

NOTA TÉCNICA

Análisis multivariado en plantaciones de eucaliptos en dos micro-ambientes de La Esperanza, provincia de Jujuy, con el método Procrustes Generalizado

Multivariate analysis of eucalyptus plantations in two microenvironments in La Esperanza, province of Jujuy, with the Generalized Procrustes method

J. M. Solís¹; S. De Tellería²; A. Montenegro³ y J. Quispe⁴

¹ Ruta Provincial Nro. 56 N 1331, Bajo La Viña, San Salvador de Jujuy, Jujuy, Argentina. E-mail: juanmasolis@fca.unju.edu.ar

² Daniel Toro S/N. La Caldera, Salta, Argentina. E-mail: santiagodetelleria@fca.unju.edu.ar

³ Esteban Salinas 487, Mayor Buratovich, Buenos Aires, Argentina. E-mail: montenegro.a@inta.gob.ar

⁴ Abra Pampa y Tres Cruces 261, San Salvador de Jujuy, Jujuy, Argentina. E-mail: quispejulian@gmail.com

Recibido en abril de 2023; aceptado en septiembre de 2023

RESUMEN

El Análisis de Procrustes Generalizado (GPA por sus siglas en inglés) constituye un método de análisis multivariado muy conveniente para el análisis de variabilidad multiambiental, ya que permite el ajuste de componentes principales parciales dentro de cada ambiente, que luego conforman una matriz de consenso que ajusta mejor a la estructura intrínseca de la variabilidad de los fenómenos observados. En este trabajo se realizó un GPA sobre plantaciones de dos años de tres genotipos de eucalipto en dos microambientes o lotes en la localidad de La Esperanza, Provincia de Jujuy, a fin de determinar la contribución relativa que sobre la variabilidad total observada tuvieron el genotipo y el ambiente, como una herramienta para el análisis y la toma de decisiones. De esta forma, se pudo identificar que el híbrido *Eucalyptus grandis* x *camaldulensis* fue el genotipo más estable, es decir, el que presentó menor variabilidad entre ambientes, como así también el de mayor productividad. *Eucalyptus saligna* presentó la mayor variabilidad entre ambientes, y *Eucalyptus camaldulensis* presentó una variabilidad intra y entre ambientes intermedias.

Palabras Claves: Análisis multiambiental, Consenso de Componentes Principales, FactoMineR

ABSTRACT

The Generalized Procrustes Analysis (GPA) is a very convenient multivariate method for the analysis of multi-environmental variability, since it allows the adjustment of partial principal components within each environment, which then form a consensus or average matrix that best fits the intrinsic structure of the variability of the observed phenomena. In this work, a GPA was carried out on two-year-old plantations of three eucalyptus genotypes in two microenvironments or plots in La Esperanza, Jujuy province, to determine the relative contribution of the genotype and the environment over the total observed variability, as a tool for analysis and decision making. In this way, it was possible to ascertain that the hybrid *Eucalyptus grandis* x *camaldulensis* was the most stable genotype and exhibited the least variability across environments, and it was also the most productive one. *Eucalyptus saligna* displayed the highest variability across environments. Lastly, *Eucalyptus camaldulensis* demonstrated intermediate intra- and inter-environment variability.

Key words: Multi-environmental analysis, Consensus of Principal Components, FactoMineR.

1. INTRODUCCIÓN

A partir de 2014, el Ingenio La Esperanza (provincia de Jujuy) decidió activar el área forestal como actividad complementaria a la caña de azúcar con una ampliación en la producción de eucaliptos. A partir de esta iniciativa, se probaron distintas variedades de eucaliptos y técnicas de plantación, incorporando material genético de buenos resultados en otras regiones del país. Los sitios

plantados fueron zonas abandonadas de cultivo de caña de azúcar, potreros ganaderos abandonados, o tierras habilitadas no aptas para la producción de la caña de azúcar, con lo cual se hace necesario contar con herramientas que permitan analizar y caracterizar diferentes genotipos como respuesta a variaciones ambientales a micro – escala, a fin de identificar aquellos con mayor aptitud o capacidad de adaptación.

