



Modelo de estimación espacial del rendimiento de papa (*Solanum tuberosum* L.) usando variables morfológicas, imágenes ópticas y algoritmos de aprendizaje de máquina

Sebastian Flórez Hormiga

Bogotá D.C, Colombia
2025

Modelo de estimación espacial del rendimiento de papa (*Solanum tuberosum* L.) usando variables morfológicas, imágenes ópticas y algoritmos de aprendizaje de máquina

Sebastian Flórez Hormiga

Tesis o trabajo de grado presentada(o) como requisito parcial para optar al título de:
Magíster en Geomática

Co-director(a):
Ph.D.Iván Alberto Lizarazon

Director(a):
Ph.D. Liliana Carolina Castillo Villamor

Línea de Investigación:
Desarrollo e implementación de tecnologías geoespaciales

Grupo de Investigación:
Nombrar el grupo en caso que sea posible

Universidad Nacional de Colombia
Facultad, Departamento (Escuela, etc.)
Bogotá D.C, Colombia
2025

(Dedicatoria o un lema)

Su uso es opcional y cada autor podrá determinar la distribución del texto en la página, se sugiere esta presentación. En ella el autor dedica su trabajo en forma especial a personas y/o entidades.

Por ejemplo:

A mis padres

o

La preocupación por el hombre y su destino siempre debe ser el interés primordial de todo esfuerzo técnico. Nunca olvides esto entre tus diagramas y ecuaciones.

Albert Einstein

Agradecimientos

Esta sección es opcional, en ella el autor agradece a las personas o instituciones que colaboraron en la realización de la tesis o trabajo de investigación. Si se incluye esta sección, deben aparecer los nombres completos, los cargos y su aporte al documento.

Resumen

El resumen es una presentación abreviada y precisa (la NTC 1486 de 2008 recomienda revisar la norma ISO 214 de 1976). Se debe usar una extensión máxima de 12 renglones. Se recomienda que este resumen sea analítico, es decir, que sea completo, con información cuantitativa y cualitativa, generalmente incluyendo los siguientes aspectos: objetivos, diseño, lugar y circunstancias, pacientes (u objetivo del estudio), intervención, mediciones y principales resultados, y conclusiones. Al final del resumen se deben usar palabras claves tomadas del texto (mínimo 3 y máximo 7 palabras), las cuales permiten la recuperación de la información.

Palabras clave: Estimación de rendimiento; Imágenes multiespectrales; Random Forest; Random Forest Geoespacial; Aprendizaje automático de máquina.

A continuación se presentan algunos ejemplos de tesauros que se pueden consultar para asignar las palabras clave, según el área temática:

Artes: AAT: Art y Architecture Thesaurus.

Ciencias agropecuarias: 1) Agrovoc: Multilingual Agricultural Thesaurus - F.A.O. y 2) GEMET: General Multilingual Environmental Thesaurus.

Ciencias sociales y humanas: 1) Tesauro de la UNESCO y 2) Population Multilingual Thesaurus.

Ciencia y tecnología: 1) Astronomy Thesaurus Index. 2) Life Sciences Thesaurus, 3) Subject Vocabulary, Chemical Abstracts Service y 4) InterWATER: Tesauro de IRC - Centro Internacional de Agua Potable y Saneamiento.

Tecnologías y ciencias médicas: 1) MeSH: Medical Subject Headings (National Library of Medicine's USA) y 2) DECS: Descriptores en ciencias de la Salud (Biblioteca Regional de Medicina BIREME-OPS).

Multidisciplinarias: 1) LEMB - Listas de Encabezamientos de Materia y 2) LCSH- Library of Congress Subject Headings.

También se pueden encontrar listas de temas y palabras claves, consultando las distintas bases de datos disponibles a través del Portal del Sistema Nacional de Bibliotecas¹, en la sección Recursos bibliográficos. opción "Bases de datos".

Abstract

Es el mismo resumen pero traducido al inglés. Se debe usar una extensión máxima de 12 renglones. Al final del Abstract se deben traducir las anteriores palabras claves tomadas del

¹ver: www.sinab.unal.edu.co

texto (mínimo 3 y máximo 7 palabras), llamadas keywords. Es posible incluir el resumen en otro idioma diferente al español o al inglés, si se considera como importante dentro del tema tratado en la investigación, por ejemplo: un trabajo dedicado a problemas lingüísticos del mandarín seguramente estaría mejor con un resumen en mandarín.

Keywords: palabras clave en inglés(máximo 10 palabras, preferiblemente seleccionadas de las listas internacionales que permitan el indizado cruzado)

Contenido

Agradecimientos	VII
Resumen	IX
Lista de símbolos	XIII
1. Introducción	1
1.1. Antecedentes	3
1.2. Planteamiento del problema	3
1.2.1. Aproximación conceptual y metodológica	3
1.2.2. Pregunta de investigación	3
1.2.3. Objetivos	3
1.2.4. Justificación	3
2. Marco Teórico y Estado del Arte	4
2.1. Uso de herramientas de fenotipado para la estimación de rendimiento	4
2.2. Avances en percepción remota para el fenotipado de alto rendimiento	5
2.3. Estado del Arte	11
2.3.1. Estudios Globales y Regionales	11
2.3.2. Percepción remota en la modelación espacial del riesgo	11
2.3.3. Monitoreo de incendios a partir de datos de percepción remota	11
2.3.4. Modelos estadísticos y de aprendizaje automático en la estimación del riesgo de incendios	11
2.3.5. Avances en Colombia y los Andes tropicales	11
3. Materiales y métodos	12
3.1. Área de estudio	12
3.2. Diseño muestral	13
3.3. Metodología	14
3.4. Adquisición de datos e imágenes	15
3.4.1. Medición de datos en campo	15
3.4.2. Adquisición de imágenes multiespectrales	16
3.5. Preprocesamiento de imágenes	17
3.5.1. Corrección radiométrica	17

3.5.2. Corrección geométrica	18
3.6. Procesamiento de datos	18
3.6.1. Densidad de siembra y establecimiento del cultivo	18
3.6.2. Altura del cultivo	18
3.6.3. Cobertura del dosel	19
3.7. Análisis de datos	19
4. Resultados	20
4.1. Resultados del preprocesamiento de los datos	20
4.2. Caracterización de las variables predictoras	20
5. Discusión	21
5.1. Cumplimiento de objetivos	21
5.2. Ventajas y limitaciones de la metodología utilizada	21
5.3. Aportes al conocimiento en Geomática	21
6. Conclusiones	22
A. Anexo: Nombrar el anexo A de acuerdo con su contenido	23
B. Anexo: Nombrar el anexo B de acuerdo con su contenido	24
C. Anexo: Nombrar el anexo C de acuerdo con su contenido	25

Listas de símbolos

Esta sección es opcional, dado que existen disciplinas que no manejan símbolos y/o abreviaturas.

Se incluyen símbolos generales (con letras latinas y griegas), subíndices, superíndices y abreviaturas (incluir sólo las clases de símbolos que se utilicen). Cada una de estas listas debe estar ubicada en orden alfabético de acuerdo con la primera letra del símbolo.

Símbolos con letras latinas

Símbolo	Término	Unidad SI	Definición
A	Área	m^2	$\int \int dx dy$
A_{BET}	Área interna del sólido	$\frac{\text{m}^2}{\text{g}}$	ver DIN ISO 9277
A_g	Área transversal de la fase gaseosa	m^2	Ec...
A_s	Área transversal de la carga a granel	m^2	Ec...
a	Coeficiente	1	Ec...
a	Contenido de ceniza	1	$\frac{m_{\text{ceniza}}}{m_{\text{bm},0}}$
c	Contenido de carbono	1	$\frac{m_{\text{C}}}{m}$
c	Longitud de la cuerda	m	Figura...
c	Concentración de la cantidad de materia	$\frac{\text{mol}}{\text{m}^3}$	$\frac{n}{V}$
D	Diámetro	m	
E_A	Energía de activación	$\frac{\text{kJ}}{\text{mol}}$	Ec....
F	Fracción de materia volátil	1	ver DIN 51720
Fr	Número de Froude	1	$\frac{\omega^2 R}{g_0}$
\vec{g}	Aceleración de la gravedad	$\frac{\text{m}}{\text{s}^2}$	$\frac{d^2 \vec{r}}{dt^2}$
H	Entalpía	J	$U + PV$
H_o	Poder calorífico superior	$\frac{\text{MJ}}{\text{kg}}$	ver DIN 51857

