

Use of satellite vegetation indices to predict yield levels of *Vicia villosa* Roth

Utilización de índices de vegetación satelitales para predecir niveles de rendimiento de *Vicia villosa* Roth

F. Marini¹, M.B. D'Amico², G.L. Calandrini², J.P. Renzi^{3,4}, and G.R. Chantre^{4,5}

¹ Instituto Nacional de Tecnología Agropecuaria - INTA
 Agencia de Extensión Bahía Blanca de la Estación Experimental Bordenave,
 Bahía Blanca, Argentina

² Instituto de Investigaciones en Ing. Eléctrica (UNS - CONICET)
 Depto. Ingeniería Eléctrica y de Computadoras, Universidad Nacional del Sur
 San Andrés 800, Bahía Blanca, Argentina

³ Instituto Nacional de Tecnología Agropecuaria - INTA
 Estación Experimental Hilario Ascasubi, Argentina

⁴ Depto. de Agronomía, Universidad Nacional del Sur, Bahía Blanca, Argentina

⁵ Centro de Recursos Naturales Renovables de la Zona Semiárida (UNS-CONICET)
 Bahía Blanca, Argentina

mbdamico@uns.edu.ar, marini.fabian@inta.gob.ar, calandri@criba.edu.ar,
renzipugni.juan@inta.gob.ar, guillermo.chantre@uns.edu.ar

Abstract. In this paper, satellite images, machine learning algorithms and field measurements are combined to analyze the possibility of generating a model that predicts the yield of the legume *Vicia villosa* Roth before harvest. In a previous study, satellite information from different dates throughout the complete phenological cycle of the legume cultivated in several plots of the Guaminí district (Buenos Aires) was used and a close relationship was found between yield level and the time series of vegetation indices. Based on these results, the possibility of predicting yield is evaluated considering nine dates of the 2021-2022 campaign between sowing and prior to harvest. They are associated in different ways, determining their impact on the accuracy of the trained model. Results show that legume yield can be adequately classified with remote monitoring of five dates. To have a prediction model could help in in-situ decisions by optimizing the use given to the crop (direct grazing, forage or seed production) based on its expected profit.

Resumen En este trabajo se combinan las imágenes satelitales, los algoritmos de aprendizaje de máquina y las mediciones de campo para analizar si es posible generar un modelo de predicción del rendimiento

Received May 2025; Accepted June 2025; Published July 2025



This work is under a Creative Commons
 Attribution – NonComercial – Share Alike 4.0 International License

de la leguminosa *Vicia villosa* Roth (VV) antes de su cosecha. En un estudio previo, se empleó información satelital de diferentes fechas a lo largo del ciclo fenológico completo de VV cultivada en varios lotes del partido de Guaminí (provincia de Buenos Aires) y se encontró una estrecha relación entre la serie temporal de los índices de vegetación y el rinde de dichos cultivos. En base a esos resultados, se evalúa la posibilidad de predecir el rendimiento a partir de nueve fechas de la campaña 2021-2022 entre la siembra y la cosecha. Las mismas se asocian de distintas maneras determinando su impacto sobre la precisión del modelo entrenado. Los resultados evidencian que con el monitoreo remoto de cinco fechas es posible clasificar adecuadamente el rendimiento de VV. Tener un modelo de predicción ayudaría en las decisiones in-situ optimizando el uso que puede darse al cultivo de VV (pastoreo directo, forraje o producción de semillas) en función del rendimiento esperado.

Keywords: model · prediction · yield · legume.

Palabras clave: modelo - predicción - rendimiento - leguminosa.

1 Introducción

En la actualidad, la tecnología desempeña un papel fundamental en el sector agrícola tanto a nivel mundial como nacional. En ese contexto, el uso de herramientas digitales está adquiriendo cada vez mayor relevancia en la planificación y toma de decisiones relacionadas con el manejo de los cultivos, logrando mejores eficiencias y rentabilidades. En los últimos años, el acceso a la información se ha vuelto más sencillo y la cantidad de datos que pueden generarse ha aumentado considerablemente. Así, por ejemplo, se encuentran disponibles, de manera libre y gratuita, imágenes satélites que ofrecen resoluciones temporales, espaciales y espectrales aptas para trabajar tanto en grandes áreas como a nivel de lote y también durante la totalidad del ciclo fenológico del cultivo [1-3]. Por otro lado, las cosechadoras más avanzadas cuentan con módulos que incorporan sensores especializados y software para relevar diversos datos georeferenciados durante la cosecha.

Sin embargo, integrar y convertir esta vasta cantidad de datos satelitales y de campo en información relevante y agronómicamente útil para la planificación y la toma de decisiones sigue representando un desafío considerable. La tendencia natural consiste en aplicar algoritmos de procesamiento avanzados, como pueden ser aquellos basados en aprendizaje de máquina (o *machine learning*), que logran simplificar y manejar eficientemente volúmenes masivos de datos. Estas herramientas elaboran modelos computacionales capaces de identificar tendencias significativas, clasificar los datos o predecir resultados de manera precisa, a pesar de la complejidad y la variabilidad inherente al entorno agrícola [4-10]. A pesar de sus conocidas capacidades, la solidez y confiabilidad del modelo obtenido dependerá de múltiple factores como la estructura interna elegida, la representatividad, la calidad de los datos, etc.

