



UNIVERSIDAD NACIONAL DE ROSARIO
FACULTAD DE CIENCIAS AGRARIAS
Campo Experimental Villarino C.C. N° 14
S2125ZAA | Zavalla | Santa Fe | Argentina
Tel/Fax: +54 0341 497-0080
www.fcagr.unr.edu.ar
agro@unr.edu.ar



FACULTAD DE CIENCIAS AGRARIAS
UNIVERSIDAD NACIONAL DE ROSARIO

“Estudio de la interacción genotipo ambiente en ensayos de soja [*Glycine max (L.) Merrill*] para la identificación de mega ambientes entre Argentina y Estados Unidos”

Autor: Ing. Agr. Carlos Agustín Basile

TESIS PARA OPTAR AL TÍTULO DE MAGISTER EN GENÉTICA
VEGETAL

DIRECTOR: Ph.D. Les Kuhlman

CO- DIRECTOR: Dr. Enrique Kreff

AÑO 2017



Universidad Nacional de Rosario
Facultad de Ciencias Agrarias
Escuela de Posgrado

Estudio de la interacción genotipo ambiente en ensayos de soja [*Glycine max (L.) Merrill*] para la identificación de mega ambientes entre Argentina y Estados Unidos

Carlos Agustín Basile

Ingeniero Agrónomo- Universidad Nacional de Lomas de Zamora

Esta tesis es presentada como parte de los requisitos para optar el grado académico de Magister en Genética Vegetal, de la Universidad Nacional de Rosario y no ha sido previamente presentada para la obtención de otro título en ésta u otra Universidad. La misma contiene los resultados obtenidos en investigaciones llevadas a cabo en ambientes de Estados Unidos y Argentina, durante el período comprendido entre 2014 y 2016, bajo la dirección de

Nombre y firma del Maestrando:

Nombre y firma del Director:

Nombre y firma del Co – Director:

Defendida: 22 de Octubre del 2018

AGRADECIMIENTOS

Quiero agradecer a las personas que de una u otra forma colaboraron, en la realización del presente trabajo de investigación.

A la empresa Pioneer Hi-Bred International por la oportunidad brindada, el soporte financiero y los recursos que me brindó para realizar la Maestría en Genética Vegetal.

Dr. Enrique Kreff por el tiempo dedicado y sus valiosos aportes en todo el proceso de la redacción de la tesis.

Ph.D. Les Kulhman por sus consejos y aportes al planteo del trabajo.

MSc. Marcelo Serre por su permanente apoyo desde antes de la postulación a la maestría, por su estímulo hacia esta aventura y por sus valiosos consejos en mi formación profesional.

Los Integrantes de la estación experimental Pioneer Pergamino por su contribución en la realización de los ensayos a campo y toma de datos

Mis padres, por haberse esforzado siempre en brindarme una educación de calidad, y haberme motivado y apoyado en mis estudios.

Mi familia Julia, Mateo y Agostina, por su soporte incondicional, el cariño y afecto brindado durante todo este tiempo.

INDICE

• LISTADO DE ABREVIATURAS.....	6
• RESUMEN.....	8
• ABSTRACT.....	9
• INTRODUCCIÓN.....	10
• HIPÓTESIS.....	18
• OBJETIVOS.....	18
• MATERIALES Y MÉTODOS.....	19
○ ZONA DE ESTUDIO.....	19
○ GENOTIPOS.....	19
○ LOCALIDADES EN ESTADOS UNIDOS.....	20
○ LOCALIDADES EN ARGENTINA.....	22
○ DATOS FENOTÍPICOS.....	23

○ ANÁLISIS ESTADÍSTICO.....	24
● RESULTADOS.....	28
○ GRUPO “CORTOS”.....	28
○ GRUPO “LARGOS”.....	38
○ CARACTERIZACIÓN DE GENOTIPOS Y AMBIENTES EN ARGENTINA.....	47
● DISCUSIÓN.....	52
○ GRUPO “CORTOS”.....	52
○ GRUPO “LARGOS”.....	54
● CONCLUSIONES.....	58
● BIBLIOGRAFÍA.....	60
● ANEXO.....	64

LISTADO DE ABREVIATURAS

Arg.	=	Argentina
Bs. As.	=	Buenos Aires
CAM	=	Coordenada ambiental media
CP	=	Componente Principal
CP1	=	Primer Componente Principal
CP2	=	Segunda Componente Principal
DBCA	=	Diseño en Bloques Completos al Azar
E	=	Ambiente
EAP	=	Eje del ambiente promedio
ECR	=	Ensayos comparativos de rendimiento
EE. UU.	=	Estados Unidos de Norteamérica
G	=	Genotipo
GA	=	Genotipo por año
GE	=	Genotipo por ambiente
GL	=	Genotipo por Localidad
GLA	=	Genotipo por Localidad por Año
GM	=	Grupo de Madurez
IL	=	Illinois
kg/ha	=	kilogramos por hectárea
km	=	kilómetros
KS	=	Kansas
MO	=	Misuri
MS	=	Misisipi
OH	=	Ohio
R	=	Repetición
REML	=	Máxima verosimilitud restringida
SF	=	Santa Fe
SVD	=	Descomposición por valor singular

- ΔG = Ganancia genética
- σ^2_g = Variancia genotípica
- σ^2_{ge} = Variancia genotípica por ambiente
- σ^2_{gl} = Variancia genotípica por localidad
- σ^2_p = Variancia fenotípica

RESUMEN

El cultivo de soja [*Glycine max* (L.) Merr.] tiene una amplia difusión y ocupa el cuarto puesto en rendimiento a nivel mundial detrás del arroz, maíz y trigo; EE. UU. y Arg, son el primer y tercer productor mundial de soja en el mundo, respectivamente. Ambos países poseen una amplitud de ambientes diversos en cuanto a clima (precipitaciones, fotoperiodo, temperaturas, etc.) y suelo que permiten evaluar un gradiente importante de GM. Todos los años se evalúan en ensayos comparativos de rendimiento multi-ambientales (ECR) diversas variedades en distintos países alrededor del mundo. El objetivo de los ECR es no solo evaluar variedades sino también la identificación de mega-ambientes. Aunque el rendimiento de un genotipo es el resultado combinado de los efectos del G, el E y la interacción GE, solo G y GE son relevantes para la evaluación de las variedades y la identificación de mega-ambientes. En este trabajo se evaluaron en ECR en ambientes de EE. UU. y Arg. variedades de soja con un rango de madurez entre 3.7 a 4.9 durante dos campañas en cada país. En EE. UU. se agruparon las variedades en GM “cortos” (3.7 a 4.4) y otro GM llamado “largos” (4.2 a 4.9). En los ambientes de Arg. también se registraron las variables días a R1, días a R8, altura en R8 (cm), vuelco (escala) y peso de 1000 granos (gr). Los objetivos del trabajo fueron 1) estimar los componentes de variancia de G, E y GE para la variable rendimiento en grano, 2) mediante gráficos biplot GGE estudiar los efectos de G y e interacción GE y existencia de mega-ambientes entre ambos países, 3) identificar genotipos estables y con buena adaptación y 4) caracterizar los ambientes de Arg. y los genotipos utilizados mediante las variables fenotípicas evaluadas. Los resultados del análisis de variancia mostraron que para ambos grupos (“cortos” y “largos”) la magnitud de la interacción GE respecto a G no justifica dividir la región en estudio en distintos mega-ambientes. Los ambientes de Estados Unidos y de Argentina mostraron un patrón de heredabilidad similar, una correlación positiva y mega-ambientes comunes, permitiendo concluir que existen ambientes comunes para seleccionar variedades para ambos países para el rango de madurez empleado. Se identificaron genotipos estables y adaptados a la región. Los resultados obtenidos para caracterizar los genotipos en los ambientes de Arg. permitieron concluir que el comportamiento de las variedades con germoplasma seleccionado para los ambientes de EE. UU. se comportó de manera similar según su GM con variedades locales de Arg.

ABSTRACT

Soybean [*Glycine max* (L.) Merr.] production is widespread and ranks fourth in global yield behind rice, maize and wheat. EE. UU. and Arg. are the world's first and third soybean producer in the world respectively. Both countries have a wide range of environments in terms of climate (precipitation, photoperiod, temperatures, etc.) and soil that allow testing a considerable MG gradient. Several yield multi-environmental trials (MET) are evaluated each year in different countries around the world. The objective of MET is not only to evaluate varieties but also to identify mega-environments. Although the measured yield is a combined result of G, E and GE interaction effects, only G and GE are relevant for crop evaluation and mega-environment identification. In this work, soybean MET with maturity range from 3.7 to 4.9 were tested during two seasons in EE. UU. and Arg. environments respectively. In the EE. UU. environments, the varieties were grouped into "short" (3.7 to 4.4) and "long" (4.2 to 4.9) MG's. In Arg. environments, the following traits were recorded: days to R1, days to R8, height at R8 (cm), lodging (scale) and weight of 1000 kernels (gr). The objectives of this study were 1) to estimate the variance components of G, E and GE interaction for yield trait, 2) to study the effects of G and GE interaction and the existence of mega-environments through biplot GGE graphs between both countries, 3) to identify stable and well-adapted genotypes, and 4) to characterize the Arg. environments and the genotypes evaluated by the phenotypic variables registered. The results of the analysis of variance showed that for both groups ("short" and "long") the magnitude of the interaction GE with respect to G does not justify dividing the region under study in different mega-environments. The environments in EE. UU. and Arg. showed a similar pattern of heritability, positive correlation and common mega-environments, allowing to conclude that there are common environments to select varieties for both countries for the range of maturity used in this study. Stable and well adapted genotypes were identified for the region. The results obtained to characterize the genotypes in Arg. environments allowed to conclude that the performance of the varieties with germplasm selected for the EE. UU. environments behaved in a similar way according to their MG with Arg. local varieties.

INTRODUCCIÓN

La soja [*Glycine max* (L.) Merr.] es una especie leguminosa originaria del este asiático, fue domesticada en China en el siglo XI a. C. y desde ahí llevada a diferentes lugares del mundo para su cultivo. Actualmente, es uno de los cultivos más antiguos de la humanidad y constituye la fuente de aceite y proteínas vegetales de mayor importancia.

Argentina es el tercer productor mundial de grano, el primer exportador de aceite y el segundo de harina de soja, constituyendo el complejo soja el mayor generador de divisas para el país. Además, el cultivo de soja es el principal cultivo de la Argentina no sólo por la producción, sino por la superficie ocupada. En la actualidad representa el 55% de los casi 37 millones de hectáreas que se siembran, seguida muy de lejos por los cultivos de maíz y trigo que en conjunto representan el 26%, menos de la mitad (Ybran y Lacelli, 2016). En Estados Unidos sobre unas 129 millones de hectáreas, poco más del 26% se destina al cultivo de soja, mientras que un 45% son sembradas con maíz y trigo (USDA, 2017).

La soja presenta un conjunto de características fisiológicas que la diferencian del resto de los cultivos más difundidos a nivel mundial y que condicionan su manejo y utilización. Entre sus principales características se puede mencionar: una importante respuesta fotoperiódica, una alta plasticidad reproductiva y producción de semillas con elevados contenidos de proteína y aceite (Baigorri et al., 1997).

Para obtener elevados rendimientos son necesarios un adecuado conocimiento del desarrollo y crecimiento del cultivo y la cuantificación de la influencia de los principales factores genéticos y ambientales que lo afectan. Esto posibilita una adecuada elección de cultivares y la aplicación de las prácticas de manejo más convenientes para cada ambiente. (Baigorri et al., 1997).

Debido a la respuesta fotoperiódica, las variedades de soja están adaptadas a bandas latitudinales no mayores a 160 a 240 km de distancia. En el hemisferio norte, al norte de estas bandas las variedades florecen más tardíamente y corren riesgos de no poder completar su ciclo antes de las

heladas, mientras que al sur de estas bandas maduran precozmente y rinden menos que las variedades adaptadas. En el hemisferio sur ocurre lo opuesto.

Las variedades están agrupadas en 13 grupos de madurez designados como 000 al X y se clasifican dentro de cada grupo de madurez de acuerdo al largo del periodo entre siembra y madurez, siendo las 000 las más precoces. Dentro de cada grupo de madurez la duración del ciclo puede variar entre 10 a 18 días. En Estados Unidos los grupos 00, 0, I, y II se siembran en las regiones del Norte del país y al Sur de Canadá. Los grupos II, III y IV están adaptados a la región central del cinturón maicero de Estados Unidos. El GM IV se adapta a una amplia gama de latitudes, incluyendo la mitad sur de los estados de Kansas, Misuri, Illinois e Indiana, así como la totalidad de los estados de Oklahoma y Kentucky. Los grupos V y VI son utilizados principalmente en la costa este y en la zona central sur de Estados Unidos. Los grupos VII y VIII están adaptados a la región sur de Estados Unidos y a lo largo del golfo.

Para la mejora continua del germoplasma y de las variedades que se inscriben en el mercado existen programas de mejoramiento del cultivo de soja en los cuales se evalúan genotipos en ensayos de rendimiento para seleccionar el de mejor comportamiento. La interacción entre el genotipo y el ambiente da como resultado diferencias significativas en el comportamiento de los genotipos cuando son evaluados en diferentes localidades (Gauch y Zobel, 1997). Para entender los efectos del genotipo y del ambiente sobre el comportamiento del cultivo de soja, se realizan ensayos multi-ambientales cada año alrededor del mundo para identificar genotipos superiores y evaluar tanto la relación entre ambientes, como la determinación de mega-ambientes (Yan et al., 2000).

La interacción Genotipo \times Ambiente (GE) es la respuesta diferencial de diferentes genotipos a través de un rango de ambientes, es una característica universal relacionada a los seres vivos desde bacterias a plantas y humanos (Kang, 1998). Baker (1988) definió la interacción GE como el fracaso de los genotipos para lograr el mismo rendimiento relativo en diferentes ambientes. La identificación de los caracteres que contribuyen al rendimiento, el conocimiento de las interacciones GE y la estabilidad del rendimiento son importantes para seleccionar nuevos

cultivares con mejor adaptación a las restricciones ambientales que prevalecen en los ambientes objetivo.

La alta estabilidad del rendimiento generalmente se refiere a la capacidad de un genotipo para tener un comportamiento consistente, ya sea a niveles de rendimiento altos o bajos, en una amplia gama de ambientes. La mayoría de las medidas de estabilidad se relacionan con cualquiera de dos conceptos contrastantes de estabilidad: "estática" (Tipo 1) y "dinámica" (Tipo 2) (Lin et al., 1986; Becker et al., 1988). La estabilidad estática es análoga al concepto biológico de homeostasis: un genotipo estable tiende a mantener un rendimiento constante a través de los ambientes. La estabilidad dinámica implica para un genotipo estable una respuesta de rendimiento en cada entorno que es siempre paralelo a la respuesta media de los genotipos probados.

En relación a la interacción GE, los efectos más importantes para caracterizar y seleccionar cultivares es la de tipo "*crossover*" afectando los genotipos de mayor rendimiento. Estos efectos implican una diferente clasificación y diferentes desviaciones entre genotipos de una localidad a otra (Baker, 1988). Sin embargo, todos los efectos de la interacción GE derivados de la falta de correlación genética entre ambientes (incluidos los relacionados con material de bajo rendimiento y no necesariamente de tipo "*crossover*") pueden ser relevantes si los resultados de un conjunto de datos dado se extrapolan para producir información sobre efectos GE que es probable que se cumplan en el mejoramiento para una región objetivo (Muir et al., 1992; Cooper et al., 1996). Se puede esperar una interacción importante cuando existe, por una parte, una amplia variación entre los genotipos para los caracteres morfo-fisiológicos que confieren tolerancia a (o escapan) uno o más tipos de estrés y, por otra parte, una amplia variación entre los ambientes de incidencia del mismo estrés (determinado por factores climáticos, de suelos, bióticos y de gestión). Los cultivares que se caracterizan por bajos niveles de heterogeneidad (por ejemplo, una línea pura, un clon, un híbrido simple) tienden a interactuar con el medio ambiente más que tipos con características opuestas (por ejemplo, población de polinización abierta, mezcla de líneas puras). Debido a que la menor riqueza en genes adaptativos implicados por su estructura genética los hace más susceptibles a la variación de las condiciones ambientales (Becker y León, 1988).

Las regiones de crecimiento deben subdividirse en mega-ambientes relativamente homogéneos y genotipos apropiados deben ser evaluados en cada mega-ambiente. Los mega-ambientes se utilizan para priorizar los recursos en los programas de mejoramiento, para utilizar de una forma más eficiente el germoplasma y el intercambio de información entre programas de mejoramiento, para incrementar la heredabilidad y el progreso genético dentro de ambientes relativamente bien definidos y predecibles, para incrementar la eficiencia en los ensayos de rendimiento de los programas de mejoramiento y para ubicar los genotipos en áreas de producción apropiadas (Gauch y Zobel, 1997).

Es necesario contar con datos de varias campañas con los mismos genotipos y ambientes para decidir si la región puede dividirse o no en mega-ambientes. Si a través de los años las localidades se agrupan de igual modo y además los genotipos superiores de cada mega-ambiente son los mismos, entonces sí se puede capitalizar la interacción seleccionando para cada mega-ambiente (Yan y Rajcan, 2002, Yan 2007a).

Los mega-ambientes son definidos como un conjunto de localidades que comparten un grupo de genotipos que representan el mejor comportamiento a través de los años. Usualmente un mega-ambiente representa un área relativamente amplia, no necesariamente continua, con condiciones similares de estrés biótico y abiótico, requerimientos de sistemas de cultivo, preferencias del consumidor y de volúmenes de producción (Braun et al., 1996).

Los mega-ambientes pueden ser simples o complejos. Un mega-ambiente simple no involucra cruces de interacción GE, mientras que un mega-ambiente complejo involucra cruces de interacción GE que no se repiten a través de los años. Para un mega ambiente simple, es suficiente uno o unos pocos ambientes para evaluar de manera precisa a los cultivares, mientras que para seleccionar cultivares superiores en mega-ambientes complejos se requieren distintas localidades de evaluación a través de toda la región durante diferentes años (Yan y Rajcan, 2002). Los resultados obtenidos de una determinada localidad de un mega-ambiente tienen una buena confianza de predictibilidad para otras localidades dentro del mismo mega-ambiente (Gauch y Zobel, 1997).

