

Análisis factorial en ensayos multiambientales comparativos de cultivares de trigo para calidad de grano

Del Vecchio E; Balzarini M y Abbate P.
Universidad Nacional de Córdoba, CONICET, INTA Balcarce.
estefaniadelvecchio@gmail.com

Resumen

*El aumento sostenido de la producción agrícola orienta a los mejoradores a alcanzar mayores rendimientos y mejores calidades de commodities y productos agroalimentarios en un marco de agricultura sostenible. La elección de un cultivar de trigo (*Triticum aestivum* L.) por sus variables de calidad es clave para definir el destino de la cosecha. A tal fin los investigadores tratan con ensayos que involucran varios ambientes (E) para comparar genotipos (G) a través de múltiples variables. Técnicas multivariadas como análisis factorial (FA), análisis de componentes principales (ACP), y análisis de conglomerados (AC) se han utilizado ampliamente para identificar grupos. FA y ACP sintetizan los datos en unos pocos componentes que retienen la cantidad máxima de información contenida en el mayor número de variables originales, haciendo que la interpretación sea más fácil. En el presente trabajo, se muestra la complementación de FA, ACP, y AC para diferenciar grupos de cultivares comerciales definidos por variables de calidad importantes en la industria y comercialización del trigo. Se analizaron matrices GxE de datos de calidad de 235 genotipos de trigo harinero, cultivados en 19 ambientes, divididos en 6 subregiones trigeras de Argentina, durante diez años. Nueve atributos de calidad importantes en la comercialización se utilizaron para determinar los subgrupos de calidad. Se obtuvieron 4 grupos bien definidos de acuerdo con estas variables de calidad.*

Palabras Clave

Análisis de Componentes Principales, Análisis de Conglomerado, mejoramiento de plantas, clasificación de genotipos.

Introducción

La interacción GxE (Kang et al., 2004) puede impactar fuertemente en el rendimiento y la calidad de las cosechas y por tanto el mercado de valores de los cultivos (Dardanelli et al., 2006; Munaro et al., 2011).

La RET (Red de ensayos comparativos de variedades de trigo) es una red de ensayos a nivel nacional con el propósito de orientar al productor sobre el comportamiento de los distintos cultivares, es decir, variedades, de trigo pan comercializados en Argentina, brindando información para elegir correctamente el cultivar más apropiado para cada localidad y sistema de producción (Abbate, 2016). Argentina es uno de los principales productores de Trigo, entre 18-20 millones de toneladas cada año, el cual tiene diferentes destinos comerciales, el consumo interno uno de los principales que en cuanto a calidad demandan una mayor concentración de Gluten, un alto peso hectolitrico y parametro W. Para exportación en lo cual se demandan un alto nivel de Peso Hectolitrico y concentración de Proteína. Y la Industria brasilera que requieren de este trigo para mejorar la calidad del trigo de exportación demandan mayor concentración de Proteína, alta estabilidad del farinograma y alto parametro W.

Material y Métodos.

Se realiza un Análisis de Componentes Principales (ACP), que tratará de representar “de forma clara y ordenada”, la variedad de los comportamientos observados en datos de calidad de cultivares comerciales de trigo, de 10 temporadas, entre 2000-2015, en 19 localidades en 6 zonas triguera de Argentina, se evaluaron 235 cultivares mediante un conjunto de 9 variables más significativas utilizadas para determinar la calidad comercial: gluten húmedo, gluten seco, proteína, parámetro W, volumen pan, estabilidad farinografica, peso hectolitrico, rendimiento harinero y Cenizas.

El estudio del ACP se realiza a través del Análisis Factorial, el cual intenta identificar variables subyacentes, o factores que expliquen la configuración de correlaciones dentro de un conjunto de variables observadas (Abdi et al., 2013). Es necesario que entre las variables de estudio haya una estructura importante de correlación. El ACP y AF se realizaron sobre la matriz de correlaciones, dando igual importancia a todas las variables originales y permitiendo utilizar el tamaño de los autovalores como un criterio para decidir si el número de factores que estarán presentes en la solución factorial. Luego se realiza un AC combinado con ACP colocando a los cultivares en grupos sugeridos por datos. El número de grupos se obtiene a través del número óptimo de clúster para la implementación de agrupamiento jerárquico a partir de los factores del ACP, utilizando la librería Nbclust.

Todos los análisis y gráficos se realizaron con el software R.

Resultados.