El Análisis de Procrustes Generalizado (GPA, por sus siglas en inglés) fue introducido por Gower (1975) como una técnica *estadística* multivariada para analizar matrices de datos multidimensionales. Este método permite abordar el análisis de datos provenientes de un conjunto de objetos o individuos en diferentes condiciones experimentales (ambientales o temporales), con un arreglo en “matrices o tablas multivías”, en las cuales cada dato es originado por tres modos o vías: individuos x variables x condiciones (Del Médico y Vitelleschi, 2015), aunque en los últimos años se han desarrollado métodos para analizar más de tres vías. El método puede describirse sintéticamente como la obtención de una matriz promedio o matriz de consenso entre dos matrices provenientes de un análisis de componentes principales X_1 y X_2 asociadas a dos ambientes 1 y 2, cada una de las cuales contiene las coordenadas de las mismas p variables. Para ello se recurre a transformaciones de las coordenadas de cada variable en la matriz de componentes principales de forma tal que la suma cuadrática de las distancias entre cada par de puntos sea la menor posible, manteniendo las distancias relativas entre puntos dentro de cada matriz (Figura 1). Las transformaciones que cumplen con la restricción impuesta pueden ser: 1) traslación, 2) reflexión, 3) escalamiento, 4) rotación (Commandeur, 1991). Es posible la construcción de *biplots reescalados por óptima rotación* (Gower, 2004).

El método se puede extender a más de dos componentes principales, y encuentra aplicación en numerosos campos (Stegmann y Delgado Gómez, 2002; Gower y Dijkstra, 2004; Grice y Assad, 2009; Torcida y Perez, 2018.)

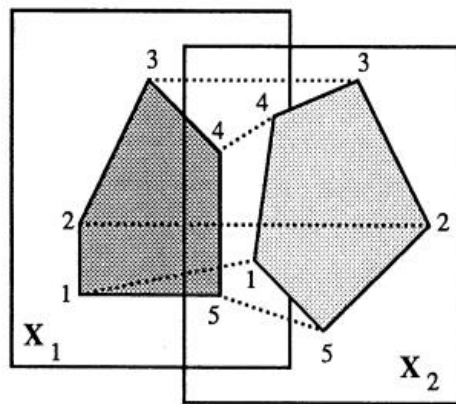


Figura 1. Matrices hipotéticas en dos dimensiones (Commandeur, 1991).

La estimación de la función de mínima distancia cuadrática se realiza por medio de la ecuación:

$$\sum_{j < k} (x_j - x_k)'(x_j - x_k) = \sum_{j=1}^n x_j'x_j - n^2 z'z$$

Donde:

$$z = \frac{1}{n} \sum_{j=1}^n x_j.$$

Obtenidas las transformaciones de las matrices originales, se obtiene la matriz de consenso como el promedio de las anteriores. La transformación Procrustes (escalamiento, rotación y traslación), en términos matriciales, pueden ser expresados del siguiente modo:

$$Y_k = \rho_k C_k H_k + T_k$$

Donde:

Y_k representa la transformación de procrustes, ρ_k el factor de escala, C_k es una matriz resultante de un Análisis de Componentes Principales (ACP) aplicado sobre una table X_k de datos conformadas por n filas (individuos) y p columnas (variables), H_k la matriz ortogonal de rotación de dimensión $p \times p$ y T_k la matriz traslación de dimensión $n \times p$ (Del Médico y Vitelleschi, 2015).

Los tres últimos elementos son encontrados minimizando la Suma de Cuadrados Residuales (SCR):

$$SCR = \sum_{j=1}^n \sum_{k=1}^K \Delta^2(P_i^{(k)}, G_i)$$

Donde:

$\Delta^2(P_i^{(k)}, G_i)$ es la distancia euclídea entre el punto $P_i^{(k)}$ y el centroide de las k puntos análogos de $P_i^{(k)}$, llamado G_i (Del Médico y Vitelleschi, 2015).

En estudios multi-ambiente, se suele asociar la matriz resultante de un análisis de componentes principales de p variables a cada uno de los ambientes analizados (Matteucci y Pla, 2009; Bruno y Balzarini, 2010; Zuliani *et al.*, 2012; Del Médico y Vitelleschi, 2015; Lavallo *et al.*, 2021), para luego obtener la matriz de consenso como un promedio de las matrices parciales de cada ambiente. Las sumas de cuadrados totales pueden ser descompuestas en sumas de cuadrados *entre ambientes* y sumas de cuadrados *intra-ambiente*.

