

**CARTOGRAFÍA DIGITAL DE PROPIEDADES DEL SUELO Y SU  
APLICACIÓN AL ESTUDIO DE LA VARIABILIDAD ESPACIAL  
DEL CULTIVO DE SOJA**

**Mauricio Castro Franco**

Trabajo de tesis para ser presentado como requisito parcial para  
optar al Título de:

**DOCTOR EN CIENCIAS AGRARIAS**

**Área de Producción Vegetal**

**PROGRAMA DE POSGRADO EN CIENCIAS AGRARIAS**

**FACULTAD DE CIENCIAS AGRARIAS**

**UNIVERSIDAD NACIONAL DE MAR DEL PLATA**

**Balcarce, Argentina**

**Diciembre de 2015**

**CARTOGRAFÍA DIGITAL DE PROPIEDADES DEL SUELO Y SU  
APLICACIÓN AL ESTUDIO DE LA VARIABILIDAD ESPACIAL  
DEL CULTIVO DE SOJA**

**Mauricio Castro Franco**

---

Director/a de Tesis (*Nombre y Títulos*)

---

Co-Director/a de Tesis (*Nombre y Títulos*)

**CARTOGRAFÍA DIGITAL DE PROPIEDADES DEL SUELO Y SU  
APLICACIÓN AL ESTUDIO DE LA VARIABILIDAD ESPACIAL  
DEL CULTIVO DE SOJA**

**Mauricio Castro Franco**

**Aprobada por:**

-----  
Evaluador/a (*Nombre y Títulos*)

-----  
Evaluador/a (*Nombre y Títulos*)

-----  
Evaluador/a (*Nombre y Títulos*)

## **DEDICATORIA**

- A mis padres
- A Marisa Domenech, por su apoyo incondicional.
- A mi hermano y su familia.

## **AGRADECIMIENTO**

- A José Luis Costa y Virginia Aparicio, por su apoyo, paciencia, enseñanzas y amistad.
- A Mónica Balzarini y Mariano Córdoba, por su apoyo, aliento y colaboración.
- A Luis Alonso y Maxi Dosantos, por su compañerismo y colaboración.
- A mi colega Natalia Izquierdo por sus opiniones, paciencia y amistad.
- A mi colega y amigo Marino Lind por su colaboración.
- Al INTA y CONICET, por darme la oportunidad de aprender tanto y por siempre tener los brazos abiertos de bienvenida.

## INDICE

### Capítulo 1

1. Introducción general.....	18
1.1. Importancia de obtener información de suelos a escala de lote.....	18
1.2. Cartografía Digital de Suelos (CDS).....	20
1.3. El cultivo de la soja en Argentina.....	23
1.4. Justificación.....	25
1.5. Hipótesis.....	26
1.6. Objetivos.....	27
1.7. Estructura de la tesis.....	28

### Capítulo 2

#### **Predicción espacial de propiedades del suelo a escala de lote integrando esquemas de muestreo de suelos basados en modelos y regresión “*Random Forest*” (RF)**

1. Introducción.....	29
2. Materiales y métodos.....	32
2.1. Zona de estudio.....	32
2.2. Variables predictoras.....	33
2.3. Desarrollo del modelo RF de predicción de propiedades del suelo...35	35
3. Resultados y discusión.....	40
3.1. Propiedades del suelo.....	41
3.2. Análisis de la estructura espacial de los predictores medidos.....	42
3.3. Análisis exploratorio de los predictores.....	44
3.4. Comparación de esquemas de muestreo de suelos basados en modelos.....	46
3.5. Importancia de los predictores y performance del modelo RF.....	50
3.6. Predicción de propiedades del suelo.....	53
3.7. Patrones espaciales de la predicción de propiedades del suelo.....	55
4. Conclusiones.....	56

### Capítulo 3

#### Desagregación de unidades cartográficas de suelo a partir de análisis de componentes principales espaciales y distancia euclidiana

1. Introducción.....	58
2. Materiales y métodos.....	63
2.1. Zonas de estudio.....	63
2.2. Fuentes de información auxiliar de covariables ambientales.....	68
2.3. Análisis de la estructura espacial de las covariables ambientales.....	70
2.4. Metodología de desagregación de unidades cartográficas de suelo.....	71
3. Resultados y discusión.....	75
3.1. Parámetros de estadística descriptiva.....	75
3.2. Análisis de la estructura espacial.....	87
3.3. Aplicación de la metodología de desagregación de unidades Cartográficas de suelo .....	91
4. Conclusiones.....	115

### Capítulo 4

#### Relación entre patrones espaciales de propiedades del suelo y cultivo de soja a escala de lote

1. Introducción.....	119
2. Materiales y métodos.....	122
2.1. Zonas de estudio.....	122
2.2. Medición de las propiedades del cultivo de soja.....	122
2.3. Metodología para determinar las relaciones entre patrones espaciales de propiedades de suelo y cultivo de soja.....	124
3. Resultados y discusión.....	127
3.1. Época de muestreo.....	127
3.2. Parámetros de estadística descriptiva de las propiedades de cultivo.....	129
3.3. Validación de zonas por tipo de suelo para propiedades de cultivo	

de soja.....	135
3.4. Importancia de las propiedades del suelo en los patrones espaciales de propiedades de cultivo de soja .....	141
3.5. Relaciones espaciales entre zonas por tipo de suelo y propiedades del cultivo de soja.....	155
4. Conclusiones.....	164
<b>Capítulo 5</b>	
Discusión general.....	168
<b>Bibliografía.....</b>	<b>178</b>

## INDICE DE TABLAS

<b>Tabla 2.1</b> Descripción y resolución espacial de los predictores y su relación con los factores de formación del suelo propuestos por <i>Jenny</i> , (1941).....	34
<b>Tabla 2.2</b> Parámetros geoestadísticos de modelos ajustados de semivariogramas experimentales para los predictores medidos.....	42
<b>Tabla 2.3</b> Resumen de parámetros de estadística descriptiva de los predictores medidos.....	45
<b>Tabla 3.1</b> Parámetros de estadística descriptiva de las covariables ambientales para los lotes de la zona 1 (Buenos Aires) .....	76
<b>Tabla 3.2</b> Parámetros de estadística descriptiva de las covariables ambientales para los lotes de la zona 2 (Córdoba).....	81
<b>Tabla 3.3</b> Parámetros de estadística descriptiva de las covariables ambientales para los lotes de la zona 3 (Entre Ríos).....	85
<b>Tabla 3.4</b> Parámetros geoestadísticos de modelos ajustados de semivariogramas experimentales para los predictores medidos.....	89
<b>Tabla 3.5</b> Estadística descriptiva de las tres primeras componentes principales generadas a partir de análisis de componentes principales (ACP) y MULTISPATI-ACP para las tres zonas estudiadas .....	92
<b>Tabla 3.6</b> Peso de las covariables ambientales (vectores propios) de los lotes de la zona 1 (Buenos Aires) en los tres primeros componentes del análisis de componentes principales (ACP) y MULTISPATI-ACP.....	93
<b>Tabla 3.7</b> Peso de las covariables ambientales (vectores propios) de los lotes de la zona 2 (Córdoba) en los tres primeros componentes del análisis de componentes principales (ACP) y MULTISPATI-ACP.....	94
<b>Tabla 3.8</b> Peso de las covariables ambientales (vectores propios) de los lotes de la zona 3 (Entre Ríos) en los tres primeros componentes del análisis de componentes principales (ACP) y MULTISPATI-ACP.....	95
<b>Tabla 4.1</b> Parámetros de estadística descriptiva de variables de rendimiento y calidad de soja para los lotes de la zona 1 (Buenos Aires).....	130
<b>Tabla 4.2</b> Parámetros de estadística descriptiva de variables de rendimiento y calidad de soja para los lotes de la zona 2 (Córdoba).....	130
<b>Tabla 4.3</b> Parámetros de estadística descriptiva de variables de rendimiento y calidad de soja para los lotes de la zona 3 (Entre Ríos).....	131

## INDICE DE FIGURAS

<b>Figura 1.1</b> Producción y área sembrada de soja en Argentina, Campañas 1995/96 - 2013/2014. (Fuente: Ministerio de Agricultura, Ganadería y Pesca de la Nación - SIIA).....	24
<b>Figura 2.1</b> Localización del área de estudio en el sudeste de la Provincia de Buenos Aires.....	32
<b>Figura 2.2</b> Distribución espacial de esquema de muestreo en grilla de tamaño 50 x 100 m ( <i>Derecha</i> ) y del esquema de muestreo generado por modelos con el conjunto de datos de entrenamiento ( <i>Izquierda</i> ).....	36
<b>Figura 2.3</b> División del conjunto completo en datos de calibración y sus respectivas sus subdivisiones (datos de entrenamiento y prueba) y datos de calibración.....	37
<b>Figura 2.4</b> Gráficos de densidad de (a) contenido de materia orgánica (MO) y (b) contenido de arcillas en el conjunto de datos completo.....	41
<b>Figura 2.5</b> Distribución espacial de cada uno de los predictores dentro del área de estudio.....	43
<b>Figura 2.6a</b> Comparación de la distribución de las muestras obtenidas a partir de Hipercubo Latino Condicionado (HCLc) (60 muestras) y Fuzzy C-Medias (FCM) (60 muestras) con respecto a la distribución total de todas las muestras (126 muestras), para todos los predictores.....	47
<b>Figura 2.6b</b> Comparación de la distribución de las muestras obtenidas a partir de Hipercubo Latino Condicionado (HCLc) (60 muestras) y Fuzzy C-Medias (FCM) (60 muestras) con respecto a la distribución total de todas las muestras (126 muestras), para todos los predictores.....	48
<b>Figura 2.6c</b> Comparación de la distribución de las muestras obtenidas a partir de Hipercubo Latino Condicionado (HCLc) (60 muestras) y Fuzzy C-Medias (FCM) (60 muestras) con respecto a la distribución total de todas las muestras (126 muestras), para todos los predictores.....	49
<b>Figura 2.7</b> Importancia de los predictores a partir de las clasificaciones realizadas con <i>Random Forest</i> con ambos esquemas de muestreo Hipercubo Latino Condicionado (HCLc) y Fuzzy C-Medias (FCM) usadas para predecir el (a) contenido de materia orgánica (MO) y (b) el contenido de arcilla a escala de lote. .....	51
<b>Figura 2.8</b> Gráfico de dispersión de puntos entre contenido de materia orgánica (MO) predicha versus MO observada utilizando (a) el esquema de muestreo de suelos Hipercubo Latino Condicionado (HCLc) y (b) Fuzzy C-Medias; y contenido de arcillas predicho versus observado utilizando (c) el esquema de muestreo de suelos Hipercubo Latino Condicionado (HCLc) y (d) Fuzzy C-Medias .....	54
<b>Figura 2.9</b> Comparaciones de la predicción espacial del contenido de materia orgánica (MO) utilizando (a) kriging ordinario a partir del conjunto de datos completo, (b) datos predichos con el modelo <i>Random Forest</i> (RF) con Hipercubo Latino Condicionado (HCLc) y RF con Fuzzy C-Medias (FCM). Comparación de la predicción espacial de contenido de arcillas utilizando (d) kriging ordinario a partir del conjunto de datos completo, (e) RF con HCLc y (f) RF con FCM).....	57
<b>Figura 3.1</b> Localización de la cartografía convencional de suelos INTA a escala 1:50.000 dentro del territorio argentino.....	58
<b>Figura 3.2</b> Localización de las tres zonas de estudio.....	64
<b>Figura 3.3</b> Distribución espacial y descripción pedológica de las unidades cartográficas de la cartografía convencional de suelos sobre el lote Aleluya – Buenos Aires.....	66
<b>Figura 3.4</b> Distribución espacial y descripción del contenido pedológico de las unidades cartográficas de la cartografía convencional de suelos sobre el lote La Beatriz – Buenos Aires.....	67

<b>Figura 3.5</b> Distribución espacial y descripción del contenido pedológico de las unidades cartográficas de la cartografía convencional de suelos sobre los lotes Mercapire (arriba) y San Esteban (abajo) – Córdoba.....	68
<b>Figura 3.6</b> Distribución espacial y descripción del contenido pedológico de las unidades cartográficas de la cartografía convencional de suelos sobre los lotes La Nueva Trinidad I (izquierda) y La Nueva Trinidad II (derecha) – Entre Ríos.....	69
<b>Figura 3.7</b> Coeficientes de correlación ( $r$ ) de las covariables ambientales de los lotes de la zona 1 (Buenos Aires). Los $r$ mostrados fueron significativos a un nivel de probabilidad $p \leq 0.05$ .....	79
<b>Figura 3.8</b> Coeficientes de correlación ( $r$ ) de las covariables ambientales de los lotes de la zona 2 (Córdoba). Los $r$ mostrados fueron significativos a un nivel de probabilidad $p \leq 0.05$ .....	83
<b>Figura 3.9</b> Coeficientes de correlación ( $r$ ) de las covariables ambientales de los lotes de la zona 3 (Entre Ríos). Los $r$ mostrados fueron significativos a un nivel de probabilidad $p \leq 0.05$ .....	88
<b>Figura 3.10</b> Índice resumen utilizado para evaluar el número óptimo de zonas por tipo de suelos en las seis lotes de las tres zonas.....	97
<b>Figura 3.11</b> Variabilidad de las covariables ambientales en cada zona delimitada en los lotes de la zona 1 (Buenos Aires).....	99
<b>Figura 3.12</b> Variabilidad de las covariables ambientales en cada zona delimitada en los lotes de la zona 2 (Córdoba).....	101
<b>Figura 3.13</b> Variabilidad de las covariables ambientales en cada zona delimitada en los lotes de la zona 3 (Entre Ríos).....	102
<b>Figura 3.14</b> Distribución espacial de las muestras de suelos en cada zona delimitada de todos los lotes en las tres zonas.....	103
<b>Figura 3.15</b> Medias de propiedades del suelo en cada zona delimitada en los lotes de la zona 1 (Buenos Aires). Barras verticales indican error estandar de la media para cada una de las zonas delimitadas.....	105
<b>Figura 3.16</b> Medias de propiedades del suelo en cada zona delimitada en los lotes de la zona 2 (Córdoba). Barras verticales indican error estandar de la media para cada una de las zonas delimitadas.....	107
<b>Figura 3.17</b> Medias de propiedades del suelo en cada zona delimitada en los lotes de la zona 3 (Entre Ríos). Barras verticales indican error estandar de la media para cada una de las zonas delimitadas.....	109
<b>Figura 3.18</b> Representación gráfica de los dos primeros ejes del análisis de componentes principales (ACP) elaborado a partir de los resultados de la Matriz de Distancia Euclidiana (MDE) entre propiedades del suelo analizadas y las propiedades de suelos de las series de suelos, para los lotes de la zona 1 (Buenos Aires).....	111
<b>Figura 3.19</b> Representación gráfica de los dos primeros ejes del análisis de componentes principales (ACP) elaborado a partir de los resultados de la Matriz de Distancia Euclidiana (MDE) entre propiedades del suelo analizadas y las propiedades de las series de suelos, para los lotes de la zona 2 (Córdoba).....	113
<b>Figura 3.20</b> Representación gráfica de los dos primeros ejes del análisis de componentes principales (ACP) elaborado a partir de los resultados de la Matriz de Distancia Euclidiana (MDE) entre propiedades del suelo analizadas y las propiedades de las series de suelos, para los lotes de la zona 3 (Entre Ríos).....	116

<b>Figura 4.1</b> Distribución espacial de las muestras de cultivo de soja en cada zona delimitada de cada uno de los lotes en las tres zonas de estudio.....	123
<b>Figura 4.2</b> Diferencia de precipitación acumulada mensual entre nov – dic de 2010 y ene – mar de 2011 (campana 2010/2011) e histórica 1990-2009 para los mismos meses.....	129
<b>Figura 4.3</b> Diferencias visuales de radiación interceptada alrededor de la etapa fenológica de soja R2-R4, en cada uno de los lotes analizados .....	133
<b>Figura 4.4</b> Medias de propiedades de cultivo en cada zona delimitada en el lote Aleluya – Buenos Aires.. Barras verticales indican error estandar de la media para cada una de las zonas delimitadas.....	135
<b>Figura 4.5</b> Medias de propiedades de cultivo en cada zona delimitada en el lote La Beatriz – Buenos Aires.. Barras verticales indican error estandar de la media para cada una de las zonas delimitadas.....	137
<b>Figura 4.6</b> Medias de propiedades de cultivo en cada zona delimitada en el lote Mercapire – Córdoba.. Barras verticales indican error estandar de la media para cada una de las zonas delimitadas.....	138
<b>Figura 4.7</b> Medias de propiedades de cultivo en cada zona delimitada en el lote San Esteban – Córdoba.. Barras verticales indican error estandar de la media para cada una de las zonas delimitadas.....	140
<b>Figura 4.8</b> Medias de propiedades de cultivo en cada zona delimitada en el lote LNT_I – Entre Ríos. Barras verticales indican error estandar de la media para cada una de las zonas delimitadas.....	141
<b>Figura 4.9</b> Medias de propiedades de cultivo en cada zona delimitada en el lote LNT_II – Entre Ríos. Barras verticales indican error estandar de la media para cada una de las zonas delimitadas.....	142
<b>Figura 4.10</b> Importancia de las propiedades del suelo para cada una de las propiedades del cultivo de soja en el lote Aleluya – Buenos Aires, a partir de las clasificaciones realizadas con el algoritmo <i>Random Forest</i> (RF).....	145
<b>Figura 4.11</b> Importancia de las propiedades del suelo para cada una de las propiedades del cultivo de soja en el lote La Beatriz – Buenos Aires, a partir de las clasificaciones realizadas con el algoritmo <i>Random Forest</i> (RF).....	146
<b>Figura 4.12</b> Importancia de las propiedades del suelo para cada una de las propiedades del cultivo de soja en el lote Mercapire – Córdoba, a partir de las clasificaciones realizadas con el algoritmo <i>Random Forest</i> (RF).....	148
<b>Figura 4.13</b> Importancia de las propiedades del suelo para cada una de las propiedades del cultivo de soja en el lote San Esteban – Córdoba, a partir de las clasificaciones realizadas con el algoritmo <i>Random Forest</i> (RF).....	149
<b>Figura 4.14</b> Importancia de las propiedades del suelo para cada una de las propiedades del cultivo de soja en el lote LNT_I – Entre Ríos, a partir de las clasificaciones realizadas con el algoritmo <i>Random Forest</i> (RF).....	152
<b>Figura 4.15</b> Importancia de las propiedades del suelo para cada una de las propiedades del cultivo de soja en el lote LNT_II – Entre Ríos, a partir de las clasificaciones realizadas con el algoritmo <i>Random Forest</i> (RF).....	153
<b>Figura 4.16</b> Comparación de medias de número de granos por metro cuadrado entre lotes, de acuerdo a las zonas delimitadas por tipo de suelo. Letras diferentes indican diferencias estadísticamente significativas ( <i>p</i>	

< 0.05). Barras verticales indican error estándar de la media para cada zona delimitada por tipo de suelo.....156

**Figura 4.17** Comparación de medias de peso de mil granos (*PI000*) entre lotes, de acuerdo a las zonas delimitadas por tipo de suelo. Letras diferentes indican diferencias estadísticamente significativas ( $p < 0.05$ ). Barras verticales indican error estándar de la media para cada zona delimitada por tipo de suelo.....159

**Figura 4.18** Comparación de medias de radiación interceptada (*RI*) entre lotes, de acuerdo a las zonas delimitadas por tipo de suelo. Letras diferentes indican diferencias estadísticamente significativas ( $p < 0.05$ ). Barras verticales indican error estándar de la media para cada zona delimitada por tipo de suelo.....160

**Figura 4.19** Comparación de medias de NDVI medido a través de GreenSekker® entre lotes, de acuerdo a las zonas delimitadas por tipo de suelo. Letras diferentes indican diferencias estadísticamente significativas ( $p < 0.05$ ). Barras verticales indican error estándar de la media para cada zona delimitada por tipo de suelo.....161

**Figura 4.20** Comparación de medias de contenido de aceite en grano de soja entre lotes, de acuerdo a las zonas delimitadas por tipo de suelo. Letras diferentes indican diferencias estadísticamente significativas ( $p < 0.05$ ). Barras verticales indican error estándar de la media para cada zona delimitada por tipo de suelo.....163

**Figura 4.21** Comparación de medias de contenido de proteína en grano de soja entre lotes, de acuerdo a las zonas delimitadas por tipo de suelo. Letras diferentes indican diferencias estadísticamente significativas ( $p < 0.05$ ). Barras verticales indican error estándar de la media para cada zona delimitada por tipo de suelo.....164

## RESUMEN

Las técnicas de cartografía digital de suelos (CDS) tienen un gran potencial para mejorar la consistencia, precisión, detalle y velocidad con la cual se produce la información espacial de propiedades del suelo a escala de lote. La información disponible relacionada con factores formadores del suelo, junto con recientes desarrollos informáticos y de análisis de datos, permitieron implementar procesos cuantitativos de generación de información acerca de las relaciones espaciales entre propiedades del suelo y cultivo de soja. El objetivo de esta tesis fue implementar y evaluar técnicas de cartografía digital de suelos (CDS) que ayuden a explicar y entender las relaciones entre patrones espaciales de propiedades del suelo y propiedades de rendimiento y calidad de soja a escala de lote. Para cumplir el objetivo, técnicas de CDS a escala de lote, fueron implementadas en 6 lotes, distribuidos en tres regiones agrícolas importantes de Argentina. La primera técnica de CDS consistió en integrar esquemas de muestreo de suelos basados en modelos (EBM) y el algoritmo Random Forest (RF) para predecir propiedades del suelo a escala de lote. Los resultados obtenidos sugieren que la integración EBM-RF puede llegar a ser una alternativa eficiente para predecir propiedades del suelo a escala de lote. Asimismo, se demostró que hipercubo latino condicionado (HCLc) es eficiente para capturar la variabilidad de predictores relacionados con factores de formación del suelo. La segunda técnica de CDS buscó aprovechar la información auxiliar digital disponible relacionada con factores formadores del suelo y la información contenida en la cartografía convencional de suelos disponible. De este modo se planteó una metodología para la desagregación de unidades cartográficas de suelo (DgUCS) y delimitación de zonas por tipo de suelo a escala de lote. Los resultados sugieren que la metodología de DgUCS fue eficiente para aumentar la escala de mapas de tipo de suelo en condiciones del sudeste bonaerense y el sur de Córdoba, mas no en condiciones del centro de Entre Ríos. Para esta zona, futuros trabajos se proponen para determinar las causas de la brechas de información entre la cartografía convencional de suelos y la encontrada en cada uno de los lotes. A partir de las zonas delimitadas por tipo de suelo, se analizaron las relaciones entre patrones espaciales de propiedades de suelo y cultivo de soja, en cada uno los lotes experimentales. Los resultados sugieren que la metodología de CDS utilizada para delimitar zonas por tipo de suelos, puede generar una base técnica y espacial para

delimitar zonas para manejo sitio específico, mas no para generar zonas para cosecha diferenciada por calidad. Una discusión general fue planteada en el capítulo final, en la cual se prioriza los impactos generales de los resultados obtenidos y se plantean trabajos futuros que requieren realizarse para corroborar los resultados de esta tesis. En general, esta tesis aportó de manera significativa, conocimientos de aplicación de técnicas de CDS en la generación rápida, fácil, precisa y económica de información de suelos a escala de lote. También demostró que los patrones espaciales complejos, junto con la presencia de tosca y el manejo inadecuado de los suelos son determinantes de la mayoría de propiedades limitantes del rendimiento de la soja. Además ratificó que la información digital disponible está siendo subutilizada y que puede aportar mucho para entender las dinámicas espacio-temporales entre propiedades del suelo y cultivos. Los resultados obtenidos en esta tesis invitan a replantear las técnicas, metodologías y uso de la información disponible, para generar información de suelos a escala de lote. Además, sugieren involucrar una dinámica de trabajos que permitan evaluar la aplicación de tecnologías emergentes en los procesos de generación de conocimientos de patrones espaciales de suelos, y su relación con propiedades del cultivo, a escala de lote.

## ABSTRACT

Digital soil mapping techniques (DSM) have great potential to improve the consistency, accuracy, detail and speediness at which soil survey information at farm-scale is produced. Easily available information related to soil-forming factors, together with recent computer developments and data analysis allowed us to implement quantitative processes in order to generate information about spatial relationship between soil properties and soybean crop. The aim of this thesis was to implement and examine DSM techniques which benefit to explain and understand the relationships among spatial patterns of soil properties and soybean yield and quality parameters at farm-scale. To accomplish this aim, DSM techniques at farm-scale were implemented in 6 agricultural fields, localized in three main agricultural regions of Argentina. The first DSM technique consisted in integrating model-based soil sampling scheme (SBM) and the Random Forest algorithm (RF) for predicting soil properties at farm-scale. Results suggest that the integration SBM-RF could be an efficient alternative to predict soil properties at farm-scale. Likewise, it was demonstrated that Conditioned Latin Hypercube (cLHS) was efficient to capture the variability of predictors related to soil-forming factors. The second DSM technique used ancillary digital information related to soil-forming factors and soil survey from conventional soil mapping. A methodology for disaggregating soil map units (DgSUM) and delineating zones according to different soil types at farm-scale was performed. Results suggest that the DgSUM was effective in the south-eastern environments of Buenos Aires province as well as in the southern of Cordoba province, but not in Entre Rios province. Particularly in this zone, further studies may be proposed in order to determine the reasons of the information gap found among conventional soil mapping and the results presented in this thesis. Based on zones delineated by soil type, the relations among spatial patterns of soil properties and soybean crop in each field were analyzed. Results suggest that DSM techniques used for delineating zones by soil type could be useful to generate a technical and spatial base in order to delineate site-specific management zones, whereas for delineating quality zone for differential harvesting, DSM techniques would not be appropriated. A general discussion was proposed in the last chapter. In this discussion, the overall impacts of the results were prioritized and further studies need to be carried out in order to corroborate the results presented. In general, this

thesis contributed to the knowledge of DSM techniques which focusing on generating soil information at farm-scale in a fast, easy, accurate and inexpensive way. It was also demonstrated that complex spatial patterns together with the presence of a tosca layer and the wrong way soil management are determinants in the distribution of yield limiting factors in soybean. Furthermore, this thesis provided that digital soil information is not being efficiently used and it may be more valuable for understanding the spatial and temporal dynamic among soil properties and crops. This thesis leads to reconsider the techniques, methodologies and use of soil available information in order to generate soil information at farm-scale. Also, it suggests further researches about the evaluation of these technologies in the generation of soil spatial patterns and its relation to crop properties at farm-scale.

# CAPÍTULO 1

## 1. Introducción General

### 1.1. Importancia de obtener información de suelos a escala de lote

Décadas de investigación han determinado que la variabilidad espacial del suelo a escala de lote (escala  $>1:2.000$ ) impacta considerablemente muchos procesos agronómicos y ambientales (Cook *et al.*, 2008). Esto significa que los resultados de experimentos agronómicos y/o ambientales pueden llegar a ser imprecisos en la medida en que el efecto de la variabilidad del suelo no se tiene en cuenta de forma explícita. Recientemente, la disponibilidad de tecnologías de agricultura de precisión ha brindado un panorama más detallado de la variabilidad del rendimiento a escala de lote, en la cual el efecto espacio temporal de las propiedades del suelo es dominante y a menudo, extremadamente grande. La significancia práctica de conocer la variabilidad espacial del suelo a escala de lote puede ser determinada en relación con la seguridad alimentaria, seguridad hídrica, seguridad energética, la eficiencia de los servicios ecosistémicos, protección a la biodiversidad y mitigación del cambio climático (Hartemink *et al.*, 2010; McBratney *et al.*, 2014).

La ausencia de información de suelos a escala de lote puede incrementar los procesos de degradación del suelo y por ende, afectar la seguridad alimentaria (Lal, 2013). Actualmente, el suelo agrícola per cápita es extremadamente bajo y en proceso de decrecimiento. La erosión, salinización, reducción de nutrientes disponibles, degradación de estructura, compactación y la reducción de la actividad y especies de la biota del suelo, son los procesos de degradación dominantes. Se estima que el 23.5 % de la superficie de la tierra es propensa a sufrir procesos de degradación (Lal y Stewart, 2009). Técnicas efectivas para obtener información de suelos pueden ser útiles para mejorar las estrategias de seguridad alimentaria, permitiendo un abastecimiento idóneo con una intensificación sustentable de los sistemas agroproductivos.

Es ampliamente conocido que el suelo es un componente clave de todos los ecosistemas terrestres. Por lo tanto, aumentar la disponibilidad de información de suelos puede incrementar la eficiencia de los servicios ecosistémicos (Bommarco *et al.*, 2013). Entre los servicios ecosistémicos que están directamente relacionados con el suelo se listan: (i) producción de biomasa y alimentos; (ii) servicios ambientales tales como ciclo de nutrientes, almacenamiento, filtración y transformación de contaminantes; (iii) hábitat biológico para organismos y mantenimiento de su base genética; (iv) fuente de materias primas no renovables; (v) herencia cultural y (vi) sustrato para mantenimiento de estructuras (Rossiter, 2012). Normalmente, estos servicios ecosistémicos son aprovechados de manera sincrónica, paralela y simultánea. Producir eficazmente información de suelo permitiría modelar funciones para evaluar la sostenibilidad de los servicios ecosistémicos. Además, mejoraría la segmentación ambiental de los paisajes y su caracterización.

El suelo es un componente clave en el ciclo de carbono (C) a escala global y tiene una influencia importante sobre los flujos globales de CO<sub>2</sub> entre la biosfera terrestre y la atmósfera (Lal, 2010). Generalmente, el suelo puede ser una fuente o disipador de C atmosférico. De acuerdo con esto, la disponibilidad de información de suelos a escala de lote puede favorecer la implementación de prácticas de manejo que reduzcan reducción de los niveles netos de CO<sub>2</sub> en la atmósfera y por ende, mitigar los efectos del cambio climático.

La escasez de agua es una amenaza inmediata (Lawford *et al.*, 2013; Misra, 2014). El efecto combinado del crecimiento de la población mundial, el cambio climático, la expansión industrial y la degradación de los suelos dan lugar a una necesidad urgente de implementar estrategias para mitigar y monitorear todos los componentes clave del ciclo del agua. El suelo es uno de ellos, el almacenamiento y la calidad del agua están directamente relacionados con las propiedades físicas, químicas y microbiológicas (McBratney *et al.*, 2014). De acuerdo con esto, la disponibilidad de información de suelos a escala de lote permite entender y modelar los procesos de retención, filtración y transformación del agua en el suelo. Además, permite determinar la contribución del

suelo en la provisión de agua para consumo humano, agricultura y necesidades de los ecosistemas.

La demanda de información de suelos a escala de lote está creciendo permanentemente (Boettinger *et al.*, 2010). Esta demanda va a ser especialmente crítica en zonas donde la información de suelos es muy limitada. En el caso de Argentina, aunque existe información valiosa, urge generar nueva información masiva de suelos a escala de lote, de modo que ayuden a mitigar el impacto del desarrollo de las zonas agrícolas actuales y la expansión de la frontera agrícola.

## **1.2. Cartografía Digital de Suelos (CDS).**

La Cartografía Digital de Suelos (CDS) es una disciplina que integra datos de campo, laboratorio, sensores del suelo y modelos digitales de elevación con métodos cuantitativos para inferir los patrones de distribución de suelos a varias escalas temporales y espaciales (Grunwald, 2010). Puntualmente, CDS se define como *la elaboración de sistemas de información espacial de suelos a través de modelos numéricos infiriendo las variaciones espacio-temporales de los tipos y propiedades del suelo a partir de observaciones del suelo y conocimiento relacionado con covariables ambientales tales como topografía, organismos, material parental, clima y uso del suelo* (Hartemink *et al.*, 2008; Thompson *et al.*, 2012).

La CDS y cartografía convencional de suelos tienen la misma base técnica. Sin embargo, la diferencia radica en la forma como el modelo deriva la información de suelos a partir de los datos disponibles (Hartemink *et al.*, 2008). En general, ambos tipos de cartografía requieren un modelo predefinido de formación del suelo y datos de propiedades del suelo o de otras covariables ambientales que tengan un impacto significativo en la formación y su distribución espacial del suelo. Es decir, tanto la CDS como la cartografía convencional de suelos requieren ingresar información de suelos y covariables ambientales que permitan caracterizar el ambiente en el que la formación del suelo tuvo lugar. Ahora bien, los modelos de cartografía convencionales se basan en estudios empíricos y cualitativos, que dependen de la correlación y la formulación mental

del espacio que tenga el cartógrafo en el momento del relevamiento. Este enfoque siempre va a requerir un intensivo y costoso trabajo de campo. Las decisiones son generalmente tomadas en el campo, donde todas las covariables ambientales pueden ser directamente observadas y la información del suelo, es el fruto de un proceso deductivo del cartógrafo (Lagacherie *et al.*, 2006).

Los mapas de suelos disponibles en Argentina han sido el resultado de la aplicación metodológica de observaciones sistemáticas de campo complementadas con análisis de laboratorio y métodos cartográficos análogos. Al respecto, Hartemink *et al.* (2008) plantea que este tipo de mapas tienen importantes desventajas, *(i)* no proporcionan información directa sobre la dinámica temporal de propiedades del suelo; *(ii)* son inflexibles para estudios cuantitativos; *(iii)* la información tabulada es escasamente tenida en cuenta; *(iv)* la información es presentada a escalas que no pueden ser utilizadas eficientemente para aplicaciones particulares; y *(v)* el modelo de datos implícito en los mapas de polígonos dificulta su integración con otros formatos de datos de covariables ambientales. Los mapas convencionales de suelos fueron una pertinente solución en la era de los “*datos análogos*”. Sin embargo, herramientas informáticas, estadísticas y tecnológicas desarrolladas en las últimas dos décadas, pueden mejorar la calidad de la información de suelos y el acceso a la misma, haciendo parte de lo que hoy se conoce como la era de los “*datos digitales*”.

Los principales modelos cuantitativos integrales para la elaboración de CDS tienen como base la ecuación de factores formadores del suelo propuesta por Jenny (1941). Esta ecuación, comúnmente conocida como “ecuación de Jenny” identificó que los principales factores formadores del suelo, son el clima (cl), los organismos (o), el relieve (r), material parental (p) y el tiempo (t):

$$S = f(cl, o, r, p, t, \dots)$$

La función de Jenny se focaliza en la predicción (S) de propiedades físicas, químicas y biológicas del suelo teniendo en cuenta la posición geográfica y su relación con las zonas aledañas. Por lo general, esta característica es utilizada en las interpolaciones

geoestadísticas para predecir las propiedades del suelo de un área determinada, a partir de puntos conocidos. Aplicando descripciones de suelos en puntos determinados para su posterior interpolación geoestadística, es conveniente considerar dentro del grupo de factores formadores del suelo la localización geográfica de cada punto (Dobos y Hengl, 2009; Lagacherie, 2008).

A partir de fuentes de información de covariables ambientales y la localización geográfica es posible predecir propiedades del suelo con una aceptable precisión.

Por su parte, McBratney *et al.* (2003) identifico 7 factores de predicción espacial de suelos:

**S:** Propiedades del suelo en una misma ubicación

**C:** Clima

**O:** Organismos

**R:** Relieve

**P:** Material Parental

**A:** Tiempo

**N:** Localización

Y formuló la siguiente ecuación llamada SCORPAN:

$$Sa = f(s, c, o, r, p, a, n) \text{ y } Scl = f(s, c, o, r, p, a, n)$$

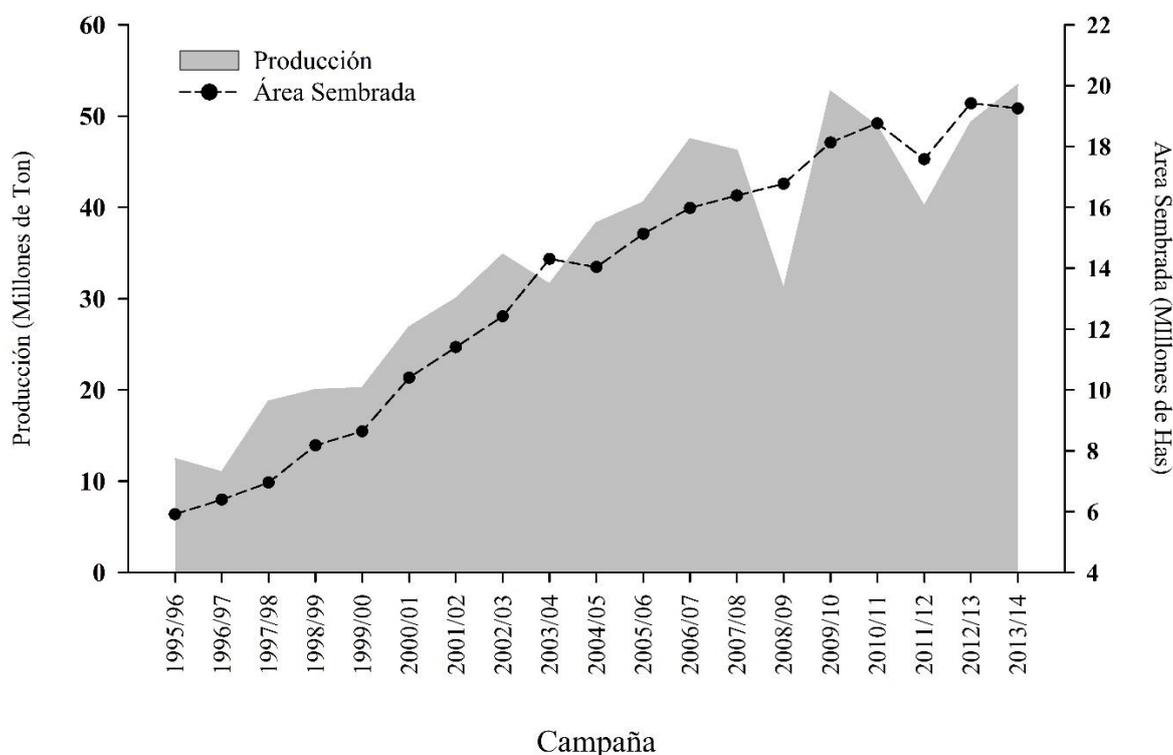
Donde  $Sa$  es el valor del atributo estimado y  $Scl$  es la clase de propiedad del suelo estimada. De acuerdo con estos autores, con este enfoque se resumen las demandas de datos digitales del suelo que se requieren para elaborar CDS.

Muchos avances de la revolución digital y tecnológica no han sido aplicados para elaborar CDS. Sin embargo, es destacable los avances parciales que han tenido algunas tecnologías aplicadas a la predicción de propiedades del suelo a partir de sensores proximales y directos del suelo, imágenes satelitales, modelos digitales de elevación y mapas de rendimiento (Adamchuk *et al.*, 2004; Lesch, 2005; Mulder *et al.*, 2011; Piikki *et al.*, 2013). Al mismo tiempo, han sido de gran utilidad las técnicas y procedimientos de estadística espacial que han sido desarrolladas y aplicadas a la predicción de propiedades del suelo en áreas con mucha, poca y ninguna información preexistente de suelos (Mucherino *et al.*, 2009; Patriche *et al.*, 2013). En esta tesis, se propone generar técnicas de CDS y evaluarlas para las condiciones de los suelos de algunas regiones de la Argentina.

### **1.3. El cultivo de la soja en Argentina**

En 1999, el consumo medio estimado de alimentos fue de 2789 Kcal/persona/día, mientras que para el 2030 y 2050 será de 3040 y 3130 Kcal/persona/día, respectivamente (FAO, 2002; FAO, 2011). Por lo tanto, para el 2050 la contribución de los cereales a la alimentación en los países en desarrollo habrá caído de 56% a 46%, y el de carnes, lácteos y aceites vegetales aumentaría de 20% a 29%. Para ese mismo año la producción anual de carne tendrá que ser de 470 millones de toneladas, frente a las 200 millones que se produjeron en 2010.

El aumento del consumo de carne y de alimentos con mayor contenido energético, requerirá de una mayor producción de alimentos concentrados, lo cual llevaría a que se tenga que incrementar en un 80% y 140% la producción de maíz y soja, respectivamente. Teniendo en cuenta estas proporciones, se requerirá producir para el 2050, 515 millones de toneladas de soja, sin contar con las que se requerirán para biocombustibles (Godfray *et al.*, 2010).



**Figura 1.1** Producción y área sembrada de soja en Argentina, Campañas 1995/96 - 2013/2014.  
(Fuente: Ministerio de Agricultura, Ganadería y Pesca de la Nación - SIIA)

En la actualidad, Argentina es el tercer exportador mundial de soja y el primer exportador de pasta y aceite de soja (Figura 1). El vertiginoso proceso de consolidación de la Argentina como productor y exportador de soja y sus subproductos se le conoce con el término “sojización”.

Es ampliamente conocido que la sojización en la Argentina ha tenido un impacto directo sobre las propiedades físicas, químicas y microbiológicas del suelo. Generalmente, la soja es un cultivo que provee una limitada cantidad de residuos impactando directamente sobre la relación carbono - nitrógeno. (C:N) (Novelli *et al.*, 2011). Este impacto promueve la degradación del suelo y lo expone a procesos erosivos intensos durante periodos de lluvia. Además, está comprobado que en rotaciones donde la soja se siembra de manera intensiva, se afecta notoriamente el almacenamiento del carbono orgánico del suelo y se reduce la macroagregación (Studdert, 2000).

A pesar de que existen numerosos estudios que han cuantificado el efecto degradativo de la sojización intensiva sobre las propiedades del suelo, explícitamente, la mayoría no han tenido en cuenta el efecto de la variabilidad espacial intrínseca del suelo (Bellaloui *et al.*, 2009; Leonardi, 2012). Los motivos de esto son: *(i)* la ausencia de métodos y técnicas que de manera rápida, económica y precisa, permitan determinarla y *(ii)* ausencia de información auxiliar relacionada con los factores formadores del suelos planteados por Jenny (1941), que permitan generar modelos cuantitativos de predicción espacial de propiedades del suelo. Específicamente, esta tesis aborda de manera directa el primer motivo y de manera indirecta, evalúa varias fuentes de información auxiliar relacionadas con los factores formadores del suelo. Lograr integrar técnicas efectivas de generación de cartografía digital de suelos a escala de lote con fuentes de información gratuitas disponibles, puntualmente, permitirá ser más específicos y precisos en las delimitaciones para manejo sitio específico y determinación de zonas para cosecha diferenciada por calidad en soja.

#### **1.4. Justificación**

La presente tesis doctoral se justifica de acuerdo a los siguientes planteamientos:

- Existen pocos trabajos que propongan el uso y calibración de esquemas de muestreo para determinar y caracterizar la variabilidad espacial de propiedades del suelo a escala de lote, a partir de información auxiliar digital disponible de factores formadores del suelo.
- Existen pocos trabajos que propongan la aplicación de técnicas de cartografía digital de suelos a escala de lote, que permitan delimitar zonas por tipo de suelo, a partir del uso simultáneo de información digital y análoga de cartografía convencional de suelos y covariables ambientales de factores formadores del suelo en condiciones de los sistemas agroproductivos de la Argentina.
- En Argentina, es necesario evaluar la aplicación de técnicas de cartografía digital de suelos a escala de lote, en la determinación de las relaciones de los patrones espaciales

de propiedades del suelo y cultivo de soja. Esto se debe a que se demanda información de suelos que permita la implementación de prácticas de manejo que mitiguen el impacto del sistemas agroproductivo sobre los recursos naturales.

- Existe escasa información que determine el uso potencial de técnicas de cartografía digital de suelos en la delimitación de zonas para la cosecha diferenciada del cultivo de soja, con énfasis en las condiciones que caracterizan los sistemas agroproductivos de la Argentina.

### 1.5. Hipótesis

Las hipótesis planteadas son:

- **Hipótesis 1.**

*“La integración de fuentes de información de covariables ambientales, esquemas de muestreo basados en modelos y el algoritmo de regresión “Randon Forest”, permite generar de manera rápida, precisa y puntual, **modelos eficientes de predicción de propiedades del suelo a escala de lote.**”*

- **Hipótesis 2.**

*“La cartografía digital de suelos elaborada a partir de la cartografía convencional de suelos disponible, fuentes de información de covariables ambientales y técnicas de desagregación espacial, permiten determinar de una manera rápida, precisa, puntual y a escala de lote, **patrones espaciales de propiedades del suelo,** en condiciones productivas de tres regiones agrícolas de Argentina.”*

- **Hipótesis 3.**

“La cartografía digital de suelos elaborada a partir de la cartografía convencional de suelos disponible, fuentes de información de covariables ambientales y técnicas de desagregación espacial, permiten determinar de una manera rápida, precisa, puntual y a escala de lote, **las relaciones entre patrones espaciales de propiedades del suelo y propiedades del rendimiento y calidad de soja**, en condiciones productivas de tres regiones agrícolas de Argentina, generando una base técnica y espacial para la delimitación de zonas para el manejo sitio específico del cultivo de la soja y zonas para cosecha diferenciada por calidad de grano de soja.”

## 1.6. Objetivos

El objetivo general de esta tesis fue implementar y evaluar técnicas de cartografía digital de suelos que ayuden a explicar y entender las relaciones entre patrones espaciales de propiedades del suelo y propiedades de rendimiento y calidad de soja a escala de lote, en tres regiones agrícolas importantes de la Argentina.

Los objetivos específicos de la presente tesis doctoral fueron:

- Predecir de manera rápida, precisa y puntual diferentes propiedades del suelo a escala de lote, a partir de la integración de fuentes de información de covariables ambientales, esquemas de muestreo basados en modelos y el algoritmo de regresión “*Random Forest*” (Capítulo 2).
- Validar para tres regiones agrícolas de Argentina, una metodología de desagregación de unidades cartográficas de suelo a escala de lote, a partir de la integración de MULTISPAT-PCA y análisis de distancia euclidiana, teniendo como fuente de información la cartografía convencional de suelos a escala 1:50.000 INTA y la información auxiliar de covariables ambientales relacionadas con suelos. (Capítulo 3).

- Caracterizar y analizar las relaciones entre patrones espaciales de propiedades del suelo y del cultivo de soja a escala de lote, a partir de la implementación de técnicas de cartografía digital de suelos (*Capítulo 4*).
- Determinar si las técnicas de cartografía digital de suelos permiten obtener una base técnica y espacial sólida para la delimitación de zonas para el manejo sitio específico del cultivo de la soja y zonas para cosecha diferenciada por calidad de grano de soja (*Capítulo 4*).

### **1.7. Estructura de la tesis**

Esta tesis doctoral se concentró en evaluar la implementación de técnicas de cartografía digital de suelos a escala de lote y su aplicación al estudio de la variabilidad espacial de propiedades de rendimiento y calidad de soja. En los capítulos 2, 3 y 4 se presentaran los resultados de los experimentos realizados para cumplir los objetivos generales. El capítulo 2 se concentró en evaluar un modelo “*Randon Forest*” para la predicción de propiedades de suelo a escala de lote, a través de la integración de información auxiliar relacionada con covariables ambientales y factores formadores del suelo; y esquemas de muestreo basados en modelos. El capítulo 3 se concentró en proponer y validar una técnica de cartografía digital de suelos para la desagregación espacial de unidades cartográficas de la cartografía convencional de INTA, a escala de lote, a partir de información auxiliar relacionada con covariables ambientales y factores formadores de suelo. La técnica de desagregación propuesta busca plantear una manera rápida, fácil y precisa para delimitar zonas por tipo de suelo a escala de lote. El capítulo 4 se concentró en implementar técnicas de cartografía digital de suelos a escala de lote, con el objetivo de caracterizar y analizar las relaciones entre patrones espaciales de propiedades del suelo y propiedades de rendimiento y calidad del grano de soja. Por último, el capítulo 5 se concentró en plantear la discusión general de todos los resultados de esta tesis y determinar las conclusiones y propuestas de investigación para futuros estudios.

## CAPITULO 2

### **Predicción espacial de propiedades del suelo a escala de lote integrando esquemas de muestreo de suelos basados en modelos y regresión “*Random Forest*”**

#### **1. Introducción**

La eficiencia de las prácticas de manejo sitio específico y de modelos hidrológicos dependen de la información espacial disponible a escala de lote (escala  $>1:2.000$ ) (Dobos y Hengl, 2009; Peralta *et al.*, 2015). Generalmente, la predicción espacial de propiedades del suelo a escala de lote es difícil (Amiotti *et al.*, 2001; Kuang y Mouazen, 2012; Zhu y Lin, 2011) debido a que los patrones espaciales del suelo dependen de una compleja interacción entre los factores formadores tales como topografía, material parental, clima, organismos, localización, tiempo y prácticas de manejo del suelo (Jenny, 1941; McBratney *et al.*, 2003). Además, medir la variabilidad espacial de suelos es laborioso y consume mucho tiempo. Por lo tanto, es necesario generar modelos que permitan predecir la distribución espacial de las propiedades del suelo a escala de lote.

La Pampa argentina es una extensa llanura plana de alrededor de 50 millones de hectáreas y es considerada una de las regiones agrícolas más importantes del mundo (Alvarez y Steinbach, 2009; Satorre y Slafer, 1999). En esta región, las bases de datos espaciales de la cartografía convencional de suelos INTA, tales como GeoINTA, han sido por décadas, la principal fuente de información de la variabilidad espacial de suelos (INTA, 2013a). En GeoINTA, los suelos son representados como unidades cartográficas con límites y contornos precisamente definidos. Como resultado de esto, los mapas de suelos de la cartografía convencional INTA son mapas “*cloropléticos*” a escala 1:50.000 donde cada polígono representa supuestamente, áreas de suelos similares. Esta representación de los suelos es de tipo discreta e imprecisa a escala de lote. Además, es incompatible con otras fuentes de información auxiliar relacionadas con la variabilidad espacial de propiedades del suelo tales como modelos digitales de elevación, imágenes satelitales y sensores del suelo (Carré *et al.*, 2007). Esta incompatibilidad es una

considerable limitante para aplicar manejo sitio específico y para calibrar modelos hidrológicos.

Técnicas de cartografía digital de suelos (CDS) o *digital soil mapping* pueden reducir los costos asociados con mejorar la información de la variabilidad espacial de propiedades del suelo a escala de lote (Cressie y Kang, 2010), debido a que logran identificar las relaciones espaciales entre propiedades del suelo e información auxiliar disponible (McBratney *et al.*, 2003). Generalmente, esta información auxiliar es obtenida a partir de diversas fuentes tales como modelos digitales de elevación, imágenes satelitales y sensores del suelo. Debido a esto, numerosos estudios han demostrado que CDS es un enfoque operativo científicamente validado para predecir propiedades del suelo (Boettinger, 2010; Hartemink *et al.*, 2010; Kidd *et al.*, 2015). En este estudio, nosotros utilizamos técnicas de CDS para predecir la variabilidad espacial de la materia orgánica (MO) y el contenido de arcilla (As) a escala de lote. Ha sido ampliamente documentado que MO y As son propiedades del suelos que determinan varios procesos físicos, químicos y microbiológicos (Lal y Stewart, 2009; Piikki *et al.*, 2013; Sumner, 1999). Además, la determinación de la variabilidad espacial de MO y As puede ser importante para la implementación de modelos hidrológicos y para la implementación de manejo sitio-específico de cultivos (Boettinger, 2010; Levi y Rasmussen, 2014).

El esquema de muestreo es un componente importante en el proceso de implementación de técnicas de CDS (Brungard y Boettinger, 2010; Hengl *et al.*, 2004b) debido a que se requiere asegurar una óptima localización y cantidad de muestras de suelos teniendo en cuenta todas las fuentes de información espacial auxiliar disponibles. Tradicionalmente, la distribución espacial de un esquema de muestreo de suelos se establecía de acuerdo a (i) las decisiones subjetivas del cartógrafo o (ii) al enfoque estadístico del muestreo utilizado tales como muestreo aleatorio simple, modelos aleatorio estratificado, muestreo sistemático o en grilla regular, modelos adaptado a cluster o muestreo compuesto (Conklin Jr, 2004; U.S. Environmental Protection Agency, 2002). Estos enfoques introducen sesgo y no aportan a lograr un sistema de clasificación cuantitativo de suelos. Además, la mayoría no pueden ser utilizados en interpolación

espacial geoestadística, no consideran la correlación espacial y no tienen en cuenta la posición relativa de las muestras de suelo (Mulder *et al.*, 2013).

En contraste, recientes estudios han demostrado que los esquemas de muestreo basados en modelos (MBM) son eficientes debido a que tienen en cuenta la variabilidad espacial de múltiples fuentes de información auxiliar, consideran su correlación espacial y la posición relativa de las muestras de suelo seleccionando la menor cantidad de estas (Rad *et al.*, 2014). Dos algoritmos para MBM que han sido ampliamente utilizados en CDS son el hipercubo latino condicionado (HCLC) (Minasny y McBratney, 2006) y el fuzzy c-medias (FCM) (Debaene *et al.*, 2014; Miklos *et al.*, 2010). Sin embargo, la aplicación de estos dos algoritmos para predecir propiedades del suelos a escala de lote ha sido escasa (Ramirez-Lopez *et al.*, 2014; Schmidt *et al.*, 2014). Adicionalmente, no existe evidencia de estudios de CDS a escala de lote en la condiciones de la Pampa Argentina y Entre Ríos.

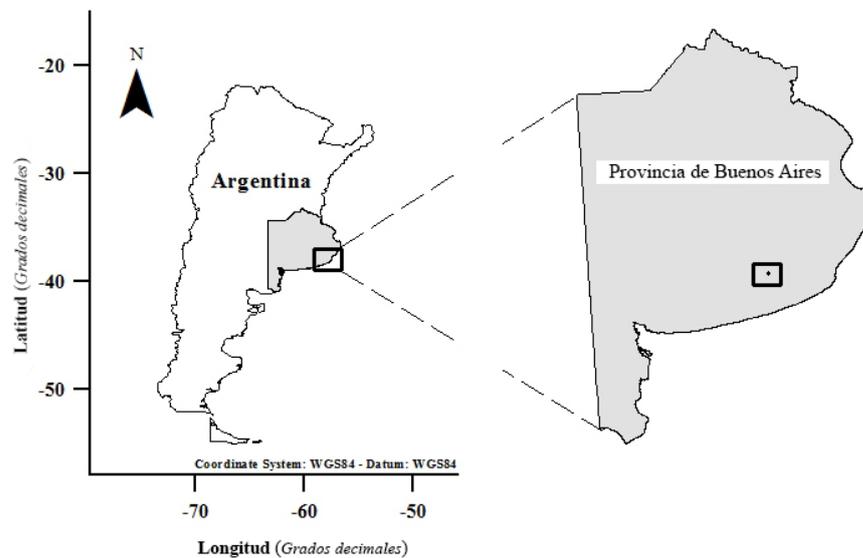
Diversas fuentes de información auxiliar de acceso rápido y económico relacionadas con factores de formación del suelo se encuentran ampliamente disponibles en la mayoría de las principales áreas agrícolas del mundo, incluyendo la pampa Argentina (Hartemink *et al.*, 2010; Peralta *et al.*, 2013). Esta disponibilidad de información auxiliar representa una oportunidad para generar modelos eficientes para predecir propiedades del suelo a escala de lote. Entre los tipos de modelos utilizados para predecir propiedades de suelos en el contexto de CDS, se encuentra regresión Kriging (Levi y Rasmussen, 2014), árboles de regresión y clasificación (Kheir *et al.*, 2010), regresión logística multinomial (Kempen *et al.*, 2009) y modelos compuestos de inferencia de tipos de suelos tales como SoLIM (Zhu *et al.*, 1997). Recientes estudios de CDS se han focalizado en utilizar el algoritmo “*Random forest*” (RF) como modelo de predicción (Breiman, 2001). Este algoritmo es un ensamble de árboles de regresión. Cada árbol de regresión es creado utilizando un muestreo bootstrap y cada nodo es dividido de acuerdo a un subconjunto aleatorio de predictores. Al final, los resultados de todos los árboles de regresión son ensamblados para proveer una predicción final (Verikas *et al.*, 2011). RF ha sido eficiente en estudios de CDS para aumentar la escala de la cartografía convencional de suelos (Rad *et al.*, 2014) y para predecir la distribución espacial de los tipos de material parental a escala regional

(Heung *et al.*, 2014), el contenido de MO del suelo (Wiesmeier *et al.*, 2011) y la textura del suelo (Ließ *et al.*, 2012). RF es un eficiente algoritmo debido a que es robusto, tiene poco sesgo, tiene alta eficiencia de predicción y no es sensible al sobreajuste (Díaz-Urriarte y De Andres, 2006). Además, RF tiene como complemento que identifica las variables con mayor capacidad predictiva y hace una clasificación de importancia de las mismas (Genuer *et al.*, 2010).

El objetivo de este estudio es evaluar un modelo RF para predecir MO y As a escala de lote utilizando información digital auxiliar de suelos y esquemas de muestreo basados en modelos. Este trabajo busca documentar un caso aplicado de técnicas de CDS a escala de lote en condiciones de la pampa Argentina y discutir una metodología práctica para mejorar la predicción de propiedades del suelo combinando MBM y regresión RF.

## 2. Materiales y métodos

### 2.1. Zona de estudio



**Figura 2.1** Localización del área de estudio en el sudeste de la Provincia de Buenos Aires

El área de estudio es un lote agrícola de 68 hectareas localizado en el sudeste de la provincia de Buenos Aires, Argentina ( $37^{\circ}55'16''\text{S}$ ,  $59^{\circ}07'59''\text{W}$ , Datum WGS84) (Figura 2.1). Este lote fue seleccionado debido a que representa la variabilidad de elevación, posición del paisaje, reflectancia de superficie, rotación de cultivos y tipos de suelos característicos del sudeste de la provincia de Buenos Aires. La rotación más utilizada en la zona incluye maíz, soja o girasol en verano y trigo o cebada en invierno (Sadras y Calviño, 2001). La media anual de precipitaciones es 913 mm con máxima precipitación (> 60%) durante el periodo de octubre a marzo. La temperatura media anual es  $14^{\circ}\text{C}$  (Pascale y Damario, 2004). Los suelos son Argiudoles típicos y Paleudoles petrocálcicos (Soil Survey Staff, USDA) y el régimen de humedad del suelo es údico (INTA, 2010a).

## 2.2. Variables predictoras

La tabla 2.1 muestra la descripción y resolución espacial de cada uno de los predictores y su relación con algunos los factores formadores del suelo propuestos por Jenny (1941). En general, todos los predictores fueron obtenidos a partir de información digital auxiliar tales como modelos digitales de elevación, imágenes satelitales, sensores del suelo y mapas de rendimiento.

Datos de elevación fueron recolectados utilizando un GPS con corrección diferencial (DGPS – Trimble® R3) en transectas separadas cada 20 m. A partir de estos, un modelo digital de elevación (MDE) fue generado mediante kriging ordinario. Índices de topografía tales como pendiente, factor de pérdida de suelos (LS Factor), pendiente de captación (PC), índice topográfico de humedad (ITH), índice de potencial de flujo (IPF) e índice de rugosidad del terreno (IRT) fueron calculados a partir de MDE utilizando el software SIG de Análisis Geo-científico Automatizado SAGA-GIS v2.1 (Olaya y Conrad, 2009; SAGA Development Team, 2012).

**Tabla 2.1** Descripción y resolución espacial de los predictores y su relación con los factores de formación del suelo propuestos por Jenny, (1941)

Factor Formador del suelo	Predictores	Resolución/ Densidad*	Descripción	Referencia
Suelo (s)	CE <sub>30</sub>	109 data/ha	Conductividad Eléctrica Aparente en mS * m <sup>-1</sup> a 0-30 cm de profundidad	Corwin y Lesch, 2003
	CE <sub>60</sub>	109 data/ha	Conductividad Eléctrica Aparente en mS * m <sup>-1</sup> a 0-90 cm de profundidad	Corwin y Lesch, 2003
Organismos (o)	NDVI - Escenas Landsat 5 TM	30 m	Índice de vegetación diferencial normalizado (NDVI) fue calculado como (Band 4 - Band 3)/(Band 4 + Band 3).	Boettinger, et al. 2008
	Mapa de rendimiento	113 data/ha	Fechas de escenas: enero/2005, Febrero/2007 and Enero/2008.	
Relieve (r)	Elevación	99 data/ha	Mapas de rendimiento de trigo 2002 y 2003; y soja 2004, 2005 y 2008.	Blackmore, 2000
	Pendiente	10 m	Elevación en metros	Hengly Reuter, 2008
	Factor de pérdida de suelo (LS Factor)	10 m	Gradiente de pendiente en radianes, para cada archivo raster en un modelo digital de elevación (MDE).	Hengly Reuter, 2008
	Pendiente de captación	10 m	Índice de erosión y deposición del suelo usando la ecuación universal de pérdida de suelos (USLE).	Hengly Reuter, 2009
	Índice Topográfico de Humedad (ITH)	10 m	Indica la tasa potencial de escorrentía superficial, densidad de drenaje y el desague del relieve.	Hengly Reuter, 2010
	Índice de rugosidad del terreno (IRT)	10 m	Describe el potencial que tiene una localización a ser saturado teniendo en cuenta las características geométricas del lote y las áreas de captación.	Olaya y Connad, 2009
	Índice del potencial de flujo (IPF)	10 m	Describe el potencial de flujo erosivo que tiene una localización específica de acuerdo al modelo digital de elevación	Flornsky, 2012
	Índice de rugosidad del terreno (IRT)	10 m	Indica la ondulación del terreno, que tan complejo es.	Hengly Reuter, 2009

\*. **Resolución:** Resolución espacial para datos raster; **Densidad:** densidad de medición de datos de punto.

Conductividad eléctrica aparente fue medida de 0-30 (CEa<sub>30</sub>) y 0-90 cm (CEa<sub>90</sub>) de profundidad, utilizando el equipo de medición móvil VERIS 3100® (VERIS Technologies, Salina, KS, USA).

