

Metodología para la estimación de productividad de un cultivo de papa mediante técnicas de percepción remota, caso de estudio municipio de Ventaquemada (Boyacá).

Elaborado por:
Camila Carmona Salazar
Jorge Andrés Vargas Pérez

Director:
Francisco Luis Hernández Torres

Trabajo de grado presentado como requisito para optar al título de ingenieros topográficos

Universidad del Valle
Escuela de Ingeniería Civil y Geomática
Programa de Ingeniería Topográfica
Cali, Colombia
2025

Nota de Aceptación

Jurado 1

Jurado 2

Dedicatoria

“El que vive sinceramente y encuentra penas verdaderas y desilusiones, que no se deja abatir por ellas, vale más que el que tiene siempre el viento en popa y que solo conocería una prosperidad relativa.”

— Vincent van Gogh

Camila Carmona Salazar

Dedico este trabajo, en primer lugar, a mi abuelo, quien, aunque no me acompañó físicamente, fue la inspiración y la valentía que me impulsaron a estudiar.

A mi abuela, por su apoyo incondicional y por acompañarme con cariño a lo largo de toda la carrera.

A mi gato, fiel compañero de largas noches en vela.

Y a mi mamá, mi hermana y mi tío, quienes, a pesar de la distancia, nunca dudaron de mí ni dejaron de brindarme su aliento y confianza.

Jorge Andrés Vargas Pérez

Dedico este trabajo a mi familia, que, a pesar de estar a 668 km de distancia, me brindaron un apoyo constante e inquebrantable. Su ejemplo de esfuerzo y perseverancia, junto con su amor y palabras de aliento, fueron la fuerza que me impulsó a culminar esta meta con gratitud y compromiso.

Agradecimiento

Agradecemos profundamente a nuestras familias por su apoyo incondicional, comprensión y paciencia durante toda nuestra formación. Su cariño y confianza fueron fundamentales para seguir adelante y culminar este proceso con éxito.

De manera especial, queremos expresar nuestro sincero agradecimiento al profesor Francisco Luis Hernández, quien nos acompañó en los últimos semestres de nuestra carrera y en el desarrollo de este trabajo de grado. Su orientación constante, sus consejos, paciencia y compromiso fueron clave para guiarnos en cada etapa, brindándonos no solo conocimiento académico, sino también motivación y confianza para afrontar los retos de esta investigación.

Finalmente, extendemos nuestro agradecimiento a la Universidad del Valle y a los docentes del programa de Ingeniería Topográfica, por su dedicación, compromiso y aporte a nuestra formación profesional y personal

Contenido

Resumen.....	9
1 Introducción	10
2 Planteamiento del problema.....	11
3 Objetivos	13
3.1 Objetivo general	13
3.2 Objetivos específicos.....	13
4 Justificación	14
5 Marco Teórico.....	15
5.1 Marco conceptual	15
5.2 Marco Teórico	19
5.3 Marco referencial.....	22
5.4 Materiales	24
6 Metodología	26
6.1 Esquema metodológico.....	26
6.2 Área de estudio	26
6.3 Obtención de datos	27
6.3.1 Datos de sensores remotos	27
6.3.2 Obtención de datos de campo.....	28
6.4 Procesamiento de información	29
6.4.1 Preprocesamiento y corrección de datos	29
6.4.2 Procesamiento y cálculo de índices.....	32
6.5 Establecimiento de modelos de regresión.....	34
6.5.1 Regresión lineal simple	34
6.5.2 Regresión lineal múltiple	34
6.5.3 Regresión Random Forest	34
6.6 Comprobación del modelo de regresión.....	35
7 Resultados	35
7.1 Adquisición de información de productividad	35
7.2 Modelos de regresión lineal simple para la estimación de productividad	36
7.2.1 NDVI.....	36
7.2.2 GNDVI.....	40
7.2.3 SAVI	43

7.2.4	EVI	46
7.2.5	Índices de Sentinel-1	49
7.3	Modelos de regresión lineal Múltiple para la estimación de la productividad	53
7.3.1	PlanetScope.	53
7.3.2	Sentinel-2	54
7.3.3	Sentinel-2 +PlanetScope	55
7.3.4	PlanetScope + Sentinel-1	56
7.3.5	Sentinel-2 + Sentinel-1.....	60
7.3.6	S2 relleno con PS y combinado con S1.....	64
7.3.7	Modelos de regresión múltiple con PlanetScope, Sentinel-2, y Sentinel-1.....	69
7.4	Modelo de regresión Random Forest.....	70
7.4.1	PlanetScope.	71
7.4.2	Sentinel-2.	72
7.4.3	Sentinel-2 + PlanetScope.	72
7.4.4	Ópticos + Radar.....	73
8	Discusión de resultados.....	75
9	Conclusiones.....	78
10	Bibliografía	80

Lista de figuras

Figura 1. Ciclo fenológico del cultivo de papa.	15
Figura 2. Diferentes señales emitidas por el objetivo en una imagen SAR.	16
Figura 3. Representación visual de una regresión Random Forest.	19
Figura 4. Esquema metodológico.....	26
Figura 5. Ubicación de Ventaquemada en Boyacá	27
Figura 6. Imágenes obtenidas para cada fase del cultivo.	28
Figura 7. Flujo de procesamiento en Sentinel-1.	29
Figura 8. Preprocesamiento información de Sentinel-1.....	29
Figura 9. Secuencia del algoritmo de detección de nubes/nieve de L2.....	30
Figura 10. Generación de máscara binaria de nubes.....	31
Figura 11. Comparación de reemplazo de información en imágenes con nubes.	32
Figura 12. Espectro de reflectancia de la hoja de papa. Características de absorción de las concentraciones químicas foliares[49].	33
Figura 13. Representación de índices de vegetación.	33
Figura 14. Localización de los 20 predios muestreados en el municipio de Ventaquemada.	35
Figura 15. Productividad real vs. estimada – NDVI (PlanetScope).....	37
Figura 16. Productividad real vs. estimada – NDVI (Sentinel-2).....	38
Figura 17. Productividad real vs. estimada – NDVI (Sentinel-2 + PlanetScope).....	39
Figura 18. Productividad real vs. estimada – GNDVI (PlanetScope).....	40
Figura 19. Productividad real vs. estimada – GNDVI (Sentinel-2).....	41
Figura 20. Productividad real vs. estimada – GNDVI (Sentinel-2 + PlanetScope).....	42
Figura 21. Productividad real vs. estimada – SAVI (PlanetScope).....	44
Figura 22. Productividad real vs. estimada – SAVI (Sentinel-2).....	45
Figura 23. Productividad real vs. estimada – SAVI (Sentinel-2 + PlanetScope).....	46
Figura 24. Productividad real vs. estimada – EVI (PlanetScope).....	47
Figura 25. Productividad real vs. estimada – EVI (Sentinel-2).....	48
Figura 26. Productividad real vs. estimada – VH (Sentinel-1).....	50
Figura 27. Productividad real vs. estimada – VV (Sentinel-1).....	51
Figura 28. Productividad real vs. estimada – VH/VV (Sentinel-1).....	52
Figura 29. Productividad real vs. estimada – RVI (Sentinel-1).....	53
Figura 30. Productividad real vs. estimada – Regresión lineal múltiple (PlanetScope).....	54
Figura 31. Productividad real vs. estimada – Regresión lineal múltiple (Sentinel-2).....	55
Figura 32. Productividad real vs. estimada – Regresión lineal múltiple (Sentinel-2 + PlanetScope).....	56
Figura 33. Productividad real vs. estimada – NDVI + RVI (PlanetScope + Sentinel-1).....	57
Figura 34. Productividad real vs. estimada – GNDVI + RVI (PlanetScope + Sentinel-1).....	58
Figura 35. Productividad real vs. estimada – SAVI + RVI (PlanetScope + Sentinel-1).....	59
Figura 36. Productividad real vs. estimada – EVI + RVI (PlanetScope + Sentinel-1).....	60
Figura 37. Productividad real vs. estimada – NDVI + RVI (Sentinel-2 + Sentinel-1).....	61
Figura 38. Productividad real vs. estimada – GNDVI + RVI (Sentinel-2 + Sentinel-1).....	62
Figura 39. Productividad real vs. estimada – SAVI + RVI (Sentinel-2 + Sentinel-1).....	63
Figura 40. Productividad real vs. estimada – EVI + RVI (Sentinel-2 + Sentinel-1).....	64
Figura 41. Productividad real vs. estimada – NDVI + RVI (Sentinel-2 + PlanetScope + Sentinel-1).....	65

Figura 42. Productividad real vs. estimada – GNDVI + RVI (Sentinel-2 + PlanetScope + Sentinel-1)	66
Figura 43. Productividad real vs. estimada – SAVI + RVI (Sentinel-2 + PlanetScope + Sentinel-1).	67
Figura 44. Productividad real vs. estimada – EVI + RVI (Sentinel-2 + PlanetScope + Sentinel-1).....	68
Figura 45. Productividad real vs. estimada – todos los índices (Sentinel-2 + PlanetScope + Sentinel-1)..	69
Figura 46. Productividad real vs. estimada – Random Forest (PlanetScope)	71
Figura 47. Productividad real vs. estimada – Random Forest (Sentinel-2).....	72
Figura 48. Productividad real vs. estimada – Random Forest (Sentinel-2 + PlanetScope).....	73
Figura 49. Productividad real vs. estimada – Random Forest + RVI (PlanetScope + Sentinel-1).....	74
Figura 50. Productividad real vs. estimada – Random Forest + RVI (Sentinel-2 + Sentinel-1).....	74
Figura 51. Productividad real vs. estimada – Random Forest + RVI (Sentinel-2 + PlanetScope + Sentinel-1)	75

Lista de tablas

Tabla 1. Bandas de PlanetScope.	24
Tabla 2. Bandas de Sentinel-2.....	25
Tabla 3. Categorías clasificación de escena SCL.....	31
Tabla 4. Variables registradas en los predios evaluados.....	36
Tabla 5. Ecuaciones de regresión lineal y R^2 por mes para el NDVI de PlanetScope.	37
Tabla 6. Ecuaciones de regresión lineal y R^2 por mes para el NDVI de Sentinel-2	38
Tabla 7. Ecuaciones de regresión lineal y R^2 por mes para el NDVI de Sentinel-2+PlanetScope	38
Tabla 8. Ecuaciones de regresión lineal y R^2 por mes para el GNDVI de PlanetScope.	40
Tabla 9. Ecuaciones de regresión lineal y R^2 por mes para el GNDVI de Sentinel-2.	41
Tabla 10. Ecuaciones de regresión lineal y R^2 por mes para el GNDVI de Sentinel-2 + PlanetScope.....	42
Tabla 11. Ecuaciones de regresión lineal y R^2 por mes para el SAVI de PlanetScope.....	43
Tabla 12. Ecuaciones de regresión lineal y R^2 por mes para el SAVI de Sentinel-2.	44
Tabla 13. Ecuaciones de regresión lineal y R^2 por mes para el SAVI de Sentinel-2+PlanetScope	45
Tabla 14. Ecuaciones de regresión lineal y R^2 por mes para el EVI de PlanetScope.	47
Tabla 15. Ecuaciones de regresión lineal y R^2 por mes para el EVI de Sentinel-2.....	47
Tabla 16. Ecuaciones de regresión lineal y R^2 por mes para el EVI de Sentinel-2+PlanetScope.....	48
Tabla 17. Ecuaciones de regresión lineal y R^2 por mes para VH.....	49
Tabla 18. Ecuaciones de regresión lineal y R^2 por mes para VV.....	50
Tabla 19. Ecuaciones de regresión lineal y R^2 por mes para VH/VV.....	51
Tabla 20. Ecuaciones de regresión lineal y R^2 por mes para RVI.....	52
Tabla 21. Regresión lineal múltiple de PlanetScope.....	54
Tabla 22. Regresión lineal múltiple de Sentinel-2.	55
Tabla 23. Regresión lineal múltiple de Sentinel-2 + PlanetScope.	55
Tabla 24. Regresión lineal múltiple NDVI (PlanetScope) y RVI (Sentinel-1).	57
Tabla 25. Regresión lineal múltiple GNDVI (PlanetScope) y RVI (Sentinel-1).	57
Tabla 26. Regresión lineal múltiple SAVI (PlanetScope) y RVI (Sentinel-1).	58
Tabla 27. Regresión lineal múltiple EVI (PlanetScope) y RVI (Sentinel-1).	59
Tabla 28. Regresión lineal múltiple NDVI (Sentinel-2) y RVI (Sentinel-1).	61
Tabla 29. Regresión lineal múltiple GNDVI (Sentinel-2) y RVI (Sentinel-1).	61
Tabla 30. Regresión lineal múltiple SAVI (Sentinel-2) y RVI (Sentinel-1).	62
Tabla 31. Regresión lineal múltiple EVI (Sentinel-2) y RVI (Sentinel-1).	63
Tabla 32. Regresión lineal múltiple NDVI (Sentinel-2+PlanetScope) y RVI (Sentinel-1).	65
Tabla 33. Regresión lineal múltiple GNDVI (Sentinel-2+PlanetScope) y RVI (Sentinel-1).	66
Tabla 34. Regresión lineal múltiple SAVI (Sentinel-2+PlanetScope) y RVI (Sentinel-1).	66
Tabla 35. Regresión lineal múltiple EVI (Sentinel-2+PlanetScope) y RVI (Sentinel-1).	67
Tabla 36. Modelo Regresión Random Forest de PlanetScope.....	71
Tabla 37. Modelo Regresión Random Forest de Sentinel-2	72
Tabla 38. Modelo de regresión Random Forest Sentinel-2 + PlanetScope.....	73

Resumen

La estimación de la productividad agrícola mediante imágenes satelitales constituye una herramienta clave para la agricultura de precisión, pues permite anticipar rendimientos y apoyar la toma de decisiones en territorios con alta variabilidad agroecológica. En el caso del cultivo de papa, este proceso presenta retos adicionales debido a su heterogeneidad intrapredial y a la frecuente nubosidad que afecta a los sensores ópticos. Ante este escenario, el presente trabajo se planteó como objetivo desarrollar una metodología para estimar la productividad de papa en el municipio de Ventaquemada (Boyacá, Colombia) a partir de la integración de sensores remotos y datos de campo. Para ello se utilizaron imágenes ópticas de Sentinel-2 y PlanetScope, complementadas con información radar de Sentinel-1 y validadas con mediciones de 20 predios productivos. A partir de estas fuentes se calcularon índices de vegetación (NDVI, GNDVI, SAVI y EVI) y parámetros SAR (VH, VV, VH/VV y RVI), que se emplearon como predictores en modelos de regresión lineal simple, múltiple y Random Forest. Los resultados mostraron que los modelos lineales simples tuvieron un desempeño limitado (R^2 máximo de 0,369), mientras que la regresión múltiple alcanzó un mejor ajuste en el mes 4 con $R^2=0,520$ y $RMSE=1,54$ t/ha. No obstante, los modelos Random Forest presentaron mayor estabilidad en los errores, con un escenario óptimo de $R^2=0,362$, $RMSE=0,663$ t/ha y $MAE=0,420$ t/ha. Estos hallazgos evidencian que la integración multifuente y el uso de enfoques de aprendizaje automático fortalecen la consistencia en la predicción de rendimientos, constituyéndose en una alternativa aplicable a la gestión agrícola en regiones con condiciones variables como Boyacá.

Palabras clave: Papa, Índices de vegetación, Teledetección, Productividad, Regresión.

1 Introducción

La agricultura en Colombia se encuentra favorecida por la ubicación geográfica del país ya que es posible encontrar una gran variedad de factores como lo es el clima, la cultura, cuencas hidrográficas, recursos naturales, flora y fauna. Según la Unidad de Planificación Rural Agropecuaria (UPRA) en Colombia se encuentran 26,5 millones de hectáreas con vocación agro, en las cuales se cultiva arroz, cacao, caucho, maíz y por último la papa, particularmente el sector agropecuario en el departamento de Boyacá cuenta con 15 cadenas productivas dentro de las cuales se encuentra el cultivo de papa [1].

La papa es uno de los alimentos de mayor consumo en el mundo, junto con el arroz y el trigo, es consumida por millones de personas alrededor del mundo. En Colombia la papa ocupa el segundo lugar de importancia alimenticia y Boyacá es el segundo mayor productor de papa en el país con un 27% de la producción nacional. El rendimiento de esta hortaliza depende de factores agrícolas (nitrógeno, fósforo, potasio, densidad de la siembra, propiedades del suelo, riego, lugar de producción, etc.) y de las condiciones climáticas de la zona considerando que estas últimas no son constantes durante el tiempo de crecimiento del cultivo.

Diversas técnicas para el monitoreo de cultivos son utilizadas, generalmente, mediante inspecciones y muestreos manuales que en algunos casos genera afectaciones en la planta y sobrecostos económicos en casos de cultivos con una extensión considerablemente grande. Por lo que utilizando percepción remota es posible proporcionar información crítica con la que se pueden generar modelos de los cultivos para mejorar las predicciones de rendimiento. Entre las distintas técnicas de percepción remota se encuentra el manejo de imágenes satelitales que facilita el mapeo eficiente en grandes áreas y que se emplea en agricultura de precisión. Debido a la respuesta espectral de la papa, que varía con base en las variedades y fases del ciclo de vida [2] es posible la implementación de este tipo de monitoreo.

La combinación de geo-información y modelado de cultivos representa una opción para fortalecer la competitividad de la cadena de papa debido a que existe una demanda de información temporal relacionada con la salud, la productividad, la biomasa y el potencial de rendimiento de los cultivos de papa analizando variables como humedad, radiación, pendiente, etc. La utilización de percepción remota permite encontrar de manera no invasiva los valores óptimos para las variables con el fin de brindar información útil para la toma de decisiones por parte de agricultores y entes gubernamentales. En el presente estudio se empleó percepción remota para calcular la productividad de los cultivos de papa mediante el análisis de la información espectral disponible en imágenes de sensores remotos que puede ser utilizada para estimar el rendimiento del cultivo.

2 Planteamiento del problema

La papa (*Solanum tuberosum*) es una de las hortalizas mayormente extendidas en el mundo, originaria de los Andes suramericanos, se produce en temperaturas de entre 15 y 25°C, precisan niveles óptimos de humedad relativa entre el 60 y 80%, suelos bien drenados y de buena estructura que permitan su crecimiento, además requiere de una precipitación pluvial óptima de 600 mm a 800 mm al año [3]. Es uno de los principales alimentos que ha permitido mitigar la hambruna después de distintos conflictos bélicos y económicos. Ocupando el tercer lugar de importancia como producto alimenticio según la demanda de consumo (FAO, 2022), la papa puede producir entre 2 y 4 veces más alimento que el arroz o el trigo y presenta numerosas propiedades nutritivas. Aproximadamente el 85% de la planta es comestible, a diferencia de los cereales, donde se consume cerca del 50% de la planta solamente (Agronegocios, 2018).

En Colombia la papa ocupa el segundo lugar de importancia alimenticia, detrás del arroz, es sembrada en 9 departamentos y al menos 300 municipios. El 90% de papa sembrada en el país se concentra en 4 departamentos: Cundinamarca (37%), Boyacá (27%), Nariño (20%) y Antioquia (6%). La cadena de papa en Colombia genera anualmente alrededor de 264 mil empleos totales, de los cuales, 75 mil son empleos directos y 189 mil son empleos indirectos [4].

El departamento de Boyacá es conocido por su vocación agrícola, destacándose la producción de tubérculos y plátanos. De este modo, se destacan las actividades agropecuarias en la participación del Producto Interno Bruto (PIB) departamental con una contribución del 5.9% en 2020. Actualmente, la papa es uno de los productos de mayor producción en el departamento con 690.714 toneladas en un área de cultivo de 32.392 Ha y un rendimiento de 21,32% en el año 2021 [5], Los municipios más representativos en el abastecimiento se encuentran fuertemente concentrados en torno a los municipios de Tunja, Ventaquemada, Toca, etc. El municipio de Ventaquemada está ubicado en la provincia Central del departamento, y es uno de los de mayor producción con cerca de 5.479 toneladas en mayo del 2021 [6]. Con el propósito de calcular la productividad del cultivo, se crean políticas gubernamentales que permiten considerar las exigencias edafoclimatológicas y su impacto sobre el desarrollo del cultivo para tomar medidas que permitan minimizar los impactos negativos que se puedan presentar.

El cultivo requiere un suelo suelto, ya que los suelos arcillosos causan deformación en los tubérculos, así mismo estos suelos no deben presentar condiciones químicas indeseables con pH ácidos (mayores a 5,5) o que contengan elementos tóxicos para las plantas, y no se recomiendan pendientes superiores al 40%. El cultivo tiene un buen desempeño en temperaturas medias de entre 12 y 14 °C y además de un requerimiento hídrico de 600 a 800 mm al año, es sensible a la escasez y al exceso de agua [7]. Es un cultivo muy susceptible a heladas, entre el mes de diciembre del 2021 y primeras semanas de enero del 2022, se registraron pérdidas de entre 5 y 6 mil Ha de papa en 25 municipios del sur de Nariño y que considerando el precio para la papa a inicios del año 2022 corresponde a pérdidas cercanas a los 230.860 millones de pesos (Semana, 2022).

Para el 2022 el departamento de Boyacá registró un descenso en el abastecimiento de papa de un 8.8% con respecto a diciembre del año previo y un -3.7% comparado con enero de 2021, con lo anterior los precios tuvieron fuertes impactos para el 2021 respecto al año 2020 con una variación promedio anual del 67% y para enero del 2022 se presentó un aumento del 34% respecto al 2021, sin embargo, al comparar los precios con el mismo mes de referencia en el año 2020, se observa una variación de 277% pasando de \$498 a \$1.876 [5].

En la actualidad, los hábitos de los agricultores fomentan una disminución de la productividad, debido a que siembran papa en todas las tierras que poseen, la estrategia de cultivo todavía se basa en las preferencias y conocimientos empíricos de los agricultores, por lo que plantan en su tierra sin ver la idoneidad de esta [8]. Las causas de las pérdidas en la explotación agrícola son numerosas y específicas de cada contexto. A menudo, influyen en ellas factores de precosecha, como las condiciones meteorológicas, la calidad de las semillas, la variedad, las prácticas de cultivo, las infestaciones por plagas y las infecciones por enfermedades [9]. Existen diversos métodos para monitorear y evaluar el crecimiento y comportamiento de los cultivos con lo que es posible estimar su comportamiento y posible rendimiento; una forma precisa de medir esto, es realizando muestreo manual, sin embargo esta práctica es en ocasiones costosa y destructiva para el cultivo ya que requiere personal capacitado que realice los muestreos periódicos y el respectivo procesamiento, que a su vez tiene la limitante de la extensión, porque no es factible en grandes extensiones de terreno debido a la cantidad de puntos de muestreo que precisa [10].

Finalmente, existe una demanda de información relacionada con la productividad y el potencial de rendimiento de los cultivos de papa [11]. Por lo que surge el problema ¿Cómo obtener datos útiles que permitan estimar la productividad del cultivo de papa?

3 Objetivos

3.1 Objetivo general

Formular una metodología que permita estimar la productividad de un cultivo de papa mediante sensores remotos en el municipio de Ventaquemada.

3.2 Objetivos específicos

- Caracterizar mediante sensores remotos las variables que influyen en la productividad del cultivo de papa.
- Establecer la relación entre las variables estimadas por medio de sensores remotos y la productividad del cultivo de papa.
- Comprobar la productividad estimada mediante sensores remotos con datos de campo.

4 Justificación

La estimación del crecimiento de cultivos y la previsión del rendimiento, permite tomar medidas en caso de procesos que causen efectos negativos que puedan influir en la seguridad alimentaria y del rendimiento económico del cultivo considerando el impacto social y económico de este sector [12]. La utilización de técnicas de teledetección mejora la potencialidad del seguimiento y previsión de rendimiento de cultivos al proporcionar información espacial de características y salud de los cultivos en todas sus etapas (inicial, desarrollo, reproducción y madurez), mediante el cálculo de índices de vegetación (VI) que registran la calidad de la actividad fotosintética de las plantas, además de que agiliza el proceso de obtención de la información lo que evita el proceso tedioso y tradicional de visitas de campo, considerando que muchas veces contiene errores que se basan en criterios del investigador y no se obtiene datos detallados del cultivo, adicionalmente, este proceso generalmente es lento y si bien se adquiere la información, en ocasiones esta llega demasiado tarde y no le da tiempo a las entidades a tomar decisiones o medidas para contrarrestar posibles efectos negativos [13].

La radiación solar es uno de los factores ambientales más importantes para los cultivos debido a que contribuye en su crecimiento y desarrollo óptimos, parámetros como cantidad, calidad y duración de luz son relevantes durante todas las fases de la cadena productora. La cantidad de luz en las plantas como fuente de energía necesaria en la que generalmente fotosintetizan esta entre 400 a 700 nm, en un intervalo conocido como radiación fotosintética activa, por lo que la duración de luz a la que se ve expuesta la planta impacta en el florecimiento [13].

Numerosos estudios han vinculado la respuesta espectral de la vegetación con parámetros biofísicos como el índice de área foliar y el contenido de clorofila en las hojas y en el dosel. Además, factores climáticos como la temperatura y la precipitación presentan variaciones significativas, lo que afecta las curvas de reflectancia a lo largo de las temporadas de cultivo. En condiciones de sequía, por ejemplo, se encuentran concentraciones más altas de clorofila en las hojas de papa, y el contenido de agua en estas puede estimarse mediante índices espectrales. Así, la reflectancia espectral proporciona información relevante para comprender la variabilidad interanual y espacial en los cultivos de papa.[2].

Mantener la humedad de suelo adecuada durante el desarrollo del cultivo representa un factor de gran importancia para lograr productividades estimadas, una planta con deficiencia de agua es más susceptible a enfermedades y, por consiguiente, una producción baja de frutos, así mismo, el exceso de agua también afecta el cultivo ya que puede generar daños en el fruto [14]. En climas húmedos y fríos, se presentan abundantes precipitaciones, lo que genera que el riego se realice de manera esporádica, por el contrario, en climas más cálidos, se debe considerar el riego como practica constante debido a la poca presencia de precipitaciones. La evaporación real en un cultivo depende de la radiación solar y temperatura. Además de ser un componente del balance hídrico, es también un requisito para la constitución de una correcta estrategia de riego [14].

Este trabajo se planteó, desde el campo de la percepción remota, mediante el uso de imágenes de sensores remotos para el estudio del comportamiento del cultivo de papa para estimar su productividad, utilizando herramientas que permitan analizar las condiciones y características de crecimiento de la planta.

5 Marco Teórico

5.1 Marco conceptual

La papa (Solanum tuberosum): es un cultivo de gran importancia agrícola y alimentaria, originario de los Andes sudamericanos. Su valor nutricional radica en su bajo contenido de grasas y su riqueza en minerales esenciales como el potasio. En Colombia, es uno de los principales productos del sector agrícola, con múltiples variedades adaptadas a diferentes condiciones agroecológicas. Entre estas, se destaca la papa superior, una variedad de alto rendimiento desarrollada por selección empírica a partir de la papa pastusa. Esta variedad ha sido ampliamente adoptada por los productores debido a su capacidad de adaptación a climas fríos y a su buen comportamiento agronómico. Se cultiva principalmente en el altiplano Cundiboyacense y en el departamento de Nariño, regiones que presentan condiciones óptimas para su desarrollo [15]. En la Figura 1 se presentan gráfica de las etapas fenológicas del cultivo de papa a lo largo de un ciclo aproximado de 120 días (4 meses).

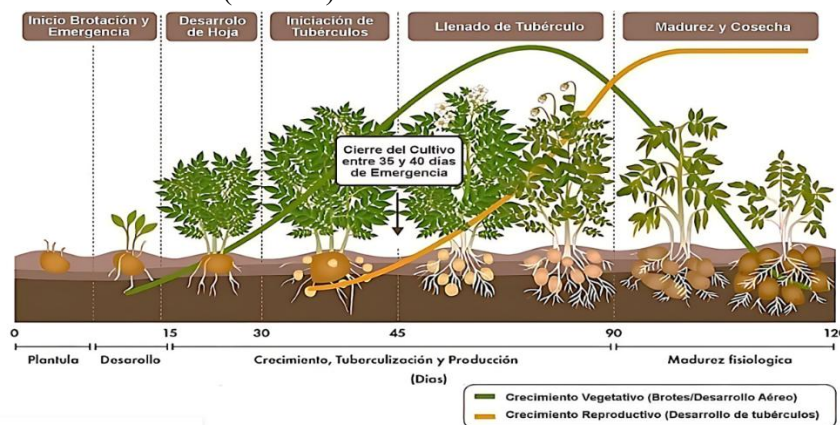


Figura 1. Ciclo fenológico del cultivo de papa.

Teledetección: También llamada observación de la tierra, se define como el arte, la ciencia y la tecnología a través de los cuales se identifican, miden y analizan las características de objetos ya sea sobre e incluso debajo de la superficie terrestre sin que exista contacto directo entre los sensores y los eventos u objetivos observados lo que permite la obtención de información de objetos y fenómenos mediante la detección y registro de energía reflejada o emitida a la cual se le realiza una serie de procesamiento y análisis para obtener productos de utilidad a diferentes disciplinas [12].

PlanetScope: La constelación PlanetScope, compuesta por más de 130 satélites tipo CubeSat, permite obtener imágenes diarias de la superficie terrestre con alta resolución espacial. Desde 2020, la generación SuperDove produce imágenes multiespectrales en ocho bandas, consolidándose como una fuente clave para el monitoreo agrícola y ambiental [16].

Sentinel-2: Es una misión de la Agencia Espacial Europea (ESA) compuesta por dos satélites gemelos (Sentinel-2A y Sentinel-2B), cuyo objetivo es proveer imágenes multiespectrales de alta resolución al programa Copernicus, contribuyendo al monitoreo ambiental y a diversas aplicaciones en agricultura, gestión territorial y recursos naturales. Ambos satélites orbitan de manera heliosíncrona a 786 km de altitud en un mismo plano orbital con un desfase de 180°, lo que garantiza una mayor frecuencia de revisita. Cada plataforma está equipada con el instrumento MSI (MultiSpectral Instrument), basado en la técnica push-broom, que proporciona una cobertura continua de la superficie terrestre con 13 bandas espectrales [17][18].

Sentinel-1: es una misión perteneciente al programa Copernicus de la Agencia Espacial Europea (ESA), está diseñada para obtener imágenes de alta resolución de la superficie terrestre y oceánica mediante un radar de apertura sintética (SAR) de banda C. Sus satélites, en órbita heliosíncrona polar, permiten adquisiciones continuas tanto de día como de noche y bajo cualquier condición climática. Los productos GRD (Ground Range Detected) corresponden a datos SAR proyectados al rango terrestre mediante el elipsoide WGS84. En el modo IW (Interferometric Wide Swath), todas las ráfagas de las subfranjas se combinan en una única imagen detectada, disponible para cada canal de polarización C.

Las polarizaciones más utilizadas son:

- VV: Polarización vertical en transmisión, polarización vertical en la recepción.
- VH: Polarización vertical en transmisión, polarización horizontal en la recepción.
- VH/VV: Relación de polarizaciones.

Speckle: El radar de apertura sintética emite un pulso que impacta en una zona del objetivo, esta zona está por lo general compuesta por objetos con distintas formas (Ver Figura 2). Cuando estos objetos son impactados, cada uno refleja una señal diferente en amplitud y fase, que puede interferir de forma constructiva o destructiva. El “Speckle” o moteado es el resultado de dicha interferencia. La señal que recibe el SAR contiene todas las diferentes señales acumuladas. Por lo anterior los píxeles pueden tener intensidades diferentes y pueden provocar que la imagen se vea granulada. La corrección de este ruido consiste en tomar la intensidad media de un área homogénea. Basándose en el supuesto de que la media de un área homogénea representa la intensidad real [19].

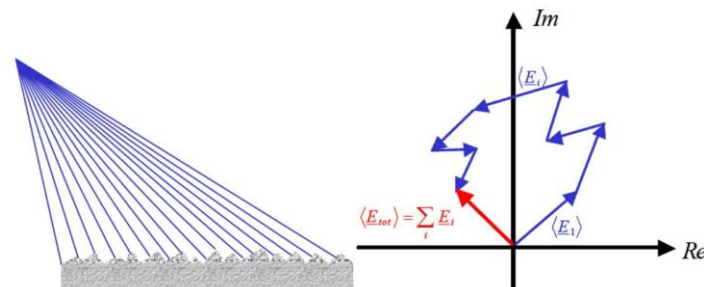


Figura 2. Diferentes señales emitidas por el objetivo en una imagen SAR.

