

Área de Consolidación: Métodos Cuantitativos para la Investigación Agropecuaria

Relación entre datos de mapas de rendimiento de cultivos de grano y el Índice de Vegetación Diferencial Normalizado (NDVI)

**De la Barrera, Guillermo
Pugliese, Edgardo
Seculin, José
Yañez, Emiliano**

**Tutor: Ing. Agr. Córdoba, Mariano Augusto
Co-tutor: Vega, Andrés**

RESUMEN

Diversas fuentes de información son utilizadas en agricultura de precisión para el estudio de la variabilidad espacial dentro del lote. Entre ellas las mediciones obtenidas a partir imágenes satelitales son ampliamente utilizadas. Los índices espectrales como el Índice de Vegetación Diferencial Normalizado (NDVI) se relacionan con variables agronómicas y resultan útiles para evaluar indirectamente la condición del cultivo. En este trabajo se evalúa la relación entre rendimiento y NDVI y la congruencia entre clasificaciones (zonificaciones) obtenidas con rendimiento y NDVI. Se utilizaron 30 mapas de rendimientos de lotes agrícolas pertenecientes a la zona central de la Provincia de Buenos Aries, a partir de las cuales se obtuvieron valores de NDVI calculados con imágenes satelitales (landsat 7 y 8). Los resultados mostraron una baja correlación entre el NDVI y el rendimiento de los cultivos, tanto cuando se analizamos los valores cuantitativos en bruto y cuando se comparan las zonificaciones obtenidas con ambas variables.

INTRODUCCIÓN

El concepto de agricultura de precisión (AP), hace referencia a la aplicación de tecnologías y principios para manejar la variabilidad espacial y temporal asociada con todos los aspectos de la producción agrícola. Se considera que existen al menos tres aspectos críticos para que su implementación sea exitosa: información, tecnología y gestión (Fernández et. al., 2016).

La mayoría de los sistemas productivos presentan variantes en los factores que los definen: topografía, suelo, régimen hídrico, materia orgánica, disponibilidad de nutrientes, entre otros; que se traducen en diferencias en el potencial de rendimiento, aún dentro de un mismo lote. Si bien este hecho es ampliamente conocido, las prácticas de manejo tradicionalmente aplicadas han sido de tipo uniforme, es decir, basadas en recomendaciones promedio. Recién a partir de las últimas décadas se ha impulsado el desarrollo de tecnologías y principios para manejar la variabilidad espacio-temporal, con el fin de mejorar la eficiencia productiva y optimizar el uso de insumos (Pierce y Nowak, 1999).

El desarrollo de la AP tuvo impulso a partir de los recientes avances tecnológicos en materia de software, sistemas de información geográficos (SIG), sistemas de posicionamiento global (GPS), sensores y controladores de aplicación. El principal desafío de los usuarios de AP ha sido conocer la variabilidad para poder manejarla. Una alternativa de manejo que ha cobrado importancia es la delimitación de zonas de manejo uniforme. Según Doerge (1999) éstas se definen como subregiones dentro de los lotes con características relativamente homogéneas, en las que una proporción de insumos única resulta apropiada. Diversas herramientas se han utilizado para la delimitación de zonas de manejo, entre ellas mediciones topográficas, mediciones georreferenciadas de conductividad eléctrica, mapas de rendimiento e información obtenida a partir de diferentes tipos de sensores remotos. Las imágenes satelitales presentan como ventaja su amplia cobertura espacial y temporal, lo que permite obtener información de campañas anteriores. Los procedimientos para la delimitación de zonas de manejo mediante imágenes satelitales y variables del cultivo, resultan de utilidad mayor para momentos de máxima cobertura y biomasa, en general, cercanos a la floración. Por lo tanto, las observaciones en dicho período permitirían maximizar el comportamiento espectral del cultivo y la delimitación de zonas de manejo (Kemerer, 2013).