Cuatro imágenes satelitales Landsat 5 TM fueron obtenidas de la página web del visualizador global del Departamento de Geología de los EEUU USGS (<http://earthexplorer.usgs.gov/>) para varias fechas (Tabla 2.1). Las imágenes fueron corregidas atmosféricamente y re proyectadas a sistemas de coordenadas planas UTM utilizando ENVI/IDL v4.7 (ITT Visual Information Solutions, 2009). Finalmente, el índice de vegetación diferencial normalizado (NDVI) fue calculado para cada imagen (Boettinger *et al.*, 2008; Thompson *et al.*, 2012).

Datos de mapas de rendimiento de cinco cosechas (dos de trigo y tres de soja) fueron recolectados y almacenados mediante sistema de monitor de rendimiento y posteriormente leídos utilizando SMS v3.0™ (AgLeader Technology, Inc.). Los datos de rendimiento fueron filtrados para eliminar los datos atípicos relacionados con rendimientos menores a 0.1 y 0.5 t ha<sup>-1</sup> y mayores a 7 y 10 ton ha<sup>-1</sup> para soja y trigo, respectivamente.

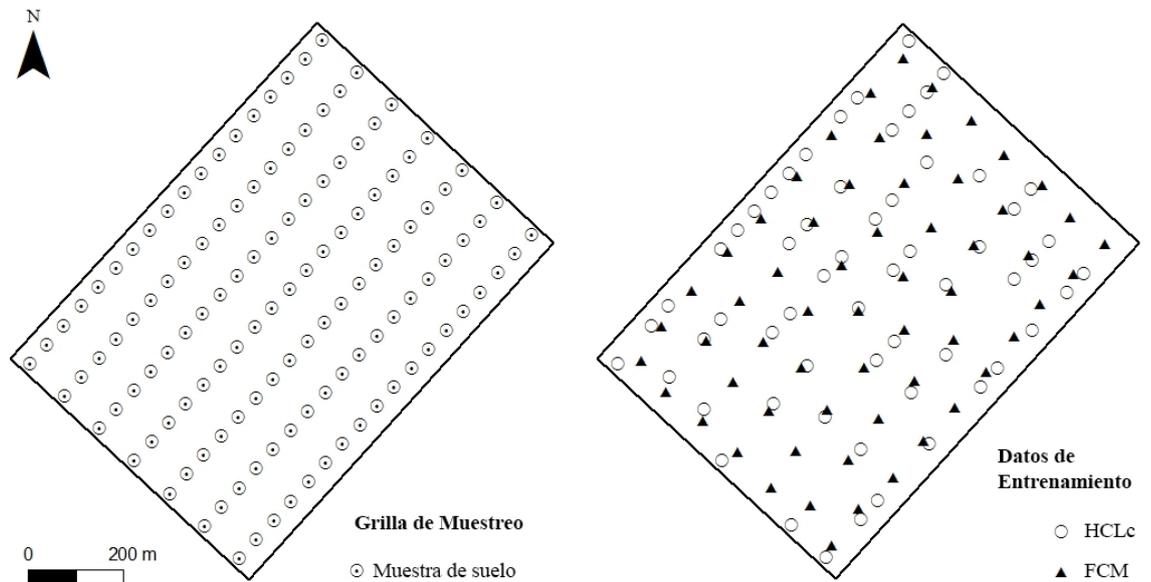
### *2.3. Desarrollo del modelo de predicción*

El modelo de predicción de MO y As fue desarrollado siguiendo un procedimiento que comprendió cinco pasos:

#### *2.3.1. Paso 1: Muestreo de suelos y análisis de laboratorio*

El muestreo de suelos fue realizado en el 2011, basado en un esquema sistemático de grilla con un tamaño de 50 x 100 m (126 muestras de suelo) (Figura 2.2). Este tamaño de grilla ha sido comúnmente utilizado por cartógrafos de suelos y agricultores en muestreos intensivos para la evaluación de la fertilidad del suelo a escala de lote. La profundidad de muestreo fue de 0 - 20 cm. Todos los análisis de suelos fueron realizados en los laboratorios de suelos de la Estación Experimental Agropecuaria Balcarce, Argentina.

Las muestras de suelos fueron analizadas para determinar MO a través de dicromato oxidación (Walkley y Black, 1934) y para determinar el contenido de As a través del método de la pipeta de Robinson (Gee y Bauder, 1986; Robinson, 1922).



**Figura 2.2** Distribución espacial de esquema de muestreo en grilla de tamaño 50 x 100 m (*Derecha*) y del esquema de muestreo generado por modelos con el conjunto de datos de entrenamiento (*Izquierda*)

Las 126 muestras fueron divididas aleatoriamente en dos grupos: 86 muestras como grupo de calibración del modelo RF (datos de calibración para ejecutar el paso 3) y 40 muestras de validación del modelo RF (datos de validación para ejecutar el paso 4). Los datos de calibración fueron a su vez divididos en dos subgrupos: 60 muestras de suelo como datos de entrenamiento del modelo RF y 26 muestras de suelo como datos de prueba del modelo RF (Figura 2.3). Los datos de validación fueron adecuadamente distribuidos por toda el área de estudio, para poder validar el modelo RF, para todas las circunstancias del suelo.

<b>División del Conjunto de Datos</b>		
<b>Muestras totales</b> 126	<b>Datos de Calibración</b> 86	<b>Datos de Entrenamiento</b> 60
		<b>Datos de prueba</b> 26
	<b>Datos de Validación</b> 40	

**Figura 2.3** División del conjunto completo de datos en datos de calibración y sus respectivas subdivisiones (datos de entrenamiento y prueba) y datos de validación

### 2.3.2. Paso 2: Análisis de la estructura espacial de los predictores

Variogramas experimentales fueron calculados y modelados para describir la estructura espacial de cada uno de las variables predictoras medidas (elevación, CEa y rendimiento). Los Variogramas experimentales fueron calculados utilizando la siguiente ecuación (Diggle y Ribeiro, 2007):

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{\alpha=1}^{N(\mathbf{h})} [z(\mathbf{U}_{\alpha}) - z(\mathbf{U}_{\alpha} + \mathbf{h})]^2 \quad (1)$$

donde  $\hat{\gamma}(\mathbf{h})$  es el valor de semivarianza a un intervalo de distancia  $h$ ,  $z(\mathbf{U}_{\alpha})$  es el valor de la muestral a partir de todos los puntos de  $\mathbf{U}_{\alpha}$  y  $N(\mathbf{h})$  es el número de pares de datos dentro de una clase dada de distancia y dirección.

El variograma experimental muestra la reducción en correlación espacial entre dos puntos cuando la distancia entre estos se incrementa (Peralta et al., 2013). Después de verificar los supuestos geoestadísticos, el modelo de variograma experimental

seleccionado y ajustado fue utilizado para interpolar elevación, CEa y rendimiento del cultivo a través de kriging ordinario. La interpolación geoestadística fue realizada utilizando el paquete “geoR” en el software estadístico R v3.1.1 (R Development Core Team, 2013). Finalmente, todos los archivos de salida se configuraron como raster (grilla regular) de 10 x 10 m, debido a que este tamaño refleja la escala de variabilidad asociada con escala de lote (Corwin y Lesch, 2005; Kitchen *et al.*, 2003).

### 2.3.3. Paso 3: Esquemas de muestreo de suelos basados en modelos (MBM)

Los algoritmos HCLc (Minasny y McBratney, 2006) y FCM (De Gruijter *et al.*, 2010) fueron ejecutados para seleccionar las 60 muestras de los datos de entrenamiento. Esta cantidad de muestras genera una apropiada densidad para el área de estudio (densidad de 0.88 muestras por hectárea). Varios estudios han demostrado que una densidad de muestreo menor a una muestra por hectárea es capaz de capturar la mayoría de la variación de los predictores a escala de lote (Debaene *et al.*, 2014; Schmidt *et al.*, 2014; Smith y Halvorson, 2011).

La ejecución del algoritmo HCLc es un procedimiento aleatorio estratificado que selecciona muestras basado en la distribución de todos los predictores (Roudier *et al.*, 2012). HCLc intenta encontrar un conjunto de valores a partir de los predictores que satisfaga un hipercubo latino, el cual busca que una muestra represente simultáneamente a todos los predictores (Brungard y Boettinger, 2010). El hipercubo latino es construido por muestreo aleatorio a partir de la distribución acumulada de cada uno de los predictores utilizando un enfoque de optimización compacta. Es decir, busca aquellas muestras dentro del conjunto de muestras totales que mejor representen la distribución de todos los predictores. Adicionalmente, también se focaliza en preservar la correlación entre los predictores en el conjunto de muestras seleccionadas (Schmidt *et al.*, 2014). HCLc en este estudio fue ejecutado utilizando el paquete “clhs” en R v3.1.1 (R Development Core Team, 2013).

La ejecución del algoritmo FCM se basa en el procedimiento de agrupamiento o clúster c-medias (De Gruijter *et al.*, 2010). FCM elabora particiones de muestreo a partir de todo

el conjunto de muestras y calcula la probabilidad de que cada muestra seleccionada pertenezca a un grupo o clúster. Lo que se espera es que las muestras que son seleccionadas para el mismo clúster tengan características similares y que las muestras de diferentes clúster tengan características diferentes. En la selección de muestras, las ubicaciones son el equivalente a los centróides de los clúster fuzzy c-medias logrados a partir de la discretización de la distribución espacial de todos los predictores. El número de muestras que se requieren es igual al número de clúster que se generan, que para este caso son 60 muestras para los datos de entrenamiento. FCM fue ejecutado utilizando el paquete “e1071” en R v3.1.1 (R Development Core Team, 2013).

#### 2.3.4. Paso 4: Calibración del modelo

El modelo RF fue calibrado utilizando los datos de entrenamiento (60 muestras) y prueba (26 muestras) (Breiman, 2001). RF utiliza numerosos árboles de regresión llamados dentro del algoritmo como  $n_{trees}$ . Un muestreo *bootstrap* diferente, el cual utiliza aproximadamente el 70% de los datos disponibles, es realizado cada vez que se ejecuta cada uno de los árboles de regresión. El 30% de datos remanentes referidos como *out-of-bag* (OOB) son utilizados para determinar el error de predicción de cada árbol de regresión (Cutler *et al.*, 2007). Cada vez que se ejecuta un árbol de regresión un subconjunto de predictores es seleccionado. A este subconjunto se le llama *mtry*. A su vez, los predictores dentro del *mtry* que permitan la mejor partición binaria dentro de cada árbol de regresión son seleccionados. Los  $n_{trees}$  y el *mtry* óptimos para la ejecución de RF deben ser observados y seleccionados por el usuario.

En este estudio, el óptimo *mtry* fue definido por iteración de valores *mtry* a partir de 1 y hasta la cantidad total de predictores. Para cada *mtry*, los valores de  $n_{trees}$  se fueron incrementando de 200 a 2000 en intervalos de 200  $n_{trees}$ . Los valores óptimos de *mtry* y  $n_{trees}$  fueron seleccionados utilizando el valor de los errores *OOB* más bajos (Heung *et al.*, 2014; Rad *et al.*, 2014). Adicionalmente, RF también estima la importancia de los predictores por permutación de los valores de cada predictor en la muestra *OOB* y reclasificando estas por permutación del predictor. Predictores que en su ausencia,

incrementen comparativamente el error *OOB*, son los más importantes (Genuer *et al.*, 2010).

El modelo RF fue ejecutado 20 veces con los 18 predictores y MO y As como variables de respuesta. La media de importancia de predictores fue calculada a partir de las 20 ejecuciones del algoritmo. Finalmente, los predictores fueron ordenados de acuerdo a la medición de su importancia.

### 2.3.5. Paso 5: Validación del modelo

El modelo RF fue validado utilizando los datos de validación (40 muestras). La validación con datos independientes es mejor que usar únicamente el error *OOB* (Rad *et al.*, 2014).

La precisión de predicción del modelo RF fue evaluada utilizando la raíz media del cuadrado del error (RMSE) de la predicción de las propiedades del suelo. RMSE es definida como:

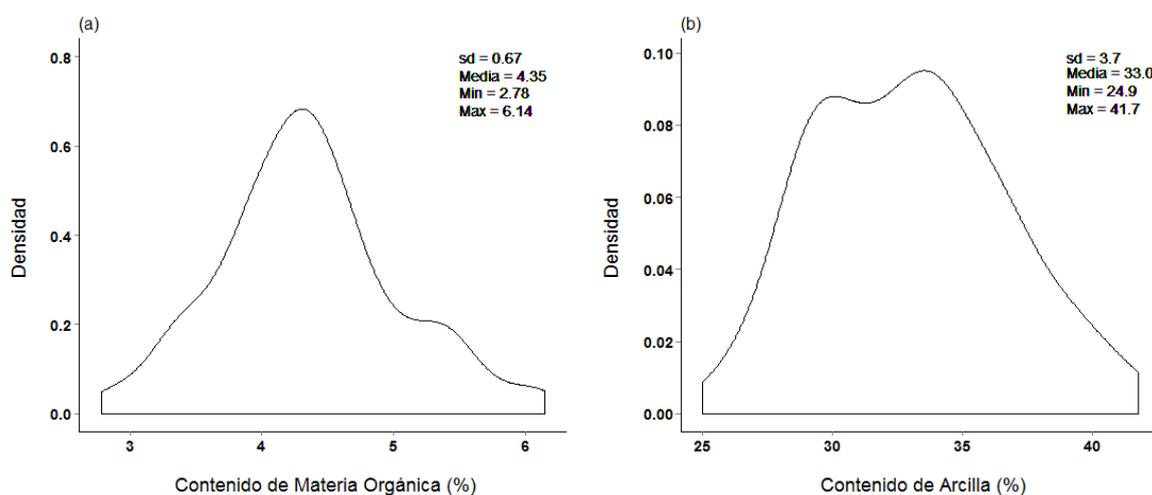
$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^N \{(X_i) - (Y_i)\}^2} \quad (2)$$

donde  $(X_i)$  son los valores observados y  $(Y_i)$  son los valores predichos. RMSE refleja la precisión de la predicción con valores de RMSE indicando el modelo que mejor ajustó. El correspondiente coeficiente de determinación ( $R^2$ ) también fue calculado (Willmott *et al.*, 1985).

Adicionalmente, las predicciones espaciales de cada modelo RF fueron analizadas para cada una de las propiedades del suelo predicha y visualmente comparadas con interpolación geoestadística a partir de conjunto de datos total (126 muestras de suelo).

## 3. Resultados y discusión

### 3.1. Propiedades del suelo



**Figura 2.4** Gráficos de densidad de (a) contenido de Materia Orgánica (MO) y (b) contenido de Arcillas (As) en el conjunto de datos completo

Las propiedades del suelo evaluadas muestran una amplia variación (Figura 2.4). Peralta et al. (2013) y Paggi et al. (2012) también encontraron similares variaciones de propiedades del suelo a escala de lote, las cuales pueden llegar a ser explicadas por el complejo patrón de suelo que existe en la pampa argentina (Borrelli et al., 2008; Peinemann et al., 2005). La distribución espacial de los suelos de la pampa es el resultado de un proceso que involucra varios periodos geológicos que determinaron la evolución del paisaje (Amiotti *et al.*, 2001). Normalmente, a escala de lote los suelos de la pampa tienen copresencia de polipedones de antigua y reciente formación. Los polipedones de antigua formación son remanentes de pretéritos ciclos de erosión, mientras que los polipedones de reciente formación son remanentes de procesos erosivos parciales y totales de polipedones de antigua formación, seguidos de pulsos sucesivos de deposiciones eólicas (Blanco y Stoops, 2007).

**Tabla 2.2** Parámetros geoestadísticos de modelos ajustados de semivariogramas experimentales para los predictores medidos

Predictor	Modelo	$C_0$ <sup>(1)</sup>	$C_0 + C$ <sup>(2)</sup>	$a$ (m) <sup>(3)</sup>
CEa <sub>90</sub>	Exponencial	3.49	31.12	10.40
CEa <sub>30</sub>	Exponencial	3.65	22.65	46.78
Rendimiento Trigo 2002	Exponencial	0.08	0.77	28.60
Rendimiento Trigo 2003	Exponencial	0.05	0.42	19.70
Rendimiento Soja 2004	Exponencial	0.01	0.09	15.60
Rendimiento Soja 2005	Exponencial	0.02	0.21	23.39
Rendimiento Soja 2008	Exponencial	0.01	0.06	20.80
Elevación	Exponencial	0.13	2.20	155.13

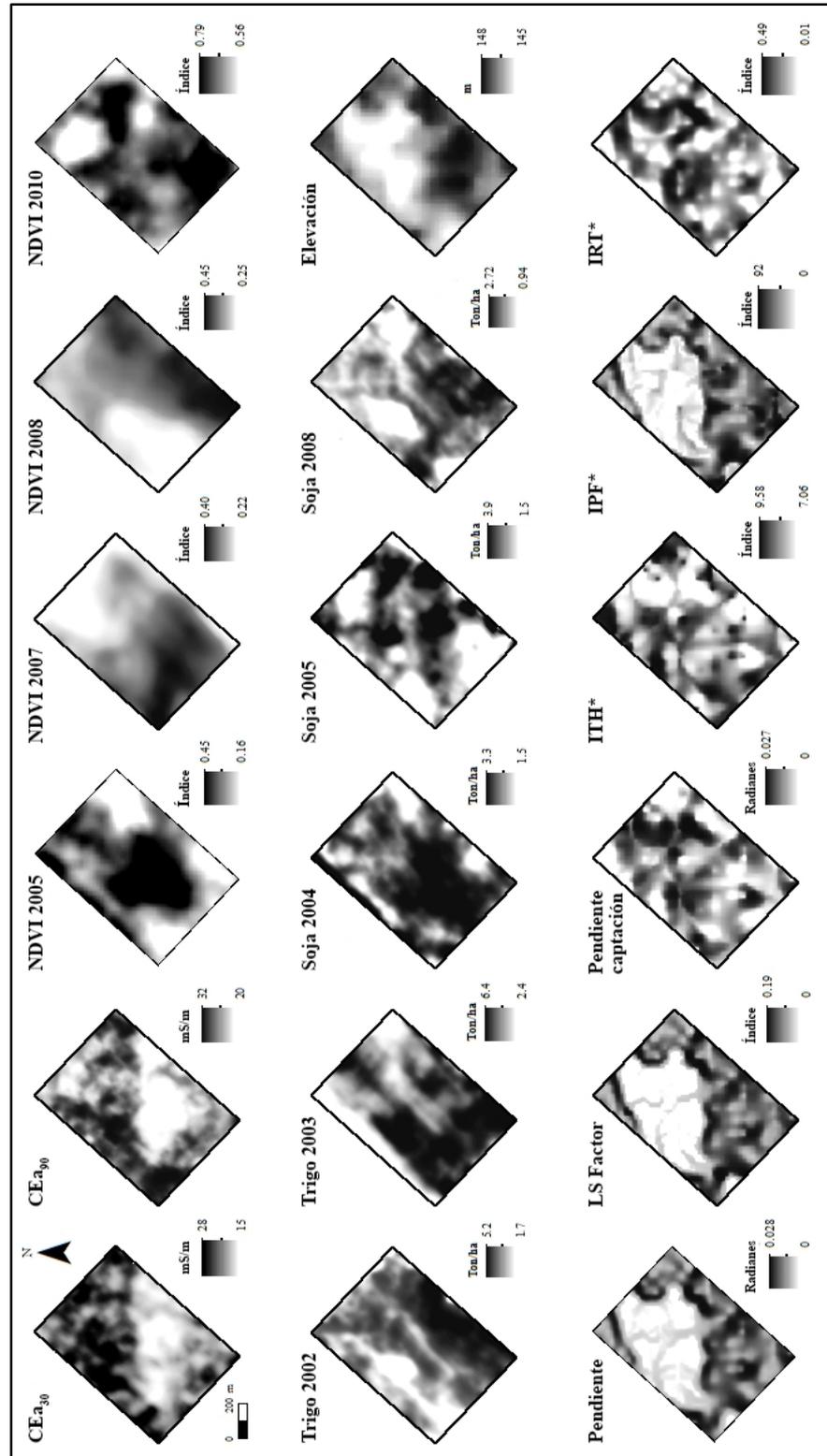
<sup>(1)</sup> Nugget varianza; <sup>(2)</sup> Silla; <sup>(3)</sup> Rango.

### 3.2. Análisis de la estructura espacial de los predictores medidos

La Figura 2.5 muestra la distribución espacial de cada predictor. Modelos teóricos exponenciales de semivariogramas tuvieron el mejor ajuste para elevación, CEa y mapas de rendimiento (Tabla 2.2). Semivariogramas exponenciales son los más adecuados cuando los predictores tienen una distribución desigual a cortas distancias (Oliver, 2010). En general, los semivariogramas exponenciales indicaron una fuerte dependencia espacial de todos los predictores (relación nugget silla-1 < 0.25) (Cambardella et al., 1994).

La semivarianza ( $C_0 + C$ ) fue alta para CEa y baja para mapas de rendimiento. CEa<sub>30</sub> tuvo mayor rango ( $a$ ) que CEa<sub>90</sub>. Por otro lado, rendimiento del trigo en 2002 tuvo mayor rango que el rendimiento de trigo en 2003, mientras que rendimiento de soja 2005 tuvo menor rango que rendimiento de soja 2004 y 2008. Similares resultados de análisis de estructura espacial fueron reportados por Basso *et al.* (2007). Como se esperaba, el rango ( $a$ ) para elevación fue alto. Normalmente, la elevación es un predictor con una fuerte estructura espacial (Hengl y Reuter, 2008).

**Figura 2.5** Distribución espacial de cada uno de los predictores dentro del área de estudio



El análisis de la estructura espacial de elevación, CEa y mapas de rendimiento confirma la presencia de un patrón complejo de suelos en el área de estudio. Valores de

CEa fueron bajos en las zonas con elevación alta, mientras que los mapas de rendimiento tuvieron una baja estabilidad temporal. Normalmente, los suelos de la pampa tienen una variabilidad que responde a las ondulaciones de la paleosuperficie petrocálcica, la que se refleja en la distribución de las taxas (Pazos y Mestelan, 2002). La variabilidad espacial de la profundidad efectiva podría afectar significativamente los patrones espaciales de otras propiedades del suelo a escala de lote. Consecuentemente, variables importantes para definir la productividad de los cultivos tales como humedad del suelo, generación de escorrentía, flujo de agua superficial y subsuperficial, capacidad de almacenamiento de agua y evapotranspiración podrían ser afectados, lo cual, a su vez afectaría la variabilidad espacial del rendimiento del cultivo y la CEa. Por otro lado, Bedogni (1996) reportó un complejo patrón de distribuciones de suelos a escala de lote, los que se caracterizan por tener en las zonas altas, poca profundidad y bajo contenido de MO y As, mientras que en las zonas bajas, ocurre lo contrario. En este estudio, variaciones de CEa en zonas altas y bajas dentro del área de estudio corroboraron las relaciones entre elevación y distribución espacial de propiedades del suelo, y se complementaron con la medición de CEa.

### *3.3. Análisis exploratorio de los predictores*

Excepto para IPF, la media y la mediana fueron similares para todos los predictores (Tabla 2.3). Las diferencias notables de media y mediana en IPF son explicadas ya que este índice topográfico es de tipo acumulativo (Hengl y Reuter, 2008; Pike et al., 2006).

Las diferencias entre las medias de NDVI 2010 y NDVI 2005 fue 0.53 (Tabla 2.3). Esta amplia variación puede deberse a tres factores (*i*) a la interferencia del suelo sobre la cálculo de NDVI cuando la canópi de los cultivos no estaban totalmente cerradas; Esto es principalmente acentuado en imágenes Landsat debido a la resolución espacial de 30 m (Dang *et al.*, 2011); (*ii*) al efecto temporal de las condiciones micrometeorológicas sobre la vegetación dentro del perímetro del área de estudio (Guo *et al.*, 2012); y (*iii*) el complejo patrón de distribución de suelos que existe dentro del área de estudio (Boettinger *et al.*, 2008).

**Tabla 2.4** Resumen de parámetros de estadística descriptiva de los predictores

Predictores	Unidades	Meida	DS <sup>(1)</sup>	Mediana	Min <sup>(2)</sup>	Max <sup>(3)</sup>	Coeficientes	
							Asimetría <sup>(4)</sup>	Curtosis <sup>(5)</sup>
CEa <sub>30</sub>	mS m <sup>-1</sup>	22.56	2.66	22.88	15.81	28.04	-0.40	-0.68
CEa <sub>90</sub>	mS m <sup>-1</sup>	26.16	2.30	26.19	20.48	32.18	-0.04	-0.39
NDVI 2005	PA	0.21	0.05	0.21	0.08	0.34	0.02	-0.45
NDVI 2007	PA	0.31	0.03	0.31	0.22	0.36	-0.58	-0.20
NDVI 2008	PA	0.42	0.05	0.42	0.31	0.55	-0.06	-0.70
NDVI 2010	PA	0.74	0.05	0.76	0.69	0.79	-0.75	0.66
Rendimiento Trigo 2002	Ton ha <sup>-1</sup>	3.49	0.50	3.43	1.71	5.15	0.24	0.36
Rendimiento Trigo 2003	Ton ha <sup>-1</sup>	4.91	0.64	5.03	2.66	6.37	-0.86	0.52
Rendimiento Soja 2004	Ton ha <sup>-1</sup>	2.44	0.22	2.43	1.63	3.17	-0.16	0.06
Rendimiento Soja 2005	Ton ha <sup>-1</sup>	2.38	0.19	2.40	1.64	3.25	-0.45	0.44
Rendimiento Soja 2008	Ton ha <sup>-1</sup>	2.30	0.24	2.32	0.10	2.73	-1.87	9.03
Elevación	m	146.95	1.10	147.08	145.10	148.38	-0.27	-0.66
Pendiente	%	0.01	0.01	0.01	0.00	0.04	0.73	-0.05
LS Factor	PA	0.04	0.04	0.03	0.00	0.27	1.42	2.88
Pendiente de captación	%	0.01	0.00	0.01	0.00	0.03	0.80	1.02
ITH	PA	8.11	0.35	8.11	7.03	10.01	0.48	1.85
IPF	PA	10.38	14.81	5.93	0.04	159.46	3.25	16.41
IRT	PA	0.16	0.08	0.14	0.02	0.50	0.71	0.48

<sup>(1)</sup> Desviación Estandar; <sup>(2)</sup> Valor Mínimo; <sup>(3)</sup> Valor Máximo; <sup>(4)</sup> Coeficiente de asimetría; <sup>(5)</sup> Coeficiente de Curtosis. CEa<sub>30</sub> and CEa<sub>90</sub> (Conductividad Eléctrica Aparente a 0-30 y 0-90 cm de profundidad, respectivamente); NDVI (Índices de Vegetación Diferencial Normalizado a partir de las escenas Landsat 5 TM); Yield (Monitor de rendimiento); LS Factor (Factor de pérdida de suelo); ITH (Índice Topográfico de Humedad); IPF (Índice de potencial de flujo) and IRT (Índice de rugosidad del terreno). PA (Predictor Adimensional).

La diferencia en la media de rendimiento de trigo entre 2002 y 2003 fue de 1.42 ton ha<sup>-1</sup>. En contraste, las medias del rendimiento de soja 2004, 2005 y 2008 fueron similares. Al igual que NDVI, estos resultados sugieren que la pequeña diferencia en la media de los rendimientos respondieron principalmente a la combinación espacio-temporal de condiciones micrometeorológicas y el complejo patrón de distribución de suelos del área de estudio (Anderson-Cook *et al.*, 2002; Blackmore, 2000).

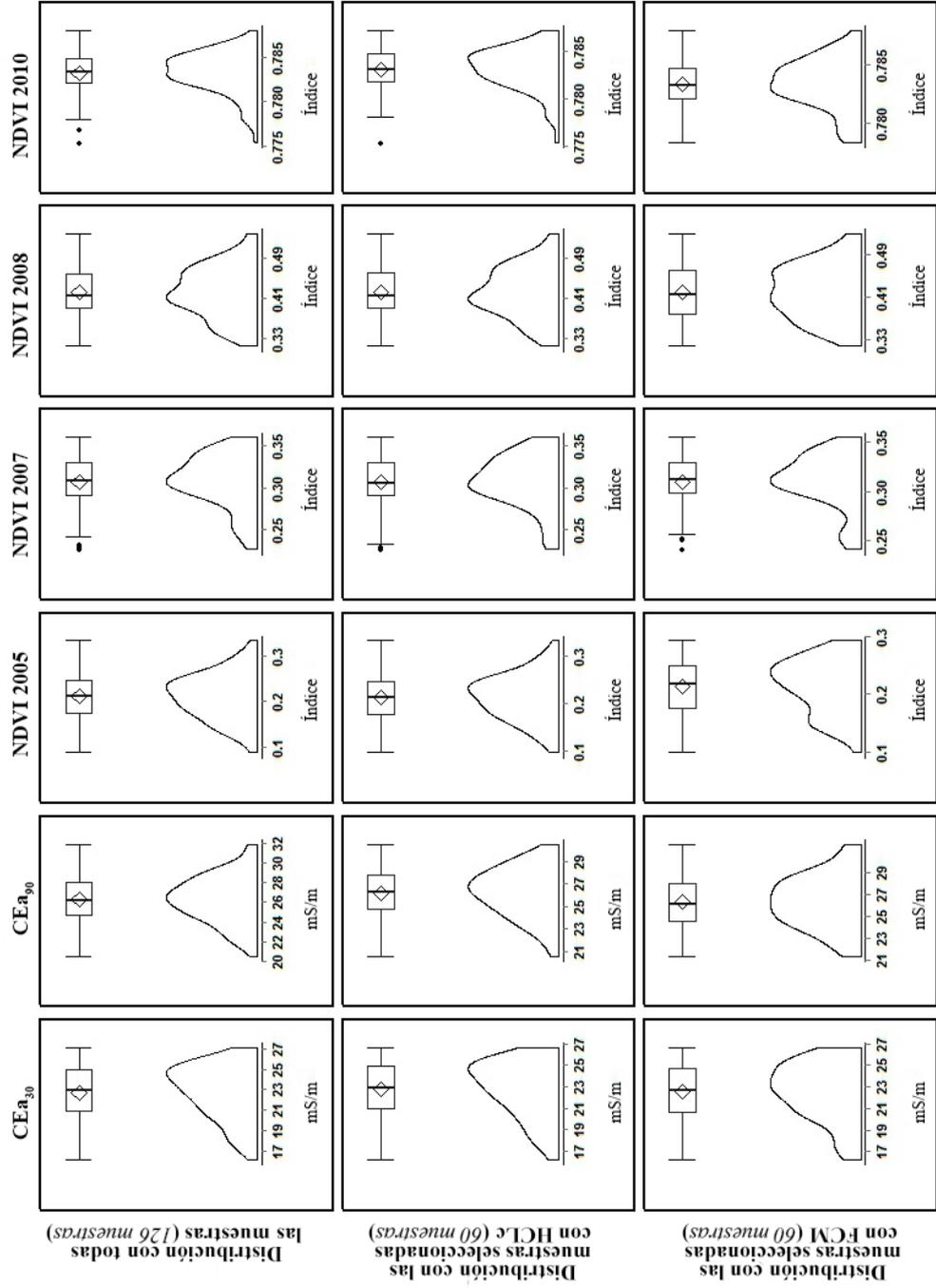
Las diferencias en elevación y pendiente fueron 3.28 m y 0.04 radianes, respectivamente (Tabla 3). En general, poca diferencia de elevación y pendiente podrían afectar la capacidad de los índices de topografía para predecir propiedades del suelo a escala de lote (Florinsky, 2012).

### 3.4. Comparación de esquemas de muestreo de suelos

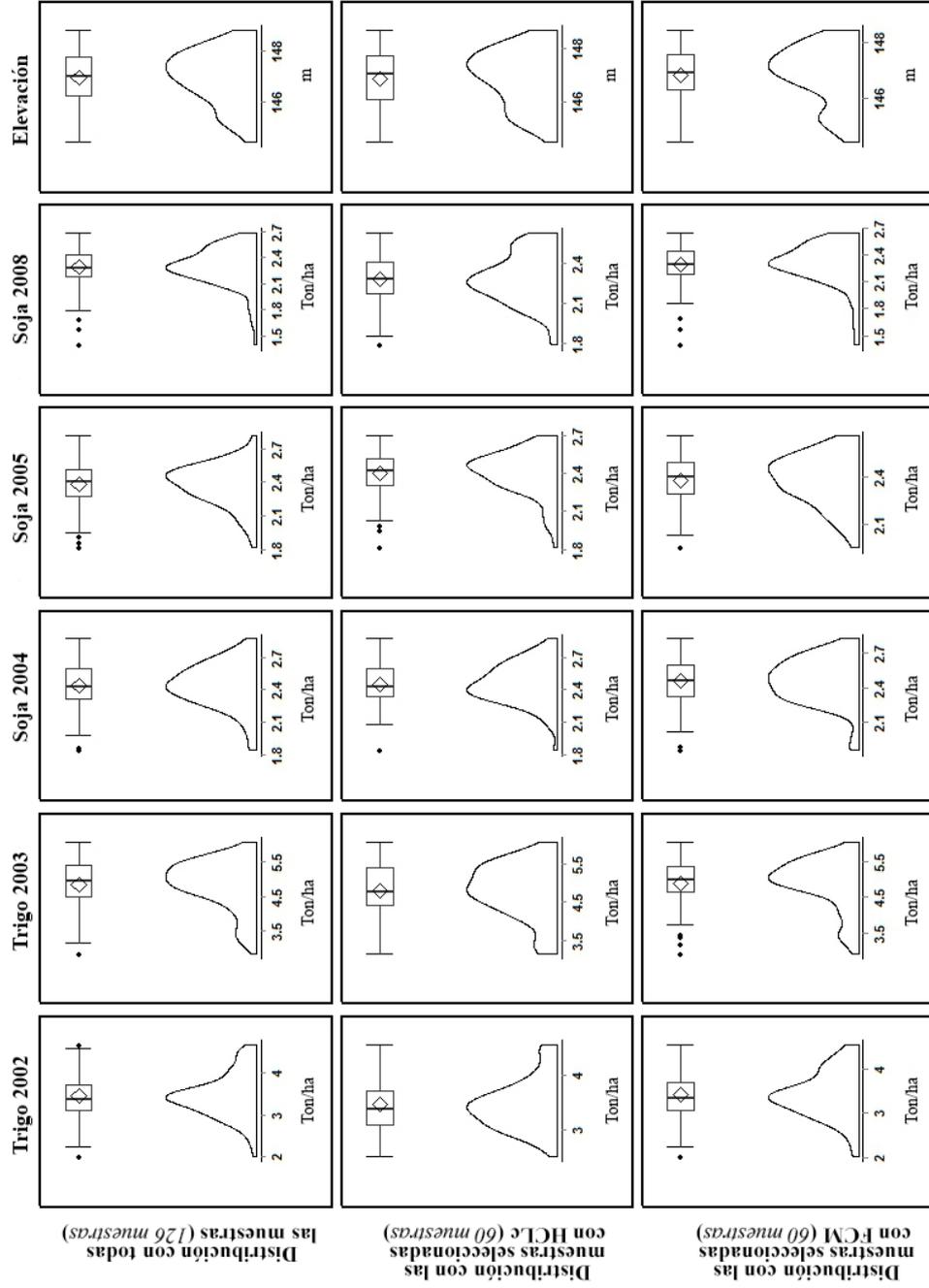
La figuras 2.6, 2.7 y 2.8 muestran las comparaciones de las curvas de distribución de frecuencia de todos los predictores en cada uno de los esquemas de muestreo de suelos utilizados, acompañado de una gráfica de caja complementaria. El tamaño de muestras usadas para cada esquema de muestreo representa el 47% (60 muestras) de todas las muestras obtenidas a partir del muestreo sistemático en grilla de tamaño 50 x 100 m (126 muestras). En general, las curvas de distribución muestran que HCLc logra obtener un conjunto de muestras que representan cerradamente la distribución de frecuencia de todos los predictores, mientras que las curvas de distribución de FCM muestran una aceptable representación.

La media, mediana y desviación estándar fueron similares entre las distribuciones de las muestras seleccionadas por HCLc y las distribuciones totales de los predictores. Excepto para IPF, rendimiento de trigo 2003 y rendimiento de soja 2008 HCLc presentó la mejor curva de ajuste en términos de asimetría y curtosis. En este trabajo se encontró que HCLc fue eficiente para capturar la variabilidad de la mayoría de los predictores de propiedades del suelo. Estos resultados coinciden con los presentados anteriormente por estudios relacionados (Brungard y Boettinger, 2010; Kidd *et al.*, 2015).

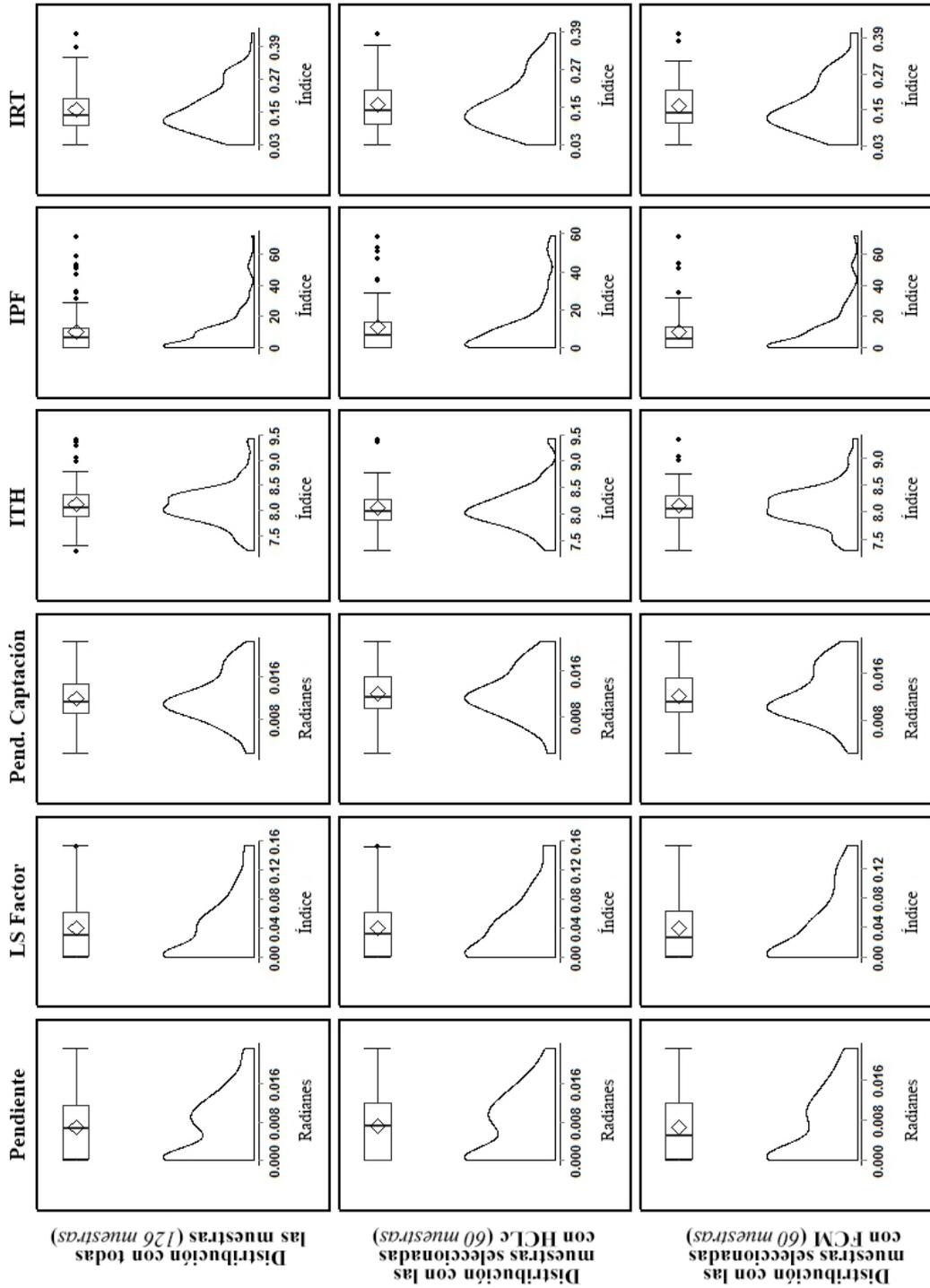
**Figura 2.6a** Comparación de la distribución de las muestras obtenidas a partir de Hipercono Latino Condicionado (HCLc) (60 muestras) y Fuzzy C-Medias (FCM) (60 muestras) con respecto a la distribución total de todas las muestras (126 muestras), para todos los predictores



**Figura 2.6b** Comparación de la distribución de las muestras obtenidas a partir de Hiperucubo Latino Condicionado (HCLc) (60 muestras) y Fuzzy C-Medias (FCM) (60 muestras) con respecto a la distribución total de todas las muestras (126 muestras), para todos los predictores



**Figura 2.6c** Comparación de la distribución de las muestras obtenidas a partir de Hipercubo Latino Condicionado (HCLc) (60 muestras) y Fuzzy C-Medias (FCM) (60 muestras) con respecto a la distribución total de todas las muestras (126 muestras), para todos los predictores



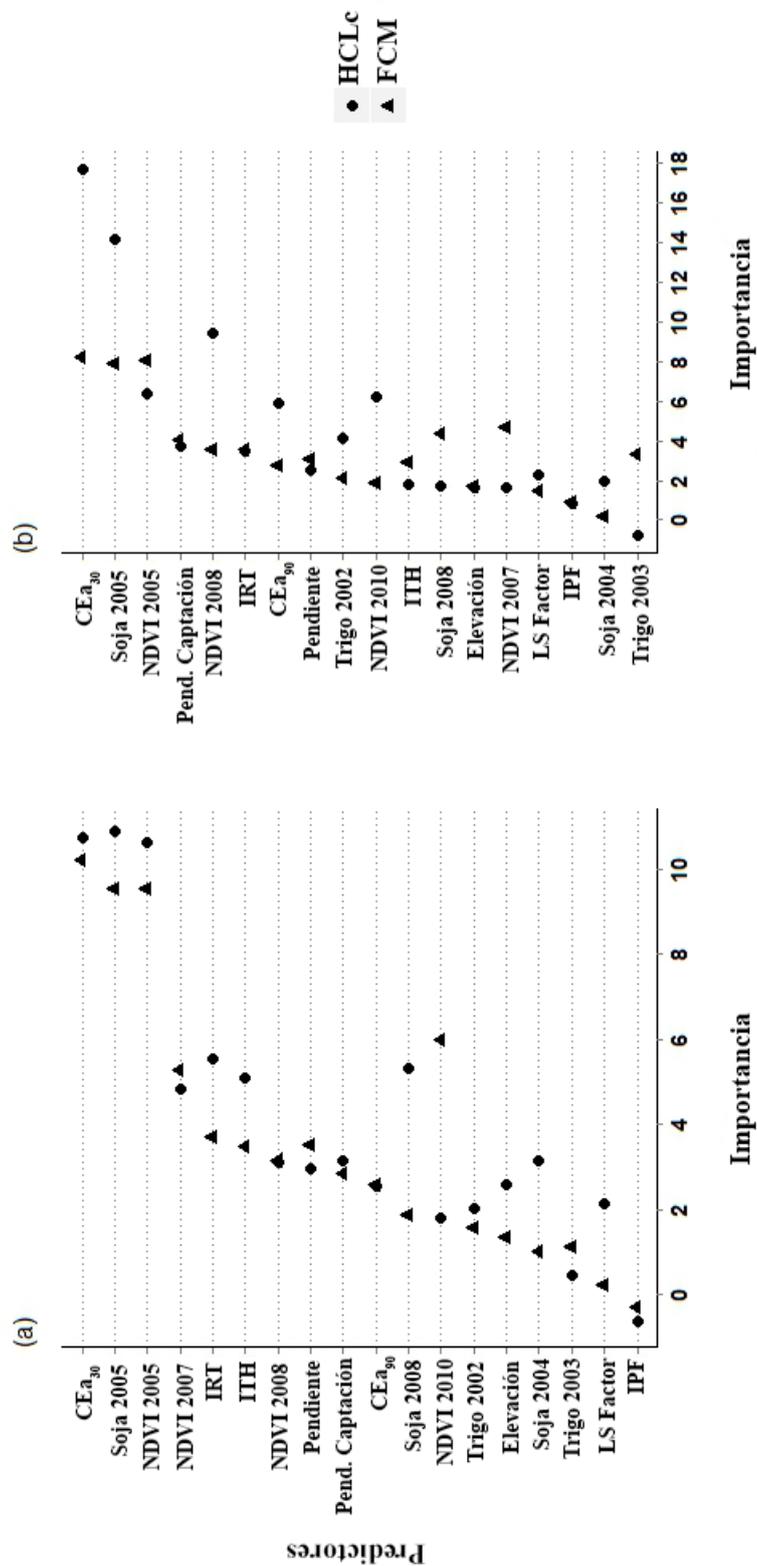
La densidad de la curva, media, mediana y desviación estándar de FCM fueron similares con la distribución total de algunos predictores (Figura 2.6, 2.7 y 2.8). Excepto para elevación, NDVI 2005, NDVI 2007 y NDVI 2010, FCM presento una razonable curva de ajuste en términos de asimetría y curtosis. Es posible que el complejo patrón de suelos haya afectado la eficiencia de FCM (Schmidt *et al.*, 2014). En zonas con complejos patrones de variabilidad espacial de suelos, las áreas clúster delimitadas por FCM pueden tener una gran variabilidad de la distribución de los predictores por unidad de área (Kidd *et al.*, 2015). Las áreas de los clúster se localizaron en zonas con topografía, CEa y NDVI diferentes. En ese contexto, es posible que las subregiones clúster delimitadas dentro del área de estudio requieran mayor o menor densidad de muestreo. Estos resultados sugieren que FCM no podría optimizar el muestreo teniendo en cuenta la distribución espacial de la totalidad de los predictores. Sin embargo, FCM podría ser muy útil para mejorar el muestreo proporcional en partes específicas de la superficie del terreno, en las cuales los predictores tengan comportamientos homogéneos. Esto podría ser importantes para optimizar el esquema de muestreo.

### 3.5. Importancia de los predictores y performance del modelo Random Forest

Modelos RF para predecir MO y As fueron ejecutados con el parámetro *mtry* por defecto y 500 *ntrees*. La Figura 2.7 muestra los predictores más importantes para MO (Figura 2.7a) y As (Figura 2.7b) a partir de ejecutar el modelo RF con los datos de entrenamiento y validación, y 18 predictores. En general, CEa<sub>30</sub>, rendimiento de soja 2005 y NDVI 2005 fueron los predictores más importantes de MO y As.

La importancia de la mayoría de los predictores de MO fueron similares tanto para HCLc, como FCM (Figura 2.7a). Por otra parte, el nivel de importancia de CEa<sub>30</sub>, rendimiento de soja 2005 y NDVI 2005 para predecir As fue mayor con HCLc (Figura 2.7b), lo cual podría incrementar la precisión del modelo RF. Los tres predictores menos importantes de MO fueron rendimiento de trigo 2003, LS Factor e IPF, mientras que los tres predictores menos importantes para As fueron rendimiento de trigo 2003, rendimiento de soja 2004 e IPF.

**Figura 2.7** Importancia de los predictores a partir de las clasificaciones realizadas con Random Forest con ambos esquemas de muestreo Hipercubo Latino Condicionado (HCLc) y Fuzzy C-Medias (FCM) usadas para predecir el (a) contenido de Materia Orgánica (MO) y (b) el contenido de Arcilla a escala de lote. Valores altos indican predictores más importantes



En general, los principales factores que influyen el comportamiento espacial de la medición de la CEa son: (i) la conductividad eléctrica de la solución del suelo, (ii) el contenido volumétrico de las propiedades del suelo, (iii) la conductividad eléctrica de las partículas del suelo, y (iv) el contenido volumétrico de agua en el suelo (Rhoades *et al.*, 1989a; Rhoades *et al.*, 1976). Los primeros tres factores están principalmente influenciados por la salinidad del suelo (Adam *et al.*, 2012; Corwin *et al.*, 2006; Li *et al.*, 2013; Rhoades, 1996), contenido de As (De Benedetto *et al.*, 2012; Lesch, 2005; Piikki *et al.*, 2013; Triantafilis y Lesch, 2005) y la capacidad de intercambio catiónico (Friedman, 2005), mientras que la humedad volumétrica del suelo está principalmente determinada por la cantidad de poros con agua retenida entre partículas del suelo. Generalmente, suelos arcillosos, los cuales tienen mayor densidad de poros continuos que almacenan agua, tienden a conducir la electricidad más fácilmente que los suelos arenosos (Peralta *et al.*, 2013).

El patrón espacio-temporal general de NDVI y de mapas de rendimiento no pudo ser identificado. Numerosos estudios han determinado que NDVI podría ser un importante predictor de la variabilidad espacial de las propiedades del suelo (Boettinger *et al.*, 2008; Guo *et al.*, 2012; Hartemink *et al.*, 2008; Takata *et al.*, 2007). Sin embargo, nuestros resultados sugieren que en áreas con patrones espaciales complejos de suelos a escala de lote, es necesario determinar el momento óptimo o lapso de tiempo para adquirir las imágenes Landsat. Al respecto, Dang *et al.* (2011) reportó que durante la antesis de soja y trigo es el momento más conveniente para adquirir imágenes satelitales Landsat. Futuros trabajos en la zona de estudio, los cuales tomen información de NDVI proveniente de imágenes satelitales Landsat como predictores de propiedades del suelo a escala de lote, deberán determinar el momento oportuno de adquisición de la imagen satelital Landsat.

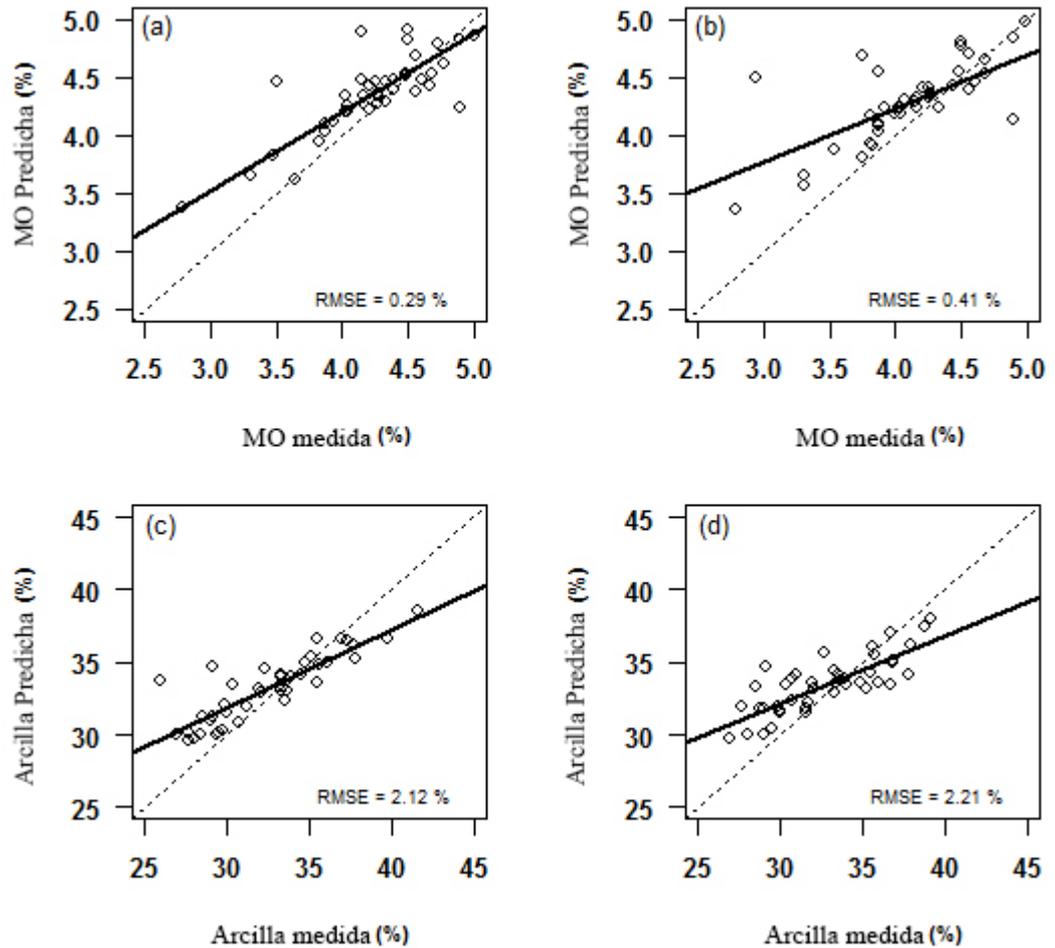
Los mapas de rendimiento presentaron una alta variabilidad espacial y una baja estabilidad temporal. Al respecto, Basso *et al.* (2007) determinó que antes de utilizar los mapas de rendimiento para predecir propiedades del suelo, es conveniente determinar la variabilidad espacial y la estabilidad temporal de al menos 5 mapas de rendimiento. Similarmente, Blackmore *et al.* (2003) estableció que los mapas de rendimiento son útiles,

siempre y cuando se logre determinar puntualmente cual es la causa de su variación espacio temporal. En el sudeste de la provincia de Buenos Aires, diferencias en la dinámica espacio temporal de los mapas de rendimiento puede llegar a ser un importante indicador para determinar y cuantificar el efecto de las condiciones micrometeorológicas sobre los patrones espaciales de suelos a escala de lote.

Nuestros resultados sugieren que los índices de topografía fueron los predictores menos importantes. En contraste, numerosos estudios ha encontrado que los índices de topografía son importantes predictores de las propiedades del suelo (Boettinger, 2010; Florinsky, 2012; Hengl y Reuter, 2008; Pei *et al.*, 2010). Estos resultados podrían deberse principalmente a las pequeñas variaciones en elevación y pendiente dentro del área de estudio (3 m y 0.028 radianes, respectivamente). Generalmente, la topografía tiene un efecto físico sobre la variabilidad espacial de las propiedades del suelo (Florinsky, 2012). Al respecto, Pei *et al.* (2010) encontró que el ITH no siempre es válido para predecir las propiedades del suelo, especialmente en áreas con poco relieve donde el movimiento de agua puede ser de tipo divergente. Wang *et al.* (2001b) encontró que el LS Factor no es un eficiente predictor de propiedades del suelo en áreas con pocas diferencias de pendiente. Similar a LS Factor, es posible que IPF e IRT puedan no ser eficientes predictores en condiciones del área de estudio. IPF está relacionado con la erosión, mientras que IRT está relacionado con la ondulación del terreno. El efecto de la erosión y ondulación sobre la variabilidad espacial de las propiedades del suelo depende puntualmente de las diferencias de elevación y pendiente dentro del área de estudio (Heung *et al.*, 2014; Moore *et al.*, 1991).

### 3.6. Predicción de propiedades del suelo

La figura 2.8 muestra una gráfica de dispersión de puntos de MO observada versus MO predicha a partir de HCLc (Figura 2.8a) y FCM (Figura 2.8b), y As observada versus As predicha a partir de HCLc (Figura 2.8c) y FCM (Figura 2.8d), para los datos de validación.



**Figura 2.8** Gráfico de dispersión de puntos entre contenido de Materia Orgánica (**MO**) predicha versus MO observada utilizando (a) el esquema de muestreo de suelos Hipercubo Latino Condicionado (HCLc) y (b) Fuzzy C-Medias; y contenido de Arcillas predicho versus observado utilizando (c) el esquema de muestreo de suelos Hipercubo Latino Condicionado (HCLc) y (d) Fuzzy C-Medias

A pesar de la compleja variación de suelos y de la inestabilidad espacio temporal de NDVI y los mapas de rendimiento, los modelos RF fueron eficientes en la predicción de propiedades del suelo evaluadas. El modelo RF con HCLc y FCM logró explicar el 65% ( $R^2=0.648$ ) y el 57% ( $R^2=0.571$ ) de la variabilidad de MO, y alrededor del 71% ( $R^2=0.714$ ) y 68% ( $R^2=0.686$ ) de la variabilidad de arcilla, respectivamente.

Una de las fortalezas de estos resultados es que los modelos RF fueron validados sobre un conjunto de datos independientes (datos de validación). Estos modelos RF, los cuales se generan a partir del uso de esquemas de muestreos de suelos basados en modelos e

información auxiliar del suelo, contribuyen a la cuantificación científica de las propiedades del suelo a escala de lote.

Los valores de RMSE indican que los modelos RF con HCLc son ligeramente mejores que los modelos RF con FCM para predecir las propiedades del suelo (Figura 2.10). Sin embargo, modelos RF con ambos esquemas de muestreo podrían sobreestimar los contenidos de MO y As en áreas con bajos contenidos de estas propiedades del suelo. Al respecto, es pertinente resaltar que los modelos RF nunca predijeron MO y As por debajo de 3.4 y 29%, respectivamente. Tesfa *et al.* (2009) reportaron inconsistencias similares en modelos RF para predecir profundidad del suelo a partir de índices de topografía y cobertura vegetal. La sobre estimación de los modelos RF puede deberse a la naturaleza discreta de los predictores en los árboles de regresión que subyacen el enfoque de predicción de los modelos RF (Heung *et al.*, 2014). La inhabilidad para predecir MO y As en zonas con bajo contenido de estas propiedades podría ser un obstáculo para la implementación de los modelos RF en condiciones de los suelos del sudeste de la provincia de Buenos Aires a escala de lote. Sin embargo, en general los modelos RF proveen una significativa mejora en la predicción de MO y As con respecto a las bases de datos existentes tales como GeoINTA. Además, los modelos RF se desarrollaron rápidamente a partir de información auxiliar, en contraste con los métodos convencionales de cartografía de suelos.

### 3.7. Predicción espacial

La Figura 2.9 muestra la predicción espacial de MO y As usando los modelos RF con HCLc (Figura 2.9a y 2.9c, respectivamente), con FCM (Figura 2.9b y 2.9e, respectivamente) y la interpolación geoestadística a partir de todo el conjunto de muestras (126 muestras) utilizando Kriging ordinario (Figuras 2.9c y 2.9f, respectivamente). En general, los patrones espaciales entre modelos RF e interpolación geoestadística fueron similares. Sin embargo, el análisis visual muestra que los modelos RF con HCLc fueron más eficientes para predecir la variabilidad espacial de las propiedades del suelo que los modelos RF con FCM. Resultados similares también fueron encontrados por Rad *et al.* (2014), quienes determinaron que la combinación de modelos RF e HCLc fueron

eficientes para mapear los grupos taxonómicos del suelo. Así mismo, Schmidt *et al.* (2014) concluyeron que la combinación óptima entre modelos RF y HCLc puede ser adoptada para predecir la variabilidad espacial de las propiedades del suelo a escala de lote.

#### 4. Conclusiones

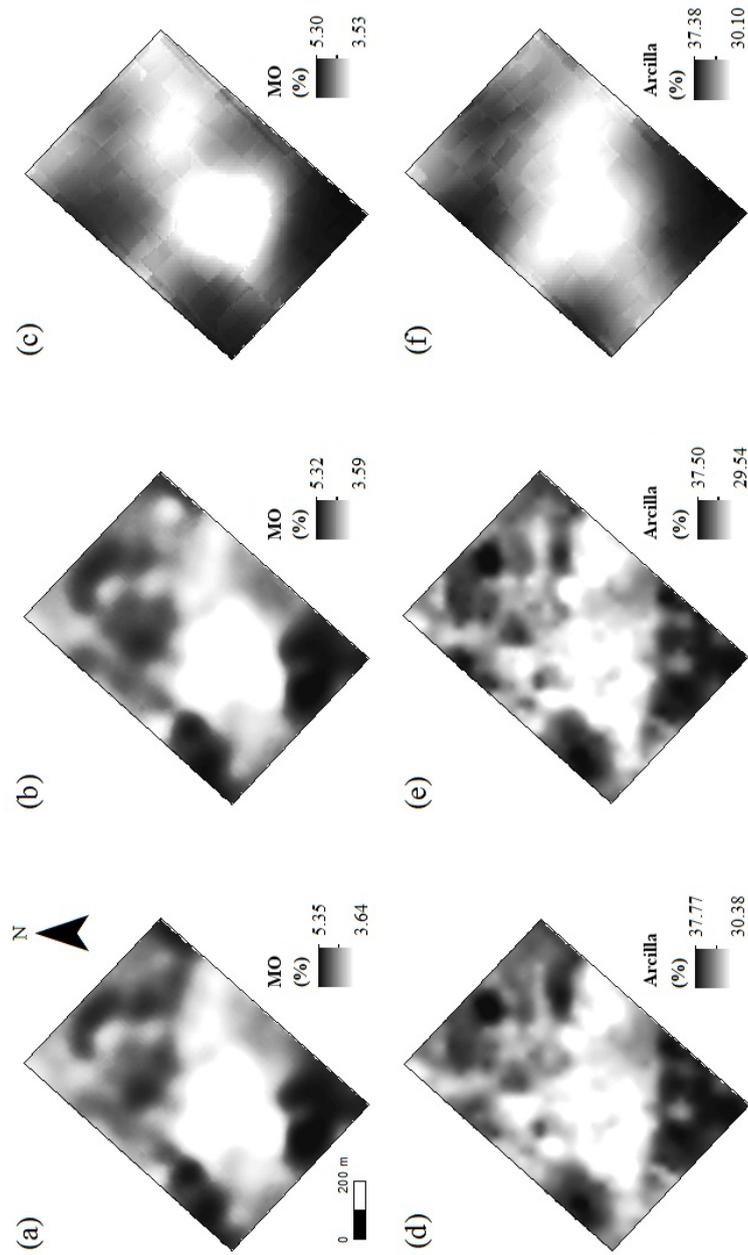
En el presente estudio, modelos RF para predecir MO y As a escala de lote fueron desarrollados a partir de dos esquemas de muestreos basados en modelos (HCLc y FCM) y predictores obtenidos a partir de información auxiliar de suelo fácilmente accesible. Este estudio valida la utilidad de los esquemas de muestreo basados en modelos como técnica de muestreo de suelos dentro del concepto de cartografía digital, en condiciones del sudeste de la provincia de Buenos Aires, y soportando el supuesto que tanto HCLc como FCM capturan eficiente y simultáneamente la distribución de frecuencia de todos los predictores. El análisis de importancia de los predictores muestra que CEa, NDVI y mapas de rendimiento son suficientes para predecir las propiedades del suelo evaluadas a escala de lote.