Consideramos que utilizar cuatro factores es lo mas adecuado, los cuales explican un 85,8% de la varianza. Como podemos observar en la figura 1, primer factor explica el 40,2% de la varianza correlacionado con las variables proteicas de calidad (gluten húmedo, gluten seco y proteína), el segundo factor explica el 22,1% de la varianza relacionado con variables que componen el potencial de masa, estabilidad, volumen pan y parámetro W. El tercer factor explica el 13,1% de la varianza correlacionado con el peso hectolitrico, y el cuarto factor explica el 10,4 % de la varianza correlacionado con el rendimiento harinero.

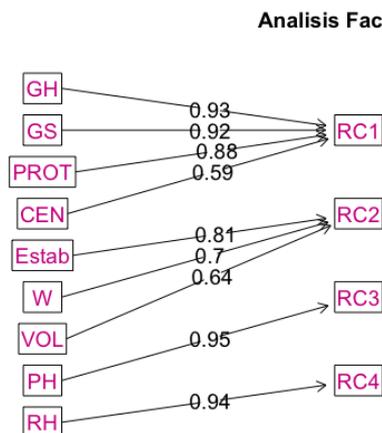


Figura 1: Relación entre las variables originales y los 4 factores del ACP.

A partir de los cuatro factores que obtuvimos con el ACP, realizamos clúster de cultivares. Para obtener el numero optimo de clúster utilizamos la librería NbClust, el cual analiza 30 índices diferentes de formación de clúster y arroja una cantidad de clúster recomendada, en la figura 2 se observa que cinco clústeres es la cantidad recomendados por la mayoría de los índices.

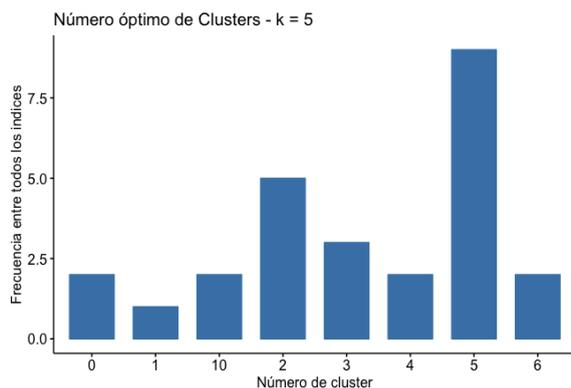


Figura 2: Número óptimo de clúster de acuerdo con 30 índices de clúster.

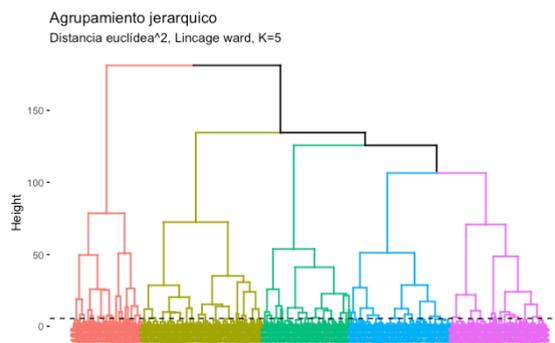


Figura 3: Dendrograma de análisis de cluster jerárquico.

Utilizamos la distancia euclídea y el criterio WARD para construir los cluster jerárquicos como se puede ver en la figura 3, hemos obtenido las clases de los cluster a partir de las variables, en la tabla 1 podemos identificar una clasificación de los cluster según los parámetros de calidad.

Tabla1: Clasificación de los cluster de acuerdo a las características de calidad.

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Alto RH, CEN, y VOL variables que componen el potencial de la masa.	Alto PH Bajo Estab, VOL, PROT, GH, GS y W.	Altos valores de todas las variables y bajo PH.	Valores de Estab, W, RH, y PH muy bajos.	Alto VOL, GH. y GS.

Conclusión.

Podemos definir que las cultivares de Trigo pan analizados pueden agruparse en 5 grupos bien definidos y clasificados de acuerdo a los parámetros de calidad más significativos para la comercialización.

Referencias.

Abbate P.E. 2016. Principales atributos de calidad de los cultivares de trigo disponibles en la Subregión triguera de Argentina. EEA Balcarce, INTA.

Dardanelli J., Balzarini M., Martínez M., Cuniberti M., Resnik S., Ramunda S., Herrero R. and Baigorri H. 2006. Soybean maturity groups, environments, and their interaction define mega-environments for seed composition in Argentina. *Crop Sci.* 46(5):1939-1947.

Abdi H., Williams L. J., Valentin D. 2013. Multiple factor analysis: Principal component analysis for multitable and multiblock data sets. *Wiley Interdisciplinary Reviews: Computational Statistics*, 5 (2), pp. 149-179.

Kang M.S. and Gauch H.G. 1996. Genotype-by-environment interaction. CRC Press, Boca Ratón, Florida, USA, 339 pp.

RStudio Team (2015). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL <http://www.rstudio.com/>.