

EVALUACIÓN DE LA COMPOSICIÓN Y PERFIL DE AMINOÁCIDOS EN GRANOS DE SOJA DE ARGENTINA

ETAPA 2

Resumen de la ETAPA 2

En esta segunda etapa del proyecto se ha realizado un estudio exploratorio sobre la calidad de soja disponible para el procesamiento en plantas de crushing proveniente de todas las áreas de siembra de la Argentina con el fin de conocer la relación entre la calidad composicional y nutricional con su respectiva geolocalización



Instituto Nacional de
Tecnología Agropecuaria



Autores de la ETAPA 2:

- Accoroni Cecilia, INTA EEA Oliveros
- Carpaneto Bárbara, INTA EEA Balcarce
- Caruso Cecilia, INTA EEA Oliveros
- Chialvo Eugenia, INTA EEA Marcos Juárez
- Conde Belén, INTA EEA Marcos Juárez
- Enrico Juan Martin, INTA EEA Oliveros
- Magnano Luciana, INTA EEA Oliveros
- Mir Leticia, INTA EEA Marcos Juárez
- Orlando Ángela, ACSOJA
- Rossi Rodolfo, ACSOJA
- Santos Diego J., INTA EEA Paraná
- Zelaya Daniela Karina, INTA EEA Balcarce

Agradecimientos especiales a:

- Tolchinsky Marcelo, Representante de INTA en ACSOJA
- Giménez Fernando, Coordinador del Programa Nacional de Cereales y Oleaginosas
- Agencias de Extensión Rural de INTA de las regiones involucradas en el estudio que participaron de la recolección de muestras
- Grupos CREA por el aporte de muestras para el estudio
- Personal técnico de Greenlab que participaron del análisis de aminoácidos y posterior interpretación de datos
- CIARA por sus aportes al trabajo y sector

ÍNDICE GENERAL

	<i>Pág.</i>
1. INTRODUCCIÓN.....	6
2. MATERIALES Y MÉTODOS.....	7
2.1. Área Geográfica.....	8
2.2. Muestreo.....	8
2.3. Caracterización del grano de soja	8
2.3.1. Determinación de la composición del grano.....	8
2.3.2. Determinación de perfil de AA.....	8
2.3.3. Clasificación cualitativa de cultivares.....	9
2.4. Análisis estadístico	9
2.5. Análisis geoestadístico.....	9
3.RESULTADOS.....	10
3.1. Caracterización de las muestras evaluadas.....	11
3.1.1. Grupo de Madurez.....	11
3.1.2. Efecto de fecha de siembra, variedad y grupo de madurez sobre proteína y aceite, zona extra pampeana.....	12
3.2. Medidas descriptivas de la composición de soja por variable: proteína, aceite y PROFAT.....	12
3.2.1. Proteína.....	12
3.2.2. Aceite.....	12
3.2.3. PROFAT.....	13
3.3 Distribuciones de probabilidad por variables.....	14
3.4 Evaluación de correlación entre las variables.....	16
3.5. Evaluación estadística de aminoácidos.....	17
3.5.1. Perfil completo de aminoácidos de la soja evaluada.....	17

3.5.2. Evaluación estadística de los key aminoácidos por separado.....	18
3.5.2.1. Cisteína.....	18
3.5.2.2. Lisina.....	19
3.5.2.3. Metionina.....	20
3.5.2.4. Treonina.....	20
3.5.2.5. Triptófano.....	21
3.5.2.6. Análisis de correlación entre los distintos aminoácidos y entre los aminoácidos y proteína.....	21
3.5.3. Análisis estadístico de los Key Aminoácidos.....	22
3.6 Análisis geoestadístico.....	26
3.6.1 Análisis de Aceite, Proteína y PROFAT.....	26
3.6.2 Análisis de Aminoácidos.....	27
3.6.3 Validación.....	27
4. DISCUSION Y CONCLUSION.....	28
4.1. Análisis de Aceite Proteína y PROFAT Argentina vs Cono Sur y USA.....	28
4.2. Variabilidad dentro de nuestro país, entre zonas, potencialidades de cada zona.	28
4.3. Análisis de Aminoácidos entre Argentina y el mundo.....	29
4.4. Conclusiones y próximos pasos.....	29
5 Conclusiones.....	30
5.1 Conclusiones generales.....	30
5.2 Próximos pasos.....	32
6. BIBLIOGRAFÍA.....	32
ANEXO I	
Detalle de los departamentos y/o partidos comprendidos en cada zona homogénea de producción.....	33

ANEXO II

Análisis químicos y por NIR.....	43
Análisis de calidad: composición fisicoquímica y aminoácidos.....	3
ANEXO III - Efectos de la fecha de siembra, la variedad y el grupo de madurez sobre proteína y aceite en muestras de grano de soja cosechada durante la campaña 2020/21 fuera de la zona Núcleo	
Efecto de la fecha de siembra declarada.....	44
Diferencias entre las variedades declaradas.....	47
Efecto de la madurez de los materiales (Grupo de madurez de las variedades declaradas).....	48
ANEXO IV. Análisis Estadístico	50
ANEXO V. Aminoácidos	55
Distribución de los AA	55
ANEXO VI. Detalle del Análisis Geoestadístico	60

1. INTRODUCCIÓN

La soja (*Glycine max*) es uno de los cultivos más valiosos, versátiles y nutricionalmente importantes a nivel mundial. Es capaz de desarrollarse en diversos ambientes, se adapta a una variedad de prácticas de manejo, y posee diversos propósitos y usos (Shea et al., 2020). Las producciones de EE.UU., Brasil y Argentina representan aproximadamente el 81 % de la producción mundial de soja, siendo Argentina el tercer productor mundial y el mayor exportador de Harina de Soja (HS) (FAO, 2018). La HS es el subproducto de la extracción del aceite del poroto, y se caracteriza por poseer un alto valor nutritivo, utilizado principalmente para la alimentación animal (Bozas et al., 2019). Además, es la principal fuente proteica en dietas para no rumiantes formulando, los nutricionistas, piensos basados en su contenido de proteína (CP) independientemente de otras características de los granos (Ibañez et. al, 2020).

Diferentes estudios han demostrado que factores tales como genotipo de la soja, tipo de suelo, manejo de cultivo (fecha de siembra, fertilización, rotaciones, etc.), condiciones ambientales durante las temporadas de crecimiento y cosecha (temperaturas promedio en llenado de grano, régimen hídrico, etc.), almacenamiento y zona geográfica de producción (altitud y latitud) afectan la calidad nutricional y el CP en soja y sus correspondientes derivados (Benavidez et al., 2008; Carrera et al., 2011; Assefa et al., 2018; Bozas et al., 2019; de Borja Reis et al., 2020; García-Rebollar et al., 2016; Lagos y Stein, 2017). Sumado a esto, Ibañez et al. (2020) muestra que HS provenientes de EE.UU., a pesar de poseer un CP inferior a Brasil, presenta un contenido de aminoácidos (AA) esenciales por unidad de proteína superior, concluyendo así que, a igual CP, la composición química, calidad proteica y valor nutritivo puede ser diferente entre las HS de distintos orígenes. Debido a ello, la calidad de la soja ha comenzado a ser monitoreada por los diferentes países involucrados. Estados Unidos posee un programa de muestreo anual de proteína, aceite, y aminoácidos que le permite realizar mapas que expresan esos valores y conocer las tendencias anuales de cada región (Assefa et al., 2018; de Borja Reis et al., 2020). El resto de los países presentan muestreos documentados en trabajos científicos o exploraciones parciales en tiempo y lugar, lo que les permite conocer la realidad en forma parcial (Carrera et al, 2011; Bozas et al, 2019). En Argentina, el INTA realiza muestreos anuales dentro de la Red Nacional de Evaluación de Cultivares de Soja (RECSO), que incluye variedades comerciales y precomerciales, y un muestreo regional (zona núcleo) anual a nivel de acopios, en el cual se releva la calidad de soja cosechada (Cuniberti et al, 2018; Chialvo et al, 2021). Además, tanto Universidades como otras instituciones cuentan con trabajos científicos y tesis de postgrado que brindan información sobre CP y las interacciones del mismo con las variables ambientales, el manejo, uso de variedades, localización, etc. (Benavidez et al, 2008; Di Mauro et al, 2018). Sin embargo, no hay antecedentes de un muestreo nacional directamente en campos de productores, que permita relacionar calidad del grano con ambiente, variedad y prácticas agronómicas.

Tanto la soja como la HS han sido valoradas históricamente por su CP. Sin embargo, es sabido que las proteínas difieren en su valor nutritivo debido a factores tales como el contenido en aminoácidos esenciales (AAE) y la digestibilidad. Los AAE (histidina, isoleucina, leucina, lisina, metionina, fenilalanina, treonina, triptófano y valina) son aquellos AA no sintetizados por los mamíferos, que deben ser adquiridos mediante la ingesta para una nutrición completa. Las proteínas de alta calidad son aquellas que contienen AAE en niveles superiores a los niveles de referencia (FAO/OMS/UNU) y una digestibilidad comparable con

la de proteínas de clara de huevo o leche (Fennema, 2008). En el caso particular de las proteínas de soja, estas se caracterizan por ser ricas en AAE tales como lisina, treonina, isoleucina, leucina, fenilalanina y valina; sin embargo, son deficientes en metionina y cisteína (Badui-Dergal, 2006). Por último, una herramienta útil para evaluar la composición de AA y comparar entre las especies o lugares de referencia son los Key AA (KAA), que se definen como la sumatoria de AA relevantes para alimentación animal, ellos son: metionina (Met), Treonina (Tre), Cisteína (Cys), Triptófano (Try) y Lisina (Lys).

A nivel internacional, se pretende demostrar mediante asociaciones (American Soybean Association y USSEC) y estudios de investigación, que los AA presentes en HS son el factor nutricional de mayor relevancia a la hora de formular piensos para monogástricos y, por ende, este valor intrínseco es el que definiría el precio diferencial de HS (Tabla 1). El estudio de Ibañez *et al* (2020), ha analizado CP, AA y otros elementos que caracterizan HS, concluyendo que HS con un valor superior de KAA (%), y no necesariamente superior CP, serían más competitivas para la formulación de piensos.

Tabla 1. Aminoácidos importantes dentro de la dieta de monogástricos

Pollos	Importancia	Cerdos
Metionina	1	Lisina
Lisina	2	Metionina
Treonina	3	Treonina
Arginina	4	Triptófano
Isoleucina	5	Valina

Argentina cuenta actualmente con información parcial de composición de granos y HS, lo cual es esencial para competir en los mercados internacionales. En las últimas décadas la composición de los granos y HS producidos en Argentina ha sido analizada (Bolsa de Comercio de Rosario, INTA). Se han medido valores medios de 37,5 % contenido proteico, 22,9 % contenido de aceite y 60,3 % de PROFAT en la composición del grano de soja de los últimos 18 años (Cuniberti *et al*, 2018; Chialvo *et al.*, 2021).

Respecto a la HS, Argentina presenta un producto de exportación con aspectos nutricionales y de inocuidad satisfactorios. Según informes pres:

- ser uniforme y homogéneo, gracias a que las plantas de crushing de Argentina presentan similitud en instalaciones de procesamiento,
- ser un ingrediente sin aditivos o demás agregados (100 % HS),
- poseer un CP medio de 46,5 %, con buena solubilidad de proteínas, lo cual indica que la proteína no fue sobre procesada; además, se registra un bajo contenido de fibra, debido a la eficiencia en descascarado
- presentar un perfil de AA comparable cuali y cuantitativamente con sus competidores,

- ser inocuo, debido a que cumple con los requisitos, controles y certificaciones correspondientes que así lo garantizan (SENASA, certificación GMP, ISO 9001, FSA, HACCP),
- poseer una calidad garantizada por análisis de laboratorios validados mediante referee program.

Por otro lado, Argentina, a pesar de ser uno de los productores de soja y HS más importantes del mundo, carece de un estudio nacional actualizado que defina la calidad y permita posicionar sus productos a nivel mundial. El objetivo general del presente trabajo ha sido realizar un estudio exploratorio sobre la calidad de soja disponible para el procesamiento en plantas de crushing proveniente de todas las áreas de siembra de la Argentina a fin de conocer la relación entre la calidad composicional (contenido de aceite, proteína y profat) y nutricional (AA) con la geolocalización y presentar dicha información en mapas.

2. MATERIALES Y MÉTODOS

2.1. Área Geográfica

En Argentina el cultivo de la soja es sembrado tanto en la región pampeana, como en otras regiones (Figura 1 a, de Abelleyra et al., 2021), especialmente en el NOA y NEA, siendo las provincias de Buenos Aires, Córdoba y Santa Fe en donde se concentra esta actividad con 10.4 millones de hectáreas (ha) equivale a un 75.4 % del área sembrada (SISA 2021). Al ser una región muy heterogénea en clima, suelos y sistemas de producción, se definieron zonas homogéneas según criterio de referentes nacionales en la temática. De esta forma delimitaron 8 zonas homogéneas de producción de soja (Figura 1 b). Los departamentos comprendidos en cada zona se detallan en el Anexo I¹.

2.2. Muestreo

La cantidad de muestras a recolectar fue definida proporcionalmente a la superficie sembrada en cada zona (MAGyP, 2021), de modo que la representación de cada área fuera equitativa. Las muestras de soja (n=680, campaña 2020/21) fueron recolectadas por profesionales de INTA de las Agencias de Extensión y la red de colaboradores de INTA (grupos AACREA, AAPRESID, otros productores), tanto en campos de productores (n=550) como en centros de acopio (n=130). El peso de las muestras fue de entre 1 a 2 Kg de grano; junto con ellas se solicitó la siguiente información: fecha de siembra, variedad, fertilización, inoculación y georreferencia del lote o punto de referencia más cercano. Cada una de estas muestras fue acondicionada, conservada en bolsas de cierre hermético e identificada con un código interno, para ser evaluada posteriormente. Aquellas muestras sin referencia geográfica (247 muestras), fueron ubicadas en un área cercana a la localidad cabecera del partido o departamento.

¹ https://docs.google.com/document/d/1Qik-AndWU7X61U_z9p-ojK29hAATac7iPrf5XUK8VgA/edit?usp=sharing

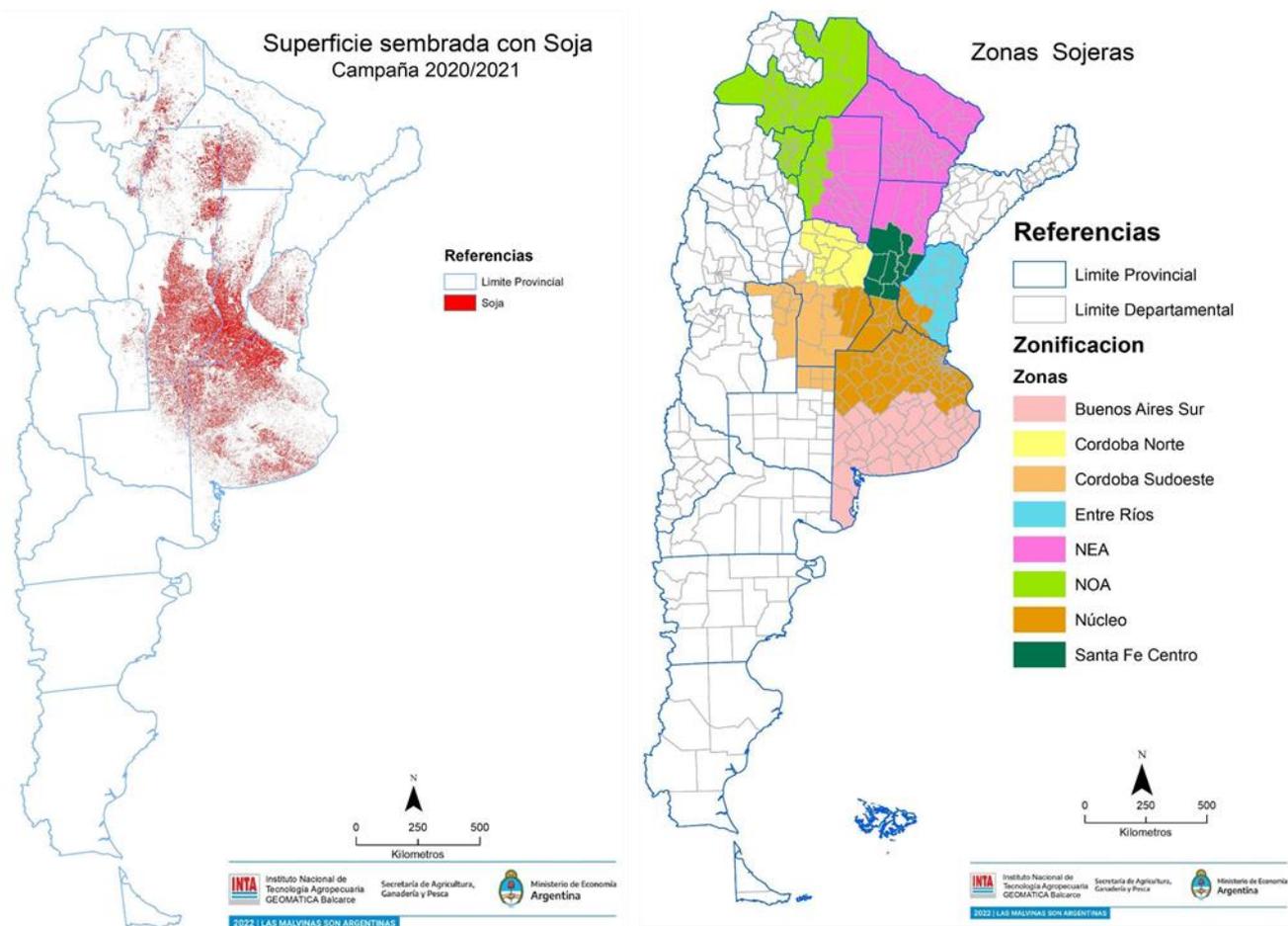


Figura 1. Izquierda: Distribución espacial del cultivo de soja en Argentina durante la campaña 2020/2021 (Fuente de Abelleyra et al., 2021. disponible en www.geointa.gob.ar), Derecha: Distribución de las 8 zonas sojeras de la República Argentina delimitadas para este trabajo.

2.3. Caracterización del grano de soja

2.3.1. Determinación de la composición del grano

Se utilizaron analizadores digitales de granos de espectroscopia infrarroja cercana, NIR DS 2500 (FOSS) para determinar algunos parámetros de calidad de poroto de soja: humedad, proteína cruda y aceite. Todos los valores fueron corregidos a % en base seca (% bs). Sumado a esto, 100 muestras seleccionadas fueron analizadas en las instalaciones de Greenlab con los métodos de referencia a fin de validar los datos obtenidos por NIR. La calibración utilizada se detalla en el Anexo II².

2.3.2. Determinación de perfil de AA

Se seleccionaron 100 muestras de soja, a las cuales se les determinó el perfil de aminoácidos completo por HPLC-MStQ. El criterio de selección de muestras y la descripción de técnica utilizada en la determinación de los aminoácidos se detalla en el Anexo II.

² <https://docs.google.com/document/d/1aoaZWP2ro-o4yh-oxiryxziBTfEV8ERhXafQxaI940/edit?usp=sharing>

2.3.3. Clasificación cualitativa de cultivares

A los fines de describir las muestras recibidas en relación a la zona de producción, se han definido los índices de “diversidad de uso” y “antigüedad media”. El índice de “diversidad de uso” indica la cantidad de cultivares diferentes respecto a la cantidad total de datos. Es decir, que a mayor cantidad de cultivares, mayor será la diversidad varietal utilizada en una macro región determinada. Se calculó de la siguiente forma:

$$\text{Diversidad de uso (\%)} = \frac{\# \text{ de Cultivares}}{\text{Total de datos}} * 100$$

Mientras que el índice de “antigüedad media” brinda información sobre la prevalencia de cultivares declarados en cada macro región, dado que es obtenido a partir de la fecha del registro de Inscripción (RNC, INASE <https://gestion.inase.gob.ar/consultaGestion/gestiones>) de los cultivares. Esto significa que la presencia de valores bajos indicará una mayor adopción de cultivares de liberación reciente. En cambio, valores más altos indicarán baja adopción de cultivares modernos y/o ambas condiciones en forma conjunta.

$$\text{Antigüedad media (años)} = \frac{(\text{Enero 1 de 2022} - \sum \text{Insc. RNC cultivar})}{365}$$

2.4. Análisis estadístico

Un estudio estadístico fue aplicado a fin de medir, tanto a nivel nacional como por zonas, el contenido de aceite, proteína y PROFAT, siendo este último la suma de los anteriores. Estas variables han sido evaluadas en su valor porcentual respecto de su base seca (% bs). Además, se evaluaron los KAA; Cys, Lys, Met, Tre, Try tanto en su valor de base seca (sss) como porcentual respecto de la proteína (%) en conjunto y por separado. Para ello se realizaron análisis de medidas de posición como promedio (media aritmética), mediana (percentil 50 o Q2), Q1 (percentil 25), Q3 (percentil 75) y medidas de variabilidad como desvío estándar y el IQR (rango intercuartílico). Además, n (cantidad de observaciones), mínimo, máximo y CV (coeficiente de variación).

2.5. Análisis geoestadístico

La geoestadística es una rama de la matemática que consiste en el desarrollo de diversos procedimientos de estimación y simulación, los cuales sirven para estudiar variables distribuidas espacialmente y de esta manera analizar el fenómeno y predecir su valor (Di Mauro et al., 2018; Bosaz et al., 2019; Radocaj et al., 2020). Se utilizaron técnicas de geoestadística (método Kriging) para extrapolar los valores de aceite, proteína, PROFAT, y KAA obtenidos en laboratorio a partir de las muestras de soja de primera siembra (1era) relevadas. Se realizó un análisis exploratorio de los datos, se estudió su simetría, outliers, tendencia espacial, semivariograma, entre otros. Se analizó la precisión del ajuste de los modelos circular, esférico, exponencial y gaussiano, realizando una validación de los mismos y se obtuvieron mapas continuos de las variables. El análisis se realizó por zonas, de acuerdo a la representatividad de muestras por zonas.

3. RESULTADOS

3.1. Caracterización de las muestras evaluadas

3.1.1. Grupo de Madurez

En base a los datos relevados, se ha identificado que el grupo de madurez (GM) más utilizado fue el IV, con una participación del 50,2 %, seguido por el V con un 17,5 %, el VI, con un 13,4 %, y el III, con 10,2 % (Tabla 2).

Tabla 2. Participación (expresada en porcentaje) de los diferentes GM respecto del total de muestras* con identificación de la variedad utilizada en el lote de producción.

GM	% de participación
II L	0.5
III C	5.0
III L	5.2
IV C	9.4
IV L	40.8
V C	14.4
V L	3.1
VI C	8.4
VI L	5.0
VII	4.5
VIII	3.7

* El # de muestras con información del GM fue 382 sobre

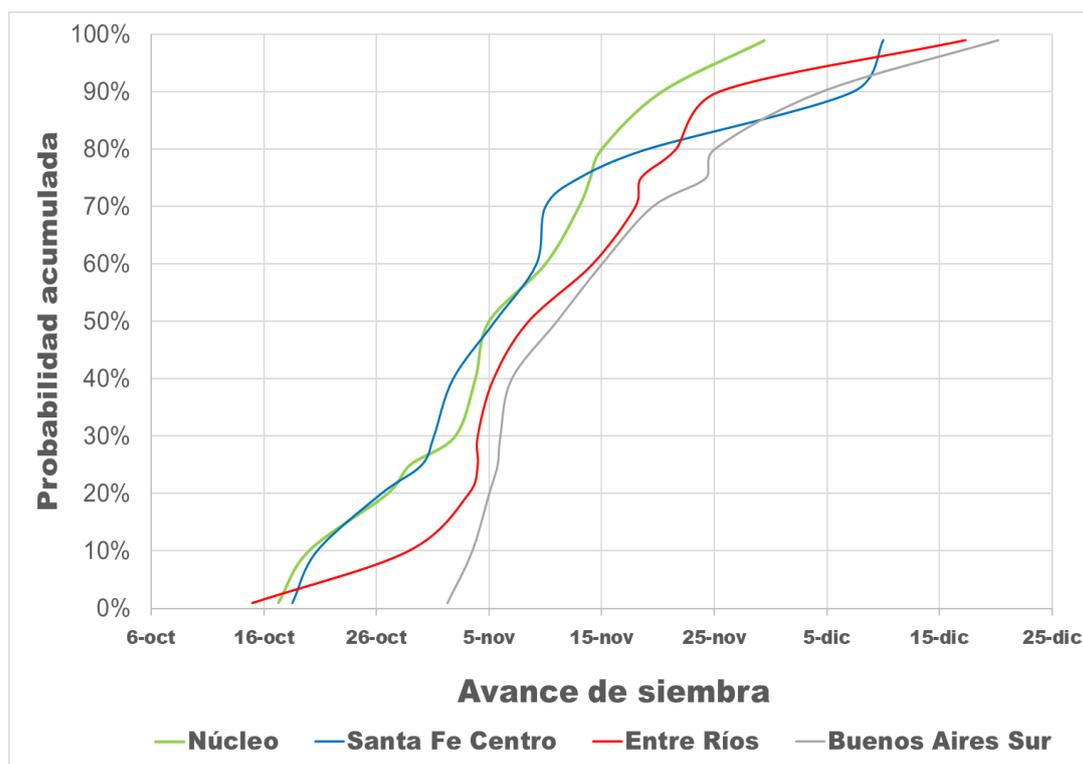
A pesar del desbalance en el número de datos registrados entre las zonas relevadas, se ha podido determinar que la antigüedad media de los cultivares sembrados a nivel nacional es de 6,5 años, siendo las zonas de Córdoba Sudoeste, NOA y Entre Ríos las que registraron antigüedad inferior a la media nacional (Tabla 3). La zona Núcleo se destaca por la alta concentración en el uso de cultivares con apenas un 28% de diversificación, le sigue la zona de Córdoba Norte, mientras que en el resto de las regiones se observaron valores altos en la diversidad de uso de cultivares (Tabla 3).

