

UNIVERSIDAD DE COSTA RICA
SISTEMA DE ESTUDIOS DE POSGRADO

SISTEMA DE MONITOREO ESPACIO-TEMPORAL DEL CULTIVO DE CAÑA DE AZÚCAR (*Saccharum officinarum*), A PARTIR DE INFORMACIÓN SATELITAL, EN COOPEVICTORIA R.L. GRECIA, COSTA RICA

Trabajo final de investigación aplicada sometido a la consideración de la Comisión del Programa de Estudios de Posgrado en Geografía para optar al grado y título de Maestría Profesional en Sistemas de Información Geográfica y Teledetección.

BRYAN ALEMÁN MONTES

Ciudad Universitaria Rodrigo Facio, Costa Rica

2020

Dedicatoria

**A Dios.
A mis padres,
Norma y Pedro,
quienes con sacrificio y esfuerzo
han allanado mi camino.**

Agradecimientos

Un especial agradecimiento a mi familia, especialmente a mis hermanos Ginneth, Yeison, Kendy, Fernando y mi sobrina Jimena, bastión fundamental en mi vida de quienes he recibido soporte en todo momento.

Agradezco a Cinthya por su ayuda y la comprensión durante este proceso, sus palabras de motivación y aliento fueron fundamentales.

Quiero externar un especial agradecimiento al Dr. Carlos Henríquez por todo el apoyo y respaldo brindado en esta investigación y durante mi desarrollo profesional en el campo de las herramientas geoespaciales aplicadas en la agricultura.

A las personas que forma parte del departamento técnico de CoopeVictoria R.L., particularmente a los ingenieros: Marvin Ovando, Gerardo Fonseca, Alfonso Cubillo, Paul Montoya y Kevin Núñez, por haberme abierto las puertas para el desarrollo de esta y otras investigaciones.

Al M.Sc. José Antonio Guzmán y el M.Sc. Rubén Martínez por sus observaciones y recomendaciones para el cumplimiento de estos objetivos.

Gracias a M.Sc. Róger Fallas, M.Sc. Gabriel Garbanzo y M.Sc. José Andrés Búcaro por las recomendaciones y consejos durante el desarrollo de este proyecto.

A todos esos profesores que han contribuido con mi formación académica, particularmente a quienes me brindaron las herramientas para la conclusión de esta investigación.

A los compañeros generación con quienes disfruté las noches de clases y los momentos extraclase, de todos me quedan aprendizajes y buenos recuerdos.

¡Muchas Gracias!

“La aplicación del SIG está limitada solo a la imaginación de quienes los usan”
Jack Dangermond

“Este trabajo final de investigación aplicada fue aceptado por la Comisión del Programa de Estudios de Posgrado en Geografía de la Universidad de Costa Rica, como requisito parcial para optar al grado y título de Maestría Profesional en **Sistemas de Información Geográfica y Teledetección**”



M.Sc. Melvin Lizano Araya
Representante del Decano Sistema de Estudios de Posgrado



PhD. Carlos Henríquez Henríquez
Profesor Guía



M.Sc. José Antonio Guzmán Álvarez
Lector



M.Sc. Rubén Martínez Barbachano
Lector



M.Sc. Jonnathan Reyes Chaves
Representante del Director del Programa de Posgrado



Bryan Alemán Montes
Sustentante

Índice General

Dedicatoria.....	ii
Agradecimientos	iii
Hoja de Aprobación	iv
Resumen.....	vii
Índice de Figuras	viii
Índice de Cuadros.....	ix
Lista de Acrónimos	xi
Capítulo 1: Introducción.....	1
1.1. Introducción.....	1
1.2. Justificación.....	2
1.3. Problema de Investigación.....	3
1.4. Objetivos.....	4
1.4.1. Objetivo general.....	4
1.4.2. Objetivos específicos.....	4
1.5. Delimitación espacial y temporal	4
1.6. Alcances del proyecto	5
1.7. Limitaciones	6
Capítulo 2: Marco Conceptual	8
2.1. Características el cultivo caña de azúcar	8
2.2. Principios de la teledetección	9
2.3. Comportamientos espectrales de las coberturas vegetales	10
2.4. Índices de vegetación como indicadores del estado de las plantas	11
2.5. Transferencia de información geoespacial rápida y oportuna	13
Capítulo 3: Marco Metodológico.....	15
3.1 Elección del sensor satelital.....	15
3.2 Relación entre índices de vegetación y variables de productividad de caña de azúcar.....	20
3.3 Construcción de modelos de estimación de variables de rendimiento.....	25

3.4	Construcción del geoportal para transferencia de información.....	29
Capítulo 4: Resultados y Discusión.....		34
4.1.	Identificación del sensor satelital de mayor viabilidad para establecer el sistema de monitoreo	34
4.2.	Índices de vegetación como indicadores de rendimiento del cultivo de la caña de azúcar.....	38
4.3	Modelos de regresión para la estimación de variables de rendimiento.....	45
4.4.	SMS-CoopeVictoria R.L.: Sistema de Monitoreo Satelital de las Plantaciones de Caña de Azúcar de CoopeVictoria R.L.....	53
Capítulo 5: Conclusiones y Recomendaciones		58
5.1.	Conclusiones.....	58
5.2.	Recomendaciones	60
Bibliografía		62
Anexos		67

Resumen

El uso de información satelital presentó resultados satisfactorios para el monitoreo y estimación de variables de rendimiento en plantaciones de caña de azúcar en CoopeVictoria R.L. Al evaluar los sensores satelitales MODIS, Landsat Oli 8 y Sentinel 2 se logró identificar este último como el sensor satelital, de acceso libre, con mayores cualidades para el desarrollo del sistema de monitoreo satelital en las plantaciones de caña de azúcar de la Cooperativa.

Con la información disponible del sensor Sentinel 2 se logró definir los meses de agosto y noviembre como los más adecuados para la estimación de variables de rendimientos a partir de índices de vegetación, ya que únicamente esos meses tenían información representativa para las zafras en estudio (2016-2017 hasta 2019-2020).

La combinación de índices de vegetación (EVI y SAVI) con variables históricas de rendimiento permitieron construir un modelo de estimación de rendimiento de campo con un RMSE de 7,10 ton/ha, un coeficiente de correlación (R^2) de 0,62 y un promedio absoluto del error (MAE) de 5,58 ton/ha. Además, se logró estimar el total de toneladas de caña de azúcar producidas con un error aproximado del 0,48%.

La variable rendimiento industrial no presentó resultados satisfactorios en la estimación, lo cual confirma lo identificado en otras investigaciones donde las variables relacionadas con la calidad del cultivo no se logran relacionar con las respuestas espectrales.

Finalmente, se logró poner a disposición de los tomadores de decisiones un mapa web con la información disponible para las fincas en estudio, asimismo en este sitio web se puede compartir información de índices de vegetación para apoyar el proceso de toma de decisiones de las actividades agrícolas. Los resultados de esta investigación son una herramienta que colabora con los esfuerzos de desarrollar los principios de agricultura de precisión dentro de las plantaciones de caña de azúcar propiedad de CoopeVictoria R.L.

Índice de Figuras

Figura 1. CoopeVictoria R.L., 2018: (A) Distribución de las plantaciones de caña de azúcar por Finca. (B) Porcentaje de área según variedad. (C) Porcentaje del área según la dimensión de la Finca.....	5
Figura 2. Evolución fenológica del cultivo de la caña de azúcar.....	8
Figura 3. Espectro electromagnético.....	9
Figura 4. Firmas espectrales de algunas coberturas sobre la superficie terrestre.	10
Figura 5. Contraste espectral de la vegetación sana, vegetación estresada y suelo en las bandas Rojo e Infrarrojo Cercano.	12
Figura 6. (A) Interacción Cliente-Servidor durante el proceso de consulta de mapas web. (B) Interacción de múltiples clientes con el servidor.....	14
Figura 7. Interfaz de descarga de MODISStp v1.3.9. La descarga se realizó definiendo la categoría, el tipo de producto, los índices espectrales (recuadro de la derecha), usuario y contraseña, periodo de descarga, localización de área de interés (Tiles), proyección, formato y ruta de salida.	17
Figura 8. Ejecución de corrección atmosférica Sen2cor a través de la línea de comandos de Windows.	19
Figura 9. Modelo de corrección atmosférica ejecutado a partir del método Sen2cor para las imágenes satelitales Sentinel 2.....	20
Figura 10. Flujo metodológico para definir los índices de vegetación (IV) que mejor correlación tiene con las variables de producción.	22
Figura 11. Protocolo de activación del localhost:8888 para el desarrollo del geoportal....	30
Figura 12. Encabezado del archivo .html tiene los detalles de la ventana home cuando se accede al geoportal.	31
Figura 13. (A) Error absoluto en hectáreas al calcular el área de las fincas a partir del número de píxeles extraídos por sensor para cada unidad productiva. (B) Comparación del área real de las fincas contra las áreas calculadas por el número de píxeles.	34
Figura 14. Variación espacial de los píxeles extraídos por finca para los tres sensores satelitales en estudio.	36
Figura 15. Variación temporal del NDVI en las fincas incluidas en la investigación. (A) Año 2017. (B) Año 2018. En ambos casos las líneas grises representan las fincas y la línea negra indica el promedio del NDVI considerando todas las fincas. Las áreas remarcadas corresponden a los meses agosto (rojo) noviembre(azul).....	45
Figura 16. (A) Coeficiente de correlación entre los rendimientos de campo y real. El rendimiento estimado se obtuvo ejecutando las regresiones sobre el conjunto de datos. (B) Comparación del valor real y los estimados en rendimiento de campo.	48

Figura 17. (A) Correlación entre las toneladas totales reales y las estimadas a partir de los modelos de regresión PROD Agosto y PROD Final, los datos se obtuvieron multiplicando la estimación de los modelos de regresión por la respectiva área. (B) Comparación del valor real y los estimados en toneladas métricas totales.	49
Figura 18. Ventana Principal al acceder al link https://bryaleman.github.io/coopevictoriarl/	53
Figura 19. Ventana SMS-CoopeVictoria R.L. donde se muestran los detalles de las capas de información compartidas a los usuarios.	54
Figura 20. Uso de la herramienta selección de finca. Acercamiento sobre la finca ASSMAN y etiquetado mediante popupText.	55
Figura 21. Panel de control de capas, donde se puede activar y desactivas las capas según sea el interés. La línea tenue de color negro representa el límite de la finca.	55
Figura 22. Disposición de los popupText según sea la capa que de esta utilizando en el panel de capas.....	56
Figura 23. Variabilidad espacial de NDVI del 22/11/2019. Se puede notar la variación en los valores NDVI, cada cuadro posee un área de 100 m ²	56
Figura 24. Ventana Contactos donde se muestra la información relacionada al proyecto de investigación y contactos de las personas participantes.	57

Índice de Cuadros

Cuadro 1. Características básicas de los sensores satelitales: MODIS, Landsat OLI 8 y Sentinel 2A. Resaltadas están las características de las bandas utilizadas para calcular los índices de vegetación.	16
Cuadro 2: Características de los índices de vegetación utilizados para evaluar la evolución espacio-temporal de las plantaciones de caña de azúcar en CoopeVictoria R.L.....	23
Cuadro 3. Fechas con imágenes satelitales disponibles del sensor Sentinel 2A y Sentinel 2B para el área de estudio por mes y año.	38
Cuadro 4. Correlación de Spearman entre las variables de rendimiento y los índices de vegetación calculados para cada imagen Sentinel 2A y Sentinel 2B, en los años 2016, 2017 y 2018. Solo se muestran las correlaciones superiores a 0,6.	40
Cuadro 5. Correlación de Spearman entre las variables de rendimiento y el valor máximo obtenido en cada índice de vegetación calculado con Sentinel 2A y Sentinel 2B, según año (2016, 2017 y 2018). Solo se muestran las correlaciones superiores a 0,6.	41
Cuadro 6. Correlación de Spearman entre las variables de rendimiento y el valor máximo obtenido en cada índice de vegetación calculado con Sentinel 2A y Sentinel 2B en los años 2016, 2017 y 2018. Solo se muestran las correlaciones superiores a 0,4.....	42

Cuadro 7. Correlación de Spearman entre las variables de rendimiento y el promedio de los índices de vegetación calculados por mes con Sentinel 2A y Sentinel 2B. Sólo se muestran las correlaciones superiores a 0,4..... 43

Cuadro 8. Resumen de los errores relativos en la estimación del total de toneladas de caña de azúcar, utilizando los modelos PROD Agosto y PROD Final. CoopeVictoria R.L. 50

Lista de Acrónimos

AIC: Akaike Information Criterion

BOA: Bottom Of Atmosphere Reflectance

CIA: Centro de Investigaciones Agronómicas

CRTM05: Costa Rica Transversal de Mercator de 2005

CSS: Cascading Style Sheets

ESA: European Spatial Agency

EVI: Enhanced Vegetation Index

GNDVI: Green Normalized Difference Vegetation Index

GNSS: Global Navigation Satellite System

HTML: HyperText Markup Language

INFOCOOP: Instituto Nacional de Fomento Cooperativo

IV: Índice de Vegetación

LOOCV: Leave-one-out cross-validation

MAE: Mean Absolute Error

MODIS: Moderate Resolution Imaging Spectroradiometer

NDVI: Normalized Difference Vegetation Index

NIR: Near Infrared

PROD: Toneladas de caña de azúcar por hectárea

REND: Kilogramos de azúcar por tonelada

RMSE: Root Mean Square Error

RSE: Residual Standard Error

RSS: Residual Sum of Squares

SAVI: Soil Adjusted Vegetation Index

SIG: Sistemas de Información Geográfica

SMS: Sistema de Monitoreo Satelital

SR: Simple Ratio

TOA: Top Of Atmosphere Reflectance

TON AZU: Total de kilogramos de azúcar

TOT TOT: Total de toneladas de caña de azúcar



Autorización para digitalización y comunicación pública de Trabajos Finales de Graduación del Sistema de Estudios de Posgrado en el Repositorio Institucional de la Universidad de Costa Rica.

Yo, Bryan Alemán Montes, con cédula de identidad 1-1366-0420, en mi condición de autor del TFG titulado Sistema de monitoreo espacio-temporal del cultivo de caña de azúcar (Saccharum officinarum), a partir de información satelital, en CoopeVictoria R.L. Grecia, Costa Rica

Autorizo a la Universidad de Costa Rica para digitalizar y hacer divulgación pública de forma gratuita de dicho TFG a través del Repositorio Institucional u otro medio electrónico, para ser puesto a disposición del público según lo que establezca el Sistema de Estudios de Posgrado. SI NO *

*En caso de la negativa favor indicar el tiempo de restricción: _____ año (s).

Este Trabajo Final de Graduación será publicado en formato PDF, o en el formato que en el momento se establezca, de tal forma que el acceso al mismo sea libre, con el fin de permitir la consulta e impresión, pero no su modificación.

Manifiesto que mi Trabajo Final de Graduación fue debidamente subido al sistema digital Kerwá y su contenido corresponde al documento original que sirvió para la obtención de mi título, y que su información no infringe ni violenta ningún derecho a terceros. El TFG además cuenta con el visto bueno de mi Director (a) de Tesis o Tutor (a) y cumplió con lo establecido en la revisión del Formato por parte del Sistema de Estudios de Posgrado.

INFORMACIÓN DEL ESTUDIANTE:

Nombre Completo: Bryan Alemán Montes

Número de Carné: A70174 Número de cédula: 1-1366-0420

Correo Electrónico: bryan.aleman@ucr.ac.cr

Fecha: 05/02/2021 Número de teléfono: 88622117

Nombre del Director (a) de Tesis o Tutor (a): Dr. Carlos Henríquez Henríquez

BRYAN
ALEMAN
MONTES
(FIRMA)

Firmado digitalmente
por BRYAN ALEMAN
MONTES (FIRMA)
Fecha: 2021.02.05
11:25:30 -06'00'

FIRMA ESTUDIANTE

Nota: El presente documento constituye una declaración jurada, cuyos alcances aseguran a la Universidad, que su contenido sea tomado como cierto. Su importancia radica en que permite abreviar procedimientos administrativos, y al mismo tiempo genera una responsabilidad legal para que quien declare contrario a la verdad de lo que manifiesta, puede como consecuencia, enfrentar un proceso penal por delito de perjurio, tipificado en el artículo 318 de nuestro Código Penal. Lo anterior implica que el estudiante se vea forzado a realizar su mayor esfuerzo para que no sólo incluya información veraz en la Licencia de Publicación, sino que también realice diligentemente la gestión de subir el documento correcto en la plataforma digital Kerwá.

Capítulo 1: **Introducción**

1.1. Introducción

En la actualidad, gran parte de las actividades cotidianas que desarrollan los seres humanos han sido influenciadas por los avances tecnológicos que buscan facilitar u optimizar las labores. Estos avances también han transformado los estilos de vida, a tal punto que muchas decisiones hoy en día se toman a partir de inteligencia artificial.

Un claro ejemplo de dichas transformaciones son las herramientas geoespaciales que en la actualidad se han convertido en insumos de importancia para apoyar los procesos de toma de decisiones de índole espacial. Estas herramientas tuvieron un auge tecnológico durante la II Guerra Mundial, la “Carrera Espacial” y la Guerra Fría, debido a la necesidad que tenían las potencias de controlar los territorios en disputas y contar con insumos actualizados para mejorar las estrategias militares.

Actualmente estos insumos continúan siendo importantes, ya que forman parte de las actividades diarias, por ejemplo, todo lo relacionado con Sistema Globales de Posicionamiento Satelital (GNSS por sus siglas en inglés) que en la actualidad son vitales en una gran cantidad de actividades cotidianas (como identificar la mejor ruta al trabajo, todos los sistemas de navegación, la ubicación del restaurante de interés, entre otras); otro insumo de uso cotidiano es la información climática obtenida a partir de satélites meteorológicos, los sistemas de monitoreo ambiental y de variabilidad climática que se utilizan globalmente.

Específicamente en temas agrícolas las herramientas geoespaciales han tomado relevancia en actividades relacionadas con agricultura de precisión porque aportan el componente de variabilidad espacial al manejo de los sistemas agrícolas, dejando de lado los preceptos de la agricultura tradicional donde los campos de cultivos son manejados con unidades homogéneas (Procisur y IICA, 2015).

Además, en este ámbito son múltiples los abordajes de las herramientas geoespaciales que abarcan desde los mapas locales de variabilidad, los GNSS instalados en maquinarias agrícolas hasta los sistemas de monitoreo de cultivos

dotados de información satelital. En los ejemplos anteriores hay cambios de escala y tecnologías, pero el objetivo se mantiene, todos persiguen el principio de tomar decisiones apoyadas con información geoespacial.

Dentro de este marco general nació la propuesta de investigación de desarrollar un sistema de monitoreo del cultivo de caña de azúcar, a partir de información satelital que permita el ajuste de labores operativas en CoopeVictoria R.L. ubicada en Grecia, Costa Rica.

1.2. Justificación

Aumentar la competitividad del sector agrícola costarricense se ha convertido en un reto tanto para los grandes como para los pequeños y medianos productores; en el caso de estos últimos particularmente debido al escaso acceso a tecnología de punta que permite tecnificar el manejo de los sistemas productivos, aumentando la rentabilidad económica, disminuyendo los riesgos de cosechas y minimizando los impactos ambientales. Es así como este proyecto de investigación se plantea como una alternativa ante la carencia de un sistema de monitoreo de cultivos a partir de información geoespacial, que permita a los productores del cultivo de la caña de azúcar, tomar decisiones en función del cuándo y dónde intervenir los cultivos.

La investigación se enmarcó en el contexto de la economía social solidaria que promueve el cooperativismo, de esta manera pequeños productores se ven beneficiados del acceso a información que les permita apoyar los procesos de toma de decisiones durante las cosechas; además esta herramienta permite a este conjunto de productores, aumentar su capacidad de respuesta ante las inminentes aperturas comerciales, la variabilidad climática y los aumentos en el precio de los insumos.

Actualmente, los procesos de estimación de cosechas en diversos cultivos, entre ellos la caña de azúcar, se realiza mediante criterio experto y valoraciones de producciones anteriores, esta herramienta propuesta permitió comprender y estimar variables de producción a escala local.

Las herramientas geoespaciales se convierten en un insumo de vital importancia para la gestión y el manejo de los espacios agrícolas, en este particular se utilizaron Sistemas de Información Geográfica y Sensores Remotos para generar un insumo que aporte información para la toma de decisiones en el sector agrícola, cumpliendo así con los objetivos de la Maestría, el cual es dar solución a problemas de índole espacial utilizando Sistemas de Información Geográfica y Teledetección.

En síntesis, este proyecto de investigación dota de una herramienta geoespacial que apoya el proceso de toma de decisiones agrícolas en CoopeVictoria R.L. Asimismo, este proyecto sirvió como un plan piloto que posteriormente podrá ser desarrollado en otras áreas de producción agrícola a nivel nacional.

1.3. Problema de Investigación

El aumento de la competitividad de los mercados agrícolas hace necesario que las empresas dedicadas a estas actividades optimicen los flujos de trabajo, con el objetivo de aumentar los réditos económicos. En el contexto costarricense los pequeños y medianos productores tiene una limitación de acceso a tecnología geoespacial que les permita tomar decisiones, contemplando la variabilidad espacial y temporal de los cultivos; esta limitación puede estar vinculada con términos económicos o bien, debido a la brecha tecnológica.

Dicha limitación está presente en la cooperativa CoopeVictoria R.L. la cual carece de un sistema de monitoreo que le permita evaluar espacial y temporalmente las plantaciones de caña de azúcar, de fincas propias y de asociados; esta limitante ha restringido los procesos de toma de decisiones, principalmente cuando se trata de intervenir las plantaciones en los momentos y lugares oportunos para mejorar la productividad. La carencia de este sistema de monitoreo implica una serie de efectos que limitan la competitividad, las labores operativas de la Cooperativa.

1.4. Objetivos

1.4.1. Objetivo general

- Generar un sistema de monitoreo del cultivo de caña de azúcar, a partir de información satelital que permita el ajuste de labores operativas en CoopeVictoria R.L. Grecia, Costa Rica.

1.4.2. Objetivos específicos

- Identificar la plataforma satelital de acceso libre que presenta la mayor viabilidad para el desarrollo del sistema de monitoreo espacio-temporal en el cultivo de la caña de azúcar en CoopeVictoria R.L.
- Determinar los índices de vegetación más apropiados para el monitoreo del cultivo de la caña de azúcar en los diferentes estados fenológicos.
- Construir una herramienta para la estimación de rendimientos comerciales de caña de azúcar, a partir de información satelital, que facilite labores operativas.
- Construir un geoportal que permita la visualización de la variabilidad espacial y temporal del cultivo de la caña de azúcar en los distintos estados fenológicos.

1.5. Delimitación espacial y temporal

El área de estudio forma parte de las propiedades de CoopeVictoria R.L., localizadas principalmente en el distrito San Isidro del cantón Grecia en la provincia Alajuela; esta cooperativa se dedica al cultivo y procesamiento de los cultivos de caña de azúcar y café, además de otros servicios. Se trabajó con 26 fincas dedicadas al cultivo de la caña de azúcar, que en total suman aproximadamente 511 hectáreas sembradas de distintas variedades caña de azúcar. El área de estudio está sembrada en un 67,3% por la variedad RB 86-7515, asimismo existe alrededor el 13% del área cuya variedad principal no está definida (Figura 1B); además las fincas de mayor extensión son Grupo Herrero La Argentina y Kooper Central con un 18,9% y un 11,8% del área total analizada (Figura 1C).

Las fincas incluidas en la investigación corresponden al total del área de producción de CoopeVictoria R.L. localizadas en Grecia y Póas (incluye fincas propias y arrendadas), estas fincas están situadas entre las coordenadas proyectadas CRTM05 459675,71 m – 474637,09 m Este y 1105406,72 m – 1121103,77 m Norte, en el marco de coordenadas geográficas el área se localiza entre las coordenadas 9,965° - 10,107° latitud Norte y 84,227° - 84,362° longitud Oeste (Figura 1A).

