

Interacción genotipo x ambiente en Soja. Un modelo espacial y dinámico para definir y localizar ambientes en Sudamérica

*Tesis presentada para optar al título de Magister de la Universidad de Buenos Aires,
Área: Producción Vegetal*

José María Clavijo

Ingeniero Agrónomo - Universidad de Buenos Aires - 2009

Lugar de trabajo: IFEVA-Don Mario Semillas



Escuela para Graduados Ing. Agr. Alberto Soriano
Facultad de Agronomía – Universidad de Buenos Aires



COMITÉ CONSEJERO

Consejero principal
Martín Oesterheld
Ing. Agrónomo (UBA)
PhD (Syracuse University)

Consejero
Gervasio Piñeiro
Ing. Agrónomo (Universidad de la República)
Doctor en Ciencias Agropecuarias (UBA)

Consejera
Adriana Kantolic
Ing. Agrónoma (UBA)
Doctora en Ciencias Agropecuarias (UBA)

JURADO DE TESIS

Jurado
José Luis Rotundo
Ing. Agrónomo (UBA)
MSc. en Recursos Naturales (UBA)
Doctor (Iowa State University)

Jurado
Pablo Antonio Calviño
Esp. en Producción Vegetal (UNMdP)
Doctor en Ciencias Agrarias (UNMdP)

Jurado
Constanza Soledad Carrera
Ing. Agrónoma (UNCor)
Doctora en Ciencias Agropecuarias (UNCor)

Fecha de defensa: 04 de Diciembre de 2018

DEDICATORIA

A Dios y a mi familia: a mi esposa Agustina, a mis padres, hermanos y sobrinos. A mis padrinos y ahijados.

A mi comité consejero.

AGRADECIMIENTOS

Al Buen Dios por la vida y por las personas que puso a mi disposición, que con su paciencia y entrega han colaborado de algún modo en esta intensa y larga carrera. A mis hermanos y a toda mi familia. A mi esposa Agustina y a mis padres Raúl y Teresa por su apoyo constante e incondicional, por sostenerme en los momentos de dificultad, por ayudarme a perseverar y por alegrarse conmigo en cada pequeño y gran avance.

A mi comité consejero por sus grandes aportes en el desarrollo del trabajo, pero sobre todo en mi formación como profesional. En especial a mi director Martín Oesterheld, por el esfuerzo y generosidad que empeña en la formación de sus estudiantes, entre los cuales tuve el privilegio de ser contado.

A todos los compañeros del LART, en especial a Gonzalo García Accinelli, con quien nos embarcamos en el extraño mundo del procesamiento y análisis de datos, por su participación activa en este trabajo, tanto en la construcción de algoritmos y bases de datos como con sus ideas geniales.

A Don Mario por el apoyo económico. A sus directores y gerentes por las provechosas rondas de discusión y a los compañeros de Desarrollo e Investigación por su valiosa tarea en la generación y preparación de datos.

DECLARACIÓN

Declaro que el material incluido en esta tesis es, a mi mejor saber y entender, original producto de mi propio trabajo (salvo en la medida en que se identifique explícitamente las contribuciones de otros), y que este material no lo he presentado, en forma parcial o total, como una tesis en esta u otra institución.

José María Clavijo
DNI: 31.877.396

ÍNDICE GENERAL

DEDICATORIA.....	III
AGRADECIMIENTOS.....	IV
DECLARACIÓN	V
ÍNDICE GENERAL	VI
ÍNDICE DE FIGURAS	VIII
ÍNDICE DE TABLAS.....	XII
ABREVIATURAS	XIII
Resumen	XV
Abstract.....	XVI
CAPÍTULO 1: Introducción.....	1
1.A. Interacción genotipo x ambiente en soja.....	1
1.B. Mapeo de la variación de rendimiento por genotipos	5
1.C. Objetivos	5
1.D. Estructura general de la tesis.....	6
CAPÍTULO 2: Materiales y métodos.....	6
2.A. Aproximación general y área de estudio.....	6
I. Área de estudio	7
II. Caracterización del área de estudio en función de la precipitación anual	10
2.B. Interacción genotipo x sitio-fecha de siembra y mega-ambientes	12
I. Información de rendimiento de genotipos de soja	12
II. Análisis de la interacción y detección de mega-ambientes	15
2.C. Caracterización ambiental de los mega-ambientes	17
I. Información ambiental	17
II. Variables ambientales asociadas a los mega-ambientes.....	20
2.D. Mapeo dinámico de mega-ambientes.....	22
I. Mapa de mega-ambientes en un escenario temporal de referencia.....	22
II. Mapeo de la variación ambiental asociada a mega-ambientes	23
III. Mapas de mega-ambientes en distintos escenarios temporales (mapeo dinámico)	23
CAPÍTULO 3: Resultados.....	24
3.A. Interacción genotipo x sitio-fecha de siembra y mega-ambientes	24
3.B. Caracterización ambiental de los mega-ambientes	28
3.C. Mapeo dinámico de mega-ambientes.....	31
I. Mapa de mega-ambientes en un escenario temporal de referencia.....	31

II. Mapeo de la variación ambiental asociada a mega-ambientes	36
III. Mapas de mega-ambientes en distintos escenarios temporales (mapeo dinámico)	38
CAPÍTULO 4: Discusión	41
4.A. Interacción genotipo x sitio-fecha de siembra y mega-ambientes	41
4.B. Caracterización ambiental de los mega-ambientes	43
4.C. Mapeo dinámico de mega-ambientes	45
CAPÍTULO 5: Conclusiones generales.....	49
Bibliografía.....	51
Apéndice.....	60
A.I. Variables ambientales meteorológicas y de índice de vegetación.....	60
A.II. Rutina automatizada de búsqueda de subconjuntos completos.....	63

ÍNDICE DE FIGURAS

Figura 2.1: Esquema metodológico general. A partir de información de rendimiento de genotipos en distintos sitios, años y fechas de siembra se identificaron grupos de sitios-años-fecha de siembra con similar ranking de rendimiento de los genotipos (mega-ambientes) (a). Sobre la base de información meteorológica, de índice de vegetación y edáfica (b), los mega-ambientes fueron caracterizados por rangos ambientales y se les asignó la categoría de “verdaderos ambientes” (c). A partir de la información ambiental en toda el área de estudio, los mega-ambientes fueron mapeados en distintos escenarios de campaña y fechas de siembra (d). “Ppt.”: precipitación. “Temp.”: temperatura. “DAS”: días antes de la siembra. “DDS”: días desde la siembra. “FS”: fecha de siembra. “ME”: mega-ambiente.

Figura 2.2: Área de estudio (cuadrícula blanca) y unidades biogeográficas pertinentes (en colores) (Morrone 2014). Líneas negras indican límites internacionales, provinciales (Argentina) y de estados (Brasil). Marcas amarillas indican sitios experimentales (n=96).

Figura 2.3: A: Anomalía de la precipitación anual para seis campañas, calculada como la diferencia entre la precipitación de cada campaña y la precipitación promedio de las campañas 2000/01 a 2012/13 (n=13), expresada en forma relativa a la precipitación promedio. En gris, se muestran las áreas excluidas del análisis. Para cada campaña, se consideró la precipitación acumulada entre junio del año inicial y mayo del siguiente. Líneas negras indican límites de las unidades biogeográficas propuestas por Morrone (2014). B: Unidades biogeográficas propuestas por Morrone (2014) dentro del área de estudio de esta tesis (Tabla 2.2).

Figura 2.4: Coeficiente de variación inter-anual de la precipitación anual de las campañas 2000/01 a 2012/13 (n=13). En gris, se muestran las áreas excluidas del análisis. Líneas negras indican límites de unidades biogeográficas propuestos por Morrone (2014) (Fig. 2.3 B; Tabla 2.2).

Figura 2.5: Fecha de siembra temprana (izquierda) y tardía (derecha) estimada por referentes. El color indica la semana de siembra. La fecha de siembra fue consultada a referentes para una grilla de 4x4° que cubría el área de estudio en su totalidad, y fue volcada a la grilla de 0.5x0.5° por interpolación espacial con el método de kriging. Las fechas de siembra consultadas representan una fecha temprana óptima y una fecha tardía frecuente, no extrema. Líneas negras indican límites de unidades biogeográficas propuestos por Morrone (2014) (Fig. 3.8, mapa 9; Tabla 2.2)

Figura 3.1: Porcentaje de la suma de cuadrados total explicada por los factores ambiental, representado por el sitio-fecha de siembra (Sitio-FS; fecha de siembra incluyendo el año y el día del año), genotípico (Genotipo), la interacción entre ambos (GxS-FS) y el residuo.

Se muestran los promedios y los desvíos estándar (n=8) de los ANOVA realizados para los subconjuntos completos (barras negras) y para los mega-ambientes (barras grises).

Fig. 3.2: Pendientes y valores p de las regresiones entre rendimiento centrado por ambiente en kg ha⁻¹ y GM de 0 a 100, para cada mega-ambiente. Se muestran todos los mega-ambientes definidos en los 8 subconjuntos completos (n=34). Valores de pendiente positivos indican mega-ambientes con rendimientos menores para GM más cortos, y mayores para GM más largos; valores negativos, lo contrario. Los puntos por debajo de la línea roja representan regresiones estadísticamente significativas (p<0,05).

Fig. 3.3: Efecto del hábito de crecimiento sobre el rendimiento cuantificado por los residuales de la regresión por GM (Fig. 3.2). Se muestran los residuales para hábito de crecimiento determinado (D) e indeterminado (I). A: promedio y error estándar de los 34 mega-ambientes definidos en los 8 subconjuntos completos; B: a partir de los mismos 34 mega-ambientes, se discriminó entre aquellos con pendiente significativa (p<0.05; n=23) y no significativa (n=11) respecto del GM.

Figura 3.4. A: Interacción genotipo por sitio-fecha de siembra (GxS-FS) entre pares de genotipos ordenados por GM creciente. La interacción fue cuantificada por la distancia euclídea estandarizada promedio para cada par de genotipos: mayor distancia, mayor interacción GxS-FS. En ambos ejes se indican nombre, grupo de madurez (GM; de 0 a 100) y hábito de crecimiento de los genotipos (I: indeterminado, D: determinado). Puntos cercanos a la diagonal representan combinaciones entre dos genotipos con GM semejante; puntos alejados de la diagonal, combinaciones entre genotipos de GM dispar. Se incluyen todos los genotipos de los 8 subconjuntos completos. **B:** Las mismas comparaciones de pares de genotipos de la figura A, fueron agrupadas en tres categorías: comparaciones entre genotipos de GM menor a VI (valores de GM menor a 60 en ambos ejes, X e Y), comparaciones entre genotipos de GM mayor a VI (valores de GM mayor a 60 en ambos ejes) y comparaciones de genotipos de GM menor a VI con genotipos de GM mayor a VI (valores de GM menor a 60 en eje X y mayor a 60 en eje Y). Se muestran las distribuciones de valores de distancia para las tres categorías definidas.

Figura 3.5: Interacción genotipo x sitio-fecha de siembra (GxS-FS) entre pares de genotipos ordenados por GM creciente circunscripta a los genotipos de hábito indeterminado. La interacción está cuantificada por la distancia euclídea estandarizada promedio para cada par de genotipos: mayor distancia, mayor interacción GxS-FS. En cada eje se indican nombre, grupo de madurez (GM; de 0 a 100) y hábito de crecimiento de los genotipos. El eje vertical contiene los mismos genotipos que el eje horizontal. Puntos cercanos a la diagonal representan combinaciones entre dos genotipos con GM semejante; puntos alejados de la diagonal, combinaciones entre genotipos de GM dispar.

Figura 3.6: Variables ambientales (Tabla 3.1) más asociadas con la interacción genotipo x sitio-fecha de siembra (GxS-FS) y su grado de asociación con dicha interacción, cuantificado por el porcentaje de explicación ponderado promedio. Se muestran las variables que fueron importantes para la mayoría de los subconjuntos completos (A;

variables persistentes) y aquellas que solo lo fueron para algunos de ellos (B; variables no persistentes). Dentro de una barra, los colores indican los distintos subconjuntos. La longitud total de una barra representa la suma de los porcentajes de explicación de los 8 subconjuntos. Se muestran los nombres abreviados, conformados por tres partes separadas por guiones. La primera indica la variable ambiental: “pef”: profundidad efectiva; “fp”: fotoperíodo; “pp”: precipitación; “crh”: capacidad de retención hídrica; “td”: temperatura diurna; “tm”: temperatura media; “tn”: temperatura nocturna; “ivn”: índice de vegetación normalizado. La segunda parte indica el estadístico de resumen temporal: “prom”: promedio; “max”: máximo; “min”: mínimo; “acum”: acumulada. La tercera parte indica el período dentro del ciclo del cultivo en días desde la fecha de siembra. El primer número indica el inicio del período y el segundo el fin. Valores negativos y positivos indican fechas anteriores y posteriores a la siembra, respectivamente.

Figura 3.7: Incidencia relativa de cada tipo de variable (A) y período temporal (B) sobre la interacción GxS-FS, calculada como el cociente entre dos proporciones: la proporción dentro de las 10 variables más persistentemente asociadas con la interacción (Fig. 3.6a) y la proporción dentro de las 105 variables ambientales totales (Tabla A.1). Valores mayores a 1 indican que las variables ambientales sintéticas que explicaron la interacción GxS-FS tendieron a incluir variables de ese grupo particular (su presencia es mayor que la esperada por azar). El grupo suelo incluyó todas las variables edáficas analizadas (Tabla 2.5). Los períodos temporales se indican con dos valores de días desde la fecha de siembra, que representan el inicio y el fin de cada período. Valores negativos indican cantidad de días antes de la fecha de siembra; valores positivos, cantidad de días después de la fecha de siembra; valores 0 indican la fecha de siembra.

Figura 3.8: 1-8: Zonas homogéneas de baja interacción GxA (mega-ambientes), para campañas neutras (cuantil 50% de precipitación anual) y fecha de siembra temprana (Fig. 2.5). Cada mapa corresponde con un subconjunto completo. En el extremo superior de cada mapa se indica el rango de GM de los genotipos incluidos ese subconjunto. Los recuadros inferiores derechos, muestran la pendiente de la regresión entre rendimiento y grupo de madurez de cada mega-ambiente, en kg ha^{-1} por unidad de GM (de 0 a 100, al igual que en la Figura 3.2). Valores positivos indican mega-ambientes con rendimientos menores para GM cortos y mayores para GM largos (favorables para GM largos); valores negativos, lo contrario. Entre paréntesis se indica el efecto del hábito de crecimiento determinado $\text{kg ha}^{-1} 10^{-3}$ sobre el rendimiento, cuantificado por los residuales de la regresión por GM, para los mega-ambientes con pendiente positiva o neutra. En gris, se muestran las áreas excluidas del análisis y las áreas no mapeadas, donde el ambiente escapó al rango ambiental explorado por los genotipos en los sitios-fecha de siembra de cada subconjunto. Líneas negras indican límites de las unidades biogeográficas propuestas por Morrone (2014). 9: Unidades biogeográficas propuestas por Morrone (2014) dentro del área de estudio de esta tesis (Tabla 2.2).

Figura 3.9: Integración de las zonas de baja interacción GxA (mega-ambientes) identificadas para los 8 subconjuntos completos, para campañas neutras (cuantil 50% de precipitación anual) y fecha de siembra temprana. Este mapa fue construido a partir de un análisis de clúster que incluyó los 8 mapas de la figura 3.8. Con un mismo color se representan zonas que, en la mayoría de los mapas, han sido mapeadas con un mismo

mega-ambiente o mega-ambientes parecidos. El color y los números indican la pendiente entre rendimiento y GM dentro de cada zona. Sobre los valores de pendiente originales de cada mapa (Fig. 3.8), se aplicó una transformación ranking en dos tramos separados, para los valores de pendiente positivos y negativos. Se muestran los valores promedio (n=8 mapas) de la variable transformada previamente. Valores positivos indican mega-ambientes favorables para GM largos y negativos para GM cortos. En gris, se muestran las áreas excluidas del análisis. Líneas negras indican límites de unidades biogeográficas propuestos por Morrone (2014) (Fig. 3.8, mapa 9; Tabla 2.2).

Figura 3.10: Coeficiente de variación promedio (CV) entre fechas de siembra temprana y tardía (Fig. 2.5) para las variables ambientales sintéticas que explicaron la interacción genotipo x sitio-fecha de siembra (GxS-FS). Se muestra el CV para campañas secas, neutras y húmedas (cuantiles 10, 50 y 90% de precipitación anual). Se incluyeron las variables ambientales sintéticas de los 8 subconjuntos completos. En gris, se muestran las áreas excluidas del análisis y las áreas donde no se calculó el CV porque el ambiente escapó al rango ambiental explorado por los genotipos en la mayoría de los sitios-fecha de siembra. Líneas negras indican límites de unidades biogeográficas propuestos por Morrone (2014) (Fig. 3.8, mapa 9; Tabla 2.2).

Figura 3.11: Coeficiente de variación promedio (CV) entre campañas (2000/01 a 2012/13; n=13) para las variables ambientales sintéticas que explicaron la interacción genotipo x sitio-fecha de siembra (GxS-FS). Se muestra el CV para fechas de siembra temprana y tardía (Fig. 2.5). Se incluyeron las variables ambientales sintéticas de los 8 subconjuntos completos. En gris, se muestran las áreas excluidas del análisis y las áreas donde no se calculó el CV porque el ambiente escapó al rango ambiental explorado por los genotipos en la mayoría de los sitios-fecha de siembra. Líneas negras indican límites de unidades biogeográficas propuestos por Morrone (2014) (Fig. 3.8, mapa 9; Tabla 2.2).

Figura 3.12: Integración de las zonas de baja interacción GxA (mega-ambientes) identificadas para los 8 subconjuntos completos, para seis escenarios que surgen de combinar campañas secas, neutras y húmedas (cuantiles 10, 50 y 90% de precipitación anual) con dos fechas de siembra (Fig. 2.5). Cada mapa fue construido a partir de un análisis de clúster que incluyó los mapas de los 8 subconjuntos completos para un escenario de campaña y fecha de siembra específico. Con un mismo color se representan zonas que en la mayoría de los subconjuntos han sido mapeadas con un mismo mega-ambiente o mega-ambientes parecidos. El color indica la pendiente entre rendimiento y GM dentro de cada zona, en promedio para los 8 subconjuntos. Se aplicó una transformación ranking sobre los valores de pendiente originales. Valores positivos indican mega-ambientes favorables para GM largos y negativos para GM cortos. Valores cercanos a cero indican mega-ambientes de baja respuesta frente al GM. En gris, se muestran las áreas excluidas del análisis. Líneas negras indican límites de unidades biogeográficas propuestos por Morrone (2014) (Fig. 3.8, mapa 9; Tabla 2.2).

ÍNDICE DE TABLAS

Tabla 2.1: Tipos de cubierta excluidos de acuerdo con MODIS Land Cover Type MCD12C1 IGBP (Type I) para el año 2012.

Tabla 2.2: Unidades biogeográficas dentro del área de estudio. A la izquierda se muestran los nombres con que son mencionadas en esta tesis; a la derecha, los nombres originales asignados por dos trabajos clave (Hüeck y Seibert 1972; Morrone 2014).

Tabla 2.3: Genotipos incluidos en el análisis, ordenados por grupo de madurez (GM) y alfabéticamente dentro de cada GM. La columna GM 0-100 muestra el grupo de madurez en la escala de 0 a 100, siendo que 10 unidades de esta escala equivalen a 1 unidad en la escala de 0, I, II... X. La columna Hábito muestra el hábito de crecimiento (I: indeterminado; D: determinado).

Tabla 2.4: Fuentes de datos meteorológicos y de índices de vegetación.

Tabla 2.5: Variables ambientales edáficas.

Tabla 3.1: Descripción de las 21 variables ambientales más asociadas con la interacción GxS-FS (Figura 3.6), ordenadas alfabéticamente. La línea horizontal divide las variables persistentes (arriba) y no persistentes (abajo).

Tabla A.1: Variables ambientales meteorológicas y de índice de vegetación.

ABREVIATURAS

acum: acumulada

ACP: análisis de componentes principales

ANOVA: análisis de la varianza (del inglés “analysis of variance”)

crh: capacidad de retención hídrica

CV: coeficiente de variación

DAS: días antes de la siembra

DDS: días después de la siembra

fp: fotoperíodo

FS: fecha de Siembra

G: genotipo

GM: grupo de madurez

GxA: interacción genotipo x ambiente

GxS-FS: interacción genotipo x sitio-fecha de Siembra

IICA: Instituto Interamericano de Cooperación para la Agricultura

ISRIC-WISE: International Soil Reference and Information Centre-World Inventory of Soil Emission Potentials

IVN: índice de vegetación normalizado

max: máximo

min: mínimo

MANOVA: análisis de la varianza multivariado (del inglés “multivariate analysis of variance”)

MODIS: Moderate Resolution Imaging Spectroradiometer

pef: profundidad efectiva

pp: precipitación

prom: promedio

SIIA: Sistema Integrado de Información Agropecuaria

td: temperatura diurna

tm: temperatura media

tn: temperatura nocturna

TRMM: Tropical Rainfall Measuring Mission

Resumen

La soja se cultiva bajo un amplio rango de condiciones ambientales y con una amplia diversidad de genotipos que responden diferencialmente al ambiente (interacción genotipo x ambiente). La variabilidad del rendimiento entre ambientes y genotipos no está suficientemente estudiada a escala regional y tampoco está claramente mapeada. El objetivo fue desarrollar un modelo espacial y dinámico de la interacción genotipo x ambiente del rendimiento de soja y generar una herramienta para definir y mapear ambientes en Sudamérica (5 a 40°S). Se analizó el rendimiento de 48 genotipos, de grupo de madurez (GM) II a IX, en 96 sitios y tres campañas. Se compiló una base de datos ambientales con 42 millones de registros de fotoperíodo, temperatura, precipitación, índice de vegetación normalizado y variables edáficas. Se identificaron grupos de ambientes con similar ranking de rendimiento de los genotipos (mega-ambientes), y se los asoció con rangos de variables ambientales que, luego, fueron mapeados en escenarios temporales combinando fechas de siembra y campañas. Las variables ambientales que más explicaron la interacción genotipo x ambiente fueron fotoperíodo, precipitación en estadios reproductivos, temperatura en estadios vegetativos y suelos. Los mapas mostraron mega-ambientes favorables para GM largos al Oeste y el Norte y favorables para GM cortos al Sur y el Este. La heterogeneidad espacial aumentó de Norte a Sur. La ubicación geográfica de los mega-ambientes fue más sensible a la fecha de siembra que a la campaña. En fecha tardía, aumentó la heterogeneidad espacial en el Norte y disminuyó en el Sur, y la región se tornó más favorable para GM largos. Al aumentar la precipitación, se amplificó el contraste de GM Norte-Sur. Los mapas de mega-ambientes podrían contribuir a una distribución espacial más eficiente de los sitios experimentales y a extrapolar espacial y temporalmente los resultados de tales experimentos. Esta tesis aporta una metodología original para mapear mega-ambientes.

Palabras clave: mega-ambientes, rendimiento, interacción genotipo x ambiente, mapeo, Soja, Sudamérica

Abstract

Soybean is grown under a wide range of environmental conditions and with a wide diversity of genotypes that differentially respond to the environment (genotype x environment interaction). Yield variability among environments and genotypes has not been successfully studied or mapped at a regional scale. The objective was to develop an spatial and dynamic model of genotype x environment interaction for soybean and set up a tool for defining and mapping environments in South America. The yield of 48 genotypes, maturity groups (MG) II to IX, was analyzed in 96 sites and three seasons. An environmental database was compiled, which included 42 million records of photoperiod, temperature, precipitation, normalized vegetation index and soil variables. Groups of environments with similar genotype yield rankings (mega-environments) were identified and associated with ranges of environmental variables. Mega-environments were then mapped under climatic scenarios that combined planting dates and years. The main environmental variables explaining genotype x environment interaction were photoperiod, precipitation during reproductive phases, temperature during vegetative phases and soils. Maps showed mega-environments in which long MGs win toward the West and the North, and mega-environments in which short MGs win toward the South and the East. Spatial heterogeneity increased from North to South. The location of mega-environments was more sensitive to planting date than to year. With late planting date, spatial heterogeneity increased in the North and decreased in the South, and the whole region became more favorable for long MGs. In humid years, the North-South MG contrast was amplified. Mega-environment maps could contribute to a more efficient spatial distribution of experimental sites and to extrapolate results of such experiments along spatial and temporal dimensions. This thesis provides an original methodology for mapping mega-environments.

Key words: mega-environments, grain yield, genotype x environment interaction, mapping, Soybean, South America

CAPÍTULO 1: Introducción

La soja se cultiva bajo un amplio rango de condiciones ambientales y con una gran diversidad de genotipos que responden diferencialmente según el ambiente (interacción genotipo x ambiente). La variabilidad del rendimiento entre ambientes y genotipos no está suficientemente estudiada a escala regional y tampoco está claramente referenciada con una base cartográfica. Este trabajo pretende generar información sobre la variación del rendimiento de distintos genotipos en respuesta al ambiente a escala regional y a mapear esa variación. De estos dos aspectos, la interacción entre genotipo y ambiente ha sido objeto de estudio en numerosos trabajos mientras que la representación cartográfica (mapeo) de tal interacción se encuentra menos desarrollada. A continuación, se resumen los antecedentes referidos a cada uno de estos dos aspectos.

1.A. Interacción genotipo x ambiente en soja

A lo largo de una amplia variación ambiental, la variación de un carácter fenotípico, como el rendimiento, está determinada por la suma de los efectos debidos al genotipo, al ambiente y a la interacción genotipo x ambiente (Yan et al. 2000). La interacción genotipo x ambiente restringe la posibilidad de extrapolar los resultados de la comparación de genotipos desde un ambiente a otros, lo cual a su vez afecta al diseño espacial de las redes de evaluación de genotipos que montan las empresas o los organismos de investigación (de la Vega y Chapman 2001; Bacigaluppo et al. 2009; Cavassim et al. 2013). Numerosos autores a través de la historia definieron modelos matemáticos para considerar la interacción genotipo x ambiente como un componente de variación al evaluar y seleccionar genotipos (Knight 1970; Gauch 1988; Kang 1998). Una forma habitual de considerarla es definir grupos de ambientes, representados por sitios, años y fechas de siembra, que se parezcan por el modo en que el rendimiento de los genotipos se ordena de mayor a menor. Estos grupos, conocidos como mega-ambientes, son definidos por el comportamiento de los genotipos entre combinaciones sitios-año-fecha de siembra y no por características estrictamente ambientales (de la Vega y Chapman 2001; Dardanelli et al. 2006). En un sentido estricto, la interacción entre genotipo y ambiente representado por sitio, año y/o fecha de siembra podría referirse como interacción genotipo x sitio, genotipo x año o genotipo x fecha de siembra para diferenciarse de la verdadera interacción genotipo x ambiente, que requiere representar a cada ambiente por sus propias características. Si bien en los capítulos siguientes se hará esta distinción, en el presente se empleará el término genotipo x ambiente (GxA) de un modo más general para referirse a todos los tipos de interacción mencionados, como lo hacen otros autores (de la Vega y Chapman 2001; Dardanelli et al. 2006; Hernandez-Segundo et al. 2009; Akbarpour et al. 2014).

La interacción GxA ha sido objeto de estudio en numerosos trabajos, a lo largo de todo el mundo. A continuación, se presentan los principales antecedentes, comenzando por Europa, Norteamérica, África y Asia, y concluyendo con la región de interés para este trabajo: Sudamérica, representada principalmente por Argentina y Brasil.

