

UNIVERSIDAD DE COSTA RICA
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA DE BIOSISTEMAS

Trabajo Final de Graduación para optar por el título de
Licenciatura en Ingeniería Agrícola y de Biosistemas

*Identificación de caña de azúcar (Saccharum officinarum) empleando el clasificador Support
Vector Machine e imágenes Sentinel-1 y Sentinel-2 en el Distrito de Riego Arenal
Tempisque, Costa Rica*

ESTUDIANTE:

ALEX FABRICIO SILES CASTILLO

CARNÉ: B77362

DIRECTORA:

ING. MELISSA ROJAS DOWNING, Ph.D.

ASESORA:

ING. ALEJANDRA ROJAS GONZÁLEZ, Ph.D.

ASESOR:

ING. SERGIO ROJAS MONTERO, M.Sc.

Proyecto de Graduación

Ciudad Universitaria Rodrigo Facio

Agosto, 2025

Acta de la Presentación Oral del Trabajo Final de Graduación ante el Tribunal Examinador para Optar por el Grado de Licenciatura en Ingeniería Agrícola y de Biosistemas

Modalidad: Proyecto de Graduación

Lugar: Sala de Audiovisuales 1	Fecha: 05/08/2025	Hora: 09:00 am	Consecutivo: 16-2025
--	-----------------------------	--------------------------	--------------------------------

Artículo 1: Presentación de los miembros del Tribunal del Trabajo Final y de la estudiante.

Se explica el procedimiento que consiste en la presentación oral de 45 min, el espacio para preguntas, la deliberación del Tribunal y la comunicación del acuerdo, según los artículos 26 y 27 del Reglamento de Trabajos Finales de Graduación.

Tribunal Examinador

MIEMBROS DEL TRIBUNAL	ROL
Ph.D. María Melissa Rojas Downing	Directora TFG
Ph.D. Alejandra Rojas González	Miembro del Comité Asesor
M.Sc. Sergio Rojas Montero	Miembro del Comité Asesor
M.Sc. Matías Chaves Herrera	Lector Invitado
M.Sc. José Francisco Aguilar Pereira	Presidente

Artículo 2: Exposición oral del estudiante

Se realiza presentación en un tiempo entre 30 a 45 min máximo.

Nombre del Trabajo Final de Graduación: Identificación de caña de azúcar (*Saccharum officinarum*) empleando el clasificador Support Vector Machine e imágenes Sentinel-1 y Sentinel-2 en el Distrito de Riego Arenal Tempisque, Costa Rica.

NOMBRE DEL POSTULANTE	CARNÉ
Alex Siles Castillo	B77362

Artículo 3: Sesión de preguntas sobre aspectos propios del trabajo presentado.

Cada uno de los miembros del Tribunal realizan preguntas a la persona sustentante.

Artículo 4: Deliberación del Tribunal.

Sale de la sala de sesiones el sustentante para la deliberación del Tribunal. Se le avisará a la persona sustentante vía llamada telefónica o medio acordado.

Observaciones:

La presentación fue clara y estructurada, evidenciando dominio del tema y cumplimiento de los objetivos propuestos. Se respondieron adecuadamente las consultas del Tribunal. Se solicita incorporar con aclaraciones solicitadas y precisión metodológicas de los resultados adicionales. Las cuales deberán ser atendidas en la versión final del documento, bajo la verificación de su directora.

Producto de la deliberación se realiza la votación pública:

Votación pública obteniendo: 5 votos a favor y
0 votos en contra.

Voto disidente: 0

Calificación del Trabajo Final de Graduación:	Aprobado	X	No aprobado	
Aprobación con distinción el Trabajo Final de Graduación:	Sí	X	No	

Artículo 5: Comunicación del resultado de la deliberación a la persona sustentante

La persona que preside el Tribunal comunica a la persona postulante el resultado de la deliberación y se le declara acreedora al grado de **Licenciatura en Ingeniería Agrícola y de Biosistemas**. Se le indica la obligación de presentarse al acto público de graduación, para ser juramentada y recibir el diploma correspondiente.

Se da por concluida la sesión y firma la persona que preside el Tribunal Examinador a las once horas y dieciocho minutos.



M.Sc. José Francisco Aguilar Pereira, persona que preside

Nota: De conformidad con los artículos 28, 29 y 36, el estudiante deberá entregar las copias con la versión final, incluyendo las modificaciones formuladas en ocasión de la presentación pública, un resumen de 200-500 palabras y la autorización de publicación del trabajo en el repositorio Kérwá. El director del trabajo final de graduación será responsable de que se realicen las correcciones propuestas en la presentación pública y de que la versión final del documento quede debidamente confeccionada.

DEDICATORIA

A Indiahlay y Orfilia que les debo todo.

AGRADECIMIENTOS

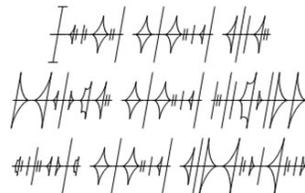
A Valeria, Jeustin y María, por su invaluable compañía y apoyo incondicional durante estos años. Su presencia fue un pilar fundamental en mi día a día.

A todos los amigos y compañeros con quienes tuve el privilegio de compartir esta etapa. Las risas, los desafíos y los momentos vividos enriquecieron enormemente mi experiencia.

A la Ing. Melissa Rojas Downing Ph.D., por su confianza constante y por brindarme siempre el apoyo y la motivación que fueron esenciales en todo este proceso.

A mis lectores, Ing. Alejandra Rojas González Ph.D. y Ing. Sergio Rojas Montero M.Sc., por sus valiosos consejos y su disposición. Su guía fue crucial para el buen desarrollo de este trabajo.

Y a Pamela Mesén, por darme la confianza que tanto necesité en los momentos clave.



ÍNDICE

DEDICATORIA.....	I
AGRADECIMIENTOS.....	II
ÍNDICE	III
ÍNDICE FIGURAS	V
ÍNDICE TABLAS.....	VIII
RESUMEN.....	1
1. INTRODUCCIÓN.....	2
1.1. Objetivos.....	5
2. MARCO TEÓRICO.....	6
2.1. Fundamentos teóricos y referencias conceptuales	6
2.1.1. Clasificación de cultivos con imágenes satelitales	6
2.1.2. Sensores activos y pasivos.....	6
2.1.2. Satélites Sentinel.....	7
2.1.3.6. Calibración de ruido granular	9
2.1.3.7. Gamma MAP	10
2.1.4.1. MSE	10
2.1.4.2. AD.....	11
2.1.4.3. SNR.....	11
2.1.4.4. PSNR.....	11
2.1.5. Desarrollo fenológico de la caña de azúcar en el DRAT.....	12
2.1.6. Máquina de soporte vectorial.....	13
2.1.6.1. Función de kernel.....	13
2.1.6.2. Valor de regularización.....	14
2.1.7.1. Índice Kappa	15
2.1.7.2. Medida F1	16
2.1.8. Índice de Vegetación de Diferencia Normalizada	16
2.2. Antecedentes del problema	17
3. METODOLOGÍA.....	22
3.1. Sitio de estudio.....	22
3.2. Descripción General.....	23

3.3.	Definición de métodos	24
3.3.1.	Preparación de datos y delimitación temporal	24
3.3.2.	Objetivo 1: Preprocesamiento de imágenes SAR	26
3.3.3.	Objetivo 2: Clasificación de caña de azúcar	27
3.3.4.	Objetivo 3: Análisis del desempeño del clasificador.....	28
4.	RESULTADOS Y DISCUSIÓN	30
4.1.	Delimitación temporal	30
4.2.	Preprocesamiento.....	31
4.3.	Evaluación de la clasificación de cultivos	35
4.4.	Análisis de sensibilidad de parámetros	38
4.5.	Desempeño del modelo en diferentes fechas.....	49
4.6.	Comparación de métodos con y sin NDVI	55
5.	CONCLUSIONES	63
6.	RECOMENDACIONES	64
7.	REFERENCIAS BIBLIOGRÁFICAS	65
8.	APÉNDICE	71
	APÉNDICE A:	72
	Objetivo específico 1	72
	APÉNDICE B:.....	87
	Objetivos específicos 2 y 3	87

ÍNDICE FIGURAS

Figura 1. Etapas fenológicas del cultivo de caña de azúcar, obtenido de (Vignola et al., 2018).	12
Figura 2. Sitio de estudio.	22
Figura 3. Flujo de trabajo empleado para el desarrollo del proyecto.	23
Figura 4. Diagrama de preprocesamiento realizado en SNAP.	26
Figura 5. Promedio del NDVI por mes en parcelas de caña de azúcar para el año 2019.	30
Figura 6. Preprocesamiento de imágenes Sentinel-1 (Polarización VH, filtro Gamma MAP y calibración radiométrica Sigma 0) del 25 de diciembre del 2019.	32
Figura 7. Promedio de los valores de MSE y AD obtenidos para los tamaños de ventana 5x5 y 7x7 del filtro Gamma Map.	34
Figura 8. Promedio de los valores de SNR y PNSR obtenidos para los tamaños de ventana 5x5 y 7x7 del filtro Gamma Map.	34
Figura 9. Polígonos de validación disponibles en la zona de estudio.	35
Figura 10. Puntos aleatorios generados para las clases “caña”, “arroz-pastos”, “bosque” y “agua”.	36
Figura 11. Índice kappa obtenido para cada una de las configuraciones de kernel estudiadas con el set de datos multitemporal y tamaño de ventana 5x5.	37
Figura 12. Índice kappa obtenido para cada una de las configuraciones de kernel estudiadas con el set de datos multitemporal y tamaño de ventana 7x7.	37
Figura 13. Análisis de sensibilidad del valor de regularización para el tamaño de ventana 5x5.	38
Figura 14. Análisis de sensibilidad para el valor de regularización para el tamaño de ventana 7x7.	39
Figura 15. Matriz de confusión para tamaño de ventana 5x5.	40
Figura 16. Matriz de confusión para tamaño de ventana 7x7.	40
Figura 17. Clasificación de la zona de estudio para las ventanas de 5x5 (a) y 7x7 (b).	42
Figura 18. Clasificación de la zona de estudio para un tamaño de ventana de 7x7.	43
Figura 19. Desempeño de la medida F1 para la ventana de 5x5.	43
Figura 20. Desempeño de la medida F1 para la ventana de 7x7.	44
Figura 21. Desempeño del modelo SVM en los escenarios con las clases de arroz y pasto combinadas	45
Figura 22. Desempeño del modelo SVM en los escenarios con las clases de arroz y pasto separadas.	45
Figura 23. Comportamiento promedio de la retrodispersión para cada una de las clases estudiadas.	46
Figura 24. Distribución del NDVI para el producto Sentinel-2 del 30 de noviembre.	47
Figura 25. Distribución del NDVI para el producto Sentinel-2 del 25 de diciembre.	49
Figura 26. Rendimiento del algoritmo para el 20 de septiembre del 2019 mediante análisis monotemporal de imágenes Sentinel-1 y ventana 7x7	50
Figura 27. Rendimiento del algoritmo para el 19 de noviembre del 2019 mediante análisis monotemporal de imágenes Sentinel-1 y ventana 7x7	50
Figura 28. Rendimiento del algoritmo para el 25 de diciembre del 2019 mediante análisis monotemporal de imágenes Sentinel-1 y ventana 7x7	50

