

CARACTERIZACIÓN REGIONAL Y MANEJO PARA AUMENTAR LA CONCENTRACIÓN DE PROTEÍNA EN SOJA

Lina Bosaz ^{1,*}, José A. Gerde ¹, Lucas Borrás ¹, Pablo Cipriotti ², Santiago Gallo ³, José L. Rotundo ¹

¹ Facultad de Ciencias Agrarias, Universidad Nacional de Rosario y CONICET

² Facultad de Agronomía, Universidad de Buenos Aires y CONICET.

³ Asociación Argentina de Consorcios Regionales de Experimentación Agrícola

* e-mail de contacto: bosaz@iicar-conicet.gob.ar

INTRODUCCIÓN

Argentina, es el tercer productor mundial de granos, y primer exportador de harinas de soja (FAO, 2017). A nivel industrial, el cultivo de soja tiene relevancia económica por ser una importante fuente de aceite y proteína. Desde 1980 a la actualidad se observó un aumento continuo en el rendimiento de las variedades liberadas al mercado, y una disminución de la concentración de proteína en grano (De Felipe, 2016). La industria Argentina ha tenido dificultades para producir harinas que cumplan con los estándares mínimos de comercialización (HI-PRO o alta proteína, 46.5%). Esto se atribuye al bajo tenor proteico de los granos de soja producidos.

Existen combinaciones de concentración de aceite y proteína en los granos que posibilitan lograr harinas con el mínimo de proteína establecido por la norma de comercialización. En campañas agrícolas recientes, dichas combinaciones sólo fueron logradas en situaciones puntuales (Fig. 1). Esto puede ser atribuido a varios factores. La concentración de proteína en grano depende de variables de manejo agronómico, por ejemplo, elección de cultivar y fecha de siembra, como así también, de variables ambientales como temperatura o disponibilidad hídrica (Mourtzinis et al, 2017; Rotundo et al., 2009).

El análisis de bases de datos provistas por los productores permite obtener información de alta calidad. Actualmente, las técnicas de minería de datos se aplican con mayor frecuencia en estudios de investigación (Liao et al., 2012). Una de las técnicas más reconocidas es el modelado de árboles de Clasificación y Regresión (Breiman et al., 1984). Este enfoque es apropiado para estudiar posibles variables determinantes de la concentración de proteína en soja según datos de productores a escala regional (De'Ath y Fabricius, 2000). Por otro lado, es posible analizar la variación en la composición de los granos de soja teniendo en cuenta patrones espaciales. Este tipo de análisis permite explorar la variabilidad a menor escala en una extensa zona productiva, como lo es la zona centro-norte Argentina (Rotundo et al., 2016).

Actualmente, se desconocen los límites precisos de los patrones zonales en Argentina en lo referente a concentración de proteína. Esto permitiría identificar zonas que consistentemente presenten altos tenores proteicos. Por otro lado, no se han utilizado datos de producción para determinar cómo el manejo particular de cada lote influye en la calidad del grano de soja. El objetivo del presente trabajo fue: (i) caracterizar zonas de producción de

Argentina según la concentración proteica de la soja allí producida e (ii) identificar variables de manejo que permitan aumentar la concentración de proteína en grano.

MATERIALES Y MÉTODOS

Sistemas bajo estudio

Los datos analizados corresponden a empresas socias de AACREA – Asociación Argentina de Consorcios Regionales de Experimentación Agrícola. Se tuvieron en cuenta tanto lotes de producción (n=1721), como también redes de ensayos que lleva adelante la institución (n=954). Las bases de datos comprenden observaciones entre las campañas 2012/13 y 2015/16 y fueron divididas en soja de primera y de segunda.