Respecto a la fecha de siembra, dentro de las zonas analizadas se observó una similar fecha de inicio de la siembra de soja de 1° (16-Oct), excepto Buenos Aires Sur (1-Nov) (Figura 2). Las macro regiones Núcleo y Santa Fe Centro mostraron curvas de avance de la siembra de soja de 1° similares entre sí y alcanzan la fecha de siembra media (50 %) alrededor del 5-Nov, mientras que las macro regiones Entre Ríos y Buenos Aires Sur lo hicieron entre el 8 y 11 de noviembre (Figura 2). En todas las regiones, entre el 25 y el 75 % de avance de las siembras transcurrieron en un período de entre 12 y 18 días. Exceptuando la zona Núcleo, que finalizó las siembras antes de diciembre, las restantes zonas analizadas lo hicieron entre el 10 y 15 de diciembre.

Tabla 3. Cantidad de variedades utilizadas por macro región, diversidad de uso (%) y antigüedad media de los materiales utilizados*

Zona	# datos	# de Cultivares	Diversidad de uso de Cvs (%)*	Antigüedad media de Cvs
Núcleo	185	52	28	-6.9 ± 3.1
Córdoba Sudoeste	41	19	46	-5.6 ± 2.8
Buenos Aires - Sur	33	19	58	-6.9 ± 4.1
NOA	30	23	77	-6.4 ± 2.6
Entre Ríos	26	20	77	-6.2 ± 2.9
NEA	20	15	75	-6.6 ± 5.1
Santa Fe Centro	19	14	74	-6.8 ± 3.6
Córdoba Norte	13	11	85	-6.2 ± 2.9
Total	367	115	65	-6.5 ± 0.5

En cuanto a la siembra de la soja de 2° la zona Núcleo tuvo aproximadamente un mes de retraso tanto en el inicio, como en el fin, respecto a la siembra de soja de 1° ocupación (Figura 3). Si bien el arranque de la siembra fue más rápido en el caso de la soja de 1°, el tiempo que transcurrió entre el 25 y 75 % de avance fue similar, con 16 días para soja de 1° (entre el 29-Oct al 14-Nov) y 12 días para soja de 2° (entre 3 al 15/Dic).

**Figura 2.** Probabilidad acumulada del avance de siembra por macro regiones. # de datos utilizados por macro región: Núcleo (78); Entre Ríos (24); Santa Fe Centro (28); Buenos Aires Sur (29).

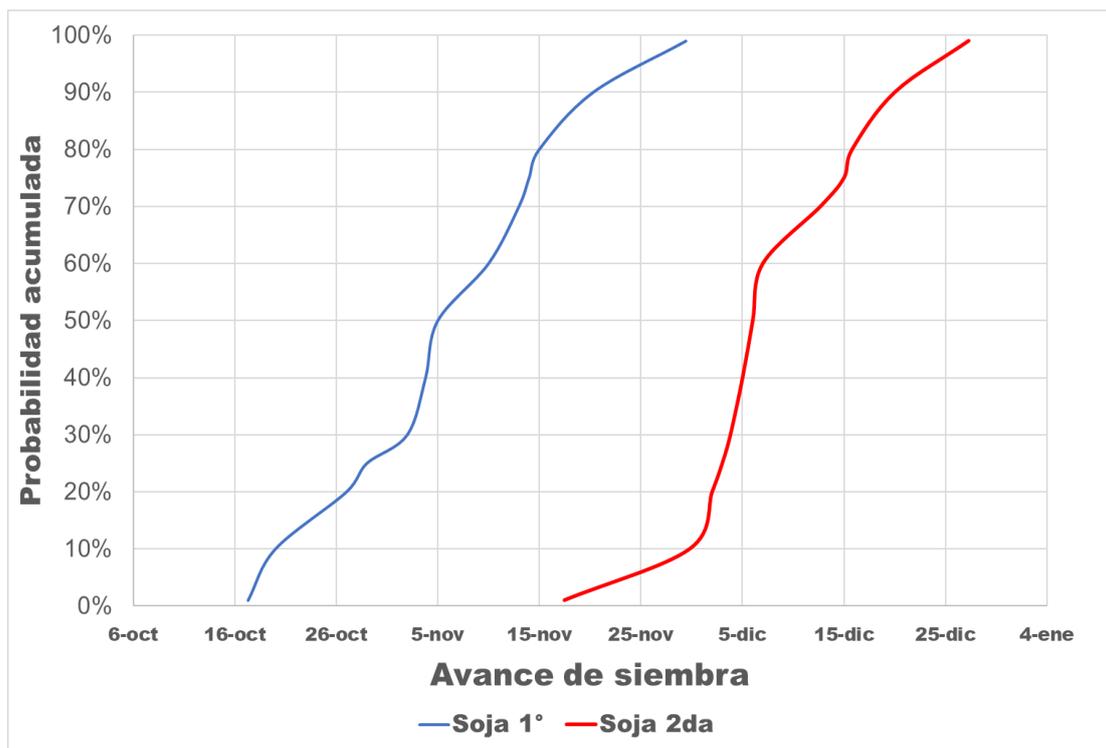


Figura 3. Probabilidad acumulada del avance de la siembra de 1º época y de 2da para la región Núcleo. # de datos utilizados por FS 1º: 78 y FS 2º: 39.

3.1.2. Efecto de fecha de siembra, variedad y grupo de madurez sobre proteína y aceite, zona extrapampeana

Considerando el bajo número de muestras y siendo una sola la campaña considerada, se presentan en forma gráfica los efectos de la fecha de siembra declarada y las diferencias entre variedades mencionadas sobre proteína, aceite y PROFAT en el Anexo III. Se incluyen allí las muestras de Entre Ríos, Córdoba Noroeste, Santa Fe Norte, La Pampa, San Luis, Santiago del Estero y Salta. Complementando a las figuras de “variedad” se agregan aquellas de Grupo de Madurez (GM).

3.2. Medidas descriptivas de composición de soja por variable: proteína, aceite y PROFAT

3.2.1. Proteína

El valor medio de proteína considerando todas las zonas fue de 36,6 % con valores mínimos y máximos de 32,0 % y 41,0 %, respectivamente (Tabla 4). El valor medio más bajo encontrado corresponde a Santa Fe Centro (35,7 %), mientras que el valor promedio más alto (38,0 %) se registró en el NOA. Se puede observar una baja variabilidad en los datos registrados para cada una de las zonas, ya que los coeficientes de variación se mantuvieron en todas las zonas inferiores a 4,6 % (Tabla 4).

3.2.2. Aceite

Al considerar todas las regiones, el valor medio de aceite fue 23,0 %, fluctuando en un rango comprendido entre 18,3 - 27,0 % (Tabla 5). Entre regiones, el valor más alto correspondió a Santa Fe Centro (24,7 %), mientras que el más bajo a Buenos Aires Sur (21,0 %). Observando los CV y RIC para las distintas zonas, los mismos fueron inferiores a 6,2 % y a 1,7 (Tabla 5).

Tabla 4. Medidas resumen por zonificación – proteína

Zona	n	Min	Max	Media	DE	CV	Mediana	Q1	Q3	RIC
Buenos Aires Sur	37	34,2	40,1	36,9	1,3	3,6	37,0	36,0	37,6	1,7
Córdoba Norte	43	34,5	39,9	37,1	1,5	4,0	37,2	35,9	38,3	2,3
Córdoba Sudoeste	58	33,5	40,2	37,2	1,3	3,6	37,4	36,7	38,1	1,4
Entre Ríos	26	34,1	38,8	36,3	1,3	3,5	36,2	35,4	37,0	1,7
NEA	24	33,3	40,1	37,2	1,7	4,6	37,3	36,4	38,0	1,6
NOA	35	36,1	41,0	38,0	1,3	3,4	38,2	37,1	38,8	1,7
Núcleo	243	32,0	40,5	36,3	1,5	4,1	36,2	35,3	37,3	2,0
Santa Fe Centro	78	32,6	39,4	35,7	1,2	3,4	35,5	34,9	36,5	1,5
Nacional	544	32,0	41,0	36,6	1,5	4,2	36,6	35,4	37,6	2,2

Tabla 5. Medidas resumen por zonificación – aceite

Zona	n	Min	Max	Media	DE	CV	Mediana	Q1	Q3	RIC
Buenos Aires Sur	37	18,3	23,4	21,0	1,2	5,7	21,0	20,0	21,8	1,7
Córdoba Norte	43	19,8	24,0	22,2	0,8	3,6	22,3	21,9	22,7	0,8
Córdoba Sudoeste	58	18,6	25,0	21,9	1,3	6,0	21,8	21,1	22,8	1,7
Entre Ríos	26	20,9	24,3	22,7	0,9	4,1	22,7	22,0	23,4	1,4
NEA	24	21,1	25,3	22,8	0,9	4,1	22,8	22,2	23,0	0,8
NOA	35	20,5	24,2	22,6	0,9	4,0	22,8	21,9	23,3	1,3
Núcleo	243	18,8	26,4	23,2	1,4	6,2	23,3	22,5	24,2	1,7
Santa Fe Centro	78	21,0	27,0	24,7	1,2	4,8	24,9	24,2	25,6	1,4
Nacional	544	18,3	27,0	23,0	1,6	6,9	23,0	21,9	24,1	2,2

3.2.3. PROFAT

El valor de PROFAT promedio general fue de 59,5 % con un rango entre 54,9 - 63,3 %, tal como se muestra en la Tabla 6. Al comparar los valores medios de las distintas zonas, Santa Fe Centro (60,4 %) y Buenos Aires Sur (57,9 %) presentaron los valores máximos y mínimos. Consistentemente con lo observado en los casos anteriores, la variabilidad dentro de las zonas fue baja y no superó al 2,5 % (Tabla 6).

Tabla 6. Medidas resumen por zonificación – PROFAT

Zona	n	Min	Max	Media	DE	CV	Mediana	Q1	Q3	RIC
Buenos Aires Sur	37	54,9	60,6	57,9	1,2	2,2	57,7	57,1	58,7	1,6
Córdoba Norte	43	56,0	62,4	59,3	1,5	2,5	59,6	58,1	60,5	2,5
Córdoba Sudoeste	58	56,2	61,4	59,2	1,0	1,8	59,3	58,5	59,9	1,3
Entre Ríos	26	57,2	60,3	59,0	0,7	1,3	58,9	58,5	59,5	1,0
NEA	24	56,1	62,4	60,0	1,4	2,4	60,2	59	60,7	1,7
NOA	35	58,2	62,8	60,6	1,1	1,7	60,6	59,8	61,4	1,6
Núcleo	243	55,9	63,3	59,5	1,3	2,2	59,6	58,7	60,4	1,7
Santa Fe Centro	78	56,9	62,6	60,4	1,1	1,8	60,5	60,0	61,1	1,1
Nacional	544	54,9	63,3	59,5	1,4	2,3	59,6	58,6	60,5	1,9

3.3. Distribuciones de probabilidad por variables

Las distribuciones de probabilidades para cada zona se ilustran utilizando el gráfico de densidad de Kernel y destacando el ordenamiento creciente de acuerdo a la mediana de cada una de las variables. Los gráficos de densidad de Kernel muestran la distribución de los datos cuantitativos en un intervalo utilizando un suavizado de Kernel que permite atenuar el ruido de las distribuciones y, de esta manera, se pueda identificar mejor la forma de la distribución.

En el caso del aceite (Figura 4 a) se observa que cinco de las ocho zonas (Buenos Aires Sur, Córdoba Sudoeste, Entre Ríos, NOA y la región Núcleo) presentan una distribución bimodal. Los valores de aceite fueron considerablemente mayores en Santa Fe Centro. Si bien no se observan marcadas diferencias en cuanto a la dispersión de los datos por zona, puede observarse que la zona Núcleo y Córdoba Sudoeste son las que presentaron mayor dispersión.

En el caso de proteína, se puede observar en la Figura 4b que las distribuciones, en general, son unimodales. Mientras se observan valores de la mediana ligeramente superiores en NOA, los valores de mediana de Buenos Aires Sur, Córdoba Norte, NEA y Córdoba Sudeste son similares. A partir de las distribuciones presentadas puede evidenciarse que la distribución es similar en todas las regiones, excepto para NEA donde se puede observar una dispersión mayor.

Respecto a la variable PROFAT (Figura 4c), se presentan valores más altos en NOA, Santa Fe Centro y NEA. Luego se observa que valores levemente inferiores se presentan en la zona Núcleo, Córdoba Norte, Córdoba Sudeste y Entre Ríos. Los valores más bajos se presentaron en Buenos Aires sur. En términos generales se observa que Entre Ríos es el que presenta menor dispersión.

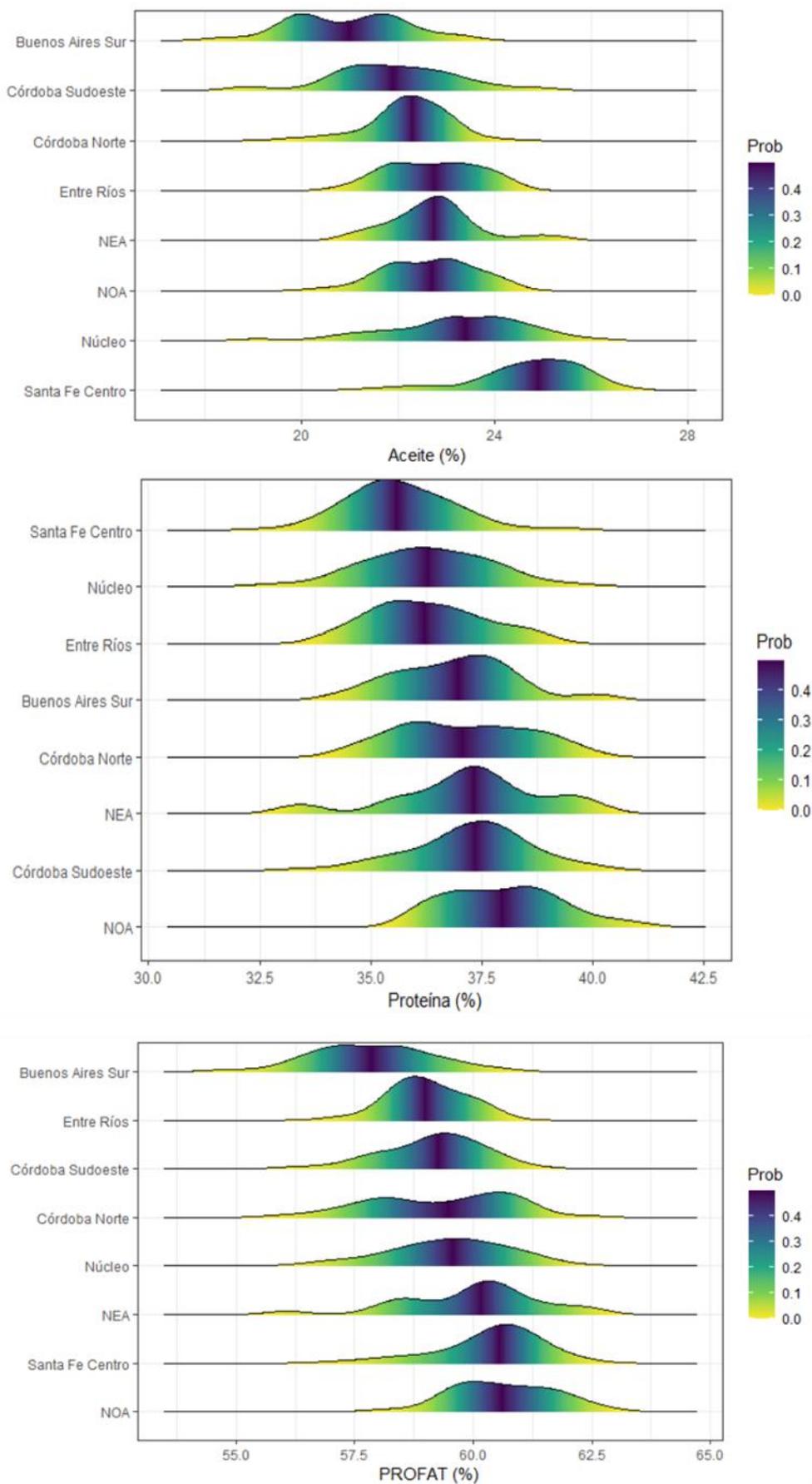


Figura 4. Gráfico de Densidad de Kernel para aceite (a), proteína (b) y PROFAT (c).

3.4. Evaluación de correlación entre las variables

Al evaluar la correlación entre el contenido de proteína con el de aceite, se observan relaciones negativas entre ellos para cada una de las zonas evaluadas. El análisis de correlación tuvo el comportamiento esperado; al aumentar el contenido de proteína, disminuye el contenido de aceite (Tabla 7). De acuerdo al análisis realizado, esta asociación es mayor en la zona de Entre Ríos. Además, puede observarse que se presentaron asociaciones relativamente altas en Córdoba Sudoeste, Núcleo, NOA y Santa Fe Centro.

Tabla 7. Coeficientes de correlación de Pearson entre porcentaje de proteína y aceite

Zonas	Correlación
Entre Ríos	-0,811
Córdoba Sudoeste	-0,691
Núcleo	-0,610
NOA	-0,575
Santa Fe Centro	-0,571
NEA	-0,534
Buenos Aires Sur	-0,521
Córdoba Norte	-0,232

A partir de los valores de los coeficientes de correlación obtenidos resulta de interés estimar una función que describa el porcentaje medio de proteína a partir de los porcentajes de aceite. Para ello, se ajustaron regresiones lineales simples por zona. En la Figura 5 se observa el gráfico de dispersión por zona con las correspondientes rectas de regresión estimadas y sus respectivas bandas de predicción. En todos los casos se presenta también la ecuación de regresión con el correspondiente coeficiente de determinación. A partir de los coeficientes de determinación puede observarse que, con excepción de Córdoba Sudoeste y Entre Ríos donde los coeficientes de determinación son de 48 % y 66 % respectivamente, los modelos ajustados explican, a lo sumo, un 37 % de la variabilidad del porcentaje de proteína. Ambientes restrictivos en aquellas dos regiones marcarían un límite superior de PROFAT difícil de superar. En muchas de las otras, en cambio, aun teniendo alto nivel de aceite, es posible tener, también, una buena concentración proteica en grano. Esto lleva a concluir, además, que en gran parte de las regiones no sería adecuado intentar predecir el valor de proteína a partir de los valores de aceite obtenidos. Mayor información estadística es posible encontrar en Anexo IV³.

³ https://docs.google.com/document/d/1JaboCVHZ8shBS-pMI3EIIiY-aaGV1dyi_ZMtLF69xg/edit?usp=sharing

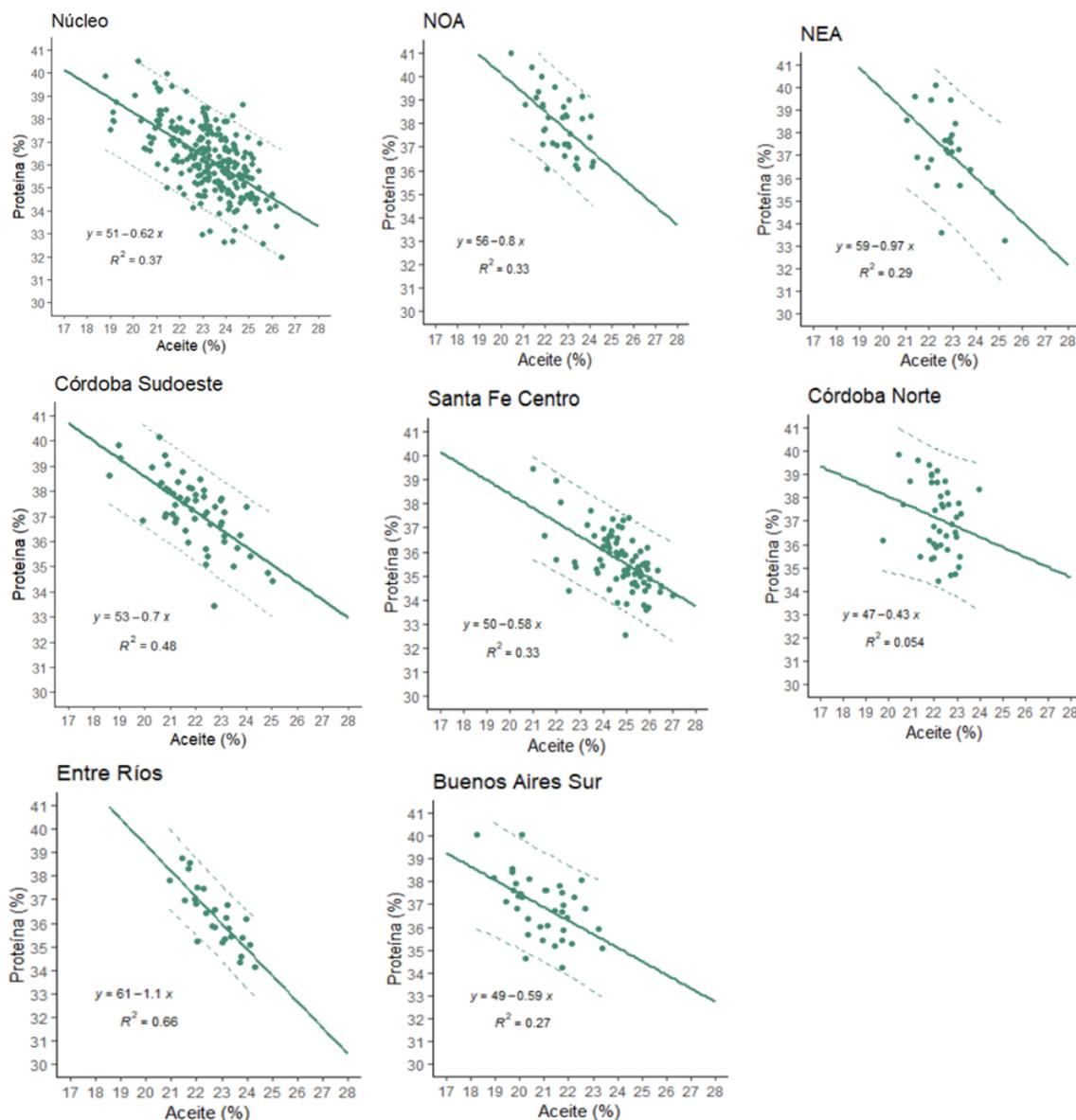


Figura 5. Regresión estimada entre Proteína y Aceite, por zonas sojeras.

3.5. Evaluación estadística de aminoácidos

3.5.1. Perfil completo de aminoácidos de la soja evaluada

En la Tabla 8 se muestran las medidas estadísticas descriptivas del total de AA (TAA) evaluados para las 100 muestras seleccionadas de soja de las diferentes zonas analizadas a lo largo de este trabajo. Del mismo modo, se pueden observar los respectivos valores para los Key (KAA) y los no Key (NKAA). Para definirlos se consideraron la suma de los valores medios (g base 12 % humedad) de cada AA; resultando KAA 5,3 g o 14,36 % de proteína y NKAA 31,62 g o 85,64 %.

Tabla 8. Valores medios de AA expresado en sustancia seca (g) y el % de proteína.