Figura 29. Clasificación de la zona de estudio para el producto Sentinel-1 del 25 de diciembre del 2019.....	51
Figura 30. Rendimiento del algoritmo para el 30 de noviembre del 2019 mediante análisis monotemporal de imágenes de NDVI.....	52
Figura 31. Rendimiento del algoritmo para el 25 de diciembre del 2019 mediante análisis monotemporal de imágenes de NDVI.....	52
Figura 32. Clasificación de la zona de estudio para el análisis de NDVI del 25 de diciembre del 2019.....	53
Figura 33. Rendimiento con NDVI 5x5.....	56
Figura 34. Rendimiento sin NDVI 5x5.....	56
Figura 35. Rendimiento con NDVI 7x7.....	57
Figura 36. Rendimiento sin NDVI 7x7.....	57
Figura 37. Índice Kappa del algoritmo SVM al excluir el NDVI del conjunto de datos.	58
Figura 38. Clasificación de la zona de estudio para el análisis multitemporal sin NDVI con una ventana de 7x7.	58
Figura 39. Métricas de desempeño del modelo utilizando únicamente datos de NDVI.....	59
Figura 40. Clasificación de la zona de estudio para el análisis multitemporal solo empleando NDVI.....	60

Figura A. 1. Imagen preprocesada filtro Gamma MAP 5x5 del 20 de septiembre del 2019.	73
Figura A. 2. Imagen preprocesada filtro Gamma MAP 5x5 del 02 de octubre del 2019.	73
Figura A. 3. Imagen preprocesada filtro Gamma MAP 5x5 del 14 de octubre del 2019.	74
Figura A. 4. Imagen preprocesada filtro Gamma MAP 5x5 del 26 de octubre del 2019.	74
Figura A. 5. Imagen preprocesada filtro Gamma MAP 5x5 del 7 de noviembre del 2019.	75
Figura A. 6. Imagen preprocesada filtro Gamma MAP 5x5 del 19 de noviembre del 2019.	75
Figura A. 7. Imagen preprocesada filtro Gamma MAP 5x5 del 01 de diciembre del 2019.	76
Figura A. 8. Imagen preprocesada filtro Gamma MAP 5x5 del 13 de diciembre del 2019.	76
Figura A. 9. Imagen preprocesada filtro Gamma MAP 5x5 del 25 de octubre del 2019.	77
Figura A. 10. Imagen preprocesada filtro Gamma MAP 7x7 del 20 de septiembre del 2019.	77
Figura A. 11. Imagen preprocesada filtro Gamma MAP 7x7 del 02 de octubre del 2019.....	78
Figura A. 12. Imagen preprocesada filtro Gamma MAP 7x7 del 14 de octubre del 2019.	78
Figura A. 13. Imagen preprocesada filtro Gamma MAP 7x7 del 26 de octubre del 2019.	79
Figura A. 14. Imagen preprocesada filtro Gamma MAP 7x7 del 07 de noviembre del 2019.	79
Figura A. 15. Imagen preprocesada filtro Gamma MAP 7x7 del 19 de noviembre del 2019.	80
Figura A. 16. Imagen preprocesada filtro Gamma MAP 7x7 del 01 de diciembre del 2019.	80
Figura A. 17. Imagen preprocesada filtro Gamma MAP 7x7 del 13 de diciembre del 2019.	81
Figura A. 18. Imagen preprocesada filtro Gamma MAP 7x7 del 25 de diciembre del 2019.	81
Figura A. 19. Imagen preprocesada sin filtro del 20 de septiembre del 2019.	82
Figura A. 20. Imagen preprocesada sin filtro del 02 de octubre del 2019.....	82
Figura A. 21. Imagen preprocesada sin filtro del 14 de octubre del 2019.....	83
Figura A. 22. Imagen preprocesada sin filtro del 26 de octubre del 2019.....	83
Figura A. 23. Imagen preprocesada sin filtro del 07 de noviembre del 2019.....	84

Figura A. 24. Imagen preprocesada sin filtro del 19 de noviembre del 2019.....	84
Figura A. 25. Imagen preprocesada sin filtro del 01 de diciembre del 2019.....	85
Figura A. 26. Imagen preprocesada sin filtro del 13 de diciembre del 2019.....	85
Figura A. 27. Imagen preprocesada sin filtro del 25 de diciembre del 2019.....	86
Figura B 1. Visualización de la clasificación multitemporal con las clases de arroz y pasto unidas para un tamaño de ventana de 5x5.	88
Figura B 2. Visualización de la clasificación multitemporal con las clases de arroz y pasto unidas para un tamaño de ventana de 5x5 sin NDVI.....	88
Figura B 3. Visualización de la clasificación multitemporal con las clases de arroz y pasto unidas para un tamaño de ventana de 7x7 sin NDVI.....	89
Figura B 4. Clasificación de la zona de estudio para el producto Sentinel-1 del 20 de septiembre del 2019.....	89
Figura B 5. Clasificación de la zona de estudio para el producto Sentinel-1 del 19 de noviembre del 2019.....	90
Figura B 6. Clasificación de la zona de estudio para el producto NDVI del 30 de noviembre del 2019.....	90
Figura B 7. Clasificación de la zona de estudio para el análisis multitemporal sin NDVI con una ventana de 5x5.	91

ÍNDICE TABLAS

Tabla 1. Listado de productos Sentinel empleados para la conformación del conjunto de datos de entrenamiento y validación.	31
Tabla 2. Parámetros de evaluación para las imágenes preprocesadas con filtro Gamma MAP. ..	33
Tabla 3. Mejores rendimientos presentados para las configuraciones analizadas.	61
Tabla B 1. Métricas de rendimiento para el análisis multitemporal para ventana 5x5.	92
Tabla B 2. Métricas de rendimiento para el análisis multitemporal para ventana 7x7.	93
Tabla B 3. Métricas de rendimiento para el análisis de sensibilidad del valor de regularización para la ventana 5x5.	94
Tabla B 4. Métricas de rendimiento para el análisis de sensibilidad del valor de regularización para la ventana 7x7.	96
Tabla B 5. Métricas de rendimiento en el análisis de clases unidas.	98
Tabla B 6. Métricas de rendimiento en el análisis de clases separadas.	98
Tabla B 7. Métricas de rendimiento del análisis monotemporal para la imagen SAR del 20 de septiembre.	98
Tabla B 8. Métricas de rendimiento del análisis monotemporal para la imagen SAR del 19 de noviembre.	99
Tabla B 9. Métricas de rendimiento del análisis monotemporal para la imagen SAR del 25 de diciembre.	99
Tabla B 10. Métricas de rendimiento del análisis monotemporal para la imagen NDVI del 30 de noviembre.	100
Tabla B 11. Métricas de rendimiento del análisis monotemporal para la imagen NDVI del 25 de diciembre.	100
Tabla B 12. Métricas de rendimiento del análisis multitemporal sin NDVI para la ventana de 5x5.	100
Tabla B 13. Métricas de rendimiento del análisis multitemporal sin NDVI para la ventana de 7x7.	101
Tabla B 14. Resumen de las principales métricas de rendimiento obtenidas para todos los escenarios evaluados.	102

RESUMEN

La clasificación de cultivos mediante imágenes satelitales es una herramienta fundamental en la gestión agrícola, ya que permite monitorear el uso del suelo y optimizar la toma de decisiones. En este estudio, se evaluó el desempeño del clasificador Support Vector Machine (SVM) para la identificación de cultivos de caña de azúcar (*Saccharum officinarum*) en el Distrito de Riego Arenal Tempisque, Costa Rica, utilizando imágenes de los satélites Sentinel-1 y Sentinel-2.

Para la clasificación, se implementó un enfoque multitemporal, con datos ópticos y de radar. Se realizó un preprocesamiento de imágenes Sentinel-1 en el software SNAP versión 10.0.0, incluyendo la reducción de ruido granular mediante el filtro Gamma MAP, evaluando ventanas de 5x5 y 7x7 píxeles. Además, se incorporó el Índice de Vegetación de Diferencia Normalizada (NDVI) de las imágenes Sentinel-2 para mejorar la discriminación de clases.

Los resultados indican que la combinación de imágenes Sentinel-1 y Sentinel-2 mejoró significativamente la precisión de clasificación del cultivo caña de azúcar. La mejor configuración se obtuvo al preprocesar la imagen Sentinel-1 mediante el filtro Gamma MAP con la ventana de 7x7 píxeles y luego en el SVM utilizando un kernel de base radial y un valor de regularización de 15, logrando un Índice kappa de 0,835. Se observó que la diferenciación entre caña de azúcar y otras coberturas agrícolas, como arroz y pastos, fue un desafío debido a similitudes espectrales.

Se concluye que el uso de imágenes Sentinel-1 y Sentinel-2, junto con el algoritmo SVM, permite una clasificación eficiente de la caña de azúcar. Este estudio demuestra el potencial de la teledetección para el monitoreo de uno de los cultivos pilares del sector agroindustrial costarricense y genera un aporte como base para futuras aplicaciones en otros cultivos y regiones.

Palabras clave: Clasificación, SAR, NDVI, Caña de azúcar, Support Vector Machine.

1. INTRODUCCIÓN

La clasificación de cultivos es una de las más importantes aplicaciones que tiene la teledetección en el campo agrícola. Conocer cuáles son los cultivos presentes en las diferentes zonas productivas que componen una región es de gran utilidad a la hora de monitorear, diseñar e implementar prácticas y políticas agrícolas para asegurar una correcta seguridad alimentaria (Arias et al., 2020).

La Organización de las Naciones Unidas proyecta que la población mundial, a partir de año 2019, aumentará en 2000 millones de personas para el año 2050 (Organización de las Naciones Unidas [ONU], 2019), lo que conlleva un mayor desafío para la intensificación de las producciones agroalimentarias. Esto resalta la importancia de implementar un manejo óptimo de los recursos agrícolas en la búsqueda de la seguridad alimentaria mundial.

Un mejor entendimiento sobre la distribución y ubicación de cultivos es esencial para tener una rápida respuesta en la búsqueda de una gestión eficiente de los servicios ecosistémicos que forman parte de la producción de alimento. Un conocimiento amplio sobre el estado de la producción permite una eficaz regulación de recursos tales como agua, suelo e insumos agrícolas como fertilizantes y fitosanitarios. Además, posibilita una mejor adaptación frente a escenarios de variabilidad climática al anticipar variaciones en el crecimiento de los cultivos, optimizar el calendario agrícola y tomar decisiones informadas sobre prácticas de manejo sostenible (Bargiel, 2017).

En adición, la obtención precisa del área de cultivo permite a los pequeños productores, la industria y entes del Estado, estimar la producción anual, lo cual conlleva múltiples beneficios para los diferentes sectores, por ejemplo, la capacidad de proyectar rendimientos, gestionar recursos y planificar estrategias de comercialización (Brinkhoff et al., 2020). Un ejemplo de esto es lo desarrollado por el gobierno de Canadá, donde el departamento de Agriculture and Agri-Food Canada dispone de una serie de mapas que utilizan el procesamiento de imágenes satelitales para clasificar y conocer la situación de su producción agrícola en gran parte de su territorio. Esta información, se considera de suma importancia a la hora de tomar decisiones en el ámbito agrícola, ya que permite reducir la incertidumbre, mitigar riesgos productivos y climáticos, y orientar políticas públicas basadas en evidencia (Dingle et al., 2020).

