



Experimentos agrícolas con medidas repetidas en el tiempo: comparación entre estrategias de análisis¹

Agricultural experiments with repeated measures over time: analysis strategies comparison

Jorge Claudio Vargas-Rojas², Alejandro Vargas-Martínez³, Eduardo Corrales-Brenes⁴

¹ Recepción: 4 de octubre, 2022. Aceptación: 7 de diciembre, 2022. Los datos utilizados para esta nota técnica derivan de un experimento realizado en Filadelfia, Guanacaste y financiado en su totalidad por la Ing. Laura Roldán Gómez.

² Universidad de Costa Rica, Sede Regional de Guanacaste, Liberia, Costa Rica. jorgeclaudio.vargas@ucr.ac.cr (autor para correspondencia, <https://orcid.org/0000-0002-1139-2148>).

³ Universidad Nacional, Heredia, Costa Rica. alejandro.vargas.martinez@una.ac.cr (<https://orcid.org/0000-0001-8039-8984>).

⁴ Centro Agronómico Tropical de Investigación y Enseñanza (CATIE), Turrialba, Costa Rica. ecorrales@catie.ac.cr (<https://orcid.org/0000-0002-7862-7546>).

Resumen

Introducción. Diversas técnicas de modelación se han utilizado para analizar experimentos con medidas repetidas en el tiempo; no obstante, algunas de estas no son pertinentes en la actualidad. **Objetivo.** Comparar cuatro estrategias de análisis que se usan para analizar experimentos agrícolas con evaluaciones a través del tiempo. **Materiales y métodos.** Se utilizaron datos de un experimento donde se evaluó el efecto de distintas fuentes de fertilizante nitrogenado sobre el contenido de clorofila en un pasto forrajero a diferentes edades de cosecha. Estos datos se analizaron con cuatro estrategias: índice de área bajo la curva (IABC), análisis de la varianza multivariado (MANOVA), efecto aleatorio de la unidad experimental y modelación de la correlación temporal. Las dos últimas estrategias se realizaron bajo la teoría de modelos lineales mixtos; en estas se ajustaron diferentes modelos, todos con la misma estructura de efectos fijos, pero con diferentes efectos aleatorios, estructura de correlación residual o estructura de varianza residual. Mediante criterios de verosimilitud penalizada [criterio de información de Akaike (AIC) y criterio de información Bayesiano (BIC)] se eligió el modelo de mejor ajuste, con este se realizaron inferencias acerca de las medias de los efectos fijos que fueron significativos y se comparó con los resultados obtenidos de las estrategias IABC y MANOVA. **Resultados.** El modelo lineal mixto de mejor ajuste tuvo una estructura de correlación de simetría compuesta y varianza heterocedástica. Este modelo permitió analizar la interacción tratamiento \times tiempo; por el contrario, tanto el MANOVA como el IABC, no permitieron analizar las tendencias temporales de los tratamientos. **Conclusión.** El modelo lineal mixto de mejor ajuste posibilitó seleccionar el mejor tratamiento en función del tiempo de evaluación. Por el contrario, tanto el MANOVA como el IABC condujeron a la selección de tratamientos que no fueron los mejores en todos los tiempos de evaluación.

Palabras clave: biometría, modelos estadísticos, análisis de la varianza, análisis multivariante, análisis estadístico.



Abstract

Introduction. Several modeling techniques have been used to analyze experiments with repeated measures over time; however, some of these are no longer relevant. **Objective.** To compare four analysis strategies that are used to analyze agricultural experiments with evaluations over time. **Materials and methods.** Data from an experiment in which the effect of different nitrogen fertilizer sources on chlorophyll content in a forage grass at different harvest ages was used. These data were analyzed using four strategies: the area under the curve index (AUC), multivariate analysis of variance (MANOVA), random effect of the experimental unit, and temporal correlation modeling. The last two strategies were performed under the theory of mixed linear models; in these different models were fitted, all with the same fixed effects structure, but with different random effects, residual correlation structure, or residual variance structure. Using penalized likelihood criteria [Akaike information criterion (AIC) and Bayesian information criterion (BIC)], the best fitting model was chosen, with which inferences were made about the means of the significant fixed effects, and compared with the results obtained from the AUC and MANOVA strategies. **Results.** The best fitting mixed linear had a compound symmetry correlation structure and heteroscedastic variances. This model allowed for the analysis of the treatment \times time interaction; on the other hand, both the MANOVA and the AUC allowed for the analysis of temporal trends of the treatments. **Conclusion.** The best fitting mixed linear model made it possible to select the best treatment based on the evaluation time. On the other hand, both the MANOVA and the AUC led to the selection of treatments that were not the best at all evaluation times.

Keywords: biometry, statistical models, analysis of variance, multivariate analysis, statistical analysis.