Aminoácido	AA expresados en porcentaje de proteína (%)			AA expresados en sustancia seca (g/100 g muestra)		
	Media	DE	CV (%)	Media	DE	CV (%)
Glicina	4,20	0,22	5,18	1,55	0,23	14,98
Alanina	4,16	0,16	3,80	1,54	0,20	13,30
Valina	4,84	0,20	4,21	1,79	0,24	13,22
Leucina	9,88	0,34	3,48	3,65	0,37	10,12
Isoleucina	5,42	0,21	3,81	2,00	0,28	14,05
Prolina	4,62	0,19	4,11	1,71	0,21	12,59
Fenilalanina	5,20	0,17	3,30	1,92	0,19	9,66
Tirosina	2,89	0,28	9,53	1,07	0,26	24,64
Triptófano	1,53	0,05	3,37	0,57	0,07	12,11
Lisina	6,16	0,27	4,33	2,27	0,31	13,59
Arginina	8,37	0,53	6,36	3,09	0,54	17,39
Histidina	2,97	0,22	7,43	1,10	0,23	20,63
Ác. Aspártico	10,12	0,51	5,07	3,74	0,51	13,70
Ác. Glutámico	18,10	0,77	4,26	6,68	0,79	11,89
Serina	4,88	0,38	7,85	1,80	0,35	19,69
Treonina	3,77	0,21	5,57	1,39	0,20	14,68
Cisteína	1,49	0,15	10,22	0,55	0,14	24,72
Metionina	1,41	0,12	8,39	0,52	0,11	20,98
TAA	100,00	0,18	0,18	36,92	0,18	0,48
KAA	14,36	0,08	0,58	5,30	0,09	1,78
NKAA	85,64	0,18	0,21	31,62	0,18	0,56

3.5.2. Evaluación estadística de los key aminoácidos por separado

A continuación, se presentan las medidas descriptivas de los KAA, tanto la sumatoria de los cinco juntos como la información de cada uno de sus AA constituyentes por separado. Para el análisis individual de los KAA se considerarán las expresiones en porcentaje (%) ya que es la más usada en la bibliografía de referencia.

3.5.2.1. Cisteína

A partir de este análisis, presentado en la Tabla 9, se ha evidenciado que el valor medio nacional de Cisteína fue 1,49 %, el mayor valor medio corresponde a la región NOA

2,62 %, mientras que el valor mínimo a la zona Núcleo (1,36 %). Se evidencia que existe diferencia en los valores medios de cada una de las zonas y, además, que las zonas NOA, Bs As Sur, Córdoba Sudoeste y Santa Fe Centro superaron el valor medio nacional.

Tabla 9. Medidas descriptivas de Cisteína (%)

Zona	n	Min	Max	Media	DE	CV	Mediana	Q1	Q3	IQR
Bs As Sur	5	1,35	1,85	1,62	0,18	10,99	1,62	1,61	1,66	0,05
Córdoba Norte	8	1,25	1,68	1,46	0,14	9,68	1,48	1,38	1,54	0,16
Córdoba Sudeste	15	1,08	2,24	1,52	0,33	21,94	1,41	1,31	1,72	0,40
Entre Ríos	6	1,27	2,12	1,47	0,33	22,22	1,34	1,30	1,44	0,15
NEA	8	1,21	1,66	1,47	0,15	10,33	1,48	1,37	1,57	0,19
NOA	7	1,68	2,62	2,22	0,35	15,59	2,27	2,03	2,45	0,42
Núcleo	45	0,91	2,12	1,36	0,29	21,50	1,29	1,08	1,58	0,50
Santa Fe Centro	6	1,24	1,79	1,51	0,16	10,82	1,47	1,40	1,58	0,17
Todas	100	0,91	2,62	1,49	0,34	23,08	1,48	1,27	1,65	0,38

3.5.2.2. Lisina

En la Tabla 10 se presentan las medidas descriptivas de la Lisina, donde se puede observar que el valor medio nacional fue igual a 6,16 %, con un valor medio máximo de 6,92 % en Santa Fe Centro y medio mínimo igual a 5,5 % en NOA. El valor máximo encontrado fue de 7,94 % y también corresponde a Santa Fe Centro. Las zonas Santa Fe Centro (6,92 %), Entre Ríos (6,65 %), Córdoba Norte (6,56 %) y NEA (6,22 %) superaron el valor medio nacional (6,16 %).

Tabla 10. Medidas descriptivas de Lisina (%)

Zona	n	Min	Max	Media	DE	CV	Mediana	Q1	Q3	IQR
Bs As Sur	5	5,44	6,49	5,93	0,40	6,79	5,83	5,73	6,15	0,41
Córdoba Norte	8	6,15	7,32	6,56	0,39	6,01	6,48	6,32	6,70	0,38
Córdoba Sudoeste	15	5,24	6,63	6,03	0,35	5,76	6,08	5,91	6,23	0,32
Entre Ríos	6	5,04	7,71	6,65	1,10	16,59	6,81	5,96	7,59	1,63
NEA	8	5,54	7,54	6,22	0,73	11,79	5,92	5,81	6,35	0,54
NOA	7	4,22	6,43	5,55	0,66	11,89	5,65	5,56	5,71	0,15
Núcleo	45	4,67	7,49	6,07	0,74	12,24	6,01	5,66	6,36	0,70
Santa Fe Centro	6	5,25	7,94	6,92	1,00	14,38	7,18	6,50	7,58	1,08
Todas	100	4,22	7,94	6,16	0,74	12,09	6,05	5,67	6,47	0,80

3.5.2.3. Metionina

En la Tabla 11 se presentan las medidas descriptivas de la Metionina. Se determinó que su valor medio nacional fue 1,41 %, siendo el valor mínimo 0,95 % y el valor máximo 2,55 %. A partir de este análisis, se puede observar que el NOA es la zona con mayores valores medios de Metionina (1,82 %) y Entre Ríos el de menor valor (1,22 %). Para el resto de las zonas sus valores medios varían entre 1,30 % y 1,55 %. Las zonas NOA, Córdoba Sudoeste, BS As Sur y Santa Fe Centro superan el valor medio nacional.

Tabla 11. Medidas descriptivas de Metionina (%)

Zona	n	Min	Max	Media	DE	CV	Mediana	Q1	Q3	IQR
Bs As Sur	5	1,20	1,75	1,49	0,26	17,74	1,64	1,21	1,66	0,44
Córdoba Norte	8	1,20	1,42	1,30	0,08	6,14	1,30	1,24	1,37	0,13
Córdoba Sudoeste	15	1,17	2,55	1,55	0,44	28,22	1,40	1,29	1,60	0,30
Entre Ríos	6	0,95	1,57	1,22	0,20	16,37	1,21	1,17	1,22	0,05
NEA	8	1,10	1,54	1,31	0,14	10,63	1,28	1,23	1,37	0,14
NOA	7	1,34	2,42	1,82	0,38	20,94	1,85	1,55	1,99	0,43
Núcleo	45	1,01	1,86	1,34	0,18	13,68	1,34	1,20	1,45	0,25
Santa Fe Centro	6	1,26	1,76	1,48	0,19	12,61	1,47	1,34	1,58	0,24
Todas	100	0,95	2,55	1,41	0,28	20,14	1,35	1,23	1,49	0,26

3.5.2.4. Treonina

En la Tabla 12 se presentan las medidas descriptivas de la Treonina. Se puede observar que su valor medio nacional es 3,77 %, siendo el valor mínimo 2,11 % y el valor máximo 5,65 %. Al evaluar este AA, se puede observar que sus valores medios se encuentran en el rango comprendido entre 3,33 - 4,78 %, el valor medio más alto corresponde a NOA mientras que el más bajo a Buenos Aires Sur. Las zonas NOA (4,78 %), Córdoba Sudoeste (3,93 %) y NEA (3,77 %) superan el valor medio nacional (3,77 %).

Tabla 12. Medidas descriptivas de Treonina (%)

Zona	n	Min	Max	Media	DE	CV	Mediana	Q1	Q3	IQR
Bs As Sur	5	2,31	3,80	3,33	0,65	19,49	3,72	3,06	3,77	0,71
Córdoba Norte	8	3,13	3,93	3,60	0,29	8,19	3,67	3,41	3,79	0,38
Córdoba Sudoeste	15	3,46	4,89	3,93	0,43	10,90	3,78	3,71	4,02	0,31
Entre Ríos	6	3,04	3,83	3,38	0,26	7,56	3,35	3,30	3,39	0,10
NEA	8	3,59	4,29	3,77	0,23	6,17	3,69	3,63	3,78	0,15
NOA	7	2,98	5,61	4,78	1,02	21,29	5,15	4,37	5,49	1,12

Núcleo	45	3,26	4,31	3,71	0,25	6,79	3,67	3,53	3,85	0,32
Santa Fe Centro	6	3,21	4,31	3,66	0,45	12,21	3,61	3,28	3,95	0,67
Total	100	2,31	5,61	3,77	0,51	13,46	3,72	3,49	3,87	0,38

3.5.2.5. Triptófano

En la Tabla 13 se presentan las medidas descriptivas del Triptófano. El valor medio nacional fue 1,53 %, con un rango de 1,16-1,99 %. Santa Fe Centro presentó el valor medio más alto de Triptófano (1,73 %) y NOA el menor valor medio (1,38 %). Sólo las zonas de las zonas NOA y Bs As Sur se encuentran por debajo del valor medio nacional.

Tabla 13. Medidas descriptivas de Triptófano (%)

Zona	n	Min	Max	Media	DE	CV	Mediana	Q1	Q3	IQR
Bs As Sur	5	1,16	1,54	1,39	0,15	10,46	1,43	1,36	1,47	0,10
Córdoba Norte	8	1,32	1,73	1,58	0,12	7,66	1,60	1,54	1,64	0,10
Córdoba Sudoeste	15	1,31	1,79	1,54	0,11	7,46	1,55	1,49	1,60	0,11
Entre Ríos	6	1,26	1,85	1,53	0,21	13,79	1,56	1,38	1,60	0,22
NEA	8	1,39	1,88	1,56	0,18	11,79	1,48	1,45	1,59	0,13
NOA	7	1,28	1,58	1,38	0,11	8,24	1,32	1,30	1,43	0,13
Núcleo	45	1,22	1,87	1,54	0,17	10,78	1,51	1,42	1,59	0,17
Santa Fe Centro	6	1,31	1,99	1,73	0,25	14,38	1,80	1,62	1,89	0,27
Total	100	1,16	1,99	1,53	0,17	11,26	1,52	1,42	1,60	0,18

3.5.2.6. Análisis de correlación entre los distintos aminoácidos y entre los aminoácidos y proteína

A continuación, en la Tabla 14, se presenta una matriz donde se pueden observar, por arriba de la diagonal de 1, las probabilidades asociadas a la hipótesis de que el coeficiente de correlación sea distinto de cero. Por lo tanto, valores menores a 0,05 indican que existe relación entre las variables. Por debajo de la diagonal de 1, se pueden observar los coeficientes de correlación estimados a partir de las muestras proporcionadas. Del mismo modo, en el Anexo V⁴, se presentan matrices de gráficos de dispersión, es decir, que se pueden observar varios gráficos de dispersión en la misma matriz. En cada uno de estos gráficos de dispersión pueden observarse las relaciones de a pares de los AA evaluados y también su relación con la proteína evaluada por tecnología NIR y kjeldahl.

4

Tabla 14. Matriz de correlación de variables KAA (%) y contenido de proteína (% bs)

	Prot. NIR	Prot. GL	Triptófano	Lisina	Treonina	Cisteína	Metionina
Proteína NIR	1	0	0,16	0,95	0,84	0,14	1
Proteína GL	0,98	1	0,24	0,72	0,76	0,16	0,89
Tryptofano	-0,14	-0,12	1	0	0,49	0,3	0,59
Lisina	0,01	0,04	0,73	1	0,02	0,36	0,44
Treonina	-0,02	-0,03	-0,07	-0,23	1	0	0
Cysteina	0,15	0,14	-0,1	-0,09	0,56	1	0
Metionina	0	-0,01	0,06	-0,08	0,62	0,58	1

Los números que se encuentran sobre la diagonal principal corresponden a las probabilidades asociadas a las correlaciones que se encuentran debajo de la misma. Se encontraron las siguientes relaciones significativas entre;

- Triptófano y Lisina ($p < 0.001 = 0$) con coeficiente 0,73; relación alta y positiva.
- Lisina y Treonina ($p < 0.001 = 0,02$) con coeficiente -0,23; relación débil y negativa.
- Treonina y Cisteína ($p < 0.001 = 0$) con coeficiente 0,56; relación débil y positiva.
- Treonina y Metionina ($p < 0.001 = 0$) con coeficiente 0,62; relación positiva.
- Metionina y Cisteína ($p < 0.001 = 0$) con coeficiente 0,58; relación débil y positiva.

3.5.3. Análisis estadístico de los Key Aminoácidos

En la Tabla 15 se presentan las medidas descriptivas de KAA presentadas por región. A partir de dichos datos, se determinó que la media nacional fue de 14,36 %, siendo el valor medio mínimo 13,70 % y el máximo 15,74 %. En la Figura 8 se muestra el valor medio de los KAA (%) y la distribución de la proporción de cada uno de los AA que lo constituyen por zonas.

Tabla 15. Medidas descriptivas de los KAA (%) por región

Zona	n	Media	Desvío standard	CV (%)
NOA	7	15,74	1,36	8,66
Santa Fe Centro	6	15,30	1,16	7,55
Córdoba Sudoeste	9	14,58	1,23	8,41
Córdoba Norte	8	14,51	0,44	3,02
Entre Ríos	6	14,24	1,02	7,18
Núcleo	53	14,01	1,15	8,22
Bs As Sur	5	13,77	0,61	4,47
NEA	6	13,70	0,35	2,53
Nacional	100	14,36	1,20	8,37

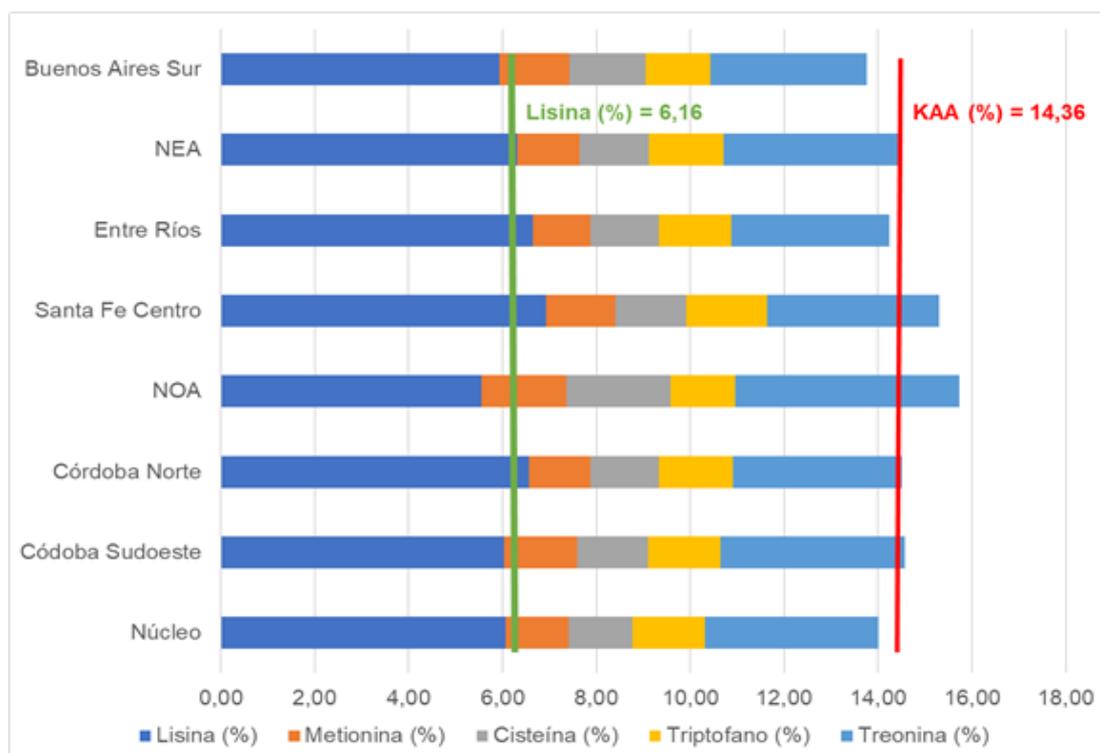


Figura 8. Distribución de los KAA (%) por región.

Al evaluar el comportamiento de las relaciones entre AA (%) y contenido proteico (% bs), se evidencia que los coeficientes de regresión no han sido significativos. Se supone que este comportamiento se debería a la proporcionalidad existente entre el porcentaje de los distintos AA, los KAA y los NKAA con proteína (% bs). No obstante, al analizar la relación existente entre el contenido de AA en sustancia seca (g), KAA (g) y NKAA (g) respecto del contenido proteico (% bs) se evidencia que los coeficientes de regresión son significativos. Para KAA (g) el R^2 fue igual a 38,72, mientras que para NKAA fue de 95,40 y mostrando, además, una tasa marcadamente superior, tal como se muestra en las Figuras 9, 10 y 11 y Tabla 16. Esto significa que al aumentar el contenido proteico de los protos de soja aumentarían los gramos de NKAA en mayor medida que los KAA.

Tabla 16. Coeficientes de correlación y regresión entre AA (% y g) con proteína (% bs)

AA	R^2	p-valor
KAA (%)	0,09	0,776
KAA (g)	38,72	0,000
NKAA (%)	0,08	0,777
NKAA (g)	95,40	0,000
Lisina (%)	0,13	0,722
Lisina (g)	25,74	0,000

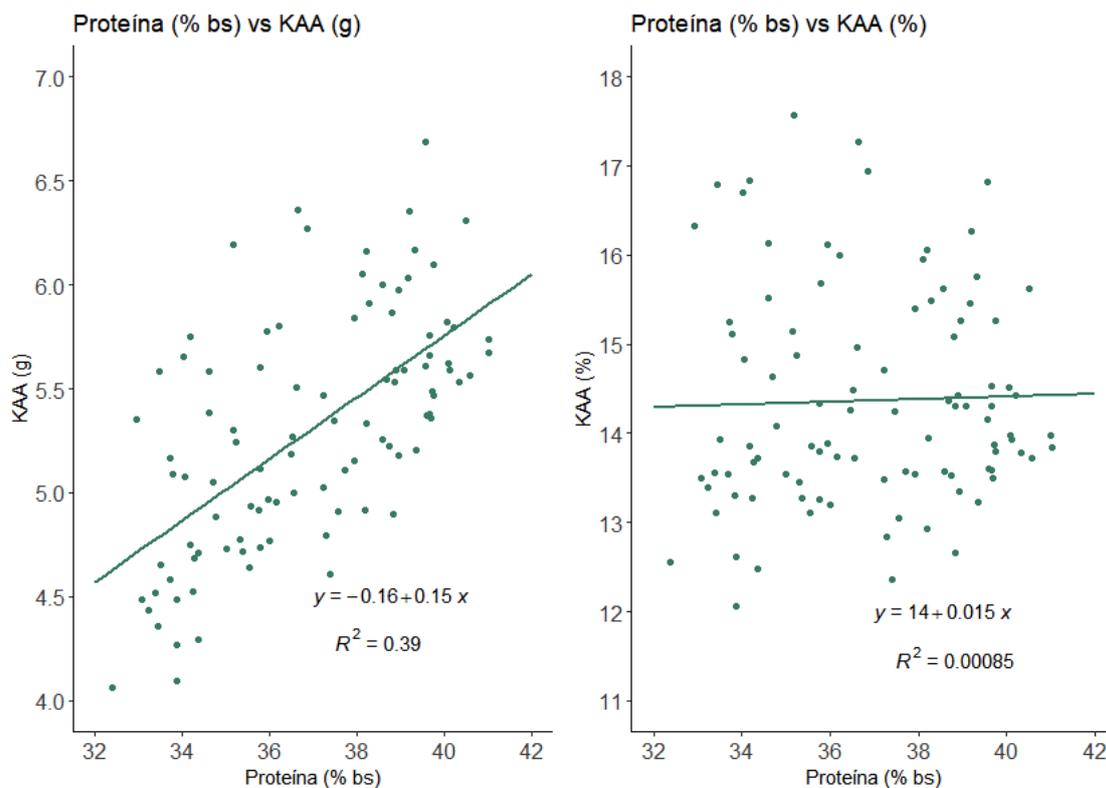


Figura 9. Gráfica de dispersión de KAA (g/100 g muestra) y KAA (%) vs proteína (%)

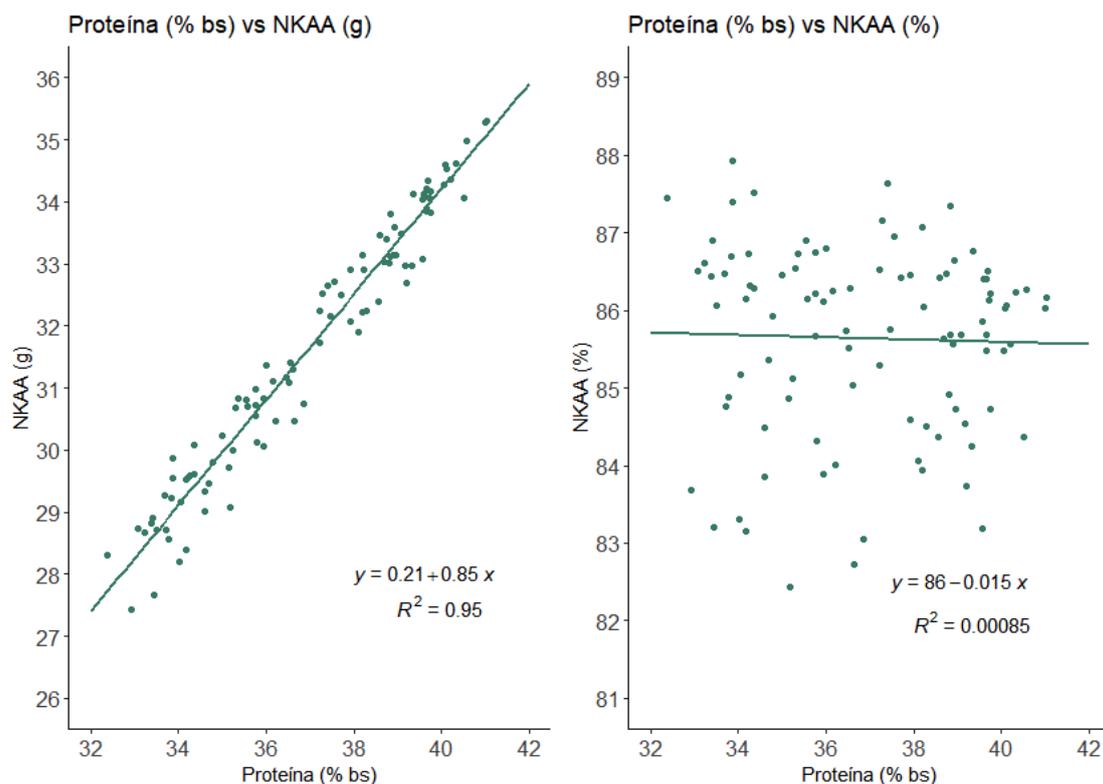


Figura 10. Gráfica de dispersión de NKAAs (g/100 g muestra) y NKAAs (%) vs proteína (%)

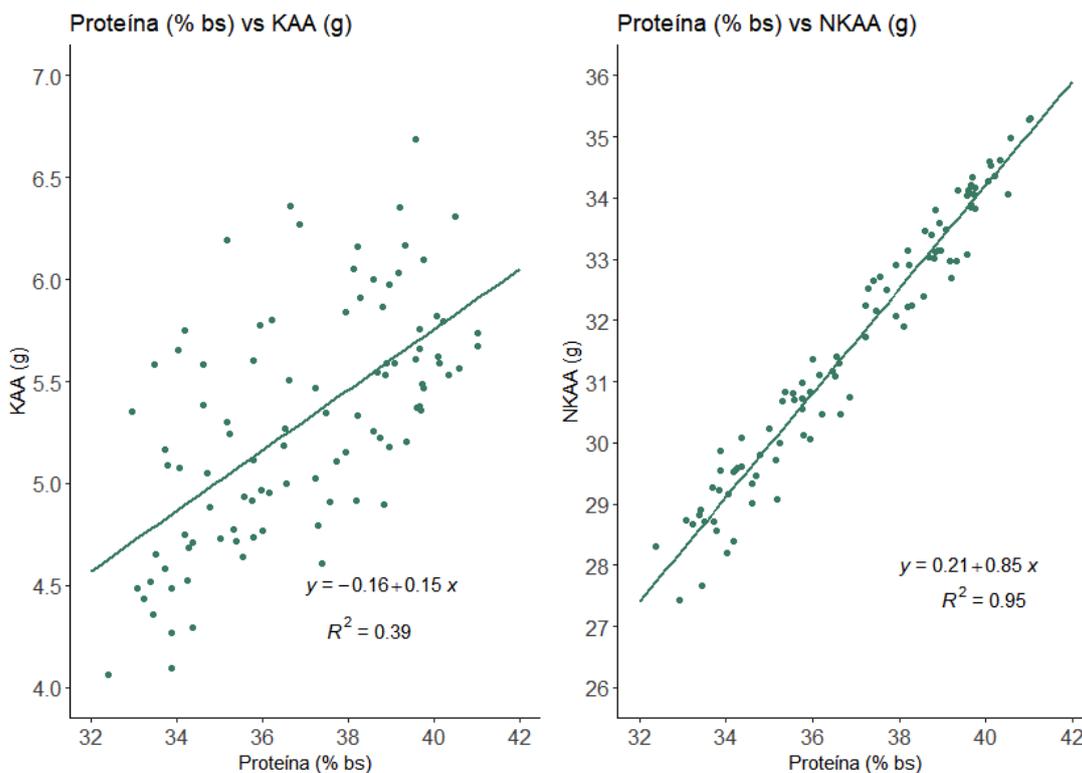


Figura 11. Gráfica de dispersión de KAA (g/100 g muestra) y NKAA (g/100 g muestra) vs proteína (%)