En Costa Rica se han desarrollado diversas investigaciones para la clasificación de la cobertura terrestre del territorio nacional con diversos objetivos (Ávila-Pérez et al., 2020; Campos, 2010; Aguilar, 2019). No obstante, los estudios centrados específicamente en la identificación de cultivos individuales mediante el procesamiento de imágenes satelitales siguen siendo relativamente escasos en comparación, lo que resalta la necesidad de continuar impulsando investigaciones aplicadas a nivel nacional. Este tipo de estudios contribuye a generar insumos clave para la toma de decisiones agrícolas, fortalecer sistemas de monitoreo y promover un manejo más eficiente y sostenible de los recursos productivos.

La provincia de Guanacaste se destaca, entre otras cosas, por su alta producción agrícola. Pese a ello, se ha visto afectada por el aumento en las temperaturas y la disminución de las precipitaciones, esto genera la necesidad de formar proyectos tales como el Distrito de Riego Arenal Tempisque (DRAT). El cual busca fomentar el desarrollo agropecuario mediante la transformación de la agricultura extensiva de secano hacia la agricultura intensiva con riego (Servicio Nacional de Aguas Subterráneas Riego y Avenamiento (SENARA), 2021).

Según datos del SENARA para el 2020, el cultivo de caña de azúcar (*Saccharum officinarum*) dentro del DRAT se desarrolló en un área aproximada de 17 000 ha, correspondiendo a un 58 % del área que conforma todo el distrito de riego. Ese mismo año, el cultivo de caña generó alrededor de 45 millones de dólares lo cual contribuye de forma significativa en la economía de la zona y a nivel nacional (SENARA, 2021).

El SENARA, encargado del DRAT, enfrenta varios desafíos, los cuales surgen debido a variabilidad climática, ocasionando bajas precipitaciones en los embalses de alimentación de canales. Considerando lo mencionado anteriormente, el desarrollo de un mapa de clasificación de cultivos dentro del DRAT significará una herramienta útil para la generación de información que podrá ser utilizada por el SENARA al momento de realizar toma de decisiones en materia de la disposición del recurso hídrico que maneja.

Entes gubernamentales de Costa Rica como el Ministerio de Agricultura y Ganadería (MAG), pueden utilizar esta información para conocer el estado de la producción. Una aplicación de la teledetección dentro del DRAT se convierte en un componente crucial para abordar diversas situaciones, ya que la combinación de la información suministrada por los satélites Sentinel-1 y Sentinel-2, ofrece un enfoque integral para la clasificación de los cultivos. Además, esta

investigación contribuiría a alcanzar la meta de una gestión eficiente de los recursos naturales, en línea con el Objetivo de Desarrollo Sostenible número doce: Producción y Consumo Responsables (ONU, 2015).

El análisis de cobertura agrícola, tradicionalmente, se ha realizado utilizando datos de campo, lo cual se convierte en una problemática. Esto se debe a que en grandes extensiones agrícolas es necesario contar con un numeroso equipo de trabajadores para realizar visitas periódicas a agricultores, con el fin de conocer los cultivos que están produciendo. Esto tiene como consecuencia un alto costo económico para los entes que deseen conocer esta información, así como una falta de datos en tiempo real debido a los intervalos de revisita que se tendría en cada producción.

La identificación de cultivos con productos satelitales puede solventar el problema de la dependencia de datos de campo, así como optimizar la gestión agrícola mediante la automatización del monitoreo de cultivos. Esta clasificación se puede realizar mediante diversos enfoques, como algoritmos de aprendizaje automático (Support Vector Machine, Random Forest, Redes Neuronales, entre otros), análisis espectral, índices de vegetación, o combinación de datos multiespectrales y radar. Sin embargo, numerosos estudios han demostrado que los rendimientos de clasificación pueden variar significativamente según las características de la zona de estudio, tales como la topografía, la cobertura nubosa, el tipo de cultivo y la calidad de los datos disponibles (Dingle et al., 2020).

Este proyecto tiene como finalidad evaluar la eficacia del clasificador Support Vector Machine (SVM) para la identificación de áreas cultivadas con caña de azúcar en el DRAT. A través de esta investigación se busca generar información geoespacial precisa que permita delimitar con mayor exactitud los cultivos de caña, facilitando su monitoreo y manejo. Asimismo, se espera contribuir al desarrollo de metodologías replicables para la clasificación de cultivos específicos en contextos productivos nacionales, fortaleciendo así las capacidades locales en el uso de tecnologías de percepción remota para la toma de decisiones agrícolas.

En el caso de esta investigación, se emplearon los satélites Sentinel-1 y Sentinel-2, de donde se obtendrán imágenes de radar e imágenes ópticas, respectivamente, para generar un set de datos más robusto y aumentar los rendimientos al momento de la clasificación (Jiang et al., 2019).

Se utilizó el modelo de SVM, el cual posee la habilidad de aprender patrones de clasificación con una eficiencia balanceada (Pisner & Schnyer, 2019). Para aumentar el rendimiento del modelo, fue necesario realizar un preprocesamiento de las imágenes satelitales. En este caso se usó el software “Sentinel Applications Platform” (SNAP) para el preprocesamiento de productos Sentinel-1 y Sentinel-2.

La investigación fue realizada en el área aproximada de 30 000 ha que conforma el DRAT. Se empleó información correspondiente al año 2019, del cual se disponían los datos de validación. Mediante un análisis del NDVI, se determinó el intervalo de tiempo análisis. Como alcance del proyecto se evaluó el algoritmo clasificador SVM para la identificación del cultivo de caña de azúcar dentro del DRAT, empleando información procedente de Sentinel-1 con dos diferentes tamaños de ventana del filtro Gamma MAP y el NDVI procedente de imágenes de Sentinel-2.

1.1. Objetivos

Objetivo general

Evaluar el clasificador Support Vector Machine para la identificación de áreas cultivadas de caña de azúcar mediante imágenes Sentinel-1 y Sentinel-2 en el Distrito de Riego Arenal Tempisque, Costa Rica.

Objetivos específicos

1. Evaluar el desempeño de dos tamaños de ventana del filtro Gamma MAP durante el preprocesamiento de imágenes Sentinel-1 para utilizarlas en la clasificación de caña de azúcar en el Distrito de Riego Arenal Tempisque.
2. Utilizar el clasificador Support Vector Machine para identificar áreas cultivadas de caña de azúcar empleando las imágenes preprocesadas Sentinel-1 y Sentinel-2 en la zona de estudio.
3. Evaluar la clasificación de caña de azúcar utilizando Support Vector Machine, NDVI de imágenes Sentinel-2 y comparando dos tamaños de ventana del filtro Gamma MAP resultantes del preprocesamiento de imágenes Sentinel-1.

2. MARCO TEÓRICO

2.1. Fundamentos teóricos y referencias conceptuales

2.1.1. *Clasificación de cultivos con imágenes satelitales*

La clasificación de cultivos permite conocer la distribución y desarrollo de diversos cultivos en una zona, lo cual tiene diversas aplicaciones, como lo puede ser la detección del cambio de uso de suelo, analizar el aumento o disminución de cultivos intensivos, identificación de zonas vulnerables a la erosión, entre otros. Lo que permite generar información para la toma de decisiones. Tradicionalmente, los métodos empleados recopilan información mediante visitas a campo, lo cual suele ser más costoso y lento si se compara con los métodos enfocados en teledetección (Barni & Ratto, 2022). Por otra parte, la teledetección es una técnica que permite obtener información sobre la superficie terrestre de una manera rápida y precisa. Esta se obtiene mediante sensores remotos como lo pueden ser los satélites o drones, lo que concede la oportunidad de realizar una clasificación de uso de suelos o de cultivos de forma eficiente (Cerezo, 2023).

2.1.2. *Sensores activos y pasivos*

Dentro de la diversidad de instrumentos utilizados para realizar observaciones a la tierra, se diferencian dos grandes grupos, los sensores activos y pasivos. La diferencia principal entre estos conjuntos de sensores se basa en la emisión de radiación. Los sensores activos emiten radiación electromagnética que se refleja en los objetos y captan la energía devuelta hacia el sensor. En el caso de los sensores pasivos, estos no emiten radiación ya que capturan la proveniente del sol reflejada por la tierra (Copernicus, 2018).

El radar de apertura sintética (SAR), es un tipo de sensor activo que permite obtener imágenes de alta resolución a larga distancia, este tipo de tecnología funciona enviando un pulso de microondas de forma lateral, lo que le permite recibir las señales de los objetos en diferentes momentos para así discriminarlos. A diferencia de los sensores ópticos, los sensores SAR permiten trabajar independientemente de las condiciones meteorológicas, lo cual le brinda una gran capacidad de trabajo en zonas de alta nubosidad. Además, este tipo de sensores trabaja a nivel macroscópico, lo que permite generar información valiosa de características como la estructura, rugosidad y contenido de humedad de la superficie terrestre (Adeli et al., 2020).

Si bien las imágenes SAR tienen la capacidad de operar bajo condiciones climáticas adversas, la presencia de lluvias al momento de la captura del producto puede influir en la señal recibida por el satélite, introduciendo variaciones en las mediciones. Esto se debe a que la retrodispersión es altamente sensible al contenido de humedad en el suelo y la vegetación, lo que puede modificar la respuesta del radar.

Por otra parte, el punto de rocío puede generar variaciones en la señal de la banda C proveniente del satélite Sentinel-1 por el mismo motivo, debido a que la presencia de humedad en la zona superficial de las hojas y el suelo puede alterar la retrodispersión. No obstante, estudios previos han indicado que el efecto del rocío sobre las imágenes SAR en aplicaciones relacionadas con el monitoreo de cultivos no genera diferencias significativas en la calidad de los datos adquiridos (Wood et al., 2002).

Las imágenes ópticas, provienen de un satélite pasivo, el cual capta las señales provenientes del sol que han sido reflejadas por la superficie terrestre. A diferencia de SAR, estas se ven afectadas por las condiciones climáticas, por lo cual en zonas de alta nubosidad generalmente se dispone de una menor cantidad de datos (Kuma et al., 2022).

2.1.2. Satélites Sentinel

Los satélites Sentinel forman parte del programa de observación de la tierra Copernicus, el cual es desarrollado por la Comisión Europea y la Agencia Espacial Europea (ESA). Sentinel es una familia de satélites los cuales disponen de diferentes características. Dentro de la gama de estos se encuentran Sentinel-1 y Sentinel-2, los cuales poseen diversas características que permiten obtener información relevante para el monitoreo de parámetros ambientales (Pereira-Sandoval, 2023).

La misión Sentinel-1 se encuentra conformada por dos satélites en órbita polar, separados por 180°, sincronizados con el sol, operan en ciclos de día y noche realizando observaciones a la Tierra mediante SAR con la banda C, lo que permite obtener visualizaciones de la superficie sin importar las condiciones meteorológicas del momento. Poseen una resolución mínima de 5 metros, no obstante, los productos Ground Range Detected (GRD) son preprocesados a una resolución de 10 por 10 metros. Cada captura abarca un área de alrededor de 400 kilómetros con un tiempo de revisita cercano a seis días, lo cual lo hace útil para análisis temporales debido a la gran cantidad de productos disponibles (ESA, s/f).