La técnica gráfica “biplot GGE” para evaluación de genotipos y ambientes fue desarrollada por Yan et al. (2000), basándose en la técnica gráfica del biplot desarrollada por Gabriel (1971). El biplot GGE, también llamado biplot SREG (Cornelius et al., 1996; Crossa y Cornelius, 1997), es una herramienta visual efectiva en el análisis de mega-ambientes que remueven el efecto de sitio y solo expresan a la respuesta en función de $G + GE$ (Yan et al., 2000). Consiste en un polígono irregular y un set de líneas dibujadas a partir del origen del biplot formando distintos ángulos. En los vértices del polígono están marcados los genotipos extremos o más alejados del origen del biplot en distintas direcciones, de tal manera que todos genotipos estudiados están contenidos dentro del polígono. La línea que parte del origen del biplot intercepta en forma perpendicular a cada lado del polígono dividiéndolo en sectores y cada sector tiene un genotipo superior (el de mayor rendimiento) ubicado en los vértices del polígono (Yan y Rajcan, 2002). Si todos los ambientes se ubican en un solo sector, esto indica que un solo genotipo es el de más alto rendimiento en todos los ambientes, por el contrario, si los ambientes estudiados se ubican en diferentes sectores significa que diferentes genotipos obtuvieron el más alto rendimiento o son superiores (patrón “which-won-were”). Los genotipos ubicados en los sectores donde no hay localidades son consideradas de pobre comportamiento en rendimiento (si esta fuese la variable estudiada) en la mayoría de las localidades evaluadas (Yan, W. 2001). Es importante destacar que este método de análisis no diferencia estadísticamente genotipos entre sí, ya que no se utilizan pruebas de hipótesis, sino que las conclusiones se obtienen por la sola observación del biplot (Yang et al., 2009).

La técnica de biplot GGE permite identificar mega-ambientes, genotipos superiores dentro de cada mega-ambiente, localidades discriminantes y representativas, y la estabilidad de los genotipos. Para evaluar la estabilidad de los genotipos Yan (2001) desarrolló una forma de visualizar el biplot GGE la cual llamó “media vs inestabilidad”. En estos gráficos se traza una línea que pasa por el centro del biplot y un ambiente virtual promedio de todos los ambientes llamada eje del ambiente promedio (EAP). Perpendicular al eje EAP y cruzando por el centro del biplot se dibuja otro eje llamado coordenada ambiental media (CAM). Los valores más alejados sobre el eje CAM corresponden a genotipos más inestables (mayor contribución a la interacción GE).

La última medida de la eficiencia de un ensayo de rendimiento es la predicción de la ganancia genética (ΔG), para un carácter simple o complejo. La predicción de la ΔG está determinada por tres factores: la intensidad de selección (i), la heredabilidad (h^2), y la raíz cuadrada de la variancia fenotípica (σ^2_p) estimada entre ambientes:

$$\Delta G = i h^2 \sigma_p$$

La intensidad de selección es la proporción inversa de la selección de la población que es seleccionada; una mayor intensidad significa una menor proporción de genotipos seleccionados. Es un número arbitrario determinado por el mejorador. La heredabilidad es la proporción de la variancia genotípica (σ^2_g) sobre la variancia fenotípica (σ^2_p) y es el concepto más importante en genética cuantitativa que respecta a ensayos de variedades:

$$h^2 = H = \sigma^2_g / \sigma^2_p$$

Todas las medidas que se toman en los ensayos de rendimiento desde el diseño, la conducción, el análisis estadístico, tienen el mismo objetivo, que es, incrementar la heredabilidad de los ensayos.

Cuando cada año-localidad combinado son tratados como un ambiente, la heredabilidad en sentido amplio estimada sobre la base media del genotipo es estimada por la siguiente fórmula (Holland et al., 2003):

$$h^2 = \sigma^2_g / [\sigma^2_g + (\sigma^2_{ge}/E) + (\sigma^2_{\epsilon}/E R)]$$

donde σ^2_g , σ^2_{ge} y σ^2_{ϵ} son estimaciones de los componentes de la varianza para el genotipo, la interacción GE y el error agrupado, E y R = número de ambientes de selección y el número de repeticiones del experimento.

Una forma de reducir el error experimental y obtener una ganancia “pura” en la heredabilidad del ensayo, sin costo adicional, es la de utilizar interacción GL (Genotipo por Localidad)

repetible a través de los años, esto se logra evaluando los genotipos en los mismos mega-ambientes.

Un mega-ambiente puede considerarse como tal si: cuando la H es alta tal que, todos los ambientes están correlacionados positivamente, donde GE, compuesto por GL, GA (Genotipo por Año) y GLA (Genotipo por Localidad por Año), es pequeño respecto a G.

Por consiguiente, una elevada H es una indicación inmediata que la región objetivo es una mega-ambiente simple. Esta es la situación ideal para un mejorador ya que puede seleccionar el mejor cultivar evaluándolo en pocas localidades y durante pocos años, y las variedades selectas pueden recomendarse con seguridad a los productores para esa región.

Si la H es baja, lo cual implica que la interacción GE es grande respecto a G, entonces los componentes de GE (GL, GA y GLA) necesitan ser examinados. Si GE es completamente aleatorio e irrepetible a través de los años, entonces se trata de un mega-ambiente complejo. En este caso se debe evaluar los cultivares en varios ambientes durante varios años para aumentar la H .

Dentro de los componentes de GE, GA y GLA son aparentemente aleatorios, ya que no es posible repetir un año, por el otro lado GL es el componente que puede repetirse a través de los años. Por ello, la relación GL/G o $\sigma_{gl}^2 / \sigma_g^2$ se vuelve el factor clave en el análisis de mega-ambientes. Si esta relación es grande, entonces la región objetivo puede dividirse en subregiones (mega-ambientes), para que parte de GL pueda convertirse en G dentro de las subregiones, y de este modo, mejorar la heredabilidad. Si la relación $\sigma_{gl}^2 / \sigma_g^2$ es baja, entonces poca ventaja se lograr al dividir la región objetivo.

Numerosos métodos estadísticos se han desarrollado para el análisis e interpretación de la interacción GE (DeLacy et al., 1996). Los modelos lineales de efectos aleatorios permiten evaluar la magnitud y la significancia de los efectos G, E y su interacción GE. El factor genotipo es aleatorio cuando el objetivo es apoyar las decisiones sobre elementos de una estrategia de mejoramiento estimando: componentes de varianza, parámetros genéticos, ganancias genéticas

esperadas de diferentes estrategias de adaptación o procedimientos de selección. El factor de localidad es definitivamente aleatorio cuando el principal interés del análisis reside en la estimación de componentes de varianza (Wricke y Weber, 1986) para sitios que son representativos de la población relevante dentro de la región objetivo.

Estos modelos permiten estimar la magnitud de los componentes de varianza y predecir efectos genotípicos y de la interacción GE a partir de bases de datos de ensayos desbalanceados en términos de genotipos, años y localidades (Patterson y Thompson, 1975). Cuando un grupo de genotipos es evaluado en un rango de ambientes y años, frecuentemente no se dispone de datos balanceados, debido a que los cultivares son continuamente reemplazados año tras año como consecuencia del proceso de evaluación y selección entre los ensayos de distinto nivel de avance dentro de un programa de mejoramiento. Tampoco el número de repeticiones por genotipo puede ser el mismo por cuestiones operativas, de manejo y/o ambiente. En estos casos los mejoradores deben trabajar y analizar datos desbalanceados. Estos modelos permiten analizar grandes cantidades de datos de ensayos multi-ambiente, donde no todos los genotipos están repetidos entre años, en los que se pueden utilizar los componentes de varianza para estimar la heredabilidad (h^2). Los supuestos clásicos de independencia y homogeneidad de varianzas del modelo lineal general (muestreo ideal) se flexibilizan en el marco del modelo mixto general para los términos aleatorios (Balzarini, 2006)

Hasta el momento, se han publicado diversos trabajos para la identificación de mega-ambientes en Argentina como en Estados Unidos para distintos cultivos (Yan et al., 2000, 2003; 2007b;; Zhang et al., 2005; Dardanelli et al., 2006; Annicchiarico et al., 2010, De la Vega y Chapman, 2010; Mohammadi et al., 2010; Yan , 2010; Rojas et al., 2011;). Sin embargo, no se ha investigado sobre el nivel de correlación de los datos de rendimiento y de características agronómicas en ambientes localizados en países de distintos hemisferios (Ejemplo, Estados Unidos de América y Argentina) para la selección de variedades en el cultivo de soja.

HIPÓTESIS

Existen ambientes en Estados Unidos y Argentina que permiten evaluar y seleccionar variedades de soja con buenas características agronómicas para ambos países.

OBJETIVOS

- Estimar los componentes de variancia de G, E y GE para la variable rendimiento en grano para las distintas campañas agrícolas.
- Estimar la correlación de rendimiento entre ambientes de Argentina y Estados Unidos, y su impacto en el ordenamiento de las variedades en cada uno.
- Evaluar la variable rendimiento en los distintos ambientes utilizando biplots GGE para determinar si los ambientes en estudio de Argentina (norte de Bs. As, y sur de Santa Fe) y Estados Unidos (desde 33.5 a 40° de latitud Norte) conforman un gran mega- ambiente.
- Caracterizar los ambientes de Argentina y los genotipos utilizados mediante las variables evaluadas (rendimiento en grano, altura de planta en R8, días desde siembra a R1, días desde siembra a R8, peso de 1000 semillas y vuelco) durante las dos campañas en Argentina.
- Detectar genotipos estables y/o con un nivel elevado de adaptación a los ambientes de Argentina y Estados Unidos para el carácter rendimiento en grano.

MATERIALES Y MÉTODOS

Zona de estudio:

La zona de estudio comprende ensayos comparativos de rendimiento realizados en distintas localidades en la zona geográfica central de la Argentina y en la zona centro-este de Estados Unidos durante un período de dos años (cuatro ciclos agrícolas). Las localidades en las que se realizaron los ensayos corresponden a ambientes normales para la siembra de soja y están comprendidas dentro de la misma clasificación climática (Cfa) según la clasificación climática mundial de Köppen-Geiger. Todos los ensayos se condujeron según las prácticas de manejo recomendadas para el cultivo y según las necesidades particulares de cada ambiente.

Genotipos

Los materiales genéticos utilizados en el presente estudio son propiedad de la empresa Pioneer Hi-Bred International e incluyeron a 17 variedades de soja comerciales inscriptas en el mercado de los Estados Unidos. El gradiente de madurez de las variedades utilizadas es de 3.7 – 4.9. Con el fin de caracterizar las variedades utilizadas, desarrolladas para el mercado de Estados Unidos, en los ambientes de Argentina se utilizaron cuatro variedades comerciales inscriptas en el mercado de Argentina, dos de la empresa Feria del Norte, uno de la empresa Nidera y uno de la empresa Don Mario (Tabla 1).

En las localidades de Estados Unidos no estuvieron presentes todas las variedades en todas las localidades. La presencia de las variedades en cada localidad correspondió a la adaptación de las mismas al ambiente agroecológico para el que fueron seleccionadas. Es por ello que los genotipos se reunieron en dos grupos según su rango de madurez, el primer grupo “cortos” comprendió las variedades de grupo de madurez (GM) entre 3.7 y 4.4; el segundo grupo “largos” correspondió a variedades de GM entre 4.2 y 4.9 (Tabla 1).

En las localidades de Argentina estuvieron presentes todas las variedades en todas las localidades.

Tabla 1: Genotipos evaluados, grupo de madurez (GM), empresa propietaria y grupo al cual se asignaron en los ambientes de Estados Unidos según su madurez.

(*) No aplican a la clasificación grupo, solo se utilizaron en ambientes de Argentina.

Genotipo	GM	Empresa	Grupo
A	3.7	Pioneer	Cortos
B	3.8	Pioneer	Cortos
C	3.9	Pioneer	Cortos
D	3.9	Pioneer	Cortos
E	4	Pioneer	Cortos
F	4.2	Pioneer	Cortos-Largos
G	4.2	Pioneer	Cortos-Largos
H	4.4	Pioneer	Cortos-Largos
I	4.4	Pioneer	Largos
J	4.5	Pioneer	Largos
K	4.6	Pioneer	Largos
L	4.6	Pioneer	Largos
M	4.7	Pioneer	Largos
N	4.8	Pioneer	Largos
O	4.8	Pioneer	Largos
P	4.9	Pioneer	Largos
Q	4.9	Pioneer	Largos
DM 4212	4.2	Don Mario	*
FN 3.85	3.8	Feria del Norte	*
FN 4.35	4.3	Feria del Norte	*
NS 4611	4.6	Nidera	*

Localidades en Estados Unidos:

Los ensayos en las localidades de Estados Unidos se realizaron durante los ciclos agrícolas 2014 y 2015; el total de ambientes evaluados fueron 34 distribuidos a lo largo de las principales regiones sojeras del país (Figura 1), desde 33.5° a 40° de latitud Norte. Las variedades comprendidas en el GM “cortos” fueron evaluadas desde 39° a 40° de latitud Norte (Tabla 2), las variedades comprendidas en el GM “largos” fueron evaluadas desde 33.5° a 38.5° de latitud Norte (Tabla 3). Durante la campaña 2014 el 2,5% de los genotipos no estuvieron presentes en todos los ambientes para los grupos “cortos”. Durante la campaña 2015 el 17,9 y el 2,8% de los genotipos no estuvieron presentes en todos los ambientes para los grupos “cortos” y “largos” respectivamente.

Los ensayos fueron conducidos en un Diseño en Bloques Completos al Azar (DBCA), con entre cuatro y ocho repeticiones. La unidad experimental fue una parcela de un surco 1,5 m de largo con una distancia entre surcos de 0,90 m (una superficie total de 1,35 m²). La siembra se realizó durante los meses de mayo y junio con una sembradora experimental ALMACO.

Tabla 2: Ambientes evaluados en ensayos comparativos de rendimiento en Estados Unidos en las campañas agrícolas 2014 y 2015 para el grupo “cortos”.

Ambiente	Campaña	Fecha de siembra	Latitud	Longitud	Estado
1	2014	20-May-14	39.00853	-95.20442	Kansas (KS)
2	2014	30-May-14	39.00853	-95.20442	Kansas (KS)
3	2014	20-May-14	38.9891	-95.2181	Kansas (KS)
4	2014	30-May-14	38.9891	-95.2181	Kansas (KS)
5	2014	21-May-14	39.2353	-93.2792	Missouri (MO)
6	2014	22-May-14	39.5012	-89.5328	Illinois (IL)
7	2014	30-May-14	39.49172	-83.76159	Ohio (OH)
8	2014	27-May-14	39.91941	-88.40588	Illinois (IL)
9	2014	24-May-14	40.147	-88.59236	Illinois (IL)
10	2014	25-May-14	40.02366	-88.40485	Illinois (IL)
11	2015	13-May-15	39.03144	-95.26599	Kansas (KS)
12	2015	10-Jun-15	39.03144	-95.26599	Kansas (KS)
13	2015	12-May-15	38.99105	-95.21993	Kansas (KS)
14	2015	9-Jun-15	38.99105	-95.21993	Kansas (KS)
15	2015	27-May-15	39.23946	-93.27696	Missouri (MO)
16	2015	21-May-15	39.48873	-83.77071	Ohio (OH)
17	2015	3-Jun-15	39.48873	-83.7711	Ohio (OH)
18	2015	11-Jun-15	39.88978	-88.43309	Illinois (IL)

Tabla 3: Ambientes evaluados en ensayos comparativos de rendimiento en Estados Unidos en las campañas agrícolas 2014 y 2015 para el grupo “largos”. Escribir los nombres de los estados en castellano

Ambiente	Campaña	Fecha de siembra	Latitud	Longitud	Estado
19	2014	5-Jun-14	33.52175	-90.9837	Mississippi (MS)
20	2014	7-Jun-14	33.42308	-90.89149	Mississippi (MS)
21	2014	6-Jun-14	33.42169	-90.89178	Mississippi (MS)
22	2014	20-May-14	38.47188	-89.78842	Illinois (IL)
23	2014	22-May-14	37.8452	-88.2493	Illinois (IL)
24	2014	17-Jun-14	35.25325	-90.33005	Arkansas (AR)
25	2014	19-Jun-14	35.14641	-90.31989	Arkansas (AR)
26	2014	6-Jun-14	36.1173	-89.8874	Arkansas (AR)
27	2015	6-Jun-15	33.52018	-90.98557	Mississippi (MS)
28	2015	6-Jun-15	33.42143	-90.89173	Mississippi (MS)
29	2015	11-Jun-15	38.47078	-89.79388	Illinois (IL)
30	2015	28-May-15	37.83533	-88.22844	Illinois (IL)
31	2015	5-Jun-15	35.24863	-90.31699	Arkansas (AR)
32	2015	8-Jun-15	35.14378	-90.32055	Arkansas (AR)
33	2015	4-Jun-15	35.19098	-90.92612	Arkansas (AR)
34	2015	11-Jun-15	37.40781	-94.58888	Missouri (MO)

Localidades en Argentina.

Los ensayos en las localidades de Argentina se ubicaron en el Norte de la provincia de Bs. As. y en el Sur de la provincia de Santa Fe (Tabla 4). Los ensayos se realizaron durante los ciclos agrícolas 2014-15 y 2015-16; el total de ambientes evaluados considerando los dos años fue de 14 (Figura 1). En cada localidad, los ensayos fueron conducidos en un Diseño en Bloques Completos al Azar (DBCA) con ocho repeticiones. La unidad experimental fue una parcela de dos surcos de 4 m de largo con una distancia entre surcos de 0,70 m (una superficie total de 5,6 m²). La siembra se realizó durante los meses de octubre, noviembre y diciembre con una sembradora experimental ALMACO.

Tabla 4: Ambientes evaluados en ensayos comparativos de rendimiento en Argentina en las campañas agrícolas 2014-5 y 2015-6 para ambos grupos de madurez (largos y cortos).

Ambiente	Campaña	Fecha de siembra	Latitud	Longitud	Provincia
35	2014-2015	5-Dec-14	-33.80345	-60.86278	Bs. As. (BA)
36	2014-2015	10-Dec-14	-33.74846	-61.30614	Santa Fe (SF)
37	2014-2015	24-Oct-14	-33.86084	-60.65849	Bs. As. (BA)
38	2014-2015	15-Dec-14	-33.85918	-60.65725	Bs. As. (BA)
39	2014-2015	28-Nov-14	-33.85587	-60.65057	Bs. As. (BA)
40	2014-2015	3-Dec-14	-34.19743	-60.78988	Bs. As. (BA)
41	2014-2015	5-Dec-14	-33.80629	-61.90655	Santa Fe (SF)
42	2014-2015	4-Dec-14	-33.69588	-61.959	Santa Fe (SF)
43	2015-2016	2-Dec-15	-33.75077	-61.31118	Santa Fe (SF)
44	2015-2016	23-Oct-15	-33.86084	-60.65849	Bs. As. (BA)
45	2015-2016	28-Dec-15	-33.85918	-60.65725	Bs. As. (BA)
46	2015-2016	26-Nov-15	-33.85538	-60.65288	Bs. As. (BA)
47	2015-2016	30-Nov-15	-34.19787	-60.78832	Bs. As. (BA)
48	2015-2016	3-Dec-15	-33.80483	-61.90621	Santa Fe (SF)

Datos fenotípicos:

Las variables evaluadas para caracterizar las variedades fueron: días desde siembra a inicio de floración (R1), días desde siembra a madurez fisiológica (R8), altura de planta (expresada en cm), vuelco, peso de 1000 semillas (expresado en gr) y rendimiento en grano (expresado en kg/ha).