Introducción

El concepto de medidas repetidas o datos longitudinales corresponde a experimentos en donde sobre una misma unidad experimental se cuantifican una o más variables respuesta en diferentes tiempos, que pueden ser o no equidistantes (Littell et al., 2006). Cuando se cuantifica la variable respuesta de una misma unidad experimental repetidamente en el tiempo, los datos no son independientes i.e. están correlacionados en el tiempo, con lo cual se viola uno de los supuestos del análisis de varianza (ANDEVA). Además, tal estructura de correlación, por lo general, induce que la varianza residual no sea constante, sino que varía en los diferentes tiempos, lo que se le conoce como heterocedasticidad (Gómez et al., 2012; Zea et al., 2014). Estas situaciones, provocan el incumplimiento de dos de los supuestos del análisis de varianza (ANDEVA) y, por tanto, la inferencia estadística no es válida (Lawal, 2014).

Para abordar el problema de la violación de los supuestos en experimentos con medidas repetidas, se han propuesto varias estrategias de análisis. Una de las primeras técnicas fue el modelado en dos etapas, donde primero se colapsan las medidas repetidas de cada unidad experimental en un estadístico o índice y luego, es este el que se analiza mediante un ANDEVA. Dentro de estas técnicas se encuentra el índice del área bajo la curva (IABC), que divide el área total que se forma entre la respuesta (eje y) y el tiempo (eje x) en pequeños trapecios y suma el área de cada uno (Shi et al., 2021). Así, el estadístico del IABC constituye una medida resumen independiente para cada unidad experimental. Entre las críticas a esta estrategia se encuentra que debido a que la variable respuesta se transforma en un índice, se imposibilita realizar una interpretación de la variable respuesta medida en su escala original, además, al ser el índice una medida resumen, se pierde la posibilidad de observar las tendencias a través del tiempo. Asimismo, el IABC puede presentar problemas dados por el desbalance cuando hay pérdida de información en el tiempo (Bell et al., 2014).

Los experimentos con medidas repetidas también se han analizado con el análisis de varianza multivariado (MANOVA). Esta estrategia es criticada, ya que requiere que se cumplan, de forma simultánea, la normalidad y homogeneidad de varianzas multivariada (Arnau & Bono, 2008). El MANOVA también requiere datos balanceados para cada tiempo, caso contrario la técnica pierde potencia (Gueorguieva & Krystal, 2004). Además, para estimar todos los parámetros se necesita que el número de mediciones en el tiempo sea igual o menor al número de repeticiones del experimento, ya que el MANOVA utiliza una matriz de covarianza sin estructura, lo que implica la estimación de todas las covarianzas que puede tener una unidad experimental en el tiempo (Gómez et al., 2012).

Otra estrategia de análisis se deriva a partir del análisis de datos provenientes del diseño experimental de parcelas divididas. Este caso es conocido como parcelas divididas en el tiempo, donde se toma la combinación entre los tratamientos y las repeticiones como parcela principal y se supone que las diferentes evaluaciones en el tiempo son las parcelas pequeñas o subparcelas. Mediante esta estrategia se puede evaluar el efecto de los tratamientos a través del tiempo y no se pierde esta relevante información. Sin embargo, la principal crítica a esta estrategia es que el factor evaluado en la subparcela (tiempo de evaluación) no está aleatorizado, sino que está ordenado, lo que hace que se viole el supuesto de independencia (Littell et al., 2006; Powers & Kozak, 2019).

Los experimentos con medidas repetidas también se han analizado bajo el enfoque de los modelos lineales mixtos que, a diferencia del enfoque de parcelas divididas, no supone la aleatorización del tiempo, sino que incluye la correlación que existe entre las evaluaciones que provienen de una misma unidad experimental. Los modelos mixtos permiten incluir diferentes estructuras de correlación y de varianzas con el fin de encontrar el modelo más verosímil que explique de forma correcta y eficiente los datos de este tipo de experimentos (Gómez et al., 2012; Verbeke & Molenberghs, 2000).

En el contexto de los modelos lineales mixtos se usan criterios de verosimilitud penalizada para comparar diferentes modelos ajustados a un mismo conjunto de datos. Uno de los criterios más utilizados es el criterio de Akaike (AIC, por sus siglas en inglés), este se basa en el negativo del logaritmo de la función de verosimilitud más una penalización según el número de parámetros estimados en el modelo. Otro criterio empleado es el bayesiano (BIC, por sus siglas en inglés), que al igual que el AIC se basa en el logaritmo de la verosimilitud, pero cobra una penalización mayor por número de parámetros estimados. Valores menores de AIC o BIC indican mejor ajuste del modelo estadístico (Demidenko, 2013; West et al., 2015).

El uso de modelos lineales mixtos provee ventajas pero aún persiste el uso de las estrategias como el IABC, MANOVA y parcelas divididas en el tiempo. El objetivo de este trabajo fue comparar cuatro estrategias de análisis que se usan para analizar experimentos agrícolas con evaluaciones a través del tiempo.