El propósito de este trabajo es explorar la posibilidad de generar un modelo computacional capaz de predecir el rendimiento de semillas de la leguminosa *Vicia villosa* Roth (VV) realizando monitoreo remoto del cultivo durante su ciclo fenológico. Para su construcción se emplea un algoritmo de clasificación supervisada basado en aprendizaje de máquina [11]. La información de entrada son los índices de vegetación calculados en base a las imágenes satelitales. Los datos de entrenamiento y validación son las mediciones de rendimiento brindadas por las cosechadoras. En trabajos previos se observó que, a diferencia de otros cultivos, las series temporales de los índices de vegetación correspondientes a VV no presentan el clásico escenario donde existe un día crítico en el que se manifiesta un valor marcadamente más alto y que además está relacionado al nivel de rinde logrado [12,13]. Sin embargo, sí se encontró que existe una sólida correlación entre el rinde y el apilamiento multitemporal de los índices para el ciclo fenológico completo, desde su siembra hasta su cosecha.

En base a estos resultados, se busca ahora analizar e identificar las fechas o conjuntos de fechas dentro del ciclo que resultan más relevantes para lograr la buena predicción de los rendimientos. La premisa principal es poder anticipar el nivel de rendimiento con suficiente antelación a la cosecha. Por ello, se considerarán las fechas hasta la floración previo al marchitamiento o secado del cultivo. Esta leguminosa posee múltiples beneficios como el aporte de nitrógeno al suelo, el pastoreo de bovinos y ovinos, la confección de rollos de forraje y la producción de semillas [14,15]. Sin embargo, su uso es resistido por los productores debido a que la cosecha de semillas puede ser muy variable de un año a otro, de lote a lote o, incluso, dentro del mismo lote. Tener un modelo de predicción que permita anticipar el rendimiento de VV de manera confiable y efectiva ayudaría en las decisiones *in-situ* optimizando el uso que puede darse al cultivo en función del rendimiento esperado [16].

El presente trabajo se organiza de la siguiente manera. En la Sección 2 se resumen las principales características del área de estudio y de la especie de *Vicia* estudiada. La información satelital utilizada, los índices de vegetación calculados, el tratamiento estadístico de los datos disponibles como así también las clasificaciones supervisadas implementadas se describen en la Sección 3. Los distintos resultados obtenidos se presentan en la Sección 4. Finalmente, las conclusiones y trabajos a futuro se incluyen en la Sección 5.

2 Area y cultivo de estudio

El partido de Guaminí se encuentra en la provincia de Buenos Aires (República Argentina) dentro de la región pampeana austral. Dicho partido es atravesado en su sector central por el Sistema de las Lagunas Encadenadas, un lineamiento estructural que ha dado lugar a una franja deprimida ocupada por un rosario de lagunas de drenaje endorreico y declive al sudoeste. El Sistema de Lagunas Encadenadas se caracteriza también por delimitar un cambio edafológico al norte y al sur del partido. De esta manera, se observan suelos con texturas gruesas (arenosas) al norte de dicho sistema lagunar y de texturas más finas (francos y

franco-arcillosos) al sur del mismo. Como consecuencia de ello, se ha evidenciado una notoria diferencia de desarrollo entre los cultivos estivales localizados al norte de los localizados al sur del mencionado Sistema [17].

El área de trabajo se localiza al sur del partido de Guaminí, una de las zonas de mayor relevancia agro productiva del país. El sistema productivo de Guaminí ha ido mutando paulatinamente de una tradicional condición ganadero-agrícola hacia una mayor proporción de la agricultura. Esta característica se verifica especialmente para los cultivos de verano, que han alcanzado una significativa y creciente expansión [18]. Tal propagación se ha ido incrementando de manera constante durante los últimos treinta años, siendo la soja (*Glicine max*) el cultivo que evidenció el mayor crecimiento, al igual que ocurrió en gran parte de la llanura pampeana.

Entre los cultivos de invierno de incipiente desarrollo en la zona, se encuentra *Vicia villosa* Roth. Dentro de la familia de las leguminosas, el género *Vicia* está compuesto por 180 a 210 especies que se cultivan con varios fines. En particular, VV es la segunda especie de *Vicia* más cultivada a nivel mundial para el consumo humano, el pastoreo directo, forraje y abono verde [14]. No es exigente en cuanto a los suelos y puede crecer tanto en ambientes semiáridos como húmedos.

En el área de trabajo, el cultivo de VV se destina a la producción de semillas. Se siembra entre fines de mayo y fines de junio (otoño en el Hemisferio Sur). Después de 120-130 días, entra en estado de prefloración. La presencia de agentes polinizantes, como las abejas, es crucial para su producción de semillas debido a su reproducción cruzada. La floración de VV es prolongada produciéndose desde noviembre hasta mediados de diciembre. Luego, comienza a secarse y maduran sus frutos.

Independientemente de que se logre la polinización, el rendimiento de semilla de VV es usualmente muy variable como se evidencia en los lotes que se encuentran dentro del área de estudio. Dichos lotes abarcan una superficie de 307,5 has y la distribución de los rindes durante la campaña 2021-2022 arrojaron niveles de 0,00 a 0,22 t/ha hasta de 0,65 a 1,71 t/ha. Al comparar los lotes, se pudo apreciar además que hay una variabilidad significativa en los rindes obtenidos. Algunos lotes parecen tener rindes predominantemente bajos, otros tienen rindes predominantemente altos, y otros tienen zonas espaciales bien definidas con diferentes niveles de producción. La distribución de los manchones por área de lote tampoco fue homogénea. Solo en algunos lotes se lograron rindes mayores a 1 t/ha, que serían los valores deseables.

3 Metodología

La información satelital se obtiene a través de la base de datos del satélite Sentinel -2 que cuenta con un total de trece bandas del espectro electromagnético: cuatro de ellas (2, 3, 4 y 8) con una resolución espacial de 10 m, seis (5, 6, 7, 8A, 11 y 12) de 20 m y las últimas tres (1, 9 y 10) de 60 m. Del total de bandas, tres de ellas se encuentran en la región del rojo cercano con longitudes de onda central de 705, 740 y 783 nm (bandas 5, 6 y 7). En virtud de estas

Cuadro 1. Fórmulas de los IV utilizados en el estudio de VV.