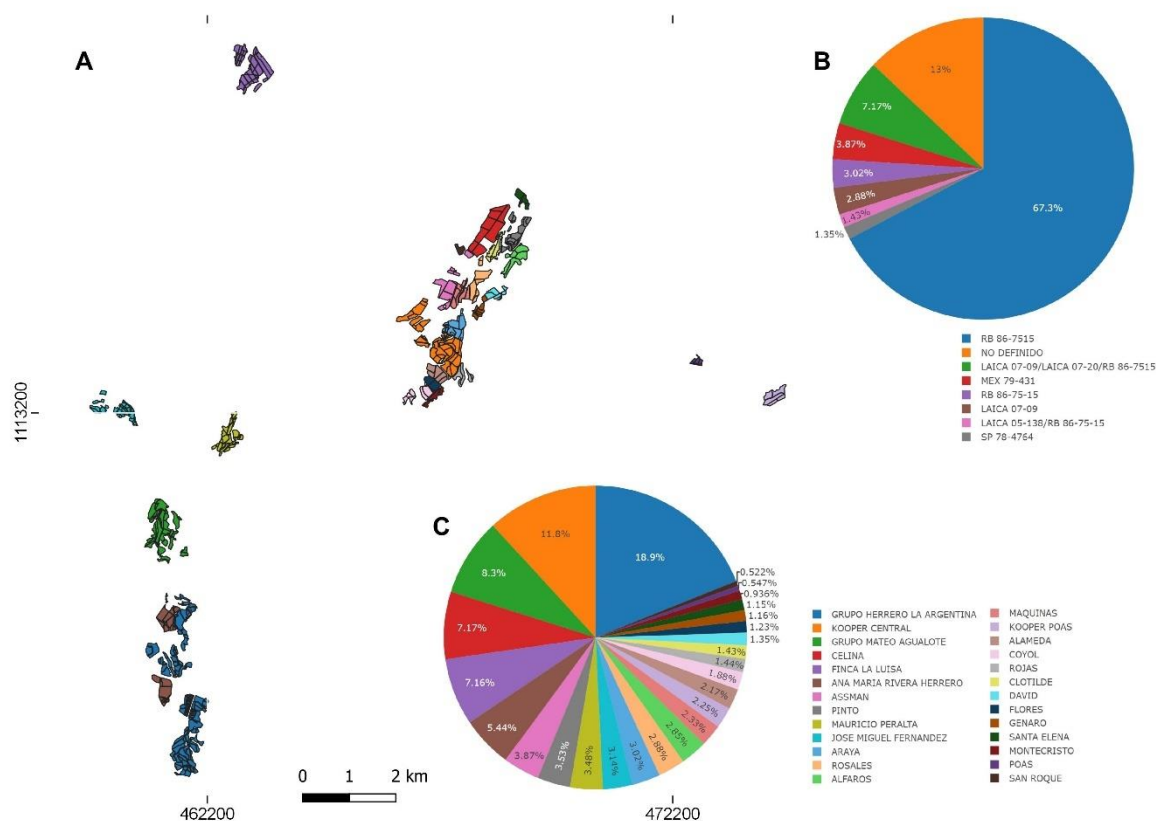


Figura 1. CoopeVictoria R.L., 2018: (A) Distribución de las plantaciones de caña de azúcar por Finca. (B) Porcentaje de área según variedad. (C) Porcentaje del área según la dimensión de la Finca.

Fuente: Base de datos de Fincas de CoopeVictoria R.L.

1.6. Alcances del proyecto

Para alcanzar los objetivos planteados en este proyecto fue necesaria la integración de metodologías y habilidades que permitieran la correcta ejecución de los procesos que desencadenaran en la construcción del sistema de monitoreo espacio-temporal

de las plantaciones de caña de azúcar propiedad de CoopeVictoria R.L., localizadas en el sector de Grecia; para lo cual fue necesario utilizar imágenes satelitales de acceso libre (Sentinel, Landsat y MODIS), software y repositorios libres (RStudio, QGIS, GitHub) y datos de campo suministrados por la Cooperativa.

Los resultados de este proyecto de investigación han sido obtenidos a partir de las variables de producción de caña de azúcar de las fincas de CoopeVictoria R.L. localizadas en el sector de Grecia, por lo cual futuras aplicaciones de los modelos de estimación deben ser aplicados sobre esa misma zona. Razón por la cual, para aplicar los modelos de estimación aquí descritos en áreas que no aportaron información para la construcción de los modelos primero se debe validar la aplicabilidad en fincas no incluidas en el proceso de modelización; por ejemplo, las fincas de productores que no son propiedad de la cooperativa y que se localizan en los alrededores de Grecia.

Para ejecutar los modelos de estimación en fincas que no aportaron en la modelización se debe de contar con la siguiente información básica: delimitación de las fincas e historiales de producción, por esta razón de existir la intención de incluir nuevas áreas en la estimación y el monitoreo se recomienda recabar la información base (áreas de producción, historiales de cosecha).

1.7. Limitaciones

Existen múltiples limitaciones, muchas de las cuales pueden convertirse en oportunidades de mejora. A continuación, se describen las limitaciones encontradas durante el desarrollo del proceso investigativo.

Las imágenes satelitales ópticas están fuertemente influenciadas por las condiciones atmosféricas, esto ocasionó que un alto porcentaje de las imágenes disponibles fueran descartadas por nubes, sombra de nubes u otras distorsiones atmosféricas, estas limitaciones pueden ser solventadas en futuras investigaciones con la incorporación de imágenes radar.

La mayor cantidad de limitaciones identificadas están relacionadas con la información base (unidad productiva). El área de las unidades productivas es muy

heterogénea, hay una finca de 2,60 ha y otra 96,6 ha, estas diferencias de área influyen negativamente en la construcción de los modelos, porque en áreas grandes se “enmascara” la variabilidad espacial de los índices de vegetación que son resumidos en un promedio, por esta razón se recomienda utilizar unidades manejo más homogéneas en relación con el área.

Además, sería importante contar con otras características cuantitativas y cualitativas de las fincas, a saber: número de socas por unidad productiva, fecha de corta de la unidad productiva y material genético sembrado, ya que estas variables se pueden incluir como factores de bloqueo en los modelos, logrando así la construcción de modelos de estimación por variedad u otra variable categórica.

Contar con la fecha de corta de la plantación permitiría construir series de tiempo que muestren la variabilidad temporal de los índices de vegetación por unidad productiva, identificando claramente todas las etapas fenológicas del cultivo, esto aumentaría sustancialmente la utilidad de las imágenes satelitales en el monitoreo de las plantaciones de caña azúcar.

Los puntos mencionados tienen gran potencial de ser mejorados o al menos evaluados, esto traería consigo un beneficio en las estimaciones y los procesos de monitoreo del cultivo de caña de azúcar en CoopeVictoria R.L.

Capítulo 2: Marco Conceptual

2.1. Características el cultivo caña de azúcar

La producción de caña de azúcar es una actividad de importancia socioeconómica para el país, en 2019 se contabilizaban 64196 ha cultivadas con caña de azúcar (INEC, 2020), de las cuales aproximadamente el 7% están sembradas de diversas variedades en el Valle Central Occidental (Chavez y Chavarría, 2012).

La principal variedad sembrada en la región es la RB 86-7515 que posee un crecimiento rápido, semi-erecto, despaje de regular a bueno y facilidades para establecerse en suelos de fertilidad baja y con déficit hídrico (Durán y Oviedo, 2012), esta variedad es anual lo cual permite compararla en términos generales con la evolución fenológica del cultivo presentada en la Figura 2.

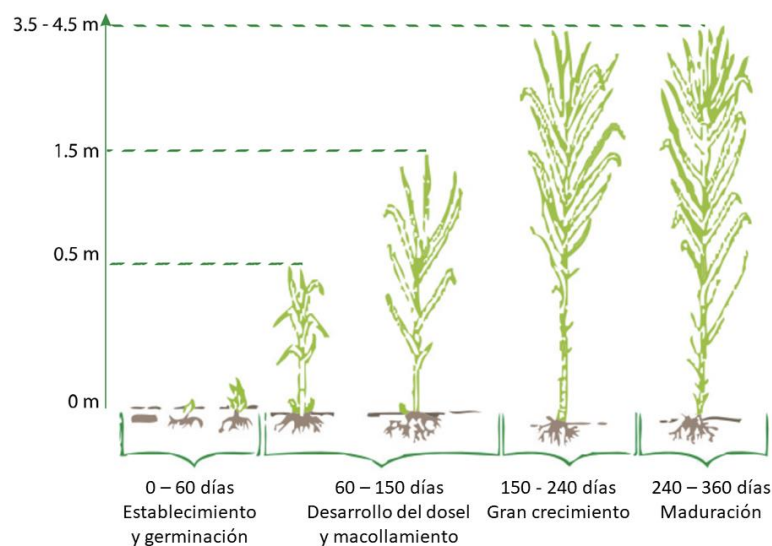


Figura 2. Evolución fenológica del cultivo de la caña de azúcar.

Fuente: Modificado de Molijn et al., 2019.

La primera etapa, establecimiento y germinación, se caracteriza por la aparición de rebrotes, lo cual permite la cobertura del suelo con vegetación (Allison et al., 2007). Posteriormente en la etapa de desarrollo del dosel y macollamiento la cobertura del suelo por vegetación aumenta, al finalizar dicha etapa inicia el periodo de gran crecimiento en la cual cultivo intercepta el 70 % de la radiación solar fotosintéticamente activa (Inman-Bamber, 1994). Finalmente, sucede un proceso

de maduración caracterizada por incremento de la concentración de sacarosa en tallos (Saez, 2017) (Figura 2).

2.2. Principios de la teledetección

El desarrollo de la aviación, la creación de la cámara fotográfica, la carrera espacial y los avances informáticos han permitido el desarrollo de los sensores remotos como herramientas para el análisis de fenómenos que tiene comportamientos espacialmente discernibles. Los sensores remotos se han convertido en insumos base para múltiples disciplinas, ya que permiten la recolección de información de elementos sobre la superficie terrestre, pero sin entrar en contacto directo (Chuvieco, 2008; Khorram et al., 2012; Paul y Mascarenhas, 1981).

Las imágenes de sensores remotos pueden ser obtenidas mediante plataformas satelitales o bien por sensores a bordo de aviones tripulados o vehículos aéreos no tripulados; estas imágenes brindan información del comportamiento espectral de los diferentes elementos sobre la superficie terrestre en distintos puntos del espectro electromagnético (Ferguson y Rundquist, 2018). El rango que abarca el espectro visible, el infrarrojo cercano y el infrarrojo medio (Figura 3) han sido ampliamente utilizados para los estudios de vegetación (Puletti et al., 2018; Zhang et al., 2017), porque permite analizar los ámbitos de mayor dinamismo de las plantas (Campbell, 2006; Chuvieco, 2008; Martínez-Martínez et al., 2018).

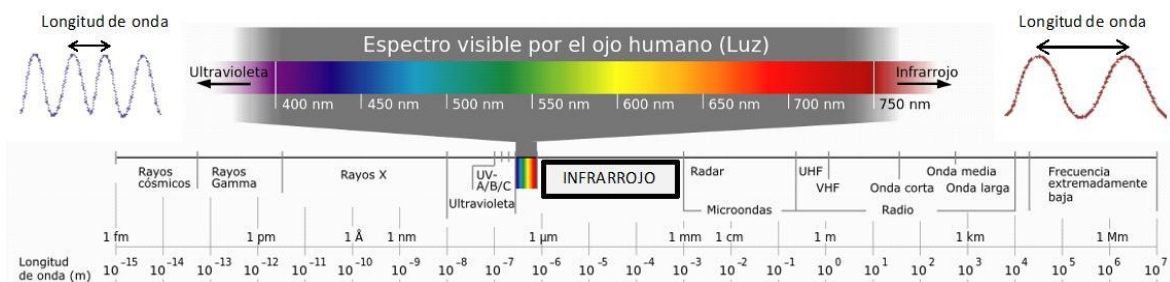


Figura 3. Espectro electromagnético

Fuente: <http://culturacientifica.com/2015/12/21/vision-infrarroja/>

Los rangos espectrales más utilizados en agricultura incluyen el espectro visible comprendido entre los rangos de 400 a 700 nm (con longitudes de onda en el azul que van de 400 a 500 nm, el verde de 500 a 600 nm y el rojo de 600 a 700 nm), el

infrarrojo cercano es de 700 a 1300 nm, el infrarrojo medio se encuentra entre los 1300 y 2500 nm y el infrarrojo térmico entre 8000 y 14000 nm (Figura 3) (Chuvienco, 2008; Ferguson y Rundquist, 2018; Lillesand y Kiefer, 1987; Moreno et al., 1999).

2.3. Comportamientos espectrales de las coberturas vegetales

En las plantas con alta densidad de hojas verdes la reflectividad en el espectro visible es controlada principalmente por la clorofila que tiene una mayor reflectividad en la región espectral del verde (500 a 600 nm), mientras tanto las regiones espectrales del azul (400 a 500 nm) y el rojo (600 a 700 nm) presentan los puntos de mayor absorción de energía incidente (Campbell, 2006; Chuvienco, 2008; Petropoulos y Kalaitzidis, 2012). En la región del infrarrojo cercano la reflectividad está controlada por la estructura interna de la hoja, donde la radiación incidente es fuertemente reflejada por los tejidos mesofilos y las cavidades internas, en esta región del infrarrojo hay un bajo porcentaje de absorción y un alto porcentaje de energía reflejada (Campbell, 2006; Chuvienco, 2008; Petropoulos y Kalaitzidis, 2012). Entre ambas regiones espectrales (visible e infrarrojo cercano) se localiza el borde de rojo que manifiesta un aumento acelerado del porcentaje de reflectancia, lo cual ha sido relacionado con vigorosidad de la vegetación (Campbell, 2006; Petropoulos y Kalaitzidis, 2012) (Figura 4).

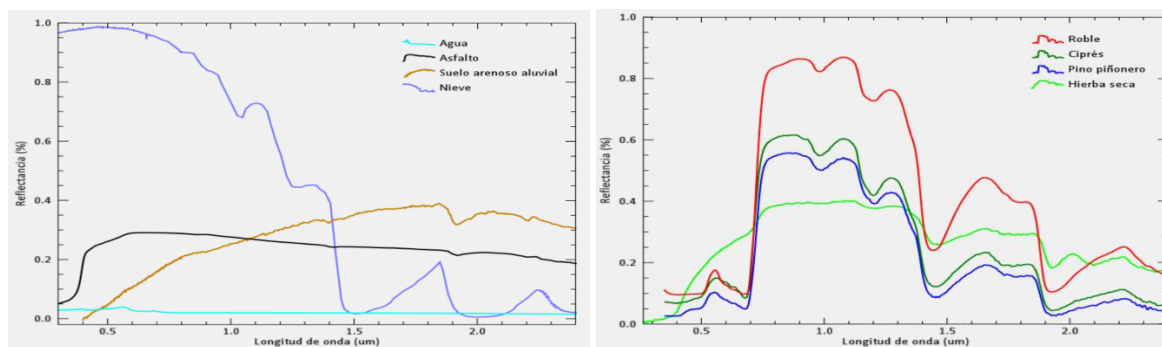


Figura 4. Firmas espectrales de algunas coberturas sobre la superficie terrestre.

Fuente: Biblioteca espectral del USGS en Solis y Martínez.R, 2016.

Con análisis espectrales, en longitudes de onda superiores al espectro visible, se ha logrado evaluar de forma indirecta características de la vegetación como: contenido de clorofila, humedad de las planta, variaciones en diferentes condiciones

de estrés de la vegetación en general, entre otras (Hogg, 2004; Thenkabail et al., 2011); estos comportamientos espectrales de la vegetación pueden manifestar propiedades biofísicas o bioquímicas de las plantas, permitiendo así la identificación temprana de alteraciones (Chuvienco, 2008). Sensores multiespectrales o hiperespectrales incorporados a plataformas satelitales, vehículos aéreos no tripulados, plataformas terrestres o radiómetros de campo ha sido de gran utilidad para la identificación temprana de manifestaciones de estrés en rangos espectrales imperceptibles para el ojo humano, posibilitando así la detección temprana de enfermedades, deficiencias, entre otras (Ferguson y Rundquist, 2018).

2.4. Índices de vegetación como indicadores del estado de las plantas

En primera instancia, es necesario dejar en claro que los sensores satelitales no tienen la capacidad de medir una propiedad física de la vegetación de forma directa, sino más bien, pueden derivar dicha propiedad a partir de las modificaciones de la señal percibida; por ejemplo, los sensores satelitales pueden medir indirectamente el contenido de clorofila de la vegetación a partir de las modificaciones en la reflectividad de las plantas (Chuvienco, 2008).

Estas propiedades físicas de la vegetación podrían ser medidas indirectamente a partir de índices de vegetación comprendidos como la combinación de distintos rangos espectrales a partir de sumas, restas, multiplicaciones o divisiones que tienen como objetivo cuantificar indirectamente el vigor de la vegetación, a partir de la fracción fotosintéticamente activa de la vegetación, donde altos valores del índice estarían relacionados con vegetación saludable (Campbell, 2006).

Múltiples análisis espectrales de la vegetación se han realizado mediante índices de vegetación entendidos como mediciones radiométricas adimensionales que generalmente funcionan como indicadores de la abundancia y actividad relativa de la vegetación (Petropoulos y Kalaitzidis, 2012). A lo largo de la historia se han desarrollado múltiples índices de vegetación combinando distintas regiones espectrales, sin embargo, el objetivo no ha cambiado y fundamentalmente se basan en acentuar signos vitales de la vegetación (humedad, contenido de clorofila,

índices de área folia, entre otros) comparando regiones espectrales contrastantes (Chuvienco, 2008; Santoso et al., 2016).

Los contrastes espectrales existentes entre el espectro visible, especialmente en la banda roja, y las bandas infrarrojo cercano, están relacionadas con las diferencias en absorción por parte de los pigmentos de la vegetación; esta característica permite establecer la diferencia entre la vegetación sana y la vegetación “no sana”. Cuando la vegetación presenta algún tipo de estrés tiende a bajar la reflectividad en la región del infrarrojo cercano y aumentar la reflectividad en el espectro visible (Figura 5) (Chuvienco, 2008).

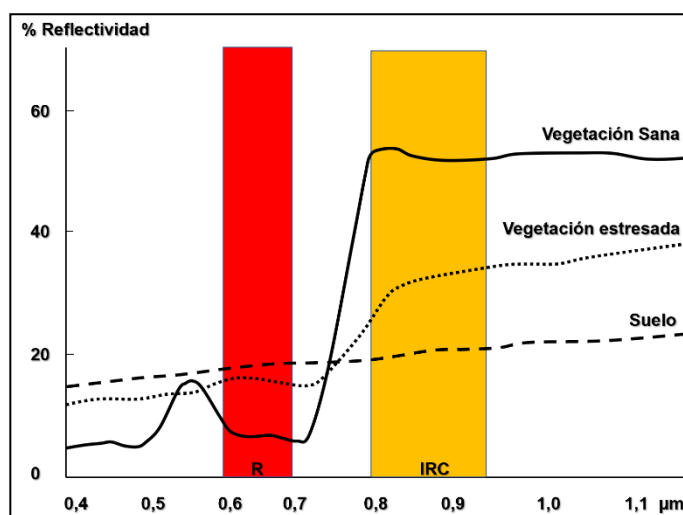


Figura 5. Contraste espectral de la vegetación sana, vegetación estresada y suelo en las bandas Rojo e Infrarrojo Cercano.

Fuente: Modificado de Chuvienco, 2008.

Algunos ejemplos en el campo agrícola son: Gröll et al., (2007) que utilizando un conjunto de índices de vegetación obtenidos a partir de información espectral lograron evaluar la utilidad de este insumo en la discriminación de plantas de manzana sanas de las afectadas por un inóculo de mancha foliar; Mahlein et al., (2013), utilizando información hiperespectral, lograron establecer diferencias entre plantas de remolacha azucarera sanas y otras afectadas por *Cercospora*. Murillo y Carbonell (2012), Rahman y Robson (2016) y Rahman y Robson (2020) lograron estimar la productividad de plantaciones de caña de azúcar utilizando índices de vegetación, extraídos de imágenes multiespectrales.

2.5. Transferencia de información geoespacial rápida y oportuna

Es claro que existe la necesidad de transferir la información geoespacial de forma rápida y oportuna, para los tomadores de decisiones. En este sentido las plataformas *GIS-Online* o mapeo web (*Web Mapping*), han facilitado esa transmisión de información de forma expedita; un ejemplo de ello en el sector agrícola, es el trabajo realizado por Papadopoulos et al., (2016) quienes desarrollaron un protocolo de trabajo para ejecutar un mapeo de atributos de suelos y su posterior transferencia a los tomadores de decisiones mediante plataformas de internet en el país Grecia.

El mapeo web posee como características fundamentales el despliegue interactivo de información geográfica en una página web, estos principios vienen a remplazar el mapeo digital en entornos cerrados, confinados y complejos de compartir (Dorman, 2020), transformando el manejo de información geográfica a través de la web en sistemas integrados y completos que pueden ser percibidos como infraestructuras SIG completas (Oyala, 2012; Veenendaal et al., 2017).

El mapeo web se fundamenta en la interactividad del usuario con los insumos cartográficos. Esto implica que el usuario pueda realizar funciones básicas activar y desactivar capas, acercamientos, revisar las propiedades de las capas y editar las características (Dorman, 2020; Sandinska, 2016). Sin embargo, dichas funciones variarán según los permisos de manipulación que posea el usuario.

El desarrollo de una arquitectura de mapeo web requiere de la integración de diferentes componentes que articulados potencien la interactividad y funcionalidad de la herramienta. El primer componente es el **Servidor** que es el repositorio que almacena la información de interés e interactúa seguidamente con el siguiente componente que son los **Clientes** (Usuarios) a través de la **Web** (Internet) (Dorman, 2020; Oyala, 2012). En la Figura 6A representa la interacción existente entre los componentes.

Cada uno de los componentes posee características específicas. El servidor puede estar alojado físicamente en un lugar con conexión a internet desde donde se accionan y ejecutan las consultas del cliente (Figura 6B) (Oyala, 2012), otra

alternativa puede ser almacenar la información en repositorios web| como GitHub, en la cual se puede almacenar el código de programación para que sea accedido y consultado por los clientes (Mudholkar y Mudholkar, 2017).

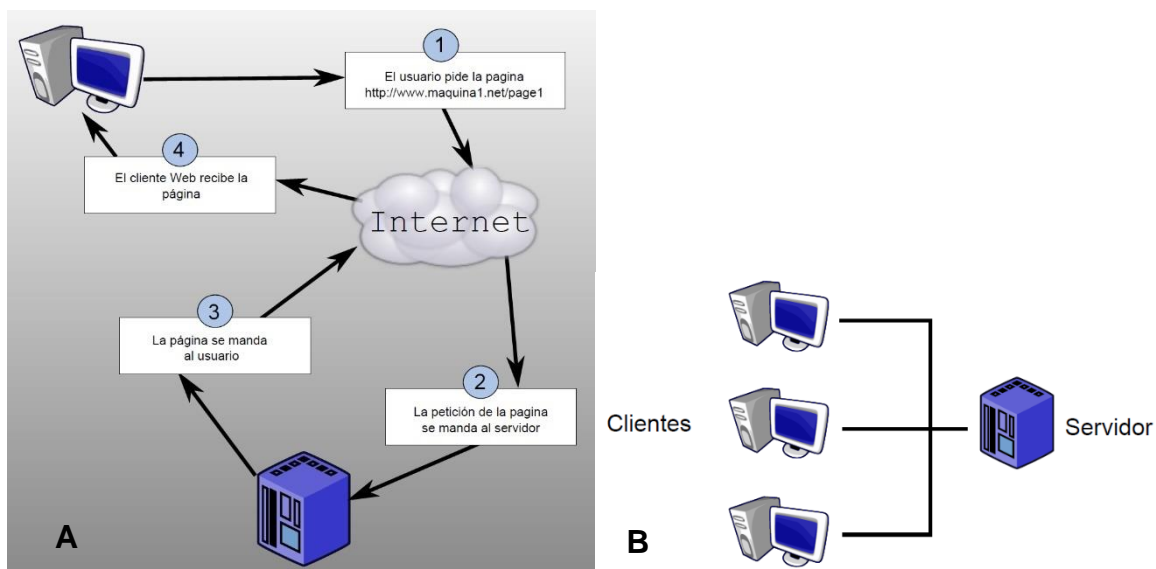


Figura 6. (A) Interacción Cliente-Servidor durante el proceso de consulta de mapas web. (B) Interacción de múltiples clientes con el servidor.