Materiales y métodos

Se utilizaron datos colectados de un experimento dispuesto en un diseño de bloques completos al azar (DBCA), donde se tuvo por objetivo evaluar el efecto de cinco distintas fuentes de fertilizante nitrogenado sobre el contenido de clorofila, para ello, se obtuvo el índice de clorofila, medido como unidades SPAD, en un pasto forrajero a diferentes edades de cosecha: 30, 45 y 70 días después del corte de uniformidad (DDU). Por motivos de confidencialidad, se resguardó la información de cada tratamiento.

Para medir las unidades SPAD, en cada unidad experimental se seleccionaron al azar treinta plantas (Peterson et al., 1993), que fueron identificadas para realizar las evaluaciones en las diferentes edades de cosecha, siempre sobre las mismas plantas. La medición se realizó en cinco puntos del tercio medio del limbo de la hoja más nueva desarrollada de cada planta. Se utilizó un medidor de clorofila Minolta® SPAD 502 (Soil Plant Analysis Development). Los datos descritos se analizaron bajo cuatro diferentes estrategias.

Estrategia 1: índice del área bajo la curva (IABC)

A los datos del experimento, se le calculó el índice del área bajo la curva (IABC), según lo describe Navarro Flores (2012), con el uso del paquete DescTools (Signorell, 2022) del lenguaje de programación R versión 3.6.3 (R Core Team, 2020). Luego, se ajustó un modelo de análisis de varianza con el IABC como variable respuesta según el modelo que se especifica en la ecuación 1. El modelo se utilizó para realizar inferencias acerca de las medias de los tratamientos, en este caso, solo para las fuentes de fertilizante nitrogenado.

$$y_{ij} = \mu + \tau_i + \beta_j + e_{ij} \quad \text{Ecuación 1}$$

Donde

y_{ij} : es el IABC del i -ésimo nivel de fertilizante nitrogenado en el j -ésimo bloque.
 μ : es la media general de la respuesta.
 τ_i : es el efecto fijo del i -ésimo nivel de fertilizante nitrogenado.
 β_j : es el efecto fijo del j -ésimo bloque.
 e_{ij} : representa el error residual, se supone que se distribuye normal e independiente con esperanza 0 y varianza $\sigma^2 \forall i, j \neq i', j'$.

Estrategia 2: análisis de la varianza multivariado (MANOVA)

El modelo utilizado para el análisis de la varianza multivariado (MANOVA) fue similar al modelo de la ecuación 1, la diferencia radica en que en lugar de una única variable respuesta, el contenido de unidades SPAD en cada tiempo se consideró como una variable respuesta independiente. Luego, el estadístico de prueba (Wilks) contrastó la hipótesis de igualdad de medias entre los niveles de fertilización nitrogenada sobre el contenido de unidades SPAD en todos los tiempos evaluados. Además, se realizó la prueba de Hotelling para la comparación de las medias de los tratamientos con un nivel de significancia (α) de 0,05. Los procedimientos se implementaron con el software estadístico InfoStat versión 2020 (Di Rienzo et al., 2020).

Estrategia 3: efecto aleatorio de la unidad experimental

La estrategia del efecto aleatorio de la unidad experimental se basa en la identificación y declaración en el modelo de la unidad experimental que es evaluada a través del tiempo y que se denomina parcela principal. El modelo (M1) de esta estrategia fue ajustado como un modelo lineal mixto (ecuación 2), donde los factores niveles de fertilización nitrogenada, tiempo e interacción entre ellos, fueron declarados efectos fijos. La correlación de las unidades experimentales que fueron medidas en el tiempo, se incluyó como un efecto aleatorio de unidad experimental (también llamado sujeto o parcela principal).

$$y_{ijkl} = \mu + t_i + \gamma_j + \tau\gamma_{ij} + b_k + s_{i(k)} + e_{ijkl} \quad \text{Ecuación 2}$$

Donde:

y_{ijkl} : es el contenido de unidades SPAD del i -ésimo sujeto en el k -ésimo bloque del i -ésimo tratamiento en el j -ésimo tiempo.
 μ : es la media general de la respuesta.
 τ_i : es el efecto fijo del i -ésimo nivel de fertilizante nitrogenado.

γ_j : es el efecto fijo del j -ésimo nivel de tiempo.

$\tau\gamma_{ij}$: representa el efecto de la interacción del i -ésimo nivel de fertilizante nitrogenado en el j -ésimo tiempo.

b_k : es el efecto aleatorio del k -ésimo bloque, distribuido normal con media cero y varianza (σ_b^2) constante entre bloques.

$s_{i(k)}$: es el efecto aleatorio de la i -ésima unidad experimental dentro del k -ésimo bloque, distribuido normal con media cero y varianza (σ_s^2) constante entre parcelas principales.

e_{ijkl} : representa el error residual, se supone que se distribuye normal con media cero. Además, los componentes aleatorios b_k , $s_{i(k)}$ y e_{ijkl} se suponen independientes.

También, se ajustó un segundo modelo (M2) que, además de tomar en cuenta la falta de independencia dada por las medidas repetidas, consideró una función de varianza (VarIdent) que permitió que la varianza en cada uno de los tiempos fuera diferente. El ajuste de los modelos se realizó con el software estadístico InfoStat versión 2020 (Di Rienzo et al., 2020) y su especificación se presenta en el Cuadro 1.