Símbolo	Término	Unidad SI	Definición
h	Contenido de hidrógeno	1	$\frac{m_H}{m}$
K	Coeficiente de equilibrio	1	Ec...
L	Longitud	m	DF
L	Longitud del reactor	m	Figura...
m	Masa	kg	DF
\dot{m}	Flujo de masa	$\frac{\text{kg}}{\text{s}}$	$\frac{m}{t}$
n	Velocidad de rotación	$\frac{1}{\text{s}}$	$\frac{\omega}{2\pi}$
n	Cantidad de materia	mol	DF
P	Presión	Pa	$\frac{\vec{F} \cdot \vec{n}}{A}$
Q	Calor	kJ	1. LT
T	Temperatura	K	DF
t	Tiempo	s	DF
x_i	Fracción de la cantidad de materia	1	$\frac{n_i}{n}$
V	Volumen	m^3	$\int dr^3$
\vec{u}	Velocidad	$\frac{\text{m}}{\text{s}}$	$(\frac{dr}{dt}, r \frac{dv}{dt}, \frac{dz}{dt})$
w_i	Fracción en masa del componente i	1	$\frac{m_i}{m_0}$
$w_{w,i}$	Contenido de humedad de la sustancia i	1	$\frac{m_{H_2O}}{m_{i,0}}$
Z	Factor de gases reales	1	$\frac{pv}{RT}$

Símbolos con letras griegas

Símbolo	Término	Unidad SI	Definición
α_{BET}	Factor de superficie	$\frac{\text{m}^2}{\text{g}}$	$(w_{F,waf})(A_{BET})$
β_i	Grado de formación del componente i	1	$\frac{m_i}{m_{bm,0}}$
γ	Wandhaftreibwinkel (Stahlblech)	1	Sección...
ϵ	Porosidad de la partícula	1	$1 - \frac{\rho_s}{\rho_w}$
η	mittlere Bettneigungswinkel (Stürzen)	1	Figura...

Símbolo	Término	Unidad SI	Definición
θ	Ángulo de inclinación de la cama	1	Figura...
θ_0	Ángulo superior de avalancha	1	Figura...
θ_U	Ángulo inferior de avalancha	1	Figura...
κ	Velocidad de calentamiento	$\frac{K}{s}$	$\frac{dT}{dt}$
ν	Coeficiente estequiométrico	1	ver DIN 13345
ρ_b	Densidad a granel	$\frac{kg}{m^3}$	$\frac{m_S}{V_S}$ (Sección...)
ρ_s	Densidad aparente	$\frac{kg}{m^3}$	$\frac{m_F}{V_P}$ (Sección...)
ρ_w	Densidad verdadera	$\frac{kg}{m^3}$	$\frac{m_F}{V_F}$ (Sección...)
τ	Tiempo adimensional	1	Ec....
Φ_V	Flujo volumétrico	$\frac{m^3}{s}$	$\frac{\Delta V}{\Delta t}$
ω	Velocidad angular	$\frac{1}{s}$	$\frac{dv}{dt}$

Subíndices

Subíndice Término

bm	materia orgánica
DR	Dubinin-Radushkevich
E	Experimental
g	Fase gaseosa
k	Condensado
Ma	Macroporos
P	Partícula
p	Poro
p	Piroizado
R	Reacción
t	Total
wf	Libre de agua
waf	Libre de agua y de ceniza
0	Estado de referencia

Superíndices

Superíndice	Término
-------------	---------

n	Coeficiente x
---	---------------

Abreviaturas

Abreviatura	Término
-------------	---------

1.LT	Primera ley de la termodinámica
DF	Dimensión fundamental
RFF	Racimos de fruta fresca

1. Introducción

La papa (*Solanum tuberosum L.*) representa el cuarto alimento más importante a nivel mundial (Cromme, 2010). El incremento en el rendimiento de este cultivo, esencial para la seguridad alimentaria, se ha basado tradicionalmente en el uso extensivo de agroinsumos químicos (Campos and Ortiz, 2019). Esta práctica ha incrementado los costos de producción y generado preocupaciones ambientales significativas (Li et al., 2020). En Colombia se cultivan aproximadamente 130 mil hectáreas y se producen cerca de 2.8 millones de toneladas de papa anualmente, con un rendimiento promedio nacional de 21.1 toneladas por hectárea (Ministerio de Agricultura y Desarrollo Rural, 2019).

Los avances en percepción remota han contribuido significativamente al desarrollo del fenotipado de alto rendimiento, al facilitar la caracterización no destructiva, automatizada y a mayor escala de variables asociadas al rendimiento de los cultivos (Lin et al., 2023). Las imágenes ópticas multiespectrales e hiperespectrales, en particular, destacan por su capacidad para reflejar las condiciones de crecimiento de los cultivos tanto a nivel espacial como temporal (Awad, 2019). En los últimos años la técnica de fenotipado de alto rendimiento mediante capturas visuales de vehículos aéreos no tripulados se ha implementado con diferentes fines como la identificación de factores de estrés abiótico, deficiencias nutricionales, tensiones bióticas, gestión de malezas, supervisión del desarrollo de las plantas y pronóstico de rendimiento (Li et al., 2019). Diversos autores (Gamon et al., 2016; Martín et al., 2014; Zarco-Tejada et al., 2005) han evidenciado que drones operados a baja altitud, equipados con cámaras RGB o multiespectrales, constituyen un medio efectivo para llevar a cabo el recuento de plántulas y estimar la cobertura del dosel en cultivos como algodón, sorgo y caña de azúcar.

La estimación temprana del rendimiento permite a los agricultores tomar decisiones informadas sobre la aplicación de fertilizantes, la planificación del riego y la gestión de pesticidas (?). Estas decisiones contribuyen a un uso más eficiente de los recursos, reducen la contaminación ambiental y pueden aumentar tanto los rendimientos como los ingresos económicos (Panda et al., 2010). En ese contexto, los desarrollos en fenotipado de alto rendimiento proveen herramientas para estimar el rendimiento, al caracterizar de manera exhaustiva y automatizada rasgos como la biomasa, el estado fitosanitario, la morfología de las plantas, la eficiencia fotosintética y la arquitectura radicular.

Estos vacíos metodológicos evidencian la necesidad de desarrollar enfoques más robustos para la estimación del rendimiento en el cultivo de papa, mediante el uso de modelos de aprendizaje automático que incorporen explícitamente la dimensión espacial y consideren variables morfológicas y fenológicas, las cuales han demostrado estar asociadas al rendimiento. Lo anterior implica establecer una metodología que utilice variables morfológicas del cultivo y que, además de ser robusta, sea asequible y replicable para su eventual implementación por parte del agricultor. Por consiguiente, los principales objetivos de este estudio fueron: 1) Determinar el aporte de la información de covariables como cantidad, altura y cobertura del dosel de plantas de papa (*S. tuberosum* L.); y 2) Implementar y evaluar los modelos Random Forest y Random Forest Espacial para la estimación de rendimiento de papa (*S. tuberosum* L.) a partir de imágenes ópticas de alta resolución espacial.

1.1. Antecedentes

cualquier cosa

1.2. Planteamiento del problema

cualquier cosa

1.2.1. Aproximación conceptual y metodológica

cualquier cosa

1.2.2. Pregunta de investigación

cualquier cosa

1.2.3. Objetivos

Objetivo General

cualquier cosa

Objetivos específicos

- cualquier cosa

- cualquier cosa

- cualquier cosa

1.2.4. Justificación

cualquier cosa

2. Marco Teórico y Estado del Arte

2.1. Uso de herramientas de fenotipado para la estimación de rendimiento

El fenotipado de plantas es el proceso de medición y análisis cuantitativo de características físicas y bioquímicas de las plantas (Akhtar et al., 2024). Las herramientas de fenotipado proporcionan datos objetivos sobre el comportamiento del cultivo (Coronado Aleans, 2022), facilitando la estimación del desempeño del rendimiento en los cultivos de plantas, ya que vincula directamente las características observadas con la productividad vegetal.

El fenotipado involucra la evaluación sistemática de diversos rasgos como densidad de plantas, altura de la planta, área foliar, biomasa aérea, estructura radicular y cobertura del dosel, características que reflejan el crecimiento, el desarrollo y el desempeño general de las plantas (Akhtar et al., 2024; Khuimphukhieo and da Silva, 2025).

La densidad de siembra, definida como el número de plantas establecidas por unidad de superficie, es un factor clave en la competencia por recursos como la luz, el agua y los nutrientes, afectando directamente el crecimiento y rendimiento de los cultivos. Una densidad óptima puede maximizar la eficiencia en la interceptación de la radiación fotosintéticamente activa y mejorar el uso del agua y los nutrientes, mientras que densidades excesivamente altas pueden generar competencia intraespecífica y reducir la biomasa individual de las plantas (Duan et al., 2025). Además, la densidad de siembra influye en la arquitectura del dosel, modificando el microclima del cultivo y afectando la incidencia de plagas y enfermedades, así como la eficiencia en el uso de la radiación y la producción de biomasa (Bussan et al., 2007).