Sentinel Application Platform (SNAP): es una plataforma de código abierto desarrollada por la ESA para el procesamiento y análisis de datos de observación terrestre provenientes de la familia de satélites Sentinel. Su arquitectura modular y extensible permite gestionar datos multifuente de forma eficiente, tanto mediante interfaz gráfica como a través de procesamiento automatizado, soportando grandes volúmenes con funcionalidades avanzadas como encadenamiento de flujos de trabajo (Graph Processing Framework), visualización interactiva y operaciones geoespaciales integradas [20].

QGIS: es un sistema de información geográfica (SIG) de código abierto, creado en 2002 como alternativa gratuita al software propietario. Es multiplataforma y cuenta con aplicaciones móviles. Actualmente se utiliza para la visualización, análisis y generación de productos cartográficos, con soporte para múltiples formatos y extensiones mediante complementos. Su distribución bajo la Licencia Pública General GNU garantiza acceso libre y modificable para todos los usuarios [21].

ArcMAP: es un componente central de ArcGIS Desktop que ha sido, durante años, una de las aplicaciones más utilizadas en el ámbito de los sistemas de información geográfica. Permite la visualización y

exploración de datos en 2D, la asignación de simbología, la creación de mapas para impresión o publicación y la edición de conjuntos de datos. Además, incorpora herramientas de geoprocésamiento que facilitan el análisis espacial, consolidándose como una herramienta de referencia en la representación y gestión de información geográfica [22].

Python: Lenguaje de programación de alto nivel, interpretado y multipropósito, desarrollado en la década de 1990 por Guido Van Rossum es un lenguaje muy compacto y de fácil aprendizaje comparado con otros lenguajes de programación, lo que permite la realización de diferentes tareas en diferentes áreas. En especial ha cobrado gran relevancia en estudios agrícolas mediante información de teledetección ya que, gracias a librerías como Rasterio, Geopandas y GDAL permite realizar el procesamiento de imágenes de satélite y extracción de información de las mismas[23].

Longitud de onda (λ): Se define como la distancia entre las crestas sucesivas de una onda.

Corrección radiométrica: Se basa en la corrección de las variaciones que se presentan en las intensidades de los niveles digitales (ND) de cada píxel de la imagen, estas variaciones pueden presentar daños debidos a errores en la calibración del sensor [24].

Fenología: La fenología de la vegetación es el estudio de las etapas que se presentan en el ciclo de vida de las plantas (brotación, floración) y es considerado además como un indicador importante del cambio ambiental y climático. Las variaciones en la fenología de la vegetación se relacionan con ciclos de carbono, energía, nutrientes y agua dentro de cada ecosistema [13].

Índices de Vegetación (IV): La predicción y estimación del rendimiento de cultivos está relacionado con la capacidad de identificar especies de cultivos y variables agronómicas como la madurez, la densidad, el vigor, etc. Que pueden utilizarse como indicadores de rendimiento. La teledetección proporciona información sobre el estado de salud de la vegetación. La reflectancia espectral de un campo de cultivo también depende de la fenología, tipo de etapa y salud del cultivo. Los índices de vegetación son combinaciones matemáticas de reflectancia en dos o más bandas espectrales, están diseñados para resaltar las propiedades de la vegetación [25]. Existen diferentes tipos de índices de vegetación (IV) derivados de la reflectancia espectral [26] en el caso del presente proyecto se consideró revisar los siguientes IV:

Índice de vegetación de diferencia normalizada (NDVI): Es un índice de vegetación ampliamente utilizado desde la década de 1970 en teledetección para el monitoreo de la biomasa y el rendimiento agrícola. Se calcula a partir de la relación entre la reflectancia del infrarrojo cercano y la del rojo, obteniendo valores entre -1 y 1, donde los más altos indican mayor actividad fotosintética. Su fundamento radica en la fuerte absorción de la luz roja por la clorofila y la alta reflectancia del infrarrojo cercano causada por la estructura foliar [27]. Al derivarse de imágenes satelitales, la precisión del NDVI depende de la resolución espacial, ya que la presencia de elementos no deseados (rocas, árboles u otros cultivos) puede introducir ruido y disminuir la calidad del índice [28].

Índice de vegetación de diferencia normalizada verde (GNDVI): es una variante del NDVI en la que se sustituye la banda roja por la verde, lo que aumenta la sensibilidad al contenido de clorofila y reduce la influencia del suelo en el índice [29].

Índice de vegetación ajustado al suelo (SAVI): Este índice se utiliza para minimizar las influencias de brillo del suelo de los índices de vegetación que utilizan longitudes de onda rojas e infrarrojas cercanas (NIR) [30].

Índice de vegetación mejorado (EVI): índice de vegetación que se desarrolló para optimizar la señal de vegetación con una sensibilidad mejorada en regiones de alta biomasa y un mejor monitoreo de la vegetación a través de un desacoplamiento de la señal de fondo de dosel y una reducción de las influencias atmosféricas.

Índice de vegetación de radar (RVI): Se trata de una serie de modelos matemáticos que se desarrollaron a partir de la combinación y procesamiento de información de retrodispersión del radar proveniente de áreas de vegetación, así como de parámetros obtenidos mediante técnicas de descomposición polarimétrica. Estos modelos permiten la caracterización tanto de aspectos fisiológicos como estructurales de la vegetación. La intensidad de la señal retro dispersada en la dirección e incidencia de la onda de radar, se expresa por lo general a través del coeficiente de retrodispersión (σ^0), que integra de forma detallada información sobre el comportamiento dispersivo de la vegetación [31].

Análisis de regresión lineal simple: Modelo estadístico más sencillo de la regresión que involucra solo dos variables (una variable independiente X y una variable dependiente Y) que están relacionadas linealmente entre sí. Los objetivos del análisis de regresión son encontrar una ecuación que conecte ambas variables para demostrar la relación entre X e Y para predecir la variable dependiente a partir de un valor específico de variable independiente [32].

Análisis de regresión lineal múltiple: Un análisis de regresión es una técnica estadística para estimar la relación entre variables que tienen relación. Un análisis de regresión lineal estudia la relación entre una variable dependiente y una variable independiente y formula la relación lineal entre estas. Un análisis de regresión múltiple considera la relación que existe entre una variable dependiente y más de una variable independiente que eventualmente se utilizan para calcular la variable dependiente [33].

Regresión Random Forest: Se trata de una técnica de aprendizaje en conjunto que integra las predicciones de múltiples árboles de decisión para construir modelos de clasificación o regresión con mayor precisión y consistencia. En el caso de la regresión, el algoritmo estima valores numéricos a través del promedio de las salidas generadas por varios árboles, lo que incrementa la capacidad predictiva y reduce la variabilidad de los resultados.

En la Figura 3 se presenta un esquema que ilustra el funcionamiento general de una regresión Random Forest. Cada árbol actúa como un predictor individual y sus resultados se combinan para obtener una predicción promedio. Este modelo es especialmente útil porque permite capturar tanto relaciones lineales como no lineales entre las variables [34].

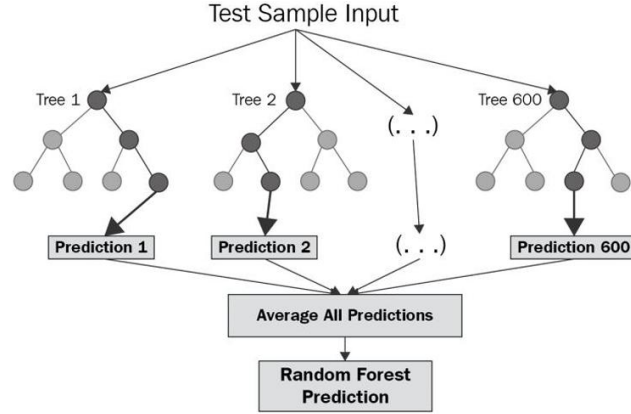


Figura 3. Representación visual de una regresión Random Forest.

Raíz Cuadrada del Error Medio Cuadrático (RMSE): Simboliza la desviación cuadrática media entre los valores simulados y los observados. Es decir que cuantifica el error existente ente dos conjuntos de datos. La precisión se caracteriza por un (RMSE) más bajo [35].

Coeficiente de correlación: Es un estadístico que proporciona información sobre la relación lineal existente entre dos variables. Básicamente se refiere dos características de la relación línea: dirección o sentido y la cercanía o fuerza. Su valor se encuentra en el rango $0 < x \leq 1$ y su interpretación varía dependiendo del concepto, sin embargo, numéricamente, mientras el valor sea más cercano a 1 significa una mayor relación entre las dos variables comparadas [36].

Error absoluto medio (MAE): es una métrica que se utiliza para evaluar la precisión de un modelo de regresión. En general, representa el promedio de los errores absolutos entre los valores predichos por un determinado modelo y los valores reales [37].

5.2 Marco Teórico

Como ya se mencionó la calibración radiométrica es un conjunto de técnicas de corrección relacionadas con la sensibilidad del sensor, ángulo solar, dispersión atmosférica y absorción. Para esto es necesario transformar los valores digitales en radiancia espectral mediante la ecuación (1) [38].

$$L_{\gamma} = Gain_{\gamma} * DN_{\gamma} + Bias_{\gamma} \quad (1)$$

Donde:

L_{γ} : Radiación espectral en la apertura del sensor, Ganancia γ

$Gain_{\gamma}$ y $Bias_{\gamma}$: Banda – Factores de re escalado específicos que se dan en el archivo de cabecera del satélite.

Los datos de radiancia se convierten en reflectancia superficial mediante la ecuación (2) [38].

$$\rho = \frac{\pi * L_{\gamma} * d^2}{ESUN_{\gamma} * Cos(\theta_s)} \quad (2)$$

Donde:

π : Constante matemática

ρ : Reflectancia planetaria sin unidades
 d : Distancia tierra – sol en unidades astronómicas
 $ESUN_{\gamma}$: Irradiancia exoatmosférica solar
 θ_s : Angulo cenital solar en grados

Para realizar el cálculo del NDVI se utiliza la ecuación (3) [27] [38][26][39].

$$NDVI = \frac{NIR - RED}{NIR + RED} \quad (3)$$

Donde:

NIR : Reflectancia de la longitud de onda en el espectro del infrarrojo cercano.

RED : Reflectancia de la longitud de onda en el rango rojo del espectro rojo.

Para realizar el cálculo del SAVI se utiliza la ecuación (4) [27][30][39].

$$SAVI = \frac{(NIR - RED)}{(NIR + RED + L)} * (1 + L) \quad (4)$$

Donde:

NIR : Reflectancia de la longitud de onda en el espectro del infrarrojo cercano.

RED : Reflectancia de la longitud de onda en el rango rojo del espectro rojo.

L : Factor de corrección de fondo del suelo normalmente establecido en 0,5.

Índice de vegetación de diferencia normalizada verde (GNDVI) se define de la siguiente manera [29]

$$GNDVI = \frac{NIR - GREEN}{NIR + GREEN} \quad (5)$$

Donde:

NIR : Reflectancia de la longitud de onda en el espectro del infrarrojo cercano.

$GREEN$: Reflectancia de la longitud de onda en el rango verde del espectro verde.

Cómo ya se mencionó anteriormente, el EVI es más sensible a las variaciones estructurales de dosel y para su cálculo se utiliza la ecuación (6) [13][39][40].

$$EVI = G * \frac{NIR - RED}{NIR + C1 * RED - C2 * BLUE + L} \quad (6)$$

Donde:

NIR : Reflectancia de la longitud de onda en el espectro del infrarrojo cercano.

RED : Reflectancia de la longitud de onda en el rango rojo del espectro rojo.

L : Factor de corrección de fondo del suelo normalmente establecido en 0,5.

$BLUE$: Reflectancia de la longitud de onda en el rango azul del espectro azul.

$C1$ y $C2$: Coeficientes que se establece en -6 y 7,5 respectivamente.

G : Factor de ganancia (establecido en 2,5).

La ecuación (7) corresponde al cálculo del RVI para estimar el estado de salud de la vegetación utilizando información de radar de la colección Sentinel-1 GRD [31].

$$RVI = \frac{4\sigma_{VH}^0}{\sigma_{VV}^0 + \sigma_{VH}^0} \quad (7)$$

Donde:

σ_{VH}^0 , σ_{VV}^0 y σ_{VH}^0 : coeficientes de retrodispersión

La ecuación (8) representa el índice empírico de la relación entre las polarizaciones VH/VV [37].

$$Rel_{\frac{VH}{VV}} = \frac{VH}{VV} \quad (8)$$

Donde:

VH: las señales de radar se transmiten verticalmente y se reciben horizontalmente.

VV: las señales de radar se transmiten y se reciben verticalmente.

Cuando el análisis de regresión usa más de una variable independiente es llamado múltiple y se calcula mediante la ecuación (10)[41].

$$y = \beta_0 + \beta_1 x_1 + \varepsilon \quad (9)$$

Donde:

y: Variable dependiente

x_i : Variable independiente

β_i : Parámetro

ε : Error aleatorio

Cuando el análisis de regresión usa más de una variable independiente es llamado múltiple y se calcula mediante la ecuación (10) [33].

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_n x_n + \varepsilon \quad (10)$$

Donde:

y: Variable dependiente

x_i : Variable independiente

β_i : Parámetro

ε : Error aleatorio

Para el cálculo del RMSE se utiliza la ecuación (11) donde se toma el cuadrado de la diferencia de cada uno de los datos comparados y realiza una suma, para finalmente dividir dicha suma entre el número de observaciones [35][37].

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (P_i - O_i)^2}{n}} \quad (11)$$

Donde:

P_i : Valores previstos

O_i : Valores observados

n : tamaño de la muestra (número de observaciones)

El cálculo del coeficiente de correlación se utiliza la ecuación (12) donde se relaciona la covarianza entre los valores de las variables (x, y) y el producto de las desviaciones típicas de cada variable [36].

$$\rho_{xy} = \frac{Cov_{xy}}{\sigma_x \sigma_y} \quad (12)$$

Donde:

Cov_{xy} : Covarianza entre los valores x e y

σ_x : Desviación típica de x

σ_y : Desviación típica de y

En la ecuación (13) se presenta el error absoluto medio con el que es posible evaluar la precisión de un modelo de regresión [37].

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (13)$$

Donde:

n es el número total de observaciones.

y_i son los valores reales.

$\hat{y}_i$ son los valores predichos.

5.3 Marco referencial

Bala & Islam (2009), utilizaron datos de reflectancia TERRA MODIS de 8 días con 500 m de resolución espacial para un periodo de entre 2005 a 2006 para estimar el rendimiento de papa en Bangladesh. La información satelital fue validada mediante datos reales de campo de 50 agricultores. Desarrollaron modelos de regresión utilizando los Índices de Vegetación (NDVI, LAI, fPAR) y el rendimiento de papa a nivel de campo para seis unidades administrativas. En sus resultados, las ecuaciones de predicción de rendimiento mostraron altos coeficientes de correlación (R^2), son 0,84, 0,72 y 0,80 para el NDVI, LAI y fPAR, respectivamente. Encontraron que un error promedio de estimación es de alrededor del 15% para la región de estudio.

Newton, et al. (2018), utilizaron imágenes satelitales de Landsat 5TM, Landsat 7 (ETM+) y Landsat 8 (OLI) para la temporada de crecimiento del cultivo de papa en un periodo del 2010 a 2011 y otras 6 temporadas de cultivo entre 2011-2017. Realizaron correcciones geométricas y radiométricas del sensor utilizado, adicionalmente realizaron entrevistas a agricultores para determinar las ubicaciones y rendimientos normales. Realizaron el cálculo de NDVI Mediante análisis de regresión entre los valores de NDVI y los rendimientos de papa a nivel de campo. Obtuvieron un perfil de cambios cronológicos del NDVI para los cultivos entre 2010-2011, el máximo coeficiente de correlación de la ecuación de predicción de rendimiento entre el NDVI medio y el rendimiento real, el error encontrado entre el rendimiento previsto y el estimado es de alrededor del 10.4% con una desviación de 3,4 t/ha [38].

Roznik et al. (2022), realizaron una investigación para determinar la influencia de la resolución espacial de imágenes satelitales MODIS en procesos de estimación de rendimientos de cultivo encontrando que una mejor resolución en las imágenes satelitales podría ayudar a mejorar en la precisión de la estimación del rendimiento de los cultivos mediante la combinación de máscaras de tierras de cultivo de alta resolución. Utilizaron el método de regresión spline cubico que proporcionó un enfoque interpretable y flexible para estimar el rendimiento del cultivo, utilizando NDVI para cada época considerada (2008-2018) en tres resoluciones (250m, 500m y 1km) para compararlos con datos de rendimiento oficiales para cuatro cultivos (maíz, soja, trigo de primavera y trigo de invierno). Obtuvieron como resultado mejoras en el modelo de rendimiento de cultivos donde el mayor aumento en la precisión fue en el trigo de primavera pasando de $R^2 = 0.190$ en resolución 1km a $R^2 = 0.367$ con resolución de 250 m lo que corresponde a un 93% de mejora de la precisión del modelo. Sin embargo, el uso de mayores resoluciones genera una mayor exigencia computacional para lo cual se plantea la utilización de plataformas de procesamiento en la nube como Google Earth Engine [28].

Ntouros et al. (2025), utilizaron datos de Sentinel-2 MSI de reflectancia superficial en nueve campos de cultivo de trigo operativos localizados en la ciudad de Ptolemaida, dentro de la unidad regional de Macedonia Occidental, Grecia. Los datos de rendimiento se obtuvieron mediante entrevistas a los agricultores. El procesamiento se realizó exclusivamente con software y marcos de programación no comerciales y de código abierto (Python, Google Earth Engine, SNAP y QGIS). Calcularon índices de vegetación como NDVI, GNDVI, EVI y LAI empleando SNAP, y posteriormente extrajeron los valores medios a nivel de campo. Para la estimación del rendimiento aplicaron análisis de regresión, relacionando el "Área bajo la curva" (AUC) de los índices de vegetación con los valores medios de rendimiento reportados. Los resultados mostraron un RMSE que oscila entre 40 y 60 kg por cada 1000 m^2 . El mejor desempeño se obtuvo con el GNDVI que alcanzó un RMSE de 41,35 para trigo duro, en el caso de trigo blando, el error osciló entre 59 y 102 kg por cada 1000 m^2 siendo el LAI el mejor predictor con un RMSE de 59,13. Adicionalmente, implementaron una aplicación web desarrollada en Python, la cual permite la búsqueda de imágenes, calcular índices, generar estadísticas de índices de vegetación y predecir el rendimiento [42].

Manivasagam et al. (2021), emplearon imágenes de reflectancia superficial Sentinel-2 y PlanetScope, corregidos atmosféricamente (nivel 2) y libres de nubes, en un campo de trigo comercial de 52 ha cercano a Saad, Israel, durante los años 2017 y 2018. A partir de estas imágenes, calcularon 13 índices de vegetación, entre ellos el NDVI y GNDVI, que fueron utilizados para generar mapas de LAI de alta resolución. Posteriormente, aplicaron el Modelo de algoritmo simple para estimación de rendimiento (SAFY). Los resultados mostraron que la fusión de imágenes PlanetScope con Sentinel-2 permitieron una mejor estimación del rendimiento en comparación con Sentinel-2. No obstante, si bien, sus resultados no capturaron bien el rango de variabilidad en el rendimiento, La estimación del rendimiento obtenida se correlacionó moderadamente con un R^2 de 0,45 demostrando así que PlanetScope y Sentinel-2 pueden proporcionar información de alta resolución para derivar la variables de cultivo y utilizarlas en modelos como el SAFY para estimar la variabilidad de rendimiento dentro del campo [43].

Kalecinski et al. (2024), emplearon imágenes de sensores ópticos y de radar, Sentinel-1 (Banda C) (GRD) de polarización dual (VV + VH), RADARSAT-2 (Banda C) y UASVAR (Banda L) para el estudio de predicción de maíz, soja y arroz en 2019 en Mississippi, EE. UU. Utilizaron imágenes ópticas de PlanetScope y Harmonized Landsat Sentinel-2 (HLS), así como datos SAR provenientes de Sentinel-1

(Banda C, productos GRD en polarización dual VV y VH), RADARSAT-2 (Banda C) y UASVAR (Banda L). Los datos de rendimiento real fueron recopilados directamente en campo a partir de una cosechadora. Calcularon índices de vegetación como NDVI, EVI, GNDVI, entre otros. Aplicaron modelos de regresión lineal y Random Forest utilizando la librería scikit-learn de Python. La evaluación de los modelos de regresión la realizaron mediante métricas como el R^2 , el R^2 Ajustado y el RMSE. Encontraron que, para el arroz y el maíz, el modelo con mejor desempeño se obtuvo con el modelo óptico basado en HLS empleando todas las bandas espectrales, alcanzando un R^2 Ajustado de 0,69 y 0,73 respectivamente. En el caso de la soja, el mejor modelo fue el que relaciona las bandas verdes, rojo y NIR. Entre los modelos SAR, el de RADARSAT-2 presentó mejor desempeño para los tres cultivos con R^2 Ajustado de 0,43 [44].

Mesas et al. (2025), estimaron el rendimiento de cultivos de espinaca a partir de imágenes de radar y ópticas. Para ello emplearon imágenes de Sentinel-1 en modo IW con polarización dual VV y VH, procesadas como productos GRD. Cada imagen fue preprocesada mediante “Sentinel-1 Toolbox” aplicando eliminación de ruido térmico, calibración radiométrica y corrección de terreno. Adicionalmente, se utilizaron imágenes de Sentinel-2. Calcularon diferentes índices a partir de Sentinel-1 (VH, VV, VH/VV, RVI, etc) y a partir de Sentinel-2 el NDVI. Para cada índice determinaron la mediana correspondiente al periodo comprendido entre octubre de 2017 a enero de 2018. Posteriormente, se aplicó un modelo de regresión Random Forest (RF-R) para estimar el NDVI a partir de los índices radiométricos Sentinel-1. Dichos índices fueron empleados para analizar la relación entre los valores de NDVI y el rendimiento del cultivo, evaluando el desempeño mediante el coeficiente de determinación (R^2) y la tasa de error. Lograron valores de R^2 en el rango de 0,63 a 0,89 [45].

Rananavare & Chitnis (2022), emplearon diversos modelos de aprendizaje automático (máquinas de soporte vectorial, redes neuronales, bosques aleatorios y regresiones lineales) utilizando datos agrícolas de cebolla y arroz del estado de Karnataka en la India. Las variables consideradas incluyeron área, producción, estado, temporada y tipo de cultivo. Para el procesamiento, construcción de modelos y evaluación de desempeño (MAE, MSE y R^2) con bibliotecas como numpy, pandas, seaborn, matplotlib y Scikitlearn. Los conjuntos de entrenamiento los dividieron en 70% entrenamiento y 30% para prueba. Obtuvieron como resultado las métricas de desempeño de cada modelo y utilizaron el MAE para comparar los diferentes modelos, encontrando así que el mejor modelo para la predicción de cultivos en su caso de estudio fueron las redes neuronales. Sin embargo es de resaltar que la regresión Random Forest obtuvo un MAE de 2,818 y la regresión lineal tuvo un MAE de 1,471 lo que indica que si bien no son los mejores predictores, tienen potencial en la predicción de la productividad [46].

5.4 Materiales

Para el desarrollo de la investigación se emplearon imágenes ópticas de la constelación PlanetScope, correspondientes a la generación SuperDove, que provee información en ocho bandas espectrales con alta resolución espacial (3 m). Estas imágenes fueron utilizadas para el cálculo de índices de vegetación y para complementar la información de Sentinel-2 en zonas con nubosidad o pérdida de datos. Las características técnicas de las bandas se presentan en la Tabla 1[16].

Tabla 1. Bandas de PlanetScope.

Banda	Resolución	Longitud de onda central (nm)	Descripción
B1	3 m	441 - 452 nm	Ultra azul (Costa y Aerosol)
B2	3 m	465 - 515 nm	Azul

B3	3 m	513 - 549 nm	Verde I
B4	3 m	547 - 583 nm	Verde
B5	3 m	600 - 620 nm	Amarillo
B6	3 m	650 - 680 nm	Rojo
B7	3 m	697 - 713 nm	Borde Rojo
B8	3 m	845 - 885 nm	Espectroscopia de infrarrojo cercano (NIR)

Asimismo, se utilizaron imágenes ópticas de Sentinel-2, que proporcionan información multispectral con resoluciones de 10, 20 y 60 metros. La Tabla 2 muestra las trece bandas que componen este sensor, abarcando desde el espectro visible hasta el infrarrojo de onda corta, con resoluciones de 10, 20 y 60 metros [17].

Tabla 2. Bandas de Sentinel-2.

Banda	Resolución	Longitud de onda central (nm)	Descripción
B1	60 m	443 nm	Ultra azul (Costa y Aerosol)
B2	10 m	490 nm	Azul
B3	10 m	560 nm	Verde
B4	10 m	665 nm	Rojo
B5	20 m	705 nm	Visible e Infrarrojo Cercano (VNIR)
B6	20 m	740 nm	Visible e Infrarrojo Cercano (VNIR)
B7	20 m	783 nm	Visible e Infrarrojo Cercano (VNIR)
B8	10 m	842 nm	Visible e Infrarrojo Cercano (VNIR)
B8a	20 m	865 nm	Visible e Infrarrojo Cercano (VNIR)
B9	60 m	945 nm	Onda Corta Infrarroja (SWIR)
B10	60 m	1375 nm	Onda Corta Infrarroja (SWIR)
B11	20 m	1610 nm	Onda Corta Infrarroja (SWIR)
B12	20 m	2190 nm	Onda Corta Infrarroja (SWIR)

Adicionalmente, se incorporaron imágenes de radar de la misión Sentinel-1, adquiridas en banda C y en modo IW (Interferometric Wide Swath), con polarizaciones VV y VH. Este tipo de datos permite analizar la estructura del dosel vegetal y complementar la información óptica en condiciones de nubosidad, mediante la generación de índices de retrodispersión como RVI y VH/VV.

Las imágenes abarcaron el área rural del municipio de Ventaquemada (Boyacá) durante el periodo comprendido entre los meses uno y cuatro posteriores a la siembra, que corresponden a las etapas fenológicas más representativas del cultivo (emergencia, máximo vigor y maduración).

Además, se dispuso de registros de productividad obtenidos en campo para los predios analizados, expresados en toneladas por hectárea (t/ha), los cuales se emplearon como variable de referencia en la evaluación de los modelos de estimación.

El procesamiento y análisis de la información se apoyó en herramientas de software especializadas, entre ellas SNAP, QGIS, ArcMap y Python, utilizadas para las etapas de preprocesamiento, análisis de índices e integración de resultados.

6 Metodología

Después de realizar la búsqueda bibliográfica de información referente al proyecto, se encontró una metodología que es comúnmente utilizada con ciertas diferencias en los modelos que se utilizan para la comprobación del modelo, sin embargo, en esencia, las metodologías encontradas comparten gran similitud. La Figura 4 muestra la metodología utilizada para el desarrollo del presente estudio, donde se presentan las diferentes fases que se planea realizar, con 5 fases principales que son: Definición del área de estudio, Obtención de datos, procesamiento, Establecimiento y optimización del modelo y comprobación del modelo de regresión.

6.1 Esquema metodológico

En la Figura 4 se presenta el diagrama que evidencia el flujo de actividades a realizar en la presente investigación.

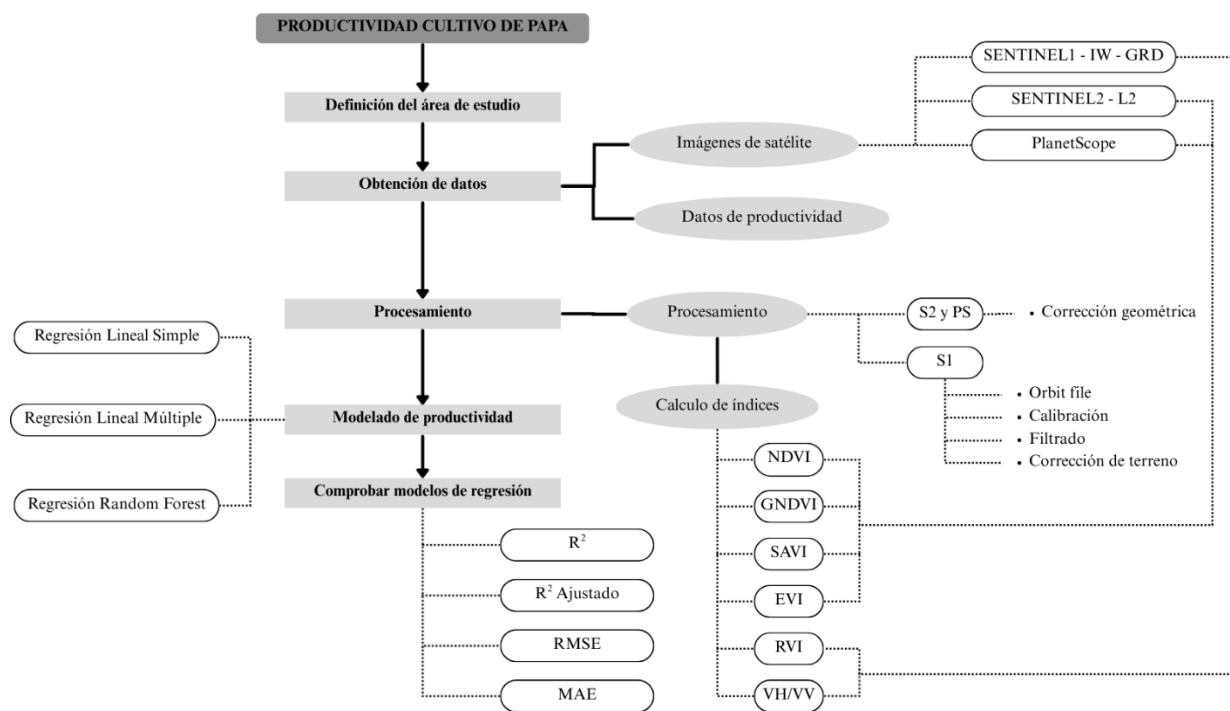


Figura 4. Esquema metodológico.

6.2 Área de estudio

El municipio de Ventaquemada (Ver Figura 5) se encuentra ubicado al suroccidente del departamento de Boyacá la provincia Centro, conocido como la “Puerta de Oro” del departamento, con una superficie de 15.100 Ha, localizado a 5°25’1” latitud Norte y 73°30’00” longitud Oeste a 3.033 msnm. El 46% de la economía del municipio corresponde a actividades agrícolas. Por las características geográficas y climáticas, en el monitoreo del año 2023 – II de la UPRA se registraron 1.217 ha cultivadas de papa en el municipio.

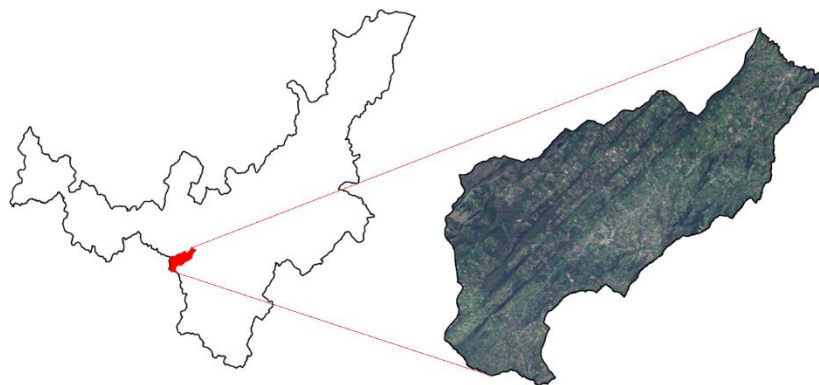


Figura 5. Ubicación de Ventaquemada en Boyacá

6.3 Obtención de datos

6.3.1 Datos de sensores remotos

Se utilizaron imágenes de sensores remotos obtenidas a partir de plataformas oficiales y de acceso gratuito. En el caso de Sentinel-1 y Sentinel-2 las imágenes fueron descargadas del portal de Copernicus Browser (<https://browser.dataspace.copernicus.eu/>) de la Agencia Espacial Europea (ESA). Se seleccionaron los productos nivel L2A para Sentinel-2 (reflectancia superficial) y GRD - IW (Ground Range Detected, Interferometric Wide Swath) para Sentinel-1, garantizando que cubrieran el área correspondiente al municipio de Ventaquemada en el periodo fenológico del cultivo de papa. En el caso de PlanetScope, las escenas fueron adquiridas mediante la plataforma de PlanetScope <https://www.planet.com/> de libre acceso gracias a la iniciativa Norway's International Climate and Forests Initiative (NICFI) por medio del cual Noruega proporciona un acceso integral a la monitorización por satélite de alta resolución con imágenes de PlanetScope, el programa permite el acceso Nivel 1 que permite la descarga de mosaicos mensuales de Reflectancia superficial (mosaicos históricos y futuros) que fueron los que se utilizaron en el presente estudio.