Los índices espectrales como el Índice de Vegetación Diferencial Normalizado (NDVI, en inglés: Normalized Difference Vegetation Index) se relacionan con variables agronómicas como biomasa, índice de área foliar, cobertura y rendimiento de los cultivos, por lo tanto,

resultan útiles para evaluar indirectamente la condición del cultivo (Wanjura y Harfield, 1987; Thenkabail et al., 1994; Zhang et al, 1998; Serrano et al., 2000; Rydberg y Söderström, 2000; Kemerer, 2003).

Este trabajo describe el proceso de obtención de imágenes satelitales para los lotes de los que se dispone de información según las bases de datos provistas para el trabajo, asimismo se evalúa la relación entre rendimiento y NDVI y la congruencia entre clasificaciones (zonificaciones) obtenidas con rendimiento y NDVI.

El objetivo del mismo es determinar si hay correlación entre el NDVI y rendimiento.

MATERIALES Y MÉTODOS

1- Datos

Se utilizaron 30 bases de datos de lotes agrícolas pertenecientes a la zona central de la Provincia de Buenos Aries. Para cada base se disponía de información de coordenadas geográficas, tipo de cultivo, rendimiento, pendiente, orientación, índices de rugosidad, relieve y fechas de cosecha. Los cultivos con los que se trabajó fueron: soja, maíz, trigo, girasol, entre otros, teniendo presente que cada lote presentó un solo cultivo. Las campañas agrícolas evaluadas fueron diversas, entre 2007 a 2013. El número de sitios georreferenciados fueron entre 4000 y 11000 sitios, con un promedio de 7000. Para la realización de este trabajo solo se utilizaron los datos referidos a las coordenadas, los cultivos y los rendimientos.

2- Procedimiento de análisis

2.1- Descarga de imágenes satelitales

Inicialmente se cargó la base de datos con el software InfoStat (Córdoba y Balzarini, 2016) y se calculó la media de las coordenadas X e Y. Posteriormente, en la página <https://earthexplorer.usgs.gov/> se cargaron las coordenadas medias obtenidas, se le asignó un rango de tiempo (4 meses anteriores a la cosecha) para obtener las imágenes y se seleccionaron los satélites Landsat 7 y Landsat 8.

Se observaron las imágenes disponibles a descargar (archivos en formato RAR), seleccionándose las de mayor calidad y cantidad de datos, las que fueron guardadas en una carpeta creada para este fin, y que contenía 30 subcarpetas que fueron generadas con los nombres de las distintas bases. Se descargaron un mínimo de 3 imágenes por base.

2.2- Conversión de imágenes satelitales

Se creó una nueva carpeta que contenía las bandas 3 y 4 de las imágenes descargadas, cambiando el nombre de cada imagen por el de la base a la que pertenecían y agregando la sigla B3_ para la banda 3 y B4_ para las de la banda 4.

Posteriormente, se utilizó el software RStudio y en su consola se utilizaron dos rutinas desarrolladas para automatizar el proceso (las mismas pueden solicitarse por email a los autores del trabajo). La primera se usó para convertir las imágenes satelitales descargadas, de tipo 'vectorial' (los puntos que forman la imagen están unidos por líneas) a una imagen *raster* (la imagen está formada por un conjunto de puntos denominados 'bits', ubicados especialmente, en este caso, según las coordenadas).

Se utilizaron 2 librerías o paquetes del software R: *sp* y *raster*, seguidamente se utilizó el primer script, en el que especificó la carpeta donde estaban las imágenes a convertir y la carpeta en la que se guardaron las imágenes convertidas. El tiempo de conversión fue de aproximadamente 5 minutos por imagen, por lo que al procesar 110 imágenes el tiempo total utilizado para esta tarea fue de 9 horas.

El paso siguiente fue usar el segundo script, que compara las bases obtenidas anteriormente (*raster*) con las originales y genera nuevas bases modificadas. Al final del proceso se obtuvieron las nuevas bases procesadas, a las que, a diferencia de las originales, se le agregaron dos columnas de datos, la columna banda 3 y la columna banda 4 con sus respectivos datos. Luego se controló que todas las bases obtenidas tuvieran los datos completos.