Cuadro 1. Efectos fijos, efectos aleatorios, estructura de correlación y varianza residual de los modelos ajustados. Filadelfia, Guanacaste. Enero 2020.

Table 1. Fixed effects, random effects, correlation structure, and residual variance of the fitted models. Filadelfia, Guanacaste. January 2020.

Modelo	Efectos fijos				Efectos aleatorios		Correlación residual	Varianza residual
M1	Media general	Tiempo	Tratamiento	Interacción	Bloque	Unidad experimental	Independencia	Homoscedástica
M2	Media general	Tiempo	Tratamiento	Interacción	Bloque	Unidad experimental	Independencia	Heterocedástica
M3	Media general	Tiempo	Tratamiento	Interacción	Bloque		Simetría compuesta	Homoscedástica
M4	Media general	Tiempo	Tratamiento	Interacción	Bloque		Simetría compuesta	Heterocedástica
M5	Media general	Tiempo	Tratamiento	Interacción	Bloque		AR1	Homoscedástica
M6	Media general	Tiempo	Tratamiento	Interacción	Bloque		AR1	Heterocedástica
M7	Media general	Tiempo	Tratamiento	Interacción	Bloque		Sin estructura	Homoscedástica
M8	Media general	Tiempo	Tratamiento	Interacción	Bloque		Sin estructura	Heterocedástica

Estrategia 4: modelación de la correlación temporal de las unidades experimentales

La modelación de la correlación temporal se realizó (al igual que la estrategia 3) mediante el uso de modelos lineales mixtos, con la diferencia que en esta estrategia la estructura de correlación de las medidas repetidas en el tiempo se indujo en la matriz de correlación de los residuos. El modelo asociado a esta estrategia es similar al descrito en la ecuación 2, con la diferencia que no se declaró el efecto aleatorio de la unidad experimental. En esta estrategia se ajustaron tres modelos diferentes: M3, M5 y M7 (Cuadro 1), los cuales se caracterizan por tener una estructura de correlación residual específica, que son simetría compuesta, autorregresiva de orden 1 y sin estructura. Luego, los modelos M4, M6 y M8 adicionan la función de varianza VarIdent para cada tiempo a la estructura de correlación específica.

La primera estructura de correlación residual se llama simetría compuesta, la cual supone que la correlación entre todos los tiempos es la misma (estima un único coeficiente de correlación (ρ)).

La segunda estructura de correlación, autorregresiva de orden 1 (AR1), se utiliza en aquellos casos donde se puede suponer que la intensidad de correlación se pierde conforme avanza el tiempo (se estima un único coeficiente de correlación, pero eleva a la potencia p , donde p es la distancia entre los diferentes tiempos; esto es, la correlación entre el tiempo 1 y 3 es el coeficiente de correlación ρ elevado a la potencia $p=2$).

La tercera estructura de correlación se conoce con el nombre de sin estructura, la cual permite que la correlación entre todos los pares de tiempos sea diferente, es decir, se tendrán tantos coeficientes de correlación como combinaciones de a -pares resulte, por ejemplo, si se evaluaron cuatro tiempos, se tendrán $\frac{4 \times 3}{2} = 6$ coeficientes de correlación diferentes. Además, para cada uno de los modelos descritos, se ajustó un modelo adicional con varianzas diferentes por cada tiempo (M6-M8, Cuadro 1). Los ocho modelos estimados (estrategia 3 y estrategia 4) tuvieron los mismos efectos fijos por lo que son comparables mediante los criterios de verosimilitud penalizada [(criterio de información de Akaike (AIC) y criterio Bayesiano (BIC)]. Bajo estos criterios, se eligió el modelo de mejor ajuste y se realizó la prueba DGC (Di Rienzo et al., 2002) para las medias de los factores que fueron significativos. Los procedimientos se hicieron con el programa estadístico InfoStat versión 2020 (Di Rienzo et al., 2020).

Comparación de estrategias

Las estrategias 3 y 4 son comparables de forma directa vía criterios de la verosimilitud penalizada (AIC y BIC) (West et al., 2015). Sin embargo, las estrategias 1 y 2 no son comparables, ya que el IABC es una transformación de la variable respuesta original (unidades SPAD) y, aunque la estrategia 2 no realiza una transformación, tampoco mantiene la estructura original de los datos, ya que supone que la respuesta en cada uno de los tiempos es una nueva variable independiente (Navarro Flores, 2012; Stroup et al., 2018). Bajo este escenario, la estrategia de comparación fue con base en las inferencias (medias ajustadas), información producida y las conclusiones obtenidas en cada una de las cuatro estrategias de modelado.

Resultados

A partir de la estrategia 1 (IABC), hubo diferencias significativas entre los tratamientos ($F = 8,40$; $p = 0,0018$) sobre el IABC. Los tratamientos T1 y T2 fueron los que presentaron mayor IABC con diferencia estadística de los tratamientos T3, T4 y T5 (Figura 1).