Para la selección de las imágenes se establecieron criterios temporales y espaciales, priorizando las fechas comprendidas entre los meses 1 y 4 posteriores a la siembra, correspondientes a las fases críticas del cultivo (emergencia, máximo vigor y tuberización). Asimismo, las descargas se delimitaron al área de estudio en el municipio de Ventaquemada (Boyacá). De esta forma se conformó un conjunto de imágenes satelitales de distintas fuentes, con resoluciones temporales y espaciales elevadas, lo que garantizó la disponibilidad de datos suficientes para el cálculo de los índices de vegetación empleados en el análisis.

En la Figura 6 se presenta una muestra de las imágenes obtenidas para cada fase del cultivo considerada en este estudio. Iniciando por una fase de emergencia, en la cual se presentan los primeros brotes del cultivo y la cobertura de la vegetación es baja y por tanto se encuentra una alta influencia del suelo. La segunda y tercera etapa corresponden a la de máximo vigor y tuberización respectivamente, en las que la cobertura de la vegetación es considerablemente mayor disminuyendo la influencia del suelo y finalmente, la etapa más tardía o cercana a la cosecha en la que la rama de la planta se corta con el fin de que el tubérculo termine su maduración, en esta etapa la vegetación presente disminuye por lo mencionado anteriormente.

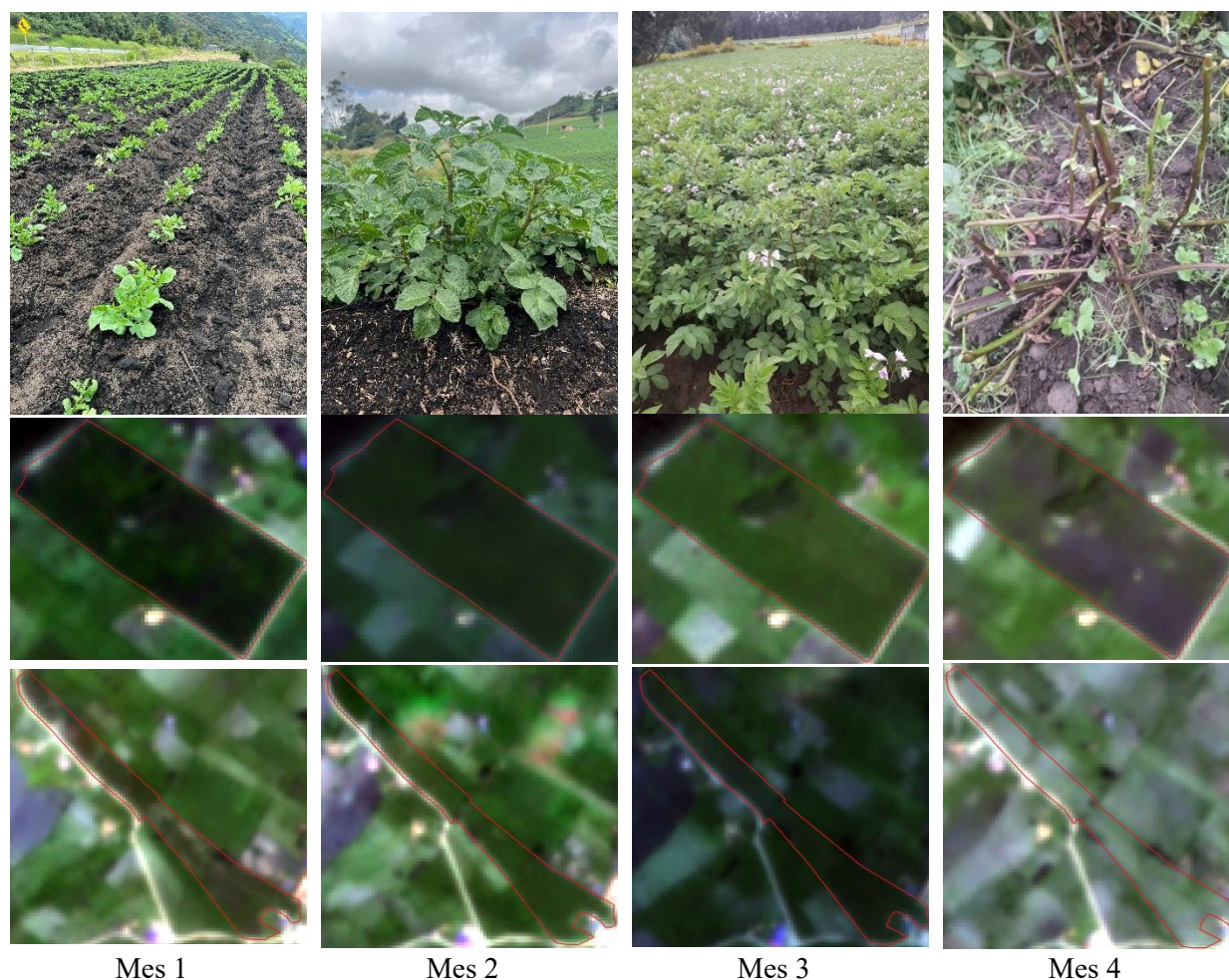


Figura 6. Imágenes obtenidas para cada fase del cultivo.

6.3.2 Obtención de datos de campo

Para validar los modelos de datos generados mediante técnicas de teledetección, fue necesario contar con información base de la zona de estudio. En este caso, uno de los parámetros fundamentales para dicha validación fue el rendimiento del cultivo, expresado en toneladas por hectárea (t/ha). Este indicador constituye una medida directa de la productividad agrícola y, por tanto, permite establecer una relación cuantificable entre las variables espectrales obtenidas por sensores remotos y la respuesta real del cultivo en campo. La obtención de esta información base no solo facilita la calibración y verificación de los modelos desarrollados, sino que también permite evaluar el grado de precisión y aplicabilidad de estos en contextos agrícolas reales. Además, el rendimiento del cultivo sirve como punto de referencia para comparar los resultados derivados de distintos índices de vegetación o algoritmos de estimación, aportando así robustez y confiabilidad al proceso de análisis. Por lo anterior, se realizaron visitas a 20 predios cultivados con papa en el municipio de Ventaquemada. En cada uno se registraron tres variables principales, proporcionadas por los agricultores: (i) productividad, expresada en toneladas por hectárea (t/ha); (ii) fecha aproximada de siembra, que indica el inicio del ciclo productivo; y (iii) fecha aproximada de cosecha, que marca el final del ciclo agrícola.

6.4 Procesamiento de información

6.4.1 Preprocesamiento y corrección de datos

Los datos de sensores satelitales están influenciados por varios factores como la dispersión y absorción atmosférica, la geometría de la iluminación del objetivo del sensor, la calibración del sensor y procedimientos de procesamiento de datos [38].

El preprocesamiento de las imágenes constituye una etapa fundamental para garantizar la calidad, homogeneidad y comparabilidad de las imágenes satelitales y lo que garantiza que la extracción de información de estas sea confiable. En esta etapa se aplicaron una serie de procedimientos según el tipo de sensor:

6.4.1.1 Radar - Sentinel-1

En la Figura 7 se representa la secuencia metodológica aplicada en el preprocesamiento de imágenes de Sentinel-1 utilizando SNAP. Donde se realizaron los procesos de lectura de datos, incorporación de órbita precisa, seguida de calibración radiométrica y reducción de ruido speckle. Posteriormente, se realizó la corrección geométrica mediante el modelo digital de elevación, recorte al área de interés y finalmente la exportación de imagen procesada en archivo ráster para posteriores procesos (Ver Figura 8).



Figura 7. Flujo de procesamiento en Sentinel-1.

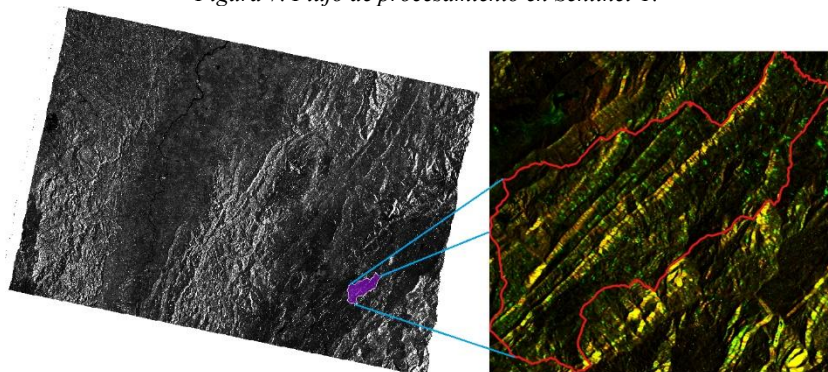


Figura 8. Preprocesamiento información de Sentinel-1.

6.4.1.2 Ópticos (Sentinel-2 – PlanetScope)

Una vez obtenida la información, es necesaria la realización de correcciones con el fin de mejorar la calidad de las imágenes, iniciando con la depuración de información de manera que la información que se utiliza sea la que presente mejor calidad, si bien se puede encontrar gran variedad de información de imágenes satelitales debido a la cantidad de misiones y satélites disponibles, es necesario seleccionar la información que presente mejor calidad, por ejemplo, imágenes con un determinado porcentaje de nubosidad e importante, de una determinada fecha. Además de la depuración de información se realizarán las siguientes correcciones:

- *Corrección radiométrica*

Este proceso corresponde a la eliminación de los efectos de factores atmosféricos como la influencia de vapor de agua y aerosoles. El proceso se realiza convirtiendo los valores digitales (ND) en reflectancia superficial a través de la conversión de radiancia mediante la ecuación (1). Los datos de radiancia se convierten en reflectancia superficial mediante la ecuación (2).

- *Creación de máscara binaria*

Tanto las imágenes de Sentinel-2 L2A como las de PlanetScope son imágenes de reflectancia superficial (SR) corregidas atmosféricamente, en el caso de Sentinel-2 son derivadas de productos nivel 1C asociados. La corrección atmosférica de las imágenes incluye la corrección de dispersión de Rayleigh, de efectos de absorción y dispersión de gases atmosféricos. Esto genera una serie de productos adicionales como el mapa de espesor óptico de aerosol (AOT), un mapa de vapor de agua (WV) y un mapa de clasificación de escenas (SCL).

Este último consiste en una máscara en la que se identifican los píxeles potencialmente nublados mediante filtrado de la región roja del espectro, estos píxeles se someten a una serie de filtros basados en umbrales de bandas, ratios e índices (NDVI, NDSI) siguiendo el esquema metodológico presentado en la Figura 9. El filtrado secuencial de un píxel potencialmente nublado se detiene en el momento de que un resultado de prueba establece su probabilidad en 0, que sería el píxel sin nubes o cielo despejado.

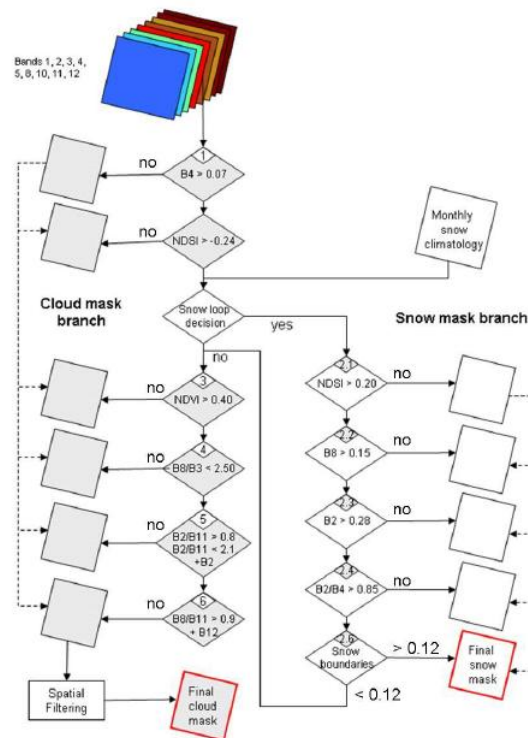


Figura 9. Secuencia del algoritmo de detección de nubes/nieve de L2.

Este proceso se realiza para cada píxel de la imagen para obtener la clasificación que se presenta en la Tabla 3 [47].

Tabla 3. Categorías clasificación de escena SCL

Etiqueta	Clasificación
0	Sin datos (Faltan datos)
1	Píxel saturado o defectuoso
2	Sombras proyectadas topográficas
3	Sombras de nubes
4	Vegetación
5	Sin vegetación
6	Agua
7	Sin clasificar
8	Probabilidad media de la nube
9	Nube de alta probabilidad
10	Cirros delgados
11	Nieve o hielo

En el presente estudio, mediante Python se utilizó la banda SCL del producto Sentinel-2 L2A con una resolución espacial de 20 m. A partir de esta capa se seleccionaron las clases: 3 (Sombras de nubes), 8 (Probabilidad media de la nube), 9 (Nube de alta probabilidad) y 10 (Cirros delgados), con el fin de generar una máscara binaria de nubes. En dicha máscara, los valores correspondientes a estas clases se reclasificaron como “1: Presencia de nubes” mientras que las demás clases se clasificaron como “0: No nubes” (Ver Figura 10) .

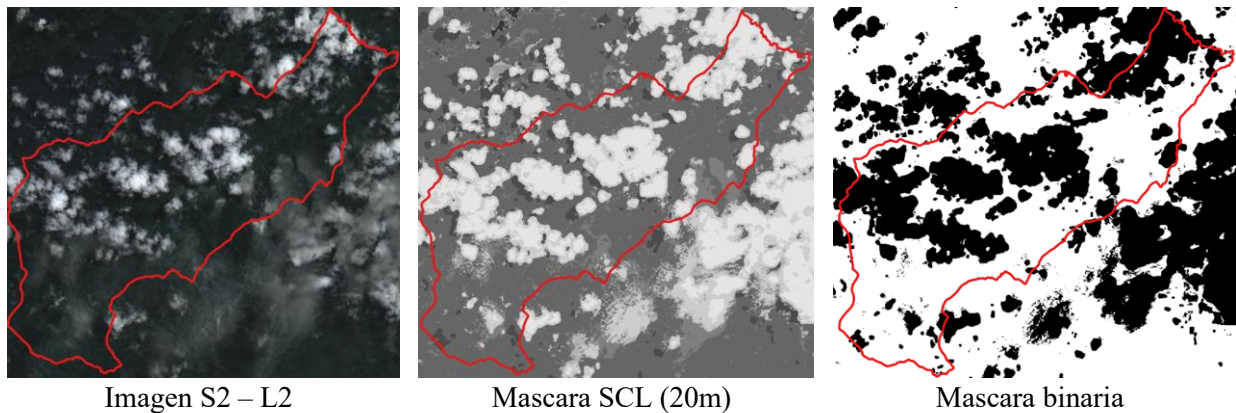


Figura 10. Generación de máscara binaria de nubes.

La capa SCL originalmente presenta una resolución espacial de 20 m, por lo que la máscara binaria generada a partir de ella también mantiene esa resolución. Posteriormente, esta máscara fue sometida a un proceso de remuestreo con el fin de homogeneizar con las otras bandas de Sentinel-2 (10m). El procedimiento se realizó utilizando Python empleando el método de vecino más cercano, lo que permitió conservar la integridad de la máscara binaria, de esta manera, puede ser utilizada de manera confiable en procesos posteriores.

6.4.1.3 Combinación de Sentinel-2 con PlanetScope.

La máscara binaria se utilizó para realizar la fusión entre imágenes de Sentinel-2 y PlanetScope, con el objetivo de reducir los efectos ocasionados por la presencia de nubes. El proceso contempla la reproyección

y recorte de la imagen de Sentinel-2 al sistema de referencia y el área de cobertura de las imágenes de PlanetScope, asegurando la compatibilidad espacial. Posteriormente, se aplicó un proceso de reescalado y alineación subpíxel de PlanetScope mediante el uso de correlación cruzada, esto garantiza un ajuste preciso de los sensores. En este procedimiento resultó fundamental el uso de la máscara binaria de nubes creada en la sección anterior. Finalmente, se aplicó un proceso de fusión ponderada por píxel, en la que los valores espectrales de Sentinel-2 y PlanetScope se combinaron proporcionalmente dependiendo de la presencia o ausencia de nubes siguiendo un modelo de mezcla lineal. Este proceso se realizó utilizando librerías de Python como rasterio, scikit-image y scipy. En la Figura 11 se presenta un ejemplo del resultado obtenido al realizar el proceso anteriormente mencionado.

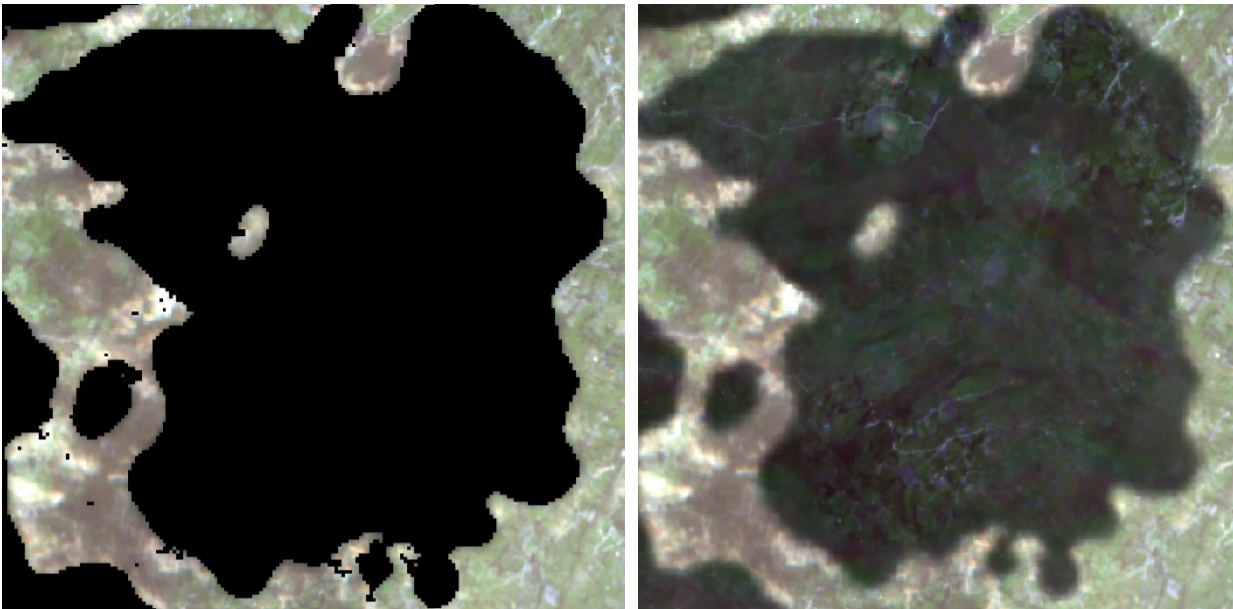


Imagen con mascara

Imagen reemplazada

Figura 11. Comparación de reemplazo de información en imágenes con nubes.

Posteriormente, todas las escenas fueron reproyectadas al sistema de referencia EPSG:3116 MAGNA-SIRGAS / Colombia Bogota zone, unificando la georreferenciación para la integración con las demás imágenes. Se estableció una resolución objetivo de 10 m, correspondiente a la del sensor MSI de Sentinel-2. Con ello se obtuvo un conjunto de imágenes homogéneas y listas para el cálculo de índices de vegetación.

6.4.2 Procesamiento y cálculo de índices

La absorción de radiación en la región de longitud de onda visible (VIS) (RGB, 400-680nm) se encuentra dominada por pigmentos fóliales como la clorofila (430 y 460 nm), carotenoides (a 450 nm) y antocianinas (a 550 nm). La clorofila -A y -B tiene una fuerte absorción de luz en las longitudes de onda azul y roja como apoyo a la fotosíntesis lo que contribuye a la coloración verde del follaje saludable. En la Figura 12 se observa una pendiente pronunciada de la reflectancia en la región de longitud de onda del borde rojo (680 – 750 nm, donde las hojas pasan de una fuerte absorción de luz por la clorofila en el rojo a una fuerte reflectancia en el infrarrojo cercano (NIR)) está controlada por los pigmentos de las hojas, la estructura de las hojas y las capas del dosel [48]. La reflectancia NIR (750-1300 nm) es impulsada principalmente por el proceso de dispersión de fotones tanto a nivel de hoja como de dosel. Las características de absorción de agua se encuentran dentro de la región de longitud de onda NIR (a 970 y 1200 nm) [25] La reflectancia

infrarroja de onda corta (SWIR, 1300-2500 nm) está dominada por el agua de las hojas. A medida que aumenta el contenido de agua de las hojas, la reflectancia SWIR disminuye considerablemente. Los constituyentes de las hojas, como el nitrógeno, la lignina, el almidón, los compuestos fenólicos y las proteínas, también afectan a la reflectancia SWIR [49].

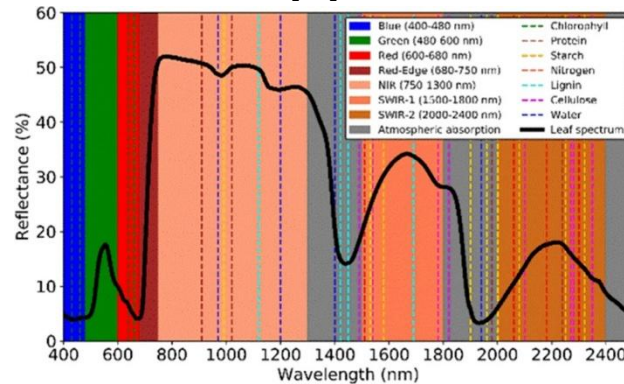


Figura 12. Espectro de reflectancia de la hoja de papa. Características de absorción de las concentraciones químicas foliares[49].

El cálculo de índices de vegetación es crucial ya que representa la salud del cultivo en base a la actividad fotosintética de la planta los índices de vegetación mencionados NDVI, SAVI, GNDVI, EVI se calculan mediante las ecuaciones(3), (4), (5) y (6) y respectivamente, los índices de Sentinel-1 RVI y la Relación VH/VV extraídos mediante las ecuaciones (7) y (8). En la Figura 13 se presentan los índices de vegetación calculados para uno de los predios estudiados.

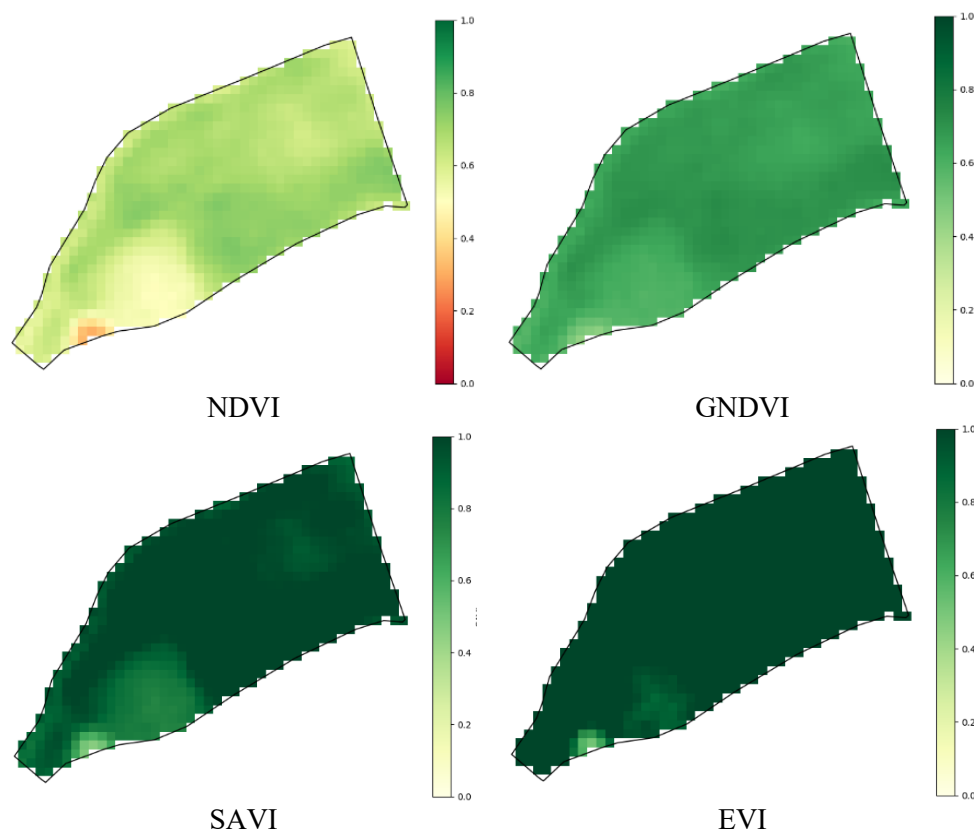


Figura 13. Representación de índices de vegetación.

Para cada uno de los 20 predios seleccionados para el estudio, se realizó la extracción de estadísticas mediante QGIS, utilizando la herramienta “Estadísticas de zona”. Esta herramienta permitió calcular los valores mínimo, máximo y medio de cada uno de los índices previamente generados. Dichas estadísticas resumen la variabilidad interna de cada índice dentro de cada predio, facilitando su posterior uso como variables posiblemente explicativas en modelos de estimación de productividad. Los resultados fueron exportados en formato .xlsx, lo que permitió su posterior organización en tablas para realizar análisis estadístico.

6.5 Establecimiento de modelos de regresión

El establecimiento de modelos de regresión constituye una parte esencial en la estimación de productividad, ya permite determinar o cuantificar la relación existente entre índices espectrales derivados de imágenes satelitales y los datos reales de productividad tomados en campo.

6.5.1 Regresión lineal simple

La regresión lineal simple se realizó utilizando la ecuación (9) y permitió analizar de manera individual (de un único predictor) la relación existente entre cada índice calculado y la productividad real, constituyéndose en una primera aproximación para evaluar la capacidad predictiva de cada variable de manera individual. En total, se realizaron alrededor 180 regresiones lineales simples, 48 por cada sensor óptico (Sentinel 2, PlanetScope y la combinación de Sentinel2 + PlanetScope) y 36 correspondientes a los parámetros de Sentinel 1. La cantidad de modelos se explica porque cada sensor aporta valores para cada índice.

6.5.2 Regresión lineal múltiple

La regresión lineal múltiple que está representada por la ecuación (10), se planteó como una herramienta útil en este estudio, dado que el desarrollo del cultivo se encuentra influenciado por múltiples factores. A diferencia de la regresión lineal simple. Que evalúa el efecto de un solo predictor, este enfoque múltiple permite integrar varios índices de vegetación y parámetros SAR en un mismo modelo, lo que puede mejorar la capacidad explicativa y mejora en las estimaciones de productividad.

En total se realizaron cerca de 282 regresiones lineales múltiples, distribuidas en diferentes grupos; 64 modelos basados en índices derivados de PlanetScope, 64 con los de Sentinel 2 y 64 con los índices de la combinación de Sentinel 2 + PlanetScope, 30 de la combinación de PlanetScope + Sentinel 1, 30 de Sentinel 2 + Sentinel 1 y 30 de la combinación de Sentinel 2 con PlanetScope + Sentinel 1. Cada conjunto incluyó diversas combinaciones de índices.

6.5.3 Regresión Random Forest

Los cultivos como el de papa, por lo general presentan dinámicas no lineales y efectos de interacción que podrían limitar la capacidad predictiva de los modelos lineales simples y múltiples. Ante estas limitaciones, se implementó un modelo de regresión con enfoque de aprendizaje automático Random Forest, que permite incorporar diferentes variables, reduce el riesgo de sobreajuste y permite capturar relaciones complejas mediante el uso de un conjunto de árboles de decisión. En este estudio se realizaron aproximadamente 126 regresiones utilizando Random Forest, implementadas a través de la librería Scikit-learn de Python.

Estos modelos se compararon con las regresiones lineales tradicionales, lo que permitió evaluar no solo la precisión de las predicciones, sino la potencialidad de los modelos de regresión Random Forest para estimar productividad de cultivos de papa a partir de índices de vegetación extraídos de imágenes de sensores remotos.

6.6 Comprobación del modelo de regresión

La generación de un modelo que represente un determinado fenómeno es importante, pero una parte importante al momento de generar un modelo matemático es comprobar su validez y exactitud de los valores que este modelo genera, este proceso de comprobación generalmente se realiza con la comparación de los valores calculados mediante la utilización del modelo y los valores considerados como verdaderos, en el presente proyecto, se realizó la comparación entre los valores de productividad estimados con los valores de productividad reales. Existen diferentes métodos para la validación de modelo. En este caso se utilizó el coeficiente de correlación (R^2) explicado en la ecuación (12), en la que, un valor más cercano a 1 indica que los conjuntos de datos tienen mayor relación, por tanto mayor fiabilidad en el modelo, además se utilizaron otras métricas como el RMSE mediante la utilización de la ecuación (11) para la relación de los conjuntos de datos, mientras más cercano sea a 0 mayor precisión tiene el modelo y el MAE expresado en la ecuación (13), que complementa la evaluación del desempeño del modelo.

7 Resultados

7.1 Adquisición de información de productividad

La Figura 14 presenta el área de estudio correspondiente al municipio de Ventaquemada, junto con la localización geográfica de los predios evaluados. En ella se observan los puntos que representan cada predio, distribuidos en diferentes sectores del municipio según su ubicación registrada con coordenadas geográficas.

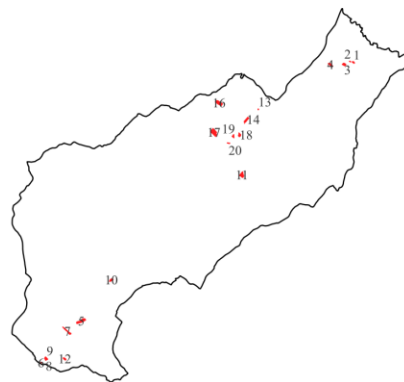


Figura 14. Localización de los 20 predios muestreados en el municipio de Ventaquemada.

En la Tabla 4 se muestran los valores registrados para cada predio, incluyendo las coordenadas de localización (latitud y longitud), el valor de precisión, la altitud o cota (Z), la productividad obtenida de los agricultores, expresada en toneladas por hectárea (t/ha) y en número de bultos, así como el área total del predio en hectáreas. La información está organizada por códigos de identificación (P1, P2, P3, etc.), permitiendo asociar de forma directa la ubicación representada en la figura con los valores productivos y de superficie correspondientes. Esta información constituye la línea base para la posterior comparación con los valores obtenidos mediante el análisis de percepción remota.

Tabla 4. Variables registradas en los predios evaluados.

<i>Predio</i>	<i>Lat.</i>	<i>Long</i>	<i>Precisión ±</i>	<i>Z</i>	<i>Producción (t/ha)</i>	<i>Producción (Bultos)</i>	<i>Área (ha)</i>
P1	5,444533	-73,442429	5	2.842	59	1180	0,794
P2	5,444339	-73,443536	4	2.840	57	1140	0,273
P3	5,442753	-73,446714	6	2.864	58	1160	1,758
P4	5,442654	-73,451803	5	2.897	60	1200	2,05
P5	5,324276	-73,567594	7	2.754	61	1220	5,334
P6	5,306011	-73,585720	7	2.910	59	1200	0,568
P7	5,320406	-73,574539	7	2.750	58	1160	2,844
P8	5,306931	-73,583753	5	2.907	57	1200	1,102
P9	5,307089	-73,583379	9	2.902	54	1080	0,642
P10	5,343197	-73,554328	3	2.901	58	1160	1,739
P11	5,391865	-73,494899	4	2.910	59	1180	3,849
P12	5,306846	-73,575029	9	2.883	57	1140	1,176
P13	5,422535	-73,486277	5	3.044	54	201	0,186
P14	5,416619	-73,491817	9	3.036	56	1120	2,356
P15	5,416568	-73,491919	9	3.036	51	1020	0,368
P16	5,424397	-73,503843	7	3.045	58	1160	5,265
P17	5,424397	-73,503843	9	3.045	57	1140	7,118
P18	5,411331	-73,495047	4	3.051	59	1180	1,528
P19	5,409373	-73,497202	4	3.054	57	1140	1,023
P20	5,406593	-73,499337	9	3.023	58	1160	0,439

Los datos de campo evidencian una variabilidad considerable en la productividad de los predios evaluados, con valores que oscilan entre 51 y 61 t/ha, así como diferencias significativas en el área sembrada (desde parcelas menores a 0,3 ha hasta extensiones superiores a 7 ha). Estas variaciones iniciales son relevantes, ya que aportan un rango amplio de referencia para validar los modelos de estimación basados en índices de vegetación.