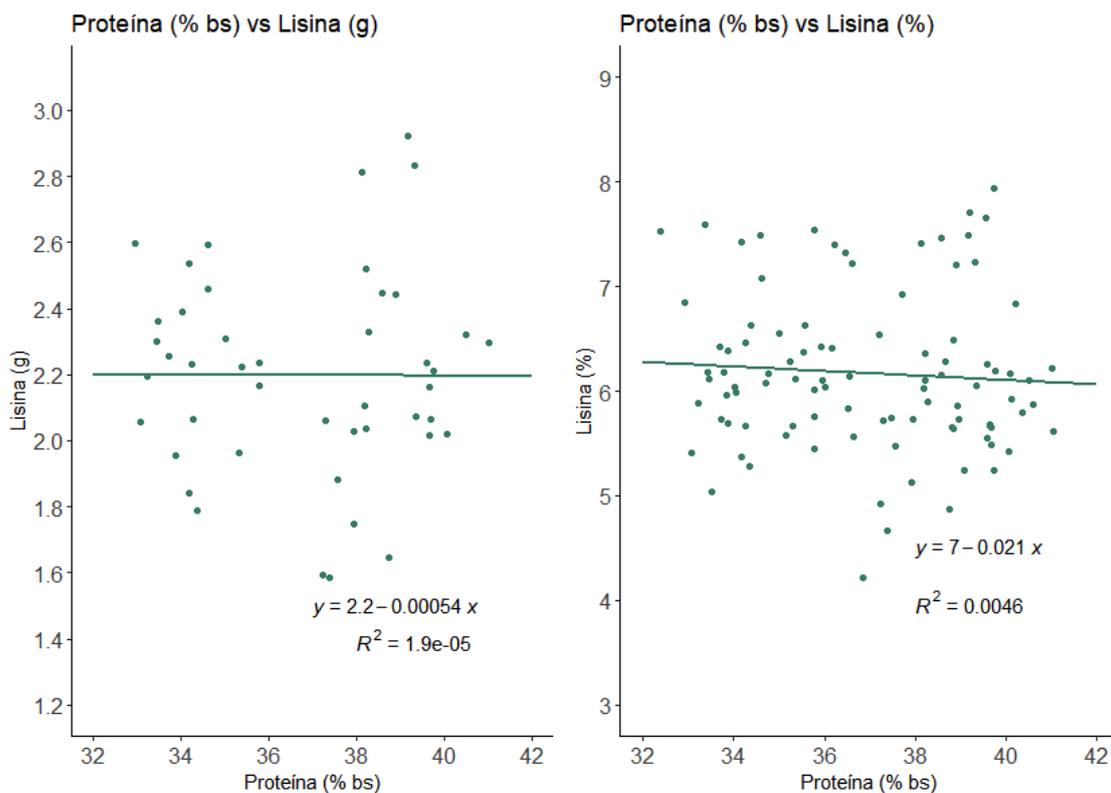


Figura 12. Gráfica de dispersión de lisina (g/100 g muestra) y lisina (g/100 g muestra) vs proteína (%)

El análisis de regresión para cada uno de los AA que constituyen los KAA por separado; lisina (%), metionina (%), cisteína (%), triptófano (%) y treonina (%) vs proteína (%) ha indicado que los coeficientes de regresión lineal no fueron significativos (Anexo V).

Considerando solo la zona núcleo (n=44), se observa un comportamiento similar al detallado anteriormente. La zona núcleo presenta valores de lisina (6,07 %) y KAA (14,01 %) inferiores a los valores medios nacionales evaluados; lisina (6,16 %) y KAA (14,36 %). En el Anexo V se detallan toda la información del resto de los AA.

3.6. Análisis geoestadístico

3.6.1. Análisis de Aceite, Proteína y PROFAT

Para el análisis geoestadístico de aceite, proteína y PROFAT, se utilizaron solamente las muestras que corresponden a la siembra de soja de 1^o (n= 313). Debido a la escasez de muestras, y a su distribución espacial se dividieron las mismas en dos regiones. Una representando a la Zona del NOA, aunque solo se graficó para 3 departamentos (Metán, Anta y Rosario de la Frontera) y la segunda abarcando las zonas Núcleo, Buenos Aires Sur, Córdoba Norte, Córdoba Sudoeste y Santa Fe Centro, y al igual que la otra región, solamente se analizaron los departamentos/partidos en los cuales se recolectaron muestras. En el Anexo VI⁵ se brinda información detallada sobre el método geoestadístico utilizado.

En la Figura 13 se observa la distribución espacial de las variables de proteína, aceite y PROFAT, en las regiones analizadas, se resalta la relación entre PROFAT y aceite, y se observa que en los departamentos de Las Colonias, Castellano y San Martín de la provincia de Santa Fe presentan mayores valores de aceite, mientras que los departamentos de Metán y Rosario de la Frontera en Salta, y Río Primero de Córdoba presentan mayores valores proteicos. Por otro lado, son los departamentos de Anta en Salta y Caseros de Santa Fe quienes presentan mayor porcentaje de PROFAT.

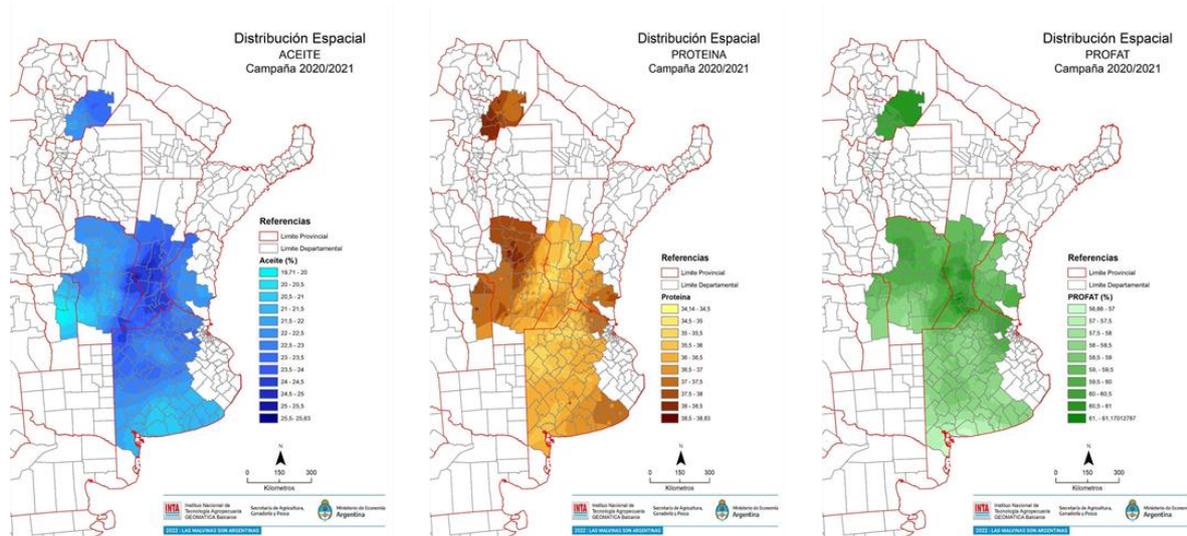


Figura 13. Distribución espacial de las variables de aceite, proteína y PROFAT en Argentina durante la campaña 2021/2022

⁵ <https://docs.google.com/document/d/1QFd1tvEIAgLyLrJIMGT8TB1kIBv9Q3VtqXsZm-Qyct8/edit?usp=sharing>

3.6.2. Análisis de Aminoácidos

Por su parte, similar metodología fue aplicada para el análisis geoestadístico de aminoácidos, con la salvedad que fue analizado solo en la zona núcleo, debido a la cantidad de muestras disponible ($n=53$). Como se aprecia en la Figura 14, los departamentos del sur de la provincia de Santa Fe son donde se expresa mayor contenido de Lisina, Metionina, Triptófano y $\Sigma 5$ Key AA (Lys, Met, Trp). Mientras que la Cisteína (Cys) concentra sus mayores valores en el sudeste de Córdoba. Por su lado, la Treonina (Thr) presenta sus máximos valores en el sur de Santa Fe y Noroeste de la provincia de Buenos Aires (Anexo VI).

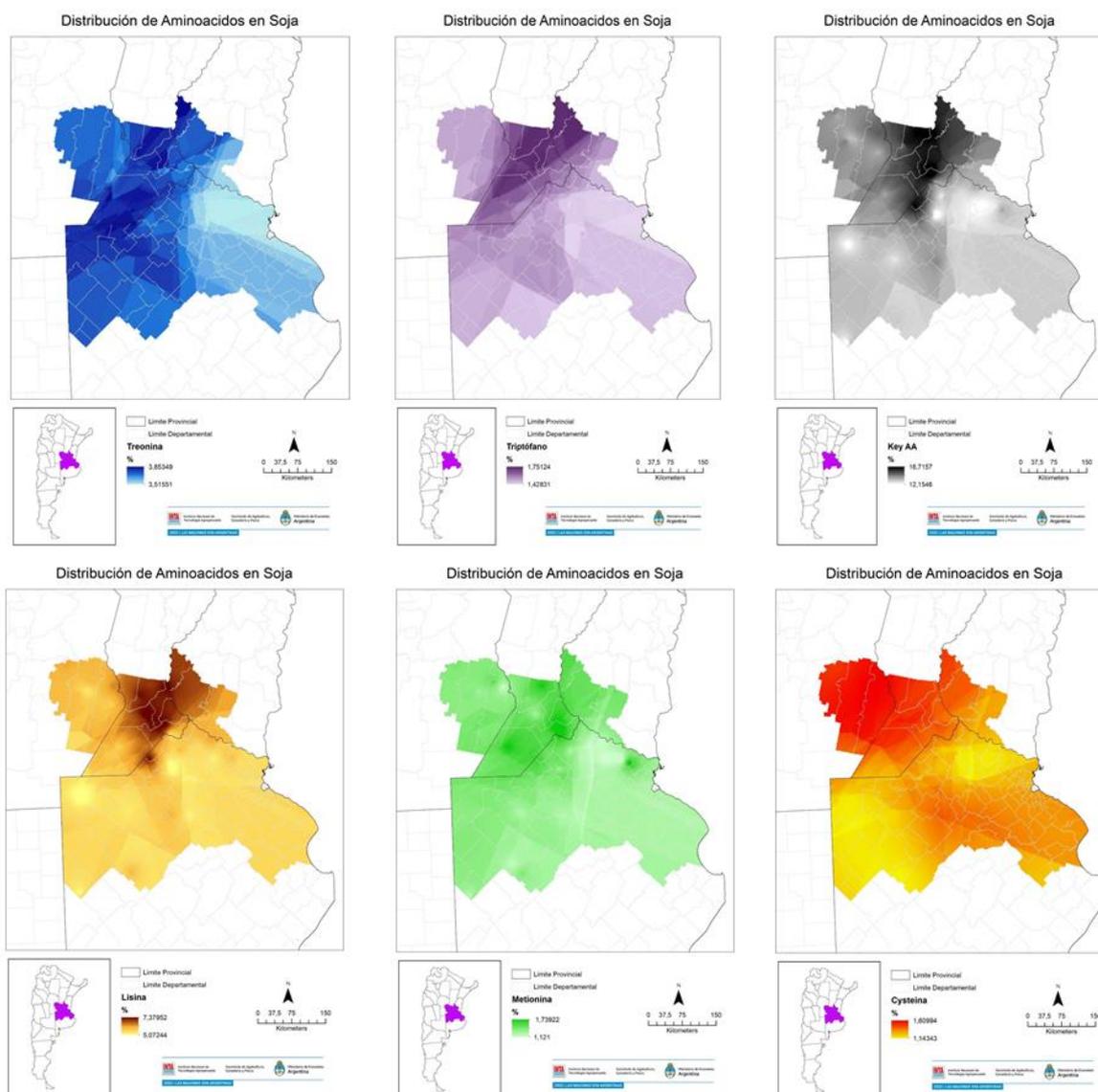


Figura 14. Distribución espacial de las 5 Key AA (Lys, Met, Cys, Thr, Trp %) y su sumatoria (%) en la Zona Núcleo de Argentina durante la campaña 2021/2022

3.6.3. Validación

La metodología utilizada presentó un ajuste aceptable, considerando el tamaño de la muestra y la distribución espacial. El modelo exponencial utilizado para el análisis de aceite,

PROFAT, triptófano (g), lisina (g), lisina (%), treonina (%), cisteína (g), metionina (g), metionina (%) y la sumatoria presentó un mejor ajuste, con un error medio cuadrático (EMC) de 1,3; 1,5; 0,006; 0,11; 0,49; 0,41; 0,02; 0,004; 0,0 y 1,11 (%), respectivamente. Mientras que el modelo circular describió mejor la distribución de proteína, triptófano % y treonina % (EMC=1,9; 0,01 y 0,07 %). Mientras que el modelo esférico fue quien mejor ajustó la distribución espacial de cisteína % (EMC=0,07 %). Mayor información en el Anexo VI.

4. DISCUSION Y CONCLUSION

4.1. Análisis de Aceite Proteína y PROFAT Argentina

Los resultados obtenidos en este estudio para los parámetros proteína, aceite y PROFAT para poroto de soja son similares a los reportados en otros estudios nacionales para la misma campaña (2020/21). No obstante, es importante remarcar que los resultados presentados corresponden a muestras de variedades recolectadas en campos de productores a diferencia de los otros estudios que se evalúan muestras compuestas de acopios o de ensayos de variedades de reciente liberación y pre-comerciales (RECSO). Esta similitud de resultados sustenta a los resultados que históricamente se vienen presentando, tal como se puede observar en la Tabla 17.

Tabla 17. Comparativa de resultados.

Parámetro	Relevamiento ETAPA 2	Relevamiento Acopios MJ – Zona Núcleo	RECSO Nacional
Proteína (% bs)	36,6	36,2	36,7
Aceite (% bs)	23,0	22,8	22,4
Profat (% bs)	59,0	59,0	59,1

Según de Borja Reis (2020) el promedio de proteína histórica 1980-2014 para EE.UU. ha sido igual a 40,11 (% bs) valor que resulta superior al encontrado en el promedio histórico argentino 37,4 (% bs) evaluado en los últimos 20 años (Chialvo et al, 2022).

4.2. Variabilidad dentro de nuestro país, entre zonas, potencialidades de cada zona

Mediante el análisis geostático, áreas con calidad diferencial pudieron ser definidas usando una interpolación kriging ordinaria siendo esta la primera aproximación para definir la distribución espacial de la composición del grano a partir de soja cosechada a campo por productores. A su vez permitió observar la variabilidad existente en Argentina y dentro de la zonificación realizada, por lo cual, sería posible ajustar esta zonificación mediante el análisis de la influencia de otros factores medioambientales y socioeconómicos. El uso de la interpolación de Kriging ordinario fue validado por Radocaj et al. (2020) con la delimitación de zonas de acuerdo a la textura de suelo para el cultivo de la soja y por Di Mauro et al. (2018) explicando la variabilidad de rendimientos. A diferencia de Bosaz et al. (2019), no se encontró autocorrelación espacial para la concentración de proteínas en soja de 1era. Esto puede deberse a las diferentes áreas de estudio utilizadas, Bosaz et al. (2019) centraron su foco de estudio en la zona núcleo, mientras que este trabajo abarcó toda las regiones productivas. Asimismo, diferencias en las zonas se demarcan, al ser este trabajo solo de una campaña

agrícola (2020/2021), mientras que el citado trabajo utilizó datos entre 2012 a 2019. Por otro lado, ambos modelos han comprobado que el modelo circular es el que mejor ajusta la distribución de proteínas.

4.3. Análisis de Aminoácidos entre Argentina y el mundo

Actualmente son escasos los antecedentes sobre perfil de AA de soja en la Argentina, es por ello, que las referencias conocidas provienen de publicaciones internacionales relacionadas con la alimentación animal, tal como en el caso de Ibañez et al (2020) y García Rebollar et al (2016). En sus estudios se comparó la composición química y KAA de HS provenientes de los principales países exportadores. En la Tabla 19 se pueden observar los valores de KAA (%) para los principales países exportadores según ambas referencias, siendo 45,5 el valor de proteína cruda (%) y 14,3 (%) de KAA para Argentina.

Tabla 19. Valores medios de KAA (%) para harinas de soja de diferentes procedencias

Referencia	Argentina	USA	Brasil
Ibañez et al (2020)	14,3 (14,0 -14,8)	14,4 (14,1 - 14,8)	14,1 (13,5 - 14,3)
García Rebollar et al (2016)	14,3 (14,0 - 14,6)	14,3 (14,0 - 14,8)	14,1 (13,8 - 14,7)

Del mismo modo, en la Tabla 20 se presentan los valores de KAA (%) y sus respectivos AA por separado para ambas etapas del proyecto en ejecución; Etapa 1 evaluación de harinas de soja y Etapa 2 porotos de soja, y las referencias citadas anteriormente.

Tabla 20. Valores de KAA (%) en las etapas evaluadas (1 y 2) y referencias bibliográficas

Referencia	ETAPA 1 – HS Arg.	ETAPA 2 - poroto	Ibañez et al (2020) – HS Arg.	García Rebollar (2016) – HS Arg.
KAA (%)	14,41	14,36 (12,06 - 17,56)	14,3 (14,0 -14,8)	14,3 (14,0 - 14,6)
Lisina (%)	6,06 (5,51 - 6,6)	6,16 (4,22 - 7,94)	6,19 (5,51 - 6,39)	6,11 (5,91 - 6,32)
Metionina (%)	1,43 (1,26 - 1,52)	1,41 (0,95 - 2,55)	1,37 (1,33 - 1,46)	1,37 (1,31 - 1,41)
Cisteína (%)	1,5 (1,26 - 1,65)	1,49 (0,91 - 2,62)	1,45 (1,28 - 1,58)	1,51 (1,41 - 1,59)
Triptófano (%)	1,35 (1,21 - 1,97)	1,53 (1,16 - 1,99)	1,39 (1,23 - 1,56)	1,37 (1,33 - 1,42)
Treonina (%)	4,08 (3,31 - 4,09)	3,77 (2,31 - 5,61)	3,89 (3,76 - 4,00)	3,94 (3,86 - 4,03)

Respecto a KAA (%), los resultados de la Etapa 2 indican que su valor medio fue igual a 14,36 (%), siendo 13,70 (%) el valor más bajo en NEA y 15,48 (%) el valor más alto en NOA. Además, podemos observar que los valores de KAA (%) de las Etapas 1 y 2, obtenidos de harinas y porotos de soja, son levemente superiores a los reportados por Ibañez et al (2020) y García Rebollar et al (2016); 14,3 % y 14,4 %, respectivamente.

En el caso de la lisina, los valores reportados en HS fueron 6,06 % para la Etapa 1, 6,19 % para Ibañez et al (2020) y 5,95 % para García Rebollar et al (2016), mientras que el valor medio de lisina en la Etapa 2 fue 6,16 %. Se encontró consistencia entre los valores de lisina (%) de la Etapa 1 (HS), valores de referencia internacional y la Etapa 2 (poroto); 6,06 %, 6,19 %, 6,11 % y 6,16 %, respectivamente. Sumado a ello, se puede observar que el valor medio de la zona núcleo (6,07 %), que es la zona que mayor cantidad de poroto aporta a la industria, es consistente con el valor medio de la etapa anterior (6,06 %). Respecto a la metionina, los valores medios obtenidos en Etapa 1 y 2 (1,43 y 1,41 %) son levemente superiores a los reportados en la bibliografía (1,37 %). El resto de los AA (%) tuvieron valores medios similares a los de referencia.

Dentro de los antecedentes nacionales, Carrera et al (2011), quien estudió el efecto de las variables climáticas en la composición de soja, plantea que la composición de aminoácidos de los granos de soja se ve fuertemente afectada por factores ambientales durante el periodo de llenado de semillas y, además, que cada AA se comporta de forma diferente según las condiciones ambientales, lo que indica efectos compensatorios entre ellos. Este último explicaría perfectamente el comportamiento de los KAA (%) por región mostrados en la Tabla 16. Es importante resaltar que este estudio considera muestra de ambientes contrastantes y, además, se han seleccionado para el estudio de AA sojas de alta y baja proteína de cada región. Es por ello que al evaluar las interacciones entre proteína y AA es difícil observar una tendencia. Se puede observar que en el rango comprendido entre 35 - 37 % de proteína, las curvas se mantienen estables. Al existir pocos datos dentro de este rango, se podría estar sesgando u ocultando una tendencia negativa.

5. CONCLUSIONES

5.1. Conclusiones generales

- i. En promedio, las muestras analizadas (n=544) mostraron $36,6 \pm 1,5$ (% bs, base seca) de proteína, $23,0 \pm 1,6$ (% bs) de aceite y $59,5 \pm 1,4$ (% bs) de PROFAT.
- ii. Entre las áreas estudiadas, Santa Fe Centro tuvo el mayor nivel promedio de aceite (24,7 % bs) y el NOA los mayores promedios de proteína (38,0 % bs) y PROFAT (60,6 % bs).
- iii. Buenos Aires Sur mostró el mínimo promedio de aceite (21,0 % bs) y PROFAT (57,9 % bs), mientras que Santa Fe Centro el mínimo promedio de proteína (35,7 % bs).
- iv. Áreas con calidad diferencial pudieron ser definidas usando una interpolación kriging, siendo esta la primera aproximación para definir la distribución espacial de la composición del grano a partir de soja cosechada a campo por productores

- v. El modelo de interpolación Kriging circular fue el que mejor ajustó para el mapeo de proteína, mientras que, para mapear aceite y PROFAT, el exponencial fue el más preciso.
- vi. Para mapear los valores de aminoácidos, específicamente en región Núcleo, el modelo de interpolación Kriging esférico mostró un mejor ajuste para Cisteína %, el circular ajustó mejor para Triptófano % y Treonina (g), y el exponencial para el resto de aminoácidos estudiados, incluida los KAA.
- vii. De acuerdo al mapa de interpolación, y coincidiendo con el análisis estadístico tradicional, el NOA, más específicamente el sudeste de Salta, y el norte de Córdoba son las regiones que presentan los valores más altos de proteína (superior al 37,0 %) y de PROFAT (superior al 59,5 %). También mostró alto PROFAT el Noroeste de la provincia de Buenos Aires. Finalmente, el más alto nivel de aceite (superior al 23,5 %) se observó en Santa Fe Centro.
- viii. Los promedios porcentuales obtenidos para KAA y no KAA fueron 14,36 % y 85,64 %, respectivamente.
- ix. El valor medio nacional de Lisina fue 6,16 %, con los promedios más altos en muestras provenientes de Santa Fe Centro y los más bajos en el NOA.
- x. Se encontraron relaciones positivas entre triptófano-lisina y metionina-cisteína. Por otra parte, existe una relación negativa entre la lisina y la treonina.
- xi. Los aminoácidos azufrados (metionina y cisteína) presentaron su valor máximo en NOA.
- xii. El análisis de regresión lineal para cada uno de los KAA; lisina (%), metionina (%), cisteína (%), triptófano (%) y treonina (%) vs proteína (% bs), ha indicado que sus coeficientes de regresión fueron no significativos. Esto indica que se mantiene la proporcionalidad de cada uno de ellos.
- xiii. Dentro del rango de proteínas explorado, al incrementar el contenido de proteínas en grano se observa que los NKAA (g/100 g muestra) aumentaron en mayor proporción que los KAA (g/100 g muestra). Este comportamiento sugiere que la variación de KAA (g/100 g muestra) estaría menos asociada a la variación del contenido de proteínas (% bs) respecto a la variación de NKAA (g/100 g muestra).
- xiv. Al considerar dos parámetros de calidad internacional de poroto de soja y harinas como lisina (%) y KAA (%), los obtenidos en este estudio para lisina (6,16 %) y KAA (14,36 %) comparados con las referencias internacionales basadas en harinas de soja procedentes de Argentina, concluimos que ambos valores son similares a los reportados por otros autores.
- xv. Respecto al concepto de calidad de grano, este es muy amplio que no solo se sesga a los objetivos puntuales de este trabajo. Por lo cual, se propone continuar explorando diferentes precursores y otras relaciones de aa que resultan útiles para la industria, la nutrición y diferentes alcances.