2.3- Depuración de datos de rendimiento

La primera operación que realizamos fue la conversión de coordenadas con módulo “Estadística Espacial” del software InfoStat (Córdoba y Balzarini, 2016), seguida por la depuración de los datos de rendimiento extremos, como por ejemplo rendimientos menores a 0.1 t y mayores a 8 t en un cultivo de soja. A continuación, se procedió la depuración de datos con el menú ‘Estadística espacial’, utilizando la opción Índice de Moran local, 2.5 desvío estándar y 20 metros de distancia máxima para la definición del vecindario.

2.4- Clasificaciones

Primero, mediante la herramienta ‘Fórmulas’ se calculó el NDVI para cada conjunto de bandas B3-B4, utilizando la fórmula $([B4] - [B3]) / ([B4] + [B3])$. Luego, se eliminaron datos de NDVI extremos. La relación entre NDVI y el Rendimiento se visualizó mediante un diagrama de dispersión.

Se realizó una interpolación espacial con los datos de rendimiento y NDVI utilizando la función ‘Automapa’ del módulo “Estadística Espacial” en función de las coordenadas transformadas a coordenadas UTM. Esta función realiza la interpolación kriging (Webster y Oliver, 2007) en este caso se usó la opción *kriging* ordinario en el que la media de la variables es estimada localmente para predecir los valores de la variable en sitios no muestreados. En el procedimiento se evaluaron tres modelos de semivariogramas teóricos (exponencial, esférico y gaussiano) para los que la función “Automapa” selecciona automáticamente el de mejor ajuste. Posteriormente, se realizó un análisis conglomerado no jerárquico (análisis multivariado) para clasificación de rendimiento y NDVI en tres conglomerados. Por último, se

realizó una tabla de contingencia para evaluar la congruencia de las dos clasificaciones: conglomerados del rendimiento y del NDVI.

RESULTADOS Y DISCUSIÓN

1- Relación entre rendimiento y NDVI

Se utilizaron las variables 'rendimiento' y 'NDVI', desplegando gráficos de dispersión y adicionando una línea de suavizado para observar la relación entre ambas (Figura 1 y 2). Puede observarse una gran variabilidad en los datos. En la Tabla 1 se presentan los Coeficientes de Correlación y sus p-valor. Se observa que la magnitud de la correlación fue baja, destacando el valor de los lotes 5 y 12 por ser los que mayor coeficiente presentaron. Dichos lotes, son representados en los gráficos de dispersión. La significancia de estos coeficientes, aun cuando la magnitud de la correlación es muy baja, puede atribuirse a la alta variabilidad observada en los datos y a la falta de independencia de las observaciones. En este tipo de datos otras opciones de análisis podrían implementarse como el cálculo de coeficientes de correlación con restricción espacial (Taylor y Bates, 2013).

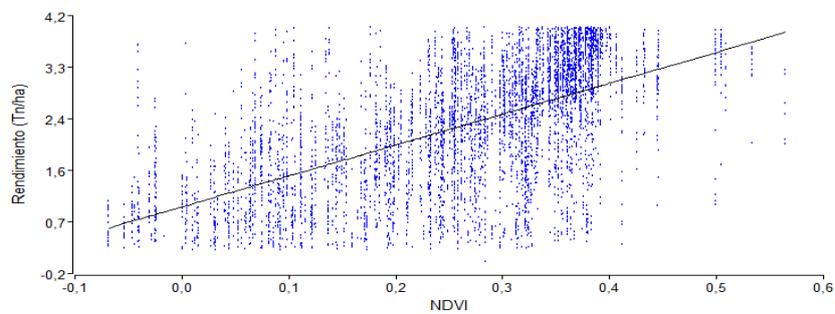


Fig. 1. Relación entre NDVI y rendimiento de girasol para el Lote 5.