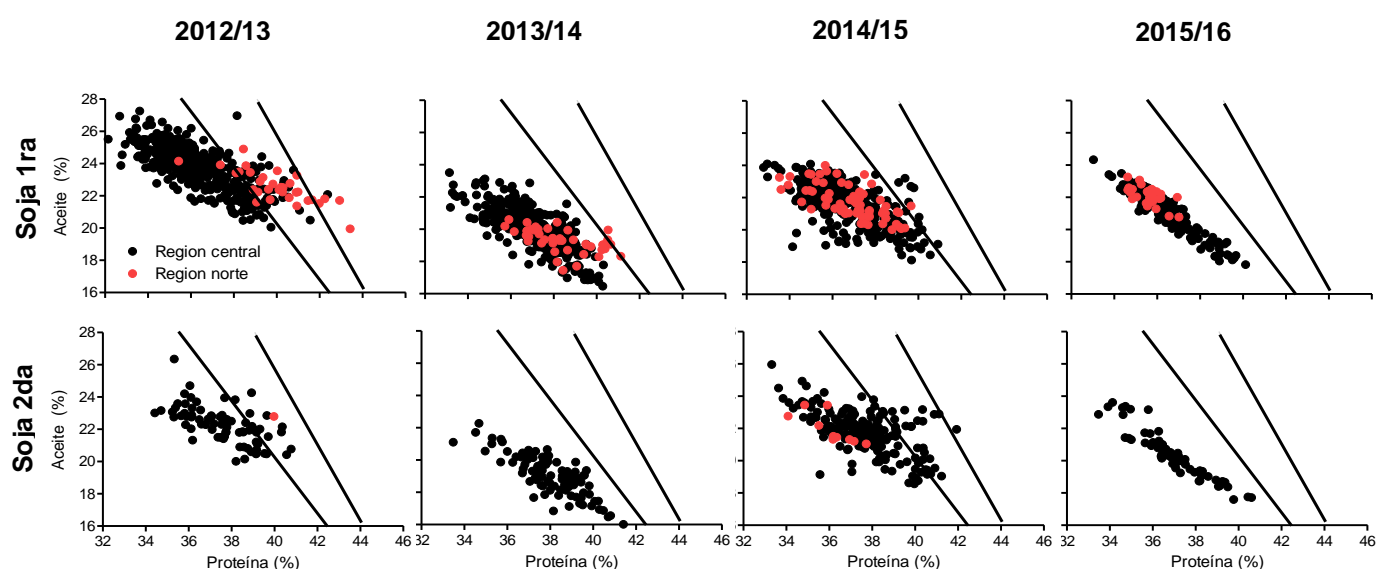


Fig. 1. Relación entre concentración de aceite y proteína en grano (% b.s.) para lotes de producción analizados. La zona delimitada entre las dos barras diagonales representa calidades de grano que posibilitan lograr harinas con el mínimo de proteína establecido por la norma de comercialización. Cada punto representa un lote de producción.

El procedimiento de toma de muestras de grano fue gestionado por la Facultad de Ciencias Agrarias de la UNR. En cada campaña, se enviaron a cada productor bolsas para el muestreo de granos a cosecha. Junto con la muestra, los productores proporcionaban datos de latitud y longitud del lote, rendimiento y el manejo agronómico realizado en el mismo. Cada muestra fue analizada mediante espectroscopia de infra-rojo cercano de reflectancia (NIR 5000, FOSS, Dinamarca), para estimar concentración de proteína y aceite en grano (% base seca). Dentro de las variables de manejo agronómico, se consideró genotipo, grupo de madurez, fecha de siembra, cultivo antecesor, distanciamiento entre hileras, inoculación, fertilización y aplicación de fungicida. Paralelamente, se estimaron los datos climáticos

utilizando datos de Estaciones Experimentales INTA. Se tuvieron en cuenta latitud, longitud, precipitación y radiación acumuladas, y temperaturas medias para cada mes del ciclo del cultivo (octubre-marzo) en cada una de las observaciones. El análisis se realizó en forma separada para lotes de producción y ensayos comparativos de rendimiento.

Lotes de producción

Se consideró un enfoque geostadístico para mapear la variabilidad en concentración de proteína de los lotes correspondientes a la zona bajo estudio. Se utilizó el método de interpolación por Krigging, que parte del supuesto que puntos próximos en el espacio tienden a tener valores de proteína en grano similares en comparación con puntos más distantes. Se realizaron variogramas que permiten explorar el grado de asociación espacial entre lotes para luego representar dichas asociaciones en un mapa de la zona bajo estudio.