- xvi. Los porotos argentinos presentan buenos valores en otros parámetros que al día de hoy no están siendo considerados, como ácido glutámico y aspártico que indica la solubilidad de proteínas.

5.2. Próximos pasos

- i. A partir de los resultados obtenidos, consideramos conveniente repetir este estudio aumentando el número de muestras evaluadas (N) y considerando otras campañas agrícolas para robustecer el trabajo
- ii. Modelar el comportamiento de las variables consideradas en diversas campañas, de manera que se pueda corroborar la subdivisión de cada zona, posibilitando la utilización de denominación de origen para diferenciar la producción en estas zonas.
- iii. Considerar variables tales como influencia de factores de manejo, biofísicos y socioeconómicos.
- iv. En la campaña 2021/22 INTA ha realizado un segundo muestreo nacional, aprovechando el interés impulsado por los resultados obtenidos, en el cual se han recolectado muestras de las principales zonas de cultivo de soja. Estas muestras fueron analizadas por NIR y se cuenta actualmente con información de proteína, aceite y PROFAT de un segundo año de evaluación.

6. BIBLIOGRAFÍA

Assefa, Y., Bajjalieh, N., Archontoulis, S. et al. (2018). Spatial Characterization of Soybean Yield and Quality (Amino Acids, Oil, and Protein) for United States. *Sci Rep* 8, 14653. <https://doi.org/10.1038/s41598-018-32895-0>

Badui Dergal, S. 2006. Química de los alimentos. Cuarta edición. México. 736 p. ISBN: 970-26-0670-5.

Badui Dergal, S. 2006. Química de los alimentos. Cuarta edición PEARSON EDUCACIÓN, México, 2006 ISBN: 970-26-0670-5. Área: Química. Páginas: 736

Benavídez, R, ME González, DM Fresoli, DJ Santos, ME Soro. 2008. Evolución del contenido de proteína y aceite en grano de soja en Argentina entre las campañas 1999-2000 y 2005-06. *Revista Análisis de Semillas Año 2 No 6*, ISSN 1851-1678. pp 102-105.

Bosaz, L., Gerde, J., Borrás, L., Cipriotti, P., Ascheri, L., Campos, M., Gallo, S., Rotundo, J. 2019. Management and environmental factors explaining soybean seed protein variability in central Argentina. *Field Crops Research* 240: 34-43.

CAC. Cámara Arbitral de Cereales, 2018. Bolsa De Comercio De Rosario. Available at URL. <http://www.cac.bcr.com.ar/pages/infojaindustrial.aspx>.

Carrera CS.; Reynoso, CM.; Funes, GJ.; Martínez, MJ.; Dardanelli, J.; Resnik, SL. 2011. Amino acid composition of soybean seeds as affected by climatic variables. *Pesq. agropec. bras.*, Brasília, v.46, n.12, p.1579-1587

Chialvo, E., Herrero, R. M., Mir, L. R., Pronotti, M. S., y Mansilla, G. J. (2021). Productividad y calidad de la soja en la zona núcleo-sojera. Campaña 2020/21. Ediciones INTA.

Chialvo, E.; Herrero, R.; Mir, L.; Berra, O.; Pronotti, M. y Mansilla, G. 2022. Productividad y calidad de la soja en la zona núcleo-sojera. Campaña 2021/22. Ediciones INTA.

Cuniberti, M., Herrero, R., Mir, L., Chialvo, E., Berra, O., Macagno, S., & Juárez, I. E. M. (2018). Productividad y calidad de la soja en la zona Núcleo-Sojera. Soja 2018, 57.

de Abellería D, Verón S, Bancharo S, Iturralde Elortegui M, Valiente S, Puig O, Murray F, Martini JP, Zelaya K, Maidana D, Varlamoff N, Peiretti J, Benedetti P, Portillo J, Melilli M, Maidana E, Goytía Y, Sapino V, Regonat A, Cracogna F, Espíndola C, Parodi M, Colombo F, Scarel J, Ayala J, Martins L, Basanta M, Rausch A, Almada G, Boero L, Calcha J, Chiavassa A, Lopez de Sabando M, Kurtz D, Marini F, Sarraillhe S, Petek M, Propato T, Ferraina A, Pezzola A, Winschel C, Muñoz S, Mesa J, Volante J. 2021. Mapa Nacional de Cultivos campaña 2020/2021 Versión 1 Publicación N°3 INTA. https://inta.gob.ar/sites/default/files/informe_mapa_nacional_de_cultivos_campana_2020_2021_0.pdf

de Borja Reis, A.F., Tamagno, S., Moro Rosso, L.H. et al. (2020). Historical trend on seed amino acid concentration does not follow protein changes in soybeans. Sci Rep 10, 17707 (2020). <https://doi.org/10.1038/s41598-020-74734-1>

Di Mauro, G., Cipriotti, P., Gallo, S., Rotundo, J. 2018. Environmental and management variables explain soybean yield gap variability in Central Argentina. European Journal of Agronomy 99: 186-194.

FAO, 2018. Food and Agricultural Organization of the United Nations. <http://faostat3.fao.org/home>.

Fennema, O.R. 2000. Química de los Alimentos. 2a Edición. Acribia, Ed. USA. ISBN 9788420009148.

Fennema, O.R. 2008. Fennema's Food Chemistry Fourth Edition (edited by Srinivisan D, Parking, K and Fennema, O). Food Science and Technology; 169

García-Rebollar P. et al. (2016). Influence of the origin of the beans on the chemical composition and nutritive value of commercial soybean meals. Animal Feed Science and Technology 221 (2016) 245–261. <http://dx.doi.org/10.1016/j.anifeedsci.2016.07.007>

Ibáñez, M.A., de Blas, C., Cámara, L., Mateos, G. 2020. Chemical composition, protein quality and nutritive value of commercial soybean meals produced from beans from different countries: A meta-analytical study. Animal Feed Science and Technology, Volume 267, 2020, 114531, ISSN 0377-8401, <https://doi.org/10.1016/j.anifeedsci.2020.114531>.

INASE. Consulta en línea <https://gestion.inase.gob.ar/consultaGestion/gestiones>.

Lagos, L.V., Stein, H.H., 2017. Chemical composition and amino acid digestibility of soybean meal produced in the United States, China, Argentina, Brazil, or India. J. Anim. Sci. 95, 1626–1636.

Medic, J., Atkinson, C. & Hurburgh, C.R. (2014). Current Knowledge in Soybean Composition. J Am Oil Chem Soc 91, 363–384. <https://doi.org/10.1007/s11746-013-2407-9>

Radocaj, D., Jurisic, M., Zebec, V., Plascak, I. 2020. Delineation of Soil texture suitability zones for soybean cultivation: a case study in continental Croatia. Agronomy 10, 823:1-22.

Shea Z., Singer M.S. & Zhang B. (2020). Soybean production, versatility and improvement. Legume crops, Prospects, Production and Uses. Chapter 3. DOI: <http://dx.doi/10.5772/intechopen.91778>

SISA 2021. https://www.argentina.gob.ar/sites/default/files/anexoi_soja2020_2021.pdf. Visitado 17-04-2023

ANEXO I**DETALLE DE LOS DEPARTAMENTOS Y/O PARTIDOS COMPRENDIDOS EN CADA ZONA HOMOGÉNEA DE PRODUCCIÓN**

1. Buenos Aires Sur: abarca 39 partidos del sur de la provincia de Buenos Aires. Esta zona se caracteriza por su actividad mixta, encontrando en el norte una región de ganadería extensiva de pastoreo, mientras que la zona del Sudeste es un sistema de lomadas, con una actividad predominantemente agrícola.
2. Córdoba Norte: corresponde 11 departamentos del norte de la provincia de Córdoba, que se encuentran al este de las Sierras.
3. Córdoba Sudoeste: corresponde a 8 departamentos del sudoeste de la provincia de Córdoba, 6 departamentos del este de la provincia de San Luis (Ayacucho, Libertadores Gral San Martín, Cnel Pringles, Chacabuco, Gral Pedernera y Junin) y 5 departamentos del norte de la provincia de La Pampa (Rancul, Chapaleufu, Realico, Trenel, Maraco).
4. Entre Ríos: Los departamentos de la citada provincia, a excepción de los del sudoeste (Diamante, Victoria y Gualeguay), se identifican en esta zona, con predominio de suelos con alto contenido de arcilla en superficie.
5. NEA. Corresponde a los departamentos de las provincias de Formosa, Chaco, 17 departamentos del este de Santiago del Estero y 4 departamentos del norte de Santa Fe (San Javier, Gral Obligado, Vera, 9 de Julio). Esta región pertenece al Chaco Seco, una región donde la agricultura ha avanzado en los últimos años y ha llevado a tener procesos de deforestación.
6. NOA. Abarca las provincias de Salta, Tucumán, 10 provincias del oeste de Santiago del Estero y el departamento de Santa Rosa en Catamarca. En esta región, al igual que en el NEA, se ha acelerado el proceso de agriculturización impulsado por la sojización, en detrimento de los bosques naturales.
7. Núcleo: abarca el norte de la provincia de Buenos Aires (92 partidos), 7 departamentos del sur de la provincia de Santa Fé, 3 departamentos del este de la provincia de Córdoba (Unión, Gral. San Martín, Marco Juárez) y 3 departamentos del sudoeste de Entre Ríos (Victoria, Diamante y Gualeguay). En esta zona es la zona por excelencia agrícola.
8. Santa Fé centro: abarcan los departamentos San Jerónimo, San Martín, Capital, Castellanos, Garay, Las Colonias, San Justo y San Cristobal de la provincia de Santa Fe.

Tabla 1. Departamentos comprendidos en cada zona

Zona	Provincia	Departamento
Buenos Aires Sur	BUENOS AIRES	LAS FLORES
		MAIPU
		GENERAL GUIDO
		GENERAL JUAN MADARIAGA
		TANDIL
		CORONEL SUARES
		ADOLFO GONZALES CHAVES

		BAHIA BLANCA
		GENERAL ALVARADO
		GENERAL PUEYRREDON
		MAR CHIQUITA
		GENERAL LAVALLE
		RAUCH
		TORNQUIST
		CORONEL PRINGLES
		SAN CAYETANO
		TRES ARROYOS
		CORONEL DORREGO
		NECOCHEA
		ADOLFO ALSINA
		PUAN
		VILLARINO
		OLAVARRIA
		GENERAL ALVEAR
		GENERAL LA MADRID
		BALCARCE
		TAPALQUE
		DOLORES
		AZUL
		DAIREAUX
		AYACUCHO
		GUAMINI
		LAPRIDA
		BENITO JUAREZ
		SAAVEDRA
		LOBERIA
		TORDILLO
		PATAGONES
		CORONEL ROSALES
Cordoba Norte	CORDOBA	TULUMBA
		SANTA MARIA
		RIO SEGUNDO
		SAN JUSTO
		RIO SECO
		SOBREMONTTE
		CAPITAL
		COLON
		ISCHILIN
		TOTAL
		RIO PRIMERO
Cordoba Sudoeste	SAN LUIS	AYACUCHO

		LIBERTADOR GENERAL SAN MARTIN	
		CORONEL PRINGLES	
		CHACABUCO	
		GENERAL PEDERNERA	
		JUNIN	
	LA PAMPA	RANCUL	
		CHAPALEUFU	
		REALICO	
		TRENEL	
		MARACO	
	CORDOBA	SAN ALBERTO	
		TERCERO ARRIBA	
		RIO CUARTO	
		PRESIDENTE ROQUE SAENZ PENA	
		GENERAL ROCA	
JUAREZ CELMAN			
SAN JAVIER			
CALAMUCHITA			
Entre Rios		ENTRE RIOS	CONCORDIA
			FEDERAL
	GUALEGUAYCHU		
	ISLAS DEL IBICUY		
	URUGUAY		
	PARANA		
	COLON		
	VILLAGUAY		
	SAN SALVADOR		
	NOGOYA		
	TALA		
	FEDERACION		
	LA PAZ		
	FELICIANO		
	NEA		CHACO
QUITILIPÍ			
PRESIDENCIA DE LA PLAZA			
O'HIGGINS			
SARGENTO CABRAL			
INDEPENDENCIA			
MAYOR LUIS J. FONTANA			
FRAY JUSTO SANTA MARIA DE ORO			
1 DE MAYO			
12 DE OCTUBRE			
GENERAL GUEMES			
9 DE JULIO			

		SAN FERNANDO
		MAIPU
		ALMIRANTE BROWN
		SAN LORENZO
		2 DE ABRIL
		CHACABUCO
		25 DE MAYO
		LIBERTAD
		COMANDANTE FERNANDEZ
		GENERAL BELGRANO
		GENERAL DONOVAN
		LIBERTADOR GENERAL SAN MARTIN
		BERMEJO
	SANTIAGO DEL ESTERO	ALBERDI
		JUAN F. IBARRA
		GENERAL TABOADA
		RIVADAVIA
		OJO DE AGUA
		COPO
		SARMIENTO
		QUEBRACHOS
		BELGRANO
		AGUIRRE
		FIGUEROA
		SALAVINA
		AVELLANEDA
		MORENO
		SAN MARTIN
		MITRE
		ATAMISQUI
	FORMOSA	PILAGAS
		LAISHI
		PIRANE
		FORMOSA
		PILCOMAYO
		PATINO
		BERMEJO
		MATACOS
		RAMON LISTA
	SANTA FE	SAN JAVIER
		GENERAL OBLIGADO
		VERA
		9 DE JULIO
NOA	SANTIAGO DEL ESTERO	CHOYA

		RIO HONDO
		BANDA
		GUASAYAN
		ROBLES
		CAPITAL
		LORETO
		PELLEGRINI
		SILIPICA
		JIMENEZ
	SALTA	CACHI
		CHICOANA
		LA VINA
		METAN
		MOLINOS
		GENERAL GUEMES
		SAN CARLOS
		CERRILLOS
		LA CALDERA
		ANTA
		ROSARIO DE LA FRONTERA
		CAPITAL
		LA CANDELARIA
		GUACHIPAS
		CAFAYATE
		SANTA VICTORIA
		GENERAL JOSE DE SAN MARTIN
		RIVADAVIA
		ROSARIO DE LERMA
		IRUYA
		LOS ANDES
		ORAN
		LA POMA
	CATAMARCA	SANTA ROSA
	TUCUMAN	LA COCHA
		GRANEROS
		JUAN BAUTISTA ALBERDI
		RIO CHICO
		SIMOCA
		CHICLIGASTA
		MONTEROS
		LEALES
		FAMAILLA
		CAPITAL
		CRUZ ALTA

Nucleo	ENTRE RIOS	LULES
		YERBA BUENA
		TAFI VIEJO
		BURRUYACU
		TRANCAS
		TAFI DEL VALLE
	BUENOS AIRES	VICTORIA
		DIAMANTE
		GUALEGUAY
		SALADILLO
		ROJAS
		ESCOBAR
		JOSE C. PAZ
		SAN ANTONIO DE ARECO
		MORENO
		LA MATANZA
		GENERAL PAZ
		MAGDALENA
		VICENTE LOPEZ
PERGAMINO		
SALLIQUELO		
BERAZATEGUI		
TIGRE		
GENERAL PINTO		
LEZAMA		
CHASCOMUS		
PUNTA INDIO		
CAMPANA		
PELLEGRINI		
GENERAL VILLEGAS		
CASTELLI		
ALMIRANTE BROWN		
LA PLATA		
PRESIDENTE PERON		
BRANDSEN		
CARLOS CASARES		
CARLOS TEJEDOR		
25 DE MAYO		
9 DE JULIO		
PEHUAJO		
TRENQUE LAUQUEN		
GENERAL BELGRANO		
MONTE		
ROQUE PEREZ		

MERLO
ESTEBAN ECHEVERRIA
ARRECIFES
CAPITAN SARMIENTO
CARMEN DE ARECO
SALTO
EXALTACION DE LA CRUZ
SAN MIGUEL
SAN ANDRES DE GILES
JUNIN
MERCEDES
PILAR
LUJAN
HURLINGHAM
MORON
CHACABUCO
MALVINAS ARGENTINAS
GENERAL SAN MARTIN
LINCOLN
ITUZAINGO
GENERAL RODRIGUEZ
SUIPACHA
TRES DE FEBRERO
CHIVILCOY
ALBERTI
BRAGADO
LANUS
GENERAL LAS HERAS
MARCOS PAZ
LOMAS DE ZAMORA
GENERAL VIAMONTE
EZEIZA
FLORENCIO VARELA
NAVARRO
CANUELAS
SAN VICENTE
LOBOS
QUILMES
PILA
BOLIVAR
HIPOLITO YRIGOYEN
TRES LOMAS
ENSENADA
AVELLANEDA

		SAN ISIDRO	
		COLON	
		BERISSO	
		SAN FERNANDO	
		ZARATE	
		RAMALLO	
		SAN NICOLAS	
		LEANDRO N. ALEM	
		GENERAL ARENALES	
		RIVADAVIA	
		BARADERO	
		SAN PEDRO	
		FLORENTINO AMEGHINO	
		CORDOBA	UNION
		GENERAL SAN MARTIN	
		MARCOS JUAREZ	
		SANTA FE	GENERAL LOPEZ
		CONSTITUCION	
		CASEROS	
		ROSARIO	
		SAN LORENZO	
Santa Fe Centro	SANTA FE	IRIONDO	
		BELGRANO	
		SAN JERONIMO	
		SAN MARTIN	
		LA CAPITAL	
		CASTELLANOS	
		GARAY	
		LAS COLONIAS	
		SAN JUSTO	
		SAN CRISTOBAL	

ANEXO II

Análisis químicos y por NIR

Se trabajó con una calibración global de la empresa FOSS que fue robustecida con muestras de soja de Argentina y ajustada por INTA Anguil y Marcos Juárez en el año 2021. Se obtuvieron datos de humedad, proteína y materia grasa tal cual, que luego fueron corregidos a base seca. El parámetro Profat es calculado a partir de los datos de proteína y materia grasa corregidos a base seca. Los análisis de laboratorio se realizaron en las unidades de INTA (EEA Marcos Juárez y EEA Oliveros). Para el análisis de aminoácidos, se seleccionaron 100 muestras (15 %) considerando su distribución geográfica según la relación con la superficie sembrada de soja a nivel nacional y los resultados de concentración proteica. (Tabla 1).

Tabla 1. Detalle de selección de muestras para determinación de aminoácidos

Provincia	Nº muestras	Nº muestras alta proteína	Nº muestras baja proteína
Santa Fe - Sur	8	4	4
Santa Fe - Centro	8	4	4
Santa Fe - Norte	2	1	1
Buenos Aires - Sur	5	3	2
Buenos Aires - Norte	25	13	12
Córdoba - NO	5	3	2
Córdoba - NE	3	2	1
Córdoba - SO	9	5	4
Córdoba - SO	10	5	5
Entre Ríos	6	3	3
Santiago del Estero	2		
Chaco (+ Norte Sta Fe)	4	2	2
Salta y Tucumán	7	4	3
San Luis	3		
La Pampa	3		

En los laboratorios de Greenlab, fueron analizadas las 100 muestras seleccionadas. Por un lado, se han realizado las técnicas analíticas de determinación de humedad y materia seca por ISO 665 y la determinación de proteínas por ISO 5983-2. La determinación del perfil de aminoácidos se analizó por la técnica AOAC official method 2018.06, mediante el uso de ultraperformance liquid chromatography -Triple quadrupole mass selective detector (UPLC / MSMS)

ANEXO III

Efectos de fecha de siembra, variedad y grupo de madurez sobre proteína y aceite en muestras de grano de soja cosechada durante la campaña 2020/21 fuera de la zona Núcleo.

Efecto de la fecha de siembra declarada

Se analizó la información proveniente de 123 muestras totales recibidas de Entre Ríos, Córdoba Noroeste, Santa Fe Norte, La Pampa, San Luis, Santiago del Estero y Salta. Se analizaron las fechas de siembra, bulk y discriminada por subregión, variedades (dato promedio) y GM.

Al analizar la fecha de siembra, se consideran un total de 74 casos, con extremos 15/Octubre/2020 y 21/Enero/2021. Se aclara que las muestras provenientes de Santiago del Estero y La Pampa no registraron la fecha de siembra. En la Figura 1 se presentan los valores de contenido proteico (a), aceite (b) y PROFAT (c), en función de la fecha de siembra, mientras que en la Figura 2, es posible visualizar la misma información según las zonas.

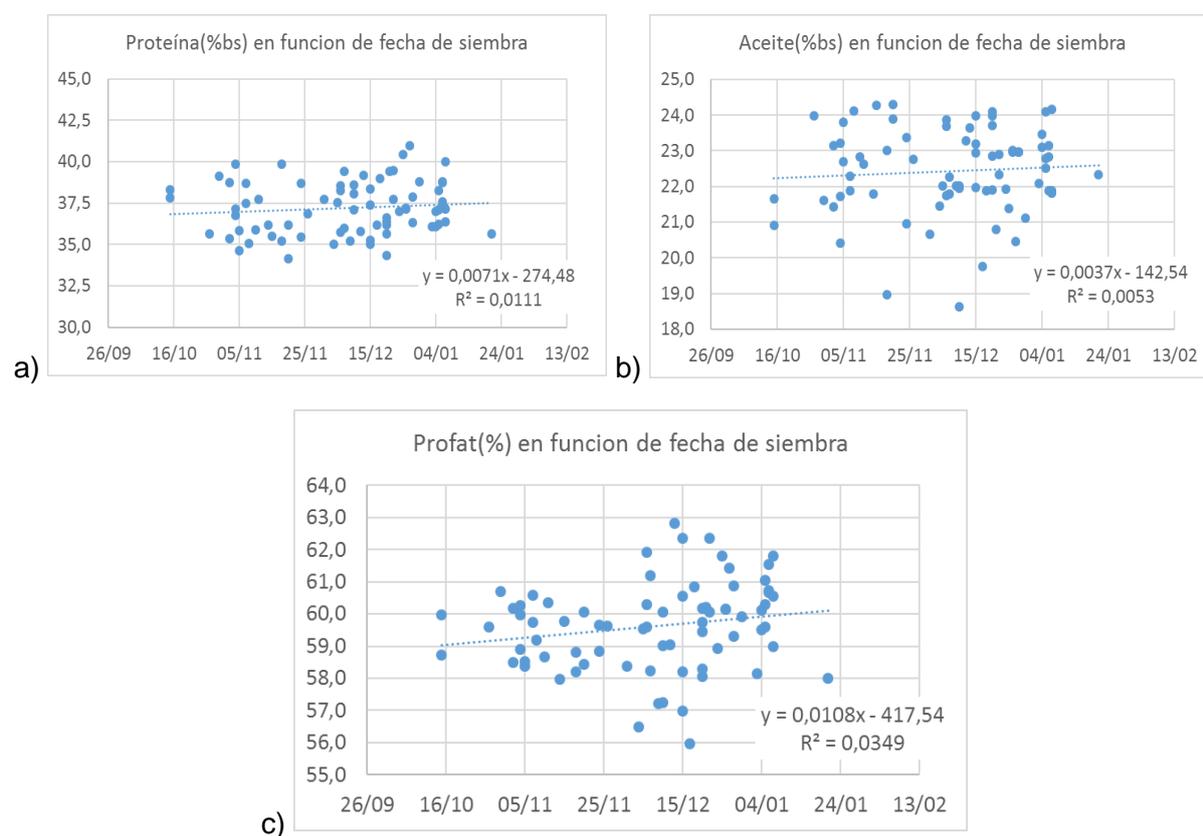


Figura 1. Concentración de proteína (a), aceite (b) y PROFAT (c) en función de la fecha de siembra declarada.