La altura de las plantas es una variable estructural esencial en los ecosistemas agrícolas, ya que influye en la competencia por la luz, la eficiencia en la fotosíntesis y la distribución de biomasa dentro del dosel. Una mayor altura puede mejorar la captura de luz en sistemas de alta densidad, pero también puede incrementar la susceptibilidad al acame y afectar la estabilidad del cultivo (Ku et al., 2015). Esta variable está determinada por factores genéticos y ambientales, incluyendo la disponibilidad de agua y nutrientes, la radiación solar y la densidad de siembra (Jing et al., 2023). Además, la altura de las plantas juega un papel

2.2 Avances en percepción remota para el fenotipado de alto rendimiento5

importante en la modelización del balance energético del cultivo, afectando la evapotranspiración y la eficiencia en el uso del agua (Aase and Siddoway, 1980).

La cobertura del dosel, definida como la proyección vertical de las copas, coronas, brotes y hojas de las plantas sobre la superficie del suelo, es una variable clave en el desarrollo de los cultivos. Se expresa como una fracción o porcentaje del área de referencia y proporciona indicadores biofísicos esenciales como la densidad de plantas, la fenología, el índice de área foliar y el rendimiento (Fang et al., 2016). Así mismo, desempeña un papel crucial en el modelado de la radiación fotosintéticamente activa interceptada, estableciendo una relación directa con la capacidad fotosintética del follaje y la productividad de la vegetación (Fang et al., 2016). El monitoreo y seguimiento de las variaciones espaciales y temporales en la cobertura vegetal de los cultivos facilita la conexión con el estado fenológico y fisiológico del cultivo, contribuyendo así a una estimación más exacta del rendimiento (García-Martínez et al., 2020). Duan et al. (2017) resalta que la medición exacta de la tasa de emergencia y la uniformidad del cultivo es esencial para el fenotipado a nivel de campo, particularmente en el ámbito del mejoramiento de cultivos y la agricultura de precisión. El mismo autor indica que la reproducibilidad de los monitoreos manuales puede ser limitada, sobre todo cuando se aplican en experimentos de gran escala en agronomía y de mejoramiento, que podrían abarcar miles de parcelas de prueba, debido a la subjetividad que existe con el muestreador en campo.

2.2. Avances en percepción remota para el fenotipado de alto rendimiento

Tradicionalmente, las mediciones de variables usadas en fenotipado se han realizado mediante toma de datos *in-situ*, lo que implica procesos demandantes en términos de tiempo y recursos, como el conteo manual de plántulas emergentes o la evaluación visual de características como la altura y la cobertura del dosel. Sin embargo, estas técnicas enfrentan desafíos significativos (De et al., 2007), ya que tienden a ser laboriosas, inefficientes, destructivas y susceptibles a errores en el muestreo, lo que puede conllevar a resultados subjetivos o inexactos (Reynolds et al., 2000). Por ejemplo, el muestreo manual y destructivo de las plantas en el campo, donde se seleccionan muestras representativas en parcelas determinadas y se miden parámetros clave como la altura, biomasa y área foliar, requiere la destrucción parcial de la planta para obtener la información. Además, los resultados obtenidos mediante el muestreo de plantas individuales son insuficientes para representar las variaciones de todas las plantas dentro del campo (Ji et al., 2022).

La estimación remota del rendimiento consiste en identificar la relación entre las características espectrales del cultivo y los datos productivos (Battude et al., 2016). La percepción

remota permite recolectar información espectral del dosel vegetal de forma no destructiva desde plataformas terrestres, aéreas o satelitales. Los sensores registran la radiancia espectral que resulta de la interacción entre el dosel y la irradiancia solar, que incluye el resultado de procesos de absorción, transmisión y reflexión (Liu et al., 2017). El espectro reflejado por el dosel está estrechamente vinculado al desarrollo del cultivo, especialmente en el rango visible, influenciado por los pigmentos, y en el infrarrojo cercano (NIR) donde intervienen el tejido celular y la estructura del dosel (Luo et al., 2020).

La percepción remota de la vegetación se fundamenta en los principios físicos de la interacción de la energía electromagnética con la materia vegetal. ? explica que cuando la radiación solar incide sobre una hoja, esta puede ser absorbida, transmitida o reflejada. En el caso de la vegetación, la mayor absorción ocurre en las longitudes de onda azul (450 nm) y roja (670 nm) debido a la presencia de pigmentos como la clorofila, mientras que el verde (alrededor de 550 nm) se refleja en mayor medida, lo que da a las plantas su color característico. En el infrarrojo cercano (700–1300 nm), la reflectancia aumenta drásticamente por la dispersión interna de la radiación en los espacios intercelulares del mesófilo esponjoso, una propiedad clave para diferenciar vegetación sana de estresada o deteriorada (Fig. 2-1).

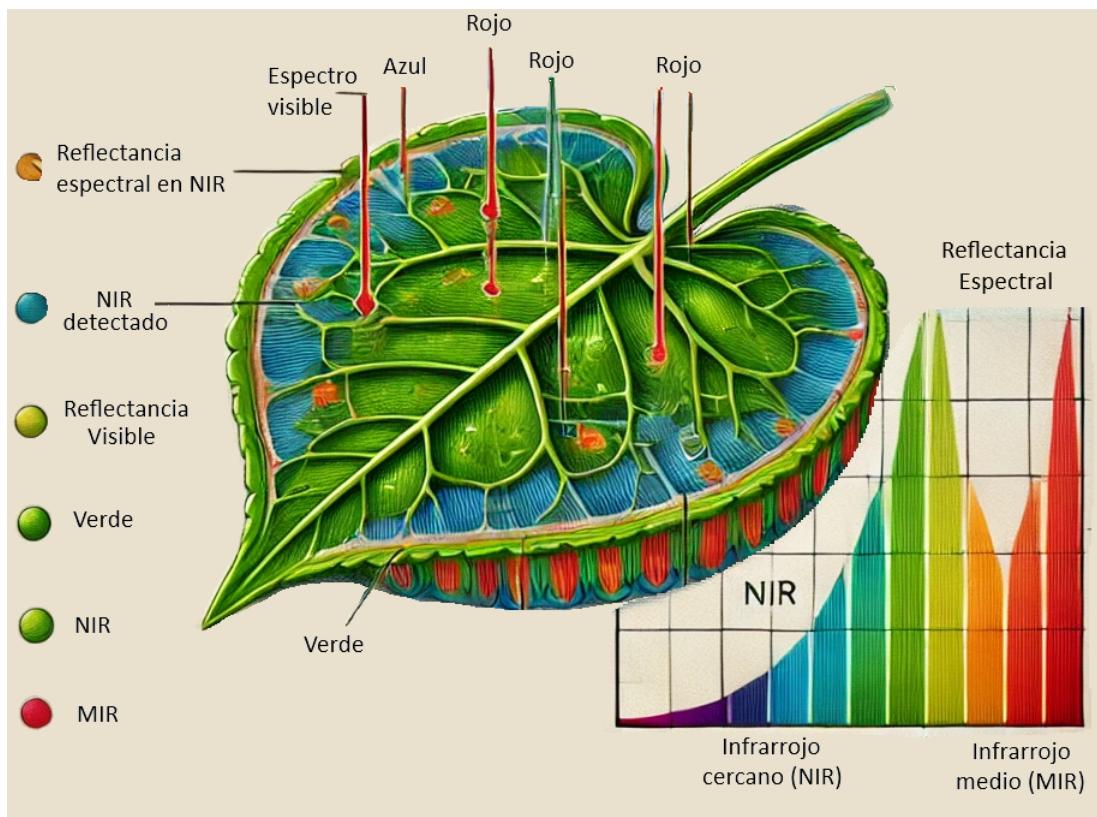


Figura 2-1.: Interacción de la energía electromagnética con la materia vegetal

2.2 Avances en percepción remota para el fenotipado de alto rendimiento⁷

? también señala que estas respuestas espectrales dependen de variables estructurales y biofísicas como la densidad foliar, el contenido de agua en las hojas y la arquitectura del dosel. Estas propiedades afectan la forma en que la vegetación interactúa con la radiación en distintas bandas espectrales. Por ejemplo, el contenido hídrico influye principalmente en el rango del infrarrojo medio (1300–2500 nm), donde se observa una fuerte absorción. Estos principios físicos permiten que sensores remotos, a bordo de satélites, aviones o vehículos aéreos no tripulados (UAV), capten respuestas espectrales que luego se traducen en información cuantitativa sobre el estado, estructura y funcionamiento de la vegetación (?).