7.2 Modelos de regresión lineal simple para la estimación de productividad

En esta sección se presentan los resultados de los modelos de regresión lineal simple aplicados para estimar la productividad de cultivos de papa a partir de índices de vegetación derivados de imágenes satelitales. Para los sensores ópticos PlanetScope y Sentinel-2 se emplearon los índices NDVI, GNDVI, SAVI y EVI, organizados de manera que cada índice incluye los resultados individuales y combinados de ambos sensores. En el caso de Sentinel-1, el análisis se estructuró a partir de los índices derivados de sus polarizaciones VH y VV, la razón VH/VV y el RVI. En las tablas se muestran, para cada mes evaluado, los dos valores más representativos entre media, mínimo y máximo, considerando las pendientes de las ecuaciones y el coeficiente de determinación R^2 .

7.2.1 NDVI

Se presentan resultados de las ecuaciones de regresión lineal obtenidas para el índice de vegetación NDVI, considerando las imágenes de PlanetScope, Sentinel-2 y la combinación Sentinel-2 + PlanetScope. Para

cada fuente de datos se calcularon los valores de media, máximo y mínimo del NDVI en diferentes meses, junto con sus respectivas ecuaciones y el coeficiente de determinación (R^2).

7.2.1.1 PlanetScope

Los resultados del NDVI de PlanetScope, presentados en la Tabla 5, muestran un comportamiento variable según la fase del cultivo. En el mes 3 se registraron los ajustes más destacados, alcanzando un coeficiente de determinación de 0,369 y una pendiente de 28,073 para el NDVI máximo, lo que refleja una relación sólida durante la etapa de mayor crecimiento. En el mes 2 también se observó un ajuste aceptable con el NDVI medio, con un R^2 de 0,130 y una pendiente de 12,695, aunque con menor consistencia. En contraste, los meses 1 y 4 presentaron pendientes negativas o reducidas y valores de R^2 bajos, lo que limita su capacidad explicativa en las fases inicial y final del ciclo.

Tabla 5. Ecuaciones de regresión lineal y R^2 por mes para el NDVI de PlanetScope.

Mes	NDVI	Ecuación	R^2
1	Media	$-4,103 NDVI_{\bar{x}} + 59,618$	0,079
	Mínimo	$-7,198 NDVI_{min} + 60,404$	0,173
2	Media	$12,695 NDVI_{\bar{x}} + 48,011$	0,130
	Máximo	$13,099 NDVI_{max} + 46,528$	0,109
3	Media	$20,696 NDVI_{\bar{x}} + 41,124$	0,261
	Máximo	$28,073 NDVI_{max} + 33,211$	0,369
4	Máximo	$9,208 NDVI_{max} + 49,985$	0,165
	Mínimo	$-4,823 NDVI_{min} + 59,713$	0,080

En la Figura 15 se presenta la comparación entre la productividad real y la estimada para el mes 3, considerando los valores máximos del NDVI. Se observa una alta dispersión en los datos, y las métricas de error indican un desempeño limitado del modelo, con un RMSE= 1,763 y un MAE= 1,344, equivalentes a un error superior a 1,3 t/ha.

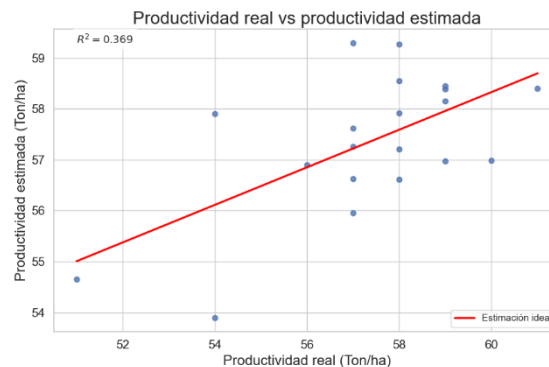


Figura 15. Productividad real vs. estimada – NDVI (PlanetScope)

7.2.1.2 Sentinel-2

Los resultados del NDVI de Sentinel-2, presentados en la Tabla 6, evidencian ajustes débiles en todas las fases del cultivo. En el mes 4 se obtuvo la pendiente más alta con un valor de 3,217 para el NDVI máximo, aunque con un coeficiente de determinación limitado de 0,086. El mejor ajuste se registró en el mes 2 con el NDVI mínimo, que alcanzó un R^2 de 0,142 acompañado de una pendiente negativa de $-0,782$, lo que

refleja una relación débil e inversa. En los meses 1 y 3, tanto las pendientes como los coeficientes de determinación se mantuvieron cercanos a cero, confirmando que los modelos presentaron un comportamiento marginal y con escasa capacidad explicativa.

Tabla 6. Ecuaciones de regresión lineal y R^2 por mes para el NDVI de Sentinel-2

Mes	Tipo	Ecuación	R^2
1	Media	$-3,521 NDVI_{\bar{x}} + 57,904$	0,056
	Mínimo	$-4,019 NDVI_{min} + 57,717$	0,047
2	Media	$-1,623 NDVI_{\bar{x}} + 57,768$	0,021
	Mínimo	$-6,299 NDVI_{min} + 58,295$	0,142
3	Media	$0,430 NDVI_{\bar{x}} + 57,226$	0,002
	Máximo	$0,957 NDVI_{max} + 56,947$	0,013
4	Media	$1,569 NDVI_{\bar{x}} + 56,900$	0,016
	Máximo	$3,217 NDVI_{max} + 56,014$	0,086

En la Figura 16 se presenta la comparación entre la productividad estimada y la productividad real para el mes 2, considerando los valores mínimos del NDVI. Se observa una alta dispersión de los valores obtenidos, lo que refleja un bajo desempeño del modelo en la estimación de productividad. Las métricas de error confirmaron este resultado, con un RMSE= 2,056 y un MAE= 1,617.

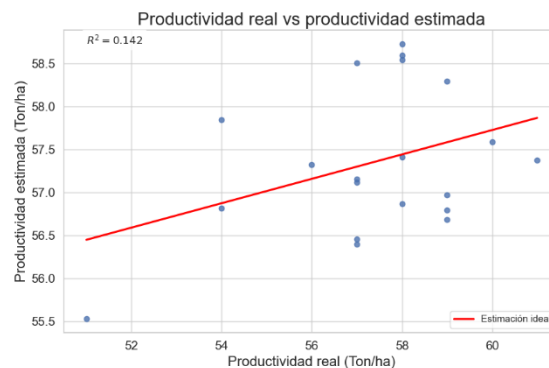


Figura 16. Productividad real vs. estimada – NDVI (Sentinel-2)

7.2.1.3 Sentinel-2 + PlanetScope

Los resultados del NDVI de la combinación Sentinel-2 + PlanetScope, presentados en Tabla 7, muestran variaciones según la fase del cultivo. En el mes 4, el NDVI máximo alcanzó un coeficiente de determinación de 0,176 con una pendiente positiva de 4,759, lo que refleja una relación directa y consistente entre el vigor del dosel y la productividad durante la etapa de maduración. En el mes 2 se registraron pendientes positivas de hasta 4,658 con ajustes moderados, lo que permite describir el crecimiento activo, aunque con menor precisión. En contraste, los meses 1 y 3 presentaron pendientes próximas a cero y valores de R^2 reducidos, lo que limita su capacidad predictiva en las fases inicial y de máximo desarrollo vegetativo, donde la combinación de resoluciones no logra compensar la heterogeneidad intraparcela.

Tabla 7. Ecuaciones de regresión lineal y R^2 por mes para el NDVI de Sentinel-2+PlanetScope

Mes	Tipo	Ecuación	R^2
1	Media	$-2,243 NDVI_{\bar{x}} + 58,203$	0,032
	Mínimo	$-3,537 NDVI_{min} + 58,350$	0,057

2	Máximo	$4,658 NDVI_{min} + 54,089$	0,071
	Mínimo	$-2,247 NDVI_{min} + 58,208$	0,027
3	Media	$3,366 NDVI_{\bar{x}} + 55,229$	0,049
	Máximo	$4,435 NDVI_{max} + 54,059$	0,078
4	Máximo	$4,759 NDVI_{max} + 54,380$	0,176
	Mínimo	$-1,329 NDVI_{min} + 57,752$	0,010

En la Figura 17 se presenta la comparación entre la productividad estimada y la productividad real para los valores máximos de NDVI en el mes 4, se observa una alta dispersión en los datos, lo que refleja un elevado nivel de error en la estimación de productividad. Este resultado se confirma con las métricas obtenidas: un RMSE de 2.014 y un MAE de 1,52.

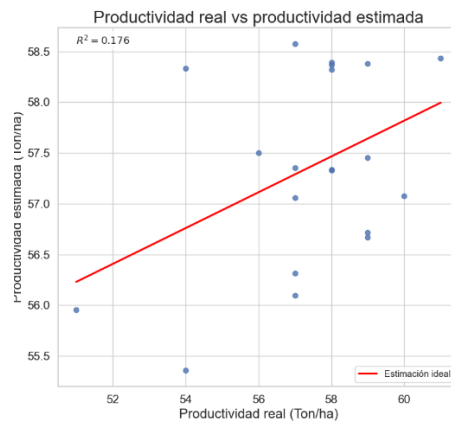


Figura 17. Productividad real vs. estimada – NDVI (Sentinel-2 + PlanetScope)

El comportamiento del NDVI reflejó con claridad la evolución del cultivo de papa a lo largo del ciclo productivo, con ajustes diferenciados según la fuente de datos utilizada. El valor más alto se alcanzó con imágenes de PlanetScope en el mes 3, donde el NDVI máximo registró un coeficiente de determinación de 0,369 y una pendiente de 28,073. Este resultado confirma la capacidad de este satélite, gracias a su alta resolución espacial, para detectar variaciones finas en etapas de cobertura foliar densa y homogénea. No obstante, la magnitud elevada de la pendiente limita su interpretación agronómica, ya que pequeñas variaciones del índice implicarían cambios desproporcionados en la productividad.

Un comportamiento más estable se observó en la combinación Sentinel-2 + PlanetScope, particularmente en el mes 4 con el NDVI máximo, que alcanzó un coeficiente de determinación de 0,176 y una pendiente de 4,759. Aunque el ajuste fue más modesto en términos estadísticos, resulta más coherente con la dinámica fenológica del cultivo en la fase de maduración, cuando la cobertura vegetal mantiene suficiente biomasa para sostener una relación consistente con el rendimiento.

En conjunto, el NDVI se consolidó como un indicador robusto del desarrollo del cultivo, al basarse en la relación entre la reflectancia en el rojo y el infrarrojo cercano, lo que le permite captar de manera sensible la densidad y cantidad de biomasa verde. Estos resultados son coherentes con la fisiología del cultivo, ya que reflejan directamente la evolución del dosel y el cierre de filas, factores determinantes en la interceptación de radiación y, en consecuencia, en el rendimiento final.

7.2.2 GNDVI

Se presentan los resultados de las ecuaciones de regresión lineal obtenidas para el GNDVI con imágenes de PlanetScope, Sentinel-2 y sus combinaciones, considerando media, máximo y mínimo por mes con sus respectivas ecuaciones y el coeficiente de determinación (R^2).

7.2.2.1 PlanetScope

Los resultados del GNDVI de PlanetScope, presentados en la Tabla 8, muestran que el mejor desempeño se alcanzó en el mes 3 con el valor máximo, que registró un coeficiente de determinación de 0,348 acompañado de una pendiente de 30,315. Esta combinación refleja una relación fuerte y estadísticamente consistente entre el vigor foliar y la productividad, en un momento donde el cultivo alcanzaba la máxima cobertura del dosel. La media de ese mismo mes también presentó un ajuste relevante, con R^2 de 0,233 y pendiente de 24,125, lo que confirma la estabilidad del índice en esta fase intermedia. En el mes 4, aunque el valor máximo alcanzó un R^2 de 0,158, su capacidad predictiva fue más moderada, indicando que la señal se debilita hacia la maduración. Por el contrario, en los meses 1 y 2 los ajustes fueron apenas discretos, con R^2 que no superaron 0,107, lo que sugiere que el índice ofrece poca capacidad explicativa en etapas tempranas debido a la heterogeneidad del establecimiento del cultivo.

Tabla 8. Ecuaciones de regresión lineal y R^2 por mes para el GNDVI de PlanetScope.

Mes	GNDVI	Ecuación	R^2
1	Media	$-3,446 \text{ GNDVI}_{\bar{x}} + 59,369$	0,049
	Mínimo	$-5,594 \text{ GNDVI}_{\min} + 60,051$	0,107
2	Media	$12,127 \text{ GNDVI}_{\bar{x}} + 48,884$	0,092
	Máximo	$11,603 \text{ GNDVI}_{\max} + 48,429$	0,065
3	Media	$24,125 \text{ GNDVI}_{\bar{x}} + 39,586$	0,233
	Máximo	$30,315 \text{ GNDVI}_{\max} + 33,329$	0,348
4	Máximo	$12,628 \text{ GNDVI}_{\max} + 47,787$	0,158
	Mínimo	$-7,650 \text{ GNDVI}_{\min} + 61,487$	0,125

En la Figura 18 se presenta la comparación entre productividad real y la productividad estimada para el mes 3, considerando los valores máximos de GNDVI. Se observa una alta dispersión en los datos, lo que refleja un elevado nivel de error en la estimación de productividad. Este resultado se confirma con las métricas del modelo, donde se obtuvo un RMSE de 1,792 y un MAE de 1,519.

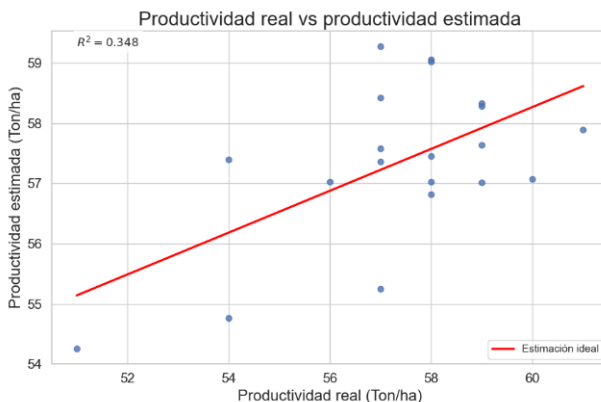


Figura 18. Productividad real vs. estimada – GNDVI (PlanetScope)

7.2.2.2 Sentinel-2

Los resultados del GNDVI con Sentinel-2, presentados en la Tabla 9, muestran desempeños en general bajos, con un mejor ajuste en el mes 2 en el valor mínimo, que alcanzó un coeficiente de determinación de 0,127 con pendiente negativa de -6,186. Este comportamiento indica que el índice logró capturar parte de la variabilidad productiva durante la fase de establecimiento, aunque en sentido inverso al rendimiento. En los meses 1 y 3 los valores de R^2 no superaron 0,037, acompañados de pendientes reducidas, lo que evidenció escasa capacidad de explicación en etapas iniciales e intermedias. Finalmente, en el mes 4 el valor máximo presentó un R^2 de 0,049 con pendiente positiva de 2,624, un ajuste muy limitado que confirma la debilidad de Sentinel-2 para consolidar relaciones estables con la productividad, posiblemente debido a su resolución espacial y a los efectos de nubosidad sobre la señal.

Tabla 9. Ecuaciones de regresión lineal y R^2 por mes para el GNDVI de Sentinel-2.

Mes	GNDVI	Ecuación	R^2
1	Media	$-2,799 \text{ GNDVI}_{\bar{x}} + 57,787$	0,037
	Mínimo	$-2,777 \text{ GNDVI}_{\min} + 57,655$	0,027
2	Media	$-1,839 \text{ GNDVI}_{\bar{x}} + 57,77$	0,024
	Mínimo	$-6,186 \text{ GNDVI}_{\min} + 58,231$	0,127
3	Máximo	$0,962 \text{ GNDVI}_{\max} + 56,997$	0,012
	Mínimo	$-1,781 \text{ GNDVI}_{\min} + 57,565$	0,012
4	Media	$1,202 \text{ GNDVI}_{\bar{x}} + 57,014$	0,008
	Máximo	$2,624 \text{ GNDVI}_{\max} + 56,35$	0,049

En la Figura 19 se presenta la comparación entre productividad real y la productividad estimada para el mes 2, utilizando los valores mínimos de GNDVI. De forma similar a los casos anteriores, se observa una alta dispersión en los datos, lo que refleja un elevado nivel de error en la estimación de productividad. Este resultado se confirma con las métricas del modelo, que reportaron un RMSE de 2,074 y un MAE de 1,604.

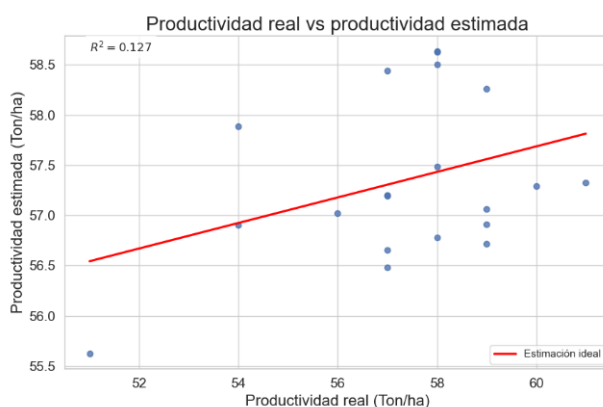


Figura 19. Productividad real vs. estimada – GNDVI (Sentinel-2)

7.2.2.3 Sentinel2 + PlanetScope

Los resultados de la combinación Sentinel-2 + PlanetScope para el GNDVI, presentados en la Tabla 10, muestran desempeños bajos en general, con mejoras puntuales en etapas específicas del cultivo. El ajuste más alto se registró en el mes 4, cuando el GNDVI máximo alcanzó un R^2 de 0,104 con pendiente positiva de 4,097, lo que indica una relación moderada y coherente con la productividad en la fase de maduración.

En contraste, los meses iniciales presentaron ajustes muy reducidos, con valores de R^2 que no superaron 0,034 y pendientes bajas o negativas, lo que refleja limitaciones en la capacidad predictiva durante la emergencia y el desarrollo temprano del cultivo.

Tabla 10. Ecuaciones de regresión lineal y R^2 por mes para el GNDVI de Sentinel-2 + PlanetScope.

Mes	GNDVI	Ecuación	R^2
1	Media	$-1,597 * GNDVI_{\bar{x}} + 57,983$	0,015
	Mínimo	$-2,720 * GNDVI_{Min} + 58,222$	0,034
2	Máximo	$4,859 * GNDVI_{Max} + 54,199$	0,073
	Mínimo	$-2,317 * GNDVI_{Min} + 58,248$	0,029
3	Media	$3,224 * GNDVI_{\bar{x}} + 55,467$	0,042
	Máximo	$4,804 * GNDVI_{Max} + 54,083$	0,082
4	Máximo	$4,097 * GNDVI_{Max} + 54,919$	0,104
	Mínimo	$-2,056 * GNDVI_{Min} + 58,063$	0,023

En la Figura 20 se presenta la comparación entre la productividad real y la productividad estimada para el mes 4, utilizando los valores mínimos de GNDVI. De forma similar a los casos anteriores, se observa una alta dispersión en los datos aun mayor, lo que refleja un mayor nivel de error en la estimación de productividad. Este resultado fue confirmado al calcular las métricas del modelo, que reportaron un RMSE de 2,102 y un MAE de 1,547.

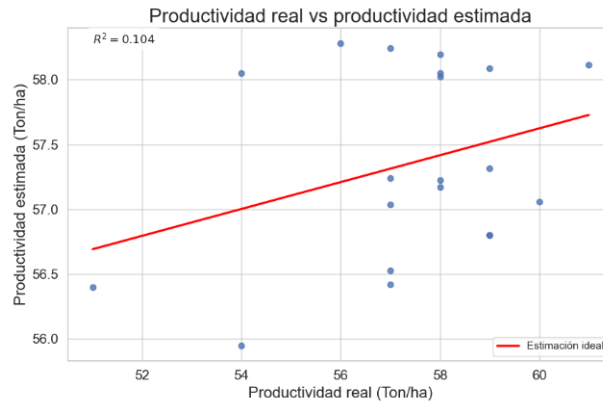


Figura 20. Productividad real vs. estimada – GNDVI (Sentinel-2 + PlanetScope)

El GNDVI mostró un comportamiento progresivamente más sólido a lo largo del ciclo productivo del cultivo de papa, destacándose especialmente en las etapas intermedias, cuando la cobertura foliar adquiere mayor relevancia para el rendimiento. El mejor ajuste se obtuvo con PlanetScope en el mes 3, con un coeficiente de determinación de 0,348 y una pendiente positiva de 30,315, lo que evidencia una relación estadísticamente consistente entre el verdor del follaje y la productividad. Este desempeño se explica por la sensibilidad del índice a la absorción de luz por la clorofila, que le permite reflejar variaciones en el vigor y la eficiencia fotosintética incluso antes de alcanzar la máxima cobertura. La alta resolución espacial de PlanetScope favoreció, además, la detección de la heterogeneidad intrapredial, consolidándolo como la fuente con mayor capacidad explicativa.

En el caso de Sentinel-2, los ajustes fueron bajos y las pendientes débiles, lo que limitó su utilidad en fases críticas, particularmente en los estadios iniciales. Las combinaciones con PlanetScope no superaron los valores máximos de este último de manera individual, pero aportaron mayor estabilidad temporal y

redujeron la dispersión de los resultados, confirmando la utilidad de integrar distintas escalas espaciales y temporales para caracterizar con mayor fidelidad la dinámica del cultivo.

Desde el punto de vista fenológico, en las fases iniciales predominaron relaciones negativas con baja significancia, reflejando la dificultad del índice para captar la heterogeneidad propia de los primeros estadios de desarrollo. Sin embargo, conforme se consolidó la estructura del dosel, los ajustes se fortalecieron y las pendientes se tornaron positivas, confirmando al GNDVI como un índice eficaz para representar el vigor del follaje en los momentos en que este resulta más determinante para la productividad.

7.2.3 SAVI

Se reportan las ecuaciones de regresión lineal del SAVI para PlanetScope, Sentinel-2 y sus combinaciones, con media, máximo y mínimo por mes, junto con sus respectivas ecuaciones y el coeficiente de determinación (R^2).

7.2.3.1 PlanetScope

Los resultados del SAVI de PlanetScope, presentados en la Tabla 11, muestran un comportamiento diferenciado según la etapa fenológica del cultivo. El mejor desempeño se registró en la fase de maduración, cuando el SAVI máximo alcanzó un coeficiente de determinación de 0,283 acompañado de una pendiente positiva de 53,455, reflejando una relación directa y sólida con la productividad. Durante el crecimiento vegetativo también se observaron ajustes importantes, particularmente con la media del índice, que en el mes 2 alcanzó un R^2 de 0,183 y en el mes 3 de 0,264, ambos con pendientes positivas que confirman la capacidad del índice para capturar la expansión del dosel y su vínculo con el rendimiento. En contraste, en la fase inicial el SAVI mínimo llegó a un R^2 de 0,126 con pendiente negativa, lo que evidenció una relación inversa y limitada en el establecimiento del cultivo.

Tabla 11. Ecuaciones de regresión lineal y R^2 por mes para el SAVI de PlanetScope.

Mes	SAVI	Ecuación	R^2
1	Media	$-2,761 * SAVI_{\bar{x}} + 59,586$	0,067
	Mínimo	$-4,087 * SAVI_{Min} + 59,959$	0,126
2	Media	$38,585 * SAVI_{\bar{x}} + 19,345$	0,183
	Mínimo	$0,0009 * SAVI_{Min} + 57,349$	0,000
3	Media	$93,914 * SAVI_{\bar{x}} - 35,815$	0,264
	Mínimo	$1,501 * SAVI_{Min} + 56,146$	0,006
4	Máximo	$53,455 * SAVI_{Max} + 4,306$	0,283
	Mínimo	$-3,263 * SAVI_{Min} + 59,718$	0,086

En la Figura 21 se presenta la comparación entre productividad real y la productividad estimada para el mes 3, utilizando los valores medios de SAVI. Se observa una dispersión considerable en los datos, aunque menor que la registrada en algunos de los modelos anteriores. Las métricas de error reportaron un RMSE de 1,904 y un MAE de 1,483.

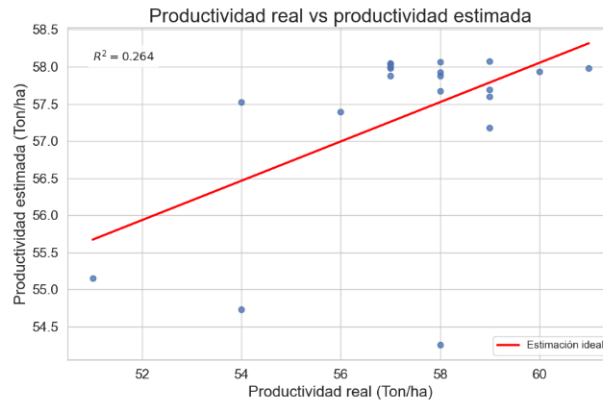


Figura 21. Productividad real vs. estimada – SAVI (PlanetScope)

7.2.3.2 Sentinel-2

Los resultados del SAVI de Sentinel-2, presentados en la Tabla 12, muestran un desempeño limitado, pero con ciertos momentos de mayor consistencia. En la fase intermedia, correspondiente al mes 2, el valor mínimo alcanzó un R^2 de 0,142 con pendiente negativa de $-4,199$, lo que indica que las zonas de menor vigor del dosel explicaron mejor la variabilidad productiva cuando el cultivo consolidaba cobertura, aunque en sentido inverso. En el mes 4, durante la maduración, el valor máximo registró un R^2 de 0,086 con pendiente positiva de $2,145$, lo que refleja una relación directa pero moderada con la productividad. En contraste, en el establecimiento y en el máximo vigor vegetativo, los ajustes fueron bajos y las pendientes cercanas a cero o negativas, evidenciando una señal poco estable. Estos resultados confirman que la menor resolución espacial de Sentinel-2 limita su capacidad para capturar variaciones intraprediales de forma consistente.

Tabla 12. Ecuaciones de regresión lineal y R^2 por mes para el SAVI de Sentinel-2.

	SAVI	Ecuación	R^2
Mes 1	Media	$-2,347 * SAVI_{\bar{x}} + 57,904$	0,056
	Mínimo	$-2,679 * SAVI_{Max} + 57,717$	0,047
Mes 2	Media	$-1,077 * SAVI_{\bar{x}} + 57,766$	0,021
	Mínimo	$-4,199 * SAVI_{Min} + 58,295$	0,142
3	Máximo	$0,641 * SAVI_{Max} + 56,947$	0,013
	Mínimo	$-0,689 * SAVI_{Min} + 57,485$	0,004
Mes 4	Media	$1,046 * SAVI_{\bar{x}} + 56,9$	0,016
	Máximo	$2,145 * SAVI_{Max} + 56,014$	0,086

En la Figura 22 se presenta la comparación entre la productividad real y la productividad estimada para el mes, 2 utilizando los valores mínimos de SAVI. Se observa una alta dispersión en los datos, lo que sugiere un elevado nivel de error en la estimación de productividad. Las métricas de este modelo reportaron un RMSE de 2,056 y un MAE 1,617.

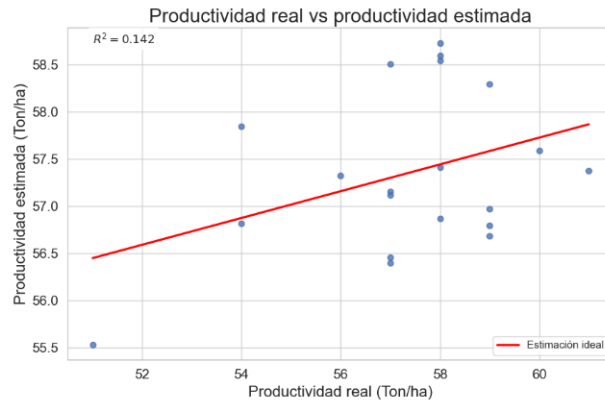


Figura 22. Productividad real vs. estimada – SAVI (Sentinel-2).

7.2.3.3 Sentinel2+Planet

Los resultados del SAVI en la combinación Sentinel-2 + PlanetScope, presentados en la

Tabla 13, reflejan un comportamiento más consistente en las etapas intermedia y final del cultivo. El mejor ajuste se alcanzó en la fase de maduración, donde el SAVI máximo obtuvo un R^2 de 0,251 con una pendiente positiva de 5,518, lo que evidencia una asociación directa y coherente entre el vigor del dosel y la productividad. Durante el crecimiento vegetativo también se observaron desempeños relevantes, especialmente en el mes 3, cuando el valor máximo alcanzó un R^2 de 0,150 y pendiente de 8,648, confirmando la sensibilidad del índice para capturar diferencias en el desarrollo del follaje. En contraste, los meses iniciales presentaron R^2 reducidos y pendientes bajas o negativas, lo que limitó la capacidad explicativa en las fases de establecimiento.

Tabla 13. Ecuaciones de regresión lineal y R^2 por mes para el SAVI de Sentinel-2+PlanetScope

	SAVI	Ecuación	R^2
Mes 1	Media	$-1,720 * SAVI_{\bar{x}} + 58,313$	0,037
	Mínimo	$-2,358 * SAVI_{Min} + 58,350$	0,057
2	Máximo	$4,969 * SAVI_{Max} + 52,645$	0,059
	Mínimo	$-1,620 * SAVI_{Min} + 58,272$	0,030
Mes 3	Media	$4,666 * SAVI_{\bar{x}} + 53,316$	0,081
	Máximo	$8,648 * SAVI_{Max} + 49,072$	0,150
Mes 4	Media	$1,475 * SAVI_{\bar{x}} + 56,350$	0,032
	Máximo	$5,518 * SAVI_{Max} + 52,699$	0,251

En la Figura 23 se presenta la relación entre la productividad real y la productividad estimada para el mes 3 con los valores máximos de SAVI, se observa una dispersión moderada alrededor de la línea de referencia que corresponde a la estimación ideal (predicción perfecta, donde la productividad estimada sería igual a la real). Las métricas para este modelo reportaron un RMSE de 1,922 y un MAE de 1,428.

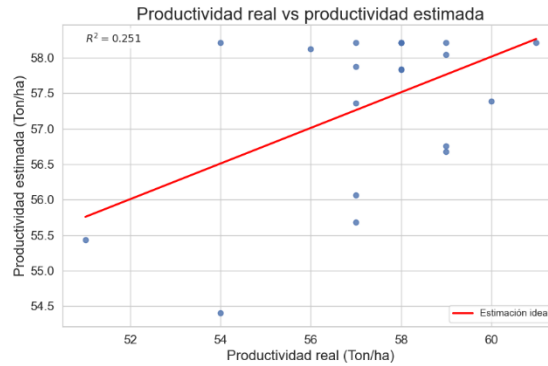


Figura 23. Productividad real vs. estimada – SAVI (Sentinel-2 + PlanetScope)

El desempeño del SAVI en la estimación de productividad mostró un comportamiento moderadamente consistente, con diferencias claras entre las fuentes de información empleadas. El mejor ajuste se alcanzó con PlanetScope en la fase de maduración, donde el SAVI máximo registró un coeficiente de determinación de 0,283 acompañado de una pendiente positiva de 5,518. Este resultado indica que valores más altos del índice se asociaron directamente con incrementos en la productividad, confirmando la capacidad de este sensor para detallar la estructura del cultivo gracias a su mayor resolución espacial.

Las combinaciones de sensores ofrecieron desempeños intermedios: aunque no superaron a PlanetScope de manera individual, aportaron mayor estabilidad temporal a la señal, con ajustes que en maduración se acercaron a R^2 de 0,25 y mantuvieron pendientes positivas. En contraste, Sentinel-2 presentó los coeficientes más bajos, rara vez por encima de 0,086, con pendientes poco pronunciadas que limitaron su capacidad explicativa cuando se utilizó de forma aislada.

Desde la perspectiva fenológica, en las fases iniciales predominaron relaciones negativas y valores de R^2 bajos. Conforme avanzó el ciclo, especialmente durante la maduración de los tubérculos, las pendientes se hicieron positivas y los ajustes se fortalecieron, confirmando que el SAVI fue más eficaz en las etapas avanzadas, cuando el dosel vegetal ya estaba consolidado e influía directamente en la productividad.