La variable días desde siembra a madurez se registró como la cantidad de días desde siembra hasta que el 95% de las vainas de la parcela alcancen el color de madurez (R8).

La altura de planta se registró como la altura desde el suelo hasta el nudo superior de la planta en los estadíos R8 (Fehr y Caviness, 1977).

El vuelco de plantas se registró en el estadio R8 utilizando un puntaje de 1 a 9 (Anexo 1).

El rendimiento se evaluó cosechando los dos surcos de cada parcela con una cosechadora experimental ALMACO de dos surcos.

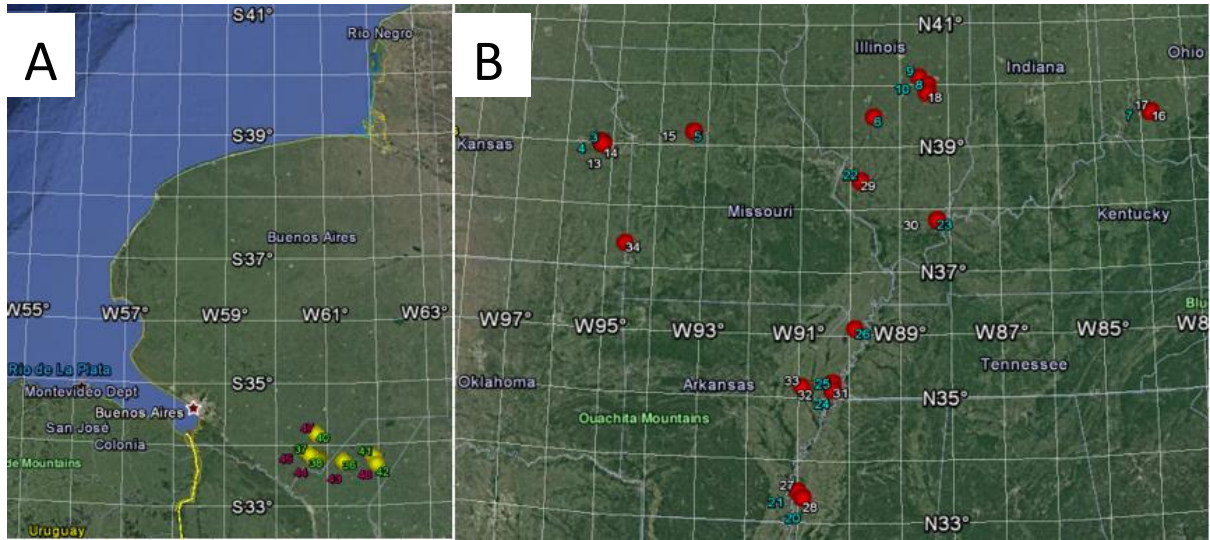


Figura 1: Ubicación geográfica de los 49 ambientes evaluados a lo largo de cuatro campañas. La figura A corresponde a los ambientes de Argentina, la figura B corresponde a los ambientes en Estados Unidos. Ambas figuras están apareadas por la misma latitud.

Análisis estadístico:

Para descomponer la varianza fenotípica del rendimiento en grano (kg/ha) en componentes debidos al efecto genotípico, al efecto ambiental, a la interacción GE, y al error experimental, se ajustó el siguiente modelo:

$$Y_{ij} = \mu + G_i + E_j + GE_{ij} + \epsilon_{ij}. \quad [1]$$

donde Y_{ij} es la respuesta (rendimiento) del genotipo i , en el ambiente j ; μ es la media general; G_i es el efecto aleatorio del genotipo i con $i=1, \dots, g$; que se supone $NID(0, \sigma_g^2)$; E_j es el efecto aleatorio del ambiente j con $j=1, \dots, t$; que se supone $NID(0, \sigma_e^2)$; GE_{ij} es el efecto aleatorio de la interacción del genotipo i con el ambiente j , que se supone $NID(0, \sigma_{ge}^2)$; y ϵ_{ij} es el término de error aleatorio asociado a la observación Y_{ij} , que se supone $NID(0, \sigma_\epsilon^2)$.

El efecto ambiente es la combinación de localidad, país y año; y se utilizó para estimar los componentes de variancia σ_e^2 combinado. Cuando se particionó el análisis por campaña el efecto ambiente correspondió a localidad.

A partir del modelo utilizado se estimaron los componentes de varianza de los efectos aleatorios (G, E e interacción GE) utilizando el paquete estadístico ASReML, que utiliza la técnica REML (máxima verosimilitud restringida), y programa estadístico R.

Mediante el z.ratio obtenido del análisis de variancia se evaluó la existencia de significancia estadística de las fuentes de variación, a mayor z.ratio, menor es la probabilidad de que los efectos se hayan producido al azar. Si el z.ratio es mayor que 1,64 y 2,33 se asume que hay diferencias significativas al 5 y 1%, respectivamente.

Se estimó la varianza fenotípica (σ_p^2) utilizando los componentes de varianza derivados del modelo [1] utilizando la fórmula:

$$\sigma_p^2 = \sigma_g^2 + (\sigma_{ge}^2/E) + (\sigma_\varepsilon^2/E R) \quad [2]$$

donde E y R son un determinado número de ambientes y repeticiones, respectivamente. Se utilizó una combinación de valores para E y R y se estimó heredabilidad en sentido amplio como la media de un genotipo a través de todos los ambientes para cada una de dichas combinaciones (Holland et al., 2003), según la expresión:

$$h^2 = \sigma_g^2 / [\sigma_g^2 + (\sigma_{ge}^2/E) + (\sigma_\varepsilon^2/E R)] \quad [3]$$

Para verificar el supuesto de distribución normal de los errores de la variable rendimiento en grano se utilizó el método gráfico “Q-Q Plot” en el cual utilizando los residuos se obtiene un diagrama de dispersión en el que, si los residuos son normales y no hay otros defectos del modelo, entonces se alinean sobre una recta a 45°.

Para visualizar la distribución del rendimiento en grano en cada ambiente y a través de los mismos se utilizó un gráfico “box-plot”, el mismo proporciona una visión general de la simetría de la distribución de los datos, con la mediana, los percentiles 25 y 75 y los valores máximos y mínimos, y también permite ver la presencia de valores atípicos.

La existencia o no de correlación fenotípica entre ambientes para la variable rendimiento se corroboró mediante un gráfico (“hetamap”) de correlación entre las mismas (Friendly, 2002)

utilizando el programa estadístico R (corrgram) y también mediante la técnica biplot GGE. Los gráficos biplots GGE se utilizaron para identificar mega-ambientes, genotipos superiores dentro de cada mega-ambiente, localidades discriminantes y representativas, y estabilidad de los genotipos.

La técnica gráfica biplot GGE, utiliza el procedimiento matemático de la descomposición por valor singular (SVD) donde una matriz $g \times e$ de rendimientos promedios de g cultivares en e ambientes puede aproximarse como el producto de una matriz de genotipos y una matriz de ambientes (Yan y Hunt, 2000), entonces el rendimiento de un genotipo i en el ambiente j , Y se aproxima a:

$$\hat{Y}_{ij} = \sum_{n=1}^r \lambda_n \xi_{in} \eta_{jn} \quad (\lambda_1 \geq \lambda_2 \geq \lambda_3 \dots \geq \lambda_r) \quad [4]$$

donde:

r es el número de CP requeridos para aproximarse al valor original,

λ_n es el valor singular para CP_n ,

ξ_{in} es el valor del genotipo i ésimo del CP_n ,

η_{jn} es el valor del ambiente j ésimo del CP_n .

La SVD permite mostrar en un gráfico la tabla de rendimientos promedios de $g \times e$, teniendo g puntos para los genotipos y e puntos para los ambientes. Se denomina biplot, ya que los genotipos y los ambientes se visualizan en un solo gráfico. Los biplots pueden ser multi-dimensionales, pero generalmente se utilizan los primeros dos CP para graficar.

Para aproximarse a una matriz con dos CP, la ecuación [4] puede reescribirse como:

$$\bar{Y}_{ij} = \xi_{i1} \lambda_1 \eta_{j1} + \xi_{i2} \lambda_2 \eta_{j2} + \varepsilon_{ij} \quad [5]$$

donde ε_{ij} es la diferencia (residuo) entre el valor real de \bar{Y}_{ij} y el estimado por los dos primeros CP.

El biplot se construye graficando $\xi_{i1} \lambda_1 \eta_{j1}$ versus $\xi_{i2} \lambda_2 \eta_{j2}$

El rendimiento promedio del genotipo i en el ambiente j se describe con el modelo lineal general:

$$\bar{Y}_{ij} = \mu + \alpha_i + \beta_j + \phi_{ji}$$

donde;

μ es la media general;

α_i es el efecto principal del genotipo i ésimo,

β_j es el efecto principal del ambiente j ésimo,

ϕ_{ji} es la interacción entre genotipo i y el ambiente j .

En la técnica biplot GGE se elimina el término α_i y somete el término ϕ_{ji} (donde ahora los rendimientos están centrados por el ambiente) a SVD:

$$\phi_{ji} = \bar{Y}_{ij} - \mu - \beta_j = \zeta_{i1} \lambda_1 \eta_{j1} + \zeta_{i2} \lambda_2 \eta_{j2} + \epsilon_{ij} \quad [6]$$

Para eliminar los errores de variancias entre ambientes, se estandarizó el modelo anterior

$$(\bar{Y}_{ij} - \mu - \beta_j) / z_j = \zeta_{i1} \lambda_1 \eta_{j1} + \zeta_{i2} \lambda_2 \eta_{j2} + \epsilon_{ij} \quad [7]$$

donde Z_j es el desvío estándar de la media del genotipo i dentro del ambiente j

Para caracterizar los ambientes y genotipos utilizados en Argentina se realizó un biplot y un análisis de conglomerados con las variables evaluadas. Las variables fueron estandarizadas ya que no todas fueron medidas en la misma escala. Para el análisis de conglomerados se utilizó el algoritmo encadenamiento promedio sobre la matriz de distancias Euclideas. El análisis de conglomerados jerárquico produce un dendograma como representación gráfica del agrupamiento resultante.

RESULTADOS

Grupo “cortos”:

Los resultados de la estimación de la magnitud y proporción relativa de los componentes de varianza para cada una de las campañas individuales analizadas y todas las campañas combinadas para rendimiento en grano se muestran en la Tabla 5. La interacción GE fue estadísticamente significativa ($p < 0,01$) en tres de las cuatro campañas evaluadas, siendo la campaña 2015 no significativa. El componente G fue no significativo en tres de las cuatro campañas, siendo en la campaña 2015-16 significativo ($p < 0,05$). El componente E fue significativo en tres de las cuatro campañas evaluados ($p < 0,05$), siendo la campaña 2015-16 no significativo. En cada una de las cuatro campañas el componente G explicó una mayor proporción de la varianza en comparación con la interacción GE.

Los resultados de todas las campañas combinadas muestran que todos los componentes de variancia (E, G y GE) fueron significativos, el componente de la interacción GE explicó 1,65 veces más variabilidad que el efecto G.

Tabla 5: Estimación de los componentes de varianza para cada una de las campañas en estudio, z.ratio, significancia y porcentaje (%) explicado sobre el total para ambiente (E), genotipo (G), genotipo por ambiente (GE), para rendimiento en grano.

“NS”: no significativo, * p<0,05, ** p<0,01.

Campaña	Fuente	Componente	z.ratio	sig	%
2014	E	463800	2.1	*	91
	G	26720	1.6	NS	5
	GE	19020	2.4	**	4
	Error	236000	19		
2014-15	E	103600	1.8	*	81
	G	12740	1.5	NS	10
	GE	11790	2.7	**	9
	Error	109400	19		
2015	E	124000	1.8	*	81
	G	14760	1.1	NS	10
	GE	14410	1.3	NS	9
	Error	262900	14		
2015-16	E	414700	1.6	NS	87
	G	42020	1.7	*	9
	GE	17290	2.4	**	4
	Error	152000	16		
Todas	E	447800	3.9	**	92
	G	16790	1.7	*	3
	GE	23520	5.7	**	5
	Error	183300	34		

Con los valores de los componentes de varianza para rendimiento en grano de la tabla 5 se estimó la Heredabilidad en sentido amplio (H), como media de un genotipo a través de todos los ambientes para cada una de las combinaciones de hasta 8 repeticiones y de hasta 40 ambientes (Locs). Las similitudes en las magnitudes de la interacción GE en relación a G en las cuatro campañas respecto a las campañas de EE. UU. solamente mostraron un patrón de respuesta de Heredabilidad similar para rendimiento en grano (Figura 2).

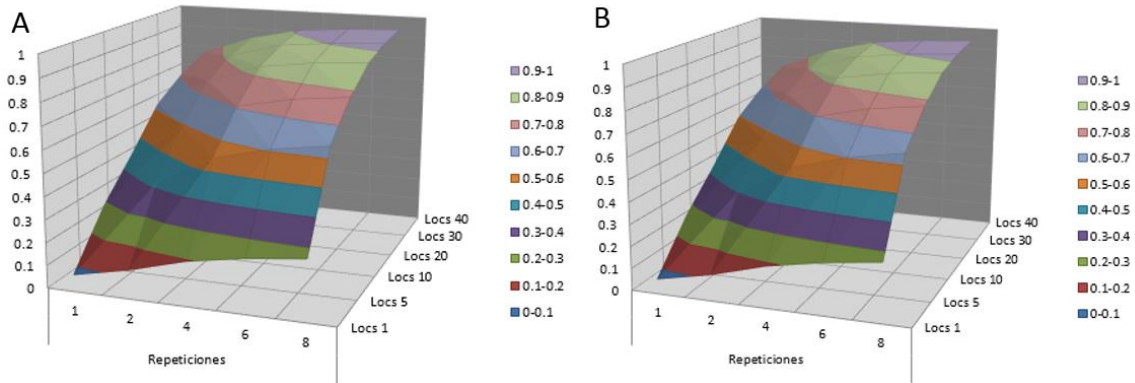


Figura 2: Heredabilidad en sentido amplio para rendimiento en grano, estimada a partir de los componentes de varianza. La figura A corresponde a las 4 campañas (EE, UU, y Arg.), la figura B corresponde a las campañas 2014 y 2015 (EE, UU.)

El supuesto de distribución normal de los errores de cada ambiente se verificó con el Q-Q Plot en la Figura 3, esto permitió incluir todas las localidades en el análisis.

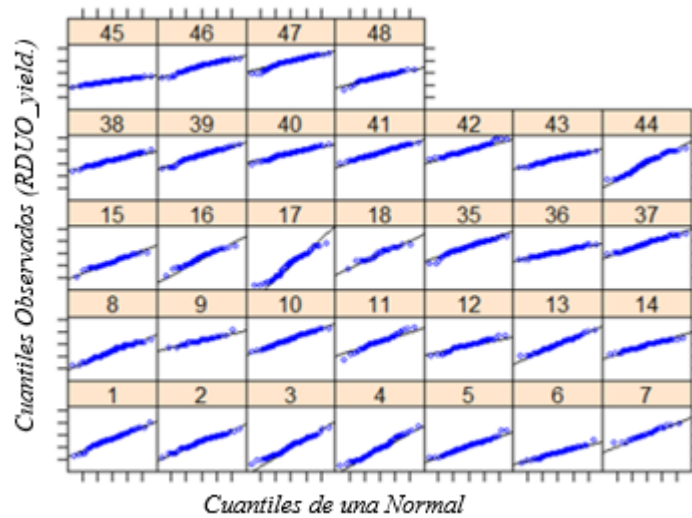


Figura 3: Q-Qplot de los residuos del rendimiento en grano para cada ambiente.

La Figura 4 muestra el Box-Plot de rendimiento en grano de cada localidad en las cuatro campañas. Se puede apreciar la variabilidad de rendimientos que se obtuvieron, siendo el

ambiente 6 de EE. UU. el de menor rendimiento promedio de ≈ 2500 kg/ha, y el ambiente 42 de Arg. el de mayor rendimiento promedio de ≈ 5000 kg/ha.

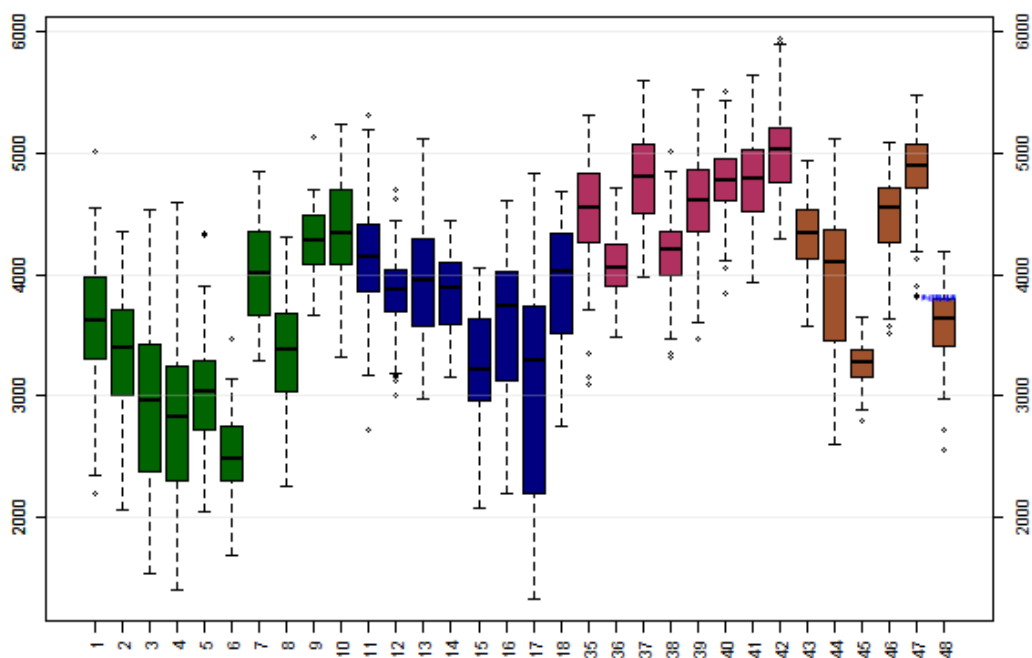


Figura 4: Box-plot de rendimiento en grano (eje y) para cada ambiente (eje x). Total 32 ambientes. Las cajas color azul, verde, rojo y marrón corresponden a las campañas 2014, 2015, 2014-15, 2015-16, respectivamente.