Las imágenes multiespectrales permiten describir el desarrollo del cultivo, la radiación fotosintéticamente activa (PAR), la acumulación de biomasa (Bio) y el índice de área foliar (LAI), por medio de los índices de vegetación (IV) como el índice de vegetación de diferencia normalizada (NDVI), el índice de vegetación mejorado (ENDVI) y el índice de mejora de cultivo perpendicular (PCEI), entre otros (Awad, 2019). García-Martínez et al. (2020) evidencia que el uso de índices de vegetación como el exceso de verde (ExG), el índice de resistencia atmosférica visible (VARI) y el índice de diferencia normalizado verde-rojo (NGRDI) ayudan a estimar la fracción de cobertura de la vegetación en maíz (*Zea mays*).

Los desarrollos en percepción remota para el monitoreo de variables de cultivo han ido de la mano con la evolución de los UAV y la miniaturización de sensores, que se presentan como alternativas innovadoras para la medición de variables relacionadas con el crecimiento y desarrollo de los cultivos. Las plataformas UAV se han equipado con cámaras RGB, multiespectrales o hiperespectrales, sensores LiDAR y ultrasónicos. La integración de UAV con estos sensores ofrece ventajas respecto a la percepción remota satelital y la captura de datos *in situ*, como operación conveniente, alta flexibilidad, adaptabilidad considerable y costos reducidos, lo que ha generado un creciente interés en la comunidad de investigación agrícola (Holman et al., 2016). Algunas de estas configuraciones se han desplegado en una variedad de cultivos, tales como maíz, trigo, arroz, sorgo y soya, abordando así las necesidades específicas de distintas especies vegetales (Ji et al., 2022).

Sin embargo, la integración de sensores en plataformas UAV ha resultado en la generación de grandes volúmenes de datos cuyo análisis requiere enfoques computacionales avanzados que sean eficientes y escalables. En este contexto, los algoritmos de aprendizaje automático ofrecen ventajas notables frente a los métodos convencionales de análisis de imágenes para procesar, interpretar y extraer valor de esta información (Sharma et al., 2021). El aprendizaje automático (ML) ha surgido junto con las tecnologías de *big data* y la informática de alto rendimiento para crear nuevas oportunidades que permiten desentrañar, cuantificar y comprender procesos intensivos en datos en entornos agrícolas. Los modelos comúnmente usados para la estimación de rendimiento son las redes neuronales (NN), bosques aleatorios (RF), árboles de decisión (T) y regresión lineal (LR) (Liakos et al., 2018).

El algoritmo de aprendizaje automático que ha demostrado mejor desempeño para estimar rendimiento en cultivos es *Random Forest* (Breiman, 2001). Este modelo se basa en el principio de ensamblaje de múltiples árboles de decisión: cada árbol se construye con una muestra aleatoria de los datos y un subconjunto aleatorio de variables, lo que permite capturar la variabilidad y las complejidades inherentes al sistema. Para obtener la predicción final en problemas de regresión, se promedian las salidas de cada árbol, reduciendo el sobreajuste y aumentando la robustez del modelo (Fig. 2-2). Este método ha demostrado ser particularmente efectivo en la predicción del rendimiento de cultivos, al integrar diversas fuentes de datos, desde variables climáticas hasta características del suelo y manejo agrícola. Estudios como (Prasad et al., 2006; Hengl et al., 2015; Nussbaum et al., 2018) han evidenciado que el método *Random Forest* resulta prometedor para aplicaciones de estimación de rendimiento. No obstante, la mayoría de estas implementaciones se basa en enfoques de análisis píxel-apíxel que no incorporan explícitamente la dependencia espacial entre observaciones, lo cual limita su capacidad para capturar patrones espaciales relevantes. En contextos agrícolas, donde variables como el suelo, la topografía y las condiciones microclimáticas presentan estructuras espaciales definidas, esta omisión puede afectar la exactitud de las estimaciones.

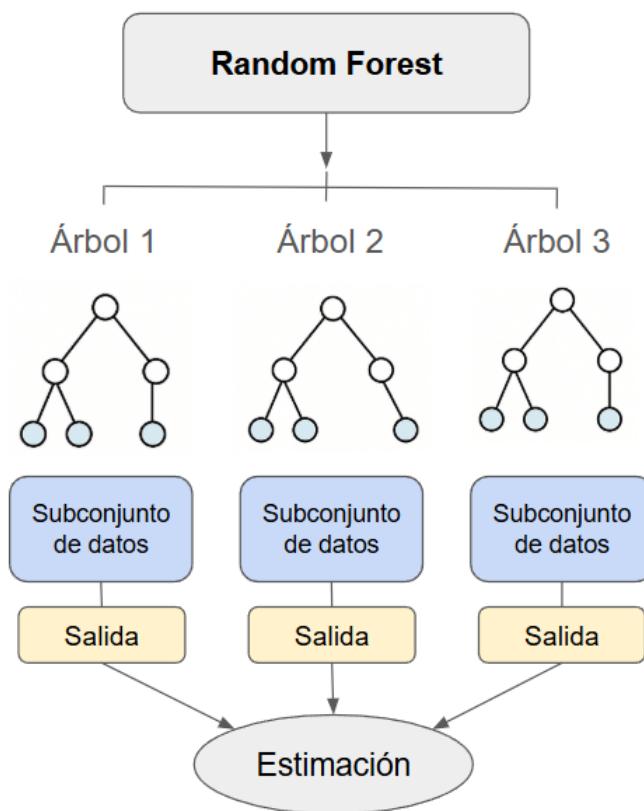


Figura 2-2.: Modelo Random Forest Tradicional

2.2 Avances en percepción remota para el fenotipado de alto rendimiento9

En respuesta a las limitaciones que presentan los algoritmos basados en análisis píxel a píxel, se han desarrollado variantes de estos modelos que integran explícitamente la dimensión espacial en el proceso de modelado. Una de las más destacadas es el *Random Forest Geoespacial* (Fig. 2-3), el cual incorpora la dependencia espacial entre observaciones para mejorar la precisión de las estimaciones (Hengl et al., 2018). Mientras que el *Random Forest* tradicional asume que las observaciones son independientes, en datos geoespaciales existe correlación espacial, lo cual determina que sitios cercanos tienden a tener valores similares. Incorporar esta información espacial puede mejorar significativamente la exactitud de las predicciones. El *Random Forest* adaptado para realizar estimaciones espaciales (RFsp) (Hengl et al., 2018; Nussbaum et al., 2018), por ejemplo, utiliza distancias entre los sitios muestreados como posibles variables explicativas, incorporando así los efectos de proximidad geográfica en el proceso de estimación. Se ha demostrado que RFsp puede generar estimaciones confiables de manera similar a las obtenidas mediante técnicas de kriging, con la ventaja de que no requiere suposiciones estadísticas rígidas sobre la distribución y estacionariedad de la variable objetivo (Hengl et al., 2018).

Random Forest Geoespacial

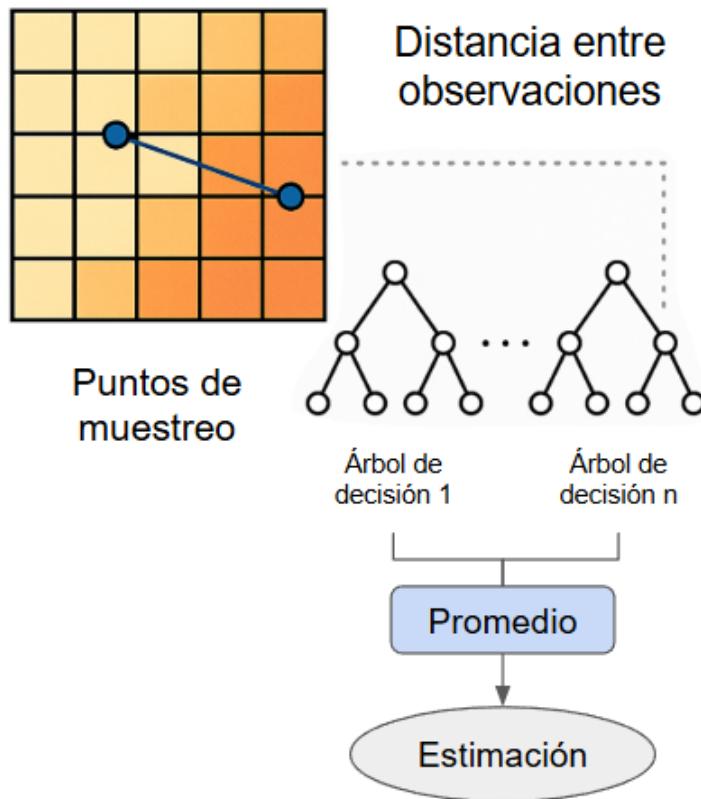


Figura 2-3.: Modelo Random Forest Geoespacial

Si bien se han documentado mejoras en tareas de estimación ambiental mediante enfoques espaciales como el *Random Forest Geoespacial* (Hengl et al., 2018), su aplicación en modelos de estimación de rendimiento en cultivos, particularmente en papa, continúa siendo limitada. La mayoría de los estudios de fenotipado de alto rendimiento orientados a la estimación de rendimiento emplea algoritmos de aprendizaje automático bajo esquemas de análisis píxel a píxel, sin incorporar de forma explícita la dependencia espacial entre observaciones. Esta omisión puede limitar la capacidad de los modelos para representar adecuadamente patrones espaciales asociados a la variabilidad de rasgos fenológicos y fisiológicos de las plantas (Fig. 2-4).