En general, estos resultados sugieren que (i) el esquema de muestreo de suelos tiene una gran influencia en la eficiencia de los modelos de predicción de propiedades del suelo; (ii) únicamente 60 muestras de suelo se requieren para integrar los múltiples predictores de propiedades del suelo; (iii) un conjunto de datos independiente es la forma más adecuada para validar un modelo RF.

Los modelos RF proporcionaron un significativo mejoramiento de información con respecto a GeoINTA. Esto es importante para la implementación de manejo sitio-específico de cultivos y para elaborar modelos hidrológicos. Sin embargo, la extrapolación de estos resultados en otras áreas dentro del sudeste de la provincia de Buenos Aires requiere ser evaluada. Los índices de topografía podrían ser importantes para predecir propiedades del suelo en otros sectores de esta región, donde las variaciones de elevación y pendientes sean mucho mayores a las que caracterizan al área de estudio

de este trabajo. Además, estudios adicionales son requeridos para evaluar nuevos predictores.

**Figura 2.9** Comparaciones de la predicción espacial del contenido de Materia Orgánica (MO) utilizando (a) kriging ordinario a partir del conjunto de datos completo, (b) datos predichos con el modelo Random Forest (RF) con Hipercono Latino Condicionado (HCLc) y RF con Fuzzy C-Medias (FCM). Comparación de a predicción espacial de contenido de arcillas utilizando (d) kriging ordinario a partir del conjunto de datos completo, (e) RF con HCLc y (f) RF con FCM)



## CAPITULO 3

### Desagregación de unidades cartográficas de suelo a partir de análisis de componentes principales espaciales y distancia euclidiana

#### 1. Introducción

En la actualidad, se estima que más de 51 millones de has de territorio argentino tienen cartografía convencional de suelos a una escala 1:50.000 (Angelini, 2010) (Figura 3.1). Esta cartografía fue elaborada dentro del Plan de Suelos de la Región Pampeana ejecutado por el INTA entre las décadas de los 60 y principios de los 90 (Moscatelli y Pazos, 2000), siguiendo los protocolos de delimitación de áreas fisiográficas y clasificación de suelos de la taxonomía americana (Soil Survey Staff, USDA). Desde entonces, esta cartografía ha sido la principal fuente de información de la variabilidad espacial de suelos en la pampa argentina (Del Valle, 2011).



**Figura 3.1** Localización de la cartografía convencional de suelos INTA a escala 1:50,000 dentro del territorio argentino.

Desde el año 2010, la mayoría de la cartografía convencional de suelos se encuentra disponible en la plataforma “WebGis” GeoINTA (Ver: <http://geointa.inta.gov.ar/visor/>) (INTA, 2013a). Espacialmente, esta cartografía está compuesta por polígonos agrupados en formato de *Unidad Cartográfica de Suelos* (UCS), la cual es considerada la unidad cartográfica básica. El agrupamiento y despliegue espacial de las UCS está guiada por la naturaleza y propiedades de los suelos en el área relevada, las cuales a su vez están determinadas por los factores formadores del suelo tales como topografía, organismos, material parental y clima (Jenny, 1941; McBratney *et al.*, 2003). La categoría más baja del sistema *soil taxonomy* es la *serie de suelos*. Esta describe los rangos límites de propiedades del suelo dentro de una UCS. Normalmente estos rangos se extienden espacialmente, más allá de los límites definidos en una serie de suelo. Por lo tanto, una UCS es el producto de una “agregación” de una o varias series de suelos, y por ende, la información de suelos descrita en cada serie de suelos, es considerada una generalización espacial (Nauman y Thompson, 2014).

La cartografía convencional de suelos en Argentina difiere en formatos y procedimientos de elaboración (Del Valle, 2011). Generalmente, la información de las propiedades del suelo de cada serie está tabulada y presentada de manera anexa al despliegue espacial de las UCS. La composición porcentual de cada serie dentro de la UCS está descrita en las tablas adjuntas a las UCS. De acuerdo a la composición de las series, las UCS se pueden clasificar como consociaciones, complejos, asociaciones y grupos indiferenciados (Casanelas, 2003). Por lo general, la descripción porcentual de cada serie no es espacialmente definida dentro de la UCS. Debido a esto, el polígono que representa una UCS es simplemente un resumen “promedio” de todas las posibilidades de tipos de suelos que se pueden encontrar dentro de un área determinada. Consecuentemente, los usuarios que utilizan la cartografía convencional de suelos, ignoran la incertidumbre relacionada con la distribución espacial de las series de suelos componentes de una UCS, llevándolos a la generalización de suelos de acuerdo a la “*serie dominante*” o condición de suelos dominante, lo cual no les permite acceder con certeza a la información de suelos que demandan.

A pesar de sus limitaciones, la información de suelos contenida en la cartografía convencional puede llegar a ser valiosa para mejorar y aumentar la resolución espacial de la información de suelos existente (Odgers *et al.*, 2013; Subburayalu *et al.*, 2014). Al respecto, muchos estudios lo han logrado utilizando técnicas de CDS (Bui *et al.*, 2009; Bui y Moran, 2001; Kerry *et al.*, 2012; Nauman y Thompson, 2014; Zhu *et al.*, 2001; Zhu *et al.*, 1997). En especial, el proyecto *GlobalSoilMap* (ver: <http://www.globalsoilmap.net/>) es un esfuerzo mundial que busca promover procedimientos estandarizados de mejoramiento y aumento de escala de la cartografía convencional de suelos existente (Hartemink *et al.*, 2010; Sanchez *et al.*, 2009). Este consorcio científico ha reconocido que las técnicas de CDS que utilizan la cartografía convencional de suelos, son eficientes para generar nuevos mapas funcionales de suelos a gran escala (Hartemink *et al.*, 2008).

En el contexto argentino, el problema que surge es *¿cómo utilizar la información contenida en la cartografía convencional de suelos INTA de manera adecuada, en orden de aumentar la escala y mejorar la información de suelos existente?* Varias técnicas de predicción espacial han sido aplicadas en el contexto de DSM en otros países (Grunwald, 2010; McBratney *et al.*, 2003). Por el contrario, en Argentina son escasos los trabajos al respecto (Angelini *et al.*, 2015; Del Valle, 2011). Las técnicas que han logrado resultados destacados (Kerry *et al.*, 2012; Nauman y Thompson, 2014; Odgers *et al.*, 2013), se han caracterizado por reestructurar la información de suelos existente, logrando *desagregar* espacialmente las UCS politáxicas, de modo que se represente mejor la distribución espacial de las series de suelo que la componen. Por lo general, las técnicas que aplican este tipo de reestructuración espacial, se les conoce como técnicas de desagregación espacial de UCS (DgUCS) (McBratney, 1998). Una breve introducción a estas técnicas, se describe a continuación.

### *1.1. Desagregación espacial de unidades cartográficas de suelos*

La DgUCS tiene como propósito generar información espacial del suelo en formato digital, a gran escala y satisfaciendo las demandas de conocimiento agronómico, a partir de la cartografía convencional de suelos existente, fuentes de información auxiliar y sobre

todo, teniendo en cuenta los conceptos pedológicos relacionados con los factores formadores del suelo planteados por Jenny (1941).

El objetivo principal de la DgUCS es lograr identificar espacialmente los tipos de suelos que componen la UCS a una escala  $>1:2,000$  (Odgers *et al.*, 2013). Al respecto, varios enfoques metodológicos han sido propuestos desde la década de los 90. Lagacherie *et al.* (1995) propuso un esquema probabilístico y cuantitativo de extrapolación de reglas de patrones espaciales de distribución de suelos para la DgUCS a partir de un área con diversa disponibilidad de cartografía convencional de suelos. Por su parte, McBratney (1998) propuso tres enfoques metodológicos para la DgUCS, utilizando funciones de transferencia espacial, análisis de fractales e interpolación picnofiláctica. De Bruin *et al.* (1999) utilizó un procedimiento escalonado (*stepwise*) de interpretación de imágenes aéreas junto con un método de aprendizaje inductivo para representar las relaciones paisaje-suelos.

En la década del 2000, en sintonía con el auge de nuevos desarrollos tecnológicos y estadísticos, otros tipos de propuestas para la DgUCS fueron planteadas. Bui y Moran (2001) y Hansen *et al.* (2009) aplicaron árboles de decisión para la desagregación y extrapolación de UCS. Por su parte, MacMillan *et al.* (2007) aplicaron técnicas híbridas de clasificación basada en lógica difusa y booleana para la DgUCS. Finalmente, a partir del año 2010 se intensificaron las propuestas de técnicas para lograr una eficiente DgUCS. Entre estas técnicas se incluye aquellas que utilizan clasificación supervisada (Nauman *et al.*, 2012), árboles de decisión probabilístico (Subburayalu *et al.*, 2014), regresión logística multinomial (Yang *et al.*, 2011), kriging área a punto (Kerry *et al.*, 2012), entre otras.

Un gran número de propuestas de DgUCS han sido planteadas en la presente década (Boettinger, 2010; Minasny *et al.*, 2012). Esto se debe a dos circunstancias: (i) nunca la demanda de suelos a escala de lote fue tan alta como ahora (Cook *et al.*, 2008; Del Valle, 2011) y (ii) nunca la oferta de tecnologías geomáticas, estadísticas, satelitales y sensores del suelo fue tan alta como ahora (Allred *et al.*, 2008; Thompson *et al.*, 2012).

La metodología que pueda lograr una DgUCS eficiente, a partir de cualquier cartografía convencional de suelos, en cualquier lugar, no existe (Minasny y McBratney, 2010; Odgers *et al.*, 2013). En el contexto de Argentina, todavía son muy escasos los trabajos que han evaluado técnicas de DgUCS (Angelini *et al.*, 2015). Sin embargo, hay que tener en cuenta que hasta el momento, la mayoría de las propuestas de DgUCS tienen todavía un alto grado de empirismo, una dificultad metodológica para predecir simultáneamente varias propiedades del suelo y por lo general, no logra generar cartografía de suelos a escala de lote. Por lo tanto, es pertinente plantear propuestas que subsanen estas limitaciones en condiciones de los suelos de Argentina.

Además de la cartografía convencional de suelos, en Argentina se dispone ampliamente de fuentes de información auxiliar de suelos a escala de lote agrícola, tales como modelos digitales de elevación, imágenes satélites, mapas de rendimiento y sensores del suelo (Bongiovanni y Lowenberg-DeBoer, 2005). Debido a esto, existe una óptima posibilidad de aplicar técnicas novedosas de DgUCS que permitan obtener mapas funcionales de suelos a escala de lote, útiles para la aplicación de agricultura de precisión, modelos hidrológicos y simulación de crecimiento y desarrollo de cultivos. Es claro que la eficiencia de la técnica de DgUCS radica no tanto en plantear un ejercicio predictivo, sino en comprobar dicha eficiencia a campo y en diversas condiciones. Debido a esto, se hipotetiza que una técnica de DgUCS que utilice simultáneamente el análisis de componentes principales espaciales (MULTISPATI-PCA) combinado con un análisis de distancia euclídea (ADE) puede generar la cartografía funcional de suelos a escala de lote que se demanda en los sistemas agroproductivos de Argentina.

El MULTISPATI-PCA es una extensión del análisis de componentes principales, el cual incorpora la dependencia espacial de las variables antes de que la componentes principales sean calculadas (Córdoba *et al.*, 2013; Dray *et al.*, 2008). La restricción espacial se impone a través del cálculo del índice de autocorrelación espacial de Moran (Oliver, 2010). Este índice mide la dependencia espacial entre observaciones y la media de las observaciones aledañas. El análisis multivariado de patrones espaciales a través de ACPs proporciona una herramienta útil para aumentar la escala espacial de información

existente (Arrouays *et al.*, 2011). Sin embargo, su aplicación en CDS y por ende en DgUCS, ha sido escasa (Cai *et al.*, 2015; Saby *et al.*, 2009).

Por su parte, el ADE ha sido ampliamente utilizado en CDS (Behrens *et al.*, 2010; Hengl *et al.*, 2004a; Minasny y McBratney, 2007). En general, ADE ha sido un algoritmo eficiente para determinar la distribución espacial de series de suelos. Sin embargo, son escasos los trabajos de CDS que aplicaron ADE a escala de lote.

El objetivo fue validar una metodología para la DegUCS a escala de lote a partir de MULTISPAT-PCA y ADE, teniendo como fuente de información la cartografía convencional de suelos a escala 1:50.000 y la información auxiliar de covariables ambientales relacionadas con suelos. Esta metodología se probó en tres zonas agrícolas de Argentina. Dentro del contexto de CDS, este trabajo se enfocó exclusivamente en la DgUCS a escala de lote. El propósito intrínseco de este trabajo es documentar tres casos donde se demuestre que la combinación de MULTISPAT-PCA y ADE es una metodología precisa, económica y rápida para lograr la DgUCS a escala de lote.

## **2. Materiales y métodos**

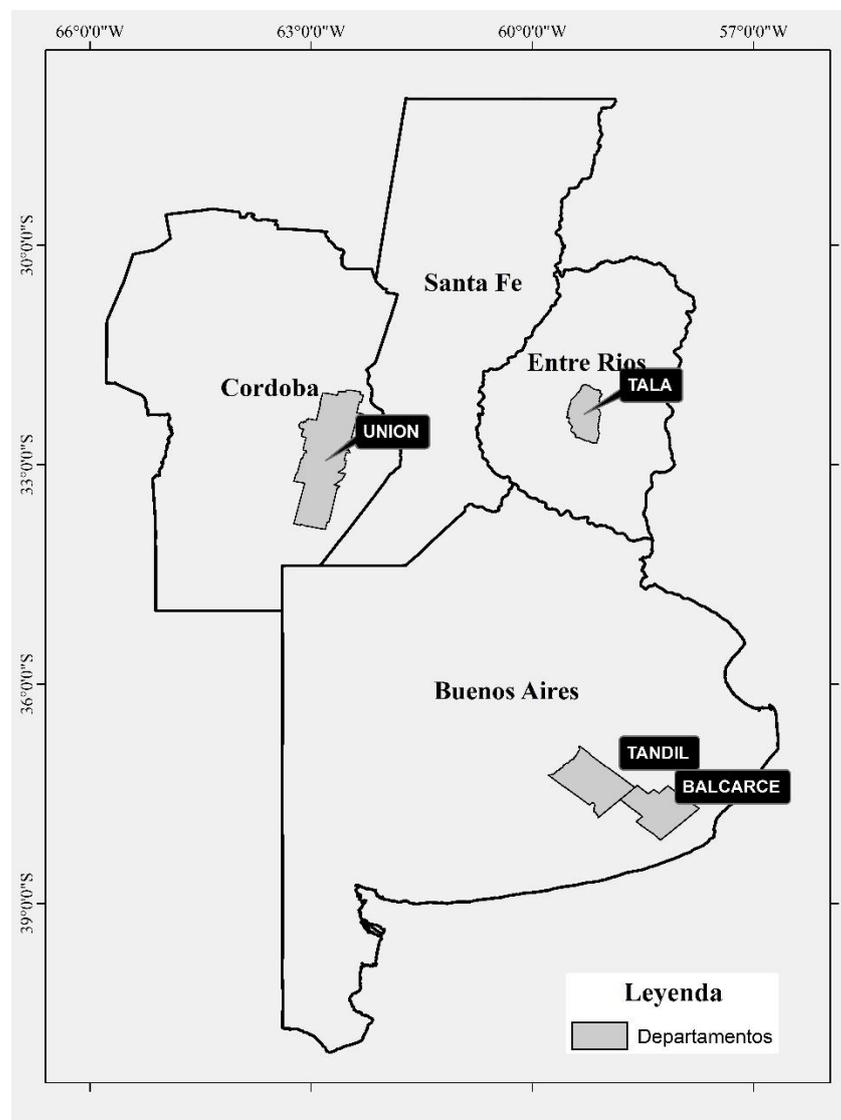
### *2.1. Zonas de estudio*

El estudio se llevó a cabo en tres zonas agrícolas de Argentina (Figura 3.2). Estas zonas se encuentran entre las latitudes 32°17'S – 37°17'S y las longitudes 58°17'O – 62°33'O. Cada zona tiene condiciones de suelo particulares y las prácticas de manejo de los cultivos van acorde con la oferta ambiental específica. Las tres zonas cuentan con cartografía convencional de suelos a escalas 1:50.000. Sin embargo, estas difieren en formato de presentación y procesos de elaboración.

#### *2.1.1. Zona 1: Sudeste de la provincia de Buenos Aires*

Dos lotes agrícolas fueron seleccionados en esta zona (Figura 3.2).

El lote *Aleluya* tiene un área de 42 has y se encuentra ubicado en el partido de Balcarce (*Lat: -37.6080, Lon: -58.6340*). De acuerdo con la cartografía convencional disponible, los suelos pertenecen a las unidades cartográficas MP11, MP16 y MP24, principalmente. Estos tienen una capacidad de uso II e, II\_es, y II\_es y un índice de productividad de 79.4\_B, 70.4\_B y 80.7\_D, respectivamente. Adicionalmente, parte del área del lote pertenece a la UCS CoAoC, la cual se está conformada en su totalidad por la serie Arroyo Coloquene.



**Figura 3.2** Localización de las tres zonas de estudio

La MP11 es una asociación de las series Mar del Plata (60%) y Balcarce (40%), ambas en fase fuertemente inclinada. La MP16 es una asociación de las series Mar del Plata

(50%), Azul (30%) y Cinco cerros (20%). La serie Mar del Plata es un Argiudol Típicos profundo, mientras que las series Balcarce, Azul y Cinco cerros, clasifican como Paleudoles Petrocálculos moderadamente profundos. La MP24 es una asociación de las series Mar del Plata (70%) y Tres Esquinas (30%). La serie Tres Esquinas es considerado un Argiudol Vértico. La UCS CoAoC fue clasificada como un complejo indiferenciado (Figura 3.3).

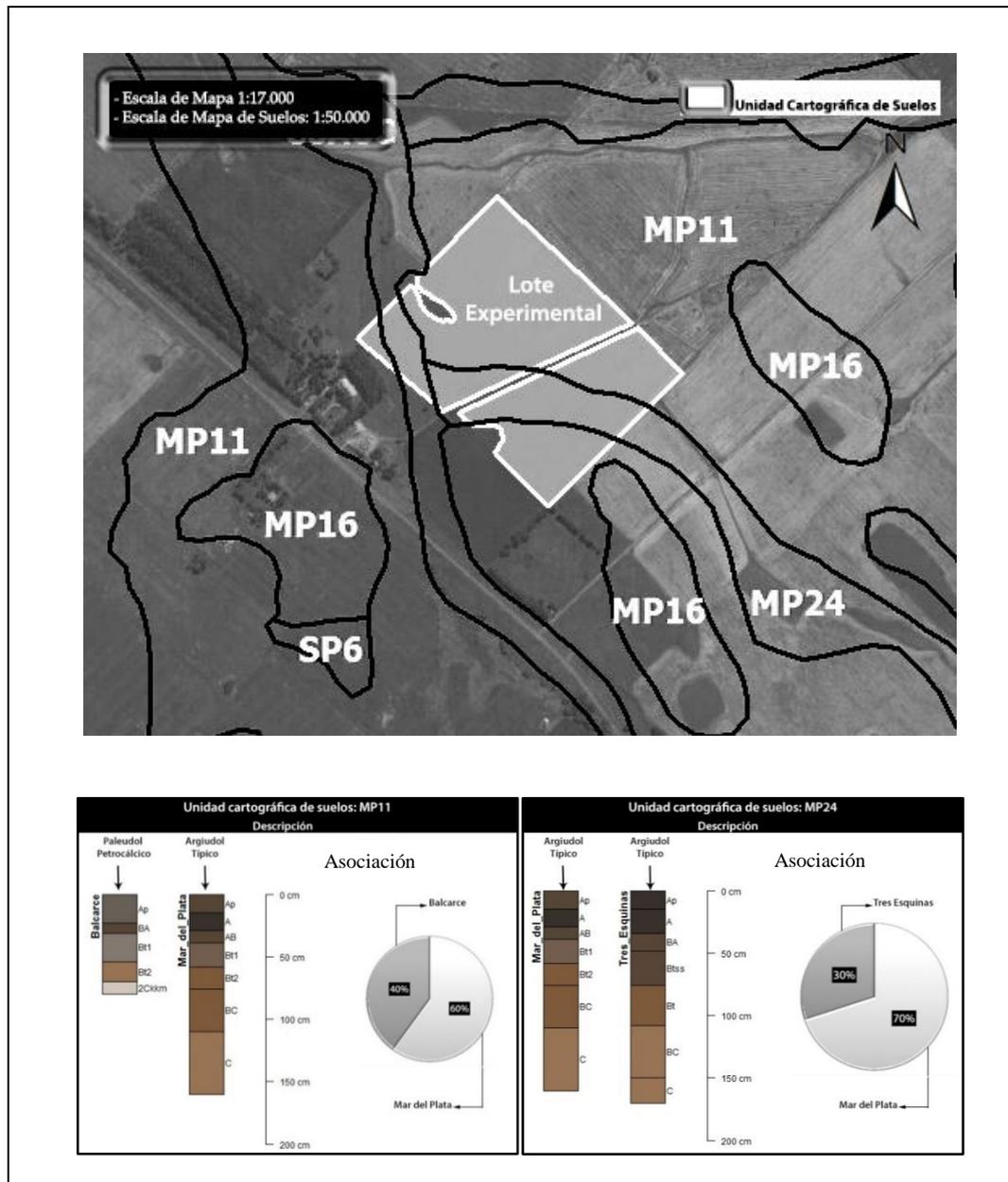
El lote *La Beatriz* tiene un área de 114 has y se encuentra ubicado en el partido de Tandil, (*Lat: -37.2070, Lon: -59.4190*). Los suelos pertenecen a las unidades cartográficas Ta19 y MP2, con una capacidad de uso III\_es y II\_es, y un índice de productividad de 67\_b y 77.2\_b, respectivamente. La Ta19 es una asociación de las series Tandil (70%) y Azul (30%) ambas en fase fuertemente inclinada. La MP2 es una consociación de las series Mar del Plata (80%) y Balcarce (20%). Las series Mar del Plata y Tandil son Argiudoles Típicos profundos, mientras que las series Balcarce y Azul son Paleudoles Petrocálculos moderadamente profundos (Figura 3.4).

El relieve general de ambos lotes se encuentra en un área de paisaje serrano en posición de loma y pendiente. Los subyacen sedimentos loésicos franco limosos bien drenados, no salinos, no alcalinos y en pendientes de 1 a 3% (INTA, 2010a).

#### 2.1.2. Zona 2: Sur de la provincia de Córdoba

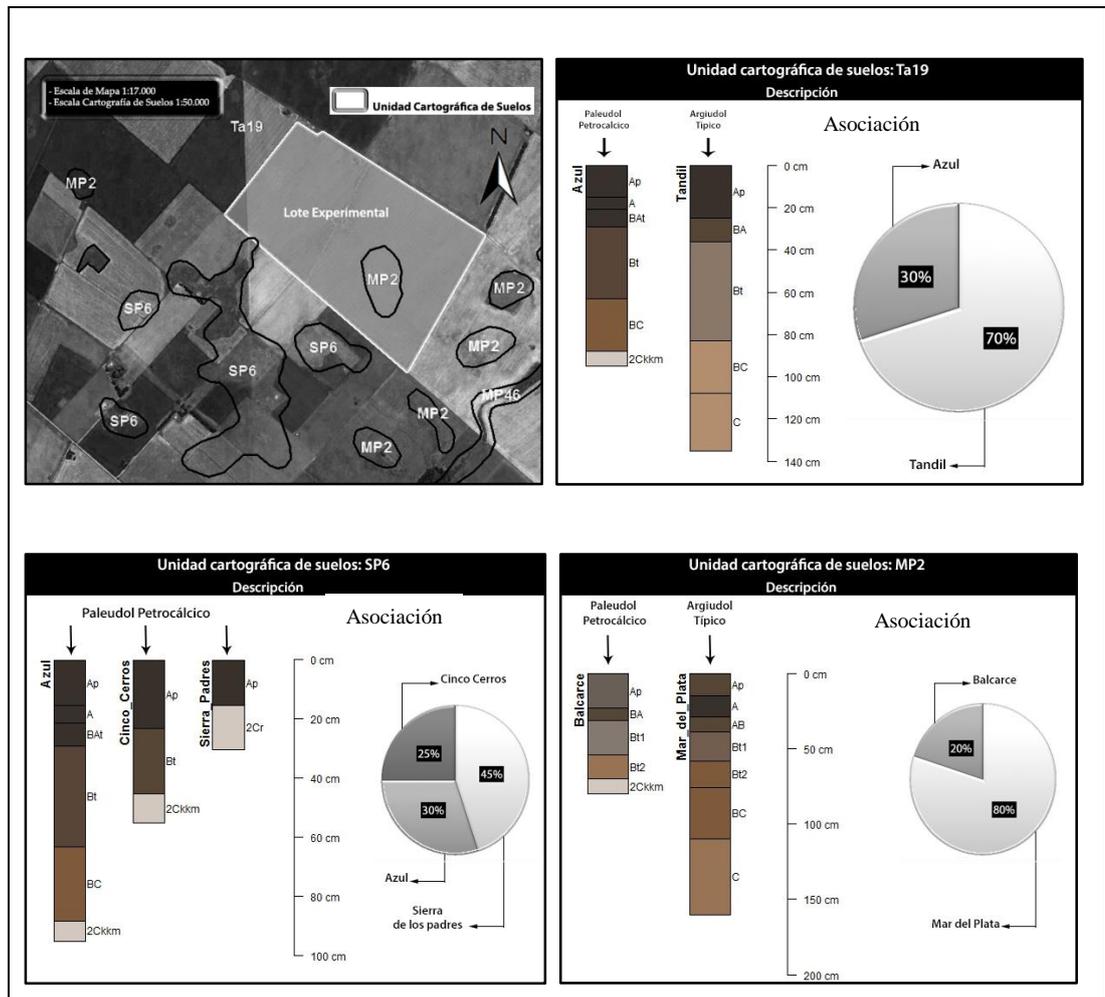
Dos lotes agrícolas fueron seleccionados en esta zona (Figura 3.2).

El lote *Mercapire* tiene un área de 42 has y se encuentra a 2.5 km al oeste de la zona urbana de Canals (*Lat: -33.5223, Lon: -62.8955*). Los suelos pertenecen a las unidades cartográficas EAo2 y Cn. La primera es una asociación de las series Los Aromos (60%), Medanitos (30%) y Canals (10%). La segunda es una consociación que está compuesta en su totalidad por la serie Canals (100%) (Figura 3.5).



**Figura 3.3** Distribución espacial y descripción pedológica de las unidades cartográficas de la cartografía convencional de suelos sobre el lote *Aleluya* – Buenos Aires

El lote San Esteban tiene un área de 44 has y se encuentra a 10.8 km al este de la zona urbana de Canals (Lat: -33.5558, Lon: -62.7623). Los suelos pertenecen a la unidad cartográfica Cn (Figura 3.5).



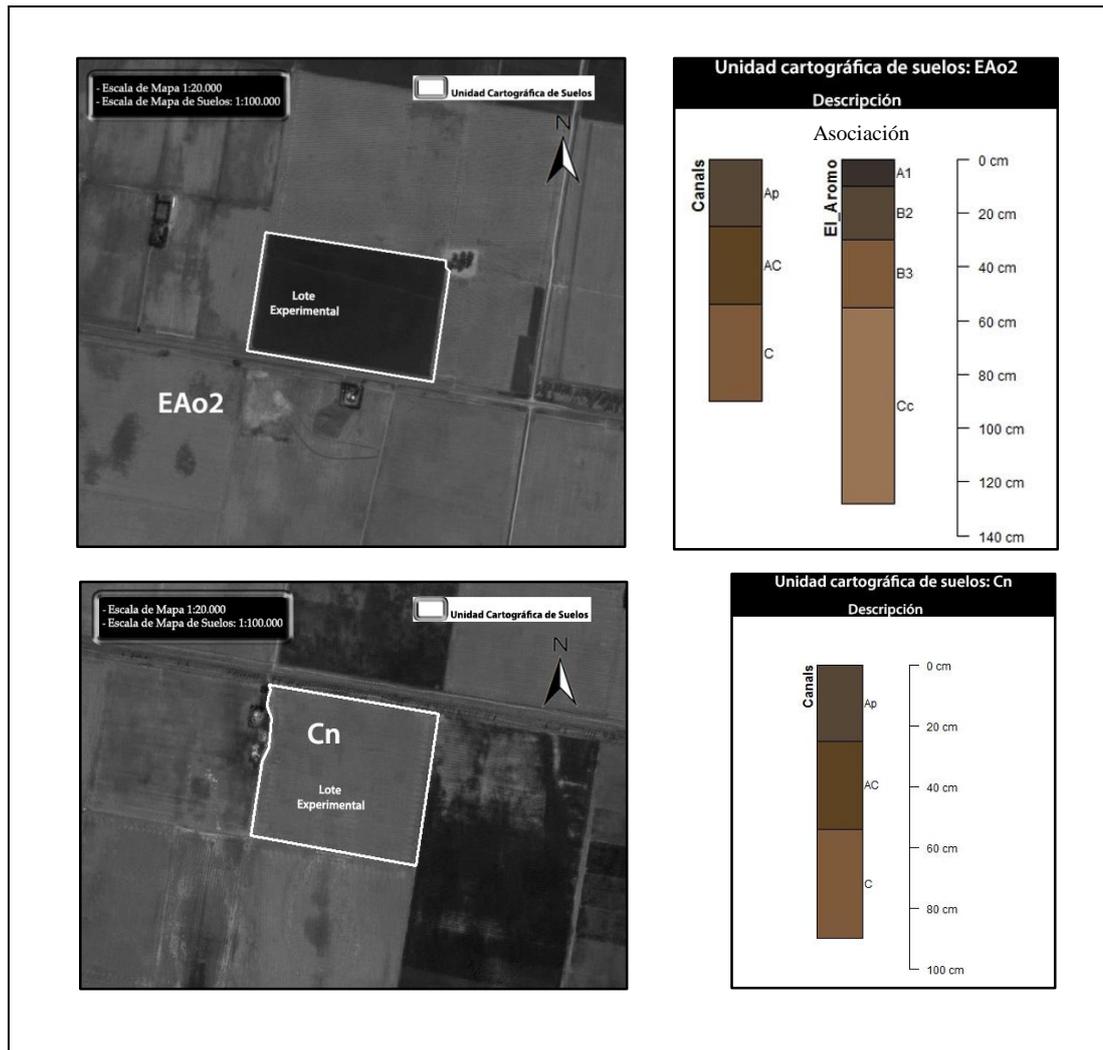
**Figura 3.4** Distribución espacial y descripción pedológica de las unidades cartográficas de la cartografía convencional de suelos sobre el lote *La Beatriz* – Buenos Aires

### 2.1.3. Zona 3: Centro de la provincia de Entre Ríos

Dos lotes fueron seleccionados en esta zona (Figura 3.2).

Ambos lotes son contiguos y pertenecen al mismo campo (*Lat: -32.5144, Lon: -59.4717*). El lote La Nueva Trinidad I (*LNT\_I*) y La Nueva Trinidad II (*LNT\_II*) tienen un área de 84 y 44 has, respectivamente. Los suelos pertenecen a las unidades cartográficas Eo2 y Ep3. La primera es una asociación de las series La Fermina (80%) y Cuatro Bocas (20%), que se caracteriza por presentarse en paisajes con planicies suavemente onduladas con manto de loess espeso. Mientras que la Ep3 es una asociación

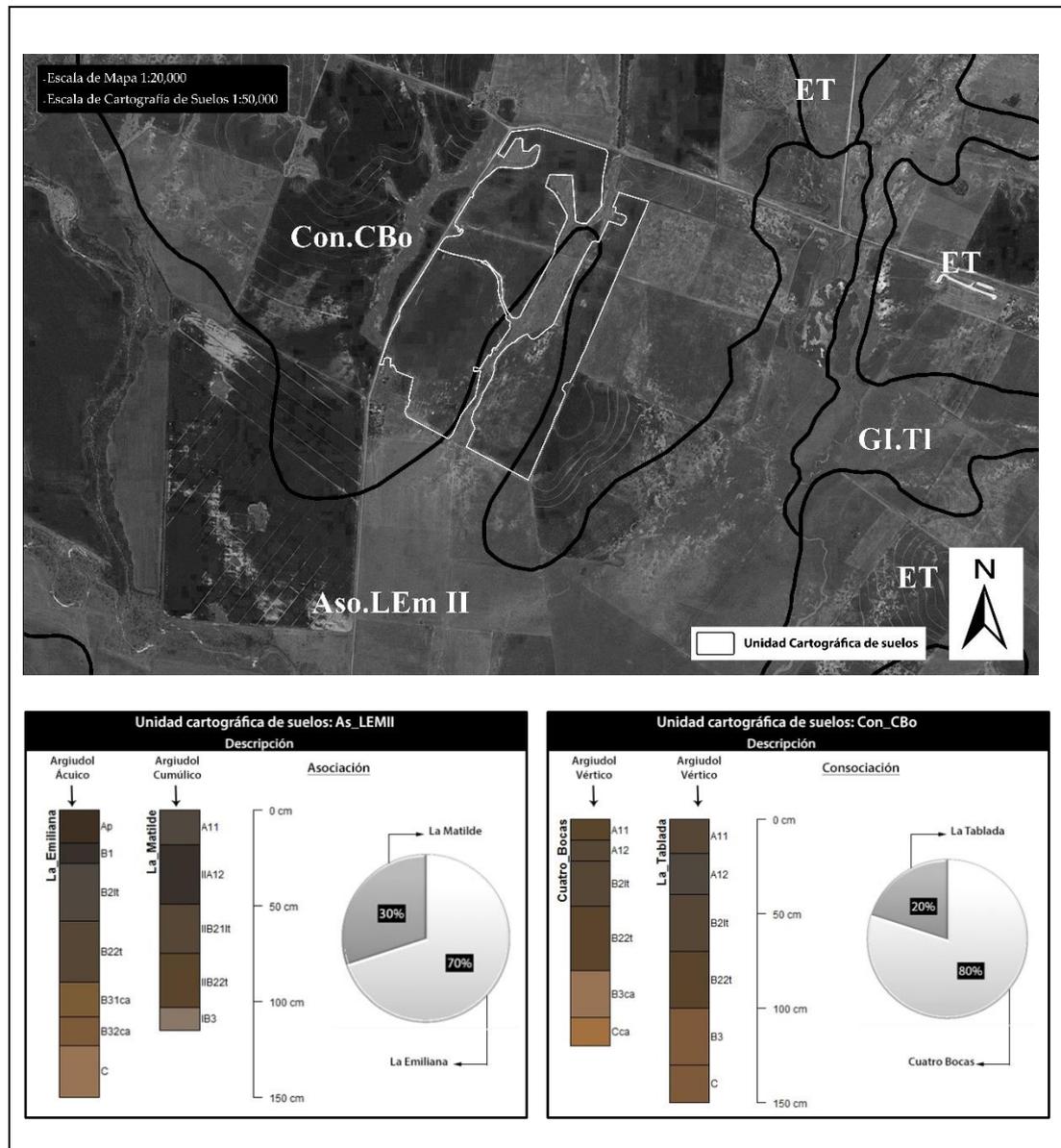
de las series Cuatro Bocas (80%) y La Tablada (20%), que se vincula a peniplanicies onduladas con manto de loess de poco espesor. Las series Cuatro Bocas y La Tablada clasifican como Argiudoles vérticos, mientras que la serie La Fermina como Argiudol típico (INTA y Gob\_Entre\_Ríos, 2012; Tasi, 2009) (Figura 3.6).



**Figura 3.5** Distribución espacial y descripción pedológica de las unidades cartográficas de la cartografía convencional de suelos sobre los lotes *Mercapire* (arriba) y *San Esteban* (abajo) – Córdoba

## 2.2. Fuentes de información auxiliar de covariables ambientales

Información auxiliar de covariables ambientales relacionadas con suelo, rendimiento y vegetación, fueron utilizadas en este estudio.



**Figura 3.6** Distribución espacial y descripción pedológica de las unidades cartográficas de la cartografía convencional de suelos sobre los lotes *La Nueva Trinidad I* (izquierda) y *La Nueva Trinidad II* (derecha) – Entre Ríos

La información auxiliar relacionada con suelo fue elevación, conductividad eléctrica aparente y profundidad de tosca. Esta última fue medida solamente en los lotes de la zona 1.

Los datos de elevación fueron recolectados utilizando un GPS con corrección diferencial (DGPS – Trimble® R3) en transectas separadas cada 20 m, en cada uno de los lotes. La Conductividad eléctrica aparente fue medida de 0-30 (CEa<sub>30</sub>) y 0-90 cm (CEa<sub>90</sub>)

de profundidad, utilizando el equipo de medición móvil VERIS 3100® (VERIS Technologies, Salina, KS, USA). Finalmente, la profundidad a la tosca fue medida sobre una grilla de muestreo de 30 x 30 m, distribuida por toda la superficie del lote. Cada medición consistió en la determinación de la profundidad a la cual, la resistencia mecánica es mayor a la equivalente aplicada con 2.5 ton de presión vertical. Para esto se adapta un barra de acero de 1.6 m de alto a un equipo hidráulico de exploración del suelos (Giddings Soil Sampling Equipment. Windsor, CO), el cual es trasladado a cada punto de medición.

La información auxiliar relacionada con rendimiento fueron los datos de monitoreo de rendimiento desde las cosechadoras. Para todos los lotes, se recolectaron, filtraron y seleccionaron la mayor cantidad posible de datos provenientes de monitores de rendimiento. Entre los lotes, estos datos tuvieron diferencia en época de recolección, cultivos, calibración, tipos y versión de monitores, e inclusive marcas de cosechadoras. Debido a esto fue necesario aplicar métodos robustos de filtrado.

La información auxiliar relacionada con vegetación fue el índice de vegetación de diferencia normalizada (NDVI) a partir de imágenes Landsat 5 TM, Landsat 8 OLI-TIRS y SPOT 5. Las imágenes Landsat fueron obtenidas de la página web del Departamento de Geología de los EEUU USGS (ver <http://earthexplorer.usgs.gov/>) para varias fechas. Por su parte, las imágenes SPOT fueron obtenidas a través de la plataforma de catálogo de imágenes de la CONAE (ver: <http://catalogos.conae.gov.ar/SPOT/>). Todas las imágenes satelitales fueron corregidas atmosféricamente y reprojectadas a sistemas de coordenadas planas UTM utilizando ENVI/IDL v4.7.

### *2.3. Análisis de la estructura espacial de las covariables ambientales*

Además del índice de autocorrelación espacial de Morán, complementariamente se analizó la estructura espacial de las covariables medidas a través del cálculo de semivariogramas teóricos (Oliver, 2010).

Semiariogramas fueron calculados y modelados para describir la estructura espacial de cada covariable ambiental medida. Estos fueron calculados utilizando la siguiente ecuación (Diggle y Ribeiro, 2007).

$$\hat{Y}(\mathbf{h}) = \frac{1}{2 N(\mathbf{h})} \sum_{\alpha=1}^{N(\mathbf{h})} [z(\mathbf{U}_{\alpha}) - z(\mathbf{U}_{\alpha} + \mathbf{h})]^2 \quad (1)$$

donde  $\hat{Y}(\mathbf{h})$  es el valor de semivarianza a un intervalo de distancia  $h$ ,  $z(\mathbf{U}_{\alpha})$  es el valor de la muestral a partir de todos los puntos de  $\mathbf{U}_{\alpha}$  y  $N(\mathbf{h})$  es el número de pares de datos dentro de una clase dada de distancia y dirección.

El semivariograma muestra la reducción en correlación espacial entre dos puntos cuando la distancia entre estos se incrementa (Peralta *et al.*, 2013). Después de verificar los supuestos geoestadísticos, el modelo de semivariograma teórico seleccionado y ajustado, fue utilizado para interpolar elevación, CEa y rendimiento a través de kriging ordinario. La interpolación geoestadística fue realizada utilizando el paquete “geoR” en el software estadístico R v3.1.1 (R Development Core Team, 2013). Finalmente, todos los archivos de salida se configuraron como raster (grilla regular) de 10 x 10 m, debido a que este tamaño refleja la escala de variabilidad asociada con escala de lote (Corwin y Lesch, 2005; Kitchen *et al.*, 2003).

#### 2.4. Metodología de desagregación de unidades cartográficas de suelo (DgUCS)

La DgUCS se realizó en tres etapas.

##### 2.4.1. Etapa I: Delimitación de zonas por tipo de suelo utilizando clasificación multivariada de covariables ambientales

La delimitación de zonas por tipo de suelo fue elaborada siguiendo la metodología propuesta por Córdoba *et al.* (2013). Este algoritmo utiliza un análisis de componentes principales espaciales (CPs) de las covariables ambientales obtenidos a través del análisis

MULTISPATI-ACP, como entrada de un análisis de cluster fuzzy c-medias (CFK) (Goktepe *et al.*, 2005).

Tomando una matriz  $\mathbf{X}$  ( $n \times p$ ) de datos que tienen muchas mediciones ( $p$ ) en cada uno de los ( $n$ ) puntos, el algoritmo MULTISPATI-ACP introduce una matriz de ponderación espacial  $\mathbf{W}$  en el análisis de componentes principales (ACP) de  $\mathbf{X}$  estandarizada. La matriz  $\mathbf{W}$  es una matriz de la suma de las filas previamente estandarizadas. Si  $\mathbf{C} = [C_{ij}]$  es la matriz de conectividad que indica las interacciones entre los puntos  $i$  y  $j$ , entonces  $\mathbf{W} = [C_{ij} / \sum_{j=1}^n C_{ij}]$ .

De acuerdo a la extensión del vector lag, una matriz lag  $\tilde{\mathbf{X}} = \mathbf{W}\mathbf{X}$  puede ser definida. Las dos tablas  $\tilde{\mathbf{X}}$  y  $\mathbf{X}$  tienen la misma cantidad de columnas (variables) y filas (observaciones). MULTISPATI-ACP tiene como objetivo identificar las estructuras multivariadas mediante el estudio de la relación entre  $\tilde{\mathbf{X}}$  y  $\mathbf{X}$  utilizando análisis de coinercia (Dray *et al.*, 2003).

La matriz lag  $\tilde{\mathbf{X}}$  está compuesta de las medias ponderadas de los valores de los puntos vecinos a través de la matriz de conexión espacial. Es decir, que únicamente los puntos vecinos son tenidos en cuenta. El puntaje de filas de este análisis maximiza el producto escalar entre la combinación lineal de las variables originales y una combinación lineal de las variables “lageadas”. La ventaja de MULTISPATI-ACP sobre ACP es que el resultado de muestras del MULTISPATI-ACP maximiza la autocorrelación espacial entre sitios, mientras que el ACP convencional maximiza la varianza total. Los resultados del MULTISPATI-ACP son, por lo tanto, “suavizados” y muestran fuertes estructuras espaciales en los primeros ejes, mientras que los resultados de ACP pueden ser robustos, suavizados o mixtos, y pueden mostrar estructuras espaciales en cualquier eje, incluso en los más lejanos (Arrouays *et al.*, 2011).

MULTISPATI-ACP fue calculado utilizando los paquetes “ade4” (función “multispati” Chessel *et al.* (2004)) y “spdep” (Bivand *et al.*, 2005) del software estadístico R v3.1.1 (R Development Core Team, 2013). El número mínimo de variables sintéticas

que lograra explicar más del 70% de la varianza fueron seleccionados y utilizados para el análisis CFK.

El análisis CFK se aplicó a partir de los componentes principales espaciales (CPs) seleccionados. Existen tres tipos de matrices involucradas en el análisis CFK. La primera contiene los datos que se quieren clasificar. Particularmente, en este trabajo la matriz de datos  $\mathbf{X}$  implica  $n$  observaciones con  $a < p$  CPs cada uno. La segunda matriz es la de los centróides  $\mathbf{V}$  de los clúster, que consta de centróides clúster  $k$  localizados en el atributo espacial definido por los CPs seleccionados. Por último, la tercera matriz de pertenencia difusa  $\mathbf{U}$ , la cual consta de valores de pertenencia para cada clúster en  $\mathbf{V}$  en cada punto muestreado en  $\mathbf{X}$ , delimitada por la restricción de que la suma de los valores de pertenencia para cada observación debe ser igual a 1. Una óptima partición difusa  $k$  es definida como una minimización de la medida ponderada de la distancia al cuadrado entre dos puntos y tipos de centróides:

$$J_m(U, v) = \sum_{j=1}^n \sum_{i=1}^k (U_{ij})^m (d_{ij})^2 \quad (2)$$

donde  $m$  es el coeficiente ponderado difuso ( $1 \leq m < \infty$ ), y  $(d_{ij})^2$  es la distancia al cuadrado en el atributo espacial entre el punto  $j$  y el centróide de clúster  $i$ .

Los índices de coeficiente de partición (*fuzziness performance index-FPI*), entropía de clasificación (*normalized classification entropy-NCE*) (Odeh *et al.*, 1992), índice de Xie-Beni (Pal y Bezdek, 1995; Xie y Beni, 1991), índice Fukuyama (Fukuyama y Sugeno, 1989; Rezaee *et al.*, 1998) y exponente de proporción (Windham, 1981) fueron utilizados para determinar el número de zonas. Para usar complementariamente la información de cada uno estos índices se realizó una combinación de los mismos, obteniendo un único índice resumen. Este índice resumen se elaboró a partir de la distancia Euclidiana de los valores de todos índices, previamente normalizados por su valor máximo a través de las diferentes clasificaciones.

Por último, diagramas de caja de cada covariable ambiental, en cada una de las zonas, fueron analizados y descritos.

#### 2.4.2. Etapa II: Validación de la delimitación de zonas

El proceso de validación de zonas se realizó a partir de los resultados del análisis de laboratorio de las muestras de suelos tomadas por zona delimitada, en cada uno de los lotes.

El número mínimo de muestras de suelos por zona delimitada fue tres y el máximo cinco, dependiendo de la representatividad espacial de cada zona. La ubicación de los puntos de muestreo de suelos por zona se hizo de manera que cubriera homogéneamente, toda el área de cada una de las zonas delimitadas. Las profundidades de muestreo de suelos fueron de 0-30 y 30-60 cm. Estas profundidades son las que habitualmente se utilizan para analizar la variabilidad espacial de suelos cuando se hacen mediciones con sensores proximales tales como el VERIS 3100® (VERIS Technologies, Salina, KS, USA) (Corwin *et al.*, 2006; Peralta *et al.*, 2013).

Las muestras de suelos fueron colectadas en bolsas plásticas. Posteriormente fueron llevadas al laboratorio de suelos de la EEA INTA Balcarce. A cada muestra se le determinó materia orgánica del suelo (MO) por dicromato oxidación (Walkley y Black, 1934), textura por el método de la pipeta de Robinson (Gee y Bauder, 1986; Robinson, 1922), capacidad de intercambio catiónico (CIC) por el método de acetato de amonio ( $\text{NH}_4 \text{COO-CH}_3$ ) (Chapman, 1965), pH en a suspensión 1:25 suelo-agua y conductividad eléctrica del extracto de saturación ( $\text{CE}_{\text{ext}}$ ) utilizando el método electrométrico (Norman, 1965).

En orden de comparar las propiedades del suelo dentro de cada zona delimitada, un modelo lineal mixto (MLM) y un ANOVA fueron ajustados para cada una de las propiedades del suelo a partir de modelo básico:

$$s_{ij} = \mu + Z_i + \varepsilon_{ij}$$

donde  $S_{ij}$  representa el valor observable de la propiedad del suelo en la zona  $i$  del lote  $j$ ;  $\mu$  representa la media general de la propiedad del suelo;  $Z_i$  es el efecto de la zona delimitada con  $i = 1, \dots, Z$ ; y  $\varepsilon_{ij}$  es el error aleatorio. Estos modelos fueron ajustados con varianzas heterogéneas para las diferentes zonas delimitadas. La selección del modelo se hizo de acuerdo al criterio de información de Akaike. Para este procedimiento se utilizó el paquete “nlme” (Pinheiro *et al.*, 2007) del software estadístico R v3.1.1 (R Development Core Team, 2013).

#### 2.4.3. Etapa III: Relaciones espaciales entre propiedades del suelo analizadas y propiedades de las series de suelos

Estas relaciones fueron analizadas y descritas a través de un biplot de análisis de componentes principales (ACP), elaborado a partir de los resultados de una Matriz de Distancia Euclidiana (MDE) entre propiedades del suelo analizadas y propiedades reportadas en las series de suelos.

### 3. Resultados y discusión

#### 3.1. Parámetros de estadística descriptiva

##### 3.1.1. Zona 1: Sudeste de la provincia de Buenos Aires

La Tabla 3.1 muestra los parámetros de estadística descriptiva de las covariables ambientales en los lotes de la zona 1.

Elevación mostró una fuerte autocorrelación espacial en ambos lotes (*Índice de Moran*  $>0.85$ ). La diferencia de elevación en *La Beatriz* fue de 26.24 m, mientras que en *Aleluya* fue de 11.07 m. En general, ambos lotes están representando los tipos de topografía a escala de lote típicos de la zona 1. *Aleluya* representa a los lotes de las partes bajas del terreno, los cuales son medianamente planos y que forman parte de zonas de transición hidrológica (Bedogni, 1996). Por su parte, *La Beatriz* integra el área agrícola de zonas

serranas, con pronunciadas ondulaciones y cambios de altura. La localización con respecto a las zonas serranas es el factor zonal principal que determina los patrones espaciales de las propiedades del suelo (Blanco y Stoops, 2007; Buschiazzo, 1986).

**Tabla 3.1** Parámetros de estadística descriptiva de las covariables ambientales para los lotes de la zona 1 (Buenos Aires)

Campo	Variables	Media	Mediana	DE*	Min*	Max*	IM*		
Aleluya Buenos Aires	Suelos	Elevación	163.56	163.31	2.61	157.71	168.78	0.85	
		Tosca	100.22	106.43	14.87	60.30	116.00	0.55	
	Rendimiento	CEa <sub>30</sub> <sup>(1)</sup>	18.21	17.55	2.97	13.32	27.73	0.68	
		CEa <sub>90</sub> <sup>(1)</sup>	26.55	26.04	3.68	18.41	36.00	0.51	
		Trigo 2008 <sup>(2)</sup>	4.47	4.45	0.86	2.01	6.72	0.56	
		Soja 2007/2008 <sup>(2)</sup>	2.31	2.31	0.35	1.29	3.38	0.55	
		Vegetación	NDVI-19/11/2009 <sup>(3a)</sup>	0.10	0.10	0.01	0.06	0.22	0.67
	NDVI-22/01/2010 <sup>(3a)</sup>		0.70	0.72	0.05	0.56	0.77	0.57	
	NDVI-04/01/2015a <sup>(3b)</sup>		0.22	0.23	0.08	0.17	0.39	0.77	
	NDVI-29/01/2015b <sup>(3b)</sup>		0.26	0.26	0.06	0.18	0.56	0.76	
			Media	Mediana	DE*	Min*	Max*	IM*	
	La Beatriz Buenos Aires	Suelos	Elevación	213.57	213.72	5.98	200.67	226.91	0.86
			Tosca	104.32	104.70	13.14	53.53	143.50	0.21
CEa <sub>30</sub> <sup>(1)</sup>			22.59	22.19	3.07	14.88	31.07	0.68	
CEa <sub>90</sub> <sup>(1)</sup>			25.86	25.78	3.20	15.72	37.61	0.59	
Rendimiento		Soja 2006/2007 <sup>(2)</sup>	1.85	1.76	0.39	1.03	2.85	0.67	
Vegetación		NDVI-19/11/2009 <sup>(3a)</sup>	0.50	0.51	0.03	0.30	0.59	0.81	
		NDVI-23/12/2010 <sup>(3a)</sup>	0.21	0.21	0.04	0.10	0.31	0.89	

<sup>(1)</sup> CEa<sub>30</sub> y CEa<sub>90</sub>: Conductividad Eléctrica Aparente de 0 - 30 cm y 0 - 90 cm, respectivamente.

<sup>(2)</sup> Mapas de rendimiento

<sup>(3)</sup> NDVI = Índice de Vegetación de Diferencia Normalizada obtenido a partir de imágenes Landsat 5 T<sup>(3a)</sup>, Landsat 8<sup>(3b)</sup> y SPOT 5<sup>(3c)</sup>.

\* DE: Desviación Estandar; Min: Valor mínimo; Max: Valor Máximo; IM: Índice de Morán.

La profundidad de tosca mostró una leve y moderada autocorrelación espacial para *La Beatriz* y *Aleluya*, respectivamente. Únicamente zonas con suelos con profundidad de tosca mayor a los 120 cm se presentaron en *La Beatriz*. En general, en la mayoría de los lotes agrícolas del sudeste bonaerense los patrones espaciales de la profundidad de tosca son complejos. Al respecto, Pazos y Mestelan (2002) determinaron que la tosca tiene gran variabilidad en profundidad, estructura, grado de dureza y contenido de CaCO<sub>3</sub> a escala de lote. Esta variabilidad por lo general es intrínseca de cada lote y depende de las

circunstancias geológicas que rodean el área de cada lote (Bedogni, 1996; Cabria y Culot, 1994).

La CEA<sub>30</sub> y CEA<sub>90</sub> presentaron moderada autocorrelación espacial en ambos lotes (*Índice de Morán*  $<0.68$ ). Estos resultados no coinciden con los resultados de otros trabajos (Carroll y Oliver, 2005; Corwin y Plant, 2005; Friedman, 2005; Heiniger *et al.*, 2003). Una explicación al respecto, es la compleja variabilidad espacial de la profundidad y dureza de la tosca en ambos lotes. Al respecto, Bork *et al.* (1998), Boettinger *et al.* (1997) y Domenech *et al.* (2012) determinaron que el patrón espacial de la tosca determina muchos parámetros hidrológicos del suelo, y por ende determina el patrón espacial de la CEA.

Los mapas de rendimiento presentaron moderada autocorrelación espacial en ambos lotes. Esto se explica por la interacción espacial entre tosca y rendimiento (Peralta *et al.*, 2015). Generalmente, la tosca determina el patrón espacio-temporal del contenido de agua y nutrientes en el suelo (Sadras y Calviño, 2001). Debido a esto, el patrón espacial del rendimiento depende de la variabilidad de la profundidad a la tosca y de las condiciones ambientales. Aunque no fue posible conseguir dos mapas de rendimiento de una misma campaña y cultivo para ambos lotes, es pertinente destacar que en campañas con muy buena disponibilidad hídrica, tales como 2007/2008, se obtuvo una media de rendimiento de soja (~2.5 ton/ha), normal para la zona (Giménez *et al.*, 2015). Por el contrario, campañas con una limitada disponibilidad hídrica, tales como 2008, el trigo rindió muy por debajo de la media para la zona. De acuerdo con esto, se estableció que en lotes con tosca, el patrón espacial del rendimiento depende de las condiciones ambientales predecesoras. Este aspecto se destacó en los datos de rendimiento disponibles para ambos lotes de la zona 1.

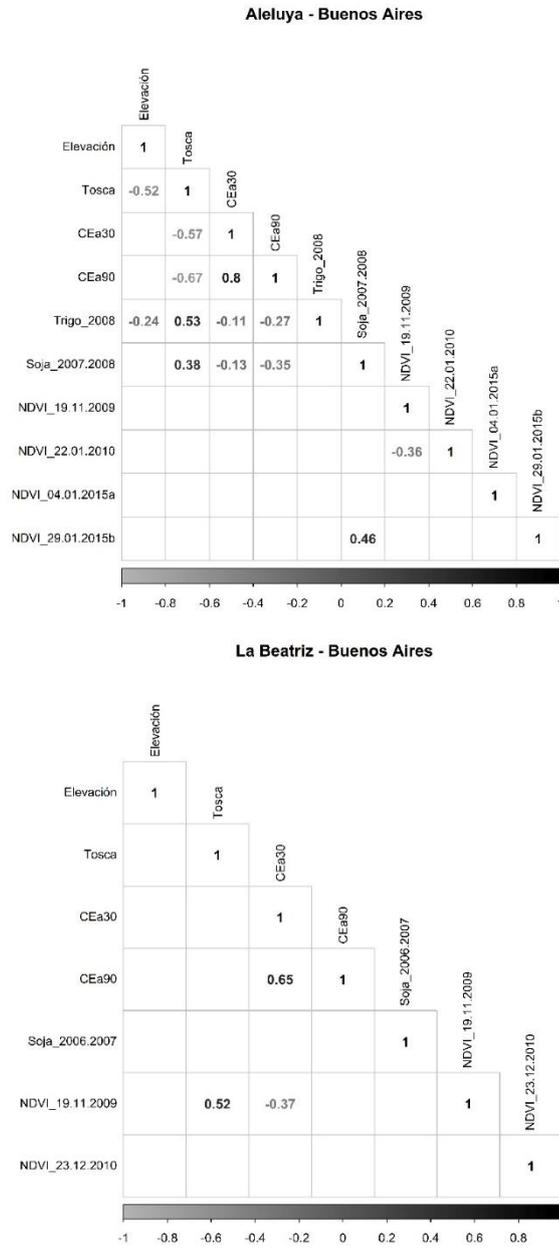
En *Aleluya*, NDVI presentó moderada y fuerte autocorrelación espacial (*Índice de Morán entre*  $0.57 - 0.77$ ). Moderada autocorrelación espacial estuvo relacionada con NDVI obtenidos en épocas de sequía, para este lote. En el momento en que se toma la imagen satelital que se utilizó para calcular el NDVI\_19/11/2009 y NDVI\_22/01/2010, la precipitación acumulada en el mes inmediatamente anterior fue un 23% y 12% inferior

al promedio histórico, respectivamente (*Fuente:* SIGA-INTA, ver: <http://siga2.inta.gov.ar/>). Por su parte, en el momento en que se toma la imagen satelital que se utilizó para calcular el NDVI\_04/01/2015, la precipitación acumulada fue 15% superior a la media histórica, mientras para NDVI\_29/01/2015 fue un 12% inferior al promedio histórico.

Por su parte, el NDVI en *La Beatriz* presentó fuerte autocorrelación espacial (*Índice de Morán*  $>0.81$ ). Sin embargo, NDVI\_19/11/2009 tuvo una media mayor que en *Aleluya*. Este resultado se debe a que proporcionalmente *La Beatriz*, tiene un área mayor de suelos con una profundidad  $>80$  cm, en comparación con *Aleluya*. Es ampliamente conocido que la tosca limita el almacenamiento de agua. Debido a esto, en este trabajo se ratifica que el complejo patrón espacial de la profundidad a la tosca y el patrón temporal de las condiciones ambientales, son los principales factores que determinan la autocorrelación espacial y valores medios del NDVI. Además, estos resultados confirman la conveniencia de seleccionar varias fechas de imágenes, con el fin de entender la relación entre covariables tales como: oferta ambiental/tosca/vegetación/rendimiento (Boettinger *et al.*, 2008).

La figura 3.7 muestra los coeficientes de correlación entre covariables ambientales para la zona 1.

En *Aleluya*, la profundidad a la tosca determinó el patrón espacial de las demás covariables ambientales. Elevación presentó correlación negativa con tosca y rendimiento de Trigo\_2008. Por su parte, tosca presentó correlación negativa con ambas CEa y positiva con rendimiento. Ambas CEa tuvieron leve correlación negativa con rendimiento. Valores altos de CEa en zonas someras puede deberse a dos causas: (i) mayor disponibilidad de agua en los horizontes más cercanos a la superficie (Amiotti *et al.*, 2008; Domenech, 2013) y (ii) mayor contenido de arcillas a profundidades entre 0-40 cm (Buschiazzo, 1986). Por otro parte, NDVI no presentó correlación con tosca y elevación. Sin embargo, NDVI\_22/01/2010 presentó correlación con rendimiento de soja 2007/2008. Los resultados confirman que el patrón espacial del NDVI y su relación con las demás covariables ambientales depende del patrón temporal de oferta hídrica.



**Figura 3.7** Coeficientes de correlación ( $r$ ) de las covariables ambientales de los lotes de la zona 1 (Buenos Aires). Los  $r$  mostrados fueron significativos a un nivel de probabilidad  $p \leq 0.05$

En *La Beatriz*, se presentó moderada correlación positiva y negativa de NDVI\_19/11/2009 con tosca y CEa<sub>30</sub>, respectivamente. Las demás correlaciones entre covariables ambientales no fueron significativas. Como se esperaba, en condiciones de déficit hídrico, el patrón espacial de NDVI está determinado por el patrón espacial de la profundidad de tosca.

### *Zona 2: Sur de la provincia de Córdoba*

La Tabla 3.2 muestra los parámetros de estadística descriptiva de las covariables ambientales en los lotes de la zona 2.

Elevación presentó una fuerte autocorrelación espacial para ambos lotes (*Índice de Morán*  $>0.88$ ). La mayor diferencia de elevación en ambos lotes fue de 3.88 m. Ambos lotes se encuentran en la zona llamada pampa loésica plana cordobesa o planicies arenosas (INTA, 2010b). En esta zona, los gradientes de pendiente son menores a 0.5%. Principalmente, las diferencias de elevación a escala de lote en esta zona, se deben a la presencia de zonas suavemente deprimidas y vías de escurrimiento.