Para obtener una primera aproximación para entender la relación entre ambientes se realizó el gráfico de correlación (Figura 5). En esta figura se puede ver aquellos ambientes que tuvieron correlación positiva y correlación negativa representados con los colores azul y rojo, respectivamente. Los colores de mayor intensidad corresponden a valores más extremos, tanto positivos como negativos. De la figura 5 se puede apreciar que los ambientes con correlación alta positiva ($>0,50$) fueron: 1, 11, 12 y 2 entre sí; 18, 38, 7, 41, 44, 48, 47, 8, 35, 4, 46, 40, 5, 10 entre sí. Los ambientes con correlación alta negativa ($> -0,50$) fueron: 1 con 39, 36 y 17; 11 con 39, 42, 36, 18; 12 con 39, 42, 36, 17, 18, 38, 7, 41 y 44; y 16 con 40, 47 y 48. El ambiente 12 fue visiblemente contrastante con el resto de los ambientes ya que la correlación fue negativa o nula (0) con casi todos los ambientes con excepción de los mencionados anteriormente.

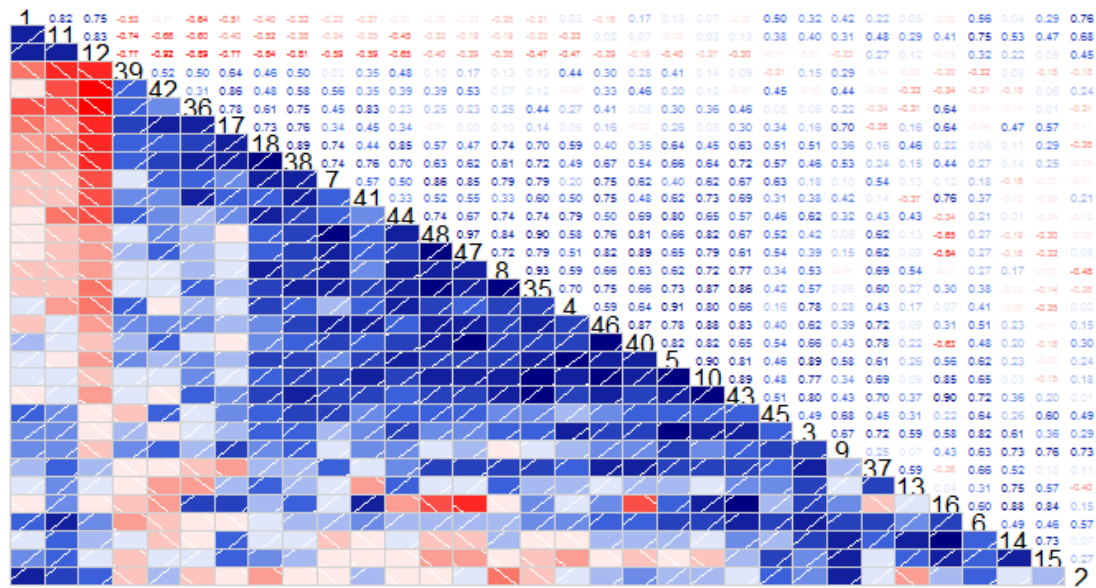
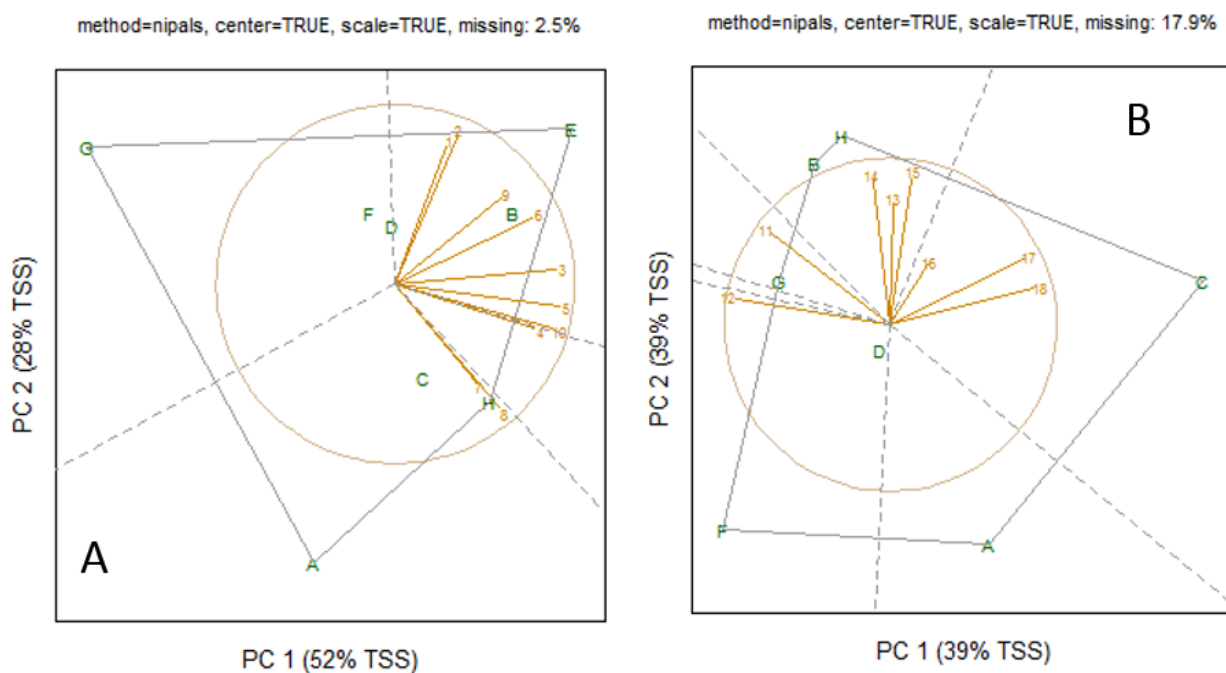


Figura 5: *Heatmap* de correlación fenotípica de rendimiento entre ambientes. Los ambientes están situados en la diagonal del gráfico.

La figura 6 muestra los gráficos biplot GGE obtenidos a partir de los dos primeros componentes principales para cada una de las cuatro campañas para el grupo “cortos”. Estas dos componentes principales (CP1 y CP2) explicaron 88, 78, 74 y 82% de la variabilidad total para rendimiento en grano para las campañas 2014, 2015, 2014-15 y 2015-16, respectivamente. En la campaña 2014 (Figura 6A) se identificaron dos mega-ambientes, uno de ellos compuesto por los ambientes 1, 2, 3, 4, 5, 6, 9 y 10; y el otro mega-ambiente conteniendo los ambientes 7 y 8; los genotipos más extremos fueron el E y H para cada mega-ambiente, respectivamente. En la campaña 2015 (Figura 6B) se identificaron cuatro mega-ambientes, el primero compuesto por el ambiente 12; el segundo compuesto por el ambiente 11; el tercero compuesto por los ambientes 13, 14 y 15; y el cuarto por los ambientes 16, 17 y 18; los genotipos más extremos fueron el F, G, H y C para cada mega-ambiente, respectivamente. Cabe aclarar que los datos faltantes en la campaña 2015 fueron del 17,9%, y debido a ello el gráfico biplot GGE puede verse alterado. Yan (2012) propone una metodología nueva para estimar datos faltantes con resultados positivos de simulación con hasta un 60% de datos faltantes. El uso exitoso del procedimiento requiere que los datos sean lo suficientemente grandes y completos para lograr una predicción correcta ya que se basa en los

valores existentes. En la campaña 2014-15 (Figura 6C) se identificaron tres mega-ambientes, el primero de ellos compuesto por el ambiente 37; el segundo por los ambientes 35 y 40; y el tercer mega-ambiente conteniendo los ambientes 36, 38, 39, 41, y 42; los genotipos más extremos fueron el B, H y C para cada mega-ambiente, respectivamente. En la campaña 2015-16 (Figura 6D) se identificaron dos mega-ambientes, uno de ellos compuesto por los ambientes 44 y 45; y el otro mega-ambiente conteniendo los ambientes 43, 46, 47 y 48; los genotipos más extremos fueron el B y H para cada mega-ambiente, respectivamente.



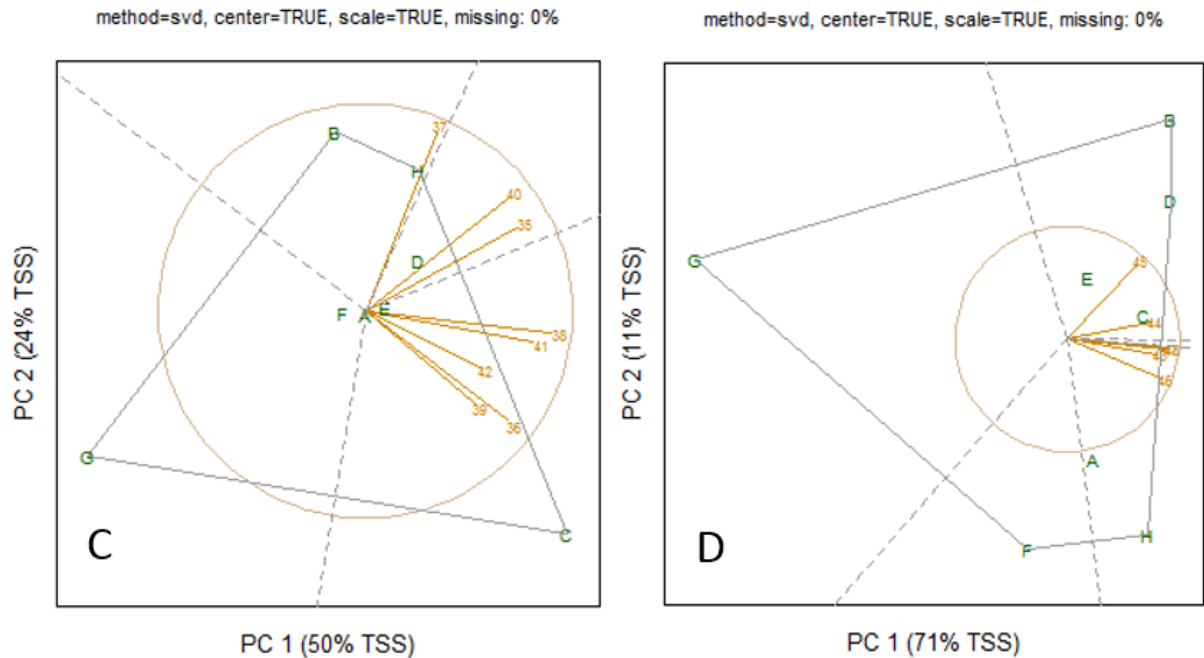
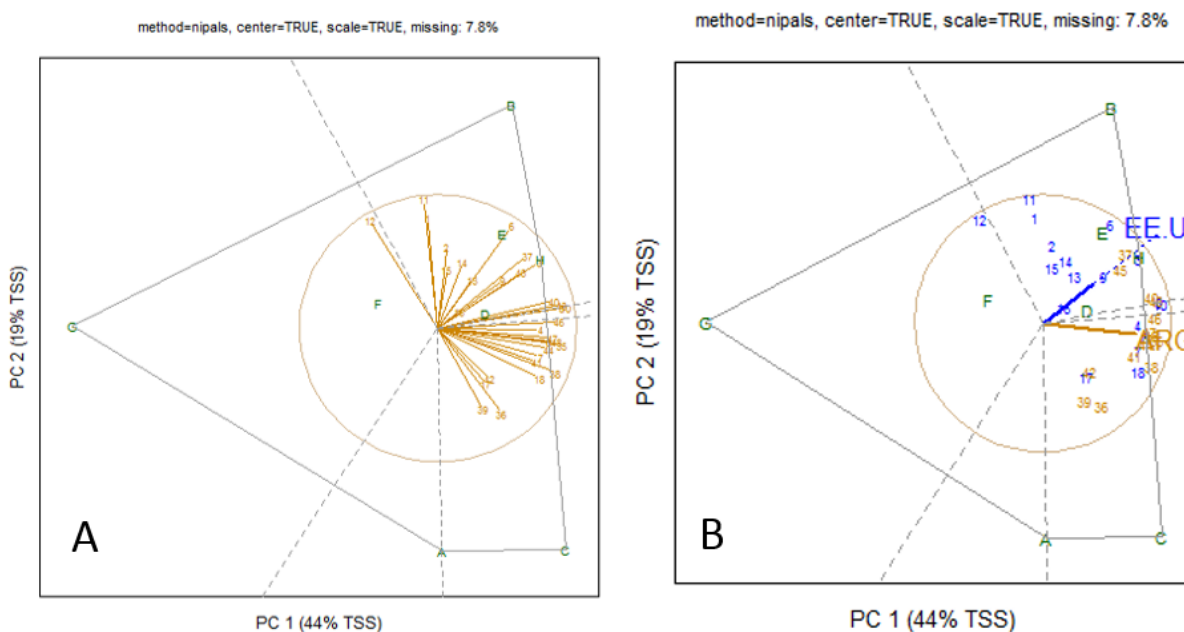


Figura 6: Biplots GGE marcando el patrón “which-won-where” para rendimiento en grano para cada una de las cuatro campañas. Los genotipos están identificados de color verde y los ambientes por vectores amarillos desde el origen. Los biplots A y B corresponden a las campañas 2014 y 2015 de Estados Unidos, respectivamente; los biplots C y D corresponden a las campañas 2014-15 y 2015-16 de Argentina, respectivamente.

La figura 7A muestra los gráficos biplots GGE de las cuatro campañas combinadas obtenidos a partir de los dos primeros componentes principales. Estas dos componentes principales (CP1 y CP2) explicaron el 63% (44 + 19%) de la variabilidad total para rendimiento en grano. Se identificaron tres mega-ambientes, el primero de ellos compuesto por el ambiente 12; el segundo mega-ambiente conteniendo los ambientes 1, 2, 3, 5, 6, 9, 10, 11, 13, 14, 15, 16, 37, 40, 43, y 45; y el tercer mega-ambiente conteniendo los ambientes 4, 7, 8, 17, 18, 35, 36, 38, 39, 41, 42, 44, 46, 47 y 48; los genotipos más extremos fueron el G, B y C para cada uno de los mega-ambientes, respectivamente. El segundo y tercer mega-ambientes contienen ambientes de EE. UU. y Arg., el 25 y el 66% corresponde al porcentaje de ambientes de Arg. comprendidos en el mega-ambiente dos y tres, respectivamente. En la figura 7B se graficó el vector promedio para cada país, se observa que ambos vectores muestran una separación entre los ambientes de EE. UU. y Arg., el ángulo que se forma entre ambos vectores es menor de 90° (agudo) y que hay ambientes de EE. UU. que son similares (correlación positiva) con ambientes de Arg.

La figura 7C muestra los resultados obtenidos en las dos campañas de ensayos de rendimiento en EE. UU. (2014 y 2015) y en Arg. (2014-15 y 2015-16) con los vectores promedio por año. Se observa que los ángulos entre los vectores de las 4 campañas son similares entre sí (aproximadamente de 30°), indicando la existencia de correlación positiva entre las campañas cuyos vectores son contiguos y baja correlación entre los vectores de los extremos donde el ángulo es cercano a los 90°. La figura 7D con los vectores promedio por Estado/Provincia se observa que las provincias de Arg. (SF y BA) mejores correlaciones tuvieron con los estados de IL y OH de EE. UU., por otra parte, el estado de KS fue el que tuvo casi nula correlación con los ambientes de Arg., donde el ángulo entre vectores fue cercano a 90°.



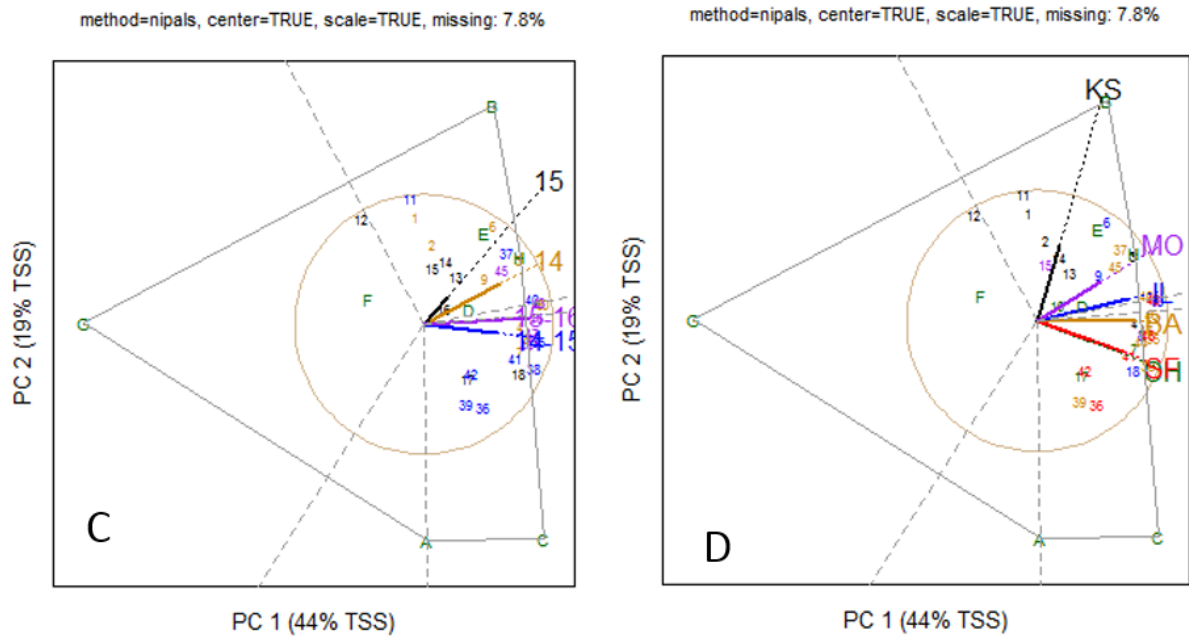


Figura 7: Biplots GGE marcando el patrón “which-won-where” para rendimiento en grano para el total de las campañas. En las figuras B, C y D los ambientes se colorean según un determinado atributo y el vector representa el promedio de cada grupo. En la figura B el atributo es países, en la figura C las respectivas campañas, y en la D las distintas provincias y/o estados. Las provincias y/o estados de la figura D corresponden a KS, Kansas; MO, Misuri; IL, Illinois; OH, Ohio; BA; Buenos Aires; SF, Santa Fe.

El gráfico biplot GGE con el vector de regresión con la media de ambientes (Figura 8) permite evaluar los genotipos y seleccionar y/o recomendar aquellas variedades de mejor comportamiento en los ensayos de rendimiento multi-ambientales. El ordenamiento de los genotipos según su rendimiento promedio a través de todos los ambientes está dado por el largo de la proyección positiva sobre el eje EAP. Los genotipos se ordenaron según su largo de la siguiente forma: B, H, E, C, D, A, F y G; siendo B el de mayor rendimiento y G el de menor rendimiento. La estabilidad de los genotipos se evalúa según el largo de su proyección sobre el eje CAM. Los genotipos se ordenaron según su largo sobre el eje CAM de la siguiente forma: D, F, H, E, G, B, A y C; siendo D el de mayor estabilidad (menor contribución a la interacción GE) y C el de menor estabilidad (mayor contribución a la interacción GE).