En adición, Sentinel-2 es una misión que se enfoca en la formación de imágenes multiespectrales, de 13 bandas que abarcan 290 kilómetros, con un periodo de revisita de cinco días, al igual que en el caso anterior se encuentra conformado por dos satélites en órbita separados por 180 °. Dentro de las bandas espectrales que lo conforman, cuatro poseen una resolución de 10 metros, seis de 20 metros y tres de 60 metros (ESA, s/f).

2.1.3. Preprocesamiento de imágenes SAR

Pese a que las imágenes SAR presentan una alta calidad y múltiples ventajas, es necesario realizarles un preprocesamiento con el fin de reducir los errores. El preprocesamiento de este tipo de productos puede ser variado, en donde algunos estudios utilizan diferentes calibraciones radiométricas, filtros de efecto granular, ventanas, entre otros, así como un diferente orden al momento de aplicar los procesos seleccionados. Sin embargo, las características propias de una zona geográfica pueden influir de diversas maneras en los resultados obtenidos del preprocesamiento (Dingle et al., 2020).

Comúnmente este preprocesamiento se desarrolla en el software SNAP, el cual es desarrollado por la ESA y posee una arquitectura común para todas las cajas de herramientas de los satélites Sentinel.

2.1.3.1. Aplicación de archivo de órbita

El uso de archivos de órbita constituye un paso esencial en el preprocesamiento de imágenes SAR, ya que permite corregir las distorsiones geométricas derivadas de las variaciones en la posición y velocidad del satélite. Estos archivos contienen información precisa sobre la trayectoria del sensor, lo que facilita la georreferenciación de las imágenes y garantiza una representación exacta de las características de la superficie terrestre. Sin embargo, los archivos de órbita no están disponibles de inmediato al momento de generar el producto; suelen publicarse días o incluso semanas después. Para abordar esta limitación, existen herramientas que permiten descargar y actualizar automáticamente los vectores de estado orbital correspondientes a cada imagen SAR (Filipponi, 2019).

2.1.3.2.Reducción de ruido térmico

Otro paso crucial es la reducción de ruido térmico, ya que la intensidad de las imágenes Sentinel-1 se ve afectada por este tipo de interferencia aditiva en el canal de polarización cruzada, lo que puede degradar significativamente la calidad del producto final. El software SNAP incluye una funcionalidad específica para su eliminación, este proceso emplea algoritmos para estimarlo y corregirlo, mejorando la claridad y precisión de la imagen (Kumar, 2021).

2.1.3.3. Reducción de ruido de borde

La generación de productos de nivel 1 del satélite Sentinel-1, puede verse afectada debido a la compresión acimutal y de rango, lo que provoca artefactos radiométricos en los bordes de la imagen. Esto genera la necesidad de ajustar el tiempo en el que se da inicio el muestreo de los datos para así compensar los efectos que provoca la curvatura terrestre. Los artefactos radiométricos pueden afectar la precisión de las mediciones y dificultar la comparación de diferentes partes de la imagen. La corrección de estos artefactos asegura que la información radiométrica sea consistente en toda la imagen (Kumar, 2021).

2.1.3.4. Calibración radiométrica

La calibración radiométrica es el proceso de convertir los valores de píxeles de la imagen SAR en unidades físicas, como lo es el coeficiente de retrodispersión. Este paso es esencial para realizar análisis cuantitativos de los datos SAR y para comparar imágenes adquiridas en diferentes momentos o por diferentes sensores (Mandal et al., 2019).

2.1.3.5.Calibración geométrica

Al momento de descargar las imágenes Sentinel-1, estas inicialmente carecen de coordenadas geográficas, lo que provoca que no se puedan emplear en diversas aplicaciones, la calibración geométrica busca proyectar las imágenes a un sistema de coordenadas. Este proceso corrige las distorsiones que se produjeron durante la adquisición, como las causadas por la topografía del terreno y las variaciones en la órbita del satélite (Moskolai et al., 2022).

2.1.3.6. Calibración de ruido granular

El ruido granular, también conocido como ruido de moteado, es un ruido multiplicativo, el cual se presenta comúnmente en los productos SAR. Este ruido surge debido a la variación aleatoria de la

señal de retorno que recibe el satélite. Existen diversas alternativas mediante las cuales se busca disminuir el efecto producido, dentro de las que se destacan las metodologías de filtrado, esto permite eliminar la mayor cantidad de ruido posible, mejorando la legibilidad de las imágenes. Diversos investigadores evalúan el rendimiento de filtros de moteado como lo son Gamma MAP, Lee, Lee Sigma, Frost, entre otros, llegando a presentar diferencias significativas según la zona de estudio (Maity et al., 2015).

2.1.3.7. Gamma MAP

Este filtro de reducción de ruido, también conocido como *Máximo A Posteriori*, se basa en un modelo de ruido multiplicativo, el cual posee parámetros de media y varianza variables. Dentro del modelo de funcionamiento del filtro, este asume que el valor del número digital (ND) presente en la imagen se encuentra entre el ND del pixel de interés y el ND medio del kernel (Meenakshi & Punitham, 2011). Este filtro ha sido utilizado en una extensa variedad de estudios a lo largo del tiempo, relacionados con la clasificación de cultivos. Debido a que, por sus características, permite un alto rendimiento de clasificación al no suprimir información importante en el proceso de filtrado (Nezry et al., 1996; McNairn et al., 2009; Dingle et al., 2018).

2.1.4. Métricas para el análisis de desempeño del filtro de ruido granular en imágenes SAR

Durante el preprocesamiento de imágenes SAR, resulta fundamental evaluar el desempeño del filtro de reducción de ruido granular aplicado. Para ello, se recurre a un conjunto de métricas que permiten cuantificar con precisión del filtrado y, en consecuencia, la calidad de la imagen resultante. Estas métricas proporcionan una base objetiva para comparar distintos métodos de filtrado y determinar su adecuación según el contexto del análisis. Entre las métricas más comúnmente utilizadas se encuentran el Error Cuadrático Medio (MSE), la Diferencia Promedio (AD), la Relación Señal Ruido (SNR) y la Relación Señal Ruido Máxima (PSNR) (Maity et al., 2015; Anand et al., 2018; Meenakshi & Punitham, 2011).

2.1.4.1. MSE

El MSE se emplea para encontrar diferencias entre la imagen original y la preprocesada, valores de mayor orden en el MSE indican una alta diferencia entre las dos imágenes estudiadas, lo que permite concluir que el preprocesamiento permitió reducir el ruido sin perder información importante (Rana & Suryanarayana, 2019).

$$MSE = \frac{1}{k} \left[\sum_{i=1}^n (I_u - I_f)^2 \right] \quad (1)$$

Donde I_u - imagen inicial sin filtro, I_f - imagen después de filtrar, k - número total de pixeles presente en la imagen.

2.1.4.2. AD

La AD, es la diferencia media entre la imagen previa a la aplicación del filtro y la filtrada, esto dividido por el tamaño que presenta la imagen. Este valor, por sus características, se relaciona en gran medida con el MSE. Valores bajos de AD reflejan un mejor rendimiento en la aplicación de filtros de moteado, lo cual también se presenta con el MSE (Jiménez & Guerrero, 2023).

$$AD = \frac{1}{MN} \sum_{j=1}^M \sum_{k=1}^N (X_{j,k} - X'_{j,k}) \quad (2)$$

Donde M – total de filas de la imagen, N – total de columnas de la imagen, X imagen original y X' la imagen filtrada.

2.1.4.3. SNR

En relación con el SNR, este es la relación de cálculo entre la potencia de la señal y la imagen con ruido granular, en este caso entre mayor sea el valor de SNR mejor es el desempeño del filtro de moteado (Santos et al., 2015).

$$SNR = 10 \log_{10} \frac{\sum_{i=1}^k S_i^2}{\sum_{i=1}^k (\bar{S}_i - S_i)^2} \quad (3)$$

Donde S_i equivale al valor de la señal y \bar{S} es el valor promedio de la señal a lo largo de todas las muestras.

2.1.4.4. PSNR

La PSNR, es la técnica de evaluación más utilizada para medir la calidad de la eliminación del ruido granular. Es la relación entre la potencia máxima posible de la señal y la potencia del ruido, la cual se calcula comúnmente en forma de decibelios. Al igual que el SNR, un valor alto de PSNR demuestra un resultado positivo en la supresión de los efectos de moteado presente en las imágenes SAR (Sara et al., 2019).

$$PSNR = 10 \log_{10} \left(\frac{P^2}{MSE} \right) \quad (4)$$

Donde P es el valor pico de la imagen y el MSE el error cuadrático medio.

2.1.5. Desarrollo fenológico de la caña de azúcar en el DRAT

El periodo de siembra de la caña de azúcar en el DRAT se desarrolla a inicios de año, ya que al contar con riego se facilita el proceso de germinación, el cual se desarrolla en los meses de enero a abril. Las fases de crecimiento y desarrollo vegetativos también comprenden un tiempo de alrededor de tres meses cada una, abarcando de mayo a octubre, por último, la fase de maduración se desarrolla entre los meses de noviembre a enero (Vignola et al., 2018). En la Figura 1 se aprecia la morfología del cultivo para cada una de sus etapas fenológicas.

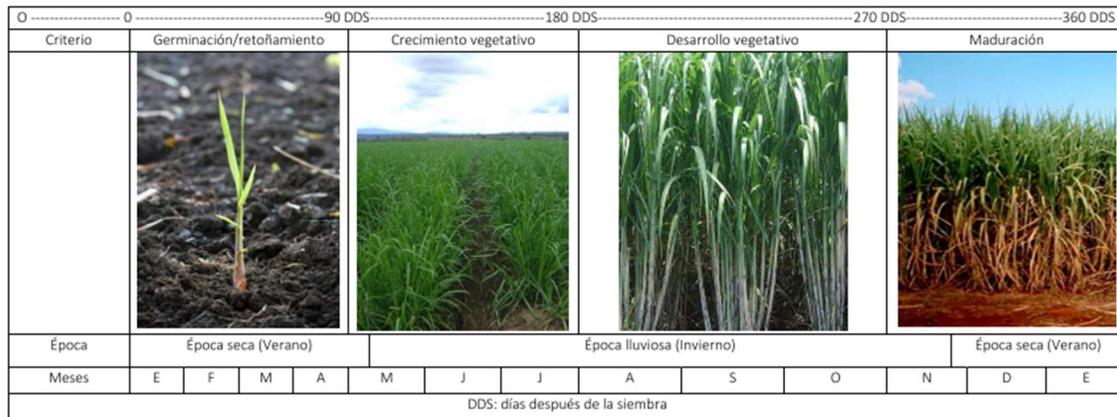


Figura 1. Etapas fenológicas del cultivo de caña de azúcar, obtenido de (Vignola et al., 2018).