La CEa<sub>30</sub> y CEa<sub>90</sub> presentaron una moderada autocorrelación espacial en ambos lotes. Este resultado se podría deber al complejo patrón espacial de las intercalaciones entre zonas deprimidas y vías de escurrimiento, a escala de lote (Peralta *et al.*, 2013). En general, las medias de CEa de la zona 2 fueron altas en comparación con la CEa de la zona 1. Especialmente, en el lote San Esteban la CEa<sub>90</sub> presentó valores muy altos, síntoma de la presencia de bajos salinos. Estos resultados podría deberse a que en la pampa loésica plana, la capa freática fluctúa entre dos a seis metros y por lo general, afecta los suelos localizados en zonas deprimidas (INTA, 2010b). Debido a esta condición, Córdoba *et al.* (2013) y Peralta *et al.* (2013) lograron concluir que la CEa tiene mucho potencial para determinar la variabilidad espacial de las propiedades del suelo en condiciones de la zona 2. Además, ha sido una herramienta eficaz para la delimitación de zonas para manejo sitio específico de cultivos.

**Tabla 3.2** Parámetros de estadística descriptiva de las covariables ambientales para los lotes de la zona 2 (Córdoba)

Campo	Variables	Media	Mediana	DE*	Min*	Max*	IM*		
Mercapire Córdoba	Suelos	Elevación	138.15	138.14	0.82	136.44	140.32	0.94	
		CEa <sub>30</sub> <sup>(1)</sup>	50.38	49.79	3.56	42.86	60.63	0.54	
		CEa <sub>90</sub> <sup>(1)</sup>	45.60	44.94	3.87	37.86	55.50	0.55	
	Rendimiento	Maiz 2005/2006 <sup>(2)</sup>	2.58	2.10	1.65	0.82	9.34	0.89	
		Soja 2009/2010 <sup>(2)</sup>	3.27	3.28	1.26	0.96	6.56	0.84	
		Soja 2011/2012 <sup>(2)</sup>	3.23	3.29	1.27	1.02	6.24	0.81	
	Vegetación	NDVI-16/02/2006 <sup>(3a)</sup>	0.36	0.34	0.07	0.23	0.60	0.83	
		NDVI-06/01/2014a <sup>(3b)</sup>	0.44	0.44	0.04	0.35	0.58	0.62	
		NDVI-25/01/2014b <sup>(3c)</sup>	0.13	0.13	0.02	0.10	0.16	0.67	
			Media	Mediana	DE*	Min*	Max*	IM*	
	San Esteban Córdoba	Suelos	Elevación	136.22	136.33	0.64	134.55	137.46	0.88
			CEa <sub>30</sub> <sup>(1)</sup>	45.64	45.49	4.68	31.61	58.73	0.62
CEa <sub>90</sub> <sup>(1)</sup>			55.86	52.73	10.98	36.54	89.88	0.58	
Rendimiento		Soja 2011/2012 <sup>(2)</sup>	1.60	1.64	0.49	0.63	2.60	0.82	
Vegetación		NDVI-16/02/2006 <sup>(3a)</sup>	0.58	0.59	0.05	0.44	0.66	0.79	
		NDVI-06/01/2014a <sup>(3b)</sup>	0.22	0.22	0.01	0.20	0.24	0.62	
		NDVI-25/01/2014b <sup>(3c)</sup>	0.10	0.10	0.04	0.04	0.14	0.74	

<sup>(1)</sup> CEa<sub>30</sub> y CEa<sub>90</sub>: Conductividad Eléctrica Aparente de 0 - 30 cm y 0 - 90 cm, respectivamente.

<sup>(2)</sup> Mapas de rendimiento

<sup>(3)</sup> NDVI = Índice de Vegetación de Diferencia Normalizada obtenido a partir de imágenes Landsat 5 T<sup>(3a)</sup>, Landsat 8<sup>(3b)</sup> y SPOT 5<sup>(3c)</sup>.

\* DE: Desviación Estandar; Min: Valor mínimo; Max: Valor Máximo; IM: Índice de Morán.

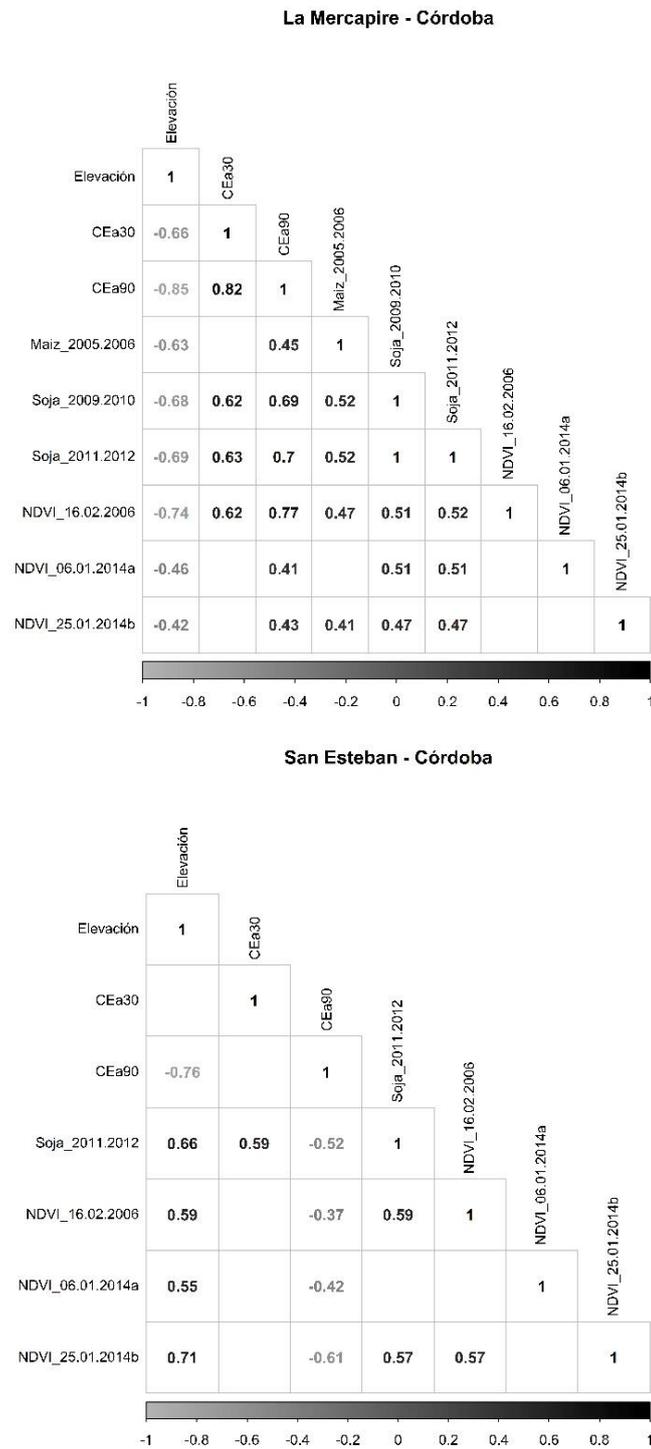
Rendimiento presentó fuerte correlación espacial en ambos lotes (*Índice de Morán* >0.81). El rendimiento general de los cultivos en la zona está determinado por la precipitación acumulada, la cual históricamente no ha sido >850 mm. Debido a esto, los rendimientos promedio de la zona suelen ser bajos, con respecto a otras zonas agrícolas de la pampa argentina. El efecto temporal de disponibilidad hídrica se reflejó en los rendimientos de los cultivos. A diciembre 31 de 2005, la precipitación acumulada de los cuatro meses inmediatamente anteriores fue 25% inferior que la precipitación acumulada histórica para ese mismo periodo (775 mm). Debido a esto, los rendimientos de maíz en la campaña 2005/2006 se vieron severamente afectados. Por otro parte, en enero de 2012 la precipitación acumulada fue de 89 mm, lo cual afectó notoriamente el rendimiento de la soja en la campaña 2011/2012. Los resultados indican que los mapas de rendimiento en la zona deben ser interpretados de acuerdo a la oferta hídrica, especialmente la que se

concentra en la precipitación acumulada, durante los meses de definición del periodo crítico del cultivo.

NDVI presentó fuerte autocorrelación espacial en ambos lotes. NDVI\_01/2014 no presentó diferencias de autocorrelación espacial entre fuentes de imágenes satelitales. Al igual que el rendimiento, el efecto temporal de disponibilidad hídrica se reflejó la variabilidad espacial y las medias de NDVI. Valores altos de medias de NDVI estuvieron relacionados con épocas con alta precipitación acumulada. NDVI\_06/01/2014 y NDVI\_25/01/2014 presentaron valores contrastantes debido a la precipitación acumulada en el mes inmediatamente anterior. Para el primer NDVI, se presentó precipitación acumulada por encima de la media, mientras que para el segundo NDVI, estuvo por debajo de la media. Estos resultados confirman el enorme efecto que tienen las variables estocásticas en la variabilidad espacial de los cultivos en la zona.

La figura 3.8 muestra los coeficientes de correlación entre covariables ambientales para la zona 2.

En *Mercapire*, elevación determinó el patrón espacial de las demás covariables ambientales. Elevación presentó moderada correlación negativa con rendimiento, NDVI y CEa<sub>30</sub> ( $r$  entre  $-0.42$  -  $-0.74$ ). Además, presentó correlación negativa fuerte con CEa<sub>90</sub>. Excepto Maíz\_2005/2006, CEa<sub>30</sub> presentó moderada correlación positiva con mapas de rendimiento y NDVI\_16/02/2006, mientras que CEa<sub>90</sub> presentó moderada correlación positiva con todos los mapas de rendimiento y NDVI\_06/01/2014, y fuerte correlación positiva con NDVI\_16/02/2006. En general, los mapas de rendimiento tuvieron moderada correlación positiva entre sí. Por su parte, la mayoría de los NDVI tuvo correlación positiva con mapas de rendimiento. Sin embargo, ninguna tuvo correlación significativa entre sí. Estos resultados determinaron que las propiedades del suelo que determinan el patrón espacial de la CEa, en su gran mayoría pueden estar determinando la generación de biomasa y rendimiento de los cultivos.



**Figura 3.8** Coeficientes de correlación ( $r$ ) de las covariables ambientales de los lotes de la zona 2 (Córdoba). Los  $r$  mostrados fueron significativos a un nivel de probabilidad  $p \leq 0.05$

En *San Esteban*, al igual que en *Mercapire*, la elevación determinó el patrón espacial de las demás covariables ambientales. Sin embargo, las relaciones de correlación no fueron similares. Correlación negativa fuerte entre elevación y  $CEa_{90}$  al igual que en *Mercapire*, mientras que correlación positiva moderada se presentó entre elevación – mapas de rendimiento y elevación – NDVI, al contrario que en *Mercapire*. Por su parte  $CEa_{90}$  tuvo moderada correlación negativa con rendimiento y NDVI, al contrario que en *Mercapire*. A pesar de que existe relación entre  $CEa_{90}$  y rendimiento en *San Esteban*, es posible que las causas de estas relaciones sean diferentes a las de *Mercapire*. Al respecto, es pertinente recordar que el valor máximo y la desviación estándar de  $CEa_{90}$  fueron muy altos en *San Esteban*, en comparación con *Mercapire*. Es posible que en zonas con valores altos de  $CEa_{90}$ , se estén delimitando áreas puntuales con alto contenido de sales. Las cuales a su vez, afectan el rendimiento. Por su parte, las relaciones espaciales entre elevación, rendimiento y NDVI en *San Esteban*, fueron en sentido contrario a las de *Mercapire*. Además de confirmar que existen zonas con presencia de sales, estas relaciones confirman que se ubican en las zonas bajas de *San Esteban*.

Estos resultados confirman lo reportado por Simón *et al.* (2013). Estos autores determinaron que la CEa tiene un alto potencial para determinar la variabilidad espacial de propiedades del suelo limitantes del rendimiento, en condiciones la zona 2. Así mismo, se confirmó que elevación es una covariable ambiental apta para definir zonas a escala de lote, de acuerdo a los tipos de suelos. Sin embargo, las relaciones espaciales entre elevación, propiedades del suelo y rendimiento deben ser determinadas con precisión. La generalización del uso de CEa y la elevación en condiciones de la zona 2, puede conllevar a malas interpretaciones y por ende a implementar prácticas inadecuadas de manejo de suelos y cultivos.

### 3.1.2. Zona 3: Centro de la provincia de Entre Ríos

La Tabla 3.3 muestra los parámetros de estadística descriptiva de las covariables ambientales en los lotes de la zona 3.

Elevación presentó una fuerte autocorrelación espacial en ambos lotes (*Índice de Morán*  $>0.86$ ). La menor diferencia de elevación para ambos lotes fue 32 m. Ambos lotes se encuentran ubicados en zonas de planicies muy onduladas (INTA y Gob\_Entre\_Ríos, 2012). Así mismo, se caracterizaron por tener pendientes largas.

**Tabla 3.3** Parámetros de estadística descriptiva de las covariables ambientales para los lotes de la zona 2 (Entre Ríos)

Campo	Variables	Media	Mediana	DE*	Min*	Max*	IM*
La Nueva	Suelos						
Trinidad I	Elevación	58.78	60.00	9.15	37.05	85.00	0.85
Entre Ríos	CEa <sub>30</sub> <sup>(1)</sup>	98.12	105.32	35.27	10.53	175.41	0.64
	CEa <sub>90</sub> <sup>(1)</sup>	120.79	122.01	23.09	60.83	200.76	0.48
	Rendimiento						
	Trigo 2009 <sup>(2)</sup>	3.93	3.94	0.58	2.54	5.17	0.65
	Maíz 2011/2012 <sup>(2)</sup>	4.36	4.41	0.52	3.03	5.67	0.63
	Vegetación						
	NDVI-10/01/2009 <sup>(3a)</sup>	0.42	0.43	0.04	0.30	0.52	0.74
	NDVI-16/01/2011 <sup>(3a)</sup>	0.33	0.33	0.04	0.24	0.41	0.81
	NDVI-16/12/2014 <sup>(3b)</sup>	0.20	0.20	0.03	0.14	0.30	0.74
		Media	Mediana	DE*	Min*	Max*	IM*
La Nueva	Suelos						
Trinidad II	Elevación	53.50	53.00	7.28	41.26	74.00	0.86
Entre Ríos	CEa <sub>30</sub> <sup>(1)</sup>	79.23	79.36	25.75	21.79	138.51	0.63
	CEa <sub>90</sub> <sup>(1)</sup>	130.51	131.69	21.51	70.70	196.79	0.66
	Rendimiento						
	Trigo 2009 <sup>(2)</sup>	3.56	3.59	0.42	2.43	4.52	0.66
	Maíz 2011/2012 <sup>(2)</sup>	4.75	4.72	0.64	3.16	6.21	0.73
	Vegetación						
	NDVI-10/01/2009 <sup>(3a)</sup>	0.29	0.29	0.06	0.12	0.42	0.81
	NDVI-16/12/2014 <sup>(3b)</sup>	0.21	0.21	0.03	0.17	0.30	0.78

<sup>(1)</sup> CEa<sub>30</sub> y CEa<sub>90</sub>; Conductividad Eléctrica Aparente de 0 - 30 cm y 0 - 90 cm, respectivamente.

<sup>(2)</sup> Mapas de rendimiento

<sup>(3)</sup> NDVI = Índice de Vegetación de Diferencia Normalizada obtenido a partir de imágenes Landsat 5 T<sup>(3a)</sup>, Landsat 8<sup>(3b)</sup> y SPOT 5<sup>(3c)</sup>.

\* DE: Desviación Estandar; Min: Valor mínimo; Max: Valor Máximo; IM: Índice de Morán.

La CEa<sub>30</sub> y CEa<sub>90</sub> presentaron moderada autocorrelación espacial para ambos lotes (*Índice de Morán entre 0.48 – 0.66*). Este resultado se explica por la presencia heterogénea de zonas salinas. Por lo general, altos contenidos de sales se encuentran en zonas cóncavas del lote. Los valores de CEa fueron muy altos para ambos lotes, en comparación con aquellos de la zona 1 y 2. Estos valores se explican por dos motivos (*i*) la presencia de horizontes del suelo con contenido de arcilla mayor al 40% y (*ii*) la presencia heterogénea de zonas salinas. Aunque no se tiene evidencia de trabajos previos

que utilicen la CEa en condiciones de la zona 3, es claro que CEa permite detallar y diferenciar espacialmente, zonas salinas a escala de lote.

Los mapas de rendimiento presentaron moderada autocorrelación espacial (*Índice de Morán entre 0.63 – 0.66*). Estos resultados se deben a (i) la presencia heterogénea de zonas salinas en ambos lotes; y (ii) al anegamiento puntual y periódico de zonas cóncavas dentro de los lotes, en época de precipitaciones intensas. En general, las medias de rendimientos fueron bajas. Principalmente, estas se debieron al anegamiento de zonas cóncavas y a la salinidad. Específicamente, en la campaña 2011/2012 la precipitación acumulada entre enero y febrero fue de 279 mm, de los cuales el 32% se debió a la intensa precipitación que se presentó en la zona, los días 5 y 6 de febrero de 2012 (ver: <http://www.hidraulica.gob.ar/>). Los suelos de la zona se caracterizan por ser moderadamente bien drenados, con escurrimiento superficial medio y con permeabilidad lenta (Tasi, 2009). Debido a esto, es normal que los rendimientos sean muy bajos en las zonas cóncavas de los lotes agrícolas, en época de precipitaciones intensas.

NDVI presentó moderada y fuerte autocorrelación espacial (*Índice de Morán entre 0.52 – 0.88*). Al igual que en las otras zonas, el efecto de las precipitaciones acumuladas y el patrón espacial de zonas cóncavas determinó el patrón espacio-temporal del NDVI en cada lote. Sin embargo, históricamente en la zona de Nogoyá, las precipitaciones son muy intensas en el verano. Debido a esto, normalmente se presentan anegamientos en zonas bajas, que asociada a la baja tasa de infiltración de los suelos, suelen afectar el crecimiento vegetativo de los cultivos. En el momento que fue tomada la imagen para calcular el NDVI\_16/01/2009, la precipitación acumulada en el mes inmediatamente anterior, estuvo por debajo de la media histórica. Por el contrario, NDVI\_16/01/2011 fue precedida de precipitación acumulada en el mes inmediatamente anterior cercana a la media. Sin embargo, presentó días en que cayeron hasta 58 mm. Esta situación afectó notoriamente la vegetación y por ende, explica los valores medios inferiores con respecto al NDVI\_10/01/2009.

La figura 3.9 muestra los coeficientes de correlación entre covariables ambientales para la zona 3.

En *La Nueva Trinidad I* (LNT\_I), elevación determinó la variabilidad espacial de las demás covariables ambientales. Elevación presentó moderada correlación con CEa<sub>30</sub>. Excepto NDVI\_12/2013, CEa<sub>30</sub> presentó baja correlación con NDVI y Trigo\_2009. Por su parte, CEa<sub>90</sub> presentó baja correlación con Maíz\_2011/2012 y NDVI\_01/2009. En general, zonas con valores altos de CEa estuvieron relacionadas con zonas de bajos rendimientos y bajo NDVI. Como se mencionó anteriormente, la topografía determina la localización heterogénea de áreas con alta presencia de sales. Estas a su vez, afectaron la medición de CEa, la vegetación y el rendimiento.

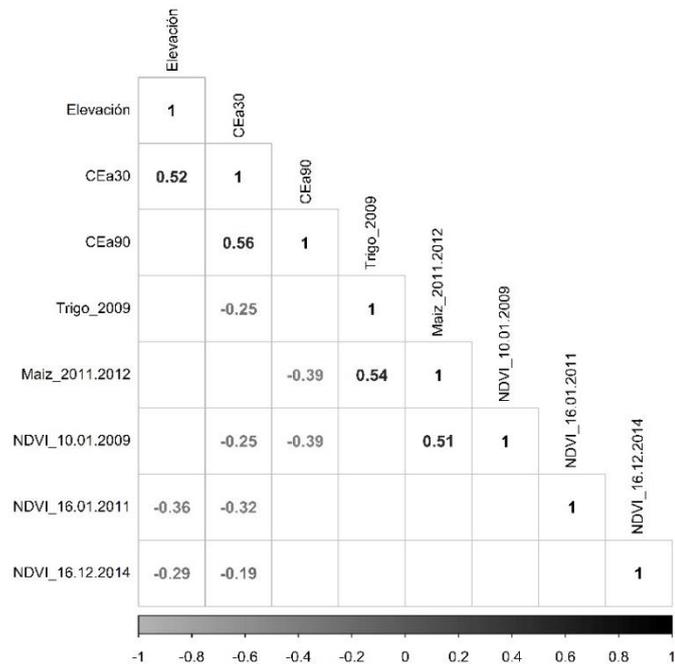
En *La Nueva Trinidad II* (LNT\_II), ocurrió algo similar a LNT\_I. Sin embargo, la mayoría del área de LNT\_II son bajos. Por lo general, estas áreas conllevan presencia nociva de sales. Esto significa, que las áreas bajas salinas afectan de manera más puntual a la vegetación y el rendimiento en LNT\_II, en comparación con LNT\_I. Como resultado de esto, CEa<sub>90</sub> presentó mayor correlación negativa con NDVI.

### 3.2. *Análisis de la estructura espacial*

La Tabla 3.4 muestra los parámetros de los modelos teóricos de semivariograma que mejor ajustaron para cada una de las variables, en cada uno de los lotes.

Como se esperaba, elevación tuvo una fuerte dependencia espacial en todos los lotes ( $Co (\%) < 9$ ). Generalmente, esta variable regionalizada es de tipo estacionaria débil o no estacionaria (Hengl y Reuter, 2008; Olaya y Conrad, 2009; Schaetzl y Anderson, 2005). Es decir, que su varianza no es finita, y mucho menos a escala de lote (Giraldo Henao, 2006).

La Nueva Trinidad I - Entre Ríos



La Nueva Trinidad II - Entre Ríos

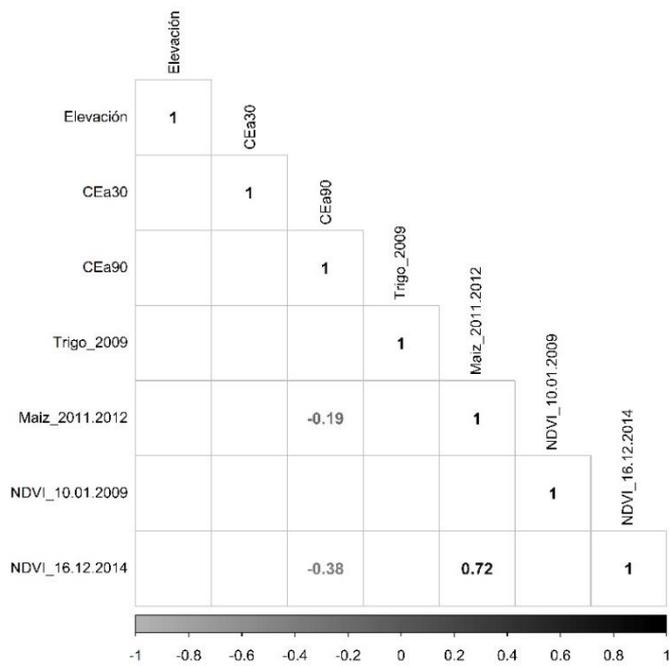


Figura 3.9 Coeficientes de correlación ( $r$ ) de las covariables ambientales de los lotes de la zona 3 (Entre Ríos). Los  $r$  mostrados fueron significativos a un nivel de probabilidad  $p \leq 0.05$

**Tabla 3.4** Parámetros geoestadísticos de modelos ajustados de semivariogramas experimentales para los predictores medidos

Predictor	Modelo <sup>(1)</sup>	Co <sup>(2)</sup>	Co + C <sup>(3)</sup>	a (m) <sup>(4)</sup>	Co (%) <sup>(5)</sup>	Dependencia Espacial <sup>(6)</sup>
<i>Alehuya - Buenos Aires</i>						
Elevación	Lineal	0.00	40.00	600.30	0.00	F
Tosca	Circular	158.83	249.50	374.09	63.66	M
CEa <sub>30</sub>	Exponencial	3.43	8.72	168.92	39.33	M
CEa <sub>90</sub>	Exponencial	12.67	14.08	85.76	89.99	L
Trigo_2008	Exponencial	0.64	0.95	207.90	67.37	M
Soja_2007/2008	Circular	0.15	0.14	228.60	107.14	L
<i>La Beatriz - Buenos Aires</i>						
Elevación	Lineal	0.00	205.06	816.80	0.00	F
Tosca	Circular	440.10	176.00	202.09	250.06	L
CEa <sub>30</sub>	Exponencial	2.42	10.43	41.58	23.20	F
CEa <sub>90</sub>	Exponencial	9.56	10.09	51.98	94.75	L
Soja_2006/2007*	Gaussiano	0.09	0.21	311.08	42.86	M
<i>Mercapire - Córdoba</i>						
Elevación	Circular	0.00	0.84	239.08	0.00	F
CEa <sub>30</sub>	Exponencial	5.39	23.11	46.78	23.32	F
CEa <sub>90</sub>	Exponencial	6.93	24.57	51.48	28.21	M
Maiz_2005/2006*	Gaussiano	0.00	3.78	89.74	0.00	F
Soja_2009/2010*	Gaussiano	0.00	1.84	58.21	0.00	F
Soja_2011/2012*	Circular	0.11	0.46	145.50	23.91	F
<i>San Esteban - Córdoba</i>						
Elevación	Lineal	0.00	0.85	441.70	0.00	F
CEa <sub>30</sub>	Exponencial	6.13	31.67	77.96	19.36	F
CEa <sub>90</sub>	Exponencial	14.74	164.40	98.75	8.97	F
Soja_2011/2012*	Circular	0.04	0.36	332.64	11.11	F
<i>La Nueva Trinidad I - Entre Ríos</i>						
Elevación	Lineal	0.00	120.30	1100.80	0.00	F
CEa <sub>30</sub>	Exponencial	612.80	1084.00	145.50	56.53	M
CEa <sub>90</sub>	Exponencial	518.60	1178.20	127.30	44.02	M
Trigo_2009*	Exponencial	0.13	0.42	155.93	30.95	M
Maiz_2011/2012*	Exponencial	0.14	0.28	51.98	50.00	M
<i>La Nueva Trinidad II - Entre Ríos</i>						
Elevación	Exponencial	4.44	47.39	187.30	9.37	F
CEa <sub>30</sub>	Exponencial	151.79	1123.20	88.36	13.51	F
CEa <sub>90</sub>	Circular	260.85	59.28	200.10	440.03	L
Trigo_2009*	Exponencial	0.08	0.29	272.80	27.59	M
Maiz_2011/2012*	Exponencial	0.04	0.42	57.17	9.52	F

<sup>(1)</sup> Todos los modelos son isotrópicos<sup>(2)</sup> Nugget<sup>(3)</sup> Silla;<sup>(4)</sup> Rango.<sup>(5)</sup> Nugget (%): (Nugget / Silla) x 100<sup>(6)</sup> **F** = Fuerte dependencia espacial (Nugget (%) <25); **M** = Moderada dependencia espacial (Nugget (%) entre 25 y 75); **L** = Leve dependencia espacial (Nugget (%) >75) (Cambardella, *et al.*, 1994).

\*. Mapas de rendimiento

Para tosca, semivariogramas teóricos circulares fueron los que mejor ajustaron. Este modelo teórico hace parte de los semivariogramas transitivos o acotados (Kerry y Oliver, 2003; Oliver, 2010). Es decir, que no alcanzan realmente a formar la silla ( $Co + C$ ). Este resultado indica que es posible que la tosca sea una variable que en alguna dirección no presenta estacionariedad. Teniendo en cuenta que la elevación tiene correlación espacial con la tosca, es posible que está también presente patrones espaciales correspondientes. En general, la dependencia espacial fue leve y moderada, confirmando los resultados encontrados con el índice de Morán (Cambardella *et al.*, 1994).

Excepto para  $CEa_{90}$  del lote LNT\_II, los semivariogramas teóricos exponenciales fueron los que mejor ajustaron para CEa. Este tipo de semivariograma es el más adecuado cuando los predictores tienen una distribución desigual a cortas distancias (Oliver, 2010). CEa es una variable muy susceptible a cambios a corta distancia. Por lo general, un efecto combinado de propiedades del suelo es lo que determina la CEa. Debido a esto, la CEa tiende a variar por diversos efectos que se presentan a distancias cortas. La dependencia espacial de la CEa no presentó un patrón general estable y varió de acuerdo a las circunstancias particulares de cada lote.

Para rendimiento en los lotes de Entre Ríos, semivariogramas teóricos exponenciales fueron lo que mejor ajustaron. Para los mapas de rendimiento de las demás zonas, no se presentó un patrón general de semivariogramas teóricos. Excepto para Soja\_2007/2008 del lote Aleluya – Buenos Aires, todos los mapas de rendimiento presentaron dependencia espacial moderada y fuerte. Estos resultados confirman lo encontrado por Guedes Filho *et al.* (2010), quienes determinaron que el efecto de las variaciones temporales de condiciones ambientales, los tipos de cultivos y las fechas de siembra determinan el efecto del patrón espacial de propiedades del suelos sobre el rendimiento. Para futuros trabajos, sería interesante analizar el efecto de los patrones espacio-temporales sobre la estructura espacial del rendimiento, utilizando semivariogramas.

En general, los resultados de este análisis de la estructura espacial confirmaron que los semivariogramas son herramientas más efectivas para detallar la estructura espacial de las covariables ambientales, en comparación con los índices de Morán. Sin embargo, el

índice de Morán es un dato puntual que permite ser incluido en modelos de regresión o análisis de datos multivariados. Una desventaja de los semivariogramas es que su análisis requiere de conocimientos previos o a priori de supuestos teóricos. Futuros trabajos deben incluir otros tipos de análisis de la estructura espacial que permitan detallar la estacionariedad, momentos de variables regionalizadas y anisotropía, tales como correlogramas y covariogramas (Giraldo Henao, 2006).

### 3.3. Aplicación de la metodología de DgUCS

#### 3.3.1. Análisis MULTISPATI-ACP de covariables ambientales

La tabla 3.5 muestra los resultados de la clasificación multivariada de covariables ambientales utilizando análisis de componentes principales (ACP) y MULTISPATI-ACP.

El índice de autocorrelación espacial o índice de Morán de las primeras tres variables sintéticas generadas con MULTISPATI-ACP para todos los lotes, en todas las zonas, fueron estadísticamente significativas.

En general, los componentes principales espaciales (CPs) obtenidos con MULTISPATI-ACP tuvieron menor varianza y mayor índice de Morán que los componentes principales (CP) obtenidos con el análisis convencional (ACP).

Las tablas 3.6, 3.7 y 3.8 muestran los autovectores o peso (*mayor contribución*) de las covariables ambientales para la construcción de los CPs, para las zonas 1, 2 y 3, respectivamente. En general, los autovectores con mayor peso fueron mayores en los CPs, en comparación con los CP. Con escasas excepciones, las variables con mayor peso no cambiaron entre CPs y CP. Cambios en el peso o en la contribución de una covariable entre CP y CPs son explicados por las diferencias en los índices de Morán de las variables sintéticas (Córdoba *et al.*, 2013; Dray y Jombart, 2011; Dray *et al.*, 2008).

**Tabla 3.5** Estadística descriptiva de las tres primeras componentes principales generadas a partir de análisis de componentes principales (ACP) y MULTISPATI-ACP para las tres zonas

Campo	Variabes	Varianza	Porcentaje	Porcentaje acumulado	IM <sup>(1)</sup>
Aleluya Buenos Aires	ACP <sup>(2)</sup>				
	CP1	2.95	29.51	29.51	0.83*
	CP2	1.95	19.51	49.02	0.65*
	CP3	1.78	17.86	66.88	0.70*
	MULTISPATI-ACP <sup>(3)</sup>				
	CPs 1	2.51	37.53	37.53	0.86*
	CPs 2	1.45	21.67	59.20	0.78*
	CPs 3	1.14	17.04	76.24	0.63*
La Beatriz Buenos Aires	ACP <sup>(2)</sup>				
	CP1	2.44	34.87	34.87	0.73*
	CP2	1.45	20.82	55.69	0.85*
	CP3	1.07	15.38	71.07	0.85*
	MULTISPATI-ACP <sup>(3)</sup>				
	CPs 1	2.36	34.76	34.76	0.78*
	CPs 2	1.41	24.12	58.88	0.91*
	CPs 3	1.08	16.84	75.72	0.83*
Mercapire Córdoba	ACP <sup>(2)</sup>				
	CP1	5.15	57.27	57.27	0.82*
	CP2	1.15	9.89	67.16	0.54*
	CP3	0.89	8.36	75.52	0.52*
	MULTISPATI-ACP <sup>(3)</sup>				
	CPs 1	5.14	68.82	68.82	0.82*
	CPs 2	1.08	10.92	79.74	0.62*
	CPs 3	0.87	8.87	88.61	0.62*
San Esteban Córdoba	ACP <sup>(2)</sup>				
	CP1	3.71	53.04	53.04	0.89*
	CP2	1.10	15.82	68.86	0.41*
	CP3	0.83	11.90	80.76	0.49*
	MULTISPATI-ACP <sup>(3)</sup>				
	CPs 1	3.70	68.33	68.33	0.90*
	CPs 2	1.03	9.96	78.29	0.47*
	CPs 3	0.77	8.98	87.27	0.56*
La Nueva Trinidad I Entre Ríos	ACP <sup>(2)</sup>				
	CP1	2.78	30.91	30.91	0.76*
	CP2	2.13	23.69	54.60	0.88*
	CP3	1.23	13.76	68.36	0.79*
	MULTISPATI-ACP <sup>(3)</sup>				
	CPs 1	2.59	33.99	33.99	0.87*
	CPs 2	2.29	26.29	60.28	0.76*
	CPs 3	1.22	14.94	75.22	0.81*
La Nueva Trinidad II Entre Ríos	ACP <sup>(2)</sup>				
	CP1	2.89	36.21	36.21	0.81*
	CP2	1.67	21.01	57.22	0.68*
	CP3	1.12	14.03	71.25	0.59*
	MULTISPATI-ACP <sup>(3)</sup>				
	CPs 1	2.85	44.44	44.44	0.84*
	CPs 2	1.63	21.96	66.40	0.72*
	CPs 3	1.04	11.75	78.15	0.60*

<sup>(1)</sup> IM = Índice de Moran<sup>(2)</sup> ACP = Análisis de Componentes Principales<sup>(3)</sup> MULTISPATI-ACP = Análisis de Componentes Principales Espaciales\* Nivel de significancia  $P < 0.05$

**Tabla 3.65** Peso de las covariables ambientales (vectores propios) de los lotes de la zona 1 (Buenos Aires) en los tres primeros componentes del análisis de componentes principales (ACP) y MULTISPATI-ACP. Los coeficientes mas importantes se encuentran subrayados.

Campo	Variables	Elevación	Tosca	CEa <sub>30</sub>	CEa <sub>90</sub>	Trigo 2008 <sup>(1)</sup>	Soja 2007/2008 <sup>(1)</sup>	NDVI-19/11/2009 <sup>(2a)</sup>	NDVI-22/01/2010 <sup>(2a)</sup>	NDVI-04/01/2015 <sup>(2b)</sup>	NDVI-29/01/2015 <sup>(2b)</sup>
<b>Aleluya</b>											
<b>ACP<sup>(3)</sup></b>											
Buenos Aires	CP1	-0.26	<u>0.49</u>	-0.45	-0.50	0.27	0.23	-0.03	0.22	-0.19	-0.01
	CP2	-0.22	-0.10	-0.14	0.01	-0.32	-0.46	-0.35	0.07	-0.46	<u>-0.50</u>
	CP3	0.48	-0.19	-0.16	-0.16	-0.30	0.18	<u>-0.49</u>	0.40	-0.05	0.36
<b>MULTISPATI-ACP<sup>(4)</sup></b>											
	CPs1	0.36	<u>-0.55</u>	0.40	0.45	-0.29	-0.20	-0.03	-0.16	0.17	0.03
	CPs2	0.56	0.03	-0.06	-0.17	0.05	0.47	-0.01	0.20	0.24	<u>0.57</u>
	CPs3	0.22	-0.21	-0.36	-0.20	-0.35	-0.14	<u>-0.54</u>	0.18	-0.50	-0.01
<b>La Beatriz</b>											
<b>ACP<sup>(2)</sup></b>											
Buenos Aires	CP1	0.08	<u>0.48</u>	-0.47	-0.45	0.21	0.44	-0.26	-0.26		
	CP2	<u>0.62</u>	-0.22	-0.26	-0.45	-0.25	-0.26	0.39			
	CP3	0.30	0.21	0.25	0.05	<u>-0.68</u>	0.00	-0.57			
<b>MULTISPATI-ACP<sup>(3)</sup></b>											
	CPs1	0.07	<u>0.58</u>	-0.33	-0.32	0.21	0.50	-0.36			
	CPs2	<u>0.80</u>	-0.06	-0.16	-0.40	-0.31	-0.11	0.21			
	CPs3	0.00	0.13	0.21	0.12	<u>-0.71</u>	-0.09	-0.63			

<sup>(1)</sup> Mapa de rendimiento

<sup>(2)</sup> NDVI = Índice de Vegetación de Diferencia Normalizada obtenido a partir de imágenes Landsat 5 T<sup>(2a)</sup>, Landsat 8<sup>(2b)</sup> y SPOT 5<sup>(2a)</sup>.

<sup>(3)</sup> ACP = Análisis de Componentes Principales

<sup>(4)</sup> MULTISPATI-ACP = Análisis de Componentes Principales Espaciales

\* Nivel de significancia  $P < 0.05$

**Tabla 3.7** Peso de las covariables ambientales (vectores propios) de los lotes de la zona 2 (Córdoba) en los tres primeros componentes del análisis de componentes principales (ACP) y MULTISPATI-ACP. Los coeficientes más importantes se encuentran subrayados.

Campo	Variables	Elevación	CEa <sub>30</sub>	CEa <sub>90</sub>	Maíz 2005/2006 <sup>(1)</sup>	Soja 2009/2010 <sup>(1)</sup>	Soja 2011/2012 <sup>(1)</sup>	NDVI- 16/02/2006 <sup>(2a)</sup>	NDVI- 06/01/2014 <sup>(2b)</sup>	NDVI- 25/01/2014 <sup>(2c)</sup>
Mercapire Córdoba	ACP <sup>(2)</sup>									
	CPI	<u>0.39</u>	-0.33	<u>-0.39</u>	-0.27	-0.38	-0.38	-0.33	-0.14	-0.23
	CP2	-0.07	0.40	0.23	-0.29	-0.19	-0.19	0.35	<u>-0.66</u>	-0.19
Córdoba	CP3	0.05	0.34	0.11	<u>-0.57</u>	0.15	0.15	-0.14	0.47	-0.49
	MULTISPATI-ACP <sup>(3)</sup>									
	CPs1	<u>0.42</u>	-0.32	-0.39	-0.30	-0.37	-0.37	-0.35	-0.13	-0.21
San Esteban Córdoba	CPs2	-0.03	0.50	0.25	<u>-0.64</u>	-0.09	-0.09	0.24	-0.36	-0.21
	CPs3	0.28	0.21	-0.16	-0.42	<u>0.43</u>	<u>0.43</u>	-0.39	0.35	0.09
	MULTISPATI-ACP <sup>(3)</sup>									
San Esteban Córdoba	ACP <sup>(2)</sup>									
	CPI	<u>0.45</u>	0.31	-0.38	0.43	0.37	-0.18	0.42		
	CP2	0.03	0.46	-0.05	0.10	-0.31	<u>0.81</u>	0.09		
Córdoba	CP3	0.28	-0.55	<u>-0.60</u>	-0.26	-0.34	0.14	0.19		
	MULTISPATI-ACP <sup>(3)</sup>									
	CPs1	<u>-0.48</u>	-0.31	0.39	-0.44	-0.36	0.13	-0.41		
Córdoba	CPs2	0.10	<u>-0.69</u>	0.02	-0.10	0.40	-0.57	-0.01		
	CPs3	0.32	-0.24	<u>-0.59</u>	-0.42	-0.51	0.04	0.17		

<sup>(1)</sup> Mapa de rendimiento

<sup>(2)</sup> NDVI = Índice de Vegetación de Diferencia Normalizada obtenido a partir de imágenes Landsat 5 T<sup>(2a)</sup>, Landsat 8<sup>(2b)</sup> y SPOT 5<sup>(2c)</sup>.

<sup>(3)</sup> ACP = Análisis de Componentes Principales

<sup>(3)</sup> MULTISPATI-ACP = Análisis de Componentes Principales Espaciales

\* Nivel de significancia  $P < 0.05$

**Tabla 3.8** Peso de las covariables ambientales (vectores propios) de los lotes de la zona 3 (Entre Ríos) en los tres primeros componentes del análisis de componentes principales (ACP) y MULTISPATI-ACP. Los coeficientes más importantes se encuentran subrayados.

Campo	Variables	Elevación	CEa <sub>30</sub>	CEa <sub>90</sub>	Trigo 2009 <sup>(1)</sup>	Maíz 2011/2012 <sup>(1)</sup>	NDVI-10/01/2009 <sup>(2a)</sup>	NDVI-16/01/2011 <sup>(2a)</sup>	NDVI-16/12/2014 <sup>(2b)</sup>
La Nueva	ACP <sup>(2)</sup>								
Trinidad I	CP1	-0.22	-0.36	<u>-0.41</u>	0.36	<u>0.41</u>	0.40	0.24	0.30
Entre Ríos	CP2	<u>-0.55</u>	-0.31	0.10	-0.02	-0.33	-0.18	0.28	0.08
	CP3	0.09	-0.25	-0.31	<u>-0.52</u>	-0.26	0.27	0.38	-0.45
	MULTISPATI-ACP <sup>(3)</sup>								
	CPs1	<u>-0.49</u>	-0.47	-0.30	0.34	0.17	0.21	0.32	0.30
	CPs2	0.35	0.08	-0.33	0.24	0.46	0.30	-0.13	0.12
	CPs3	0.05	-0.32	-0.39	<u>-0.55</u>	-0.22	0.17	0.35	-0.38
La Nueva	ACP <sup>(2)</sup>								
Trinidad II	CP1	-0.40	-0.15	0.06	-0.41	<u>-0.51</u>	-0.35	-0.48	
Entre Ríos	CP2	-0.27	-0.09	<u>-0.68</u>	-0.11	0.09	-0.27	0.23	
	CP3	0.21	<u>0.79</u>	-0.16	-0.06	-0.17	-0.43	0.14	
	MULTISPATI-ACP <sup>(3)</sup>								
	CPs1	-0.46	-0.16	-0.05	-0.42	<u>-0.49</u>	-0.37	-0.42	
	CPs2	-0.11	0.19	<u>-0.73</u>	-0.07	0.12	-0.34	0.35	
	CPs3	<u>-0.54</u>	-0.43	0.07	0.47	0.22	0.14	-0.11	

<sup>(1)</sup> Mapa de rendimiento

<sup>(2)</sup> NDVI = Índice de Vegetación de Diferencia Normalizada obtenido a partir de imágenes Landsat 5 T<sup>(2a)</sup> y Landsat 8<sup>(2b)</sup>

<sup>(2)</sup> ACP = Análisis de Componentes Principales

<sup>(3)</sup> MULTISPATI-ACP = Análisis de Componentes Principales Espaciales

\* Nivel de significancia  $P < 0.05$

Para los lotes de la zona 1, tosca y CEa fueron las covariables ambientales con mayor peso en las CPs1. En el lote *Aleluya*, NDVI presentó mayor peso en las CPs2 y CPs3, mientras que en *La Beatriz* presentó mayor peso en elevación y rendimiento (Tabla 3.6).

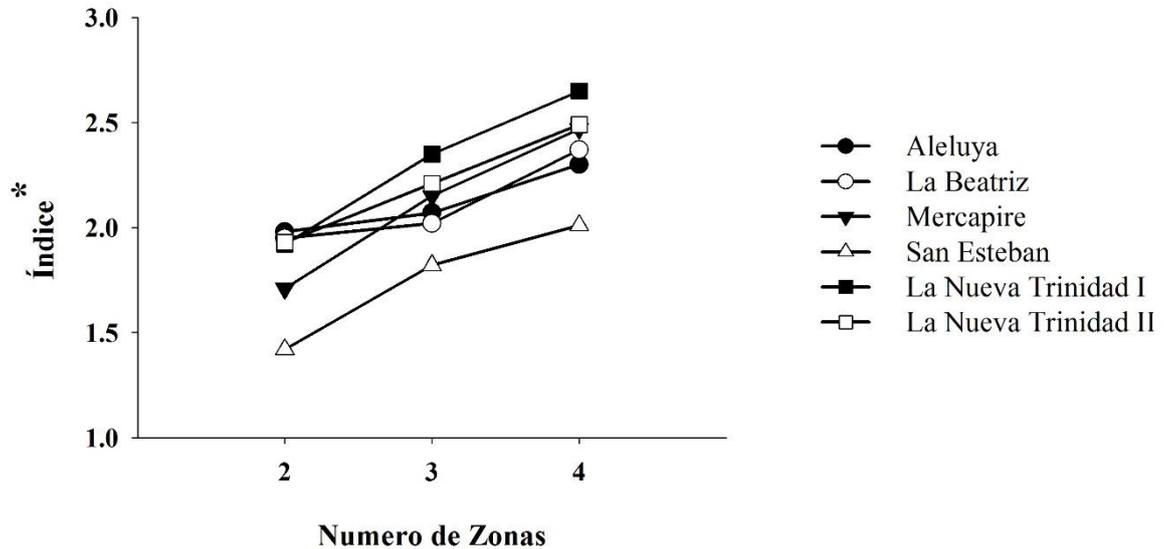
Para los lotes de la zona 2, elevación, CEa y mapas de rendimiento, en ese orden, presentaron mayor peso en las CPs1. Por su parte, mapas de rendimiento presentaron mayor peso en las CPs2 y CPs3 (Tabla 3.7).

En *La Nueva Trinidad I* de la zona 3, elevación presentó mayor peso en las CPs1, NDVI\_12/2013 en las CPs2 y Trigo\_2009 en las CPs3. Por su parte, en *La Nueva Trinidad II*, mapas de rendimiento seguidos por NDVI\_12/2014 presentaron mayor peso en las CPs1, CEa90 en las CPs2 y elevación seguida por Trigo\_2009 en las CPs3 (Tabla 3.8).

Los resultados confirman lo reportado por Córdoba *et al.* (2013) y Balzarini *et al.* (2011), quienes determinaron que los CPs obtenidos a través de MULTISPATI-ACP presentan mayor autocorrelación espacial con respecto a los CPs obtenidos con ACP. Puntualmente, MULTISPATI-ACP permitió identificar las covariables ambientales que más aportan a la variación espacial global y que tienen mayor autocorrelación espacial. Además, MULTISPATI-ACP se enfocó en la varianza multidimensional que tiene estructura espacial. Por el contrario, ACP puede ser ineficaz para detectar la estructura espacial, si esta estructura no está asociada con la varianza (Dray *et al.*, 2008).

### 3.3.2. Etapa I: Delimitación de zonas a partir de los CPs

La figura 3.10 muestra la variación del índice resumen sobre el número potencial de zonas determinadas a partir del análisis CFK de los CPs. El número de zonas para cada lote se determinó a partir del índice resumen. Índices resumen bajos indican mayor diferencia entre los clúster y menor diferencia entre datos dentro de cada clúster. De acuerdo con esto, el número de zonas que presentó el menor índice resumen fue elegido para cada lote (Lark y Stafford, 1997). En definitiva, para los lotes de la zona 1 y para *Mercapire* de la zona 2, se determinaron 3 zonas, mientras que para los demás lotes, se determinaron 2 zonas.



\*. Índice resumen a partir de la distancia euclídea de los siguientes índices:

Índice de coeficiente de partición (*fuzziness performance index-FPI*),  
 Índice de entropía de clasificación (*normalized classification entropy-NCE*)  
 Índice de Xie-Beni  
 Índice Fukuyama  
 Índice exponente de proporción

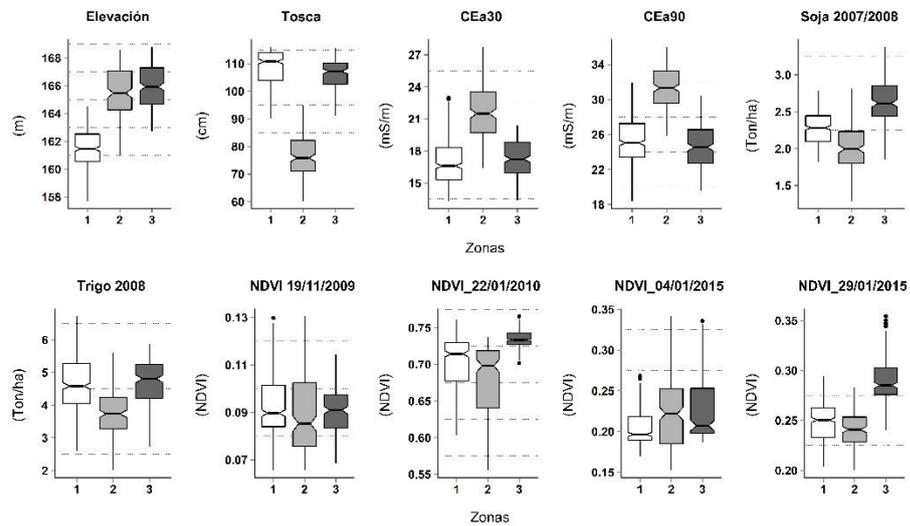
**Figura 3.10** Índice resumen utilizado para evaluar el número óptimo de zonas por tipo de suelos en las seis lotes de las tres zonas

La figura 3.11 muestra los gráficos de caja de cada covariable ambiental en cada zona, para los lotes de Buenos Aires o zona 1.

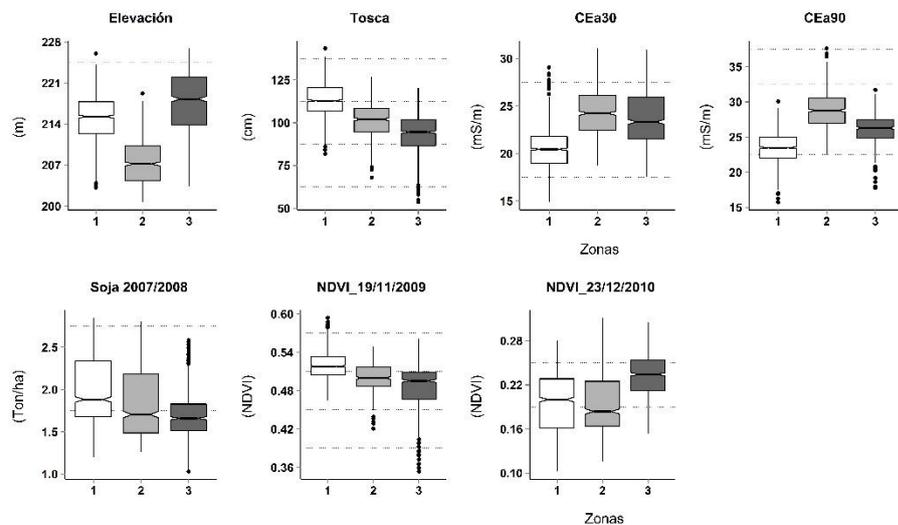
En general, tosca y CEa fueron las covariables que más aportaron en la construcción de zonas delimitadas en ambos lotes de Buenos Aires. Sin embargo, las relaciones entre estas covariables y las zonas delimitadas no fueron homogéneas. En *Aleluya*, tosca profunda y valores bajos de CEa definieron las zonas delimitadas 1 y 3. Por el contrario, la zona 2 fue definida por tosca somera y valores altos de CEa. En *La Beatriz*, tosca profunda y valores bajos de CEa definieron únicamente la zona delimitada 1, mientras que tosca somera y valores medios de CEa definieron la zona delimitada 3. En ambos lotes, NDVI y mapas de rendimiento se definieron en cada zona delimitada de acuerdo a la circunstancia ambiental predecesora que los influyó. Teniendo en cuenta el peso que

tuvieron CEa y tosca en la construcción del CPs1 y que a su vez, este fue el que mayor varianza explicó de todos los CPs, estos resultados eran esperables. Los mismos coinciden con lo reportado por Peralta *et al.* (2015) y Peralta *et al.* (2013), quienes determinaron que CEa y tosca en un análisis multivariado, son las covariables ambientales que definen la delimitación de zonas para manejo sitio específico en condiciones del sudeste de la provincia de Buenos Aires.

### Aleluya - Buenos Aires



### La Beatriz - Buenos Aires



**Figura 3.11** Variabilidad de las covariables ambientales en cada zona delimitada en los lotes de la zona 1 (Buenos Aires)

La figura 3.12 muestra los gráficos de caja de cada covariable ambiental en cada zona, para los lotes del sur de Córdoba o zona 2.

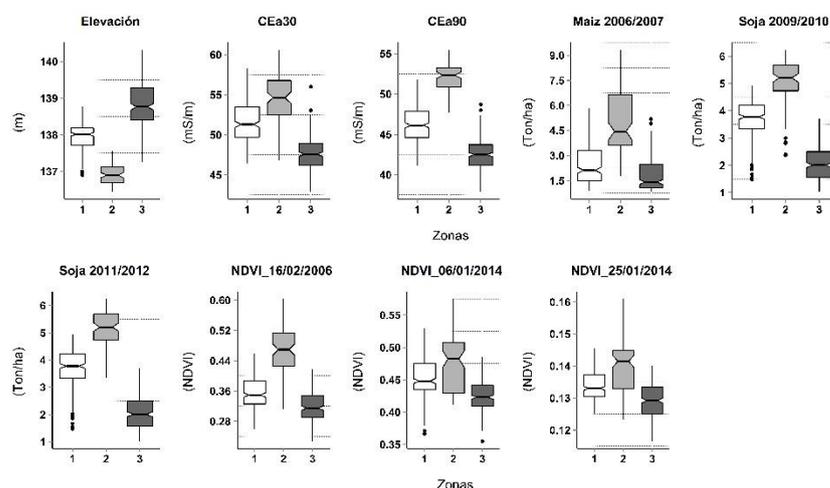
En *Mercapire*, todas las variables tuvieron un comportamiento particular en cada zona delimitada. La zona delimitada 2 estuvo determinada por zonas bajas, CEa alta, rendimientos altos y NDVI altos. La zona delimitada 3 fue lo contrario a la zona delimitada 2. Por su parte, la zona delimitada 1 estuvo determinada por valores medios de todas las covariables. Teniendo en cuenta el peso que tuvieron CEa, elevación y mapas de rendimiento en los CPs, estos resultados eran esperables. Por su parte, en *San Esteban*, la zona delimitada 1 estuvo determinada por zonas bajas, CEa<sub>30</sub> alta, CEa<sub>90</sub> baja, rendimiento bajo y NDVI bajo. La zona delimitada 2 fue lo contrario de la zona delimitada 1. El contraste entre CEa<sub>30</sub> y CEa<sub>90</sub> se explica por la presencia de sales a una profundidad mayor a 30 cm, en zonas bajas. Algunos trabajos han utilizado CEa para la delimitación de zonas para manejo sitio específico en condiciones de la zona (Simón *et al.*, 2012; Simón *et al.*, 2013). Sin embargo, estos trabajos no han incluido ningún tipo de análisis multivariado que incluya simultáneamente elevación, mapas de rendimiento y NDVI. En ese aspecto, los resultados reportados en esta tesis son novedosos y pueden ser importantes para la delimitación de zonas de manejo sitio-específico o para aplicar otro tipo de técnicas de cartografía digital de suelos a escala de lote.

La figura 3.13 muestra los gráficos de caja de cada covariable ambiental en cada zona, para los lotes de la zona 3.

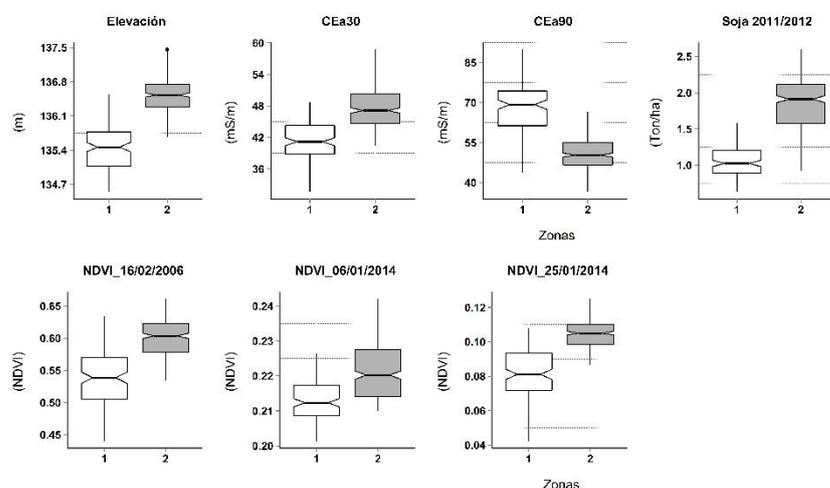
En *LNT\_I*, la zona delimitada 1 estuvo determinada por zonas bajas, CEa baja, rendimiento alto y NDVI alto. La zona delimitada 2 fue lo contrario de la zona delimitada 1. Por su parte *en LNT\_II*, la zona delimitada 1 estuvo determinada por zonas altas, CEa<sub>30</sub> alta, rendimiento alto y NDVI alto. En general, elevación y CEa<sub>30</sub> determinaron de manera contradictoria las zonas delimitadas 1 y 2 en *LNT\_I* y *LNT\_II*. En ambos lotes, el patrón espacial de la presencia de sales, alteró, especialmente, el patrón espacial de la CEa<sub>90</sub>. Al respecto, Moral *et al.* (2010) estableció que en un análisis multivariado que tenga como objetivo la delimitación de zonas, la CEa es la covariable que define las propiedades del suelo limitantes del rendimiento. Sin embargo, los resultados de este trabajo, demuestran

que el efecto del patrón espacial del alto contenido de sales, es decir, de una propiedad del suelo muy influyente de la CEa, no permite determinar con claridad las relaciones espaciales entre CEa y las demás covariables ambientales. Por el contrario, elevación, rendimiento y NDVI demostraron aportar más a la delimitación de zonas mediante la metodología CPs\_CFK, en comparación con la CEa. De este modo, se concluye que en presencia de alto contenido de sales, es posible que la CEa pueda perder su potencial de delimitación de zonas por tipo de suelos. Sin embargo, delimitar zonas con altos contenidos de sales, puede llegar a ser útil para la aplicación de agricultura de precisión, mas no para generar y desagregar unidades cartográficas de suelo.