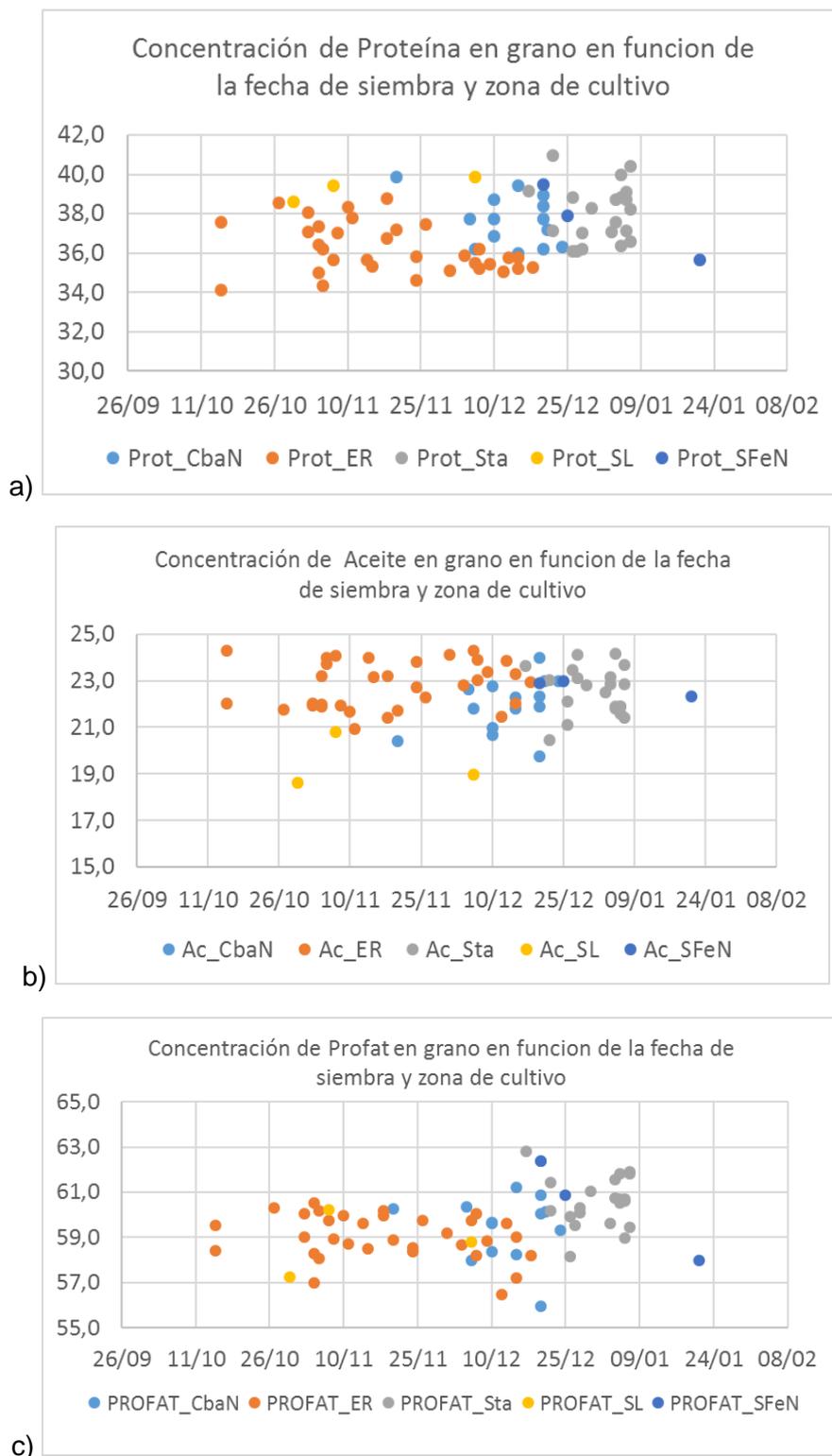


Figura 2. Concentración de proteína (a), aceite (b) y PROFAT (c) en grano de soja, en función de la fecha de siembra en muestras provenientes de Córdoba Norte (CbaN), Entre Ríos (ER), Salta (Sta), San Luis (SL), Santa Fe Norte (SFeN).

Diferencias entre las variedades declaradas

Respecto a las variedades, de las 110 muestras consideradas; la frecuencia de aparición, de mayor a menor, fue: DM 60i62IPRO, n=10; DM 62R63STS, NS 8282, y NS 8288STS, n=5; NS 6483, n=4; DM 63i64IPROSTS, DM 67i70IPROSTS, DM 8473RSF, M 6410IPRO, NS 5258, NS 6248, NS 7809, y RA 655, n=3. El resto de las variedades fueron mencionadas una o dos veces. Sumado a esto, en las Figuras 3, 4 y 5 se presentan los valores medios de contenido de aceite, proteína y PROFAT para cada una de las variedades.

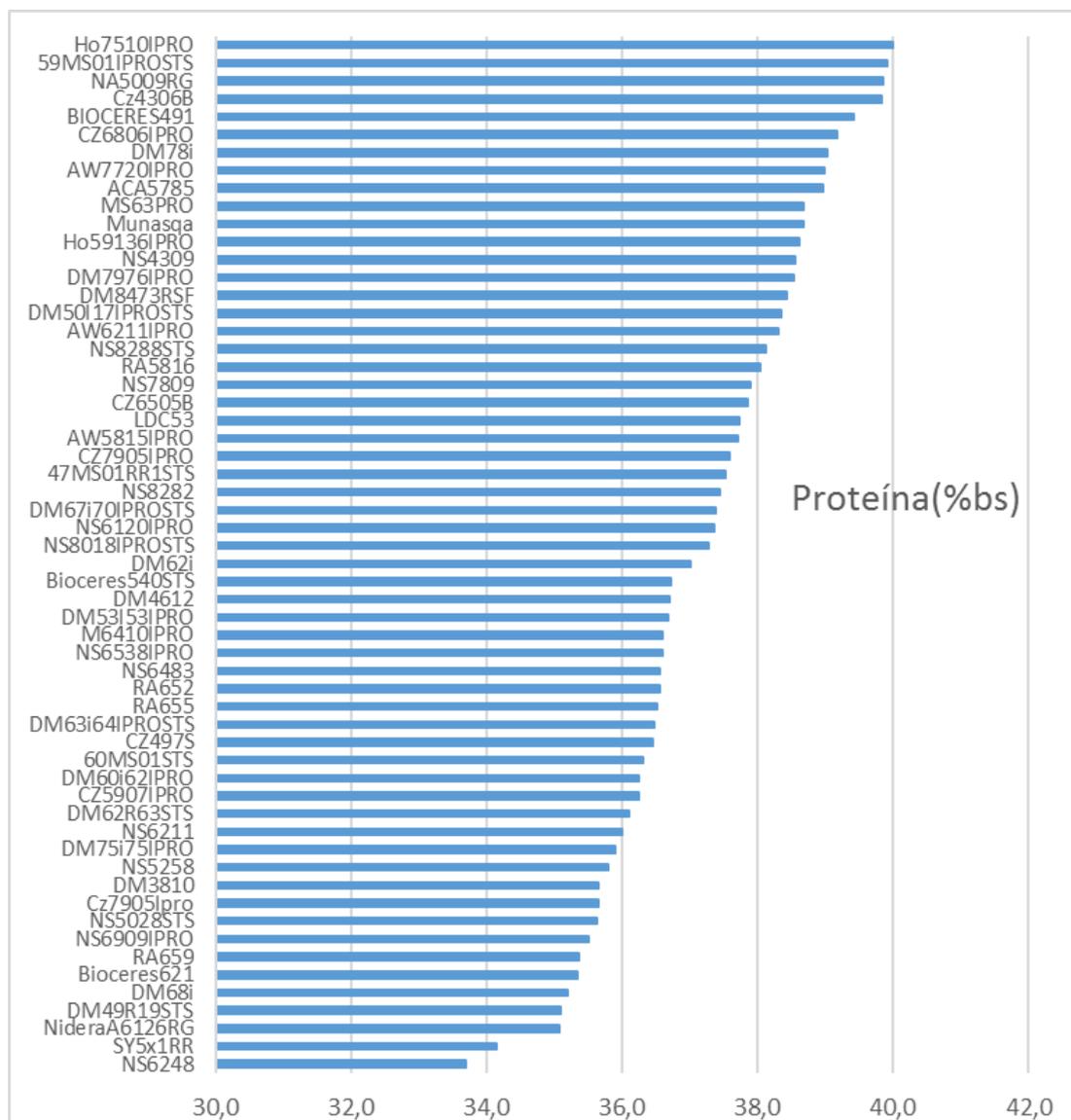


Figura 3. Concentración de proteína en grano de soja de muestras procedentes de zonas extra pampeanas, campaña 20/21; dato promedio por variedad (N variable)

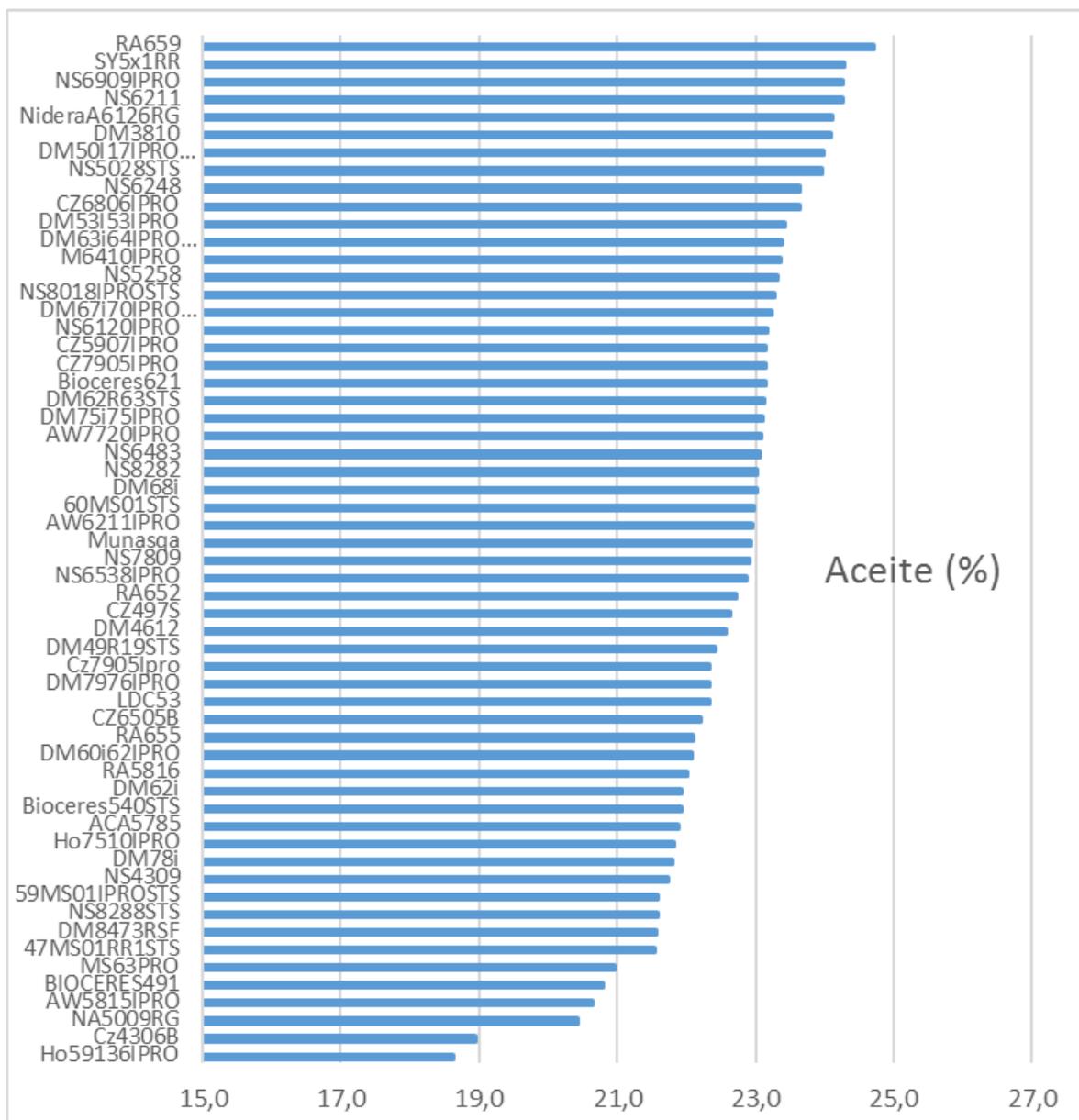


Figura 4. Concentración de aceite en grano de soja de muestras procedentes de zonas extra pampeanas, campaña 20/21; dato promedio por variedad (N variable)

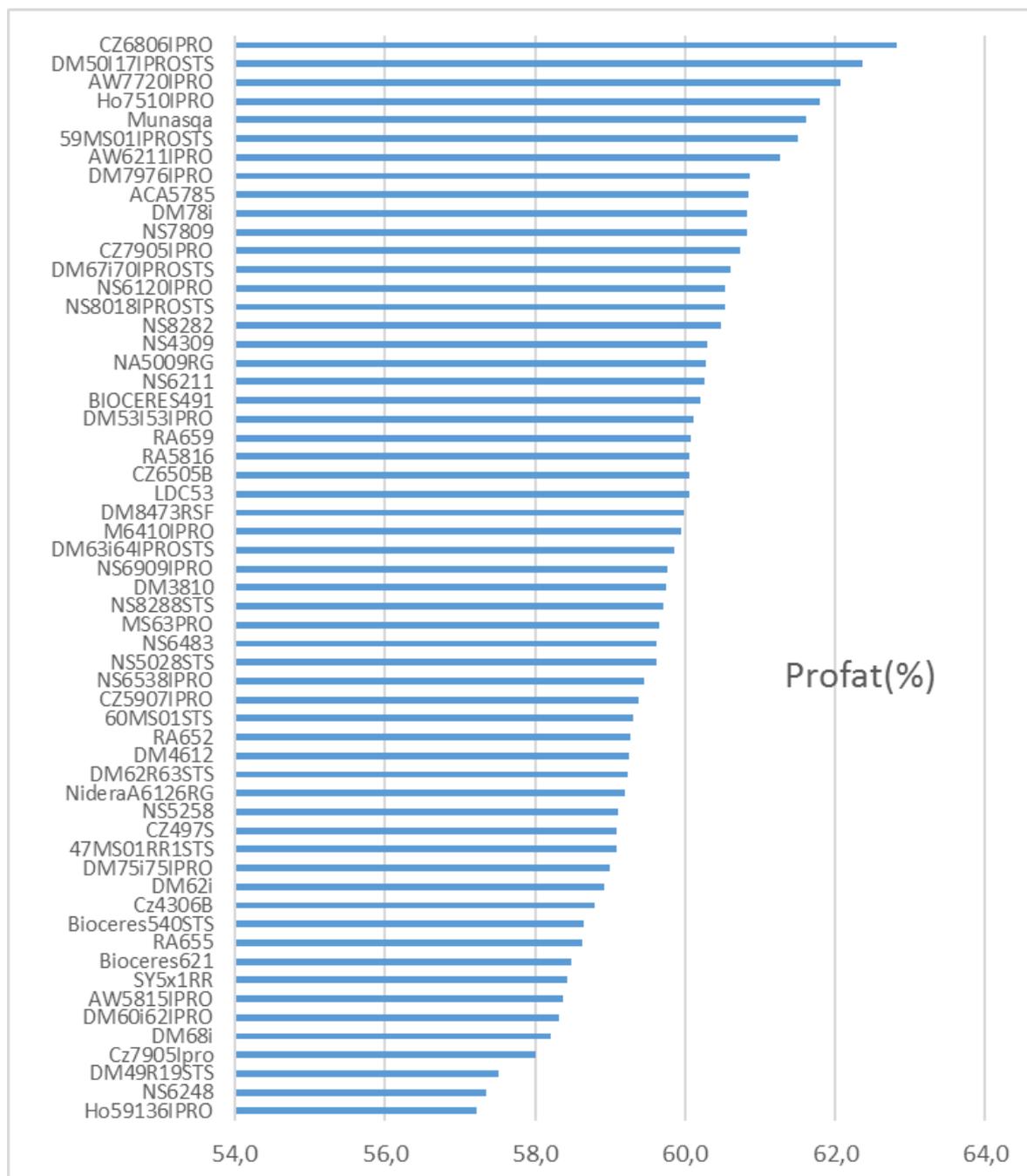


Figura 5. Concentración de profat en grano de soja de muestras procedentes de zonas extra pampeanas, campaña 2020/21; dato promedio por variedad (N variable)

Efecto de la madurez de los materiales (Grupo de madurez de las variedades declaradas)

De las muestras con información sobre grupo de madurez (GM, n=121), se observa que el GM VI fue el más utilizado (n=39), seguido por el GM V (n=36) y por el GM VII (n=24). En la Figura 6 se presentan los valores promedio de concentración de proteína (a), aceite (b) y PROFAT (c), organizados por GM creciente y, dentro de cada GM, por cultivares según valores crecientes.

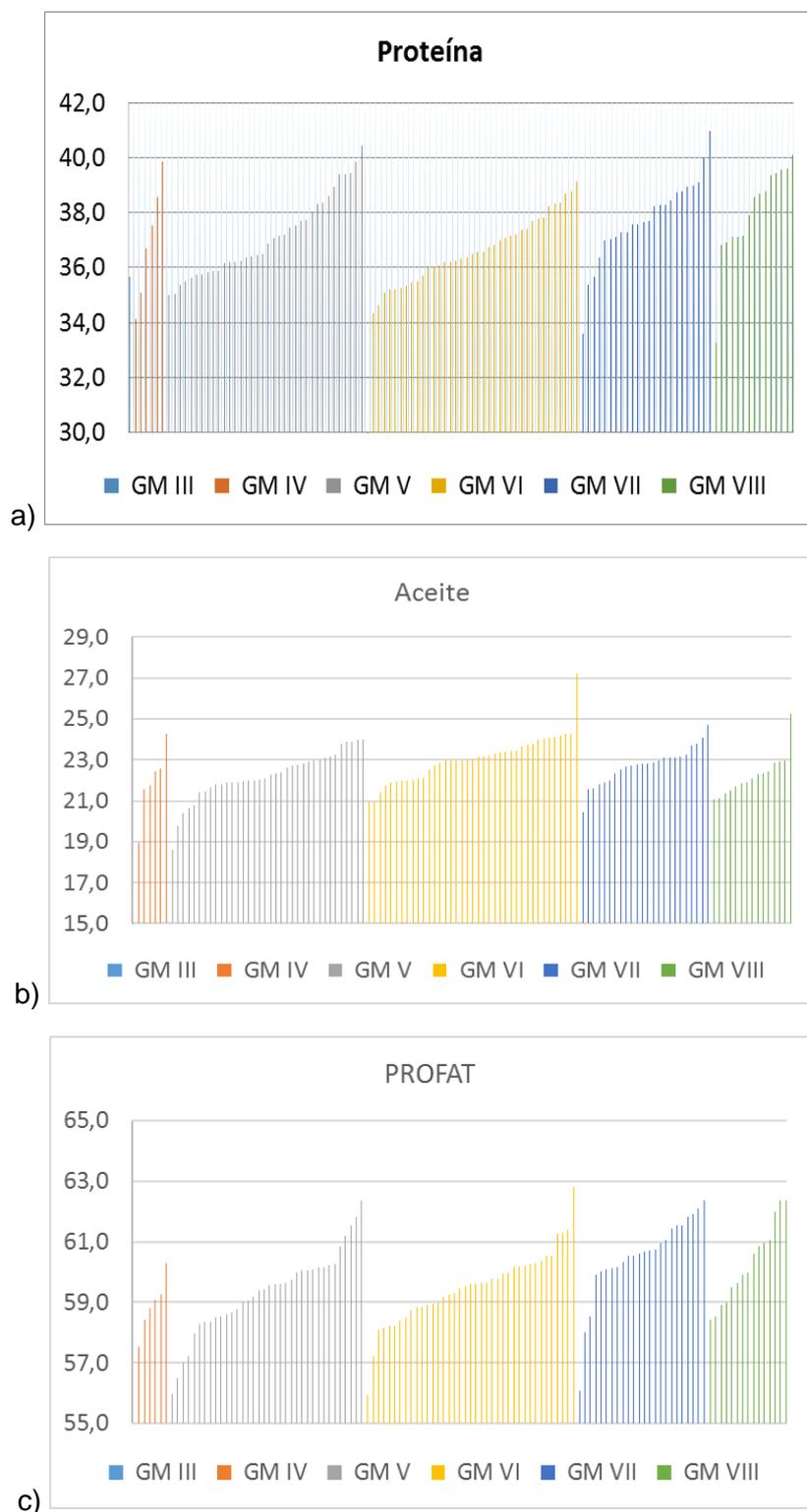


Figura 6. Concentración de proteína (a), aceite (b) y PROFAT (c) en grano de soja, en función del grupo de madurez en muestras provenientes de zonas extra pampeanas, campaña 2020/21.

ANEXO IV

Análisis Estadístico

Además de las medidas resúmenes se realizaron diagramas de caja (boxplots) que nos indican cómo fueron variando las distribuciones en cada zona. Pudiendo observar las variabilidades y medidas de posición central. La figura 1 ilustra los boxplots de proteína (%), en donde se detalla que NOA es la zona con mayores valores medios de porcentaje de proteína, luego NEA, Córdoba sudeste, Córdoba Norte y Buenos Aires Sur y por último, Entre Ríos, Zona núcleo y Santa Fe centro. En cuanto a variabilidades, no se observan grandes diferencias entre las zonas, sólo en NEA se destacan algunos valores muy bajos, consideramos extremos. En lo referido al aceite (%) se observa que Santa Fe Centro es la zona con mayores valores medios de porcentaje de aceite, luego zona núcleo, NOA, NEA, Entre Ríos, Córdoba Norte y por último Córdoba sudeste, y Buenos Aires Sur. En cuanto a variabilidades, no se observan grandes diferencias entre las zonas, sólo que en NEA se destacan algunos valores muy altos, consideramos extremos y en Santa Fe Centro y Córdoba Norte algunos valores muy bajos para su distribución (Figura 2). Asimismo, los boxplot de PROFAT (Figura 3) describen que NOA es la zona con mayores valores medios, compartiendo casi este primer lugar con NEA y Santa Fe Centro. Siguiendo en orden descendente por zonas Núcleo, Córdoba Norte, Córdoba Sudeste y Entre Ríos. Por último, quedó con menores valores de PROFAT Buenos Aires Sur. En cuanto a variabilidades, no se observan grandes diferencias entre las zonas, sólo que en NEA, Santa Fe centro y Córdoba sudeste se destacan algunos valores muy bajos, consideramos extremos y en zona núcleo un valor muy bajo y otro alto para su distribución.

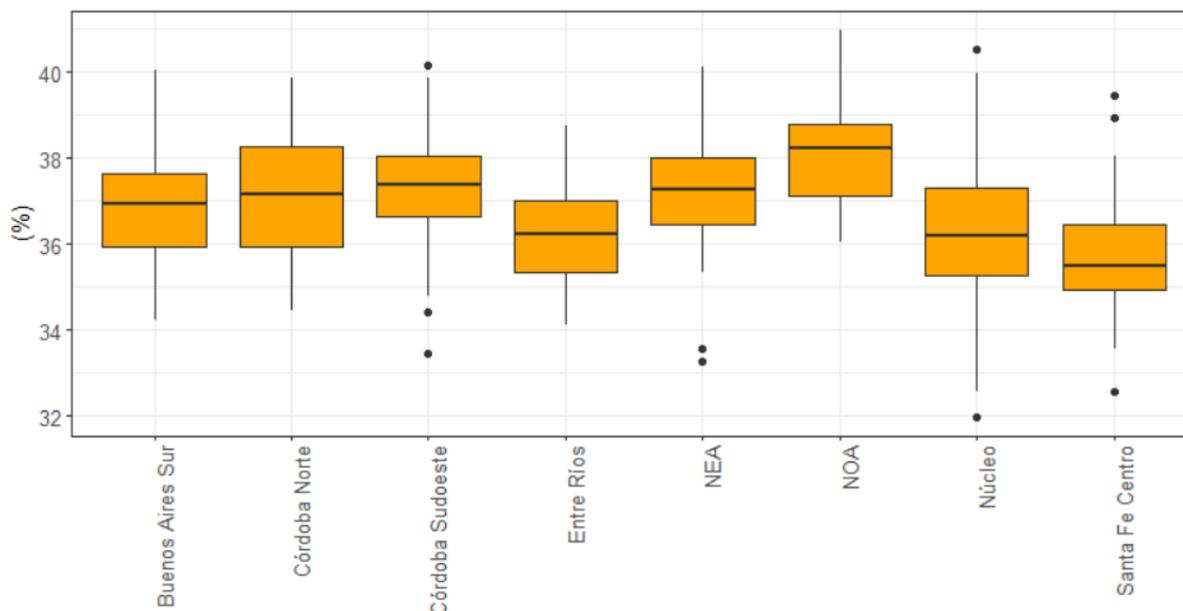


Figura 1. Boxplot de proteínas.