7.2.4 EVI

Se presentan los resultados de las ecuaciones de regresión lineal obtenidas para el índice de vegetación EVI, considerando las imágenes de PlanetScope, Sentinel-2, la combinación PlanetScope + Sentinel-2 y la combinación Sentinel-2 + PlanetScope. Para cada fuente de datos se calcularon los valores de media, máximo y mínimo en diferentes meses, junto con sus respectivas ecuaciones y el coeficiente de determinación (R^2).

7.2.4.1 PlanetScope

Los resultados del EVI de PlanetScope, presentados en la Tabla 14, muestran un desempeño muy limitado a lo largo del ciclo. El mejor ajuste se registró en el mes 1 con el valor mínimo, que alcanzó un coeficiente de determinación de 0,080 acompañado de una pendiente negativa de $-4,291$, lo que refleja una relación débil, pero con cierta capacidad explicativa en la fase de establecimiento. Sin embargo, en los meses siguientes los ajustes fueron prácticamente nulos, con R^2 cercanos a cero y pendientes muy variables, lo que confirma que el índice no logró capturar de manera consistente la relación entre el vigor del follaje y la productividad en las etapas de crecimiento y maduración.

Tabla 14. Ecuaciones de regresión lineal y R^2 por mes para el EVI de PlanetScope.

	EVI	Ecuación	R^2
Mes 1	Media	$-8,861 * EVI_{\bar{x}} + 66,034$	0,028
	Mínimo	$-4,291 * EVI_{Min} + 61,038$	0,080
Mes 2	Media	$-136,665 * EVI_{\bar{x}} + 193,935$	0,004
	Mínimo	$-3,365 * EVI_{Min} + 60,417$	0,034
Mes 3	Media	$-41,637 * EVI_{\bar{x}} + 98,971$	0,0005
	Mínimo	$-0,063 * EVI_{Min} + 57,41$	0,00001
Mes 4	Media	$3,600 * EVI_{\bar{x}} + 53,892$	0,011
	Mínimo	$0,0291 * EVI_{Min} + 57,325$	0,000004

En la Figura 24 se presenta la relación entre la productividad real y la productividad estimada para el mes 1 en los valores mínimos de EVI. Se observa una dispersión amplia y los datos no se ajustan de manera consistente a la línea de referencia que sería el ajuste ideal entre estimada y real. En consecuencia, las estimaciones de productividad presentan un grado de error elevado, lo que limita la fiabilidad de este modelo para la predicción. Las métricas asociadas a este modelo reportaron un RMSE de 2,129 y un MAE de 1,659.

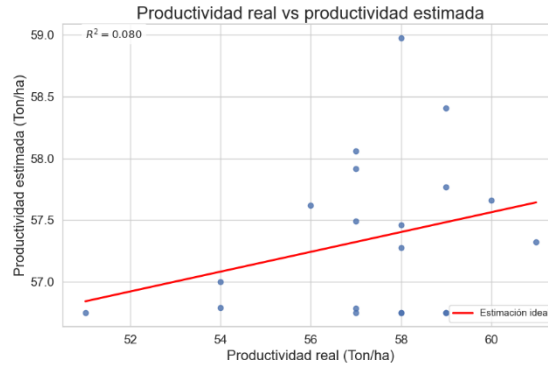


Figura 24. Productividad real vs. estimada – EVI (PlanetScope)

7.2.4.2 Sentinel-2

Los resultados del EVI de Sentinel-2, presentados en la Tabla 15, reflejan un comportamiento débil en todo el ciclo, con ajustes bajos y pendientes poco consistentes. Los valores más destacables se dieron en el mes 1 con el EVI medio, que alcanzó un coeficiente de determinación de 0,078 acompañado de una pendiente negativa de $-1,228$, y en el mes 4 con el EVI mínimo, donde también se obtuvo un R^2 de 0,078 con pendiente negativa de $-0,765$. Estos resultados indican una ligera capacidad explicativa en las fases de establecimiento y maduración, aunque siempre con relaciones débiles y en sentido inverso. En los meses intermedios, los coeficientes de determinación fueron muy reducidos y las pendientes cercanas a cero, lo que confirma la baja utilidad del índice para describir la variabilidad en el crecimiento vegetativo activo.

Tabla 15. Ecuaciones de regresión lineal y R^2 por mes para el EVI de Sentinel-2.

Mes	EVI	Ecuación	R^2
1	Media	$-1,228 * EVI_{\bar{x}} + 57,866$	0,078
	Mínimo	$-0,827 * EVI_{Min} + 57,406$	0,070
2	Media	$-0,414 * EVI_{\bar{x}} + 57,468$	0,014
	Mínimo	$-0,288 * EVI_{Min} + 57,283$	0,011
3	Media	$0,181 * EVI_{\bar{x}} + 57,268$	0,002

	Mínimo	$-0,460 * EVI_{Min} + 57,213$	0,028
4	Media	$-0,979 * EVI_{\bar{x}} + 57,875$	0,069
	Mínimo	$-0,765 * EVI_{Min} + 57,409$	0,078

En la Figura 25 se presenta la relación entre la productividad real y la productividad estimada para el mes 4 usando los valores medios de EVI. Se observa una dispersión de los datos amplia y no sigue una tendencia claramente definida en torno a la línea de referencia que sería la estimación ideal de productividad, lo que indica que las estimaciones presentan un alto grado de error. Este modelo obtuvo un RMSE de 2,141 y un MAE de 1,543.

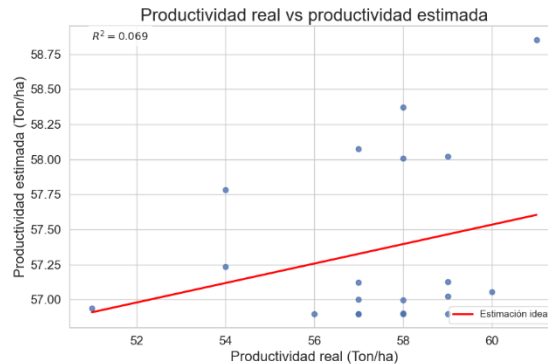


Figura 25. Productividad real vs. estimada – EVI (Sentinel-2)

7.2.4.3 Sentinel2+ PlanetScope

Los resultados del EVI en la combinación Sentinel-2 + PlanetScope, presentados en la Tabla 16, muestran un comportamiento irregular con escasa capacidad explicativa en la mayoría de las fases. El ajuste más alto se registró en el mes 4 con el EVI máximo, que alcanzó un coeficiente de determinación de 0,120 junto a una pendiente muy elevada de 96,442, que, por su magnitud, podría estar influenciado por la presencia de datos atípicos más que por una relación consistente. En la etapa intermedia, el EVI mínimo del mes 2 destacó con un R^2 de 0,090 y pendiente negativa de $-0,947$, lo que sugiere un vínculo débil e inverso con la productividad. En contraste, los meses iniciales y el mes 3 reflejaron ajustes muy bajos, con coeficientes de determinación cercanos a cero y pendientes poco significativas, lo que limita la utilidad de esta combinación para explicar la variabilidad productiva en fases tempranas y de crecimiento vegetativo activo.

Tabla 16. Ecuaciones de regresión lineal y R^2 por mes para el EVI de Sentinel-2+PlanetScope.

Mes	EVI	Ecuación	R^2
1	Media	$-1,744 * EVI_{\bar{x}} + 58,965$	0,024
	Mínimo	$-0,196 * EVI_{Min} + 57,486$	0,003
2	Media	$-8,850 * EVI_{\bar{x}} + 65,950$	0,077
	Mínimo	$-0,947 * EVI_{Min} + 57,968$	0,090
3	Media	$3,716 * EVI_{\bar{x}} + 53,666$	0,002
	Mínimo	$0,022 * EVI_{Min} + 57,335$	0,00005
4	Media	$2,311 * EVI_{\bar{x}} + 55,192$	0,011
	Máximo	$96,442 * EVI_{Min} - 38,916$	0,120

El desempeño del EVI en la estimación de productividad mostró una respuesta menos consistente de lo esperado a lo largo del ciclo del cultivo. El ajuste más alto se alcanzó con la combinación Sentinel-2 + PlanetScope durante el mes 4, en el valor máximo del índice, con un coeficiente de determinación de 0,120

acompañado de una pendiente extremadamente elevada de 96,442. Este comportamiento, aunque estadísticamente destacable, limita su interpretación agronómica, ya que una variación mínima en el índice implicaría incrementos poco realistas en la productividad.

En términos de estabilidad, se identificaron desempeños discretos en las fases iniciales e intermedias, con valores de R^2 inferiores a 0,100 tanto en PlanetScope como en Sentinel-2 de forma individual. Estos resultados reflejan que el EVI no logró capturar de manera consistente la dinámica del cultivo, probablemente debido a la menor sensibilidad del índice en escenarios de alta cobertura foliar y a la influencia de factores externos como el suelo y la atmósfera.

En general, el EVI mostró limitaciones importantes como predictor de productividad, ya que la mayoría de los ajustes fueron bajos y en algunos casos estuvieron condicionados por pendientes extremas. Aun así, la integración de Sentinel-2 y PlanetScope evidenció un potencial de mejora en fases avanzadas, siempre que los resultados estadísticos mantengan correspondencia con los procesos fisiológicos y fenológicos del cultivo.

7.2.5 Índices de Sentinel-1

En esta sección se presentan los resultados de los modelos de regresión lineal simple aplicados a variables derivadas de la información radar del satélite Sentinel-1, considerando las polarizaciones VH y VV, la razón VH/VV y el índice RVI (Radar Vegetation Index). Para cada índice se calcularon valores mínimos, medios y máximos en distintos meses del ciclo productivo, y en las tablas se reportan las dos regresiones lineales más representativas por mes, seleccionadas de acuerdo con el mayor coeficiente de determinación R^2 , incluyendo en cada caso la ecuación correspondiente y el valor de ajuste.

7.2.5.1 VH

Los resultados del índice VH, presentados en la Tabla 17, evidencian contrastes importantes entre las etapas iniciales y avanzadas del cultivo. En el mes 1 se observó el mejor ajuste, con una pendiente de $-128,97$ y un R^2 de 0,178, lo que refleja una relación inversa pero estadísticamente significativa con la productividad en la fase de establecimiento. En el mes 4, ya en etapas más avanzadas, el índice mostró pendientes positivas y moderadas, destacando un valor máximo con pendiente de $51,864$ y R^2 de 0,121, lo que indica una relación débil pero más coherente con el crecimiento vegetativo. En los meses intermedios los ajustes fueron menores, con valores de R^2 bajos y pendientes de menor magnitud, lo que limitó la capacidad explicativa del índice en estas fases.

Tabla 17. Ecuaciones de regresión lineal y R^2 por mes para VH.

Mes		Ecuación	R^2
1	Media	$-39,325 * VH + 58,603$	0,032
	Mínimo	$-128,97 * VH + 60,091$	0,178
2	Mínimo	$-69,272 * VH + 58,887$	0,078
	Media	$-34,605 * VH + 58,483$	0,029
3	Máximo	$34,689 * VH + 55,904$	0,107
	Media	$25,882 * VH + 56,639$	0,019
4	Máximo	$51,864 * VH + 55,148$	0,121
	Media	$38,835 * VH + 56,203$	0,04

En la Figura 26 se presenta la relación entre la productividad real y la productividad estimada para el mes 4 usando los valores máximos de VH. Se observa que los valores tienden a agruparse en torno a la franja superior de la gráfica, con una dispersión moderada alrededor de la línea de referencia. reflejando un ajuste relativamente bajo con errores considerables en la predicción. Las métricas de este modelo reportaron un RMSE de 2,081 y un MAE de 1,687.

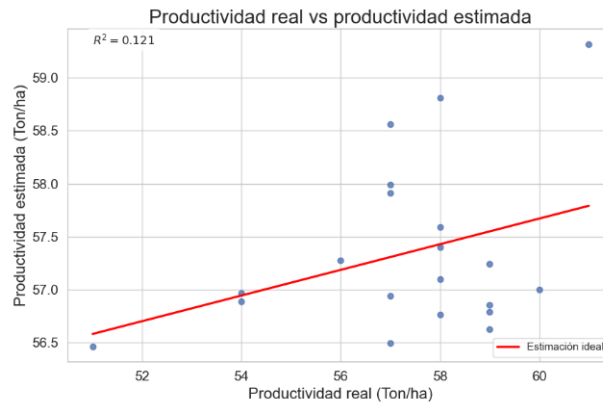


Figura 26. Productividad real vs. estimada – VH (Sentinel-1)

7.2.5.2 VV

Los resultados de la polarización VV, mostrados en la Tabla 18, evidencian un desempeño muy limitado a lo largo del ciclo productivo. Los valores de R^2 se mantuvieron bajos en todos los meses, con un máximo de apenas 0,049 en el mes 4, acompañado de una pendiente positiva de 3,012. Este resultado, aunque representa el mejor ajuste para VV, sigue siendo insuficiente para establecer relaciones significativas entre la señal de radar y la productividad. En los demás meses, las pendientes fueron reducidas y los coeficientes de determinación cercanos a cero, lo que confirma la escasa sensibilidad de esta polarización para explicar variaciones en el cultivo de papa.

Tabla 18. Ecuaciones de regresión lineal y R^2 por mes para VV.

Mes		Ecuación	R^2
1	Mínimo	$-6,097 * VV + 57,958$	0,024
	Máximo	$1,281 * VV + 56,948$	0,018
2	Mínimo	$-3,508 * VV + 57,701$	0,012
	Máximo	$0,295 * VV + 57,248$	0,005
3	Máximo	$1,772 * VV + 56,926$	0,024
	Media	$1,982 * VV + 57,076$	0,005
4	Máximo	$3,012 * VV + 56,662$	0,049
	Media	$6,501 * VV + 56,499$	0,003

En la Figura 27 se muestra la comparación entre la productividad real y la productividad estimada para el mes 4 utilizando los valores máximos de VV. Se observa una alta dispersión de alrededor de la línea de referencia que sería la estimación óptima, lo que sugiere un escaso poder predictivo y por tanto un alto error en la estimación. Las métricas para este modelo reportaron un RMSE de 2,164 y un MAE de 1,589.

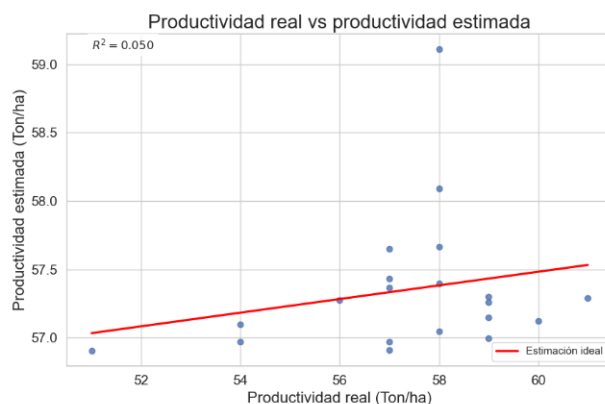


Figura 27. Productividad real vs. estimada – VV (Sentinel-1)

7.2.5.3 VH/VV

Los resultados del índice VH/VV, presentados en la Tabla 19, evidencian una mejora progresiva en la capacidad explicativa a lo largo del ciclo. En las etapas iniciales, las asociaciones fueron débiles, con coeficientes de determinación inferiores a 0,02. A medida que avanzó el desarrollo del cultivo, el índice mostró un comportamiento más estable: en el mes 3 alcanzó un R^2 de 0,113 con pendiente negativa de $-14,804$, reflejando una relación inversa. El mejor desempeño se registró en la fase de maduración, donde el valor máximo obtuvo un R^2 de 0,213 y pendiente positiva de $12,679$, consolidándose como el ajuste más alto de este índice y confirmando su mayor capacidad explicativa en etapas avanzadas del cultivo.

Tabla 19. Ecuaciones de regresión lineal y R^2 por mes para VH/VV

Mes		Ecuación	R^2
1	Máximo	$5,733 * (VH/VV) + 55,298$	0,008
	Mínimo	$-4,313 * (VH/VV) + 57,903$	0,014
2	Máximo	$3,973 * (VH/VV) + 55,851$	0,060
	Mínimo	$5,915 * (VH/VV) + 58,245$	0,023
3	Mínimo	$-14,804 * (VH/VV) + 59,707$	0,113
	Máximo	$5,454 * (VH/VV) + 55,207$	0,097
4	Máximo	$12,679 * (VH/VV) + 52,369$	0,213
	Mínimo	$-9,399 * (VH/VV) + 58,828$	0,044

En la Figura 28 se muestra la relación entre la productividad real y la productividad estimada para el mes 4 usando los valores máximos de VH/VV. La dispersión de los datos alrededor de la línea de referencia refleja que las predicciones presentan un margen de error considerablemente algo. Sin embargo, la inclinación de la recta sugiere cierta correspondencia entre la productividad medida en campo y la estimada por el modelo. Las métricas para este modelo reportaron un RMSE de 1,970 y un MAE de 1,506.

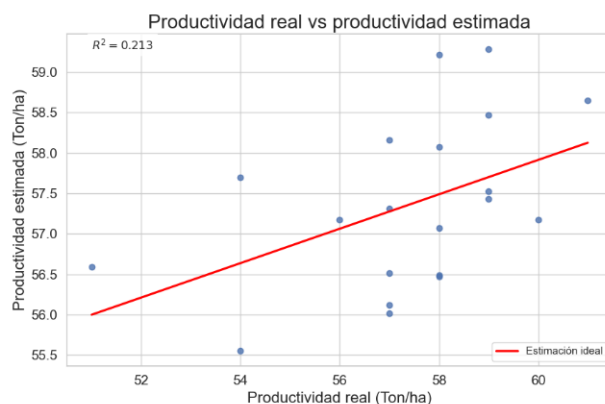


Figura 28. Productividad real vs. estimada – VH/VV (Sentinel-1)

7.2.5.4 RVI

Los resultados del RVI, presentados en la Tabla 20, muestran un patrón de fortalecimiento progresivo a lo largo del ciclo. En las fases iniciales, las asociaciones fueron débiles, con valores de R^2 que no superaron 0,065 en el mes 1. Hacia las etapas intermedias, el índice comenzó a reflejar mayor sensibilidad, alcanzando en el mes 3 un R^2 de 0,109 con pendiente negativa de $-4,569$. El mejor desempeño se observó en la fase de maduración, cuando el valor máximo alcanzó un R^2 de 0,217 acompañado de una pendiente positiva de $6,251$, lo que constituye el ajuste más alto de este índice. Este comportamiento confirma que el RVI adquiere mayor relevancia en etapas avanzadas del cultivo, cuando la acumulación de biomasa y la estructura del dosel permiten establecer relaciones más consistentes con la productividad.

Tabla 20. Ecuaciones de regresión lineal y R^2 por mes para RVI

Mes		Ecuación	R^2
1	Máximo	$2,334 * RVI + 54,935$	0,065
	Mínimo	$-1,463 * RVI + 58$	0,016
2	Máximo	$1,768 * RVI + 55,464$	0,054
	Mínimo	$-1,946 * RVI + 58,357$	0,023
3	Mínimo	$-4,569 * RVI + 59,827$	0,109
	Máximo	$2,485 * RVI + 54,601$	0,076
4	Máximo	$6,251 * RVI + 50,357$	0,217
	Mínimo	$-3,202 * RVI + 59,070$	0,046

En la Figura 29 se presenta la relación entre la productividad real y la productividad estimada para el mes 4 usando los valores máximos de RVI. Se observa una alta dispersión de los datos alrededor de la línea de referencia, lo que muestra que las predicciones aún presentan importante margen de error. Las métricas para este modelo representaron un RMSE de 1,964 y un MAE de 1,495.

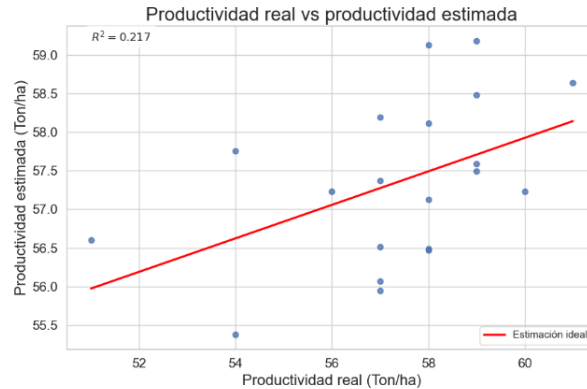


Figura 29. Productividad real vs. estimada – RVI (Sentinel-1)

Los resultados obtenidos mediante regresión lineal simple mostraron que el uso de un único índice de vegetación no es suficiente para describir de manera adecuada la variabilidad productiva del cultivo, especialmente durante las primeras etapas de crecimiento, cuando la cobertura foliar es reducida y las señales espectrales presentan una alta inestabilidad. Aunque algunos índices evidenciaron un desempeño ligeramente superior en los meses 3 y 4, el ajuste general permaneció limitado, lo que indica que la información univariante no capta la complejidad biofísica del dosel ni su relación con el rendimiento. Estas limitaciones indican la necesidad de incorporar enfoques que integren una mayor cantidad de información relevante para mejorar el proceso de estimación.

7.3 Modelos de regresión lineal Múltiple para la estimación de la productividad

En esta sección se presentan los resultados de los modelos de regresión lineal múltiple aplicados para la estimación de la productividad del cultivo de papa. Los modelos fueron construidos empleando como variables explicativas los índices de vegetación NDVI, GNDVI, SAVI y EVI, calculados a partir de imágenes de sensores ópticos (PlanetScope y Sentinel-2) y de sus combinaciones. Adicionalmente, se incluyeron escenarios en los que estos índices se integraron con variables derivadas del radar de Sentinel-1, lo que permitió evaluar el aporte de la información de retrodispersión en la mejora de los ajustes.

Para cada índice se consideraron valores medios, máximos y mínimos en diferentes momentos del ciclo fenológico, con el fin de capturar tanto la variabilidad temporal como espacial del cultivo. Las tablas correspondientes muestran las ecuaciones de regresión obtenidas y los valores de los coeficientes de determinación (R^2) y R^2 ajustado, diferenciados por mes y tipo de valor considerado. Este enfoque permite identificar no solo el sensor o combinación con mayor capacidad explicativa, sino también las fases del desarrollo del cultivo en las que los modelos resultan más estables y consistentes.

7.3.1 PlanetScope.

Los resultados de la regresión lineal múltiple con PlanetScope, presentados en la Tabla 21, muestran un desempeño variable a lo largo del ciclo fenológico del cultivo. El mejor ajuste se registró en el mes 3, cuando el valor máximo alcanzó un coeficiente de determinación de 0,517 y un R^2 ajustado de 0,389, acompañado de una pendiente positiva en las variables, lo que indica una relación sólida entre los índices de vegetación y la productividad en la fase de crecimiento pleno. En el mes 2 se observaron ajustes moderados, con R^2 de 0,328 y 0,333 en la media y el mínimo, lo que refleja una mejora progresiva respecto a etapas anteriores. En contraste, en el mes 1 los valores de R^2 se mantuvieron bajos, con un máximo de

0,265 en el mínimo, señal de una limitada capacidad explicativa en las fases iniciales de establecimiento. Finalmente, en el mes 4, aunque el valor máximo obtuvo un R^2 de 0,387 con un R^2 ajustado de 0,224, se evidenció una disminución en la capacidad predictiva, probablemente asociada a la saturación espectral y a los cambios fisiológicos propios de la maduración del cultivo.

Tabla 21. Regresión lineal múltiple de PlanetScope.

Mes		Ecuación	R^2	R^2_{Aj}
1	Máx.	$60,587 + 11,186 NDVI_{max} - 3,046 GNDVI_{max} - 9,890 SAVI_{max} + 0 EVI_{max}$	0,189	-0,028
	Mín.	$61,857 - 23,868 NDVI_{min} + 5,713 GNDVI_{min} + 8,063 SAVI_{min} - 2,660 EVI_{min}$	0,265	0,070
2	Media	$355,590 + 7,532 NDVI_{\bar{x}} - 5,079 GNDVI_{\bar{x}} + 30,510 SAVI_{\bar{x}} - 331,575 EVI_{\bar{x}}$	0,328	0,149
	Mín.	$54,182 - 108,966 NDVI_{min} + 36,388 GNDVI_{min} + 60,306 SAVI_{min} - 7,180 EVI_{min}$	0,333	0,156
3	Media	$782,392 + 0,053 NDVI_{\bar{x}} + 5,546 GNDVI_{\bar{x}} + 132,774 SAVI_{\bar{x}} - 861,222 EVI_{\bar{x}}$	0,440	0,290
	Máx.	$32,641 + 20,885 NDVI_{max} + 8,518 GNDVI_{max} + 0 SAVI_{max} + 0 EVI_{max}$	0,517	0,389
4	Máx.	$2,833 - 2,289 NDVI_{max} + 2,834 GNDVI_{max} + 54,621 SAVI_{max} + 0 EVI_{max}$	0,387	0,224
	Mín.	$57,502 - 0,839 NDVI_{min} - 2,189 GNDVI_{min} - 4,880 SAVI_{min} + 5,756 EVI_{min}$	0,179	-0,040

En la Figura 30 se presenta la relación entre la productividad real y la productividad estimada para el mes 3, utilizando los valores máximos de los índices NDVI y GNDVI. El modelo presenta una capacidad de predicción moderada, reflejada en la dispersión de los puntos con respecto a la línea de productividad estimada ideal. Al evaluar las métricas de desempeño, se obtuvo un RMSE de 1,759 y un MAE de 1,373, lo que confirma la presencia de errores de predicción que limitan la precisión del modelo.

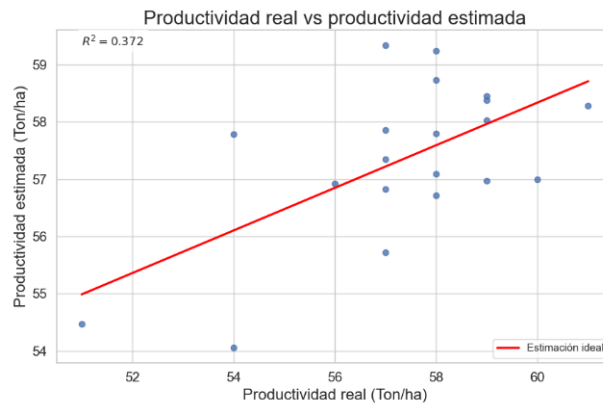


Figura 30. Productividad real vs. estimada – Regresión lineal múltiple (PlanetScope)

7.3.2 Sentinel-2

Los resultados de la regresión lineal múltiple con Sentinel-2, presentados en la

Tabla 22, reflejan un comportamiento diferenciado a lo largo del ciclo del cultivo. El mejor ajuste se alcanzó en el mes 2 con el valor medio de los índices, que registró un coeficiente de determinación de 0,545 y un R^2 ajustado de 0,424, lo que evidencia una relación sólida durante la fase intermedia del crecimiento vegetativo. En la etapa de maduración, el mes 4 mostró también un desempeño destacable, especialmente en el valor máximo, con un R^2 de 0,513 y un R^2 ajustado de 0,384, señal de una asociación positiva en la fase final del ciclo. En contraste, el mes 1 evidenció un ajuste moderado, mientras que en el mes 3 los valores se redujeron de manera considerable, con R^2 por debajo de 0,162, lo que refleja limitaciones en la capacidad explicativa del sensor en ese momento del desarrollo del cultivo.

Tabla 22. Regresión lineal múltiple de Sentinel-2.

Mes		Ecuación	R^2	R^2_{Aj}
1	Media	$58,106 - 103290,558 NDVI_{\bar{x}} + 39,175 GNDVI_{\bar{x}} + 68841,084 SAVI_{\bar{x}} - 2,137 EVI_{\bar{x}}$	0,227	0,021
	Máx	$58,079 - 48768,759 NDVI_{max} + 50,863 GNDVI_{max} + 32474,877 SAVI_{max} + 4,752 EVI_{max}$	0,275	0,081
2	Media	$56,163 - 5767,439 NDVI_{\bar{x}} - 77,386 GNDVI_{\bar{x}} + 3895,586 SAVI_{\bar{x}} + 0,279 EVI_{\bar{x}}$	0,545	0,424
	Mín.	$57,452 + 234935,168 NDVI_{min} + 21,461 GNDVI_{min} - 156652,486 SAVI_{min} - 1,087 EVI_{min}$	0,338	0,161
3	Media	$56,494 - 4687,056 NDVI_{\bar{x}} - 45,855 GNDVI_{\bar{x}} + 3154,345 SAVI_{\bar{x}} - 0,160 EVI_{\bar{x}}$	0,035	-0,222
	Mín.	$57,180 - 41693,325 NDVI_{min} - 56,800 GNDVI_{min} + 27833,844 SAVI_{min} - 0,098 EVI_{min}$	0,162	-0,062
4	Media	$56,383 - 4560,002 NDVI_{\bar{x}} + 15,823 GNDVI_{\bar{x}} + 3036,084 SAVI_{\bar{x}} - 3,080 EVI_{\bar{x}}$	0,262	0,066
	Máx	$55,599 + 192644,332 NDVI_{max} - 30,977 GNDVI_{max} - 128413,439 SAVI_{max} - 3,524 EVI_{min}$	0,513	0,384

En la Figura 31 se presenta la comparación entre la productividad real y la productividad estimada usando los valores medios del mes 2 mediante regresión lineal múltiple, basándose en los índices de vegetación NDVI, GNDVI, SAVI y EVI. Se observa una dispersión más moderada, con mayor concentración en torno a la línea de estimación ideal. lo que refleja un ajuste más consistente del modelo en comparación con los casos menos ajustados, el modelo presenta un desempeño aceptable, con un RMSE de 1,497 y un MAE de 1,107.

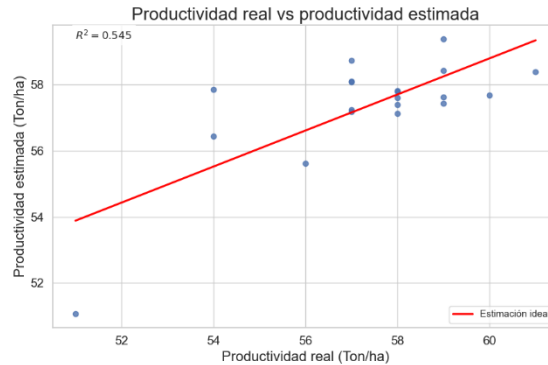


Figura 31. Productividad real vs. estimada – Regresión lineal múltiple (Sentinel-2)

7.3.3 Sentinel-2 +PlanetScope

Los resultados de la combinación Sentinel-2 + PlanetScope, presentados en la Tabla 23, muestran un comportamiento variable a lo largo del ciclo del cultivo. Los ajustes más destacados se alcanzaron en los valores máximos de los meses 1 y 4, con coeficientes de determinación de 0,375 y 0,361 respectivamente, lo que evidencia asociaciones moderadas en las fases inicial y final del ciclo. En contraste, los meses 2 y 3 reflejaron desempeños bajos, con valores mínimos y medios que no superaron un R^2 de 0,198, lo cual limita la capacidad explicativa en estas etapas. En conjunto, esta integración de sensores ofreció relaciones más consistentes en los extremos del ciclo fenológico, mientras que en los periodos intermedios su poder predictivo fue más reducido.

Tabla 23. Regresión lineal múltiple de Sentinel-2 + PlanetScope.

Mes		Ecuación	R^2	R^2_{Aj}
1	Máx	$61,887 + 19,509 NDVI_{max} - 0,922 GNDVI_{max} - 19,272 SAVI_{max} + 0 EVI_{max}$	0,375	0,208
	Mín.	$55,815 - 86680,047 NDVI_{min} + 27,262 GNDVI_{min} + 57777,818 SAVI_{min} - 0,560 EVI_{min}$	0,284	0,093
2	Media	$65,994 + 24,013 NDVI_{\bar{x}} - 17,705 GNDVI_{\bar{x}} - 3,351 SAVI_{\bar{x}} - 10,437 EVI_{\bar{x}}$	0,167	-0,055
	Máx	$53,696 + 0,048 NDVI_{max} + 4,010 GNDVI_{max} + 1,077 SAVI_{max} + 0 EVI_{max}$	0,154	-0,072
3	Media	$41,687 + 42,218 NDVI_{\bar{x}} - 54,665 GNDVI_{\bar{x}} + 15,415 SAVI_{\bar{x}} + 7,734 EVI_{\bar{x}}$	0,174	-0,046

	Máx	$48,927 - 17,807 NDVI_{max} + 17,435 GNDVI_{max} + 10,219 SAVI_{max} + 0 EVI_{max}$	0,198	-0,016
4	Máx	$60,195 + 6,804 NDVI_{Max} - 13,731 GNDVI_{Max} + 10,238 SAVI_{Max} - 7,589 EVI_{Max}$	0,361	0,190
	Mín.	$57,219 - 51,902 NDVI_{Min} - 3,431 GNDVI_{Min} + 37,854 SAVI_{Min} + 0,117 EVI_{Min}$	0,157	-0,068

En la Figura 32 se presenta la relación entre la productividad real y la productividad estimada para el conjunto de datos evaluado mediante regresión lineal múltiple para el mes mediante los valores máximos de los índices de vegetación NDVI, GNDVI, SAVI y EVI. Se observa cierta dispersión, lo que indica una alta potencialidad para cometer errores en la estimación de la productividad. Las métricas asociadas a este modelo reportaron un RMSE de 1,775 y un MAE de 1,292.