### Mercapire - Córdoba

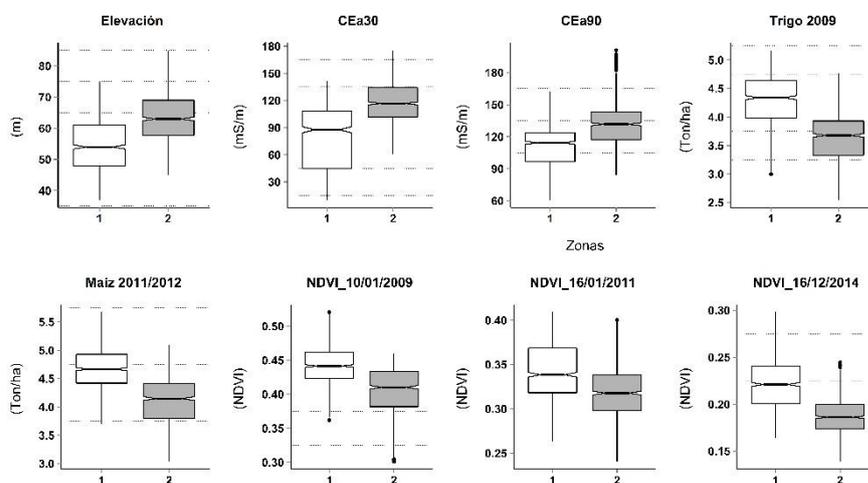


### San Esteban - Córdoba

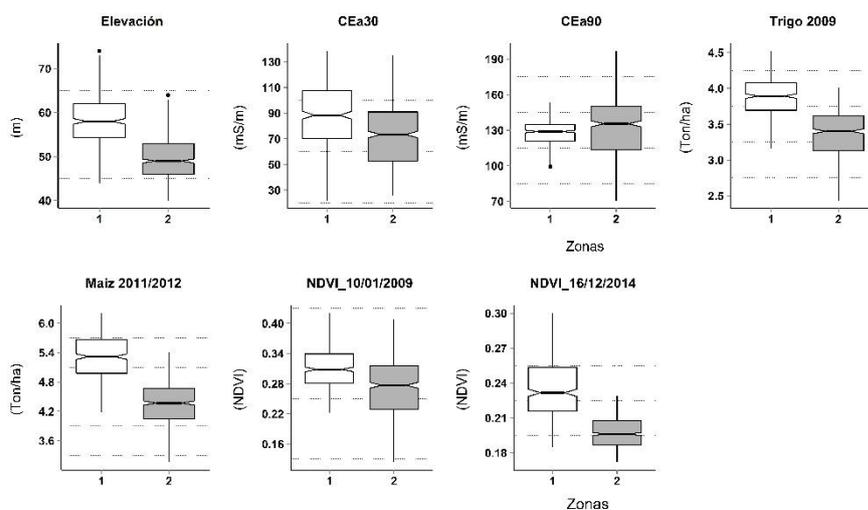


**Figura 3.12** Variabilidad de las covariables ambientales en cada zona delimitada en los lotes de la zona 2 (Córdoba)

### La Nueva Trinidad I - Entre Ríos



### La Nueva Trinidad II - Entre Ríos

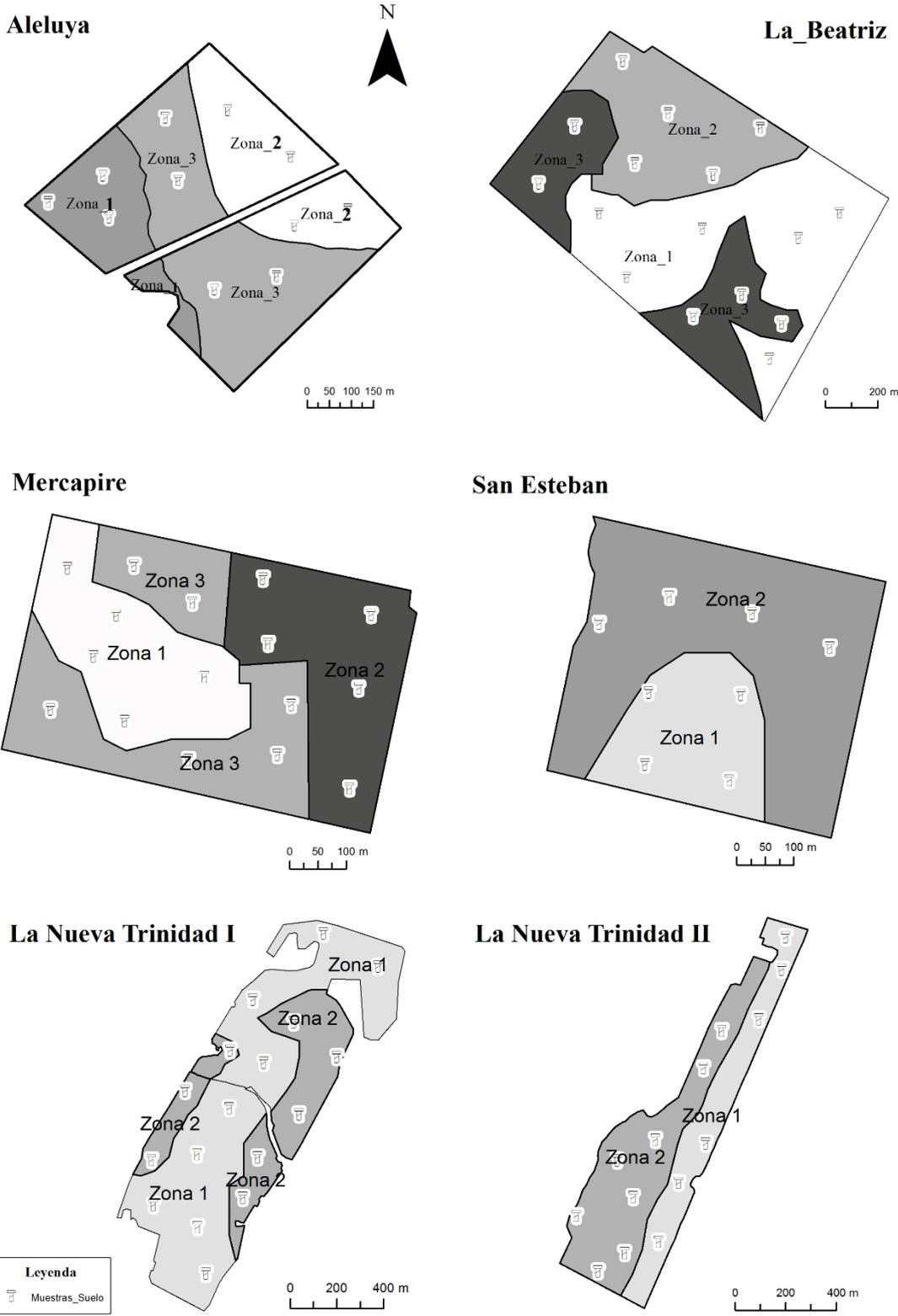


**Figura 3.13** Variabilidad de las covariables ambientales en cada zona delimitada en los lotes de la zona 3 (Entre Ríos)

#### 3.3.3. Etapa II: Validación de la delimitación de zonas

La figura 3.14 muestra la localización de las zonas delimitadas por lote y puntos de muestreo de suelos por zona delimitada.

La figura 3.15 muestra la respuesta de cada propiedad del suelo analizada en función de las zonas delimitadas en cada lote de la zona 1.



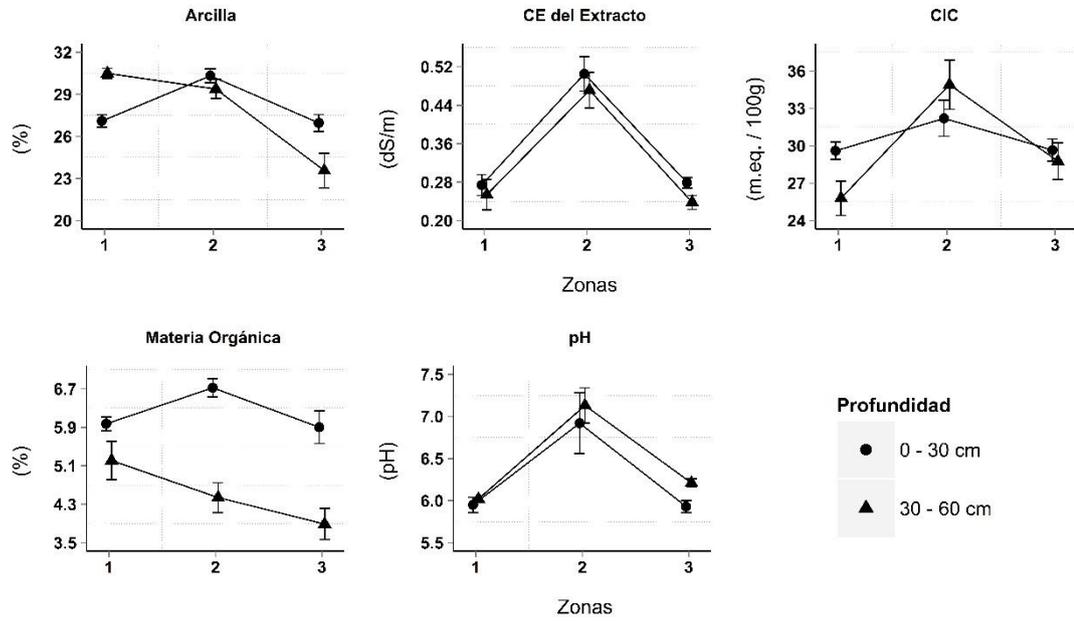
**Figura 3.14** Distribución espacial de las muestras de suelos en cada zona delimitada de todos los lotes en las tres zonas

En *Aleluya*, valores altos de arcilla, CE del extracto, CIC y pH entre 0 – 30 cm y 30 – 60 cm de profundidad estuvieron asociados con tosca somera y CEa alta (*zona delimitada 2*). De la misma manera, valores altos de materia orgánica 0 – 30 cm estuvieron asociados con tosca somera y CEa alta. Sin embargo, valores altos de materia orgánica entre 30 – 60 cm estuvieron asociados a zonas bajas (*zona delimitada 1*). Excepto MO, los valores de las propiedades del suelo en la zona delimitada 3, tuvieron una asociación similar a los valores de las propiedades del suelo en la zona delimitada 1.

En *La Beatriz*, el patrón de valores de propiedades del suelo por zona delimitada presentó una tendencia similar a *Aleluya*. Excepto pH, valores bajos de todas las propiedades del suelo entre 0 – 30 cm de profundidad, estuvieron asociados a tosca profunda y CEa baja (*zona delimitada 1*). Excepto MO, valores bajos de todas las propiedades entre 30 – 60 cm de profundidad, estuvieron asociados a zonas bajas con tosca medianamente profunda y CEa alta (*zona delimitada 2*). Valores bajos de MO entre 30 – 60 cm estuvieron asociados a tosca profunda (*zona delimitada 1*). Valores altos de pH en ambas profundidades estuvieron asociados con tosca somera (*zona delimitada 3*).

El efecto que tiene la profundidad de tosca sobre el patrón espacial de las propiedades del suelo en condiciones del sudeste de la provincia de Buenos Aires ha sido poco estudiado. Cabria y Culot (1994) determinaron que la presencia de tosca, o un manto calcáreo cementado, tiene influencia sobre las propiedades del suelo que manifiestan el grado de desarrollo del horizonte B. Tosca somera a poca profundidad generalmente conlleva a horizontes B más delgados, más ricos en arcilla y con mayor proporción de mortmorillonita. Por su parte, Pazos (1984) determinó que la presencia de tosca inmediatamente por debajo del horizonte B<sub>2t</sub> provoca la formación de mayor cantidad de argilanes y de mayor espesor que en suelos similares pero sin capas limitantes. Estos efectos de la tosca sobre el contenido de arcilla, especialmente en el horizonte B, posiblemente determinaron cambios en sentidos similares en otras variables del suelo tales como capacidad de almacenamiento de agua, MO, CIC y pH (Doolittle *et al.*, 1994; Heil y Schmidhalter, 2012; Piikki *et al.*, 2013).

### Aleluya - Buenos Aires



### La Beatriz - Buenos Aires

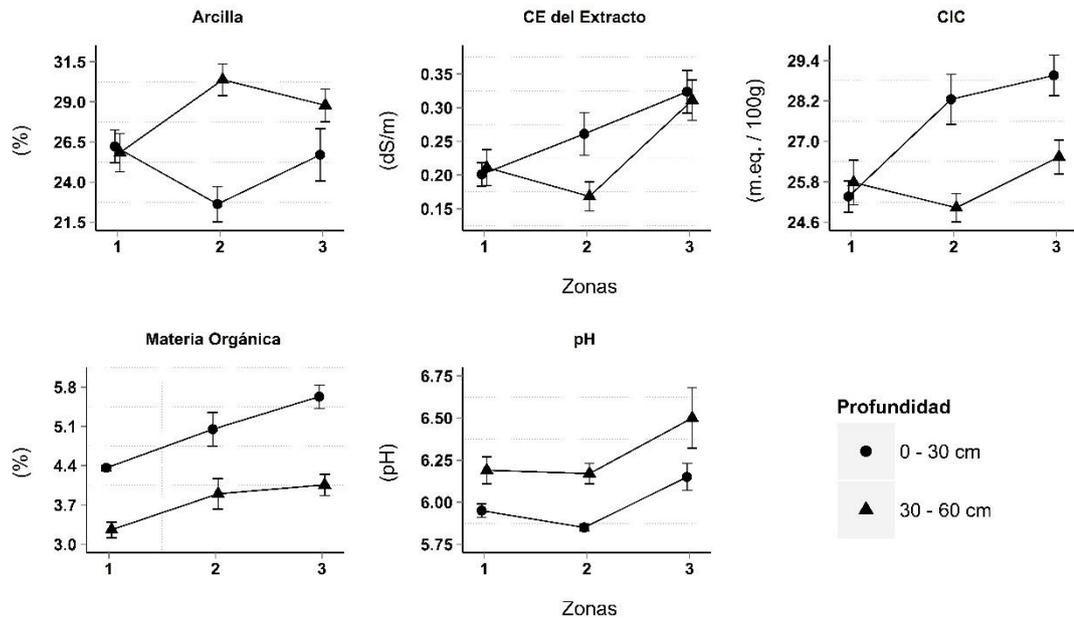


Figura 3.15 Medias de propiedades del suelo en cada zona delimitada en los lotes de la zona 1 (Buenos Aires). Barras verticales indican error estandar de la media para cada una de las zonas delimitadas

En condiciones de la zona 1, la metodología CPs\_CFK logró diferenciar eficientemente zonas de acuerdo a las propiedades del suelo analizadas. Contenido de arcilla,  $CE_{ext}$ , CIC y pH fueron las propiedades del suelo determinantes en la delimitación de zonas. El patrón espacial de estas propiedades depende del patrón espacial de la profundidad de tosca. Es importante que en futuros trabajos se incluya estudiar el efecto de la variabilidad espacial vertical y horizontal de la profundidad de tosca, sobre el patrón espacial de propiedades del suelo.

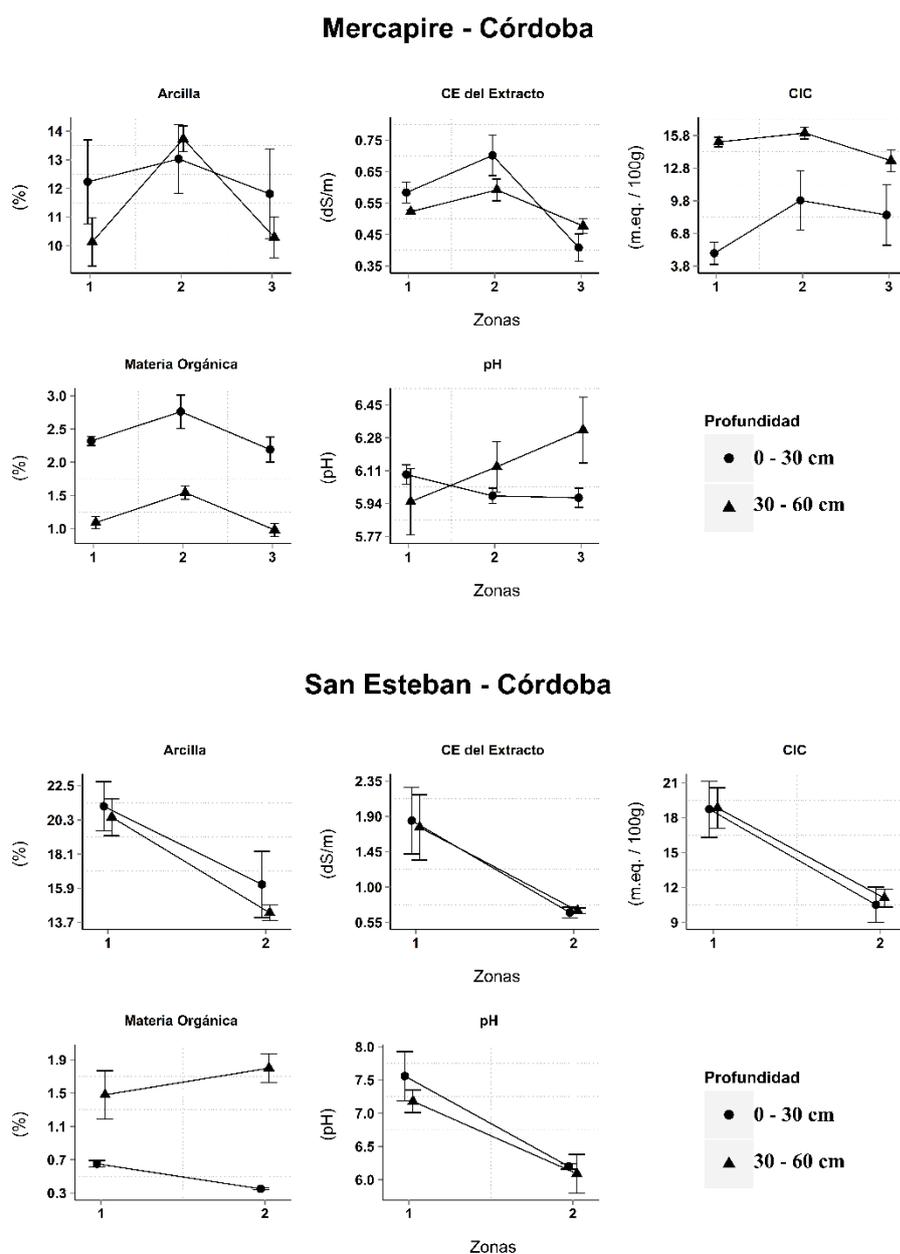
La figura 3.16 muestra la respuesta de cada propiedad del suelo analizada en función de las zonas delimitadas en cada lote de la zona 2.

En *Mercapire*, valores bajos de arcilla, CIC y MO entre 0 – 30 cm de profundidad, estuvieron asociados con zonas medias de elevación, CEa, rendimiento y NDVI (*zona delimitada 1*). Excepto pH, valores altos de todas las propiedades del suelo entre 0 – 30 cm y 30 – 60 cm de profundidad, estuvieron asociados con zonas bajas, CEa alta, rendimiento alto y NDVI alto (*zona delimitada 2*). Valores altos de pH entre 30 – 60 cm de profundidad, estuvieron asociados con zonas altas, CEa baja, rendimiento bajo y NDVI bajo (*zona delimitada 3*).

En *San Esteban*, valores bajos de MO entre 0 – 30 cm de profundidad y valores altos de MO entre 30 – 60 cm, estuvieron asociados a zonas bajas, rendimiento bajo y NDVI bajo (*zona delimitada 1*). Excepto MO, valores bajos de todas las propiedades del suelo entre 0 – 30 cm y 30 – 60 cm de profundidad, estuvieron asociados con zonas altas, CEa<sub>30</sub> alta, CEa<sub>90</sub> baja, rendimiento alto y NDVI alto (*zona delimitada 2*).

El patrón de variabilidad espacial de propiedades del suelo a escala de lote en condiciones del sur de la provincia de Córdoba ha sido poco estudiado. Los resultados de esta tesis ratifican lo reportado por Peralta *et al.* (2013) y Simón *et al.* (2013). Estos autores determinaron que zonas bajas estuvieron relacionadas con zonas con CEa<sub>90</sub> altas, lo cual coincidió plenamente con los resultados de esta tesis. Así mismo, encontraron que valores altos de MO y valores altos de pH y CIC estuvieron relacionados con zonas con CEa<sub>90</sub> baja y alta, respectivamente, lo cual también coincidió plenamente con los

resultados de esta tesis. En relación con valores de propiedades del suelo y elevación, los resultados también ratifican lo reportado por Peralta *et al.* (2013). Por su parte, Simón *et al.* (2013) encontraron que valores altos de arcilla, MO, pH y CE del extracto estuvieron asociados a zonas de CEa<sub>90</sub> alta.



**Figura 3.16** Medias de propiedades del suelo en cada zona delimitada en los lotes de la zona 2 (Córdoba). Barras verticales indican error estándar de la media para cada una de las zonas delimitadas

En condiciones de la zona 2, la metodología CPs\_CFK logró diferenciar eficientemente zonas de acuerdo a los patrones de las propiedades del suelo analizadas. Además, se ratificó que el contenido de arcilla, CE extracto y CIC fueron las propiedades del suelo determinantes en la delimitación de zonas. Es ampliamente conocido que el patrón espacial de estas propiedades está determinado por el patrón espacial de la elevación, lo cual es consistente con los CPs. Por lo tanto, la CEa correspondió espacialmente a estos patrones (Corwin *et al.*, 2003; Friedman, 2005; Sudduth *et al.*, 2005). MO y pH presentaron patrones contradictorios. Ambas propiedades son susceptibles a cambios en sus patrones espaciales debido a factores temporales y al tipo de manejo de suelos y cultivos (Bogaert y D'Or, 2002; Kravchenko y Bullock, 1999).

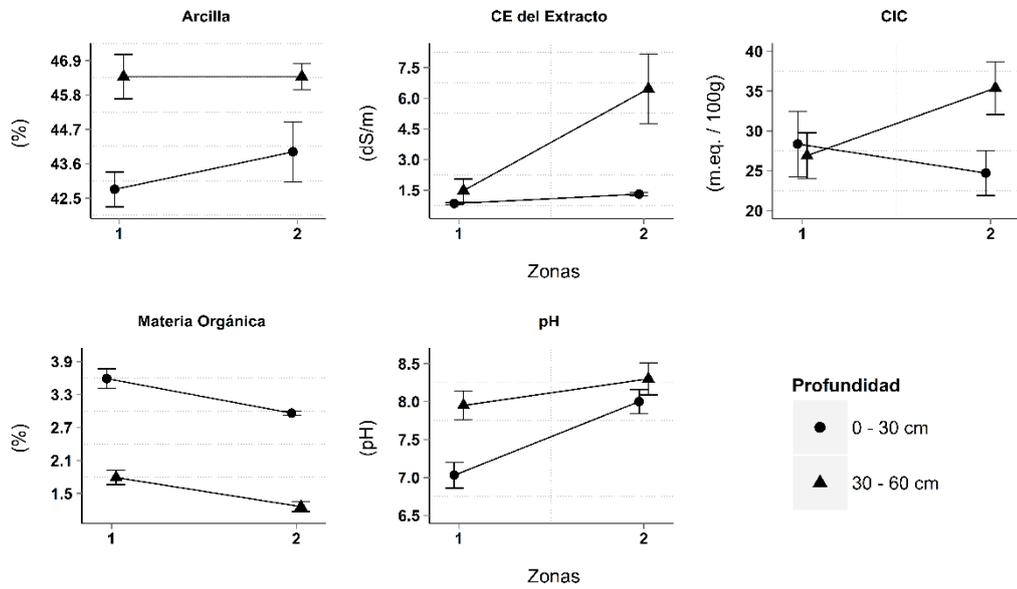
La figura 3.17 muestra la respuesta de cada propiedad del suelo analizada en función de las zonas delimitadas en cada lote de la zona 3.

En *LNT\_I*, valores bajos de contenido de arcilla y pH entre 0 – 30 cm de profundidad estuvieron asociados con zonas bajas, CEa baja, rendimientos altos y NDVI altos (*zona delimitada 1*). Por su parte, valores altos de CE<sub>ext</sub>, CIC y pH entre 30 – 60 cm de profundidad, estuvieron asociados con zonas altas, CEa alta, rendimientos bajos y NDVI bajos (*zona delimitada 2*). Valores altos de materia orgánica entre 0 – 30 cm y 30 – 60 cm de profundidad, tuvieron tendencia a presentarse en zonas bajas, CEa baja, rendimientos altos y NDVI altos (*zona delimitada 1*). CE<sub>ext</sub> entre 0 – 30 cm de profundidad no presentó diferencia significativa entre zonas delimitadas.

En *LNT\_II*, valores altos de contenido de arcilla, CIC y MO entre 0 – 30 cm y 30 – 60 cm de profundidad, estuvieron asociados con zonas altas, CEa<sub>30</sub> alta, rendimientos altos y NDVI altos (*zona delimitada 1*). Por su parte, valores altos de CE<sub>ext</sub> entre 0 – 30 cm y 30 – 60 cm estuvieron asociados a zonas bajas, CEa<sub>90</sub> alta, rendimientos bajos y NDVI bajos. Valores altos de pH entre 0 – 30 cm de profundidad, estuvieron asociados a zonas bajas, CEa<sub>30</sub> baja, rendimientos bajos y NDVI bajos (*zona delimitada 2*).

El patrón de variabilidad espacial de las propiedades del suelo a escala de lote en condiciones de la zona 3, indican la presencia de suelos salinos en la zona delimitada 2.

### La Nueva Trinidad I - Entre Ríos



### La Nueva Trinidad II - Entre Ríos

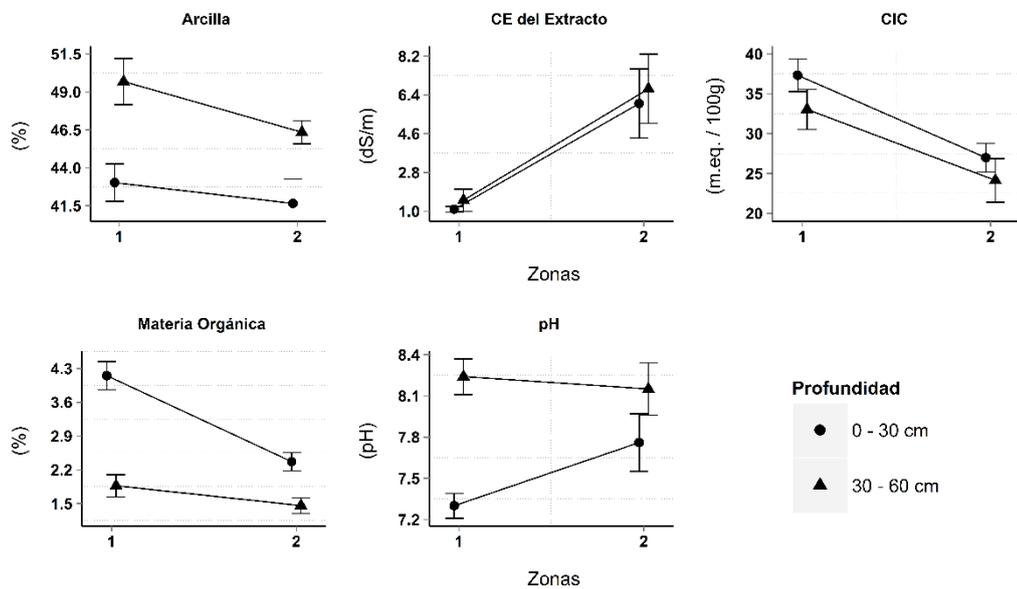


Figura 3.17 Medias de propiedades del suelo en cada zona delimitada en los lotes de la zona 3 (Entre Ríos). Barras verticales indican error estandar de la media para cada una de las zonas delimitadas

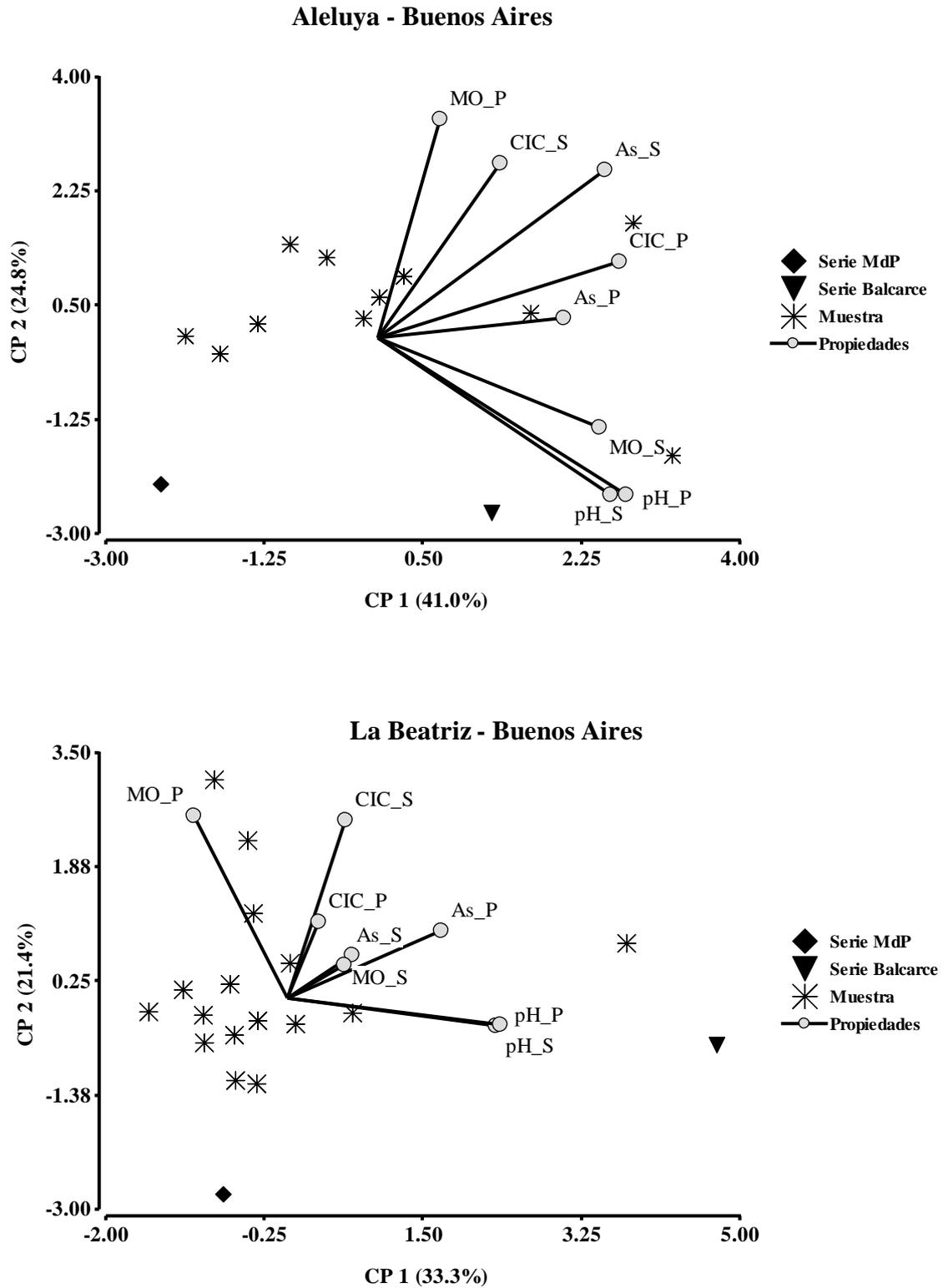
Este tipo de suelos se caracteriza por tener  $CE_{ext}$  mayor a 4 dS/m, pH menor a 8.5, porcentaje de sodio intercambiable (PSI) mayor al 15% y la relación de adsorción de sodio (RAS) mayor a 13 (Abrol *et al.*, 1988; Rhoades *et al.*, 1989b). Es ampliamente conocido que la concentración alta de sales afecta el crecimiento de las plantas. Debido a esto, en la zona delimitada 2 de ambos lotes o zonas con suelos salinos, se presentaron de manera temporalmente constante rendimientos y NDVI bajos. Esta situación tiene un efecto importante en el proceso de elaboración de cartografía digital de suelos. Actualmente, el efecto antrópico sobre el patrón espacial de suelos, determinan que se planteen nuevas formas de generar cartografía digital de suelos, en las cuales la dinámica espacio-temporal y tridimensional de las propiedades del suelo, deben ser incluidas (Castrignano *et al.*, 2002; Veronesi *et al.*, 2012).

Al igual que en la zona 1 y 2, la metodología CPs\_CFK logró diferenciar eficientemente zonas de acuerdo a los patrones espaciales de las propiedades de suelos analizadas en la zona 3. CE del extracto y sus relaciones espaciales con pH y capacidad de intercambio catiónico fueron las propiedades determinantes en la delimitación de zonas.

#### *3.3.4. Etapa III Relaciones entre propiedades del suelo: zonas delimitadas vs. unidades cartográficas de suelo*

Excepto en la zona 3, no se encontraron reportados valores de la  $CE_{ext}$  en la cartografía convencional de suelos disponible. Debido a esto,  $CE_{ext}$  no fue tomada en cuenta en la MDE y por ende, en el ACP. En la zona 3, a pesar de que se contaba con el dato de  $CE_{ext}$ , se decidió no tenerlo en cuenta debido a que la correspondencia actual del valor de  $CE_{ext}$  reportado, posiblemente está desactualizado.

La figura 3.18 muestra el *biplot* de análisis de componentes principales (ACP), elaborado a partir de los resultados de una Matriz de Distancia Euclidiana (MDE) entre propiedades del suelo analizadas y las propiedades de suelos de las series, en cada uno de los lotes de la zona 1.



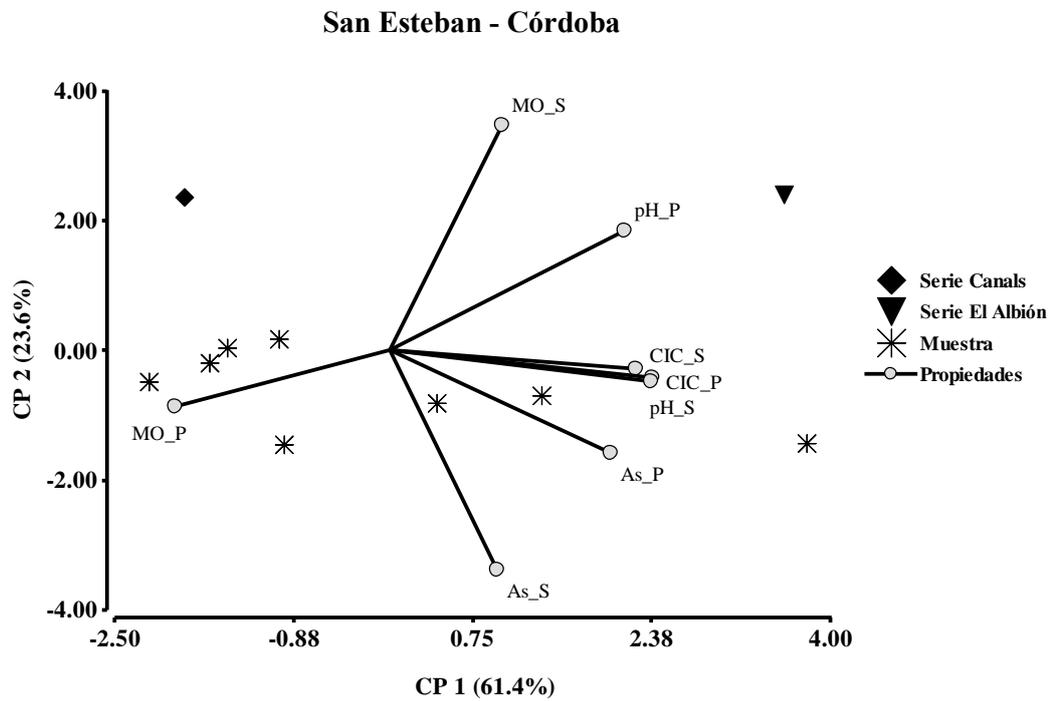
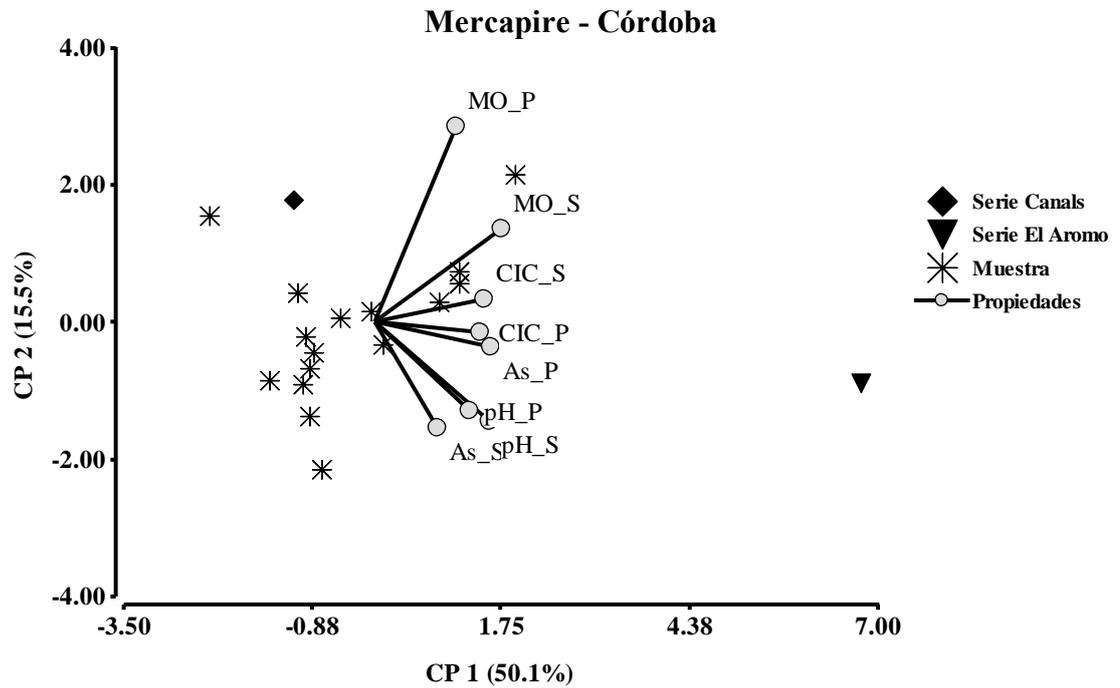
**Figura 3.18** Representación gráfica de los dos primeros ejes del análisis de componentes principales (ACP) elaborado a partir de los resultados de la Matriz de Distancia Euclidiana (MDE) entre propiedades del suelo analizadas y las propiedades de suelos de las series, para los lotes de la zona 1 (*Buenos Aires*)

En *Aleluya*, la zona delimitada 2 (33% de área total) podría estar en correspondencia espacial y edáfica con la serie Balcarce, mientras que la zona delimitada 3 (46% del área total) con la serie MdP. Las proporciones de área entre zonas delimitadas tipo serie de suelos Balcarce y serie MdP, corresponden con la descrita en la cartografía de suelos. Por su parte, la zona delimitada 1 (21% del área total) podría corresponder espacialmente a la unidad cartográfica de suelos CoAoC. Esta unidad está conformada en su totalidad por la serie Arroyo Coloquene, la cual fue clasificada como un complejo indiferenciado de suelos (INTA, 2010a). Sin embargo, la zona delimitada 1 tiende a tener una correspondencia edáfica con la zona delimitada 3. De acuerdo con esto, se podría asumir que en la zona delimitada 1, predominan Argiudoles típicos. Es posible que esta zona sea una transición de la serie MdP o la serie Tres esquinas.

En *La Beatriz*, la zona delimitada 1 (45% del área total) presentó correspondencia espacial y edáfica con la serie MdP, mientras que la zona delimitada 3 (27% del área total) con la serie Balcarce. Las proporciones de área entre zonas delimitadas tipo serie de suelos Balcarce y serie MdP, corresponden con la descrita en la cartografía de suelos. Por su parte, La zona delimitada 2 (28% del área total) presentó correspondencia espacial y edáfica con una transición entre las series MdP y Balcarce.

Los resultados para ambos lotes de la zona 1, ratifican lo planteado por Cabria y Culot (1994). Estos autores determinaron las características edáficas que diferencian las series Balcarce y MdP en condiciones del sudeste bonaerense. De acuerdo con ellos, la serie Balcarce tiene mayor CIC y una profundidad menor del techo del horizonte B, con respecto a la serie MdP. Ambas características, ratifican la relación entre las zonas delimitadas en ambos lotes y las series MdP y Balcarce.

La figura 3.19 muestra el *biplot* de análisis de componentes principales (ACP), elaborado a partir de los resultados de una Matriz de Distancia Euclidiana (MDE) entre propiedades del suelo analizadas y las propiedades de suelos de las series, en cada uno de los lotes de la zona 2.



**Figura 3.19** Representación gráfica de los dos primeros ejes del análisis de componentes principales (ACP) elaborado a partir de los resultados de la Matriz de Distancia Euclidiana (MDE) entre propiedades del suelo analizadas y las propiedades de suelos de las series de suelos, para los lotes de la zona 2 (Córdoba)

En *Mercapire*, todas las zonas delimitadas tienen correspondencia espacial y edáfica con la serie Canals. Sin embargo, las variaciones de las propiedades edáficas encontradas entre zonas delimitadas, plantea que estas son potenciales a ser incluidas en la elaboración de cartografía digital de suelos a escala de lote. Además, dichas variaciones de suelo tienen efecto sobre el rendimiento de los cultivos y sobre el NDVI.

En *San Esteban*, la zona delimitada 1 presentó correspondencia espacial con la serie Canals, mientras que la zona delimitada 2 con la serie El Aromo.

La figura 3.20 muestra el *biplot* de análisis de componentes principales (ACP), elaborado a partir de los resultados de una Matriz de Distancia Euclidiana (MDE) entre propiedades del suelo analizadas y las propiedades de suelos de las series de suelos, en cada uno de los lotes de la zona 3.

En *LNT\_I* y *LNT\_II*, las zonas delimitadas no presentaron correspondencia espacial y edáfica con las series de suelos de la cartografía disponible. Principalmente, estos resultados se deben a dos factores (*i*) al considerable efecto que tuvo la presencia de altos contenidos de sales sobre las covariables ambientales CEa, rendimiento y NDVI; y (*ii*) la diferencia de valores del contenido de arcilla entre las muestras analizadas para cada zona, y las reportadas para las series de suelos. Sin embargo, es pertinente remarcar que las condiciones actuales de manejo de suelos y los patrones temporales de precipitación, en gran medida, son los responsables de la degradación de los suelos en ambos lotes. A pesar del efecto de las sales, propiedades del suelo estables en el tiempo tales como arcilla, no presentaron relación alguna con la información de las series de suelos.

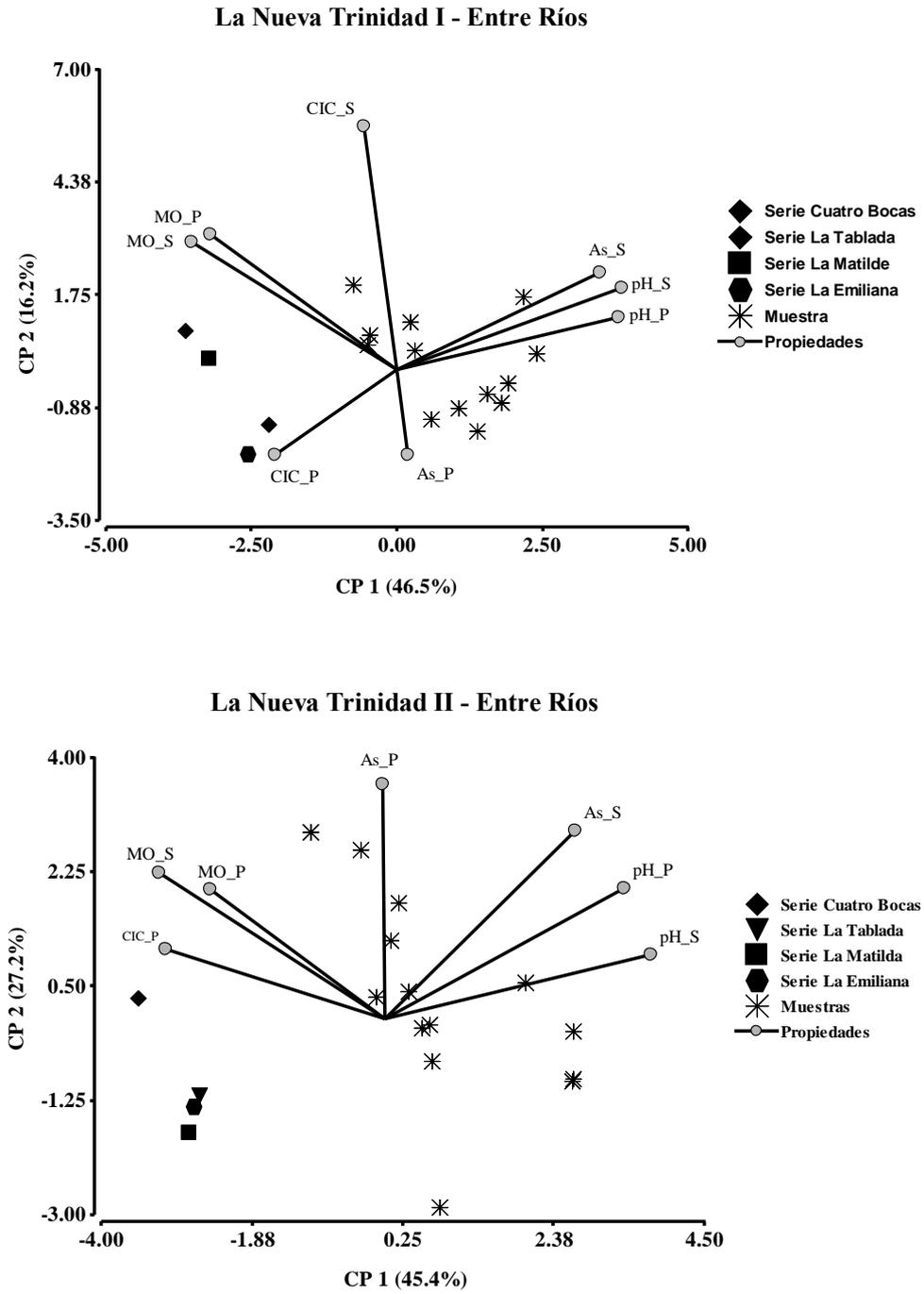
Para los lotes de la zona 1 y 2, la metodología de DgUCS propuesta fue eficiente para determinar el patrón espacial de los tipos de suelos y su relación con las series de suelos descritas en la cartografía convencional. A través de la aplicación de cada una de las etapas de la metodología de DgUCS se plantea un prospecto efectivo para mejorar la información de suelos a escala de lote, a partir del relevamiento de covariables ambientales en formato digital y la información disponible en la cartografía convencional de suelos INTA, en condiciones del sudeste de la provincia de Buenos Aires y el sur de

córdoba. Para las zonas 1 y 2, los resultados también confirman que la selección de covariables ambientales relacionados con la formación y definición de patrones espacio-temporales de propiedades del suelo, fue acertada. En contraste, en la zona 3 la metodología fue eficiente para determinar el patrón espacial de los tipos de suelos. Sin embargo, la relación de los resultados obtenidos con la información de las series de suelo no pudo ser definida. A pesar de esto, la metodología DgUCS probó ser un prospecto metodológico efectivo para la actualización y aumento de escala de la cartografía convencional disponible.

Las variaciones de propiedades del suelo en el tiempo como consecuencia del efecto antrópico o por procesos de degradación, puede tener un efecto importante en la aplicación de la metodología de DgUCS. Varias de las covariables propuestas son susceptibles a cambios en las propiedades de suelos. Para futuros trabajos se recomienda complementar las covariables del suelo con otras fuentes de información relacionada con los factores formadores del suelo, tales como índices de topografía, datos de georradar y otros tipos de índices de vegetación que incluya la banda del infrarrojo medio. Así mismo, se recomienda utilizar técnicas de morfometría digital de suelos para determinar las variaciones espaciales verticales de los suelos. Por último, la comparación de la metodología de DgUCS propuesta con otras metodologías, sería importante para determinar si existen algoritmos de clúster más eficientes para las condiciones de suelos de la pampa argentina (Nauman *et al.*, 2012; Odgers *et al.*, 2013).

#### **4. Conclusiones**

La sostenibilidad de los servicios sistémicos del suelo en Argentina depende de la información disponible de suelo a grandes escalas. Debido a esto, se propuso y discutió una técnica que permita generar información espacial de tipos de suelos de manera rápida, precisa y económica, a partir de la cartografía existente y de la medición de covariables ambientales.



**Figura 3.20** Representación gráfica de los dos primeros ejes del análisis de componentes principales (ACP) elaborado a partir de los resultados de la Matriz de Distancia Euclidiana (MDE) entre propiedades del suelo analizadas y las propiedades de suelos de las series de suelos, para los lotes de la zona 3 (*Entre Ríos*)

Específicamente, en este capítulo se presentó un algoritmo de desagregación de unidades cartográficas de suelo a escala de lote agrícola a partir de análisis de componentes principales y análisis de distancia euclidiana (DgUCS).

En general, la metodología propuesta demostró ser eficiente para determinar el patrón espacial de tipos de suelos a escala de lote. Sin embargo, las relaciones espaciales entre tipos de suelos a escala de lote y series de suelos descritas en las cartas de suelos fueron consistentes para el sudeste bonaerense y sur de Córdoba, pero no para el centro de Entre Ríos. En el sudeste bonaerense, se encontró que el patrón espacial de suelos a escala de lote depende de los patrones espaciales de tosca y elevación, principalmente. Por su parte, en el sur de Córdoba, se encontró que el patrón espacial de suelos a escala de lote depende de los patrones espaciales de elevación. Así mismo, para esta zona se ratificó que la CEa es una variable potencial para determinar tipos de suelos a escala de lote en condiciones de suelos del sur y sudeste de Córdoba. Por último, en el centro de Entre Ríos, el patrón temporal de precipitaciones intensas, junto con manejos inadecuados del suelo, hicieron que las variables estudiadas se encontraran seriamente afectadas. Esta situación alteró la medición de varias covariables ambientales. Así mismo, se encontró que la cartografía disponible y espacialmente correspondiente con los lotes del centro de Entre Ríos, no tenía información edáfica similar a la determinada por este estudio.

En general, el NDVI tuvo alta dependencia espacial de los patrones climáticos predominantes. Debido a esto, se sugiere utilizar el patrón espacial del NDVI de acuerdo a una escala temporal caracterizada, en la cual los patrones de precipitación y su intensidad, sean tenidos en cuenta. En este trabajo se ratificó el potencial que tiene la CEa para determinar los tipos de suelos a escala de lote. Esta variable fue útil para delimitar detalladamente, áreas afectadas con altos contenidos de sales. Así mismo, las relaciones entre los patrones espaciales de CEa y tosca, permiten determinar que CEa tiene un alto potencial para identificar los patrones espaciales de tosca a escala de lote. Con la metodología propuesta, se demostró el potencial que tienen los mapas de rendimiento para delimitar tipos de suelos y DgUCS.

La extrapolación de estos resultados en otros lotes dentro y fuera de las zonas estudiadas, requiere ser evaluado. Así mismo, otras fuentes de información de covariables ambientales deben ser incorporadas y evaluadas al proceso de DgUCS, utilizando la metodología propuesta. Estudios adicionales son requeridos para (i) comparar otros tipos de algoritmos de agrupamiento espacial de suelos; (ii) determinar la importancia de las covariables ambientales en la predicción y delimitación de tipos de suelos; (iii) validar las áreas desagregadas con modelos de simulación de crecimiento y desarrollo de cultivos; y (iv) evaluar la metodología propuesta a otras escalas espaciales.

## CAPITULO 4

### **Relación entre patrones espaciales de propiedades del suelo y cultivo de soja, a escala de lote**

#### **1. Introducción**

En la campaña 2014/2015, se estima que la producción argentina de soja fue de 60.1 millones de toneladas, producidas en 20.2 millones de has (Bolsa de Comercio de Rosario, 2015; SIIA, 2015). Esto equivale al 19.2% de la producción mundial de soja (USDA, 2015). Es ampliamente conocido que la soja es una importante fuente para satisfacer la demanda mundial de proteínas y aceite vegetal. Debido a esto, es importante entender los factores que limitan el rendimiento y la calidad de soja. Numerosos trabajos han determinado que los principales factores limitantes del rendimiento y la calidad de soja son (i) la oferta ambiental y (ii) el potencial genético de las semillas (Andrade y Sadras, 2002; Satorre et al., 2003). Sin embargo, limitados trabajos han contribuido a entender el efecto de los patrones espaciales del suelo sobre el patrón espacial del rendimiento y la calidad de soja, a escala de lote.

El efecto de los patrones espaciales de algunos factores formadores del suelo sobre el rendimiento y calidad de soja, han sido previamente estudiados (Kravchenko *et al.*, 2005; Martin *et al.*, 2007; Mondo *et al.*, 2012). Al respecto, Kravchenko y Bullock (2002b) determinaron que el patrón espacial de la topografía puede ser utilizada como un indicador del patrón espacial del contenido de aceite y proteína. Estos autores encontraron altos contenidos de proteína en zonas altas, con alta pendiente y con zonas convexas durante periodos con suficiente o excesiva precipitación, mientras bajos contenidos de proteína fueron observados en las mismas condiciones topográficas durante periodos con déficit hídrico. Por su parte, Leonardi (2012) y Sadras y Calviño (2001) determinaron que la profundidad efectiva delimitada por la presencia de horizontes petrocálcicos pudo explicar el patrón espacial del rendimiento de soja a escala de lote en condiciones del sudeste de la provincia de Buenos Aires. Así mismo, Martín *et al.* (2005) determinaron que propiedades del suelo relacionadas con la retención de agua en el perfil tales como

contenido de arcilla, materia orgánica, pH y conductividad eléctrica del extracto estuvieron espacial y consistentemente asociadas con propiedades del rendimiento y calidad de la soja. La mayoría de estos trabajos coinciden en establecer que pueden existir patrones espaciales comunes entre propiedades del suelo y propiedades del cultivo de soja. Por lo tanto, en este capítulo se plantea que estas estructuras espaciales pueden ser analizadas mediante técnicas cuantitativas que involucren de manera simultánea análisis estadístico, geoestadístico y determinación de la importancia de predictores para modelos de predicción de propiedades de suelo o cultivos.

En la actualidad son ampliamente conocidas las interrelaciones que existen entre las propiedades de rendimiento y calidad de soja. Por ejemplo, la concentración de proteína a menudo está inversamente relacionada con concentración de aceite y rendimiento (Burton, 1985; Clarke y Wiseman, 2000; Helms y Orf, 1998; Hu y Wiatrak, 2012). Por lo general, rendimiento y calidad de soja están limitadas por factores tales como temperatura, radiación y propiedades físicas y químicas de suelos (Anthony *et al.*, 2012; Hu y Wiatrak, 2012; Kravchenko y Bullock, 2002a). Al respecto, la pampa Argentina, se caracteriza por tener patrones espaciales complejos de propiedades físicas y químicas del suelo, a escala de lote (Amiotti *et al.*, 2001; Blanco y Stoops, 2007; Castro Franco *et al.*, 2015; Pazos y Mestelan, 2002). Debido a esto, es pertinente entender cuál es el rol de los patrones espaciales de propiedades del suelo en los patrones espaciales de propiedades del cultivo de soja, a escala de lote.

Técnicas de cartografía digital de suelos (CDS) generadas a partir de la disponibilidad de información relacionada con topografía, profundidad de tosca, índices de vegetación, sensores del suelo y mapas de rendimiento, han sido eficientes para determinar los patrones de variabilidad espacial de propiedades del suelo a escala de lote, en condiciones de la pampa argentina (Castro Franco *et al.*, 2015; Domenech, 2013; Peralta *et al.*, 2013; Peralta *et al.*, 2015; Simón *et al.*, 2012). Sin embargo, en estas mismas condiciones, no se tiene evidencia de trabajos que implementen CDS para estudiar el efecto del patrón espacial del suelo sobre las propiedades de rendimiento y calidad de soja. Las técnicas de CDS se caracterizan por tener en cuenta varias fuentes de información digital cuantitativa, que definen la distribución de propiedades del suelo tan importantes como textura,

materia orgánica, entre otras (McBratney *et al.*, 2003; Schmidt *et al.*, 2014). Por lo tanto, CDS puede jugar un rol importante para determinar las relaciones espaciales entre propiedades del suelo y propiedades de rendimiento y la calidad de la soja. Una ventaja a destacar de CDS es que permite obtener información de propiedades del suelo de manera rápida, precisa y a bajo costo (Cook *et al.*, 2008). Además, toda la información proviene de información cuantitativa relacionada con los factores de formación del suelo planteados por Jenny (1941). Como ya se mencionó, es posible que la mayoría de estos factores formadores del suelo y su patrón espacial determinen sustancialmente la variabilidad espacial de las propiedades del rendimiento y la calidad de los cultivos (Kravchenko y Bullock, 2000; Kravchenko y Bullock, 2002b; Kravchenko *et al.*, 2005; Martín *et al.*, 2005). Por lo tanto, es pertinente determinar la relación propiedades del suelo y propiedades del cultivo de soja, mediante técnicas de CDS.

En el capítulo tres de esta tesis, se detectó que los rendimientos de los cultivos, obtenidos a partir de mapas de rendimiento, tienen moderada y fuerte dependencia espacial. Así mismo, se detectó que los patrones espaciales del suelo afectan los patrones espaciales del rendimiento. De acuerdo con esto, en este capítulo se plantean dos hipótesis (i) el patrón espacial de propiedades de rendimiento y calidad de soja a escala de lote, están determinados por el patrón espacial de las propiedades del suelo; y (ii) la delimitación de zonas por tipo de suelos, permiten generar una base técnica y espacial para delimitar zonas para manejo sitio específico y zonas de cosecha diferenciada por calidad (Martin *et al.*, 2007).

El objetivo de este capítulo fue caracterizar y analizar las relaciones entre patrones espaciales de propiedades del suelo y cultivo de soja a escala de lote, a partir de la implementación de técnicas de CDS. Este capítulo busca documentar 6 casos de estudio en 3 regiones agrícolas en condiciones de la pampa argentina, discutiendo (i) la pertinencia y oportunidad de utilizar técnicas de CDS en estudio que determinen las relaciones entre patrones espaciales de propiedades del suelo y cultivo de soja a escala de lote, (ii) los tipos de relaciones y dependencias espaciales entre propiedades del suelo y cultivo de soja y (iii) la pertinencia y oportunidad de plantear zonas para manejo sitio específico y zonas para cosecha diferenciada por calidad.

## 2. Materiales y Métodos

### 2.1. Zonas de estudio

Las zonas y lotes de estudio fueron los mismos que se utilizaron en el capítulo tres de esta tesis. Estos ya se encuentran debidamente descritos y caracterizados. Por lo tanto, para efectos prácticos, recomendamos dirigirse a la sección 2.1 del Capítulo Tres.

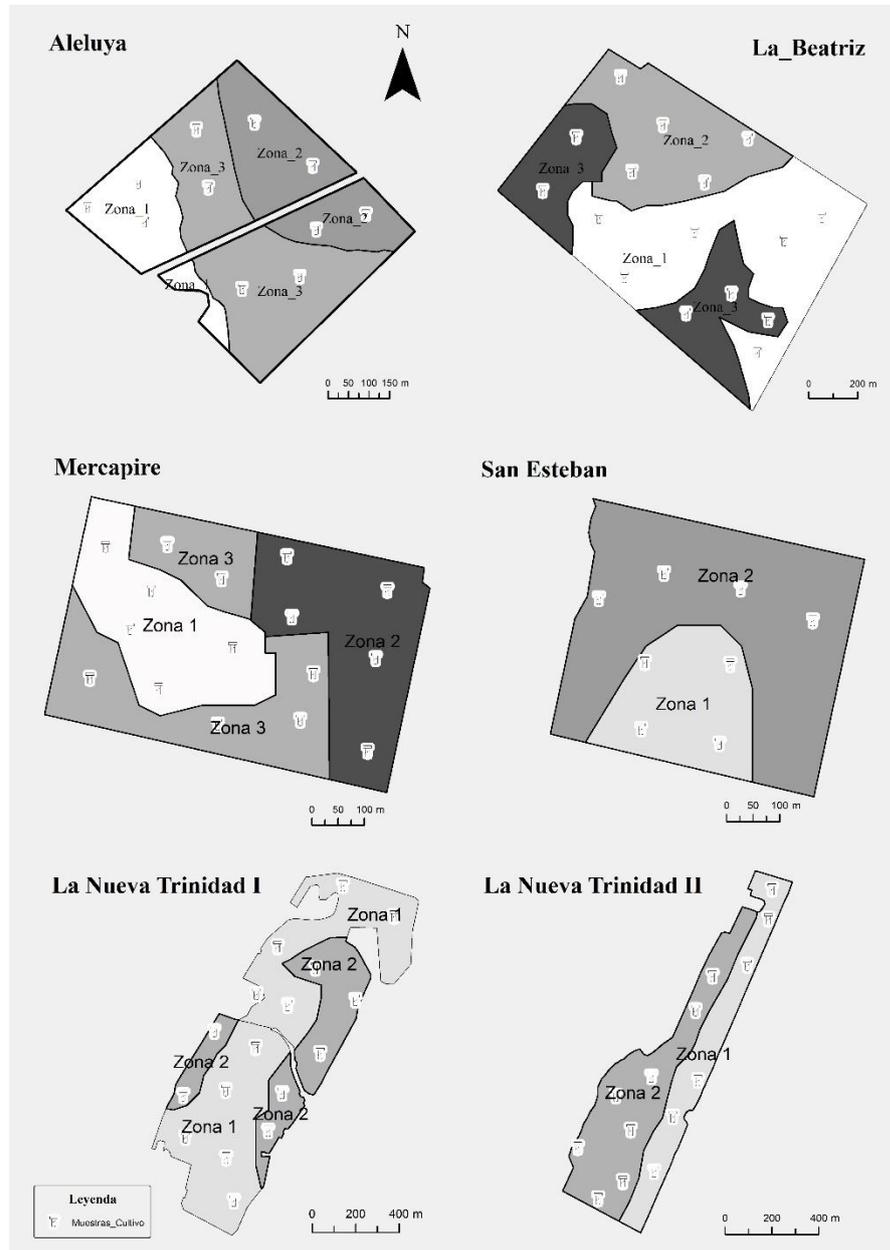
### 2.2. Medición de las propiedades del cultivo de soja

La figura 4.1 muestra los puntos de muestreo de variables de cultivo de soja en cada uno de los lotes. Como se observa, estos fueron exactamente los mismos donde se tomaron las muestras de suelo, utilizadas para hacer la validación de zonas por tipo de suelo en el capítulo 3.

Las variables de cultivo que se midieron estuvieron divididas de acuerdo a variables de rendimiento del cultivo y variables de calidad de grano. Las primeras fueron granos por metro cuadrado ( $\text{granos/m}^2$ ), peso de mil granos, intercepción de radiación e índice de vegetación de diferencia normalizada (NDVI). Por su parte, las segundas variables fueron contenido de aceite y proteína en grano de soja.

Las mediciones de radiación fotosintéticamente activa incidente ( $\text{RFA}_{\text{in}}$ ) se realizaron con un sensor lineal cuántico (Ceptómetro BAR-RAD DUAL USB, Cadadevices®, Bs.As. Argentina) colocado diagonalmente entre los surcos del dosel. La radiación incidente interceptada ( $\text{RFA}_{\text{int}}$ ) se midió durante la etapa fenológica R4 con el mismo sensor, debajo del dosel, en cada uno de los lotes (Gallo y Daughtry, 1986). En total se tomaron 15 mediciones de  $\text{RFA}_{\text{int}}$  localizadas al azar dentro de un radio de 10 m alrededor de cada punto de muestreo. Estas se promediaron, obteniendo la medición general para cada punto. Todas las mediciones se tomaron entre las 11:30 y 13:30, con día despejado (Andrade y Sadras, 2002; Satorre *et al.*, 2003). El porcentaje de  $\text{RFA}_{\text{int}}$  (RI) se calculó como:

$$RI = [(RFA_{inc} - RFA_{Int})/RFA_{inc}] \times 100 \quad (1)$$



**Figura 4.1** Distribución espacial de las muestras de cultivo de soja en cada zona delimitada de cada uno de los lotes en las tres zonas de estudio

Las mediciones del NDVI se realizaron con el sensor óptico activo GreenSeeker® (NTech Industries, Inc., CA, USA). Las mediciones se realizaron a un metro de altura de la canópia y sobre una transecta de 2 m lineales, durante la etapa fenológica R4, en cada uno de los lotes (Grove y Navarro, 2013a). GreenSeeker® tiene iluminación propia en las bandas del rojo (656 nm) e infrarrojo cercano (774 nm). GreenSeeker® calcula el NDVI como:

$$NDVI = \frac{(F_{NIR} - F_{Rojo})}{(F_{NIR} + F_{Rojo})} \quad (2)$$

donde  $F_{NIR}$  y  $F_{Rojo}$ , son las fracciones de las bandas infrarrojo cercano y rojo, respectivamente, reflejadas por el área sensada. GreenSeeker® tiene una alta tasa de mediciones (aproximadamente 1000 mediciones por segundo) y los resultados que se obtienen son el resultado de la media de NDVI del total de mediciones realizadas. Se estima que a una velocidad de medición de  $0.5 \text{ m s}^{-1}$ , GreenSeeker® puede obtener hasta 10 mediciones promedio por segundo (Ali *et al.*, 2015).