Nombre	Fórmula	Referencia
NDVI	$\frac{B8 - B4}{B8 + B4}$	[20]
EVI	$\frac{2,5(B8 - B4)}{B8 + C_1 B4 - C_2 B2 + L}$	[21]
NDVI red-edge 5	$\frac{B8 - B5}{B8 + B5}$	[22]
LAI SeLi	$\frac{B8A - B5}{B8A + B5}$	[6]
SR	$\frac{B8}{B11}$	[4]
SD	$B8 - B12$	[4]

ventajas en su resolución multiespectral, dichas bandas se emplearon para la obtención de varios índices de vegetación (IV). Los cálculos se realizaron para varias fechas del año 2021. Las imágenes fueron luego procesadas utilizando la plataforma de geomática basada en la nube *Google Earth Engine* [19], que permite realizar análisis geoespaciales con una gran capacidad de procesamiento. A fin de contar con parámetros comparables en el tiempo se trabajó con el producto de reflectancia en nivel 2A corregido atmosféricamente a nivel del suelo (BOA).

3.1 Obtención de los Índices de vegetación

Las fórmulas de los seis IV utilizados en este trabajo se presentan en el Cuadro 1. El NDVI (Índice de Vegetación de Diferencia Normalizada), es uno de los que mayor difusión ha alcanzado en los estudios que involucran a la vegetación. No obstante, puede verse afectado por los diversos niveles de humedad del suelo y también suele saturarse ante elevadas densidades de cultivo [23]. Para atenuar estas desventajas, se han propuesto diversas alternativas como el Índice de Vegetación Mejorada (EVI), que posee una mayor sensibilidad en regiones de alta biomasa, así como una reducción de las influencias atmosféricas y del suelo. Los coeficientes C_1 y C_2 están relacionados a la resistencia de la atmósfera y L ajusta el NIR diferencial y la transferencia de radiación roja a través del dosel. Los valores por defecto son $C_1 = 6$, $C_2 = 7,5$ y $L = 1$.

Por su parte, aprovechando la existencia de la región del rojo cercano en satélites como el Sentinel-2, se han desarrollado los índices de Diferencia Normalizada de Borde Rojo, que emplean la reflectancia en dicha franja del espectro electromagnético (entre 0,7 y 0,8 μm) y permiten estimar el contenido de clorofi-

la y nitrógeno de la vegetación y de esa manera monitorear su salud y desarrollo. Otro índice de vegetación muy difundido es el Índice de Área Foliar (LAI), que representa el área total de la hoja fotosintéticamente activa por unidad de suelo [24]. En tal sentido, se ha desarrollado un índice alternativo denominado SeLi (Sentinel-2 LAI Green Index) empleado con éxito para la estimación de LAI verde a partir de datos del satélite Sentinel-2 [6]. Paralelamente, se han empleado con buenos resultados otros índices simples como el cociente SR y el de diferencias entre bandas SD [25].

3.2 Elección de las fechas

La Fig. 1 muestra a modo de ejemplo la variación promedio de la serie temporal del índice NDVI en uno de los lotes del área de estudio [12,13]. Observando la curva a lo largo del año, se evidencia que el índice depende del ciclo fenológico de VV. Los valores son bajos durante la siembra y la emergencia hasta los últimos días de agosto. Luego, presentan un crecimiento abrupto en septiembre para mantenerse en niveles altos durante octubre hasta los últimos días de noviembre cuando se encuentra en floración. La curva desciende para la época en la que el cultivo comienza a secarse y maduran los frutos, mostrando una caída abrupta en enero donde se aplica un desecante que reduce la humedad y acelera el secado previo a la cosecha. A diferencia de otros cultivos no se observa el típico escenario donde existe un día crítico en el que los índices manifiestan un valor marcadamente más alto y que además esté relacionado al nivel de rinde logrado. Teniendo en cuenta este comportamiento y considerando que el objetivo es encontrar un modelo de predicción del rendimiento de VV, se seleccionaron todas las fechas disponibles y con ausencia de nubes unos 60 días luego de la siembra del cultivo hasta unos 60 días antes de la cosecha.

Las escenas HND del año 2021 que se recolectaron corresponden a las fechas entre los meses de agosto y noviembre que se detallan en Fig. 1. Una vez obtenidos los valores de los diferentes IV para cada fecha, se los trabajó como una multibanda, agrupándolos en una única imagen. Este apilamiento (*stacking*) se realizó para los cinco índices y considerando diferentes combinaciones de las fechas elegidas dentro del ciclo fenológico de VV. Las combinaciones se hicieron buscando indentificar cuál de los casos resulta más relevante para lograr una buena predicción de los rendimientos. En un primer ensayo (Caso 1), se tuvieron en cuenta las nueve fechas recolectadas (4/8 al 22/11). Las pruebas siguientes consistieron en reducir progresivamente el número de imágenes en el apilamiento para así obtener un modelo que funcione correctamente con la menor cantidad de fechas posibles y a su vez analizar la influencia que cada instancia del ciclo fenológico puede tener sobre la clasificación del rendimiento. Inicialmente, se eliminó la fecha del 4/8 (Caso 2) pues de acuerdo a los resultados previos de [12,13] los IV crecen significativamente a partir de mediados de agosto. Luego, se consideraron las alternativas de trabajar con las cinco últimas fechas del período considerado (Caso 3), las cuatro últimas fechas (Caso 4), las tres últimas fechas (Caso 5) y las dos últimas fechas (Caso 6).