Para la soja, como para otros cultivos, se ha evaluado la interacción GxA y se ha analizado su importancia para la selección de genotipos por rendimiento y calidad química de granos, con resultados contrastantes (Braun y Rajaram 1996; Gauch y Zobel 1996; de la Vega y Chapman 2001; Yan y Rajcan 2003; Hernandez-Segundo et al. 2009; Miranda et al. 2009; Yan et al. 2011; Badu-Apraku et al. 2012; Carrera et al. 2014; Gedif y Yigzaw 2014; Akbarpour et al. 2014). Se observaron efectos despreciables para la interacción GxA en relación al factor genotípico al evaluar el comportamiento de 14 genotipos de los grupos de madurez (GM) 0 y I en tres sitios del este de Croacia durante cinco campañas (Sudaric et al. 2006) y 14 genotipos del GM II en ocho sitios de Iowa durante tres campañas (Fehr et al. 2003). Resultados semejantes se obtuvieron para 12 genotipos en cuatro sitios de Estados Unidos durante cuatro campañas, para rendimiento y algunos de sus componentes (Rao et al. 2002). Sin embargo, Asfaw et al. (2009) obtuvieron un efecto de interacción GxA seis veces mayor que el efecto genotípico (34,13% y 4,79% sobre la varianza total explicada por la suma de ambiente, genotipo e interacción), al analizar el rendimiento de 11 genotipos de grupo de madurez intermedio evaluados en cuatro sitios de Etiopía durante tres campañas. Primomo et al. (2005) encontraron efectos significativos para la interacción genotipo x año para definir el perfil de ácidos grasos de 14 líneas provenientes de tres cultivares en cuatro sitios del sur de Ontario durante tres campañas a la vez que la interacción genotipo x sitio fue despreciable frente al efecto del año y la interacción genotipo x año. En Argentina, al estudiar el comportamiento de 6 genotipos no transgénicos de GM IV-V en 23 ambientes (24-38°S), la interacción GxA fue un 50% mayor que el efecto genotípico para contenido de aceite (16.2 y 10.8% de la varianza total, respectivamente) y un 30% menor para proteína (22,4 y 31,3%) (Carrera et al. 2014). Arslanoglu y Aytac (2010) describieron respuestas diferenciales de 8 genotipos de los GM I a III, en ocho sitios durante tres campañas, para rendimiento en grano y algunos de sus componentes, como consecuencia de una interacción GxA significativa en la región central del Mar Negro, Turquía. También Yan y Rajcan (2002), al analizar el rendimiento de 125 genotipos en cuatro sitios de Ontario (42 a 45°N) durante seis campañas, detectaron efectos de interacción GxA que variaron desde semejante hasta un 30% mayor que el efecto genotípico en 5 de los 6 años analizados, pero concluyeron que la interacción no era suficiente para dividir la región en mega-ambientes. Estos antecedentes muestran que se ha estudiado el problema en un número limitado de sitios, en rangos ambientales estrechos y en rangos acotados de genotipos, y que se han obtenido resultados diversos en cuanto a la importancia de la interacción GxA.

Entre los antecedentes más relevantes para este trabajo, dentro de Sudamérica, en Argentina se han realizado trabajos de distinta complejidad: algunos evaluaron el efecto del ambiente sobre el rendimiento de distintos genotipos de soja en un mismo sitio en varios años, otros hicieron lo mismo en distintos sitios durante un año, mientras que unos pocos estudiaron la variación entre sitios y entre años conjuntamente. Entre el primer tipo de estudios, en la localidad de Oliveros, Martignone et al. (2008) encontraron efectos importantes de la interacción genotipo x año respecto del factor genotípico (21 y 8% de la varianza total, respectivamente) para 100 genotipos de los GM II a VII en siete campañas, al igual que Enrico et al. (2013) para siete genotipos de GM III a VI en distintas fechas de siembra durante dos campañas (la interacción equiparó al factor genotípico en 26% de la varianza total). En cambio, en Zavalla la interacción genotipo x año para rendimiento no fue significativa para ocho genotipos del GM IV en tres campañas (Rosbaco et al. 2004). Además, Manlla y Salvaghiotti (2012) encontraron que la interacción GxA fue importante en relación a las otras fuentes de variación al evaluar el

rendimiento y sus componentes para tres genotipos de los GM III a V en tres fechas de siembra y dos condiciones hídricas, durante dos campañas.

Dentro del segundo tipo de estudios, trabajos realizados en distintos sitios y dentro de un único año mostraron efectos importantes de la interacción genotipo x sitio para genotipos de los GM III a VI creciendo en un número variable de 10 a 13 localidades en el sur de Santa Fe, y en algunos casos los autores propusieron localidades clave para evaluar la respuesta de genotipos de soja al ambiente (Bacigaluppo et al. 2008; 2009; 2010).

Dentro del tercer tipo de estudios, que incluyen la respuesta de genotipos de soja entre sitios y años, se mostró para el Noroeste Argentino que la alta interacción GxA provocó cambios en los patrones de respuesta al ambiente explorado por 50 genotipos de GM IV a VIII en 17 localidades (22 a 28°S, 63 a 65°O) durante seis campañas (Sánchez et al. 2009). Dos trabajos adicionales definieron mega-ambientes, uno para rendimiento en grano (Luquez et al. 2010) y otro para calidad química (Dardanelli et al. 2006). En el primer caso, Luquez et al. (2010) estudiaron el comportamiento de 19 genotipos del GM IV largo en 23 localidades ubicadas entre los 26 y 39°S y los 58 y 65°O en tres campañas, que representaban 27 ambientes en total. Dependiendo de la metodología de análisis, agruparon los ambientes en tres mega-ambientes (grupos) de 1, 11 y 15 ambientes, o en cuatro mega-ambientes de 4, 4, 9 y 10 ambientes. En el segundo caso, Dardanelli et al. (2006) analizaron 189 genotipos de los GM II a IX, en 24 sitios del área sojera argentina en tres años e identificaron dos o tres mega-ambientes (dependiendo de la campaña) para proteína y para proteína más aceite. En uno de ellos, genotipos de GM VI mostraron valores máximos para estos componentes mientras que en otro los máximos valores fueron alcanzados por GM II-III (Dardanelli et al. 2006).

En Brasil varios trabajos analizaron la interacción GxA para rendimiento y sus componentes en distintas regiones. A continuación, se detallan dos grupos de trabajos por su objetivo de estudio: un grupo orientado a caracterizar genotipos por su capacidad de respuesta o su estabilidad en rendimiento frente a distintos ambientes (Cucolotto et al. 2007; Dias et al. 2009; Palaver y Fagundes 2009; Rocha et al. 2009; Barros et al. 2010; Meotti et al. 2012; Yokomizo et al. 2013; Polizel et al. 2013; Cavalcante et al. 2014) y otro a identificar sitios de importancia (sitios clave, del inglés “key locations”) para la evaluación de genotipos de soja (Rocha y Vello 1999; Carvalho et al. 2002; Pacheco et al. 2009). Dentro del primer grupo, Cucolotto et al. (2007) evaluaron 30 genotipos de GM precoz, semiprecoz y medio en 16 sitios del estado de Paraná durante tres años, encontraron efectos significativos de la interacción GxA e identificaron el genotipo de mayor adaptabilidad y predictibilidad de rendimiento a través de los ambientes, dentro de cada GM. Genotipos con distinta estabilidad del rendimiento a través de los ambientes y distinta capacidad de respuesta a la calidad del ambiente fueron identificados en varios sectores de Brasil, abarcando entre 1 y 4 años, entre 1 y 4 sitios y hasta 38 genotipos (Dias et al. 2009; Palaver y Fagundes 2009; Rocha et al. 2009; Barros et al. 2010; Meotti et al. 2012; Yokomizo et al. 2013; Polizel et al. 2013; Cavalcante et al. 2014).

En el segundo grupo, Carvalho et al. (2002) analizando la respuesta al ambiente de genotipos de los GM precoz, semi-precoz y medio en 13 sitios del estado de Paraná

durante 10 años, también encontraron efectos significativos de la interacción genotipo x sitio sobre el rendimiento en grano e identificaron sitios de elevada y reducida importancia para la evaluación de genotipos de soja de cada GM. Resultados semejantes se obtuvieron para 18 sitios en la región de Cerrado durante tres años con genotipos de los GM corto y medio (Pacheco et al. 2009) y para tres localidades del municipio de Piracicaba durante un año con genotipos de los GM precoz, semi-precoz, medio y semi-tardío (Rocha y Vello 1999).

A pesar de la gran cantidad de trabajos realizados en varias regiones del mundo, todos ellos abarcaron escalas espaciales relativamente pequeñas y no evaluaron las características ambientales que determinan la interacción GxA. En primer lugar, la mayoría de los trabajos se basaron en un número reducido de genotipos pertenecientes a unos pocos GM, o bien analizaron solo un año en particular, y los resultados obtenidos fueron de alcance local. Particularmente en la región de Sudamérica, es escasa la información acerca de la existencia de ambientes que abarquen sitios análogos en los distintos países principales productores y de cómo distintos genotipos pueden aprovechar diferencialmente esos ambientes.

En segundo lugar, estos trabajos han evaluado la interacción GxA como una interacción entre genotipos, sitios, fechas de siembra y/o años sin identificar claramente las características ambientales relevantes de los sitios, fechas de siembra y años que determinan la interacción. Trabajos previos en este sentido abordaron el problema a una escala temporal amplia, con períodos de tiempo calendario (por Ej. temperatura media de Noviembre a Mayo; Dardanelli et al. 2006) o bien para variables ambientales específicas en forma aislada (precipitación o balance hídrico; Bacigaluppo et al. 2008; 2009; 2010) y regiones geográficas acotadas (Martignone et al. 2010). Ya que la respuesta del rendimiento a la disponibilidad de recursos varía entre ventanas de tiempo pequeñas dentro del ciclo del cultivo (Sionit y Kramer 1977; Mathew et al. 2000; Bacigaluppo et al. 2011), un nivel de detalle mayor, combinando las diversas variables ambientales con períodos de tiempo acotados y relacionados con el ciclo del cultivo (días desde la siembra en vez de meses durante el año), permitiría identificar las características ambientales (rangos de variables ambientales) que causan las respuestas diferenciales de los genotipos entre grupo de sitios, fechas de siembra y años, es decir, definir la interacción GxA en términos ambientales. Dada la gran disponibilidad de información ambiental, especialmente de sensores remotos, con cobertura temporal y espacial continua, estos rangos ambientales pueden ser mapeados en el espacio y en el tiempo (por Ej. en años secos y húmedos). Contar con mapas de la interacción GxA (más adelante: mapas de mega-ambientes) permitiría detectar vacíos y redundancias en las redes de evaluación a campo y extrapolar en el espacio y en el tiempo los resultados obtenidos en los experimentos comparativos de rendimiento.

1.B. Mapeo de la variación de rendimiento por genotipos

Trabajos previos han propuesto zonificaciones para cultivos de granos en Sudamérica, generalmente para un país o una región particular dentro de un país. Algunos autores definieron ambientes de importancia para los principales cultivos de granos incluida la soja (Frascarolo 2008; Pizzio y Pacín, 2008) y otros lo hicieron para el cultivo de soja en particular (Pascale y Damario 2004; Kaster y Farias 2012). En el primer caso, Pizzio y Pacín (2008) y Frascarolo (2008), basados en la experiencia personal y sin especificar criterios metodológicos, propusieron zonificaciones para el análisis de resultados y la planificación de cultivos en distintas provincias de Argentina.

En el segundo caso, se han delimitado ambientes de acuerdo a la capacidad del entorno para satisfacer los requerimientos bioclimáticos y edáficos del cultivo de soja en distintas zonas. Por un lado, Pascale y Damario (2004) a partir de información generada en 59 sitios durante siete años de evaluación para 40 genotipos exóticos, definieron y mapearon tipos macroagroclimáticos para la soja en Argentina. Estos tipos macroagroclimáticos representaban rangos de temperatura, disponibilidad hídrica y fotoperíodo con distinto grado de limitación para el cultivo en general. Por otro lado, en Brasil el área de interés para la soja fue dividida en macrorregiones de importancia para la evaluación y selección de genotipos por un modelo basado en la temperatura, el fotoperíodo y el régimen de precipitación, que a su vez fueron subdivididas en regiones edafoclimáticas al contemplar la variación de temperatura debida a la altitud y las condiciones edáficas (Kaster y Farias 2012).

Sin embargo, no se encuentran trabajos que demarquen ambientes para la soja en Sudamérica a una escala mayor que la de país y las zonificaciones existentes no describen la interacción GxA, léase no parten de la respuesta diferencial propia de cada genotipo o grupo de genotipos (Ej. grupos de madurez) y, por tanto, no revelan su distinta capacidad para discriminar ambientes (Carvalho et al. 2002). Finalmente, en escasas oportunidades los trabajos existentes muestran cómo las condiciones climáticas de un año en particular pueden modificar la disponibilidad de recursos en un área determinada, desplazando los límites geográficos de los ambientes explorados por los cultivos (Martignone et al. 2010; Bacigaluppo et al. 2011).

1.C. Objetivos

El objetivo general de esta tesis es desarrollar un modelo espacial y dinámico de la interacción genotipo x ambiente en soja y generar una herramienta para definir y localizar ambientes en Sudamérica. La naturaleza espacial y dinámica del modelo indica que es geográficamente explícito y cambia con el tiempo (campañas y fechas de siembra).

Objetivos específicos

(a) Identificar rangos y combinaciones de variables climáticas, edáficas y de índices de vegetación (características ambientales) que permitan discriminar ambientes con comportamiento homogéneo para los genotipos en el área de estudio

(b) Mapear las áreas geográficas coincidentes con dichas características ambientales, en campañas con diferente régimen de precipitación, desde húmedas hasta secas, combinadas con fechas de siembra tempranas y tardías (mapeo dinámico).

1.D. Estructura general de la tesis

Esta tesis se encuentra organizada en cuatro capítulos. En el capítulo 1 (Introducción; este capítulo) se presentan antecedentes bibliográficos, el problema abordado y los objetivos (general y específicos). En el capítulo 2 (Materiales y métodos) se presentan, (a) el esquema metodológico general de la tesis y una caracterización ambiental del área de estudio y la metodología aplicada para (b) analizar la interacción genotipo x sitio-fecha de siembra y detectar los mega-ambientes, (c) caracterizar ambientalmente los mega-ambientes a partir de la respuesta de los genotipos a la variación ambiental y (d) mapear los mega-ambientes en distintos escenarios temporales. El capítulo 3 presenta los resultados, el 4 la discusión y el 5 las conclusiones generales.

CAPÍTULO 2: Materiales y métodos

2.A. Aproximación general y área de estudio

A partir de información de rendimiento de genotipos de soja en distintos sitios, años y fechas de siembra se identificaron grupos de sitios-fecha de siembra (fecha de siembra incluía el año y el día del año) con similar ranking de rendimiento de los genotipos (Fig. 2.1a) y, por lo tanto, con mínima interacción genotipo x sitio-fecha de siembra (GxS-FS). En adelante, a estos grupos se los denominará mega-ambientes (Yan et al. 2000; de la Vega y Chapman 2001). Luego, sobre la base de información meteorológica, de índice de vegetación y edáfica (Fig. 2.1b), se identificaron características ambientales propias de los mega-ambientes detectados en el paso anterior. De esta manera, los mega-ambientes inicialmente identificados por los rankings de genotipos en función del rendimiento, quedaron definidos por rangos ambientales y se les asignó la categoría de “verdaderos ambientes” (Fig. 2.1c). Esto permitió considerar a la interacción genotipo x sitio-fecha de siembra como interacción genotipo x ambiente (GxA). Finalmente, a partir de la información ambiental de toda el área de estudio, se generaron mapas de distribución de los mega-ambientes en distintos escenarios temporales definidos por campañas y fechas de siembra (Fig. 2.1d).

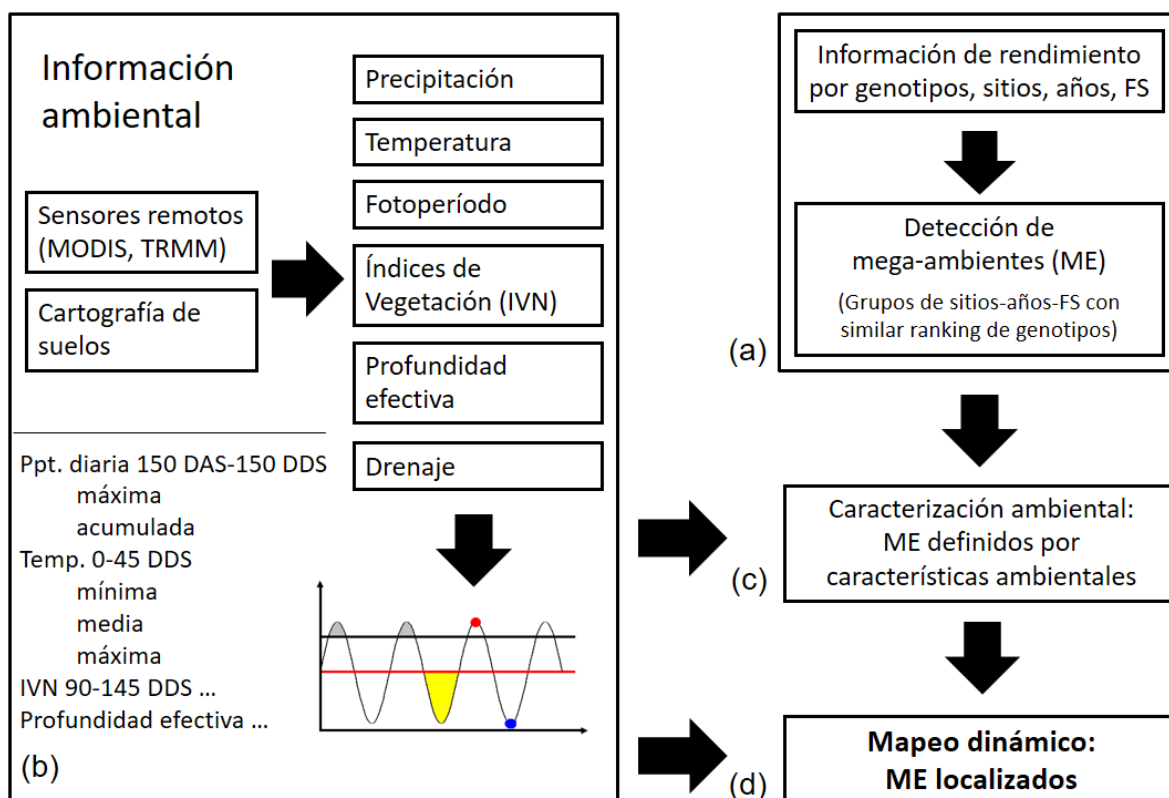


Figura 2.1: Esquema metodológico general. A partir de información de rendimiento de genotipos en distintos sitios, años y fechas de siembra se identificaron grupos de sitios-años-fecha de siembra con similar ranking de rendimiento de los genotipos (mega-ambientes) (a). Sobre la base de información meteorológica, de índice de vegetación y edáfica (b), los mega-ambientes fueron caracterizados por rangos ambientales y se les asignó la categoría de “verdaderos ambientes” (c). A partir de la información ambiental en toda el área de estudio, los mega-ambientes fueron mapeados en distintos escenarios de campaña y fechas de siembra (d).

“Ppt.”: precipitación. “Temp.”: temperatura. “DAS”: días antes de la siembra. “DDS”: días desde la siembra. “FS”: fecha de siembra. “ME”: mega-ambiente.

I. Área de estudio

El área de estudio comprendió las regiones sojeras de los principales países productores de Sudamérica: Brasil, Argentina, Paraguay, Uruguay y Bolivia, y abarcó un rango de latitudes entre 2 y 39°S y un rango de longitudes entre 44 y 66°O. Por su biogeografía, esta área corresponde con la totalidad o parte de 6 unidades de norte a sur: Provincia de Rondonia, Provincia de Cerrado, Bosque Paranense, Bosque de Araucaria, Provincia Chaqueña y Provincia Pampeana (Morrone 2014) (Fig. 2.2; Tabla 2.2).

Se excluyeron las áreas con muy baja proporción de cultivos de soja (Fig. 2.2), según el conocimiento de expertos y estadísticas agropecuarias (SIIA, 2014). En toda el área de estudio se excluyeron áreas sin cultivos agrícolas según el producto Land Cover Type

MCD12C1 de MODIS para el año 2012 (Tabla 2.1). Además, para Argentina se excluyeron las provincias sin ningún sitio experimental de la base de datos y/o con menos de 100.000 ha de soja en promedio para las campañas 2009/10-2013/14 (SIIA, 2014). Para Brasil se excluyeron los estados sin ningún sitio experimental y/o las zonas no mapeadas por Kaster y Farias (2012). Finalmente, para Paraguay se excluyeron los departamentos con menos de 100.000 ha de soja en la campaña 2012/13 (IICA, 2014). Por falta de información local, para Uruguay y Bolivia no se excluyeron otras áreas más que las definidas por el producto Land Cover Type MCD12C1. En el caso de Uruguay esto resultó en incluir la totalidad de su territorio.

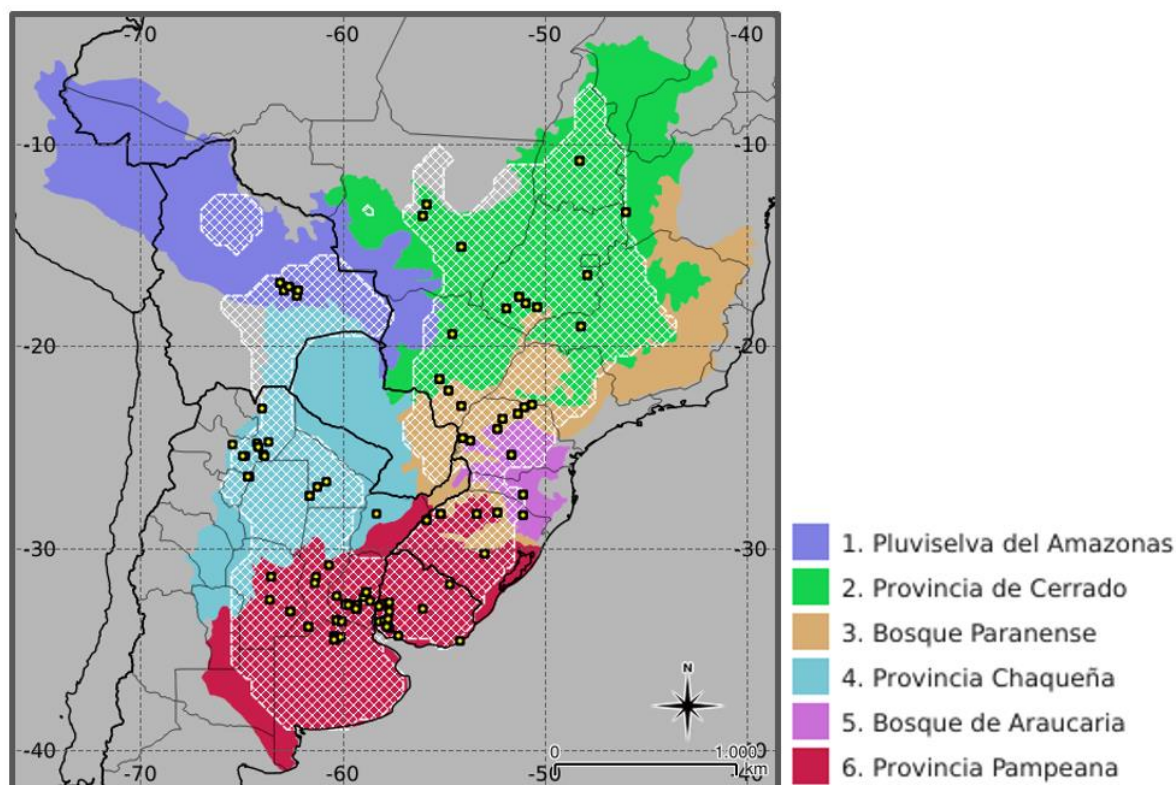


Figura 2.2: Área de estudio (cuadrícula blanca) y unidades biogeográficas pertinentes (en colores) (Morrone 2014). Líneas negras indican límites internacionales, provinciales (Argentina) y de estados (Brasil). Marcas amarillas indican sitios experimentales (n=96).

Tabla 2.1

Tipos de cubierta excluidos de acuerdo con MODIS Land Cover Type MCD12C1 IGBP (Type I) para el año 2012.

Clase	Descripción
0	Agua
2	Bosque latifoliado siempre-verde
7	Arbustales abiertos
11	Humedales permanentes
13	Urbano
16	Desiertos y semidesiertos

Tabla 2.2

Unidades biogeográficas dentro del área de estudio. A la izquierda se muestran los nombres con que son mencionadas en esta tesis; a la derecha, los nombres originales asignados por dos trabajos clave (Hüeck y Seibert 1972; Morrone 2014).

Unidad biogeográfica	Referencias bibliográficas
Bosque de Araucaria	Provincia del Bosque de Araucaria (1) Bosque de Araucaria angustifolia (2) Campos Limpios, praderas densas de las alturas meridionales de Brasil (2)
Bosque de las Sierras de Chiquitos	Bosque mesofítico deciduo, en parte siempreverde, de las sierras de Chiquitos (2)
Bosque Paranaense	Provincia del Bosque de Parana (1) Bosques subtropicales deciduos y mesofíticos del Brasil oriental y meridional, en parte con alta proporción de especies siempreverdes (2)
Chaco Central	Bosque seco del Chaco central (2)
Chaco Occidental	Bosques mesofíticos de transición del Chaco marginal occidental (2)
Chaco Oriental	Bosque desde seco hasta semihúmedo del Chaco Oriental (2)
Espinal	Bosques secos de los Espinales (2) Transición del bosque seco de los Espinales a la vegetación del Monte (2)
Pantanal	Vegetación del Pantanal (2)
Pluviselva del Amazonas	Provincia de Rondonia (1) Pluviselva tropical siempreverde del Amazonas y Orinoco (2)
Provincia Chaqueña	Provincia Chaqueña (1)
Provincia de Cerrado	Provincia de Cerrado (1) Campos Cerrados, Chaparrales y sabanas semejantes (2)
Provincia Pampeana	Provincia Pampeana (1)
Sabana de Palmeras de Bolivia	Sabana de palmeras de Santa Cruz y Trinidad (Bolivia) (2)

1: Morrone (2014)

2: Hüeck y Seibert (1972)

II. Caracterización del área de estudio en función de la precipitación anual

Se caracterizó el área de estudio por la variación espacial y temporal de la precipitación anual. Se emplearon datos de precipitación diaria estimados a partir de imágenes satelitales derivadas de la misión TRMM (Su et al. 2008). Se calcularon las anomalías de precipitación anual para seis campañas (2000/01, 2001/02, 2007/08, 2008/09, 2009/10 y 2012/13), como la diferencia entre la precipitación de cada campaña individual y la precipitación promedio de las campañas 2000/01 a 2012/13 ($n=13$), expresada en forma relativa a la precipitación promedio. Se calculó el coeficiente de variación inter-anual de la precipitación para las campañas 2000/01 a 2012/13 ($n=13$). El período de tiempo abarcado por una campaña fue desde junio de un año hasta mayo del siguiente (junio-mayo). La anomalía de la precipitación anual varió entre áreas geográficas y campañas (Fig. 2.3). La variación entre áreas fue muy marcada en 2000/01, 2001/02 y 2012/13 y débil en 2008/09 (anomalías predominantemente negativas) y 2009/10 (positivas). En relación con la variación entre campañas, la fluctuación relativa de la precipitación anual fue menor hacia el Norte, sobre Pluviselva del Amazonas y Sabana de Palmeras de Bolivia, Bosque de las Sierras de Chiquitos, provincia de Cerrado, Bosque Paranaense y Bosque de Araucaria, y mayor hacia el Sur, en las provincias Chaqueña y Pampeana (Fig. 2.4, Tabla 2.2).