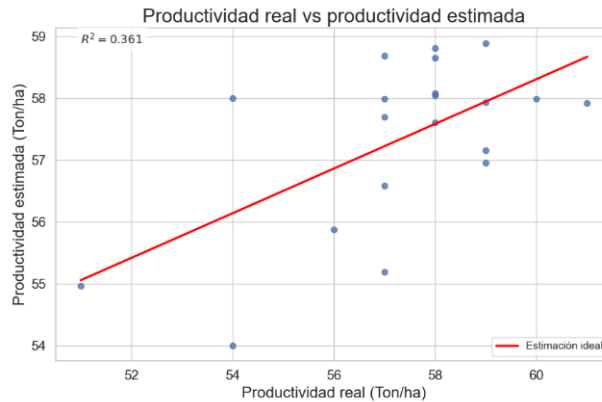


Figura 32. Productividad real vs. estimada – Regresión lineal múltiple (Sentinel-2 + PlanetScope)

7.3.4 PlanetScope + Sentinel-1

En este apartado se presentan los resultados de la regresión lineal múltiple que combina la información óptica de PlanetScope con el índice RVI derivado de Sentinel-1. El objetivo de esta integración es explorar la complementariedad entre los índices espectrales clásicos de vegetación (NDVI, GNDVI, SAVI y EVI) y la respuesta de retrodispersión de radar, con el fin de evaluar si la inclusión de datos SAR permite mejorar el ajuste de los modelos de productividad del cultivo de papa. Los resultados se organizan según el índice empleado en combinación con RVI, mostrando las ecuaciones correspondientes y los valores de R^2 y R^2 ajustado para cada mes. Se realizó la regresión lineal múltiple usando la información de cada uno de los índices y se presentan a continuación los modelos con mejor rendimiento para cada uno de los índices:

7.3.4.1 NDVI+RVI

Los resultados de la combinación entre NDVI de PlanetScope y RVI de Sentinel-1, presentados en la

Tabla 24, evidencian un patrón contrastante a lo largo del ciclo. En los primeros meses, los valores mínimos y medios mostraron desempeños limitados, con ajustes que no superaron un R^2 de 0,132, lo que refleja una capacidad reducida para explicar la variabilidad productiva en fases iniciales. A partir del mes 3 se observa una mejora considerable, especialmente en los valores máximos, donde se alcanzaron coeficientes de determinación de 0,458 y 0,495 en los meses 3 y 4 respectivamente, con R^2 ajustados superiores a 0,390. Estos resultados sugieren que la integración de ambos índices ofrece un mayor poder explicativo en las etapas avanzadas del cultivo, cuando la estructura del dosel está más consolidada. En contraste, los valores mínimos mantuvieron un desempeño más débil, confirmando la menor estabilidad de esta medida frente a la productividad.

Tabla 24. Regresión lineal múltiple NDVI (PlanetScope) y RVI (Sentinel-1).

Mes		Ecuación	R^2	R^2_{Aj}
1	Media	$60,909 - 1,586 RVI - 4,524 NDVI$	0,095	-0,012
	Mínimo	$61,377 - 1,934 * RVI - 7,466 NDVI$	0,120	0,106
2	Media	$47,524 + 0,507 RVI + 12,833 NDVI$	0,132	0,030
	Máximo	$46,733 + 1,137 RVI + 11,383 NDVI$	0,130	0,027
3	Media	$39,647 + 0,985 RVI + 21,587 NDVI$	0,265	0,178
	Máximo	$29,811 + 2,686 RVI + 28,570 NDVI$	0,458	0,394
4	Máximo	$38,69 + 7,914 RVI + 12,262 NDVI$	0,495	0,435
	Mínimo	$61,815 - 3,591 RVI - 5,175 NDVI$	0,138	0,036

En la Figura 33 se presenta la relación entre la productividad real y la productividad estimada mediante un modelo de regresión lineal múltiple, empleando como variables predictoras los valores máximos de los índices de vegetación NDVI (Sentinel-2) y RVI (Sentinel-1) correspondientes al mes 4. Se observa una dispersión moderada en los datos, con una mayor tendencia de los puntos a alinearse en torno a la línea de estimación ideal, lo que sugiere más adecuado en comparación con otros modelos. Las métricas de error asociadas a este modelo corresponden a RMSE de 1,578 y MAE de 1,259.

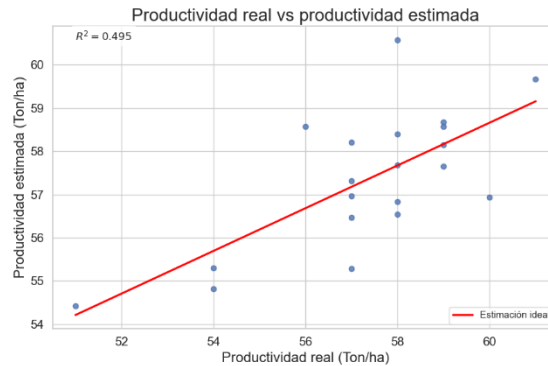


Figura 33. Productividad real vs. estimada – NDVI + RVI (PlanetScope + Sentinel-1).

7.3.4.2 GDVI + RVI

Los resultados de la combinación entre GNDVI de PlanetScope y RVI de Sentinel-1, presentados en la Tabla 25, muestran un comportamiento progresivo a lo largo del ciclo. Durante los primeros meses, los desempeños fueron limitados, con R^2 que no superaron 0,150 en los valores mínimos y 0,072 en los valores medios, lo que refleja baja capacidad explicativa en la etapa inicial del cultivo. A partir del mes 3 se observó una mejora importante, especialmente en los valores máximos, donde se alcanzaron coeficientes de determinación de 0,426 y 0,462 en los meses 3 y 4 respectivamente, con R^2 ajustados cercanos a 0,40. Estos resultados evidencian que la combinación gana solidez en las fases intermedias y finales, cuando la estructura del dosel es más consistente. En contraste, los valores mínimos y medios se mantuvieron con desempeños más bajos, confirmando que los ajustes más representativos se concentran en los extremos máximos.

Tabla 25. Regresión lineal múltiple GNDVI (PlanetScope) y RVI (Sentinel-1).

Mes		Ecuación	R^2	R^2_{Aj}
1	Mínimo	$61,574 - 2,493 RVI - 6,453 GNDVI$	0,150	0,050
	Media	$61,242 - 1,999 RVI - 4,364 GNDVI$	0,072	-0,037
2	Máximo	$48,044 + 1,477 RVI + 10,055 GNDVI$	0,101	-0,004
	Media	$48,253 + 0,585 RVI + 12,391 GNDVI$	0,094	-0,012

3	Máxima	$30,444 + 2,526 RVI + 30,428 GNDVI$	0,426	0,359
	Media	$38,424 + 0,730 RVI + 24,920 GNDVI$	0,235	0,145
4	Máximo	$36,771 + 7,541 RVI + 16,035 GNDVI$	0,462	0,398
	Mínimo	$62,901 - 2,889 RVI - 7,395 GNDVI$	0,163	0,064

En la Figura 34 se presenta la relación entre la productividad real y la productividad estimada mediante un modelo de regresión lineal múltiple, correspondiente al mes 4, utilizando los valores máximos de los índices RVI (Sentinel-1) y GNDVI (PlanetScope). Se observa una dispersión considerable, lo que evidencia una capacidad limitada del modelo para estimar la productividad de manera precisa. Las métricas obtenidas para este modelo fueron un RMSE de 1,629 y un MAE de 1,308.

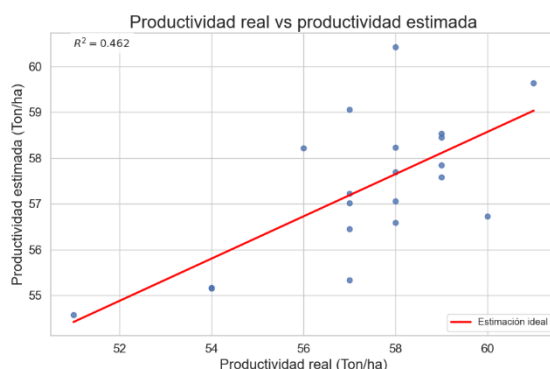


Figura 34. Productividad real vs. estimada – GNDVI + RVI (PlanetScope + Sentinel-1)

7.3.4.3 SAVI+RVI

Los resultados de la combinación entre SAVI de PlanetScope y RVI de Sentinel-1, presentados en la Tabla 26, muestran un patrón de mejora progresiva a lo largo del ciclo. En las fases iniciales, los ajustes fueron limitados, con R^2 de 0,153 en el mínimo y 0,085 en la media durante el mes 1, mientras que en el mes 2 el mejor desempeño apenas alcanzó 0,186 en la media. A partir del mes 3, la combinación comenzó a ganar consistencia, destacándose el valor medio con un R^2 de 0,302 y un R^2 ajustado de 0,220, lo que evidenció mayor capacidad para capturar la variabilidad productiva en la etapa de crecimiento activo. El mejor ajuste se alcanzó en el mes 4 con el valor máximo, donde el R^2 llegó a 0,520 y el R^2 ajustado a 0,463, constituyéndose en la relación más sólida observada para esta combinación. En contraste, los valores mínimos y medios mantuvieron desempeños bajos e inestables, lo que confirma que la capacidad explicativa se concentró principalmente en los valores extremos máximos durante las fases avanzadas del cultivo.

Tabla 26. Regresión lineal múltiple SAVI (PlanetScope) y RVI (Sentinel-1).

Mes		Ecuación	R^2	$R^2 Aj$
1	Mínimo	$60,948 - 1,932 RVI - 4,293 SAVI$	0,153	0,053
	Media	$61,119 - 1,770 RVI - 3,194 SAVI$	0,085	-0,022
2	Media	$18,254 + 0,672 RVI + 39,173 SAVI$	0,186	0,091
	Mínimo	$57,906 - 2,138 RVI + 0,711 SAVI$	0,025	-0,090
3	Media	$-62,283 + 3,672 RVI + 117,668 SAVI$	0,302	0,220
	Mínimo	$58,310 - 4,676 RVI + 1,963 SAVI$	0,120	0,017
4	Máximo	$-4,866 + 6,535 RVI + 55,331 SAVI$	0,520	0,463
	Mínimo	$61,380 - 3,140 RVI - 3,229 SAVI$	0,130	0,028

En la Figura 35 se presenta la comparación entre la productividad real y la productividad estimada para el mes 4, utilizando los valores máximos de los índices SAVI y RVI. Se observa que los puntos se encuentran dispersos, pero en menor medida que en otros modelos, este modelo alcanzó un RMSE de 1,538 y un MAE de 1,226 lo que refleja un desempeño relativamente más estable.

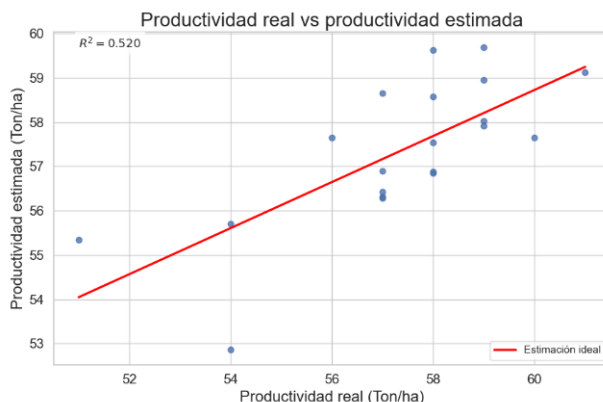


Figura 35. Productividad real vs. estimada – SAVI + RVI (PlanetScope + Sentinel-1)

7.3.4.4 EVI + RVI

Los resultados de la combinación entre EVI de PlanetScope y RVI de Sentinel-1, presentados en la Tabla 27, evidencian un desempeño muy limitado en la estimación de productividad. A lo largo del ciclo, los valores de R^2 fueron bajos, sin superar 0,110 en el mes 3, mientras que en los demás periodos se mantuvieron por debajo de 0,05 e incluso con ajustes negativos. En conjunto, esta combinación se consolidó como la de menor capacidad explicativa, indicando que el EVI integrado con el RVI no logra representar adecuadamente la relación entre el desarrollo del cultivo y la productividad.

Tabla 27. Regresión lineal múltiple EVI (PlanetScope) y RVI (Sentinel-1).

Mes		Ecuación	R^2	R^2_{Aj}
1	Mínimo	$61,383 + 1,212 RVI - 5,320 EVI$	0,086	-0,021
	Media	$66,482 + 0,345 RVI - 9,552 EVI$	0,029	-0,085
2	Mínimo	$61,146 - 1,763 RVI - 3,164 EVI$	0,052	-0,059
	Media	$191,745 + 0,084 RVI - 134,539 EVI$	0,004	-0,113
3	Mínimo	$59,539 - 4,581 RVI + 0,308 EVI$	0,110	0,005
	Media	$197,041 - 2,388 RVI - 137,857 EVI$	0,021	-0,094
4	Mínimo	$59,56 - 3,292 RVI - 0,515 EVI$	0,047	-0,065
	Media	$53,544 + 0,657 RVI + 3,414 EVI$	0,013	-0,104

En la Figura 36 se presenta la comparación entre la productividad real y la productividad estimada para el mes 3 en los valores mínimos de los índices RVI y EVI, se observa que los puntos presentan una alta dispersión, lo que evidencia una capacidad limitada del modelo para estimar la productividad. El modelo alcanzó un RMSE de 2,094 y un MAE de 1,594.

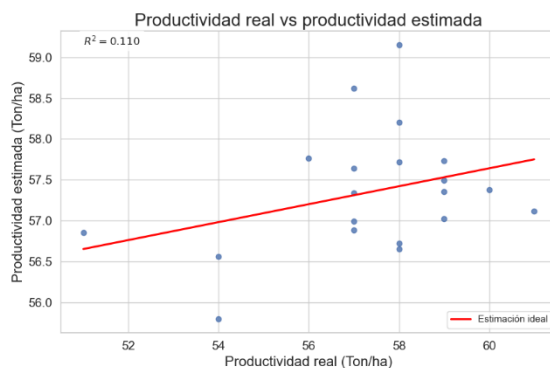


Figura 36. Productividad real vs. estimada – EVI + RVI (PlanetScope + Sentinel-1)

Los resultados de la combinación entre PlanetScope y el índice RVI de Sentinel-1 evidencian un comportamiento heterogéneo según el índice de vegetación y la medida estadística considerada. En general, las combinaciones con NDVI, GNDVI y SAVI mostraron mejores desempeños en los meses 3 y 4, especialmente en los valores máximos, donde los coeficientes de determinación superaron 0,426 y alcanzaron hasta 0,520, con R^2 ajustados superiores a 0,39. Esto refleja una mayor capacidad explicativa en las etapas intermedias y finales del cultivo, cuando la biomasa es más abundante y el radar capta mejor los contrastes estructurales.

En las fases iniciales, en cambio, los valores mínimos y medios registraron desempeños bajos, con R^2 generalmente inferiores a 0,15 e incluso con R^2 ajustados negativos, lo que indica baja estabilidad y escasa capacidad predictiva. Entre las combinaciones evaluadas, SAVI + RVI alcanzó el mayor ajuste con un R^2 de 0,520 y un R^2 ajustado de 0,463 en el mes 4, mientras que EVI + RVI fue la menos favorable, sin superar un R^2 de 0,110 en todo el ciclo.

En conjunto, la integración óptica–radar mostró potencial para mejorar la estimación de la productividad, aunque con limitaciones en los valores medios y mínimos y en las fases tempranas, lo que resalta la necesidad de explorar modelos más complejos o incorporar nuevas variables que fortalezcan la capacidad explicativa.

7.3.5 Sentinel-2 + Sentinel-1

En este apartado se presentan los resultados de los modelos de regresión lineal múltiple que combinan los índices ópticos de Sentinel-2 con el índice de radar RVI de Sentinel-1. Se evalúan las parejas NDVI + RVI, GNDVI + RVI, SAVI + RVI y EVI + RVI, reportando por mes y por medida estadística media, máximo y mínimo los valores de R^2 y R^2 ajustado.

7.3.5.1 NDVI + RVI

Los resultados de la combinación NDVI de Sentinel-2 con el RVI de Sentinel-1, presentados en la Tabla 28, evidencian un comportamiento irregular a lo largo del ciclo del cultivo. En la etapa final, el valor máximo destacó con un R^2 de 0,247 y un R^2 ajustado de 0,159, lo que sugiere una relación moderada y relativamente consistente entre el índice y la productividad. En las fases intermedias, los valores mínimos ofrecieron cierta capacidad explicativa, con R^2 de 0,149 en el mes 2 y 0,133 en el mes 3, mientras que los promedios resultaron prácticamente nulos. Estos resultados indican que, dentro de esta combinación, los

valores extremos aportan más información que los medios, y que la señal se vuelve más representativa en los momentos de consolidación y maduración del cultivo.

Tabla 28. Regresión lineal múltiple NDVI (Sentinel-2) y RVI (Sentinel-1).

Mes		Ecuación	R^2	R^2_{Aj}
1	Máximo	$55,506 + 2,080 RVI - 1,158 NDVI$	0,076	-0,033
	Mínimo	$58,301 - 1,337 RVI - 3,910 NDVI$	0,060	-0,050
2	Mínimo	$58,839 - 1,125 RVI - 6,045 NDVI$	0,149	0,049
	Máximo	$55,679 + 1,962 RVI - 1,245 NDVI$	0,071	-0,038
3	Mínimo	$60,462 - 5,131 RVI - 2,527 NDVI$	0,133	0,031
	Media	$59,284 - 2,304 RVI - 0,391 NDVI$	0,016	-0,099
4	Máximo	$50,256 + 5,601 RVI + 1,994 NDVI$	0,247	0,159
	Mínimo	$59,295 - 3,367 RVI - 0,832 NDVI$	0,048	-0,064

En la Figura 37 se presenta la relación entre la productividad real y la productividad estimada a partir de un modelo de regresión lineal múltiple, utilizando como variables predictoras los valores máximos de los índices NDVI (Sentinel-2) y RVI (Sentinel-1) correspondientes al mes 4. Se observa una dispersión considerable, lo que indica una capacidad limitada del modelo para estimar la productividad. Se obtuvo un RMSE de 1,926 y un MAE de 1,437.

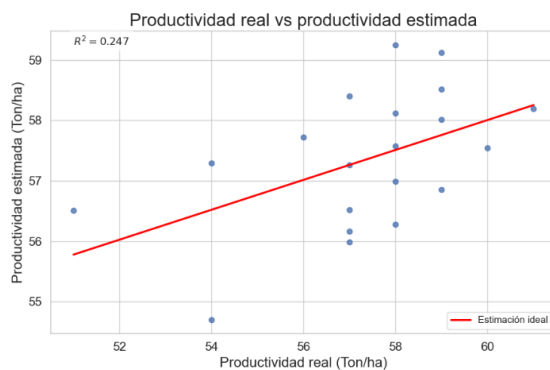


Figura 37. Productividad real vs. estimada – NDVI + RVI (Sentinel-2 + Sentinel-1)

7.3.5.2 GNDVI+RVI

Los resultados de la combinación GNDVI de Sentinel-2 con el RVI de Sentinel-1, presentados en la Tabla 29, muestran un desempeño progresivo a lo largo del ciclo. El ajuste más consistente se alcanzó en la fase de maduración, cuando el valor máximo registró un R^2 de 0,226 y un R^2 ajustado de 0,134, reflejando una relación más estable entre el índice y la productividad. En etapas intermedias, el valor mínimo del mes 3 aportó una señal moderada con R^2 de 0,145, mientras que los meses iniciales evidenciaron desempeños débiles en todas las medidas. Este patrón confirma que los valores máximos adquieren mayor relevancia en las fases de mayor desarrollo del cultivo, mientras que los mínimos y las medias presentan una capacidad predictiva reducida en los estadios tempranos.

Tabla 29. Regresión lineal múltiple GNDVI (Sentinel-2) y RVI (Sentinel-1).

Mes		Ecuación	R^2	R^2_{Aj}
1	Máximo	$55,374 + 2,130 RVI - 0,952 GNDVI$	0,071	-0,037
	Mínimo	$58,352 - 1,544 RVI - 2,869 GNDVI$	0,044	-0,068
2	Mínimo	$58,856 - 1,279 RVI - 5,929 GNDVI$	0,136	0,034

	Máximo	$55,670 + 1,975 RVI - 1,436 GNDVI$	0,074	-0,035
3	Mínimo	$60,534 - 5,174 RVI - 3,152 GNDVI$	0,145	0,045
	Máximo	$53,697 + 2,798 RVI + 1,516 GNDVI$	0,104	-0,001
4	Máximo	$50,333 + 5,880 RVI + 1,151 GNDVI$	0,226	0,134
	Mínimo	$59,346 - 3,391 RVI - 0,971 GNDVI$	0,049	-0,062

En la Figura 38 se presenta la relación entre la productividad real y la productividad estimada a partir de un modelo de regresión lineal múltiple, empleando como variables predictoras los valores máximos de los índices GNDVI (Sentinel-2) y RVI (Sentinel-1) en el mes 4. Se observa dispersión de los datos, lo que refleja una capacidad limitada del modelo para representar adecuadamente la productividad. Las métricas de error obtenidas para este ajuste fueron un RMSE de 1,953 y un MAE de 1.476.

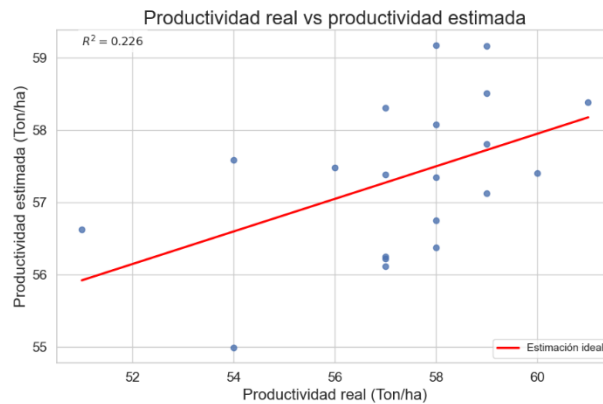


Figura 38. Productividad real vs. estimada – GNDVI + RVI (Sentinel-2 + Sentinel-1).

7.3.5.3 SAVI+RVI

Los resultados de la combinación del SAVI de Sentinel-2 con el RVI de Sentinel-1, consignados en la Tabla 30, muestran desempeños modestos a lo largo del ciclo. El mejor ajuste se registró en el mes 4 con un R^2 de 0,159 en el valor máximo, seguido por el mes 2 con un R^2 de 0,149 en el valor mínimo. Estos hallazgos indican que la maduración y el crecimiento intermedio ofrecen cierta capacidad explicativa, principalmente en medidas extremas, mientras que en las fases iniciales y en valores medios los ajustes fueron bajos e inestables, limitando su utilidad.

Tabla 30. Regresión lineal múltiple SAVI (Sentinel-2) y RVI (Sentinel-1).

Mes		Ecuación	R^2	$R^2 Aj$
1	Máximo	$55,506 + 2,080 RVI - 0,772 SAVI$	0,076	-0,033
	Mínimo	$58,301 - 1,337 RVI - 2,607 SAVI$	0,060	-0,050
2	Mínimo	$58,839 - 1,125 RVI - 4,031 SAVI$	0,149	0,049
	Máximo	$55,680 + 1,963 RVI - 0,842 SAVI$	0,071	-0,038
3	Mínimo	$60,462 - 5,131 RVI - 1,685 SAVI$	0,133	0,031
	Máximo	$53,632 + 2,797 RVI + 0,992 SAVI$	0,105	0,000
4	Máximo	$50,256 + 5,601 RVI + 1,329 SAVI$	0,159	0,247
	Mínimo	$59,295 - 3,367 RVI - 0,554 SAVI$	0,048	-0,064

En la se presenta la comparación entre la productividad real y la productividad estimada para el mes 4 en sus valores máximos, se observa una dispersión considerablemente alta de los puntos respecto a la línea de estimación ideal. El modelo registró un RMSE de 1,926 y un MAE de 1,437.

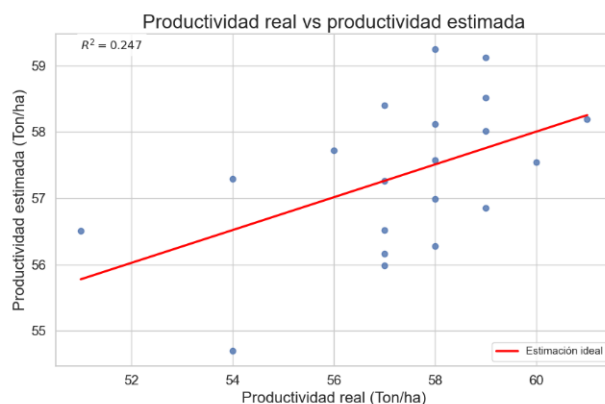


Figura 39. Productividad real vs. estimada – SAVI + RVI (Sentinel-2 + Sentinel-1).

7.3.5.4 EVI+RVI

Los resultados de la combinación del EVI de Sentinel-2 con el RVI de Sentinel-1, presentados en la Tabla 31, reflejan una tendencia de mejora hacia el final del ciclo. El mejor ajuste se alcanzó en el mes 4 con el valor máximo, que registró un R^2 de 0,219 y un R^2 ajustado de 0,127, indicando una asociación más consistente en la etapa de maduración. En el mes 3, el valor mínimo alcanzó un R^2 de 0,118, mientras que los meses 1 y 2 mostraron desempeños modestos en todas las medidas, con valores de R^2 por debajo de 0,102. En conjunto, la combinación se mostró más informativa en los valores máximos durante fases avanzadas del cultivo, mientras que mínimos y medias ofrecieron escasa capacidad explicativa en las etapas iniciales.

Tabla 31. Regresión lineal múltiple EVI (Sentinel-2) y RVI (Sentinel-1).

Mes		Ecuación	R^2	$R^2 Aj$
1	Media	$59,302 - 2,010 RVI - 1,453 EVI$	0,102	-0,004
	Mínimo	$58,314 - 2,029 RVI - 0,917 EVI$	0,099	-0,007
2	Máximo	$54,998 + 1,964 RVI + 0,397 EVI$	0,064	-0,046
	Mínimo	$58,376 - 2,137 RVI - 0,340 EVI$	0,038	-0,075
3	Mínimo	$59,581 - 4,262 RVI - 0,266 EVI$	0,118	0,015
	Máximo	$54,142 + 2,622 RVI + 0,387 EVI$	0,080	-0,028
4	Máximo	$50,363 + 6,463 RVI - 0,286 EVI$	0,219	0,127
	Mínimo	$59,220 - 3,369 RVI - 0,789 EVI$	0,129	0,027

En la Figura 40 se presenta la comparación entre productividad real y productividad estimada para el mes 4 en sus valores máximos. Se observa una alta dispersión, lo que evidencia una baja capacidad del modelo para estimar productividad de cultivos de papa. El modelo registró un RMSE de 1,962 y un MAE de 1,492

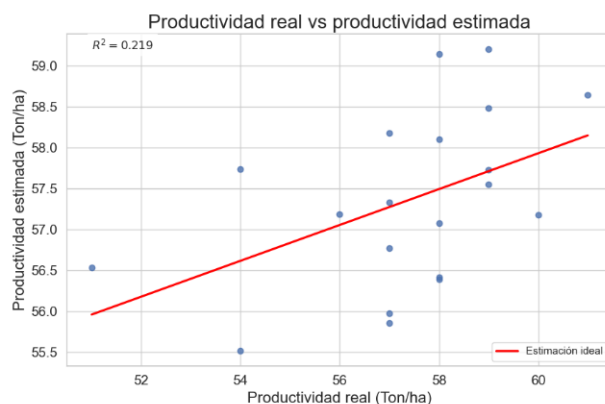


Figura 40. Productividad real vs. estimada – EVI + RVI (Sentinel-2 + Sentinel-1).

Los resultados de la combinación de la información óptica de Sentinel-2 con el índice RVI de Sentinel-1 mostraron un comportamiento diferenciado a lo largo del ciclo del cultivo. En las etapas intermedias y de maduración (meses 3 y 4) se alcanzaron los ajustes más relevantes, con valores máximos de NDVI, GNDVI y EVI que lograron coeficientes de determinación superiores a 0,219 y hasta 0,247, con R^2 ajustados cercanos a 0,16. Esto sugiere que, cuando el dosel se encuentra consolidado, la interacción entre la señal óptica y la de radar permite capturar de manera más efectiva las variaciones estructurales de la biomasa asociadas a la productividad.

Por el contrario, en las fases iniciales predominó una baja capacidad explicativa: los valores mínimos y medios registraron R^2 reducidos e incluso ajustados negativos, lo que refleja la dificultad de ambos sensores para representar la heterogeneidad del cultivo en los estadios de establecimiento. Entre las combinaciones evaluadas, NDVI + RVI y GNDVI + RVI mostraron mayor consistencia en sus ajustes, mientras que SAVI + RVI evidenció menor estabilidad y valores más discretos en todo el ciclo.

En conjunto, los resultados confirman que la integración de Sentinel-2 con Sentinel-1 aporta mejoras puntuales en la estimación de la productividad, particularmente al considerar los valores máximos de los índices en fases avanzadas del cultivo. Sin embargo, la utilidad de esta integración se reduce en los valores mínimos y medios, lo que señala la necesidad de explorar modelos más robustos o la inclusión de variables complementarias para captar de mejor manera las dinámicas productivas en etapas tempranas.

7.3.6 S2 relleno con PS y combinado con S1

La combinación de la información de Sentinel-2, complementada con PlanetScope en las áreas cubiertas por nubes y posteriormente integrada con Sentinel-1, se evaluó a través de los índices NDVI, GNDVI, SAVI y EVI junto con el RVI. Los resultados se organizaron considerando los valores mínimos, medios y máximos de cada índice en los diferentes meses de análisis, con el fin de identificar los patrones de ajuste y capacidad explicativa de cada combinación.

7.3.6.1 NDVI+RVI

Los resultados de la combinación del NDVI de Sentinel-2 + PlanetScope con el RVI de Sentinel-1, presentados en la Tabla 32, evidencian un desempeño limitado en los primeros meses, donde los valores máximos y mínimos apenas alcanzaron R^2 de 0,066 y ajustes negativos en el R^2 ajustado. En el mes 2 se observó una ligera mejoría, destacando el valor máximo con R^2 de 0,130, aunque sin estabilidad suficiente.

En el mes 3, tanto el mínimo como el máximo se ubicaron en torno a 0,120, mostrando un incipiente aumento en la capacidad explicativa. El mayor ajuste se registró en el mes 4 con el valor máximo, que alcanzó un R^2 de 0,401 y un R^2 ajustado de 0,330, lo que confirma que esta combinación adquiere mayor relevancia en la fase de maduración, mientras que en etapas tempranas mantiene un aporte reducido.

Tabla 32. Regresión lineal múltiple NDVI (Sentinel-2+PlanetScope) y RVI (Sentinel-1).

Mes	NDVI	Ecuación	R^2	R^2_{Aj}
1	Mínimo	$58,797 - 1,123 RVI - 3,354 NDVI$	0,066	-0,043
	Máximo	$55,088 + 2,333 RVI - 0,277$	0,066	-0,044
2	Máximo	$52,017 + 1,843 RVI + 4,809 NDVI$	0,130	0,0277
	Mínimo	$58,910 - 1,596 RVI - 1,923 NDVI$	0,041	-0,071
3	Mínimo	$59,535 - 5,128 RVI + 1,440 NDVI$	0,119	0,015
	Máximo	$52,780 + 1,869 RVI + 3,372 NDVI$	0,117	0,013
4	Máximo	$47,199 + 6,362 RVI + 4,861 NDVI$	0,401	0,330
	Mínimo	$59,243 - 3,032 RVI - 0,873 NDVI$	0,050	-0,061

En la Figura 41 se presenta la relación entre la productividad real y la productividad estimada a partir de un modelo de regresión lineal múltiple, empleando como variables predictoras los valores máximos de los índices RVI (Sentinel-1) y NDVI (Sentinel-2 + PlanetScope) en el mes 4. Se observa una alta dispersión de los datos, lo que evidencia una baja capacidad del modelo para realizar estimaciones precisas de la productividad.