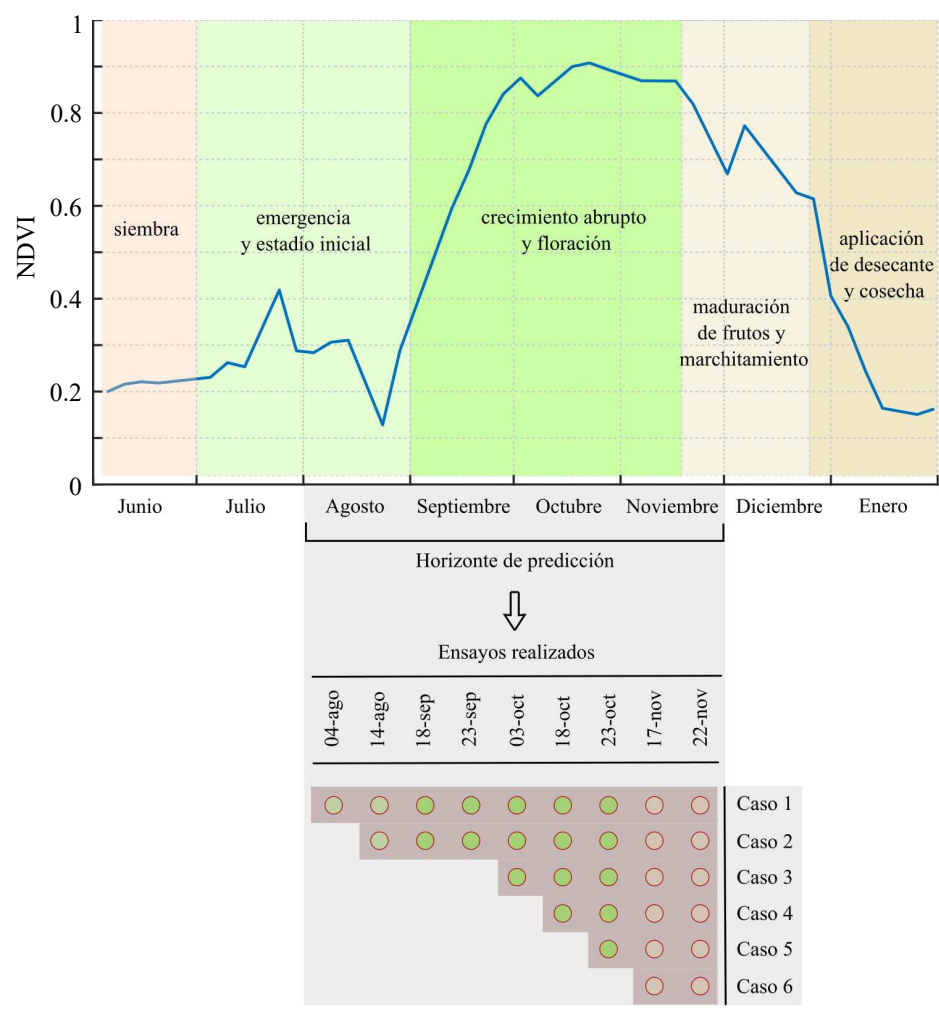


Figura 1. Ejemplo del comportamiento que presenta en promedio la serie temporal del índice NDVI a lo largo de un año dentro de uno de los lotes que se encuentra en el área de estudio. Además se incluyen las fechas elegidas y las combinaciones realizadas para entrenar el modelo de predicción de rinde de VV.

3.3 Análisis estadístico

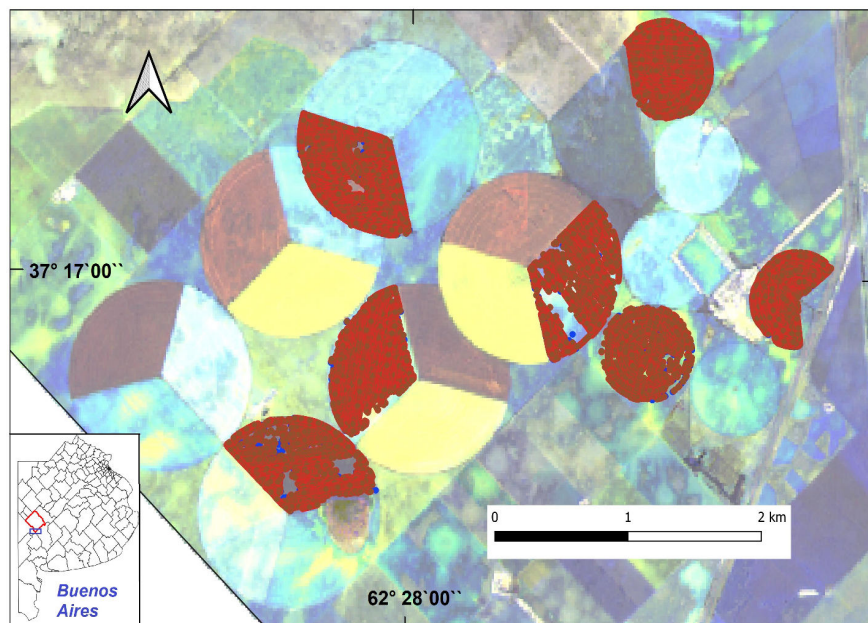


Figura 2. Localización de los siete pivots de riego de VV en el sur del partido de Guaminí. En rojo se aprecian los puntos de registro de rindes.