Los genotipos ubicados en los sectores donde no hay localidades son considerados de pobre comportamiento en rendimiento (si esta fuese la variable estudiada) en la mayoría de las

localidades evaluadas (Yan, 2001, Yan y Rajcan, 2002). En la Figura 8 se observa que estos genotipos fueron el A, F y G.

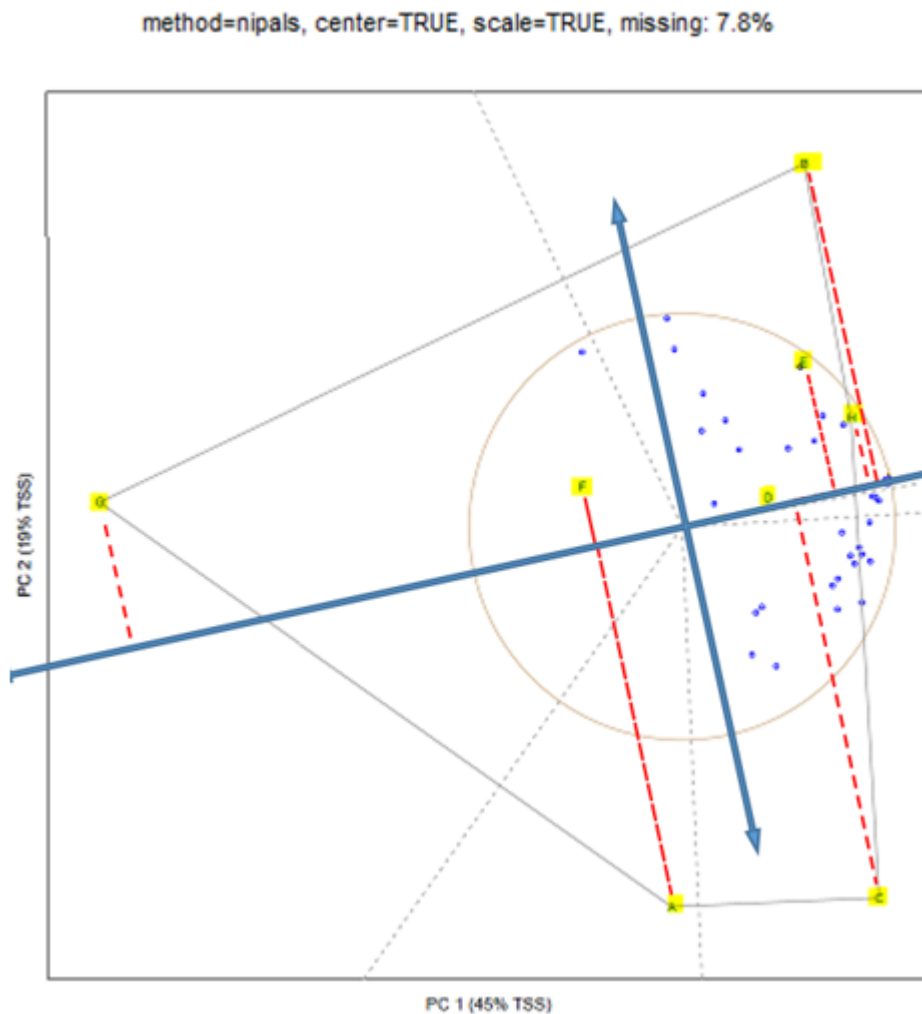


Figura 8: Biplot GGE con la visualización “media vs inestabilidad” para rendimiento en grano para las cuatro campañas. Los genotipos están resaltados en color amarillo, los ambientes corresponden a los puntos azules. El eje azul con punta de flechas corresponde al eje CAM, perpendicular a este se encuentra el eje EAP.

Grupo “Largos”:

Los resultados de la estimación de la magnitud y proporción relativa de los componentes de varianza para cada una de las campañas individuales analizadas y todas las campañas combinadas para rendimiento en grano se muestran en la Tabla 6. La interacción GE fue estadísticamente significativa ($p < 0,01$) en las cuatro campañas evaluadas, el componente G fue significativo ($p < 0,05$) en tres de las cuatro campañas, siendo la campaña 2015 no significativa. El componente E fue significativo en tres de las cuatro campañas evaluados ($p < 0,05$), siendo la campaña 2015-16 no significativa. En tres de las cuatro campañas el componente G explicó una mayor proporción de la varianza en comparación con la interacción GE.

Los resultados de las campañas combinadas muestran que todos los componentes de variancia fueron significativos, el componente de la interacción GE explicó 1,4 veces más variabilidad que el efecto G.

Tabla 6: Estimación de los componentes de varianza para cada una de las campañas en estudio, z.ratio, significancia y porcentaje (%) explicado sobre el total para ambiente (E), genotipo (G), genotipo por ambiente (GE), para rendimiento en grano.

Campaña	Fuente	Componente	z.ratio	sig	%
2014	E	421100	1.8	*	85
	G	31690	1.7	*	6
	GE	45180	3.3	**	9
	Error	295400	19		
2014-15	E	110900	1.8	*	67
	G	34420	2.1	*	21
	GE	21210	4.3	**	13
	Error	106200	23		
2015	E	418100	1.7	*	79
	G	58480	1.6	NS	11
	GE	50480	3	**	10
	Error	269700	16		
2015-16	E	349000	1.6	NS	76
	G	86780	2.2	*	19
	GE	21960	3.9	**	5
	Error	79820	19		
Todas	E	429300	3.8	**	83
	G	38600	2.2	*	7
	GE	51210	8.8	**	10
	Error	179000	39		

Con los valores de los componentes de varianza para rendimiento en grano de la tabla 6 se estimó la Heredabilidad en sentido amplio (H) como media de un genotipo a través de todos los ambientes para cada una de las combinaciones de hasta 8 repeticiones y de hasta 40 ambientes (Locs). Al igual que en los grupos “cortos” se encontraron magnitudes similares de la interacción GE en relación a G en las cuatro campañas combinadas respecto a las campañas de EE. UU. solamente mostrando un patrón de Heredabilidad similar de respuesta para rendimiento en grano.

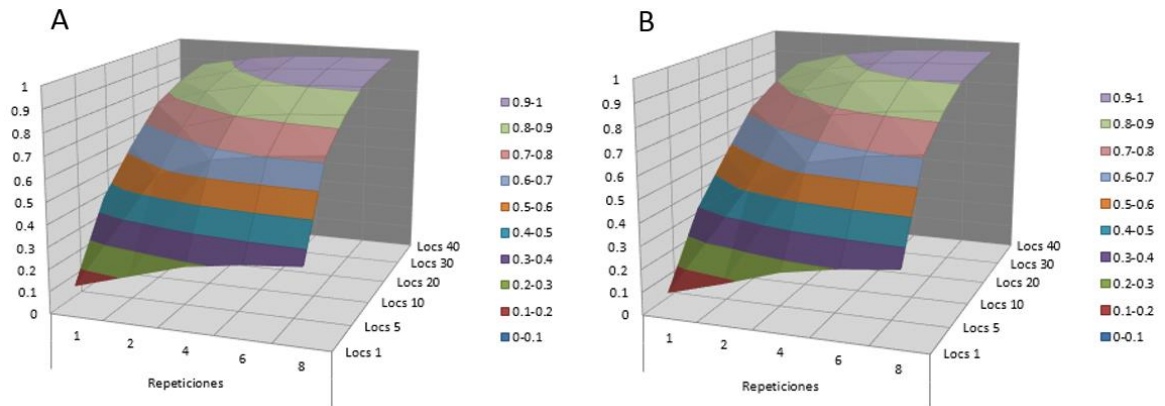


Figura 9: Heredabilidad en sentido amplio para rendimiento en grano, estimada a partir de los componentes de varianza. La figura A corresponde a las 4 campañas (EE. UU. y Arg.), la figura B corresponde a las campañas 2014 y 2015 (EE. UU.)

El supuesto de distribución normal de los errores de cada ambiente se verificó con el Q-Q Plot en la Figura 10, esto permitió incluir todas las localidades en el análisis.

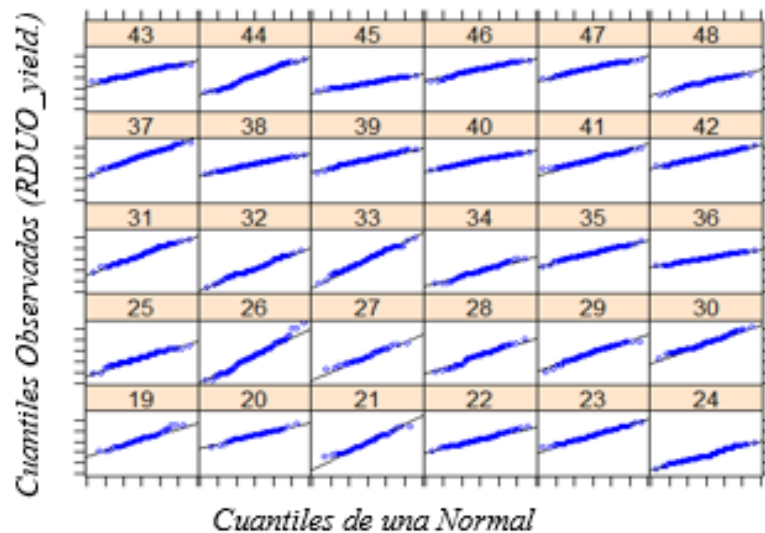


Figura 10: Q-Q plot de cada ambiente para los errores de rendimiento en grano.

La Figura 11 muestra el Box-Plot de rendimiento en grano de cada localidad en las cuatro campañas. Se puede apreciar la variabilidad de rendimientos que se obtuvieron, siendo el

ambiente 24 de EE. UU. el de menor rendimiento promedio de ≈ 2600 kg/ha, y el ambiente 37 de Arg. el de mayor rendimiento promedio de ≈ 5000 kg/ha.

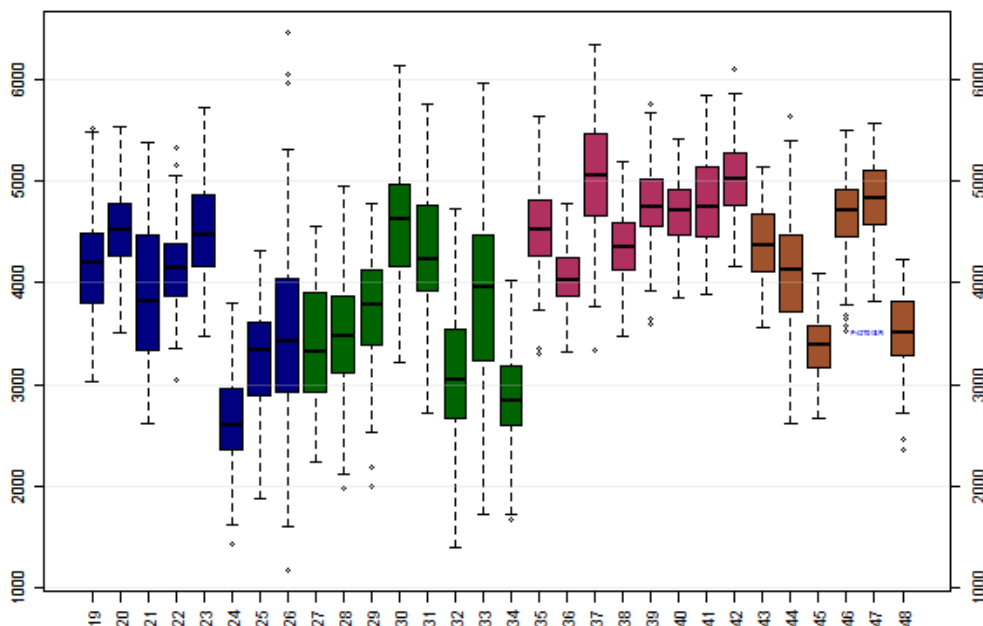


Figura 11: Box-plot de rendimiento en grano (eje y) para cada ambiente (eje x). Total 30 ambientes. Las cajas color azul, verde rojo y marrón corresponden a las campañas 2014, 2015, 2014-15, 2015-16, respectivamente.

En la figura 12 se muestran los resultados del gráfico de correlación entre ambientes. En ella se observa que la mayoría de los ambientes evaluados correlacionaron positivamente y en general esta correlación fue alta ($>0,50$). No se encontraron ambientes de correlación alta negativa ($> -0,50$), solo los ambientes 26, 25 y 24 tuvieron correlación negativa con algunos pocos ambientes.

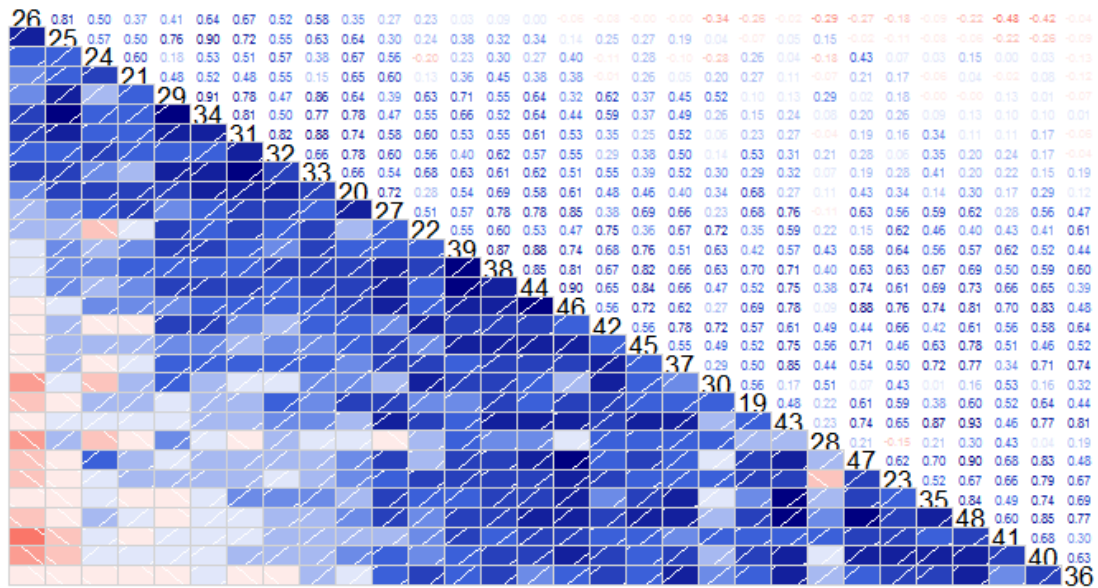


Figura 12: *Heatmap* de correlación de rendimiento entre ambientes. Los ambientes están situados en la diagonal del gráfico.

La figura 13 muestra los gráficos biplots GGE obtenidos a partir de los dos primeros componentes principales para cada una de las cuatro campañas para el grupo “largos”. Estas dos componentes principales (CP1 y CP2) explicaron 70, 76, 78 y 90% de la variabilidad total para rendimiento en grano para las campañas 2014, 2015, 2014-15 y 2015-16, respectivamente. En la campaña 2014 se identificaron tres mega-ambientes, uno de ellos compuesto por los ambientes 19, 22 y 23; el otro mega-ambiente conteniendo los ambientes 20 y 21; y el tercer mega-ambiente conteniendo los ambientes 24, 25 y 26; los genotipos más extremos fueron el J, K y P, respectivamente. En la campaña 2015 se identificaron dos mega-ambientes, uno de ellos compuesto por los ambientes 28, 29 y 30; y el otro mega-ambiente conteniendo los ambientes 27, 31, 32, 33 y 34; los genotipos más extremos fueron el M y Q, respectivamente. En la campaña 2014-15 se identificaron dos mega-ambientes, uno de ellos compuesto por los ambientes 35, 37, 38, 39, 40, 41 y 42; y el otro mega-ambiente conteniendo el ambiente 36 solamente; los genotipos más extremos fueron el M y Q, respectivamente. En la campaña 2015-16 se identificó solo un mega-ambientes compuesto por los ambientes 43, 44, 45, 46, 47 y 48, el genotipo más extremo fue el K.

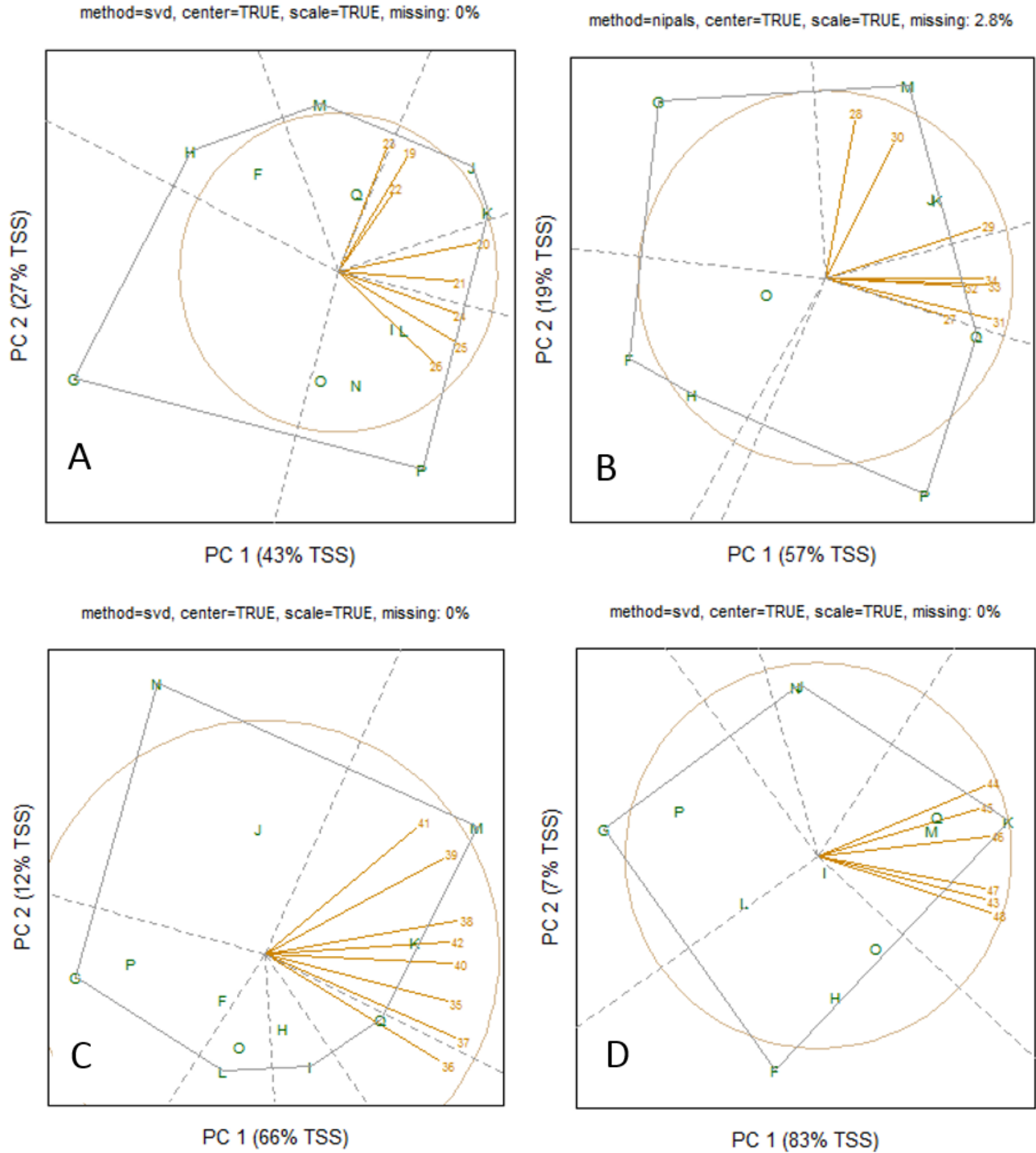


Figura 13: Biplots GGE marcando el patrón “which-won-where” para rendimiento en grano para cada una de las cuatro campañas. Los genotipos están identificados de color verde y los ambientes por vectores amarillos desde el origen. Los biplots A y B corresponden a las campañas 2014 y 2015 de Estados Unidos respectivamente; los biplots C y D corresponden a las campañas 2014-15 y 2015-16 de Argentina, respectivamente.