Hasta el momento, son escasos los estudios que integran esta dimensión espacial en la estimación de rendimiento en papa, y no se ha establecido de forma sistemática cómo se comparan estos enfoques espaciales con los modelos tradicionales en términos de exactitud temática. Esta limitación metodológica persiste junto con otros desafíos, como el uso de una única variable predictora, la escasa evaluación del efecto de la resolución espacial y de la integración de múltiples fuentes de datos sobre el desempeño de los modelos, así como la utilización exclusiva de imágenes satelitales con resoluciones espaciales limitadas (Makanza et al., 2018; Holman et al., 2016).

Además, muchas investigaciones se basan en datos adquiridos en un único momento del ciclo del cultivo, sin aprovechar la evolución temporal de los rasgos fenológicos, lo cual podría mejorar la precisión de las estimaciones (Ji et al., 2022).

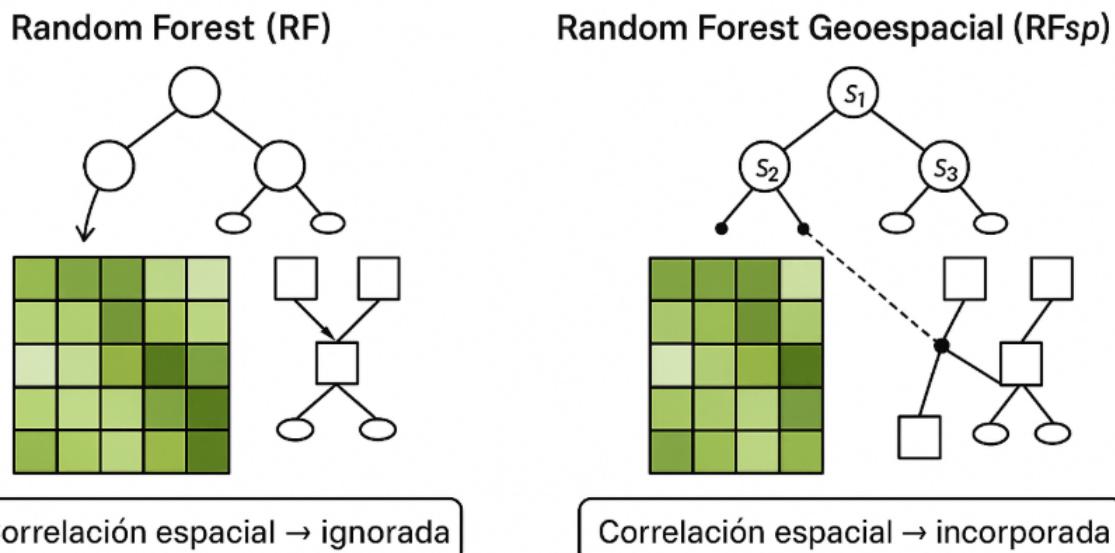


Figura 2-4.: Comparación RF con respecto al Rfsp

2.3. Estado del Arte

2.3.1. Estudios Globales y Regionales

- Tendencias en modelación del riesgo de incendios en regiones tropicales y andinas.
- Revisión de investigaciones recientes (2015-2025) sobre incendios en ecosistemas altoandinos.

2.3.2. Percepción remota en la modelación espacial del riesgo

- Percepción remota en la gestión de incendios
- Modelos espaciales de riesgo: probabilísticos, empíricos y basados en machine learning
- Escalas espacial y temporal en el análisis de incendios.

2.3.3. Monitoreo de incendios a partir de datos de percepción remota

- Integración de datos multifuente: MODIS, Landsat 8, Sentinel-1/2, GEDI
- Uso de índices espectrales y series temporales para monitoreo post-fuego
- Comparación entre enfoques ópticos, SAR y LiDAR

2.3.4. Modelos estadísticos y de aprendizaje automático en la estimación del riesgo de incendios

- Aplicación de Random Forest y regresiones espaciales para estimar riesgo.
- Validación de modelos con métricas de precisión espacial (AUC, Kappa, RMSE).
- Avances en modelación con GEE y Python

2.3.5. Avances en Colombia y los Andes tropicales

- Experiencias del IDEAM, Parques Nacionales y universidades nacionales
- Estudios sobre el Santuario de Fauna y Flora de Iguaque y páramos adyacentes.
- Vacíos de información y oportunidades para la modelación predictiva del riesgo.

3. Materiales y métodos

3.1. Área de estudio

La investigación se desarrolló en un lote cultivado con papa (*Solanum tuberosum* L.) ubicado en el municipio de Madrid, Cundinamarca, Colombia, en terrenos administrados por la empresa McCain Colombia. El área de estudio se localiza en las coordenadas 4.7758772° N y 74.2495896° W, a una altitud promedio de 2561 metros sobre el nivel del mar (Fig. 3-1). La recolección de datos *in situ* y de teledetección se realizó en un área total de 5 hectáreas.

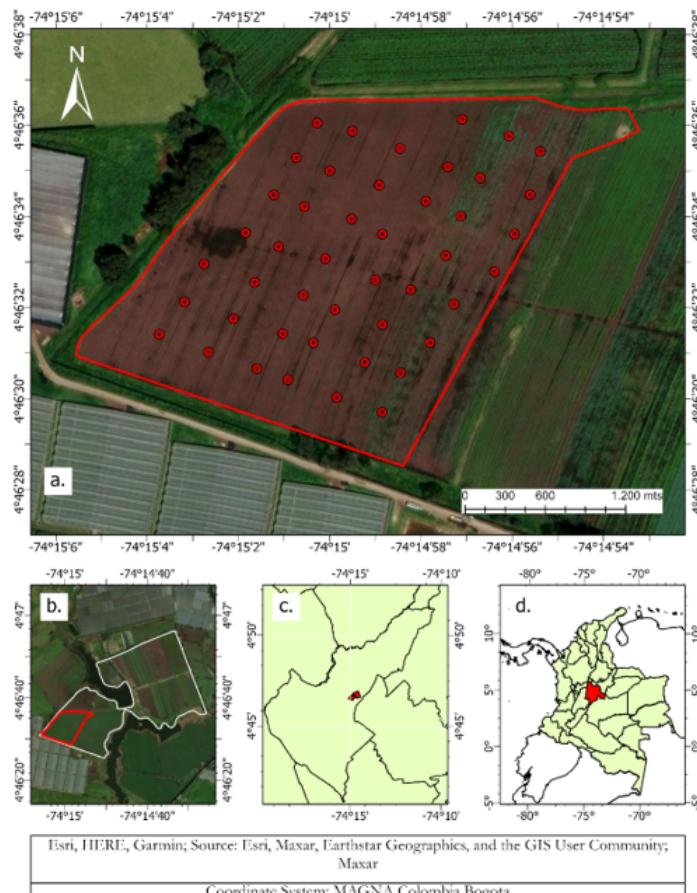


Figura 3-1.: Ubicación del área de estudio: (a) Área de estudio del cultivo de papa; (b) Finca comercial Palo Alto; (c) Ubicación de la finca en el municipio de Madrid; y (d) Ubicación en el departamento de Cundinamarca.

Los suelos del sitio de estudio corresponden a formaciones de origen fluviolacustre, con textura franco-arcillosa, caracterizada por una mezcla de partículas de arena, limo y arcilla. El clima de la zona corresponde a un régimen templado húmedo de montaña, con una temperatura media anual de aproximadamente 14 °C y una precipitación acumulada anual estimada en 562 mm. El ciclo del cultivo monitoreado se extendió entre el 8 de abril y el 6 de agosto de 2024. La variedad de papa empleada fue Andina, de uso común en los sistemas productivos de la región.

3.2. Diseño muestral

La recolección de datos de campo se realizó durante seis campañas. En cada una de ellas se capturaron imágenes multiespectrales (IMS) con el sensor a bordo del vehículo aéreo no tripulado (UAV) y se recolectaron datos *in situ* asociados a tres variables morfológicas (Tabla 3-1).

Tabla 3-1.: Fecha y datos *in situ* recolectados en cada campaña de muestreo.