El área de estudio está formada por siete pivots de riego de los cuales se tienen los mapas de rendimiento de VV obtenidos a través de las cosechadoras durante la campaña 2021-2022. En total, se cuenta con 88929 puntos de muestreo que se distribuyen como se muestran en la Fig. 2. Mediante el empleo del software gratuito QGis, se seleccionaron aleatoriamente el 70 % de los puntos de muestreo para clasificación y el 30 % para validación. Siguiendo los desarrollos realizados en [13], se mantuvo la clasificación de los valores de rendimiento en tres categorías pero corregidos levemente en la forma:

- Categoría 1: de 0 a 200 kg/ha (rindes bajos)
- Categoría 2: de 200 a 600 kg/ha (rindes medios)
- Categoría 3: mayores a 600 kg/ha (rindes altos)

Así, las muestras se distribuyeron con las siguientes proporciones: 21, 16 % en la Categoría 1, 44, 10 % en la Categoría 2 y 34, 74 % en la Categoría 3.

Como se cuenta con varios puntos de rinde dentro de un mismo píxel, se optó además por realizar el remuestreo (resampling) de las imágenes llevando el tamaño del píxel de cada banda involucrada en los cálculos artificialmente a 1 metro [13]. Como se ilustra en la Fig. 3, este procedimiento permite que

cada punto de muestreo quede asignado a un único píxel y que entonces puedan utilizarse la totalidad de los datos del mapa en la generación de los modelos.

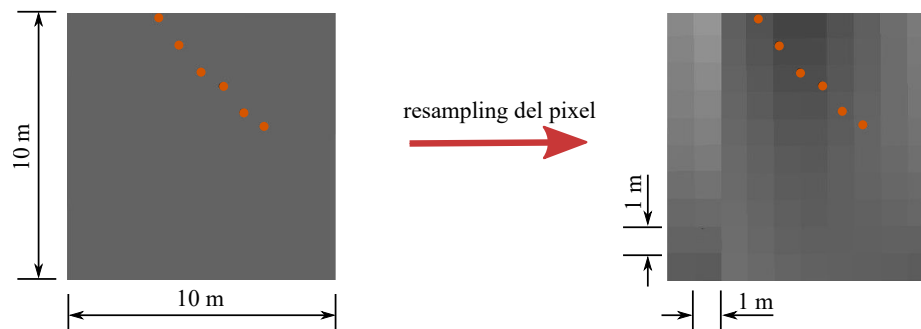


Figura 3. Ejemplo de distribución de los puntos de muestreo dentro de un píxel de 10 m de resolución espacial del satélite Sentinel-2 (izquierda). Efecto sobre dichos puntos al aplicar el remuestreo (resampling) a 1 m de resolución (derecha).

3.4 Clasificación y validación de las imágenes

Para generar los posibles modelos de predicción, se empleó el método de clasificación supervisada *Random Forest* (RF). Este método ha sido aplicado con éxito en problemas de clasificación y regresión [11] y se basa en la definición de un conjunto de árboles de decisión, con la incorporación de aleatoriedad en la construcción de cada clasificador individual. Dicha aleatoriedad se introduce en la partición del espacio de características o variables tanto durante la construcción del árbol de decisión como en la muestra de entrenamiento. Los nodos de cada árbol se dividen utilizando la mejor variable escogida de un subconjunto de variables predictivas seleccionadas al azar. Los errores individuales de cada árbol se compensan al realizar la asignación final por mayoría [5].

Para conocer la precisión en cada una de las clasificaciones realizadas, se obtuvieron las correspondientes matrices de confusión sobre los datos test representativos de las verdades de campo de cada cubierta. A partir de estas matrices se derivan diversos índices de exactitud como el índice de precisión global (P), que proviene del cociente entre el total de píxeles correctamente clasificados y el total de píxeles correspondientes a las verdades de campo [26]. Otro índice es el coeficiente Kappa global (K), que representa el grado de ajuste entre la imagen clasificada y la realidad observada en el terreno [27]. Valores de K cercanos a 1 indican un excelente acuerdo entre la realidad observada y el mapa obtenido.

4 Resultados

Las clasificaciones supervisadas se realizaron empleando los apilamientos estacionales (*stacks*) descritos anteriormente y considerando cada uno de los seis

IV seleccionados. El algoritmo de RF se implementó considerando 50 árboles de decisión. En lo que respecta al resto de los parámetros, el tamaño mínimo del nodo terminal utilizado (Min Leaf Population) fue de 1 y la proporción de observaciones a ser muestreada en cada árbol (bag fraction) fue de 0.5. Los resultados obtenidos se resumen en el Cuadro 2. Como podría esperarse, los mejores desempeños se obtienen, en general, cuando la clasificación involucra la mayor cantidad de fechas posibles. De hecho, los valores de exactitud global para el Caso 1 se encuentran entre 82 y 86 % con índices Kappa de alrededor de 0.7. Al eliminar la fecha de principios de agosto quedando así la información del ciclo fenológico desde el 14/8 al 22/11, las precisiones tienden a disminuir pero no de manera significativa como para que se justifique considerar dicha fecha. Incluso con uno de los índices la respuesta de la clasificación para el Caso 2 mejora. Esto ocurre con LAI SeLi, que es el índice con el que se obtuvieron los mejores valores de P y K (86 % y 0.72 respectivamente).

Las clasificaciones obtenidas en el Caso 3 que considera cinco fechas del ciclo fenológico (3/10 al 22/11) presentan disminuciones en P y K pero tampoco de forma significativa. Algo similar se repite con las clasificaciones que resultaron en el Caso 4 que tiene en cuenta cuatro fechas (18/10 al 22/11). Salvo por los cálculos realizados con EVI, el resto de las experiencias muestran valores de P que se mantienen cercanos al 80 %. En particular, el índice LAI SeLi presenta una exactitud global de 83,61 % cuando se toman las últimas cinco fechas, disminuyendo al 80 % al utilizar las últimas cuatro. Estos desempeños se degradan notoriamente en las clasificaciones de los Casos 5 y 6 que involucran sólo las últimas tres y dos fechas, respectivamente. Las precisiones globales llegan a valores por debajo del 60 % con índices de confianza que apenas llegan a 0.16.