La figura 14A muestra los gráficos biplots GGE de las cuatro campañas combinadas obtenidos a partir de los dos primeros componentes principales. Estas dos componentes principales (CP1 y CP2) explicaron el 66% (47 + 19%) de la variabilidad total para rendimiento en grano. Se identificaron tres mega-ambientes, el primero de ellos compuesto por los ambientes 21, 24, 25 y 26; el segundo mega-ambiente conteniendo los ambientes 19, 20, 22, 27, 29, 30, 31, 32, 33, 34, 37, 38, 39, 42, 44, 45 y 46; y el tercer mega-ambiente conteniendo los ambientes 23, 28, 35, 36, 40, 41, 43, 47 y 48; los genotipos más extremos fueron el P, K y M para cada uno de los mega-ambientes, respectivamente. El mega-ambiente 1 contiene solamente ambientes de EE. UU., los otros dos mega-ambientes contienen ambientes de EE. UU. y Arg. En la figura 14B se graficó el vector promedio para cada país, y al igual que en los grupos “cortos” se puede apreciar que ambos vectores muestran una separación entre los ambientes de EE. UU. y Arg., que el ángulo que se forma entre ambos vectores es menor de 90° (agudo) y que también hay ambientes de EE. UU. que son similares (correlación positiva) con ambientes de ARG principalmente los ambientes 19, 22, 23, 27, 28, y 30.

La figura 14C muestra los resultados obtenidos en las dos campañas de ensayos de rendimiento en EE. UU. (2014 y 2015) y en Arg. (2014-15 y 2015-16) con los vectores promedio por año. Se observa que los vectores de las campañas 2014 y 2015; y 2014-15 y 2015-16 son casi paralelos (nunca pueden ser paralelos) entre sí, ello indica que existe una correlación muy alta cercana a 1. La figura 14D con los vectores promedio por Estado/Provincia se observa que las provincias de Arg. (SF y BA) tuvieron mejor correlación con los estados IL y MS de EE. UU., por otra parte, el estado de AR fue el que tuvo poca o nula correlación con Arg., donde el ángulo entre vectores fue cercano a 90°. El estado de MO no se tuvo en cuenta ya que solo contenía un solo ambiente (34).

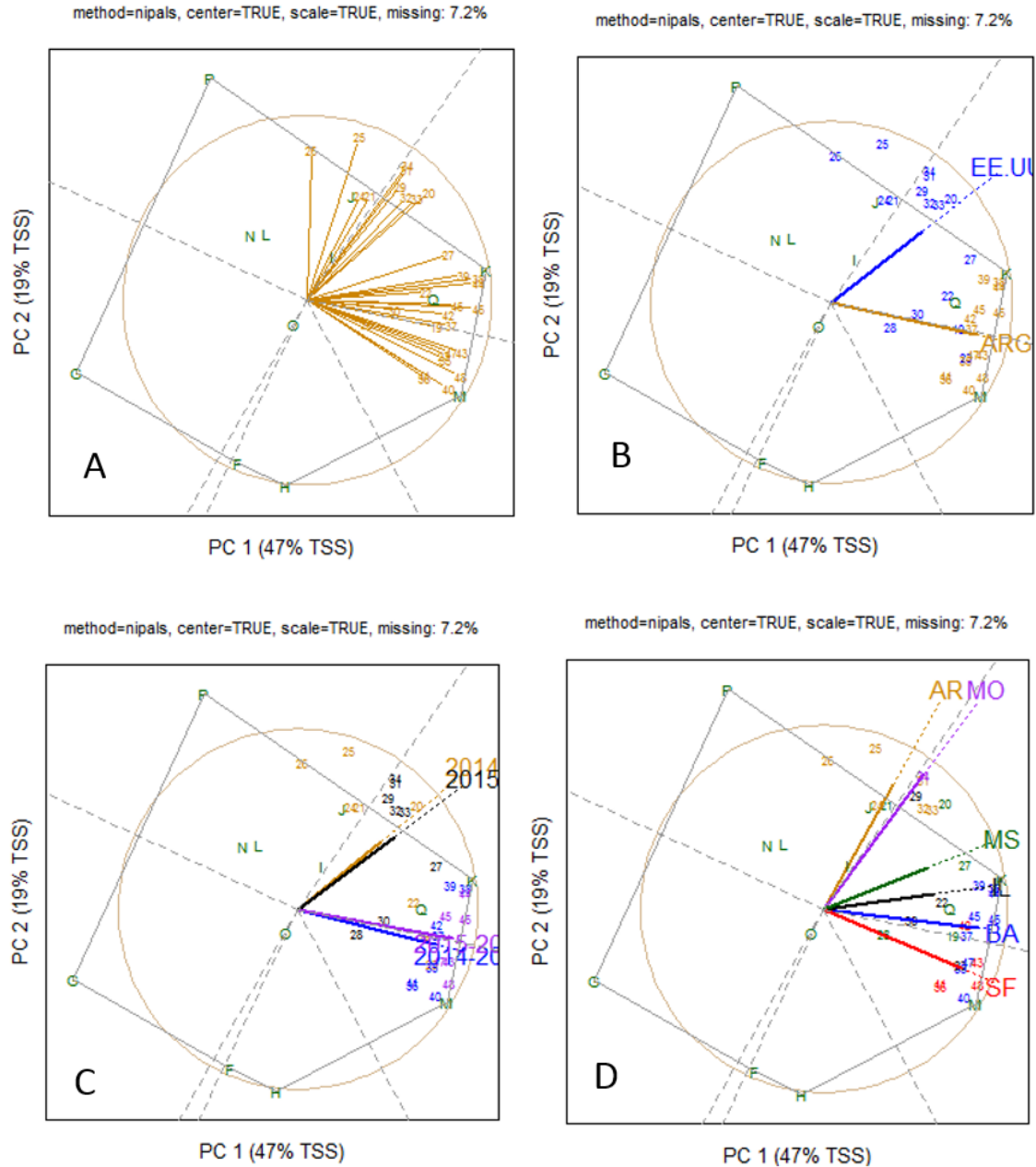


Figura 14: Biplots GGE marcando el patrón “which-won-where” para rendimiento en grano para el total de las campañas. En las figuras B, C y D los ambientes se colorean según un determinado atributo y el vector representa el promedio de cada grupo. En la figura B el atributo es países, en la figura C las respectivas campañas, y en la D las distintas provincias y/o estados. Las provincias y/o estados de la figura D corresponden a AR, Arkansas; MO, Misuri; MS, Misisipi; IL, Illinois; BA; Buenos Aires; SF, Santa Fe.

El gráfico biplot GGE (Figura 15) con el vector de regresión con la media de ambientes muestra el ordenamiento de los genotipos según su rendimiento promedio en todos los ambientes

evaluados. Los genotipos se ordenaron de la siguiente forma: K, M, Q, J, I, O, L, N, P, H, F Y G; siendo K el de mayor rendimiento y G el de menor rendimiento. Los genotipos se ordenaron según su estabilidad de la siguiente forma: K, Q, O, G, I, L, N, J, M, F, H y P; siendo K el de mayor estabilidad (menor contribución a la interacción GE) y P el de menor estabilidad (mayor contribución a la interacción GE).

Los genotipos H, F, G y O se ubicaron en los sectores del biplot GGE donde no hay ambientes representados, por lo tanto, fueron los de menor comportamiento en rendimiento.

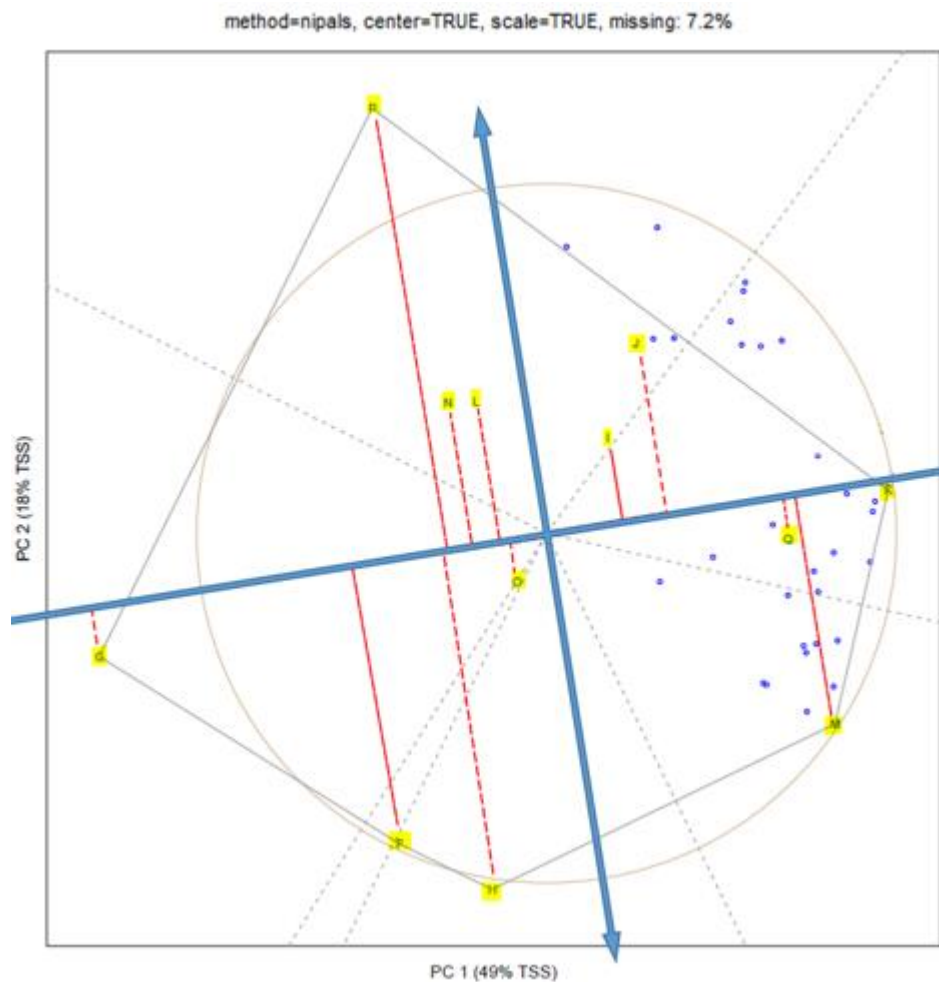


Figura 15: Biplot GGE con la visualización “media vs inestabilidad” para rendimiento en grano para las cuatro campañas. Los genotipos están resaltados en color amarillo, los ambientes corresponden a los puntos azules. El eje azul con punta de flechas corresponde al eje CAM, perpendicular a este se encuentra el eje EAP.

Caracterización de genotipos y ambientes de Argentina:

La relación entre los genotipos y las variables evaluadas en las dos campañas de ensayos en Argentina se observa en el biplot correspondiente al análisis de componentes principales de la Figura 16. El plano conformado por las dos primeras componentes principales explica el 66,1% de la variabilidad total entre genotipos, siendo el aporte de la CP1 de 42,9 y de la CP2 de un 23,2%. Este biplot para el estudio de correlaciones fenotípicas permite entender las relaciones entre las variables y también entender cómo se relacionan los genotipos con las variables evaluadas.

Los caracteres que marcaron la mayor diferenciación a nivel de la CP1 (derecha del gráfico) es días a R8, altura en R8 y días a R1 asociadas a los genotipos I, J, K, L, M, N, O, P, Q, FN4.35 y NS4611 contenidos dentro del círculo rojo dibujado en el gráfico. Estos genotipos pertenecen al rango de madurez de entre 3.7 a 4.4.

Los caracteres Vuelco y P1000 se correlacionaron positivamente con los genotipos A, B, C, D, E, F, G, H y FN3.85 (izquierda del gráfico) contenidos dentro del círculo azul dibujado en el gráfico. Estos genotipos pertenecen al rango de madurez de entre 4.3 a 4.9.

El genotipo DM4212 demarcado con un círculo amarillo se ubicó en el biplot alejado del resto de los genotipos, tuvo un comportamiento contrastante para las variables evaluadas con los demás genotipos.

La variable rendimiento en grano mostró correlación fenotípica positiva determinada con días a R8, altura en R8 y en menor medida días a R1.

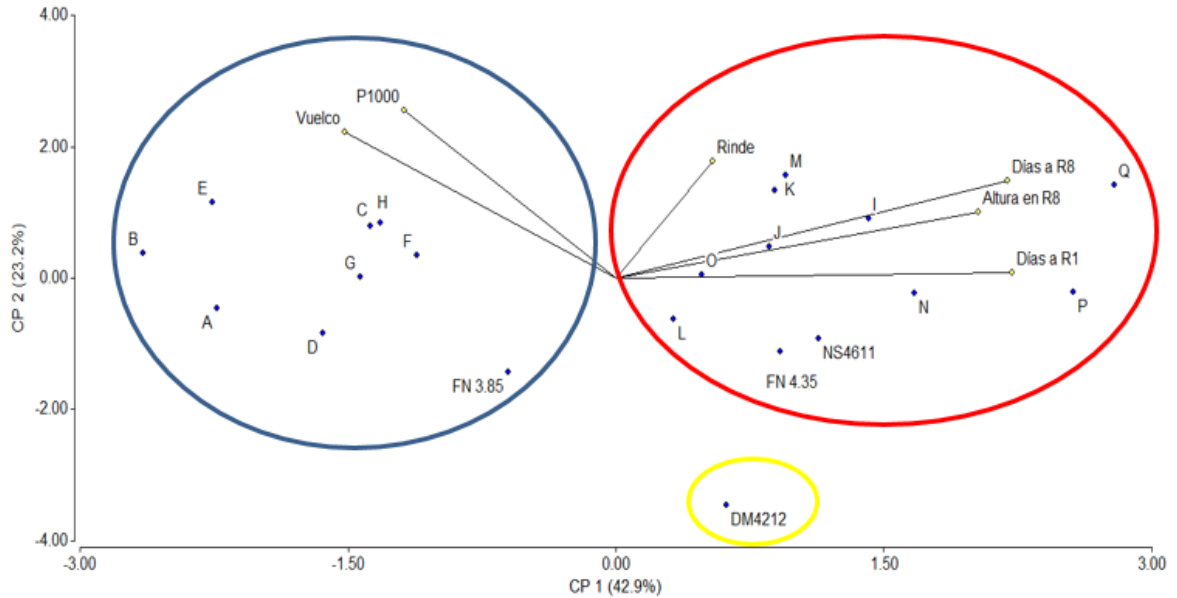


Figura 16: Biplot según el plano conformado por las dos primeras componentes principales (CP1 y CP2). Los puntos representan los genotipos en los ambientes de Argentina, campañas 2014-15 y 2015-16, y los vectores las variables evaluadas (estandarizadas).

Para verificar el agrupamiento obtenido en el biplot de correlaciones fenotípicas (Figura 16) se realizó un dendograma (Figura 17) para el análisis de conglomerados aplicado a los 21 genotipos. Se utilizó el algoritmo encadenamiento promedio sobre la matriz de distancias Euclídeas, la elección del método de encadenamiento promedio se debió a que brindó el mayor coeficiente de correlación cofenética (0,724), respecto a los otros métodos.

En el dendograma de la Figura 17 se obtuvo un agrupamiento similar al observado en el biplot de la Figura 16, mostrando tres grupos. El grupo 1 compuesto por el genotipo DM 4212, el grupo 2 por los genotipos con rango de madurez de entre 4.4 a 4.9, y el grupo 3 por los genotipos con rango de madurez de entre 3.7 a 4.4.

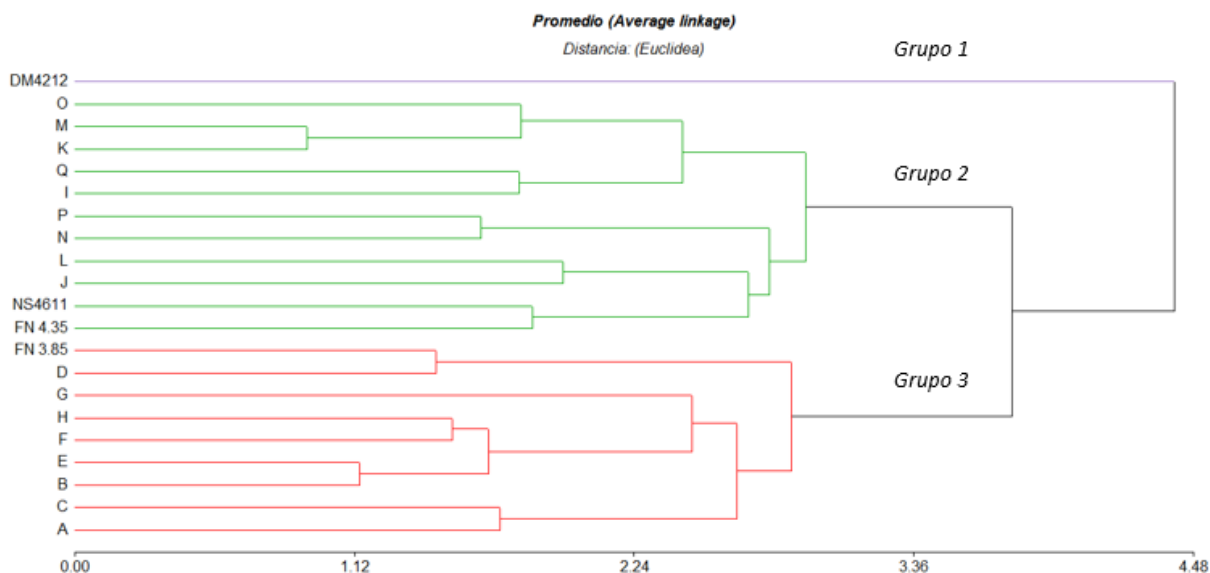


Figura 17. Dendrograma resultante del análisis de conglomerado aplicado sobre los 21 genotipos evaluados en los ambientes de Argentina, campañas 2014-15 y 2015-16, utilizando el algoritmo encadenamiento promedio sobre la matriz de distancias Euclídeas. Las variables estandarizadas utilizadas fueron Rinde, P1000, altura en R8, días a R1, días a R8 y vuelco).