La segunda estrategia (MANOVA) mostró diferencias significativas entre tratamientos ($F = 11,08$ y $p < 0,0001$) sobre las unidades SPAD. En esta, la comparación se realizó para todos los tiempos de evaluación de forma conjunta y no individual. Así, los tratamientos T3, T1 y T2 fueron los que presentaron mayores valores de SPAD con diferencia estadística de los tratamientos T4 y T5 (Cuadro 2).

Con respecto a las estrategia 3 (efecto aleatorio de la unidad experimental) y estrategia 4 (modelación de la correlación temporal de las unidades experimentales), el menor valor de AIC se obtuvo en el modelo M8, mientras el modelo con menor BIC fue M4 (Cuadro 3). En estos casos donde los criterios de información no son contundentes para seleccionar un modelo, se recomienda seleccionar el más simple i.e. parsimonioso, que en modelación estadística concuerda con el modelo con menos parámetros a estimar, en este caso el mejor fue el modelo M4.

La interacción tratamiento \times tiempo del modelo M4 fue significativa ($F = 7,73$ y $p < 0,0001$). Esta interacción indicó que existe dependencia entre el tratamiento y el tiempo, es decir, que el efecto de la fuente de fertilizante nitrogenada sobre las unidades SPAD no fue el mismo para los diferentes tiempos (DDU) de evaluación (Figura 2). Todos los tratamientos presentaron una tendencia a disminuir las unidades SPAD conforme aumentó la fecha de evaluación, a excepción del T4 que en la última evaluación aumentó de manera significativa las unidades SPAD.

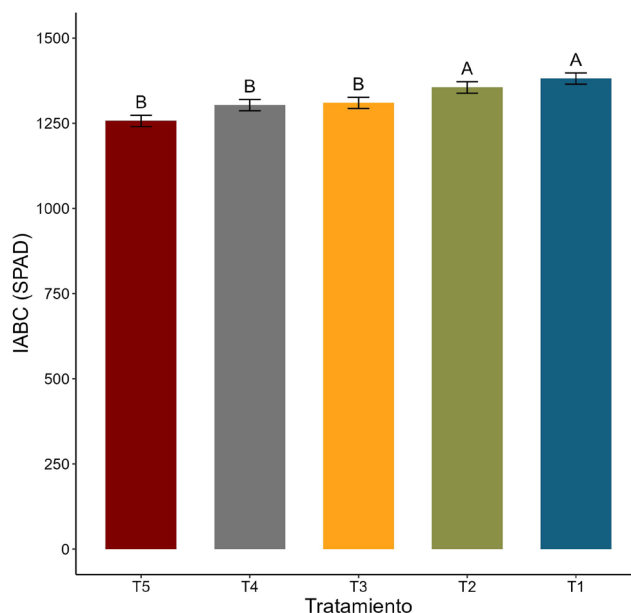


Figura 1. Gráfico de barras de las medias del índice del área bajo la curva (IABC) en función del tratamiento. Filadelfia, Guanacaste. Enero 2020.

Medias con letra común no son estadísticamente diferentes ($p>0,05$).

Figure 1. Bar graph of the area under the curve index (IABC) means according to treatment. Filadelfia, Guanacaste. January 2020.

Means with common letter are not statistically different ($p>0.05$).

Cuadro 2. Separación de perfiles promedio de unidades SPAD para los distintos tratamientos. Filadelfia, Guanacaste. Enero 2020.

Table 2. Average SPAD units' separation profiles for the different treatments. Filadelfia, Guanacaste. January 2020.

Tratamiento	DDU			Clasificación
	30	45	70	
T1	36,52	34,23	33,83	A
T2	36,48	33,87	32,33	A
T3	36,00	32,27	31,55	A
T4	32,42	31,85	33,87	B
T5	31,37	31,77	30,88	B

*DDU: días después del corte de uniformidad; medias con letra común no son estadísticamente diferentes ($p>0,05$). / *DDU: days after uniformity cut; means with common letter are not statistically different ($p>0.05$).

A los 30 DDU los tratamientos T1, T2 y T3, fueron los que presentaron mayores unidades SPAD. A los 45 DDU el T1 y el T2 se mantuvieron como los tratamientos con mayores unidades SPAD, pero el T3 disminuyó significativamente esta variable y no se diferenció del T4 y el T5. Para la evaluación de los 70 DDU, el T2 disminuyó significativamente las unidades SPAD, los tratamientos T1, T3 y T5 se mantuvieron sin cambios con respecto a la evaluación a los 45 DDU, pero el tratamiento T4 presentó un incremento significativo en las unidades SPAD en comparación con la fecha anterior.

Cuadro 3. Criterios de verosimilitud penalizada [criterio de información de Akaike (AIC) y criterio bayesiano (BIC)] de los modelos lineales mixtos ajustados. Filadelfia, Guanacaste. Enero 2020.