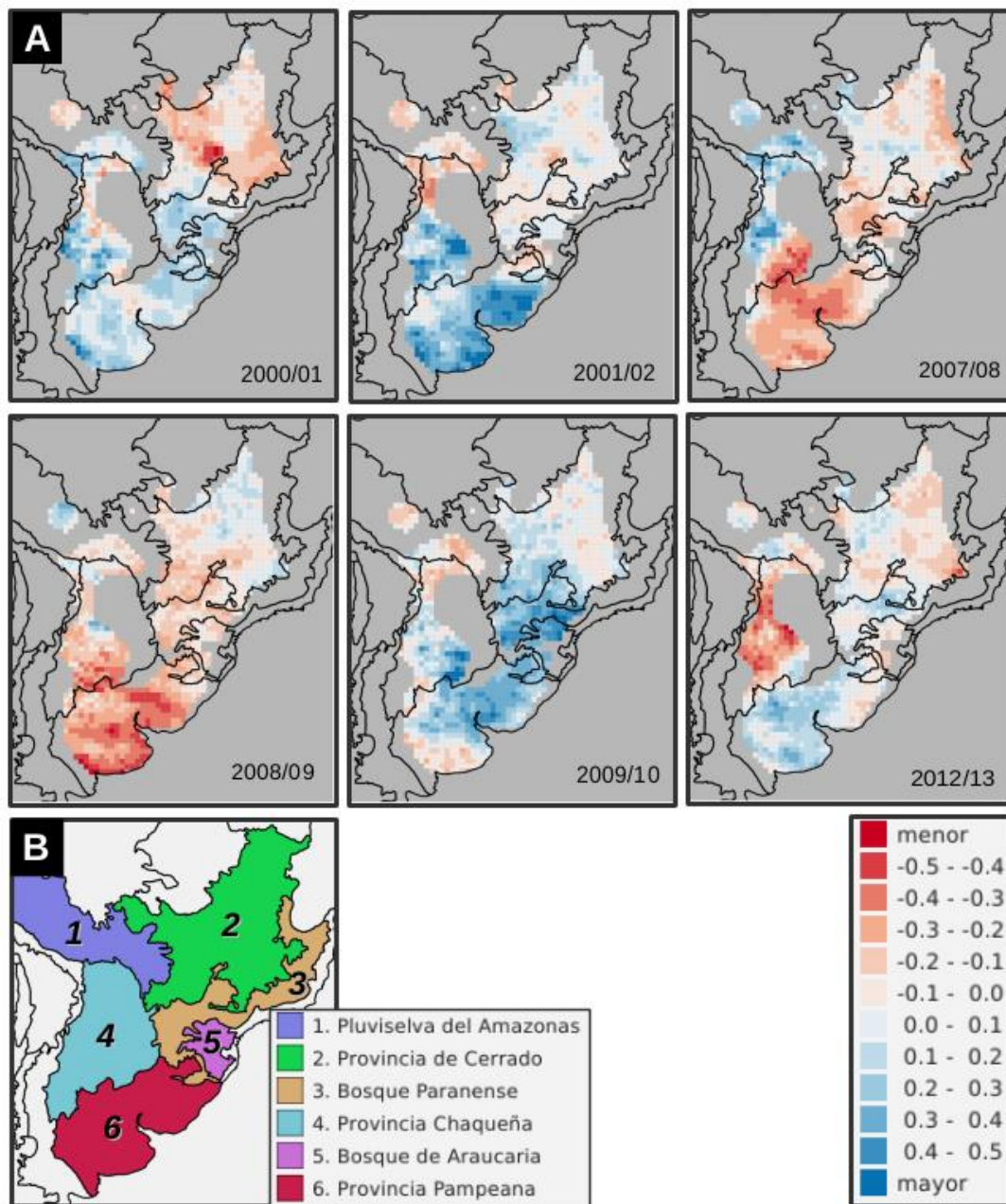


Figura 2.3: A: Anomalía de la precipitación anual para seis campañas, calculada como la diferencia entre la precipitación de cada campaña y la precipitación promedio de las campañas 2000/01 a 2012/13 ($n=13$), expresada en forma relativa a la precipitación promedio. En gris, se muestran las áreas excluidas del análisis. Para cada campaña, se consideró la precipitación acumulada entre junio del año inicial y mayo del siguiente. Líneas negras indican límites de las unidades biogeográficas propuestas por Morrone (2014). B: Unidades biogeográficas propuestas por Morrone (2014) dentro del área de estudio de esta tesis (Tabla 2.2).

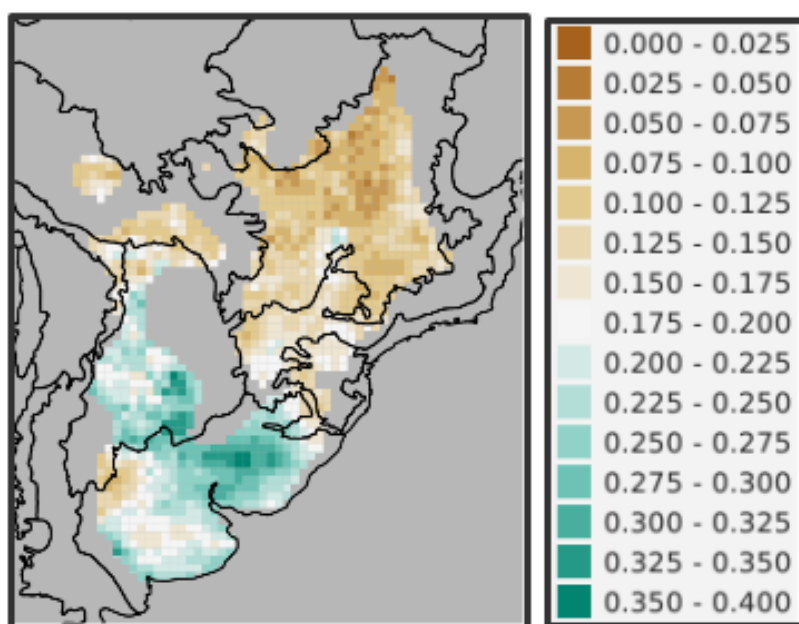


Figura 2.4: Coeficiente de variación inter-anual de la precipitación anual de las campañas 2000/01 a 2012/13 (n=13). En gris, se muestran las áreas excluidas del análisis. Líneas negras indican límites de unidades biogeográficas propuestos por Morrone (2014) (Fig. 2.3 B; Tabla 2.2).

2.B. Interacción genotipo x sitio-fecha de siembra y mega-ambientes

I. Información de rendimiento de genotipos de soja

Se emplearon datos de rendimiento de 48 genotipos de soja de los grupos de madurez II largo a IX corto, de hábito de crecimiento indeterminado y determinado (Fig. 2.1a; Tabla 2.3), generados por la red de evaluación de cultivares de GDM Seeds (Grupo Don Mario). Los datos provinieron de 133 experimentos conducidos durante las campañas 2010/11 (6% de los datos), 2011/12 (12 %) y 2012/13 (82%), en 96 sitios experimentales ubicados en ambientes de buena calidad y bajo condiciones de manejo nutricional, de plagas y enfermedades recomendadas de acuerdo a la zona. Estos datos que conformaron la base de datos final analizada, fueron previamente seleccionados dentro una base de datos de mayor tamaño de acuerdo con dos criterios, asociados a la fecha de siembra y a la completud de los datos, que se detallan en los párrafos siguientes. La base de datos final contó con 7735 datos de rendimiento correspondientes con sendas combinaciones de genotipos y sitios-fechas de siembra. El rango de rendimientos explorado fue 0-6130 kg ha⁻¹, con los percentiles 5 y 95% de 950 a 4400 kg ha⁻¹, respectivamente. Dada la escala regional adoptada, el interés radicó exclusivamente en estudiar patrones a nivel de rendimiento total, más allá de sus componentes individuales (número y peso de granos).

Tabla 2.3

Genotipos incluidos en el análisis, ordenados por grupo de madurez (GM) y alfabéticamente dentro de cada GM. La columna GM 0-100 muestra el grupo de madurez en la escala de 0 a 100, siendo que 10 unidades de esta escala equivalen a 1 unidad en la escala de 0, I, II... X. La columna Hábito muestra el hábito de crecimiento (I: indeterminado; D: determinado).

GM	Nombre	GM 0-100	Hábito
II	DM2200	25	I
III	DM0314	38	I
	DM3070	33	I
	DM3312	35	I
	DM3700	39	I
IV	ADM4800	48	I
	DM3810	40	I
	DM4210	41	I
	DM4212	41	I
	DM4612	48	I
	DM4670	46	I
	DM4712	48	I
	FN4.50	45	I
	NA4009	40	I
	NA4613	49	I
	NA4725	47	I
	NA4990	49	I
	SPS4X4	45	I
V	BMX_TITAN	56	I
	DM4970	50	I
	DM5.1i	50	I
	DM5.5i	56	I
	DM5.8i	58	I
	DM5.9i	59	I
	DM6.2i	59	I
	NA4910	51	I
	NA5009	52	I

Tabla 2.3 (segunda parte)

GM	Nombre	GM 0-100	Hábito
VI	BMX_FORCA	64	I
	DM6.8i	68	I
	NA5909	60	I
	NA6126	61	D
	NA6411	64	D
	NK7059RR	64	I
	SYN9070RR	68	I
VII	BMX_POTENCIA	70	I
	DM7.0i	70	I
	DM7.8i	78	I
	DM7869RSF	78	I
	DM7874RSF	72	I
	MSOY7211	72	I
VIII	DM1092	80	I
	DM8002	80	D
	DM8473RSF	86	I
	DM8576RSF	87	I
	NA8000	80	D
	NA8009	80	I
	TOB7800	80	I
IX	MUNASQA	90	D

Cada experimento constó de un diseño en bloques, con tres (94% de los datos) o dos repeticiones por genotipo, con parcelas de 200x3,5 m o 5x2 m. Cada experimento representó un sitio-fecha de siembra. El 97% de los datos de rendimiento contó con coordenadas geográficas del campo donde se ubicaron los experimentos, mientras que el 3% restante contó con coordenadas de una localidad cercana. El rendimiento de cada genotipo en cada combinación sitio-fecha de siembra fue calculado como el promedio de las repeticiones. Fueron descartados previamente aquellos datos con fecha de siembra tardía extrema respecto de experimentos vecinos en una ventana de 3x3° geográficos alrededor de cada experimento. Se consideró que la fecha era tardía extrema cuando superó el cuartil 75% de la distribución en más de una vez y media del rango intercuartil 25-75%.

La interacción GxA fue abordada a partir del análisis de mega-ambientes (Yan et al. 2000; de la Vega y Chapman 2001). Identificar mega-ambientes requiere manejar subconjuntos completos de datos de rendimiento, es decir donde los genotipos hayan sido evaluados en todos los sitios-fechas de siembra. Se identificaron siete subconjuntos completos con al menos 37 sitios-fecha de siembra (filas) y 6 genotipos (columnas) mediante una rutina automática de búsqueda por rotación de filas y columnas especialmente diseñada para esta tesis (ver apéndice II). Como esta rutina automática identifica solo una muestra de todos los subconjuntos completos posibles, se identificó un subconjunto adicional por búsqueda manual intentando incluir la máxima cantidad posible de genotipos y sitios-fecha de siembra. Todos los subconjuntos (8 en total) contaron con al menos un 80% de presencias dentro de cada fila y columna, lo que resultó en al menos un 97% de presencias a nivel global. El 3% de datos faltantes fue rellenado con los promedios de genotipos (columnas) correspondientes. Estos 8 subconjuntos completos, que representaron la base de datos final analizada, difirieron en más de un 50 % de sus datos, a excepción de dos de ellos que difirieron en un 37% entre sí. Cada subconjunto completo fue tratado como una unidad de análisis independiente para la cual se analizó la interacción genotipo x sitio-fecha de siembra y se identificaron mega-ambientes.

II. Análisis de la interacción y detección de mega-ambientes

Para cada subconjunto completo, se identificaron grupos de sitios-fechas de siembra (fecha de siembra incluía el año) con similar ranking de rendimiento de los genotipos (mega-ambientes; Fig. 2.1a). Dentro de cada subconjunto, se ejecutó un análisis de clúster jerárquico que agrupó los sitios-fechas de siembra por el rendimiento de los genotipos centrado por ambiente (Yan et al. 2000). La matriz tenía a los sitios-fechas de siembra como casos a agrupar (filas) y a los genotipos como variables (columnas). Cada celda representó el rendimiento centrado por ambiente (considerando ambiente a cada combinación sitio-fecha de siembra) de un genotipo en un sitio-fecha de siembra. El rango entre percentiles 5 y 95% para el rendimiento centrado por ambiente fue de más de 1200 kg ha⁻¹ en todos los subconjuntos (entre 1287 y 2066 kg ha⁻¹, con un promedio de 1653; n=8), lo cual significa que hubo más de 1200 kg ha⁻¹ de variación atribuible a los genotipos y a su interacción. Se empleó el método de clúster jerárquico de Ward (1963; implementado en Python, módulo “hierarchy” del paquete “scipy.cluster”; Jones et al. 2001), con distancia euclídea como medida de similitud/disimilitud (de la Vega y Chapman 2001), con un número de grupos (mega-ambientes) tal que el porcentaje de información retenida fuera cercano al 55%. Todos los subconjuntos resultaron en al menos 4 mega-ambientes con 5 o más sitios-fecha de siembra cada uno.

Por definición, dentro de los mega-ambientes la interacción genotipo x sitio-fecha de siembra es menor que dentro del conjunto original de datos (de la Vega y Chapman 2001; Dardanelli et al. 2006). Esto fue verificado por medio de ANOVA, considerando el modelo

$$Y_{ij} = G_i + A_j + (G \times A)_{ij} + e_{ij}$$

que descompone el rendimiento de cada genotipo i en cada sitio-fecha de siembra j Y_{ij} en la suma de los efectos principales de genotipo G_i y sitio-fecha de siembra A_j , más la interacción genotipo x sitio-fecha de siembra $(G \times A)_{ij}$ y un residual e_{ij} (Dardanelli et al. 2006). Para cada subconjunto, se realizó un ANOVA para el total de los datos y ANOVAs para cada uno de los mega-ambientes. Luego, se estimó el promedio de la variación debida a cada factor (G, A y $G \times A$) antes de identificar mega-ambientes y se lo comparó con la variación aportada por los factores dentro de los mega-ambientes ($n=8$ subconjuntos). La variación debida a cada factor dentro de los mega-ambientes fue calculada en dos pasos: primero se calculó el promedio dentro de cada subconjunto considerando todos los mega-ambientes que generó ese subconjunto ($n=\text{número de mega-ambientes del subconjunto}$), y luego se promedió entre subconjuntos ($n=8$). Se emplearon las rutinas estadísticas “lm” y “anova” del paquete “stats” de R (R Core Team 2013).

Se analizó en qué medida la interacción genotipo x sitio-fecha de siembra se asoció con el grupo de madurez (GM) y el hábito de crecimiento a través de dos enfoques. El primero consistió en evaluar la asociación entre el rendimiento de los genotipos y el GM dentro de cada mega-ambiente. Para cada mega-ambiente se ajustó una regresión lineal en la que la variable dependiente fue el rendimiento promedio centrado por ambiente y la independiente el GM (se consideró una escala de 0 a 100, siendo que 10 unidades de esta escala equivalen a 1 unidad en la escala de 0, I, II... X). Una asociación fuerte y positiva, significaría que, en el mega-ambiente en cuestión, los genotipos de GM cortos tuvieron rendimientos menores que los de GM largos, y viceversa. Una asociación débil, significaría que en ese mega-ambiente el ranking de rendimiento de los genotipos no estuvo ligado a los GM. El efecto del hábito de crecimiento fue estimado a partir de los residuos no explicados por el GM. Primero, dentro de cada mega-ambiente, los efectos del hábito determinado e indeterminado fueron estimados con el promedio de residuales de los genotipos determinados e indeterminados, respectivamente (n variable entre subconjuntos), obteniéndose un valor para cada hábito en cada mega-ambiente. Un valor negativo para hábito determinado, indica que el mega-ambiente es desfavorable para genotipos determinados, y viceversa. Luego, para cada hábito se promediaron los valores de todos los mega-ambientes ($n=34$ mega-ambientes), y discriminando entre mega-ambientes con pendiente significativa ($n=23$) y no significativa para la regresión por GM ($n=11$). Si los promedios discriminando por significancia difieren entre sí y respecto del promedio de todos los mega-ambientes, significa que el efecto del hábito sobre la interacción depende del efecto del GM.

El segundo enfoque consistió en cuantificar el efecto de la interacción genotipo x sitio-fecha de siembra sobre el rendimiento para todos los pares posibles de genotipos y asociarlo con los GM. Dentro de cada subconjunto completo, se calcularon las distancias euclídeas entre genotipos en el hiperplano de los sitios-fecha de siembra para todos los pares de genotipos posibles. Estas distancias fueron estandarizadas por la media y el desvío estándar dentro de cada subconjunto. Luego, se calculó la distancia euclídea estandarizada promedio para cada par de genotipos, como el promedio de las distancias euclídeas estandarizadas de dicho par en todos los subconjuntos en que se encontró. Una distancia promedio grande, indicaría que un par de genotipos generó una interacción genotipo x sitio-fecha de siembra también grande. Para el cálculo de las distancias, se consideraron rendimientos centrados por ambientes. Se empleó el algoritmo “pdist” del módulo “scipy.spatial.distance” de Python (Jones et al. 2001).

2.C. Caracterización ambiental de los mega-ambientes

I. Información ambiental

Las variables ambientales consideradas para analizar las respuestas de los genotipos al ambiente fueron: fotoperíodo, precipitación, temperatura, índice de vegetación normalizado (IVN) y una serie de variables edáficas: contenido de arcilla y arena, conductividad eléctrica, densidad aparente, pH y sodio intercambiable, todas ellas en los estratos 0-30 y 30-100 cm, y capacidad de retención hídrica, clase de drenaje y profundidad efectiva, para el perfil completo (Fig. 2.1b). Los datos originales provinieron de distintas fuentes públicas, y tuvieron diversas resoluciones espacio-temporales (Tablas 2.4 y 2.5). La precipitación, la temperatura y el IVN provinieron de imágenes generadas a partir de sensores remotos, para el período 2000 a 2013 (14 años). Se calculó la temperatura media diaria como el promedio de las temperaturas diarias diurna y nocturna. El fotoperíodo fue calculado en función de la latitud y el día del año. Las variables edáficas fueron obtenidas de una compilación de datos de campo almacenados en la base de datos ISRIC-WISE (Batjes 2005; Tabla 2.5). Para las variables temperatura diurna y nocturna y para IVN, se rellenaron los datos faltantes en cada imagen por interpolación lineal, dentro de cada píxel, entre las imágenes con datos válidos más próximas en el tiempo (anterior y posterior).

Las variables ambientales fueron utilizadas de dos maneras. Por un lado, se utilizaron las correspondientes a los sitios experimentales para generar las relaciones entre los grupos de sitios-fecha de siembra (mega-ambientes) y el ambiente propiamente dicho. Por otro lado, fueron volcadas a celdas de una grilla de 0.5x0.5 grados de latitud y longitud para generar los mapas que se describen más adelante.

Previamente, las variables ambientales originales fueron resumidas temporal y espacialmente. Para fotoperíodo, IVN, precipitación y temperatura, se calcularon medidas de resumen temporal (máximo, mínimo, promedio, desvío estándar y suma acumulada) en diferentes períodos del ciclo de cultivo, definidos en función de la fecha de siembra: 150 días antes (DAS) a 150 días después de la siembra (DDS), 30 DAS a 145 DDS, 0 a 145 DDS, 0 a 55 DDS, 55 a 145 DDS, 90 a 145 DDS, dependiendo de la variable considerada (ver apéndice, Tabla A.1). Para los sitios experimentales se emplearon los registros reales de fecha de siembra (fecha de siembra conocida), mientras que para las celdas 0.5x0.5° en la totalidad del área de estudio se consideraron fechas de siembra tempranas y tardías hipotéticas, estimadas por referentes para una grilla de 4x4° geográficos de resolución (Fig. 2.5). Como el interés de este trabajo no era evaluar la fenología, se definieron períodos de tiempo fijos, expresados en días desde la fecha de siembra. A modo orientativo, se mencionan las fases de un cultivo de soja promedio que podrían coincidir con los períodos de tiempo elegidos para este trabajo (Tabla A.1).

Las variables edáficas (que carecieron de dimensión temporal) fueron integradas espacialmente hasta obtener un valor por celda para el perfil completo y dos estratos (0-30 y 30-100 cm; Tabla 2.5). La integración a nivel de celda se logró promediando los

valores de todos los suelos presentes en cada celda, ponderando por la proporción de superficie ocupada por cada suelo.

Como consecuencia de todos estos procedimientos, se obtuvieron en total 105 variables ambientales, considerando las combinaciones de las variables originales (meteorológicas, índices de vegetación y edáficas) con los períodos de tiempo definidos en días desde la siembra. En adelante, se empleará el término “variable ambiental” para referirse a estas 105 variables que resultan de haber resumido las variables originales al grano temporal y espacial correspondiente.

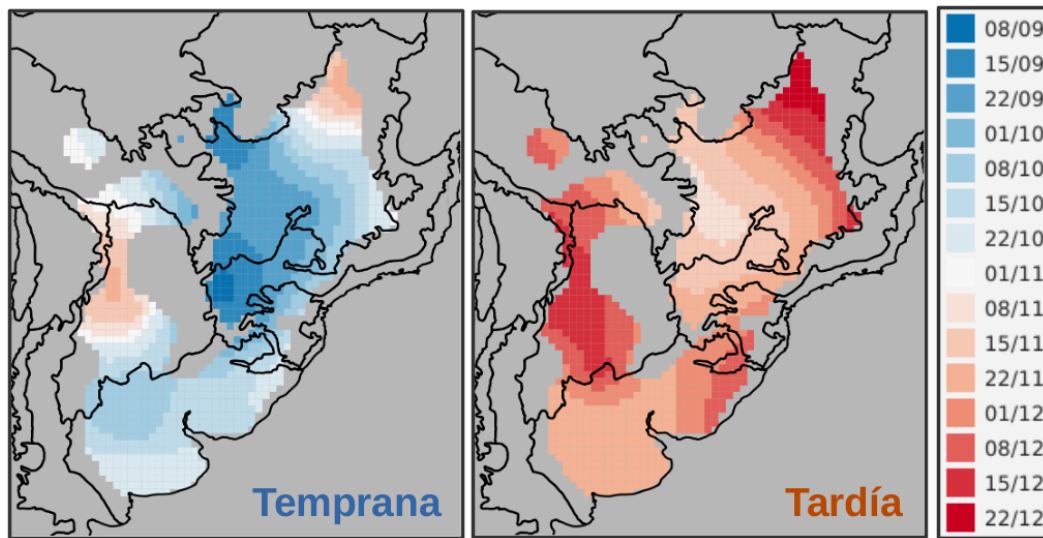


Figura 2.5: Fecha de siembra temprana (izquierda) y tardía (derecha) estimada por referentes. El color indica la semana de siembra. La fecha de siembra fue consultada a referentes para una grilla de $4 \times 4^\circ$ que cubría el área de estudio en su totalidad, y fue volcada a la grilla de $0.5 \times 0.5^\circ$ por interpolación espacial con el método de kriging. Las fechas de siembra consultadas representan una fecha temprana óptima y una fecha tardía frecuente, no extrema. Líneas negras indican límites de unidades biogeográficas propuestos por Morrone (2014) (Fig. 3.8, mapa 9; Tabla 2.2)

Tabla 2.4

Fuentes de datos meteorológicos y de índices de vegetación.

Variable	Fuente	Producto	Resolución espacial	Resolución temporal (días entre escenas)	Referencia
Fotoperíodo	Calculado	-	0.5x0.5°	1	Keisling, 1982
IVN	MODIS	13 C1	5x5 km	16	Salomonson et al. 1989
Precipitación	TRMM	3B42 v7	0.25x0.25°	1	Su et al. 2008
Temp. diaria diurna	MODIS	11 C2	5x5 km	8	Salomonson et al. 1989
Temp. media diaria	Calculado		5x5 km	8	-
Temp. diaria nocturna	MODIS	11 C2	5x5 km	8	Salomonson et al. 1989

Tabla 2.5

Variables ambientales edáficas.

Variable	Horizonte de suelo
Contenido de Arcilla (%)	0-30 cm
Contenido de Arena (%)	0-30 cm
Conductividad eléctrica	0-30 cm
Densidad aparente	0-30 cm
pH	0-30 cm
Sodio intercambiable	0-30 cm
Contenido de Arcilla (%)	30-100 cm
Contenido de Arena (%)	30-100 cm
Conductividad eléctrica	30-100 cm
Densidad aparente	30-100 cm
pH	30-100 cm
Sodio intercambiable	30-100 cm
Capacidad de retención hídrica	0-100 cm
Clase de drenaje (FAO)	0-100 cm
Profundidad efectiva	0-100 cm

Para cada variable se detalla el horizonte de suelo y la descripción asignada por Batjes (2005). La resolución espacial es de 0.5x0.5° para todas las variables.

* No considera impedancias químicas, como podrían ser concentraciones elevadas de aluminio.

II. Variables ambientales asociadas a los mega-ambientes

Los mega-ambientes, inicialmente identificados por el ranking de rendimiento de los genotipos en distintos sitios-fechas de siembra, fueron caracterizados por verdaderas variables ambientales (Fig. 2.1c). Para ello, las variables ambientales anteriores fueron combinadas en variables sintéticas (variables que surgen como combinación lineal de otras variables) que explicaron la interacción genotipo x sitio-fecha de siembra. Esto se logró mediante dos pasos: (1) la generación de modelos discriminantes canónicos partiendo de distintos grupos de variables ambientales tomadas al azar y (2) la selección de los modelos que mejor explicaron las diferencias de ranking de genotipos entre mega-ambientes, dentro de cada subconjunto completo. Los ejes canónicos de estos modelos discriminantes son combinaciones lineales de las variables ambientales (es decir: variables ambientales sintéticas) que explican la interacción GxS-FS y se pueden mapear. En los dos párrafos que siguen se detallan los procedimientos empleados para generar estos modelos y seleccionar los mejores de ellos.

Para cada subconjunto completo, se generaron 5000 modelos discriminantes canónicos para los mega-ambientes definidos previamente. Se generaron modelos discriminantes para combinaciones distintas de variables ambientales, tomadas al azar, hasta obtener un número de 5000 modelos de MANOVA significativos ($p < 0.10$). La matriz a analizar en cada caso estuvo compuesta por los sitios-fecha de siembra en las filas, las variables ambientales en las columnas, y una columna categórica adicional con el identificador del mega-ambiente al que pertenecía cada sitio-fecha de siembra. La cantidad de variables a incluir en cada modelo se fijó en un tercio de la cantidad de sitios-fecha de siembra presentes en cada subconjunto, excepto cuando esa fracción fue mayor a 15, en cuyos casos se empleó la cantidad máxima de 15 variables. Ya que el análisis discriminante puede estar influenciado por grupos pequeños, se eliminaron previamente los mega-ambientes con menos de 5 sitios-fecha de siembra. Se emplearon las rutinas estadísticas “lm” del paquete “stats”, y “Anova” y “candisc” del paquete “candisc” de R (R Core Team 2013).

Para cada subconjunto completo, se seleccionaron los 50 modelos discriminantes que mejor explicaron las diferencias de ranking de rendimiento de los genotipos entre mega-ambientes. Para cada subconjunto, se ejecutó un análisis de componentes principales (ACP) sobre la misma matriz empleada por el análisis de clúster para generar los mega-ambientes, con el rendimiento de los genotipos (columnas) en los sitios-fechas de siembra (filas; se empleó el algoritmo “ACP” del módulo “sklearn.decomposition” de Python; Pedregosa et al. 2011). Luego, para cada modelo discriminante entre los 5000 generados, se calculó un índice de explicación integrado por las proporciones de variabilidad explicada (autovalores) de los dos primeros componentes principales del ACP y de los dos primeros ejes canónicos del modelo discriminante, y por las correlaciones entre el componente principal 1 y el eje canónico 1, y entre el componente principal 2 y el eje canónico 2 para los sitios-fechas de siembra. Valores elevados de este índice implican que las variables ambientales sintéticas explicaron buena parte de las diferencias de ranking de los genotipos entre mega-ambientes. Para cada subconjunto completo, se seleccionaron los 50 modelos con mayor valor para este índice. Cada modelo fue representado por los dos primeros ejes canónicos (variables ambientales sintéticas), que

acumularon al menos un 75 % de la variabilidad entre mega-ambientes. En total, se obtuvieron las 800 variables ambientales sintéticas que mejor explicaron la interacción GxS-FS, representadas por los dos primeros ejes canónicos de 50 modelos discriminantes para cada uno de los 8 subconjuntos completos.