La relación entre los ambientes y las variables evaluadas en las dos campañas en Argentina se observa en el biplot correspondiente al análisis de componentes principales de la Figura 18. Las dos primeras componentes principales explican el 80,4% de la variabilidad total entre ambientes, siendo el aporte de la CP1 de 53,1 y de la CP2 de un 27,3%.

La CP1 separa los ambientes extremos en cuanto a su fecha de siembra, los ambientes 37 y 44 con siembras de octubre están asociados con las variables días a R8 y días a R1 (derecha del gráfico), en el extremo opuesto del gráfico se encuentra el ambiente 45 cuya fecha de siembra fue el 28 de diciembre del 2015.

Los ambientes 39, 40, 41, 42 y 47 mostraron correlación positiva a las variables Rinde, altura y P1000, en el extremo opuesto (correlación negativa) los ambientes 36, 48 y 45. También se diferencian los ambientes 35, 38, 46 y 43 correlacionados positivamente con la variable vuelco (poco) y negativamente con las variables R1 y R8; la asociación para estas variables fue inversa para los ambientes 37 y 44.

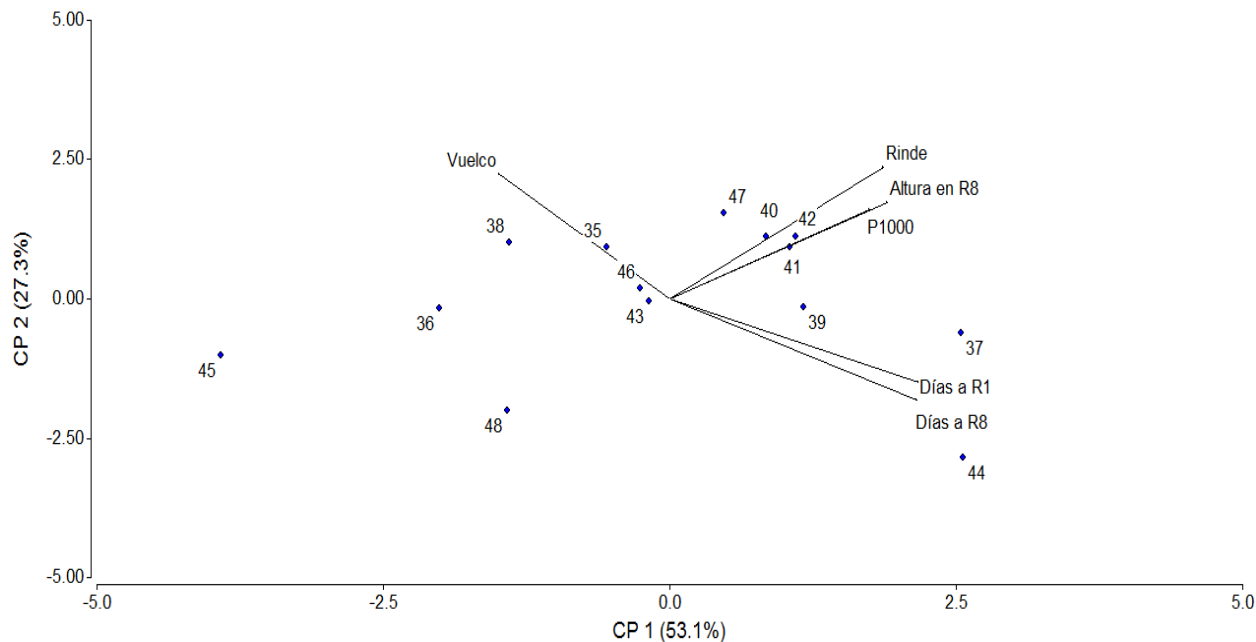


Figura 18: Biplot según el plano conformado por las dos primeras componentes principales (CP1 y CP2). Los puntos representan los ambientes de Argentina, campañas 2014-15 y 2015-16, y los vectores las variables evaluadas (estandarizadas).

El agrupamiento obtenido en el biplot de correlaciones ambientales (Figura 18) se verificó mediante un dendograma (Figura 19) para el análisis de conglomerados aplicado a los 14 ambientes. Se utilizó el algoritmo encadenamiento promedio sobre la matriz de distancias Euclídeas, la elección del método de encadenamiento promedio se debió a que brindó el mayor coeficiente de correlación cofenética (0,79), respecto a los otros métodos.

En el dendograma se encontraron similitudes en el agrupamiento de ambientes con relación a lo observado en el biplot de la Figura 19, se evidenciaron tres grupos. El grupo 1 compuesto por los ambientes 44 y 37 con fecha de siembra de octubre, el grupo 2 por los ambientes 48 y 45 de menor calidad ambiental (menor rendimiento), y el grupo 3 conteniendo los demás ambientes.

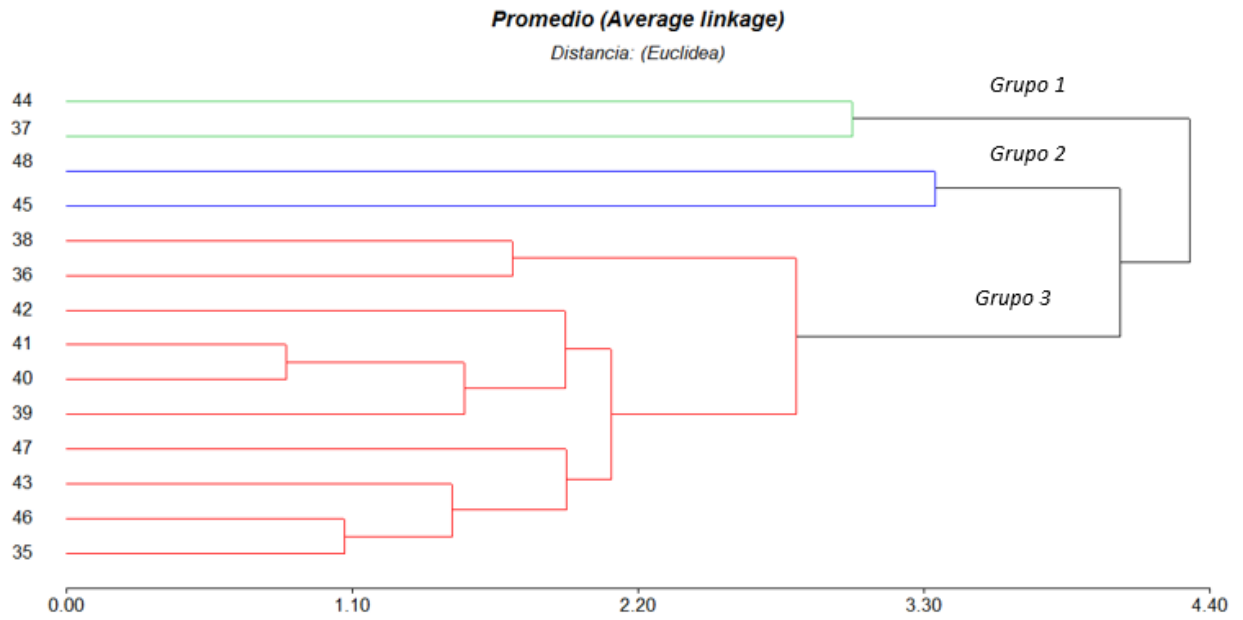


Figura 19. Dendrograma resultante del análisis de conglomerado aplicado sobre los 14 ambientes de Argentina, campañas 2014-15 y 2015-16, utilizando el algoritmo encadenamiento promedio sobre la matriz de distancias Euclídeas. Las variables estandarizadas utilizadas fueron Rinde, P1000, altura en R8, días a R1, días a R8 y vuelco).

DISCUSIÓN

El principal objetivo de los ensayos multi-ambientales es identificar los genotipos superiores para un ambiente objetivo, medido por el desempeño de un carácter en particular o varios. El ambiente objetivo está determinado por la suma de ambientes similares que se encuentran a través de localidades y años dentro de una región. En general la información proveniente de distintas campañas está desbalanceada debido a cambios en los genotipos y de las localidades evaluadas cada año. DeLacy et al. (1996) propuso cuatro estrategias para abordar este problema. La estrategia más conveniente parece ser aquella en la que los datos de varias localidades se analizan anualmente y se resumen a lo largo de los años. Solo los patrones que son consistentes a través de los años se pueden utilizar para delinear mega-ambientes, seleccionar localidades superiores y descartar aquellas que no son informativos o son redundantes.

Una medida para alcanzar este objetivo es determinar la heredabilidad del carácter en estudio. El número óptimo de localidades, años y repeticiones que permita alcanzar una heredabilidad cercana a 0,75. Si la heredabilidad obtenida es mucho mayor a 0,75, esto sugiere que se han utilizado muchos ambientes (combinación de años y localidades) o repeticiones (Yan, 2014).

La técnica biplot GGE permite seleccionar aquellas localidades más representativas y discontinuar aquellas que no lo son. El número de localidades que se requieren para obtener una heredabilidad aceptable es menor si se realizan los ensayos dentro de un mega-ambiente, logrando disminuir los costos que implican conducir los ensayos en más localidades.

Grupo Cortos

En el análisis de variancia realizado en los ensayos de rendimiento en grano en los ambientes de Argentina y Estados Unidos el componente G explicó una mayor proporción respecto a la interacción GE en cada una de las cuatro campañas individuales. El componente G fue solo significativo en la campaña 2015-16, y el componente GE fue solo no significativo en la campaña 2015. Dado que la significancia del término de interacción debe interpretarse primero antes de evaluar la significancia de los efectos principales de cada factor (E y G) es que se observa que la interacción GE fue estadísticamente significativa entre E y G, indicando que las diferencias en el rendimiento promedio entre los genotipos evaluados no se mantuvieron a través de los diferentes ambientes. El componente E fue el que más explicó la varianza total (>80%) en coincidencia con

otros autores (Dardanelli et al., 2006; Rojas et al., 2011). La relación GE/G para cada una de las campañas fue del 71, 92, 98 y 41% para las campañas 2014, 2014-15, 2015, y 2015-16, respectivamente, esta relación es una forma de evaluar la correlación genética a través de los ambientes, si es grande se puede pensar en subdividir la región en mega-ambientes (ej. >150 o 200%) (Annicchiarico, 2002). Los resultados obtenidos muestran que la relación GE/G no fue superior al 150% como sugiere Annicchiarico para pensar en subdividir la región en mega-ambientes.

En los gráficos GGE con el patrón “which won where” para cada una de las campañas individuales (Figura 6) se identificaron entre 2 y 4 mega-ambientes, estos mega-ambientes no mostraron un patrón de agrupamiento de forma extremo para cada campaña. En los gráficos biplots GGE esta relación se observa en los ángulos obtusos que se formaron entre los vectores que representan las localidades. La significancia de la interacción GE encontrada en el ANOVA se puede interpretar gráficamente en el biplot GGE de la Figura 6 con los genotipos más extremos para cada mega-ambiente.

Cuando se realizó en el ANOVA conjunto de los cuatro campañas, donde los ambientes correspondieron a la combinación de años y localidades, todas las fuentes de variancia fueron significativas, siendo los componentes GE y E altamente significativos. Aquí la relación GE/G fue del 140%, siendo mayor que para cada una de las campañas individuales aunque aún su magnitud no fue importante para considerar dividir a la región evaluada en mega-ambientes. Otra forma de evaluar la existencia de mega-ambientes es mediante el cálculo de la heredabilidad del carácter rendimiento. Yan (2014) indica que cuando el valor es muy alto (>0,9) es muy probable que la región en estudio constituya un mega-ambiente homogéneo simple. El resultado de heredabilidad obtenido con los 8 genotipos y los 32 ambientes evaluados en las cuatro campañas es de 0,92, este resultado contribuye a pensar que la región evaluada constituya un mega-ambiente simple. Cuando se realizó el cálculo de heredabilidad para los ambientes solo de EE. UU. (2 campañas) el mismo mostró un valor muy alto de 0,87, aunque el mismo fue menor que para las cuatro campañas.

Sin embargo, tanto en los gráficos biplots GGE de los cuatros campañas combinadas como en los gráficos de las campañas individuales, si se diferenciaron mega-ambientes. Los gráficos biplot GGE contiene la información de G e interacción GE, donde la interacción GE está compuesta por GL, GA y GLA en ensayos multi-ambientales (ambientes como combinación de años y localidades). Como se observa en la figura 7C el componente de interacción GA tuvo un peso de moderada importancia en la discriminación de localidades, aquí los ángulos entre los vectores de regresión de cada campaña formaron un ángulo agudo entre sí, siendo las campañas de EE. UU. (14 y 15) y Arg. (14-15 y 15-16) más parecidas entre sí respectivamente.

Cuando se realizaron los gráficos GGE con las regresiones de cada estado/provincia (Figura 7D) para entender si los ambientes de ellos mostraban algún patrón de correlación, se observó que efectivamente hay estados/provincias que correlacionaron mejor entre ambos países. Los resultados obtenidos mostraron a los ambientes evaluados en los estados de IL y OH como de mejor correlación con los ambientes de Arg. Por otro lado, los ambientes de KS tuvieron menor correlación con los ambientes de Arg. para las campañas evaluadas. Los ambientes del estado de KS a diferencia de IL y OH son de menor potencial de rendimiento en grano. El estado de KS presenta un clima continental más acentuado, en donde las precipitaciones se distribuyen casi exclusivamente durante el ciclo estival, esto sumado a que los suelos son de textura más arenosa, los convierte en ambientes más inestables. En Arg. ambientes similares a los de KS se pueden encontrar al oeste de la provincia de Bs. As., este de La Pampa y en Córdoba.

Grupo Largos

Los resultados obtenidos en el análisis de variancia para cada una de las campañas individuales muestran que el componente G explicó una mayor proporción que el componente GE en tres de las cuatro campañas evaluadas. El componente GE fue altamente significativo en los cuatros campañas, en tanto que el componente G fue significativo en tres de las cuatro campañas. Al igual que en el grupo “cortos” dado que la interacción es significativa, el ordenamiento de los genotipos no se mantuvo en las distintas localidades, siendo no recomendado utilizar las comparaciones de medias entre genotipos a través de las localidades para seleccionar genotipos. El componente E fue el que más explicó la varianza total (>67%). La relación GE/G para cada

una de las campañas fue del 142, 62, 86, y 25% para las campañas 2014, 2014-15, 2015 y 2015-16, respectivamente. La campaña 2014 tuvo una mayor magnitud, aunque en ninguna de las cuatro campañas individuales esta relación fue importante ($>150\%$) que sugiera dividir la región en mega-ambientes.

En los gráficos GGE para cada una de las campañas individuales (Figura 13) se verificó gráficamente las magnitudes de la relación de GE/G obtenidas en el análisis de variancia con genotipos extremos en cada mega-ambiente. En estos gráficos si se diferenciaron mega-ambientes. En la campaña 2014 se diferenció el mayor número de mega-ambientes (tres), y los ambientes de la campaña 2015-2016 conformaron un único mega-ambiente.

En el análisis de variancia para las cuatro campañas combinadas, y al igual que para el grupo “cortos” todas las fuentes de variación fueron significativas, siendo el componente E y GE, altamente significativos. La relación GE/G fue del 133%, aunque fue mayor que para cada una de las campañas individuales, no muestra una tendencia clara para considerar dividir la región en mega-ambientes. Tanto la heredabilidad obtenida que fue del 0,94 y la Figura 12 mostrando la correlación positiva y alta entre casi todos los ambientes, fortalecen la idea de que los ambientes evaluados pertenecen a un mismo mega-ambiente.

En los gráficos GGE de las campañas combinadas (Figura 14) se diferenciaron tres mega-ambientes, el vector de regresión (Figura 14C) mostró una separación clara entre ambos países, aunque el ángulo entre vectores fue menor a 90° . Los ambientes de EE. UU. mostraron una mayor dispersión (mayor interacción GE) que los ambientes de Arg, esto se debió a que los ambientes evaluados en EE. UU. abarcaron una amplia región desde los 33.5° a 38.5° de latitud Norte con condiciones climáticas y edáficas más diversas que los ambientes de Arg.

Sin embargo, se puede apreciar que hay ambientes entre ambos países que se agruparon en el mismo mega-ambiente, siendo su correlación alta y positiva. Como se puede apreciar en la figura 14C los ambientes de los estados de IL y MS fueron los que más correlacionaron con los ambientes de Arg, en cambio los ambientes de AR no mostraron correlación (ángulo de 90° entre vectores). La principal diferencia que tuvieron los ambientes de AR fue el tipo de suelo, ya que

estas localidades en cercanías al río Misisipi contienen suelos del orden Vertisoles. Estos suelos están compuestos por un horizonte superficial y sub superficial de textura arcillo, arcillo-limoso donde el contenido de arcilla es mayor al 30% en la fracción tierra-fina. Este orden de suelos también se encuentra en zonas agrícolas de Arg., ampliamente distribuidos en la provincia de Entre Ríos y también en algunas localidades de la provincia de Bs.As. Esto brinda la oportunidad de ampliar el área de la red de ensayos hacia estos ambientes con suelos similares para estudiar la existencia de correlación entre ambos países.

Con los resultados obtenidos para ambos grupos (“cortos” y “largos”) no se pretende extrapolar sus conclusiones al total de la superficie que comprende cada estado/provincia, ni cada país ya que el número de ambientes y genotipos evaluados en cada caso fue acotado. De todas formas, permite tener una primera aproximación para futuras investigaciones y de confirmarse este patrón, identificar con mayor precisión cuáles son las variables bióticas y abióticas que en mayor medida explican las diferencias entre ambientes.