Table 3. Penalized likelihood criteria [Akaike’s information criterion (AIC) and Bayesian criterion (BIC)] of the fitted linear mixed models. Filadelfia, Guanacaste. January 2020.

Modelo	Criterio de verosimilitud penalizada	
	AIC	BIC
M1	202,89	235,41
M2	196,90	233,03
M3	202,89	235,41
M4	196,80	232,94
M5	203,38	235,90
M6	197,51	233,64
M7	204,48	240,62
M8	194,75	234,50

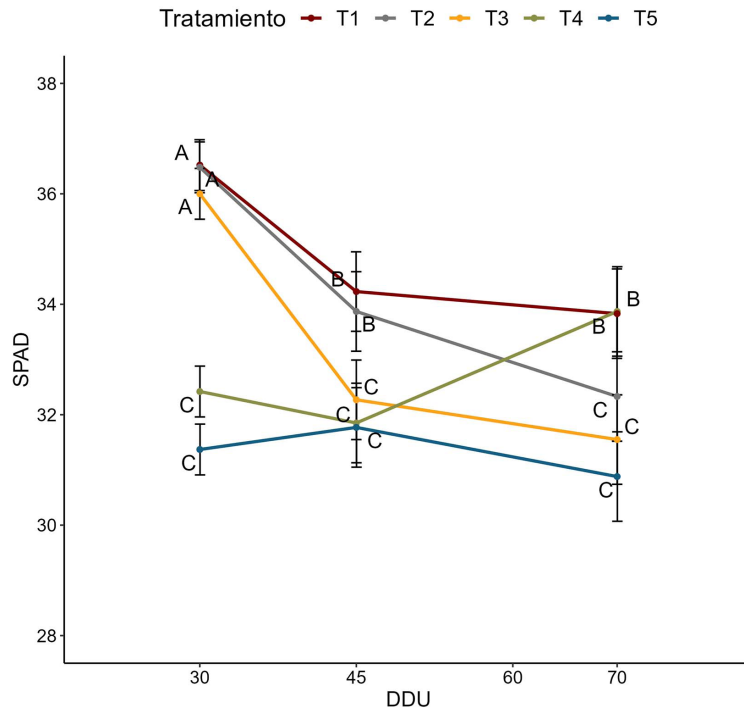


Figura 2. Separación de medias obtenidas con el modelo M4 para cada combinación de niveles de los factores tratamiento y días después del corte de uniformidad (DDU). Filadelfia, Guanacaste. Enero 2020.

Figure 2. Means separation obtained with the M4 model for each combination of treatment and days after the uniformity cutoff (DDU). Filadelfia, Guanacaste. January 2020.

Discusión

Es fundamental hacer énfasis sobre la estructura de los datos en experimentos con medidas repetidas. Por definición, al menos, existen dos factores, tratamiento y tiempo; por tanto, son experimentos factoriales y en estos, es de primordial interés analizar el término de la interacción entre factores (Littell et al., 2006). Como la interacción del modelo M4 fue significativa, la diferencia en la respuesta entre los niveles de tratamiento depende del DDU (Montgomery, 2020), por lo que no se puede dar una recomendación de tratamiento de forma independiente al tiempo de evaluación, sino que ambos factores se deben analizar de manera conjunta para llegar a una conclusión correcta.

El T2 mostró altas unidades SPAD en la primera y segunda fecha de evaluación, no obstante, este tratamiento tuvo una tendencia a disminuir conforme aumentó el tiempo, esa tendencia no se pudo analizar con el uso de técnicas que no consideraron el tiempo como un factor dentro del modelo. Motivo por el cual, tanto en el análisis del IABC como en el MANOVA, el T2 quedó bien posicionado, debido a dos evaluaciones con valores altos. Así, al concluir con base en los resultados del IABC o MANOVA, se podría incurrir en el error de recomendar el T2 para cualquier fecha que se establezca para cortar el pasto, sin considerar que a los 70 DDU este tratamiento presentó una disminución de unidades SPAD.

Con el modelo lineal mixto (M4) se pudo seleccionar el tratamiento en función de la fecha que se definió para la corta del pasto. Las conclusiones cambiaron de una estrategia a la otra. La principal razón es porque el IABC eliminó el tiempo al colapsar todas las medidas en un solo índice y esto imposibilitó evaluar las tendencias temporales que el modelo lineal mixto de la estrategia 3 y 4 sí pusieron en evidencia (Powers & Kozak, 2019).

El M1 se puede considerar como un enfoque moderno del análisis de parcelas divididas en el tiempo bajo la teoría de modelos lineales mixtos. La principal diferencia entre este enfoque y el análisis de parcelas divididas en el tiempo bajo el enfoque clásico, es que en este último se asume que las medidas repetidas sobre una misma unidad experimental no están correlacionadas o que están correlacionadas con la misma magnitud (Barrios Maestre & Silva-Acuña, 2019; Gómez et al., 2012), además, asume que las varianzas de los diferentes tiempos son constantes. El modelo lineal mixto permite incluir distintas estructuras para modelar la correlación entre unidades experimentales y modelar la heterogeneidad de varianzas. Un modelo que de forma incorrecta asuma la independencia de las observaciones conlleva a la distorsión del estadístico F.