Camp	Fecha	DDS IMS	Muestreo en tierra	Conteo de plantas	Altura de plantas	Cobertura del dosel	Escala BBCH
1	13 de abril	10	Sí	No	No	No	Brotación
2	04 de mayo	31	Sí	Sí	Sí	Sí	Desarrollo de las hojas
3	25 de mayo	52	Sí	Sí	Sí	Sí	Tuberización y floración
4	15 de junio	73	Sí	Sí	No	Sí	Formación del fruto
5	29 de junio	87	Sí	Sí	No	Sí	Maduración de frutos y semillas
6	03 de agosto	122	Sí	Sí	No	No	Senescencia

La primera campaña se realizó con el objetivo de capturar la información necesaria para la obtención del modelo digital de elevación del área de estudio. Esta campaña se llevó a cabo

después de la preparación del terreno y la siembra del cultivo, y antes de la emergencia de las plantas.

La segunda campaña se ejecutó cuando se observó una emergencia del cultivo superior al 70 % en el área de estudio; corresponde a la etapa de desarrollo vegetativo según la escala BBCH.

La tercera campaña se desarrolló durante el inicio de la tuberización y la floración de las plantas de papa.

En la cuarta y quinta campaña no se realizó el conteo de plantas, debido a que el cultivo ya había crecido lo suficiente como para que se solaparan las hojas entre plantas y se hubieran llenado las calles, dificultando la diferenciación individual en las imágenes capturadas con *UAV*. Por tal motivo, se decidió no tomar más datos de esa covariable en tierra. Durante la cuarta campaña el cultivo se encontraba en el estado de llenado de tubérculo; durante la quinta, en el estado de maduración, según la escala BBCH, con plantas en senescencia.

En la sexta y última campaña se recolectaron los datos de rendimiento del cultivo: número de plantas, número de tallos, tubérculos con longitud mayor a 90 mm, tubérculos con longitud igual o menor a 90 mm, peso de los tubérculos según su longitud y presencia de daños en los tubérculos muestreados. Estas mediciones se realizaron en los 45 puntos de muestreo.

3.3. Metodología

Para el desarrollo de la investigación se seleccionó un lote que tenía establecido el cultivo de papa, donde se optó por un área de aproximadamente 5 hectáreas para realizar la colección de los datos, por motivos de operatividad y costos. Se realizaron muestreos durante todo el desarrollo y maduración del tubérculo hasta llegar a la cosecha, con el fin de obtener información del estado del cultivo durante todas sus etapas de desarrollo.

De manera general, la metodología comprende seis etapas de trabajo (Fig. 3-2):

1. Recolección de datos
2. Preprocesamiento de datos
3. Procesamiento de datos
4. Análisis exploratorio de los datos
5. Implementación de los modelos

6. Validación de los modelos

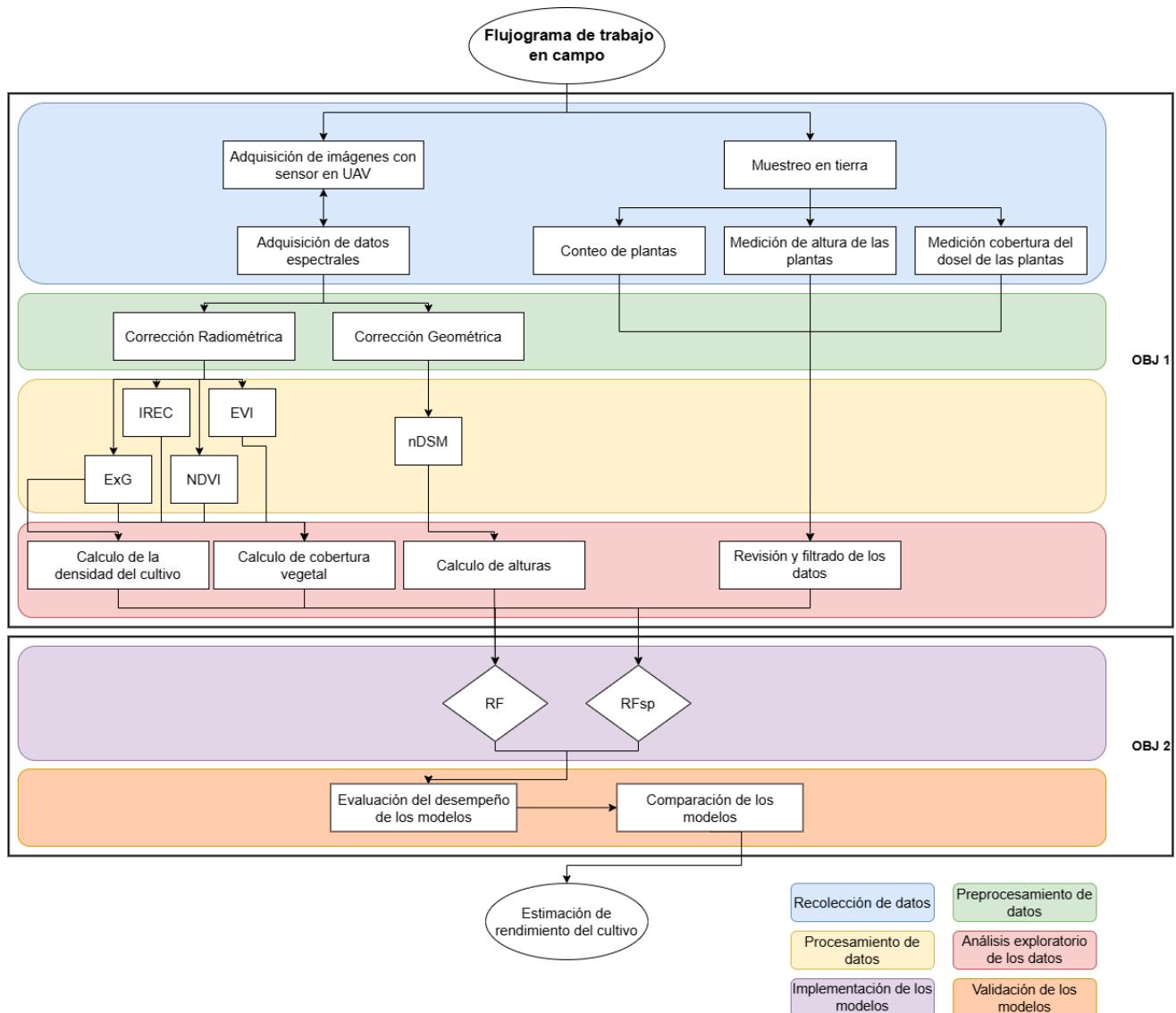


Figura 3-2.: Diagrama de flujo de las actividades desarrolladas.

3.4. Adquisición de datos e imágenes

3.4.1. Medición de datos en campo

La recolección de datos en campo y las variables morfológicas en tierra estuvieron a cargo del equipo de trabajo. Estas variables comprenden el establecimiento, la altura y la cobertura del dosel del cultivo de papa. Se ubicaron 45 sitios de muestreo (unidades espaciales de muestra) distribuidos en cuadrícula a lo largo del lote (Fig. 1). Cada unidad espacial de

muestra estuvo delimitada por un recuadro de PVC con dimensiones de 2 m por 2 m.

Para la medición del número de plantas establecidas en el cultivo, en cada unidad espacial de muestra se contó el total de plantas delimitadas por el recuadro. Para la medición de la altura de las plantas, se utilizó una cinta métrica para determinar la distancia desde el suelo hasta el ápice de la planta, sin estirarla. Se midieron las cuatro plantas más próximas a las esquinas del recuadro de PVC (Fig. 3-3), con el fin de obtener un valor promedio de altura.

Adicionalmente, se tomaron registros fotográficos de las plantas ubicadas en cada sitio de muestreo utilizando cámaras de uso común. Estas imágenes fueron analizadas mediante el software *ImageJ*, con el propósito de determinar la cobertura del dosel de las muestras.



Figura 3-3.: Unidad espacial de muestra de 2 x 2 metros. Los marcadores rojos indican las plantas a las que se les midió la altura.

3.4.2. Adquisición de imágenes multiespectrales

Figura 3-4.: Unidad espacial de muestra de 2 x 2 metros. Los marcadores rojos indican las plantas a las que se les midió la altura.

La cámara utilizada en este estudio cuenta con seis sensores de 1.6 MP y 12 bits, con filtros discretos de banda estrecha que capturan imágenes en longitudes de onda del visible al infrarrojo cercano (NIR), entre 475 y 860 nm, en las bandas multiespectrales, y 5.1 MP en la banda pancromática.

Además, la cámara está equipada con un sensor de luz descendente (*Downwelling Light Sensor*, DLS), el cual registra la irradiancia solar incidente para cada conjunto de imágenes y en las diferentes bandas espectrales, proporcionando una medición hemisférica sobre la superficie del detector.