La técnica gráfica biplot GGE también permitió evaluar los genotipos de ambos grupos (“cortos” y “largos”) y seleccionar aquellos de mejor comportamiento en cuanto a su rendimiento y estabilidad utilizando la técnica desarrollada por Yan et al. (2007). Esta técnica permite seleccionar los genotipos de adaptación general o en sentido amplio (a través de los ambientes) o seleccionar para ambientes específicos (para cada ambiente o grupo de ambientes). Mejorar por alta estabilidad de rendimiento puede considerarse un objetivo útil cuando la variación en la interacción GE es grande (Balzarini et al., 2005) Dado que en este trabajo la interacción GE fue significativa pero la proporción GE/G fue moderada a baja en el análisis de campañas combinadas, es probable que la región objetivo esté compuesta por uno o pocos mega-ambientes. Con estos resultados la estrategia de selección sería la de elegir un grupo de genotipos para toda la región (ambos países) basados en su rendimiento y estabilidad. Atlin et al. (2000), encontraron en el cultivo de cebada que los programas de mejoramiento que prueban las variedades por adaptación amplia pueden lograr mejores rendimientos que aquellos que lo hacen en regiones específicas.

Con el objetivo de caracterizar todos los genotipos utilizados en este trabajo es que se registraron distintas variables fenotípicas en los ambientes de Argentina durante las campañas 2014-15 y 2015-16. Los 17 genotipos desarrollados para el mercado de EE. UU. y los 4 genotipos desarrollados para el mercado de Arg. se agruparon de manera similar según su GM. Siendo los genotipos de grupo de madurez largos los que tuvieron correlación positiva con las variables, altura de planta en R8, días a R8, días a R1 y rendimiento en grano. En cambio, los grupos de madurez cortos tuvieron correlación positiva con las variables peso de granos y vuelco (menor). Los grupos que se diferenciaron en los gráficos biplot (Figura 16) y dendograma (Figura 17) coinciden con los genotipos evaluados en los grupos “cortos” y “largos” agrupados en los ambientes de EE. UU. según su adaptación al ambiente agroecológico para el que fueron seleccionadas. Estos resultados muestran que dentro del germoplasma originado para EE. UU. hay genotipos con características y comportamientos similares a las variedades comerciales del mercado Arg., esto permitirían utilizarlo como material base para el desarrollo de un programa de mejoramiento en Arg.

También mediante los biplot y dendograma con las variables evaluadas permitieron diferenciar los ambientes utilizados en Argentina. Las localidades cuya fecha de siembra fue temprana, sembradas el mes de octubre (ambas campañas) fueron las de ciclo de crecimiento más largo ya que tuvieron correlación positiva con las variables días a R1, días a R8 y también tuvieron mayor vuelco. Los ambientes de mayor rendimiento fueron los que mayor correlación tuvieron con las variables altura a R8 y peso de 1000 granos. Los ambientes de menor potencial de rendimiento fueron aquellas en las que el ciclo de crecimiento fue más corto, tuvieron menor altura de planta en R8, menor peso de 1000 granos y menor vuelco. Estos resultados concuerdan con diversos trabajos presentados (Baigorri et al., 1997; De Battista y Arias, 2009; Perez y Perez, 2010).

CONCLUSIONES

En este trabajo se pudo corroborar la hipótesis planteada y concluir que existen ambientes comunes para seleccionar variedades para ambos países para el rango de madurez empleado. Diferentes estrategias se utilizaron para entender el efecto de la interacción GE en el rendimiento de los genotipos a través de las distintas campañas y ambientes. Se utilizó la metodología gráfica de biplot GGE para visualizar el agrupamiento y correlación de los ambientes evaluados permitiendo también diferenciar simultáneamente por comportamiento y estabilidad de genotipos.

Los resultados obtenidos en los ensayos de Argentina con variedades locales muestran que dentro de una base de germoplasma (GM entre III y IV largo) originado y seleccionada para ambientes de Estados Unidos hay varios genotipos que presentan una amplia adaptación y un muy buen rendimiento. Esto permite visualizar la potencialidad del pool genético para ser utilizado como material base para el desarrollo de un programa de mejoramiento de la empresa Pioneer en la Argentina; y, por otro lado, de maximizar el uso de los recursos priorizando la selección de variedades de amplia adaptación.

Para un programa de mejoramiento la existencia de mega-ambientes comunes entre distintos hemisferios (ej. Estados Unidos y Argentina) brinda la posibilidad de seleccionar los mejores genotipos en ambientes de contra-estación, además permite aumentar la ganancia genética por año, ya que los ambientes evaluados en ambos países se complementan para lograr dos ciclos de cultivo por año.

La metodología empleada abre la posibilidad de extender el análisis utilizando más ambientes, incorporar un número mayor de genotipos, y la investigación adicional de variables climáticas, edáficas y factores bióticos para relacionarlas con el rendimiento en grano. Ello permitiría enriquecer el conocimiento de las interacciones GE entre ambos países y hacer un uso más eficiente del germoplasma de la empresa. Una mejor comprensión de los efectos GE también

contribuirá al estudio y la aplicación de métodos de selección molecular para el mejoramiento genético de este importante cultivo.

BIBLIOGRAFÍA

- Annicchiarico, P. 2002. Genotype x Environment Interactions - Challenges and Opportunities for Plant Breeding and Cultivar Recommendations. FAO Plant Production and Protection publ. 174.
- Atlin, G., Baker, R., McRae, K. y Lu, X. 2000. Selection response in subdivided target regions. *Crop Sci.* 40: 7-13.
- Baigorri, H.E., Giorda, L. 1997. El cultivo de la soja en Argentina. INTA-Centro Regional Córdoba EEA Marcos Juárez / EEA Manfredi / Coordinación Subprograma Soja / Editar, p.448.
- Baker, R.J. 1988. Tests for crossover genotype-environmental interactions. *Can. J. Plant Sci.*, 68: 405-410.
- Balzarini, M., Bruno, C., Arroyo, A. 2005. Análisis de ensayos agrícolas multi-ambientales: Ejemplos con Info-Gen. Universidad Nacional de Córdoba p. 21-29.
- Balzarini, M., Macciavelli, R. Casanoves, F. 2006. Aplicaciones de modelos mixtos en Agricultura y forestería. Notas de Curso Ofrecido en CATIE (Turrialba, Costa Rica). p. 1-58.
- Becker, H.C., León, J. 1988. Stability analysis in plant breeding. *Plant Breed.* 101: 1-23.
- Braun, H., Pfeiffer, W., Pollmer, W. 1996. CIMMYT's approach to breeding for wideadaptation. *Euphytica* 92: 175-183.
- Cooper, M., DeLacy, I.H., Basford, K.E. 1996. Relationships among analytical methods used to study genotypic adaptation in multi-environment trials. In M. Cooper & G.L. Hammer, eds. *Plant adaptation and crop improvement*, p. 193-224. Wallingford, UK, CABI.
- Cornelius, P.L., Crossa, J., Seyedsadr, MS. 1996. Statistical test and estimators of multiplicative models for genotype-by-environment interaction. In M.S. Kang y H.G. Gauch, eds. *Genotype-by-environment interaction*, p. 199-234. Boca Raton, FL, CRC Press.
- Crossa, J., Cornelius, P.L. 1997. Sites regression and shifted multiplicative model clustering of cultivar trials sites under heterogeneity of variantes. *Crop Sci.* 37: 406-415.
- Dardanelli, J.L., Balzarini, M., Martinez, M.J., Cuniberti, M., Resnik, S., Ramunda, S.F., Herrero, R., Baigorri, H. 2006. Soybean maturity groups, environments, and their interaction define mega-environments for seed composition in Argentina. *Crop Sci.* 46: 1939-1947.
- De Battista, J., Arias, N. 2009. Comportamiento de cultivares de soja en fechas de siembra. Resultados 2008/09. En: *Cultivo de soja en el centro este de Entre Ríos. Resultados 2008/09.* Bol. Téc. Serie Prod. Veg. N° 50. INTA EEA C. del Uruguay. Pág. 49- 65.

De la Vega, A.J., Chapman, S.C. 2010. Mega-Environment differences affecting genetic progress for yield and relative value of component traits. *Crop Sci.* 50: 574-583.

DeLacy I.H., Basford K.E., Cooper M., Bull J.K., McLauren C.G. 1996. Analysis of multi-environment trials - An historical perspective. En: M. Cooper, G.L. Hammer (Ed.). *Plant Adaptation and Crop Improvement*. CAB INTERNATIONAL, p. 39-124.

Fehr, W.R., Caviness C.E. 1977. Stages of soybean development. Iowa State University, Special Report 80, Ames, Iowa, p. 12.

Friendly, Michael. 2002. "Corrgrams: Exploratory Displays for Correlation Matrices." *The American Statistician* 56: 316–24. doi:10.1198/000313002533.

Gabriel, K.R. 1971. The biplot graphic display of matrices with application to principal component analysis. *Biometrika* 58: 453-467.

Gauch, H.G, Zobel, R. 1997. Identifying mega-environments and targeting genotypes. *Crop Sci.* 37: 311-326.

Holland, J.; Nyquist, W., Cervantes, C. 2003. Estimating and interpreting heritability for plant breeding. *Plant Breeding Reviews*. 22: . 9-112.

Kang, M.S. 1998 Using genotype-by-environment interaction for crop cultivar development. *Advances in Agronomy* 62: 199–252.

Lin C, Binns M, Levkowitch L, 1986. Stability analysis: where do we stand? *Crop Sci* 26:894-900.

Mohammadi, R., Haghparast, R., Amri, A., Ceccarelli, S. 2010. Yield stability of rainfed durum wheat and GGE biplot analysis of multi-environment trials. *Crop Pasture Sci* 61: 92-101.

Muir, W., Nyquist, W.E., Xu, S. 1992. Alternative partitioning of the genotype-by environment interaction. *Theor. Appl. Genet.* 84: 193-200.

Patterson, H.D., Thompson, R. 1975. Maximum likelihood estimation of components of variance. In *Proc. of the 8th Internatl. Biometrics Conf.* p. 197–207.

Pérez, L., Pérez, M. 2010. Comportamiento de cultivares de soja según fecha de siembra. *Memoria Técnica 2009-2010 EEA INTA General Villegas* p. 61-63.

Rojas, E, Bologna, S., Di Renzo, M., Soldini, D., Salines, L. 2011. Interacción genotipo ambiente en líneas avanzadas de soja. Mercosoja 2011 p. 1-4.

USDA. 2017. Crop production 2016 summary. ISSN: 1057-7823.

Wricke, G., Weber, W.E. 1986. Quantitative genetics and selection in plant breeding. Berlin, W. de Gruyter. ISBN: 311007561X.

Yan, W. 2001. GGE Biplot. A Windows application for graphical analysis of multienvironment trial data and other types of two-way data. Agron. J. 93: 1111-1118.

Yan, W. 2012. Biplot analysis of incomplete two-way data. Crop Sci. 53:48-57.

Yan, W. 2014. Crop Variety Trials, Data Management and Analysis. Ottawa, John Wiley y Sonic, Inc. ISBN: 9781118688649.

Yan, W., Rajcan, I. 2002. Biplot analysis of test sites and trait relations of soybean in Ontario. Crop Sci. 42: 11-20.

Yan, W., Rajcan, I. 2003. Prediction of cultivar performance based on single-versus multiple-year tests in soybean. Crop Sci. 43: 549-555.

Yan, W., Hunt, L. A., Sheng Q., Szlavnic, Z. 2000. Cultivar evaluation and mega-environment investigation based on the GGE biplot. Crop Sci. 40: 597-605.

Yan, W., Kang, M.S., Ma, B., Woods, S., Cornelius P.L. 2007a. GGE Biplot vs. AMMI analysis of genotype-by-environment data. Crop Sci. 47: 641-653.

Yan, W., Tinker, A.A., Molnar, S., Fregeau-Reid J., and McElroy, A. 2007b. Associations among oat traits and their response to the environment in North America. Journal of Crop Improvements 20: 1-30.

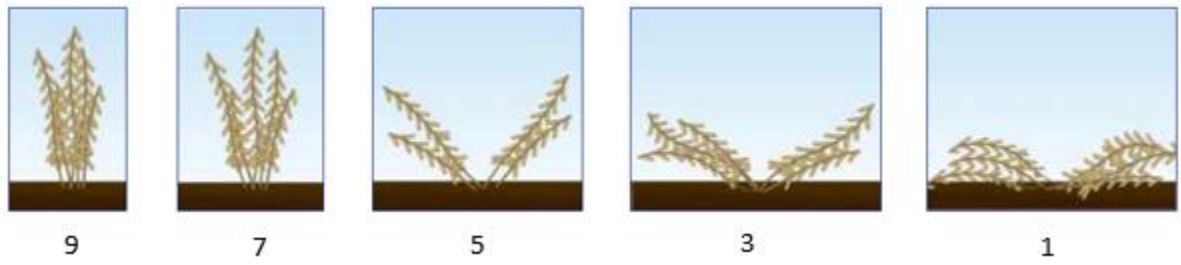
Yan, Z., Lauer, J.G, Borges, R., De Leon, N. 2010. Effects of genotype \times environment interaction on agronomic traits in soybean. Crop Sci. 50: 696-702.

Yang, R.C., J. Crossa, P. L. Cornelius, and J. Burgueño. 2009. Biplot analysis of genotype \times environment interaction: Proceed with caution. Crop Sci. 49: 1564-1576.

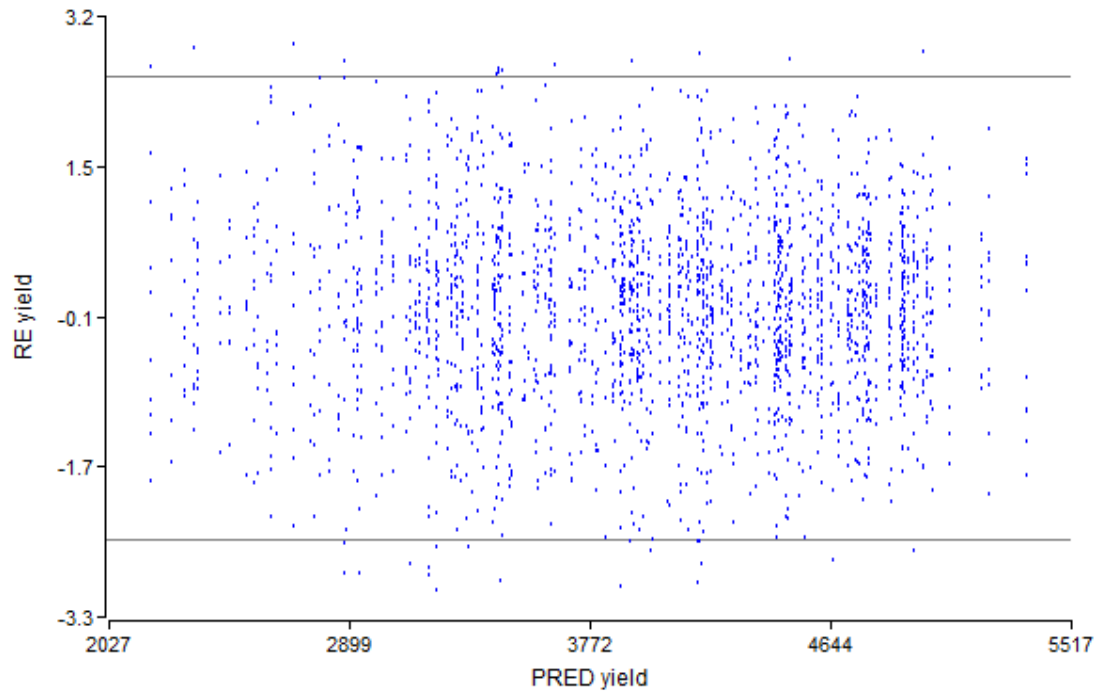
Ybran, R., Lacelli, G. 2016. Informe estadístico del mercado de la soja. Informe INTA- E.E.A. Reconquista. p. 5

Zhang, M., Kang, S.M., Reese, Jr.P.F., Bhardwaj, H. 2005. Soybean cultivar evaluation via GGE biplot analysis. *Journal of New Seeds*7(4): 37-50.

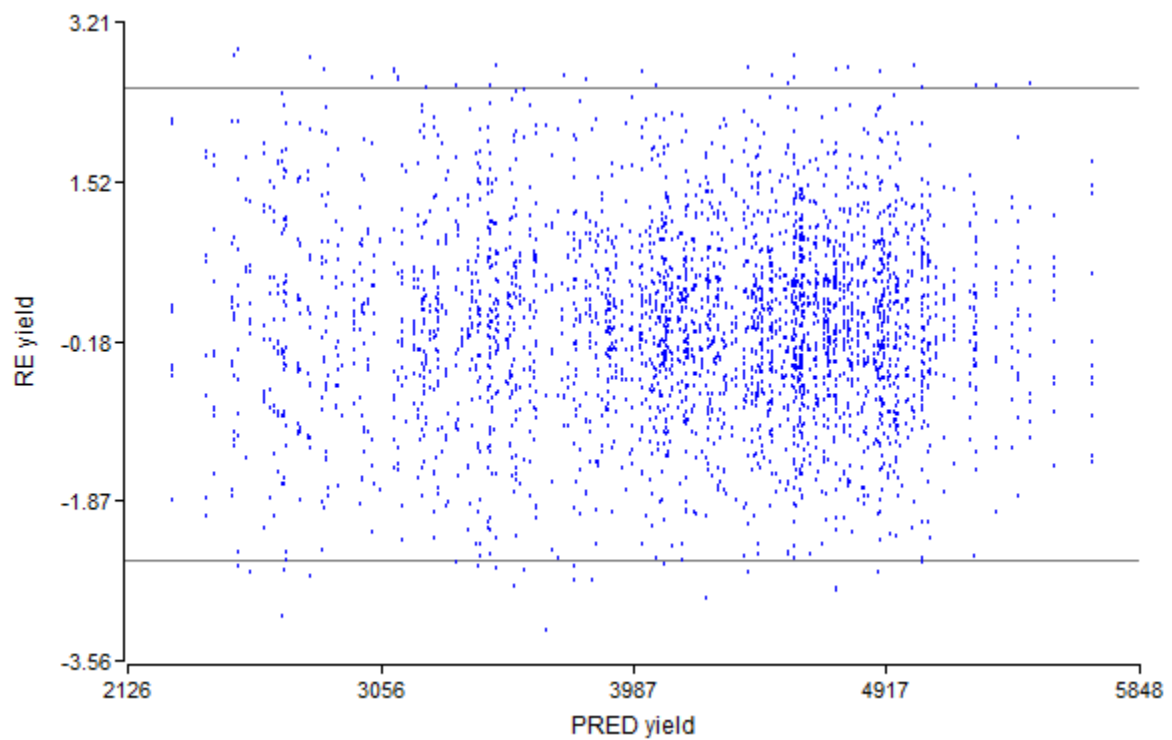
ANEXO



Anexo 1: Puntaje de vuelco, siendo 9 el valor para parcelas con plantas erectas y 1 para parcelas con plantas totalmente volcadas



Anexo 2: Gráfico de residuos en función de predichos para el grupo de Madurez “cortos”.



Anexo 3: Gráfico de residuos en función de predichos para el grupo de Madurez “largos”.