En presencia de correlación los modelos que suponen independencia subestiman la varianza (Legendre & Fortin, 1989) e inflan el efecto de los tratamientos (Mets et al., 2017), lo que conduce a la pérdida de precisión del modelo y a un aumento en la tasa de error tipo I (Beale et al., 2010). En comparaciones de modelos para analizar datos correlacionados, los modelos sin estructura de correlación estimaron varianzas más pequeñas y tuvieron mayor tasa de error tipo I que los modelos con estructura de correlación (Mets et al., 2017). Además, en estudios de simulación, se encontró que cuando existe correlación entre las observaciones, los métodos clásicos arrojaban resultados significativos por encima de ocho veces la tasa nominal (Dale & Fortin, 2014). Esta situación puede conducir a errores en la toma de decisiones basadas en los resultados de la prueba de hipótesis, en los ensayos agrícolas. En estos se busca un tratamiento que sea mejor que un estándar y se puede incurrir en pérdidas de recursos valiosos si se cambia o se invierte en un “buen” tratamiento que en realidad es un falso positivo. Por esta razón, se recomienda el uso de modelos lineales mixtos para abordar la correlación entre observaciones y realizar estimaciones más precisas (Dale & Fortin, 2014).

La ventaja de modelar la heterocedasticidad entre los diferentes tiempos quedó reflejada al comparar tanto el AIC como el BIC entre los modelos M3 y M4. El M4 tuvo mejor ajuste debido a la función VarIdent (Pinheiro & Bates, 2000; Zuur et al., 2009) que le permitió a cada fecha de evaluación tener su propia varianza, de manera que la violación al supuesto de homogeneidad de varianzas se corrigió. Se debe tener presente que si existen varianzas heterogéneas entre grupos y esto no se modela, se alteran las tasas nominales de error tipo I y tipo II (Rosopa et al., 2013; Wilcox, 1986).

Es imperativo, en la etapa de planificación de un experimento, que el investigador considere la pertinencia de realizar medidas en el tiempo. Es importante analizar aspectos prácticos, si los cambios de la respuesta en el tiempo son relevantes para el tratamiento a evaluar o si solo interesa la respuesta final, si es así, las evaluaciones en el tiempo no son necesarias. Caso contrario, si el investigador sospecha que la respuesta al tratamiento y al tiempo no se puede explicar como la suma de los efectos principales, el modelo factorial con evaluaciones en el tiempo se hace necesario (Di Rienzo et al., 2005) y en la etapa de análisis se debe realizar un modelado que tenga en consideración que las evaluaciones que se realizan sobre una misma unidad experimental no son independientes.

Conclusiones

El modelo lineal mixto de mejor ajuste permitió seleccionar el mejor tratamiento en función del tiempo de evaluación, debido a que permitió analizar de forma correcta las tendencias temporales de los tratamientos. Por el contrario, tanto el MANOVA como el IABC, enmascararon el efecto del tiempo, lo que condujo a la selección de tratamientos que que no fueron los mejores en todos los tiempos de evaluación.

Tanto la inclusión de la estructura de correlación como la modelación de la heterocedasticidad mejoraron la bondad de ajuste del modelo lineal mixto.

Referencias

- Arnau, J., & Bono, R. (2008). Estudios longitudinales: modelos de diseño y análisis. *Escritos de Psicología*, 2(1), 32–41.
- Barrios Maestre, R., & Silva-Acuña, R. (2019). Pertinencia de uso del análisis estadístico de medidas repetidas en la investigación agrícola. *Agronomía Tropical*, 69, 9–17.
- Beale, C. M., Lennon, J. J., Yearsley, J. M., Brewer, M. J., & Elston, D. A. (2010). Regression analysis of spatial data. *Ecology Letters*, 13, 246–264. <https://doi.org/10.1111/J.1461-0248.2009.01422.X>
- Bell, M. L., King, M. T., & Fairclough, D. L. (2014). Bias in area under the curve for longitudinal clinical trials with missing patient reported outcome data: Summary measures versus summary statistics. *SAGE Open*, 4(2). <https://doi.org/10.1177/2158244014534858>
- Dale, M. R. T., & Fortin, M. J. (2014). *Spatial analysis. A guide for ecologists* (2nd ed.). Cambridge University Press.
- Demidenko, E. (2013). *Mixed models: Theory and applications with R* (2nd ed.). Wiley.
- Di Rienzo, J. A., Casanoves, F., Balzarini, M. G., Gonzalez, L., Tablada, M., & Robledo, C. W. (2020). *Infostat* (Versión 2020) [programa de computadora]. Centro de Transferencia InfoStat. <https://www.infostat.com.ar/index.php?mod=page&id=15>
- Di Rienzo, J. A., Casanoves, F., Gonzalez, L., Tablada, E., Díaz, M., Robledo, C., & Balzarini, M. (2005). Diseño de experimentos. In J. Di Rienzo., & F. Casanoves (Eds.), *Estadística para las ciencias agropecuarias* (7^a ed., pp. 245–274). Editorial Brujas.
- Di Rienzo, J. A., Guzmán, A. W., & Casanoves, F. (2002). A multiple-comparisons method based on the distribution of the root node distance of a binary tree. *Journal of Agricultural, Biological, and Environmental Statistics*, 7(2), 129–142. <https://doi.org/10.1198/10857110260141193>