El Cuadro 3 muestra un detalle de las precisiones obtenidas por categoría en el caso de las clasificaciones supervisadas que mejores resultados arrojaron en base a los valores de P y K. Como se mencionó antes, esto ocurre cuando se considera el índice de vegetación LAI SeLi y los cuatro apilamientos temporales que surgen en los Casos 1 a 4 al tener en cuenta entre 4 y 9 fechas disponibles. Como puede observarse, las Categorías 2 y 3 que corresponden a los rendimientos medios y altos, respectivamente, son los que mejor se clasifican. Los porcentajes de precisión son superiores al 80 %, llegando incluso al 90 % en el caso de la Categoría 3. Por otro lado, los rindes bajos (Categoría 1) son los que mayor imprecisión presentan al momento de ser detectados. Notoriamente, los porcentajes de precisión son inferiores al 67 %, llegando incluso a 0 % cuando el apilamiento de imágenes se construye con todas las fechas disponibles.

Finalmente, la Fig. 4 presenta la clasificación de los siete pivots del área de estudio en base al índice LAI SeLi para el Caso 2 que arrojó los mejores resultados (precisión del 86 % y $K = 0.7268$). Las áreas de rojo corresponden a la Categoría 1 (niveles bajos) de rinde, las de color amarillo a la Categoría 2 (niveles medios) y las de color verde a los rindes de Categoría 3 (niveles altos). El modelo se ejecutó considerando el perímetro que encierra a todos los lotes y por ello aparecen bordes y zonas espúreas en rojo, amarillo o verde que no deben tenerse en cuenta. Observando los pivots bajo estudio (el resto se pintó de color

Cuadro 2. Resultados generales de las clasificaciones efectuadas considerando los diferentes IV en cada uno de los seis casos de ensayo diagramados.

Indices	Caso	Fechas	Exactitud	Global Kappa
NDVI	1	4/8 al 22/11	83.64 %	0.6837
	2	14/8 al 22/11	83.25 %	0.6684
	3	3/10 al 22/11	80.47 %	0.6128
	4	18/10 al 22/11	77.94 %	0.5610
	5	23/10 al 22/11	72.61 %	0.4468
	6	17/11 al 22/11	63.83 %	0.2499
NDVI red-edge 5	1	4/8 al 22/11	84.14 %	0.6876
	2	14/8 al 22/11	82.54 %	0.6437
	3	3/10 al 22/11	81.95 %	0.6439
	4	18/10 al 22/11	77.58 %	0.5556
	5	23/10 al 22/11	71.79 %	0.4315
	6	17/11 al 22/11	59.55 %	0.1600
EVI	1	4/8 al 22/11	82.29 %	0.6379
	2	14/8 al 22/11	81.65 %	0.6249
	3	3/10 al 22/11	76.40 %	0.5147
	4	18/10 al 22/11	70.20 %	0.4004
	5	23/10 al 22/11	58.55 %	0.1977
	6	17/11 al 22/11	61.19 %	0.1948
LAI SeLi	1	4/8 al 22/11	84.92 %	0.6935
	2	14/8 al 22/11	86.00 %	0.7268
	3	3/10 al 22/11	83.61 %	0.6783
	4	18/10 al 22/11	80.00 %	0.6052
	5	23/10 al 22/11	73.61 %	0.4568
	6	17/11 al 22/11	60.59 %	0.1831
SR	1	4/8 al 22/11	85.80 %	0.7228
	2	14/8 al 22/11	84.22 %	0.6783
	3	3/10 al 22/11	82.68 %	0.6500
	4	18/10 al 22/11	79.05 %	0.5864
	5	23/10 al 22/11	73.73 %	0.4733
	6	17/11 al 22/11	60.66 %	0.1853
SD	1	4/8 al 22/11	84.12 %	0,6764
	2	14/8 al 22/11	84.81 %	0,7019
	3	3/10 al 22/11	81.68 %	0,6380
	4	18/10 al 22/11	77.83 %	0,5604
	5	23/10 al 22/11	71.74 %	0.4300
	6	17/11 al 22/11	61.43 %	0,1862

Cuadro 3. Precisiones obtenidas por categoría obtenidas al clasificar considerando el índice LAI SeLi y cuatro de los apilamientos estacionales considerados.

Caso	Fechas	Precisión (%)			Exactitud Global Kappa	
		Categoría 1	Categoría 2	Categoría 3	(%)	
1	4/8 al 22/11	0	82.44	90.10	84.92	0.6935
2	14/8 al 22/11	66.99	85.42	89.43	86.00	0.7268
3	3/10 al 22/11	65.34	83.25	86.37	83.61	0.6783
4	18/10 al 22/11	62.31	80.16	81.47	80.00	0.6052

gris difuso), puede apreciarse que las estimaciones que brinda la clasificación corroboran y reflejan con precisión la variabilidad que presentan los niveles de rinde de VV dentro de un lote y entre lotes.

5 Conclusiones

Los resultados obtenidos permiten conjeturar que la predicción del rendimiento de la leguminosa VV mediante técnicas de teledetección es viable. Este hecho se corrobora con el empleo de diferentes índices de vegetación obtenidos de imágenes satelitales Sentinel -2 durante distintos períodos del ciclo fenológico del cultivo. En tal sentido, el índice de vegetación LAI SeLi mostró los mejores índices de correlación (P 86 % y K 0.72). Otros índices de vegetación como el SR y el NDVI presentaron valores levemente inferiores convirtiéndolos también en buenas opciones para predecir los niveles de rendimiento antes de la cosecha de la VV.