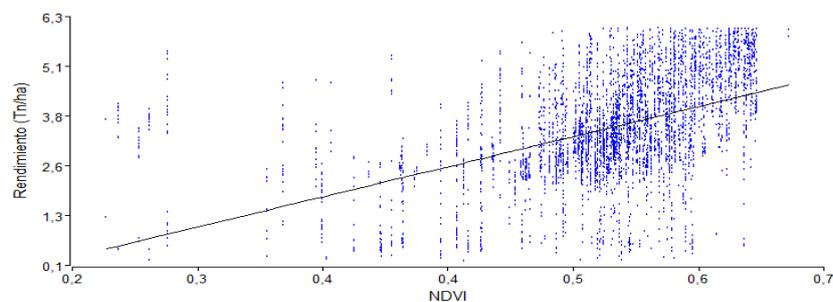


Fig. 2. Relación entre NDVI y rendimiento de cebada para el Lote 12.

Tabla 1. Coeficientes de correlación y p-valor por lote

Lote	Coef. Corr.	P-valor
1	0,36	<0.0001
2	0,04	<0.0291
3	0,27	<0.0001
4	-0.15	<0.0001
5	0,6	<0.0001
6	SD	SD
7	0,37	<0.0001
8	0,45	<0.0001
9	0,20	<0.0001
10	0,42	<0.0001
11	-0,06	<0.0001
12	0,48	<0.0001
13	0,03	<0.236
14	0,06	<0.0001
15	-0,30	<0.0001
16	-0,27	<0.0001
17	0,21	<0.0001
18	0,27	<0.0001
19	0,14	<0.0001
20	0,35	<0.0001
21	0,28	<0.0001
22	0,13	<0.0001
23	0,20	<0.0001
24	0,45	<0.0001
25	0,31	<0.0001
26	-0,17	<0.0001
27	-0,08	<0.0001
28	0,35	<0.0001
29	0,26	<0.0001
30	0,29	<0.0001

2- Predicción y mapeo de la variabilidad espacial

Se visualizaron las variaciones espaciales de los rendimientos en cada lote, con una escala de referencia en toneladas. Los resultados obtenidos mostraron patrones suavizados y no se observaron zonas de cambios bruscos que oscilaran de altos a bajos rendimientos repentinamente (Figura 3).

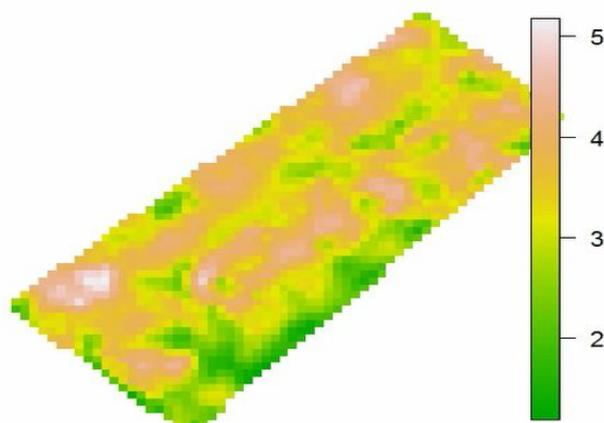


Fig. 3: mapa de predicción de rendimiento, lote 12, cultivo de cebada.

3- Clasificación de los datos

Se realizó un gráfico de dispersión con las coordenadas X e Y particionado por la columna conglomerado del rendimiento. De la misma forma, para los NDVI, se hizo el análisis de conglomerados con 3 zonas, se visualizó en un diagrama de dispersión y se observó que muchas bandas no contaban con todos los datos del NDVI por lo que se optó por elegir los lotes con los valores de NDVI completos.

4- Análisis de congruencia

Se unen el análisis de congruencia de los conglomerados del rendimiento con los del NDVI en valores absolutos, se observó que en el total de los lotes la congruencia del conglomerado del rendimiento y del NDVI fue baja (Tabla 2). Esto también puede observarse visualmente en los mapas de las Figuras 4 y 5. En algunos lotes solo se observó una congruencia entre zonas clasificadas como de Alto o Medio (lotes 4, 10, 13, 14, 20, 21).