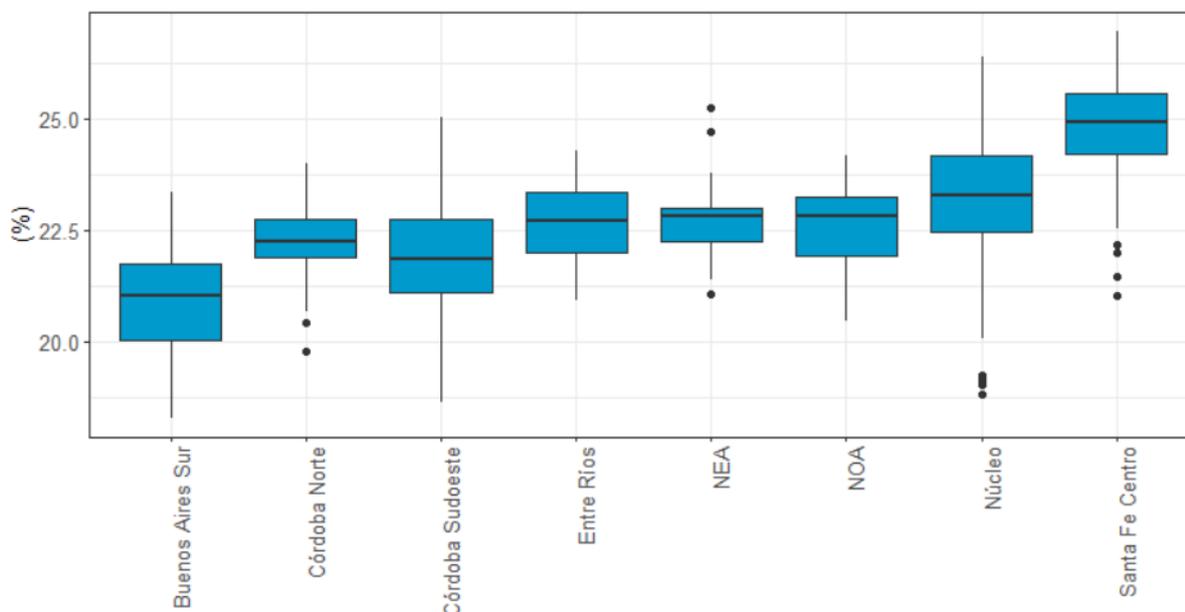


Figura 2. Boxplot de aceites.

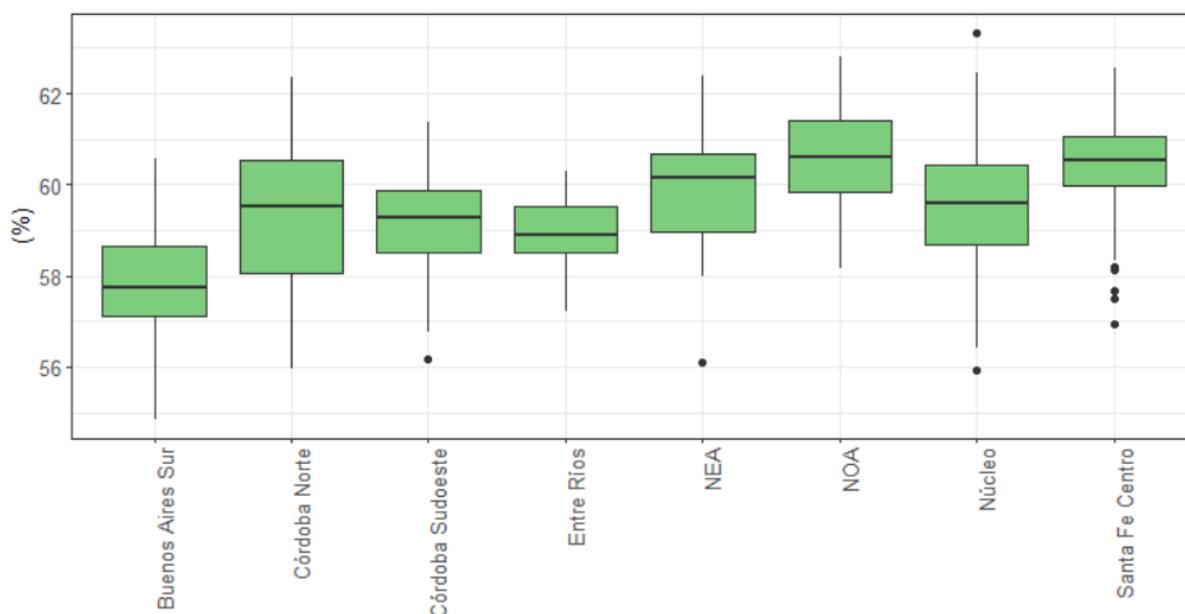


Figura 3. Boxplot de PROFAT

Los boxplot por zonas y por fechas de siembra, destacan que hay algunas zonas, como Córdoba Sudoeste, NEA y NOA, que sólo poseen una fecha de siembra. El resto de las zonas tienen información de ambas fechas, observándose grandes diferencias en el porcentaje de proteína en Buenos Aires Sur a favor de la fecha de siembra de 2, en cambio en Córdoba Norte, Entre Ríos y Núcleo, la fecha de siembra 1 es la que posee mayores porcentajes de proteína, por otro lado, en Santa Fe Centro no se observan diferencias en los valores medios para las fechas de siembra (Figura 4). Con respecto al porcentaje de aceite, se observa grandes diferencias en Buenos Aires Sur, Córdoba Norte, Núcleo y Santa Fe

Centro a favor de la fecha de siembra de 1, en cambio en Entre Ríos la fecha de siembra 2 es la que posee mayores porcentajes de aceite (Figura 5), cuyo patrón se repite para el porcentaje de PROFAT (Figura 6).

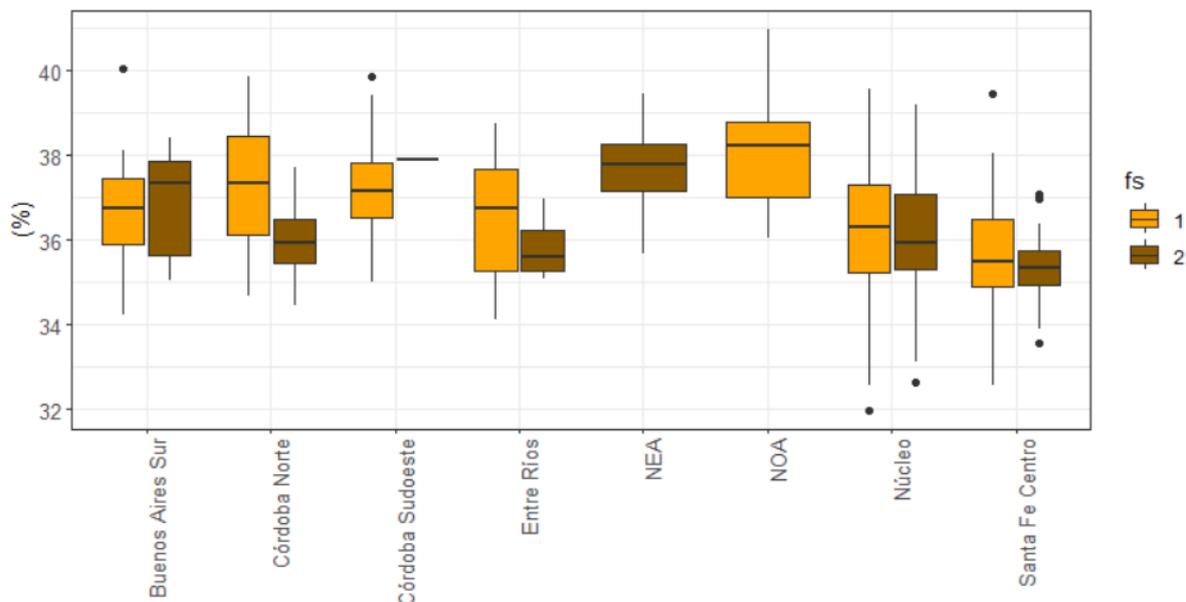


Figura 4. Relación Porcentaje de Proteína y Fecha de Siembra

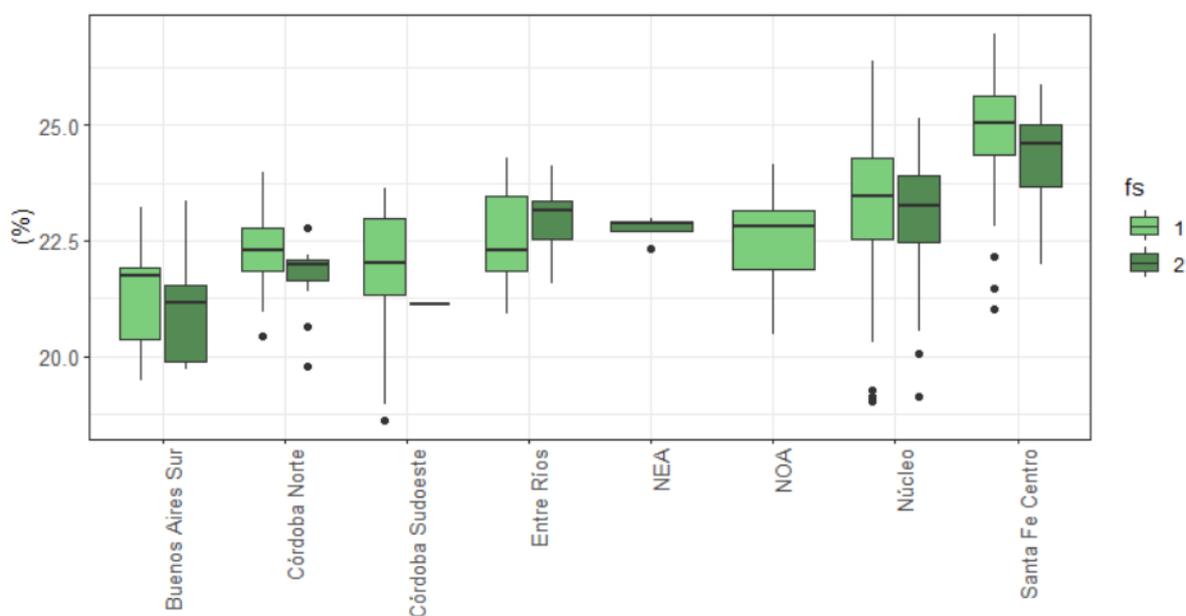


Figura 5. Relación Porcentaje de Aceite y Fecha de Siembra

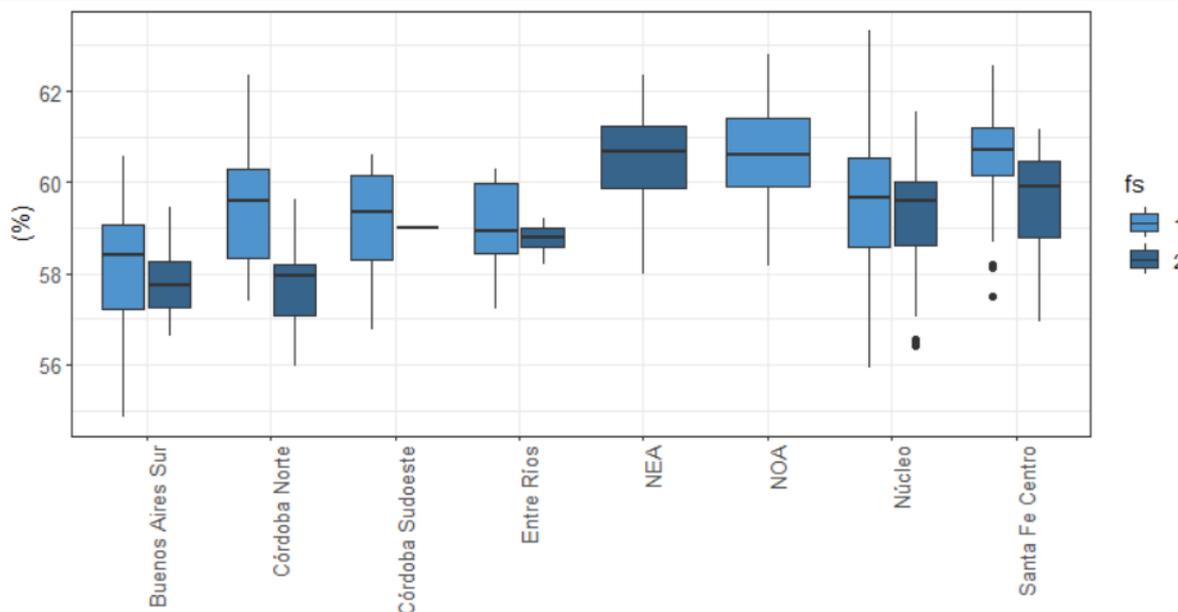


Figura 6. Relación PROFAT y Fecha de Siembra

Al analizar, solo la zona Núcleo, ya que es la zona con mayor número de muestras, de acuerdo a las fechas de siembra, se observa que existe una diferencia a favor de la fecha de siembra de 1 en los valores promedios de las tres variables analizadas: proteína (%), aceite (%) y PROFAT (%).

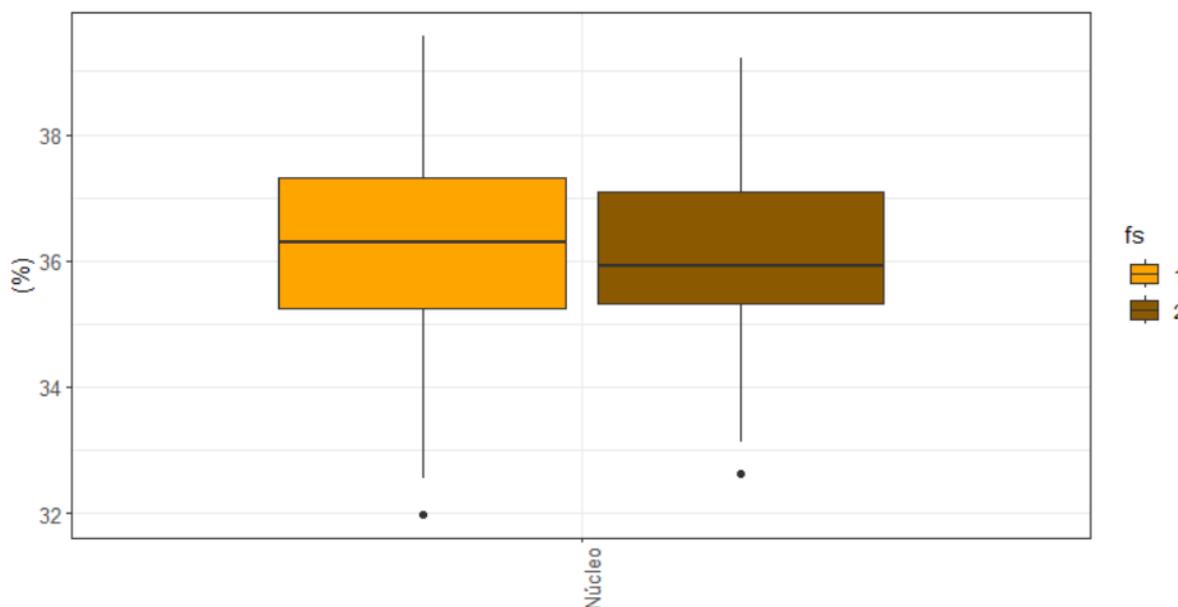


Figura 7. Relación Proteína y Fecha de Siembra en zona Núcleo.

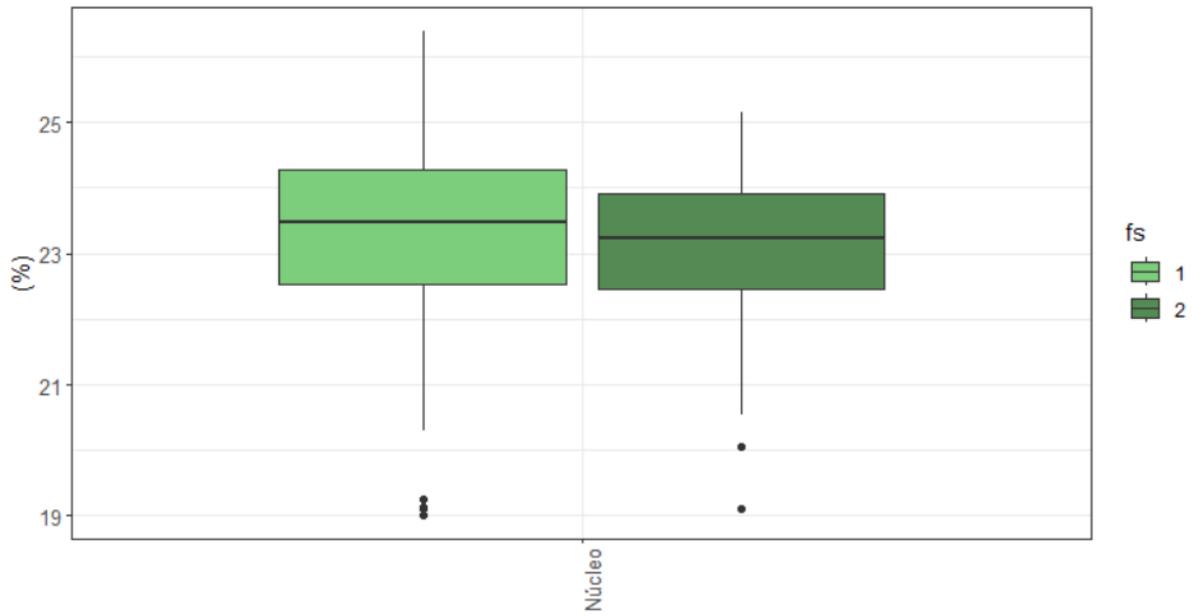


Figura 8. Relación Aceite y Fecha de Siembra en zona Núcleo

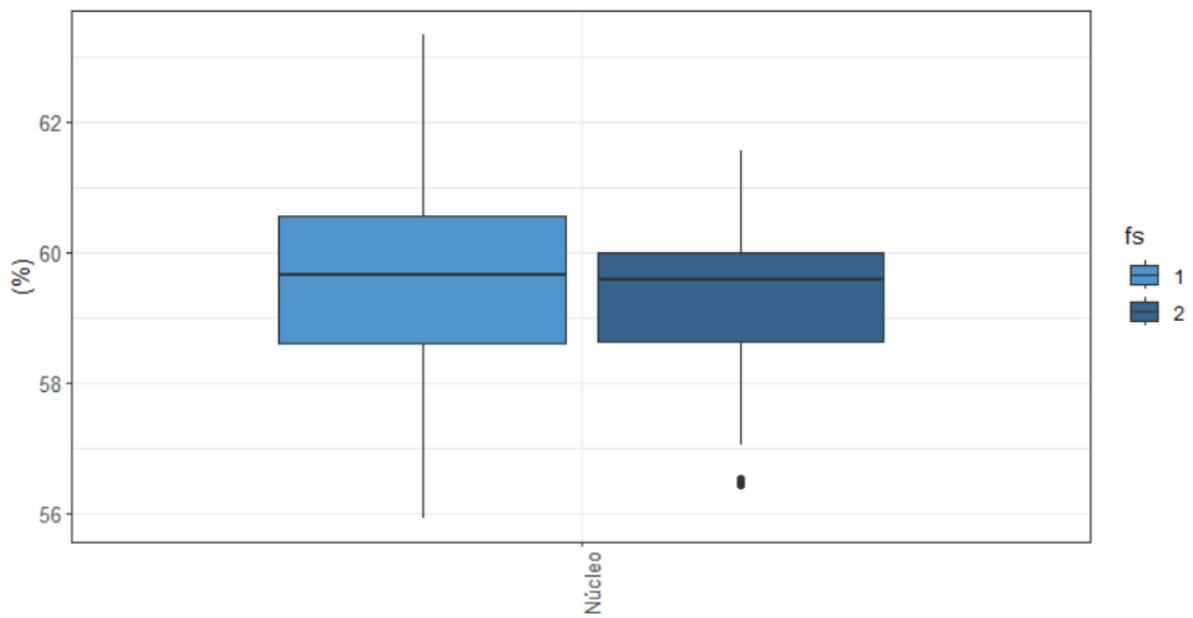


Figura 9. Relación PROFAT y Fecha de Siembra en zona Núcleo

ANEXO V AMINOÁCIDOS

Distribución de los AA

En la Figura 1, se muestra el gráfico de boxplots de los cinco AA (%) considerando el total de muestras (n=100). Allí, puede observarse que al evaluar AA (%), el mayor valor corresponde a Lisina (%), seguido por Treonina (%), y en menores, Cisteína (%), Metionina (%) y Triptófano (%). Además, es interesante resaltar que Lisina (%) posee una distribución simétrica. A partir de los boxplots construidos para cada AA (Figura 2) con las observaciones de la zona núcleo (n=45), se observó un comportamiento similar al nacional. En las Figuras 3, 4, 5 y 6 se muestran las gráficas de los análisis de correlación entre los distintos aminoácidos y entre los aminoácidos y proteína para diferentes criterios.

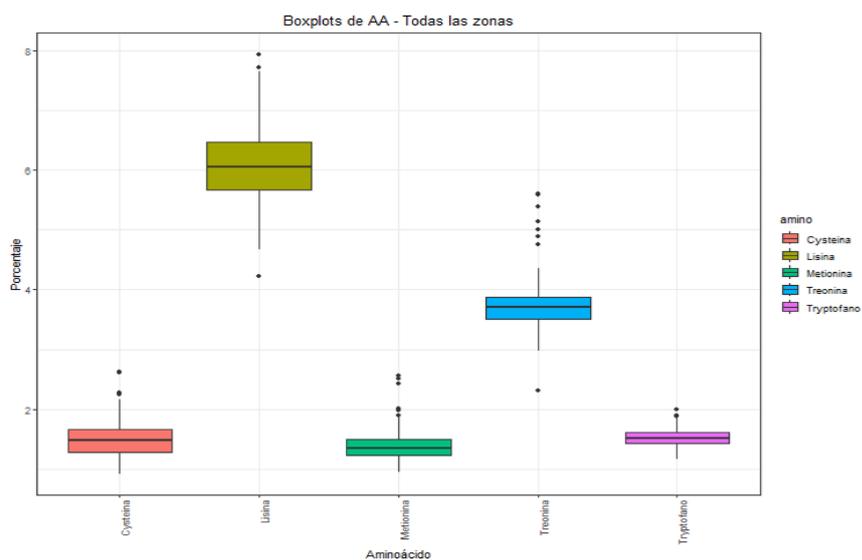


Figura 1. Boxplot KAA para todas las regiones

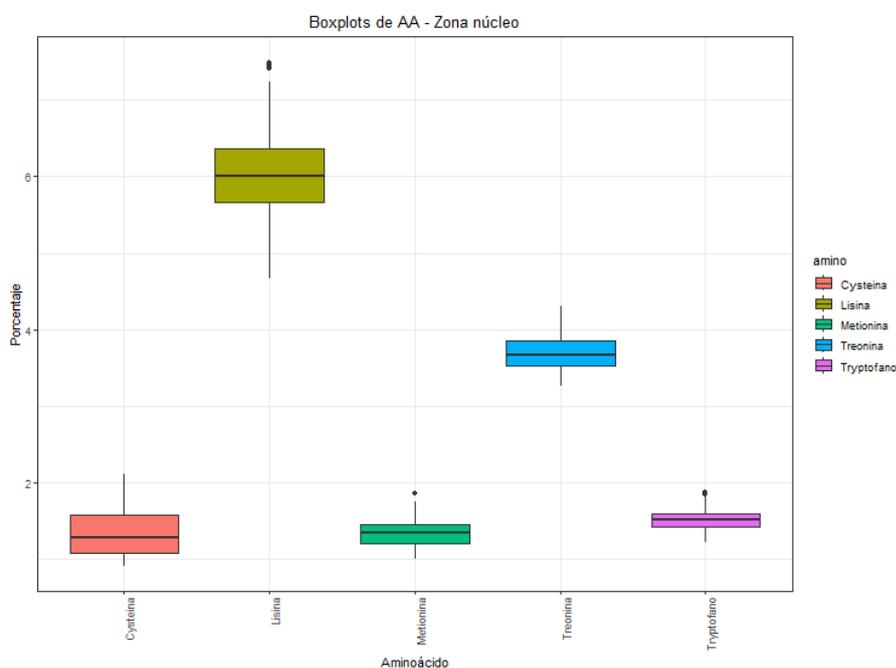


Figura 2. Boxplot KAA para la región núcleo.