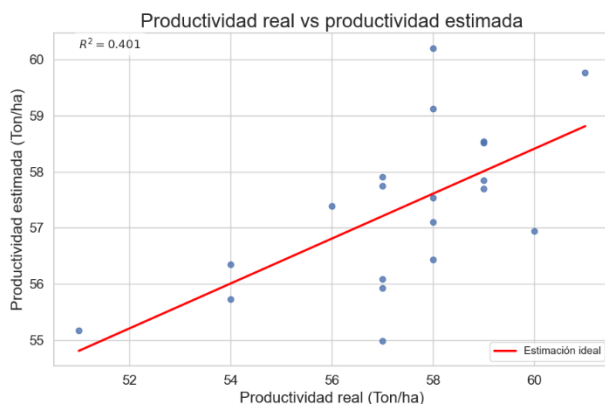


Figura 41. Productividad real vs. estimada – NDVI + RVI (Sentinel-2 + PlanetScope + Sentinel-1)

7.3.6.2 GNDVI+RVI

Los resultados de la combinación del GNDVI de Sentinel-2 + PlanetScope con el RVI de Sentinel-1, presentados en la

Tabla 33, muestran un desempeño limitado en los primeros meses, donde los valores se mantuvieron bajos y sin estabilidad, con R^2 que no superaron 0,065 y ajustes negativos en el R^2 ajustado. En el mes 2 se observó un incremento en el valor máximo, que alcanzó un R^2 de 0,136 con un R^2 ajustado de 0,035, lo que marca una leve mejora, aunque todavía con baja capacidad explicativa. El mes 3 mantuvo una tendencia similar, con el mínimo en 0,123 y el máximo en 0,119, ambos con ajustes moderados. La fase más destacada se dio en el mes 4, cuando el valor máximo alcanzó un R^2 de 0,327 con un R^2 ajustado de 0,248, confirmando que la combinación adquiere mayor relevancia hacia la etapa de maduración, mientras que los valores mínimos se mantuvieron débiles y poco útiles para explicar la variabilidad productiva.

Tabla 33. Regresión lineal múltiple GNDVI (Sentinel-2+PlanetScope) y RVI (Sentinel-1).

Mes	GNDVI	Ecuación	R^2	R^2_{Aj}
1	Máximo	$55,025 + 2,337 RVI - 0,177 GNDVI$	0,065	-0,045
	Mínimo	$58,858 - 1,443 RVI - 2,704 GNDVI$	0,050	-0,062
2	Máximo	$51,970 + 1,911 RVI + 5,154 GNDVI$	0,136	0,035
	Mínimo	$58,929 - 1,566 RVI - 1,985 GNDVI$	0,043	-0,070
3	Mínimo	$59,513 - 5,213 RVI + 1,594 GNDVI$	0,123	0,019
	Máximo	$52,838 + 1,825 RVI + 3,666 GNDVI$	0,119	0,015
4	Máximo	$47,731 + 6,353 RVI + 4,235 GNDVI$	0,327	0,248
	Mínimo	$59,425 - 2,866 RVI - 1,544 GNDVI$	0,058	-0,052

En la Figura 42 se presenta la relación entre la productividad real y la productividad estimada a partir de un modelo de regresión lineal múltiple, utilizando como variables predictoras los valores máximos de los índices RVI (Sentinel-1) y GNDVI (Sentinel-2 + PlanetScope) en el mes 4. Los resultados muestran una alta dispersión de los valores, lo que evidencia una baja capacidad del modelo para estimar la productividad. Las métricas de error obtenidas reportaron un RMSE de 1,820 y un MAE de 1,520.

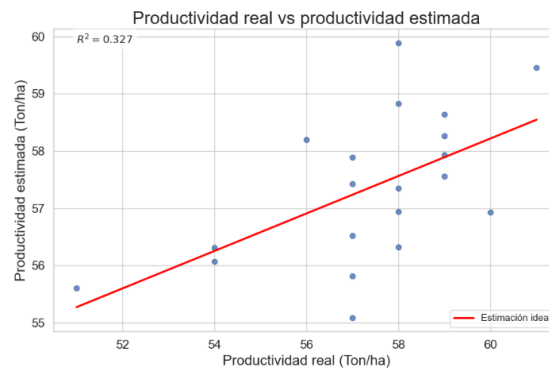


Figura 42. Productividad real vs. estimada – GNDVI + RVI (Sentinel-2 + PlanetScope + Sentinel-1)

7.3.6.3 SAVI+RVI

Los resultados de la combinación del SAVI de Sentinel-2 + PlanetScope con el RVI de Sentinel-1, presentados en la Tabla 34, muestran un comportamiento discreto en los dos primeros meses, con valores de R^2 por debajo de 0,110 y R^2 ajustados cercanos o inferiores a cero, lo que evidencia una capacidad explicativa limitada en las etapas iniciales. A partir del mes 3 se observó una mejora clara, destacando tanto en el valor máximo con un R^2 de 0,189 y un R^2 ajustado de 0,094, como en el mínimo con 0,130 y 0,028 respectivamente, lo que sugiere que el índice comenzó a captar mejor la dinámica del dosel en pleno crecimiento. El mejor desempeño se alcanzó en el mes 4, cuando el valor máximo logró un R^2 de 0,488 con un R^2 ajustado de 0,428, constituyéndose en el ajuste más sólido de esta combinación y evidenciando una relación consistente entre el vigor del cultivo y la productividad en la etapa de maduración, mientras que los valores mínimos permanecieron sin relevancia.

Tabla 34. Regresión lineal múltiple SAVI (Sentinel-2+PlanetScope) y RVI (Sentinel-1).

Mes	SAVI	Ecuación	R^2	R^2_{Aj}
1	Máximo	$56,319 + 2,443 RVI - 1,947 SAVI$	0,090	-0,017
	Mínimo	$58,797 - 1,123 RVI - 2,236 SAVI$	0,067	-0,043
2	Máximo	$50,938 + 1,717 RVI + 4,837 SAVI$	0,110	0,006
	Mínimo	$58,949 - 1,558 RVI - 1,393 SAVI$	0,044	-0,069

3	Máximo	$47,999 + 1,813 RVI + 7,675 SAVI$	0,189	0,094
	Mínimo	$59,346 - 5,361 RVI + 1,489 SAVI$	0,130	0,028
4	Máximo	$45,186 + 6,546 RVI + 5,745 SAVI$	0,488	0,428
	Mínimo	$59,179 - 3,117 RVI - 0,347 SAVI$	0,047	-0,064

En la Figura 43 se presenta la comparación entre la productividad real y la productividad estimada a partir de un modelo de regresión lineal múltiple aplicado a los valores máximos de los índices RVI (Sentinel-1) y SAVI (Sentinel-2 + PlanetScope) en el mes 4. Se observa una dispersión moderada de los datos, lo que evidencia un margen de error considerable en la estimación de la productividad. Las métricas asociadas a este modelo reportaron un RMSE de 1,588 y un MAE de 1,330.

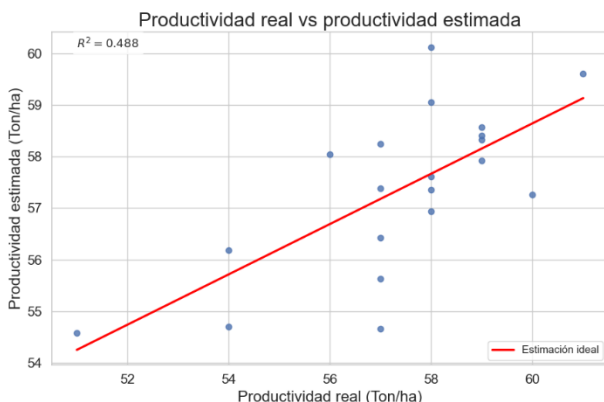


Figura 43. Productividad real vs. estimada – SAVI + RVI (Sentinel-2 + PlanetScope + Sentinel-1).

7.3.6.4 EVI+RVI

Los resultados de la combinación EVI de Sentinel-2 con el RVI de Sentinel-1, presentados en la Tabla 35, muestran un desempeño irregular a lo largo del ciclo. Durante los primeros meses los valores fueron bajos y sin estabilidad, con R^2 cercanos a cero y varios ajustes negativos. Una ligera mejora se observó en el mes 3 en el valor mínimo, donde el R^2 alcanzó 0,167 y el R^2 ajustado 0,069, lo que sugiere un inicio de mayor correspondencia con la productividad. El avance más significativo se registró en la fase final, con el valor máximo del mes 4, que obtuvo un R^2 de 0,369 y un R^2 ajustado de 0,295, constituyendo el ajuste más alto dentro de esta combinación. En conjunto, aunque los resultados son limitados en la mayoría de los escenarios, los valores máximos en etapas avanzadas aportaron mayor capacidad explicativa.

Tabla 35. Regresión lineal múltiple EVI (Sentinel-2+PlanetScope) y RVI (Sentinel-1).

Mes	EVI	Ecuación	R^2	$R^2 Aj$
1	Media	$59,563 - 0,832 RVI - 1,790 EVI$	0,028	-0,086
	Mínimo	$58,076 - 1,410 RVI - 0,142 EVI$	0,017	-0,098
2	Mínimo	$58,963 - 1,925 RVI - 0,944 EVI$	0,112	0,007
	Media	$67,859 - 1,058 RVI - 9,984 EVI$	0,084	-0,023
3	Mínimo	$60,462 - 6,888 RVI + 0,920 EVI$	0,167	0,069
	Media	$54,916 - 1,972 RVI + 4,029 EVI$	0,018	-0,098
4	Máximo	$-59,008 + 6,725 RVI + 109,034 EVI$	0,369	0,295
	Mínimo	$59,067 - 3,204 RVI + 0,007 EVI$	0,046	-0,066

En la Figura 44 se presenta la comparación entre productividad real y productividad estimada. Se observa una dispersión considerable de los valores, lo que indica que, si bien es posible estimar la productividad, el

modelo presenta un nivel de error significativo en las predicciones. En este caso, el cálculo de las métricas arrojó un RMSE de 1,764 y un MAE de 1,297 t/ha.

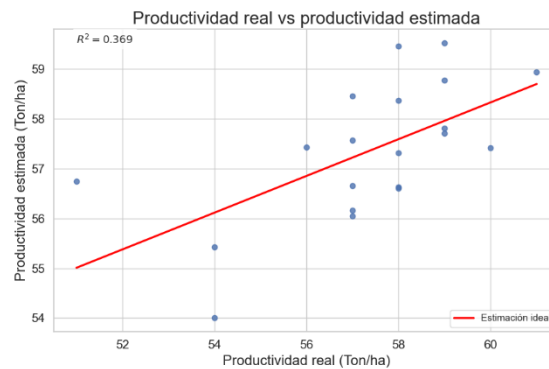


Figura 44. Productividad real vs. estimada – EVI + RVI (Sentinel-2 + PlanetScope + Sentinel-1)

Los resultados de la combinación entre los índices ópticos de Sentinel-2, rellenados con PlanetScope, y el índice de radar RVI de Sentinel-1 muestran una tendencia heterogénea a lo largo de los meses evaluados. En términos generales, los valores correspondientes a los primeros meses fueron bajos, con coeficientes de determinación que en la mayoría de los casos no superaron el 0,15 y con R^2 ajustados frecuentemente negativos, lo que refleja poca capacidad explicativa y estabilidad limitada en las etapas iniciales del ciclo del cultivo.

A partir del mes 3 se observa un cambio en el comportamiento de todas las combinaciones. Aunque los valores aún eran moderados, algunos índices como SAVI y NDVI comenzaron a registrar desempeños más consistentes, con R^2 que se acercaron o superaron el 0,12 y R^2 ajustados positivos en ciertos casos. Esta evolución evidencia una mejora en la capacidad predictiva cuando el cultivo entra en fases de mayor desarrollo vegetativo, en las que los contrastes estructurales de la biomasa pueden ser captados con mayor efectividad por la combinación de la información óptica y de radar.

El mes 4 representó el punto más favorable de la serie, especialmente en las combinaciones que incluyeron SAVI y NDVI. En particular, la combinación SAVI + RVI alcanzó un R^2 de 0,488 y un R^2 ajustado de 0,428 en el valor máximo, constituyéndose en la relación con mejor ajuste dentro de este conjunto. También el NDVI + RVI presentó un desempeño destacable, con un R^2 de 0,401 y un R^2 ajustado de 0,330, mientras que GNDVI + RVI y EVI + RVI mostraron desempeños intermedios, con R^2 máximos de 0,327 y 0,369 respectivamente. Estos resultados sugieren que la combinación de SAVI y NDVI con la información radar es más adecuada para explicar la productividad en etapas avanzadas del cultivo.

En conjunto, los hallazgos muestran que la integración de información óptica de Sentinel-2, complementada con PlanetScope para cubrir zonas nubladas, y el radar de Sentinel-1 tiene potencial para la estimación de la productividad de cultivos de papa, particularmente cuando se emplean valores máximos de índices de vegetación en fases avanzadas del ciclo. Sin embargo, las limitaciones evidenciadas en los primeros meses y en los valores mínimos y medios señalan la necesidad de explorar enfoques metodológicos más complejos que permitan capturar con mayor precisión la variabilidad productiva en todas las etapas del desarrollo del cultivo de papa.

7.3.7 Modelos de regresión múltiple con PlanetScope, Sentinel-2, y Sentinel-1.

Se desarrollaron modelos de regresión lineal múltiple empleando diferentes combinaciones de índices espectrales (NDVI, GNDVI, SAVI y EVI) e índices de radar (VH, VV, VH/VV y RVI), a partir de datos ópticos (PlanetScope y Sentinel-2) y de radar (Sentinel-1). Estos modelos se ajustaron tanto de manera individual para cada sensor como mediante la integración de información, destacando el proceso de relleno de nubes en Sentinel-2 con imágenes de PlanetScope. Los resultados evidenciaron que el mejor desempeño se obtuvo con la combinación de índices provenientes de PlanetScope y Sentinel-2 (con relleno de nubes) durante el mes 4 en sus valores máximos, alcanzando un R^2 de 0,707, un R^2 ajustado de 0,493 y un coeficiente de correlación múltiple de 0,841.

La superioridad del modelo (ecuación (14)) puede explicarse por la complementariedad entre la información óptica y de radar. Mientras que los índices derivados de PlanetScope y Sentinel-2 capturan la actividad fotosintética y el vigor de la vegetación (NDVI, SAVI, GNDVI, EVI), los parámetros provenientes de Sentinel-1 (VH, VV, VH/VV y RVI) incorporan la estructura del dosel, aspectos que resultan relevantes en las etapas intermedias del ciclo fenológico. Sin embargo, su aplicación implica la necesidad de disponer de un conjunto más amplio de índices y variables, lo cual puede aumentar los requerimientos de procesamiento.

$$-1,25 + 6,88\text{NDVI} - 15,01\text{ GNDVI} + 9,94\text{ SAVI} + 26,53\text{ EVI} + 45,35\text{ VH} + 0,05\text{ VV} - 126,21\text{ VH/VV} + 67,94\text{ RVI} \quad (14)$$

En la Figura 45 se presenta la relación entre la productividad real y la productividad estimada a partir de un modelo de regresión lineal múltiple, utilizando como variables predictoras mencionadas para el modelo con mejor rendimiento. Se observa una dispersión moderada, aunque con una mayor tendencia de los puntos a alinearse en torno a la línea de estimación ideal, lo que refleja un ajuste más consistente en comparación con otros modelos. Las métricas de error asociadas a este modelo corresponden a un RMSE de 1,202 y un MAE de 0,926.

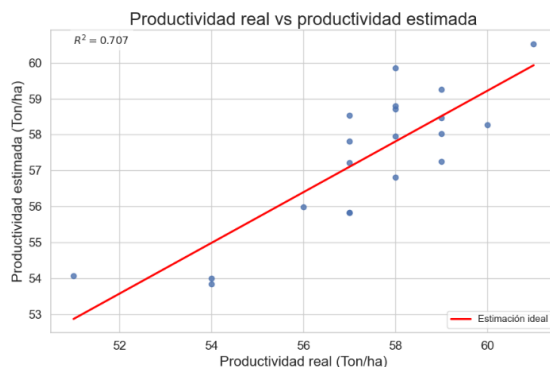


Figura 45. Productividad real vs. estimada – todos los índices (Sentinel-2 + PlanetScope + Sentinel-1)

Adicionalmente, a partir de la combinación de información de los sensores mencionados se obtuvo un modelo representado por la ecuación (15), con un R^2 de 0,639, un R^2 Ajustado de 0,472 y un coeficiente de correlación múltiple de 0,799, correspondiente a los valores máximo de los índices para el mes 4.

$$14,32 + 9,04\text{ NDVI} - 15,16\text{ GNDVI} + 9,31\text{ SAVI} + 17,35\text{ EVI} - 90,56\text{ VH/VV} + 50,77\text{ RVI} \quad (15)$$

De manera complementaria, se obtuvo un modelo basado en los valores máximos de los índices para el mes 4, en el que se consideraron dos sensores ópticos (NDVI y GNDVI) y dos derivados de radar (VH/VV y

RVI). Este modelo alcanzó R^2 de 0,530, un R^2 Ajustado de 0,405 y un coeficiente de correlación múltiple de 0,728, tal como se presenta en la ecuación (16).

$$31,579 + 18,186 \text{ NDVI} - 14,901 \text{ GNDVI} - 106,471 \text{ VH/VV} + 58,184 \text{ RVI} \quad (16)$$

Por otra parte, al considerar únicamente los índices extraídos de PlanetScope el mejor modelo identificado se corresponde con la ecuación (17), también ajustada a los valores máximos en el mes 4. Este modelo alcanzó un R^2 de 0,691, un R^2 Ajustado de 0,466 y un coeficiente de correlación múltiple de 0,831.

$$-26,09 + 7,43 \text{ NDVI} - 12,57 \text{ GNDVI} + 56,975 \text{ SAVI} + 46,29 \text{ VH} + 1,90 \text{ VV} - 135,38 \text{ VH/VV} + 72,65 \text{ RVI} \quad (17)$$

Asimismo, se identificaron modelos derivados de PlanetScope que presentaron un buen desempeño utilizando un número reducido de índices. Entre ellos, se destacan los representados en las ecuaciones (18) y (19). El modelo de la ecuación (18) alcanzó un R^2 de 0,62, un R^2 Ajustado de 0,443 y un coeficiente de correlación múltiple de 0,79, mientras que el de la ecuación (19) obtuvo un R^2 de 0,532, un R^2 Ajustado de 0,408 y un coeficiente de correlación múltiple de 0,73.

$$-8,34 + 14,45 \text{ NDVI} - 12,31 \text{ GNDVI} + 40,83 \text{ SAVI} - 96,08 \text{ VH/VV} + 54,24 \text{ RVI} \quad (18)$$

$$29,19 + 24,082 \text{ NDVI} - 15,745 \text{ GNDVI} - 76,176 \text{ VH/VV} + 45,368 \text{ RVI} \quad (19)$$

En el caso de los índices extraídos de Sentinel-2, el mejor modelo se obtuvo a partir de los valores máximos correspondientes al mes 4, alcanzando un(20) R^2 de 0,690, un R^2 Ajustado de 0,464 y un coeficiente de correlación múltiple de 0,83, como se observa en la ecuación (20).

$$33,81 + 152202,77 \text{ NDVI} - 25,0 \text{ GNDVI} - 101454,6 \text{ SAVI} - 3,72 \text{ EVI} + 46,09 \text{ VH} + 0,78 \text{ VV} - 103,84 \text{ VH/VV} + 54,99 \text{ RVI} \quad (20)$$

Finalmente, se identificaron modelos adicionales con un menor número de variables, los cuales obtuvieron desempeños intermedios. La ecuación (21)(21) registró un R^2 de 0,616, un R^2 Ajustado de 0,438 y un coeficiente de correlación múltiple de 0,785, mientras que la ecuación (22) alcanzó un R^2 de 0,532, un R^2 Ajustado de 0,408 y un coeficiente de correlación múltiple de 0,730.

$$40,10 + 130684,35 \text{ NDVI} - 39,44 \text{ GNDVI} - 87100,98 \text{ SAVI} - 2,31 \text{ EVI} - 71,97 \text{ VH/VV} + 39,73 \text{ RVI} \quad (21)$$

$$37,66 + 41,34 \text{ NDVI} - 43,56 \text{ GNDVI} - 75,62 \text{ VH/VV} + 43,65 \text{ RVI} \quad (22)$$

Los modelos de regresión lineal múltiple mostraron un mejor desempeño cuando se integró de manera simultánea información óptica y de radar. Los ajustes más altos se obtuvieron en los meses 3 y 4, coincidiendo con la fase de maduración del cultivo, periodo en el que la estructura del dosel es más completa y genera señales espectrales y de retrodispersión más estables y representativas del estado del cultivo. En cambio, los valores mínimos y medios registrados en los meses tempranos presentaron desempeños más inestables, lo que evidencia la limitada capacidad de estos índices para describir la productividad cuando la cobertura foliar aún es incipiente.

7.4 Modelo de regresión Random Forest

En este apartado se presentan los resultados obtenidos al aplicar el modelo de regresión Random Forest para la estimación de la productividad de los cultivos de papa. Al igual que en los apartados anteriores, el modelo se construyó empleando los índices de vegetación calculados a partir de las imágenes satelitales, considerando distintas fuentes de información (PlanetScope, Sentinel-2 y combinaciones de ambas). Se

evaluaron diferentes valores estadísticos de los índices (media, mínimo y máximo) en los meses de seguimiento 2 y 3, con el objetivo de describir el desempeño del modelo en cada escenario.

Los resultados se expresan mediante tres métricas: el coeficiente de determinación (R^2), que refleja el grado de ajuste del modelo; la raíz del error cuadrático medio (RMSE), que indica la magnitud promedio del error en las predicciones; y el error absoluto medio (MAE), que representa el promedio de las diferencias absolutas entre los valores estimados y los reales. A continuación, se presentan los resultados organizados por subapartados según la fuente de datos utilizada.

7.4.1 PlanetScope.

Los resultados del modelo Random Forest aplicado a PlanetScope, presentados en la Tabla 36, muestran desempeños limitados en general. Los valores medios y mínimos no aportaron capacidad explicativa, con R^2 negativos y errores elevados, que en el caso de los mínimos superaron las 2 toneladas por hectárea de RMSE en ambos meses. En contraste, los valores máximos ofrecieron un mejor ajuste relativo, especialmente en el mes 3, donde el R^2 alcanzó $-0,819$ acompañado de un RMSE de 1,118 y un MAE de 1,023. Aunque los ajustes siguen siendo bajos, este comportamiento indica que los valores extremos máximos de los índices concentran un mayor potencial para capturar la relación con la productividad, mientras que las medias y mínimos resultaron poco informativos.

Tabla 36. Modelo Regresión Random Forest de PlanetScope

VI	Mes	Métricas		
		R^2	RMSE	MAE
Mean	2	-3,658	1,789	1,543
	3	-2,474	1,545	1,410
Min	2	-8,745	2,588	2,063
	3	-8,199	2,515	1,996
Max.	2	-5,812	2,164	1,810
	3	-0,819	1,118	1,023

En la Figura 46 se presenta la comparación entre la productividad estimada y la productividad real (t/ha), utilizando el modelo que presentó el mejor desempeño, correspondiente a los valores máximos de los índices de vegetación en el mes 3. En general, aunque se observa una dispersión considerable en los valores, el modelo logra un rendimiento aceptable en la predicción. Esto coincide con lo reportado en la Tabla 36, donde las métricas de error arrojaron valores de RMSE y MAE cercanos a 1 t/ha.

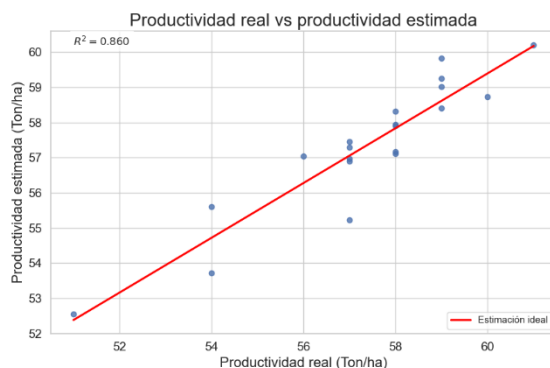


Figura 46. Productividad real vs. estimada – Random Forest (PlanetScope)

7.4.2 Sentinel-2.

Los resultados del modelo Random Forest con Sentinel-2, resumidos en la Tabla 37, muestran un comportamiento irregular y en general poco satisfactorio. En los valores medios y mínimos, los R^2 permanecieron en rangos negativos y los errores superaron con frecuencia las 2 toneladas por hectárea, lo que refleja una capacidad predictiva deficiente. Sin embargo, los valores máximos del mes 3 ofrecieron un desempeño relativamente más estable, alcanzando un R^2 de $-0,221$ junto con un RMSE de 0,916 y un MAE de 0,686. Aunque este ajuste sigue siendo bajo en términos absolutos, representa la aproximación más cercana a la productividad real obtenida con este sensor, indicando que el uso de valores extremos es más informativo que las medias o los mínimos.

Tabla 37. Modelo Regresión Random Forest de Sentinel-2

VI	Mes	Métricas		
		R^2	RMSE	MAE
Mean	2	-1,314	1,261	0,897
	3	-8,372	2,538	2,022
Min	2	-8,015	2,490	2,125
	3	-4,390	1,925	1,712
Max.	2	-4,397	1,926	1,265
	3	-0,221	0,916	0,686

En la Figura 47 se presenta la comparación entre la productividad real y la productividad estimada, utilizando la información de los valores máximos de los índices de vegetación correspondientes al mes 3. Se observa que, aunque los valores presentan cierta dispersión, el modelo mantiene un buen desempeño en la estimación de productividad, lo cual coincide con lo observado en la Tabla 37, donde las métricas de error (RMSE y MAE) fueron inferiores a 1 t/ha.

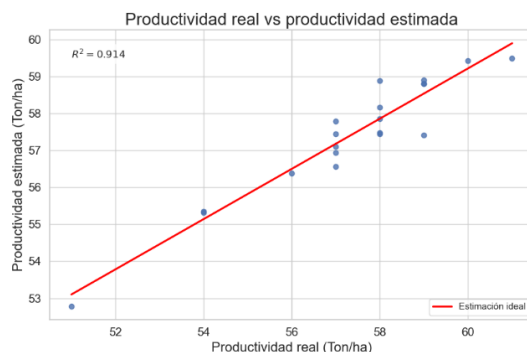


Figura 47. Productividad real vs. estimada – Random Forest (Sentinel-2).

7.4.3 Sentinel-2 + PlanetScope.

Los resultados de la combinación Sentinel-2 + PlanetScope, mostrados en la Tabla 38, evidencian un comportamiento irregular, aunque con algunos escenarios más estables que los obtenidos con cada fuente por separado. El caso más destacado se dio en el valor medio del mes 2, con un R^2 de 0,148 y errores relativamente bajos, RMSE de 0,765 y MAE de 0,604, lo que refleja una señal más consistente en esta etapa del ciclo. También resultó relevante el valor máximo del mes 3, que alcanzó un R^2 cercano a cero (0,009) pero con los errores más bajos de toda la serie, RMSE de 0,825 y MAE de 0,637, lo que sugiere una aproximación más ajustada pese al bajo nivel explicativo del coeficiente. En contraste, los valores mínimos y otros escenarios mantuvieron R^2 negativos y errores altos, limitando la capacidad predictiva del modelo.

En conjunto, esta integración mostró que los valores medios y máximos en fases intermedias ofrecen mayor estabilidad para representar la productividad que los valores mínimos.

Tabla 38. Modelo de regresión Random Forest Sentinel-2 + PlanetScope.

VI	Mes	Métricas		
		R^2	RMSE	MAE
Mean	2	0,148	0,765	0,604
	3	-3,365	1,732	1,286
Min	2	-5,434	2,103	1,615
	3	-2,780	1,612	1,334
Max.	2	-2,976	1,653	1,345
	3	0,009	0,825	0,637

En la Figura 48 se presenta la comparación entre la productividad real y la estimada, utilizando los valores medios de los índices de vegetación correspondientes al mes 2. En este caso, el modelo muestra un desempeño relativamente superior en comparación con los demás modelos obtenidos, lo que confirma la viabilidad de estimar la productividad a partir de esta información. Esto coinciden con lo consignado en la Tabla 38, donde las métricas de error (RMSE y MAE) fueron inferiores a 0,8 t/ha.

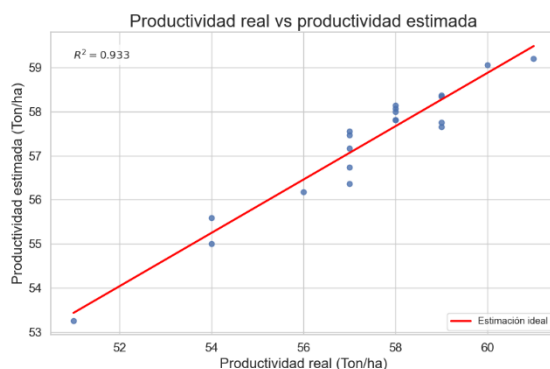


Figura 48. Productividad real vs. estimada – Random Forest (Sentinel-2 + PlanetScope)

7.4.4 Ópticos + Radar

Se aplicaron modelos de Random Forest a los mejores ajustes identificados en la sección 0 los cuales se construyeron a partir de índices derivados de sensores ópticos y de radar. El propósito fue evaluar si la utilización de Random Forest permite mejorar el desempeño de la regresión en comparación con los modelos lineales múltiples previamente estimados.

7.4.4.1 PlanetScope + S1

El mejor modelo obtenido con PlanetScope correspondió al mes 3, utilizando los valores medios de los índices NDVI, GNDVI y VHVV. Este modelo alcanzó un R^2 de -0,578, un RMSE de 1,042 y un MAE de 0,933. Aunque el valor de R^2 evidencia un ajuste muy bajo, al realizar un análisis gráfico de dispersión entre la productividad real (X) y la productividad calculada (Y) muestra que las diferencias entre ambos conjuntos de valores no son tan pronunciadas (Ver Figura 49).

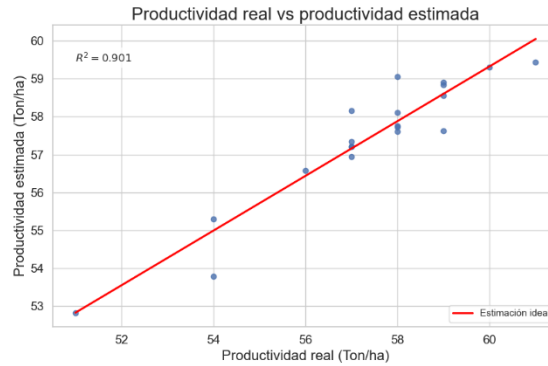


Figura 49. Productividad real vs. estimada – Random Forest + RVI (PlanetScope + Sentinel-1)

7.4.4.2 Sentinel-2 + Sentinel-1

En el caso de la información derivada de Sentinel-2, el mejor modelo se obtuvo a partir de la combinación de valores máximos de NDVI, GNDVI, VHVV y RVI calculados para el mes 3. Este modelo alcanzó un R^2 de -0,197, un RMSE de 0,907 y un MAE de 0,712. De forma similar a lo observado en la combinación de Sentinel-1 con PlanetScope, aunque el ajuste estadístico del modelo es bajo, la comparación entre productividad real y la productividad estimada, muestra una correspondencia relativa entre los valores. (Ver Figura 50).