Fuente: Oyala, 2012.

Por otra parte, los códigos de programación alojados en el Servidor pueden ser programados en diversos lenguajes como Python y JavaScript, que a su vez son puestos a disposición de los clientes a través de protocolos como HTML (Hypertext Markup Language) para la creación de páginas web, CSS (Cascading Style Sheets) para la edición de estilos en las páginas web (W3.CSS, 2019a, 2019b), estos dos últimos en conjunto con JavaScript son las tres tecnologías principales para el desarrollo de mapas web (Sandinska, 2016).

Capítulo 3: Marco Metodológico

Para cumplir con los objetivos de esta investigación la información fue manejada integralmente, ya que la información generada en cada apartado estuvo vinculada directamente con los siguientes. A continuación, se explica el procedimiento metodológico abordado para el desarrollo y el cumplimiento de los objetivos planteados.

3.1 Elección del sensor satelital

El punto de partida fue la determinación la plataforma satelital de acceso libre más adecuada para monitorear las plantaciones de caña de azúcar en CoopeVictoria R.L.; para esto se utilizaron los sensores MODIS, Landsat 8 y Sentinel 2A o Sentinel 2B. La determinación del sensor óptimo se realizó a partir de la evaluación cualitativa y cuantitativa de resolución: espacial, espectral y temporal (Cuadro 1).

Los criterios cualitativos y cuantitativos de resolución fueron analizados tomando en consideración las características del área de estudio, principalmente el área de las unidades de análisis (Fincas), ya que una mayor cantidad de píxeles dentro de las Fincas permitió tener más representatividad. Por tal razón, se determinó, con base en cada sensor, el número de píxeles que había dentro cada finca, la extracción del número de píxeles se realizó con una función *extract()* del paquete ráster integrado en RStudio.

Una vez calculado el número de píxeles por finca y para cada sensor, se calculó el área de las fincas a partir de la multiplicación del número de píxeles por el área del píxel; el área del píxel varía según la resolución espacial, para el caso del sensor MODIS cada píxel tiene un área de 293946 m², mientras tanto cada píxel de Landsat 8 cubre 900 m² y un píxel de Sentinel 2 tiene un área de 100 m². A partir de dicha información se analizó cual sensor ajustaba en mejor medida los datos reales. Adicionalmente, se consideró como criterio cualitativo que los píxeles dentro de cada finca permitieran interpretar variabilidad espacial.

Cuadro 1. Características básicas de los sensores satelitales: MODIS, Landsat OLI 8 y Sentinel 2A. Resaltadas están las características de las bandas utilizadas para calcular los índices de vegetación.

Sensor	Fuente	Resolución				
		Espacial	Espectral (nm)		Radiométrica	Temporal
MODIS ¹	NASA	250 metros (bandas 1-2) 500 metros (bandas 3-7) y 1000 metros (bandas 8-36)	B1: 620-670 B2: 841-876 B3: 459-479 B4: 545-565 B5: 1230-1250 B6: 1628-1652 B7: 2105-2155 B8: 405-420 B9: 438-448 B10: 483-493 B11: 526-536 B12: 546-556 B13: 662-672 B14: 673-683 B15: 743-753 B16: 862-877 B17: 890-920 B18: 931-941	B19: 915-965 B20: 3660-3840 B21: 3929-3989 B22: 3929-3989 B23: 4020-4080 B24: 4433-4498 B25: 4482-4549 B26: 1360-1390 B27: 6535-6895 B28: 7175-7475 B29: 8400-8700 B30: 9580-9880 B31: 10780-11280 B32: 11770-12270 B33: 13185-13485 B34: 13485-12785 B35: 13785-14085 B36: 14085-14385	12 bits	8 días
Landsat OLI 8 ²	NASA	15 metros (banda 8) 30 metros (bandas 1-7, 9) y 100 metros (bandas 10-11)	B1: 435-431 B2: 452-512 B3: 533-590 B4: 636-673 B5: 851-879 B6: 1566-1651	B7: 2107-2294 B8: 503-676 B9: 1363-1384 B10: 10600-11190 B11: 11500-12510	12 bits	16 días
Sentinel 2A y 2B ³	Agencia Espacial Europea (ESA)	10 metros (bandas 2-4, 8) 20 metros (bandas 5-7, 8A, 11-12) 60 metros (bandas 1, 9-10)	B1: 443 B2: 490 B3: 560 B4: 665 B5: 705 B6: 740	B7: 783 B8: 842 B9: 945 B10: 1375 B11: 1610 B12: 2190	12 bits	5 días

Nota: ¹ Información obtenida en <https://modis.gsfc.nasa.gov/about/specifications.php>. Los índices de vegetación del sensor MODIS fueron obtenidos directamente mediante el paquete de R MODISstsp v 1.3.9. ² Información obtenida en <https://landsat.gsfc.nasa.gov/landsat-data-continuity-mission/>.

³ Información obtenida en <https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi/resolutions/radiometric>.

La información satelital fue descargada con diferentes niveles de procesamiento (MODIS nivel 3, Landsat OLI 8 nivel 2 y Sentinel 2 nivel 1) y utilizando plataformas y métodos distintos, a continuación, se describe el proceso de descarga de cada sensor utilizado para esta investigación.

- MODIS

Estas imágenes fueron descargadas en nivel 3 de procesamiento, lo cual implicó el uso directo de la información, sin necesidad de realizar correcciones, calibraciones técnicas ni el cálculo de los índices de vegetación (este nivel de procesamiento permitió obtener los índices de vegetación calculados). La descarga de los productos MODIS se realizó mediante el paquete de R MODISstp desarrollado por Busetto y Ranghetti (2016) (Anexo 1); el cual posee una interfaz intuitiva para descargar diferentes productos MODIS y con distintos niveles de procesamiento, como ya se indicó, para esta investigación se descargaron imágenes con los índices de vegetación calculados del área de estudio en los meses de abril a diciembre de cada año en estudio (Figura 7).

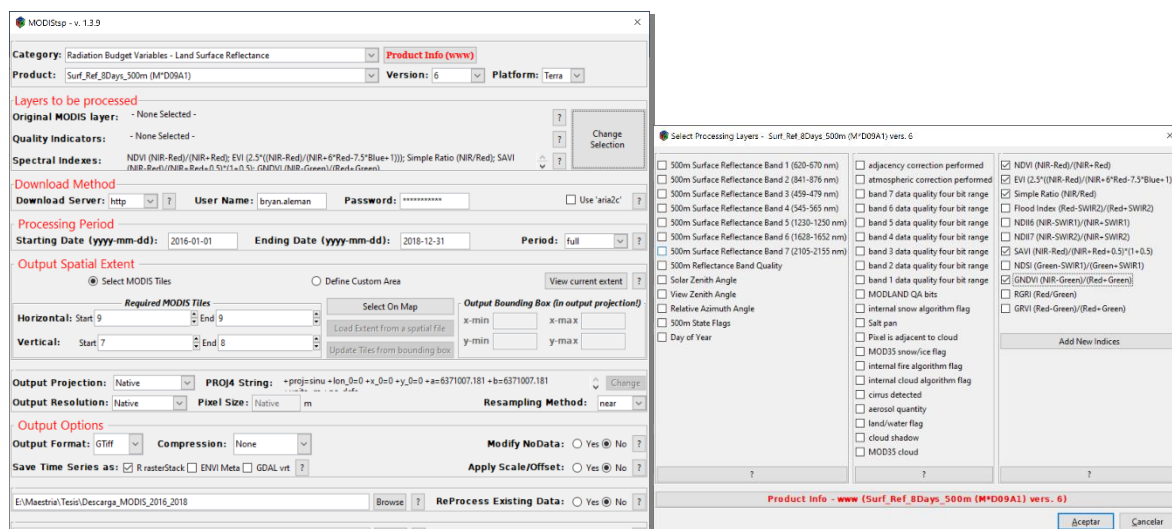


Figura 7. Interfaz de descarga de MODISstp v1.3.9. La descarga se realizó definiendo la categoría, el tipo de producto, los índices espectrales (recuadro de la derecha), usuario y contraseña, periodo de descarga, localización de área de interés (Tiles), proyección, formato y ruta de salida.

- **LandSat OLI 8**

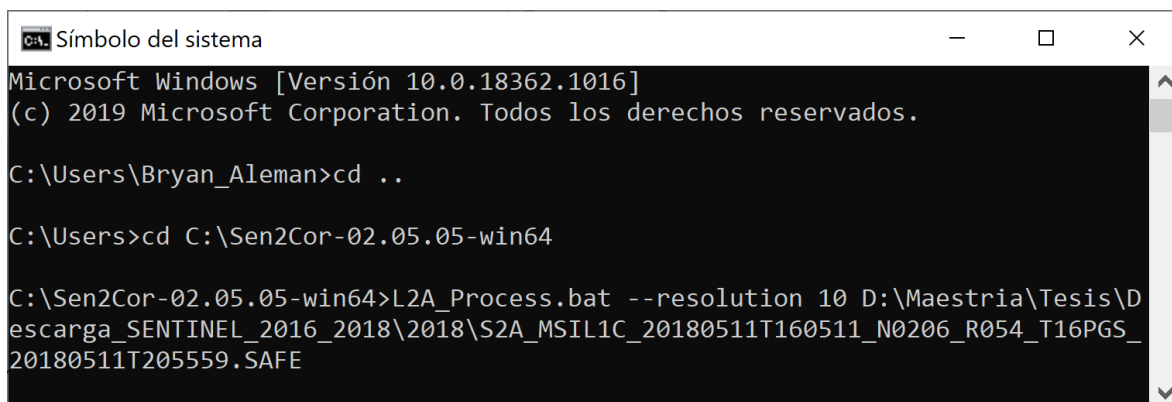
El nivel de procesamiento de los productos Landsat OLI 8 permitió descargar las imágenes corregidas a reflectancia de la superficie (nivel 2 de procesamiento). La descarga se realizó utilizando el paquete de R rLandsat desarrollado por Sikaria, 2018 y sistematizado en un código de RStudio por Nino Bravo (Anexo 2). En el cual se debe definir la temporalidad de interés (de abril a diciembre), el Path y Row, y el porcentaje de nubosidad permitido.

Una vez descargadas y seleccionadas las imágenes funcionales para el área de estudio, se calcularon los índices de vegetación, del Cuadro 2, utilizando un código de programación sistematizado en RStudio, el cual permite obtener un ráster del producto del álgebra de mapas (Anexo 3).

- **Sentinel 2**

Las imágenes Sentinel 2A y 2B fueron seleccionadas y descargadas directamente desde la plataforma EarthExplorer (<https://earthexplorer.usgs.gov/>) donde se definió: el área de interés, la temporalidad y el tipo de sensor (Sentinel 2A y 2B según la temporalidad). Para la descarga se evaluó visualmente la nubosidad en el área de interés y solo se descargaron las imágenes que estaban libres de nubes en el área de estudio, además únicamente fueron consideradas imágenes entre abril y diciembre periodo en el cual no había actividades de zafra.

Posteriormente las imágenes fueron corregidas atmosféricamente mediante el método Sen2cor, a través de la línea de comando de Windows (Figura 8). El procesador de corrección atmosférica Sen2Cor fue desarrollado por Telespazio VEGA Deutschland GmbH a nombre de la Agencia Espacial Europea y tiene como principal objetivo corregir los productos Sentinel 2 nivel 1C (TOA -*Top Of Atmosphere Reflectance*-) a productos Sentinel 2 nivel 2A (BOA -*Bottom Of Atmosphere Reflectance*-), dicha calibración atmosférica corrige los efectos de la atmosfera para cada fecha en específico (Main-Knorn et al., 2017).



```
Símbolo del sistema
Microsoft Windows [Versión 10.0.18362.1016]
(c) 2019 Microsoft Corporation. Todos los derechos reservados.

C:\Users\Bryan_Aleman>cd ..

C:\Users>cd C:\Sen2Cor-02.05.05-win64

C:\Sen2Cor-02.05.05-win64>L2A_Process.bat --resolution 10 D:\Maestria\Tesis\D
escarga_SENTINEL_2016_2018\2018\S2A_MSIL1C_20180511T160511_N0206_R054_T16PGS_
20180511T205559.SAFE
```

Figura 8. Ejecución de corrección atmosférica Sen2cor a través de la línea de comandos de Windows.

La corrección atmosférica permite una mayor calidad de información de las imágenes satelitales, porque basa su procesamiento en modelos de transferencia radiativa de la atmósfera terrestre (Richter et al., 2012). Los modelos de transferencia radiativa se usan para calcular el campo de radiación en múltiples condiciones atmosféricas y de superficie tomando en consideración la radiación del sol y de la tierra, la interacción con nubes, las partículas de aerosol y los gases traza (Mayer y Kylling, 2005).

Particularmente, el método Sen2cor utiliza los modelos de transferencia radiativa definidos por Mayer y Kylling (2005). En la Figura 9 se describe el flujo metodológico propuesto por Richter et al. (2012) para el método de corrección atmosférica Sen2cor, en el cual se pueden observar procesos de clasificación de cubiertas, proceso de remoción de elementos atmosféricos corrección con los modelos de transferencia radiativa para la obtención de reflectancias en el piso de la atmósfera a través de distintos procesos de corrección.

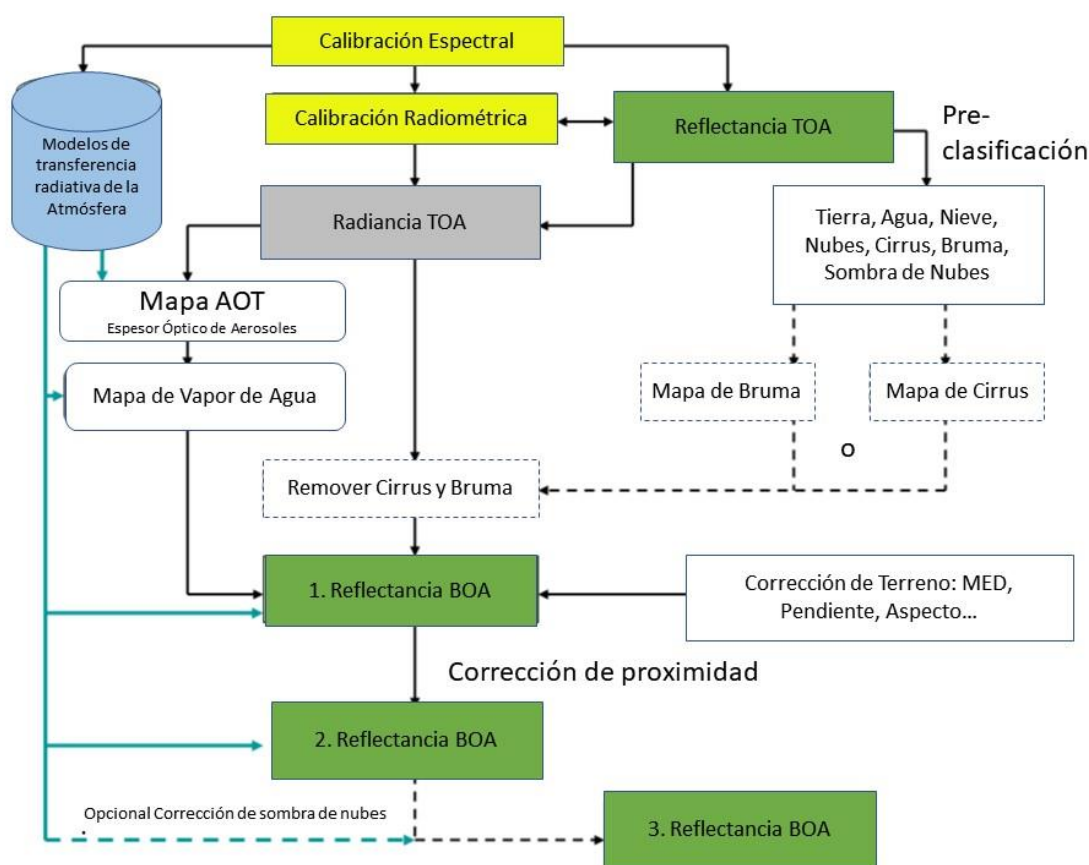


Figura 9. Modelo de corrección atmosférica ejecutado a partir del método Sen2cor para las imágenes satelitales Sentinel 2.

Fuente: Modificado de Richter et al. (2012)

Una vez corregidas atmosféricamente las imágenes Sentinel 2 los índices de vegetación fueron calculados a partir del código de programación del Anexo 4.

3.2 Relación entre índices de vegetación y variables de productividad de caña de azúcar

Una vez determinado el sensor más conveniente para realizar la investigación, se evaluaron los índices de vegetación que mejor representaban la variabilidad temporal y de producción de las fincas. Como variables de respuesta se tomaron los registros de producción que brindó la Cooperativa para este estudio, de las zafras 2016-2017, 2017-2018, 2018-2019 y 2019-2020

Los registros de producción disponibles fueron:

- **Toneladas de caña de azúcar por hectárea (PROD):** es la división del total de toneladas métricas recolectadas en una finca dividido entre el número de hectáreas de esta (Rendimiento de campo).
- **Kilogramos de azúcar por tonelada (REND):** es el valor promedio de los kilogramos de azúcar que se obtuvieron en cada tonelada cosechada por finca (Rendimiento industrial).
- **Total de toneladas de caña de azúcar (TON TOT):** es la sumatoria de las toneladas de caña de azúcar cosechadas por finca.
- **Total de kilogramos de azúcar (TOT AZU):** es la suma del total kilogramos de azúcar obtenidos en las distintas fincas.

A partir de dicha información se realizaron análisis de correlación entre los índices de vegetación de cada zafra y los respectivos datos de producción. Asimismo, se realizó el análisis integrando el total de los datos de los tres años de cosecha, con el objetivo de capturar la tendencia general y no las tendencias en cada año. Estos procesos permitieron determinar los índices que mejor se relacionan con la producción, que a la vez tiene una relación directa con el vigor del cultivo.

Los índices de vegetación utilizados fueron EVI (*Enhanced Vegetation Index*), GNDVI (*Green Normalized Difference Vegetation Index*), NDVI (*Normalized Difference Vegetation Index*), SAVI (*Soil Adjusted Vegetation Index*) y SR (*Simple Ratio*), estos índices fueron seleccionados porque han sido los más utilizados en investigaciones similares. En el Cuadro 2 se definen las principales características de los índices de vegetación utilizados para el desarrollo de esta investigación, los cuales fueron seleccionados a partir de revisión de literatura.

Como se indicó en el apartado anterior el cálculo de estos índices para los sensores Sentinel 2 y Landsat 8 se realizó mediante la ejecución de un código de programación en RStudio, el cual permitió sistematizar las fórmulas, reduciendo los posibles errores asociados a la tabulación de los datos.

De igual manera, mediante un código de programación en RStudio se extrajo el valor promedio de los índices de vegetación para cada finca y para cada una de las

fechas seleccionadas. En este punto es importante resaltar que la extracción del valor promedio de cada índice se realizó excluyendo 10 metros de borde en cada finca, para evitar ruido ocasionado por el efecto borde. Esto permitió crear una tabla de atributos con los valores promedio de cada índice, en cada finca y durante los meses de abril a diciembre (cuando hubo imágenes disponibles). La estimación del valor promedio se basa en la sumar del valor de los píxeles dentro de cada finca y la posterior división entre la cantidad de píxeles (Anexo 5). En la Figura 10 se sintetiza el flujo metodológico de este apartado.

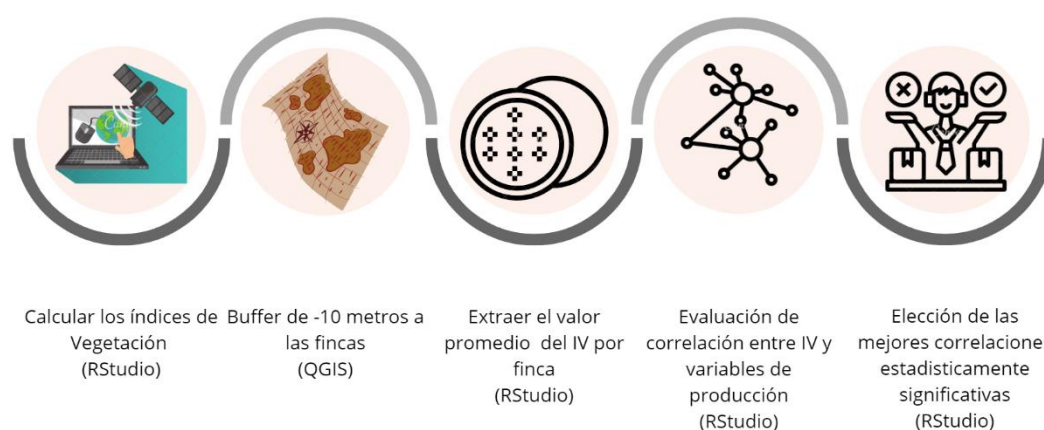


Figura 10. Flujo metodológico para definir los índices de vegetación (IV) que mejor correlación tiene con las variables de producción.

Cuadro 2: Características de los índices de vegetación utilizados para evaluar la evolución espacio-temporal de las plantaciones de caña de azúcar en CoopeVictoria R.L.

Índice	Formula	Utilidad	Referencia
EVI: Enhanced Vegetation Index	$2,5 \times \frac{NIR - Red}{(NIR + 6Red - 7,5Blue + 1)}$	Este índice al igual que el NDVI son buenos estimadores lineales de la fracción fotosintéticamente activa interceptada por la vegetación.	Huete et al. (2002)
GNDVI: Green Normalized Difference Vegetation Index	$\frac{Green - Red}{Green + Red}$	Similar al NDVI pero mide a partir del espectro verde entre 540-570 nm en lugar del espectro rojo. Es más sensible a la concentración de clorofila que el NDVI.	Gitelson et al. (1996)
NDVI: Normalized Difference Vegetation Index	$\frac{NIR - Red}{NIR + Red}$	Mide la vegetación verde y su estado de salud. El uso de las regiones de mayor absorción y reflectancia de la clorofila, así como la diferencia normalizada, lo hace un índice robusto para diferentes condiciones. El rango de valores va entre -1 y 1, donde los más comunes para vegetación verde están entre 0,25 y 0,85.	Rouse et al. (1973)
SAVI: Soil-Ajusted Vegetation Index	$\frac{NIR - Red}{(NIR + Red + 0,5) \times (1 + 0,5)}$	Similar al NDVI pero suprime el efecto ruido de los píxeles de suelo. Utiliza un valor de ajuste a partir de los valores del dosel (L), que es una función de la densidad de vegetación. Se sugiere un valor óptimo de L=0.5 para dar cuenta de las variaciones del suelo de primer orden en el fondo. El mejor empleo de este índice es en áreas con escasa vegetación donde el suelo sea visible.	Huete, 1988
SR: Simple Ratio	$\frac{NIR}{Red}$	Es el cociente simple entre el infrarrojo cercano y el rojo. El infrarrojo cercano es la longitud de onda con mayor reflectancia para la vegetación y el rojo es la longitud de onda con mayor absorción de clorofila. Es capaz de mejorar el contraste entre suelo y vegetación y minimizar efectos en las condiciones de iluminación.	Birth y McVey, 1968

Debido al alto volumen de datos y la heterogeneidad temporal de estos, se optó por analizar la información agrupándola de las siguientes formas:

- a) **Separando por años, por imágenes y por índices:** El análisis se realizó separando los datos de producción de las tres zafas y correlacionándolos contra los distintos índices de vegetación obtenidos a partir de las imágenes que estuvieron disponibles en los años estudiados.
- b) **Separando el valor máximo de los índices en cada año:** En este caso, el análisis se realizó correlacionando las variables de producción y el valor máximo de cada índice de vegetación extraído de las fincas (se analizaron los tres años por separado). El procedimiento anterior se justifica porque todas las fincas fueron cortadas en fechas distintas y los datos suministrados por la Cooperativa no contienen la fecha de corta, por lo cual cada imagen obtenida representa una cantidad diferente de días después de corta.

