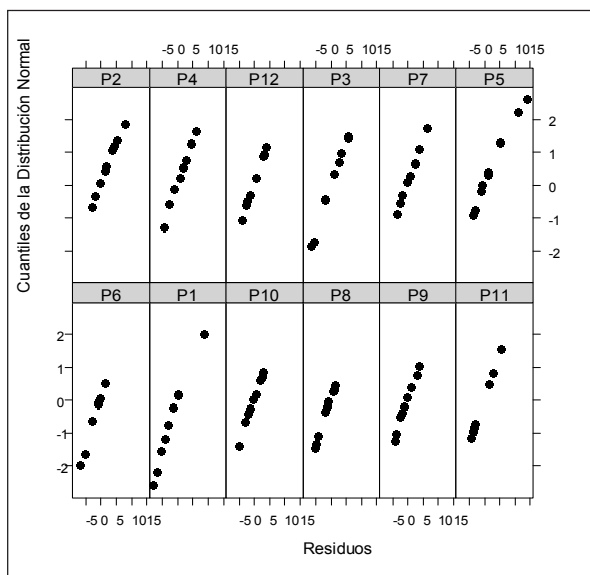


Figura 7. Normalidad de los residuos por parcela.

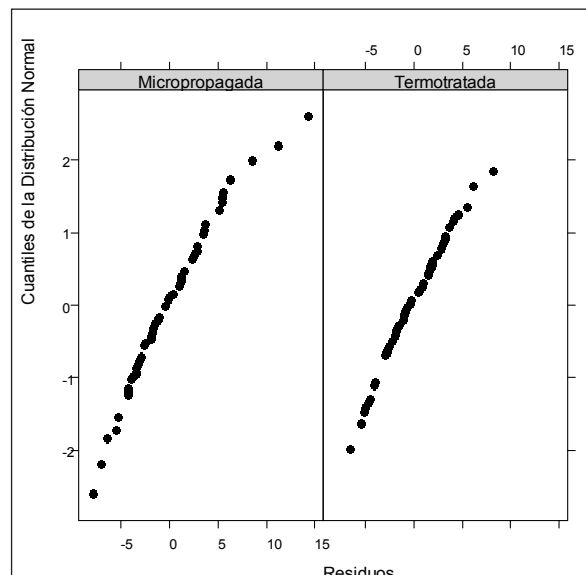


Figura 9. Normalidad de residuos por metodología.

Agradecemos los valiosos aportes de los evaluadores externos anónimos.

Referencias bibliográficas

Bates D.M., Watts D.G. (1988). Nonlinear regression analysis and its applications. Wiley and Sons, N.Y., USA.
 Cochran W.S. (1947). Some consequences when the assumptions for the analysis of variance are not satisfied. *Biometrics* 3: 22-38.
 Crawley M. (2007). *The R Book*. John Wiley & Sons Ltd, Chichester, England.
 Davidian M., Giltinan D.M. (1995). *Nonlinear models for repeated measurement data*. Chapman & Hall/CRC, Boca Raton FL, USA.

Davidian M. (2005). *Applied Longitudinal Data Analysis. Lecture Notes*. En: Department of Statistics, North Carolina State University, www4.stat.ncsu.edu/~davidian/papers.html, consulta: julio 2007.
 Diggle P.J., Heagerty P.J., Liang K.Y., Zeger S.L. (2002). *Analysis of Longitudinal Data*. Oxford University Press, Oxford, Great Britain.
 Digonzelli P.A., Giardina, J.A., Fernández de Ullivarri, J., Casen S.D., Tonatto M.J., Leggio Neme M.F., Romero E.R., Alonso L.G. (2009). Caña semilla de alta calidad. Obtención y manejo. En: *Manual del cañero*. Romero E, Digonzelli P, Scandaliaris, J. (Eds.). Estación Experimental Agroindustrial Obispo Colombes, Las Talitas, Argentina. Pp. 45-56.
 Digonzelli P.A., Giardina J., Ponce de León R., Sánchez Ducca, A., Fernández de Ullivarri, J., Scandaliaris J., Romero E. (2010). *Producción de caña se-*

- milla de alta calidad (Proyecto Vitroplantas): logros y desafíos. *Publicación Especial N° 40 Estación Experimental Agroindustrial*: 7-10.
- Faraway J. (2006). *Extending the linear model with R. Generalized linear, mixed effects and nonparametric regression models*. Chapman & Hall/CRC, Boca Raton FL, USA.
- Glasbey C.A. (1988). Examples of regression with serially correlated errors. *The Statistician* 37: 277-292.
- Khirsagar A., Smith B.W. (1995). *Growth curves*. Marcel Dekker Inc., N.Y., USA.
- Laird N.M., Ware J.H. (1982) Random-Effects models for longitudinal data. *Biometrics* 38: 963-974.
- Littel R.C., Milliken G.A., Stroup W.W., Wolfinger R.D. (2000). *SAS® System for Mixed Model*. 4th edition, SAS Institute Inc, Cary NC, USA.
- Maindonald J., Braun J. (2003). *Data analysis and graphics using R. An example-base approach*. Cambridge University Press, NY, USA.
- McCulloch C.E., Searle S.R., Neuhaus J.M. (2008). *Generalized, Linear, and Mixed Models*. Wiley & Sons Inc., NY, USA.
- Pinheiro J.C., Bates D.M. (2000). *Mixed-Effects Models in S and S-PLUS*. Springer-Verlag NY, USA.
- Pinheiro J., Bates D., DebRoy S., Sarkar D., Core Team (2013). *nlme: Linear and Nonlinear Mixed Effects Models*. R package. En: <http://CRAN.R-project.org/package=nlme>, consulta: mayo 2013.
- R Core Team (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. En: <http://www.r-project.org/>, consulta: mayo 2013.
- Sarkar D. (2008) *Lattice. Multivariate Data Visualization with R*. Springer Science+Business Media, N Y, USA.
- Venables W.N., Ripley B.D. (2002). *Modern Applied Statistics with S*. 4ª ed., Springer, NY, USA.
- Vonesh E.F., Chinchilli V.M. (1997). *Linear and non-linear models for the analysis of repeated measurements*. Marcel Dekker, Inc., NY, USA.
- Twisk J.W.R. (2003). *Applied longitudinal data analysis for epidemiology. A practical guide*. Cambridge University Press, Cambridge, United Kingdom.
- West B., Welch K., Galecki A. (2015). *Linear mixed models. A practical guide using statistical software*. 2nd edition, Chapman & Hall/CRC, Boca Raton FL, USA
- Wu L. (2010). *Mixed effects models for complex data*. Chapman & Hall/CRC, Boca Raton FL, USA.
- Zeger G.O., Liang K.Y, Albert P.S. (1988). Models longitudinal data: A generalized for estimating equation approach. *Biometrics* 44: 1049-1060.