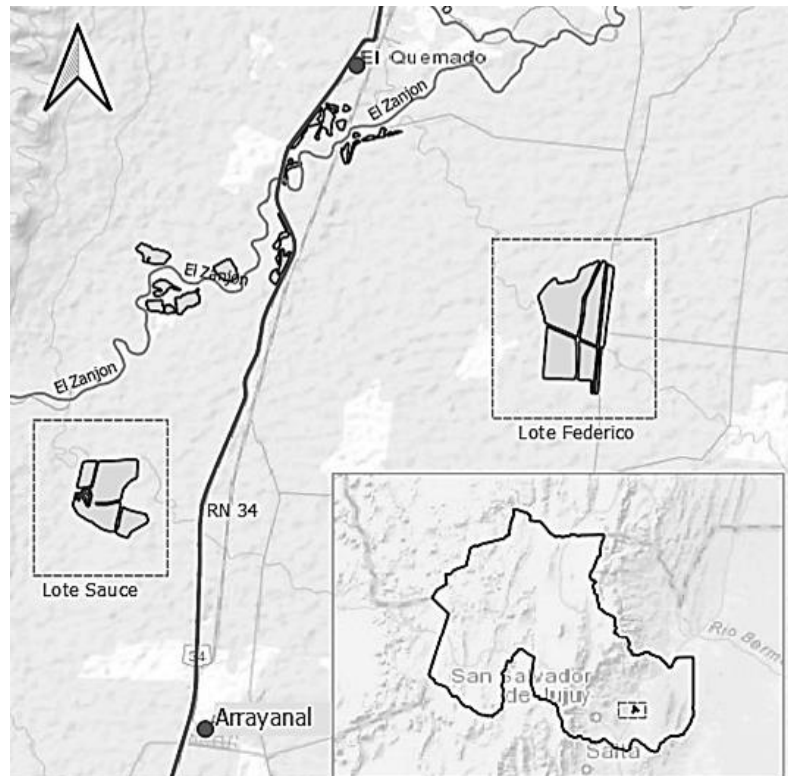
El objetivo de este trabajo fue comparar la variabilidad *dentro y entre ambientes* de diferentes variables en 3 genotipos de eucalipto en plantaciones de dos años realizadas en dos microambientes de la localidad de La Esperanza, provincia de Jujuy por medio de un análisis GPA.

2. METODOLOGÍA

Los datos analizados provienen de un trabajo de monitoreo de plantaciones forestales de eucaliptos realizados en agosto de 2017 en dos microambientes o lotes de plantación pertenecientes al Ingenio la Esperanza (San Pedro de Jujuy): lote *Federico* y Lote *Sauce* (Figura 2), con los genotipos *Eucalyptus camaldulensis*, *Eucalyptus saligna* y el híbrido *Eucalyptus grandis x camaldulensis*, todos clones plantados en 2015 con un marco de plantación 3×3 . El monitoreo incluyó 3.686 plantas distribuidas según detalle en Tabla 1.

Tabla 1. Sitios monitoreados y distribución de plantas por lote, según genotipo.

Sitio	Superficie (ha)	A.s.n.m. promedio (m)	Pendiente promedio (%)	Suelo (FAO)	Plantas/genotipo		
					<i>E. camaldulensis</i>	<i>E. saligna</i>	<i>E. g x c</i>
<i>Federico</i>	86	443	2	Fluvisol éutrico (Asociación La Esperanza)	182	1442	127
<i>Sauce</i>	45	482	2,5		72	1186	677