Al seleccionar el valor máximo de cada índice se estaría considerando la cúspide de la evolución temporal del índice, este momento coincide con una mayor cantidad de área foliar y por lo tanto la mayor fracción de la radiación fotosintéticamente activa absorbida por la plantación (Wiegand et al., 1991). Este además sería el único punto común que se puede identificar en todas las fincas.

- c) **Integrando el valor máximo de cada índice:** Se ordenó toda la información de producción, sin distinción de año, y se correlacionó contra el valor máximo de los índices de vegetación según corresponda. Esto permitió capturar toda la variabilidad temporal de los datos y un aumento en el número de observaciones.
- d) **Integrando el valor promedio de cada índice por mes:** Se obtuvo el valor promedio de cada índice para los meses en análisis (abril a diciembre), ese valor fue correlacionado con los datos de producción. Este proceso permite identificar cuáles son los meses que tienen mejores condiciones para la

estimación. En meses con dos o más imágenes satelitales se calculó el valor promedio de los índices de vegetación disponibles.

En síntesis, los índices fueron evaluados mediante análisis de correlación y evaluación de las significancias estadísticas, esto permitió definir el índice de vegetación y la temporalidad con el mayor potencial para realizar las estimaciones de las variables de productividad en la plantación de caña de azúcar.

3.3 Construcción de modelos de estimación de variables de rendimiento

El desarrollo de los modelos de estimación de variables de rendimiento se realizó tomando en consideración los índices de vegetación más apropiados para evaluar el cultivo, según el apartado anterior. Se utilizaron regresiones lineales múltiples para estimar las variables de rendimiento **PROD** (Toneladas métricas de caña de azúcar por hectárea) y **REND** (Kilos de azúcar por tonelada de caña de azúcar cosechada); a partir de estas variables modeladas se calcularon las **TON TOT** (Total de toneladas métricas de caña de azúcar) y **TOT AZU** (Total de kilogramos de azúcar). En análisis de correlación entre los índices de vegetación y las variables de producción permitió determinar que la mejor agrupación de la información era integrando el valor promedio de cada índice de vegetación por mes. Por lo anterior, para la construcción de los modelos de estimación se utilizó el valor promedio de los índices de vegetación calculados para las fincas en los meses con información disponible en los años en estudio (covariables) y los datos de rendimiento en las mismas zafras (variables de respuesta).

Adicionalmente para aumentar el “n” en la construcción de los modelos de estimación se incluyeron los índices de vegetación y las variables de rendimiento del año 2019. Se incluyeron únicamente los índices de vegetación de agosto y noviembre, porque el análisis estadístico del apartado anterior determinó que esos índices de vegetación tenían mayor correlación con las variables de rendimiento, además eran representativos y constantes para los años en estudio.

Con el fin de aumentar y probar otras combinaciones se incluyeron dos covariables adicionales, el promedio y el máximo de las variables de rendimiento de las zafras en años anteriores, esto con el objetivo de tener una referencia histórica del potencial productivo de cada finca. La adición de nuevas covariables también permitió tomar la decisión de excluir la zafra 2016-2017 del periodo de modelación, ya que solo se obtuvieron índices de vegetación de noviembre, sin embargo, la información de la zafra 2016-2017 sirvió de soporte para la construcción de las covariables adicionales.

Así las cosas, el conjunto de datos utilizados para realizar la modelación contaba con 50 registros (fincas) distribuidos de la siguiente manera: 14 de la zafra 2017-2018, 18 de la zafra 2018-2018 y 18 de la zafra 2019-2020. Este fue el resultado de un proceso de depuración de datos donde se excluyeron las fincas que estuvieron localizadas dentro de imágenes afectadas por nubes en alguno de los meses, así como las fincas en las cuales se detectó la cosecha del cultivo entre agosto y noviembre y finalmente las fincas con menos de 70 ton/ha (eran 5 registros) porque se consideraron fincas bajo condiciones anormales que afectaban negativamente las regresiones.

Para realizar los modelos de estimación de las variables de rendimiento se utilizaron como referencia los índices de vegetación que se encontró tenían mayor relación con las variables de producción (EVI Agosto y NDVI Noviembre) y las covariables adicionales. Sin embargo, también se probaron nuevas combinaciones integrando todas las covariables disponibles que pudiesen explicar las variables de rendimiento en función de los índices de vegetación.

Las regresiones lineales múltiples se construyeron en el entorno de RStudio, utilizando la función *lm()* instalada como función base en R. Este tipo de regresión permite la predicción de una variable de respuesta en función de dos o más predictor (James et al., 2013). La regresión lineal múltiple se puede resumir en la siguiente formula:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n$$

Donde Y es la variable de respuesta, β_0 es el intercepto con el eje y , β_1 es el coeficiente de la covariable X_1 , β_2 es el coeficiente de la covariable X_2 y así sucesivamente para las “ n ” covariables existentes.

La estimación de los coeficientes se basa en el principio de minimizar la suma de los errores al cuadrado, a partir de la siguiente formula:

$$RSS = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$= \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2} - \dots - \hat{\beta}_p x_{ip})^2$$

Donde RSS es la suma de los residuos al cuadrado (por las siglas en inglés), y_i es la variable observada, \hat{y}_i es la variable estimada, $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_p$ son los múltiples coeficientes que minimizan los errores al cuadrado.

- Elección de las covariables utilizadas

Con el objetivo de evaluar las covariables adicionales se construyó una regresión lineal inicial que integraba la totalidad de covariables, esto permitió la seleccionar la mejor combinación de covariables para estimar de forma más ajustada las variables de respuesta.

La selección de covariables, para cada modelo, se realizó utilizando la función *stepwise()* integrada en el paquete StepReg desarrollado por Li et al. (2020), dicha función se ejecutó en RStudio. Esta función permite medir la integralidad del modelo de regresión a partir de criterios de información específicos, la significancia estadística (p-value <0,05) y el Criterio de Información de Akaike (AIC por sus siglas en inglés).

Para la selección de covariables predictoras se utilizó el criterio de selección *bidirection*, el cual va agregando y quitando covariables según sean los efectos sobre el modelo, si el efecto es positivo la covariable se agrega al modelo, por el contrario si el efecto es nulo o negativo la covariable no se incluye (Li et al., 2020).

Una vez definidas las covariables que aportan en la solución de la ecuación se ajustó el modelo tomando en consideración los parámetros error residual estándar (RSE por las siglas en inglés) y coeficiente de determinación ajustado (R^2_a), donde un R^2_a cercano a 1 indica que el modelo explica un gran porcentaje de la varianza de la variable de respuesta (Li et al., 2020).

- Validación de los modelos seleccionados

El ajuste de los modelos seccionados (para PROD y REND) fue validado mediante una validación cruzada (*cross-validation*). Esta metodología resuelve la ecuación del modelo “n”-1 veces. Durante cada corrida se extrae un “n” de conjunto de datos, luego se compara el valor calculado por el modelo para la posición “n” con el valor “n” extraído (Li et al., 2020).

Para realizar este proceso de validación se utilizó el paquete *caret* desarrollado por (Max et al., 2020) y ejecutable en RStudio. La función utilizada fue *trainControl()* y método de validación LOOCV (Leave-one-out cross-validation). Una vez evaluado “n”-1 veces el modelo se calcula el promedio de los siguientes estadísticos RMSE (error cuadrático medio), R^2 (Coeficiente de Correlación) y MAE (Promedio Absoluto del Error), dichos estadísticos permiten evaluar y comparar la calidad de los modelos empleados (Chai y Draxler, 2014; Li et al., 2020; Wang y Lu, 2018).

Aprovechando la variabilidad temporal de los datos (índices en agosto y noviembre) se decidió realizar dos estimaciones de las variables de rendimiento, la primera en agosto, aproximadamente cinco meses antes de inicio cosecha, y una en noviembre, aproximadamente dos meses antes de inicio cosecha, esto con el objetivo evaluar el potencial que tienen las determinaciones temporales de índices de vegetación en la estimación de variables rendimientos y el potencial de intervención en las fincas, esto bajo el principio de monitoreo (Mahlein et al., 2013; Rahman y Robson, 2020).

En síntesis, para elegir los mejores modelos de estimación de las variables PROD y REND se valoraron en primera instancia las covariables identificadas en el apartado anterior, pero también se incluyeron la totalidad de las covariables, esto

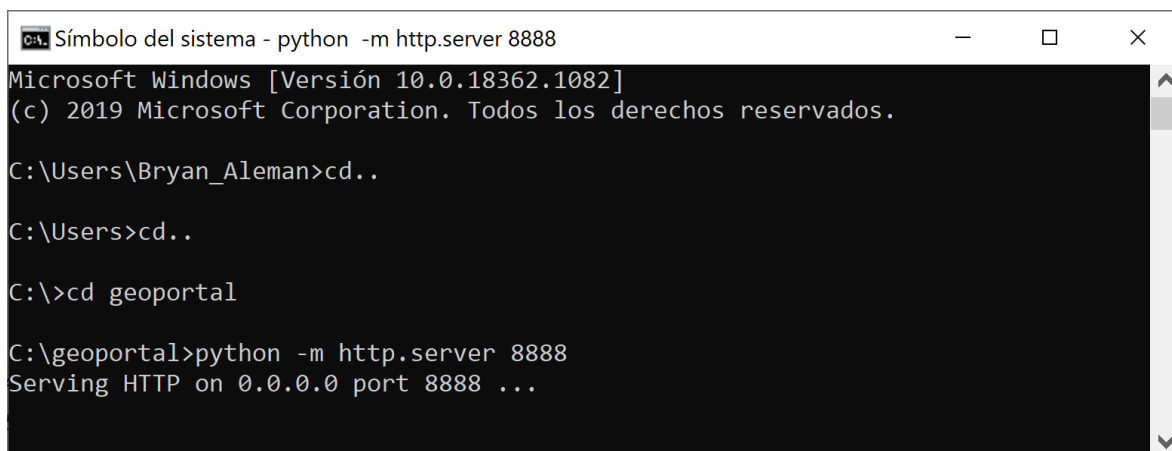
con el objetivo de encontrar relaciones más fuertes entre las variables respuesta y las covariables. Posteriormente, se ejecutaron las regresiones lineales múltiples con las covariables influyentes, para finalmente evaluar el desempeño de los modelos mediante la validación cruzada.

3.4 Construcción del geoportal para transferencia de información

Ante la necesidad de transferir la información generada en esta investigación a los tomadores de decisiones y productores quienes son finalmente los beneficiados de este tipo de herramientas, se planteó construir y poner a disposición un portal web que permita visualizar y hacer consultas básicas de las áreas de interés de la Cooperativa.

La construcción del geoportal se realizó mediante la integración de los lenguajes de programación HTML, CSS y JavaScript. Además, se utilizó el repositorio de datos GitHub y la librería Leaflet para el desarrollo intuitivo, interactivo y ligero del portal de información geográfica (Agafonkin, 2019; Tutorialspoints, 2015).

Como punto de partida se creó un archivo con extensión .html, que contaba con los elementos (tags) necesarios para la programación de la página web, dicho archivo tenía relaciones directas con otros, permitiendo la integración de los elementos en un entorno web. La programación se realizó utilizando como “servidor” de prueba el localhost:8888 a partir de la configuración que se muestra en la Figura 11.



```
Símbolo del sistema - python -m http.server 8888
Microsoft Windows [Versión 10.0.18362.1082]
(c) 2019 Microsoft Corporation. Todos los derechos reservados.

C:\Users\Bryan_Aleman>cd..

C:\Users>cd..

C:\>cd geoportal

C:\geoportal>python -m http.server 8888
Serving HTTP on 0.0.0.0 port 8888 ...
```

Figura 11. Protocolo de activación del localhost:8888 para el desarrollo del geoportal.

El uso del localhost:8888 facilitó las pruebas realizadas porque no dependía estrictamente de internet para el desarrollo de las actualizaciones. Los archivos programados en .html, y .css se presentan en los Anexo 6 y Anexo 7, sin embargo, en la Figura 12 se muestra en encabezado del archivo .html que define las características de la ventana Home en el geoportal; aquí se pueden detallar aspectos como el elemento que define las características del fondo de la página (background-image: url("canal.jpg");), los links existentes entre las ventanas (Home), los link con las características de estilo, archivos .css (<link rel="stylesheet" type="text/css" href="style.css">), entre otros detalles.

```

1 <!DOCTYPE html>
2 <html lang="es">
3 <meta name="viewport" content="width=device-width, initial-scale=1.0">
4
5 <head>
6
7 <title>Plantaciones de caña de azúcar en CoopeVictoria R.L.</title>
8
9 <style>
10   body {
11     background-image: url("canal.jpg");
12     background-size: 1269px 843px;
13     background-repeat: no-repeat;
14   }
15 </style>
16 <meta charset="UTF-8">
17 <!-- Hoja de estilos CSS adicionales -->
18 <link rel="stylesheet" type="text/css" href="style.css">
19 <link rel="stylesheet" type="text/css" href="./style_barra.css">
20
21 </script>
22
23 </head>
24 <body>
25
26
27 <div class="row">
28 <div class="topnav">
29 <a class="active" href="index.html">Home</a>
30 <a href="smscopevictoria.html">SMS-CoopeVictoria R.L.</a>
31 <a href="contactos.html">Contactos</a>
32 </div>
33
34 <h1>Sistema de Monitoreo Satelital de las Plantaciones de Caña de Azúcar en CoopeVictoria R.L. </h1>
35

```

Figura 12. Encabezado del archivo .html tiene los detalles de la ventana home cuando se accede al geoportal.

Los detalles de las capas de información geográfica que se colocaron a disposición de los usuarios tienen características importante de resaltar; los archivos fueron almacenados en formato .geoJson para disminuir el tamaño en megabits y aumentar la velocidad de transferencia de información.

Se colocaron para consulta en el repositorio web las siguientes capas de información:

- **Áreas de Producción:** este archivo corresponde a los límites de cada unidad productiva definidos por la Cooperativa, se incluyeron únicamente las fincas propias y arrendadas. Dicho archivo había sido corregido y editado en las etapas previas de esta investigación. Con base en las áreas de producción se presentan las siguientes características de las fincas: nombre, número de lote, variedad sembrada y área.
- **Distritos de Influencia:** este archivo permite la visualización de los límites político-administrativos dentro de los cuales se sitúan las fincas de la

Cooperativa. Aquí se muestran los detalles: Provincia, Cantón, Distrito y Área.

- **Historial de cosecha por finca:** permite visualizar el historial del rendimiento de campo de cada finca definida por la Cooperativa, se presenta la información de las zafras: 2016-2017, 2017-2018, 2018-2019 y 2019-2020.
- **NDVI calculado con el sensor Sentinel el 22/11/2019:** este archivo tiene como objetivo evidenciar la factibilidad de transferir información a los tomadores de decisiones mediante este geoportal.

En primera instancia se calculó el NDVI el cual posteriormente fue recordado con base en los límites de las fincas. Posteriormente se realizó una vectorización del ráster, ya que la plataforma no soporta archivos ráster. Ese nuevo archivo vectorial fue reescalado con base a 100 y posteriormente convertido a valor entero para disminuir el peso.

En la plataforma se podrá observar el valor del índice NDVI, el cual debe ser dividido entre 100 para obtenerlo en rangos de valor de 0 a 1.

Las cuatro acciones descritas fueron programadas en un archivo JavaScript (Anexo 8), en dicho archivo fueron programadas las funciones de selección y despliegue de información de las capas de información.

Los mapas base utilizados se obtuvieron de la librería de código abierto Leaflet, la cual está programada en JavaScript y facilita interactividad y conectividad con mapas base debido a su reducido peso de almacenamiento en los dispositivos (Agafonkin, 2019).

Finalmente, una vez programado y configurado el geoportal se puso a disposición de los usuarios a través de internet utilizando el repositorio de datos GitHub, el cual provee de espacio de almacenamiento, programación y ejecución de funciones. Este repositorio posee una serie de características de almacenamiento, en este caso se utilizó una versión gratuita que cuenta con las siguientes características: repositorios públicos y privados ilimitados, número de colaboradores ilimitados,

máximo 2000 minutos por mes, 500 megasbits de almacenamiento y soporte comunitario (GitHub, 2020).

Hasta aquí y con base en todo lo anterior, los procesos descritos hasta este momento marcaron la guía procedimental de esta investigación para el cumplimiento de las metas planteadas. En los apartados posteriores se mostrarán los resultados obtenidos a partir de este desarrollo metodológico.

Capítulo 4: Resultados y Discusión

4.1. Identificación del sensor satelital de mayor viabilidad para establecer el sistema de monitoreo

Los resultados del Cuadro 1, evidencia que los tres sensores evaluados cubren los rangos espectrales: visible e infrarrojo cercano, razón por la cual, todos poseen la resolución espectral adecuada para el cálculo de índices de vegetación. Sin embargo, al evaluar la resolución espacial de dichos sensores existen diferencias marcadas, así se demuestra en las Figura 13A y Figura 13B donde se realizó una comparación del área real de las fincas y el área calculada a partir del número de pixeles de los tres sensores inmersos en las fincas.

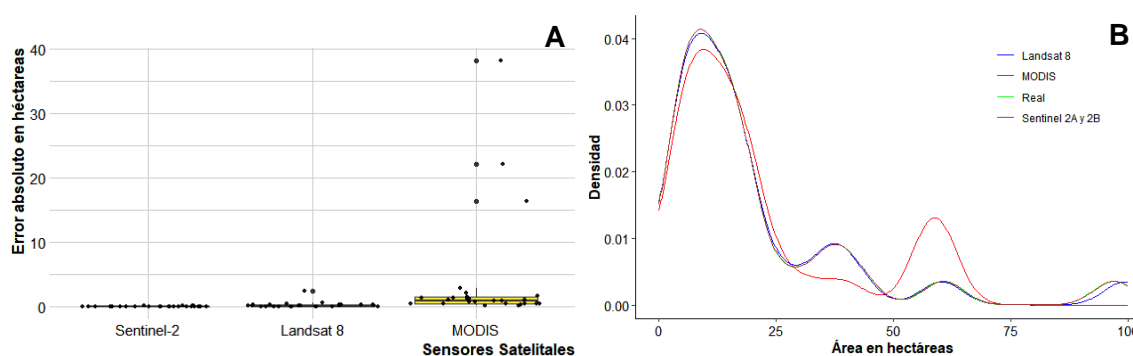


Figura 13. (A) Error absoluto en hectáreas al calcular el área de las fincas a partir del número de pixeles extraídos por sensor para cada unidad productiva. (B) Comparación del área real de las fincas contra las áreas calculadas por el número de pixeles.

La Figura 13A deja en evidencia que los errores en la estimación del área a partir del número de pixeles aumentan conforme incrementa la resolución espacial, siendo que, en dos fincas, el área calculada a partir del número de pixeles de MODIS tiene errores de aproximadamente 38 hectáreas, para el caso de las áreas calculadas con el sensor Landsat 8 se evidencian errores de alrededor de tres hectáreas y finalmente el sensor Sentinel 2 ajusta de mejor manera el cálculo de área a partir del número de pixeles, ya que no se observan errores que superen una hectárea (Figura 13A).

Así mismo, al contrastar el comportamiento de las áreas calculadas mediante el número de pixeles (para los tres sensores) y el área real, se puede observar que el

sensor MODIS presentan la mayor disparidad con respecto al área real de las fincas, siendo más acentuados los errores en fincas de 10, 35, 60 y 99 hectáreas (Figura 13A). Por otra parte, el sensor Landsat 8 presenta las mayores disparidades en fincas con áreas de aproximadamente 10 y 95 hectáreas, mientras que existe un ajuste casi perfecto entre las áreas calculadas a partir del sensor Sentinel 2 y el área real (Figura 13B).

Las imágenes Sentinel 2 tiene aproximadamente 100,2 píxeles por hectárea, mientras las Landsat 8 llegan a 11,3 píxeles por hectárea y las imágenes MODIS sólo presentan 0,04 píxeles por hectárea; la anterior evidencia que el sensor Sentinel 2 es el más adecuado para analizar la variabilidad espacial del cultivo a partir de índices de vegetación, mientras que el sensor MODIS presenta serias limitaciones para evaluar la variabilidad espacial de cultivos de poca estación, como es el caso de las fincas de CoopeVictoria R.L. y sus asociados.

La Figura 14 evidencia las disparidades existentes entre los tres sensores en resolución espacial, es claro que la imagen Sentinel 2B se ajusta de mejor manera a límites de las fincas, caso contrario sucede con la imagen MODIS donde sólo se pueden observar 2 píxeles para la misma área, por otro lado la imagen Landsat se considera una condición intermedia.

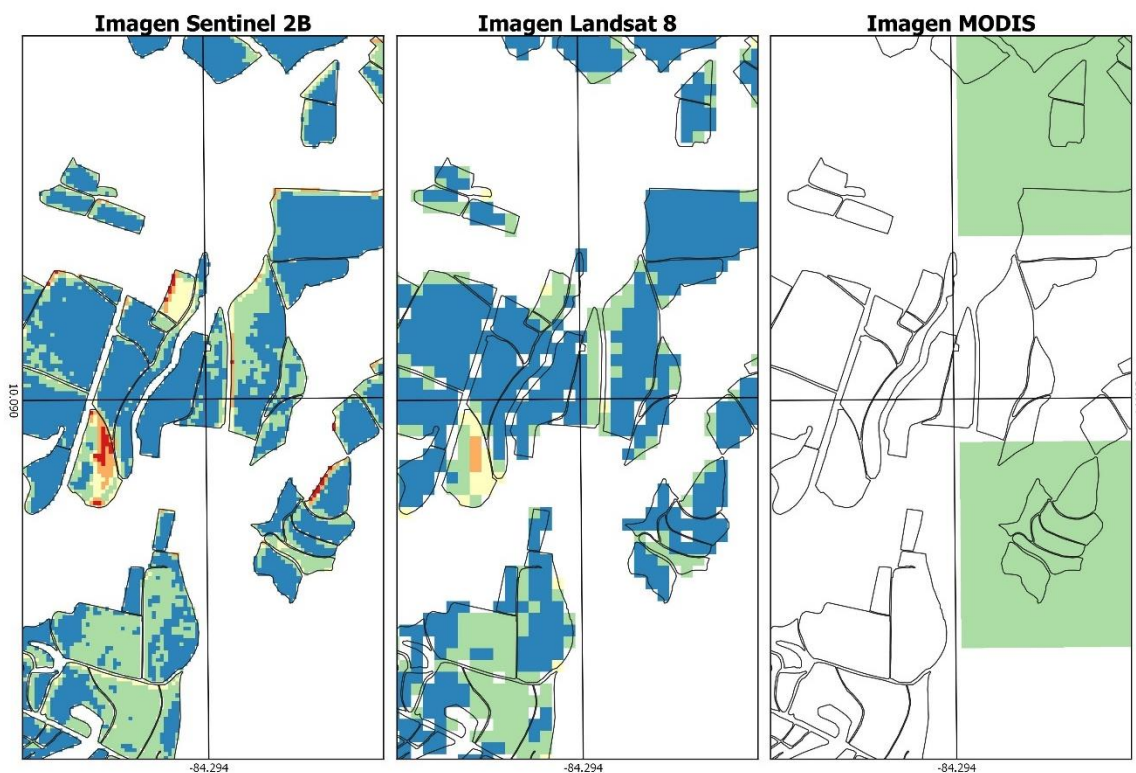


Figura 14. Variación espacial de los pixeles extraídos por finca para los tres sensores satelitales en estudio.

Las tres secciones de la Figura 14 corresponde a la extracción de los pixeles de contenidos dentro de los límites de cada finca, específicamente corresponde al cálculo del índice de vegetación NDVI de finales de diciembre de 2018 (27 y 30 de diciembre) lo cual hace esperar un comportamiento similar en la variación espacial del índice de vegetación, sin embargo, es claro como la imagen Sentinel 2B representa de mejor manera los resultados.