En relación con la identificación de caña de azúcar, en etapas muy tempranas posterior a la siembra, se presenta una pequeña cobertura de dosel, que en las imágenes satelitales puede llegar a ser interpretada como suelo desnudo, lo cual puede incurrir en un error de clasificación (Vignola et al., 2018). Esto debido a que los polígonos de validación utilizados podrían indicar la presencia del cultivo de interés, cuando no es así.

En el caso de la cosecha sucede algo similar, donde algunas parcelas pueden ya encontrarse cosechadas al momento de obtener los productos Sentinel, provocando discrepancias en los datos de estudio lo que reduce el rendimiento del algoritmo SVM. Además, el periodo de cosecha de

caña de azúcar es sumamente susceptible a las precipitaciones, por lo que no se desarrolla en una fecha establecida a lo largo de los años, esta puede variar entre los meses de noviembre a enero.

Para garantizar una clasificación precisa de las plantaciones de caña de azúcar, es fundamental contar con un conjunto de datos depurado. Esto ayuda a evitar que elementos ajenos al cultivo de interés sean identificados incorrectamente como caña de azúcar u otras de las clases establecidas.

2.1.6. Máquina de soporte vectorial

La máquina de soporte vectorial o Support Vector Machine (SVM, por sus siglas en inglés), es un algoritmo de aprendizaje supervisado de inteligencia artificial. Debido a sus características, es comúnmente conocido como un clasificador, ya que es empleado para la separación de múltiples conjuntos de datos de diversas áreas. Este algoritmo funciona al generar un hiperplano el cual separa una serie de datos según las clases que se le hayan establecido. Estas separaciones se desarrollan según patrones de información introducidas al algoritmo para que lo pueda analizar (Pisner & Schnyer, 2019). Una vez este algoritmo es entrenado con un conjunto de datos representativo, se convierte en un modelo capaz de realizar nuevas predicciones para datos con características similares.

En algunos casos, dependiendo del conjunto de datos que se disponga, una separación lineal no presentará altos rendimientos en la clasificación. En estos casos, es posible modificar el kernel lineal que posee por defecto SVM, para así encontrar un hiperplano que divida mejor los datos en grupos, y proyectar características en una mayor dimensión. Adicionalmente, se puede definir un valor de regularización, el cual busca definir la sensibilidad del modelo de inteligencia artificial.

2.1.6.1. Función de kernel

En los modelos SVM, la función de kernel es esencial para abordar problemas de clasificación no lineal. Su objetivo principal es realizar transformaciones a los datos originales en un espacio de mayor dimensionalidad. Este enfoque se conoce como “truco de kernel” y permite que SVM encuentre hiperplanos que separen las clases de interés en un nuevo espacio. Existen diversos tipos de funciones de kernel de las cuales destacan:

Kernel Lineal: Este kernel presenta buenos rendimientos cuando los datos se pueden separar linealmente en su espacio original.

(5)

$$K(x, y) = xy$$

Kernel Polinomial: Este kernel es útil en casos donde la relación entre las clases sigue una función polinómica. Permite que el clasificador considere tanto las características individuales como las combinaciones hasta el grado d , lo que permite capturar interacciones más complejas entre las características.

(6)

$$K(x, y) = (xy + c)^d$$

Kernel de Base Radial: Este kernel es eficaz en situaciones donde la frontera de decisión es altamente no lineal, lo que lo hace ideal cuando no se tiene conocimiento previo sobre la distribución de los datos ya que puede adaptarse a infinitas formas de fronteras de decisión.

(7)

$$K(x, y) = e^{(-\gamma\|x-y\|^2)}$$

Para todas las ecuaciones de kernel, x y y son vectores de las características, c es una constante que controla la influencia de los términos de menor grado, d es el grado del polinomio y γ es un parámetro que define el alcance de influencia de un solo ejemplo de entrenamiento (López Cabrera & Pereira-Toledo, 2018).

2.1.6.2. Valor de regularización

Dentro de la configuración que se le puede dar al modelo SVM, se presenta el valor de regularización, también conocido como valor “C”. La función que posee es controlar el equilibrio entre minimizar el margen y reducir los errores de clasificación en el conjunto de entrenamiento. Valores altos de C provocan que el modelo priorice clasificar correctamente todos los puntos de entrenamiento, lo que puede llevar a un sobreajuste. Por otra parte, un valor bajo permite que se presenten más errores en la clasificación, lo que reduce la posibilidad de un sobreajuste y mejora la generalización del modelo (Lopez-Martinez, 2017).

2.1.7. Métricas para la evaluación de la clasificación de cultivos

En los procesos de clasificación de cultivos, es necesario realizar una evaluación que permita determinar si la metodología aplicada cumple con los objetivos propuestos. Aunque la generación

de mapas de visualización puede ofrecer ciertos criterios cualitativos, como la homogeneidad de los polígonos, para ingerir el desempeño relativo de los métodos utilizados, resulta imprescindible complementar este análisis con evaluaciones cuantitativas. La aplicación de métricas objetivas permite valorar de manera precisa el rendimiento del modelo de clasificación bajo diferentes configuraciones. Entre los indicadores más empleados en este tipo de estudios se encuentra el índice kappa y la medida F1, ampliamente utilizados en el ámbito de la clasificación de cultivos debido a su capacidad para reflejar la precisión y consistencia de los resultados obtenidos (Rojas & Medina, 2020; Zhang et al., 2024; Bargiel, 2017).

2.1.7.1. Índice Kappa

El índice kappa, muestra la diferencia entre la exactitud lograda en la clasificación, y la probabilidad que tiene una clasificación aleatoria de lograr la clasificación correcta. Un valor de 0 del índice kappa significa que la clasificación obtenida se dio de forma completamente al azar, mientras que un valor de uno sugiere una clasificación perfecta (Belmonte & Caamaño, 2009). Para obtener este valor se hace uso de la ecuación 8.

$$k = \frac{N \sum_{i=1}^r x_{(i,i)} - \sum_{i=1}^r (x_{(i,+)} x_{(+,i)})}{N^2 - \sum_{i=1}^r (x_{(i,+)} x_{(+,i)})} \quad (8)$$

Donde r es el número de filas de la matriz, $x_{(i,i)}$ el acuerdo observado, $x_{(i,+)}$ las observaciones por fila i, $x_{(+,i)}$ total de observaciones presentes en la columna i y N el total de observaciones incluidas dentro de la matriz (Belmonte & Caamaño, 2009).

El valor que debe poseer el índice kappa para que la clasificación sea exitosa puede variar según los alcances que se le quiera dar al proyecto. No obstante, la escala de valoración descrita por Landis & Koch (1977) es una de las más utilizadas por diversos autores y se ha mantenido vigente hasta la actualidad, empleándose en investigaciones de diversas temáticas donde se incluye la clasificación de cultivos (Rojas & Medina, 2020; Zhang et al., 2024; Andrade et al., 2024). La escala se presenta a continuación:

$K = 0$	Sin acuerdo
$0 < K \leq 0,20$	Acuerdo insignificante

0,20 < K = < 0,40	Acuerdo mediano
0,40 < K = < 0,60	Acuerdo moderado
0,60 < K = < 0,80	Acuerdo sustancial
0,80 < K = < 1	Acuerdo casi perfecto

2.1.7.2. Medida F1

En relación con la medida F1 este, al igual que en el caso anterior, posee un rango entre cero y uno, donde a mayor valor, mejor es la precisión de la clasificación (Lipton et al., 2014). La ecuación 9 presenta el cálculo de este parámetro.

$$F1 = \frac{2tp}{2tp + fp + fn} \quad (9)$$

Relacionado a este parámetro, también se puede obtener las métricas de precision y recall, los cuales equivalen a la precisión y sensibilidad de la clasificación. La precisión mide la proporción de predicciones positivas que son realmente correctas, se obtienen con la siguiente ecuación:

$$Precision = \frac{tp}{tp + fp} \quad (10)$$

En el caso de la sensibilidad, esta indica la proporción de elementos realmente positivos que fueron identificados de manera correcta. Se calcula como:

$$Recall = \frac{tp}{tp + fn} \quad (11)$$

Para las ecuaciones 9, 11 y 12, tp equivale a los verdaderos positivos, fp los falsos positivos y fn los falsos negativos (Lipton et al., 2014).

2.1.8. Índice de Vegetación de Diferencia Normalizada

El Índice de Vegetación de Diferencia Normalizada (NDVI) es un índice multiespectral comúnmente utilizado para conocer el estado de salud de las plantas. Este índice se relaciona con la actividad fotosintética y la estructura foliar que presentan las plantas. Valores altos de NDVI se relacionan con una vegetación sana, mientras que valores bajos pueden representar la presencia de enfermedades o la no presencia de vegetación en las imágenes. Este parámetro también es

comúnmente utilizado para la identificación de zonas con presencia de suelo desnudo u objetos que no presenten vegetación (Hidalgo, 2022).

2.2. Antecedentes del problema

La clasificación de cultivos mediante imágenes satelitales se puede desarrollar de diversas formas, no obstante, la metodología utilizada para su preprocesamiento y clasificación puede resultar en diferencias significativas en la calidad de la detección de cultivos. La siguiente información se centra en conocer cuáles han sido las metodologías utilizadas por diferentes autores, así como sus principales conclusiones.

Dingle et al. (2018), examina la relación entre el tamaño de las parcelas agrícolas y la selección de filtros de moteado y tamaños de ventana en imágenes SAR, con el objetivo de optimizar la clasificación de cultivos. El estudio se propuso determinar si la selección del tipo de filtro y el tamaño de ventana influye en el desempeño de los modelos de clasificación en regiones con diferentes dimensiones de parcela. Para ello, utilizaron un modelo de árboles de decisión y se analizaron tres zonas geográficas: Canadá, Argentina e Iowa, representando campos agrícolas pequeños (promedio de 3 ha), medianos (28 ha) y grandes (60 ha), respectivamente, en los cuales se encontraban los cultivos de maíz, soya y trigo. En cuanto al preprocesamiento de las imágenes, se evaluaron tres tipos de filtros de moteado: Gamma MAP, Touzi y Multi-temporal, cada uno aplicado con diferentes configuraciones de tamaño de ventana. El filtro Gamma MAP fue evaluado con siete tamaños de ventana, desde 3x3 hasta 15x15; el filtro Multi-temporal con nueve dimensiones, de 3x3 a 19x19; y el filtro Touzi con doce tamaños distintos, desde 11x11 hasta 33x33.

De este estudio se puede concluir que no existe una relación directa entre el tamaño de las parcelas y el rendimiento de la clasificación, no obstante, el rendimiento presentado dependerá de las características propias de la zona de estudio. Esto debido a que ninguno de los filtros analizados tuvo una diferencia significativa con los demás evaluados. Siguiendo esta misma línea, Dingle et al. (2020), realizaron un estudio para conocer las mejores técnicas de preprocesamiento de imágenes SAR para la identificación de cultivos. Dentro de sus objetivos se encontraba determinar el orden en el que se deben aplicar las herramientas de preprocesamiento, el impacto de diferentes filtros de moteado y diversas configuraciones de ventanas. Para esto, utilizaron la misma configuración empleada por estos autores en el estudio de Dingle et al. (2018).