**Figura 2.** Ubicación de los lotes Sauce y Federico

El lote *Sauce* se encuentra ubicado en los límites de los faldeos orientales de la Sierra de Zapla en dirección NO de la localidad de Arrayanal, zona de selva montana de transición Yungas-Parque Chaqueño, transformada por cultivo de caña de azúcar, próximo a la confluencia del arroyo Zanjón y la Ruta Nacional N° 34. El lote *Federico* se halla emplazado a unos 3 km hacia el E de la Ruta Nacional N° 34, cercano a un macizo de bosque chaqueño atravesado por el río Lavayén, con presencia de afloraciones salinas y síntomas de toxicidad por salinidad en algunos árboles y arbustos. Ambos lotes se emplazan en un bloque edáfico clasificado como Fluvisol éutrico (FAO), con suelos jóvenes originados por depósitos aluviales y horizontes A y C, de textura media, bajo contenido de materia orgánica y un alto porcentaje de saturación de bases. La zona, de relieve suavemente ondulado, se encuentra completamente transformada conformando un mosaico de lotes destinados al cultivo de caña de azúcar de productividad variable por zonas, algunos de los cuales fueron abandonados por baja productividad, con presencia de afloraciones salinas más notorias en las zonas más bajas, hacia el E, donde se ubica el lote Federico.

Para el muestreo se realizaron transectas diagonales delimitadas por caminos o cortafuegos. Para la distribución de las transectas se tuvo en cuenta la forma específica de cada lote y la dirección de plantación. El espaciamiento entre ellas y la cantidad dependió de la forma y superficie de cada lote. En el inicio de cada transecta se midieron los 10 primeros individuos sobre una hilera,

luego a la misma altura se pasó a otra hilera, con una de por medio sin medir, para seguir con los siguientes 10 individuos a registrar. Esta metodología se siguió hasta llegar al final de la transecta. Esto determinó un avance de forma diagonal a las hileras de árboles, con lo que se intentó abarcar la mayor diversidad de situaciones posibles dentro del lote.

Las variables registradas en dicho monitoreo sobre cada planta fueron:

1. *Diámetro a la altura del pecho* (DAP), cuyo valor fue utilizado para estimar el *área basimétrica*. Se utilizó un calibre con vernier que se dispuso de forma perpendicular al eje del tallo a una altura de 1,3 m.
2. *Altura total de la planta*.
3. *Estado sanitario*:
 - a) Sano (s)
 - b) Muerto (m)
 - c) Falla (f): cuando no se encontró el árbol
 - d) Decrépito (de): con síntomas de marchitez

Se registraron otras variables sanitarias que no se incluyeron en este trabajo por representar una proporción relativamente baja de individuos.

A partir de las mediciones anteriores, se construyeron las siguientes variables para cada lote y variedad/híbrido:

1. *Área basimétrica* por ha en m²/ha, obtenida sobre individuos no muertos (AB).
2. *Altura promedio* en metros de árboles no muertos (H).
3. *Porcentaje de individuos sanos* (S).
4. *Porcentaje de individuos muertos* (M).
5. *Porcentaje de fallas* (F).
6. *Porcentaje de individuos con síntomas de marchitez* (DE).

Las variables AB y H constituyen atributos de productividad, y las variables S, M, F y DE pueden considerarse atributos de adaptación al sitio.

La variabilidad asociada a cada variedad de eucalipto y/o ambiente (lote) fue analizada por medio de un análisis de Procrustes Generalizado empleando la función *GPA* de la librería *FactoMineR* del programa estadístico R (versión 4.2.1), la cual considera casos perdidos (Lê, 2008). Dado que las variables presentaron diferentes escalas de medición, se trabajó con las mismas de forma escalada. Los grupos estuvieron constituidos por cada uno de los microambientes o lotes: *Federico* y *Sauce*. Para analizar las sumas de cuadrados, se recurrió al atributo "PANOVA" o Análisis de la Varianza del objeto procrustes. Por ejemplo, es posible analizar en qué medida cada variedad aportó variabilidad sobre los datos observados. Además, se realizó una lectura e interpretación del gráfico biplot correspondiente.

3. RESULTADOS

En líneas generales, *E. grandis x camaldulensis* presentó los mayores valores de área basimétrica por ha y altura media en ambos microambientes (Tabla 2).

Tabla 2. Tabulación de los valores de las variables analizadas por genotipo de eucalipto según microambiente o lote.