Según lo indicado, las imágenes del sensor Sentinel 2 presentan las mayores virtudes para desarrollar el sistema de monitoreo satelital en CoopeVictoria R.L., esto porque cumple con los requerimientos espectrales, temporales y principalmente los espaciales. Este sensor presenta la mejor resolución espacial y temporal, en relación con la resolución espectral tiene similitud de condiciones con los otros sensores.

Por otra parte, el sensor MODIS presentó las capacidades más limitadas principalmente por resolución espacial, ya que las áreas de interés (Fincas) requiere

de unidades de análisis (píxeles) pequeños para poder ser evaluados, por esta razón se descarta este sensor, aunque presente otras bondades como el nivel de procesamiento.

El sensor Landsat 8 presenta características relevantes, sin embargo, la resolución espacial limita su uso en áreas con unidades productivas pequeñas como lo son las fincas de CoopeVictoria R.L., además la mediana resolución temporal disminuye las posibilidades de encontrar imágenes satelitales de calidad para el área de interés. Por lo tanto, este sensor al igual que MODIS, no será considerado en el siguiente apartado en la evaluación de la correlación entre índices de vegetación calculados a partir de sus bandas y las variables de rendimiento del cultivo de la caña de azúcar.

Con base en todo lo anterior, el sensor Sentinel 2 es el seleccionado para realizar el sistema de monitoreo satelital en CoopeVictoria R.L. en este estudio, debido a que presenta la combinación de las mejores características en términos de resolución espacial, temporal y espectral.

4.2. Índices de vegetación como indicadores de rendimiento del cultivo de la caña de azúcar

Los índices de vegetación se han convertido en un insumo de importancia para el monitoreo de las plantaciones agrícolas (Hatfield y Prueger, 2010) porque permiten la evaluación, no invasiva, de la cobertura, el vigor y la dinámica de las plantaciones (Xue y Su, 2017); estas características fundamentales de los índices de vegetación fueron relacionadas con las variables de rendimiento facilitadas por la Cooperativa para esta investigación.

Debido a las condiciones atmosféricas y la resolución temporal del sensor Sentinel 2A y Sentinel 2B se obtuvo un número limitado de imágenes satelitales en el área de estudio (Cuadro 3) para la temporalidad de interés. A partir de dichas imágenes satelitales se calcularon los índices de vegetación planteados en el Cuadro 2.

Cuadro 3. Fechas con imágenes satelitales disponibles del sensor Sentinel 2A y Sentinel 2B para el área de estudio por mes y año.

Mes	2016	2017	2018
Abril			
Mayo			6/5/2018 11/5/2018
Junio		25/6/2017	
Julio	20/7/2016		
Agosto		14/8/2017	19/8/2018
Septiembre			13/9/2018
Octubre			
Noviembre	7/11/2016	17/11/2017	7/11/2018
Diciembre	7/12/2016 27/12/2016	2/12/2017 22/12/2017	7/12/2018 12/12/2018 17/12/2018 27/12/2018

Los resultados del Cuadro 3 evidencian poca constancia en la disponibilidad de imágenes Sentinel 2 por año y por mes, lo cual implicó la necesidad de aplicar distintos análisis para identificar los índices de vegetación y las fechas que tienen mayores correlaciones con las variables de productividad.

En primera instancia, la prueba de normalidad Shapiro-Wilk aplicada a las variables de rendimiento, mostró que los datos no presentaban una distribución normal, se realizaron pruebas de normalización de los datos, pero los resultados no fueron satisfactorios; por lo tanto, fue necesario aplicar, para todos los análisis de

correlación, la prueba de correlación Spearman utilizada para conjuntos de datos no-paramétricos (Rebekić et al., 2015), como lo eran las variables de rendimiento de esta investigación.

Los distintos análisis de correlación se agruparon en cuatro criterios según se describe a continuación:

- **Separando por años, por imágenes y por índices:**

La evaluación de los coeficientes de correlación entre las variables de rendimiento y los índices de vegetación mostraron poca constancia. Las variables RENDI, TON TOT y TOT AZU únicamente presentaron correlaciones superiores a 0,6 en la zafra 2016-2017, caso contrario sucedió con la variable PROD que presentó correlaciones superiores a 0,6 en las tres zafras analizadas y en combinación con diferentes índices de vegetación (Cuadro 4).

Las mejores correlaciones entre la variable PROD y los índices de vegetación, en las zafras 2016-2017 y 2017-2018, se presentaron con los índices de vegetación calculados en noviembre e inicios de diciembre; en la zafra 2018-2019 las mejores correlaciones se encontraron con la información del mes setiembre (Cuadro 4).

Cuadro 4. Correlación de Spearman entre las variables de rendimiento y los índices de vegetación calculados para cada imagen Sentinel 2A y Sentinel 2B, en los años 2016, 2017 y 2018. Solo se muestran las correlaciones superiores a 0,6.

Variables de Rendimiento	Valor Índice de Vegetación	R²	p-value
PROD 2016	GNDVI 20161107	0,72	0,000110606
PROD 2016	NDVI 20161107	0,60	0,002425316
PROD 2016	SAVI 20161107	0,61	0,00204798
PROD 2017	EVI 20171202	0,65	0,000745479
PROD 2017	NDVI 20171202	0,63	0,001290381
PROD 2017	SAVI 20171117	0,64	0,004159551
PROD 2017	SAVI 20171202	0,64	0,000906699
PROD 2018	EVI 20180819	0,71	5,99E-05
PROD 2018	EVI 20180913	0,74	2,22E-05
PROD 2018	GNDVI 20180819	0,66	0,000290265
PROD 2018	GNDVI 20180913	0,69	0,000125786
PROD 2018	NDVI 20180819	0,66	0,000368564
PROD 2018	NDVI 20180913	0,70	0,000105786
PROD 2018	SAVI 20180819	0,71	7,59E-05
PROD 2018	SAVI 20180913	0,76	1,27E-05
PROD 2018	SR 20180819	0,69	0,00014903
PROD 2018	SR 20180913	0,70	0,000105786
RENDI 2016	EVI 20161227	-0,71	0,000275422
RENDI 2016	SAVI 20161227	-0,70	0,000396943
TON TOT 2016	EVI 20161227	0,70	0,000457041
TON TOT 2016	NDVI 20161107	0,66	0,000625251
TON TOT 2016	SAVI 20161227	0,65	0,001488734
TON TOT 2016	SR 20161107	0,76	2,50E-05
TON TOT 2016	SR 20161207	0,68	0,000978282
TOT AZU 2016	EVI 20161227	0,63	0,002401601
TOT AZU 2016	NDVI 20161107	0,64	0,00095138
TOT AZU 2016	SR 20161107	0,73	7,63E-05
TOT AZU 2016	SR 20161207	0,62	0,003681419

n = 23 (2016 y 2017) n = 26 (2018)

Los índices de vegetación que tuvieron las mejores correlaciones con PROD fueron GNDVI (2016-2017), EVI (2017-2018) y SAVI (2018-2019), lo cual confirma que los índices de vegetación no tuvieron correlaciones constantes con las variables de rendimiento.

Estos resultados demuestran la necesidad de buscar otras formas de agrupar la información para obtener resultados representativos para las tres zafras en análisis y no una correlación específica para una zafra con condiciones particulares en el cultivo.

- **Separando el valor máximo de los índices en cada año:**

Este abordaje identificando el valor máximo del cada índice buscó determinar la fecha con mayor actividad fotosintética del cultivo (Wiegand et al., 1991), ya que fue el único punto en común que se pudo identificar para todas las fincas, a partir de la disponibilidad de imágenes satelitales y de la información base.

Al evaluar la correlación del valor máximo de cada índice de vegetación por año, con el respectivo valor de las variables de rendimiento, se identificó que solo las variables PROD 2016, PROD 2017, TON TOT 2016 y TOT AZU 2016 presentaron correlaciones superiores a 0,6 (Cuadro 5), lo cual hace ver la poca constancia en patrones de correlación.

Cuadro 5. Correlación de Spearman entre las variables de rendimiento y el valor máximo obtenido en cada índice de vegetación calculado con Sentinel 2A y Sentinel 2B, según año (2016, 2017 y 2018). Solo se muestran las correlaciones superiores a 0,6.

Variables de Rendimiento	Valor Máximo Índice de Vegetación	R²	p-value
PROD 2016	MAX GNDVI	0,68	0,00033344
PROD 2017	MAX EVI	0,71	0,00014996
PROD 2017	MAX NDVI	0,77	1,55E-05
PROD 2017	MAX SAVI	0,76	2,22E-05
PROD 2017	MAX SR	0,75	3,92E-05
TON TOT 2016	MAX SR	0,66	0,00062525
TOT AZU 2016	MAX SR	0,63	0,00129038

n=23 (2016 y 2017). n=26 (2018)

Las correlaciones muestran que la variable PROD de las zafras 2016-2017 y 2017-2018 tuvo buenas correlaciones con el MAX GNDVI y MAX NDVI respectivamente. Asimismo, las variables TON TOT 2016 y TOT AZU 2016 tuvieron las mejores correlaciones con el MAX SR (Cuadro 5).

Es imposible identificar un patrón que permita definir el valor máximo de algún índice de vegetación como indicador de buenas correlaciones para las tres zafras en investigación. La variable de rendimiento PROD no presentó buena correlación con los valores máximos de los índices de vegetación del año 2018.

- **Integrando el valor máximo de cada índice:**

Evaluando de forma integral el valor máximo de cada índice de vegetación y sus respectivas variables de rendimiento se obtiene que el MAX NDVI presenta el mejor R^2 (0,53), el cual es superior al 0,47 obtenido con el MAX SR (Cuadro 6). El MAX EVI no presentó un coeficiente de correlación superior a 0,4.

Cuadro 6. Correlación de Spearman entre las variables de rendimiento y el valor máximo obtenido en cada índice de vegetación calculado con Sentinel 2A y Sentinel 2B en los años 2016, 2017 y 2018. Solo se muestran las correlaciones superiores a 0,4.

Variables de Rendimiento	Valor Máximo Índice de Vegetación	R²	p-value
PROD	MAX GNDVI	0,44	9,71E-05
PROD	MAX NDVI	0,53	1,95E-06
PROD	MAX SAVI	0,41	0,00031063
PROD	MAX SR	0,47	2,57E-05

n=72

Este método de agrupación requiere una serie de tiempo para identificar el valor máximo de los índices de vegetación por finca, ya que todas las áreas presentaron una cantidad diferente de días después de siembra de zafra, esto hace más compleja la adquisición de la información, ya que se requiere una búsqueda continua de información para identificar la cúspide en el valor del índice de vegetación para todas las Fincas.

- **Integrando el valor promedio de cada índice por mes:**

Al observar la poca constancia que existe al evaluar los índices de vegetación de forma individual, se optó por analizar la correlación entre las variables de rendimiento y el promedio de cada índice de vegetación por mes. La interpretación de los resultados del Cuadro 7 debe ser analizada tomando como referencia la

información del Cuadro 3, porque en este se identifica la disponibilidad de imágenes por mes en cada año.

Los meses de noviembre y diciembre presentaron disponibilidad de imágenes en los tres años analizados, seguido del mes agosto que tuvo disponibilidad imágenes en los años 2017 y 2018, los otros meses presentaron disponibilidad de imágenes solo en uno de los años o bien en ninguno (Cuadro 3). Por tanto, solo las correlaciones obtenidas en noviembre y diciembre serán representativas para el periodo de estudio.

Cuadro 7. Correlación de Spearman entre las variables de rendimiento y el promedio de los índices de vegetación calculados por mes con Sentinel 2A y Sentinel 2B. Sólo se muestran las correlaciones superiores a 0,4.

Variables de Rendimiento	Índice de Vegetación y Mes	R²	p-value
TON TOT	SR Noviembre	0,42	0,000569872
TOT AZU	NDVI Noviembre	0,43	0,000438825
TOT AZU	SR Noviembre	0,45	0,000191243
PROD	GNDVI Julio	0,46	0,048562673
PROD	GNDVI Agosto	0,49	0,000603037
PROD	SAVI Agosto	0,50	0,00050307
PROD	EVI Agosto	0,53	0,000195438
PROD	NDVI Agosto	0,55	0,000107668
PROD	SR Agosto	0,60	1,60E-05
PROD	GNDVI Septiembre	0,69	0,000125786
PROD	NDVI Septiembre	0,70	0,000105786
PROD	SR Septiembre	0,70	0,000105786
PROD	EVI Septiembre	0,74	2,22E-05
PROD	SAVI Septiembre	0,76	1,27E-05
PROD	SR Noviembre	0,52	9,92E-06
PROD	GNDVI Noviembre	0,53	6,03E-06
PROD	EVI Noviembre	0,54	5,25E-06
PROD	SAVI Noviembre	0,54	4,89E-06
PROD	NDVI Noviembre	0,55	2,56E-06
PROD	SAVI Diciembre	0,41	0,000367976

n=72

Los resultados del Cuadro 7 evidencian correlaciones medias y positivas entre los índices de vegetación y las variables de rendimiento TON TOT, TOT AZU y PROD

en el mes noviembre, las correlaciones de dichas variables de rendimiento en otros meses no son representativas para el periodo de análisis.

La variable de rendimiento TON TOT tuvo la mejor correlación con el índice de vegetación SR Noviembre (0,42). La variable TOT AZU presentó las mejores correlaciones con los índices de vegetación NDVI Noviembre (0,43) esta correlación media coincide con lo encontrado por Wiegand et al. (1991), quienes lograron obtener buenas correlaciones entre el NDVI y el total de toneladas de azúcar. Por otra parte, el índice SR Noviembre tuvo un coeficiente de correlación de 0,45 con la variable TON AZU (Cuadro 7).

En relación con la variable PROD los 5 índices de vegetación calculados en noviembre presentaron correlaciones superiores a 0,5, siendo el coeficiente de correlación de 0,55 la correlación más fuerte entre NDVI Noviembre y PROD (Cuadro 7), lo anterior es consistente con lo encontrado en el Cuadro 6, donde el MAX NDVI presenta el mayor coeficiente de correlación con la variable PROD.

En la Figura 15 es evidente que la plantación alcanza el punto máximo de NDVI alrededor de noviembre, esto permite relacionar el punto de mayor porción fotosintéticamente activa de la vegetación con el mejor momento para realizar la estimación de PROD, también es evidente el comportamiento que sucede en agosto, donde se observa un cambio en la pendiente del crecimiento del NDVI, dicho comportamiento podría estar relacionado con el final de la etapa de macollamiento y desarrollo del dosel, como se observa en la Figura 2.

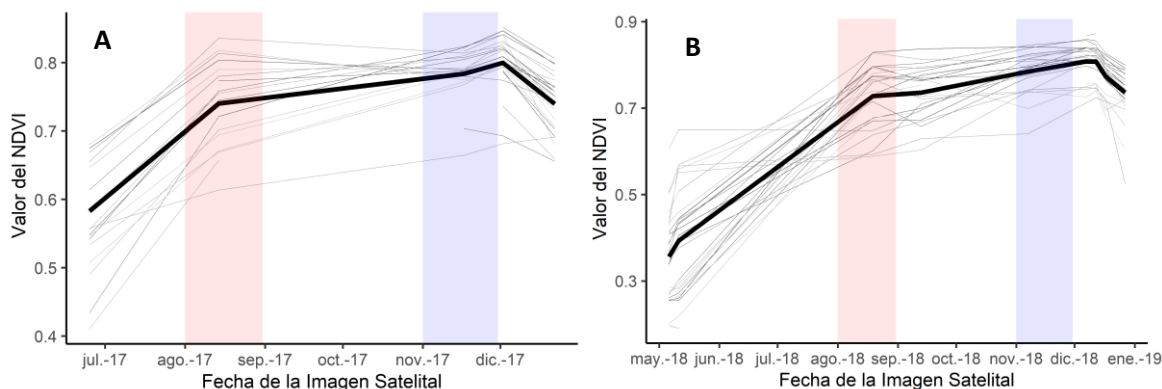


Figura 15. Variación temporal del NDVI en las fincas incluidas en la investigación. (A) Año 2017. (B) Año 2018. En ambos casos las líneas grises representan las fincas y la línea negra indica el promedio del NDVI considerando todas las fincas. Las áreas remarcadas corresponden a los meses agosto (rojo) noviembre(azul).

La información del Cuadro 7 permite concluir que la mejor fecha para realizar estimación de variables de rendimiento es el mes de noviembre, aproximadamente dos meses antes de iniciada la zafra; lo anterior es coincidente con los identificado por Rahman y Robson (2020).

Es de resaltar que en los casos anteriores la variable de rendimiento REND no presentó correlaciones buenas o medias con algún índice de vegetación, lo cual es congruente con los encontrado por (Shukla et al., 2017).

Con los resultados anteriores se puede concluir que la agrupación de índices de vegetación por mes es la más representativa para al conjunto de datos. Por tanto, en el análisis de regresión del siguiente apartado, se modelará, como punto de partida, la variable de rendimiento PROD con base en los valores de los índices de vegetación obtenidos para el mes de noviembre. Además, las variables TON TOT y TOT AZU serán modeladas en función del índice de vegetación SR calculado para noviembre. Sin embargo, también se valorarán otras posibles combinaciones de covariables.

4.3 Modelos de regresión para la estimación de variables de rendimiento

Se construyeron los modelos lineales únicamente con las covariables definidas como promisorias del apartado anterior, sin embargo, los resultados de la

modelación no fueron satisfactorios. Lo anterior se puede explicar por la naturaleza de los métodos de correlación, estos métodos representan una relación lineal entre dos variables; sin embargo, cuando se integran otras variables en la regresión se pasa de una regresión lineal simple a una regresión lineal múltiple donde la interacción entre variables independientes puede afectar las relaciones lineales simples (Li et al., 2020).

Por lo tanto, se evaluaron todas las covariables en función de las variables de rendimiento. A continuación, se presentan los resultados de la selección de covariables y la modelización de las variables PROD y REND, la variable de rendimiento TON TOT se calculó a partir de las regresiones realizadas para PROD; además no se realizó la regresión TOT AZU por detalles que se expondrán posteriormente.

- **Modelización de la variable rendimiento de campo (PROD)**

El proceso de selección de covariables de estimación definió que las covariables que aportan para el modelo más parsimonioso (menor AIC) en el mes agosto fueron: **EVI Agosto, SAVI Agosto y PROD promedio de años anteriores**. Una vez ejecutada la regresión lineal múltiple con los datos de agosto los resultados fueron los siguientes:

$$\begin{aligned}
 & \textit{PROD Agosto} \\
 & = 38,9975390 + (374,8274384 \times \textit{EVI Agosto}) \\
 & + (-371,0561799 \times \textit{SAVI Agosto}) \\
 & + (0,3953307 \times \textit{PROD promedio de años anteriores})
 \end{aligned}$$

Donde PROD Agosto es el rendimiento de campo estimado para el mes de agosto, EVI Agosto es el promedio del índice de vegetación EVI calculado en ese mes, SAVI Agosto es el índice de vegetación SAVI calculado en el mismo mes y PROD promedio de años anteriores es el promedio del rendimiento de campo obtenido en años anteriores.

Los coeficientes fueron estadísticamente significativos (p-value<0,01), la regresión PROD Agosto presentó un error residual estándar de 8,38 ton/ha, un coeficiente de

determinación ajustado de 0,48 y el p-value de $2,779 \times 10^{-07}$. La validación del modelo PROD Agosto presentó un RMSE de 8,86 ton/ha, un coeficiente de correlación (R^2) de 0,40 y un MAE de 6,94 ton/ha.

Por otra parte, la selección de covariables para la construcción del modelo de estimación de noviembre identificó que **EVI Agosto**, **SAVI Agosto**, **PROD promedio de años anteriores** y **SR Noviembre** en conjunto integraban el modelo más parsimonioso al presentar el AIC más bajo.

Al utilizar las covariables seleccionadas para el mes de noviembre la regresión lineal múltiple permitió definir el siguiente modelo:

$$\begin{aligned} \text{PROD Final} = & 11,7382375 + (271,6416259 \times \text{EVI Agosto}) \\ & + (-243,2335433 \times \text{SAVI Agosto}) \\ & + (0,2863842 \times \text{PROD promedio de años anteriores}) \\ & + (3,3586580 \times \text{SR Noviembre}) \end{aligned}$$

Donde PROD Final es el rendimiento de campo estimado integrando los índices de vegetación de ambos meses, EVI Agosto es el promedio del índice de vegetación EVI calculado en agosto, SAVI Agosto es el índice de vegetación SAVI calculado en el mismo mes, PROD promedio de años anteriores es el promedio del rendimiento de campo obtenidos en años anteriores y finalmente SR Noviembre es el índice de vegetación SR calculado en noviembre.

Los parámetros de ajuste de este modelo PROD Final indican que todos los coeficientes son estadísticamente significativos ($p\text{-value} < 0,05$), el error residual estándar del modelo PROD Final es de 6,64 ton/ha, con un coeficiente de determinación ajustado de 0,67 y el p-value de $2,951 \times 10^{-11}$.

La validación cruzada del modelo PROD Final tuvo como parámetros un RMSE de 7,10 ton/ha, con un coeficiente de correlación de 0,62 y un promedio absoluto del error (MAE) de 5,58 ton/ha.

Los resultados obtenidos de los modelos anteriores permiten definir que existe un aporte aproximado de 20% en la explicación de la variable dependiente cuando se incluye la información del mes noviembre, específicamente el índice de vegetación

SR. Al comparar ambos modelos se identificó que todos los parámetros estadísticos del modelo PROD Final mejoraron en comparación con el modelo PROD Agosto.

La validación de los modelos también es contundente al demostrar que los parámetros, RMSE, R^2 y MAE, son mejores cuando se integra la información de ambos meses. Se ejecutaron las regresiones lineales múltiples calibradas al conjunto de datos, lo cual permitió confirmar que se obtiene un mejor coeficiente de correlación cuando se integra la totalidad de la información en la construcción del modelo (Figura 16A).

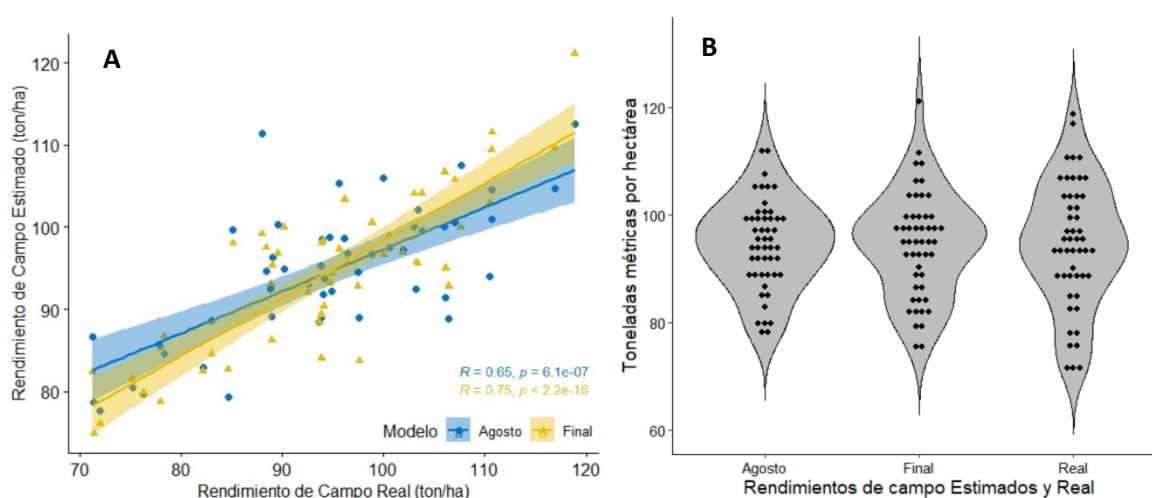


Figura 16. (A) Coeficiente de correlación entre los rendimientos de campo y real. El rendimiento estimado se obtuvo ejecutando las regresiones sobre el conjunto de datos. (B) Comparación del valor real y los estimados en rendimiento de campo.