- Gómez, S., Torres, V., García, Y., & Navarro, J. A. (2012). Procedimientos estadísticos más utilizados en el análisis de medidas repetidas en el tiempo en el sector agropecuario. *Revista Cubana de Ciencia Agrícola*, 46(1), 1–7.
- Georgueva, R., & Krystal, J. H. (2004). Move over ANOVA: progress in analyzing repeated-measures data and its reflection in papers published in the Archives of General Psychiatry. *Archives of General Psychiatry*, 61(3), 310–317. <https://doi.org/10.1001/ARCHPSYC.61.3.310>
- Lawal, B. (2014). *Applied statistical methods in agriculture, health and life sciences*. Springer Cham. <https://doi.org/10.1007/978-3-319-05555-8>
- Legendre, P., & Fortin, M. J. (1989). Spatial pattern and ecological analysis. *Vegetatio*, 80(2), 107–138. <https://doi.org/10.1007/BF00048036>
- Littell, R. C., Milliken, G. A., Stroup, W. W., Wolfinger, R. D., & Schabenberger, O. (2006). *SAS for mixed models* (2nd ed.). SAS Institute Inc.
- Mets, K. D., Armenteras, D., & Dávalos, L. M. (2017). Spatial autocorrelation reduces model precision and predictive power in deforestation analyses. *Ecosphere*, 8(5), Article e01824. <https://doi.org/10.1002/ECS2.1824>
- Montgomery, D. C. (2020). *Design and analysis of experiments* (10th ed.). John Wiley & Sons, Inc.
- Navarro Flores, J. R. (2012). *La técnica del área bajo la curva*. Centro de Protección de Cultivos. <https://www.kerwa.ucr.ac.cr/handle/10669/585>
- Peterson, T. A., Blackmer, T. M., Francis, D. D., & Schepers, J. S. (1993). *G93-1171 using a chlorophyll meter to improve N management*. University of Nebraska. <https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=2349&context=extensionhist>
- Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. Springer-Verlag.
- Powers, S. J., & Kozak, M. (2019). Repeated measures: There's added value in modelling over time. *Annals of Applied Biology*, 175(2), 129–135. <https://doi.org/10.1111/aab.12534>
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Rosopa, P. J., Schaffer, M. M., & Schroeder, A. N. (2013). Managing heteroscedasticity in general linear models. *Psychological Methods*, 18(3), 335–351. <https://doi.org/10.1037/A0032553>
- Shi, L., Hatsukami, D. K., Koopmeiners, J. S., Le, C. T., Benowitz, N. L., Donny, E. C., & Luo, X. (2021). A mixed effects model for analyzing area under the curve of longitudinally measured biomarkers with missing data. *Pharmaceutical Statistics*, 20(6), 1249–1264. <https://doi.org/10.1002/pst.2146>
- Signorell, A. (2022). *Tools for descriptive statistics and exploratory data analysis* (Version 0.99.44) [Computer software]. R Core Team. <https://cran.r-project.org/package=DescTools>
- Stroup, W. W., Milliken, G. A., Claassen, E. A., & Wolfinger, R. D. (2018). *SAS for Mixed models: Introduction and basic applications* (2nd ed.). SAS Institute Inc.
- Verbeke, G., & Molenberghs, G. (2000). *Linear mixed models for longitudinal data*. Springer. <https://doi.org/10.1007/978-1-4419-0300-6>

- West, B. T., Welch, K. B., & Galecki, A. T. (2015). *Linear mixed models. A practical guide using statistical Software* (2nd ed.). CRC Press.
- Wilcox, R. R. (1986). Controlling power in a heteroscedastic ANOVA procedure. *British Journal of Mathematical and Statistical Psychology*, 39(1), 65–68. <https://doi.org/10.1111/J.2044-8317.1986.TB00845.X>
- Zea, J. F., Murcia, M. A., & Poveda, F. E. (2014). Modelos mixtos aplicados a la productividad de hojarasca. *Comunicaciones En Estadística*, 7(2), 173–189. <https://doi.org/10.15332/S2027-3355.2014.0002.04>
- Zuur, A. F., Ieno, E. N., Walker, N., Saveliev, A. A., & Smith, G. M. (2009). Dealing with heterogeneity. In M. Gail., K. Krickeberg., J. M. Samet., A. Tsiatis., & W. Wong (Eds.), *Mixed effects models and extensions in ecology with R*. Springer. <https://doi.org/10.1007/978-0-387-87458-6>