Genotipo/ híbrido	Lote Federico						Lote Sauce					
	AB m ²	H m	S %	M %	F %	DE %	AB m ²	H m	S %	M %	F %	DE %
<i>E. camaldulensis</i>	1,04	4,25	53,3	0,55	41,21	4,95	0,16	2,06	93,06	0	0	0
<i>E. grandis x camaldulensis</i>	2,73	7,25	28,35	8,66	41,73	0,79	4,88	7,44	55,83	1,48	22,16	0,44
<i>E. saligna</i>	0,77	4,37	21,98	4,16	64,22	2,64	4,19	6,81	65,51	4,3	23,27	2,11

Nota: para el análisis del estado sanitario se consideraron únicamente las variables *porcentaje de individuos sanos (S)*, *porcentaje de individuos muertos (M)*, *porcentaje de fallas (F)* y *porcentaje de individuos con síntomas de marchitez (DE)*. Es posible que la sumatoria de los valores de estas columnas no resulte exactamente 100, en virtud de otras variables que no se tuvieron en cuenta para el análisis.

E. camaldulensis presentó los mayores valores de porcentaje de individuos sanos y los menores de mortalidad y falla. Mientras que *E. saligna* presentó un comportamiento muy variable, con un alto porcentaje de falla en lote Federico y un alto porcentaje de individuos sanos en lote Sauce. Además, presentó un valor relativamente bajo de área basimétrica en lote Federico, pero alto en lote Sauce.

Una vez creado el objeto “procrustes” en R, se construyó un biplot de las dos dimensiones principales de la matriz de consenso, incluyendo los puntos correspondientes a cada matriz transformada (Figura 3).

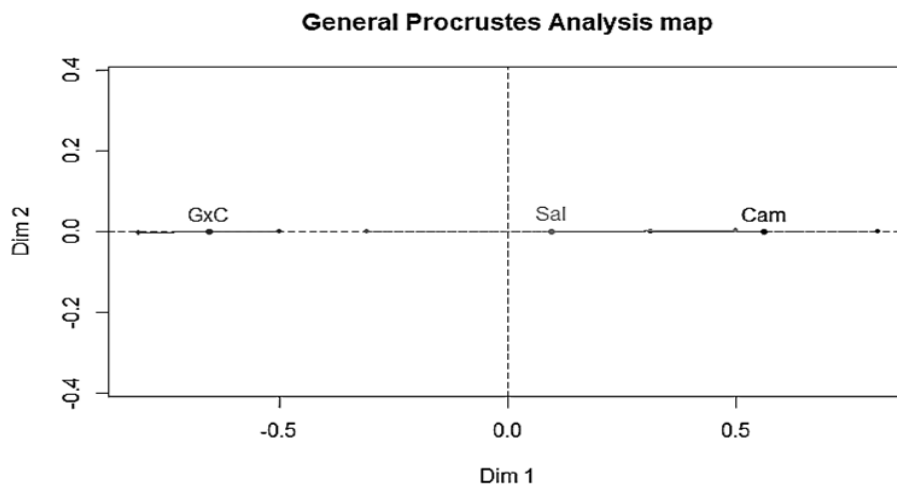


Figura 3. Biplot de matriz consenso y configuraciones transformadas

En este gráfico es posible reconocer cada punto G_i como el centroide de los correspondientes puntos $P_i^{(k)}$, habiendo minimizado la suma de cuadrados residual (SCR). El biplot muestra que la dimensión 1 explica prácticamente toda la variación observada en el grupo de variables

analizadas. De hecho, la suma de cuadrados explicada por esta dimensión representa un 99 % de toda la variación observada. En consecuencia, es suficiente tomar como referencia el eje asociado a la dimensión 1 para realizar el análisis de la variabilidad por variedad y/o lote. El híbrido *E. grandis x camaldulensis* muestra un comportamiento multivariado diferente (y opuesto) a *E. saligna* y *E. camaldulensis*. Lo anterior puede ser observado a través de la posición relativa y la distancia entre el centroide de cada matriz de consenso, y el par coordenado (0,0) del eje dimensional 1. Asimismo, la variabilidad observada para *E. grandis x camaldulensis* es la menor, de acuerdo a la longitud de los vectores que ligan los puntos de cada una de las matrices con el centroide determinado por la matriz de consenso dentro de cada grupo. En términos del modelo, hubo una buena concordancia entre las matrices parciales de cada matriz de datos. La variabilidad observada en *E. saligna* fue la mayor.