Se cuantificó la importancia de las distintas variables ambientales para explicar la interacción GxS-FS, a través de su grado de asociación con las variables ambientales sintéticas seleccionadas en el paso anterior. Además, se identificaron variables ambientales que fueron importantes para la mayoría (persistentes) o solo para algunos subconjuntos completos (no persistentes). Para ello, se consideraron los coeficientes de peso de las variables ambientales en los ejes canónicos (autovectores) de los 50 modelos seleccionados para cada subconjunto, y las proporciones de variabilidad explicada (autovalores) por los dos primeros componentes principales del ACP y de cada uno de los 50 modelos. Se generó una matriz vacía con 400 filas, correspondientes con los 50 modelos discriminantes para cada uno de los 8 subconjuntos, y con las 105 variables ambientales en las columnas. Para cada modelo (fila) se calculó el porcentaje de explicación ponderado de cada variable ambiental incluida (columna) como el promedio de los pesos de la variable en los dos primeros ejes canónicos, ponderado por las proporciones de explicación de esos mismos ejes y de los dos primeros componentes principales del ACP. Las celdas correspondientes con variables no incluidas en cada modelo, se rellenaron con valores 0. Luego, para cada variable (columna), se calculó el promedio de los porcentajes calculados en el paso anterior agrupando por subconjunto (porcentaje de explicación ponderado promedio; $n=50$ modelos), obteniéndose para cada variable 8 valores, uno por cada subconjunto. Finalmente, para cada variable, se calculó la suma y el desvío estándar de los porcentajes de explicación ponderados promedios ($n=8$) y se seleccionó el 20% de variables con mayor suma (21 variables), y se las dividió en mitades con mayor y menor cociente suma/desvío estándar. Una variable con un valor de suma elevado, significa que tuvo una importancia elevada para explicar la interacción GxS-FS de algunos o todos los subconjuntos. Si, además, esa variable tuvo un cociente suma/desvío elevado, significa que fue importante en la mayoría los subconjuntos (variables persistentes). Por el contrario, valores reducidos del cociente suma/desvío indican que fue importante solo en algunos subconjuntos (variables no persistentes).

Se analizó la incidencia del tipo de variable ambiental y del período temporal sobre la interacción GxS-FS. Las variables ambientales fueron clasificadas por tipo en 5 grupos: fotoperíodo, temperatura, precipitación, suelo y IVN; y por el período temporal alrededor de la siembra en 6 grupos: 150 DAS a 150 DDS, 30 DAS a 145 DDS, 0 a 145 DDS, 0 a 55 DDS, 55 a 145 DDS y 90 a 145 DDS. Para cada grupo de variables (ya sea por tipo o por período), se calculó su proporción entre las 105 variables ambientales posibles y también entre las 10 variables persistentes, y se calculó el cociente entre ambas proporciones. Un valor mayor a 1 para un grupo de variables, indica que las variables ambientales sintéticas que explicaron la interacción GxS-FS tienden a incluir variables de ese grupo particular (su presencia es mayor que la esperada por azar) y significa que ese tipo de variable o período temporal está asociado con la interacción GxS-FS.

2.D. Mapeo dinámico de mega-ambientes

Los mega-ambientes, inicialmente identificados por el ranking de rendimiento de los genotipos y luego caracterizados ambientalmente, fueron mapeados a partir de localizar los rangos ambientales que los definen (rangos de las variables ambientales sintéticas identificadas en el apartado anterior). Se generó el mapa de mega-ambientes en un escenario de referencia, representado por una campaña neutra y fecha de siembra temprana. El proceso consistió en dos pasos: primero, mapear los mega-ambientes de cada subconjunto por separado y, luego, resumir los 8 mapas generados en un único mapa consenso. Finalmente, los dos pasos anteriores fueron repetidos en otros escenarios temporales para completar las seis combinaciones posibles entre campañas (seca, neutra, húmeda) y fechas de siembra (temprana y tardía). Haber definido a los mega-ambientes como rangos ambientales en lugar de grupos de sitios-fechas de siembra permitió asignarles la categoría de verdaderos ambientes y referirse a la interacción en términos de interacción GxA en lugar de genotipo x sitio-fecha de siembra.

I. Mapa de mega-ambientes en un escenario temporal de referencia

Los mega-ambientes fueron mapeados en un escenario temporal de referencia representado por campaña neutra (cuantil 50% de precipitación anual junio-mayo) y fecha de siembra temprana (Fig. 2.1d y 2.5). Se generaron 8 mapas, uno por cada subconjunto. Para cada uno, se mapearon las 100 variables ambientales sintéticas (50 modelos discriminantes x 2 variables sintéticas) en las celdas de 0.5x0.5° a partir de las variables ambientales multiplicadas por los coeficientes de peso provenientes de los modelos discriminantes (autovectores). Para cada modelo, se determinaron los rangos ambientales que definen a los mega-ambientes por los polígonos de Thiessen formados por los centroides de los mega-ambientes en el plano de las dos variables ambientales sintéticas (Thiessen y Alter 1911), y considerando la máxima distancia euclídea entre un sitio-fecha de siembra y el centroide del mega-ambiente al que pertenecía. Así, por ejemplo, dadas dos variables sintéticas a y b: sea la variable ‘a’ una combinación de temperatura 0-45 DDS y precipitación 90-145 DDS y la variable ‘b’ una combinación de densidad aparente 0-30 cm y profundidad efectiva; el rango ambiental correspondiente con un mega-ambiente se representó como un polígono de Thiessen en el plano de los dos ejes ‘a’ y ‘b’, que implica un conjunto de combinaciones de valores para las 4 variables involucradas (temperatura, precipitación, densidad y profundidad). Si una celda de 0.5x0.5° tomaba valores de las variables sintéticas dentro del rango definido para un mega-ambiente, se le asignaba como categoría ese mega-ambiente, y si tomaba valores fuera de rango para todos los mega-ambientes, se le asignaba la categoría “no mapeada”. De este modo, las celdas no mapeadas representaron condiciones ambientales distintas a las exploradas por los mega-ambientes de un subconjunto completo. Para cada subconjunto, este proceso generó 50 mapas de mega-ambientes (uno por cada modelo discriminante) que fueron resumidos en un único mapa, representado por el mega-ambiente más frecuente en cada celda 0.5x0.5°. Zonas con un mismo mega-ambiente tienen similar ranking de rendimiento de los genotipos (baja interacción GxA). Los mapas resultantes fueron finalmente refinados mediante la técnica del vecino más cercano, para una ventana de 3x3 celdas. Se empleó el algoritmo r. neighbours de Grass Gis (GRASS Development Team, 2016).

Se resumieron los 8 mapas de mega-ambientes de los subconjuntos completos en un único mapa consenso. El proceso requirió previamente rellenar las celdas no mapeadas de cada subconjunto asignándoles el valor de mega-ambiente de la celda más cercana (se empleó la rutina `gdal_grid` de Geospatial Data Analysis Library, GDAL, con el algoritmo `Nearest Neighbor`). Se realizó un análisis de cluster sobre una matriz con las celdas de mapas como casos a clasificar (filas) y los mapas de subconjuntos como variables de clasificación (8 columnas). El cuerpo de esta matriz contenía el valor de pendiente de la regresión rendimiento-GM de cada celda $0.5 \times 0.5^\circ$ para cada subconjunto completo. Dado que los distintos subconjuntos exploraron distintos rangos de GM, los valores de pendiente fueron estandarizados previamente por una transformación ranking para hacerlos comparables entre subconjuntos. El resultado fue un mapa consenso donde las zonas son grupos de celdas $0.5 \times 0.5^\circ$ con el mismo mega-ambiente o mega-ambientes similares entre subconjuntos. Se emplearon las rutinas `i.cluster` y `i.maxlik` de Grass Gis, para un número de clases igual a 6 (GRASS Development Team, 2016).

II. Mapeo de la variación ambiental asociada a mega-ambientes

Las variables ambientales sintéticas que explicaron la interacción GxS-FS fueron mapeadas en distintas campañas y fechas de siembra (Fig. 2.5). Luego, se generaron mapas de CV promedio entre campañas y entre fechas de siembra. La variación entre campañas fue analizada en dos escenarios de fecha de siembra: temprana y tardía, mientras que la variación entre fechas de siembra fue analizada en tres escenarios de campaña: seca, neutra y húmeda (cuantil 10, 50 y 90% de precipitación anual junio-mayo). El CV promedio fue obtenido a partir de CV calculados para los 50 modelos de los 8 subconjuntos completos ($n=400$). Para el cálculo del CV de cada modelo, se promediaron los CV de las dos variables ambientales sintéticas ($n=13$ entre campañas, $n=2$ entre fechas de siembra), ponderando por las proporciones de varianza explicadas (autovalores).

III. Mapas de mega-ambientes en distintos escenarios temporales (mapeo dinámico)

Se analizó la variación temporal del mapa de mega-ambientes al cambiar entre campañas y fechas de siembra. Del mismo modo que para el escenario de referencia, fueron generados mapas consenso de mega-ambientes para otros escenarios temporales contrastantes (Fig. 2.1d). En total, se obtuvieron seis mapas consenso correspondientes a tres niveles de precipitación anual junio-mayo (campañas secas, cuantil 10%; neutra, 50%; y húmeda, 90%) combinados con dos fechas de siembra (temprana y tardía; Fig. 2.5). Se analizaron visualmente los cambios de ubicación de los mega-ambientes entre los distintos escenarios.

CAPÍTULO 3: Resultados

3.A. Interacción genotipo x sitio-fecha de siembra y mega-ambientes

Dentro de cada subconjunto completo de datos de rendimiento, el factor ambiental sitio-fecha de siembra (Sitio-FS) explicó la mayor parte de la variación del rendimiento (62%), seguido por la interacción genotipo x sitio-fecha de siembra (GxS-FS; 23%) y por el factor genotípico (5%). Dentro de cada mega-ambiente (grupo de sitios-fecha de siembra con similar ranking de rendimiento de los genotipos), el factor ambiental continuó explicando la mayor parte de la variación (cerca del 60%), pero la interacción GxS-FS se redujo a la mitad, alcanzando un valor de 12%, y el factor genotípico se multiplicó por 3,5, alcanzando un valor de 18%. En consecuencia, dentro de cada mega-ambiente el factor genotípico explicó mayor proporción de la variación de rendimiento que la interacción GxS-FS (Fig. 3.1).

La respuesta diferencial de los genotipos frente a los distintos mega-ambientes estuvo asociada con el GM y el hábito de crecimiento. El 70% de los mega-ambientes mostró una relación significativa entre el rendimiento y el GM de los genotipos y un 50% mostró incrementos o disminuciones de más de 15 kg ha⁻¹ por cada unidad de GM (GM en la escala de 0 a 100; $p < 0,05$; Fig. 3.2). El hábito de crecimiento explicó parte de la respuesta de los genotipos que no fue explicada por el GM (Fig. 3.3). En general, los genotipos con hábito determinado (minoritarios) mostraron rendimientos 500 kg ha⁻¹ por debajo de la media de su GM, mientras que aquellos con hábito indeterminado (mayoritarios) fueron superiores a la media en 100 kg ha⁻¹ (Fig. 3.3 A). Al discriminar entre mega-ambientes por su respuesta frente al GM, el efecto del hábito de crecimiento fue máximo en los mega-ambientes donde no hubo una relación significativa entre rendimiento y GM, alcanzando valores de 700 kg ha⁻¹ por debajo y 180 kg ha⁻¹ por encima de la media para genotipos con hábito determinado e indeterminado respectivamente (Fig. 3.3 B, barras negras).

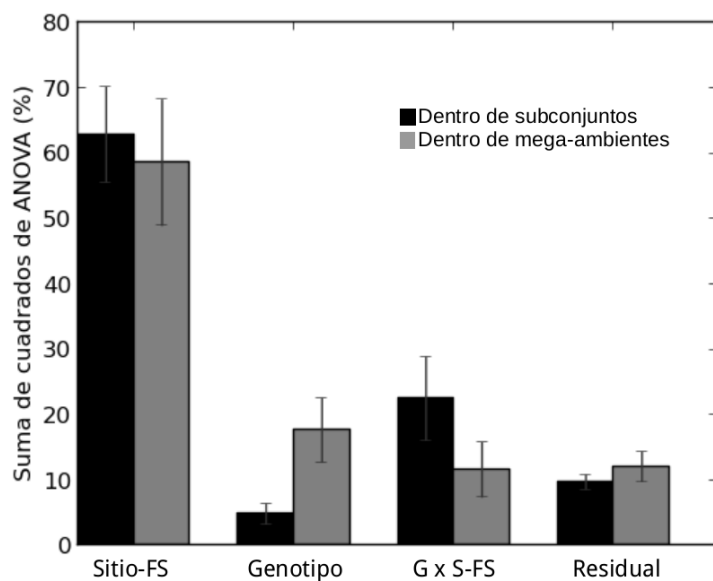


Figura 3.1: Porcentaje de la suma de cuadrados total explicada por los factores ambiental, representado por el sitio-fecha de siembra (Sitio-FS; fecha de siembra incluyendo el año y el día del año), genotípico (Genotipo), la interacción entre ambos (GxS-FS) y el residuo. Se muestran los promedios y los desvíos estándar ($n=8$) de los ANOVA realizados para los subconjuntos completos (barras negras) y para los mega-ambientes (barras grises).

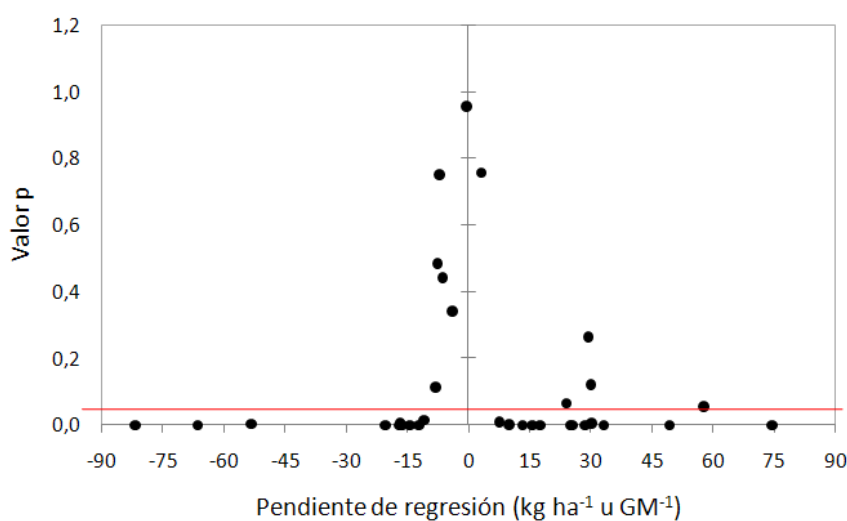


Fig. 3.2: Pendientes y valores p de las regresiones entre rendimiento centrado por ambiente en kg ha^{-1} y GM de 0 a 100, para cada mega-ambiente. Se muestran todos los mega-ambientes definidos en los 8 subconjuntos completos ($n=34$). Valores de pendiente positivos indican mega-ambientes con rendimientos menores para GM más cortos, y mayores para GM más largos; valores negativos, lo contrario. Los puntos por debajo de la línea roja representan regresiones estadísticamente significativas ($p < 0,05$).

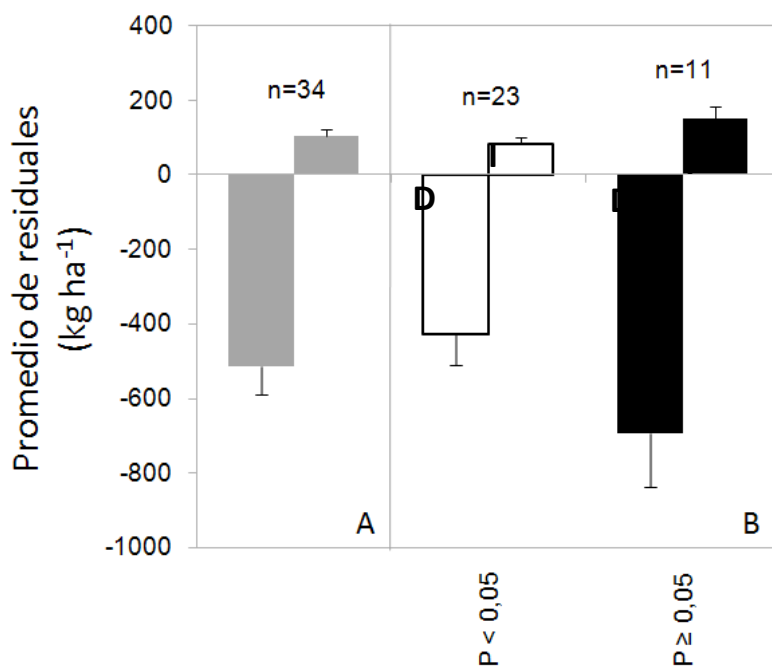


Fig. 3.3: Efecto del hábito de crecimiento sobre el rendimiento cuantificado por los residuales de la regresión por GM (Fig. 3.2). Se muestran los residuales para hábito de crecimiento determinado (D) e indeterminado (I). A: promedio y error estándar de los 34 mega-ambientes definidos en los 8 subconjuntos completos; B: a partir de los mismos 34 mega-ambientes, se discriminó entre aquellos con pendiente significativa ($p < 0.05$; $n = 23$) y no significativa ($n = 11$) respecto del GM.

La interacción GxS-FS entre pares de genotipos mostró dos patrones consistentes y algunas excepciones. El patrón más importante fue la diferencia entre los GM menores y mayores a VI. La interacción, medida por la distancia euclídea estandarizada, fue mínima entre genotipos de GM menor a VI (-1.0 en promedio), mayor entre genotipos de GM mayor a VI (0.1) y máxima entre genotipos GM menores y mayores a VI (0.4) (Fig. 3.4 A y B). Un segundo patrón fue que, en general, los genotipos de GM dispares (Fig. 3.4 A; puntos distantes de la diagonal) mostraron mayor interacción que aquellos de GM semejantes (puntos próximos a la diagonal). Algunas excepciones se debieron al comportamiento particular de ciertos genotipos (ADM 4800, BMX TITAN, NA 6411 y SYN 9070 RR, algunos de ellos con hábito de crecimiento determinado) o de pares de genotipos (FN 4.50 con DM 6.8i, DM 6.2i con DM 7.8i). Al considerar únicamente genotipos de hábito indeterminado, muchas de estas excepciones desaparecieron y se hicieron más evidentes los patrones descritos anteriormente (Fig. 3.5; GM mayores a VI).

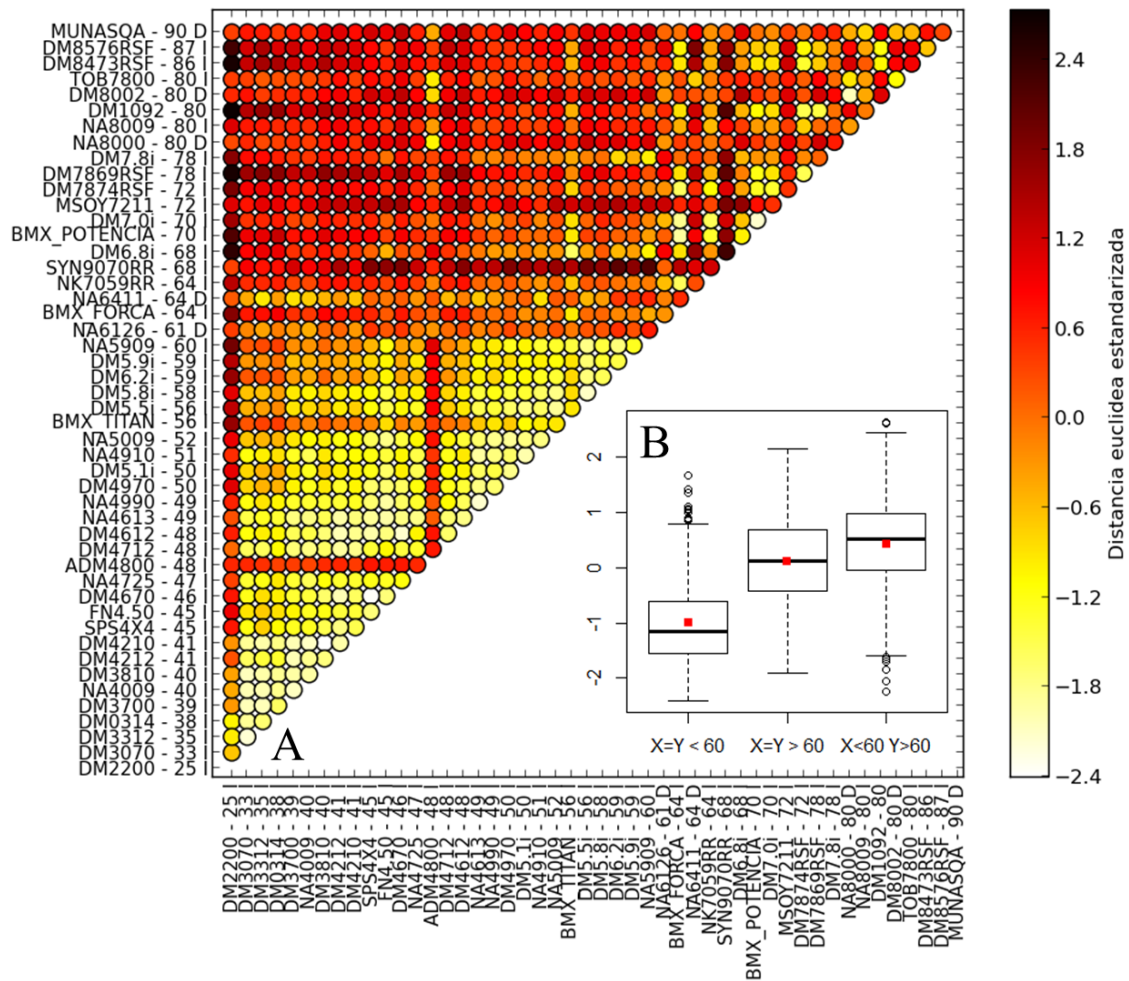


Figura 3.4. A: Interacción genotipo por sitio-fecha de siembra (GxS-FS) entre pares de genotipos ordenados por GM creciente. La interacción fue cuantificada por la distancia euclídea estandarizada promedio para cada par de genotipos: mayor distancia, mayor interacción GxS-FS. En ambos ejes se indican nombre, grupo de madurez (GM; de 0 a 100) y hábito de crecimiento de los genotipos (I: indeterminado, D: determinado). Puntos cercanos a la diagonal representan combinaciones entre dos genotipos con GM semejante; puntos alejados de la diagonal, combinaciones entre genotipos de GM dispar. Se incluyen todos los genotipos de los 8 subconjuntos completos. **B:** Las mismas comparaciones de a pares de genotipos de la figura A, fueron agrupadas en tres categorías: comparaciones entre genotipos de GM menor a VI (valores de GM menor a 60 en ambos ejes, X e Y), comparaciones entre genotipos de GM mayor a VI (valores de GM mayor a 60 en ambos ejes) y comparaciones de genotipos de GM menor a VI con genotipos de GM mayor a VI (valores de GM menor a 60 en eje X y mayor a 60 en eje Y). Se muestran las distribuciones de valores de distancia para las tres categorías definidas.

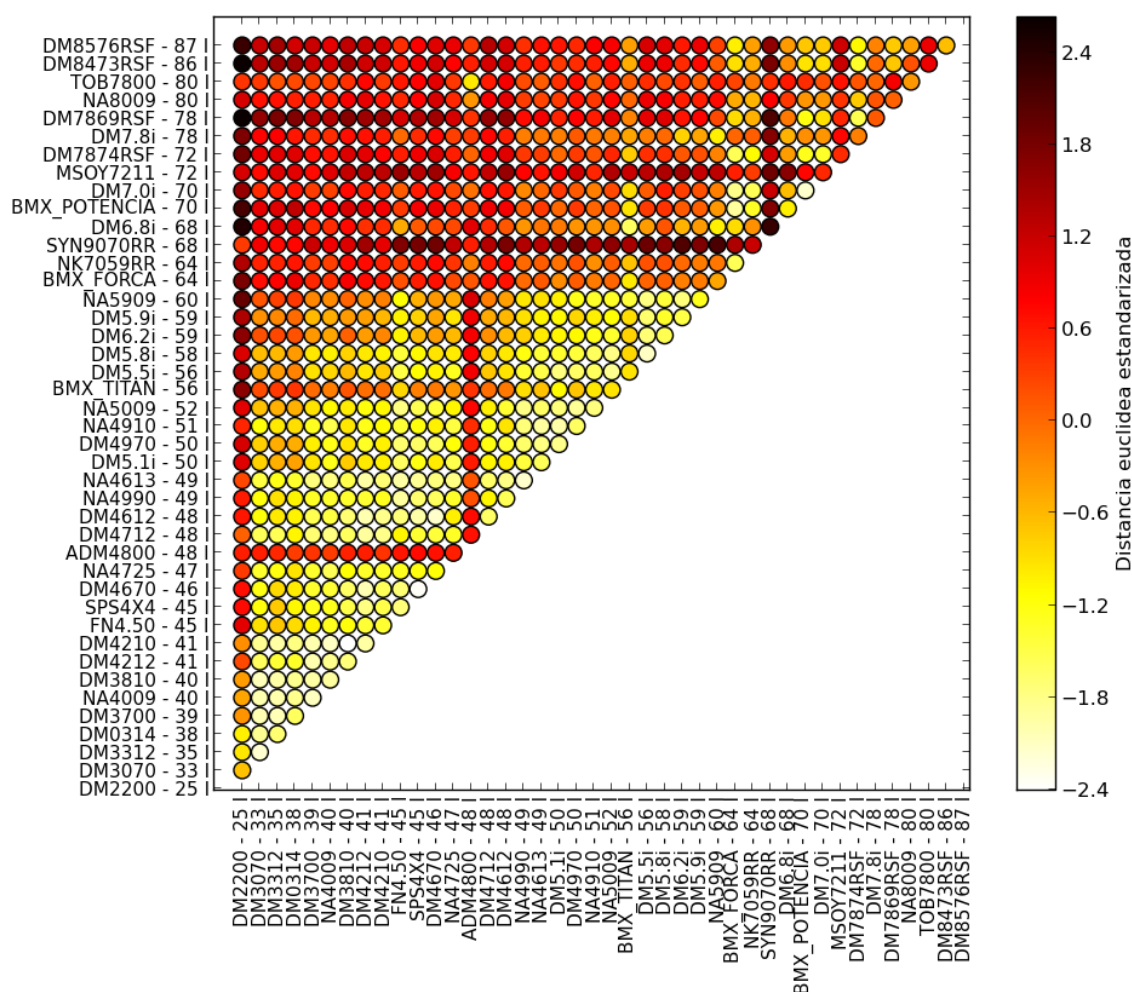


Figura 3.5: Interacción genotipo x sitio-fecha de siembra (GxS-FS) entre pares de genotipos ordenados por GM creciente circunscripta a los genotipos de hábito indeterminado. La interacción está cuantificada por la distancia euclídea estandarizada promedio para cada par de genotipos: mayor distancia, mayor interacción GxS-FS. En cada eje se indican nombre, grupo de madurez (GM; de 0 a 100) y hábito de crecimiento de los genotipos. El eje vertical contiene los mismos genotipos que el eje horizontal. Puntos cercanos a la diagonal representan combinaciones entre dos genotipos con GM semejante; puntos alejados de la diagonal, combinaciones entre genotipos de GM dispar.

3.B. Caracterización ambiental de los mega-ambientes

Los mega-ambientes, previamente definidos por el ranking de rendimiento de los genotipos, fueron caracterizados a través de variables ambientales (meteorológicas, IVN y edáficas) que explicaron la interacción genotipo x sitio-fecha de siembra (Fig. 3.6; Tabla 3.1). Algunas variables fueron relevantes en la mayoría de los subconjuntos completos (variables persistentes), mientras que otras solo lo fueron en algunos subconjuntos específicos (variables no persistentes). Entre las variables persistentes, las más importantes fueron el fotoperíodo durante todo el ciclo del cultivo y particularmente en post-floración, la precipitación durante floración, fructificación y llenado de granos, la temperatura nocturna durante los estadios vegetativos, la profundidad efectiva del perfil

y la capacidad de almacenamiento hídrico (Fig. 3.6a). Entre las no persistentes, las más importantes fueron la temperatura durante floración, fructificación y llenado de granos, la precipitación de la campaña y el IVN durante llenado de granos (Fig. 3.6b). Una síntesis de la información anterior muestra que las variables que explicaron la interacción GxS-FS describían aspectos del fotoperíodo y del suelo (Fig. 3.7a) y que las fases reproductivas tardías fueron las más relevantes (Fig. 3.7b).