Las clasificaciones realizadas también evidencian que la predicción del rendimiento puede llevarse a cabo considerando períodos de tiempo significativamente menores que el ciclo fenológico completo. Tomando como límite para la predicción unos 60 días antes de la cosecha, se pudo constatar que con cinco fechas del ciclo (3/10 al 22/11) o incluso cuatro fechas (18/10 al 22/11), se obtienen muy buenos valores de correlación para la mayoría de los índices de vegetación empleados. No ocurre lo mismo si se consideran tres (23/10 al 22/11) o dos (17/11 y 22/11) fechas. De este último caso se desprende especialmente que las fechas del mes de noviembre donde las plantas de VV se encuentran en plena floración no parecen ser determinantes para poder estimar el rendimiento de semillas que logrará el cultivo.

Del detalle de las precisiones logradas por categoría con el índice LAI SeLi se ilustra que los rindes medios y altos (Categorías 2 y 3) son los que mejor se detectan en este caso. Los porcentajes de precisión superan el 80 %. Sin embargo, con los rendimientos más bajos (Categoría 1) no se logra llegar al 67 %. Consecuentemente, esta última categoría es la que mayor error introduce en el análisis estadístico. Cabe mencionar que el conjunto de datos que corresponde a

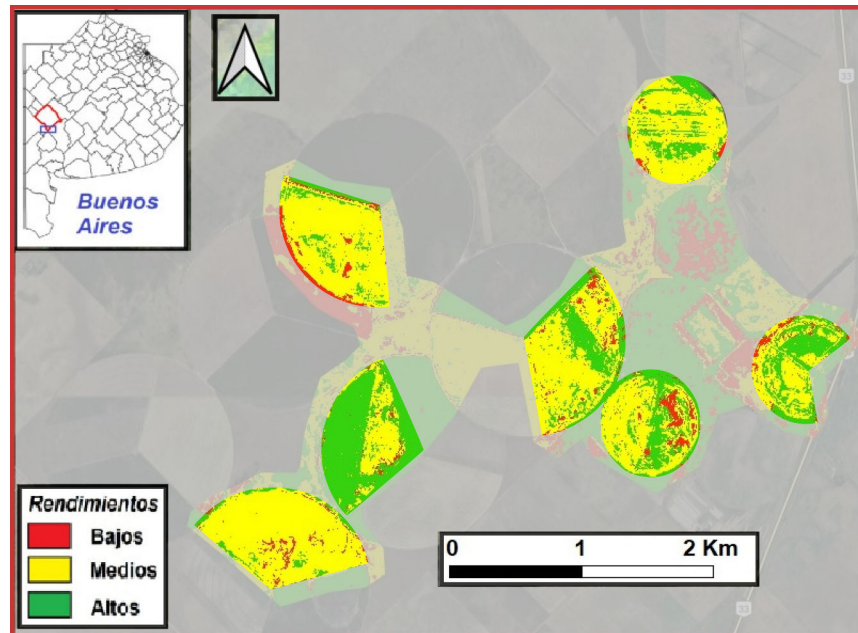


Figura 4. Clasificación del área de estudio considerando el índice LAI SeLI y las imágenes de las fechas del 14/08 al 22/11 (Caso 2).

la Categoría 1 es alrededor de un 50 a 100 % menor que los conjuntos de las Categorías 2 y 3 y se estima que la falta de representatividad es una de las posibles causas de error.

Finalmente, resulta relevante subrayar la potencialidad de los modelos obtenidos a partir del uso de imágenes del satélite Sentinel-2A para poder predecir los niveles de rendimiento de VV. Se seguirá trabajando a futuro para determinar de manera precisa las fechas que pueden ser más críticas en la predicción. Además, se intentará lograr valores de precisión más balanceados profundizando en el tratamiento estadístico y distribución por categorías de los datos. El análisis se completará evaluando la robustez de los modelos obtenidos a partir de la validación del desempeño con datos de otras campañas o de otros lotes diferentes y suficientemente alejados del área de estudio.

Agradecimientos

Los autores agradecen los datos de campo brindados por Marcelo Verdinelli y Facundo Bilbao y también el financiamiento recibido por parte de la UNS (PGI 24/K097), de CONICET (PIP 11220200102076) y de la CIC-PBA (RESO-2023-456, RIDEE No. 1).