Para detectar qué variables ambientales y de manejo influyen en la calidad de los granos, se utilizó el modelado mediante árboles de regresión. Los árboles de regresión son técnicas de partición recursiva que permiten ir separando la variabilidad total de los datos de en grupos más homogéneos.

Ensayos comparativos de rendimiento

Para desglosar el efecto del genotipo, del ambiente y su interacción sobre la concentración de proteína se optó por analizar los datos mediante modelos mixtos. A su vez, para predecir el verdadero efecto genotípico de cada cultivar en particular, independientemente de los efectos ambientales, el uso de los mejores predictores lineales no sesgados (BLUP, por su sigla en inglés) proporcionó una estimación adecuada del efecto cultivar en diferentes ambientes, lo cual permitió identificar genotipos que presentan consistentemente altos valores proteicos. Los BLUPs de cada cultivar se obtuvieron usando estimaciones en un modelo que incluyó ambiente, genotipo e interacción genotipo por ambiente.

RESULTADOS

Lotes de producción

El variograma indicó una fuerte auto-correlación espacial para concentración de proteína en soja de primera (Fig. 2a). No hubo auto-correlación espacial en soja de segunda (Fig. 2b). La variación espacial de concentración de proteína para soja de primera se predijo usando Krigging ordinario. El mapa predicho indicó zonas con baja concentración de proteína (<35%) y zonas bien definidas con altos valores de proteína en grano (>38%) (Fig. 3). Se pudo observar un claro gradiente norte-sur para la variable analizada. En cuanto a soja de segunda, al no haber encontrado un patrón espacial determinado (Fig. 2b), no fue posible realizar un mapa que refleje algún tipo de asociación entre concentración de proteína y ubicación del lote.

No encontrar asociación espacial no implica que no haya variabilidad en los valores de proteína entre localidades.

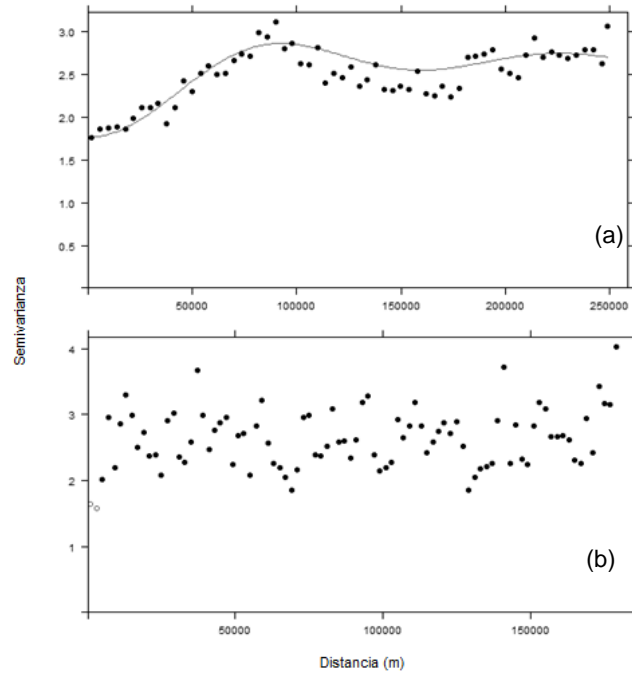


Fig. 2. Variogramas para soja de primera (a), soja de segunda (b) para describir la auto-correlación espacial local en concentración de proteína.

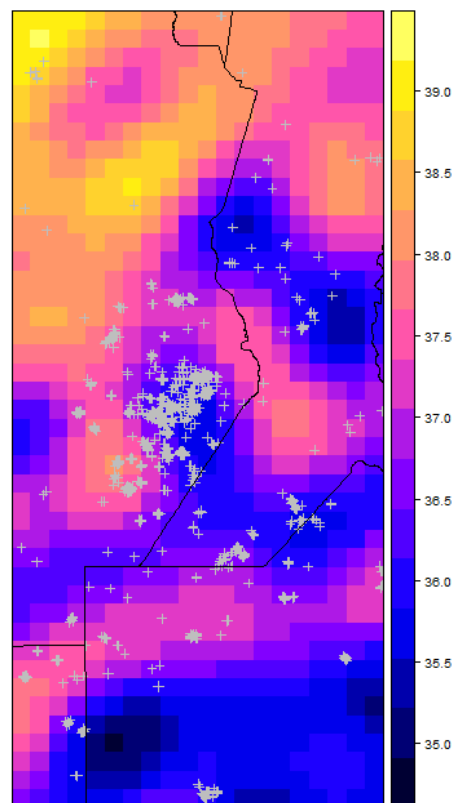


Fig. 3. Mapa de concentración de proteína para soja de primera. Cada cruz corresponde a un lote de producción. Los valores de concentración de proteína se refieren al valor medio para cada lote a través de las campañas analizadas.