Los muestreos de grano se localizaron dentro de un radio de 10 m alrededor del punto de muestreo. Este muestreo se realizó cuando el cultivo se encontraba en madurez fisiológica, en cada uno de los lotes. A partir de todos los puntos de muestreo, se recolectó una muestra general equivalente a 500 g, la cual fue embalada y etiquetada. Los análisis de contenido de aceite y proteína fueron realizados en el laboratorio de Calidad Industrial de Cereales y Oleaginosas de la EEA INTA Marcos Juárez, utilizando un equipo de tecnología infrarroja NIR (Infratec 1241), de acuerdo a la norma AACC 39-21 (Cuniberti *et al.*, 2013).

### 2.3. Metodología para determinar las relaciones espaciales entre patrones espaciales de propiedades del suelo y cultivo de soja

Esta metodología se implementó en tres etapas:

### 2.3.1. Etapa I: Validación de la delimitación de zonas para propiedades de cultivo de soja.

Las zonas por tipo de suelo dentro de cada lote, delimitadas en los procesos de desagregación espacial de unidades cartográficas de suelo, descritos en el capítulo tres, fueron utilizadas para determinar las relaciones entre zona delimitada por tipo de suelo y propiedades del cultivo de soja.

El proceso de validación de zonas para propiedades del cultivo de soja, se realizó a partir de los resultados de los análisis de las variables de rendimiento y calidad de grano de soja, en cada uno de los lotes.

En orden de comparar las propiedades de rendimiento y calidad de cultivo de soja, dentro de cada zona delimitada, un modelo lineal mixto (MLM) y un ANOVA fueron ajustados para cada una de las propiedades de cultivo de soja, a partir del modelo básico:

$$S_{ij} = \mu + Z_i + \varepsilon_{ij}$$

donde  $S_{ij}$  representa el valor observable de la propiedad del cultivo de soja en la zona  $i$  del lote  $j$ ;  $\mu$  representa la media general de la propiedad del suelo;  $Z_i$  es el efecto de la zona delimitada con  $i = 1, \dots, Z$ ; y  $\varepsilon_{ij}$  es el error aleatorio. Estos modelos fueron ajustados con varianzas heterogéneas para las diferentes zonas delimitadas. La selección del modelo se hizo de acuerdo al criterio de información de Akaike. Para este procedimiento se utilizó el paquete “nlme” (Pinheiro *et al.*, 2007) del software estadístico R v3.1.1 (R Development Core Team, 2013).

### 2.3.2. Etapa II: Determinación de la importancia de propiedades del suelo sobre los patrones espaciales de propiedades del cultivo de soja

El algoritmo *Random Forest* (RF) fue utilizado para determinar la importancia de las propiedades del suelo sobre los patrones espaciales de las propiedades de rendimiento y

calidad del cultivo de soja (Breiman, 2001). La descripción del funcionamiento y principios matemáticos de RF fueron debidamente descritos en la sección 2.3 del capítulo 2 de esta tesis.

Durante el ensamble de árboles de clasificación, RF crea una estimación interna de la importancia de las variables mediante el método de estimación con los conjuntos “*Out of Bag*” (OOB) (Burgos, 2009). Los OOB son utilizados para obtener una estimación no sesgada del error de clasificación mientras los árboles se agregan al “bosque”. Al mismo tiempo, los OOB son utilizados para obtener una estimación de la importancia de las variables. Primero se estima el error de clasificación del ensamble sobre los OOB. Luego, se toma una variable predictora  $m$  y se permutan sus valores dentro del conjunto de datos de entrenamiento, decorrelacionando esa variable de lo aprendido en el modelo. Se itera la estimación del error OOB. Si la variable no es importante para el ensamble, el error OOB no debe cambiar mucho al desordenarla. Por el contrario, un gran cambio en el error OOB implica una gran importancia de la variable. Este proceso se itera y se ordenan las variables de acuerdo al cambio relativo en la estimación del error OOB que produce cada una (Archer y Kimes, 2008; Grömping, 2009). Numerosos estudios han determinado que RF es un algoritmo eficaz para determinar la importancia de variables predictores de propiedades del suelo (Brungard *et al.*, 2015; Castro Franco *et al.*, 2015; Genuer *et al.*, 2010; Ließ *et al.*, 2012; Rad *et al.*, 2014).

La importancia de cada propiedad del suelo sobre cada propiedad del cultivo de soja fue graficada para cada lote.

### *2.3.3. Etapa III: Relaciones espaciales entre zonas delimitadas por tipo de suelo y propiedades de cultivo*

En orden de comparar las propiedades de rendimiento y calidad de soja por zona delimitada, entre lotes, todos los datos de variables de rendimiento y cultivo de soja fueron normalizados por los valores máximos por lote. Posteriormente, un modelo lineal mixto (MLM) y un ANOVA fueron ajustados para cada una de las propiedades de cultivo de soja, a partir del modelo básico:

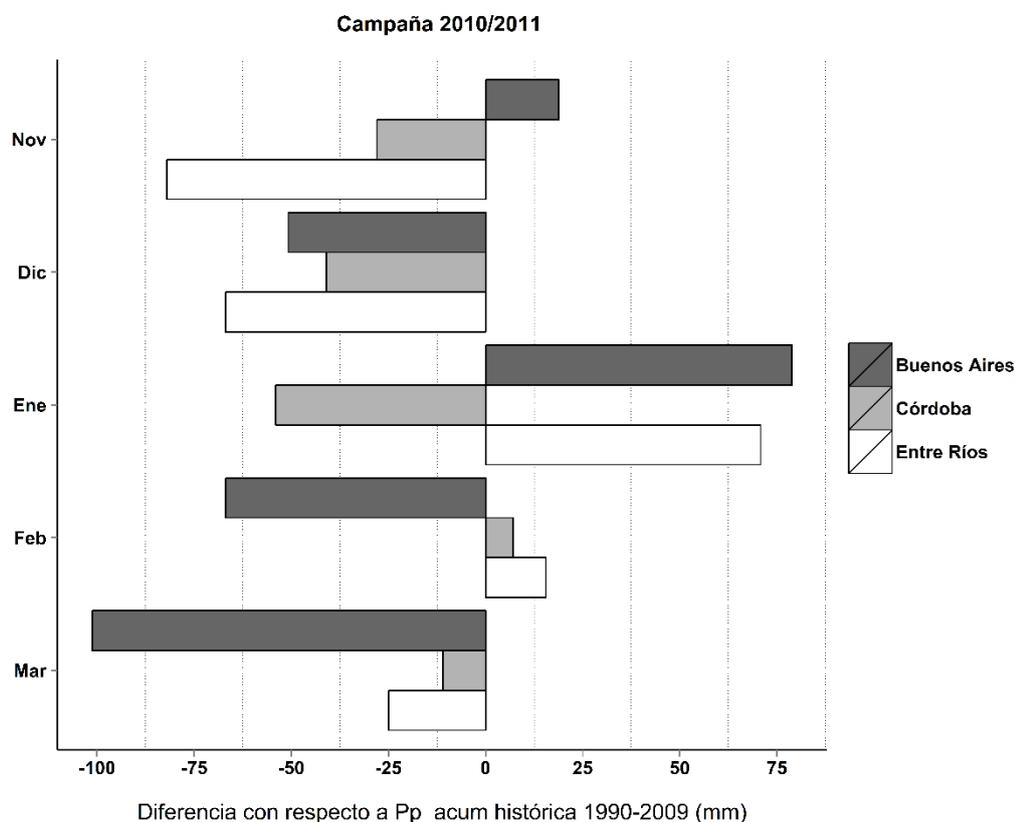
$$C_{ijk} = \mu + T_i + Z_{j(i)} + \varepsilon_{ijk}$$

donde  $C_{ijk}$  representa el valor observable de la propiedad del cultivo de soja en la zona  $i$  del lote  $j$ ;  $\mu$  representa la media general de la propiedad del cultivo;  $z_i$  es el efecto de la zona delimitada con  $i = 1, \dots, z$ , anidada dentro del lote  $j$ ; y  $\varepsilon_{ijk}$  es el error aleatorio. Estos modelos fueron ajustados con varianzas heterogéneas para las diferentes zonas delimitadas. La selección del modelo se hizo de acuerdo al criterio de información de Akaike. Las diferencias entre medias de propiedades del cultivo fueron comparadas utilizando el test de diferencias mínimas significativas. Para este procedimiento se utilizó el paquete “nlme” (Pinheiro *et al.*, 2007) del software estadístico R v3.1.1 (R Development Core Team, 2013).

### 3. Resultados y discusión

#### 3.1. Época de muestreo

El muestreo de las propiedades de rendimiento y calidad del cultivo de soja se realizaron en la campaña 2010/2011, en todos los lotes. Esta campaña fue la tercera con mayor déficit de precipitación en los últimos 20 años (INTA, 2013b). La figura 4.2 muestra las diferencias entre precipitación acumulada mensual de noviembre de 2010 hasta marzo de 2011, y el promedio de precipitación acumulada mensual para los mismos meses entre 1990 – 2009. Como se observa, la campaña 2010/2011 se caracterizó por un fuerte déficit hídrico, especialmente en la etapa de llenado de grano (febrero-marzo). Ha sido ampliamente documentado que el déficit hídrico en las etapas de llenado de grano de la soja (R4 – R5.5), afectan variables de rendimiento tales como número de granos y, de acuerdo a la duración del déficit hídrico, el peso de grano (Andrade y Sadras, 2002; Andriani, 2002; Dornbos *et al.*, 1989; Satorre *et al.*, 2003). Así mismo, afecta variables de calidad tales como concentración de aceite (Bellaloui *et al.*, 2015). Con excepción del lote San Esteban, no se encontraron granos de soja verde en los muestreos. Estos granos son un indicador de daño fisiológico y de poder germinativo en la soja (Del Valle Gallo, 2008).



**Figura 4.2** Diferencia de precipitación acumulada mensual entre nov – dic de 2010 y ene – mar de 2011 (*campaña 2010/2011*) e histórica 1990-2009 para los mismos meses.

### 3.2. Parámetros de estadística descriptiva de las propiedades de cultivo

Las tablas 4.1, 4.2 y 4.3 muestran los resultados de parámetros de estadística descriptiva de variables de rendimiento y calidad de soja, en cada uno de los lotes.

En la zona 1, los rendimientos medios de soja fueron 3425 y 2402 kg ha<sup>-1</sup>, para los lotes Aleluya y Mercapire, respectivamente. Por su parte, en la zona 2, los rendimientos medios fueron 1223 y 1191 kg ha<sup>-1</sup>, para los lotes Mercapire y San Esteban, respectivamente. Por último, en la zona 3 los rendimientos medios fueron 1946 y 2100 kg ha<sup>-1</sup>, para los lotes LNT\_I y LNT\_II, respectivamente. En general, estos rendimientos medios fueron bajos en relación con los históricos para cada zona (Andrade y Sadras, 2002; Baigorri, 2004; Sadras y Calviño, 2001). Esto se explica por el déficit hídrico presentado durante la campaña 2010/2011.

**Tabla 4.1** Parámetros de estadística descriptiva de variables de rendimiento y calidad de soja para los lotes de la zona 1 (Buenos Aires)

Campo	Variables	Media	DE*	Min*	Max*	
Aleluya Buenos Aires	Rendimiento	Granos/m <sup>2</sup>	2616.00	377.25	2155.00	3174.00
		Peso de 1000 granos (g)	130.12	7.83	1222.20	143.33
		Int. de Radiación (%)	95.74	3.72	87.85	99.63
		NDVI <sup>(1)</sup>	0.89	0.02	0.86	0.93
	Calidad	Aceite <sup>(2)</sup> (%)	22.06	0.99	20.90	23.80
		Proteína <sup>(2)</sup> (%)	40.20	0.44	39.40	40.73
		Media	DE*	Min*	Max*	
La Beatriz Buenos Aires	Rendimiento	Granos/m <sup>2</sup>	1981.00	279.78	1367.00	2425.00
		Peso de 1000 granos (g)	121.18	9.69	90.83	132.67
		Int. de Radiación (%)	95.13	3.08	88.56	98.89
		NDVI <sup>(1)</sup>	0.89	0.02	0.86	0.92
	Calidad	Aceite <sup>(2)</sup> (%)	23.60	1.32	19.33	24.63
		Proteína <sup>(2)</sup> (%)	38.26	1.66	36.13	42.13

<sup>(1)</sup> NDVI = Índice de Vegetación de Diferencia Normalizada obtenido con equipo Green Seeker  
<sup>(2)</sup> Porcentaje de aceite y proteína en semilla de soja determinado en Laboratorio  
\* DE : Desviación Estandar; Min : Valor mínimo; Max : Valor Máximo.

**Tabla 4.2** Parámetros de estadística descriptiva de variables de rendimiento y calidad de soja para los lotes de la zona 2 (Córdoba)

Campo	Variables	Media	DE*	Min*	Max*	
Mercapire Córdoba	Rendimiento	Granos/m <sup>2</sup>	925.00	315.73	381.00	1496.00
		Peso de 1000 granos (g)	132.31	6.53	120.26	143.35
		Int. de Radiación (%)	61.29	12.76	39.23	84.77
		NDVI <sup>(1)</sup>	0.65	0.07	0.55	0.74
	Calidad	Aceite <sup>(2)</sup> (%)	22.69	1.43	20.43	25.60
		Proteína <sup>(2)</sup> (%)	41.65	0.76	40.25	43.57
		Media	DE*	Min*	Max*	
San Esteban Córdoba	Rendimiento	Granos/m <sup>2</sup>	910.00	117.66	743.00	1112.00
		Peso de 1000 granos (g)	130.89	6.02	123.41	140.55
		Int. de Radiación (%)	52.18	8.75	42.42	67.66
		NDVI <sup>(1)</sup>	0.64	0.07	0.56	0.76
	Calidad	Aceite <sup>(2)</sup> (%)	23.43	0.79	22.20	24.90
		Proteína <sup>(2)</sup> (%)	40.90	0.69	39.50	41.70

<sup>(1)</sup> NDVI = Índice de Vegetación de Diferencia Normalizada obtenido con equipo Green Seeker  
<sup>(2)</sup> Porcentaje de aceite y proteína en semilla de soja determinado en Laboratorio  
\* DE : Desviación Estandar; Min : Valor mínimo; Max : Valor Máximo.

**Tabla 4.3** Parámetros de estadística descriptiva de variables de rendimiento y calidad de soja para los lotes de la zona 3 (Entre Ríos)

Campo	Variables	Media	DE*	Min*	Max*
La Nueva	Rendimiento				
Trinidad I	Granos/m <sup>2</sup>	1098.00	205.10	839.00	1702.00
Entre Ríos	Peso de 1000 granos (g)	177.32	15.59	130.48	196.74
	Int. de Radiación (%)	90.70	6.26	80.02	98.47
	NDVI <sup>(1)</sup>	0.79	0.04	0.70	0.85
	Calidad				
	Aceite <sup>(2)</sup> (%)	24.18	0.98	22.97	26.30
	Proteína <sup>(2)</sup> (%)	38.51	1.97	33.90	40.80
		Media	DE*	Min*	Max*
La Nueva	Rendimiento				
Trinidad II	Granos/m <sup>2</sup>	1313.00	412.34	590.00	2040.00
Entre Ríos	Peso de 1000 granos (g)	165.02	7.28	156.60	179.25
	Int. de Radiación (%)	90.34	6.84	79.42	99.50
	NDVI <sup>(1)</sup>	0.75	0.10	0.54	0.85
	Calidad				
	Aceite <sup>(2)</sup> (%)	25.93	0.69	24.10	26.70
	Proteína <sup>(2)</sup> (%)	35.04	1.64	33.80	39.90

<sup>(1)</sup> NDVI = Índice de Vegetación de Diferencia Normalizada obtenido con equipo Green Seeker

<sup>(2)</sup> Porcentaje de aceite y proteína en semilla de soja determinado en Laboratorio

\* DE: Desviación Estandar; Min: Valor mínimo; Max: Valor Máximo.

El déficit hídrico presentado durante la campaña 2010/2011 fue más intenso en la zona 2 (Figura 4.2). A diferencia de las otras zonas, la zona 2 presentó un déficit de precipitaciones muy fuerte en enero de 2011. Teniendo en cuenta que para todas las zonas, el déficit hídrico se venía acentuando desde noviembre de 2010, el hecho de tener un fuerte déficit en enero, época determinante del rendimiento en soja (Andrade y Sadras, 2002), generó que los rendimientos en la zona 2 fueron muy bajos, con respecto a los rendimientos de las campañas predecesoras.

Como se esperaba, la sequía en la campaña 2010/2011 tuvo un efecto considerable en la definición de granos m<sup>-2</sup>. En general, el efecto de la sequía sobre los componentes del rendimiento de soja varía de acuerdo al estado fenológico y cantidad de tiempo en que el cultivo está sometido a estrés (Borrás *et al.*, 2004; Oya *et al.*, 2004). Al respecto, Andrade (1995) y Andrade y Ferreiro (1996) determinaron que una modificación de capacidad de fuente durante la etapa de llenado de grano de soja tiene un impacto importante en granos m<sup>-2</sup> y por ende, en el rendimiento. Por su parte, Andriani y Bodrero (1995) determinaron

que deficiencias hídricas en etapas fenológicas R4 y R5.5 en soja, afectan principalmente el número de granos y en estados fenológicos posteriores afecta el peso de los granos. A pesar de que el efecto de la sequía sobre granos  $m^{-2}$  fue generalizado, las desviaciones estándar en todos los lotes fueron altas para esta variable de rendimiento. Este resultado indica que las variaciones intra-lote en granos  $m^{-2}$ , posiblemente estén relacionadas a patrones espaciales ambientales o edáficos dentro de cada lote, que las determinan (Anthony *et al.*, 2012).

La sequía en la campaña 2010/2011 no tuvo un efecto generalizado en la definición del peso de 1000 granos (P1000). Además de las condiciones ambientales, el P1000 puede variar entre lotes de acuerdo al tipo de semilla, grupo de madurez y fecha de siembra (Hu y Wiatrak, 2012). Como se mencionó anteriormente, el efecto de la sequía sobre el P1000 se da cuando los periodos de estrés hídrico son muy prolongados (Andriani y Bodrero, 1995; Andriani, 2002). En general, un periodo de estrés hídrico prolongado no se presentó en ninguna de las zonas. Un estrés hídrico prolongado durante las etapas reproductivas afecta la duración y tasa del llenado de grano (Hu y Wiatrak, 2012), lo cual está directamente correlacionado con acumulación de materia seca en grano (De Souza *et al.*, 1997; Desclaux y Roumet, 1996). En la zona 3, se encontraron los valores más altos de P1000, mientras que en la zona 1 y 2 se encontraron valores similares. Sin embargo, las desviaciones estándar para P1000 fueron mayores para la zona 3. Las diferencias en P1000, posiblemente se deben al uso de semillas de GM VI ciclo corto (**DM 6.2i STS**) (ver: <http://www.donmario.com>), con promedio potencial de P1000 de 175 -185 g en los lotes de la zona 3 (Anthony *et al.*, 2012; De Souza *et al.*, 1997). Mientras que en los lotes de la zona 1 y 2 se utilizaron semilla de GM IV ciclo corto y GM IV ciclo largo (**CQ4.9-AGD**), respectivamente, los cuales tienen un potencial de P1000 entre 156-178 g.

La figura 4.3 muestra las diferencias visuales de RI alrededor de la etapa fenológica R2-R4, en cada uno de los lotes. En general, RI tuvo correlación y correspondencia con los resultados de granos  $m^{-2}$ . Los menores valores de RI se presentaron en los lotes de la zona 2, pero especialmente fueron bajos en el lote San Esteban.



**Figura 4.32** Diferencias visuales de radiación interceptada alrededor de la etapa feonológica de soja R2-R4, en cada uno de los lotes.

Estos resultados se explican por el déficit hídrico presentado en la campaña 2010/2011 sumado al efecto de baja retención de humedad por parte de los suelos con alto contenido de arenas, predominantes en la zona 2. Es ampliamente conocido que en estrés hídrico, la turgencia de los tejidos vegetales de la soja disminuyen, afectando la tasa de expansión foliar. Por ende, este efecto a su vez, genera otro efecto combinado de reducción de índice de área foliar, intercepción de radiación, biomasa total, altura de plata y número de vainas (Andrade, 1995; Hu y Wiatrak, 2012).

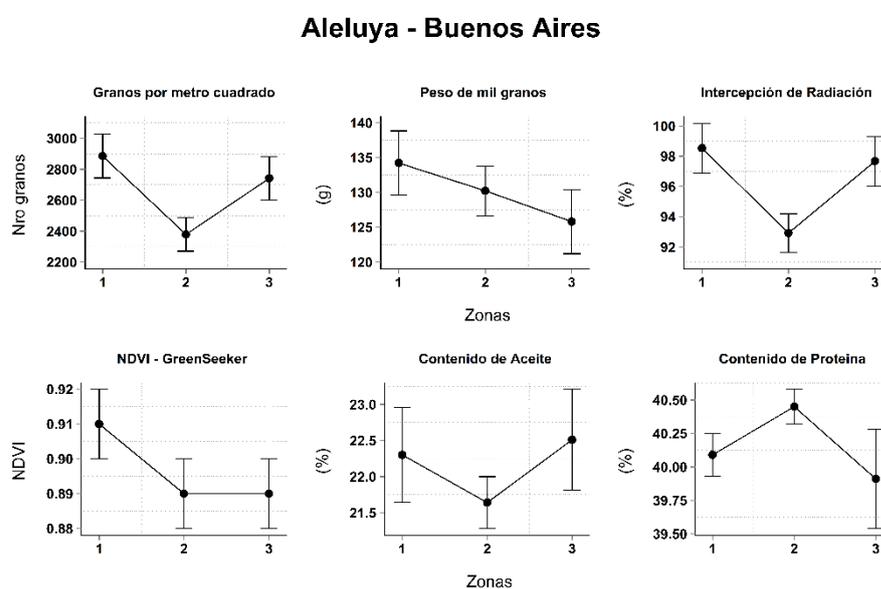
GreenSeeker® permitió obtener NDVI del cultivo de soja en cada punto de muestreo de manera precisa, fácil, económica y rápida. En general, NDVI tuvo correlación y correspondencia con los resultados de granos  $m^{-2}$  e RI. Como se esperaba, los NDVI más bajos se presentaron en los lotes de la zona 2, seguido por los lotes de la zona 3 y 1, respectivamente. NDVI en el cultivo de soja está asociado cerradamente con capacidad fotosintética y absorción de energía por parte de las canópia del cultivo (Mulla, 2013; Zhitao *et al.*, 2014). Al respecto, Grove y Navarro (2013b) reportaron que NDVI en el cultivo de soja está relacionado con rendimiento en grano y con contenido de nitrógeno y fosforo foliar. Además, los resultados de NDVI corroboran lo planteado por Ma *et al.* (2001), quienes determinaron que la medición no destructiva de la reflectancia de la canópia del cultivo de soja entre los estados fenológicos R4 – R5, discrimina adecuadamente los rendimientos potenciales y proporciona un indicador holístico de la generación del rendimiento, aprovechando las cerradas asociaciones entre rendimiento en grano y NDVI. Las desviaciones estándar de los NDVI en cada lote, posiblemente están sugiriendo que factores ambientales o edáficos dentro del lote, pueden estar afectando los patrones espaciales del NDVI.

La concentración de proteína fue moderadamente variable, mientras que la concentración de aceite fue levemente variable, entre lotes. Un limitado número de estudios ha reportado los efectos de los factores ambientales y/o edáficos dentro de un lote, que determinan cerradamente los cambios en los contenidos de proteína (Bellaloui *et al.*, 2009; Kravchenko y Bullock, 2002b). Al respecto, Anthony *et al.* (2012) determinaron relaciones consistentes entre patrones espaciales de contenido de proteína, mineralización de nitrógeno y potasio en el suelo. De acuerdo con estos autores, el

contenido de proteína en grano responde positivamente a la fertilización nitrogenada. Sin embargo, esta relación no fue consistente temporalmente. Por su parte, el potasio del suelo tuvo una significativa correlación negativa con contenido de proteína y positiva con contenido de aceite. Aunque es ampliamente conocido que los factores ambientales y genéticos son los principales factores que determinan las relaciones proteína/aceite en soja, las desviaciones estándar, valores mínimos y máximos reportados en este capítulo, sugieren que en los lotes de las tres zonas podrían existir factores intrínsecos ambientales y/o edáficos que afectan los patrones espaciales de los contenidos de aceite y proteína en soja.

### 3.3. Etapa I: Validación de zonas por tipo de suelo para propiedades del cultivo de soja

Las figuras 4.4 y 4.5 muestran la respuesta de cada propiedad del cultivo analizada en función de las zonas delimitadas en cada lote de la zona 1.

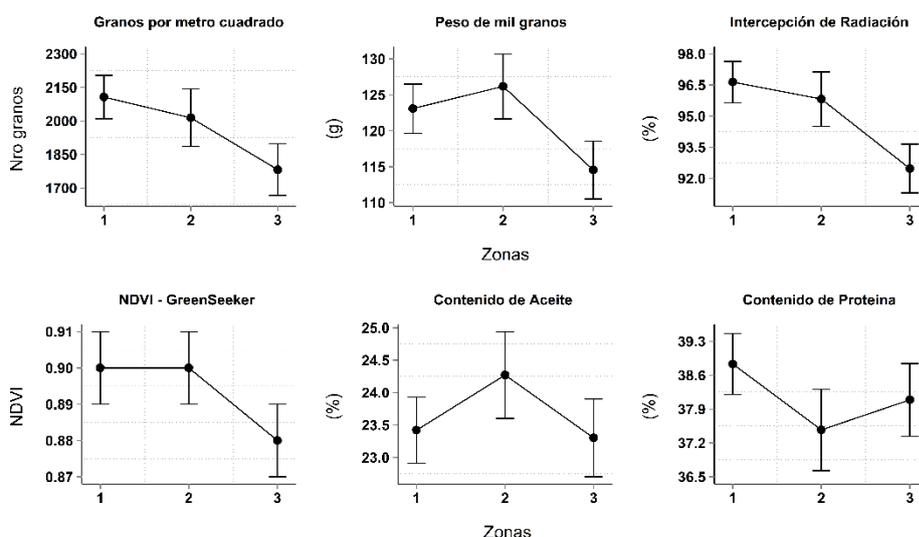


**Figura 4.4** Medias de propiedades de cultivo en cada zona delimitada en el lote Aleluya – Buenos Aires.. Barras verticales indican error estandar de la media para cada una de las zonas delimitadas

En *Aleluya*, valores bajos de granos  $m^{-2}$ , RI y contenido de aceite estuvieron asociados a zonas con tosca somera, CEa alta, mayor contenido de arcilla y MO, mayor CE del extracto y pH (*zona delimitada 2*). Valores altos de P1000 y NDVI estuvieron asociados a zonas de elevación media con profundidad mayor a 100 cm (*zona delimitada 1*).

Contenido de proteína presentó respuesta inversa a contenido de aceite en todas las zonas delimitadas. En general, las zonas delimitan las propiedades del rendimiento y calidad de granos de soja de manera acertada y consistente con las propiedades del suelo relacionadas en cada zona. Los resultados de este capítulo sugieren que al igual que los patrones espaciales de propiedades del suelo, los patrones espaciales de cultivo están relacionados con la profundidad de la tosca. Limitada cantidad de trabajos han determinado el efecto de la profundidad de la tosca sobre las propiedades de rendimiento y calidad de la soja. Al respecto, Sadras y Calviño (2001) determinaron que la tosca afecta los componentes de rendimiento de la soja a escala de lote, debido a que limita la capacidad de almacenamiento de agua en el suelo. Por su parte, Leonardí (2012) determinó una merma en granos  $m^{-2}$  en zonas con tosca somera, coincidiendo con los resultados de este trabajo. En general, los patrones de tosca somera en condiciones de déficit hídrico potenciaron mayores contenidos de proteína, corroborando que las zonas dentro del lote más propensas a estrés hídrico, son las que menor contenido de aceite presentan (Andrade y Sadras, 2002; Andriani, 2002; Cuniberti *et al.*, 2013). Corroborar este principio, es importante para plantear zonas de cosecha diferenciada, utilizando la metodología de delimitación de zonas por tipo de suelos, propuesta en el capítulo 3.

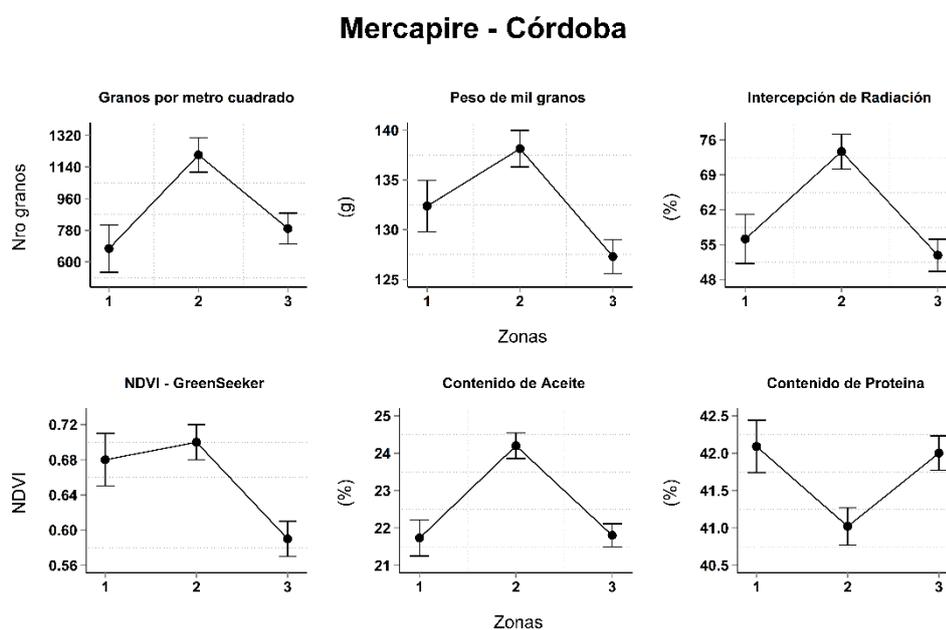
### La Beatriz - Buenos Aires



**Figura 4.5** Medias de propiedades de cultivo en cada zona delimitada en el lote La Beatriz – Buenos Aires.. Barras verticales indican error estandar de la media para cada una de las zonas delimitadas

En La Beatriz, excepto contenido de proteína, valores bajos de todas las variables de rendimiento y calidad de soja, estuvieron asociados a zonas altas con tosca somera, mayor contenido de arcilla, CE extracto, CIC, MO y pH (zona delimitada 3). Los valores más bajos de contenido de proteína estuvieron asociados a zonas bajas con profundidades del suelo >120 cm (zona delimitada 1). Al igual que en Aleluya, en La Beatriz contenido de proteína presentó respuesta inversa a contenido de aceite en todas las zonas delimitadas. En general, el patrón espacial de las propiedades del rendimiento y calidad de granos de soja estuvo delimitado de manera acertada y consistente con las propiedades del suelo en cada zona. A partir de esto, se reitera que la metodología para delimitar zonas por tipo de suelos, puede ser potencialmente aplicable como delimitación de zonas para cosecha diferenciada.

Las figuras 4.6 y 4.7 muestran la respuesta de cada propiedad del cultivo analizada en función de las zonas delimitadas en cada lote de la zona 2.



**Figura 4.6** Medias de propiedades de cultivo en cada zona delimitada en el lote Mercapire – Córdoba.. Barras verticales indican error estándar de la media para cada una de las zonas delimitadas

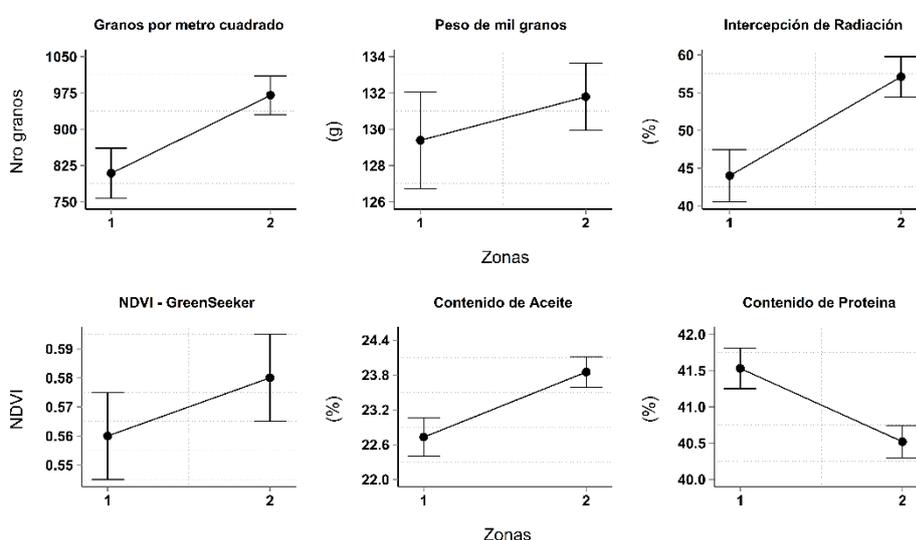
En *Mercapire*, excepto contenido de proteína, valores altos de todas las variables de rendimiento y calidad estuvieron asociadas a zonas bajas con suelos con mayores contenidos de arcilla, CE del extracto, CIC y MO (*zona delimitada 2*). Por su parte,

valores altos de contenido de proteína estuvieron asociados a zonas altas (*zonas delimitadas 1 y 3*). Estos resultados corroboran lo planteado por Kravchenko y Bullock (2002b), quienes determinaron que la topografía del lote tiene un efecto importante en la calidad de grano en soja y que puede ser utilizada como un indicador del patrón espacial del contenido de aceite y proteína en soja. Sin embargo, estos autores plantearon que altos contenidos de proteína estaban asociados a zonas de mayor altura, solamente en condiciones de buena disponibilidad hídrica, y que por el contrario, esta asociación cambiaba en épocas de déficit hídrico. Los resultados presentados en este trabajo contradicen las asociaciones de patrones espaciales entre topografía y propiedades de calidad del grano presentados por Kravchenko y Bullock (2002b). Teniendo en cuenta la severa sequía de la campaña 2010/2011, las asociaciones encontradas ratificaron que en zonas más altas, los contenidos de proteínas son mayores en condiciones de déficit hídrico. Esto se puede explicar debido a las condiciones generales de textura y contenido de MO de los suelos de la zona 2. En las zonas altas predominan mayores contenidos de arenas. Es ampliamente conocido que la capacidad de retención de agua es menor en suelos con mayores contenidos de arenas. Por lo que la tendencia del cultivo de soja es a que el estrés sea más severo en zonas altas, en épocas de sequía y por lo tanto, los contenidos de proteína son mayores. Una vez más, la metodología de delimitación de zonas por tipo de suelos planteada en el capítulo tres, ratifica que es posible utilizarla para plantear zonas de cosecha diferenciada por tipo de calidad de grano en soja.

En *San Esteban*, excepto proteína, valores altos de todas las propiedades de cultivos estuvieron asociados con zonas con menor CE del extracto y mayor contenido de MO (*zona delimitada 2*). Por su parte, valores altos de contenido de proteína estuvieron asociados a zonas con CE del extracto  $>1 \text{ dS m}^{-1}$  (*zonas delimitada 1*). Claramente, todas las variables de rendimiento y calidad de soja fueron afectadas por la presencia de zonas salinas dentro del lote. Es ampliamente conocido que la presencia de zonas salinas dentro del lote genera dos tipos de estrés en los cultivos, (i) aumenta la presión osmótica en la solución del suelo, lo cual limita el consumo de agua de los cultivos, y (ii) aumenta la acumulación de  $\text{Na}^+$  y  $\text{Cl}^-$  en la planta (Abrol *et al.*, 1988; Casanellas, 2003; Metternicht y Zinck, 2003; Rhoades, 1993). Generalmente, el estrés por salinidad afecta la RI, la conducción estomatal y la eficiencia en el uso de la radiación (fotosíntesis) (Salisbury y

Ross, 1992). Debido a esto, los síntomas de estrés salino se expresan fenotípicamente como hojas cloróticas y necróticas. Por su parte, Wang *et al.* (2001a) plantearon que la salinidad reduce la RI y su eficiencia en cultivos de soja. Estos autores observaron un menor tamaño de área foliar y presencia de zonas necróticas en las hojas, los cuales son causante de una reducción de índice de área foliar. Claramente, estos resultados demuestran que la delimitación de zonas salinas (CE del extracto  $>1 \text{ dS m}^{-1}$ ) dentro del lote, puede ser un fuerte indicador de los patrones espaciales de rendimiento y calidad en soja.

### San Esteban - Córdoba



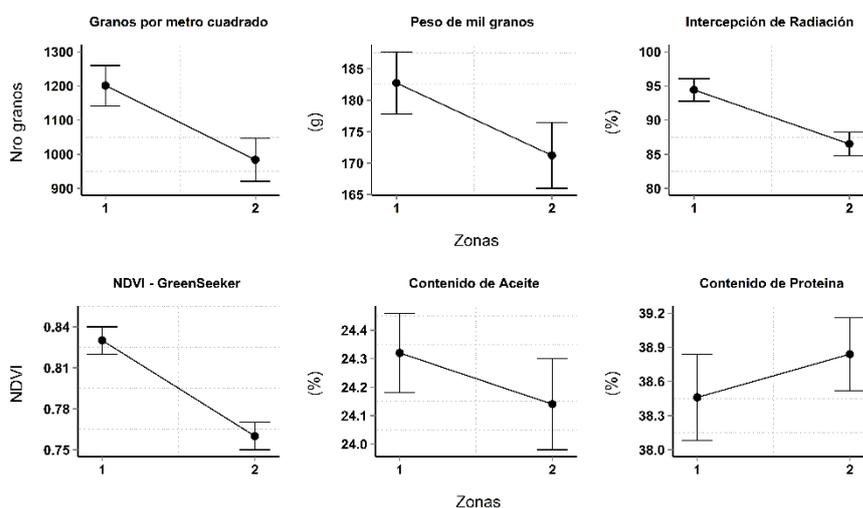
**Figura 4.7** Medias de propiedades de cultivo en cada zona delimitada en el lote San Esteban – Córdoba.. Barras verticales indican error estandar de la media para cada una de las zonas delimitadas

Las figuras 4.8 y 4.9 muestran la respuesta de cada propiedad del cultivo analizada en función de las zonas delimitadas en cada lote de la zona 3.

En *LNT\_I* y *II*, las variables de rendimiento de soja fueron afectadas por la presencia de zonas salinas dentro del lote. Excepto contenido de proteína y P1000 en *LNT\_II*, valores bajos de todas las variables de rendimiento y calidad de soja estuvieron asociados con zonas con CE del extracto  $>3 \text{ dS m}^{-1}$  (*zona delimitada 2 en LNT\_I y II*). Diferencias sustanciales de P1000 se presentaron entre zonas delimitadas en *LNT\_I*, mientras que en *LNT\_II*, no se presentaron. Contenidos de proteína levemente alto estuvo asociado a zonas con alto contenido de sales. Al igual que *San Esteban*, estos resultados indican que

la delimitación de zonas salinas dentro del lote, puede ser un fuerte indicador de los patrones espaciales de rendimiento de soja, mas no de los patrones espaciales de calidad de soja. Así mismo, ante presencia de zonas salinas ( $CE$  del extracto  $>1.5 \text{ dS m}^{-1}$ ), la topografía no fue un indicador veraz de los patrones espaciales de rendimiento y calidad en soja. En estos casos, los sensores de CEa pueden ser una herramienta eficaz para la delimitación de zonas para manejo sitio específico, mas no para delimitación de zonas de cosecha diferenciada (Rhoades *et al.*, 1976; Rhoades, 1993).

### La Nueva Trinidad I - Entre Ríos



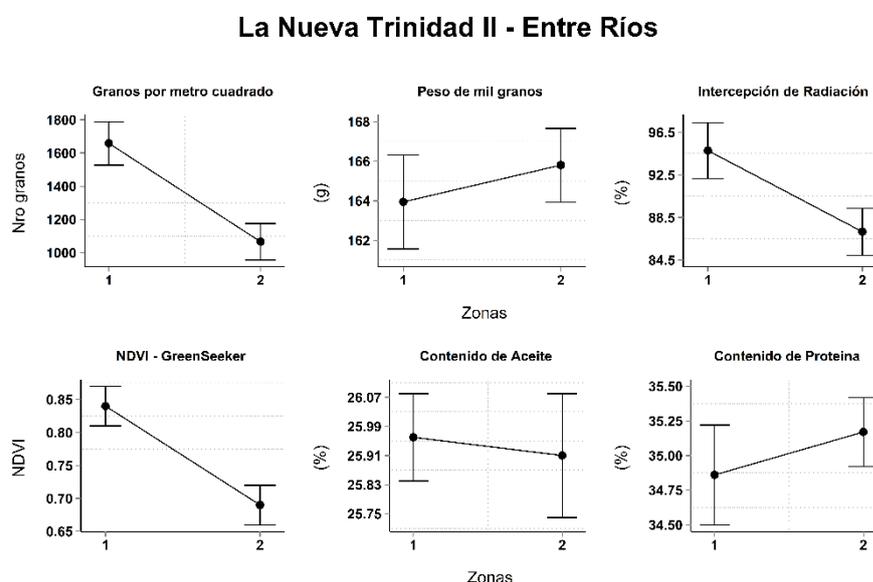
**Figura 4.8** Medias de propiedades de cultivo en cada zona delimitada en el lote LNT\_I – Entre Ríos. Barras verticales indican error estandar de la media para cada una de las zonas delimitadas

#### 3.4. Etapa II: Importancia de propiedades del suelo en los patrones espaciales de propiedades del cultivo de soja.

Las relaciones espaciales puntuales ente propiedades del suelo y propiedades de rendimiento y calidad de soja son analizadas en esta etapa,

Las figuras 4.10 y 4.11 muestran la importancia de propiedades del suelo, para cada una de las propiedades de rendimiento y calidad de soja en los lotes de la zona 1, determinado a partir de la función de importancia de variables del algoritmo RF. A través de esta función se logró explicar el 67 y 54% de la varianza de las propiedades del cultivo,

para *Aleluya* y *La Beatriz*, respectivamente. Variables con valores negativos de importancia significa que tienen una escasa o nula importancia (Genuer *et al.*, 2010).

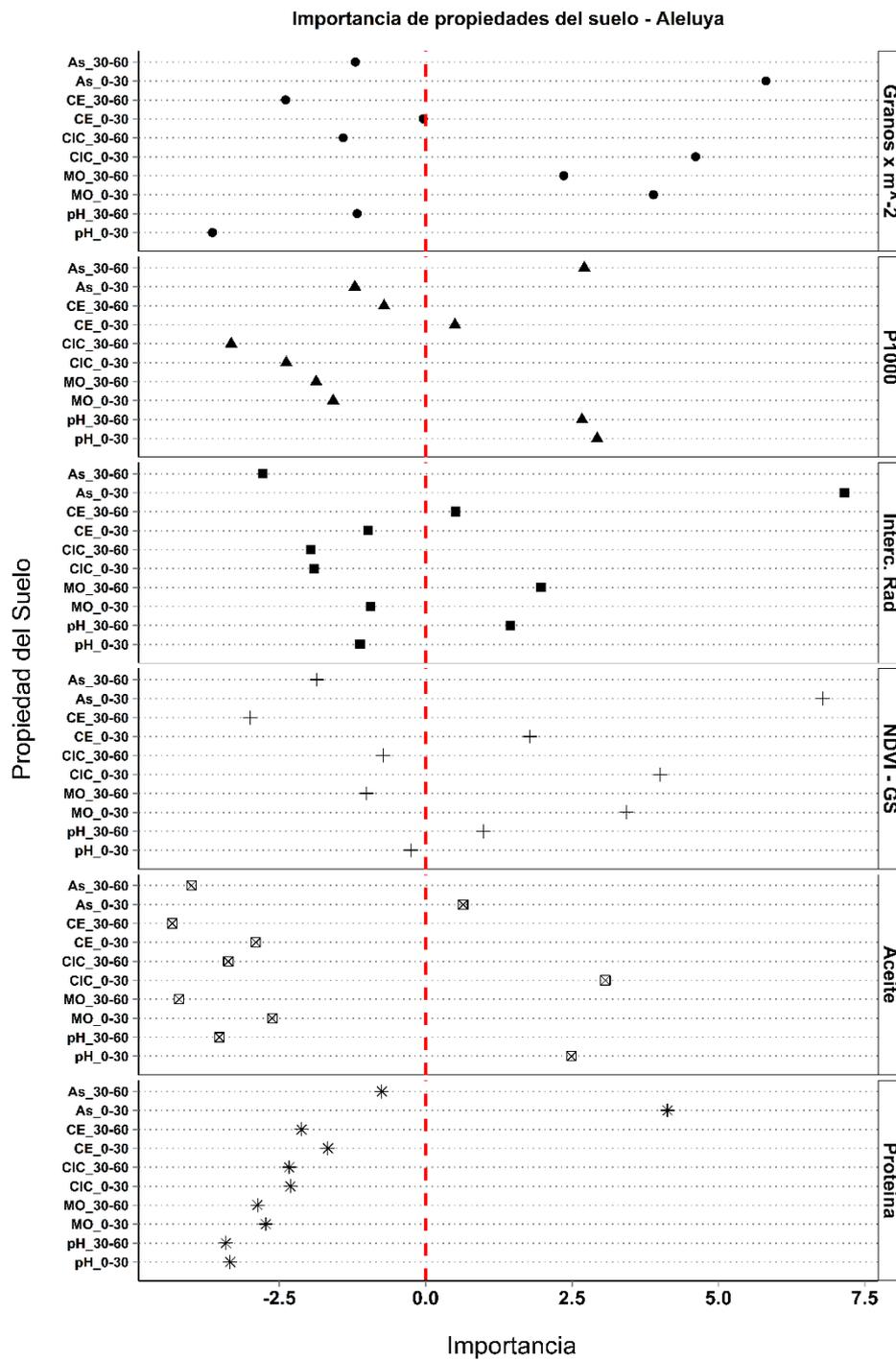


**Figura 4.9** Medias de propiedades de cultivo en cada zona delimitada en el lote LNT\_II – Entre Ríos. Barras verticales indican error estandar de la media para cada una de las zonas delimitadas

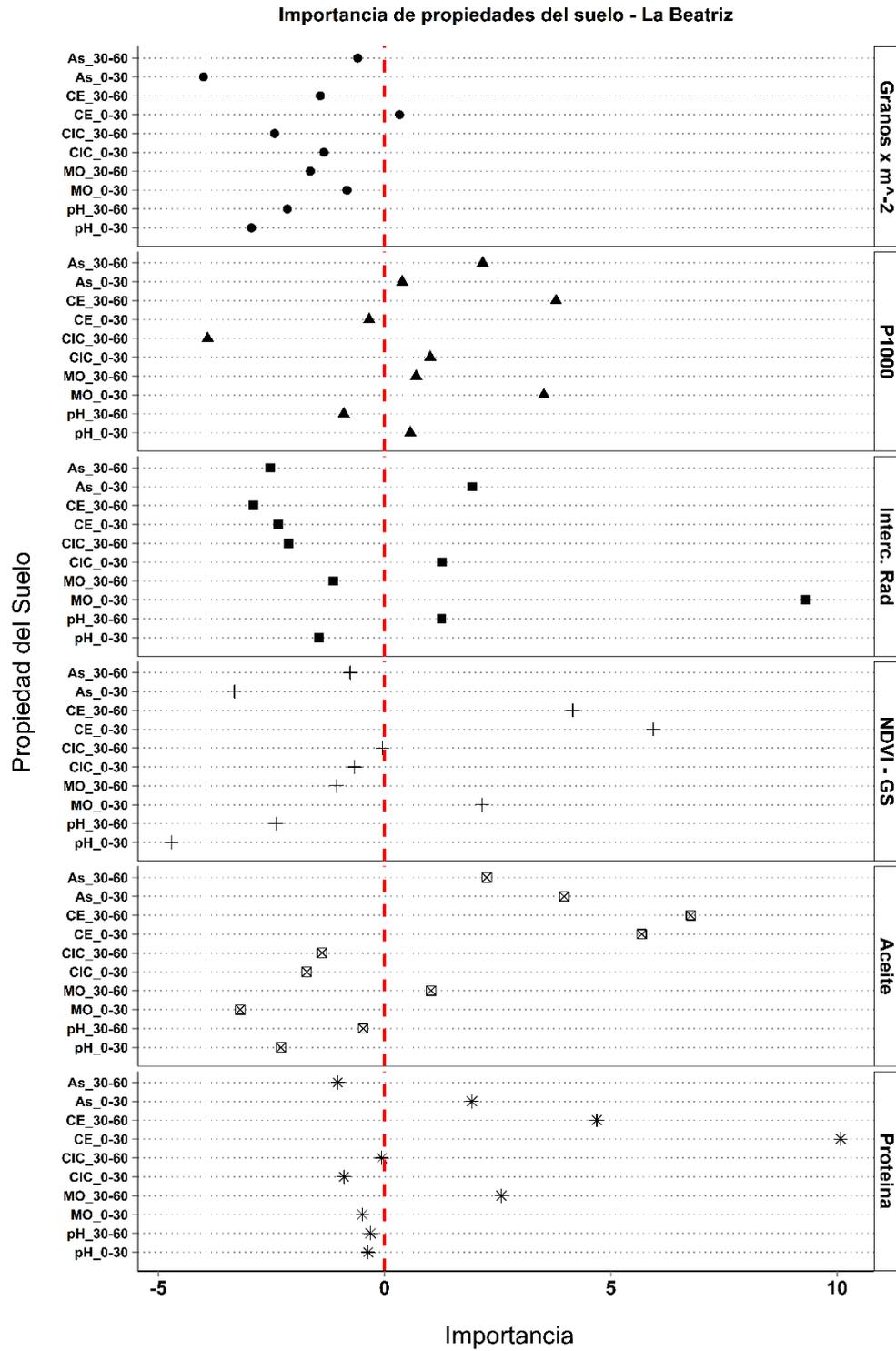
En *Aleluya*, contenido de arcilla superficial (*As 0-30 cm*) fue la variable más importante para granos  $m^{-2}$ , RI, NDVI y contenido de proteína. Arcilla profunda (*As 30-60 cm*) y pH en las dos profundidades, fueron las propiedades del suelos más importantes para P1000, mientras que CIC somero (*CIC 0-30 cm*) fue importante para contenido de aceite. Estos resultados son consistentes con lo reportado anteriormente en la validación de zonas por tipo de suelos para propiedades del cultivo de soja (Sección 3.3 - Etapa I). Valores altos de arcilla estuvieron relacionados con valores bajos de granos  $m^{-2}$  e RI y valores bajos de contenido de proteína. En las zonas delimitadas con altos contenido de arcilla se relacionaban con contenidos mayores de MO y tosca somera. Debido a esto, MO superficial (*MO 0-30 cm*) fue la tercera variable en importancia para granos  $m^{-2}$  y NDVI. De manera similar, Martín *et al.* (2005) determinaron que mayor cantidad MO y contenidos de arcilla explican la variación de los componentes del rendimiento en soja. Por su parte, Schepers *et al.* (2004) determinó que factores edáficos relacionados con disponibilidad hídrica están relacionados con los patrones espaciales de propiedades de rendimiento del cultivo de soja. Los resultados de relaciones entre propiedades del suelo y contenido de proteína coinciden parcialmente con lo planteado por Anthony *et al.*

(2012). Estos autores determinaron que pH superficial estuvo correlacionado con la tasa proteína/aceite en granos de soja. Los resultados para *Aleluya* muestran que pH superficial fue importante para contenido de aceite, pero no para contenido de proteína. Este resultado era inesperado ya que en general, la síntesis de proteínas suele ser particularmente sensible a los cambios en pH. Además, en zonas delimitadas con valores diferentes de pH estuvieron relacionadas con cambios en los contenidos de proteína. Estos resultados sugieren que es posible que variables temporales relacionados con la oferta ambiental, determinan las interacciones entre propiedades del suelo y cultivo de soja. Por lo tanto, en condiciones del sudeste bonaerense en épocas de déficit hídrico, los contenidos de proteína pueden ser más sensibles a la disponibilidad de agua que al pH del suelo.

En *La Beatriz*, la mayoría de las propiedades del suelo no tuvieron importancia considerable para granos  $m^{-2}$ . Por su parte, CE del extracto superficial (0 – 30 cm) fue la más importante para granos  $m^{-2}$ , NDVI y contenido de proteína, mientras que CE del extracto profunda (30 – 60 cm) fue la propiedad del suelo más importante para P1000 y contenido de aceite. MO superficial (0 – 30 cm) fue la propiedad del suelo más importante para RI. Contenido de arcilla en el suelo también fue importante para contenido de aceite, P1000 e RI. Estos resultados no son consistentes con los reportados para *Aleluya*. Los resultados se pueden explicar a qué CE del extracto, MO y pH presentaron variaciones importantes entre zonas delimitadas. Los patrones espaciales de estas propiedades del suelo podrían estar asociadas a los patrones espaciales de la profundidad de tosca y elevación (Bedogni, 1996; Cabria y Culot, 1994). A pesar de la importancia de la CE del extracto en la mayoría de las propiedades de rendimiento y calidad de la soja, no se excluye el efecto de la humedad del suelo. La importancia de contenido de arcilla en contenido de aceite y proteína sugiere que la disponibilidad hídrica podría ser un factor determinante de la calidad de grano en soja. Esto es importante, si se tiene en cuenta que los patrones espaciales de arcilla están cerradamente relacionados con profundidad de tosca. Zonas con suelos profundos facilitan la infiltración, incrementan la capacidad de almacenamiento de agua, facilitando la disponibilidad de nutrientes y aumentando la eficiencia en el uso del agua. Esto a su vez, tiene un impacto determinante en los componentes de rendimiento y calidad de la soja (Bellaloui *et al.*, 2009).



**Figura 4.10** Importancia de las propiedades del suelo para cada una de las propiedades del cultivo de soja en el lote Aleluya – Buenos Aires, a partir de las clasificaciones realizadas con el algoritmo *Random Forest* (RF)

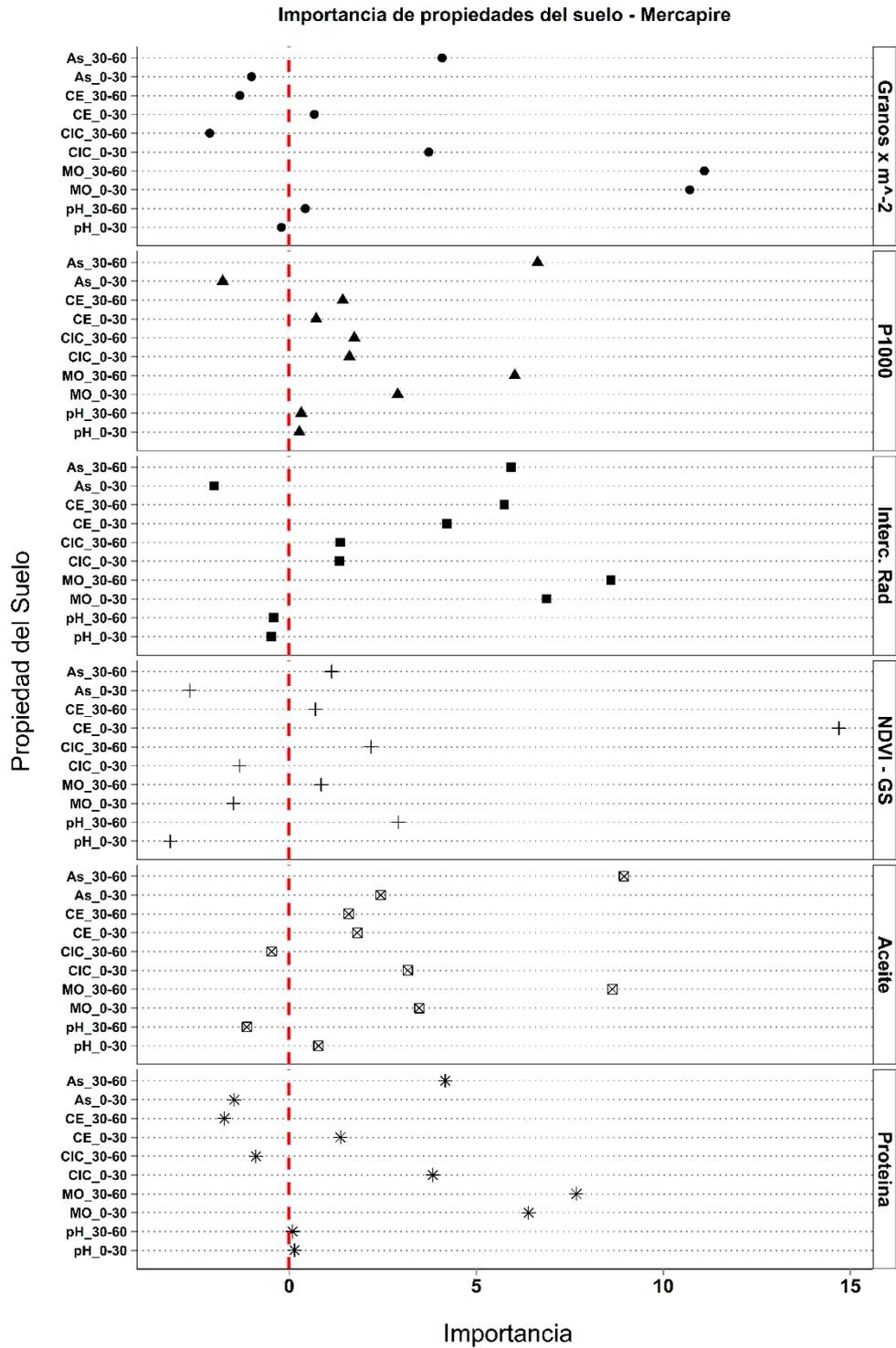


**Figura 4.11** Importancia de las propiedades del suelo para cada una de las propiedades del cultivo de soja en el lote La Beatriz – Buenos Aires, a partir de las clasificaciones realizadas con el algoritmo *Random Forest* (RF)

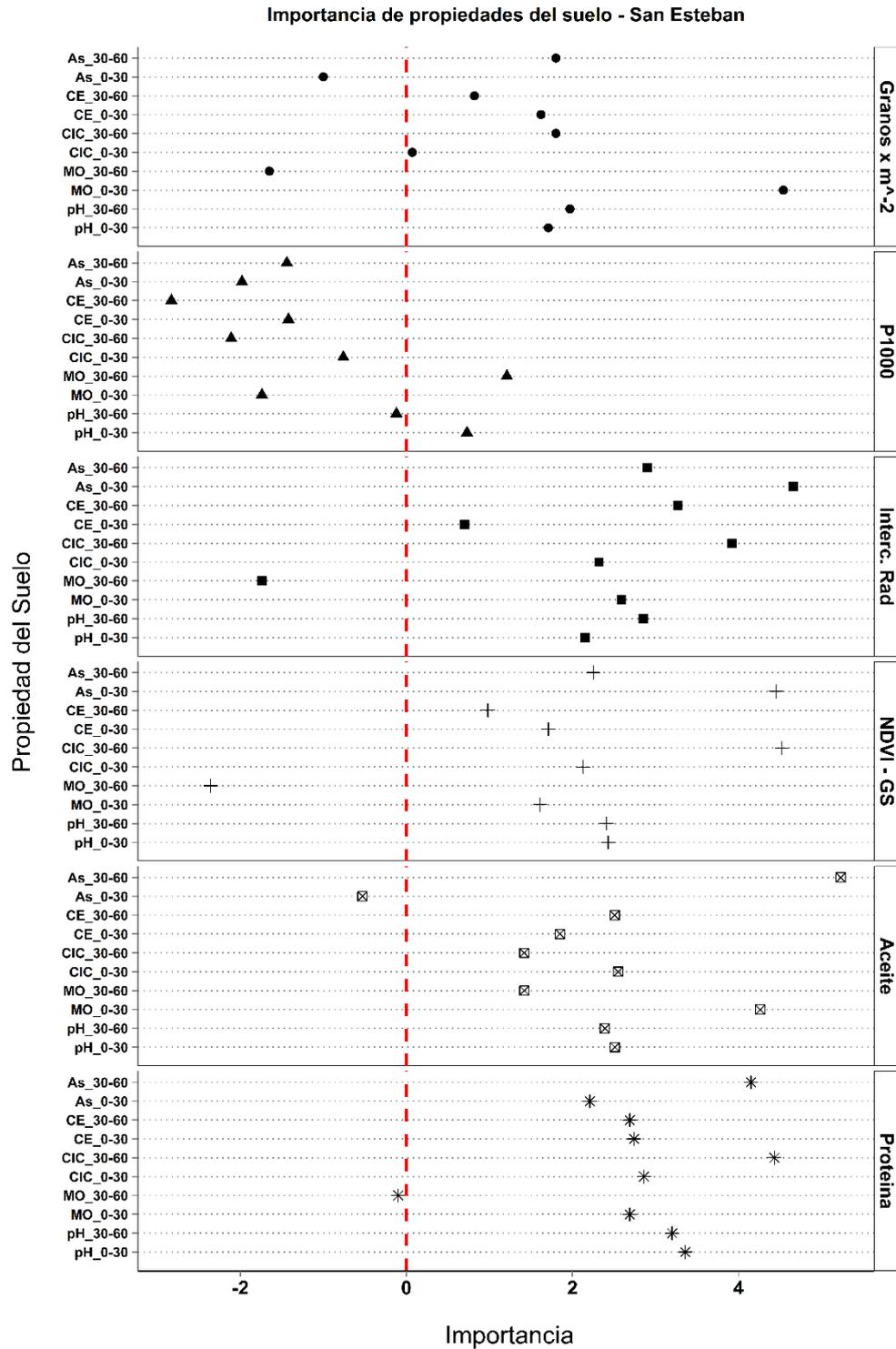
Los resultados reportados para la zona 1 corroboraron que la topografía no es un indicador veraz de las relaciones espaciales entre propiedades del suelo y cultivo en soja (Martin *et al.*, 2007). Estas relaciones están fuertemente influenciadas por el patrón espacial de la tosca, que a su vez determina los patrones de disponibilidad de agua en el suelo. Información obtenida a partir de sensores del suelo (CEa), mapas de rendimiento e imágenes satelitales, pueden complementar eficientemente la caracterización de los patrones espaciales de tosca y por ende, ayudar a entender las relaciones espaciales entre propiedades del suelo y el cultivo.