Como principales conclusiones encontraron que el filtro que presentó los mejores resultados fue el Touzi seguido por el Gamma MAP. Sin embargo, se menciona que el uso de Gamma MAP con una ventana de 7x7 presentó los valores más altos de clasificación para todos los sitios de estudio y tipos de cultivo. Para la totalidad de los casos estudiados, las eficiencias fueron superiores al 85 %, además, concluyen que en ordenes altos de tamaño de ventana, se observa una disminución de la calidad de identificación de cultivos.

A continuación, se presentarán los estudios enfocados en la propia clasificación de cultivos y no en las metodologías de preprocesamiento. Bargiel (2017) tuvo como objetivo en su investigación clasificar diversos cultivos en Alemania, dentro de los cuales se encuentran pasto, maíz, remolacha, papa, entre otros. Este autor utilizó imágenes provenientes del satélite Sentinel-1 y los algoritmos de clasificación “Random Forest” y “Maximun Likelihood”, teniendo como resultado un valor de F1 superior a 0,8 para ambos clasificadores, lo cual sugiere una alta eficiencia al momento de desarrollar la clasificación. Este rendimiento el autor lo atribuye a aplicar un enfoque multitemporal lo que permite evaluar según las etapas fenológicas de los cultivos.

En el estudio realizado por Marini (2021), se analiza la identificación de cultivos de invierno, específicamente el trigo y la cebada, mediante teledetección. Estos cultivos presentan características fenológicas muy similares lo que genera un desafío para su clasificación. Para ello, el autor comparó diversas metodologías de clasificación con el objetivo de identificar aquella que ofreciera el mejor rendimiento. Su enfoque metodológico se basó en la aplicación de tres modelos de clasificación: Random Forest, Maximun Likelihood y SVM. Cada modelo fue evaluado utilizando tres conjuntos de datos distintos: el primero compuesto únicamente por productos del satélite Sentinel-1, el segundo, por imágenes provenientes de Sentinel-2; y el tercero por la combinación de datos de ambos satélites.

Como conclusión principal de la investigación, se tiene que los mejores resultados de todas las clasificaciones realizadas se obtuvieron al combinar imágenes ópticas e imágenes SAR, llegando a obtener una precisión global de 89,37 %. Lo que permite asegurar que la suma de estos productos puede presentar beneficios en cultivos donde sus características fenológicas dificultan su identificación.

Pese a que la clasificación utilizando imágenes Sentinel-1 presenta altos rendimientos (Dingle et al., 2018; Bargiel, 2017; Marini, 2021), diversos estudios han buscado maximizar la eficiencia de

clasificación incluyendo la aplicación de imágenes ópticas provenientes de los satélites Sentinel-2. Un ejemplo de esto se observa en la investigación realizada por Brinkhoff et al. (2020), donde utilizaron las polarizaciones VV y VH provenientes de Sentinel-1, y de los productos Sentinel-2, se emplearon las bandas azul, roja, verde, infrarrojo cercano y el índice NDVI, como set de datos para el modelo de clasificación. Dicho modelo se realizó con el algoritmo K-means para filtrar puntos de entrenamiento de algunas zonas específicas del mapa, así como SVM con un kernel de base radial para la clasificación. Se obtuvieron rendimientos superiores al 84,8 % en la identificación de cultivos perennes tales como oliva, almendra, cereza, entre otros.

En relación con la identificación del cultivo de caña de azúcar, a nivel mundial se han desarrollado varios estudios que buscan determinar las mejores metodologías para aumentar la eficiencia de clasificación a través de diversos algoritmos. Sin embargo, estas metodologías pueden derivar en un aumento de los recursos necesarios tanto computacionales como de datos para su ejecución.

Mishra et al. (2023) realizaron una clasificación de varios cultivos al norte de India, dentro de los principales cultivos clasificados se encontraron el arroz, caña de azúcar, maíz y sorgo. Los autores aplicaron los algoritmos de clasificación Random Forest y Support Vector Machine, con un enfoque multitemporal, obteniendo índices kappa de 0,923 y 0,918, respectivamente. Para esto utilizaron las polarizaciones VV y VH provenientes del satélite Sentinel-1 y el software SNAP para realizar el preprocesamiento de estas. Dentro de los pasos de preprocesamiento calcularon los índices Normalized Ratio Procedure Between Bands (NRPB), Dual-Pol Diagonal Distance (DPDD), Inverse Dual-Pol Diagonal Distance (IDPDD) y el Vertical Dual De-Polarization Index (VDDPI) con el fin de generar un set de datos más robusto.

Los mayores rendimientos obtenidos fueron presentados en el arroz, llegando a presentar una precisión de usuario de hasta 98,30 %, no obstante, el rendimiento en el cultivo de caña de azúcar también fue alto, llegando a un 89,56 %. Los autores destacan el uso de índices bajo metodologías multitemporales, la cual presentó un alto rendimiento para todos los cultivos analizados.

En un estudio realizado en el sur de China, Wang et al. (2019) desarrollaron una metodología para mapear caña de azúcar en paisajes montañosos, utilizando imágenes multitemporales de Sentinel-2 en combinación con información fenológica del cultivo. Se generaron imágenes NDVI libres de nubes para tres etapas clave en el desarrollo que ellos definieron como plántula, elongación y cosecha, y se evaluó el desempeño de cinco clasificadores. El mejor resultado se obtuvo con SVM

de kernel polinomial, alcanzando un índice kappa de 0,88. El estudio destacó la importancia de incorporar la dinámica fenológica en los procesos de clasificación, no obstante, se identificaron dificultades al clasificar áreas sombreadas, lo que podría ser mitigado mediante la incorporación de texturas o datos de radar.

En esta misma línea Lozano-Garzon et al. (2022) realizaron una investigación para la identificación de caña de azúcar en el departamento de Bocayá, Colombia. Para esto usaron 28 polígonos de datos validados e imágenes satelitales provenientes de Landsat-8 y Sentinel-2. Con esto se construyó un conjunto de datos para entrenar los modelos de clasificación estudiados (Random Forest, K-Nearest Neighbors y SVM), los cuales fueron evaluados mediante métricas como la medida F1. Los mejores resultados se obtuvieron con el modelo K-Nearest Neighbors logrando un F1 de 98 %. Dentro de sus conclusiones, este artículo destaca que los modelos permiten aumentar la precisión de la identificación de caña de azúcar, no obstante, se requiere mejorar la recolección de datos de campo, aumentar la variedad de coberturas y considerar el uso de imágenes SAR para superar limitaciones de nubosidad.

En contraste, Mandal et al. (2022) exploraron el uso de datos SAR provenientes de RADARSAT-2 para la clasificación de cultivos en India. El enfoque combinó componentes principales con funciones de kernel para transformar las firmas temporales de retrodispersión, seguido por la clasificación de SVM. Esta estrategia no solo generó una precisión global de 89 %, sino que también permitió una mejor detección de falsos negativos en el cultivo de caña. Sin embargo, se observaron dificultades para separar cultivos con alta biomasa, como plátano, caña de azúcar y bosques debido a la saturación del sensor en la banda C.

El estudio realizado por Yeasin et al. (2022) tiene como objetivo principal el evaluar el potencial de los productos Sentinel-1 y Sentinel-2, así como su uso combinado para identificar el cultivo de caña de azúcar mediante aprendizaje automático. El estudio se realizó en 40 parcelas de donde se obtuvieron los datos de validación. Se aplicaron siete modelos de clasificación dentro de los cuales se encuentra, Random Forest, Neural Networks, Naive Bayes y SVM, aplicados a parámetros derivados de Sentinel-1 e índices de Sentinel-2. Los resultados indicaron que el uso de conjuntos de datos Sentinel-1 y Sentinel-2 produjo mejores resultados que su uso individual, alcanzando un índice kappa de 0,83. Entre los modelos, Random Forest, Neural Networks y SVM destacaron su

rendimiento, además, se concluyó que los índices NDVI y NDWI y las polarizaciones VV y VH fueron las variables más relevantes para predecir las fases fenológicas.

Por otra parte, Jiang et al. (2019) centraron su estudio en la clasificación de caña de azúcar, utilizando información multitemporal proveniente de Sentinel-1 y Sentinel-2, donde obtuvieron un rendimiento aproximado de 86,3 % en la clasificación de 598.95 km² de caña de azúcar. La metodología desarrollada consistió en primera instancia utilizar las polarizaciones VV y VH de las imágenes SAR, para realizar una clasificación utilizando Random Forest y XGBoost como algoritmos clasificadores.

Los autores mencionan que tuvieron tres limitaciones principales, la primera es que los estudios previos realizados para el mapeo de caña de azúcar empleaban sensores ópticos, lo cual limitaba las observaciones por la alta presencia de nubosidad en la zona de estudio; la segunda limitación sigue esta misma línea, ya que al ser las nubes ser una limitante, provocaban que los productos que no presentaran nubosidad no cubran la totalidad del ciclo de cultivo llegando a que en ocasiones solo se obtuvieran imágenes en fechas en las que la caña de azúcar ya fue cosechada.

La tercera limitante se debió a errores por superficies no vegetadas. El mapeo inicial podía identificar de forma errónea zonas sin vegetación como caña de azúcar, esto incluyendo zonas urbanas y cuerpos de agua, por este motivo los autores recomiendan incorporar en futuras investigaciones metodologías para suprimir del análisis zonas sin vegetación. Ligada a esta última limitante, con las bandas presentes en las imágenes de Sentinel-2, los autores procedieron a calcular el NDVI máximo para todo el año en el que se desarrolló el estudio. Esto con el fin de discriminar de una forma más eficiente las zonas donde no se presenta vegetación y así generar una máscara que le fue aplicada al mapeo inicial obtenido de las imágenes SAR.

3. METODOLOGÍA

3.1. Sitio de estudio

El Distrito de Riego Arenal Tempisque, donde se desarrolla este proyecto, se encuentra ubicado en la provincia de Guanacaste, Costa Rica, y dispone de un área de 29 500 ha habilitadas con el servicio de riego (SENARA, 2021). Dentro del área que abarca este proyecto, se encuentran localidades pertenecientes a los cantones de Cañas, Bagaces, Carrillo y Abangares, contribuyendo con el desarrollo de actividades agrícolas y generación de empleo para los habitantes de estas localidades. El cultivo de caña de azúcar abarcó para el 2020 un total de 17 000 ha en esta zona.

Este cultivo posee diversas etapas fenológicas las cuales son germinación, crecimiento y desarrollo vegetativo, y maduración (Cruz-Sanabria et al., 2020). Para maximizar el rendimiento de la clasificación, se realizó la selección de las imágenes Sentinel-1 y Sentinel-2 cuando el cultivo se encontró en etapa de desarrollo vegetativo, lo cual se definió mediante el cálculo del NDVI anual.

La investigación se llevó a cabo utilizando datos correspondientes al año 2019, debido a que para la aplicación de algoritmos supervisados como SVM es fundamental contar con datos de validación, los cuales estaban disponibles para este periodo. La ubicación del sitio de estudio se presenta en la Figura 2.