En la Figura 16B también es posible observar que los modelos estimados presentaron una mayor densidad de datos entre el rango de 95 ton/ha y 105 ton/ha, sin embargo, también es preciso indicar que el modelo final tiene mayor similitud con el rendimiento de campo real, el modelo final también permitió modelar de mejor manera los datos más alejados de la región intercuartil.

Posteriormente se calculó el total de toneladas de caña de azúcar producidas por finca (TON TOT) multiplicando el resultado de la PROD Agosto y la PROD Final por el área de cada finca.

Al aplicar ambos modelos (PROD Agosto y PROD Final) en la estimación TON TOT se obtuvo un coeficiente de correlación muy bueno entre los datos observados y los datos modelados (Figura 17A). Sin embargo, las estimaciones presentaron errores en el cálculo del total de toneladas producidas. En la Figura 17B se observa la distribución de los datos estimados, con lo que se podría deducir que la mayoría de las fincas producen menos de 2500 toneladas totales de caña de azúcar, además muy pocas fincas tiene producciones totales superior a 5000 toneladas. Es claro que la modelización permitió capturar la tendencia global del total de toneladas de caña de azúcar cosechadas de las fincas estudiadas.

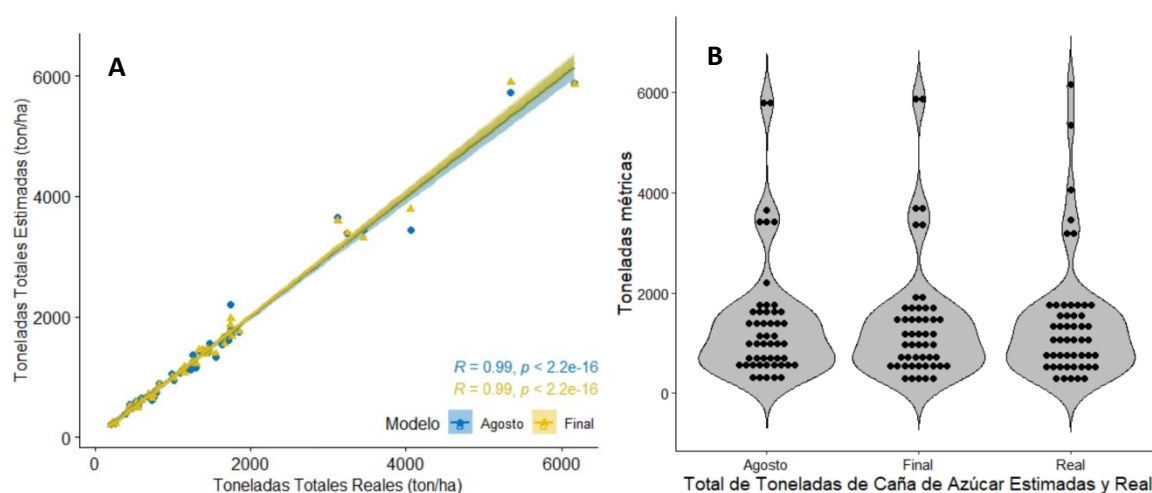


Figura 17. (A) Correlación entre las toneladas totales reales y las estimadas a partir de los modelos de regresión PROD Agosto y PROD Final, los datos se obtuvieron multiplicando la estimación de los modelos de regresión por la respectiva área. (B) Comparación del valor real y los estimados en toneladas métricas totales.

El total de toneladas de caña de azúcar producidas en las fincas estudiadas durante el periodo de estudio según los registros de la Cooperativa fue 69413,21, mientras las toneladas totales de caña de azúcar calculadas con el modelo de estimación de PROD Agosto fue de 69559,2 toneladas y utilizando el modelo de estimación PROD Final se estimaron 69778,16 toneladas. En términos relativos la estimación en el mes agosto tiene un error promedio de estimación de 0,75%, mientras el modelo PROD Final presentó 0,48% de error promedio en la estimación (Cuadro 8). Lo anterior sugiere una muy buena estimación con el modelo usado.

Cuadro 8. Resumen de los errores relativos en la estimación del total de toneladas de caña de azúcar, utilizando los modelos PROD Agosto y PROD Final. CoopeVictoria R.L.

Modelo	Mínimo	1° cuartil	Mediana	Promedio	3° Cuartil	Máximo
PROD Agosto	-16,60%	-5,30%	0,08%	0,75%	6,47%	26,60%
PROD Final	-14,26%	-4,58%	0,38%	0,48%	4,68%	15,62%

El Cuadro 8 reafirma que el modelo final tuvo un mejor desempeño; es posible observar una reducción de aproximadamente 11% de error en el valor máximo y 2% en el valor mínimo. En términos generales ambos modelos sobreestimaron la producción total.

La diversidad de abordaje metodológicos ha permitido que muchos autores hayan logrado obtener diferentes tipos de conclusiones en investigaciones similares que incluyen el cultivo caña de azúcar e índices de vegetación (Abdel-Rahman y Ahmed, 2008). Un claro ejemplo es la correlación de 0,75 entre contenidos de azúcar y NDVI obtenido por Wiegand et al. (1991), asimismo Rahimi-Jamnani et al. (2019) que encontraron relación entre la concentración de sacarosa y el índices de vegetación GNDVI; dichos datos son contrarios a los encontrados en este estudio.

Los resultados de esta investigación son coincidentes con los encontrados Mutanga et al. (2013), Rahman y Robson, (2020), Wiegand et al. (1991) y Gunnula et al. (2011) donde indican que aproximadamente dos meses antes del inicio del periodo de cosecha es la mejor fecha para realizar la estimación de variables de rendimiento en el cultivo caña de azúcar, en este momento apenas ha iniciado el estrés por maduración.

Los modelos de estimación PROD Agosto y PROD Final obtuvieron RMSE de 8,86 ton/ha y 7,10 ton/ha respectivamente, ambos parámetros son superiores a los identificados por Rahman y Robson, (2020) que utilizando el índice de vegetación GNDVI, para estimar el rendimiento de campo, alcanzaron un RMSE de 11,33 ton/ha, de igual manera Lisboa et al., (2018) obtuvieron un RMSE de 10,34 ton/ha utilizando regresiones lineales del índice NDVI calculado con el espectroradiómetro

FieldSpec 4 y Morel et al., (2014) utilizando regresiones lineales e índices NDVI calculados con imágenes Sentinel 2 alcanzaron un RMSE de 10,4 ton/ha . Otros autores han reportado mejores parámetros de ajuste que los encontrados en este estudio Jamnani et al., (2019) utilizando redes neuronales y el índices de vegetación GVI calculado con Landsat 8 consiguieron un RMSE de 6,1 ton/ha cuando estimaron el rendimiento de campo y Rahman y Robson, (2016) obtuvieron un RMSE de 6,2 ton/ha utilizado el índice de vegetación GNDVI calculado con imágenes Landsat 8 y modelos cuadráticos.

- **Modelización de la variable rendimiento de industrial (REND)**

El análisis de correlación del apartado anterior había demostrado que los índices de vegetación no tenían relación con la variable rendimiento industrial (REND). Sin embargo, en este apartado se realizó nuevamente la evaluación de todas las covariables utilizando la selección *bidirection* de la función *stepwise()*; los resultados en la estimación del REND fueron deficientes, lo cual confirma que no existe una relación directa entre la concentración de azúcares y la fracción fotosintéticamente activa de la vegetación (Shukla et al., 2017).

Las covariables que conformaron el modelo más parsimonioso para la estimación del rendimiento industrial fueron: **GNDVI Agosto**, **REND promedio de años anteriores** y **REND máximo de años anteriores**. Es importante resaltar que en ambos meses (Agosto y Noviembre) el método de selección determinó las mismas covariables. El modelo de regresión integrado de dichas covariables es el siguiente:

$$\begin{aligned} REND\ Final &= 78,3532 + (76,0547 \times GNDVI\ Agosto) \\ &+ (-1,9185 \times REND\ promedio\ de\ años\ anteriores) \\ &+ (1,7813 \times REND\ máximo\ de\ años\ anteriores) \end{aligned}$$

Donde REND Final es la estimación del rendimiento industrial, GNDVI Agosto, es el índices de vegetación GNDVI calculado para el mes agosto, REND promedio de años anteriores es el promedio del rendimiento industrial según el historial de producción y REND máximo de años anteriores es el rendimiento industrial máximo obtenidos en años anteriores.

Los parámetros del modelo REND Final tienen coeficientes con una significancia estadística $< 0,10$, el error estándar residual es de 9,63 kg de azúcar/ton, un coeficiente de determinación ajustado de 0,2 y un p-value del modelo de 0,003333. Los parámetros anteriores permiten indicar que este modelo es deficiente para realizar la estimación deseada. Lo anterior se confirmó con la validación cruzada de la cual se obtuvo un RMSE de 10,81 kg de azúcar/ton, un coeficiente de correlación (R^2) de 0,09 y promedio absoluto del error (MAE) de 6,83 kg de azúcar/ton.

Los resultados anteriores reafirman la poca capacidad predictiva que tienen los índices de vegetación para estimar el rendimiento industrial en caña de azúcar, razón por la cual no se calculó la cantidad de toneladas de azúcar totales (TON AZU). Quizá la incorporación de variables climáticas, variedad, características de suelos, entre otras permitan mejorar la estimación de esta variable.

Los análisis realizados en el apartado de estimadores de variables de rendimiento permiten indicar que los índices de vegetación EVI y SAVI resultaron ser más adecuados para estimar el rendimiento de campo, además sus aportes en la predicción fueron constantes en ambos modelos, lo cual permite enfocar los esfuerzos en el uso de índices de vegetación que incluyen dentro de los algoritmos correcciones para disminuir la saturación por alta densidad de la vegetación.

Las estimaciones de PROD Agosto y PROD Final fueron satisfactorias, permitiendo el uso de forma integral y complementaria, esto permite que el proceso de estimación de variables de rendimiento también sea una herramienta de uso potencial para el monitoreo de la plantaciones de caña de azúcar.

Por otra parte, el modelo de estimación REND Final confirma la poca relación entre el rendimiento industrial y las respuestas espectrales obtenidas con imágenes satelitales. Sin embargo, importante probar nuevos métodos de estimación utilizando covariables adicionales que tengan mayor relación con la concentración de azúcares en la planta, a saber: clima, suelos, variedad, entre otras.

4.4. SMS-CoopeVictoria R.L.: Sistema de Monitoreo Satelital de las Plantaciones de Caña de Azúcar de CoopeVictoria R.L.

Los resultados de este apartado responden técnicamente a una necesidad de transferir información geoespacial a productores y tomadores de decisiones, estos se presentarán de una forma descriptiva caracterizando las funcionalidades de la herramienta.

La herramienta puesta a disposición de los tomadores de decisiones de la Cooperativa se puede acceder mediante un link de internet, en el cual cuando los usuarios accedan se encontrarán con la pantalla principal que se presenta en la Figura 18, la cual está compuesta por un encabezado que contiene tres opciones Principal, SMS-CoopeVictoria R.L. y Contactos, al seleccionar cada uno de ellos se trasladará a nuevas ventanas con información y herramientas distintas.

Principal SMS-CoopeVictoria R.L. Contactos

Sistema de Monitoreo Satelital de las Plantaciones de Caña de Azúcar en CoopeVictoria R.L.

SMS-CoopeVictoria R.L.

Sistema de Monitoreo Satelital de las Plantaciones de Caña de Azúcar en CoopeVictoria R.L. pretende apoyar la gestión de las áreas de cultivo mediante información satelital. Esta iniciativa desarrolla de forma colaborativa entre el Centro de Investigaciones Agronómicas (CIA) de la Universidad de Costa Rica (UCR) y CoopeVictoria R.L. surge por la necesidad de dotar de herramientas de agricultura de precisión al sector cooperativo costarricense.

La pestaña SMS-CoopeVictoria R.L. cuenta con información histórica y actual sobre las plantaciones de caña de azúcar administradas por la Cooperativa, donde usted encontrará la siguiente información:

- Áreas de Producción.
- Distritos de Influencia.
- Historial de Cosecha por Finca.
- NDVI Calculado con el Sensor Sentinel 2 .

Características de la información suministrada.

El NDVI es un índice de vegetación que está fuertemente relacionado con la porción fotosintéticamente activa de la vegetación. En caña de azúcar los valores rondan entre 0,25 (cuando inician los rebrotes) y 0,8 (previo al inicio de estrés de la planta por maduración). Esos valores varían según: manejo, variedad, condiciones atmosféricas, entre otras.

Los valores que encontrará en la pestaña SMS-CoopeVictoria R.L. deberán ser divididos entre 100 para reescalarlos de 0 - 1. Variaciones temporales abruptas en el índice pueden estar relacionadas factores atmosféricos (nubes o sombras de nubes).

El estimado de rendimiento fue calculados con base a regresiones lineales múltiples que buscan explicar la variación del rendimiento de campo en función al a índices de vegetación y variables históricas del cultivo.

UNIVERSIDAD DE COSTA RICA CIA Centro de Investigaciones Agronómicas COOPE VICTORIA

Figura 18. Ventana Principal al acceder al geoportal.

Posteriormente se presentan el título y dos subtítulos con información relevante, la información de la izquierda (SMS-CoopeVictoria R.L.) hace una descripción del interés y los aliados de esta página, además se enumeran los productos disponibles en la ventana SMS-CoopeVictoria R.L. En la columna de la izquierda (Características de la Información Suministrada) se facilitan conceptos técnicos en un lenguaje comprensible y características de los insumos que verá en la ventana

SMS-CoopeVictoria R.L. y para finalizar se insertaron los logos de las instituciones participantes (Figura 18).

La siguiente ventana, SMS-CoopeVictoria R.L., presentan como elementos centrales el encabezado (nótese que ahora está de color verde la ventana seleccionada), además está el título y como elemento nuevo y central en este trabajo aparece la ventana de un visor cartográfico, el cual está compuesto por una herramienta de búsqueda en la parte superior (Seleccione una Finca), la escala, el controlador de “zoom” (en la parte superior izquierda), el control de capas (en la parte superior izquierda) y la leyenda del NDVI (en rango de valores de 20 a más de 90, se excluyen los valores menores a 20 por considerarlos suelos, distorsiones atmosféricas u otras anomalías) (Figura 19).

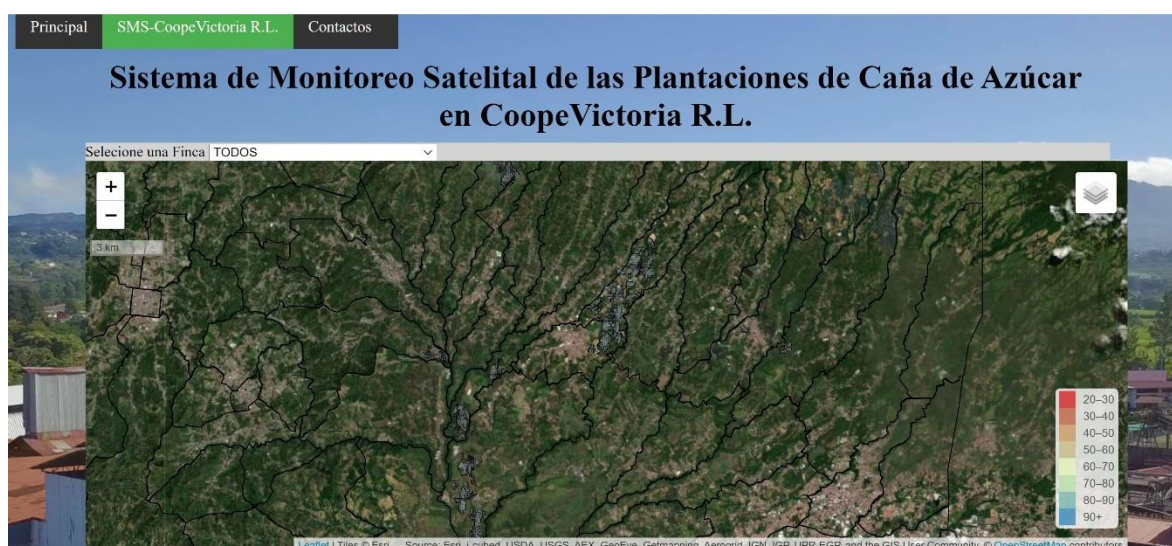


Figura 19. Ventana SMS-CoopeVictoria R.L. donde se muestran los detalles de las capas de información compartidas a los usuarios.

La herramienta de búsqueda está programada para que el usuario seleccione una finca de interés y un algoritmo de selección lo acerca a la posición. La Figura 20 muestra el resultado de la ejecución de selección de finca, además del acercamiento y selección, se programó un etiquetado a través de un popupText.



Figura 20. Uso de la herramienta selección de finca. Acercamiento sobre la finca ASSMAN y etiquetado mediante popupText.

La opción de control de capas permite la visualización de las capas e información disponibles, la Figura 21 presenta las capas disponibles, éstas se activarán y desactivarán según sea la necesidad de usuario. Además, cada capa activa permitirá la visualización de las características mediante un popupText.



Figura 21. Panel de control de capas, donde se puede activar y desactivas las capas según sea el interés. La línea tenue de color negro representa el límite de la finca.

La Figura 22 muestra la selección de un mismo lote, pero utilizando dos capas distintas del panel de capas, por un lado, se muestra el histórico de las zafras y en

el otro las características de cada finca, estos detalles son relevantes para conocer rápidamente las características de las distintas fincas.



Figura 22. Disposición de los popupText según sea la capa que de esta utilizando en el panel de capas

En esta ventana también se incluyó el índice de vegetación NDVI para una fecha en particular (22/11/2019), el cual podría ser actualizado fácilmente y puesto a disposición de los usuarios; este índice de vegetación pone en evidencia la variabilidad espacial que sucede al interno de los lotes, razón por la cual esta herramienta permitiría detectar anomalías a lo interno de las plantaciones (Figura 23).



Figura 23. Variabilidad espacial de NDVI del 22/11/2019. Se puede notar la variación en los valores NDVI, cada cuadro posee un área de 100 m².

Finalmente, la ventana Contactos presenta información referente a los resultados de esta investigación que estuvo adscrita proyecto de investigación de la

Universidad de Costa Rica, además se incorporan los contactos de las personas participantes en el desarrollo del proyecto (Figura 24).



Figura 24. Ventana Contactos donde se muestra la información relacionada al proyecto de investigación y contactos de las personas participantes.

Los resultados presentados en este apartado confirman que las metas fueron cumplidas, se logró diseñar un geoportal que cumple las características para transferir información relevante, de las plantaciones de caña de azúcar, a los tomadores de decisiones de la Cooperativa.

Las herramientas informáticas utilizadas para el diseño y ejecución de este portal web son en su totalidad de acceso libre y open source, lo cual permite reafirmar que actualmente existen los medios disponibles para desarrollar una agricultura que cumpla los principios fundamentales de la agricultura de precisión; en este sentido las herramientas geoespaciales aportan información fundamental para el cumplimiento de estos principios básicos.

Capítulo 5: Conclusiones y Recomendaciones

5.1. Conclusiones

El abordaje metodológico planteado para el desarrollo de esta investigación permitió el cumplimiento de las metas planteadas en cada uno de los objetivos, ya que se logró la generación de herramientas e insumos que pueden colaborar en la toma de decisiones de la Cooperativa con respecto al manejo de las plantaciones de caña de azúcar.

Las plantaciones agrícolas del contexto costarricense poseen características de extensión que limitan el uso de plataformas satelitales de baja resolución espacial, porque la homogenización de píxeles impide la identificación de variabilidad espacial en unidades de producción que muchas veces son menores que el tamaño de un píxel de un sensor como MODIS e incluso Landsat.

De igual manera los índices de vegetación analizados presentaron desempeños variados, algunos de ellos tuvieron fuertes correlaciones en momentos particulares, sin embargo, el comportamiento no fue constante, lo cual obligó a utilizar diversas técnicas de agrupación para intentar conseguir información representativa para el periodo de estudio.

Los diversos métodos de agrupación y las limitaciones por causas atmosféricas permitieron definir que agosto y noviembre representaban las mejores fechas para la estimación de las variables de rendimiento, sin embargo, queda presente la incógnita de las variaciones y manifestaciones que tiene el cultivo durante el periodo de septiembre y octubre.

Los procesos de modelización permitieron definir un reordenamiento de las covariables explicativas de los modelos de estimación, además queda demostrado que correlaciones lineales fuertes no implican buenos ajustes en regresiones múltiples. Las covariables asociadas al historial de producción resultaron fundamentales para la construcción de los modelos de estimación, ya que permiten caracterizar el potencial productivo del cultivo.

Se confirma que las variables de producción vinculadas con la calidad del producto (rendimiento industrial) no presentan relación con los índices de vegetación del cultivo, ya que tanto los coeficientes de correlación como las regresiones lineales múltiples indicaron muy bajas relaciones. Caso contrario sucede con variables de producción vinculadas directamente con la porción fotosintéticamente activa de cultivo, específicamente el rendimiento de campo, el cual presentó buenos coeficientes de correlación y ajustes en las regresiones lineales.

El ajuste en la estimación del rendimiento de campo y el posterior cálculo del total de tonelada permite validar el uso de imágenes satelitales para la gestión de plantaciones agrícolas, específicamente la caña de azúcar. Los resultados obtenidos son en muchos casos más ajustados a los presentados por otros autores en otras partes del mundo.

Es posible y viable la transferencia de la información a los productores y tomadores de decisiones que gestionan las plantaciones de caña de azúcar. La herramienta construida en este trabajo brinda un marco general de uso y aprovechamiento, que puede guiar las actividades de los agricultores.

Las técnicas de programación y las herramientas open source facilitan y ayudan en el procesamiento de amplios volúmenes de información, así se confirmó en esta investigación en la cual, más del 80% de los recursos utilizados son de acceso libre y en la cual hay una comunidad de desarrolladores-colaboradores que facilitan el aprendizaje de las técnicas.

Finalmente, los SIG y la Teledetección tienen un alto potencial para la agricultura costarricense a través de herramientas que permitan la toma de decisiones informadas y apoyadas en conceptos fundamentales de la agricultura de precisión como lo son, el dónde y el cuándo realizar acciones de manejo en las plantaciones agrícolas.

5.2. Recomendaciones

Al finalizar las actividades planteadas en esta investigación se pueden definir algunas recomendaciones para futuras investigaciones o el seguimiento y actualización de estos resultados por parte de la Cooperativa, por tal razón las recomendaciones serán divididas en dos grupos.

A la Cooperativa

En primera instancia se recomienda mantener constante y en aumento el interés en incorporar el uso de las tecnologías geoespaciales en el manejo de las plantaciones agrícolas, tanto en las plantaciones de caña de azúcar como en las de café. Por tal razón, se insta a dar seguimiento a estos resultados y validarlos en la zafra 2020-2021.

Valorar la posibilidad de utilizar unidades de manejo más específicas en las fincas propias y arrendadas, por ejemplo, lotes en el interior de las fincas, ya que es probable que información más específica pueda brindar resultados más específicos. Además, incorporar otras variables como número de socas, variedades y fechas de corta o de siembra.

Asimismo, se recomienda incorporar las propiedades de los productores asociados a la Cooperativa (información que fue construida en el proyecto desarrollado de forma colaborativa entre CoopeVictoria R.L., la Universidad de Costa Rica y el Instituto Nacional de Fomento Cooperativo), la incorporación de dicha información vendría a aumentar el número de beneficiados y mejorar el desempeño de los modelos.

A futuras investigaciones

El monitoreo de las plantaciones agrícolas mediante información satelital en Costa Rica es una necesidad, por tal razón se recomienda el desarrollo de investigaciones en esta línea que puedan ayudar a mejorar la competitividad de los pequeños y medianos productores.

Las condiciones atmosféricas de la región dificultan el uso de imágenes satelitales ópticas, por tal razón se debe abordar de forma más intensiva el uso de información radar en la evaluación de plantaciones agrícolas y ecosistemas naturales.