Las figuras 4.12 y 4.13 muestran la importancia de propiedades del suelo, para cada una de las propiedades de rendimiento y calidad de soja en los lotes de la zona 2, determinado a partir de la función de importancia de variables del algoritmo RF. A través de esta función se logró explicar el 49 y 29% de la varianza de las propiedades del cultivo, para *Mercapire* y *San Esteban*, respectivamente.

En *Mercapire*, excepto para NDVI, MO y contenido de arcilla en ambas profundidades fueron más importantes para todas las variables de rendimiento y calidad de soja. CE del extracto fue la variable de suelo más importante para NDVI. pH tuvo muy poca importancia para todas las propiedades de rendimiento y calidad de soja. Esto se puede explicar por la sequía predominante durante el momento del muestreo de grano. En general, las propiedades del suelo relacionadas con contenido de humedad, fueron las más importantes para propiedades de cultivo. En condiciones de la zona 2, MO y contenido de arcilla están relacionadas cerradamente con disponibilidad de agua. Zonas delimitadas con mayor contenido de arcilla se localizaron en zonas bajas, corroborando que para *Mercapire*, la topografía puede ser un indicador eficaz de las relaciones espaciales entre propiedades del suelo y propiedades del cultivo de soja (Kravchenko y Bullock, 2002b). Así mismo, la importancia de CE del extracto en RI y NDVI, corrobora que el patrón espacial de la CEa también puede ser un indicador eficaz de las relaciones espaciales entre propiedades del suelo y propiedades del cultivo de soja.



**Figura 4.12** Importancia de las propiedades del suelo para cada una de las propiedades del cultivo de soja en el lote Mercapire – Córdoba, a partir de las clasificaciones realizadas con el algoritmo *Random Forest* (RF)

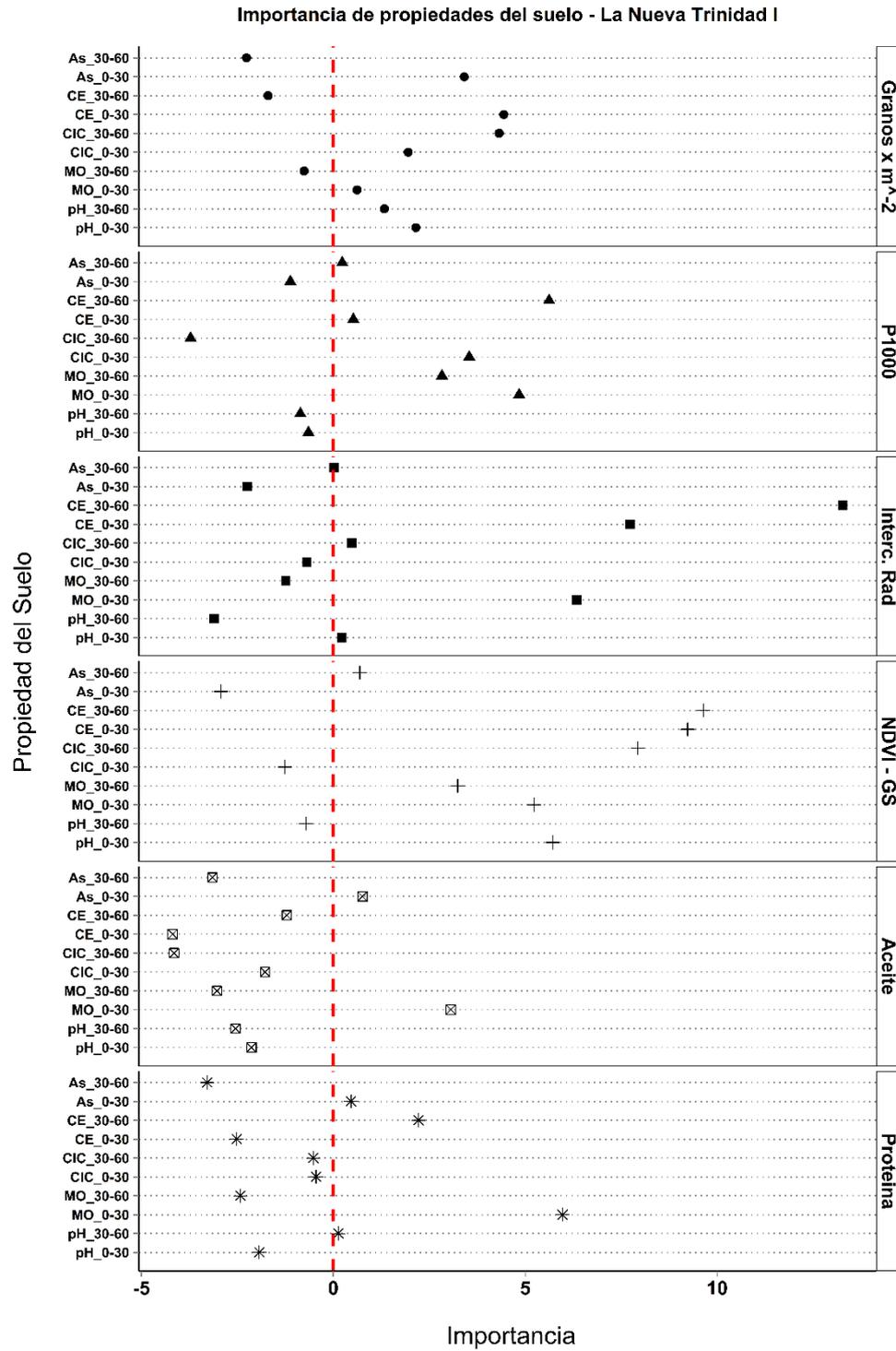


**Figura 4.13** Importancia de las propiedades del suelo para cada una de las propiedades del cultivo de soja en el lote San Esteban – Córdoba, a partir de las clasificaciones realizadas con el algoritmo *Random Forest* (RF)

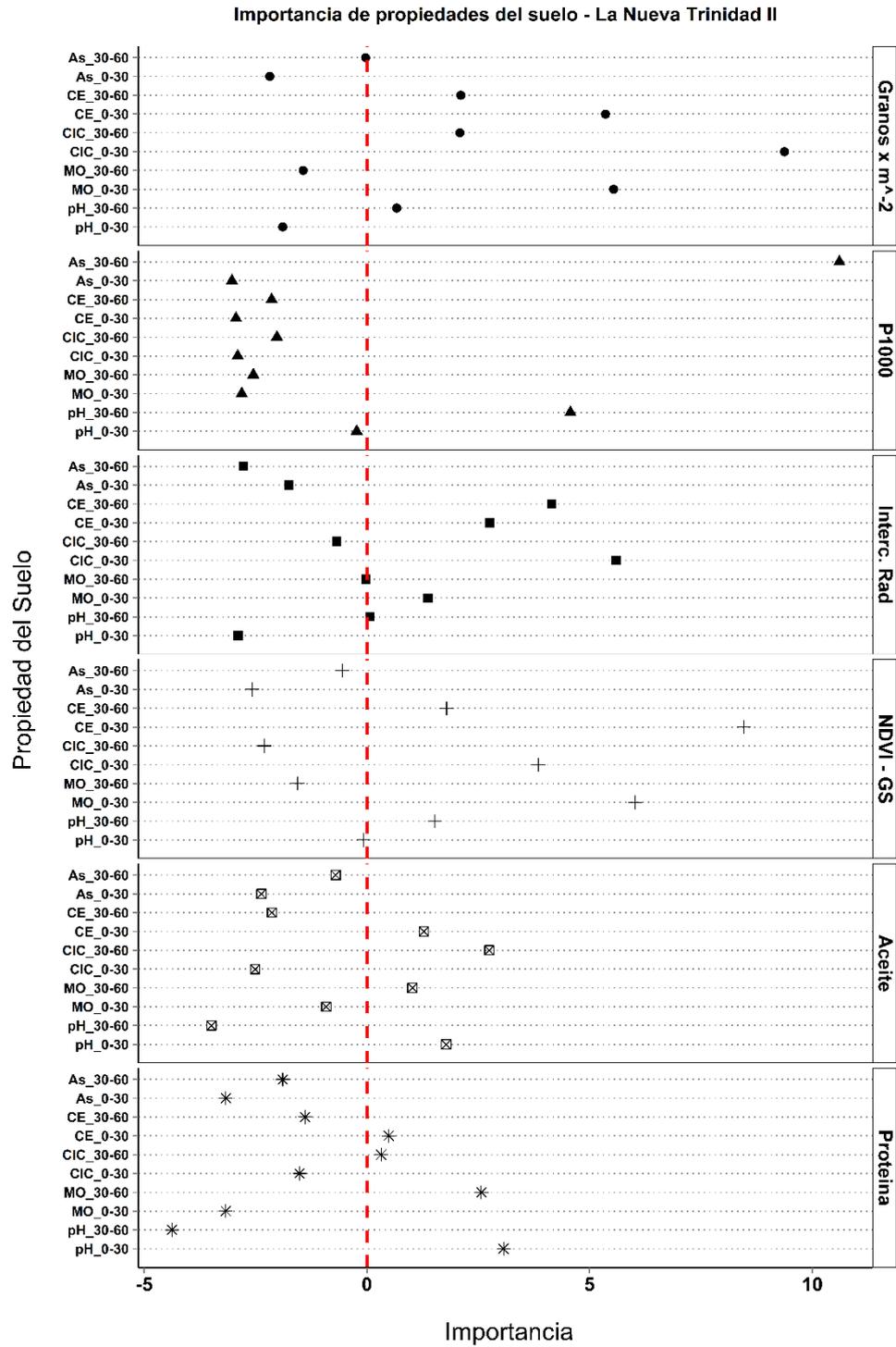
En *San Esteban*, excepto P100, contenido de arcilla, CE del extracto y CIC fueron las variables más importantes para todas las variables de rendimiento y calidad de soja. pH también fue importante para granos  $m^{-2}$ , RI, NDVI y P1000. Este resultado coincide con lo planteado por (Anthony *et al.*, 2012). Como se mencionó anteriormente, estos autores determinaron que pH determina la disponibilidad de nutrientes y por ende puede ser importante en la determinación de la variabilidad de contenido de aceite y proteína. Estos resultados sugieren que debido a la sequía predominante durante la época de recolección de muestras de granos, es posible que el efecto combinado de déficit hídrico y contenido de sales en zonas delimitadas, haya determinado los patrones espaciales de la mayoría de las variables de rendimiento y calidad de soja. Para *San Esteban*, la topografía fue un indicador eficaz de las relaciones espaciales entre propiedades del suelo y propiedades del cultivo de soja. Por su parte, los resultados corroboran que el patrón espacial de la CEa puede ser un indicador eficaz de las relaciones espaciales entre propiedades del suelo y cultivo de soja. A pesar que la función de importancia de variable de RF logró explicar un porcentaje bajo de la varianza de las propiedades de cultivo, estos resultados indican que esta función es eficaz para determinar las propiedades del suelo más relevantes, de acuerdo a cada variable de rendimiento y calidad de soja.

Los resultados reportados para la zona 2 corroboraron que la topografía es un indicador veraz de las relaciones espaciales entre propiedades del suelo y de cultivo de soja, siempre y cuando no existan áreas dentro del lote con altos contenidos de sales. En lotes con presencia de este tipo de zonas salinas, las relaciones espaciales entre propiedades del suelo y cultivo de soja pueden ser determinados mediante la medición de la CEa. Así mismo, la CEa demostró ser un excelente complemento de la topografía para entender las relaciones espaciales entre propiedades del suelo y cultivo, en ausencia de zonas salinas.

Las figuras 4.14 y 4.15 muestran la importancia de propiedades del suelo, para cada una de las propiedades de rendimiento y calidad de soja en los lotes de la zona 23 determinado a partir de la función de importancia de variables del algoritmo RF. A través de esta función se logró explicar el 56 y 54% de la varianza de las propiedades del cultivo, para *Mercapeire* y *San Esteban*, respectivamente.



**Figura 4.14** Importancia de las propiedades del suelo para cada una de las propiedades del cultivo de soja en el lote LNT\_I – Entre Ríos, a partir de las clasificaciones realizadas con el algoritmo *Random Forest* (RF)



**Figura 4.15** Importancia de las propiedades del suelo para cada una de las propiedades del cultivo de soja en el lote LNT\_II – Entre Ríos, a partir de las clasificaciones realizadas con el algoritmo *Random Forest* (RF)

En *LNT\_I*, excepto para las propiedades de calidad de soja, CE del extracto fue la variable del suelo más importante para todas las propiedades de cultivo de soja. Por su parte, MO superficial (0 – 30 cm) fue la variable más importante para las propiedades de calidad del cultivo. Estos resultados se esperaban, ya que las diferencias de valores de cada propiedad de calidad fueron leves entre zonas delimitadas. La presencia de zonas con altos contenidos de sales determinó el patrón espacial de las propiedades de rendimiento, pero no el patrón espacial de las propiedades de calidad. MO está relacionado con la disponibilidad de nitrógeno, lo cual a su vez es importante para la síntesis de proteínas (Anthony *et al.*, 2012). Para *LNT\_I*, estos resultados sugieren que la topografía no es un indicador veraz de las relaciones espaciales entre propiedades del suelo y cultivo de soja. Además, estos resultados ratifican que la CEa solamente, puede llegar a ser un indicador eficaz de las relaciones entre propiedades del suelo y propiedades del rendimiento del cultivo de soja. Las relaciones entre patrones de propiedades de suelo y propiedades de calidad del cultivo no fueron claramente establecidas. Por lo tanto, una vez más se ratifica que las zonas delimitadas por tipo de suelo mediante la metodología propuesta en el capítulo 3, pueden ser más funcionales a manejo sitio específico que a cosecha diferenciada por calidad de soja.

En *LNT\_II*, excepto P1000 y propiedades de calidad, CE del extracto superficial (0 – 30 cm) o profunda (30 – 60 cm) fueron las variables del suelo más importantes para todas las propiedades de cultivo. pH y MO superficial (0 – 30 cm) fueron las propiedades del suelo más importantes para contenido de proteína, mientras que CIC profundo (30 – 60 cm) y pH superficial (0 – 30 cm) lo fueron para contenido de aceite. Por su parte, contenido de arcilla profunda (30 – 60 cm) fue la propiedad del suelo más importante para P1000. Al igual que en *LNT\_I*, estos resultados indican que la presencia de zonas con altos contenidos de sales determinó el patrón espacial de las propiedades de rendimiento, pero no el patrón espacial de las propiedades de calidad. Los resultados de importancia de propiedades del suelo para calidad no fueron consistentes con los reportados para *LNT\_I*, pero si fueron consistentes con los reportados por Anthony *et al.* (2012). En general, las relaciones entre propiedades del suelo y propiedades de calidad del cultivo no fueron claramente establecidas. En *LNT\_II* se ratifica que la topografía no es un indicador veraz de las relaciones espaciales entre propiedades del suelo y cultivo de

soja. Así mismo, se establece que la CEa solamente es un indicador eficaz de propiedades del rendimiento, mas no de calidad de soja. Por lo tanto, se establece que las zonas delimitadas podrían ser más funcionales a manejo sitio específico que a cosecha diferenciada por calidad de soja.

Los resultados reportados para la zona 3 determinaron que la topografía no es un indicador veraz de las relaciones espaciales entre propiedades del suelo y cultivo de soja. Ambos lotes tienen zonas con suelos que tienen altos contenidos de sales. Sin embargo, los patrones espaciales de las propiedades del suelo están relacionadas únicamente con propiedades de rendimiento de soja. Los patrones espaciales de propiedades de suelo no presentaron relaciones espaciales similares con propiedades de rendimiento y con propiedades de calidad de soja, Por lo tanto, la CEa puede ser una herramienta eficaz para delimitar zonas para manejo sitio específico, mas no para delimitar áreas para cosecha diferenciada.

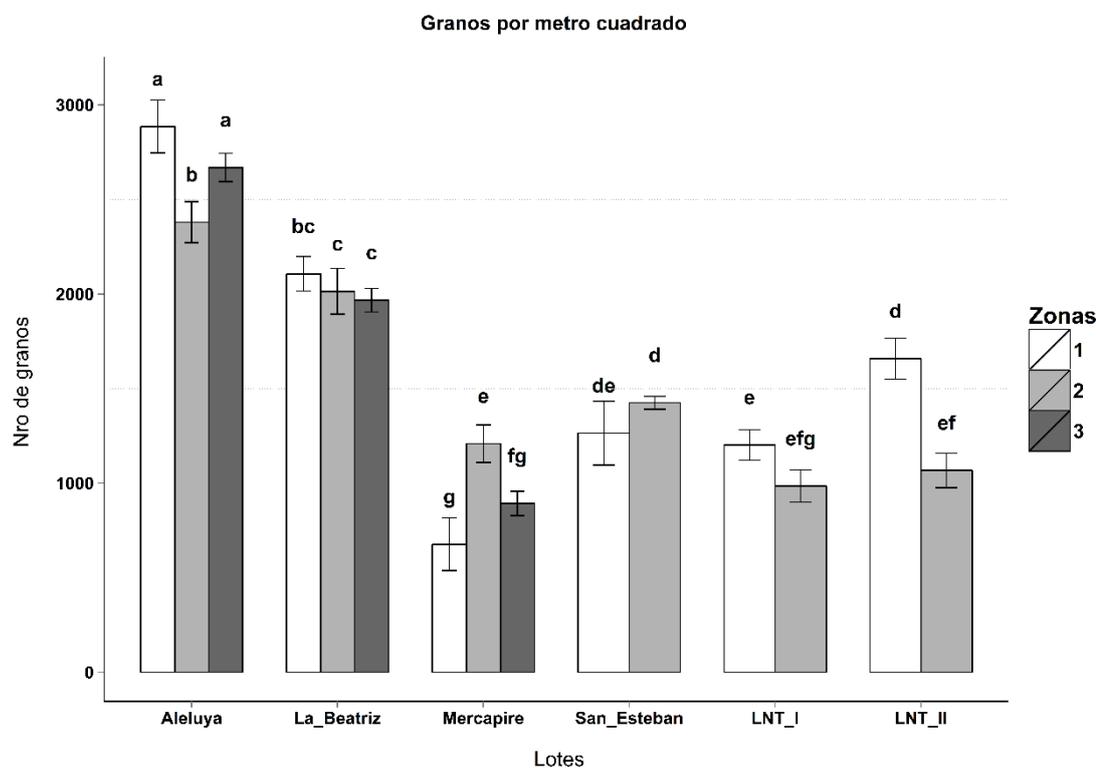
### *3.5. Etapa III: Relaciones espaciales entre zonas por tipo de suelo y propiedades del cultivo*

#### *3.5.1. Granos por metro cuadrado (granos m<sup>-2</sup>)*

La figura 4.16 muestra las diferencias entre medias de granos m<sup>-2</sup> comparadas entre todos los lotes, a través del test de diferencias mínimas significativas.

En general, los resultados corroboran la importancia del efecto ambiental sobre la definición de granos m<sup>-2</sup> en soja (Andrade y Sadras, 2002; Andriani, 2002). Asimismo, se corroboran diferencias significativas de granos m<sup>-2</sup> entre zonas, dentro de cada lote. Los máximos valores de granos m<sup>-2</sup> se presentaron en la zona 1. Este resultado se explica por la adecuada disponibilidad hídrica que tuvieron los cultivos durante enero de 2011 en dicha zona. Los valores de granos m<sup>-2</sup> fueron similares en los lotes de la zona 2 y 3. Esto se debe a diferentes causas. En la zona 2, el efecto de la sequía fue drástico sobre granos m<sup>-2</sup>. Por su parte, en la zona 3, el efecto de presencia de sales y situaciones puntuales de inundaciones, determinaron los valores bajos de granos m<sup>-2</sup>. Estos resultados de granos

m<sup>-2</sup> en tres regiones agrícolas de Argentina, plantean la viabilidad de implementar técnicas de CDS para la delimitación de zonas para manejo sitio específico de cultivos.



**Figura 4.16** Comparación de medias de *número de granos por metro cuadrado* entre lotes, de acuerdo a las zonas delimitadas por tipo de suelo. Letras diferentes indican diferencias estadísticamente significativas ( $p < 0.05$ ). Barras verticales indican error estándar de la media para cada zona delimitada por tipo de suelo

En la zona 1, *Aleluya* presentó los máximos valores de granos m<sup>-2</sup>. Dentro de este lote, en la zona delimitada 2 se presentaron los valores mínimos de granos m<sup>-2</sup>. Como se esperaba, existe diferencia significativa ( $p < 0.05$ ) entre las zonas delimitadas 2 y las demás. La profundidad efectiva determinó los patrones espaciales de las propiedades del suelo y la capacidad efectiva de almacenamiento de agua. De este modo, la zona delimitada 2 presentó los suelos más someros. Debido a esto, estos suelos limitaron la capacidad de definición de granos m<sup>-2</sup>. Las zonas delimitadas 1 y 3 presentaron suelos profundos, que le permitieron almacenar mayor cantidad de agua que la zona delimitada 2. Por ende, en estas zonas, se definieron la mayor cantidad de granos m<sup>-2</sup>. Por otro lado, en *La Beatriz*, el patrón de diferencias significativas entre zonas, corroboró el efecto de la presencia de suelos somero en la definición de granos m<sup>-2</sup>. En este lote, la zona

delimitada 3 estuvo relacionada con suelos someros. En general para la zona 1, la topografía no fue un indicador veraz de la variabilidad de granos  $m^{-2}$ . CEa presentó patrones espaciales relacionados con granos  $m^{-2}$ , indicando que es un potencial indicador veraz del patrón espacial de granos  $m^{-2}$ . Mapas de rendimiento tuvo una correspondencia espacial con el patrón espacial de granos  $m^{-2}$ . Este resultado difiere de lo planteado por Leonardi (2012), quien no encontró relación entre mapas de rendimiento de soja y granos  $m^{-2}$  en zonas delimitadas por profundidad de tosca y topografía en lotes del sudeste bonaerense. NDVI no presentó un comportamiento estable temporalmente y en general, la dependencia de la resolución espacial limita su potencial como indicador veraz del patrón espacial granos  $m^{-2}$ . De acuerdo con estos resultados, para los lotes de la zona 1, el patrón espacial de granos  $m^{-2}$  tiene una alta dependencia del patrón espacial de profundidad del suelo. Información complementaria a partir de la medición de CEa, mapas de rendimiento y en menor medida NDVI a partir de imágenes satelitales ayudó a caracterizar el patrón espacial de tosca y por ende, el patrón espacial de granos  $m^{-2}$ .

En la zona 2, *Mercapire* presentó los menores valores de granos  $m^{-2}$ . Dentro de este lote, en la zona delimitada 2 se presentaron los mayores valores de granos  $m^{-2}$ . Existen diferencia significativa ( $p < 0.05$ ) entre la zona delimitada 2 y las demás. La topografía determina los patrones espaciales de contenido de arcilla y MO. Como se demostró, estas dos propiedades son las más importantes para determinar granos  $m^{-2}$ . La zona delimitada 2 estuvo relacionada con suelos con mayor contenido de arcilla y MO, los cuales en condiciones de sequía, permitieron mayores valores de granos  $m^{-2}$ . Por su parte, en *San Esteban* el patrón de diferencias significativas entre zonas, corroboró el efecto de la presencia de suelos salinos en la definición de granos  $m^{-2}$ . En este lote, la zona delimitada 1 estuvo relacionada con suelos salinos. En general, para la zona 2 la topografía fue un indicador veraz de la variabilidad de granos  $m^{-2}$ . CEa30 tuvo patrones espaciales inversos a CEa90 en *San Esteban*, pero no en *Mercapire*. Por lo tanto, se determinó que únicamente CEa puede ser un indicador veraz en la medida de acuerdo a las proporciones espaciales de presencia de suelos salinos y no salinos dentro de un lote. Mapas de rendimiento y NDVI fueron consistentes con los patrones espaciales de granos  $m^{-2}$ . De acuerdo con estos resultados, para los lotes de la zona 1, el patrón espacial de la topografía, complementado con mapas de rendimiento, NDVI y CEa, en presencia de

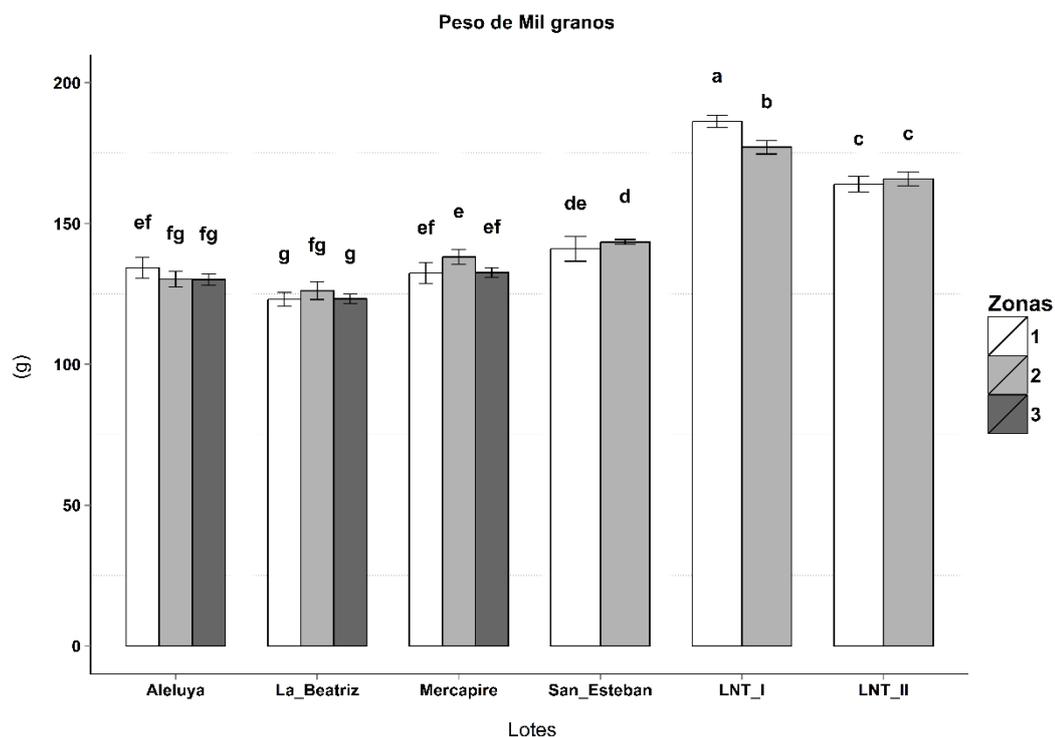
suelos salinos, ayudó a caracterizar el patrón espacial del contenido de arcilla y MO y por ende, el patrón espacial de granos  $m^{-2}$ .

En la zona 3, en ambos lotes, valores altos de granos  $m^{-2}$  estuvieron relacionados a las zonas delimitadas 2. Estas zonas a su vez, estuvieron relacionadas con presencia de suelos salinos. En general, para los lotes de la zona 3 la topografía no fue un indicador veraz del patrón espacial de granos  $m^{-2}$ . CEa presentó patrones espaciales relacionados a los patrones espaciales de granos  $m^{-2}$ . Así mismo, mapas de rendimiento y NDVI obtenido a partir de imágenes satelitales tuvieron una correspondencia espacial con el patrón espacial de granos  $m^{-2}$ . De acuerdo con estos resultados, para los lotes de la zona 3, el patrón espacial de la CEa, complementado con mapas de rendimiento y NDVI, ayudó a caracterizar los patrones espaciales de suelos salinos y por ende, el patrón espacial de granos  $m^{-2}$ .

### 3.5.2. *Peso de mil granos (P1000)*

La figura 4.17 muestra las diferencias entre medias de P1000 comparadas entre todos los lotes, a través del test de diferencias mínimas significativas.

En general, los resultados corroboran que P1000 no es una variable que depende de la oferta ambiental (Baigorri, 2004; Cober y D Voldeng, 2000; Hu y Wiatrak, 2012). Excepto en LNT\_I, se corroboró que existen escasas diferencias de P1000 entre zonas delimitadas, dentro de todos los lotes. Valores altos de P1000 en los lotes de la zona 3, se explican por el tipo de semilla que se utilizó, la cual se caracterizó por tener alto potencial de P1000. Las diferencias significativas entre las zonas delimitadas en LNT\_I, se explican por la excesiva cantidad de sales que se encuentran en los suelos de la zona delimitada 2. Sin embargo, no se encontraron trabajos que hayan estudiado específicamente, el efecto de suelos salinos en P1000 en soja. La duración del estrés en soja, especialmente en la etapa vegetativa del cultivo, es la que determina el patrón de variabilidad espacial de P1000 (Dornbos et al., 1989). De acuerdo con esto, es posible que el efecto simultaneo de la presencia de sales y la sequía, determinaran las diferencias de P1000 entre las zonas delimitadas en LNT\_I.



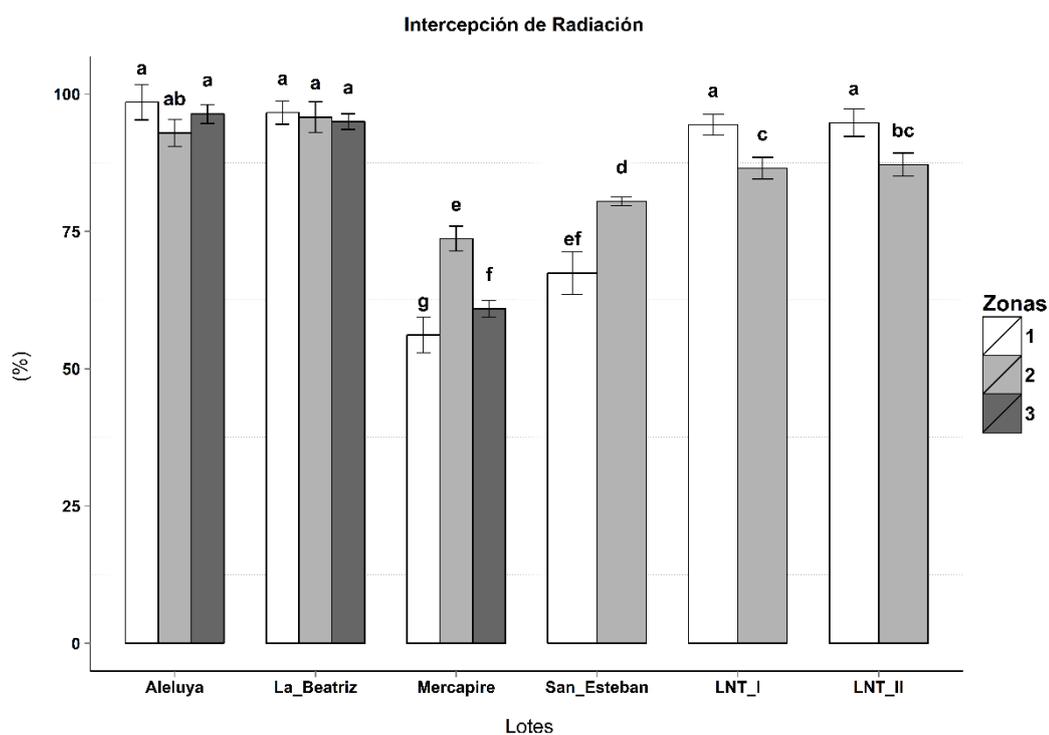
**Figura 4.17** Comparación de medias de *peso de mil granos* ( $P1000$ ) entre lotes, de acuerdo a las zonas delimitadas por tipo de suelo. Letras diferentes indican diferencias estadísticamente significativas ( $p < 0.05$ ). Barras verticales indican error estándar de la media para cada zona delimitada por tipo de suelo

### 3.5.3. Radiación interceptada (RI)

La figura 4.18 muestra las diferencias entre medias de RI comparadas entre todos los lotes, a través del test de diferencias mínimas significativas.

En general, RI presentó una correspondencia espacial y de magnitud similar a granos  $m^{-2}$ , en todas las zonas. En este trabajo se asume que las causas de las variaciones de RI están cerradamente relacionadas con las causas que determinaron las variaciones de granos  $m^{-2}$ . Ha sido ampliamente documentado que la RI determina la generación de biomasa total (Andrade, 1995; Andrade y Ferreiro, 1996) y por ende condiciona la relación fuente/destino. Una disminución en los valores de RI reduce el rendimiento del cultivo en la medida que el cultivo sea afectado en la etapa más crítica de definición de rendimiento, que es a partir de la etapa fenológica R4 (Andrade y Sadras, 2002; Satorre *et al.*, 2003). Así mismo, granos  $m^{-2}$  es el componente de rendimiento que más se afecta

durante un estrés en el periodo crítico en el cultivo de soja (Andrade y Ferreiro, 1996). Los resultados de RI corroboran lo planteado para granos  $m^{-2}$ .



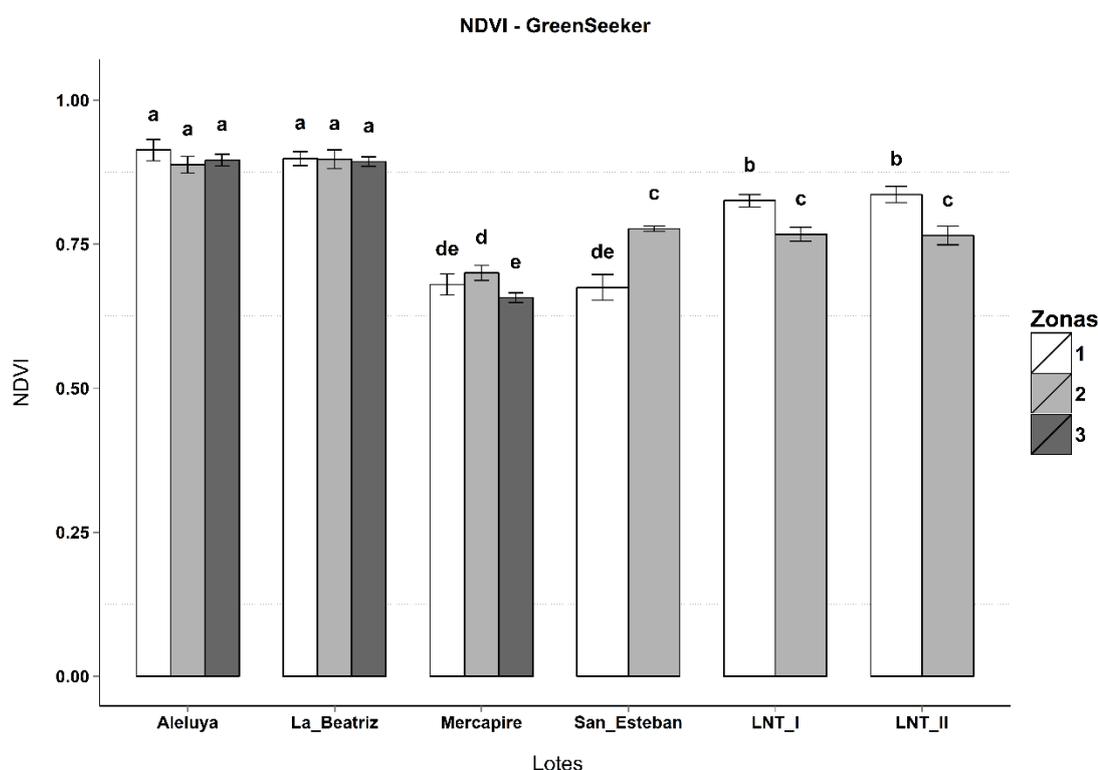
**Figura 4.18** Comparación de medias de *radiación interceptada (RI)* entre lotes, de acuerdo a las zonas delimitadas por tipo de suelo. Letras diferentes indican diferencias estadísticamente significativas ( $p < 0.05$ ). Barras verticales indican error estándar de la media para cada zona delimitada por tipo de suelo

#### 3.5.4. NDVI – GreenSeeker

La figura 4.19 muestra las diferencias entre medias de NDVI comparadas entre todos los lotes, a través del test de diferencias mínimas significativas.

En general, NDVI presentó una correspondencia espacial similar a granos  $m^{-2}$  y RI. Sin embargo, en las magnitudes se presentaron diferencias puntuales. Específicamente, en los lotes de la zona 1, las medias de NDVI no presentaron diferencias significativas ( $p < 0.05$ ) entre zonas delimitadas. Aunque espacialmente tuvieron una tendencia similar a RI. Estos resultados sugieren que la efectividad de la medición de NDVI utilizando el equipo GreenSeeker® depende de la severidad del estrés del cultivo de soja. Como se observa, en lotes sometidos a estrés hídrico y salino severos, el NDVI tuvo una

correspondencia espacial con RI y granos  $m^{-2}$ . Estos resultados se pueden explicar debido a la plasticidad vegetal del cultivo de soja. La gran producción de ramas secundarias y terciarias y su despliegue de área foliar, hace que la soja tenga una susceptibilidad moderada a cambios en el índice de área foliar y por ende a las lecturas puntuales de NDVI (Dornbos *et al.*, 1989; Fehr *et al.*, 1971; Zhitao *et al.*, 2014). De acuerdo con esto, es posible que el uso del equipo GreenSeeker® tenga una aplicación práctica para el cultivo de soja, de acuerdo al tipo y severidad de estrés al que está sometido.



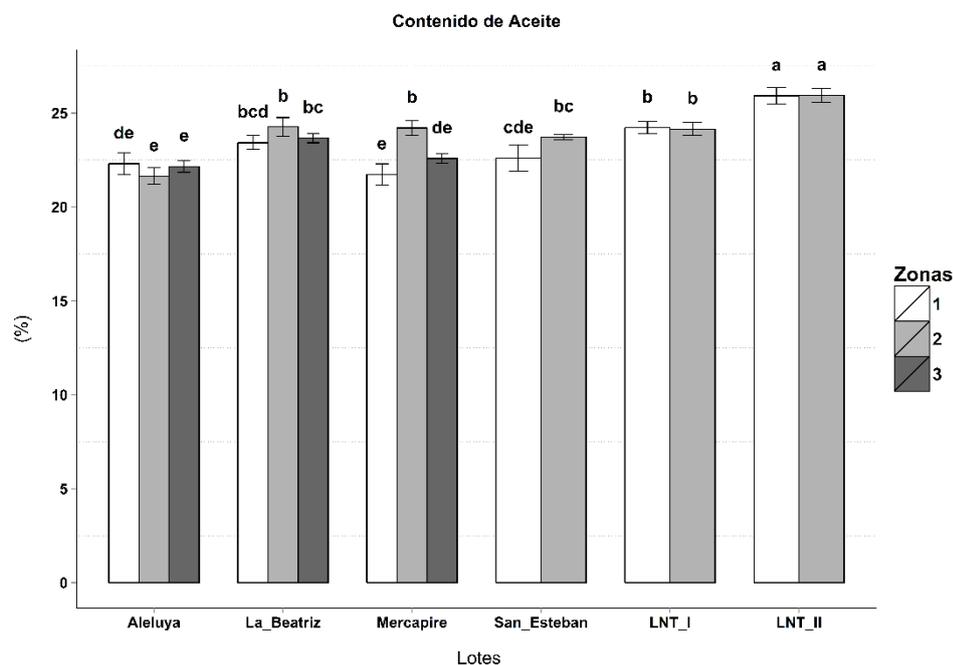
**Figura 4.19** Comparación de medias de *NDVI* medido a través de GreenSekker® entre lotes, de acuerdo a las zonas delimitadas por tipo de suelo. Letras diferentes indican diferencias estadísticamente significativas ( $p < 0.05$ ). Barras verticales indican error estándar de la media para cada zona delimitada por tipo de suelo

### 3.5.5. Contenido de Aceite

La figura 4.20 muestra las diferencias entre medias de contenido de aceite comparadas entre todos los lotes, a través del test de diferencias mínimas significativas.

En general, contenido de aceite presentó una correspondencia espacial similar a granos  $m^{-2}$ , RI y NDVI. En este trabajo se asume que las causas de las variaciones de granos  $m^{-2}$

<sup>2</sup>, RI y NDVI podrían estar cerradamente relacionadas con las causas que determinaron las variaciones de contenido de aceite. Las magnitudes de las diferencias de medias de contenido de aceite presentaron diferencias puntuales con respecto a granos m<sup>-2</sup> y RI. Específicamente en los lotes de la zona 3, las medias de contenido de aceite no presentaron diferencia significativa ( $p < 0.05$ ) entre zonas delimitadas. Sin embargo, *LNT\_II* tuvo medias de contenido de aceite significativamente mayores ( $p < 0.05$ ) que *LNT\_I*. Estos resultados sugieren que el efecto por estrés salino afecta más a las propiedades del rendimiento que a propiedades de calidad de soja. Además se ratifica lo planteado en la determinación de la importancia de propiedades del suelo en las propiedades de rendimiento y calidad de soja. CE del extracto fue una variable levemente importante para contenido de aceite. Propiedades tales como MO y pH fueron más importantes. Aunque existieron claras diferencias de MO superficial entre zonas delimitadas para los lotes de la zona 3, para pH estas diferencias no fueron claras. De acuerdo con esto, los resultados para los lotes de la zona 3, ratifican que la metodología de delimitación de zonas por tipo de suelo propuesta en el capítulo tres de esta tesis, es más útil para delimitar zonas para manejo sitio específico de cultivos y no para delimitar zonas de cosecha diferenciada por calidad. Por el contrario, en los lotes de las zonas 1 y 2, la metodología propuesta puede llegar a ser útil para delimitar ambas, zonas para manejo sitio específico y zonas para cosecha diferenciada por calidad.



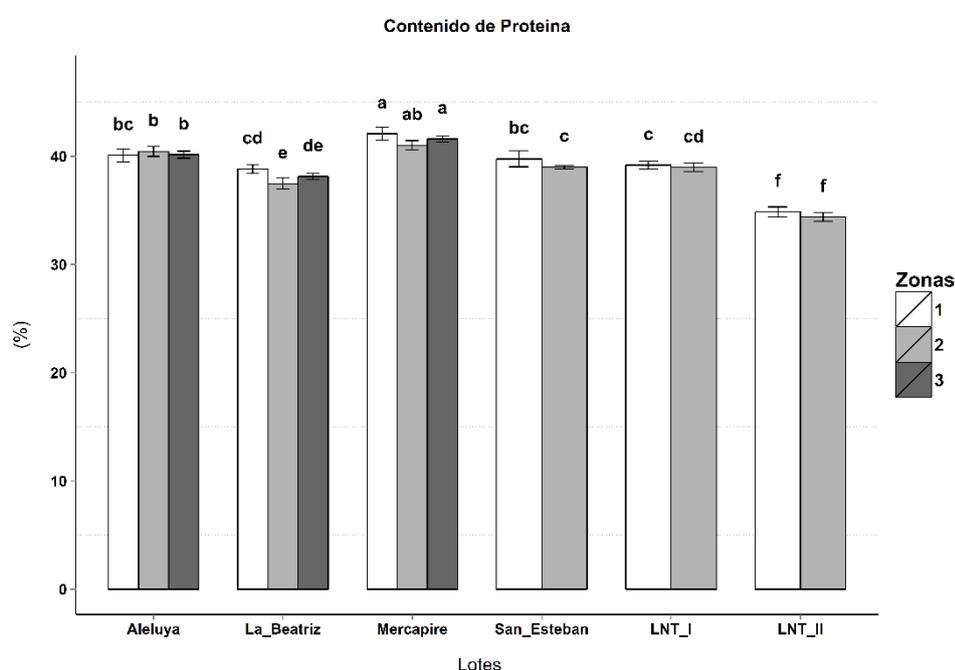
**Figura 4.20** Comparación de medias de *contenido de aceite en grano de soja* entre lotes, de acuerdo a las zonas delimitadas por tipo de suelo. Letras diferentes indican diferencias estadísticamente significativas ( $p < 0.05$ ). Barras verticales indican error estándar de la media para cada zona delimitada por tipo de suelo

### 3.5.6. Contenido de Proteína

La figura 4.21 muestra las diferencias entre medias de contenido de proteína comparadas entre todos los lotes, a través del test de diferencias mínimas significativas.

En general, contenido de proteína presentó una contraria correspondencia espacial a contenido de aceite. En este trabajo se asume que las causas de variaciones de contenido de aceite están cerradamente relacionadas con las causas de variaciones de contenido de proteína. Al igual que contenido de aceite, contenido de proteína presentó diferencias puntuales con respecto a su relación inversa con granos  $m^{-2}$  y RI. Específicamente en los lotes de la zona 3, las medias de contenido de proteína no presentaron diferencia significativa ( $p < 0.05$ ) entre zonas delimitadas. Los resultados para los lotes de la zona 1 y 2 sugieren que los efectos de las propiedades del suelo que estuvieron relacionados con contenido de aceite, también estuvieron relacionados con contenido de proteína. Es ampliamente conocido que (i) el contenido de aceite tiene una relación cerradamente inversa con contenido de proteína en soja; y (ii) el contenido de proteína aumenta cuando

el cultivo de soja ha sufrido algún tipo de estrés abiótico, sobre todo en la etapa de llenado de grano. De acuerdo con esto, para las zonas 1 y 2 únicamente, la metodología de delimitación de zonas por tipo de suelos puede llegar a ser útil para delimitar zonas para manejo sitio específico y zonas de cosecha diferenciada por calidad. La situación de los componentes de calidad de grano en la zona 3, deja una brecha de conocimiento que debe ser específicamente estudiada. El efecto del estrés salino sobre los componentes de calidad de grano de soja, no pudieron ser claramente determinados en este trabajo.



**Figura 4.21** Comparación de medias de *contenido de proteína en grano de soja* entre lotes, de acuerdo a las zonas delimitadas por tipo de suelo. Letras diferentes indican diferencias estadísticamente significativas ( $p < 0.05$ ). Barras verticales indican error estándar de la media para cada zona delimitada por tipo de suelo

#### 4. Conclusiones

Caracterizar y analizar las relaciones espaciales entre propiedades del suelo y propiedades de rendimiento y calidad de soja a escala de lote es importante para delimitar zonas para manejo sitio específico y para zonas de cosecha diferenciada de calidad de grano. Debido a esto se propuso y discutió una técnica de CDS que permita generar

información espacial veraz de las relaciones entre patrones espaciales de propiedades del suelo de rendimiento y calidad de soja.

Específicamente, en este capítulo se evaluó la pertinencia de utilizar las zonas delimitadas por tipo de suelo en cada lote, elaboradas a partir de la implementación del algoritmo de DgUCS propuesto en el capítulo 3. Estas zonas se utilizaron como un marco de conocimiento de los patrones espaciales de propiedades del suelo, que permitiera plantear delimitaciones eficientes de zonas para manejo sitio específico y para cosecha diferenciada por calidad.

En general, la metodología de DgUCS demostró ser eficiente para delimitar zonas para manejo sitio específico para todos los lotes. Las relaciones entre patrones espaciales de propiedades del suelo y propiedades de rendimiento de soja pudieron ser caracterizadas y entendidas a través de la delimitación de zonas por tipo de suelo. Sin embargo, no ocurrió lo mismo con los patrones de propiedades de calidad de grano. Especialmente en los lotes del centro de Entre Ríos, el efecto de estrés salino e hídrico no permitió determinar con claridad sus efectos sobre componentes de calidad de grano. Debido a esto, no se demostró que la metodología de DgUCS pudiera ser eficaz, para la delimitación de zonas para cosecha diferenciada por calidad. Estos resultados invitan a seguir estudiando metodologías alternativas de delimitación de zonas por tipo de suelo, que aclaran y permitan caracterizar las relaciones espaciales entre patrones espaciales de estrés salino y patrones espaciales de calidad de grano de soja.

Así mismo se encontró que la topografía no demostró ser un indicador veraz de las relaciones de patrones espaciales de propiedades de suelo y cultivo, para todos los lotes, tal y como lo planteaban trabajos preliminares. En los lotes del sudeste bonaerense, las relaciones entre elevación y propiedades del cultivo de soja no fueron claras. Sin embargo, las relaciones entre profundidad de tosca y propiedades del cultivo tuvieron una similar y coherente correspondencia y magnitud espacial. En los lotes del sur de Córdoba, la topografía, complementada con CEa, mapas de rendimiento y NDVI demostraron ser un indicador eficaz de las relaciones entre patrones espaciales de propiedades de suelo y cultivo, incluso ante presencia de zonas salinas. Por último, en los lotes del centro de

Entre Ríos, la CEa, complementada en algunos casos por mapas de rendimiento y NDVI demostró ser un indicador veraz de las relaciones entre patrones espaciales de propiedades del suelo y rendimiento de soja, únicamente.

La metodología para determinar la importancia de propiedades del suelo sobre las propiedades de cultivo permitió determinar con puntualidad, las propiedades del suelo que limitaron las propiedades de rendimiento y calidad del cultivo. Por ejemplo, gracias a esta metodología, se pudo determinar que propiedades del suelo fueron importantes para las propiedades de calidad de grano en los lotes de la zona 3. En las demás zonas, la metodología para determinar la importancia de propiedades del suelo, permitió aclarar las relaciones de los patrones espaciales entre topografía CEa, propiedades del suelo y de cultivo.

De acuerdo con esto, la hipótesis (i) el patrón espacial de propiedades de rendimiento y calidad de soja a escala de lote, están distribuidos de acuerdo al patrón espacial de las propiedades del suelo, fue no aceptada. Por su parte, la hipótesis (ii) la delimitación de zonas por tipo de suelos, permiten generar una base técnica y espacial para delimitar zonas para manejo sitio específico y zonas de cosecha diferenciada por calidad, fue aceptada parcialmente.

La extrapolación de estos resultados en otros lotes dentro y fuera de las zonas estudiadas, requiere ser evaluado. Como se observó, las relaciones entre propiedades del suelo y cultivo son complejas. En la actualidad, son muchos los factores de degradación del suelo y de efecto del cambio climático, los que determinan los patrones espaciales de propiedades del cultivo. En este capítulo se determinó que otras fuentes de información auxiliar relacionadas con factores formadores del suelo y que involucren factores temporales de oferta ambiental, son necesarios para delimitar zonas para cosecha diferenciada por calidad. Estudios adicionales son requeridos para (i) comparar otros tipos de algoritmos de delimitación de zonas, que tengan en cuenta de manera simultánea tipos de suelos y propiedades de cultivo. (ii) comparar otros tipos de algoritmos para la determinación de importancia de propiedades del suelo en la definición de los patrones

espaciales de propiedades del cultivo. *(iii)* determinar el efecto del patrón espacial de los suelos salinos sobre los componentes de calidad de soja.

## CAPITULO 5

### Discusión General

Las técnicas de cartografía digital de suelos (CDS) tienen un gran potencial para mejorar la consistencia, precisión, detalle y velocidad con la cual se produce la información espacial de propiedades del suelo a escala de lote (Hempel et al., 2008). Para esto, estas técnicas utilizan (i) técnicas estadísticas, minería de datos y/o aprendizaje automático; y (ii) información obtenida a partir de la cartografía convencional de suelos disponible y fuentes de información auxiliar de covariables ambientales o factores formadores del suelo. CDS puede mejorar notablemente la aplicación de la agricultura de precisión y el uso de modelos de simulación hidrológicos (Dobos y Hengl, 2009; McBratney *et al.*, 2005). Por ende, puede generar simultáneamente procesos que mitiguen el impacto ambiental dentro del sistema agroproductivo y el mejoramiento del rendimiento económico. Sin embargo, limitada cantidad de trabajos han estudiado la aplicación de técnicas de CDS en condiciones de la pampa argentina (Angelini *et al.*, 2015). La bibliografía disponible de técnicas de CDS para generar información de suelos a escala de lote no es abundante (Schmidt *et al.*, 2014; Smith y Halvorson, 2011). Además, no existe un consenso general de una técnica de CDS que sea útil para todas las condiciones de suelo. En ese contexto, los experimentos realizados en esta tesis doctoral fueron novedosos en tanto permitieron, mediante la propuesta de uso de esquemas de muestreo basado en modelo, técnicas de desagregación espacial de unidades cartográficas de suelo y evaluación de las relaciones entre patrones espaciales de suelos y cultivo de soja, explorar el potencial de las técnicas de CDS en condiciones de tres de las zonas agroproductivas más importantes de la Argentina.

La integración de esquemas de muestreo de suelos basados en modelo (EBM) y Random Forest, a partir de información auxiliar de factores formadores del suelo fue validada como técnica de generación de información de suelos a escala de lote. El uso de EBM permitió capturar de manera eficiente la distribución de los predictores determinados a partir de la información auxiliar disponible. A partir de esa captura, fue posible aumentar la capacidad predictiva de la regresión Random Forest. Estos resultados

son novedosos, ya que hasta hace unos años, no se contaba ampliamente ni con los algoritmos, ni los programas y ni mucho menos con las computadoras para hacerlo. Tradicionalmente la selección de sitios de muestreo se basaba en la experiencia de los productores o pedólogos y/o algoritmos que no tenían en cuenta la dependencia espacial, ni mucho menos la información multivariada. En el capítulo 2 se demostró que los EBM requieren menor cantidad de muestras que los esquemas tradicionales, para poder capturar de manera eficiente toda la variabilidad de las propiedades del suelo a escala de lote. Además, los EBM demostraron que pueden aprovechar toda la información auxiliar relacionada con factores formadores del suelo, que hoy en día es abundante para todos los lotes agrícolas de la pampa Argentina. Por otra parte, en el capítulo 2 se corroboró que Random Forest es un algoritmo de regresión que permite generar modelos predictivos más eficientes y que simultáneamente, permiten entender y desglosar la importancia predictiva que tienen la información auxiliar disponible, relacionada con factores formadores del suelo. Random Forest proporcionó un significativo mejoramiento de información de patrones espaciales de propiedades del suelo con respecto a lo que existe en la actualidad. Lo que es pertinente destacar es que el proceso de predicción de propiedades del suelo se realizó de manera rápida, fácil, precisa y a bajo costo. Esto es importante si se quiere implementar en aplicaciones de agricultura de precisión, o específicamente en manejo sitio específico. Aunque es claro que los resultados obtenidos deben ser validados en otras condiciones dentro y fuera de las condiciones del sudeste bonaerense, lo demostrado en el capítulo 2 permite corroborar el potencial que tienen las técnicas de CDS para mejorar la información de suelo a escala de lote.

A pesar de los resultados obtenidos en el capítulo 2, es pertinente aclarar que en la aplicación de técnicas de CDS para obtener información de los patrones espaciales de propiedades del suelo, queda mucho por evaluar, corroborar y explorar. En ese contexto, es conveniente que futuros estudios implementen y comparen otros tipos de algoritmos de EMB y de regresión, en condiciones similares a las tenidas en cuenta en este estudio o en otras zonas agroproductivas de la Argentina. Otros algoritmos de regresión, comúnmente utilizados en aprendizaje automático o “*machine learning*”, tales como máquinas vectoriales de soporte, árboles de regresión y clasificación, o redes neuronales han sido eficientes para predecir los patrones espaciales de propiedades del suelo,

principalmente a escala regional (Ahmad *et al.*, 2010; Brungard *et al.*, 2015). Aunque Random Forest es eficiente en aplicaciones de CDS (Grimm *et al.*, 2008; Heung *et al.*, 2014; Ließ *et al.*, 2012; Rad *et al.*, 2014), es cierto que estas determinaciones se han realizado en condiciones de suelos particulares de otras zonas agrícolas del mundo y con información auxiliar que en muchos casos, fue mucho más limitada que la que se dispone actualmente en la Argentina. El diseño de futuros experimentos que incluyan la integración de varios algoritmos de EBM y regresión, ayudará a corroborar los resultados obtenidos o a replantear nuevos tipos de integraciones específicas de algoritmos más eficientes para la predicción de propiedades del suelo a escala de lote. Será importante realizar experimentos en lotes con diversas condiciones y con complejos patrones espaciales de suelos, de manera que se determine la robustez de las predicciones. Además que permitan determinar el potencial predictivo de cada fuente de información auxiliar. Por ejemplo, en el sudeste bonaerense va a ser importante probar y evaluar la eficiencia productiva de diversas combinaciones de algoritmos para plantear la predicción del patrón espacial de la tosca a escala de lote. Como se observó, la profundidad de la tosca determina los patrones espaciales de la mayoría de las propiedades del suelo, del rendimiento y la calidad de los cultivos.

En futuros trabajos también se deben evaluar nuevas fuentes de información de factores formadores del suelo o de covariables ambientales. Al respecto, queda mucho por explorar (Minasny y McBratney, 2010). Sensores del suelo tales como índice de penetración de cono, georradars, LIDAR, NIR, imágenes de radar, entre otros, han demostrado tener un potencial importante en técnicas de CDS (Adamchuk *et al.*, 2004; Allred *et al.*, 2008; Doolittle *et al.*, 2006; Rossel *et al.*, 2010). Nuevas fuentes de imágenes satelitales y aerotransportadas, hasta hace un tiempo muy limitadas de conseguir o inexistentes, emergen actualmente como una nueva fuente de información de factores formadores del suelo. En la actualidad, gracias a la gestión de CONAE e INTA, es posible acceder a imágenes de alta resolución SPOT 7 (ver catálogo de imágenes: <http://www.conae.gov.ar>). Asimismo, recientemente fue lanzado el satélite *Sentinel 2*, por parte de la Agencia Espacial Europea (ESA), el cual promete complementar y mejorar la información relacionada con diversos componentes ambientales. Esta información promete ser revolucionaria para la agricultura debido a que la distribución espectral de

las bandas está pensada para estudiar específicamente la vegetación. Por otro lado, los aviones no tripulados (UAV) o drones también prometen revolucionar la manera en que se adquiere información espectral. La autonomía temporal y la altísima resolución espacial son demandas que hasta el momento la información satelital no ha podido cubrir. El panorama tecnológico venidero nos invita a replantear la manera en que estamos generando información de suelo y sin duda, nos está abriendo una puerta a un mundo de conocimientos edafológicos, que apenas intuimos.

Los resultados obtenidos en el Capítulo 2 determinaron que la integración entre EBM y Random Forest, permite generar de manera eficiente, rápida, precisa y a bajo costo, información de los patrones espaciales de propiedades del suelo, a escala de lote. En ese contexto, **en esta tesis se acepta el planteamiento de hipótesis 1**, la cual enunciaba que “*la integración de fuentes de información de covariables ambientales, esquemas de muestreo basados en modelos y el algoritmo de regresión “Random Forest”, permite generar de manera rápida, precisa y puntual, modelos eficientes de predicción de propiedades del suelo a escala de lote*”. Estos resultados complementan trabajos previos de CDS realizados en otras regiones agrícolas del mundo. Por ejemplo, Rad *et al.* (2014) determinó que la combinación de hipercubo latino condicionado y Random Forest, permitió mejorar y aumentar la información de suelos loessicos localizados en el norte de Iran. Por su parte, Schmidt *et al.* (2014) determinaron que la combinación de hipercubo latino condicionado y Random Forest es óptima para obtener información de suelos a escala de lote. Estos autores concluyeron que las complejidades de los patrones espaciales de propiedades del suelo, pueden ser descritas y resueltas mediante la integración de hipercubo latino condicionado y Random Forest. Por su parte, puntualmente este trabajo ratificó que la integración de estos algoritmos, es eficiente para predecir contenido de arcilla y MO a escala de lote, a partir de la disponibilidad de información de CEa, NDVI, topografía y mapas de rendimiento.

Un hallazgo muy importante en el Capítulo 3 fue que la metodología propuesta de desagregación de unidades cartográficas de suelo (DgUCS), fue eficiente para la delimitación de zonas por tipo de suelos. Sin embargo, no permitió determinar la correspondencia espacial entre la información proveniente de la cartografía convencional

y la obtenida mediante CDS, para todas las zonas estudiadas. Específicamente, esto se presentó para los lotes del centro de Entre Ríos. En esta zona, se generó un vacío en la validación de la metodología de DgUCSk. Este vacío pueden deberse a dos motivos (i) imprecisión de la información tabulada para la cartografía convencional de suelos; y/o (ii) efecto de la degradación del suelo por sales sobre las mediciones que se realizaron para obtener la información auxiliar de covariables ambientales (CEa, NDVI y mapas de rendimiento) que se utilizó en el método de DgUCS. Al respecto, se sugiere que futuros trabajos evalúen la misma metodología de DgUCS en otros lotes del centro entrerriano. Preferiblemente, en lotes que no presenten problemas de degradación de suelos por sales. Estos futuros trabajos deberán ratificar o no los resultados encontrados en el Capítulo 3. Por su parte, en el sudeste bonaerense y sur de Córdoba, es pertinente ejecutar nuevos trabajos que ratifiquen los resultados encontrados. Por ejemplo, en el sudeste bonaerense el patrón espacial de la tosca a escala de lote suele ser muy complejo y particular. A pesar de que la metodología de DgUCS tuvo en cuenta prioritariamente, el patrón espacial de la tosca, es válido corroborar si efectivamente esta técnica tiene la robustez necesaria para delimitar zonas por tipo de suelos, en condiciones de otros tipos de patrones espaciales de tosca.