Tabla 2. Análisis de congruencia entre rendimiento y NDVI de los conglomerados. Lote 1 a 15.

Lote	Clasificación de Rendimiento	Clasificación de NDVI (%)		
		Alto	Medio	Bajo
1	Alto	31	42	27
	Medio	46	38	16
	Bajo	13	30	57
2	Alto	18	81	1
	Medio	26	72	2
	Bajo	18	81	1
3	Alto	17	-	83
	Medio	39	-	61
	Bajo	8	-	92
4	Alto	85	11	4
	Medio	94	5	1
	Bajo	90	9	1
5	Alto	49	34	17
	Medio	17	47	36
	Bajo	4	17	79
7	Alto	27	55	18
	Medio	13	50	36
	Bajo	5	39	56
8	Alto	33	53	14
	Medio	15	46	40
	Bajo	4	28	68
9	Alto	32	56	12
	Medio	10	66	23
	Bajo	11	49	39
10	Alto	38	61	1
	Medio	14	85	1
	Bajo	4	95	1
11	Alto	1	73	27
	Medio	1	95	4
	Bajo	4	76	20
12	Alto	70	28	2
	Medio	29	65	6
	Bajo	18	54	28
13	Alto	1	90	9
	Medio	5	82	13
	Bajo	1	77	22
14	Alto	6	85	8
	Medio	9	83	9
	Bajo	11	80	9
15	Alto	1	6	93
	Medio	1	21	78
	Bajo	1	45	54

Tabla 2. Análisis de congruencia entre rendimiento y NDVI de los conglomerados. Lote 16 a 30.

Lote	Clasificación de Rendimiento	Clasificación de NDVI (%)		
		Alto	Medio	Bajo
16	Alto	1	11	88
	Medio	2	29	69
	Bajo	9	28	63
17	Alto	75	21	4
	Medio	69	26	5
	Bajo	51	41	8
18	Alto	65	25	10
	Medio	44	42	14
	Bajo	31	52	17
19	Alto	2	91	7
	Medio	1	95	4
	Bajo	1	88	11
20	Alto	83	14	3
	Medio	65	28	6
	Bajo	38	47	15
21	Alto	77	19	4
	Medio	64	24	10
	Bajo	42	48	10
22	Alto	39	27	34
	Medio	26	23	51
	Bajo	25	25	50
23	Alto	56	38	6
	Medio	45	46	9
	Bajo	33	46	21
24	Alto	66	31	3
	Medio	49	40	11
	Bajo	16	40	44
25	Alto	71	24	4
	Medio	42	54	5
	Bajo	27	48	25
26	Alto	49	-	51
	Medio	75	-	25
	Bajo	71	-	29
27	Alto	13	72	15
	Medio	10	52	38
	Bajo	43	36	21
28	Alto	1	89	10
	Medio	1	76	24
	Bajo	-	74	25
29	Alto	64	26	9
	Medio	53	19	28
	Bajo	45	22	33
30	Alto	72	21	7
	Medio	65	23	12
	Bajo	35	36	29