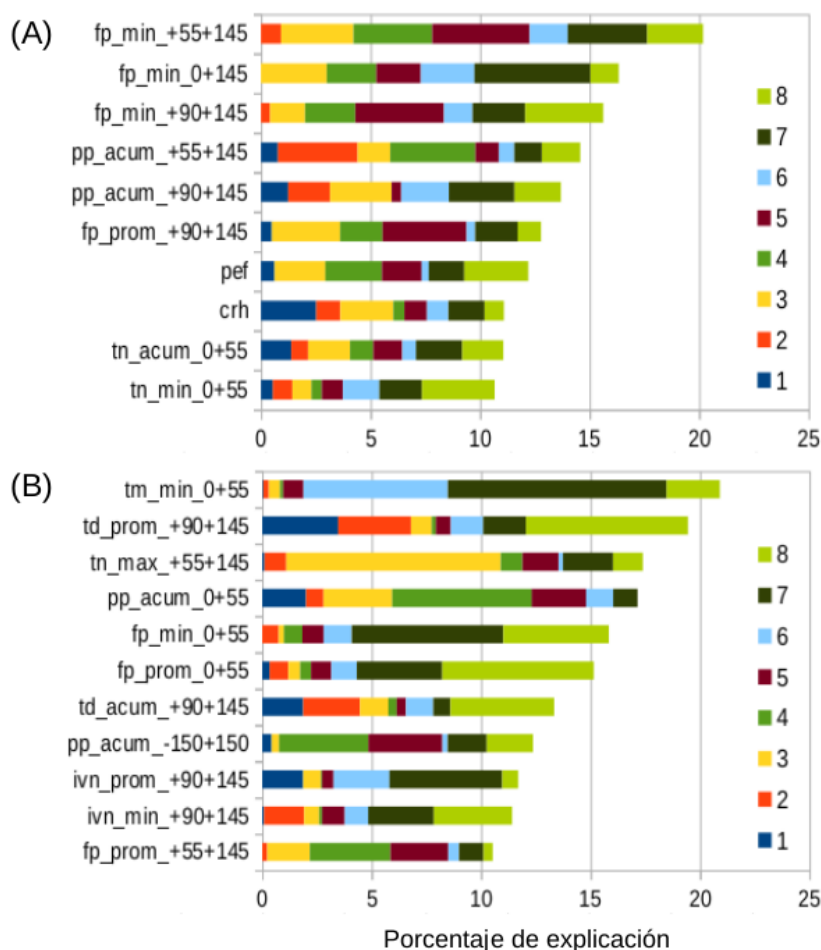


Figura 3.6: Variables ambientales (Tabla 3.1) más asociadas con la interacción genotipo x sitio-fecha de siembra (GxS-FS) y su grado de asociación con dicha interacción, cuantificado por el porcentaje de explicación ponderado promedio. Se muestran las variables que fueron importantes para la mayoría de los subconjuntos completos (A; variables persistentes) y aquellas que solo lo fueron para algunos de ellos (B; variables no persistentes). Dentro de una barra, los colores indican los distintos subconjuntos. La longitud total de una barra representa la suma de los porcentajes de explicación de los 8 subconjuntos. Se muestran los nombres abreviados, conformados por tres partes separadas por guiones. La primera indica la variable ambiental: “pef”: profundidad efectiva; “fp”: fotoperíodo; “pp”: precipitación; “crh”: capacidad de retención hídrica; “td”: temperatura diurna; “tm”: temperatura media; “tn”: temperatura nocturna; “ivn”: índice de vegetación normalizado. La segunda parte indica el estadístico de resumen temporal: “prom”: promedio; “max”: máximo; “min”: mínimo; “acum”: acumulada. La tercera parte indica el período dentro del ciclo del cultivo en días desde la fecha de siembra. El primer número indica el inicio del período y el segundo el fin. Valores negativos y positivos indican fechas anteriores y posteriores a la siembra, respectivamente.

Tabla 3.1

Descripción de las 21 variables ambientales más asociadas con la interacción GxS-FS (Figura 3.6), ordenadas alfabéticamente. La línea horizontal divide las variables persistentes (arriba) y no persistentes (abajo).

Abreviatura	Variable	Estadístico	Inicio	Fin	Fase fenológica	Horizonte de suelo
crh	Almacenaje	-	-	-	-	0-100 cm
fp_min_0+145	Fotoperíodo	mínimo	0	145	S – R8	-
fp_min_+55+145	Fotoperíodo	mínimo	55	145	R1 – R8	-
fp_min_+90+145	Fotoperíodo	mínimo	90	145	R5 – R8	-
fp_prom_+90+145	Fotoperíodo	promedio	90	145	R5 – R8	-
pef	Prof. Efectiva	-	-	-	-	0-100 cm
pp_acum_+55+145	Precipitación	Suma	55	145	R1 – R8	-
pp_acum_+90+145	Precipitación	Suma	90	145	R5 – R8	-
tn_min_0+55	Temp. nocturna	mínimo	0	55	S – R1	-
tn_acum_0+55	Temp. Nocturna	Suma	0	55	S – R1	-
fp_min_0+55	Fotoperíodo	mínimo	0	55	S – R1	-
fp_prom_0+55	Fotoperíodo	promedio	0	55	S – R1	-
fp_prom_+55+145	Fotoperíodo	promedio	55	145	R1 – R8	-
ivn_min_+90+145	IVN	mínimo	90	145	R5 – R8	-
ivn_prom_+90+145	IVN	promedio	90	145	R5 – R8	-
pp_acum_-150+150	Precipitación	Suma	-150	150	S – R8 (*)	-
pp_acum_0+55	Precipitación	Suma	0	55	S – R1	-
td_prom_+90+145	Temp. Diurna	promedio	90	145	R5 – R8	-
td_acum_+90+145	Temp. Diurna	Suma	90	145	R5 – R8	-
tm_min_0+55	Temp. Media	mínimo	0	55	S – R1	-
tn_max_+55+145	Temp. nocturna	máximo	55	145	R1 – R8	-

Se detallan: la variable, el estadístico de resumen temporal para los datos incluidos dentro del período definido por las columnas inicio y fin (en días desde la fecha de siembra) y una fase fenológica (Fehr y Kaviness, 1977) que podría coincidir con ese período temporal. Para las variables edáficas se indica el horizonte de suelo.

* Incluye un período de barbecho previo a la siembra.

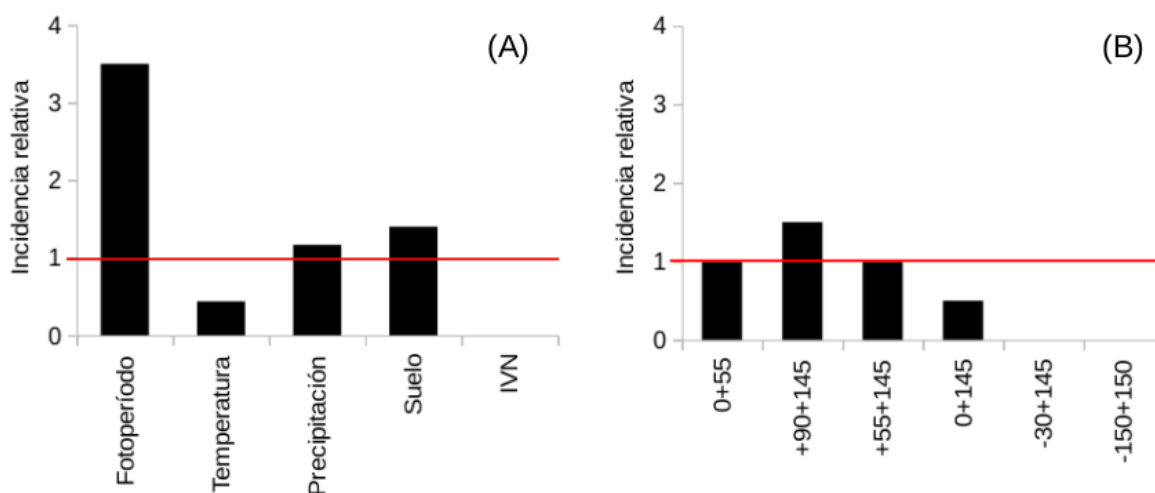


Figura 3.7: Incidencia relativa de cada tipo de variable (A) y período temporal (B) sobre la interacción GxS-FS, calculada como el cociente entre dos proporciones: la proporción dentro de las 10 variables más persistentemente asociadas con la interacción (Fig. 3.6a) y la proporción dentro de las 105 variables ambientales totales (Tabla A.1). Valores mayores a 1 indican que las variables ambientales sintéticas que explicaron la interacción GxS-FS tendieron a incluir variables de ese grupo particular (su presencia es mayor que la esperada por azar). El grupo suelo incluyó todas las variables edáficas analizadas (Tabla 2.5). Los períodos temporales se indican con dos valores de días desde la fecha de siembra, que representan el inicio y el fin de cada período. Valores negativos indican cantidad de días antes de la fecha de siembra; valores positivos, cantidad de días después de la fecha de siembra; valores 0 indican la fecha de siembra.

3.C. Mapeo dinámico de mega-ambientes

Los mega-ambientes, inicialmente definidos por el ranking de rendimiento de los genotipos y luego caracterizados ambientalmente, fueron mapeados en diferentes escenarios temporales dados por la campaña y la fecha de siembra. En primer lugar, se muestran los mapas en un escenario de referencia (precipitación intermedia y fecha de siembra temprana). Luego se analiza la variabilidad ambiental y se muestran los mapas en escenarios de campañas con mayores y menores precipitaciones y siembras tempranas y tardías.

1. Mapa de mega-ambientes en un escenario temporal de referencia

Para campañas con precipitación intermedia (neutras, cuantil 50%) y fecha de siembra temprana, los mapas mostraron patrones espaciales similares para la mayoría de los subconjuntos (Fig. 3.8). Entre las similitudes, hubo un patrón Noreste-Sudeste y cierta correspondencia con las unidades biogeográficas y de vegetación de América del Sur (Huëck y Seibert 1969; Morrone 2014): los mega-ambientes donde los GM largos superaron en rendimiento a los cortos (favorables para los GM largos), se ubicaron generalmente hacia el Norte, sobre la provincia de Cerrado, y hacia el Oeste sobre la

provincia Chaqueña, mientras que los mega-ambientes donde los GM cortos superaron a los largos (favorables para los GM cortos), se ubicaron hacia el Sur sobre la provincia Pampeana y hacia el Este sobre el Bosque de Araucaria. A su vez, la heterogeneidad espacial aumentó en sentido Norte a Sur. La provincia de Cerrado fue cubierta generalmente por un único mega-ambiente, que llegó a ocupar además extensiones variables de la provincia Chaqueña, mientras que la provincia Pampeana contuvo generalmente dos o más mega-ambientes simultáneamente. Algunos sectores aislados sobre el límite Oeste de la provincia Chaqueña fueron cubiertos por mega-ambientes favorables para los GM largos que ocuparon también sectores de la provincia Pampeana.

Los mapas mostraron algunos patrones particulares entre subconjuntos. Algunos de estos patrones respondieron al rango de GM explorado por cada subconjunto. El patrón general Noroeste-Sudeste fue más claro para los subconjuntos de rango de GM más amplio (Fig. 3.8, mapas 1-5), mientras que se hizo más difuso o no existió para los rangos de GM muy acotados (Fig. 3.8, mapas 6-8), para los cuales la mayoría de los mega-ambientes no tuvieron respuestas en rendimiento claramente asociadas con el GM. A su vez, la heterogeneidad espacial disminuyó en el Norte y aumentó en el Sur del área de estudio al disminuir el rango de GM desde amplio a muy acotado.

Como se mencionó anteriormente, la ubicación de los mega-ambientes guardó un alto grado de asociación con las unidades biogeográficas y de vegetación (Huëck y Seibert 1969; Morrone 2014; Tabla 2.2). El área que comprende la provincia de Cerrado y el Bosque Paranaense fue cubierta mayoritariamente por un solo mega-ambiente excepto en los mapas 1 y 8 de la Fig. 3.8, elaborados con rangos de GM amplio y muy acotado, respectivamente. En los casos mayoritarios, que incluyeron rangos intermedios de GM, esta área resultó favorable para los GM largos (Fig. 3.8, mapas 2-7), al igual que la transición entre la Pluviselva del Amazonas y la Sabana de Palmeras de Bolivia sobre el Noroeste del área de estudio (Fig. 3.8, mapas 4, 5, 6). Para el subconjunto con el rango más amplio (Fig. 3.8, mapa 1), la provincia de Cerrado y el Bosque Paranaense incluyeron tres mega-ambientes: uno muy favorable para los GM largos y desfavorable para el hábito determinado al Oeste y al Norte, otro moderadamente favorable para los GM largos al Este, y el último débilmente favorable para los GM largos y desfavorable para el hábito determinado al Sur. En el caso que incluyó solo genotipos de GM V y VI (Fig. 3.8, mapa 8), hubo dos mega-ambientes que se diferenciaron principalmente por la respuesta del hábito de crecimiento: uno muy desfavorable para el hábito determinado en el centro y otro ligeramente desfavorable en el Norte y el Sudeste.

La provincia Chaqueña resultó favorable para los GM largos en toda su área, excepto en el subconjunto de rango de GM más amplio (Fig. 3.8, mapa 1). Por un lado, en los casos mayoritarios (Fig. 3.8, mapas 2-8), mostró dos mega-ambientes que se diferenciaron por la respuesta del hábito de crecimiento. En los subconjuntos con rangos de GM acotados (Fig. 3.8, mapas 4-7), esta provincia mostró un mega-ambiente muy desfavorable para el hábito determinado hacia el Este (por el Sur o por el Norte) que se extendió sobre la provincia de Cerrado, y otro ligeramente desfavorable al Oeste. En el subconjunto con rango GM intermedio (Fig. 3.8, mapa 2), mostró un mega-ambiente ligeramente desfavorable para el hábito determinado que se extendió sobre el Espinal. Por otro lado, en el subconjunto con rango de GM más amplio (mapa 1), la provincia Chaqueña mostró

dos mega-ambientes con respuestas de los GM contrarias: uno favorable para GM largos en el Noroeste, que ocupó también el Este de la provincia de Cerrado, y otro favorable para los GM cortos en el Sur, que se extendió sobre la provincia Pampeana.

La provincia Pampeana mostró una gran heterogeneidad espacial, excepto para el subconjunto con rango de GM más amplio (Fig. 3.8, mapa 1). En los casos mayoritarios (Fig. 3.8, mapas 2-8), la heterogeneidad espacial determinó que el Norte y el Este de esta provincia mostraran los mismos mega-ambientes que las provincias Chaqueña y de Cerrado. En los subconjuntos con rangos de GM intermedio y acotado (Fig. 3.8, mapas 2-5), estos mega-ambientes compartidos entre provincias se debieron al patrón general Noroeste-Sudeste descripto previamente, con mega-ambientes favorables para los GM largos hacia el Norte de la provincia Pampeana y favorables para los cortos hacia el Sur. En cambio, en los subconjuntos con rangos GM más acotados (Fig. 3.8, mapas 6-8), los mega-ambientes compartidos entre provincias se debieron a la respuesta del hábito determinado: el Noreste de la provincia Pampeana mostró mega-ambientes muy desfavorables para el hábito determinado, asociados con la mayor parte de la provincia de Cerrado (a excepción de algunas zonas hacia el Norte y el Este de esta provincia), mientras que el centro y el Oeste mostró mega-ambientes moderada o levemente desfavorables para el hábito determinado. En los casos minoritarios, en subconjuntos con rangos de GM intermedio y amplio, el Noreste (Fig. 3.8, mapa 2) o toda la provincia Pampeana (Fig. 3.8, mapa 1) mostraron mega-ambientes que ocuparon también el Bosque de las Sierras de Chiquitos y el Norte de la provincia Chaqueña, favorables para los GM cortos.

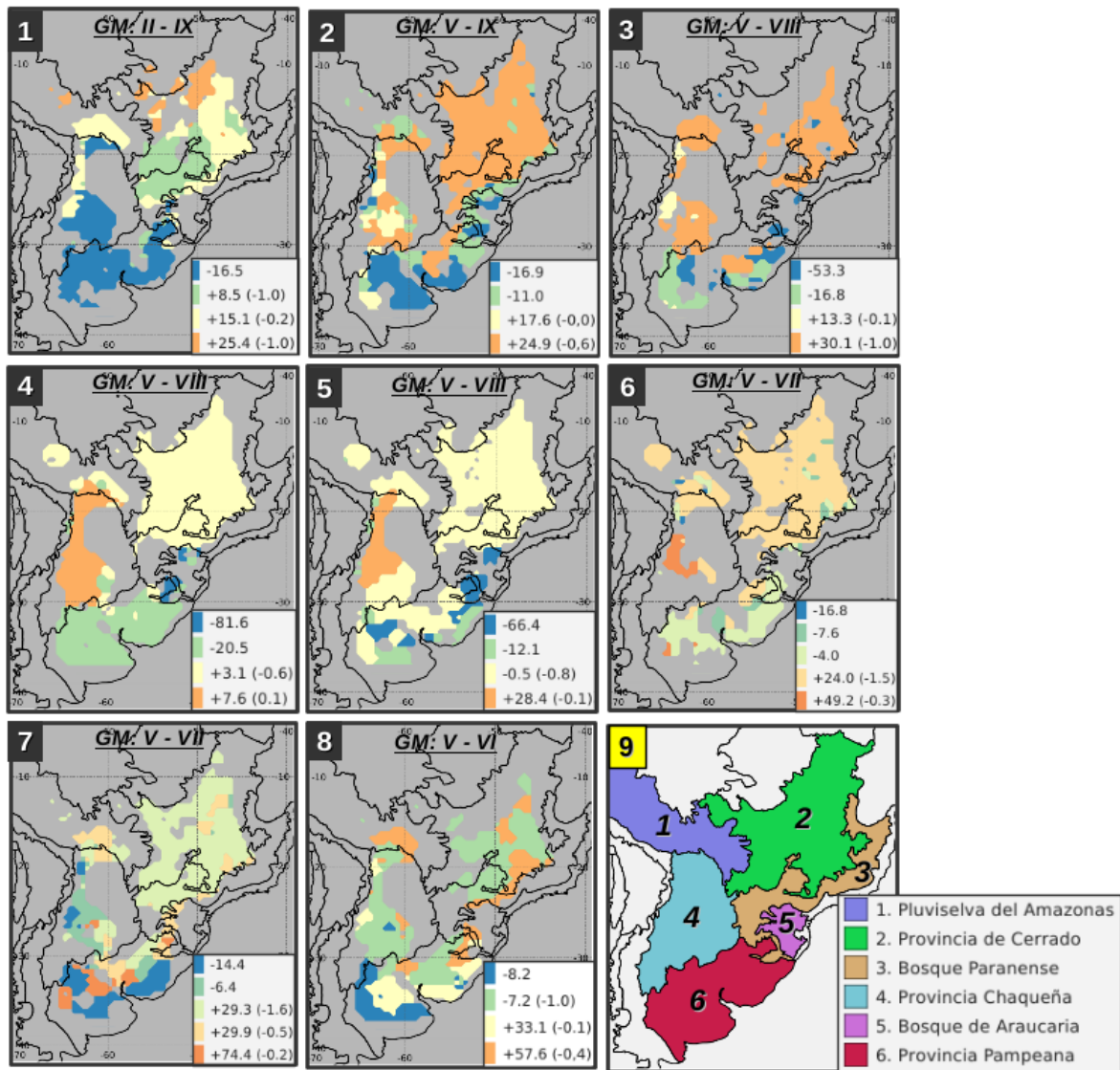


Figura 3.8: 1-8: Zonas homogéneas de baja interacción GxA (mega-ambientes), para campañas neutras (cuantil 50% de precipitación anual) y fecha de siembra temprana (Fig. 2.5). Cada mapa corresponde con un subconjunto completo. En el extremo superior de cada mapa se indica el rango de GM de los genotipos incluidos ese subconjunto. Los recuadros inferiores derechos muestran la pendiente de la regresión entre rendimiento y grupo de madurez, en kg ha⁻¹ por unidad de GM (de 0 a 100). Valores positivos indican mega-ambientes con rendimientos menores para GM cortos y mayores para GM largos (favorables para GM largos); valores negativos, lo contrario. Entre paréntesis se indica el efecto del hábito de crecimiento determinado sobre el rendimiento en kg ha⁻¹ 10⁻³, cuantificado por los residuales de la regresión por GM, para los mega-ambientes con pendiente positiva o neutra. En gris, se muestran las áreas excluidas del análisis y las áreas no mapeadas, donde el ambiente escapó al rango ambiental explorado por los genotipos en los sitios-fecha de siembra de cada subconjunto. Líneas negras indican límites de las unidades biogeográficas propuestas por Morrone (2014). 9: Unidades biogeográficas propuestas por Morrone (2014) dentro del área de estudio de esta tesis (Tabla 2.2).

El mapa consenso de los 8 subconjuntos mostró más claramente el patrón Noroeste-Sudeste descrito anteriormente, con mega-ambientes cuya respuesta en rendimiento varió gradualmente desde más favorable para GM largos (valor de pendiente más positivo) a más favorable para GM cortos (más negativo). Siguiendo este patrón, mostró dos grandes áreas, centro-Norte y Sur, con respuestas contrastantes (Fig. 3.9). El área centro-Norte (zonas con valor positivo, en tonos de naranja), que abarcó la Pluviselva del Amazonas y la Sabana de Palmeras de Bolivia, el Bosque de las Sierras de Chiquitos, la provincia Chaqueña, la provincia de Cerrado, el Bosque Paranaense y el centro-Norte de la provincia Pampeana, mostró mega-ambientes favorables para los GM largos. En esta área, los genotipos DM 7.8i, DM 7869 RSF y Munasqa superaron a DM 3810, DM 4212, NA 5009, NA 6126 y NA 6411. El área S (zonas con signo negativo, en tonos de verde a azul), que abarcó desde la provincia Pampeana hasta el Bosque de Araucaria, mostró mega-ambientes favorables para los GM cortos. Los genotipos DM 3810, DM 4212 y NA 5909 superaron a DM 7.8i, DM 7869 RSF y Munasqa en esta área, contrariamente a lo ocurrido en el área centro-Norte.

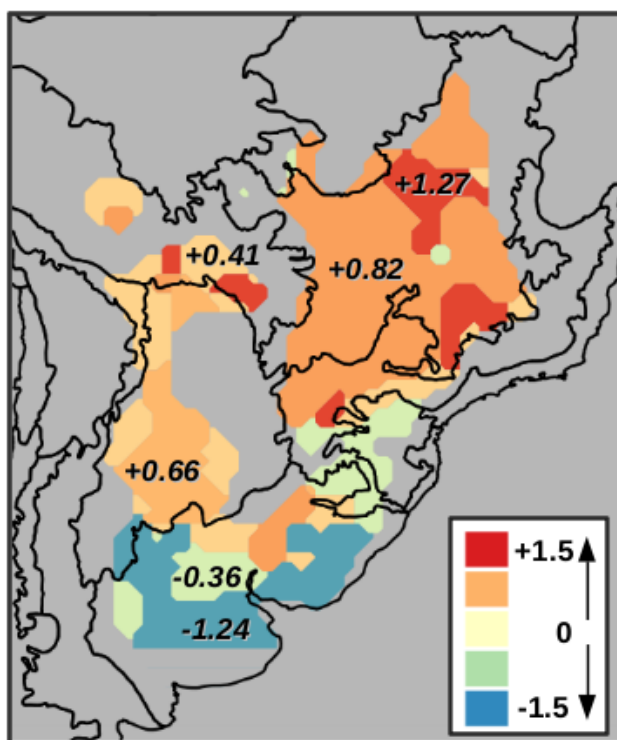


Figura 3.9: Integración de las zonas de baja interacción GxA (mega-ambientes) identificadas para los 8 subconjuntos completos, para campañas neutras (cuantil 50% de precipitación anual) y fecha de siembra temprana. Este mapa fue construido a partir de un análisis de clúster que incluyó los 8 mapas de la figura 3.8. Con un mismo color se representan zonas que, en la mayoría de los mapas, han sido mapeadas con un mismo mega-ambiente o mega-ambientes parecidos. El color y los números indican la pendiente entre rendimiento y GM dentro de cada zona. Sobre los valores de pendiente originales de cada mapa (Fig. 3.8), se aplicó una transformación ranking en dos tramos separados, para los valores de pendiente positivos y negativos. Se muestran los valores promedio ($n=8$ mapas) de la variable transformada previamente. Valores positivos indican mega-ambientes favorables para GM largos y negativos para GM cortos. En gris, se muestran las áreas excluidas del análisis. Líneas negras indican límites de unidades biogeográficas propuestos por Morrone (2014) (Fig. 3.8, mapa 9; Tabla 2.2).

Dentro del área centro-Norte, favorable para GM largos, se identificaron cuatro zonas por las respuestas del GM, el hábito de crecimiento y algunos genotipos específicos. La zona del Noroeste y el Sudeste de la provincia de Cerrado mostró los mega-ambientes con respuesta más favorable para estos GM (Fig. 3.9, pendiente +1.27; Fig. 3.8, mapas 1-3 de rango GM amplio a intermedio) y moderadamente desfavorables para el hábito de crecimiento determinado (Fig. 3.8, mapas 1 y 7 de rango amplio y acotado). Además de los genotipos mencionados anteriormente, en esta zona NA 8009 superó a DM 8002, NA 8000 y DM 5.9i. La zona que abarca el resto de la provincia de Cerrado y el Bosque Paranaense mostró mega-ambientes moderadamente favorables para GM largos (Fig. 3.9, pendiente +0.82; Fig. 3.8, mapas 1-6 de rango amplio a acotado) y muy desfavorables para el hábito determinado (Fig. 3.8, mapas 4-7 de rango acotado). En esta zona, DM 5.9i no mostró un comportamiento diferencial como sí lo hizo en otras zonas. La mayor parte de la provincia Chaqueña mostró mega-ambientes débilmente favorables para GM largos (Fig. 3.9, pendiente +0.66; Fig. 3.8, mapas 1-3 de rango amplio a intermedio) y débilmente desfavorables para el hábito determinado (Fig. 3.8, mapas 4-6 de rango acotado). Al contrario que en las anteriores, en esta zona DM 8002 y NA 8000 superaron a NA 8009. Finalmente, algunos sectores aislados en el Sudeste del Bosque Paranaense, en el Norte de la provincia Pampeana, en el centro-Sur y el Oeste de la provincia Chaqueña y en el centro y el Sudeste de la Pluviselva del Amazonas conformaron una zona de transición con mega-ambientes favorables para GM largos o cortos, dependiendo de los subconjuntos completos (Fig. 3.9, pendiente +0.41; Fig. 3.8, mapas 1-4 de rango amplio a acotado).

Dentro del área Sur, favorable para GM cortos, se identificaron dos zonas por las respuestas del GM y algunos genotipos específicos. Los sectores Noreste, centro y Este de la provincia Pampeana conformaron una zona con mega-ambientes muy favorables para esos GM (Fig. 3.9, pendiente -1.24; Fig. 3.8, mapas 1-2, 7-8, de rango amplio a acotado) donde, además de los genotipos mencionados anteriormente, DM 5.9i superó a NA 4910 y DM 8473 RSF. El centro-Norte y el Noreste de la provincia Pampeana y el Bosque de Araucaria conformaron una zona con mega-ambientes débilmente favorables para GM cortos (Fig. 3.9, pendiente -0.36; Fig. 3.8, mapas 1-2, 7-8, de rango amplio a acotado) donde NA 4910 y DM8473 RSF superaron a DM 5.9i, al contrario que en la zona anterior.

II. Mapeo de la variación ambiental asociada a mega-ambientes

Las variables ambientales sintéticas que explicaron la interacción GxS-FS variaron temporalmente de forma similar a lo descripto para la precipitación anual (Fig. 2.4). El coeficiente de variación (CV) entre fechas de siembra (Fig. 3.10) y entre campañas (Fig. 3.11) fue menor hacia el Norte (0.09 ± 0.03 y 0.08 ± 0.01 , respectivamente), en la Pluviselva del Amazonas y la Sabana de Palmeras de Bolivia, el Bosque de las Sierras de Chiquitos, la provincia de Cerrado, el Bosque Paranaense y el Bosque de Araucaria, y mayor hacia el Sur (0.14 ± 0.04 entre fechas y 0.10 ± 0.01 entre campañas), en las provincias Chaqueña y Pampeana. Esta diferencia entre zonas fue máxima entre la zona desde el Sur de la provincia Chaqueña hasta el E de la Pampeana y la zona del Norte de Cerrado.