En la actualidad el uso de herramientas open source y de acceso libre permiten el desarrollo de aplicaciones de calidad y de bajo costo, por lo cual se insta a explotar estos recursos disponibles donde colaboradores-desarrolladores están facilitando nuevo conocimiento constantemente.

Bibliografía

- Abdel-Rahman, E. M., y Ahmed, F. B. (2008). The application of remote sensing techniques to sugarcane (*Saccharum spp. hybrid*) production: A review of the literature. *International Journal of Remote Sensing*, 29(13), 3753–3767. <https://doi.org/10.1080/01431160701874603>
- Agafonkin, V. (2019). *Leaflet*. <https://leafletjs.com/reference-1.3.4.html>
- Allison, J. C. S., Pammenter, N. W., y Haslam, R. J. (2007). Why does sugarcane (*Saccharum sp. hybrid*) grow slowly? *South African Journal of Botany*, 73(4), 546–551. <https://doi.org/10.1016/j.sajb.2007.04.065>
- Birth, G. S., y McVey, G. R. (1968). Measuring the Color of Growing Turf with a Reflectance Spectrophotometer 1. *Agronomy Journal*, 60(6), 640–643. <https://doi.org/10.2134/agronj1968.00021962006000060016x>
- Busetto, L., y Ranghetti, L. (2016). MODISStsp : A Tool for Automatic Preprocessing of. *Computers y Geosciences*, 97, 40–48. <https://github.com/ropensci/MODISStsp>
- Campbell, J. (2006). *Introduction to remote sensing* (G. Press (ed.); Fourth Edi).
- Chai, T., y Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7(3), 1247–1250. <https://doi.org/10.5194/gmd-7-1247-2014>
- Chavez, M., y Chavarría, E. (2012). ¿Cómo Se Distribuye Y Dónde Se Cultiva Territorialmente La Caña Destinada a La Fabricación De Azúcar En Costa Rica? (Número 506).
- Chuvieco, E. (2008). Teledetección ambiental. *International Journal of Remote Sensing*, 1, 594.
- Dorman, M. (2020). Introduction to Web Mapping. *Introduction to Web Mapping*. <https://doi.org/10.1201/9780429352874>
- Durán, J., y Oviedo, M. (2012). *Catalogo de variedades de caña de azúcar cultivadas en algunas regiones cañeras de Costa Rica*.
- Ferguson, R., y Rundquist, D. (2018). *Remote Sensing for Site-Specific Crop Management*. 103–117. <https://doi.org/10.2134/precisionagbasics.2016.0092>
- Gitelson, A. A., Kaufman, Y. J., y Merzlyak, M. N. (1996). Use of a green channel in remote sensing of global vegetation from EOS- MODIS. *Remote Sensing of Environment*, 58(3), 289–298. [https://doi.org/10.1016/S0034-4257\(96\)00072-7](https://doi.org/10.1016/S0034-4257(96)00072-7)
- GitHub, I. (2020). *Pricing Plans for all developers*. <https://github.com/pricing>
- Gröll, K., Graeff, S., y Claupein, W. (2007). Use of vegetation indices to detect plant diseases. *GIL Jahrestagung*, 340, 95–98.
- Gunnula, W., Manit, K. M., Manit, R., Mayura, W., y Prabban, P. (2011). Normalized difference vegetation index relationships with rainfall patterns and yield in small plantings of rain-fed sugarcane. *Australian Journal of Crop Science*, 13(5), 1845–1851. <https://doi.org/10.17235/reed.2016.4048/2015>

- Hatfield, J. L., y Prueger, J. H. (2010). *Value of Using Different Vegetative Indices to Quantify Agricultural Crop Characteristics at Different Growth Stages under Varying Management Practices*. 562–578. <https://doi.org/10.3390/rs2020562>
- Hogg, J. (2004). Quantitative remote sensing of land surfaces. En *The Photogrammetric Record* (Vol. 19, Número 108). https://doi.org/10.1111/j.0031-868x.2004.295_1.x
- Huete, A. (1988). A Soil-Adjusted Vegetation Index (SAVI). *Remote Sensing of Environment*, 25(1), 295–309. [https://doi.org/https://doi.org/10.1016/0034-4257\(88\)90106-X](https://doi.org/https://doi.org/10.1016/0034-4257(88)90106-X)
- Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., y Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83(12), 195–213. [https://doi.org/https://doi.org/10.1016/S0034-4257\(02\)00096-2](https://doi.org/https://doi.org/10.1016/S0034-4257(02)00096-2)
- INEC (Instituto Nacional de Estadística y Censos). (2020). *Encuesta Nacional Agropecuaria 2019: Resultados Generales de la Actividad Agrícola y Forestal*.
- Inman-Bamber, N. G. (1994). Temperature and seasonal effects on canopy development and light interception of sugarcane. *Field Crops Research*, 36(1), 41–51. [https://doi.org/10.1016/0378-4290\(94\)90051-5](https://doi.org/10.1016/0378-4290(94)90051-5)
- James, G., Witten, D., Trevor, H., y Tibshirani, R. (2013). *An Introduction to Statistical Learning - with Applications in R | Gareth James | Springer*. <https://www.springer.com/gp/book/9781461471370%0Ahttp://www.springer.com/us/book/9781461471370>
- Jamnani, M. R., Liaghat, A., y Loyeh, N. S. (2019). *Sugarcane yield prediction at farm scale using remote sensing and artificial neural network Sugarcane yield prediction at farm scale using remote sensing and artificial neural network*. June 2020, 0–10.
- Khorram, S., Koch, F. ., van der Wiele, C. ., y Nelson, S. A. . (2012). *Remote Sensing*.
- Li, J., Lu, X., Cheng, K., y Liu, W. (2020). Regression and Time Series Model Selection. *Regression and Time Series Model Selection, 1968*. <https://doi.org/10.1142/3573>
- Lillesand, T. M., y Kiefer, R. . (1987). *Remote sensing and image interpretation* (2ª ed.). John Wiley y Sons.
- Lisboa, I. P., Damian, M., Cherubin, M. R., Barros, P. P. S., Fiorio, P. R., Cerri, C. C., y Cerri, C. E. P. (2018). Prediction of sugarcane yield based on NDVI and concentration of leaf-tissue nutrients in fields managed with straw removal. *Agronomy*, 8(9). <https://doi.org/10.3390/agronomy8090196>
- Mahlein, A. K., Rumpf, T., Welke, P., Dehne, H. W., Plümer, L., Steiner, U., y Oerke, E. C. (2013). Development of spectral indices for detecting and identifying plant diseases. *Remote Sensing of Environment*, 128, 21–30. <https://doi.org/10.1016/j.rse.2012.09.019>
- Main-Knorn, M., Pflug, B., Louis, J., Debaecker, V., Müller-Wilm, U., y Gascon, F. (2017). *Sen2Cor for Sentinel-2*. October 2018, 3. <https://doi.org/10.1117/12.2278218>
- Martínez-Martínez, V., Gomez-Gil, J., Machado, M. L., y Pinto, F. A. C. (2018). Leaf and canopy reflectance spectrometry applied to the estimation of angular leaf spot

- disease severity of common bean crops. *PLoS ONE*, 13(4), 1–18.
<https://doi.org/10.1371/journal.pone.0196072>
- Max, A., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., Ziem, A., Scrucca, L., Hunt, T., y Kuhn, M. M. (2020). *Package ‘ caret ’ R* (p. 223).
- Mayer, B., y Kylling, A. (2005). The libRadtran software package for radiative transfer calculations - description and examples of use. *Atmospheric Chemistry and Physics*, 5(7), 1855–1877. <https://doi.org/10.5194/acp-5-1855-2005>
- Molijn, R. A., Iannini, L., Rocha, J. V., y Hanssen, R. F. (2019). Sugarcane productivity mapping through C-band and L-band SAR and optical satellite imagery. *Remote Sensing*, 11(9), 1–27. <https://doi.org/10.3390/rs11091109>
- Morel, J., Todoroff, P., Bégué, A., Bury, A., Martiné, J. F., y Petit, M. (2014). Toward a satellite-based system of sugarcane yield estimation and forecasting in smallholder farming conditions: A case study on reunion island. *Remote Sensing*, 6(7), 6620–6635. <https://doi.org/10.3390/rs6076620>
- Moreno, V., Alonso, C., y Rodríguez, E. (1999). Determinación experimental de la firma espectral de la vegetación. Una sencilla práctica de introducción a la teledetección. *Teledetección*, 429–432.
- Mudholkar, M., y Mudholkar, P. (2017). A Beginner ' s Guide to Git and GitHub A Beginner ' s Guide to Git and GitHub Department of Master of Computer Application. *International Journal for Scientific Research y Development*, January 2017, 1403–1406. <https://doi.org/10.13140/RG.2.2.20126.89927>
- Murillo, P. J., y Carbonell, J. (2012). *Principios y aplicaciones de la percepción remota en el cultivo de la caña de azúcar en Colombia* (1ª ed.).
- Mutanga, S., Schoor, C. van, Olorunju, P. L., Gonah, T., y Ramoelo, A. (2013). Determining the Best Optimum Time for Predicting Sugarcane Yield Using Hyper-Temporal Satellite Imagery. *Advances in Remote Sensing*, 02(03), 269–275. <https://doi.org/10.4236/ars.2013.23029>
- Oyala, V. (2012). *Sistemas de Información Geográfica* (1ª ed.).
http://wiki.osgeo.org/wiki/Libro_SIG
- Papadopoulos, A., Papadopoulos, F., Tziachris, P., Metaxa, I., y Iatrou, M. (2016). *Site specific management with the use of a digitized soil map for the regional unit of Kastoria Site Specific Management with the Use of a Digitized Soil Map for the Regional Unit of Kastoria*. February 2014.
- Paul, C. K., y Mascarenhas, A. C. (1981). Remote sensing in development. *Science*, 214(4517), 139–145. <https://doi.org/10.1126/science.214.4517.139>
- Petropoulos, G., y Kalaitzidis, C. (2012). Multispectral vegetation index in remote sensing: An overview. En W. ZHANG (Ed.), *Ecological Modeling* (1ª ed., p. 405).
- Procisur, y IICA. (2015). Manual de Agricultura de Precisión. En E. Chartuni y C. Magdalena (Eds.), *Agromensajes - Facultad de Ciencias Agrarias . Universidad Nacional del Rosario* (Vol. 41, Número 1).
- Puletti, N., Chianucci, F., y Castaldi, C. (2018). Use of Sentinel-2 for forest classification in

- Mediterranean environments. *Annals of Silvicultural Research*, 42(1), 32–38.
<https://doi.org/10.12899/ASR-1463>
- Rahimi-Jamnani, M., Liaghat, A., y Mirzaei, F. (2019). Optimization of sugarcane harvest using remote sensing. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 42(4/W18), 857–861.
<https://doi.org/10.5194/isprs-archives-XLII-4-W18-857-2019>
- Rahman, M. M., y Robson, A. (2020). *Integrating Landsat-8 and Sentinel-2 Time Series Data for Yield Prediction of Sugarcane Crops at the Block Level*. 33, 1–15.
- Rahman, M. M., y Robson, A. (2016). Multi-temporal remote sensing for yield prediction in sugarcane crops. *19th Precision Agriculture Symposium in Australasia, September 2016*. http://sydney.edu.au/agriculture/pal/documents/2016/PA_Symposium_16.pdf
- Rebekić, A., Lončarić, Z., Petrović, S., y Marić, S. (2015). Pearson'S or Spearman'S Correlation Coefficient - Which One To Use? *Poljoprivreda*, 21(2), 47–54.
<https://doi.org/10.18047/poljo.21.2.8>
- Richter, R., Louis, J., Berthelot, B., Louis, J., y Berthelot, B. (2012). Sentinel-2 MSI – Level 2A Products Algorithm Theoretical Basis Document. *European Space Agency, (Special Publication) ESA SP, 49(0)*, 1–72.
https://earth.esa.int/c/document_library/get_file?folderId=349490&name=DLFE-4518.pdf
- Rouse, J. W., Hass, R. H., Schell, J. A., y Deering, D. W. (1973). Monitoring vegetation systems in the great plains with ERTS. *Third Earth Resources Technology Satellite (ERTS) symposium*, 1, 309–317.
<https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19740022614.pdf>
- Saez, J. V. (2017). *Dinámica de acumulación de sacarosa en tallos de caña de azúcar (Saccharum spp .) modulada por cambios en la relación fuente-destino*.
- Sandinska, Y. (2016). Technological Principles and Mapping Applications of Web Gis. *6Th International Conference on Cartography and Gis, Vols 1 and 2, June*, 287–296.
- Santoso, H., Tani, H., y Wang, X. (2016). A simple method for detection and counting of oil palm trees using high-resolution multispectral satellite imagery. *International Journal of Remote Sensing*, 37(21), 5122–5134.
<https://doi.org/10.1080/01431161.2016.1226527>
- Shukla, S., Felderhoff, T. J., Saballos, A., y Vermerris, W. (2017). The relationship between plant height and sugar accumulation in the stems of sweet sorghum (*Sorghum bicolor* (L.) Moench). *Field Crops Research*, 203, 181–191.
<https://doi.org/10.1016/j.fcr.2016.12.004>
- Sikaria, H. (2018). *Package ‘ rLandsat ’*.
- Solis, G., y Martínez.R. (2016). *Uso de espectroscopía de campo en la caracterización espectral de cultivos de palma africana “Elaeis guineensis” en Parrita, Puntarenas, Costa Rica*. Universidad de Costa Rica.
- Thenkabail, P., Lyon, J., y Huete, A. (2011). Advances in Hyperspectral Remote Sensing of Vegetation and Agricultural Croplands. En P. Thenkabail, J. Lyon, y A. Huete (Eds.), *Hyperspectral Remote Sensing of Vegetation* (p. 705).

<https://doi.org/10.1201/b11222-3>

- Tutorialspoints. (2015). *LeafletJs* (p. 2). <https://doi.org/10.1017/CBO9781107415324.004>
- Veenendaal, B., Brovelli, M. A., y Li, S. (2017). Review of web mapping: Eras, trends and directions. En *ISPRS International Journal of Geo-Information* (Vol. 6, Número 10). <https://doi.org/10.3390/ijgi6100317>
- W3.CSS. (2019a). *CSS Tutorial*. <https://www.w3schools.com/css/>
- W3.CSS. (2019b). *HTML Tutorial*. <https://www.w3schools.com/html/default.asp>
- Wang, W., y Lu, Y. (2018). Analysis of the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) in Assessing Rounding Model. *IOP Conference Series: Materials Science and Engineering*, 324(1). <https://doi.org/10.1088/1757-899X/324/1/012049>
- Wiegand, C. L., Richardson, A. J., Escobar, D. E., y Gerbermann, A. H. (1991). Vegetation indices in crop assessments. *Remote Sensing of Environment*, 35(2–3), 105–119. [https://doi.org/10.1016/0034-4257\(91\)90004-P](https://doi.org/10.1016/0034-4257(91)90004-P)
- Xue, J., y Su, B. (2017). Significant remote sensing vegetation indices: A review of developments and applications. *Journal of Sensors*, 2017. <https://doi.org/10.1155/2017/1353691>
- Zhang, Z., Werner, F., Cho, H. M., Wind, G., Platnick, S., Ackerman, A. S., Di Girolamo, L., Marshak, A., y Meyer, K. (2017). A framework for quantifying the impacts of sub-pixel reflectance variance and covariance on cloud optical thickness and effective radius retrievals based on the bi-spectral method. *AIP Conference Proceedings*. <https://doi.org/10.1063/1.4975502>

Anexos

Anexo 1: Código de programación en RStudio del paquete MODISrsp v1.3.9

```
##ESTE SCRIPT PERMITE DESCARGAR PRODUCTO MODIS POR SERIES DE  
TIEMPO
```

```
##PREPARADO POR BRYAN ALEMÁN Montes
```

```
##SE DEBE DESCARGAR E INSTALAR EL PAQUETE
```

```
install.packages("MODISrsp")
```

```
library(MODISrsp)
```

```
MODISrsp()
```

```
##Category: debe incluir el tipo de información que se desea descargar
```

```
##Product: debe seleccionar la resolución espacial y temporal del producto
```

```
##Plataform: Terra
```

```
##Change Selection: seleccionar la información que desea descargar (Original  
Modis, Quality Indicators, Spectral Indexes)
```

```
##Download Method: Download Server (http), User Name y Password (debe tener  
una cuenta en: https://urs.earthdata.nasa.gov/)
```

```
##Starting and Ending data: definir la temporalidad de los datos de interés.
```

```
##Select MODIS tiles: seleccionar los tiles de MODIS para el área de interés (en  
este caso se usó Horizontal (Start 9- End 9) y Vertical (Start 7 - End 8))
```

```
##Output Projection: User Define (4326)
```

```
##Output format: formato de salida.
```

```
##Save Time Series as: R rasterStack
```

```
##Definir las carpetas donde se guardará la información
```


Anexo 2: Código de programación en RStudio del paquete rLandsat

```
## Curso Teledeteccion y PDI con R + RSTUDIO + QGIS
```

```
## Autor Nino Bravo Morales
```

```
## Descarga de imagenes landsat 8 con PATH y ROW
```

```
## Instalar Paquete
```

```
#install.packages("devtools")
```

```
#install.packages("rLandsat")
```

```
#install.packages("sp")
```

```
#install.packages("raster")
```

```
devtools::install_github("socialcopsdev/rLandsat", force = TRUE)
```

```
## Cargar las librerias necesarias
```

```
library(sp)
```

```
library(raster)
```

```
library(rLandsat)
```

```
## Identificar el directorio de trabajo
```

```
getwd()
```

```
setwd("E:/Maestria/I_2019/Proyecto_II/Descarga_LANDSAT_2016_2018/Imagene  
s")
```

```
# 1 Identificar el producto Landsat por PATH y ROW
```

```
resultado_Path = landsat_search(min_date = "2017-01-01", max_date = "2017-12-  
31", path_master = (15), row_master = (53))
```

```
# 2 Introducir su Usuario y contraseña de Earth Explorer
```

```
espa_creds("bryan.aleman", "xxxxxxxxxx")
```

```
## 3 Seleccionar imagenes menor porcentaje de nubes
```

```
#producto_100nube = subset(resultado_Path, resultado_Path$cloud_coverage <
100)
```

```
producto_100nube = resultado_Path
```

```
## 4 Obtener los productos disponibilidad Imagen
```

```
producto = espa_products(producto_100nube$product_id)
```

```
producto = producto$master
```

```
## 5 Ordenar pedido por imagenes en su correo
```

```
result_orden_imagen = espa_order(producto_100nube$product_id, product =
c("toa","stats"),
```

```
file_format = "gtiff",
```

```
projection = "lonlat",
```

```
host = 'https://espa.cr.usgs.gov/api/v1/',
```

```
order_note = "Zona de estudio en menor 10 porciento nube")
```

```
#, "sr", "sr_ndvi", "sr_evi", "sr_savi", "sr_nbr", "bt"
```

```
result_orden_imagen$response
```

```
orden_imagen = result_orden_imagen$order_details$orderid
```

```
## 6 Obtener los ordenes y descarga imagen
```

```
des_url = espa_status(order_id = orden_imagen, getSize = TRUE)
```

```
dowlo_imagen = des_url$order_details
```

```
## 7 Descargar Landsat imagen en una carpeta
```

```
landsat_download(download_url = dowlo_imagen$product_dload_url, dest_file =
"E:/Maestria/I_2019/Proyecto_II/Descarga_LANDSAT_2016_2018/Imagenes")
```

Anexo 3: **Código de programación para el cálculo de los índices de vegetación (EVI, GNDVI, NDVI, SAVI y SR) a partir de las imágenes Landsat 8.**

```
# Cálculo de Índices de Vegetación Landsat
```

```
# Desarrollado por Bryan Alemán Montes
```

```
rm(list = ls())
```

```
library(raster)
```

```
library(dplyr)
```

```
library(rgdal)
```

```
library(gdalUtils)
```

```
library(rgeos)
```

```
library(RColorBrewer)
```

```
library(stringr)
```

```
#####
```

```
memory.size(max = TRUE)
```

```
## Definir la carpeta donde están las bandas de LandSat
```

```
setwd ("C:/LandSat_filtro")
```

```
dirBase <- "C:/LandSat_filtro"
```

```
#####
```

```
dirs <- list.dirs()
```

```
dirs <- c(grep(pattern = "R30m", dirs, value = TRUE))
```

```
for (i in 1:length(dirs)){
```

```
  #print(dirs[i])
```

```
  dirs_temp <- substr(dirs[i],2,nchar(dirs[i]))
```

```
  dirs_temp <- paste( dirBase,dirs_temp,sep="")
```

```

print(dirs_temp)
setwd(dirs_temp)

#####

  ## Seleccionar las bandas de LandSat
  Bandas_LandSat <- list.files(pattern = "*.tif$")
  ## Construir el raster_stack con las bandas seleccionadas anteriormente
  raster_stack <- raster::stack(paste0(Bandas_LandSat))

  ## Plotear las bandas en RGB para confirma el correcto orden.
  #plotRGB(raster_stack, r = 3, g = 2, b = 1, scale=800, stretch = "hist")

  ## Definir SRC en al raster
  crs(raster_stack) <- CRS('+init=EPSG:32616')

  ## Definir el nombre para los archivos guardados
  producto <- "RS"
  fecha <- str_match(Bandas_LandSat, "3_(.*?)_2")
  fecha <- fecha[,2]
  fecha <- substr(fecha[1],1,8)
  fnames <- paste( producto,fecha,sep="_")

#####

  ## Convertir el ráster a rasterbrick para acelerar los procesos
  imagen <- brick(raster_stack)
  ## Confirmar las características del rasterbrick
  #imagen[[4]]

```

```

b1<-imagen[[2]]/10000
b3<-imagen[[4]]/10000
b4<-imagen[[5]]/10000

#####

# Calcular el EVI: blue (band 2), red (band 4) y nir (band 5)
evi <-2.5*((b4-b3)/(b4+6*b3-7.5*b1+1))

## Plot el evi
#plot(evi, main="EVI")

## Guardar el ráster como evi
producto <- "EVI"
fnames <- paste( producto,fecha,sep="_")

writeRaster(evi,file.path('E:/Indices_Landsat/',filename = fnames), format="GTiff",
overwrite=TRUE)

#####

# Calcular el gndvi: green (band 3)y nir (band 5)
gndvi <- ((imagen[[5]]-imagen[[3]])/(imagen[[5]]+imagen[[3]]))

## Plot el gndvi
#plot(gndvi, main="GNDVI")

## Guardar el ráster como gndvi
producto <- "GNDVI"
fnames <- paste( producto,fecha,sep="_")

writeRaster(gndvi,file.path('E:/Indices_Landsat/',filename = fnames), format="GTiff",
overwrite=TRUE)

```

```
#####
```

```
# Calcular el NDVI: red (band 4) y nir (band 5)
ndvi <- (imagen[[5]] - imagen[[4]]) / (imagen[[5]] + imagen[[4]])
##Plot el ndvi
#plot(ndvi, main="NDVI")
```

```
## Guardar el ráster como ndvi
producto <- "NDVI"
fnames <- paste( producto,fecha,sep="_ ")
```

```
writeRaster(ndvi,file.path('E:/Indices_Landsat/',filename = fnames), format="GTiff",
overwrite=TRUE)
```

```
#####
```

```
# Calcular el SAVI: red (band 4) y nir (band 5)
savi <- (((imagen[[5]]/10000) - (imagen[[4]]/10000)) / (((imagen[[5]]/10000) +
(imagen[[4]]/10000))+0.5))*(1+0.5)
## Plot el savi
#plot(savi, main="SAVI")
```

```
## Guardar el ráster como savi
producto <- "SAVI"
fnames <- paste( producto,fecha,sep="_ ")
```

```
writeRaster(savi,file.path('E:/Indices_Landsat/',filename = fnames), format="GTiff",
overwrite=TRUE)
```

```
#####
```

```
# Calcular el SR: red (band 4) y nir (band 5)
sr <- imagen[[5]] / imagen[[4]]
## Plot el sr
```

```

#plot(sr, main="SR")

## Guardar el ráster como sr
producto <- "SR"
fnames <- paste( producto,fecha,sep="_")

writeRaster(sr,file.path('E:/Indices_Landsat/',filename = fnames), format="GTiff",
overwrite=TRUE)

rm(sr,savi,raster_stack,ndvi,imagen,gndvi,evi)
}