5- Evaluación de Concordancia de Clasificaciones

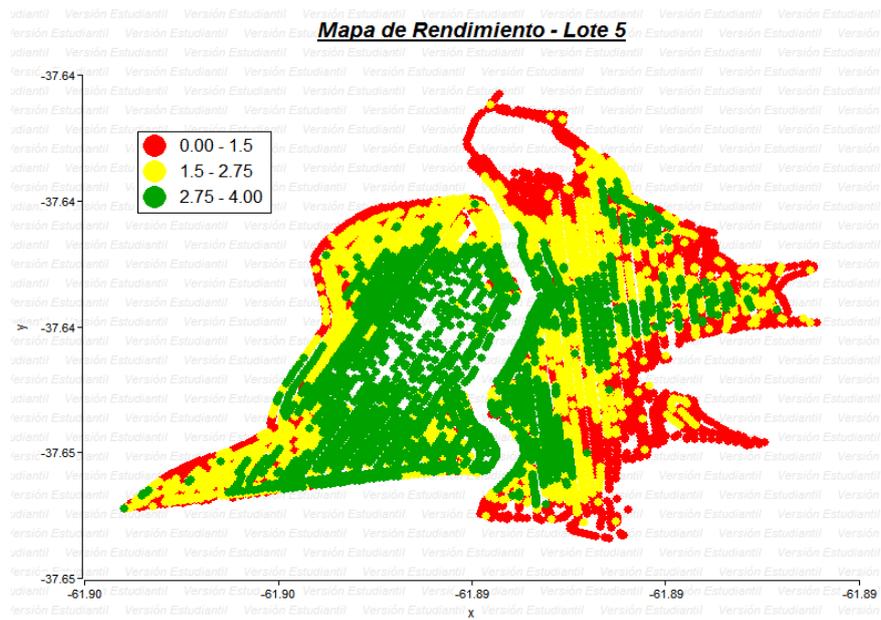


Fig. 4: Mapa de rendimiento del lote 5. En la leyenda los rangos de rendimiento en Tn/ha de cada zona definida por cada conglomerado.

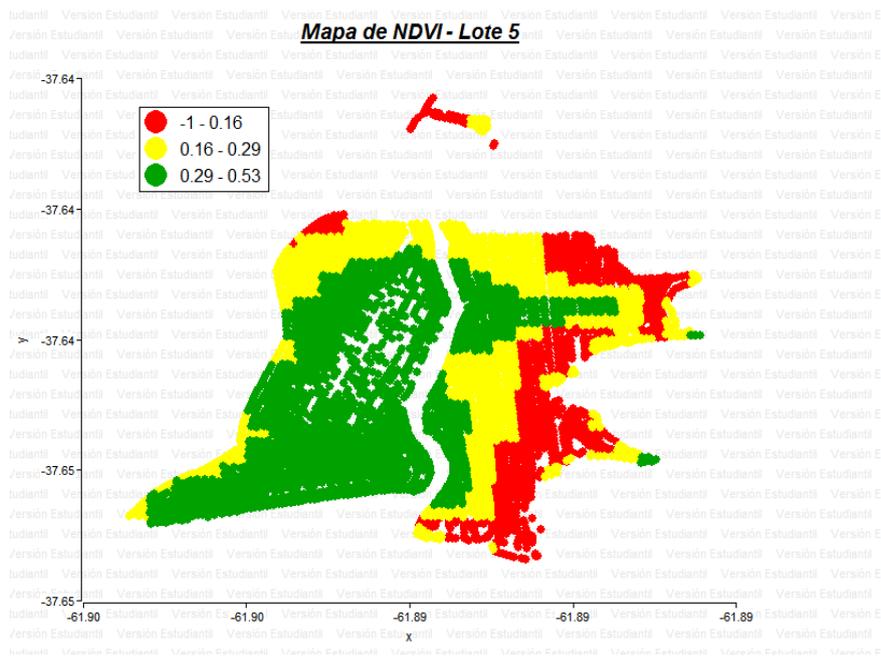


Fig. 5: Mapa de NDVI del lote 5. En la leyenda los rangos de NDVI de cada zona definida por cada conglomerado.