La variedad con mayor suma de cuadrados del modelo en el Análisis de la Varianza de Procrustes (PANAVA) es la que mayor variabilidad aportó al consenso, marcando su carácter diferencial con el resto (Tabla 3).

Tabla 3. PANAVA de la matriz de consenso por cada genotipo de eucalipto.

Variedad/híbrido	Suma cuadrados modelo	Suma cuadrados residuales	Suma cuadrados totales
<i>E. camaldulensis</i>	31,43	6,17	37,60
<i>E. grandis x camaldulensis</i>	42,86	2,40	45,26
<i>E. saligna</i>	0,89	16,26	17,14
Suma	75,17	24,83	100,00

Las variedades con mayor suma de cuadrados residual son las que mayores diferencias presentaron *entre* ambientes, y las que tienen menor valor de suma de cuadrados residual son las de comportamiento más homogéneo. El “peso” de la suma de cuadrados del modelo asociado a *E. grandis x camaldulensis* fue el mayor, indicando que fue el genotipo que mayor aportó a la variabilidad observada *dentro* de cada microambiente, seguido por *E. camaldulensis*. En contrapartida, *E. saligna* realizó la mayor contribución relativa a la suma de cuadrados residuales, lo que puede interpretarse como que fue la variedad menos estable o más variable *entre* microambientes.

4. DISCUSIÓN

Una característica deseable al momento de evaluar la adopción de un genotipo para un determinado ambiente es su *estabilidad*, es decir, su capacidad para expresar determinados atributos de productividad y adaptación (o simplemente *atributos*) de forma consistente en diferentes ambientes. Analizar la contribución relativa que tienen el genotipo y el ambiente sobre un conjunto de atributos supone también identificar ambientes potencialmente más favorables para determinados genotipos.

Desde el punto de vista del factor ambiente, el lote *Sauce*, ubicado hacia el O de la zona en estudio con una pendiente media mayor y con menor evidencia de afloramientos salinos que en el lote Federico, estuvo asociado a mayores valores de atributos de adaptación para todos los genotipos, y mayores valores de atributos de productividad en dos de los tres genotipos, por lo que se puede afirmar que fue el lote con mejor *calidad de sitio*.

Desde el punto de vista del factor genotipo, *E. grandis x camaldulensis* fue el material más estable con la menor contribución de variabilidad residual o entre ambientes en el PANOVA, seguido de *E. camaldulensis*. En este punto se debe notar que el híbrido presentó los mayores valores medios de atributos productivos, pero *E. camaldulensis* presentó los mejores valores en atributos de adaptación, lo cual era esperable por sus cualidades de rusticidad y plasticidad (Pardos, 2007). De todas formas, realizando un sencillo cálculo comparativo en términos de volumen, esto es:

$$Vol_{m^3/100\ ind} = \frac{(AB \cdot H)}{1111} \cdot S$$

Se observa que *E. grandis x camaldulensis* fue el genotipo con mayor productividad: 0,50 $m^3/100$ individuos y 1,82 $m^3/100$ individuos en los lotes Federico y Sauce respectivamente, debido principalmente a los atributos de productividad aportados por *E. grandis*. La productividad de *E. camaldulensis* fue de 0,21 $m^3/100$ individuos y 0,03 $m^3/100$ individuos en los lotes Federico y Sauce respectivamente, y la productividad de *E. saligna* fue de 0,07 $m^3/100$ individuos y 1,68 $m^3/100$ individuos en los lotes Federico y Sauce respectivamente.

E. saligna fue el genotipo menos estable con un marcado contraste de sus atributos entre sitios, superando ampliamente en productividad a *E. camaldulensis* en el lote Sauce. Este resultado es coherente con la literatura especializada, la cual indica que suelos drenados sin acumulación de sales, en laderas, son adecuados para *E. saligna* (Pardos, 2007; Speranza *et al.*, 2020).

Como principal limitación de este estudio se puede señalar que, al haber aplicado el análisis a plantaciones realizadas en solo dos sitios, no se aprovechan las ventajas que tiene el método Procrustes para discernir las contribuciones que sobre la variabilidad general observada tienen el genotipo y el ambiente, cuando se incluyen más ambientes.