```

Anexo 4. Código de programación para el cálculo de los índices de vegetación (EVI, GNDVI, NDVI, SAVI y SR) a partir de las imágenes Sentinel 2.

```

# Calculo de índices de vegetación para imágenes Sentinel 2
# Desarrollado por Bryan Alemán Montes

rm(list = ls())

library(raster)
library(dplyr)
library(rgdal)
library(gdalUtils)
library(rgeos)
library(RColorBrewer)
library(stringr)

#####

memory.size(max = TRUE)

## Definir la carpeta donde están las bandas corregidas con sen2cor
setwd ("C:/Descarga_SENTINEL_2016_2018_sen2cor")
dirBase <- "C:/Descarga_SENTINEL_2016_2018_sen2cor"

```

```
#####
dirs <- list.dirs()
dirs <- c(grep(pattern = "R10m", dirs, value = TRUE))
for (i in 1:length(dirs)){
  #print(dirs[i])
  dirs_temp <- substr(dirs[i],2,nchar(dirs[i]))
  dirs_temp <- paste( dirBase,dirs_temp,sep="")
  print(dirs_temp)
  setwd(dirs_temp)
#####

  ## Hacer una selección de las bandas .jp2 para convertirlas en .tiff
  jp2_to_raster <- function(jp2_path, ref_crs) {
    # leer ".jp2" y convertirlo a raster
    sen2_GDAL <- readGDAL(jp2_path)
    sen2_GDAL %>%
    as.matrix() %>%
    t() %>%
    raster(crs=ref_crs,
           xmn=xmin(sen2_GDAL), xmx=xmax(sen2_GDAL),
           ymn=ymin(sen2_GDAL), ymx=ymax(sen2_GDAL)) -> tmp_raster
    # Obtener el número de las bandas de los archivos en el ruta
    strsplit(jp2_path, "_")[[1]][length(strsplit(jp2_path, "_")[[1]])] %>%
    strsplit(".jp2") %>%
    as.character() -> names(tmp_raster)

    return(tmp_raster)
  }

  ## Seleccionar las bandas B2, B3, B4, B8
```



```

SEN2_PATHS <- list.files(pattern = "B0",full.names = T)

## Construir el raster_stack con las bandas seleccionadas anteriormente
raster_stack <- raster::stack(lapply(SEN2_PATHS, jp2_to_raster, ref_crs =
crs("+init=epsg:32616")))

## Plotear las las bandas en RGB para confirma el correcto orden.
#plotRGB(raster_stack, r = 3, g = 2, b = 1, scale=800, stretch = "hist")

## Definir SRC en al raster
crs(raster_stack) <- CRS('+init=EPSG:32616')

## Definir el nombre para los archivos guardados
producto <- "RS"

fecha <- str_match(SEN2_PATHS, "GS_(.*?)T")
fecha <- fecha[,2]
fecha <- substr(fecha[1],1,8)

fnames <- paste( producto,fecha,sep="_ ")

## Guardar el ráster como raster_stack

#writeRaster(raster_stack,file.path('E:/Indices_Sentinel/',filename = fnames),
format="GTiff", overwrite=TRUE)

#####

## Convertir el ráster a rasterbrick para acelerar los procesos
imagen <- brick(raster_stack)

## Confirmar las características del rasterbrick
#imagen[[4]]

b1<-imagen[[1]]/10000
b3<-imagen[[3]]/10000
b4<-imagen[[4]]/10000

#####

# Calcular el EVI: blue (band 1), red (band 3) y nir (band 4)
evi <-2.5*((b4-b3)/(b4+6*b3-7.5*b1+1))

```

```

## Plot el evi
#plot(evi, main="EVI")
## Guardar el ráster como evi
producto <- "EVI"
fnames <- paste( producto,fecha,sep="_")
writeRaster(evi,file.path('E:/Indices_Sentinel/',filename = fnames), format="GTiff",
overwrite=TRUE)
#####

# Calcular el gndvi: green (band 2)y nir (band 4)
gndvi <- ((imagen[[4]]-imagen[[2]])/(imagen[[4]]+imagen[[2]]))
## Plot el gndvi
#plot(gndvi, main="GNDVI")
## Guardar el ráster como gndvi
producto <- "GNDVI"
fnames <- paste( producto,fecha,sep="_")
writeRaster(gndvi,file.path('E:/Indices_Sentinel/',filename = fnames), format="GTiff",
overwrite=TRUE)
#####

# Calcular el NDVI: red (band 3) y nir (band 4)
ndvi <- (imagen[[4]] - imagen[[3]]) / (imagen[[4]] + imagen[[3]])
##Plot el ndvi
#plot(ndvi, main="NDVI")
## Guardar el ráster como ndvi
producto <- "NDVI"
fnames <- paste( producto,fecha,sep="_")
writeRaster(ndvi,file.path('E:/Indices_Sentinel/',filename = fnames), format="GTiff",
overwrite=TRUE)

```

```
#####

# Calcular el SAVI: red (band 3) y nir (band 4)

savi <- (((imagen[[4]]/10000) - (imagen[[3]]/10000)) / (((imagen[[4]]/10000) +
(imagen[[3]]/10000))+0.5))*(1+0.5)

## Plot el savi
#plot(savi, main="SAVI")

## Guardar el ráster como savi

producto <- "SAVI"

fnames <- paste( producto,fecha,sep="_")

writeRaster(savi,file.path('E:/Indices_Sentinel/',filename = fnames), format="GTiff",
overwrite=TRUE)

#####

# Calcular el SR: red (band 3) y nir (band 4)

sr <- imagen[[4]] / imagen[[3]]

## Plot el sr
#plot(sr, main="SR")

## Guardar el ráster como sr

producto <- "SR"

fnames <- paste( producto,fecha,sep="_")

writeRaster(sr,file.path('E:/Indices_Sentinel/',filename = fnames), format="GTiff",
overwrite=TRUE)

rm(sr,savi,raster_stack,ndvi,imagen,gndvi,evi)

}
```

Anexo 5: Código de programación para la extracción del valor promedio de cada índice dentro de cada finca para los tres años de análisis.

```
## Extracción de las estadísticas para cada uno de los años según los límites de cada finca.

## Desarrollado por Bryan Alemán Montes

setwd("D:/Maestria/Tesis/Descarga_SENTINEL_2016_2018_sen2cor/Indices_Sentinel")

#install.packages("openxlsx")

library(raster)
```

```

library(openxlsx)
library(shapefiles)

## Extracción las imágenes Sentinel

#####

## 2016

## Lista de los índices del 2016
se_in_2016 <- list.files(path = './2016/', pattern="tif$")
se_in_2016 <- list.files(path = './2016/', pattern="tif$", full.names=TRUE)

se_in_2016_evi <- list.files(path = './2016/', pattern="^EVI", full.names=TRUE)
se_in_2016_gndvi <- list.files(path = './2016/', pattern="^GNDVI", full.names=TRUE)
se_in_2016_ndvi <- list.files(path = './2016/', pattern="^NDVI", full.names=TRUE)
se_in_2016_savi <- list.files(path = './2016/', pattern="^SAVI", full.names=TRUE)
se_in_2016_sr <- list.files(path = './2016/', pattern="^SR", full.names=TRUE)

## Raster Stack de los Índices
stack_se_in_2016 <- stack(se_in_2016)
stack_se_in_2016_evi <- stack(se_in_2016_evi)
stack_se_in_2016_gndvi <- stack(se_in_2016_gndvi)
stack_se_in_2016_ndvi <- stack(se_in_2016_ndvi)
stack_se_in_2016_savi <- stack(se_in_2016_savi)
stack_se_in_2016_sr <- stack(se_in_2016_sr)

## Polígono de límite
fincas <-
shapefile("D:/Maestria/Tesis/Shapes_CV/FINCAS_PROPIAS_ARRENDADAS_COOPEVICTORIA_UNIO
N_CORREGIDA_UTM_buffer.shp")

## Extracción de la MEDIA
mean_se_in_2016 <- extract(stack_se_in_2016, fincas, fun='mean', na.rm=TRUE, df=TRUE)

```

```

mean_se_in_2016f <- data.frame(ID=fincas$qc_id, mean_se_in_2016)
## Guardar como xlsx la MEDIA para cada finca
write.xlsx(mean_se_in_2016f, "./mean_se_in_2016.xlsx")
#####
## 2017
## Lista de los índices del 2017

se_in_2017 <- list.files(path = './2017/', pattern="tif$")
se_in_2017 <- list.files(path = './2017/', pattern="tif$", full.names=TRUE)

se_in_2017_evi <- list.files(path = './2017/', pattern="^EVI", full.names=TRUE)
se_in_2017_gndvi <- list.files(path = './2017/', pattern="^GNDVI", full.names=TRUE)
se_in_2017_ndvi <- list.files(path = './2017/', pattern="^NDVI", full.names=TRUE)
se_in_2017_savi <- list.files(path = './2017/', pattern="^SAVI", full.names=TRUE)
se_in_2017_sr <- list.files(path = './2017/', pattern="^SR", full.names=TRUE)

## Raster Stack de los Índices
stack_se_in_2017 <- stack(se_in_2017)
stack_se_in_2017_evi <- stack(se_in_2017_evi)
stack_se_in_2017_gndvi <- stack(se_in_2017_gndvi)
stack_se_in_2017_ndvi <- stack(se_in_2017_ndvi)
stack_se_in_2017_savi <- stack(se_in_2017_savi)
stack_se_in_2017_sr <- stack(se_in_2017_sr)

## Extracción de la MEDIA
mean_se_in_2017 <- extract(stack_se_in_2017, fincas, fun='mean', na.rm=TRUE, df=TRUE)
mean_se_in_2017f <- data.frame(ID=fincas$qc_id, mean_se_in_2017)
## Guardar como xlsx la MEDIA para cada finca

```

```

write.xlsx(mean_se_in_2017f, "./mean_se_in_2017.xlsx")

#####

## 2018

## Lista de los índices del 2018

se_in_2018 <- list.files(path = './2018/', pattern="tif$")
se_in_2018 <- list.files(path = './2018/', pattern="tif$", full.names=TRUE)

se_in_2018_evi <- list.files(path = './2018/', pattern="^EVI", full.names=TRUE)
se_in_2018_gndvi <- list.files(path = './2018/', pattern="^GNDVI", full.names=TRUE)
se_in_2018_ndvi <- list.files(path = './2018/', pattern="^NDVI", full.names=TRUE)
se_in_2018_savi <- list.files(path = './2018/', pattern="^SAVI", full.names=TRUE)
se_in_2018_sr <- list.files(path = './2018/', pattern="^SR", full.names=TRUE)

## Raster Stack de los Índices

stack_se_in_2018 <- stack(se_in_2018)
stack_se_in_2018_evi <- stack(se_in_2018_evi)
stack_se_in_2018_gndvi <- stack(se_in_2018_gndvi)
stack_se_in_2018_ndvi <- stack(se_in_2018_ndvi)
stack_se_in_2018_savi <- stack(se_in_2018_savi)
stack_se_in_2018_sr <- stack(se_in_2018_sr)

## Extracción de la MEDIA

mean_se_in_2018 <- extract(stack_se_in_2018, fincas, fun='mean', na.rm=TRUE, df=TRUE)
mean_se_in_2018f <- data.frame(ID=fincas$qc_id, mean_se_in_2018)

## Guardar como xlsx la MEDIA para cada finca

write.xlsx(mean_se_in_2018f, "./mean_se_in_2018.xlsx")

```

Anexo 6. Código de programación de .html del geoportal desarrollado en la investigación.

```

<!DOCTYPE html>

<html lang="es">

```

```

<meta name="viewport" content="width=device-width, initial-scale=1.0">
    <!-- Hoja de estilos CSS adicionales -->
    <link rel="stylesheet" type="text/css" href="style.css">
    <link rel="stylesheet" type="text/css" href="/style_barra.css">
</head>
<title>Plantaciones de caña de azúcar en CoopeVictoria R.L.</title>
<style>
    body {
        background-image: url("canal.jpg");
        background-size: 1269px 843px;
        background-repeat: no-repeat;
    }
</style>
    <meta charset="UTF-8">

</script>
</head>
<body>
<div class="row">
<div class="topnav">
    <a class="active" href="index.html">Principal</a>
    <a href="smscopevictoria.html">SMS-CoopeVictoria R.L.</a>
    <a href="contactos.html">Contactos</a>
</div>
<h1>Sistema de Monitoreo Satelital de las Plantaciones de Caña de Azúcar en CoopeVictoria R.L.
</h1>
<div class="row">
    <div class="column" style="background:rgba(255,255,255,.85)">
        <h2>SMS-CoopeVictoria R.L.</h2>

```

<p>Sistema de Monitoreo Satelital de las Plantaciones de Caña de Azúcar en CoopeVictoria R.L. pretende apoyar la gestión de las áreas de cultivo mediante información satelital. Esta iniciativa desarrolla de forma colaborativa entre el Centro de Investigaciones Agronómicas (CIA) de la Universidad de Costa Rica (UCR) y

CoopeVictoria R.L. surge por la necesidad de dotar de herramientas de agricultura de precisión al sector cooperativo costarricense.</p>

<p>La pestaña SMS-CoopeVictoria R.L. cuanta con información histórica y actual sobre las plantaciones de caña de azúcar administradas por la Cooperativa, donde usted encontrará la siguiente información:</p>

<p>- Áreas de Producción.</p>

<p>- Distritos de Influencia.</p>

<p>- Historial de Cosecha por Finca.</p>

<p>- NDVI Calculado con el Sensor Sentinel 2 .</p> <!-- CAMBIE LA FECHA DE LA DE NDVI -->

</div>

<!-- Aquí inicia la segunda columna -->

<div class="column" style="background:rgba(255,255,255,.85)">

<h2> Características de la información suministrada.</h2>

<p> El NDVI es un índice de vegetación que está fuertemente relacionado con la porción fotosintéticamente activa de la vegetación. En caña de azúcar los valores rondan entre 0,25 (cuando inician los rebrotes) y 0,8 (previo al inicio de estrés de la planta por maduración). Esos valores variaran según: manejo, variedad, condiciones atmosféricas, entre otras.</br>

</br>

Los valores que encontrará en la pestaña SMS-CoopeVictoria R.L. deberán ser divididos entre 100 para reescalarlos de 0 - 1. Variaciones temporales abruptas en el índice pueden estar relacionadas factores atmosféricos (nubes o sombras de nubes). </br>

</br>

El estimado de rendimiento fue calculados con base a regresiones lineales múltiples que buscan explicar la variación del rendimiento de campo en función al a índices de vegetación y variables históricas del cultivo.

</p>


```

</div>
</div>

<div class="row" style="background:rgba(255,255,255,.85);margin-right: 1.5px;">
  <center>
    
  </center>

</div>
</body>
</html>

```

Anexo 7. Código de programación del archivo .css vinculado con .html del geoportal desarrollado en la investigación.

```

#mapid
{
    height:420px;
    width: 1120px;
    margin-left: 75px;
    margin-right: 100px;
}

/* Crear dos columnas */
.column {
    float: left;
    width: 608px;
    padding:7px;
    height: 340px;
    text-align:left;
    border:0.2px solid;
    background-color:#CCD1D1;
    opacity: 0.99;
}

```

```
/* */
.row:after {
  content: "";
  display: table;
  clear: both;
}

h1 {
  font-size: 220%;
  text-align: left;
  margin: 10px;
}

h2 {
  font-size: 25px;
  margin: 10px;
}

p {
  font-size: 15px;
  text-align: justify;
}

.legend {
  color: #555;
  padding: 3px 4px;
  font: Verdana;
  background: white;
  background: rgba(255,255,255,0.8);
  box-shadow: 0 0 15px rgba(0,0,0,0.2);
  border-radius: 5px;
}

.legend i {
  width: 18px;
  height: 18px;
  float: left;
  margin-right: 8px;
  opacity: 0.75;
}
```

Anexo 8. Código de programación en JavaScript para definir las funciones integradas en el geoportal

```
//

// Mapa base
var map = L.map("mapid");

// Centro del mapa y nivel de acercamiento
var mapacoopevi = L.latLng([10.06773512, -84.2996532494]);
var zoomLevel = 12;

// Definición de la vista del mapa
map.setView(mapacoopevi, zoomLevel);

//Control de escala
L.control.scale({position:'topleft',imperial:false}).addTo(map);

// Adición de las capas base
esri = L.tileLayer.provider("Esri.WorldImagery").addTo(map);
osm = L.tileLayer.provider("OpenStreetMap.Mapnik").addTo(map);

// Indice de Vegetacion diciembre 2019
var ndvi = L.imageOverlay("ndvi.png",
    [[10.1387257937176312, -84.3679128707571806],
    [9.9967152737829750, -84.2313861376834154]],
    {opacity:0.8}
); //addTo(map); se eliminó esta parte para que no aparezca en la vista

//function updateOpacity() {
//    document.getElementById("span-opacity").innerHTML =
document.getElementById("sld-opacity").value;
//    ndvi.setOpacity(document.getElementById("sld-opacity").value);
//}

////////////////////////////////////
/

// Conjunto de control de Capas Base
var baseMaps = {
    "OpenStreetMap": osm,
```

```

        "ESRI World Imagery": esri
    };

    ///////////////////////////////////////////////////
    /

    // Conjunto de capas overlay
    var overlayMaps = {

    };

    ///////////////////////////////////////////////////
    /

    // Fincas de CoopeVictoria
    $.getJSON("lotes_coopevictoriar1.geojson", function(geodata) {
        var layer_geojson_lotes_coopevictoriar1 = L.geoJson(geodata, {
            style: function(feature) {
                return {'color': "black", 'weight': 1, 'fillOpacity':
0.0}
            },
            onEachFeature: function(feature, layer) {
                var popupText = "Finca: " + feature.properties.FINCA +
"<br>" + "Lote: " + feature.properties.LOTE +
                "<br>" + "Variedad: " + feature.properties.VARIEDAD +
"<br>" + "Área: " + feature.properties.AREA + " ha";
                layer.bindPopup(popupText);
            }
        }).addTo(map);
        control_layers.addOverlay(layer_geojson_lotes_coopevictoriar1, 'Áreas
de Producción');
    });

    ///////////////////////////////////////////////////
    /

    // Mapa de Coropletas

    // Paleta de colores capa NDVI en base a 100

```



```

var legend = L.control({position: 'bottomright'});
legend.title = "Rangos del NDVI"
legend.onAdd = function (map) {
    var div = L.DomUtil.create('div', 'info legend'),
        grades = [20, 30, 40, 50, 60, 70, 80, 90],
        labels = ['Categories'],
        from, to;
    for (var i = 0; i < grades.length; i++) {
        div.innerHTML +=
            '<i style="background:' + getColor(grades[i] + 1) + '></i> ' +
            grades[i] + (grades[i + 1] ? '&ndash;' + grades[i + 1] + '<br>' :
'+');
    }
    return div;
};
legend.addTo(map);

```

```

////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////

```

```

// Selección de fincas según interés

```

```

var fincas_coope = L.layerGroup().addTo(map);

function colorFincas(d) {
    return d == "TODOS" ? '#FFFF00' :
        d == "ALAMEDA" ? '#FF0000' :
        d == "ALFAROS" ? '#00FF00' :
        d == "ANA MARIA RIVERA HERRERO" ? '#FFFF00' :
        d == "ARAYA" ? '#FFFF00' :
        d == "ASSMAN" ? '#FFFA0' :
        d == "CELINA" ? '#FFFF00' :
        d == "CLOTILDE" ? '#FFFF00' :
        d == "COYOL" ? '#FFFF00' :
        d == "DAVID" ? '#FFFF00' :
        d == "FINCA LA LUISA" ? '#FFFF00' :
        d == "FLORES" ? '#FFFF00' :
        d == "GENARO" ? '#FFFF00' :
        d == "GRUPO HERRERO LA ARGENTINA" ? '#FFFF00' :
        d == "GRUPO MATEO AGUALOTE" ? '#FFFF00' :
        d == "JOSE MIGUEL FERNANDEZ" ? '#FFFF00' :
        d == "KOOPER CENTRAL" ? '#FFFF00' :

```

```

d == "MAQUINAS" ? '#FFFF00' :
d == "MAURICIO PERALTA" ? '#FFFF00' :
d == "MONTECRISTO" ? '#FFFF00' :
d == "PINTO" ? '#FFFF00' :
d == "POAS" ? '#FFFF00' :
d == "ROSALES" ? '#FFFF00' :
d == "SAN ROQUE" ? '#FFFF00' :
d == "SANTA ELENA" ? '#FFFF00' :
'#000000';
};

function estilo_fincas (feature) {
    return{
        fillColor: colorFincas(feature.properties.FINCA),
    };
};

function myFunction() {
    $.getJSON("lotes_coopectivictoriar1.geojson", function(geodata){
        var layer_geojson_lotes_coopectivictoriar1 = L.geoJson(geodata, {
            style: estilo_fincas,
            onEachFeature: function(feature, layer) {
                var popupText = "Finca: " +
feature.properties.FINCA;
                layer.bindPopup(popupText);
            }
        });
        fincas_coope.addLayer(layer_geojson_lotes_coopectivictoriar1);
        control_layers.addOverlay(layer_geojson_lotes_coopectivictoriar1,
'Fincas');
        layer_geojson_lotes_coopectivictoriar1.remove();
    });
};

function estiloSelect() {
    var miSelect = document.getElementById("estilo").value;

    $.getJSON("lotes_coopectivictoriar1.geojson", function(geodata){
        var layer_geojson_lotes_coopectivictoriar1 = L.geoJson(geodata, {
            filter: function(feature, layer) {

```

```

        if(miSelect != "TODOS")
        return (feature.properties.FINCA == miSelect );
        else
        return true;
    },
    style: estilo_fincas,
    onEachFeature: function(feature, layer) {
        var popupText = "Finca: " +
feature.properties.FINCA;
        layer.bindPopup(popupText);
        map.fitBounds(layer.getBounds());
    }
    });
    fincas_coope.clearLayers();
    fincas_coope.addLayer(layer_geojson_lotes_coopevictoriar1);
});
};
// Distritos de Influencia
$.getJSON("distritos_influencia.geojson", function(geodata) {
    var layer_geojson_distritos_influencia = L.geoJson(geodata, {
        style: function(feature) {
            return {'color': "#000000", 'weight': 1, 'fillOpacity':
0.0}
        },
        onEachFeature: function(feature, layer) {
            var popupText = "Provincia: " +
feature.properties.provincia + "<br>" + "Cantón: " + feature.properties.canton
+
            "<br>" + "Distrito: " + feature.properties.distrito +
"<br>" + "Área: " + feature.properties.area;
            layer.bindPopup(popupText);
        }
    }).addTo(map);
    control_layers.addOverlay(layer_geojson_distritos_influencia,
'Distritos de Influencia');
    layer_geojson_ndvi.remove();
});

// Fincas de CoopeVictoria
$.getJSON("rendimientohistorico.geojson", function(geodata) {
    var layer_geojson_historial = L.geoJson(geodata, {

```



```

        style: function(feature) {
            return {'color': "black", 'weight': 1, 'fillOpacity':
0.0}
        },
        onEachFeature: function(feature, layer) {
            var popupText = "Zafra 2016-2017: " +
feature.properties.PROD_16 + " Ton/ha" + "<br>" + "Zafra 2017-2018: " +
feature.properties.PROD_17 + " Ton/ha" + "<br>" + "Zafra 2018-2019: " +
feature.properties.PROD_18 + " Ton/ha" + "<br>" + "Zafra 2019-2020: " +
feature.properties.PROD_19+ " Ton/ha" ;
            layer.bindPopup(popupText);
        }
    }).addTo(map);
    control_layers.addOverlay(layer_geojson_historial, 'Historial de
Cosecha por Finca');
});

// Ubicación del control de capas
control_layers = L.control.layers(baseMaps, overlayMaps, {position:'topright',
"autoZIndex": true, collapsed:true}).addTo(map);

```