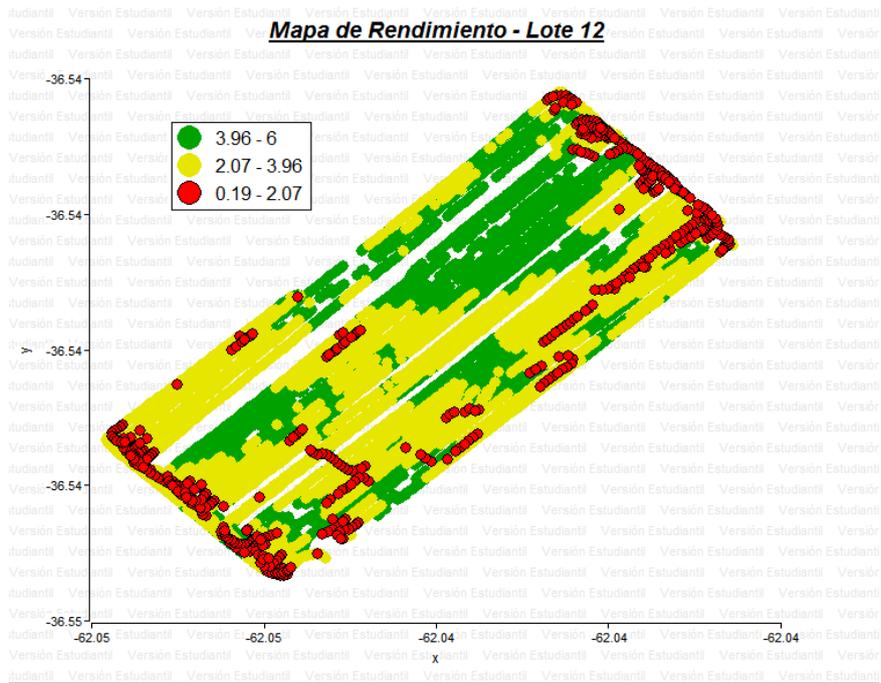


Fig. 6: Mapa de rendimiento del lote 12. En la leyenda los rangos de rendimiento en Tn/ha de cada zona definida por cada conglomerado.

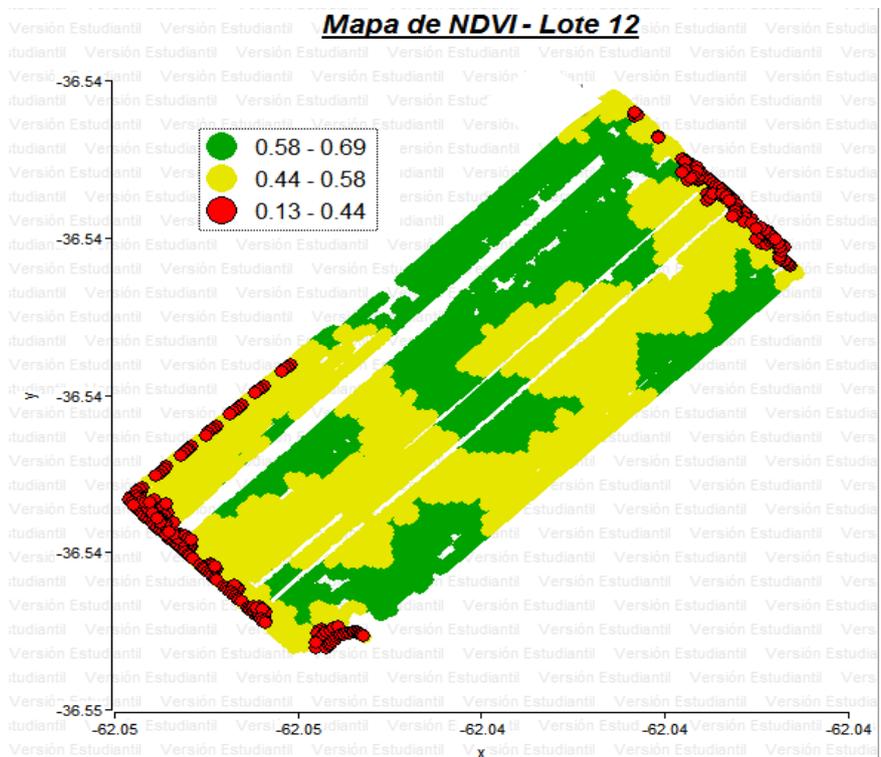


Fig. 7: Mapa de NDVI del lote 12. En la leyenda los rangos de NDVI de cada zona definida por cada conglomerado.