Las características espectrales de las bandas de la cámara *RedEdge-P* se presentan en la Tabla 3-2.

Tabla 3-2.: Longitudes de onda y anchos de banda de la cámara *MicaSense RedEdge-P*.

Nombre de la banda	Centro (nm)	Ancho de banda (nm)
Azul	475	32
Verde	560	27
Rojo	668	16
Red Edge	717	12
Infrarrojo cercano (NIR)	842	57
Pancromático	634.2	463

3.5. Preprocesamiento de imágenes

Se aplicaron pasos de preprocesamiento radiométrico y geométrico a las imágenes con el objetivo de corregir los efectos relacionados con las características del sensor y las condiciones de adquisición, y de obtener mosaicos posicionalmente exactos.

3.5.1. Corrección radiométrica

Para realizar esta corrección se usó un panel de calibración de reflectancia (CRP) antes y después de realizar el vuelo. Adicionalmente se utilizó el sensor de luz descendente (DLS) para capturar la variación de la radiación durante el vuelo. Esto permitió convertir los píxeles de las imágenes en valores de radiancia espectral y reflectancia de superficie, ayudando a describir con precisión el material que componen los objetos de interés compensando las características del sensor, las condiciones climáticas y la iluminación.

3.5.2. Corrección geométrica

Para la corrección geométrica se ubicaron cinco puntos de control en tierra (GNSS), con el fin de poder ensamblar el ortomosaico a partir de las imágenes capturadas, esto se dio a través de un proceso de ortorrectificación que combina la técnica estructura a partir de movimiento (SfM) y multivisión estereoscópica (MVS). Adicionalmente, durante el vuelo se instaló una antena RTK, la cual permitió obtener coordenadas de cada centro de fotografía con alta exactitud geométrica en tiempo real.

3.6. Procesamiento de datos

3.6.1. Densidad de siembra y establecimiento del cultivo

El cálculo del índice verde excesivo (ExG) y el índice rojo excesivo (ExR) se usan para el recuento de plantas de papa en etapas de emergencia, ya que ayuda a diferenciar las plantas con respecto al suelo. El ExG se calcula de la siguiente manera:

$$ExG_{x,y} = 2 GREEN_{x,y} - RED_{x,y} - BLUE_{x,y} \quad (3-1)$$

Para diferenciar los objetos correspondientes a las plantas objetivo en la imagen, se aplicó el método de umbral de Otsu (Li et al., 2019), con el fin de transformar la imagen en escala de grises en una imagen binaria, asignando el valor 0 a los píxeles correspondientes al suelo desnudo. Otra manera de distinguir las plantas del suelo consiste en el uso del índice *ExG-ExR*, el cual convierte las intensidades de los píxeles de las plantas en valores positivos y el fondo en valores negativos, utilizando el valor 0 como umbral de separación.

Este último índice se calcula mediante la Ecuación 3-2.

$$(ExG - ExR)_{x,y} = 3 GREEN_{x,y} - 2,4 RED_{x,y} - BLUE_{x,y} \quad (3-2)$$

Donde X y Y son las coordenadas del pixel especificado.

3.6.2. Altura del cultivo

El cálculo de la altura del cultivo se realizó utilizando los modelos digitales de superficie (MDS) y de terreno (MDT) obtenidos en la etapa de preprocesamiento, los cuales permitieron generar el modelo digital de superficie normalizado (nDSM), que se asocia directamente con la altura del cultivo. La altura del cultivo se estimó a partir de la diferencia entre el modelo digital de superficie y el modelo digital de terreno derivados de las imágenes adquiridas con el UAV.

El cálculo se realizó mediante la Ecuación 3-3.

$$nDSM = DSM - DEM \quad (3-3)$$

3.6.3. Cobertura del dosel

A partir de las imágenes se generaron superficies con diferentes índices de vegetación en función del estado de desarrollo del cultivo. Lo anterior se debe a que algunos índices tienden a saturarse en condiciones de alta cobertura del dosel.

Según la base de datos de índices (*Index Database*) y la literatura revisada, los índices utilizados para el cálculo de la cobertura del dosel en las plantas se presentan en la Tabla 3-3.

Tabla 3-3.: Índices espectrales.

Índice espectral	Siglas	Fórmula	Referencia
Índice de vegetación de diferencia normalizada	NDVI	$\frac{NIR - RED}{NIR + RED}$	(Rouse et al., 1974)
Índice verde excesivo	ExG	$2GREEN - RED - BLUE$	(Woebbecke et al., 1995)
Índice de clorofila RedEdge	IREC	$\frac{NIR}{REDEDGE} - 1$	(Gitelson et al., 2003)
Índice de vegetación mejorado	GCI	$\frac{NIR}{GREEN} - 1$	(Gitelson et al., 2003)

NIR: reflectancia en la banda del infrarrojo cercano, RED: reflectancia en la banda del rojo, GREEN: reflectancia en la banda del verde, BLUE: reflectancia en la banda del azul.

3.7. Análisis de datos

4. Resultados

4.1. Resultados del preprocesamiento de los datos

4.2. Caracterización de las variables predictoras

5. Discusión

- 5.1. Cumplimiento de objetivos**
- 5.2. Ventajas y limitaciones de la metodología utilizada**
- 5.3. Aportes al conocimiento en Geomática**

6. Conclusiones

A. Anexo: Nombrar el anexo A de acuerdo con su contenido

Los Anexos son documentos o elementos que complementan el cuerpo de la tesis o trabajo de investigación y que se relacionan, directa o indirectamente, con la investigación, tales como acetatos, cd, normas, etc.

B. Anexo: Nombrar el anexo B de acuerdo con su contenido

A final del documento es opcional incluir índices o glosarios. Éstos son listas detalladas y especializadas de los términos, nombres, autores, temas, etc., que aparecen en el mismo. Sirven para facilitar su localización en el texto. Los índices pueden ser alfabéticos, cronológicos, numéricos, analíticos, entre otros. Luego de cada palabra, término, etc., se pone coma y el número de la página donde aparece esta información.

C. Anexo: Nombrar el anexo C de acuerdo con su contenido

MANEJO DE LA BIBLIOGRAFÍA: la bibliografía es la relación de las fuentes documentales consultadas por el investigador para sustentar sus trabajos. Su inclusión es obligatoria en todo trabajo de investigación. Cada referencia bibliográfica se inicia contra el margen izquierdo.

La NTC 5613 establece los requisitos para la presentación de referencias bibliográficas citas y notas de pie de página. Sin embargo, se tiene la libertad de usar cualquier norma bibliográfica de acuerdo con lo acostumbrado por cada disciplina del conocimiento. En esta medida es necesario que la norma seleccionada se aplique con rigurosidad.

Es necesario tener en cuenta que la norma ISO 690:1987 (en España, UNE 50-104-94) es el marco internacional que da las pautas mínimas para las citas bibliográficas de documentos impresos y publicados. A continuación se lista algunas instituciones que brindan parámetros para el manejo de las referencias bibliográficas:

Institución	Disciplina de aplicación
Modern Language Association (MLA)	Literatura, artes y humanidades
American Psychological Association (APA)	Ambito de la salud (psicología, medicina) y en general en todas las ciencias sociales
Universidad de Chicago/Turabian	Periodismo, historia y humanidades.
AMA (Asociación Médica de los Estados Unidos)	Ambito de la salud (psicología, medicina)
Vancouver	Todas las disciplinas
Council of Science Editors (CSE)	En la actualidad abarca diversas ciencias
National Library of Medicine (NLM) (Biblioteca Nacional de Medicina)	En el ámbito médico y, por extensión, en ciencias.
Harvard System of Referencing Guide	Todas las disciplinas
JabRef y KBibTeX	Todas las disciplinas

Para incluir las referencias dentro del texto y realizar lista de la bibliografía en la respectiva sección, puede utilizar las herramientas que Latex suministra o, revisar el instructivo desarrollado por el Sistema de Bibliotecas de la Universidad Nacional de Colombia¹, disponible en la sección "Servicios", opción "Trámites" enlace .Entrega de tesis".