Así los resultados obtenidos en el Capítulo 3 ***permiten aceptar parcialmente la hipótesis 2***, la cual enunciaba que “*La cartografía digital de suelos elaborada a partir de la cartografía convencional de suelos disponible, fuentes de información de covariables ambientales y técnicas de desagregación espacial, permiten determinar de una manera rápida, precisa, puntual y a escala de lote, patrones espaciales de propiedades del suelo, en condiciones productivas de tres regiones agrícolas de Argentina*”. Si bien no se pudo realizar la DgUCS, la delimitación de zonas por tipo de suelos, si demostró ser válida para cada uno de los lotes. Además, se demostró que permite, no solo determinar los patrones espaciales de propiedades del suelo, sino también los patrones de afectación de los suelos por factores abióticos (suelos salinos). En la mayoría de los trabajos que proponen metodologías de DgUCS, estos buscan aprovechar la información convencional de suelos, lo más eficientemente posible. Por lo general, dichos trabajos partieron del principio que toda la cartografía convencional disponible fue elaborada bajo estándares metodológicos similares (Kerry et al., 2012; Nauman et al.,

2012; Subburayalu et al., 2014). Sin embargo, en Argentina existe heterogeneidad en las metodologías, escalas y procesos de elaboración de la mayor parte de la cartografía de suelos disponible (Del Valle, 2011). En ese contexto era infructuoso plantear una metodología de DgUCS que permitiera simplemente aumentar la escala de toda la cartografía convencional de suelos disponible. Por el contrario, en esta tesis se consideró que el aumento de la escala de la información de suelos disponible se debe hacer desde la escala de cada lote en particular y no pretendiendo abarcar todos los lotes en cada región. Es posible que para estudios ambientales o hidrológicos, metodologías de DgUCS que involucre espacialmente todos los lotes agrícolas de una región agroproductiva puedan ser convenientes y suficientes. Sin embargo, para lotes agrícolas la demanda de información de suelos es puntual y espacialmente precisa, tal y como se presentaron en los resultados obtenidos con la metodología propuesta.

En Argentina, el efecto de la sojización en la sostenibilidad de sistema agroproductivo a escala de lote ha sido ampliamente estudiado y discutido. Específicamente, el cultivo de la soja genera un impacto ambiental significativo en el suelo, debido a que tiene un bajo aporte de residuos y una alta capacidad extractiva de nutrientes (Echeverría y García, 2005). Puntualmente, en la medida que se desconozca como son las relaciones entre patrones espaciales de propiedades del suelo y propiedades de rendimiento y calidad de soja, es posible que el impacto de la sojización no pueda mitigarse en los planes de manejo de suelos implementados por los técnicos y productores.

En el Capítulo 4 se propuso una metodología de CDS que permitiera explorar y analizar las relaciones entre patrones de suelo y cultivo de soja. Teniendo en cuenta los resultados obtenidos en el Capítulo 3, en relación con la delimitación de zonas por tipo de suelo, se consideró que era pertinente validar si estas zonas delimitadas dentro de cada lote, podrían ser útiles para crear una base técnica y espacial para implementar zonas para manejo sitio específico y/o zonas para cosecha diferenciada por calidad. Para lograr esto, se midieron propiedades de rendimiento y calidad de soja, en los mismos puntos donde se realizó el muestreo de suelos. La idea fue explorar y analizar en cada zona delimitada, como eran las relaciones entre patrones espaciales de propiedades de suelo y cultivo.

A partir de los resultados obtenidos en el Capítulo 4 fue posible explicar y analizar la dinámica de las relaciones entre patrones espaciales de propiedades del suelo y de rendimiento de soja en todos de los lotes experimentales. Sin embargo, lo mismo no ocurrió con la dinámica espacial entre propiedades de suelo y propiedades de calidad de soja. En zonas con propiedades de suelo limitantes de rendimiento, variables de rendimiento tales como granos por metros cuadrado, radiación interceptada y NDVI fueron afectadas de manera homogénea y con correspondencia espacial de acuerdo a las limitaciones del suelo, en la mayoría de los lotes experimentales. Por su parte, el contenido de aceite y proteína no presentó correspondencia espacial con zonas delimitadas en los lotes del centro de Entre Ríos. En estos lotes, MO y pH fueron las variables más importantes para contenido de proteína. Es ampliamente conocido que MO y pH tienen mucha relación con disponibilidad de nitrógeno y que a su vez, este nutriente es determinante en la producción de proteínas. Sin embargo, el efecto de la presencia abundante de sales sobre el contenido de proteína en los lotes entrerrianos, no quedó claro. Por el contrario, en el lote *San Esteban* el efecto de la presencia de sales en el suelo, si fue importante sobre contenido de proteína. Esta incongruencia en las relaciones de los patrones espaciales de suelos salinos y contenido de proteína deberá ser objeto de estudio de futuros trabajos. Estos resultados **permiten aceptar parcialmente la hipótesis 3**, la cual enunciaba que “*La cartografía digital de suelos elaborada a partir de la cartografía convencional de suelos disponible, fuentes de información de covariables ambientales y técnicas de desagregación espacial, permiten determinar de una manera rápida, precisa, puntual y a escala de lote, las relaciones entre patrones espaciales de propiedades del suelo y propiedades del rendimiento y calidad de soja, en condiciones productivas de tres regiones agrícolas de Argentina, generando una base técnica y espacial para la delimitación de zonas para el manejo sitio específico del cultivo de la soja y zonas para cosecha diferenciada por calidad de grano de soja*”. Si bien existían trabajos previos que estudiaron los efectos de la topografía y los patrones espaciales del suelo sobre los patrones de calidad de la soja, los resultados de esta tesis sugieren que queda mucho por explorar (Anthony *et al.*, 2012; Kravchenko y Bullock, 2002b; Martín *et al.*, 2005). La presencia de tosca, los complejos patrones espaciales de suelos a escala de lote, los patrones espaciales de suelos salinos y las prácticas de manejo de suelos, son algunas particularidades importantes que, como se demostró, pueden llegar a tener un

efecto importante en las propiedades del cultivo de soja. Estas particularidades en su conjunto, no se encuentran en las zonas donde se han realizado trabajos previos. Además, no se encontró evidencia de trabajos que utilizaran técnicas de CDS para entender las relaciones espaciales entre patrones espaciales de suelos y calidad del cultivo de soja. Estos antecedentes y los resultados encontrados, sugieren que las técnicas de CDS pueden ayudar a resolver las brechas de conocimientos que existen para entender las relaciones entre patrones espaciales de propiedades del suelo y calidad del cultivo de soja.

Por último, los resultados del Capítulo 4 definieron que las técnicas de CDS tienen todo el potencial para ser utilizadas en procesos de delimitación de zonas para manejo sitio específico. Sin embargo, futuros estudios son necesarios para ajustar similares técnicas de CDS para delimitar zonas para cosecha diferenciada por calidad del cultivo.

En conclusión, esta tesis realizó aportes importantes relacionados con:

- La aplicación de técnicas de CDS al estudio de los patrones espaciales de propiedades del suelo a escala de lote en condiciones de los suelos de regiones agroproductivas importantes de la Argentina.
- Demostró que los esquemas de muestreo de suelos basados en modelos son eficientes para reconocer las distribuciones de todas los predictores de propiedades del suelo y que mejoran el rendimiento predictivos de los algoritmos de regresión para propiedades del suelo.
- Ratificó que el algoritmo Random Forest es eficiente para agrupar gran cantidad de predictores y adecuado para predecir propiedades del suelo a escala de lote.
- Ratificó que el algoritmo Random Forest es eficiente para determinar la importancia de propiedades del suelo sobre las propiedades de rendimiento y calidad del cultivo de soja.

- Generó conocimientos acerca del uso de técnicas de CDS en el estudio de las relaciones espaciales entre propiedades del suelo y cultivo de la soja, en condiciones de tres zonas agrícolas de Argentina.
- Demostró que la combinación de los algoritmos MULTISPATI-ACP y análisis de distancia euclídea permite delimitar eficientemente zonas por tipo de suelos a escala de lote.
- Generó conocimientos para aprovechar la información contenida en la cartografía convencional de suelos, para aumentar la información de suelos a escala de lote.
- Es el estudio más completo con respecto a la aplicación de técnicas de CDS para delimitar zonas para manejo sitio específico y para cosecha por calidad de grano de soja, en condiciones de suelos de la Argentina.
- Encontró diferencias en los efectos de la topografía sobre las propiedades del cultivo de soja, con respecto a trabajos preliminares.
- Demostró que la información auxiliar disponible de manera gratuita y fácil de obtener, tiene mucho potencial para ser utilizada en trabajos cuyo objeto de estudio sea la generación de conocimientos acerca de las relaciones entre patrones espaciales de propiedades del suelo y cultivo.
- Determinó de manera puntual el potencial de la información auxiliar de propiedades del suelo tales como CEa, imágenes satelitales y mapas de rendimientos, para predecir los patrones espaciales de propiedades del suelo y cultivo de soja.

Por su parte, los resultados obtenidos invitan a replantear las técnicas, metodologías y uso de la información disponible, para generar información de suelos a escala de lote. Además, sugieren involucrar una dinámica de trabajos que permitan evaluar la aplicación de tecnologías emergentes en los procesos de generación de conocimientos de patrones espaciales de suelos, y su relación con propiedades del cultivo. Inclusive, deben explorarse otras técnicas de CDS que puedan incorporarse a aplicaciones importantes

tales como modelos de simulación de crecimiento y desarrollo de cultivos, de modo que facilite la modelación de las relaciones entre propiedades del suelo y cultivo de soja. Por último, lo más importante, los resultados obtenidos en esta tesis doctoral generaron una gran cantidad de interrogantes que fueron debidamente acompañados de propuestas para futuros estudios, que darán lugar a futuras líneas de investigación.

## Bibliografía

- ABROL I.P., YADAV J.S.P., MASSOUD F.I. 1988. Salt-affected Soils and Their Management Food and Agriculture Organization of the United Nations FAO, Rome. 375 p.
- ADAM I., MICHOT D., GUERO Y., SOUBEGA B., MOUSSA I., DUTIN G., WALTER C. 2012. Detecting soil salinity changes in irrigated Vertisols by electrical resistivity prospecting during a desalinisation experiment. *Agric. Water Manage.* 109:1-10.
- ADAMCHUK V.I., HUMMEL J.W., MORGAN M.T., UPADHYAYA S.K. 2004. On-the-go soil sensors for precision agriculture. *Comput. Electron. Agric.* 44:71-91.
- AHMAD S., KALRA A., STEPHEN H. 2010. Estimating soil moisture using remote sensing data: A machine learning approach. *Advances in Water Resources.* 33:69-80.
- ALI A.M., THIND H.S., VARINDERPAL S., BIJAY S. 2015. A framework for refining nitrogen management in dry direct-seeded rice using GreenSeeker™ optical sensor. *Comput. Electron. Agric.* 110:114-120.
- ALVAREZ R., STEINBACH H.S. 2009. A review of the effects of tillage systems on some soil physical properties, water content, nitrate availability and crops yield in the Argentine Pampas. *Soil Tillage Res.* 104:1-15.
- ALLRED B., DANIELS J.J., EHSANI M.R. 2008. Handbook of Agricultural Geophysics. CRC press, Florida, USA. 410 p.
- AMIOTTI N., BRAVO O., IBARGUREN G., SCHMIDT E. 2008. Relaciones suelo-paisaje y su incidencia en la productividad del cultivo de girasol en el sudoeste bonaerense. Actas del XXI Congreso Argentino de la Ciencia del Suelo. Potrero de los Funes, 13 al 16 de mayo de 2008. Resumen expandido en CD
- AMIOTTI N., DEL C. BLANCO M., SANCHEZ L.F. 2001. Complex pedogenesis related to differential aeolian sedimentation in microenvironments of the southern part of the semiarid region of Argentina. *Catena.* 43:137-156.
- ANDERSON-COOK C.M., ALLEY M.M., ROYGARD J.K.F., KHOSLA R., NOBLE R.B., DOOLITTLE J.A. 2002. Differentiating Soil Types Using Electromagnetic Conductivity And Crop Yield Maps. *Soil Sci. Soc. Am. J.* 66:1562-1570.
- ANDRADE F., SADRAS V. 2002. Bases para el manejo del maíz, el girasol y la soja, Balcarce - Argentina. 450 p.
- ANDRADE F.H. 1995. Analysis of growth and yield of maize, sunflower and soybean grown at Balcarce, Argentina. *Field Crops Res.* 41:1-12.

- ANDRADE F.H., FERREIRO M.A. 1996. Reproductive growth of maize, sunflower and soybean at different source levels during grain filling. *Field Crops Research*. 48:155-165.
- ANDRIANI J., BODRERO M. 1995. Respuesta de cultivares de soja a la disponibilidad hídrica, 1er Congreso Nacional de Soja y Segunda Reunión Nacional de Oleaginosos. AIANBA, Pergamino. Buenos Aires. pp. 81 - 87.
- ANDRIANI J.M. 2002. Estrés hídrico en soja. *IDIA XXI*. 3:48-51.
- ANGELINI M. 2010. La cartografía de suelos en la Argentina, presente y futuro de la disciplina, Taller Regional GlobalSoilMap.net, nodo Latinoamérica, Rio de Janeiro. pp. 68-93.
- ANGELINI M., HEUVELINK G.B., KEMPEN B., MORRAS H., RODRIGUEZ D. 2015. Digital soil mapping of an argentinian pampa region using structural equation modelling. *Wageningen soil conference 2015 Wageningen, Netherlands*. pp. 256-260.
- ANTHONY P., MALZER G., SPARROW S., ZHANG M. 2012. Soybean yield and quality in relation to soil properties. *Agr. J.* 104:1443-1458.
- ARCHER K.J., KIMES R.V. 2008. Empirical characterization of random forest variable importance measures. *Comput. Stat. Data. Anal.* 52:2249-2260.
- ARROUAYS D., SABY N.P.A., THIOULOUSE J., JOLIVET C., BOULONNE L., RATIÉ C. 2011. Large trends in French topsoil characteristics are revealed by spatially constrained multivariate analysis. *Geoderma*. 161:107-114.
- BAIGORRI H. 2004. Criterios generales para la elección y el manejo de cultivares en el cono sur. *Manual práctico para la producción de soja*. M. Díaz Zorita y G. Duarte, Eds. Buenos Aires. pp. 39-77.
- BALZARINI M., GILI A., CÓRDOBA M. 2011. Análisis in situ: Análisis de Variabilidad Espacial Multivariada en Agricultura de Precisión. 16º Reunión Científica Grupo Argentino de Biometría Salta Argentina. 31 p.
- BASSO B., BERTOCCO M., SARTORI L., MARTIN E.C. 2007. Analyzing the effects of climate variability on spatial pattern of yield in a maize-wheat-soybean rotation. *Eur. J. Agron.* 26:82-91.
- BEDOJNI S. 1996. Caracterización de las acumulaciones calcareas en la zona periserrana de Tandilia. Tesis para optar el título de Ingeniero Agrónomo. UNMP - Facultad de Ciencias Agriarias. Balcarce, Argentina. 174 p.
- BEHRENS T., ZHU A., SCHMIDT K., SCHOLTEN T. 2010. Multi-scale digital terrain analysis and feature selection for digital soil mapping. *Geoderma*. 155:175-185.

- BELLALLOUI N., HANKS J.E., FISHER D.K., MENGISTU A. 2009. Soybean seed composition is influenced by within-field variability in soil nutrients. *Crop management*. 8:10-25.
- BELLALLOUI N., REDDY K.N., MENGISTU A. 2015. Chapter 45 - Drought and Heat Stress Effects on Soybean Fatty Acid Composition and Oil Stability. En: *Processing and Impact on Active Components in Food*, Academic Press, San Diego, USA. pp. 377-384.
- BIVAND R., BERNAT A., CARVALHO M., CHUN Y., DORMANN C., DRAY S., HALBERSMA R., LEWIN-KOH N., MA J., MILLO G. 2005. The spdep package. *Comprehensive R Archive Network*. 1:3-13.
- BLACKMORE S. 2000. The interpretation of trends from multiple yield maps. *Comput. Electron. Agric.* 26:37-51.
- BLACKMORE S., GODWIN R.J., FOUNTAS S. 2003. The Analysis of Spatial and Temporal Trends in Yield Map Data over Six Years. *Biosyst. Eng.* 84:455-466.
- BLANCO M.D.C., STOOPS G. 2007. Genesis of pedons with discontinuous argillic horizons in the Holocene loess mantle of the southern Pampean landscape, Argentina. *J. S. Am. Earth Sci.* 23:30-45.
- BOETTINGER J., RAMSEY R., BODILY J., COLE N., KIENAST-BROWN S., NIELD S., SAUNDERS A., STUM A. 2008. Landsat spectral data for digital soil mapping. En: *Digital Soil Mapping with limited data*, Springer. pp. 193-202.
- BOETTINGER J.L. 2010. *Digital Soil Mapping: Bridging Research, Production, and Environmental Application* Springer Science & Business Media. 462 p.
- BOETTINGER J.L., DOOLITTLE J.A., WEST N.E., BORK E.W., SCHUPP E.W. 1997. Nondestructive assessment of rangeland soil depth to petrocalcic horizon using electromagnetic induction. *Arid. Soil. Res. Rehab.* 11:375-390.
- BOETTINGER J.L., HOWELL D., MOORE A., HARTEMINK A.E., KIENAST-BROWN S. 2010. *Digital Soil Mapping: Bridging Research, Environmental Application, and Operation* Springer Science & Business Media, Dordrecht, Heidelberg, London, New York. 268 p.
- BOGAERT P., D'OR D. 2002. Estimating Soil Properties from Thematic Soil Maps. *Soil Sci. Soc. Am. J.* 66:1492-1500.
- BOLSA DE COMERCIO DE ROSARIO. 2015. Estimaciones de Producción. [En línea] <[www.bcr.com.ar/Pages/gea/estimaProd.aspx](http://www.bcr.com.ar/Pages/gea/estimaProd.aspx)>. [Consulta:30/05/2015]
- BOMMARCO R., KLEIJN D., POTTS S.G. 2013. Ecological intensification: harnessing ecosystem services for food security. *Trends Ecol Evol.* 28:230-238.

- BONGIOVANNI R., LOWENBERG-DEBOER J. 2005. Precision Agriculture in Argentina. 3er Simpósio Internacional de Agricultura de Precisão. Sete Lagoas, MG. 56 p.
- BORK E.W., WEST N.E., DOOLITTLE J.A., BOETTINGER J.L. 1998. Soil depth assessment of sagebrush grazing treatments using electromagnetic induction. *J. Range. Manage.*:469-474.
- BORRÁS L., SLAFER G.A., OTEGUI M.A.E. 2004. Seed dry weight response to source-sink manipulations in wheat, maize and soybean: a quantitative reappraisal. *Field Crops Res.* 86:131-146.
- BORRELLI N., OSTERRIETH M., MARCOVECCHIO J. 2008. Interrelations of vegetal cover, silicophytolith content and pedogenesis of Typical Argiudolls of the Pampean Plain, Argentina. *Catena* 75:146-153.
- BREIMAN L. 2001. Random Forests. *Machine Learning.* 45:5-32.
- BRUNGARD C., BOETTINGER J. 2010. Conditioned latin hypercube sampling: Optimal sample size for digital soil mapping of arid rangelands in Utah, USA. En: *Digital soil mapping*, Springer. pp 67-75.
- BRUNGARD C.W., BOETTINGER J.L., DUNIWAY M.C., WILLS S.A., EDWARDS JR T.C. 2015. Machine learning for predicting soil classes in three semi-arid landscapes. *Geoderma.* 239-240:68-83.
- BUI E., HENDERSON B., VIERGEVER K. 2009. Using knowledge discovery with data mining from the Australian Soil Resource Information System database to inform soil carbon mapping in Australia. *Global Biogeochemical Cycles.* 23.
- BUI E.N., MORAN C.J. 2001. Disaggregation of polygons of surficial geology and soil maps using spatial modelling and legacy data. *Geoderma.* 103:79-94.
- BURGOS A.C. 2009. Selección de variables en problemas multiclase. Tesis para optar el título de Tesis de Grado para optar el título de Universidad Nacional de Rosario - Facultad de Ciencias Exactas, Ingeniería y Agrimensura. Rosario - Argentina. 58 p.
- BURTON J. 1985. Breeding soybeans for improved protein quantity and quality. *World Soybean Research Conference, 3rd*, Ames, IA. Boulder, CO. pp 61-367.
- BUSCHIAZZO D. 1986. Estudio sobre la Tosca. Parte I: Evidencias de un movimiento descendente del carbonato en base a la interpretación de características macro y geomorfológicas. *Ciencia del suelo.* 4:55-66.
- CABRIA F., CULOT J.P. 1994. Selección y utilización de características edáficas para discriminar series de Argiudoles en el sudeste bonaerense. *Ciencia del suelo.* 12:41-45.

- CAI L., XU Z., BAO P., HE M., DOU L., CHEN L., ZHOU Y., ZHU Y.-G. 2015. Multivariate and geostatistical analyses of the spatial distribution and source of arsenic and heavy metals in the agricultural soils in Shunde, Southeast China. *J Geochem Explor.* 148:189-195.
- CAMBARDELLA C.A., MOORMAN T.B., PARKIN T.B., KARLEN D.L., NOVAK J.M., TURCO R.F., KONOPKA A.E. 1994. Field-Scale Variability of Soil Properties in Central Iowa Soils. *Soil Sci. Soc. Am. J.* 58:1501-1511.
- CARRÉ F., MCBRATNEY A.B., MINASNY B. 2007. Estimation and potential improvement of the quality of legacy soil samples for digital soil mapping. *Geoderma.* 141:1-14.
- CARROLL Z.L., OLIVER M.A. 2005. Exploring the spatial relations between soil physical properties and apparent electrical conductivity. *Geoderma.* 128:354-374.
- CASANELLAS J.P. 2003. *Edafología para la Agricultura y el Medio Ambiente Mundi-* Prensa, Madrid, España. 929 p.
- CASTRIGNANO A., MAIORANA M., FORNARO F., LOPEZ N. 2002. 3D spatial variability of soil strength and its change over time in a durum wheat field in Southern Italy. *Soil Tillage Res.* 65:95-108.
- CASTRO FRANCO M., COSTA J.L., PERALTA N., APARICIO V. 2015. Prediction of Soil Properties at Farm Scale Using a Model-Based Soil Sampling Scheme and Random Forest. *Soil Sci.* 180:1-12.
- CLARKE E., WISEMAN J. 2000. Developments in plant breeding for improved nutritional quality of soya beans I. Protein and amino acid content. *The Journal of Agricultural Science.* 134:111-124.
- COBER E., D VOLDENG H. 2000. Developing high-protein, high-yield soybean populations and lines. *Crop Sci.* 40:39-42.
- CONKLIN JR A.R. 2004. *Field sampling: principles and practices in environmental analysis* CRC Press. 347 p.
- COOK S., JARVIS A., GONZALEZ J. 2008. A new global demand for digital soil information. En: *Digital Soil Mapping with Limited Data*, Springer. pp 31-41.
- CÓRDOBA M., BRUNO C., COSTA J., BALZARINI M. 2013. Subfield management class delineation using cluster analysis from spatial principal components of soil variables. *Comput. Electron. Agric.* 97:6-14.
- CORWIN D.L., LESCH S.M. 2005. Characterizing soil spatial variability with apparent soil electrical conductivity: I. Survey protocols. *Comput. Electron. Agric.* 46:103-133.

- CORWIN D.L., LESCH S.M., OSTER J.D., KAFFKA S.R. 2006. Monitoring management-induced spatio-temporal changes in soil quality through soil sampling directed by apparent electrical conductivity. *Geoderma*. 131:369-387.
- CORWIN D.L., LESCH S.M., SHOUSE P.J., SOPPE R., AYARS J.E. 2003. Identifying Soil Properties that Influence Cotton Yield Using Soil Sampling Directed by Apparent Soil Electrical Conductivity. *Agron. J.* 95:352-364.
- CORWIN D.L., PLANT R.E. 2005. Applications of apparent soil electrical conductivity in precision agriculture. *Comput. Electron. Agric.* 46:1-10.
- CRESSIE N., KANG E.L. 2010. High-resolution digital soil mapping: Kriging for very large datasets. En: *Proximal Soil Sensing*, Springer, Netherlands. pp 49-63.
- CUNIBERTI M., HERRERO R., MIR L., BERRA O., MACAGNO S. 2013. Rendimiento y calidad comercial e industrial de la soja en la Región Núcleo-Sojera. Cosecha 2012-13. Estación Experimental Agropecuaria. INTA Marcos Juárez.
- CUTLER D.R., EDWARDS JR T.C., BEARD K.H., CUTLER A., HESS K.T., GIBSON J., LAWLER J.J. 2007. Random forests for classification in ecology. *Ecology*. 88:2783-2792.
- CHAPMAN H. 1965. Cation-exchange capacity. En: *Methods of soil analysis. Part 2. Chemical and microbiological properties*, Soil Science Society of America, New York, EEUU. pp 891-901.
- CHESEL D., DUFOUR A.B., THIOULOUSE J. 2004. The ade4 package-I-One-table methods. *R news*. 4:5-10.
- DANG Y., PRINGLE M., SCHMIDT M., DALAL R., APAN A. 2011. Identifying the spatial variability of soil constraints using multi-year remote sensing. *Field Crop Res.* 123:248-258.
- DE BENEDETTO D., CASTRIGNANO A., SOLLITTO D., MODUGNO F., BUTTAFUOCO G., PAPA G.L. 2012. Integrating geophysical and geostatistical techniques to map the spatial variation of clay. *Geoderma*. 171-172:53-63.
- DE BRUIN S., WIELEMAKER W., MOLENAAR M. 1999. Formalisation of soil-landscape knowledge through interactive hierarchical disaggregation. *Geoderma*. 91:151-172.
- DE GRUIJTER J., MCBRATNEY A., TAYLOR J. 2010. Sampling for high-resolution soil mapping. En: *Proximal Soil Sensing*, Springer. pp 3-14.
- DE SOUZA P., EGLI D.B., BRUENING W.P. 1997. Water stress during seed filling and leaf senescence in soybean. *Agr. J.* 89:807-812.

- DEBAENE G., NIEDŹWIECKI J., PECIO A., ŹUREK A. 2014. Effect of the number of calibration samples on the prediction of several soil properties at the farm-scale. *Geoderma*. 214:114-125.
- DEL VALLE GALLO C. 2008. Calidad Fisiológica y Efecto de la Presencia de Semillas Verdes de Soja en Lotes destinados a simiente. Tesis para optar el título de Magister en Ciencias Agropecuarias. Universidad Nacional de Córdoba - Facultad de Ciencias Agrarias. Córdoba - Argentina. 134 p.
- DEL VALLE H. 2011. La Cartografía Digital de Suelos en la Argentina, entre la ciencia y el empirismo, 1er Seminario Nacional de Cartografía Digital, San Luis - Argentina. pp 1-18.
- DESCLAUX D., ROUMET P. 1996. Impact of drought stress on the phenology of two soybean (*Glycine max* L. Merr) cultivars. *Field Crops Resea*. 46:61-70.
- DÍAZ-URIARTE R., DE ANDRES S.A. 2006. Gene selection and classification of microarray data using random forest. *BMC bioinformatics*. 7:3.
- DIGGLE P.J., RIBEIRO P.J. 2007. Model based Geostatistics. Springer Series in Statistics, New York. 232 p.
- DOBOS E., HENGL T. 2009. Digital Soil Mapping as a support to production of functional maps Institute for environmental and Sustainability, Luxemburg. 68 p.
- DOMENECH M. 2013. Caracterización de la variabilidad espacial de propiedades de suelo a partir de la conductividad eléctrica aparente. Tesis para optar el título de Tesis para optar el título de Universidad Nacional del Sur - Facultad de Agronomía. Bahía Blanca - Argentina. 75 p.
- DOMENECH M., COSTA J., CASTRO FRANCO M. 2012. Uso de la Conductividad Eléctrica Aparente para diferenciar series de suelo. Actas del 19º Congreso Latinoamericano y 23º Congreso Argentino de la Ciencia del Suelo. Mar del Plata, 16 al 20 de abril de 2012. Resumen expandido en CD.
- DOOLITTLE J.A., JENKINSON B., HOPKINS D., ULMER M., TUTTLE W. 2006. Hydropedological investigations with ground-penetrating radar (GPR): Estimating water-table depths and local ground-water flow pattern in areas of coarse-textured soils. *Geoderma*. 131:317-329.
- DOOLITTLE J.A., SUDDUTH K.A., KITCHEN N.R., INDORANTE S.J. 1994. Estimating depths to claypans using electromagnetic induction methods. *J. Soil Water Conserv*. 49:572-575.
- DORNBOS D., MULLEN R., SHIBLES R. 1989. Drought stress effects during seed fill on soybean seed germination and vigor. *Crop Sci*. 29:476-480.
- DRAY S., CHESSEL D., THIOULOUSE J. 2003. Co-inertia analysis and the linking of ecological data table. *Ecology*. 84:3078-3089.

- DRAY S., JOMBART T. 2011. Revisiting guerry's data: introducing spatial constraints in multivariate analysis. *Ann Appl Stat.* 5:2278-2299.
- DRAY S., SAÏD S., DÉBIAS F. 2008. Spatial ordination of vegetation data using a generalization of Wartenberg's multivariate spatial correlation. *J Veg Sci.* 19:45-56.
- ECHEVERRIA H., GARCIA F. 2005. Fertilidad de suelos y fertilización de cultivos. INTA - INPOFOS, Balcarce - Argentina. 526 p.
- FAO. 2002. World agriculture : towards 2015/2030 : summary report Food and Agriculture Organization of the United Nations, Rome, Italy. 444 p.
- FAO. 2011. Looking ahead in world food and agriculture : perspectives to 2050 / edited by Piero Conforti FAO, Rome, Italy. 56 p.
- FEHR W., CAVINESS C., BURMOOD D., PENNINGTON J. 1971. Stage of development descriptions for soybeans, *Glycine max (L.) Merrill.* *Crop Sci.* 11:929-931.
- FLORINSKY I.V. 2012. Digital Terrain Analysis in Soil Science and Geology. Academic Press. 379 p.
- FRIEDMAN S.P. 2005. Soil properties influencing apparent electrical conductivity: a review. *Comput. Electron. Agric.* 46:45-70.
- FUKUYAMA Y., SUGENO M. 1989. A new method of choosing the number of clusters for the fuzzy c-means method, *Proc. 5th Fuzzy Syst. Symp.* pp. 247-250.
- GALLO K., DAUGHTRY C. 1986. Techniques for measuring intercepted and absorbed photosynthetically active radiation in corn canopies. *Agr. J.* 78:752-756.
- GEE G., BAUDER J. 1986. Particle-size analysis. A. Klute (Ed.), En: *Methods of soil analysis. Part 1. Physical and mineralogical methods*, Madison, WI. pp. 383 - 411.
- GENUER R., POGGI J.-M., TULEAU-MALOT C. 2010. Variable selection using random forests. *Pattern Recognit Lett.* 31:2225-2236.
- GIMÉNEZ F., LÚQUEZ J., SUÁREZ J. 2015. Estabilidad y adaptabilidad de cultivares de soja para rendimiento en el sudeste de la provincia de Buenos Aires. *Revista de la Facultad de Agronomía, La Plata.* 104:93-103.
- GIRALDO HENAO R. 2006. *Introducción a la Geoestadística.* Universidad Nacional de Colombia., Bogotá Colombia. 94 p.
- GODFRAY H.C.J., BEDDINGTON J.R., CRUTE I.R., HADDAD L., LAWRENCE D., MUIR J.F., PRETTY J., ROBINSON S., THOMAS S.M., TOULMIN C.

2010. Food Security: The Challenge of Feeding 9 Billion People. *Science*. 327:812-818.
- GOKTEPE A.B., ALTUN S., SEZER A. 2005. Soil clustering by fuzzy c-means algorithm. *Adv Eng Softw.* 36:691-698.
- GRIMM R., BEHRENS T., MÄRKER M., ELSENBEER H. 2008. Soil organic carbon concentrations and stocks on Barro Colorado Island — Digital soil mapping using Random Forests analysis. *Geoderma*. 146:102-113.
- GRÖMPING U. 2009. Variable importance assessment in regression: linear regression versus random forest. *Am Stat.* 63.
- GROVE J., NAVARRO M. 2013a. The problem is not N deficiency: Active canopy sensors and chlorophyll meters detect P stress in corn and soybean. En: *Precision agriculture'13*, Springer. pp 137-144.
- GROVE J.H., NAVARRO M.M. 2013b. The problem is not N deficiency: Active canopy sensors and chlorophyll meters detect P stress in corn and soybean. En: *Precision agriculture '13*, Wageningen Academic Publishers. pp 137-144.
- GRUNWALD S. 2010. Current State of Digital Soil Mapping and What Is Next. En: *Digital Soil Mapping*, Springer Netherlands. pp 3-12.
- GUEDES FILHO O., VIEIRA S.R., CHIBA M.K., GREGO C.R. 2010. Geostatistical analysis of crop yield maps in a long term no tillage system. *Bragantia*. 69:9-18.
- GUO W., MAAS S., BRONSON K. 2012. Relationship between cotton yield and soil electrical conductivity, topography, and Landsat imagery. *Precis Agric.* 13:678-692.
- HARTEMINK A., HEMPEL J., LAGACHERIE P., MCBRATNEY A., MCKENZIE N., MACMILLAN R., MINASNY B., MONTANARELLA L., MENDONÇA SANTOS M., SANCHEZ P., WALSH M., ZHANG G.-L. 2010. *GlobalSoilMap.net – A New Digital Soil Map of the World*. En: *Digital Soil Mapping*, Springer Netherlands. pp 423-428.
- HARTEMINK A.E., MCBRATNEY A.B., DE LOURDES MENDONÇA-SANTOS M. 2008. *Digital Soil Mapping with Limited Data* Springer Science+Business Media B.V. 446 p.
- HEIL K., SCHMIDHALTER U. 2012. Characterisation of soil texture variability using the apparent soil electrical conductivity at a highly variable site. *Comput. Geosci.* 39:98-110.
- HEINIGER R.W., MCBRIDE R.G., CLAY D.E. 2003. Using Soil Electrical Conductivity to Improve Nutrient Management. *Agron. J.* 95:508-519.
- HELMS T., ORF J. 1998. Protein, oil, and yield of soybean lines selected for increased protein. *Crop Sci.* 38:707-711.

- HEMPEL J., HAMMER R., MOORE A., BELL J., THOMPSON J., GOLDEN M. 2008. Challenges to digital soil mapping. En: Digital Soil Mapping with Limited Data, Springer. 81-90 p.
- HENGL T., HEUVELINK G., STEIN A. 2004a. A generic framework for spatial prediction of soil variables based on regression-kriging. *Geoderma*. 120:75-93.
- HENGL T., REUTER H.I. 2008. *Geomorphometry: Concepts, Software, Applications*. Elsevier Science, Amsterdam - Países Bajos. 765 pp.
- HENGL T., ROSSITER D.G., STEIN A. 2004b. Soil sampling strategies for spatial prediction by correlation with auxiliary maps. *Soil Res*. 41:1403-1422.
- HEUNG B., BULMER C.E., SCHMIDT M.G. 2014. Predictive soil parent material mapping at a regional-scale: A Random Forest approach. *Geoderma*. 214–215:141-154.
- HU M., WIATRAC P. 2012. Effect of planting date on soybean growth, yield, and grain quality: Review. *Agro. J*. 104:785-790.
- INTA. 2010a. Carta de Suelos de la Provincia de Buenos Aires Instituto Nacional de Tecnología Agropecuaria, Buenos Aires - Argentina. 525 p.
- INTA. 2010b. Carta de Suelos de la Provincia de Córdoba. Instituto Nacional de Tecnología Agropecuaria, Buenos Aires - Argentina. 321 p.
- INTA. 2013a. Sistema de geoinformación para la consulta y procesamiento de datos gerreferenciados de Argentina GEOINTA. [En línea]. <[www.geointa.inta.gob.ar](http://www.geointa.inta.gob.ar)>. [Consulta:11/2014].
- INTA. 2013b. Sistema de información y gestión agrometeorológico SIGA. [En línea] <[www.siga2.inta.gob.ar](http://www.siga2.inta.gob.ar)>. [Consulta: 11/ 2013]
- INTA, Gob\_Entre\_Ríos. 2012. Plan Mapa de Suelos de Entre Ríos Instituto Nacional de Tecnología Agropecuaria, Paraná - Argentina. 525 p.
- ITT Visual Information Solutions. 2009. ENVI/IDL v4.7, Boulder, CO, USA. 184 p.
- JENNY H. 1941. *Factors of soil formation, A System of Quantitative Pedology*. McGraw-Hill Book Company, New York, London. 191 p.
- KEMPEN B., BRUS D.J., HEUVELINK G.B., STOORVOGEL J.J. 2009. Updating the 1: 50,000 Dutch soil map using legacy soil data: a multinomial logistic regression approach. *Geoderma*. 151:311-326.
- KERRY R., GOOVAERTS P., RAWLINS B.G., MARCHANT B.P. 2012. Disaggregation of legacy soil data using area to point kriging for mapping soil organic carbon at the regional scale. *Geoderma*. 170:347-358.

- KERRY R., OLIVER M.A. 2003. Variograms of ancillary data to aid sampling for soil surveys. *Precis Agric.* 4:261-278.
- KHEIR R.B., GREVE M.H., BØCHER P.K., GREVE M.B., LARSEN R., MCCLOY K. 2010. Predictive mapping of soil organic carbon in wet cultivated lands using classification-tree based models: The case study of Denmark. *J. Environ. Manage.* 91:1150-1160.
- KIDD D., MALONE B., MCBRATNEY A., MINASNY B., WEBB M. 2015. Operational sampling challenges to digital soil mapping in Tasmania, Australia. *Geoderma Regional.* 4:1-10.
- KITCHEN N.R., DRUMMOND S.T., LUND E.D., SUDDUTH K.A., BUCHLEITER G.W. 2003. Soil Electrical Conductivity And Topography Related To Yield For Three Contrasting Soil-crop Systems. *Agron. J.* 95:483-495.
- KRAVCHENKO A., BULLOCK D. 2002a. Spatial variability of soybean quality data as a function of field topography 42:804-815.
- KRAVCHENKO A., BULLOCK D.G. 1999. A Comparative Study of Interpolation Methods for Mapping Soil Properties. *Agron. J.* 91:393-400.
- KRAVCHENKO A.N., BULLOCK D.G. 2000. Correlation of Corn and Soybean Grain Yield with Topography and Soil Properties. *Agron. J.* 92:75-83.
- KRAVCHENKO A.N., BULLOCK D.G. 2002b. Spatial Variability of Soybean Quality Data as a Function of Field Topography. *Crop Sci.* 42:804-815.
- KRAVCHENKO A.N., ROBERTSON G.P., THELEN K.D., HARWOOD R.R. 2005. Management, Topographical, and Weather Effects on Spatial Variability of Crop Grain Yields. *Agron. J.* 97:514-523.
- KUANG B., MOUAZEN A.M. 2012. Influence of the number of samples on prediction error of visible and near infrared spectroscopy of selected soil properties at the farm scale. *Eur. J. Agron.* 63:421-429.
- LAGACHERIE P. 2008. Digital Soil Mapping: A State of the Art. En: *Digital Soil Mapping with Limited Data*, Springer Netherlands. 3-14 p.
- LAGACHERIE P., LEGROS J.P., BURFOUGH P.A. 1995. A soil survey procedure using the knowledge of soil pattern established on a previously mapped reference area. *Geoderma.* 65:283-301.
- LAGACHERIE P., MCBRATNEY A., VOLTZ M. 2006. *Digital Soil Mapping: An Introductory Perspective*. Elsevier Science. 658 p.
- LAL R. 2010. Beyond Copenhagen: mitigating climate change and achieving food security through soil carbon sequestration. *Food Sec.* 2:169-177.

- LAL R. 2013. Food security in a changing climate. *Ecohydrology & Hydrobiology*. 13:8-21.
- LAL R., STEWART B.A. 2009. *Food Security and Soil Quality* Taylor and Francis. 430 p.
- LARK R.M., STAFFORD J.V. 1997. Classification as a first step in the interpretation of temporal and spatial variation of crop yield. *Annals Appl Biol*. 130:111-121.
- LAWFORD R., STRAUCH A., TOLL D., FEKETE B., CRIPE D. 2013. Earth observations for global water security. *Curr. Opin. Environ. Sustain*. 5:633-643.
- LEONARDI M. 2012. Comportamiento sitio específico de la soja en el centro sur bonaerense. Tesis para optar el título de Ingeniero Agrónomo. Universidad Nacional del Sur - Facultad de Agronomía. Bahía Blanca - Provincia de Buenos Aires. 46 p.
- LESCH S.M. 2005. Sensor-directed response surface sampling designs for characterizing spatial variation in soil properties. *Comput. Electron. Agric*. 46:153-179.
- LEVI M.R., RASMUSSEN C. 2014. Covariate selection with iterative principal component analysis for predicting physical soil properties. *Geoderma*. 219:46-57.
- LI H.Y., SHI Z., WEBSTER R., TRIANTAFILIS J. 2013. Mapping the three-dimensional variation of soil salinity in a rice-paddy soil. *Geoderma*. 195-196:31-41.
- LIEß M., GLASER B., HUWE B. 2012. Uncertainty in the spatial prediction of soil texture: Comparison of regression tree and Random Forest models. *Geoderma*. 170:70-79.
- MA B.L., DWYER L.M., COSTA C., COBER E.R., MORRISON M.J. 2001. Early Prediction of Soybean Yield from Canopy Reflectance Measurements. *Agr. J*. 93:1227-1234.
- MACMILLAN R.A., MOON D.E., COUPÉ R.A. 2007. Automated predictive ecological mapping in a forest region of BC, Canada, 2001-2005. *Geoderma*. 140:353-373.
- MARTIN N.F., BOLLERO A.G., BULLOCK D.G. 2007. Relationship between secondary variables and soybean oil and protein concentration American Society of Agricultural Engineers, St. Joseph, MI, ETATS-UNIS. 8 p.
- MARTÍN N.F., BOLLERO G., BULLOCK D.G. 2005. Associations between field characteristics and soybean plant performance using canonical correlation analysis. *Plant Soil*. 273:39-55.

- MCBRATNEY A., FIELD D.J., KOCH A. 2014. The dimensions of soil security. *Geoderma*. 213:203-213.
- MCBRATNEY A., WHELAN B., ANCEV T., BOUMA J. 2005. Future Directions of Precision Agriculture. *Precis Agric*. 6:7-23.
- MCBRATNEY A.B. 1998. Some considerations on methods for spatially aggregating and disaggregating soil information. En: *Soil and water quality at different scales*, Springer, Netherlands. pp 51-62.
- MCBRATNEY A.B., MENDONÇA SANTOS M.L., MINASNY B. 2003. On digital soil mapping. *Geoderma*. 117:3-52.
- METTERNICHT G.I., ZINCK J.A. 2003. Remote sensing of soil salinity: potentials and constraints. *Remote Sens Environ*. 85:1-20.
- MIKLOS M., SHORT M.G., MCBRATNEY A.B., MINASNY B. 2010. Mapping and comparing the distribution of soil carbon under cropping and grazing management practices in Narrabri, north-west New South Wales. *Soil Res*. 48:248-257.
- MINASNY B., MALONE B.P., MCBRATNEY A.B. 2012. *Digital Soil Assessments and Beyond: Proceedings of the 5th Global Workshop on Digital Soil Mapping 2012*, Sydney, Australia CRC Press. 482 p.
- MINASNY B., MCBRATNEY A.B. 2006. A conditioned Latin hypercube method for sampling in the presence of ancillary information. *Comput. Geosci*. 32:1378-1388.
- MINASNY B., MCBRATNEY A.B. 2007. Incorporating taxonomic distance into spatial prediction and digital mapping of soil classes. *Geoderma*. 142:285-293.
- MINASNY B., MCBRATNEY A.B. 2010. *Methodologies for Global Soil Mapping*. En: *Digital Soil Mapping*, Springer Netherlands. pp 429-436.
- MISRA A.K. 2014. Climate change and challenges of water and food security. *International Journal of Sustainable Built Environment*. 3:153-165.
- MONDO V.H.V., GOMES JUNIOR F.G., PINTO T.L.F., MARCHI J.L.D., MOTOMIYA A.V.D.A., MOLIN J.P., CICERO S.M. 2012. Spatial variability of soil fertility and its relationship with seed physiological potential in a soybean production area. *Revista Brasileira de Sementes*. 34:193-201.
- MOORE I.D., GRAYSON R., LADSON A. 1991. Digital terrain modelling: a review of hydrological, geomorphological, and biological applications. *Hydrol. Process*. 5:3-30.
- MORAL F.J., TERRÓN J.M., SILVA J.R.M.D. 2010. Delineation of management zones using mobile measurements of soil apparent electrical conductivity and multivariate geostatistical techniques. *Soil Tillage Res*. 106:335-343.

- MOSCATELLI G., PAZOS M.S. 2000. Soils of Argentina: nature and use. 22 p.
- MUCHERINO A., PAPAJORGJI P.J., PARDALOS P.M. 2009. Data Mining in Agriculture Springer. 274 p.
- MULDER V., DE BRUIN S., SCHAEPMAN M.E. 2013. Representing major soil variability at regional scale by constrained Latin Hypercube Sampling of remote sensing data. *Int. J. Appl. Earth. Obs.* 21:301-310.
- MULDER V.L., DE BRUIN S., SCHAEPMAN M.E., MAYR T.R. 2011. The use of remote sensing in soil and terrain mapping - A review. *Geoderma*. 162:1-19.
- MULLA D.J. 2013. Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosyst Engineeg.* 114:358-371.
- NAUMAN T., THOMPSON J., ODGERS N., LIBOHOVA Z. 2012. Fuzzy disaggregation of conventional soil maps using database knowledge extraction to produce soil property maps. En: *Digital Soil Assessments and Beyond: Proceedings of the 5th Global Workshop on Digital Soil Mapping*, CRC Press, Sydney, Australia. pp 203-207.
- NAUMAN T.W., THOMPSON J.A. 2014. Semi-automated disaggregation of conventional soil maps using knowledge driven data mining and classification trees. *Geoderma*. 213:385-399.
- NORMAN A.G. 1965. *Methods of Soil Analysis. Part 2. Chemical and Microbiological Properties* American Society of Agronomy, Soil Science Society of America, Madison, WI. 456 p.
- NOVELLI L.E., CAVIGLIA O.P., MELCHIORI R.J.M. 2011. Impact of soybean cropping frequency on soil carbon storage in Mollisols and Vertisols. *Geoderma*. 167-168:254-260.
- ODEH I.O.A., CHITTLEBOROUGH D.J., MCBRATNEY A.B. 1992. Soil Pattern Recognition with Fuzzy-c-means: Application to Classification and Soil-Landform Interrelationships. *Soil Sci. Soc. Am. J.* 56:505-516.
- ODGERS N.P., SUN W., MCBRATNEY A.B., MINASNY B., CLIFFORD D. 2013. Disaggregating and harmonising soil map units through resampled classification trees. *Geoderma*. 214:91-100.
- OLAYA V., CONRAD O. 2009. *Geomorphometry in SAGA. Developments in Soil Science.* 33:293-308.
- OLIVER M.A. 2010. *Geostatistical Applications for Precision Agriculture.* Springer, United Kingdom. 337 pp.
- OYA T., NEPOMUCENO A.L., NEUMAIER N., FARIAS J., EACUTE, RENATO B., CCEDIL, AS, TOBITA S., ITO O. 2004. Drought Tolerance Characteristics of Brazilian Soybean Cultivars. *Plant Prod Sci.* 7:129-137.

- PAGGI M.V., PERALTA N.R., CALANDRONI M.B., COSTA J.L. 2012. Relación entre la conductividad eléctrica aparente y la profundidad efectiva en la diferenciación de suelos. Actas del 19º Congreso Latinoamericano y 22º Congreso Argentino de la Ciencia del Suelo. Mar del Plata, 16 al 20 de abril de 2012. Resumen expandido en CD.
- PAL N.R., BEZDEK J.C. 1995. On cluster validity for the fuzzy c-means model. *Fuzzy Systems, IEEE Transactions on.* 3:370-379.
- PASCALE A.J., DAMARIO E.A. 2004. *Bioclimatología agrícola y agroclimatología* Editorial Facultad de Agronomía. Universidad de Buenos Aires, Buenos Aires, Argentina. 550 p.
- PATRICHE C., PÎRNĂU R., ROȘCA B. 2013. Statistical Spatial Models for Soil Parameters: A Comparative Analysis of Several Methods Applied to Dobrovăț Basin (Eastern Romania). *Commun. Soil Sci. Plant Anal.* 44:104-112.
- PAZOS M. 1984. Relación arcilla Iluvial/arcilla total en Molisoles del sudeste de la provincia de Buenos Aires. *Ciencia del suelo.* 2:131-136.
- PAZOS M.S., MESTELAN S.A. 2002. Variability of Depth to Tosca in Udolls and Soil Classification, Buenos Aires Province, Argentina. *Soil Sci. Soc. Am. J.* 66:1256-1264.
- PEI T., QIN C.-Z., ZHU A.X., YANG L., LUO M., LI B., ZHOU C. 2010. Mapping soil organic matter using the topographic wetness index: A comparative study based on different flow-direction algorithms and kriging methods. *Ecol. Indic.* 10:610-619.
- PEINEMANN N., GUGGENBERGER G., ZECH W. 2005. Soil organic matter and its lignin component in surface horizons of salt-affected soils of the Argentinian Pampa 60:113-128.
- PERALTA N.R., COSTA J.L., BALZARINI M., ANGELINI H. 2013. Delineation of management zones with measurements of soil apparent electrical conductivity in the southeastern pampas. *Can. J. Soil Sci.* 93:205-218.
- PERALTA N.R., COSTA J.L., BALZARINI M., CASTRO FRANCO M., CÓRDOBA M., BULLOCK D. 2015. Delineation of management zones to improve nitrogen management of wheat. *Computers and Electronics in Agriculture.* 110:103-113.
- PIIKKI K., SÖDERSTRÖM M., STENBERG B. 2013. Sensor data fusion for topsoil clay mapping. *Geoderma.* 199:106-116.
- PIKE A.C., MUELLER T.G., MIJATOVIC B., KOOSTRA B.K., POULETTE M.M., PREWITT R.M., SHEARER S.A. 2006. Topographic Indices: Impact of Data Source. *Soil Sci.* 171:809.
- PINHEIRO J., BATES D., DEBROY S., SARKAR D. 2007. Linear and nonlinear mixed effects models. *R package version.* 3:57.

- R DEVELOPMENT CORE TEAM. R: A language and Environment for Statistical Computing. Vienna, Austria. Programa computacional.
- RAD M.R.P., TOOMANIAN N., KHORMALI F., BRUNGARD C.W., KOMAKI C.B., BOGAERT P. 2014. Updating soil survey maps using random forest and conditioned Latin hypercube sampling in the loess derived soils of northern Iran. *Geoderma*. 232:97-106.
- RAMIREZ-LOPEZ L., SCHMIDT K., BEHRENS T., VAN WESEMAEL B., DEMATTÉ J.A., SCHOLTEN T. 2014. Sampling optimal calibration sets in soil infrared spectroscopy. *Geoderma*. 226:140-150.
- REZAEI M.R., LELIEVELDT B.P., REIBER J.H. 1998. A new cluster validity index for the fuzzy c-mean. *Pattern Recognit Lett*. 19:237-246.
- RHOADES J. 1996. Salinity: Electrical conductivity and total dissolved solids 3:417-435.
- RHOADES J., MANTEGHI N., SHOUSE P., ALVES W. 1989a. Soil electrical conductivity and soil salinity: new formulations and calibrations. *Soil Sci. Soc. Am. J.* 53:433-439.
- RHOADES J., RAATS P., PRATHER R. 1976. Effects of liquid-phase electrical conductivity, water content, and surface conductivity on bulk soil electrical conductivity. *Soil Sci. Soc. Am. J.* 40:651-655.
- RHOADES J.D. 1993. Electrical Conductivity Methods for Measuring and Mapping Soil Salinity. En: *Adv Agron*, Academic Press. pp 201-251.
- RHOADES J.D., MANTEGHI N.A., SHOUSE P.J., ALVES W.J. 1989b. Soil Electrical Conductivity and Soil Salinity: New Formulations and Calibrations. *Soil Sci. Soc. Am. J.* 53:433-439.
- ROBINSON G.W. 1922. A new method for the mechanical analysis of soils and other dispersions. *J. Agric. Sci.* 12:306-321.
- ROSSEL R.A.V., MCKENZIE N.J., GRUNDY M.J. 2010. Using Proximal Soil Sensors for Digital Soil Mapping. En: *Digital Soil Mapping*, Springer Netherlands. 79-92 p.
- ROSSITER D. 2012. A pedometric approach to valuing the soil resource, *Digital Soil Assessments and Beyond: Proceedings of the 5th Global Workshop on Digital Soil Mapping 2012*, Sydney, Australia, CRC Press. pp. 25.
- ROUDIER P., BEAUDETTE D., HEWITT A. 2012. A conditioned Latin hypercube sampling algorithm incorporating operational constraints. *Budiman Minasny, et al.* (Eds), En: *Digital soil assessments and beyond*. pp 227-231.
- SABY N., THIOULOUSE J., JOLIVET C., RATIÉ C., BOULONNE L., BISPO A., ARROUAYS D. 2009. Multivariate analysis of the spatial patterns of 8 trace

- elements using the French soil monitoring network data. *Sci Total Environ.* 407:21-25.
- SADRAS V.O., CALVIÑO P.A. 2001. Quantification of Grain Yield Response to Soil Depth in Soybean, Maize, Sunflower, and Wheat. *Agron. J.* 93:577-583.
- SAGA Development Team. System for Automated Geoscientific Analysis (SAGA). 2.0: Programa computacional.
- SALISBURY F.B., ROSS C.W. 1992. *Plant Physiology* Wadsworth Publishing Company, California USA. 682 p.
- SANCHEZ P.A., AHAMED S., CARRÉ F., HARTEMINK A.E., HEMPEL J., HUISING J., LAGACHERIE P., MCBRATNEY A.B., MCKENZIE N.J., DE LOURDES MENDONÇA-SANTOS M. 2009. Digital soil map of the world. *Science.* 325:680-681.
- SATORRE E.H., BENECH ARNOLD R., SLAFER G., DE LA FUENTE E., MIRALLES D., OTEGUI M.E., SAVIN R. 2003. *Producción de granos: Bases funcionales para su manejo.* Facultad de Agronomía, Universidad de Buenos Aires. 783 p.
- SATORRE E.H., SLAFER G.A. 1999. *Wheat: Ecology and Physiology of Yield Determination.* Taylor & Francis. 503 p.
- SCHAETZL R.J., ANDERSON S. 2005. *Soils: Genesis and Geomorphology* Cambridge University Press. 833 p.
- SCHEPERS A.R., SHANAHAN J.F., LIEBIG M.A., SCHEPERS J.S., JOHNSON S.H., LUCHIARI A. 2004. Appropriateness of management zones for characterizing spatial variability of soil properties and irrigated corn yields across years. *Agr. J.* 96:195-203.
- SCHMIDT K., BEHRENS T., DAUMANN J., RAMIREZ-LOPEZ L., WERBAN U., DIETRICH P., SCHOLTEN T. 2014. A comparison of calibration sampling schemes at the field scale. *Geoderma.* 232:243-256.
- SIIA. 2015. Sistema Integrado de Información Agropecuaria. [En línea]. <[www.siiia.gov.ar](http://www.siiia.gov.ar)>. [Consulta:01/05/2015].
- SIMÓN M., PERALTA N., CASTRO FRANCO M., COSTA J.L. 2012. Estimación de propiedades del suelo con la conductividad eléctrica aparente para delimitar zonas de manejo. 19º Congreso Latinoamericano de la ciencia del suelo. 23º Congreso Argentino de la ciencia del suelo Mar del Plata: 6 p.
- SIMÓN M., PERALTA N., COSTA J.L. 2013. Relación entre la conductividad eléctrica aparente con propiedades del suelo y nutrientes. *Ciencia del suelo.* 31:45-55.

- SMITH J.L., HALVORSON J.J. 2011. Field Scale Studies on the Spatial Variability of Soil Quality Indicators in Washington State, USA. *Appl Environ Soil Science*. 2011:7.
- STUDDERT G.A. 2000. Crop rotations and nitrogen fertilization to manage soil organic carbon dynamics. *Soil Sci. Soc. Am. J.* 64:1496-1503.
- SUBBURAYALU S.K., JENHANI I., SLATER B.K. 2014. Disaggregation of component soil series on an Ohio County soil survey map using possibilistic decision trees. *Geoderma*. 213:334-345.
- SUDDUTH K.A., KITCHEN N.R., WIEBOLD W.J., BATCHELOR W.D., BOLLERO G.A., BULLOCK D.G., CLAY D.E., PALM H.L., PIERCE F.J., SCHULER R.T., THELEN K.D. 2005. Relating apparent electrical conductivity to soil properties across the north-central USA. *Comput. Electron. Agric.* 46:263-283.
- SUMNER M.E. 1999. *Handbook of soil science* CRC press. 2148 p.
- TAKATA Y., FUNAKAWA S., AKSHALOV K., ISHIDA N., KOSAKI T. 2007. Spatial prediction of soil organic matter in northern Kazakhstan based on topographic and vegetation information. *Soil Sci. Plant Nutr.* 53:289-299.
- TASI H. 2009. Aplicación de las Cartas de Suelos de Entre Ríos, Argentina, para evaluar índices de productividad específicos para los principales cultivos agrícolas. Tesis para optar el título de Doctor. Universidad de la Coruña - Facultad de Ciencias. La Coruña - España. 619 p.
- TESFA T.K., TARBOTON D.G., CHANDLER D.G., MCNAMARA J.P. 2009. Modeling soil depth from topographic and land cover attributes. *Water Resour. Res.* 45.
- THOMPSON J.A., ROECKER S., GRUNWALD S., OWENS P.R. 2012. Chapter 21 - Digital Soil Mapping: Interactions with and Applications for Hydropedology. En: *Hydropedology*, Academic Press, Boston. pp 665-709.
- TRIANAFILIS J., LESCH S.M. 2005. Mapping clay content variation using electromagnetic induction techniques. *Comput. Electron. Agric.* 46:203-237.
- U.S. ENVIRONMENTAL PROTECTION AGENCY. 2002. *Guidance on Choosing a Sampling Design for Environmental Data Collection (QA/G-5S)*, Washington, D.C. 178 p.
- USDA. 2015. *World Agricultural Supply and Demand Estimates*. [En línea]. [www.usda.gov/oce/commodity/wasde/latest.pdf](http://www.usda.gov/oce/commodity/wasde/latest.pdf). [Consulta:30/05/2015].
- VERIKAS A., GELZINIS A., BACAUSKIENE M. 2011. Mining data with random forests: A survey and results of new tests. *Pattern Recognit.* 44:330-349.
- VERONESI F., CORSTANJE R., MAYR T. 2012. Mapping soil compaction in 3D with depth functions. *Soil Tillage Res.* 124:111-118.

- WALKLEY A., BLACK I.A. 1934. An examination of the Degtjareff method for determining soil organic matter, and a proposed modification of the chromic acid titration method. *Soil Science*. 37:29-38.
- WANG D., SHANNON M., GRIEVE C. 2001a. Salinity reduces radiation absorption and use efficiency in soybean. *Field Crops Research*. 69:267-277.
- WANG G., GERTNER G., PARYSOW P., ANDERSON A. 2001b. Spatial prediction and uncertainty assessment of topographic factor for revised universal soil loss equation using digital elevation models. *ISPRS J. Photogramm. Remote Sens.* 56:65-80.
- WIESMEIER M., BARTHOLD F., BLANK B., KÖGEL-KNABNER I. 2011. Digital mapping of soil organic matter stocks using Random Forest modeling in a semi-arid steppe ecosystem. *Plant Soil*. 340:7-24.
- WILLMOTT C.J., ACKLESON S.G., DAVIS R.E., FEDDEMA J.J., KLINK K.M., LEGATES D.R., O'DONNELL J., ROWE C.M. 1985. Statistics for the evaluation and comparison of models. *J. Geoph Res.* 90:8995-9005.
- WINDHAM M.P. 1981. Cluster validity for fuzzy clustering algorithms. *Fuzzy Sets and Systems*. 5:177-185.
- XIE X.L., BENI G. 1991. A Validity Measure for Fuzzy Clustering. *IEEE Trans. Pattern Anal. Mach. Intell.* 13:841-847.
- YANG L., JIAO Y., FAHMY S., ZHU A., HANN S., BURT J.E., QI F. 2011. Updating conventional soil maps through digital soil mapping. *Soil Sci. Soc. Am. J.* 75:1044-1053.
- ZHITAO Z., LAN Y., PUTE W., WENTING H. 2014. Model of soybean NDVI change based on time series. *International J. Agricul Biolo Eng.* 7:64-70.
- ZHU A.-X., HUDSON B., BURT J., LUBICH K., SIMONSON D. 2001. Soil mapping using GIS, expert knowledge, and fuzzy logic. *Soil Sci. Soc. Am. J.* 65:1463-1472.
- ZHU A., BAND L., VERTESSY R., DUTTON B. 1997. Derivation of soil properties using a soil land inference model (SoLIM). *Soil Sci. Soc. Am. J.* 61:523-533.
- ZHU Q., LIN H. 2011. Influences of soil, terrain, and crop growth on soil moisture variation from transect to farm scales. *Geoderma*. 163:45-54.