Los árboles de regresión explicaron aproximadamente el 50% de la variabilidad total observada para soja de primera y soja de segunda. En la Fig. 4 se observa el árbol de regresión para soja de primera. El criterio de división primario fue genotipo, identificando dos grupos que difieren en un 2% en proteína en grano. Para genotipos con valores cercanos a 36% de proteína promedio, las condiciones térmicas durante el mes de diciembre fueron importantes; observando mayores valores proteicos ante mayores temperaturas en dicho mes. En este grupo de observaciones, grupos de madurez III mostraron menores valores respecto a grupos más largos. Por otro lado, genotipos con mayores valores de proteína (~38%), su magnitud está dada en base a la longitud del lote dónde son cultivados. En resumen, las variables de manejo explicaron un 73% de las variaciones observadas en concentración de proteína para soja de primera (Tabla 1).

En la Fig. 5 puede observarse el árbol de regresión para soja de segunda. El criterio de división primario fue nuevamente genotipo, diferenciando dos grupos (diferencia de aproximadamente 2% en proteína). Para genotipos con menores valores de proteína promedio, las condiciones ambientales durante el mes de enero fueron determinantes. En líneas generales, mayor disponibilidad hídrica y temperaturas medias disminuyen la concentración de proteína, observando mayores valores proteicos ante menores temperaturas en dicho mes. Contrariamente, ante escasas precipitaciones en enero, genotipos de alto potencial sembrados tardíamente (luego del 10 de diciembre) manifestaron altos valores de proteína en grano. Paralelamente, el valor protéico de los granos de ciertos genotipos dependieron de las precipitaciones totales de la campaña. Condiciones hídricas más restrictivas presentaron más proteína en grano que mejores ambientes. Finalmente, las variables de manejo explicaron un 71% de las variaciones observadas en concentración de proteína para soja de segunda (Tabla 1).