Las variables ambientales sintéticas variaron más entre fechas de siembra que entre campañas, con valores de CV máximos de 0.30 entre fechas (Fig. 3.10) y 0.15 entre campañas (Fig. 3.11). El CV entre fechas de siembra dependió de la campaña y viceversa. En general, a lo largo de toda el área de estudio, el CV entre fechas de siembra fue menor para las campañas húmedas (0.101 ± 0.04) que para las neutras y secas (0.124 ± 0.04 ; Fig. 3.10), mientras que el CV entre campañas fue menor en fechas de siembra tempranas (0.080 ± 0.01) y mayor en tardías (0.094 ± 0.02 ; Fig. 3.11). Estas diferencias en el CV entre fechas al cambiar de campaña y en el CV entre campañas al cambiar de fecha fueron máximas en las provincias Chaqueña y Pampeana.

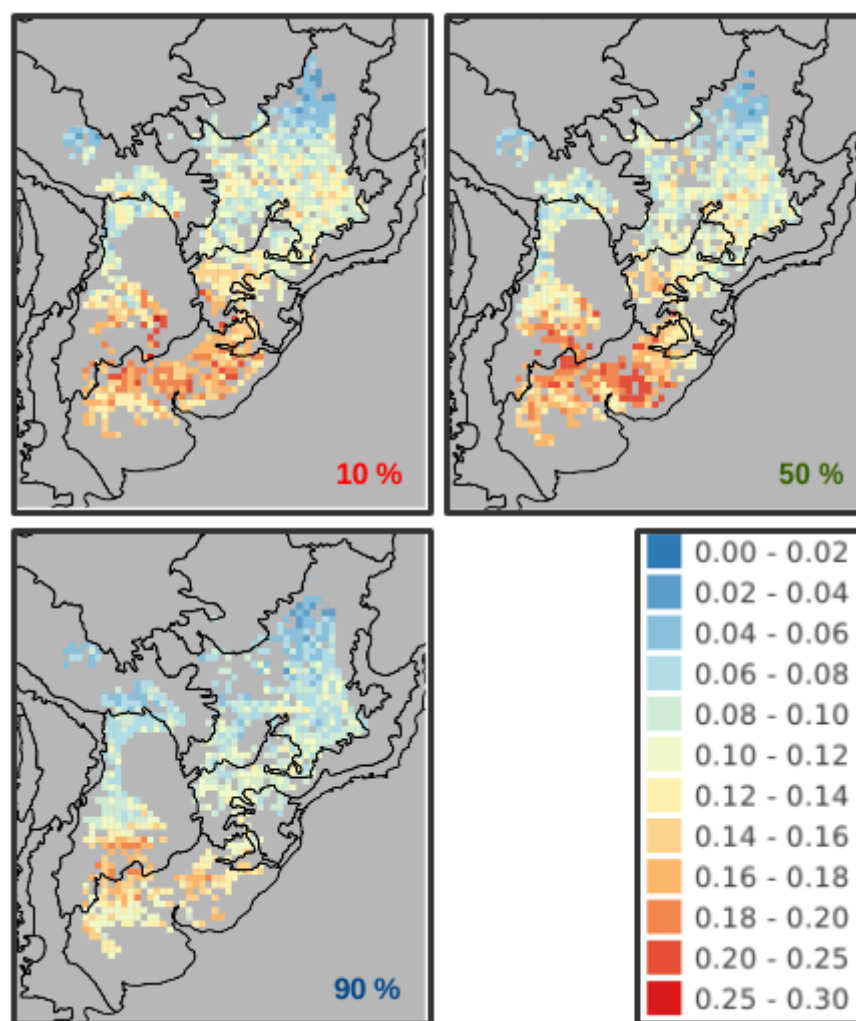


Figura 3.10: Coeficiente de variación promedio (CV) entre fechas de siembra temprana y tardía (Fig. 2.5) para las variables ambientales sintéticas que explicaron la interacción genotipo x sitio-fecha de siembra (GxS-FS). Se muestra el CV para campañas secas, neutras y húmedas (cuantiles 10, 50 y 90% de precipitación anual). Se incluyeron las variables ambientales sintéticas de los 8 subconjuntos completos. En gris, se muestran las áreas excluidas del análisis y las áreas donde no se calculó el CV porque el ambiente escapó al rango ambiental explorado por los genotipos en la mayoría de los sitios-fecha de siembra. Líneas negras indican límites de unidades biogeográficas propuestos por Morrone (2014) (Fig. 3.8, mapa 9; Tabla 2.2).

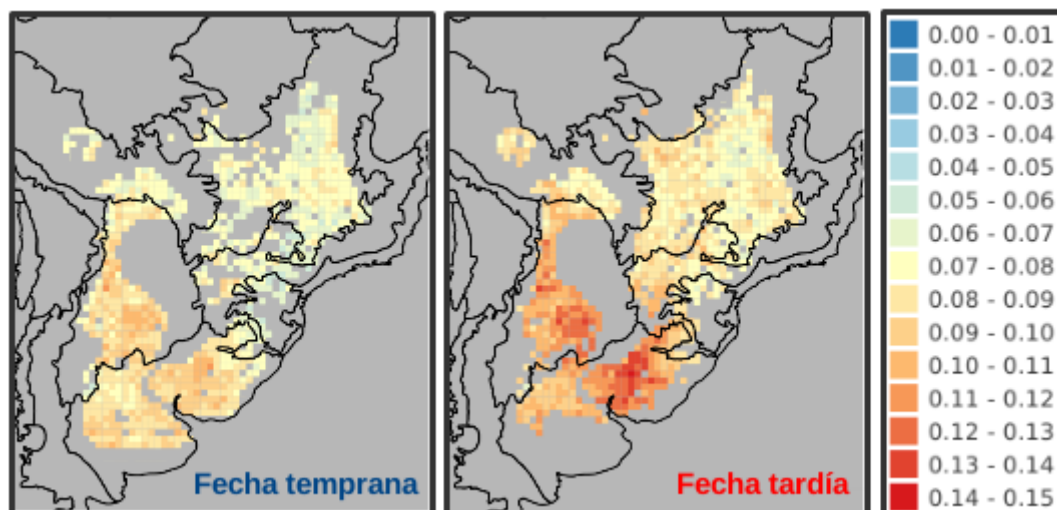


Figura 3.11: Coeficiente de variación promedio (CV) entre campañas (2000/01 a 2012/13; $n=13$) para las variables ambientales sintéticas que explicaron la interacción genotipo x sitio-fecha de siembra (GxS-FS). Se muestra el CV para fechas de siembra temprana y tardía (Fig. 2.5). Se incluyeron las variables ambientales sintéticas de los 8 subconjuntos completos. En gris, se muestran las áreas excluidas del análisis y las áreas donde no se calculó el CV porque el ambiente escapó al rango ambiental explorado por los genotipos en la mayoría de los sitios-fecha de siembra. Líneas negras indican límites de unidades biogeográficas propuestos por Morrone (2014) (Fig. 3.8, mapa 9; Tabla 2.2).

III. Mapas de mega-ambientes en distintos escenarios temporales (mapeo dinámico)

La distribución espacial de los mega-ambientes cambió con la campaña y la fecha de siembra (Fig. 3.12). Los mayores cambios se observaron hacia el Sur, aunque persistió el patrón Noroeste-Sudeste mencionado previamente. Como era de prever a partir de las Fig. 3.10 y 3.11, las diferencias fueron mayores entre fechas de siembra que entre campañas (Fig. 3.12). Además, la interacción entre fechas de siembra y campañas hizo que las diferencias entre fechas dependieran en gran medida de la campaña considerada y viceversa.

Al cambiar de fecha temprana a tardía, cambió el gradiente de heterogeneidad espacial Norte-Sur y el área de estudio en general mostró una tendencia hacia mega-ambientes favorables para GM largos. Por un lado, la heterogeneidad espacial disminuyó en el Sur y aumentó en el Norte del área de estudio, revirtiendo el gradiente de incremento de la heterogeneidad de Norte a Sur mencionado previamente para fecha temprana (Fig. 3.12). La mayor parte de la provincia Pampeana mostró 5 a 7 sectores discontinuos (con mega-ambientes distintos) en fecha temprana y solo 4 en tardía, con gran parte del área abarcada por un único mega-ambiente mayoritario, mientras que el sector desde el Bosque Paranaense hasta la provincia de Cerrado mostró 2 a 4 sectores discontinuos en temprana, gran parte con un único mega-ambiente mayoritario, y 5 a 6 en tardía. Este cambio en el gradiente de heterogeneidad Norte-Sur fue más notorio para campañas neutras y secas. Por el otro lado, la tendencia general hacia mega-ambientes favorables para GM largos en fecha tardía se compuso de una fuerte reducción en la superficie favorable para los

GM cortos hacia el Sur, en la provincia Pampeana, que se tornó una zona de baja respuesta frente al GM pero con una tendencia favorable para los GM largos, acompañada de una respuesta más pronunciada a favor de los GM largos hacia el Norte, en el sector desde el Bosque Paranaense hasta la provincia de Cerrado (Fig. 3.12). La reducción de la superficie favorable para los GM cortos en el Sur fue muy evidente en campañas secas a neutras, mientras que la respuesta más pronunciada a favor de los GM largos en el Norte fue más evidente en campañas secas y húmedas.

La ubicación de los mega-ambientes también se modificó entre campañas. Al aumentar las precipitaciones desde campañas secas hasta húmedas, el Norte y el Sur del área de estudio tendieron hacia mega-ambientes más favorables para GM largos y cortos, respectivamente (Fig. 3.12). En el Sur, el sector desde la provincia Pampeana hasta el Bosque de Araucaria en general fue más favorable para GM cortos, aumentando la superficie con mega-ambientes moderada a fuertemente favorables para GM cortos. Esta diferencia entre campañas fue más evidente en fecha temprana, ya que aproximadamente la mitad del área pasó de un mega-ambiente de baja respuesta frente al GM en campañas secas a mega-ambientes favorables para GM cortos en las húmedas. En el Norte, el sector desde el Bosque Paranaense hasta provincia de Cerrado, mostró una mayor superficie con mega-ambientes fuertemente favorables para GM largos al aumentar las precipitaciones, especialmente en fechas tardías.

El efecto de la fecha de siembra dependió de la campaña y viceversa. Gran parte de esta interacción entre fecha de siembra y campaña se asoció con el comportamiento particular de los mega-ambientes en el escenario campaña neutra-fecha tardía. Respecto de los patrones entre fechas de siembra, el cambio en el gradiente de heterogeneidad espacial Norte-Sur, dado por el incremento de la heterogeneidad en el Norte y su disminución en el Sur para fechas tardías, fue mayor para campañas neutras y secas que para húmedas. Al cambiar de fecha temprana a tardía, en campañas húmedas y secas, el sector desde el Sudoeste de la provincia de Cerrado hasta el centro-Norte del Bosque Paranaense mostró mega-ambientes más favorables para GM largos, mientras que no hubo diferencias marcadas entre fechas en campañas neutras (Fig. 3.12). En cuanto a los patrones entre campañas, en fecha temprana el sector que incluye la provincia de Cerrado y el Bosque Paranaense mostró mega-ambientes más favorables para GM largos al aumentar la precipitación, mientras que en fecha tardía campañas neutras mostraron mega-ambientes menos favorables para GM largos que las secas y húmedas, especialmente al Este del sector. Finalmente, la respuesta más favorable para GM cortos desde la provincia Pampeana hasta el Bosque de Araucaria al aumentar las precipitaciones fue más evidente en fecha temprana, ya que en tardía la mayor parte de este sector mostró una tendencia ligeramente más favorable para GM largos en campañas neutras que en secas (Fig. 3.12).

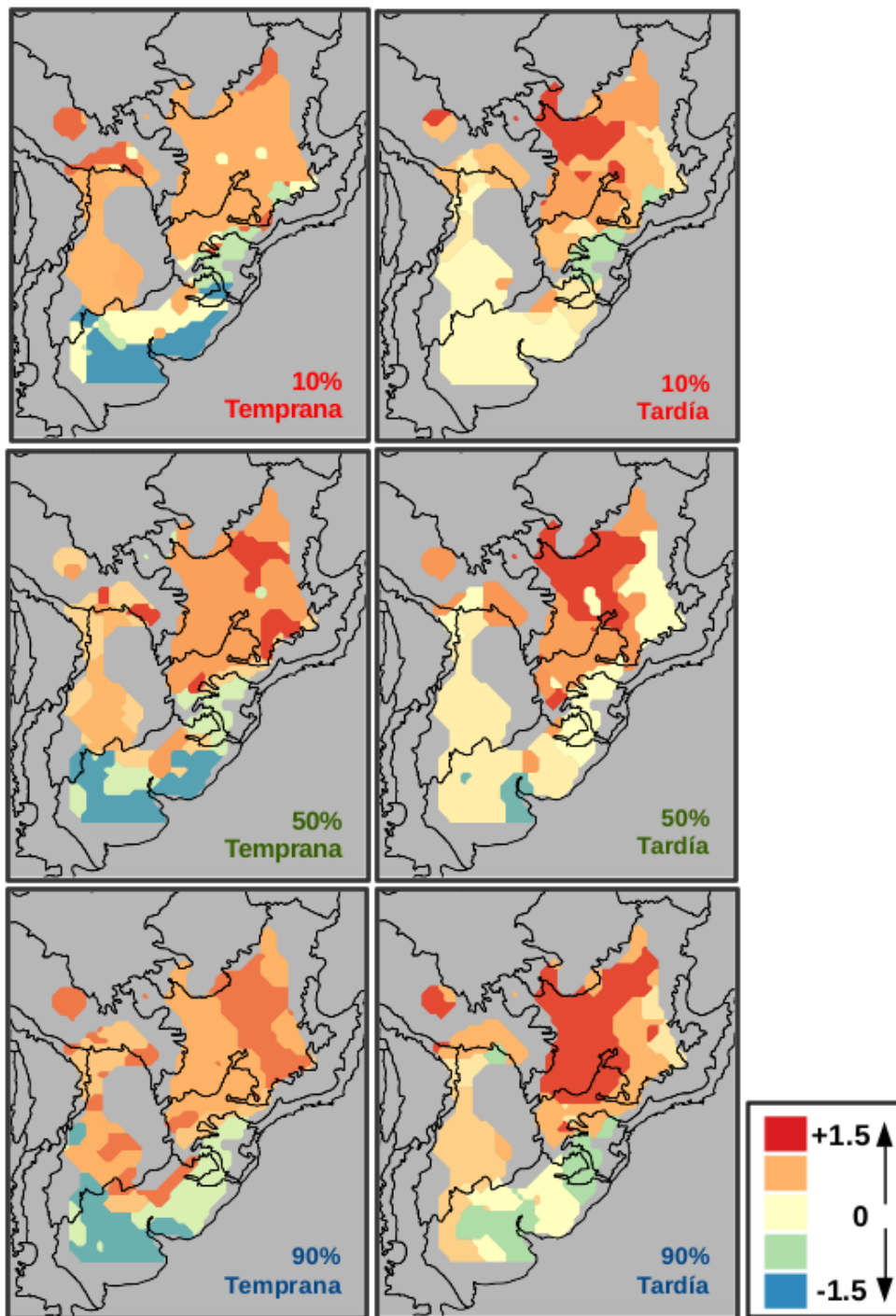


Figura 3.12: Integración de las zonas de baja interacción GxA (mega-ambientes) identificadas para los 8 subconjuntos completos, para seis escenarios que surgen de combinar campañas secas, neutras y húmedas (cuantiles 10, 50 y 90% de precipitación anual) con dos fechas de siembra (Fig. 2.5). Cada mapa fue construido a partir de un análisis de clúster que incluyó los mapas de los 8 subconjuntos completos para un escenario de campaña y fecha de siembra específico. Con un mismo color se representan zonas que en la mayoría de los subconjuntos han sido mapeadas con un mismo mega-ambiente o mega-ambientes parecidos. El color indica la pendiente entre rendimiento y GM dentro de cada zona, en promedio para los 8 subconjuntos. Se aplicó una transformación ranking sobre los valores de pendiente originales. Valores positivos indican mega-ambientes favorables para GM largos y negativos para GM cortos. Valores cercanos a cero indican mega-ambientes de baja respuesta frente al GM. En gris, se muestran las áreas excluidas del análisis. Líneas negras indican límites de unidades biogeográficas propuestos por Morrone (2014) (Fig. 3.8, mapa 9; Tabla 2.2).

CAPÍTULO 4: Discusión

4.A. Interacción genotipo x sitio-fecha de siembra y mega-ambientes

Las fuentes de variación de ANOVA para rendimiento dentro de los subconjuntos completos originales mostraron valores comunes para la escala de trabajo adoptada: el factor ambiental sitio-fecha de siembra (Sitio-FS) explicó la mayor parte de la variación (62%), seguido por la interacción genotipo x sitio-fecha de siembra (GxS-FS; 23%) y por el factor genotípico (5%) (Fig. 3.1). Trabajos previos muestran que entre el 60 y el 90% de la variación del rendimiento se debe al factor ambiental y el resto a la interacción GxA y en menor proporción al factor genotípico, tanto en soja (Martignone et al. 2008; Bacigaluppo et al. 2010; Yokomizo et al. 2013) como en otros cultivos (de la Vega y Chapman 2001; Arkbarpour y Dehghani 2014). Dentro de los mega-ambientes la interacción se redujo a la mitad, alcanzando un valor de 12%, y el factor genotípico se multiplicó por 3,5, alcanzando un valor de 18%, mayor a la interacción (Fig. 3.1), tal como se ha visto en trabajos anteriores (de la Vega y Chapman 2001; Dardanelli et al. 2006). Esto significa que los mega-ambientes capturaron el comportamiento diferencial de los genotipos frente al ambiente. Cada mega-ambiente reunió sitios-fecha de siembra con rankings de rendimiento de los genotipos relativamente homogéneos y a su vez contrastantes con otros mega-ambientes. El factor ambiental se mantuvo en valores cercanos al 60% tanto en los subconjuntos completos originales como dentro de los mega-ambientes.

Dentro de la mayoría de los mega-ambientes, el rendimiento de los genotipos se asoció significativamente con el GM, en algunos de manera positiva (mega-ambientes favorables para los GM largos) y en otros negativa (favorables para los cortos). Los mega-ambientes minoritarios (10 sobre un total de 32) que no mostraron una relación significativa entre el rendimiento y el GM (Fig. 3.2), pertenecían generalmente a subconjuntos con rangos de GM acotado y mostraron un mayor efecto del hábito determinado sobre el rendimiento que los mega-ambientes mayoritarios (Fig. 3.3). Trabajos previos también estudiaron la interacción entre genotipos y ambientes, representados estos últimos por sitios, campañas y/o fechas de siembra, pero no analizaron cómo varía la interacción al tomar rangos de GM distintos. No obstante, en concordancia con esta tesis, los trabajos que abarcaron rangos de GM amplios mostraron interacciones importantes (Sánchez et al. 2009) y definieron mega-ambientes (Dardanelli et al. 2006), mientras que los que abarcaron rangos de GM más acotados mostraron interacciones despreciables (Fehr et al. 2003; Sudaric et al. 2006).

Los resultados de esta tesis indican que a escala regional los mega-ambientes se definen principalmente por el GM. Es sabido que la adaptabilidad del cultivo está asociada con el GM (Remussi y Pascale 1977) que define si el ciclo de cultivo será suficientemente largo para garantizar una cantidad de biomasa vegetativa que permita una cierta captura de recursos pero suficientemente corto como para ajustarse a la estación térmica e hídricamente apta (Egli et al. 1993; 2011). Asimismo, se ha establecido que, en términos generales y dentro de una determinada zona, los genotipos de ciclo más corto tienden a

expresar un mayor potencial en ambientes hídrica y edáficamente favorables, mientras que los más largos tienen mejor comportamiento en ambientes más restrictivos (Baigorri 2008; Kantolic 2014). Esta tesis ha puesto en evidencia que estas características, generalmente identificadas a escala local, son también prioritarias en un análisis regional.

La magnitud de la interacción varió con el GM. La interacción entre GM menores a VI (GM cortos-medios) fue menor que entre GM mayores a VI (largos), mientras que la interacción entre GM mayores y menores a VI fue máxima (Fig. 3.4 A y B). Coincidentemente, algunos trabajos que estudiaron la interacción entre genotipos y ambientes representados por sitios y campañas también mostraron baja interacción entre GM cortos-medios (Fehr et al. 2003; Sudaric et al. 2006), y alta entre GM largos (Cucolotto et al. 2007). Sin embargo, en contraposición a lo observado en esta tesis, otros trabajos mostraron importantes interacciones dentro de rangos acotados para los GM cortos-medios (Arslanoglu y Aytac 2010; Salmeron et al. 2014). Los resultados de esta tesis para subconjuntos con distintos rangos de GM y para comparaciones entre genotipos de GM similares o disímiles, sugieren que los cambios de ranking más importantes en el rendimiento de los genotipos ocurren dentro de los GM largos y entre los GM largos y los cortos-medios, mientras que dentro de los GM cortos-medios el ranking se mantiene más estable. Esto implica que la zonificación por mega-ambientes para la soja cobra especial relevancia al evaluar genotipos de GM largos, y rangos amplios que abarcan desde GM cortos hasta largos.

En términos generales, el hábito determinado afectó negativamente al rendimiento, particularmente en los mega-ambientes en los que el rendimiento no estuvo asociado significativamente con el GM (Fig. 3.3). Trabajos previos a escala local mostraron efectos del hábito de crecimiento detrás de la respuesta diferencial de los genotipos al ambiente (Kilgore-Norquest y Sneller 2000; Enrico et al. 2013). En Oliveros, Argentina, Enrico et al. (2013) para dos genotipos de GM VI con distinto hábito de crecimiento, en dos campañas y 4 fechas de siembra, mostraron que el genotipo indeterminado rindió más que el determinado en promedio, aunque en algunas combinaciones de campaña y fecha de siembra se observó lo opuesto. Para la región de Tohoku, Japón, Kato et al. (2015) mostraron que el hábito indeterminado produjo un incremento en el rendimiento de los genotipos de ciclo más corto en los ambientes de Akita pero no en Kyoto. El período vegetativo más prolongado y la mayor acumulación de biomasa en post-floración en los genotipos de hábito indeterminado confieren ventajas en ambientes que presentan restricciones al crecimiento vegetativo, ya sea por períodos de estrés transitorios o por acortamiento del ciclo del cultivo (Slafer et al. 2015). En efecto, el período de formación de vainas más extendido en el tiempo y la distribución de las vainas en un mayor número de nudos (Egli y Bruening 2006) otorgan a los genotipos indeterminados una mayor capacidad de establecer granos, principalmente en ambientes restrictivos (Nico et al. 2015; 2016). Nuestros resultados muestran que en la región y campañas exploradas el hábito determinado rindió menos que el indeterminado para un mismo GM (Fig. 3.3). Sin embargo, la interacción entre el hábito y el ambiente permitió discriminar mega-ambientes con distintos grados de limitación para el hábito determinado.

Algunos genotipos mostraron patrones de interacción genotipo x sitio-fecha de siembra particulares que permitieron identificar mega-ambientes más allá del GM y el hábito de crecimiento. Los genotipos TOB 7800 y DM 8473 RSF, ambos de GM VIII y con hábito de crecimiento indeterminado, mostraron un comportamiento disímil frente a los mega-ambientes, lo cual se vio reflejado en la fuerte interacción entre ellos (Fig. 3.5). Dentro de los mega-ambientes favorables para los GM largos, se diferenciaron algunos que resultaron desfavorables para TOB 7800, mientras que todos fueron favorables para DM 8473 RSF, lo cual sugiere una mayor estabilidad para este último genotipo (Annicchiarico 2002). Los genotipos NA 4910 y DM 5.9i, ambos de GM V y con hábito indeterminado también mostraron comportamientos disímiles entre sí. Estas diferencias de comportamiento a nivel genotipo permitieron discriminar dentro de los mega-ambientes favorables para GM cortos entre unos que resultaron favorables para NA 4910 y otros que, por el contrario, resultaron favorables para DM 5.9, lo que sugiere que existe interacción con cross-over (Annicchiarico 2002). Si bien, como se mencionó anteriormente, las respuestas diferenciales de los genotipos al ambiente fueron explicadas en muchos trabajos casi exclusivamente por el GM (Dardanelli et al. 2006; Meotti et al. 2012; Salmeron et al. 2014), es frecuente encontrar respuestas diferenciales dentro de un mismo GM y hábito de crecimiento (Rosbaco 2004 et al.; Sudaric et al. 2006; Arslanoglu y Aytac 2010) y, en algunos casos, se han identificado mega-ambientes (Asfaw et al. 2009; Luquez et al. 2010). Los resultados de esta tesis a escala regional revelan que, aun para un mismo GM-hábito, los genotipos tienen comportamientos disímiles frente al ambiente que deben ser considerados al definir los mega-ambientes.

4.B. Caracterización ambiental de los mega-ambientes

La interacción genotipo x sitio-fecha de siembra se asoció con las variables ambientales, lo cual permitió definir a los mega-ambientes en términos de rangos ambientales. Las variables que explicaron la mayor parte de la interacción describían aspectos del fotoperíodo y del suelo y, en muchos casos, estaban relacionadas con las etapas reproductivas tardías. Entre las variables persistentes, las más importantes fueron el fotoperíodo durante todo el ciclo del cultivo y particularmente en post-floración, la precipitación en floración, fructificación y llenado de granos, la temperatura nocturna durante los estadios vegetativos, la profundidad efectiva del perfil y la capacidad de almacenamiento hídrico (Fig. 3.6a y 3.7). Varios autores analizaron la interacción entre sitios, campañas y/o fechas de siembra (Rao et al. 2002; Yan y Rajcan 2002; Dardanelli et al. 2006; Lúquez et al. 2010; Branquinho et al. 2014) pero muy pocos definieron los mega-ambientes a partir de variables ambientales (Dardanelli et al. 2006; Hernández-Segundo et al. 2009). En contraste con esta tesis, donde el fotoperíodo fue la variable más importante (Fig. 3.7), Hernández-Segundo et al. (2009) definieron mega-ambientes de cebada a nivel mundial que se diferenciaron por la temperatura y la precipitación durante la estación de crecimiento. Los autores emplearon variables climáticas, sin variación inter-anual, y un único período temporal para toda la estación de crecimiento, sin discriminar entre fases del cultivo. Con un nivel de detalle temporal similar pero a escala de país, Dardanelli et al. (2006) no identificaron variables climáticas que explicaran el comportamiento diferencial de los GM entre los mega-ambientes definidos para contenido de aceite y proteína de soja. Con mayor detalle temporal, considerando la variación entre años y entre fases del cultivo, otros autores encontraron relaciones entre las variables ambientales y el comportamiento diferencial de genotipos de soja entre

ambientes representados por campañas en la localidad de Oliveros (Martignone et al. 2008). A diferencia de esta tesis, donde la temperatura fue importante en los estadios vegetativos, el fotoperíodo durante todo el ciclo y la precipitación en los estadios reproductivos, para Martignone et al. (2008) las variables más asociadas con la interacción fueron temperatura y radiación durante todo el ciclo y fotoperíodo en estadios reproductivos, seguidas por la precipitación durante todo el ciclo. En resumen, ya sea por tratarse de distintos cultivos (Hernández-Segundo et al. 2009), variables respuesta, escalas espaciales o niveles de detalle temporal (Dardanelli et al. 2006; Martignone et al. 2008), la fuerte asociación del fotoperíodo con la interacción GxA en soja a escala regional encontrada en el presente trabajo no había sido demostrada anteriormente.