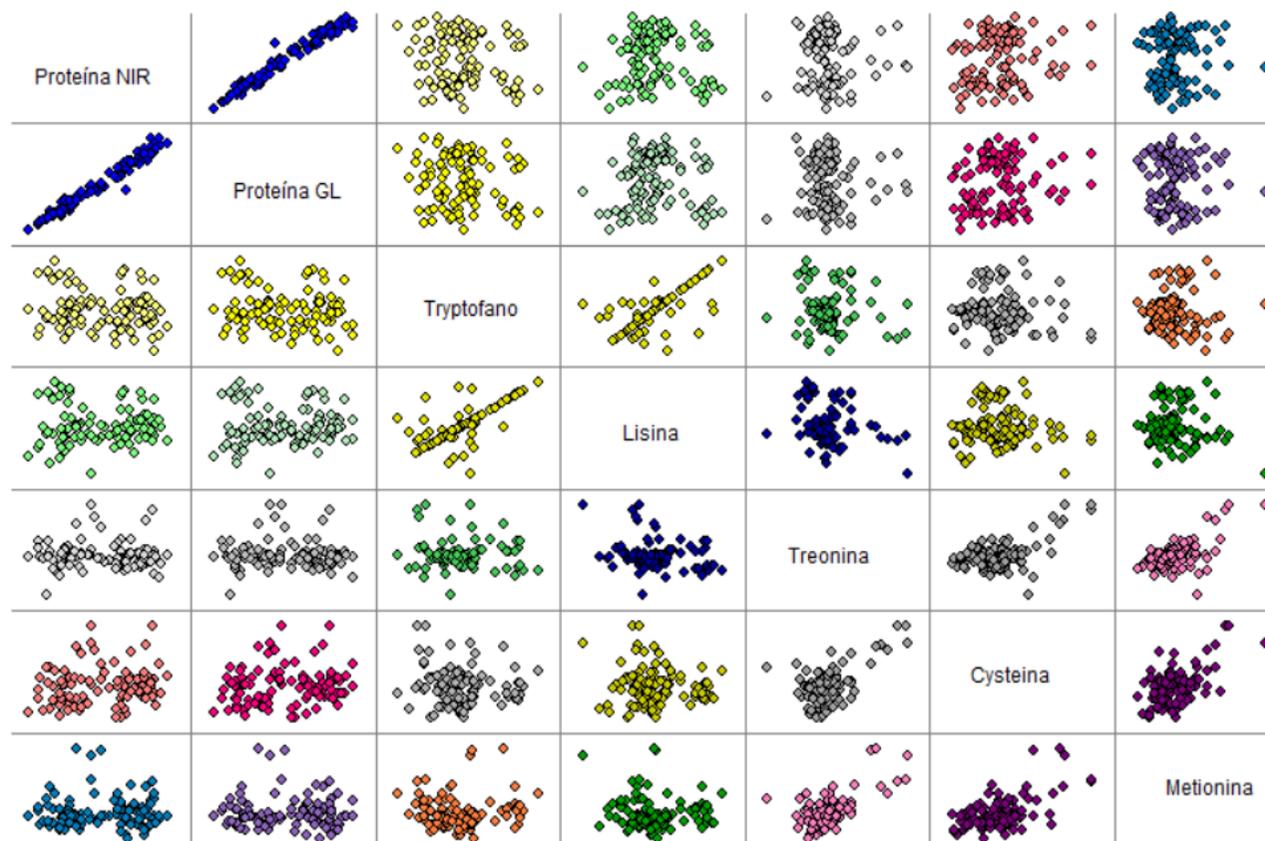


Figura 3. Matriz de dispersión entre los distintos aminoácidos y entre los aminoácidos y proteína de todas las muestras

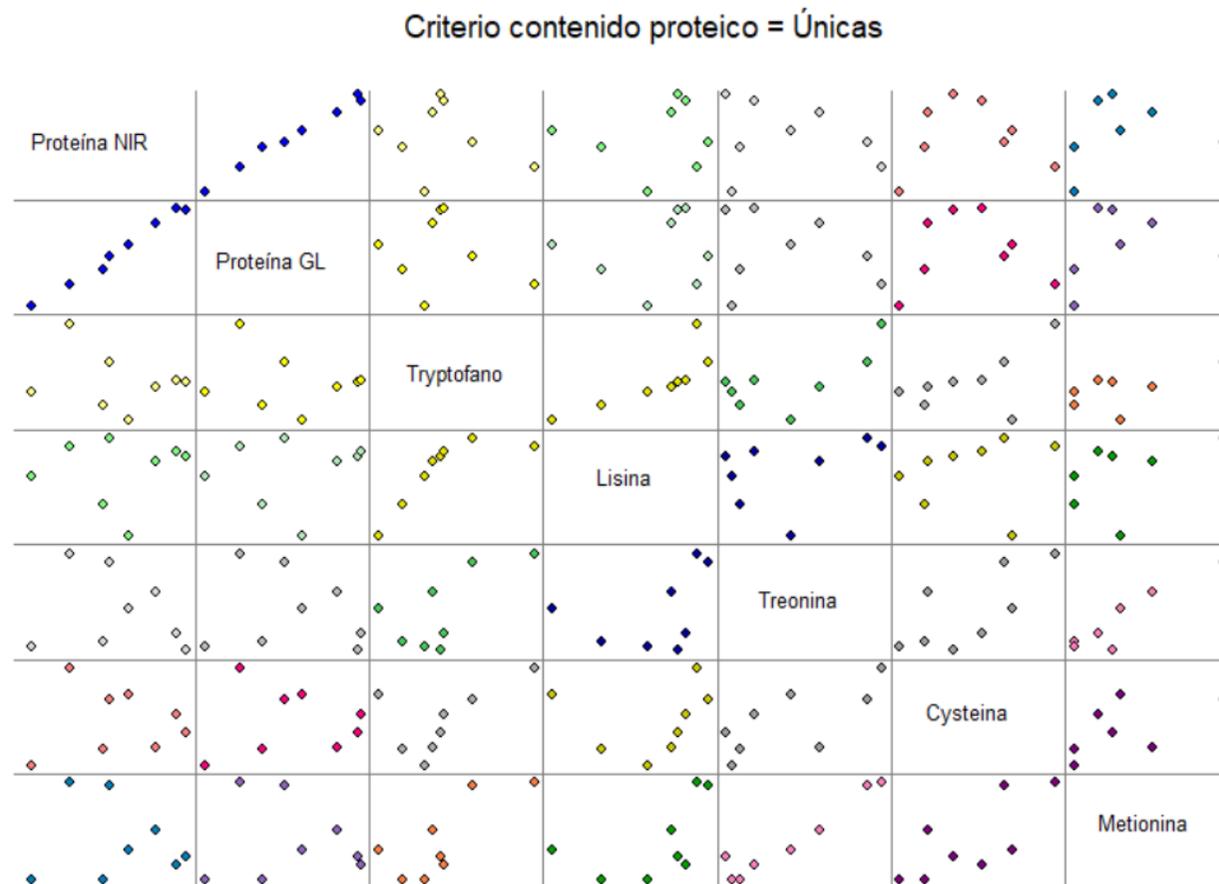


Figura 4. Matriz de dispersión entre los distintos aminoácidos y entre los aminoácidos y proteína de muestras única

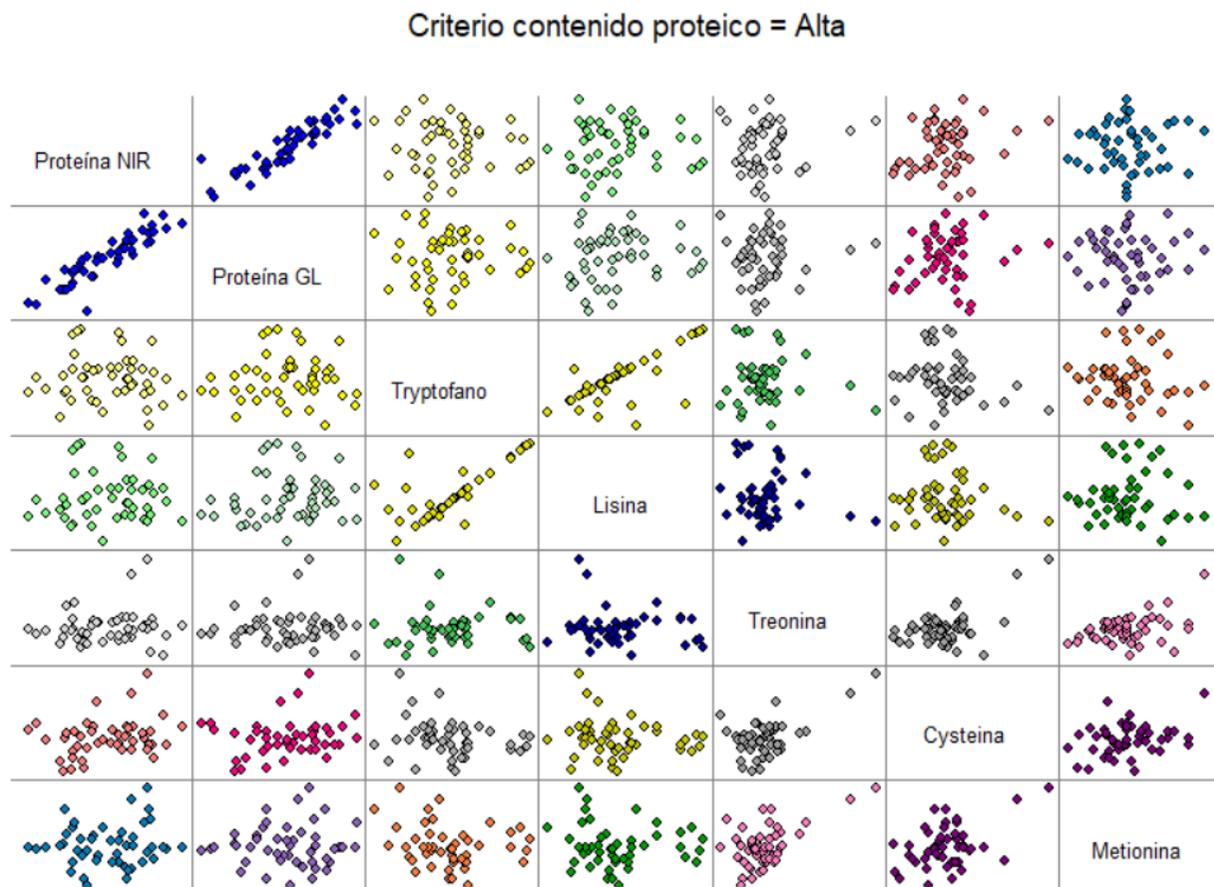


Figura 5. Matriz de dispersión entre los distintos aminoácidos y entre los aminoácidos y proteína de muestras con alta proteína.

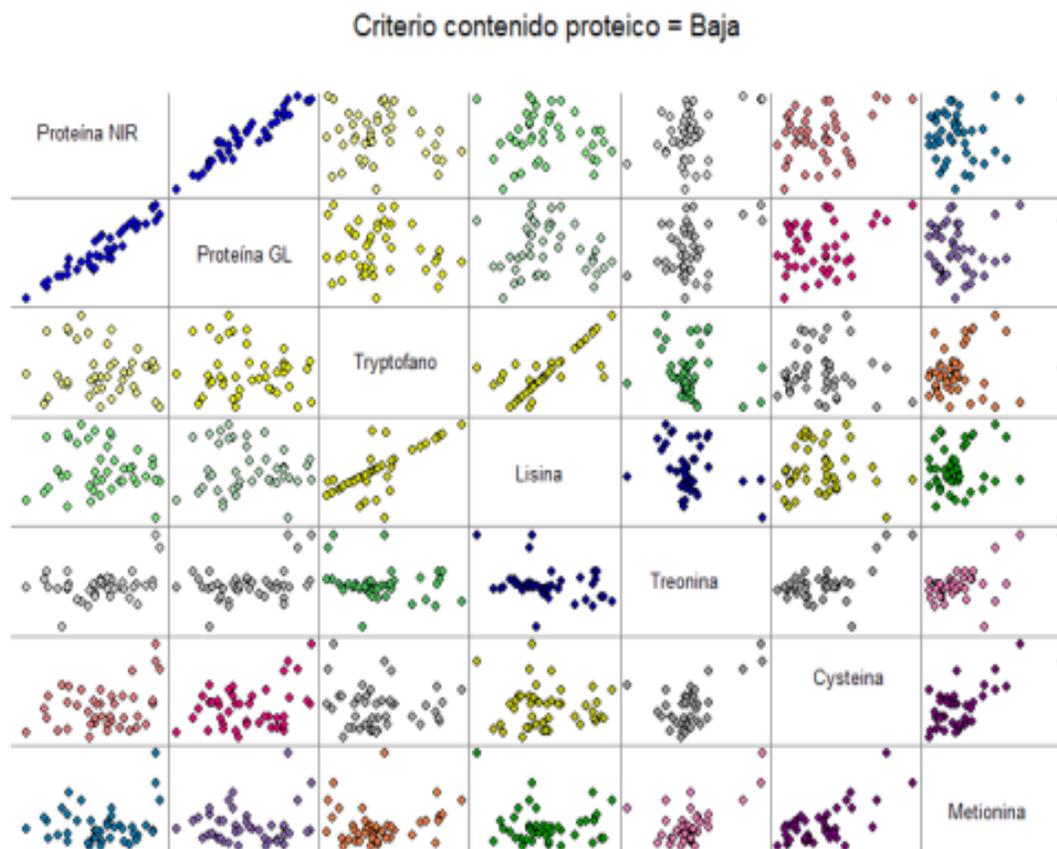


Figura 6. Matriz de dispersión entre los distintos aminoácidos y entre los aminoácidos y proteína de muestras con baja proteína

ANEXO VI

DETALLE DEL ANÁLISIS GEOESTADÍSTICO

Se reservó el 5% de las muestras para validar el modelo (Figura 1), y con las restantes se realizó un análisis exploratorio (Figura 2), el cual comprobó que no existía tendencia espacial entre los valores de las muestras, y se aplicó el método Kriging, con una distancia (lag) de 12000 m y un pixel de 4 km². Asimismo, se analizó el modelo de ajuste (Exponencial, Esférico, Circular y Gaussiano) para cada una de las variables, mediante el error cuadrático medio R², a fin de seleccionar el modelo que mejor ajuste los datos (Tabla 1). En la misma se observa que el modelo exponencial es el que mejor ajuste brinda a las variables de aceite y PROFAT, mientras que el modelo circular es el que mejor ajusta para proteína.

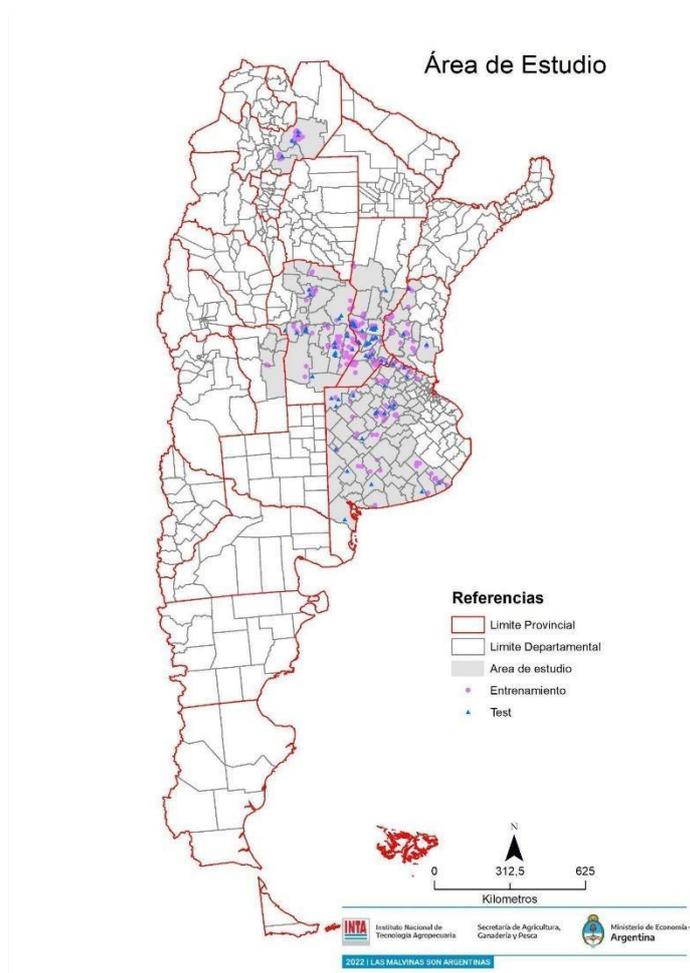


Figura 1. Regiones analizadas y distribución espacial de las muestras utilizadas para la interpolación como para la validación.

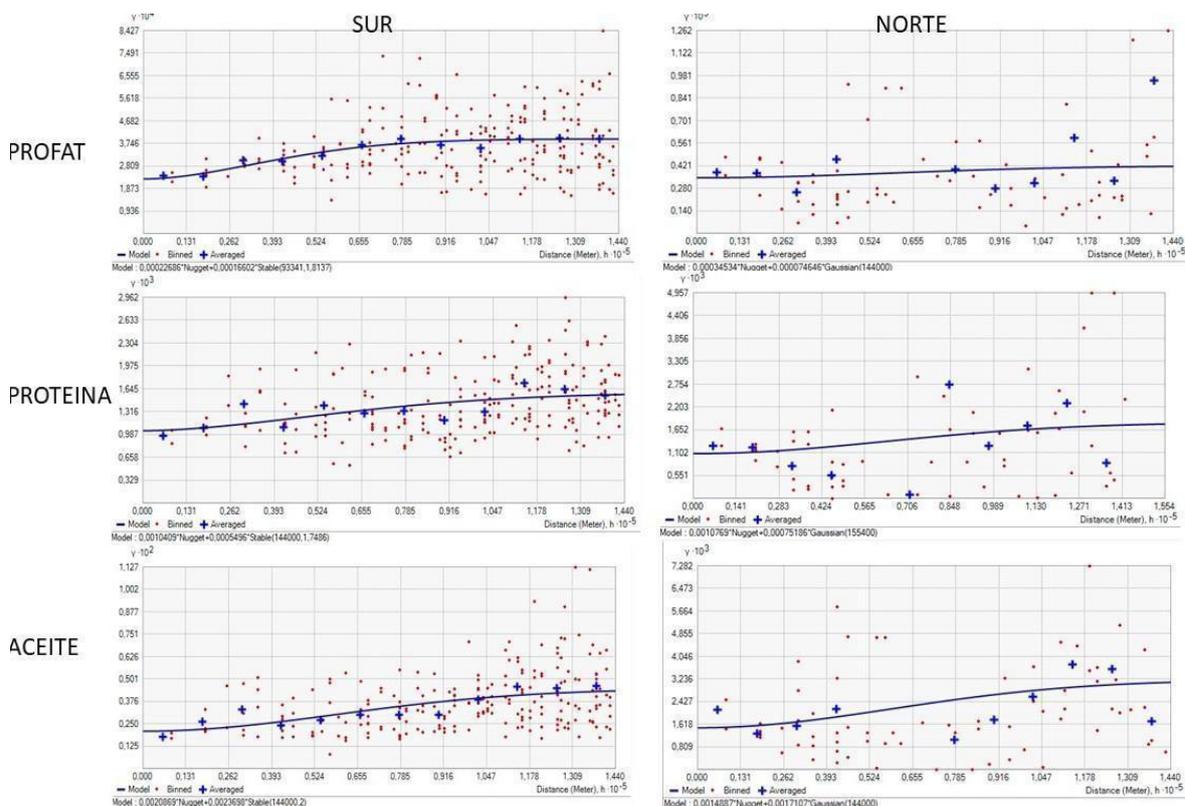


Figura 2. Semivariograma de las variables de proteína, aceite y PROFAT.

Tabla 1. Error cuadrático medio de los 4 modelos de ajuste para las variables de proteína, aceite y PROFAT

	Gaussiano	Exponencial	Esférico	Circular
Proteína	1,98906511	1,96250031	1,9052847	1,89940054
Aceite	1,40756619	1,27438915	1,35585325	7,38720846
PROFAT	1,6161821	1,50920921	1,51868697	1,52536012

Para el análisis de aminoácidos se utilizó similar metodología que, para Proteína, Aceite y PROFAT. Se reservó el 5 % de las muestras, de la zona núcleo, para validar el modelo, y con las restantes se realizó un análisis exploratorio (Figura 3), el cual comprobó que no existía tendencia espacial entre los valores de las muestras, y se aplicó el método Kriging, con una distancia (lag) de 12000 m y un pixel de 4km². Asimismo, se analizó el modelo de ajuste (Exponencial, Esférico, Circular y Gaussiano) para cada una de las variables, mediante el error cuadrático medio R², a fin de seleccionar el modelo que mejor ajuste los datos (Tabla 2). En la misma se observa que el modelo exponencial es el que mejor ajuste brinda a la mayoría de las variables de aminoácidos.

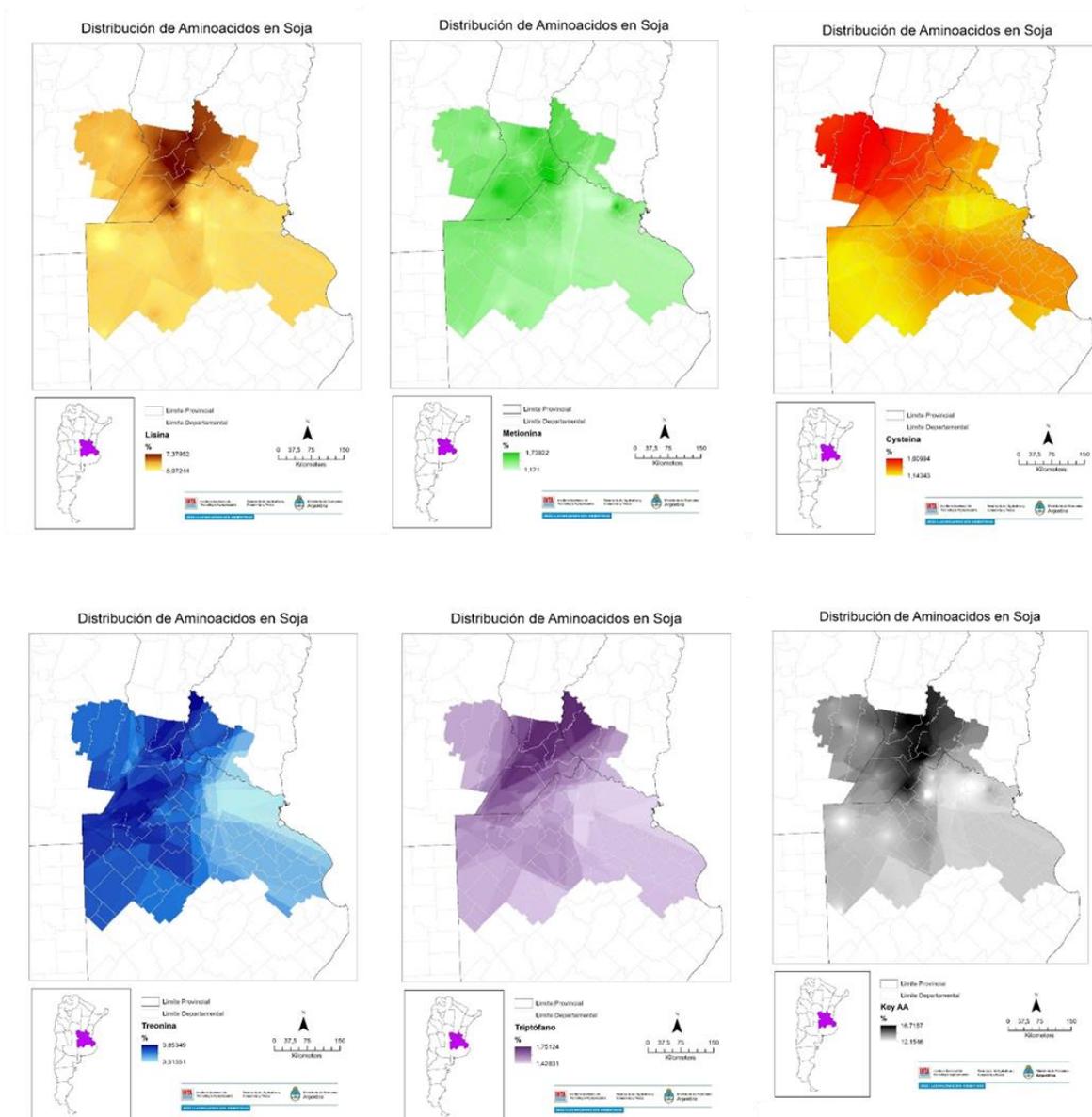


Figura 3. Distribución espacial de los 5 Key AA (g) en la Zona Núcleo de Argentina durante la campaña 2021/2022

Tabla 2. Error cuadrático medio de los 4 modelos de ajuste para las variables de aminoácidos

	Gauss	Exponencial	Esférico	Circula
Tryptofano SSS	0,008351578	0,005950431	0,007534313	0,007513533
Tryptofano %	0,01658135	0,015527882	0,015091074	0,014948502
Lisina SSS	0,134327518	0,118477612	0,126850268	0,128291918
Lisina %	0,524010404	0,491138612	0,50155615	0,510832604
Treonina SSS	0,072606136	0,075785822	0,076280942	0,071267758
Treonina %	0,41869976	0,413876152	0,415676016	0,415687902

Cysteina SSS	0,021319257	0,019409434	0,021243187	0,021156117
Cysteina %	0,19239757	0,171323904	0,18783465	0,189747322
Metionina SSS	0,00483302	0,003962168	0,004356346	0,004577202
Metionina %	0,05417241	0,051172142	0,052163498	0,082557152
Sumatoria %	1,277416796	1,113980048	1,120540372	1,128626846

Tabla 1. Error cuadrático medio de los 4 modelos de ajuste para las variables de proteína, aceite y PROFAT

	Gaussiano	Exponencial	Esférico	Circular
Proteína	1,98906511	1,96250031	1,9052847	1,89940054
Aceite	1,40756619	1,27438915	1,35585325	7,38720846
PROFAT	1,6161821	1,50920921	1,51868697	1,52536012