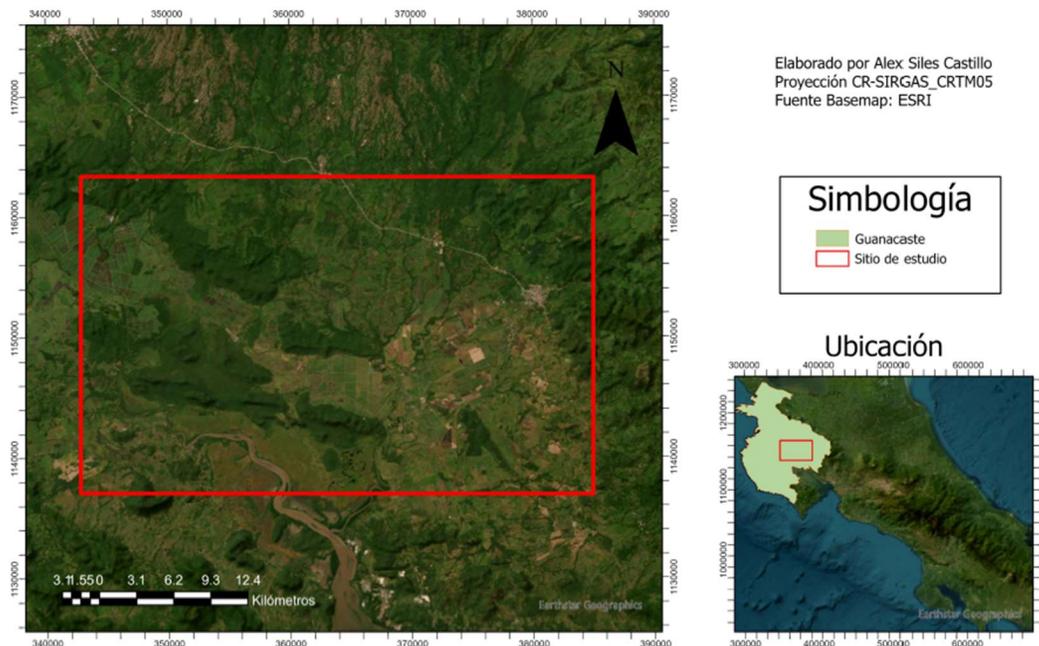


Figura 2. Sitio de estudio.

3.2. Descripción General

Para alcanzar los objetivos planteados en esta investigación, inicialmente se seleccionaron productos Sentinel-2 con una frecuencia mensual, abarcando la totalidad del año 2019. Estas imágenes fueron utilizadas para calcular el NDVI, con el fin de identificar con mayor precisión el periodo en el que el cultivo de caña de azúcar se encontraba en sus últimas etapas fenológicas. Por lo que este análisis previo funcionó como base para la determinación del periodo de estudio y la selección de las imágenes Sentinel-1 que serían sometidas al preprocesamiento (objetivo específico 1).

Una vez definido el periodo de estudio, se obtuvieron nuevas imágenes Sentinel-2 correspondientes a ese intervalo temporal, que presentaran condiciones de nubosidad favorables para el cálculo del NDVI. Los productos Sentinel-1 preprocesados, combinado con el NDVI de Sentinel-2 correspondiente al año 2019, se emplearon para generar 1880 puntos aleatorios, para cada una de las 11 imágenes satelitales estudiadas, teniendo un total de 20 680 muestreos, los cuales se dividieron en una distribución 70/30 para conformar los datos de entrenamiento y validación, respectivamente. La clasificación se realizó mediante la aplicación del algoritmo SVM, el cual permitió identificar las zonas de interés (objetivo específico 2).

Finalmente, se evaluó la eficiencia de la metodología propuesta mediante la obtención de índices de confiabilidad, con el fin de determinar su capacidad para procesar y clasificar con precisión las áreas cultivadas con caña de azúcar (objetivo específico 3). El flujo de trabajo detallado se encuentra representado en el diagrama en la Figura 3.

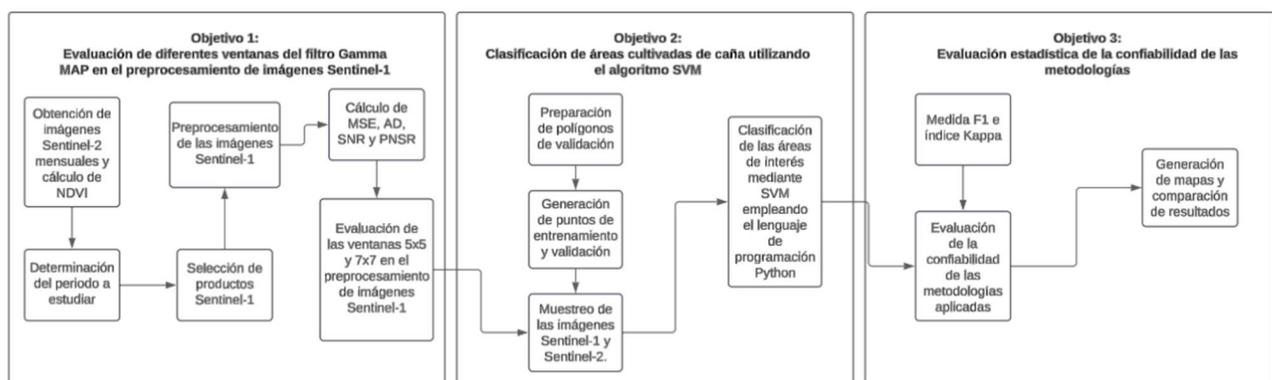


Figura 3. Flujo de trabajo empleado para el desarrollo del proyecto.

3.3. Definición de métodos

3.3.1. Preparación de datos y delimitación temporal

Como punto de partida, se contó con una base de datos de 591 polígonos georreferenciados dentro del DRAT. Esta información fue generada en el marco del proyecto C0214-22, inscrito en la Vicerrectoría de Investigación, de la Universidad de Costa Rica, y titulado “*Capacidades de SAR Sentinel-1 para la clasificación de cultivos en el Distrito de Riego Arenal Tempisque utilizando Random Forest*”. Este proyecto se enfocó en la evaluación del algoritmo Random Forest, mediante diversas configuraciones de ventana, datos satelitales, polarizaciones, entre otros, para un periodo de estudio que comprendió del 2017 al 2020.

En colaboración con SENARA, dicho proyecto recopiló información sobre las áreas agrícolas del DRAT y los cultivos presentes durante el intervalo de tiempo de análisis de esta investigación. Cada uno de estos polígonos contenían información relacionada con el cultivo que se encontraba cultivado en la parcela de referencia, esto para el periodo de estudio en el que se desarrolló esta investigación.

Los polígonos georreferenciados fueron consultados mediante el software ArcGIS Pro 2.8.0. A partir de los 591 polígonos originales, se descartaron aquellos que presentaban en las imágenes Sentinel-2 nubosidad o presencia de suelo desnudo, resultando en un total de 528 polígonos con condiciones adecuadas para ser utilizados en el proceso de clasificación. Además, se seleccionó un subgrupo de 14 polígonos que no presentaban nubosidad en la mayoría de las imágenes Sentinel-2 analizadas a lo largo del año. Este subgrupo fue clave para delimitar el intervalo temporal del estudio, ya que permitió identificar el periodo en el que la caña de azúcar alcanzaba su máximo desarrollo, condición fundamental para la selección de las imágenes satelitales Sentinel-1.

Una vez definidas las áreas de interés, se procedió a generar puntos de entrenamiento y validación para todas las clases consideradas. Estos puntos fueron generados de forma aleatoria con un buffer mínimo de 80 metros entre ellos, con el objetivo de evitar muestreos múltiples dentro de un mismo píxel al aplicar tamaños de ventana de 7x7.

La clasificación se realizó considerando cuatro clases: caña de azúcar, agua, arroz-pastos y bosque. Cabe destacar que, como actividad adicional a los objetivos de esta investigación, se realizó un

análisis exploratorio, donde los cultivos de arroz y pasto se agruparon en una misma clase para mejorar el desempeño en la identificación de la caña de azúcar. Además, se realizaron múltiples pruebas, como por ejemplo la separación de dichas clases y análisis monotemporales, para evaluar el impacto de esta agrupación en la precisión de la clasificación.

Para cada clase se generaron 470 puntos por clase, totalizando 1880 puntos por imagen. Esta cantidad se definió considerando que los polígonos correspondientes a cuerpos de agua eran de menor tamaño y no permitían extraer un mayor número de puntos. Dado que se desarrolló un análisis multitemporal, las coordenadas de los puntos se mantuvieron constantes en todas las imágenes Sentinel-1 y Sentinel-2 utilizados en el estudio, generando un total de 20 680 muestreos. Finalmente, el conjunto de datos fue dividido en un 70 % para entrenamiento y 30 % para validación.

3.3.1.1. Delimitación temporal

Se calculó el NDVI para el año 2019 utilizando datos provenientes del satélite Sentinel-2 y el software ArcGIS Pro. Con esto se determinó el intervalo de tiempo en el que la caña de azúcar se encontraba con un desarrollo óptimo para seleccionar las imágenes Sentinel-1 presentes en dicha ventana de tiempo.

Es importante tomar en consideración que las imágenes Sentinel-2 deben pasar por un proceso de corrección atmosférica para su correcta aplicación. La plataforma *Copernicus Browser* de la ESA presenta la característica que permite descargar los productos Sentinel-2 ya corregidos atmosféricamente, los cuales se identifican como productos L2A (Instituto Nacional de Técnica Aeroespacial, 2018). Por consecuencia, se descendieron estos productos de la plataforma, con los que fue posible realizar el cálculo del NDVI de forma directa.

Para la selección de las imágenes Sentinel-1 utilizadas en este estudio, se emplearon únicamente aquellas que presentaban una órbita descendente, las cuales son adquiridas alrededor de las 11:30 UTC (5:30 am hora local en Costa Rica). La elección de productos descendentes se fundamentó en factores como el punto de rocío y las condiciones meteorológicas prevalentes durante el periodo de estudio.

3.3.2. Objetivo 1: Preprocesamiento de imágenes SAR

Para aplicar el preprocesamiento a las imágenes Sentinel-1, se hizo uso del programa SNAP, el cual dispone de una serie de herramientas especializadas para el desarrollo de este tipo de procesos. Tomando esto a consideración el orden de aplicación de las herramientas de SNAP se desarrolló según lo mostrado en la Figura 4.

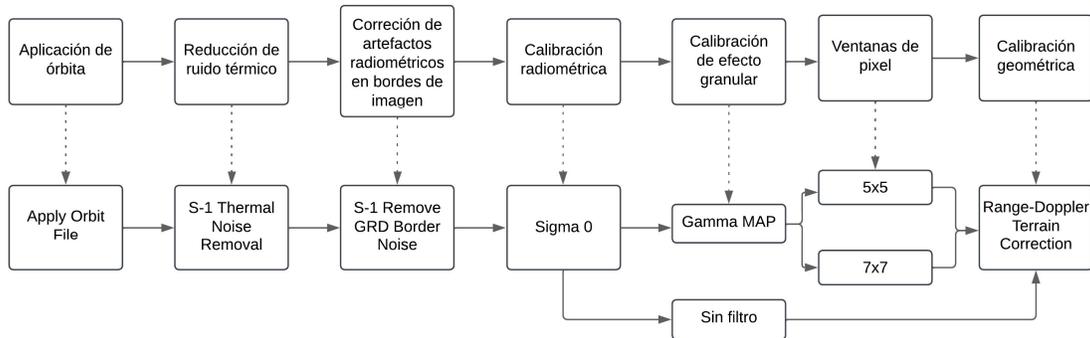


Figura 4. Diagrama de preprocesamiento realizado en SNAP.