Las variables persistentes que definieron los mega-ambientes en este estudio son también definitorias del rendimiento a escala de lote, por su influencia en el desarrollo y el crecimiento del cultivo. Respecto del desarrollo, el fotoperíodo es el principal regulador de la duración del ciclo y de las fases del cultivo. El GM está asociado a la sensibilidad fotoperiódica, siendo los GM largos más sensibles que los cortos (Cober et al. 2001; Boote et al. 2003), lo cual se manifiesta en que su distribución geográfica sigue un gradiente latitudinal (Baigorri y Giorda 1997). La temperatura también controla la duración de las fases, principalmente antes de la floración (Setiyono et al. 2007; Bastidas et al. 2008), condicionando el tamaño y la duración del área foliar. Respecto del crecimiento, entre las variables persistentes surgieron algunas condicionantes de la disponibilidad hídrica a partir de los 55 días desde la siembra y, particularmente, a partir de los 90 días desde la siembra, períodos donde ocurren las etapas más críticas del cultivo, que incluyen la fructificación y parte del llenado de granos (Egli y Yu 1991; Egli 2010). En estas etapas, la tasa de crecimiento, que depende en gran medida de la disponibilidad hídrica, es la principal definitoria del número de granos y del rendimiento del cultivo (Board y Tan 1995; Debruin y Pedersen 2009). Los resultados obtenidos sugieren que los mega-ambientes discriminarían a los genotipos según (i) la duración del período vegetativo, que debe ser suficientemente largo como para capturar los recursos disponibles y lograr altas tasas de crecimiento en el período crítico (Egli 2011), (ii) la duración del período crítico, asociado a las condiciones fotoperiódicas que se exploran (Kantolic et al. 2001; 2005), que define el rendimiento potencial (Nico et al. 2015) y (iii) la tasa de crecimiento durante el período crítico, asociada a las precipitaciones durante ese período y a la capacidad de almacenamiento hídrico del suelo, que define el rendimiento logrado.

Esta tesis es el trabajo que más exhaustivamente ha traducido en términos de variables ambientales a los mega-ambientes identificados por la respuesta de los genotipos. Definir los mega-ambientes por rangos de variables ambientales abrió la posibilidad de localizar los mega-ambientes y analizar su variación espacial ante distintos escenarios de fecha de siembra y campaña (mapeo dinámico). Los resultados obtenidos sugieren que, a escala regional, la mayor parte de la variación en el ranking de rendimiento de los genotipos entre mega-ambientes se debe a la respuesta diferencial de los GM al fotoperíodo. En consecuencia, la distribución espacial de los mega-ambientes guardó una fuerte relación con el gradiente latitudinal de los GM en respuesta al fotoperíodo, como se discutirá en la sección siguiente.

4.C. Mapeo dinámico de mega-ambientes

Hasta el momento, se han generado zonificaciones por diversos criterios agroecológicos pero no existen mapas que surjan directamente de asignar una referencia cartográfica a la interacción GxA. Además, los trabajos anteriores consideran modelos estancos en el tiempo que no contemplan la variación de los ambientes y sus límites geográficos entre campañas y fechas de siembra. Por ejemplo, para maíz se mapearon mega-ambientes a escala mundial por un modelo basado exclusivamente en variables ambientales (climáticas y edáficas) pero sin considerar la respuesta diferencial de los genotipos al ambiente (Hartkamp et al. 2001). Para la soja, dos trabajos pre-existentes generaron zonificaciones a escala de país. Para Argentina, se definieron tipos macroagroclimáticos representados por rangos de temperatura, disponibilidad hídrica y fotoperíodo con distinto grado de limitación para el cultivo de soja, que una vez mapeados resultaron en 20 zonas (Pascale y Damario 2004). Los índices bioclimáticos y los umbrales que definieron los tipos macroagroclimáticos surgieron de revisión bibliográfica y de la interpretación subjetiva de la información generada por una red de experimentos a campo con 59 sitios y siete años de evaluación para 40 genotipos exóticos. Para Brasil, Kaster y Farias (2012) generaron un mapa con cinco macrorregiones para las cuales recomendaron distintos GM, que fueron subdivididas en 20 regiones edafoclimáticas de importancia para la evaluación y la selección de genotipos. Su trabajo consistió en recopilar información de distintas fuentes, como una zonificación macroagroecológica, una clasificación climática, recomendaciones técnicas, evaluaciones de genotipos y consultas a referentes. Los factores considerados para delimitar las macrorregiones fueron la latitud y sus efectos sobre el fotoperíodo y la temperatura, y el régimen de precipitaciones. Para delimitar las regiones edafoclimáticas se consideraron altitud y tipo de suelo. A diferencia de estos antecedentes, esta tesis es el primer trabajo basado en datos de rendimiento de genotipos que propone una metodología objetiva y repetible para mapear mega-ambientes a partir de su respuesta diferencial al ambiente. También es el primero en analizar la interacción GxA a escala mayor que la de país y en generar mapas dinámicos, que reflejan la variación de los mega-ambientes ante distintas campañas y fechas de siembra.

A escala regional, los mega-ambientes mostraron similitudes y diferencias respecto de los ambientes mapeados a escalas más reducidas en trabajos anteriores, tanto en la distribución espacial de las unidades mapeadas como en el nivel de heterogeneidad. En cuanto a la distribución espacial, a escala regional, los mega-ambientes mostraron un gradiente Noroeste-Sudeste, con mega-ambientes favorables para los GM largos hacia el Noroeste y favorables para los GM cortos hacia el Sudeste (Fig. 3.8 y 3.9), coincidentemente con trabajos a escalas menores que mostraron gradientes principalmente latitudinales (Pascale y Damario 2004; Kaster y Farias 2012). Las variables ambientales más importantes para caracterizar a los mega-ambientes describieron aspectos del fotoperíodo (Fig. 3.7), estrechamente relacionado con el GM y con la latitud (Boote et al. 2003). Esta relación estrecha entre GM, fotoperíodo y latitud ha dado lugar a zonificaciones para la soja en ocasiones anteriores (American Soybean Association 1964-69; INTA 1997). En trabajos a escala de país, se definieron cinco macrorregiones para Brasil y cinco regiones fotoperiódicas para Argentina, en ambos casos con un patrón de distribución de Norte a Sur, con ambientes favorables para los GM largos al Norte y cortos al Sur (Pascale y Damario 2004; Kaster y Farias 2012). Sin embargo, el GM para obtener un ciclo de cultivo intermedio en una zona específica,

mostró algunas incongruencias entre países: para Brasil fue desde IX corto en el Norte de la provincia de Cerrado hasta VI corto en el Este de la Pampeana, mientras que para Argentina fue desde IX en el centro de la provincia Chaqueña hasta IV en el centro-Sur de la Pampeana. En tanto, a una misma latitud el GM óptimo fue IX en el centro-Oeste de la Provincia Chaqueña (Argentina; Pascale y Damario 2004) y VII en el Oeste del Bosque Paranaense (Brasil; Kaster y Farias 2012). En esta tesis, al analizar la región sudamericana en forma conjunta, el patrón Noroeste-Sudeste mostró que los GM largos (hasta IX) rindieron más que los cortos hacia el Oeste de la provincia Chaqueña y el Norte de Cerrado, mientras que los GM cortos (hasta V o hasta II, dependiendo del subconjunto) rindieron más hacia el Sur y el Sudeste de la provincia Pampeana. La componente longitudinal de este patrón espacial, no comunicada hasta el momento, determinó que ambientes distintos se ubicaran en una misma latitud, y que un rango amplio de latitudes mostrara ambientes similares. Por ej., en una misma latitud de 27°S se encontraron mega-ambiente favorables para los GM cortos desde el Noreste de la provincia Pampeana hasta el Bosque de Araucaria, y mega-ambientes favorables para GM largos en la provincia Chaqueña (Fig. 3.9, pendientes -0.36 y +0.66). En tanto, una zona extensa que abarca desde el Norte de Cerrado hasta el Norte de provincia Pampeana mostró mega-ambientes moderadamente favorables para los GM largos y muy desfavorables para el hábito de crecimiento determinado en toda su extensión (Fig. 3.9, pendiente +0.82).

La heterogeneidad espacial a la escala regional en este estudio aumentó de Norte a Sur, con menor cantidad de mega-ambientes más extensos hacia el Norte y mayor cantidad y menos extensos hacia el Sur. Generalmente, la provincia de Cerrado fue cubierta por uno o dos mega-ambientes, mientras que la provincia Pampeana contuvo tres o más mega-ambientes (Fig. 3.8 y 3.9). En este sentido, a escala de país, Kaster y Farias (2012) mostraron una leve tendencia definida por macrorregiones más extendidas en el Norte de Brasil y menos en el Sur. En su trabajo, la provincia de Cerrado fue cubierta en su mayor parte por dos de las cinco macrorregiones, en concordancia con esta tesis que mostró generalmente uno o dos mega-ambientes (Fig. 3.8 y 3.9). Sin embargo, para el sector Noreste de la provincia Pampeana abarcado por Brasil, también incluyeron dos macrorregiones (Kaster y Farias 2012), mientras que esta tesis mostró generalmente tres o cuatro mega-ambientes. Para Argentina, el gradiente de heterogeneidad espacial Norte-Sur de los mega-ambientes se contrapuso con la heterogeneidad de los tipos macroagroclimáticos, que aumentó principalmente en el sentido de Este a Oeste (Pascale y Damario 2004). Los resultados de esta tesis indican que la eficiencia de una red de evaluación de cultivares para capturar las respuestas diferenciales frente al ambiente depende de la distribución de sus sitios experimentales en la dirección latitudinal más que en la longitudinal, y que la mayor densidad de sitios debería corresponderse con la zona centro-Sur de la región, donde la heterogeneidad espacial es máxima (Fig. 3.9).

Si bien los mapas de mega-ambientes de los diferentes subconjuntos completos mostraron patrones espaciales en común, hubo también diferencias entre ellos (Fig. 3.8). Por definición, la interacción entre genotipos y ambientes depende de ambos factores (de la Vega y Chapman 2001; Dardanelli et al. 2006). Ya se ha descripto cómo distintos genotipos generaron patrones distintos de interacción y que el GM explicó una parte importante de la interacción (Fig. 3.2 a 3.5). Los subconjuntos completos difirieron en cuanto a los genotipos y los ambientes (representados por los sitios-fecha de siembra). Esto implicó que los mega-ambientes fueran diferentes entre subconjuntos, se asociaran

con variables ambientales distintas (variables no persistentes; Fig. 3.6b) y se localizaran en distintas zonas, dependiendo del subconjunto. De hecho, a medida que los subconjuntos incluyeron rangos de GM más acotados, los mega-ambientes respondieron en menor medida al GM y en mayor medida al hábito de crecimiento (Fig. 3.3). Para rangos amplios, el patrón más relevante fue el definido por mega-ambientes favorables para los GM largos al Noreste y cortos al Sudeste (Fig. 3.8, mapas 1 y 2), mientras que para rangos acotados las principales diferencias entre mega-ambientes se debieron al hábito de crecimiento (mapa 7), que definió dos zonas de Este a Oeste: una donde el impacto negativo del hábito determinado fue mayor, desde la provincia de Cerrado hasta el Sudeste de la provincia Chaqueña, y otra donde el impacto fue menor, sobre el centro y el Oeste de la última. Rangos de intermedios a acotados mostraron tanto el gradiente Noroeste-Sudeste por GM como las zonas Oeste-Este definidas por el hábito de crecimiento determinado (mapas 4-7). Los resultados obtenidos ponen de manifiesto que los mega-ambientes a escala regional dependen del GM y, en menor medida, del hábito de crecimiento de los genotipos considerados. Sin embargo, cuando este último fue importante, en todos los casos las ventajas fueron a favor del hábito indeterminado, sugiriendo una mayor adaptabilidad de los genotipos con este atributo.

La ubicación geográfica de los mega-ambientes varió considerablemente entre fechas de siembra, y en menor medida entre campañas, dando origen a distintos mapas de mega-ambientes dependiendo del escenario temporal. Esta variación en la localización de los mega-ambientes respondió a la variación ambiental (Fig. 3.9 a 3.11) y fue mayor hacia el Sur que hacia el Norte (Fig. 3.12). Al cambiar de fecha temprana a tardía, cambió el gradiente de heterogeneidad espacial Norte-Sur y el área de estudio en general mostró una tendencia hacia mega-ambientes favorables para GM largos. Al aumentar las precipitaciones desde campañas secas hasta húmedas, el Norte y el Sur del área de estudio tendieron hacia mega-ambientes más favorables para GM largos y cortos, respectivamente. En Midsouth, EEUU, en latitudes equivalentes a las de la provincia Pampeana, para los GM III a VII en dos campañas, las fechas tardías resultaron favorables para GM cortos (III-IV) frente a otros largos (IV-V) (Salmerón et al. 2014), al contrario que en esta tesis. Sobre el límite Sudoeste del Bosque de Araucaria, para GM V largo a VI largo, los largos mostraron mayor productividad que los cortos ante atrasos en la fecha de siembra (Meotti et al. 2012), como ocurrió en esta tesis en campañas neutras. Experimentos conducidos en la provincia Pampeana mostraron diversos resultados. En San Justo, Córdoba, sobre el límite Noroeste de la provincia Pampeana, para los GM IV a VII, las fechas de siembra tempranas fueron consistentemente favorables para los GM cortos, mientras que en fechas tardías los GM no se diferenciaron (Cortés 2015). En Oliveros, Enrico et al. (2013), analizando dos campañas y cuatro fechas de siembra, encontró que la respuesta diferencial entre GM cortos (III-IV) y largos (V-VI) dependió de la campaña y no de la fecha de siembra, mientras que para Martignone et al. (2010), con dos campañas contrastantes en disponibilidad hídrica, la respuesta de los GM fue similar en ambas campañas, con mayor rendimiento para los GM IV-V respecto de los III y VII. Cabe aclarar que en esta tesis sólo un subconjunto incluyó genotipos de los GM II a IV. Actualmente, la mayor parte de la provincia Pampeana se siembra con genotipos del GM IV y, en menor medida, con GM menores, estando los GM V limitados a escenarios de baja productividad, principalmente en el límite norte de la región (Kantolic 2014). Futuros estudios, que incluyan una mayor proporción de GM inferiores a V permitirían comprender en mayor medida el comportamiento de los diferentes GM en estos mega-ambientes y en las diferentes fechas de siembra y campañas. De todas formas,

y pesar de algunas inconsistencias entre los antecedentes mencionados, todos ellos muestran una evidente interacción genotipo x fecha de siembra y genotipo x campaña, frecuentemente asociada a los GM, que se contrapone a los mega-ambientes (Yan y Rajcan 2002; Dardanelli et al. 2006) y a las zonificaciones estancos en el tiempo (Pascale y Damario 2004; Kaster y Farias 2012).

Los resultados de esta tesis son novedosos en mostrar cómo al cambiar de fecha de siembra y campaña, los mega-ambientes se localizan en distintas zonas y, en consecuencia, cómo algunas zonas con mega-ambientes similares en un escenario temporal pasan a tener mega-ambientes contrastantes en otro. En cuanto a la localización de los mega-ambientes por ej., en campañas neutras a húmedas, los mega-ambientes muy favorables para GM largos se extendieron desde el centro-Norte hacia el Sudeste de la provincia de Cerrado en fechas tempranas, mientras que en fechas tardías se extendieron desde el centro-Norte hacia el Sudoeste. En tanto, los ambientes muy favorables para GM cortos abarcaron áreas considerables desde el Sudeste hasta el Oeste de la provincia Pampeana en fechas tempranas, mientras que en fechas tardías prácticamente desaparecieron. Como ejemplo de las zonas que pasaron de mega-ambientes similares a contrastantes, en fecha tardía, especialmente campañas secas y neutras, la mayor parte de la provincia Pampeana y la Chaqueña mostraron en conjunto mega-ambientes de baja respuesta frente al GM. En cambio, en fecha temprana estas provincias mostraron mega-ambientes contrastantes: la provincia Pampeana, favorables para GM cortos y la Chaqueña, para GM largos (Fig. 3.12).

Como era esperable de esta variación de los mega-ambientes entre escenarios temporales, algunas coincidencias entre las zonificaciones previas y el mapa de mega-ambientes en fecha de siembra temprana y campaña neutra desaparecieron al considerar otros escenarios. Para la provincia Pampeana, Pascale y Damario (2004) definieron los tipos macroagroclimáticos BDB y BDC (clima templado-frío, subhúmedo-húmedo, con fotoperíodo mediano a largo) donde recomendaron los GM IV y V, en concordancia con los mapas de fecha temprana de esta tesis, que mostraron mega-ambientes muy favorables para los mismos GM (cortos). Por el contrario, esta tesis indicó que en fecha tardía la misma zona mostró mayoritariamente mega-ambientes favorables para GM largos (Fig. 3.12). Para el Oeste de la provincia de Cerrado, Kaster y Farias (2012) definieron la macrorregión 4, de régimen megatérmico con invierno seco, donde los GM VII y VIII tienen un largo de ciclo intermedio, de acuerdo con los mapas de campañas neutra y húmeda en fecha temprana que mostraron mega-ambientes moderadamente favorables para los mismos GM (Fig. 3.12). En fecha tardía esta zona también mostró mega-ambientes favorable para esos GM pero con una mayor respuesta en rendimiento. Todos estos resultados ponen de relieve la importancia de la fecha de siembra y las precipitaciones pronosticadas para la campaña en el diseño y el análisis de los resultados de las redes experimentales que abarcan muchos sitios y ambientes. Los resultados de experimentos a campo en una zona para un escenario temporal, podrían emplearse para la toma de decisiones en otra zona para otro escenario, aun cuando ambas zonas nunca fueran mapeadas con un mismo mega-ambiente simultáneamente.

Resumiendo los antecedentes aquí revisados, surge que por un lado muchos trabajos definieron mega-ambientes como conjuntos de sitios y fechas de siembra a partir del comportamiento diferencial de los genotipos. Por otro lado, otros trabajos mapearon zonificaciones a partir del comportamiento del cultivo en general. Esta tesis contribuye una metodología original que hizo posible asociar los mega-ambientes definidos a partir del comportamiento de los genotipos con variables ambientales de manera objetiva y mapear áreas de mínima interacción GxA para rendimiento (zonas con rankings de genotipos relativamente estables). Al menos tres mensajes clave para técnicos e investigadores surgen de este análisis: (i) si bien la campaña es importante, la fecha de siembra tiene mayor efecto sobre la respuesta diferencial, y por lo tanto sobre el ranking, de los genotipos. Ajustar la elección del genotipo ante un atraso en la fecha de siembra podría compensar e incluso sobre-compensar el efecto adverso de una mala campaña. (ii) Ya que este mapa surge del análisis conjunto de toda la región sudamericana, las empresas mejoradoras, generalmente presentes en más de un país, deberían considerarlo para distribuir sus programas, estaciones y sitios experimentales de manera más eficiente, ahorrando costos entre zonas geográficas que representan un mismo ambiente e identificando aquellas zonas que realmente merecen una atención particular, por representar ambientes distintos. (iii) La metodología aquí puesta en práctica, repetible y de bajo costo, basada en información de rendimiento de genotipos e información ambiental de uso libre, podría ser ajustada para mapear a escalas espacial y genética más detalladas que permitirían predecir los ambientes donde un genotipo candidato a lanzamiento podría destacarse respecto de sus competidores (por ejemplo, dentro de un mismo GM).

CAPÍTULO 5: Conclusiones generales

El objetivo general de esta tesis fue desarrollar un modelo espacial y dinámico de la interacción GxA en soja y generar una herramienta para definir y localizar ambientes en Sudamérica. El modelo desarrollado permitió definir los mega-ambientes por rangos de variables ambientales concretas, pasando de conjuntos de sitios-fechas de siembra (de la Vega y Chapman 2001; Yan y Rajcan 2002; Dardanelli et al. 2006; Sánchez et al. 2009; Hernández-Segundo et al. 2009) a conjuntos definidos por variables ambientales. La herramienta de mapeo generada permitió localizar los mega-ambientes en distintos escenarios temporales, considerando variaciones en la fecha de siembra y en el régimen de precipitaciones de la campaña. Además, la misma herramienta se encuentra preparada para mapear los mega-ambientes en cualquier otro escenario temporal distinto de los explorados en esta tesis. A continuación, se detallan las principales conclusiones en referencia a los dos objetivos específicos que fueron necesarios para cumplir con el objetivo general.

(a) identificar características climáticas y edáficas que permitan definir ambientes de comportamiento homogéneo para los genotipos en el área de estudio

- Los mega-ambientes fueron definidos principalmente por la respuesta diferencial de los GM frente al ambiente, en menor medida por la respuesta al hábito de crecimiento y, finalmente, por el comportamiento particular de algunos genotipos.
- Las variables ambientales más importantes describieron aspectos del fotoperíodo y del suelo y están preferentemente relacionadas con las etapas reproductivas. Entre las variables persistentes, las más relevantes fueron las relacionadas con el desarrollo (el fotoperíodo durante todo el ciclo del cultivo y la temperatura nocturna durante los estadios vegetativos) y con la oferta hídrica (la precipitación durante los estadios reproductivos, la profundidad efectiva del perfil y la capacidad de almacenamiento hídrico).

(b) identificar áreas geográficas y campañas donde se encuentran esas características ambientales (mapeo dinámico)

- A escala regional, los mega-ambientes mostraron un gradiente Noroeste-Sudeste, con mega-ambientes favorables para los GM largos hacia el Noroeste y favorables para los GM cortos hacia el Sudeste.
- La heterogeneidad espacial aumentó de Norte a Sur, con menor cantidad de mega-ambientes más extensos hacia el Norte y mayor cantidad y menos extensos hacia el Sur.
- La localización de los mega-ambientes varió entre fechas de siembra y, en menor medida, entre campañas. Al cambiar de fecha temprana a tardía, cambió el gradiente de heterogeneidad espacial Norte-Sur y el área de estudio en general mostró una tendencia hacia mega-ambientes favorables para GM largos. Al aumentar las precipitaciones de campañas secas a húmedas, el Norte y el Sur del área de estudio tendieron hacia mega-ambientes más favorables para GM largos y cortos, respectivamente.

Bibliografía

- Akbarpour, O., Dehghani, H., Sorkhi, B. y Gauch, H.G. 2014. Evaluation of genotype x environment interaction in Barley (*Hordeum vulgare* L.) based on AMMI model using developed SAS program. Journal of Agricultural Science and Technology, 16: 909-920.
- Alliprandini, L. F., Abatti, C., Bertagnolli, P. F., Cavassim, J. E., Gabe, H. L. et al. 2009. Understanding soybean maturity groups in Brazil: environment, cultivar classification, and stability. Crop Science, 49: 801-808.
- Annicchiarico, P. 2002. Genotype x Environment interactions - Challenges and opportunities for plant breeding and cultivar recommendations. FAO Plant production and protection paper – 174, FAO, Rome.
- Arslanoglu, F. y Aytac, S. 2010. Determination of stability and genotype x environment interactions of some agronomic properties in the different soybean (*Glycine max.* (L) Merrill) cultivars. Bulgarian Journal of Agricultural Science, 16: 181-195.
- Asfaw, A., Alemayehu, F., Gurum, F. y Atnaf, M. 2009. AMMI and SREG GGE biplot analysis for matching varieties onto soybean production environments in Ethiopia. Scientific Research and Essay, 4: 1322-1330.
- Bacigaluppo, S., Bodrero, M., Andriani, J., De Emilio, M., Enrico, J. et al. 2008. Evaluación de cultivares de soja de los grupos de madurez III, IV, V y VI en siembras de primera época en diferentes ambientes del sur de Santa Fe. Campaña 2007/08. Para mejorar la producción, 39: 14-20.
- Bacigaluppo, S., Bodrero, M., Andriani, J., De Emilio, M., Enrico, J. et al. 2009. Evaluación de cultivares de soja de los grupos de madurez III, IV y V, en siembras de primera época en diferentes ambientes del sur de Santa Fe. Campaña 2008/09. Para mejorar la producción, 42: 7-13.
- Bacigaluppo, S., Bodrero, M., Albrecht, R.E., Andriani, J.M., Gentili, O.A. et al. 2010. Evaluación en Red de cultivares de soja de los grupos de madurez III, IV y V, en diferentes ambientes del centro-sur de Santa Fe, campaña 2009/10. Para mejorar la producción, 45: 7-12.

- Bacigaluppo, S., Bodrero, M., Balzarini, M., Gerster, G.R., Andriani, J.M. et al. 2011. Main edaphic and climatic variables explaining soybean yield in Argiudolls under no-tilled systems. *European Journal of Agronomy*, 35: 247-254.
- Badu-Apraku, B., Oyekunle, M., Obeng-Antwi, K., Osuman, A.S., Ado, S.G. et al. 2012. Performance of extra-early maize cultivars based on GGE biplot and AMMI analysis. *The Journal of Agricultural Science*, 150: 473-483.
- Baigorri, H. 2008. Elección de variedades. En: Satorre, E.H. (Coord.): *Producción de Soja*. AACREA, Buenos Aires, 33-43.
- Baigorri, H. y Giorda, L. 1997. *El cultivo de la soja en Argentina*. INTA. Buenos Aires. 448 pp.
- Barros, H. B., Sedyama, T., Teixeira, R. de C., Fidelis, R.R., Cruz, C.D. et al. 2010. Adaptabilidade e estabilidade de genótipos de soja avaliados no estado do Mato Grosso. *Ceres*, 57: 359-366.
- Bastidas, A.M., Setiyono, T.D., Dobermann, A., Cassman, K.G., Elmore, R.W. et al. 2008. Soybean sowing date: The vegetative, reproductive, and agronomic impacts. *Crop Science*, 48: 727-740.
- Batjes N. H. 2005. ISRIC-WISE global data set of derived soil properties on a 0.5 by 0.5 degree grid (Version 3.0). Report 2005/08, ISRIC – World Soil Information, Wageningen (with data set).
- Board, J.E. y Tan, Q. 1995. Assimilatory capacity effects on soybean yield components and pod number. *Crop Science*, 35: 846-851.
- Boote, K.J., Jones, J.W., Batchelor, W.D., Nafziger, E.D. y Myers, O. 2003. Genetic coefficients in the CROPGRO–Soybean Model: Links to field performance and genomics. *Agronomy Journal*, 95: 32-51.
- Branquinho, R.G., Duarte, J.B., De Souza, P.I.M., Da Silva Neto, S.P. y Pacheco, R.M. 2014. Estratificação ambiental e otimização de rede de ensaios de genótipos de soja no Cerrado. *Pesquisa Agropecuária Brasileira*, 49: 783-795.
- Braun, H. y Rajaram, S. 1996. CIMMYT's approach to breeding for wide adaptation. *Euphytica*, 92: 175-183.

- Carrera, C.S., Dardanelli, J. L. y Soldini, D.O. 2014. Genotypic and environmental variation in seed nutraceutical and industrial composition of non-transgenic soybean (*Glycine max*) genotypes. *Crop & Pasture Science*, 65(12): 1311-1322.
- Cavalcante, A.K., Hamawaki, O.T., Hamawaki, R.L., Sousa, L.B., Oliveira Nogueira, A.P. et al. 2014. Fentopic adaptability and stability in soybean genotypes in Porto Alegre Do Norte, State of Mato Grosso. *Bioscience Journal*, 30(4): 942-949.
- Cavassim, J.E., Bessalho Filho, J.C., Alliprandini, L.F., De Oliveira, R.A., Daros, E. et al. 2013. AMMI analysis to determine relative maturity groups for the classification of soybean geotypes. *Journal of Agronomy*, 12: 168-178.
- Cober, E.R., Stewart, D.W. y Voldeng, H.D. 2001. Photoperiod and temperature responses in early-maturing, near-isogenic soybean lines. *Crop Science*, 41: 721-727.
- Cortés, E.J. 2015. Desarrollo de estrategias de siembra del cultivo de soja [*Glycine max* (L.) Merrill] en el departamento San Justo, Córdoba, Argentina. Universidad Nacional de Córdoba.
- Cucolotto, M., Garbuglio, D., Pipolo, V., Da Silva Fonseca, N. y Kamikoga, M.K. 2007. Genotype x environment interaction in soybean: evaluation through three methodologies. *Crop Breeding and Applied Biotechnology*, 7: 270-277.
- Dardanelli, J.L., Balzarini, M., Martínez, M.J., Cuniberti, M., Resnik, S., et al. 2006. Soybean maturity groups, environments, and their Interaction define mega-environments for seed composition in Argentina. *Crop Science*, 46: 1939-1947.
- De Bruin J. L. y Pedersen, P. 2009. Growth, yield, and yield component changes among old and new soybean cultivars. *Agronomy Journal*, 101: 124–130.
- De Carvalho, C.G.P., Arias, C.A.A., De Toledo, J.F.F., De Almeida, L.A., De Souza Kiihl, R.A. et al. 2002. Interação genótipo x ambiente no desempenho produtivo da soja no Paraná. *Pesquisa Agropecuária Brasileira*, 37: 989–1000.
- De la Vega, A. y Chapman, S. 2001. Genotype by environment interaction and indirect selection for yield in sunflower: I. Two-mode pattern analysis of oil and biomass yield across environments. *Field Crops Research*, 72: 17-38.