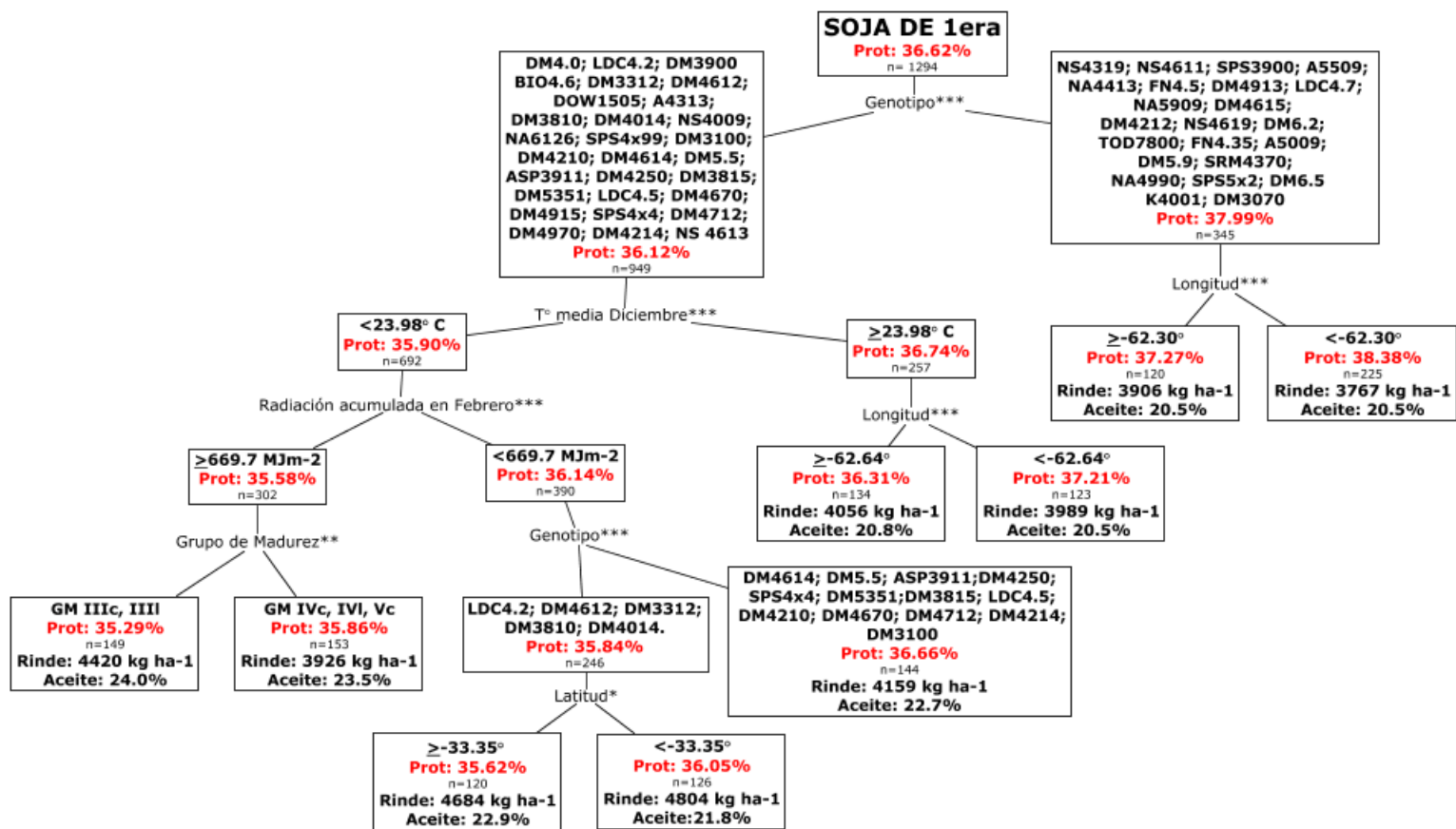


Fig. 4. Modelo de árbol de regresión para concentración de proteína (% b.s.) en soja de primera. *** $p < 0.001$. n, representa el número de datos en cada división. En nodos terminales además se presenta los valores de rendimiento promedio (kg ha^{-1}) y concentración de aceite (% b.s.).