¹Ver: www.sinab.unal.edu.co

Bibliografía

- Aase, J. K. and Siddoway, F. H. (1980). Stubble height effects on seasonal microclimate, water balance, and plant development of no-till winter wheat. *Agricultural Meteorology*, 21(1):1–10.
- Akhtar, M. S., Zafar, Z., Nawaz, R., and Fraz, M. M. (2024). Unlocking plant secrets: A systematic review of 3d imaging in plant phenotyping techniques. *Computers and Electronics in Agriculture*, 222.
- Awad, M. M. (2019). Toward precision in crop yield estimation using remote sensing and optimization techniques. *Agriculture (Switzerland)*, 9(3).
- Battude, M., Al Bitar, A., Morin, D., Cros, J., Huc, M., Marais Sicre, C., Le Dantec, V., and Demarez, V. (2016). Estimating maize biomass and yield over large areas using high spatial and temporal resolution sentinel-2 like remote sensing data. *Remote Sensing of Environment*, 184:668–681.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1):5–32.
- Bussan, A. J., Mitchell, P. D., Copas, M. E., and Drilias, M. J. (2007). Evaluation of the effect of density on potato yield and tuber size distribution. *Crop Science*, 47(6):2462–2472.
- Campos, H. and Ortiz, O. (2019). *The Potato Crop: Its Agricultural, Nutritional and Social Contribution to Humankind*. Springer International Publishing.
- Coronado Aleans, V. (2022). Fenotipado de alto rendimiento mediante el análisis de imágenes digitales en raíces de maíz (*zea mays l.*).
- Cromme, N. (2010). *Strengthening potato value chains: Technical and policy options for developing countries*. Food and Agriculture Organization of the United Nations and the Common Fund for Commodities.
- De, A., Casa, L., Ovando, G., Bressanini, L., Martínez, J., Ibarra, E., and Rodríguez, (2007). El índice de área foliar en papa estimado a partir de la cobertura del follaje (leaf area index in potato estimate from canopy cover). *Revista de Investigaciones Agropecuarias*. Artículo en español.

- Duan, M., Han, C., Zhang, X., Wei, Z., Wang, Z., and Zhang, B. (2025). Spatial and temporal dynamics of photosynthetically active radiation in crops: Effects of canopy structure on yield. *Agronomy*, 15(4).
- Duan, T., Zheng, B., Guo, W., Ninomiya, S., Guo, Y., and Chapman, S. C. (2017). Comparison of ground cover estimates from experiment plots in cotton, sorghum and sugarcane based on images and ortho-mosaics captured by uav. *Functional Plant Biology*, 44(1):169–183.
- Fang, S., Tang, W., Peng, Y., Gong, Y., Dai, C., Chai, R., and Liu, K. (2016). Remote estimation of vegetation fraction and flower fraction in oilseed rape with unmanned aerial vehicle data. *Remote Sensing*, 8(5).
- Gamon, J. A., Huemmrich, K. F., Wong, C. Y. S., Ensminger, I., Garrity, S., Hollinger, D. Y., Noormets, A., and Peñuelas, J. (2016). A remotely sensed pigment index reveals photosynthetic phenology in evergreen conifers. *Proceedings of the National Academy of Sciences of the United States of America*, 113(46):13087–13092.
- García-Martínez, H., Flores-Magdaleno, H., Khalil-Gardezi, A., Ascencio-Hernández, R., Tijerina-Chávez, L., Vázquez-Peña, M. A., and Mancilla-Villa, O. R. (2020). Estimación de la fracción de cobertura de la vegetación en maíz (*zea mays*) mediante imágenes digitales tomadas por un vehículo aéreo no tripulado (uav). *Revista Fitotecnia Mexicana*, 43(4):399–407. Artículo en español.
- Hengl, T., Heuvelink, G. B. M., Kempen, B., Leenaars, J. G. B., Walsh, M. G., Shepherd, K. D., Sila, A., MacMillan, R. A., De Jesus, J. M., Tamene, L., and Tondoh, J. E. (2015). Mapping soil properties of africa at 250 m resolution: Random forests significantly improve current predictions. *PLoS ONE*, 10(6).
- Hengl, T., Nussbaum, M., Wright, M. N., Heuvelink, G. B. M., and Gräler, B. (2018). Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables. *PeerJ*, 6(8).
- Holman, F. H., Riche, A. B., Michalski, A., Castle, M., Wooster, M. J., and Hawkesford, M. J. (2016). High throughput field phenotyping of wheat plant height and growth rate in field plot trials using uav based remote sensing. *Remote Sensing*, 8(12).
- Ji, Y., Chen, Z., Cheng, Q., Liu, R., Li, M., Yan, X., Li, G., Wang, D., Fu, L., Ma, Y., Jin, X., Zong, X., and Yang, T. (2022). Estimation of plant height and yield based on uav imagery in faba bean (*vicia faba l.*). *Plant Methods*, 18(1).
- Jing, F., Shi, S., A, Y., Guan, J., Lu, B., Wu, B., Wang, W., Ma, R., and Nan, P. (2023). Analysis of phenotypic and physiological characteristics of plant height difference in alfalfa. *Agronomy*, 13(7).

- Khuimphukhieo, I. and da Silva, J. A. (2025). Unmanned aerial systems (uas)-based field high throughput phenotyping (htp) as plant breeders' toolbox: A comprehensive review. *Smart Agricultural Technology*, 11.
- Ku, L., Zhang, L., Tian, Z., Guo, S., Su, H., Ren, Z., Wang, Z., Li, G., Wang, X., Zhu, Y., Zhou, J., and Chen, Y. (2015). Dissection of the genetic architecture underlying the plant density response by mapping plant height-related traits in maize (*zea mays l.*). *Molecular Genetics and Genomics*, 290(4):1223–1233.
- Li, B., Xu, X., Han, J., Zhang, L., Bian, C., Jin, L., and Liu, J. (2019). The estimation of crop emergence in potatoes by uav rgb imagery. *Plant Methods*, 15(1).
- Li, B., Xu, X., Zhang, L., Han, J., Bian, C., Li, G., Liu, J., and Jin, L. (2020). Above-ground biomass estimation and yield prediction in potato by using uav-based rgb and hyperspectral imaging. *ISPRS Journal of Photogrammetry and Remote Sensing*, 162:161–172.
- Liakos, K. G., Busato, P., Moshou, D., Pearson, S., and Bochtis, D. (2018). Machine learning in agriculture: A review. *Sensors (Switzerland)*, 18(8).
- Lin, Y., Li, S., Duan, S., Ye, Y., Li, B., Li, G., Lyv, D., Jin, L., Bian, C., and Liu, J. (2023). Methodological evolution of potato yield prediction: A comprehensive review. *Frontiers in Plant Science*, 14.
- Liu, N., Budkewitsch, P., and Treitz, P. (2017). Examining spectral reflectance features related to arctic percent vegetation cover: Implications for hyperspectral remote sensing of arctic tundra. *Remote Sensing of Environment*, 192:58–72.
- Luo, S., He, Y., Li, Q., Jiao, W., Zhu, Y., and Zhao, X. (2020). Nondestructive estimation of potato yield using relative variables derived from multi-period lai and hyperspectral data based on weighted growth stage. *Plant Methods*, 16(1).
- Makanza, R., Zaman-Allah, M., Cairns, J. E., Magorokosho, C., Tarekegne, A., Olsen, M., and Prasanna, B. M. (2018). High-throughput phenotyping of canopy cover and senescence in maize field trials using aerial digital canopy imaging. *Remote Sensing*, 10(2).
- Martín, P., Zarco-Tejada, P. J., González, M. R., and Berjón, A. J. (2014). Using hyperspectral remote sensing to map grape quality in “tempranillo” vineyards affected by iron deficiency chlorosis. Disponible en: <https://www.researchgate.net/publication/39569434>.
- Ministerio de Agricultura y Desarrollo Rural (2019). Estrategia de ordenamiento de la producción cadena productiva de la papa y su industria. Bogotá, Colombia.

- Nussbaum, M., Spiess, K., Baltensweiler, A., Grob, U., Keller, A., Greiner, L., Schaepman, M. E., and Papritz, A. (2018). Evaluation of digital soil mapping approaches with large sets of environmental covariates. *SOIL*, 4(1):1–22.
- Panda, S. S., Ames, D. P., and Panigrahi, S. (2010). Application of vegetation indices for agricultural crop yield prediction using neural network techniques. *Remote Sensing*, 2(3):673–696.
- Prasad, A. M., Iverson, L. R., and Liaw, A. (2006). Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. *Ecosystems*, 9(2):181–199.
- Reynolds, C. A., Yitayew, M., Slack, D. C., Hutchinson, C. F., Huete, A., and Petersen, M. S. (2000). Estimating crop yields and production by integrating the fao crop specific water balance model with real-time satellite data and ground-based ancillary data. *International Journal of Remote Sensing*, 21(18):3487–3508.
- Sharma, A., Jain, A., Gupta, P., and Chowdary, V. (2021). Machine learning applications for precision agriculture: A comprehensive review. *IEEE Access*, 9:4843–4873.
- Zarco-Tejada, P. J., Berjón, A., López-Lozano, R., Miller, J. R., Martín, P., Cachorro, V., González, M. R., and De Frutos, A. (2005). Assessing vineyard condition with hyperspectral indices: Leaf and canopy reflectance simulation in a row-structured discontinuous canopy. *Remote Sensing of Environment*, 99(3):271–287.