- Dias, F.T.C., Bitombeira, J.B., Teófilo, E.M. y Barbosa, F. 2009. Adaptabilidade e estabilidade fenotípica para o caráter rendimento de grãos em cultivares de soja para o Estado do Ceará. *Revista Ciência Agronômica*, 40: 129-134.
- Egli, D.B. 1993. Cultivar maturity and potential yield of soybean. *Field Crops Research*, 32: 147-158.
- Egli D.B. 2010. Soybean reproductive sink size and short-term seductions in photosynthesis during flowering and pod set. *Crop Science*, 50: 1971–1977.
- Egli D.B. 2011. Time and the productivity of agronomic crops and cropping systems. *Agronomy Journal*, 103: 743–750.
- Egli, D.B. y Yu Z-W. 1991. Crop growth rate and seeds per unit area in soybeans. *Crop Science*, 31: 439-442.
- Egli, D.B. y Bruening, W.P. 2006. Temporal profiles of pod production and pod set in soybean. *European Journal of Agronomy*, 24: 11–18.
- Enrico, J., Conde, M., Martignone, R. y Bodrero, M. 2013. Soja: evaluación de la estabilidad del rendimiento según fechas de siembra. Para mejorar la producción, 50: 71-78.
- Fehr, W., Hoeck, J., Johnson, S. y Murphy, P. 2003. Genotype and environment influence on protein components of soybean. *Crop Science*, 43: 511-514.
- Frascarolo, D. 2008. Modelos zonales de producción en el Movimiento CREA. Región NOA. En: Satorre, E. Producción de Soja. 1ra ed. AACREA. 176 pp.
- Gauch, H.G. 1988. Model Selection and Validation for Yield Trials with Interaction. *Biometrics*, 44(3): 705–715.
- Gauch, H.G. y Zobel, R.W. 1996. Identifying mega-environments and targeting genotypes. *Crop Science*, 37: 311-326.
- Gedif, M. y Yigzaw, D. 2014. Genotype by environment interaction analysis for tuber yield of potato (*Solanum tuberosum* L.) using a GGE biplot method in Amhara region, Ethiopia. *Agricultural Sciences*, 5: 239-249.

- GRASS Development Team, 2016. Geographic Resources Analysis Support System (GRASS) Software, Version 7.0. Open Source Geospatial Foundation. Electronic document: <http://grass.osgeo.org>
- Hartkamp, A.D., White, J., Rodríguez, A. y Banziger, M. 2000. Maize production environments revisited: a GIS-based approach. 37 pp.
- Hernandez-Segundo, E., Capettini, F., Trethowan, R., Van Ginkel, M., Mejia, A. et al. 2009. Mega-environment identification for barley based on twenty-seven years of global grain yield data. *Crop Science*, 49: 1705-1718.
- Hüeck, K. y Seibert, P. Vegetationskarte von Sudamerika: Mapa de la vegetación de America del Sur. Fischer, G. Stuttgart, V. 1972. 71 pp.
- IICA. Instituto Interamericano de Cooperación para la Agricultura. Disponible en línea: <http://www.iica.org.py>. [consulta: 31 de diciembre de 2014].
- Jones E, Oliphant E y Peterson P. 2014. SciPy: Open Source Scientific Tools for Python. <http://www.scipy.org/> [consulta: 01 de diciembre de 2015].
- Kantolic, A.G. 2014. Criterios para la elección de variedades y fecha de siembra. En: Satorre, E.H. (Coord.) Soja: Claves para una producción rentable y sostenible. AACREA, 23-28.
- Kantolic, A.G. y Slafer, G.A. 2001. Photoperiod sensitivity after flowering and seed number determination in indeterminate soybean cultivars. *Field Crops Research*, 72: 109-118.
- Kantolic, A.G. y Slafer, G.A. 2005. Reproductive development and yield components in indeterminate soybean as affected by post-flowering photoperiod. *Field Crops Research*, 93: 212-222.
- Kang, M.S. 1998. Using genotype-by-environment interaction for crop cultivar development. *Advances in Agronomy*, 62.
- Kaster, M. y Farias, J.R.B. 2012. Regionalização dos testes de Valor de Cultivo e Uso e da indicação de cultivares de soja - Terceira aproximação. Londrina, PR: Embrapa Soja.

- Kato, S., Fujii, K., Yumoto, S., Ishimoto, M., Shiraiwa, T. et al. 2015. Seed yield and its components of indeterminate and determinate lines in recombinant inbred lines of soybean. *Breeding science*, 65(2): 154-60.
- Keisling, T.C. 1982. Calculation of the Length of Day. *Agronomy Journal*, 74: 758-759.
- Kilgore-Norquest, L. y Sneller, C.H. 2000. Effect of stem termination on soybean traits in southern U.S. production systems. *Crop Science*, 40: 83-90.
- Knight, R. 1970. The measurement and interpretation of genotype-environment interactions. *Euphytica*, 19(2): 225–235.
- Luquéz, J., Capurro, M. y Erazzú, L.E. 2010. Genotype-environment interactions, megaenvironments and winner genotypes and environments for soybean grain yield in Argentina. *Revista de la Facultad de Agronomía, La Plata*, 109: 31-36.
- Manlla, A. y Salvagiotti, F. 2012. Rendimiento potencial y brechas de producción en soja bajo condiciones de campo, en Oliveros, sur de Santa Fe. Para mejorar la producción, 48: 85-88.
- Martignone, R.A., Enrico, J.M., Bodrero, M.L. y Andriani, J.M. 2010. Factores asociados con la variabilidad de rendimientos entre grupos de madurez en soja. Para mejorar la producción, 45: 25-34.
- Martignone, R.A., Bodrero, M.L., Enrico, J.M., Quijano, A. y Andriani, J.M. 2008. Caracterización de ambientes y variabilidad de rendimientos de soja de primera época de siembra, en Oliveros (Santa Fe). Para mejorar la producción, 39: 48-54.
- Mathew, J.P., Herbert, S.J., Zhang, S., Rautenkranz, A.A. y Litchfield, G.V. 2000. Differential Response of Soybean Yield Components to the Timing of Light Enrichment. *Agronomy Journal*, 92: 1156–1161.
- Meotti, G., Benin, G. y Silva, R. 2012. Épocas de semeadura e desempenho agrônômico de cultivares de soja. *Pesquisa Agropecuária Brasileira*, 47: 14-21.
- Miranda, G.V., De Souza, L.V., Guimarães, L.J.M., Namorato, H., Oliveira, L.R. et al. 2009. Multivariate analyses of genotype x environment interaction of popcorn. *Pesquisa Agropecuária Brasileira*, 44: 45-50.

- Morrone, J. 2014. Biogeographical regionalisation of the Neotropical region. *Zootaxa*, 3782: 1-110.
- NASA Land Processes Distributed Active Archive Center (LP DAAC). MODIS MCD12C1. USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota, 2013.
- Nico, M., Miralles, D.J. y Kantolic, A.G. 2015. Post-flowering photoperiod and radiation interaction in soybean yield determination: Direct and indirect photoperiodic effects. *Field Crops Research*, 176: 45–55.
- Nico, M., Mantese, A.I., Miralles D.J. y Kantolic, A.G. 2016. Soybean fruit development and set at the node level under combined photoperiod and radiation conditions. *Journal of Experimental Botany*, 67: 365-377.
- Pacheco, R.M., Duarte, J.B., De Souza, P.I.M., Da Silva, S.A. y Nunes Junior, J. 2009. Key locations for soybean genotype assessment in Central Brazil. *Pesquisa Agropecuária Brasileira*, 44: 478-486.
- Palaver, V. y Fagundes, R. S. 2009. Comportamento de genótipos de soja em quatro ambientes no estado do Paraná. *Cascavel*, 4: 1-10.
- Pascale, A.J. y E.A. Damario. *Bioclimatología agrícola e agroclimatología*. Editorial Facultad Agronomía. Buenos Aires. 2004. 550 pp.
- Pedregosa, F, Varoquaux, G, Gramfort, A, Michel, V, Thirion, B. et al. 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12: 2825-2830.
- Pizzio, M. y F. Pacín. 2008. Modelos zonales de producción en el Movimiento CREA. Región CREA Sudoeste. En: Satorre, E. *Producción de Soja*. 1ra ed. AACREA. 176 pp.
- Polizel, A., Juliatti, F. y Hamawaki, O. 2013. Phenotypical adaptability and stability of soybean genotypes in the state of Mato Grosso. *Bioscience Journal*, 29: 910-920.
- Primomo, V.S., Falk, D.E., Ablett, G.R., Tanner, J.W. y Rajcan, I. 2002. Genotype x environment interactions, stability, and agronomic performance of soybean with altered fatty acid profiles. *Crop science*, 42: 37-44.

- R Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Remussi C. y Pascale, A.J. 1977. La soja. Cultivo, mejoramiento, comercialización, usos. Enciclopedia Argentina de Agricultura y Jardinería, Tomo 2, fascículo 12-4, Segunda Edición. Buenos Aires. 108 pp.
- Rao, M.S.S., Mullinix, B.G., Rangappa, M., Cebert, E., Bhagsari, A.S. et al. 2002. Genotype x environment interactions and yield stability of food-grade soybean genotypes. *Agronomy Journal*, 94: 72-80.
- Rocha, M.D.M. y Vello, N.A. 1999. Interação genótipos e locais para rendimento de grãos de linhagens de soja com diferentes ciclos de maturação. *Bragantia*, 58: 69-81.
- Rocha, M.D.M., Vello, N.A., De Almeida Lopes, Â.C. y Maia, M.C.C. 2009. Estabilidade e adaptabilidade produtiva em linhagens de soja de ciclo médio. *Ceres*, 56: 764-771.
- Rosbaco, I., Bisaro, V. y Romagnoli, M. 2004. Métodos estadísticos para la caracterización de ambientes climáticos y evaluación de respuestas de cultivares de soja. *Revista de la Facultad de Ciencias Agrarias - UNR*, 6: 27-40.
- Salmeron, M., Gbur, E.E., Bourland, F.M., Buehring, N.W., Earnest, L. et al. 2014. Soybean maturity group choices for early and late plantings in the Midsouth. *Agronomy Journal*, 106: 1893-1901.
- Salomonson, V., Barnes, W., Maymon, P.W., Montgomery, H.E. y Ostrow, H. 1989. MODIS: Advanced facility instrument for studies of the Earth as a system. *IEEE Transactions on Geoscience and Remote Sensing*, 27: 145-153.
- Sánchez, J.R., Devani, M.R., Ledesma, F., Pérez, D.R., Gamboa, D.E. et al. 2009. Análisis del comportamiento de cultivares comerciales de soja en el noroeste argentino. *Revista industrial y agrícola de Tucumán*, 86: 17-22.
- SIIA. Sistema Integrado de Información Agropecuaria. Disponible en línea: <http://www.sii.gov.ar>. [consulta: 31 de diciembre de 2014].

- Sionit, N. y Kramer, P. J. 1977. Effect of Water Stress During Different Stages of Growth of Soybean. *Agronomy Journal*, 69: 274–278.
- Slafer, G. a, Kantolic, A.G., Appendino, M.L., Tranquilli, G., Miralles, D.J. et al. 2015. Genetic and environmental effects on crop development determining adaptation and yield. En: Sadras, V.O. y Calderini, D.F (Eds.): *Crop physiology applications for genetic improvement and agronomy*, 2nd ed, Elsevier, Amsterdam, 285-319.
- Su, F., Hong, Y. y Lettenmaier, D. P. 2008. Evaluation of TRMM Multisatellite Precipitation Analysis (TMPA) and its utility in hydrologic prediction in the La Plata Basin. *Journal of Hydrometeorology*, 9: 622-640.
- Sudaric, A., Simic, D. y Vrataric, M. 2006. Characterization of genotype by environment interactions in soybean breeding programmes of southeast Europe. *Plant breeding*, 125: 191-194.
- Thiessen, A.J. y Alter J.C. 1911. Precipitation Averages for Large Areas. *Monthly Weather Review*, 39: 1082-1084.
- Ward, J.H. 1963. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58, 236-244.
- Yan, W. y Rajcan, I. 2002. Biplot analysis of test sites and trait relations of soybean in Ontario. *Crop Science*, 42: 11-20.
- Yan, W. y Rajcan, I. 2003. Prediction of Cultivar Performance Based on Single-versus Multiple-Year Tests in Soybean. *Crop Science*, 43: 549–555.
- Yan, W., Hunt, L., Sheng, Q. y Szlavnics, Z. 2000. Cultivar evaluation and mega-environment investigation based on the GGE biplot. *Crop Science*, 40: 597-605.
- Yan, W., Pageau, D., Frégeau-Reid, J., Lajeunesse, J., Goulet, J. et al. 2011. Oat mega-environments and test-locations in Quebec. *Canadian Journal of Plant Science*, 91: 643-649.
- Yokomizo, G.K.-I., Duarte, J.B., Vello, N.A. y Unfried, J.R. 2013. Análise AMMI da produtividade de grãos em linhagens de soja selecionadas para resistência à ferrugem asiática. *Pesquisa Agropecuária Brasileira*, 48: 1376-1384.

Apéndice

A.I. Variables ambientales meteorológicas y de índice de vegetación

Tabla A.1 (Primera parte)

Variables ambientales meteorológicas y de índice de vegetación.

Variable	Inicio	Fin	Fase fenológica	Estadístico
Fotoperíodo	0	55	S – R1	máximo
Fotoperíodo	0	55	S – R1	mínimo
Fotoperíodo	0	55	S – R1	promedio
Fotoperíodo	0	145	S – R8	máximo
Fotoperíodo	0	145	S – R8	mínimo
Fotoperíodo	0	145	S – R8	promedio
Fotoperíodo	55	145	R1 – R8	máximo
Fotoperíodo	55	145	R1 – R8	mínimo
Fotoperíodo	55	145	R1 – R8	promedio
Fotoperíodo	90	145	R5 – R8	máximo
Fotoperíodo	90	145	R5 – R8	mínimo
Fotoperíodo	90	145	R5 – R8	promedio
IVN	0	55	S – R1	máximo
IVN	0	55	S – R1	mínimo
IVN	0	55	S – R1	promedio
IVN	0	145	S – R8	máximo
IVN	0	145	S – R8	mínimo
IVN	0	145	S – R8	promedio
IVN	55	145	R1 – R8	máximo
IVN	55	145	R1 – R8	mínimo
IVN	55	145	R1 – R8	promedio
IVN	90	145	R5 – R8	máximo
IVN	90	145	R5 – R8	mínimo
IVN	90	145	R5 – R8	promedio

Se detallan: la variable, el período dentro del ciclo del cultivo en días desde la fecha de siembra (Inicio, Fin), una fase fenológica (Fehr y Kaviness, 1977) que podría coincidir con ese período, y el estadístico de resumen temporal para los datos incluidos dentro de ese período. Números negativos de Inicio indican días antes de la siembra.

Tabla A.1 (Segunda parte)

Variables ambientales meteorológicas y de índice de vegetación.

Variable	Inicio	Fin	Fase fenológica	Estadístico
Precipitación	-150	150	S – R8 *	desvío estándar
Precipitación	-150	150	S – R8 *	máximo
Precipitación	-150	150	S – R8 *	suma acumulada
Precipitación	-30	145	S – R8 *	desvío estándar
Precipitación	-30	145	S – R8 *	máximo
Precipitación	-30	145	S – R8 *	suma acumulada
Precipitación	0	55	S – R1	desvío estándar
Precipitación	0	55	S – R1	máximo
Precipitación	0	55	S – R1	suma acumulada
Precipitación	0	145	S – R8	desvío estándar
Precipitación	0	145	S – R8	máximo
Precipitación	0	145	S – R8	suma acumulada
Precipitación	55	145	R1 – R8	desvío estándar
Precipitación	55	145	R1 – R8	máximo
Precipitación	55	145	R1 – R8	suma acumulada
Precipitación	90	145	R5 – R8	desvío estándar
Precipitación	90	145	R5 – R8	Máximo
Precipitación	90	145	R5 – R8	suma acumulada
Temp. diurna	0	55	S – R1	Máximo
Temp. diurna	0	55	S – R1	Mínimo
Temp. diurna	0	55	S – R1	Promedio
Temp. diurna	0	55	S – R1	suma acumulada
Temp. diurna	0	145	S – R8	Máximo
Temp. diurna	0	145	S – R8	Mínimo
Temp. diurna	0	145	S – R8	Promedio
Temp. diurna	0	145	S – R8	suma acumulada
Temp. diurna	55	145	R1 – R8	Máximo
Temp. diurna	55	145	R1 – R8	Mínimo
Temp. diurna	55	145	R1 – R8	Promedio
Temp. diurna	55	145	R1 – R8	suma acumulada
Temp. diurna	90	145	R5 – R8	Máximo
Temp. diurna	90	145	R5 – R8	Mínimo
Temp. diurna	90	145	R5 – R8	Promedio
Temp. diurna	90	145	R5 – R8	suma acumulada

* Incluye un período de barbecho previo a la siembra.

Tabla A.1 (Tercera parte)

Variables ambientales meteorológicas y de índice de vegetación.

Variable	Inicio	Fin	Fase fenológica	Estadístico
Temp. media	0	55	S – R1	Máximo
Temp. media	0	55	S – R1	Mínimo
Temp. media	0	55	S – R1	Promedio
Temp. media	0	55	S – R1	suma acumulada
Temp. media	0	145	S – R8	Máximo
Temp. media	0	145	S – R8	Mínimo
Temp. media	0	145	S – R8	Promedio
Temp. media	0	145	S – R8	suma acumulada
Temp. media	55	145	R1 – R8	Máximo
Temp. media	55	145	R1 – R8	Mínimo
Temp. media	55	145	R1 – R8	Promedio
Temp. media	55	145	R1 – R8	suma acumulada
Temp. media	90	145	R5 – R8	Máximo
Temp. media	90	145	R5 – R8	Mínimo
Temp. media	90	145	R5 – R8	Promedio
Temp. media	90	145	R5 – R8	suma acumulada
Temp. nocturna	0	55	S – R1	Máximo
Temp. nocturna	0	55	S – R1	Mínimo
Temp. nocturna	0	55	S – R1	Promedio
Temp. nocturna	0	55	S – R1	suma acumulada
Temp. nocturna	0	145	S – R8	Máximo
Temp. nocturna	0	145	S – R8	Mínimo
Temp. nocturna	0	145	S – R8	Promedio
Temp. nocturna	0	145	S – R8	suma acumulada
Temp. nocturna	55	145	R1 – R8	Máximo
Temp. nocturna	55	145	R1 – R8	Mínimo
Temp. nocturna	55	145	R1 – R8	Promedio
Temp. nocturna	55	145	R1 – R8	suma acumulada
Temp. nocturna	90	145	R5 – R8	Máximo
Temp. nocturna	90	145	R5 – R8	Mínimo
Temp. nocturna	90	145	R5 – R8	Promedio
Temp. nocturna	90	145	R5 – R8	suma acumulada

A.II. Rutina automatizada de búsqueda de subconjuntos completos

```
## Cargar modulos y librerias Python
import psycpg2 as pg
import numpy as np
import pandas as pd

def seleccionar_subconjuntos_completos(pdataf, umbral_completud, n_min_filas, n_min_columnas):
    """
    Identificar subconjuntos (submatrices) de datos GxA completos, para un umbral de completud
    de filas y columnas, y con una cantidad minima de filas y de columnas

    pdataf = pandas.DataFrame
            Rendimientos por genotipo (G) y ambiente (E)

    umbral_completud = numerico
            Umbral de completud de filas y columnas. Proporción de celdas con dato (presencias)
            dentro de cada fila y columna de la matriz 2D GxA. Se incluyen unicamente
            subconjuntos      cuyas filas y columnas superan este umbral de completud

    n_min_filas = numerico
            Cantidad minima de filas (E; ambientes). Subconjuntos con una cantidad de filas
            menor este umbral, se descartan

    n_min_columnas = numerico
            Cantidad minima de columnas (G; genotipos). Subconjuntos con una cantidad de
            columnas menor este umbral, se descartan
    """

    print '\n', 'rotar_filas_columnas(...)', umbral_completud, n_min_filas, n_min_columnas

    def rotaciones(matriz, umbral_completud=umbral_completud,
                    n_min_filas=n_min_filas, n_min_columnas=n_min_columnas):
        """
        Para cada fila de la matriz GxA de presencias/ausencias, con dimensiones n*m:

            (1) descarta las columnas con 0 (ausencia) para la fila actual,
            obteniendo una matriz intermedia n*m'

            (2) descarta las filas con una completud menor al umbral, considerando
            unicamente las columnas remanentes m' del paso anterior (1),
            obteniendo una nueva matriz n'*m'

            (3) evalua las exigencias de minimo de filas y columnas para la nueva
            matriz n'*m' obtenida en el paso anterior (2)
        """

        n_filas, n_columnas = np.shape(matriz)
        submatrices_idx = []
        for i in range(n_filas):

            # SELECCIONAR SUBINDICES DE COLUMNAS DONDE LA FILA   i='1':

            fila_con_id = np.array([range(n_columnas), matriz[i]])
            col_idx = fila_con_id[0, fila_con_id[1,:]==1]
            col_idx = np.array(col_idx).astype(int)
            #-----

            # SELECCIONAR SUBINDICES DE FILA
            #CON COMPLETUD > UMBRAL %:
            suma_cols = np.sum(matriz[:,col_idx],1)
```

```

porc_cols = (suma_cols*1.0) / len(col_idx)

porc_con_id = np.array([range(len(porc_cols)), porc_cols])
fil_idx = porc_con_id[0, porc_con_id[1,:]>umbral_completud]
fil_idx = np.array(fil_idx).astype(int)
#-----

# SELECCIONAR SUBINDICES DE COLUMNAS
# CON COMPLETUD > UMBRAL %:
mat = matriz[fil_idx, :]
mat = mat[:, col_idx]      # subset de la matriz original por filas y columnas
                           #preseleccionadas

suma_filas = np.sum(mat,0)
porc_filas = (suma_filas*1.0) / len(fil_idx)

porc_con_id = np.array([range(len(porc_filas)), porc_filas])
completas = porc_con_id[0, porc_con_id[1,:]>umbral_completud]
completas = np.array(completas).astype(int)
col_idx = np.array(col_idx[completas]).astype(int)
#-----

# Si cumple con umbrales minimos de filas E y columnas G:
if len(list(fil_idx)) >= n_min_filas:
    if len(list(col_idx)) >= n_min_columnas:

        if not (list(fil_idx), list(col_idx)) in submatrices_idx:

            submatrices_idx.append((list(fil_idx), list(col_idx)))
            # contiene un par de listas por cada submatriz completa:
            #      1. lista de los subindices de FILAS de la
            #      de presencias/ausencias
            #      2. ... COLUMNAS ...

return submatrices_idx

def eliminar_ambientes_cero(posiciones_numpy, pdataf):
    """
    Para cada subconjunto completo, eliminar subindices de filas (E; ambientes)
    para las cuales TODAS las columnas (G; genotipos) tienen rendimiento=0
    """

    # Matriz 2D, rendimiento (cuerpo) para cada combinacion de ambiente (filas) - variedad
    #(columnas):
    rend_2D      =      apaisar_stvar(pdataf=pdataf,      lista_fields=identificador_de_ensayo,
matriz
cuerpo='mean')

    resultado = []
    for i in range(len(posiciones_numpy)):

        sm = posiciones_numpy[i]

        # Filtrar filas E, con rend=0 para todas las columnas G:
        sm[0]  =  np.array(sm[0])[rend_2D.ix[sm[0],  sm[1]].values.sum(axis=1)  !=
0.0].tolist()

        resultado.append(sm)

    return resultado

def indices_pandas(posiciones_numpy, pdataf):
    """
    Transformar posiciones de filas y columnas en subindices de pandas
    (.index y .columns)
    """

```



```

        for i in range(len(posiciones_numpy)):

            sm = posiciones_numpy[i]

            # Seleccionar celdas GxA del subconjunto completo:
            select = pdataf.ix[sm[0], sm[1]]

            ix_trats = select.values.flatten()
            filas_E = select.index.values
            columnas_G = select.columns.values

            posiciones_numpy[i] = {
                'ix_trats': ix_trats, 'filas_E': filas_E, 'columnas_G': columnas_G
            }

        return posiciones_numpy

#####

# Formateo de datos para obtener una matriz 2D GxA, con filas ambiente (E) y columnas genotipo (G)
trats = apaisar_stvar(pdataf=pdataf, lista_fields=identificador_de_ensayo, cuerpo='ix_trat')

# Transformar a matriz de presencias/ausencias:
# (mismas filas y columnas que matriz original GxA, pero rellena con 1 y 0)
presaus = (trats / trats).values

# Búsqueda subconjuntos GxA completos
# -> posiciones de filas y columnas numpy

print '\n', 'búsqueda por FILAS'
busqueda_por_filas = rotaciones(matriz=presaus)

print '\n', 'búsqueda por COLUMNAS'
busqueda_por_columnas = rotaciones(matriz=presaus.transpose())
# Seleccionar solo subconjuntos nuevos:
busqueda_por_columnas = [i for i in busqueda_por_columnas if not (i[1], i[0]) in busqueda_por_filas]

# Merge busqueda_por_filas-busqueda_por_columnas
posiciones_fila_columna = np.array([np.array(busqueda_por_filas[:,0]).tolist() +
np.array(busqueda_por_columnas[:,1]).tolist(), np.array(busqueda_por_filas[:,1]).tolist() +
np.array(busqueda_por_columnas[:,0]).tolist() ] ).transpose().tolist()

# Eliminar filas (E; ambientes) con rend=0 para todas las columnas (G; genotipos)
# para cada subconjunto GxA completo:
posiciones_fila_columna = eliminar_ambientes_cero(posiciones_fila_columna, pdataf)

# Transformar posiciones numpy a subíndices pandas
# -> multi-index identificador_de_ensayo (.index); nombres de genotipos (.columns)
posiciones_pandas = indices_pandas(posiciones_fila_columna, trats)

return posiciones_pandas

def consulta_to_PD(conn_args, tabla, clausulas=""):
    """
    Consultar una tabla de la DB y almacenar el resultado en un variable
    del tipo pandas.DataFrame
    """

    conn = pg.connect(conn_args)
    cursor = conn.cursor()

    select = 'SELECT * FROM ' + tabla + ' ' + clausulas
    cursor.execute(select)

```

```

data = cursor.fetchall()

# EXTRAER NOMBRES DE COLUMNAS Y ASIGNAR UN SUBINDICE:
nombres = np.array(cursor.description)[:0]

return pd.DataFrame(data, columns=nombres)

def apaisar_stvar(pdataf, lista_fields, cuerpo):
    """
    Formteo de datos. Tranforma una matriz con formato
    dim1 | dim2 | ... dim n | valor
    al formato de doble entrada, con registros definidos por los campos
    [lista_fields] y columnas definidas por los campos [cols]
    """

    # Pivot table
    pdataf = pd.pivot_table(data=pdataf, rows=lista_fields, cols='stvar', values=cuerpo)

    # Ordenar por columnas de lista_fields (filas) y por variedad (columnas)
    pdataf.sortlevel(axis=0, ascending=True, inplace=True)
    pdataf.sort(axis=1, ascending=True, inplace=True)

    return pdataf

if __name__=='__main__':

    ## Parametros de conexion a la DB PostgreSQL
    conn_args = "host='localhost' dbname='GxA_v100' user='postgres' password='postgres'"

    ## Campos de identificacion de ensayos (ambientes; E)
    identificador_de_ensayo = ['area_dm', 'id_sitio', 'f_siembra', 'anyo', 'loc_num', 'tipo_ens', 'nro_ens',
    'cv_ens']

    #####

    # 1. Leer datos de rendimiento grano tratamiento desde DB
    # (promedio de repeticiones)
    d = consulta_to_PD(conn_args=conn_args, tabla='rendimiento_stats')

    # 2. Búsqueda iterativa de subconjuntos completos
    subconjuntos = seleccionar_subconjuntos_completos(d, umbral_completud=0.8, n_min_filas=5,
    n_min_columnas=5)

```