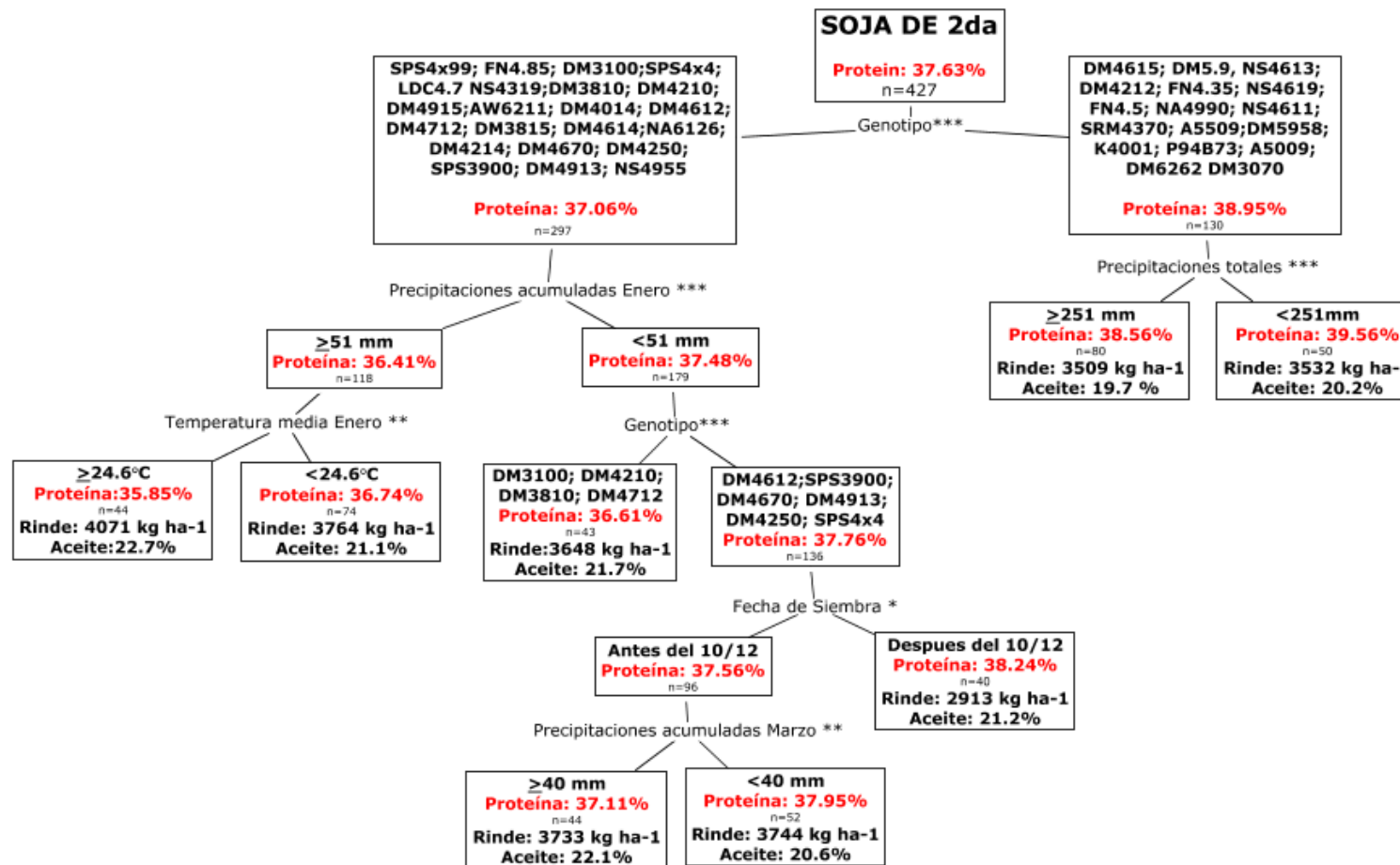


Fig. 5. Modelo de árbol de regresión para concentración de proteína (% b.s.) en soja de segunda. *** $p < 0.001$. n, representa el número de datos en cada división. En nodos terminales además se presenta los valores de rendimiento promedio (kg ha^{-1}) y concentración de aceite (% b.s.).

Tabla 1. Variabilidad explicada relativa (%) por las variables de manejo y ambiente asociadas a concentración de proteína en soja utilizando árboles de regresión. La variabilidad explicada relativa se refiere a la variabilidad explicada por cada variable respecto a la total explicada por el modelo para cada cultivo. Los árboles de regresión explicaron 50% de la variación total en los valores de proteína para soja de primera y segunda.

Variable	Soja de 1ra	Soja de 2da
Genotipo	71.5	68.9
Grupo de madurez	1.8	0.0
Fecha de Siembra	0.0	2.5
Σ Manejo	73.3	71.4
Temperatura media diciembre	10.3	0.0
Longitud	11.3	0.0
Radiación acumulada febrero	4.1	0.0
Latitud	1.0	0.0
Precipitación acumulada enero	0.0	15.5
Precipitación total acumulada	0.0	5.9
Temperatura media enero	0.0	4.0
Precipitación acumulada marzo	0.0	3.2
Σ Ambiente	26.7	28.6

Ensayos comparativos de rendimiento

El factor ambiental tuvo un fuerte peso relativo en la expresión diferencial de la concentración proteica. Es relevante observar que el genotipo presentó un mayor valor que la interacción genotipo x ambiente para porcentaje de proteína en grano (17.2 y 15.9%, para soja de primera y de segunda, respectivamente) (Tabla 2). Este resultado implica la existencia de genotipos que a través de ambientes presentan consistentemente un alto valor proteico (Tabla 3).

Tabla 2. Resultados análisis mediante modelos mixtos. Se presenta la variabilidad total observada para la concentración de proteína utilizando datos de ensayos comparativos de rendimiento. El porcentaje de la variación total representa la variabilidad explicada por cada uno de los factores.