La baja correlación entre NDVI y Rendimiento de los cultivos analizados, puede deberse principalmente a la falta de exactitud de las imágenes disponibles que se extrajeron de los servidores. Tal vez debería considerarse repetir el método utilizado, pero con una fuente distinta que proporcione datos de B3 y B4 con un chequeo previo de los procesos de calibración y corrección de imágenes, y en fecha deseada. En otro trabajo donde se analizan la correlación entre rendimiento y NVDI, Bocco et. al. 2009, afirma que es mejor contar con datos de más de un satélite (en su caso uso LANSAT y SPOT) para obtener un mejor ajuste al estimar rendimientos a través de NDVI, no obstante la frecuencia temporal de estos satélites y la posibilidad de nubosidad durante sus pasadas son, a veces, una limitante.

Otra consideración necesaria es que, en el caso de cultivos como la soja, el rendimiento no tiene una alta correlación con la variable NDVI, que en definitiva se relaciona con aspectos tales como biomasa, IAF, intercepción de radiación fotosintéticamente activa (rfa), entre otros, a diferencia de otros cultivos, como por ejemplo el maíz y la caña de azúcar, en los que sí se advierte esta correlación porque dependen en mayor medida de ellas.

CONCLUSIÓN

Es notable la importancia de esta herramienta geoestadística en el análisis de datos para la recolección de información y toma de decisiones a nivel predial. Si bien se considera que los procedimientos utilizados para analizar los datos geoestadísticos fueron adecuados no se pudo encontrar una relación importante entre el NDVI y el rendimiento de los diferentes lotes cuando analizamos los valores cuantitativos en bruto y tampoco al observar las zonificaciones obtenidas con ambas variables.

BIBLIOGRAFIA

Bocco, M., Heredia, S., Ovando, G., Sayago, S., Willington, E.: Estimación de cobertura vegetal con redes neuronales a partir de imágenes de distintos satélites. In: X Congreso Argentino de Ingeniería Rural y II del MERCOSUR, Rosario, pp. 1250–1256 (2009).

Córdoba M., Balzarini M. 2016. Análisis de datos espaciales en el software InfoStat. III Jornadas Académicas de Estadística – Geoestadística – Sociedad Argentina de Estadística. Neuquén, Agosto de 2016.

Doerge, T.A. 1999. Management zone concepts. SSMG-2. En: Information Agriculture Conference, 9 al 11 de Agosto de 1999, Stewart Center, Purdue, IN.
Fernandez et. al. 2016. Frontiers in Plant Science- Supplementary Material.

Kemerer, A.C. 2003. Utilización de imágenes satelitales en agricultura de precisión para delimitar zonas de manejo uniforme. Tesis de grado. Facultad de agronomía, Universidad Nacional de Entre Ríos, Oro Verde.

Luchiari et al. 2000. Strategies for establishing management zones for site specific nutrient management. Precision Agriculture Center, ASA, CSSA and SSSA, Madison.

Pierce y Nowak, 1999. Advance in Agronomy Nro. 67. Academic Press.

Richard Webster y Margaret A. Oliver , 2007 .Geostatistics for Environmental Scientists. Second Edition. Wiley, Chichester.

Rydberg y Söderström, 2000. Potential crop growth assessment from remotely sensed images compared to ordinary yield maps. In: Proc. 5th International Conference on Precision Agriculture. Precision Agriculture Center, ASA, CSSA and SSSA, Minesota.

Serrano et al., 2000. Remote Sensing of Biomass and Yield of Winter Wheat under Different Nitrogen Supplies Crop Science.

Taylor, J.A. & Bates .(2013).A discussion on the significance associated with Pearson's correlation in precision agriculture studies.T.R. Precision Agriculture 14: 558.

Thenkabail et al., 1994. Landsat-5 Thematic Mapper Models of Soybean and Corn Crop Characteristics. International Journal of Remote Sensing.

Wanjura y Harfield, 1987. Sensitivity of spectral vegetative indices to crop biomass. Transaction of the ASSE.

Zhang et al, 1998. Corn and Soybean Yields Indicators Using Remotely Sensed Vegetation Index. In: Proc. 4th International Conference on Precision Agriculture. St. Paul. MN.