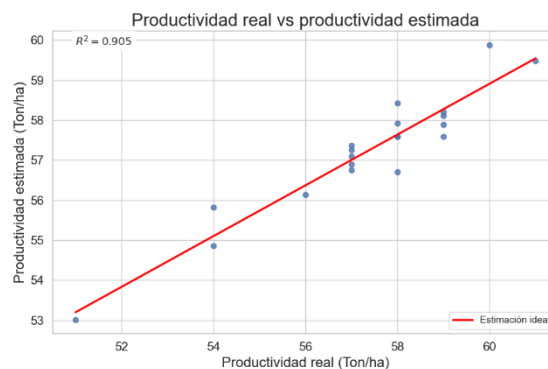


Figura 50. Productividad real vs. estimada – Random Forest + RVI (Sentinel-2 + Sentinel-1)

7.4.4.3 Sentinel-2 + Planet+S1

Finalmente, el modelo que presentó el mejor ajuste estadístico correspondió al obtenido con los valores mínimos del mes tres de los índices NDVI, GNDVI, SAVI, EVI, VHVV y RVI, a partir de la combinación de Sentinel-2 con el relleno de zonas de nubes mediante información de PlanetScope. Este modelo alcanzó un R^2 de 0,362, un RMSE de 0,663 y un MAE de 0,42. Al realizar el gráfico comparativo (Ver Figura 51) entre la productividad real y la productividad estimada mediante regresión Random Forest, se observó una correspondencia relativamente elevada entre ambos valores.

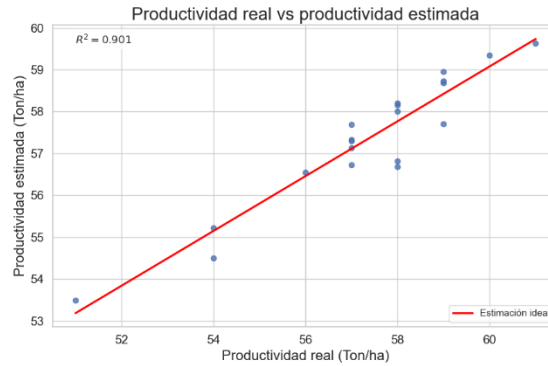


Figura 51. Productividad real vs. estimada – Random Forest + RVI (Sentinel-2 + PlanetScope + Sentinel-1)

El modelo de regresión Random Forest presentó valores de R^2 más bajos en comparación con las regresiones lineales; sin embargo, mostró una mejora evidente en las métricas de error, particularmente en RMSE y MAE. Esto indica que, aunque su capacidad para explicar la variabilidad estadística es limitada, logra generar predicciones más precisas y cercanas a los valores observados. Desde un punto de vista analítico, el modelo no representa completamente la estructura del proceso productivo, pero sí consigue ajustar relaciones no lineales entre los índices de vegetación y la productividad, lo que mejora su desempeño predictivo y revela patrones que los modelos lineales no logran capturar.

En conjunto, los resultados muestran que el desempeño de los modelos varía en función del índice empleado, el sensor utilizado y el momento fenológico evaluado. Las regresiones lineales múltiples presentaron coeficientes de determinación bajos, especialmente en los primeros meses del ciclo del cultivo, debido a que la cobertura foliar insuficiente limita la capacidad de los índices espectrales para reflejar la variabilidad productiva. A medida que el cultivo avanza y aumenta la biomasa, los índices se vuelven más representativos del estado del cultivo y muestran un comportamiento más consistente. El desempeño observado en Random Forest complementa este comportamiento, ya que evidencia que, aun cuando la capacidad explicativa es reducida, la información contenida en los índices sí permite generar estimaciones razonables cuando se emplean métodos capaces de capturar relaciones no lineales.

8 Discusión de resultados

Un paso esencial en la construcción de modelos orientados a la estimación de productividad agrícola corresponde a la obtención de datos reales. Estos constituyen la base para la calibración y validación de predicciones generadas a partir de índices derivados de información satelital, mejorando así la confiabilidad de los resultados. Sin esta etapa, los modelos carecerían de un punto de referencia que permita evaluar su desempeño frente a datos cercanos a la realidad. La recolección de información real puede realizarse mediante visitas de campo —como en el presente estudio— o a través de entrevistas con agricultores, tal como lo hicieron Bala & Islam (2009) [26] con 50 agricultores, Newton et al. (2018) [38] en 35 campos, igualmente que Ntouro et al. (2025) [42] en trigo. De igual forma, Mesas et al. (2025) recopilaban información de 28 parcelas comerciales de espinaca y Kalecinski et al. (2024) utilizaron datos obtenidos de una cosechadora [44], otro método consiste en el uso de fuentes oficiales, como reportes estadísticos nacionales o regionales. Sin embargo, estas bases suelen carecer de detalle espacial por lo que ofrecen valores globales por municipio o departamento, limitando su utilidad en análisis de precisión. Ejemplo de ello son los trabajos de Roznik et al. (2022) [28] quienes emplearon la base NASS del USDA y Rananavare & Chitnis (2022)[46], que recurrieron a registros agrícolas del gobierno de India.

Los hallazgos de Bala & Islam (2009), reportaron una correlación importante entre los índices de vegetación derivados de MODIS y el rendimiento de la papa, particularmente el NDVI basado en valores medios ($R^2 = 0,84$). En contraste, en este estudio las regresiones lineales simples con NDVI derivado de PlanetScope y Sentinel 2, alcanzaron un máximo de $R^2=0,369$. Esta diferencia puede atribuirse a factores metodológicos. En primer lugar, el sensor de MODIS, tiene una resolución espacial mucho menor (500 m) y la resolución de los sensores utilizados en el presente estudio fluctuaron entre 5 y 10 m, lo que genera mayor sensibilidad a la heterogeneidad de elementos presentes en el predio como otras clases de vegetación, en segundo lugar, la presencia de nubes en los sensores ópticos puede haber introducido una mayor variabilidad en las relaciones. Además, Bala e Islam utilizaron únicamente sensores ópticos, mientras que en este trabajo se integró información de radar ampliando los predictores, pero también la complejidad de las relaciones estadísticas. Sin embargo, en ambos estudios coinciden en la utilidad de los índices de vegetación derivados de sensores remotos como variables clave en la estimación de productividad de cultivos de papa, aunque el desempeño de cada modelo generado dependería del entorno geográfico, la resolución espacial, temporal de las imágenes utilizadas y los métodos de modelado empleados [26].

Newton et al. (2018), mostraron que el NDVI máximo derivado de series temporales de imágenes de Landsat 5, 7 y 8 permitía predecir el rendimiento de cultivos de papa con una precisión considerable de $R^2=0,81$ y un error máximo del 10,4 %, confirmando la relevancia de este índice en la fase de máximo vigor del cultivo. En contraste, los mejores resultados de este estudio con NDVI alcanzaron $R^2=0,369$, lo que en estas condiciones significa una menor capacidad explicativa de los modelos lineales simples que se utilizaron [38]. Del mismo modo, Roznik et al. (2022), utilizaron NDVI derivado de imágenes satélites del sensor MODIS con resolución de 250m para estimar rendimiento de cultivos de maíz, soja y trigo, mediante modelos de regresión spline cubica no lineal, reportando ajustes de $R^2=0,615$ para el maíz y $R^2=0,49$ para la soja [28]. Una diferencia clave con el presente estudio radica en la resolución espacial de los sensores empleados, mientras MODIS y Landsat tienen resoluciones de 250 y 30 m respectivamente, aquí se trabajó con resoluciones de 5 y 10 m, lo que junto a la nubosidad que tiene el municipio de Ventaquemada pudo influir en menores ajustes alcanzados.

Ntouros et al. (2025) desarrollaron un marco en Python para la estimación de productividad de trigo duro y blando empleando imágenes Sentinel-2. En el caso del trigo duro, los modelos alcanzaron R^2 entre 0,97 y 0,99, con RMSE entre 41,35 y 60,03 kg/1000 m² (0,41–0,60 t/ha); para el trigo blando, los ajustes fueron igualmente altos (R^2 entre 0,97 y 0,99), aunque con errores mayores de hasta 102,22 kg/1000 m² ($\approx 1,02$ t/ha). En contraste, los modelos aplicados en el presente estudio sobre papa presentaron R^2 más bajos (máximo 0,520 en la regresión lineal múltiple con SAVI (PlanetScope) + RVI (Sentinel-1) y 0,362 en Random Forest con Sentinel-2 + PlanetScope + Sentinel-1), aunque con errores comparables o incluso inferiores (RMSE=0,66 t/ha; MAE=0,42 t/ha). Estas diferencias pueden explicarse por la mayor uniformidad fenológica del trigo, que facilita la captura de la variabilidad con índices espectrales, mientras que la papa presenta heterogeneidad intrapredial asociada a la densidad de cobertura del dosel, pendientes y manejo agrícola. Además, Ntouros et al. utilizaron únicamente sensores ópticos, en tanto que aquí se incorporó información radar, lo que aporta estabilidad bajo nubosidad, pero añade complejidad a las relaciones estadísticas. Aun así, en ambos estudios se reconoce que, más allá de las diferencias en el R^2 , errores bajos hacen que los modelos sean útiles en la práctica y se resalta la importancia de flujos reproducibles en Python para escalar a distintos cultivos y regiones[42].

Por su parte, Manivasagam et al. (2021) demostraron que la integración de imágenes Sentinel-2 y PlanetScope en el modelo SAFY para trigo permitía mejorar la predicción de rendimientos, alcanzando un $R^2 \approx 0,45$ y un RMSE de 0,69 t/ha en su mejor escenario. Este hallazgo coincide con lo observado en el presente estudio, donde la fusión de fuentes satelitales también mejoró la capacidad explicativa respecto al uso de un único sensor. Sin embargo, existen diferencias relevantes en el enfoque: mientras que SAFY es un modelo de asimilación fisiológica que depende exclusivamente de la información óptica y se beneficia de la uniformidad fenológica del trigo, el presente trabajo incorporó índices ópticos y variables de radar para papa, lo que permitió ampliar el rango de predictores, pero también introdujo mayor complejidad en las relaciones estadísticas. Esta diferencia metodológica explica que los errores lineales fueran más altos en papa que en trigo, aunque confirma que la integración multifuente constituye una estrategia robusta para la estimación de productividad agrícola [43].

La importancia del momento fenológico también se confirmó en este estudio, fue el mes 4, correspondiente al máximo vigor vegetativo y la fase de llenado de tubérculos, el que presentó el mejor ajuste lineal múltiple. Este resultado evidencia que la etapa de mayor desarrollo del dosel es determinante para la precisión de los modelos de predicción. Una conclusión semejante fue reportada por Kalecinski et al. (2024), quienes al analizar cultivos de maíz, soja y arroz mediante la sinergia de datos SAR y ópticos observaron que los mayores desempeños se alcanzaban precisamente durante las fases de crecimiento más avanzadas, con valores de R^2 superiores a 0,60 y errores controlados. Ambos estudios coinciden, por tanto, en que la selección del momento fenológico adecuado constituye un factor decisivo en la calidad de los modelos. En el caso de papa, factores propios de la zona de estudio, como la variabilidad climática en Boyacá, la nubosidad frecuente en imágenes ópticas y la complejidad topográfica de los predios, pudieron influir en que los ajustes fueran más bajos respecto a otros cultivos. No obstante, la consistencia alcanzada en los modelos basados en aprendizaje automático confirma que, más allá del coeficiente de determinación, métricas de error y estabilidad en la predicción son esenciales para garantizar la aplicabilidad de los resultados en agricultura de precisión. [44].

Mesas et al. (2025), al aplicar un enfoque de fusión de Sentinel-1 y Sentinel-2 para la estimación de rendimiento en espinaca, evidenciaron que los modelos de Random Forest a escala de parcela ofrecieron los mejores desempeños, alcanzando R^2 de hasta 0,89 y errores muy bajos ($\approx 1,4$ %). Estos resultados confirman la importancia de trabajar con información multiescala y multifuente para mejorar la precisión de las predicciones, especialmente en escenarios de nubosidad. En el presente estudio sobre papa, aunque los valores de R^2 fueron más modestos, la incorporación de radar junto con índices ópticos permitió reducir la incertidumbre en comparación con los modelos basados solo en datos ópticos. Ambos estudios coinciden en que la sinergia SAR-óptico constituye una estrategia robusta para fortalecer la estabilidad y la aplicabilidad de los modelos, demostrando su utilidad tanto en cultivos de dosel homogéneo como la espinaca como en sistemas más heterogéneos y complejos como la papa [45].

Finalmente, los modelos de Random Forest fueron los de mejor desempeño en este estudio, destacándose un modelo óptimo que alcanzó $R^2=0,362$ al integrar Sentinel-2 fusionado con PlanetScope y Sentinel-1 en el mes 3, con RMSE=0,663 t/ha y MAE=0,420 t/ha. El rendimiento de estos modelos varió en función de la cantidad de variables predictoras empleadas, lo que refleja la flexibilidad del enfoque. Estos resultados confirman parcialmente lo reportado por Rananavare & Chitnis (2022), quienes, aunque no utilizaron índices de vegetación, mostraron que Random Forest alcanzó un desempeño sobresaliente en la estimación de productividad con un MAE=2,818 y $R^2=0,82$, frente a un MAE=1,471 en la regresión lineal. Aunque los

valores absolutos difieren, ambos estudios coinciden en que Random Forest ofrece mayor estabilidad y capacidad predictiva, consolidándose como un enfoque adecuado para la estimación de productividad de cultivos de papa.[46].

Los coeficientes de determinación obtenidos en el presente estudio se ubicaron, en términos generales, por debajo de los valores reportados por otros autores que han modelado la productividad agrícola mediante índices derivados de sensores remotos. Esto se debe a que gran parte de esos estudios fueron desarrollados en condiciones experimentales controladas o en cultivos con mayor homogeneidad fenológica. En contraste, esta investigación se realizó a escala predial en un entorno de alta complejidad fisiográfica, climática y topográfica, donde es común, particularmente en agricultura andina, que los R^2 sean más bajos debido a la heterogeneidad espacial, la variabilidad intrapredial y la frecuente presencia de nubosidad. Factores como las diferencias en manejo agrícola, el efecto del suelo y la heterogeneidad del dosel influyen directamente en la estabilidad de los modelos. Por tanto, los valores de R^2 obtenidos deben interpretarse dentro de la variabilidad inherente del paisaje agrícola andino y no como indicadores de deficiencia del modelo.

Adicionalmente, es necesario considerar que la integración de sensores con características radiométricas y resoluciones espaciales diferentes introduce complejidades estadísticas que también pueden reducir el R^2 . En el presente estudio, la combinación de información óptica y radar permitió una caracterización más completa del cultivo y una mayor robustez en presencia de nubosidad, condición frecuente en Boyacá; sin embargo, esta integración también genera mayor dispersión en los datos y, por ende, disminuye el ajuste lineal. A diferencia de estudios realizados en entornos controlados, donde existe homogeneidad temporal y espectral, los paisajes de montaña presentan variabilidad en pendiente, iluminación, estructura del dosel y humedad del suelo, lo que incrementa la dispersión residual y limita la capacidad explicativa de los modelos.

En relación con la calidad del modelo, los valores bajos de R^2 indican que los índices empleados capturan solo una fracción de la variabilidad del rendimiento; sin embargo, esto no invalida los resultados obtenidos. Además, los valores bajos de RMSE y MAE demuestran que, pese a los R^2 moderados, los modelos siguen siendo operativamente útiles para la estimación de productividad a escala predial. En síntesis, los coeficientes de determinación obtenidos reflejan la complejidad biológica y paisajística de la zona de estudio y evidencian la necesidad de incorporar información adicional en futuros trabajos, como variables agroclimáticas, modelos no lineales o series multitemporales más densas, con el fin de incrementar la robustez y el desempeño predictivo en escenarios agrícolas de alta variabilidad.

9 Conclusiones

La metodología formulada en este trabajo permitió integrar información de sensores ópticos (Sentinel-2 y PlanetScope) y de radar (Sentinel-1) para estimar la productividad de papa en el municipio de Ventaquemada, superando las limitaciones individuales de cada fuente satelital y ofreciendo una caracterización más completa del ciclo fenológico. PlanetScope aportó mayor detalle espacial para capturar la heterogeneidad intrapredial, mientras que Sentinel-2 garantizó una cobertura regional estable, complementando la capacidad del radar para operar bajo condiciones de nubosidad frecuentes en la zona de estudio.

En términos estadísticos, los modelos de regresión lineal múltiple mostraron un mejor desempeño que los modelos simples, confirmando la utilidad de combinar distintos índices de vegetación. El mejor ajuste se obtuvo con la combinación de SAVI de PlanetScope y RVI de Sentinel-1 en el mes 4, alcanzando un R^2 de 0,520, un R^2 ajustado de 0,463, un RMSE de 1,538 t/ha y un MAE de 1,226 t/ha. Estos valores permiten concluir que, aunque la capacidad explicativa es moderada, la calidad del modelo es adecuada para representar tendencias productivas a escala predial, dado que los errores absolutos se mantienen dentro de rangos manejables para aplicaciones agrícolas.

Por otra parte, los resultados obtenidos mediante Random Forest mostraron un comportamiento diferente. Aunque sus valores de R^2 fueron menores (máximo de 0,362 con la combinación de Sentinel-2 + PlanetScope con relleno + Sentinel-1 en el mes 3), los errores fueron considerablemente más bajos, con un RMSE de 0,663 t/ha y un MAE de 0,420 t/ha. Esto indica que, pese a explicar menos variabilidad estadística, la calidad predictiva del modelo es alta en términos absolutos, lo que confirma su utilidad operativa en escenarios de alta variabilidad espacial como Ventaquemada.

De manera general, los índices de vegetación reflejaron adecuadamente la dinámica fenológica del cultivo de papa. Los mejores desempeños se alcanzaron en las fases de crecimiento pleno y maduración, cuando la biomasa aérea y el vigor foliar ejercen mayor influencia sobre el rendimiento final, mientras que las etapas iniciales presentaron mayor dispersión debido a la heterogeneidad del establecimiento.

Desde una perspectiva aplicada, la metodología propuesta constituye una herramienta replicable para el monitoreo agrícola en regiones con condiciones similares. Para agricultores y asociaciones productoras, la información satelital representa un complemento útil a los registros tradicionales, facilitando decisiones sobre fertilización, manejo fitosanitario y planificación de la cosecha. Para instituciones gubernamentales y de investigación, esta aproximación ofrece un insumo valioso para el seguimiento de cultivos, la planificación territorial y el diseño de políticas orientadas a la seguridad alimentaria.

Los coeficientes de determinación obtenidos fueron inferiores a los reportados en otros estudios desarrollados en condiciones experimentales o en cultivos más homogéneos, lo que evidencia una menor capacidad explicativa en el contexto fisiográfico y productivo del altiplano cundiboyacense. Sin embargo, esta menor capacidad se relaciona directamente con la complejidad espacial, climática y de manejo que caracteriza a la agricultura andina. El comportamiento de los errores (RMSE y MAE) demuestra que, pese a la limitada capacidad explicativa, los modelos mantienen precisión suficiente para ser utilizados en estimaciones reales de productividad, lo que confirma su aplicabilidad en contextos de alta variabilidad.

Finalmente, se concluye que los modelos estadísticos tradicionales y los algoritmos de aprendizaje automático son enfoques complementarios. Mientras la regresión lineal múltiple identificó los índices con mayor capacidad explicativa (SAVI y RVI), Random Forest aportó una precisión superior en términos de error absoluto. En conjunto, estos resultados permiten afirmar que la metodología planteada es apta, válida y suficientemente robusta para cumplir el objetivo general de estimar la productividad de papa mediante integración multifuente de información satelital, abriendo la puerta al desarrollo de sistemas de monitoreo agrícola escalables y útiles para la gestión productiva en Colombia.

10 Bibliografía

- [1] F. R. Jimenez-Lopez, I. A. Ruge-Ruge, and A. F. Jimenez-Lopez, “Deep Learning techniques applied to predict the irrigation prescription for potato crops in Boyacá,” *2021 IEEE Chil. Conf. Electr. Electron. Eng. Inf. Commun. Technol. CHILECON 2021*, 2021, doi: 10.1109/CHILECON54041.2021.9703064.
- [2] D. Gómez, P. Salvador, J. Sanz, and J. L. Casanova, “New spectral indicator Potato Productivity Index based on Sentinel-2 data to improve potato yield prediction: a machine learning approach,” *Int. J. Remote Sens.*, vol. 42, no. 9, pp. 3430–3448, 2021, doi: 10.1080/01431161.2020.1871102.
- [3] A. F. Velez Betancourt, *Estado del arte de la cadena de valor de la papa en Colombia*. 2020.
- [4] Ministerio de Agricultura, “Cadena de la papa: Indicadores e instrumentos,” *Minist. Agric. y Desarro. Rural*, vol. unknown, no. 52, pp. 167–194, 2019.
- [5] Fedepapa, “Boletín regional Boyacá, Vol. 6.” pp. 2–10, 2022.
- [6] Fedepapa-FNFP, “Boletín regional Boyaca,” *Fondo Nac. la Papa Fed. Nac. la Papa*, vol. 5, p. 10, 2021.
- [7] A. Porras, P., Espitia, E., Rodriguez, G., Polo, S., Vásquez, *Modelo Productivo de la papa(Solanum tuberosum L.) de la variedad pastusa para el altiplano cundiboyacense*. 2011.
- [8] R. Yusianto, Marimin, Suprihatin, and H. Hardjomidjojo, “Spatial analysis for crop land suitability evaluation: A case study of potatoes cultivation in wonosobo, Indonesia,” *Proc. - 2020 Int. Semin. Appl. Technol. Inf. Commun. IT Challenges Sustain. Scalability, Secur. Age Digit. Disruption, iSemantic 2020*, pp. 313–319, 2020, doi: 10.1109/iSemantic50169.2020.9234284.
- [9] FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS., *STATE OF FOOD AND AGRICULTURE 2019 (SPANISH EDITION) : moving forward on food loss and waste ... reduction*. FOOD & AGRICULTURE ORG, 2019.
- [10] Y. Liu *et al.*, “Remote-sensing estimation of potato above-ground biomass based on spectral and spatial features extracted from high-definition digital camera images,” *Comput. Electron. Agric.*, vol. 198, no. May, p. 107089, 2022, doi: 10.1016/j.compag.2022.107089.
- [11] N. J. Rabe, D. R. Peddle, and A. M. Smith, “Spectral mixture analysis of potato crops under different irrigation regimes,” *Int. Geosci. Remote Sens. Symp.*, vol. 5, no. August 1998, pp. 2784–2786, 2002, doi: 10.1109/igarss.2002.1026774.
- [12] J. Awange and J. Kiema, *Light Detection And Ranging (LiDAR) BT - Environmental Geoinformatics: Extreme Hydro-Climatic and Food Security Challenges: Exploiting the Big Data*. 2019.
- [13] E. F. Berra and R. Gaulton, “Remote sensing of temperate and boreal forest phenology: A review of progress, challenges and opportunities in the intercomparison of in-situ and satellite phenological metrics,” *For. Ecol. Manage.*, vol. 480, no. August 2020, p. 118663, 2021, doi: 10.1016/j.foreco.2020.118663.
- [14] H. Wang, Yao; Zhao, Chuanyan; Xu, Zhonglin; Peng, “ESTIMATE CROP

EVAPOTRANSPIRATION OF SPRING WHEAT IN THE MIDDLE REACHES OF HEIHE RIVER BASIN , NORTHWESTERN CHINA (1 . Key Laboratory of Arid and Grassland Agroecology with the Ministry of Education , Lanzhou University , Lanzhou 730000 , China 2 . Key La,” *Ieee*, vol. 2, pp. 1537–1540, 2010.

- [15] Ruiz Barros Liliana, “Historia, variedades y propiedades de la papa en Colombia,” pp. 1–13, 2024.
- [16] Planet Labs, “Planet Imagery Product Specifications,” *Planet Labs Inc*, no. December, pp. 1–92, 2023, [Online]. Available: <https://www.planet.com/products/planet-imagery/>.
- [17] European Space Agency, “Sentinel-2 Products Specification Document. Issue 13.1. ESA REF: S2-PDGS-TAS-DI-PSD,” 2024, [Online]. Available: https://sentiwiki.copernicus.eu/_attachments/1692737/S2-PDGS-CS-DI-PSD - S2 Product Specification Document 2024 - 15.0.pdf?inst-v=1d5d1a47-ceec-4a48-8345-fc161fc5e777.
- [18] European Space Agency, “S2 MPC Level-1 Algorithm Theoretical Bases Document,” *Https://Sentinel.Esa.Int/Web/Sentinel/Technical-Guides/Sentinel-2-Msi/Msi-Instrument*, vol. 1, pp. 8–12, 2023.
- [19] K. Sorensen and K. Sorensen, “Speckle Reduction of Synthetic Aperture Radar Images,” 2017.
- [20] European Space Agency, “SNAP (Sentinel Application Platform) – Science Toolbox Exploitation Platform.” 2025, [Online]. Available: <https://step.esa.int/main/toolboxes/snap/>.
- [21] QGIS, “About QGIS,” *QGIS Website*, 2021, [Online]. Available: <https://qgis.org/en/site/about/index.html>.
- [22] E. Cañon, W. Vargas, and C. Gomez, *Manual ArcGIS ArcMap Desktop avanzado*. 2023.
- [23] P. Spector, “Introduction to Python Programming Course Notes,” *Univ. Calif. Berkeley*, 2005.
- [24] H. Aguilar, R. Mora, and C. Vargas, “Atmospheric Correction Methodology for Aster, Rapideye, Spot 2 and Landsat 8 Images With Envi Flaash Module Software,” *Rev. Geográfica América Cent.*, vol. 2, no. 53, pp. 39–59, 2014.
- [25] A. Alkhaled, P. A. Townsend, and Y. Wang, “Remote Sensing for Monitoring Potato Nitrogen Status,” *Am. J. Potato Res.*, vol. 100, no. 1, pp. 1–14, 2023, doi: 10.1007/s12230-022-09898-9.
- [26] S. K. Bala and A. S. Islam, “Correlation between potato yield and MODIS-derived vegetation indices,” *Int. J. Remote Sens.*, vol. 30, no. 10, pp. 2491–2507, 2009, doi: 10.1080/01431160802552744.
- [27] E. Farg, S. M. Arafat, M. S. Abd El-Wahed, and A. M. El-Gindy, “Estimation of Evapotranspiration ETc and Crop Coefficient K c of Wheat, in south Nile Delta of Egypt Using integrated FAO-56 approach and remote sensing data,” *Egypt. J. Remote Sens. Sp. Sci.*, vol. 15, no. 1, pp. 83–89, 2012, doi: 10.1016/j.ejrs.2012.02.001.
- [28] M. Roznik, M. Boyd, and L. Porth, “Improving crop yield estimation by applying higher resolution satellite NDVI imagery and high-resolution cropland masks,” *Remote Sens. Appl. Soc. Environ.*, vol. 25, no. October 2021, p. 100693, 2022, doi: 10.1016/j.rsase.2022.100693.
- [29] A. Nițu, C. Florea, M. Ivanovici, and A. Racoviteanu, “NDVI and Beyond: Vegetation Indices as Features for Crop Recognition and Segmentation in Hyperspectral Data †,” *Sensors*, vol. 25, no. 12, pp. 1–26, 2025, doi: 10.3390/s25123817.

- [30] N. Karthikeyan, M. C. Shashikkumar, and J. Ramanamurthy, "A study on vegetation vigour as affected by soil properties using remote sensing approach," *Proc. Int. Conf. "Recent Adv. Sp. Technol. Serv. Clim. Chang. - 2010", RSTS CC-2010*, pp. 107–110, 2010, doi: 10.1109/RSTSCC.2010.5712811.
- [31] X. Hu *et al.*, "Radar vegetation indices for monitoring surface vegetation: Developments, challenges, and trends," *Sci. Total Environ.*, vol. 945, no. June, 2024, doi: 10.1016/j.scitotenv.2024.173974.
- [32] N. Roustaei, "Application and interpretation of linear-regression analysis," *Med. Hypothesis, Discov. Innov. Ophthalmol.*, vol. 13, no. 3, pp. 151–159, 2024, doi: 10.51329/mehdiophthal1506.
- [33] G. K. Uyanık and N. Güler, "A Study on Multiple Linear Regression Analysis," *Procedia - Soc. Behav. Sci.*, vol. 106, pp. 234–240, 2013, doi: 10.1016/j.sbspro.2013.12.027.
- [34] L. Lemeš and A. Akagic, "Prediction of Real Estate Market Prices with Regression Algorithms," *Lect. Notes Networks Syst.*, vol. 539 LNNS, no. October, pp. 401–411, 2023, doi: 10.1007/978-3-031-17697-5_32.
- [35] S. M. M. Abdou, B. A. Engel, S. M. Emam, and K. M. A. El-Latif, "AquaCrop Model Validation for Simulation Wheat Productivity under Water Stress Condition," *Commun. Soil Sci. Plant Anal.*, vol. 53, no. 3, pp. 281–292, 2022, doi: 10.1080/00103624.2021.2016806.
- [36] E. Lahura, "El coeficiente de correlacion y correlaciones espureas," vol. 3, pp. 103–111, 2013.
- [37] T. O. Hodson, "Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not," *Geosci. Model Dev.*, vol. 15, no. 14, pp. 5481–5487, 2022, doi: 10.5194/gmd-15-5481-2022.
- [38] I. H. Newton, A. F. M. Tariqul Islam, A. K. M. Saiful Islam, G. M. Tarekul Islam, A. Tahsin, and S. Razzaque, "Yield Prediction Model for Potato Using Landsat Time Series Images Driven Vegetation Indices," *Remote Sens. Earth Syst. Sci.*, vol. 1, no. 1–2, pp. 29–38, 2018, doi: 10.1007/s41976-018-0006-0.
- [39] E. P. Glenn, C. M. U. Neale, D. J. Hunsaker, and P. L. Nagler, "Vegetation index-based crop coefficients to estimate evapotranspiration by remote sensing in agricultural and natural ecosystems," *Hydrol. Process.*, vol. 25, no. 26, pp. 4050–4062, 2011, doi: 10.1002/hyp.8392.
- [40] A. Huete, K. Didan, T. Miura, E. P. Rodriguez, X. Gao, and L. G. Ferreira, "Overview of the radiometric and biophysical performance of the MODIS vegetation indices," *Inorganica Chim. Acta*, vol. 95, no. 1, pp. 1–3, 1984, doi: 10.1016/S0020-1693(00)85959-9.
- [41] S. Nagini, T. V. R. Kanth, and B. V. Kiranmayee, "Agriculture yield prediction using predictive analytic techniques," *Proc. 2016 2nd Int. Conf. Contemp. Comput. Informatics, IC3I 2016*, pp. 783–788, 2016, doi: 10.1109/IC3I.2016.7918789.
- [42] K. Ntouros, K. Papatheodorou, G. Gkologkinas, and V. Drimzakas-Papadopoulos, "A Python Framework for Crop Yield Estimation Using Sentinel-2 Satellite Data," *Earth (Switzerland)*, vol. 6, no. 1, 2025, doi: 10.3390/earth6010015.
- [43] V. S. Manivasagam, Y. Sadeh, G. Kaplan, D. J. Bonfil, and O. Rozenstein, "Studying the feasibility of assimilating sentinel-2 and planetscope imagery into the SAFY crop model to predict within-field wheat yield," *Remote Sens.*, vol. 13, no. 12, 2021, doi: 10.3390/rs13122395.
- [44] N. I. Kalecinski *et al.*, "Crop yield estimation at different growing stages using a synergy of SAR

- and optical remote sensing data,” *Sci. Remote Sens.*, vol. 10, no. July, p. 100153, 2024, doi: 10.1016/j.srs.2024.100153.
- [45] F. J. Mesas-Carrascosa, J. T. Arosemena-Jované, S. Cantón-Martínez, F. Pérez-Porras, and J. Torres-Sánchez, “Enhancing Crop Yield Estimation in Spinach Crops Using Synthetic Aperture Radar-Derived Normalized Difference Vegetation Index: A Sentinel-1 and Sentinel-2 Fusion Approach,” *Remote Sens.*, vol. 17, no. 8, pp. 1–16, 2025, doi: 10.3390/rs17081412.
 - [46] L. B. Rananavare and S. Chitnis, “Crop Yield Prediction Using Temporal Data,” *2022 IEEE Int. Conf. Electron. Comput. Commun. Technol. CONECCT 2022*, pp. 1–5, 2022, doi: 10.1109/CONECCT55679.2022.9865791.
 - [47] ESA, “S2 MPC Level-2A Algorithm Theoretical Basis Document,” *ESA Doc. Libr.*, pp. 1–78, 2021.
 - [48] J. R. Jensen, *Remote sensing of the environment: An earth resource perspective 2/e*. Pearson Education India, 2009.
 - [49] P. J. Curran, “Remote sensing of foliar chemistry,” *Remote Sens. Environ.*, vol. 30, no. 3, pp. 271–278, 1989, doi: [https://doi.org/10.1016/0034-4257\(89\)90069-2](https://doi.org/10.1016/0034-4257(89)90069-2).