Soja de primera	Porcentaje de la variación total	Soja de segunda	Porcentaje de la variación total
Ambiente (A)	48.7	Ambiente (A)	56.4
Genotipo (G)	17.2	Genotipo (G)	15.9
G x A	11.7	G x A	7.2
Residual	22.4	Residual	20.6
G / G x A	1.5	G / G x A	2.2

Los BLUPs permitieron identificar qué genotipos son los que se destacan a través de los ambientes respecto a la concentración de proteína en grano. En la Tabla 3 se puede observar un ranking de genotipos que mantienen los valores de proteína en la mayoría de los sitios de producción. Estos resultados surgieron a partir del aislamiento del efecto ambiental, pudiendo así separar el efecto genotípico *per se*, y determinar cuáles son las variedades que producen mayor proteína potencial.

Tabla 3. Ranking de genotipos, los tres primeros de altos valores proteicos, los tres centrales de valores medios y los tres últimos de valores bajos. Dichas concentraciones de proteína (% b.s.) son consistentes a través de los ambientes, y en algunos casos se repiten los mismos genotipos para los dos cultivos estudiados.

Soja de 1ra		Soja de 2da	
Genotipo	Proteína (% b.s.)	Genotipo	Proteína (% b.s.)
NS 4619	38.15	NS 4619	39.01
A 5009	37.57	NS 6248	38.34
NS 4955	37.46	NS 4955	38.29
LDC 5.9	37.16	LDC 5.9	37.86
Syn 5x2	37.06	DM 4712	37.58
DM 3810	36.71	Syn 5x1	37.49
LDC 5.3	36.58	LDC 5.3	37.09
DM 4014	36.41	DM 4612	36.57
DM 4612	36.18	DM 4014	36.66

Mediante éste análisis, no se observó una correlación negativa entre rendimiento y concentración de proteína. Sin embargo, se reconoció una frontera en la exportación de nitrógeno en granos de alrededor de 290 kg N ha⁻¹ (Fig. 6).

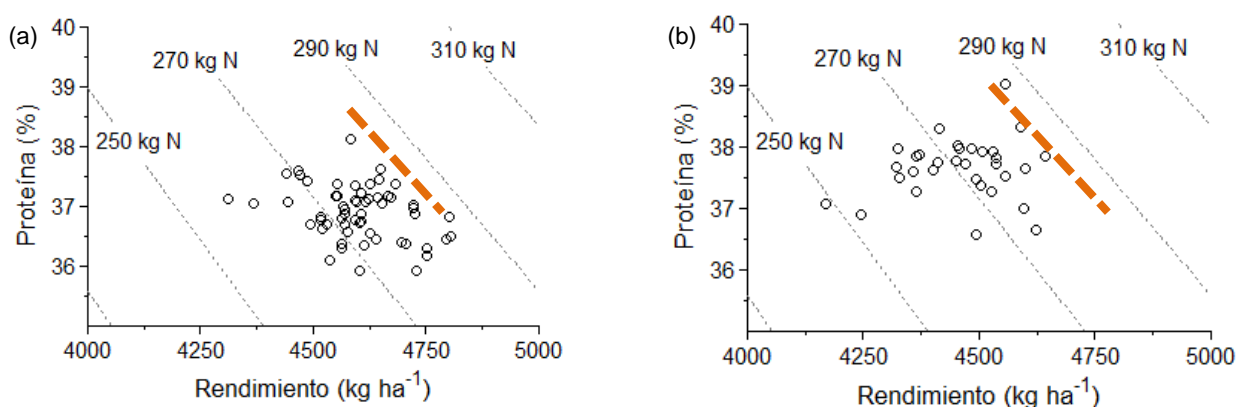


Fig. 6. Relación entre concentración de proteína en grano (%) y rendimiento (kg ha⁻¹) para soja de primera (a) y para soja de segunda (b). Las líneas punteadas indican kilogramos de nitrógeno por hectárea exportados por el cultivo. Cada punto representa valores de BLUPs para cada genotipo evaluado, es decir, valores por encima o por debajo de la media con los que se identifica cada variedad en los ambientes productivos en los que se desarrollan.