Referencias

1. Hatfield J. L. y Prueger J. H.: Value of using different vegetative indices to quantify agricultural crop characteristics at different growth stages under varying management practices. *Remote Sensing* **2**(2), 562–578 (2010).
2. Segarra J., Buchailot M. L., Araus J. L. y Kefauver S. C.: Remote sensing for precision agriculture: Sentinel-2 improved features and applications. *Agronomy* **10**(5), (2020).
3. Patil P., Biradar C., Atassi L., Moussadek R., Kharrat M., Singh M., Andaloussi F. y Agrawal S. K.: Mapping and monitoring of food legumes and dryland cereal production systems. En: *Anales del Fourth International Conference on Agro-Geoinformatics (Agro-geoinformatics)*, 407–413, Estambul, Turquía (2015).
4. Lu S., Lu X., Zhao W., Liu Y., Wang Z., Omasa K.: Comparing vegetation indices for remote chlorophyll measurement of white poplar and Chinese elm leaves with different adaxial and abaxial surfaces. *Journal of Experimental Botany* **6**(18), 5625–5637 (2015).
5. Larrañaga Urien A.: Clasificación de cultivos a partir de observaciones de teledetección radar. Aplicación a zonas de secano y regadío de Navarra. Tesis Doctoral. En: Programa de Doctorado en Ingeniería y Arquitectura. Ediciones Universidad Pública de Navarra, España (2016).
6. Pasqualotto Vicente N., Delegido J., Van Wittenberghe S., Rinaldi M. y Moreno J.: Estimación de LAI verde de diversos cultivos mediante el nuevo índice SeLI para Sentinel-2. En: *Anales del XVIII Simposio Internacional de la Sociedad Latinoamericana de Percepción Remota y Sistemas de Información Espacial (SELPER)*, La Habana, Cuba (2018).
7. Marini F.: Discriminación de trigo y cebada empleando imágenes satelitales ópticas y radar. Estudio de caso: partido de Coronel Rosales (Argentina). *Investigaciones Geográficas UNAM*, 104 (2021).
8. Tenreiro T. R., García-Vila M., Gómez J. A., Jiménez-Berni J. A. y Fereres E.: Using NDVI for the assessment of canopy cover in agricultural crops within modelling research. *Computers and Electronics in Agriculture* **182**, art. 106038 (2021).
9. Gumma M. K., Tummala K., Dixit S., Collivignarelli F., Holecz F., Kolli R. N. y Whitbread A. M.: Crop type identification and spatial mapping using sentinel-2 satellite data with focus on field-level information. *Geocarto International*, **37**(7), 1833–1849 (2022).
10. Paredes Gómez V., del Blanco Medina V., Gutiérrez García A. y Nafría García D. A.: Seguimiento y evaluación de la capacidad de discriminación de cultivos herbáceos en regadío a partir de imágenes de satélite en el periodo 2016-2018. *Tierras de Castilla y León: Agricultura* **282**, 36–43 (2020).
11. Breiman L.: Random Forests. *Machine Learning* **45**, 5–32 (2001).
12. Regalado Galván J. A., D'Amico M. B. y Calandrini G. L.: Análisis temporal de índices de vegetación satelitales en *Vicia villosa* Roth. En: *Anales del XXVIII Congreso Argentino de Control Automático*, 97–102, Buenos Aires, Argentina (2023).
13. D'Amico M. B., Marini F., Calandrini G. L., Renzi J. P. y Chantre G. R.: Relación entre los índices de vegetación satelitales y el nivel de rinde de la leguminosa *Vicia villosa* Roth. En: *Anales de la XX Reunión de Trabajo en Procesamiento de la Información y Control (RPIC2023)*, Oberá, Argentina (2023).
14. Renzi J. P., Chantre G. R., Smýkal P., Presotto A. D., Zubiaga L., Garayalde A. F. y Cantamutto M. A.: Diversity of naturalized hairy vetch (*vicia villosa* Roth) populations in central Argentina as a source of potential adaptive traits for breeding. *Frontiers in Plant Science* **11**, art. 189 (2020).

15. Renzi J. P.: Caracterización Agroecológica de Poblaciones Naturales de *Vicia villosa* Roth (FABACEAE) de Argentina. Tesis de Doctorado en Agronomía, Universidad Nacional del Sur (2020).
16. Renzi J. P., Chantre G. R., González-Andújar J. L. y Cantamutto M. A.: Development and validation of a simulation model for hairy vetch (*Vicia villosa* Roth) self-regeneration under different crop rotations. *Field Crops Research* **235**, 79–86 (2019).
17. Marini M. F.: Discriminación de cultivos de distinto desarrollo utilizando imágenes satelitales MODIS. *GeoFocus (Informes y aplicaciones)*, **13**(1) 48-60 (2013)
18. Datos abiertos del Ministerio de Agricultura, Ganadería y Pesca de la Nación. <http://www.magyp.gob.ar/datosabiertos/>. Accedido en abril de 2024.
19. Página web de Google Earth Engine. <https://earthengine.google.com/>. Accedido en abril de 2024.
20. Rouse J. W., Haas R. H., Schell J. A., Deering D. W. y Harlan J. C.: Monitoring the vernal advancement of retrogradation of natural vegetation. Reporte final, NASA/OSFC, 371 (1974).
21. Justice C. O., Vermote E., Townshend J. R., Defries R., Roy D. P., Hall D. K. y Lucht W.: The Moderate Resolution Spectroradiometer (MODIS): Land remote sensing for global change research. *IEEE Transactions on Geosciences and Remote Sensing* **36**(4), 1228–1249 (1998).
22. Huang H., Roy D. P., Boschetti B., Zhang H. K., Yan L., Kumar S. S., Gomez-Dans J. y Li J.: Separability analysis of Sentinel-2A multi-spectral instrument (MSI) data for burned area discrimination. *Remote Sensing* **8**(10), 873 (2016).
23. Gilabert M. A., González-Piqueras J. y García-Haro F. J.: Acerca de los índices de vegetación. *Revista de Teledetección* **8** (1997).
24. Daughtry C. S. T., Gallo K. P., Goward S. N., Prince S. D. y Kustas W. P.: Spectral estimates of absorbed radiation and phytomass production in corn and soybean canopies. *Remote Sensing Environment* **39**(2), 141–152 (1992).
25. Ovando G., de la Casa A., Díaz G., Díaz P., Bressanini L. y Miranda C.: Desempeño de diferentes índices de vegetación de Sentinel-2A para estimar el rendimiento de soja en agricultura de precisión. *AgriScientia* (38), 1–12 (2021).
26. Jensen J. R.: *Introductory Digital Image Processing*. Prentice-Hall, Nueva Jersey, Estados Unidos (1986).
27. Bishop Y., Fienberg S. y Holland P.: *Discrete Multivariate Analysis. Theory and Practice*. MIT Press, Cambridge, Inglaterra (1975).