5. CONCLUSIONES

El Análisis de Procrustes Generalizados (GPA) permitió sintetizar y describir de forma completa la variabilidad observada de dos especies y un híbrido de eucalipto en dos microambientes de la Esperanza (provincia de Jujuy), por medio de ajustes de matrices parciales y la generación de una matriz de consenso.

Constituye una alternativa muy recomendable para el análisis multivariado de la variabilidad por ambientes, ya que permite comparar los aportes ponderados que cada genotipo realiza *dentro* y *entre* ambientes.

Se pudo observar que el clon híbrido *E. grandis x camaldulensis* de dos años fue el genotipo más estable *entre* ambientes, y que contribuyó en mayor medida con la variabilidad *dentro* de cada ambiente, además de presentar los mayores valores de AB por ha y altura promedio.

E. camaldulensis también presentó una relativa estabilidad entre ambientes. Además, fue el que presentó los menores valores de mortalidad y tasas de falla.

E. saligna fue el genotipo que presentó mayor variabilidad *entre* ambientes, es decir, fue el menos estable.

Se espera que este trabajo pueda contribuir al reconocimiento del método de GPA como una técnica óptima para el análisis multivariado de componentes de variabilidad en estudios multiambientales en el campo forestal.

6. REFERENCIAS BIBLIOGRÁFICAS

- Bruno, C. y Balzarini, M. 2010. Ordenaciones de material genético a partir de información multidimensional. *Revista FCA UNCUYO*. 42(2): 183-200.
- Commandeur, J. 1991. *Matching configurations*. Department of Psychometrics and Research Methodology. DSWO Press, Leiden University. Países Bajos. 190 p.
- Del Médico, A. y M. Vitelleschi. 2015. *Análisis de Procrustes Genetalizado. Una aplicación en el área agrícola*. Actas de las Vigésimas Jornadas "Investigaciones en la Facultad" de Ciencias Económicas y Estadística, Universidad Nacional de Rosario.
- Gower, J. C. 1975. Generalized Procrustes Analysis. *Psychometrika* 40: 33-50.
- Gower, J. y G. Dijksterhuis. 2004. *Procrustes problems*. Oxford Press: Oxford Statistical Sciences Series. 194 p.
- Gower, J. 2004. The geometry of biplot scaling. *Biometrika*. 91(3): 705-714.
- Grice, J. y K. Assad. 2009. Generalized Procrustes analysis: a tool for exploring aggregates and persons. *Applied Multivariate Research*, 13(1): 93-112.
- Lavalle, A. et al. 2021. Propuesta metodológica para la caracterización de accesiones en Bancos de Germoplasma mediante Análisis de Procrustes Generalizado aplicado a ensayos incompletos pero conectados. *Revista FCA UNCUYO*. 53(1): 35-45.
- Lê, S.; J. Josse y F. Housson. 2008. FactorMineR: an R package for multivariate analysis. *Journal of Statistical Software*. *Journal of Statistical Software*. 25(1): 1-18.
- Matteucci, S. y L. Pla. 2009. Procrustes analysis as a tool for land management. *Elsevier: Ecological Indicators*. 10: 516-526.
- Pardos, J. 2007. Perspectiva fisiológica en la producción y mejoramiento del Eucalipto (con énfasis en *Eucaliptus globulus* Labill). *Boletín del CIDEU* (3): 7-55.
- Speranza, F.; L. Cosimi y C. Badiner et al. 2020. *Guía de Aptitud de Especies Forestales en Yungas*. Tapirus Estudios. 49 p.
- Stegmann, L. y D. Delgado Gómez. 2002. *A Brief Introduction to Statistical Shape Analysis. Informatics and Mathematical Modelling*. Technical University of Denmark. Marzo de 2002. 15 p.
- Torcida, S. y S. Perez. 2018. Análisis de Procrustes y el estudio de la variación morfológica. *Revista Argentina de Antropología Biológica*. 14(1): 131-141.
- Zuliani, P. et al. 2012. Caracterización de poblaciones nativas de maíz mediante Análisis de Procrustes Generalizado y Análisis Factorial Múltiple. *Revista FCA UNCUYO*. 44(1): 49-64.