CONCLUSIONES

Se observó gran variación en el porcentaje de proteína en grano a través de lotes de producción. Se pudo determinar un patrón espacial para soja de primera pero no para soja de segunda. La elección del genotipo es la herramienta de manejo que surge con mayor consistencia en la determinación de la concentración de proteína en granos. Que el genotipo sea la variable que más influye en la concentración de proteína no implica que no haya efecto ambiental, se vio que variables ambientales como temperatura y precipitaciones contribuyeron también a la explicación de la variación observada.

Respecto a ensayos comparativos de rendimiento, se observó un fuerte efecto ambiental en la expresión del porcentaje de proteína. Sin embargo, el efecto genotipo fue mayor a la interacción genotipo por ambiente, lo que evidencia que determinados genotipos tienen un alto valor genético para producir elevados niveles proteicos. La elección del genotipo surge como la herramienta más importante sobre la cual los productores agrícolas tienen control para la determinación de proteína.

No hubo evidencias que demostraran que genotipos de alto rendimiento presenten necesariamente bajos niveles de proteína en grano, como lo indica la falta de correlación negativa entre rinde y proteína. Sin embargo, lo que se encontró fue una frontera de exportación de nitrógeno por el cultivo de alrededor de 290 kilogramos por hectárea. El gran desafío para aumentar el porcentaje de proteína con niveles altos de rinde radica en desplazar la frontera de exportación de N mediante aumentos en la captura de N del cultivo, lo que invariablemente lleva a estudiar cuestiones de absorción directa y fijación biológica de N en los cultivos de soja de la Rep. Argentina (Santachiara et al., 2017).

ACCIONES A FUTURO

Una de las actividades a realizar es modelar el comportamiento de algunos genotipos en relación a variables ambientales y de manejo. De esta manera se podría predecir la variabilidad causada por el ambiente respecto a concentración de proteína en grano.

Como otra acción a futuro surge la exploración de otros aspectos de la calidad de grano, más allá del uso asociado a la alimentación animal, por ejemplo propiedades funcionales y caracterización de la proteína.

AGRADECIMIENTOS

Los autores agradecen a los productores y asesores de AACREA por las muestras enviadas y proporcionar los datos de manejo agronómico. Este trabajo fue financiado por la Agencia Nacional de Promoción Científica y Tecnológica, Argentina (PICT 2014-3177) y por AACREA.

REFERENCIAS

- Breiman, L., Friedman, J., Stone, C. J., y Olshen, R. A. (1984). Classification and regression trees. CRC press.
- de Felipe, M., Gerde, J. A., y Rotundo, J. L. (2016). Soybean genetic gain in maturity groups III to V in Argentina from 1980 to 2015. *Crop Science*, 56(6), 3066-3077.
- De'ath, G. y Fabricius, K. Classification and regression trees: A powerful yet simple technique for ecological data analysis. *Ecology*. 2000. 81:3178–3192.
- FAOSTAT. (2015). Production: Crops. Food and Agriculture Organization of the United Nations, Statistics Division. <http://faostat3.fao.org/>.
- Liao, S. H., Chu, P. H., y Hsiao, P. Y. (2012). Data mining techniques and applications—A decade review from 2000 to 2011. *Expert systems with applications*, 39(12), 11303-11311.
- Mourtzinis, S., Gaspar, A. P., Naeve, S. L., y Conley, S. P. (2017). Planting date, maturity, and temperature effects on soybean seed yield and composition. *Agronomy Journal*, 109(5), 2040-2049.
- Rotundo, J. L., Miller-Garvin, J. E., y Naeve, S. L. (2016). Regional and temporal variation in soybean seed protein and oil across the United States. *Crop Science*, 56(2), 797-808.
- Rotundo, J. L., y Westgate, M. E. (2009). Meta-analysis of environmental effects on soybean seed composition. *Field Crops Research*, 110(2), 147-156.
- Santachiara, G., Borrás, L., Salvagiotti, F., Gerde, J.A., y Rotundo, J.L. (2017) Relative importance of biological nitrogen fixation and mineral uptake in high yielding soybean cultivars. *Plant & Soil*, 418(1), 191-203.