En primera instancia se aplicó un archivo de órbita para poseer una ubicación precisa de las imágenes Sentinel-1. Este archivo se obtiene del propio software SNAP ejecutando la herramienta de “Apply Orbit File”. El siguiente paso es la aplicación del comando “Sentinel-1 Thermal Noise Removal”. Para realizar este proceso es necesario determinar el parámetro de procesamiento a tomar en cuenta, para lo cual se utilizó la polarización VH.

Se hizo uso de la herramienta “Sentinel-1 TOPS S-1 Remove GRD Border Noise” para los productos Sentinel-1. Al igual que en el caso de la reducción del ruido térmico el parámetro a utilizar fue la polarización VH, en relación con el límite de margen de borde, este se asigna por defecto dentro del software SNAP.

El siguiente paso fue la calibración radiométrica. En esta investigación se seleccionó la calibración Sigma 0 basándose en investigaciones previas donde se ha demostrado su eficacia en la clasificación de cultivos y principalmente el cultivo de caña de azúcar.

En relación con los filtros para la eliminación del moteado, en el caso de esta investigación se empleó el filtro Gamma MAP. En este caso, pese a ser un análisis multitemporal, se utilizó la herramienta “Single Product Speckle Filtering” dentro del software SNAP. En relación con el

tamaño de ventana, se trabajó con valores de 5x5 y 7x7 esto debido a que investigaciones previas realizadas por otros autores, muestran que entre mayor sea el tamaño de ventana, se puede dar un aumento en la pérdida de información que contiene el producto SAR (Dingle et al. 2020).

El último paso de preprocesamiento consistió en la aplicación de la herramienta “Range-Doppler Terrain Correction”. Este proceso requiere de un modelo de elevación digital (DEM), para lo cual se utilizó el SRTM de un segundo de arco. Además, como método de remuestreo se utilizó la interpolación bilineal y el sistema de coordenadas aplicado fue el WGS84.

Para comparar el efecto del filtro Gamma MAP en los tamaños de ventana 5x5 y 7x7, se aplicó el preprocesamiento descrito previamente, presentado en la Figura 4, dando como resultado la generación de tres conjuntos de imágenes. Los primeros dos grupos incluyeron todas las etapas de preprocesamiento descritas, incorporando el filtro Gamma MAP con los respectivos tamaños de ventana. En contraste, el tercer grupo pasó por el mismo preprocesamiento exceptuando la aplicación del filtro de moteado.

Una vez realizado el preprocesamiento de las imágenes Sentinel-1, se realizó una evaluación de su desempeño. La evaluación se dio entre las imágenes sin filtro y las filtradas, para lo cual se calculó los valores de MSE, AD, SNR y PNSR los cuales son comúnmente utilizados para la evaluación del desempeño de los filtros de efecto granular en imágenes SAR (Maity et al., 2015; Anand et al., 2018; Meenakshi & Punitham, 2011). Para eso se empleó el lenguaje de programación *Python* en la plataforma de *Google Colab*.

3.3.3. *Objetivo 2: Clasificación de caña de azúcar*

Para la clasificación del cultivo de caña de azúcar, se utilizó el clasificador Support Vector Machine, el cual se implementó en el entorno de *Google Colab*, el cual trabaja bajo el lenguaje de programación *Python*. Esto permitió, mediante librerías especializadas como *NumPy 1.26.4*, *GeoPandas 0.14.3*, *Rasterio 1.3.9* y *Scikit-Learn 1.4.0*, el procesamiento y análisis de imágenes satelitales, para luego realizar la clasificación.

Al momento de realizar la clasificación con SVM, evaluaron diferentes configuraciones de kernel para cada una de las ventanas estudiadas. Específicamente, se utilizaron los kernel: Lineal, Base Radial, Polinomial Grado 2, Polinomial Grado 3 y Polinomial Grado 4. Con un valor de

regularización de uno para todos los casos, el cual posteriormente se fue graduando hasta un máximo de 20.

Una vez entrenado el modelo, se generaron mapas temáticos, en ArcGis Pro, para evaluar de manera gráfica el rendimiento de la clasificación. Adicionalmente, se realizó un análisis estadístico para evaluar la precisión del clasificador, el cual también se desarrolló en *Python* (Objetivo específico 3).

3.3.4. Objetivo 3: Análisis del desempeño del clasificador

El análisis estadístico empleado se basa en el cálculo de dos índices de confiabilidad, los cuales son la medida F1 y el Índice kappa presentes en las Ecuaciones 8, 9, 10 y 11. Estos índices son comúnmente utilizados a la hora de analizar el desempeño de algoritmos de clasificación en cultivos. La obtención de estos parámetros se realizó utilizando el lenguaje de programación *Python*.

Además, se evaluaron múltiples escenarios para la conformación del set de datos, esto se realizó debido a que se deseaba conocer cual configuración, tanto del modelo como del conjunto de datos de entrenamiento, presentaba el mejor desempeño. Dentro de las diferentes configuraciones estudiadas, se evaluaron escenarios con diferentes cantidades de clases de datos, kernels y valores de regularización.

3.3.5. Pruebas adicionales realizadas

Tomando en consideración la disponibilidad de un conjunto de datos amplio y diverso, además de los procedimientos definidos para dar cumplimiento a los objetivos de investigación, se realizaron análisis complementarios con el propósito de enriquecer la evaluación del clasificador. Estos análisis consistieron principalmente en modificaciones tanto en los conjuntos de datos de evaluación como en la metodología aplicada.

En cuanto a los conjuntos de datos, se llevaron a cabo pruebas utilizando exclusivamente información proveniente de imágenes SAR, y de manera independiente, pruebas empleando únicamente el índice NDVI derivados de Sentinel-2. Esta estrategia permitió comparar el desempeño del clasificador bajo escenarios con fuentes de información diferenciadas, y así valorar la eficiencia relativa de cada tipo de producto para la identificación de áreas cultivadas de caña de azúcar.

Por otra parte, se introdujeron variaciones metodológicas orientadas a la evaluación de escenarios monotemporales. En el caso de las imágenes SAR, se seleccionaron tres fechas distintas, espaciadas aproximadamente un mes entre sí, con el fin de analizar la estabilidad y consistencia del modelo en diferentes momentos del ciclo agrícola. En lo que respecta al NDVI, se consideraron dos fechas específicas, igualmente separadas por un intervalo de un mes, lo cual permitió observar la variación en el desempeño del clasificador en función del estado fenológico del cultivo.

Además, se llevó a cabo un análisis orientado a la agrupación de las clases en estudio, particularmente arroz y pastos. El objetivo fue evaluar de qué manera la unión o separación de estas categorías influía en el desempeño del modelo, tanto en términos generales como en la precisión alcanzada para cada clase específica incluida en el conjunto de datos. De esta manera, las pruebas adicionales aportaron evidencia complementaria sobre el comportamiento del modelo bajo diferentes condiciones de entrada y temporalidad, contribuyendo a una valoración más integral de la metodología propuesta.

4. RESULTADOS Y DISCUSIÓN

4.1. Delimitación temporal

Para la determinación del espacio temporal donde se desarrolló el proyecto, se siguió el procedimiento establecido en el apartado de metodología, donde se obtuvo el NDVI para cada uno de los meses del 2019. El resultado obtenido se encuentra en la figura 5. Es importante aclarar que en los meses de julio, agosto y septiembre no fue posible obtener productos Sentinel-2 sin presencia de nubes en los polígonos seleccionados por lo que se excluyeron del análisis.

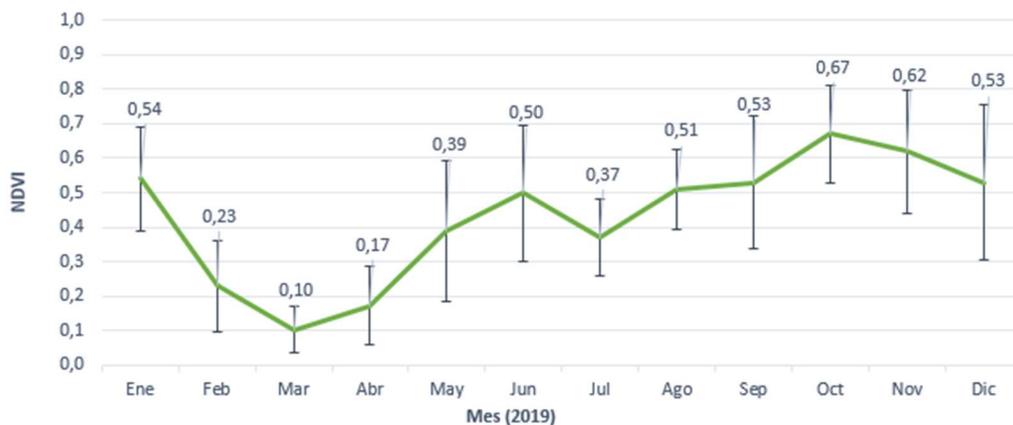


Figura 5. Promedio del NDVI por mes en parcelas de caña de azúcar para el año 2019.

El mes donde se presentó un NDVI mayor fue octubre, por lo cual se generó un radio de tiempo alrededor de este mes. Con esta información se determina que el periodo de tiempo en donde se dio un pico de desarrollo del cultivo es el que va de septiembre al mes de diciembre. Lo cual permitió seleccionar los productos Sentinel 1 y 2, obteniendo así 9 imágenes SAR y 2 ópticas, respectivamente.

En la Tabla 1 se detallan los productos empleados, en el caso de Sentinel-2, se utilizó una imagen Sentinel-2A (S2A) y otra Sentinel-2B (S2B). Esto se debe a que cada una proviene de un satélite diferente, sin embargo, su combinación es viable debido a que ambos productos comparten la misma órbita con una separación de 180° con el fin de reducir el tiempo de revisita y mejorar la disponibilidad de datos (ESA, s.f.).

Tabla 1. Listado de productos Sentinel empleados para la conformación del conjunto de datos de entrenamiento y validación.

Satélite	Producto	Fecha de obtención
Sentinel-1	S1A_IW_GRDH_1SDV_20190920T113042_20190920T113107_029104_034DB6_1F08	2019-09-20
	S1A_IW_GRDH_1SDV_20191002T113043_20191002T113108_029279_0353BA_F96D	2019-10-02
	S1A_IW_GRDH_1SDV_20191014T113042_20191014T113107_029454_0359C7_F8BC	2019-10-14
	S1A_IW_GRDH_1SDV_20191026T113042_20191026T113107_029629_035FC5_3E4A.	2019-10-26
	S1A_IW_GRDH_1SDV_20191107T113043_20191107T113107_029804_0365F6_4843	2019-11-07
	S1A_IW_GRDH_1SDV_20191119T113042_20191119T113107_029979_036C05_2FC0	2019-11-19
	S1A_IW_GRDH_1SDV_20191201T113042_20191201T113107_030154_037216_004C	2019-12-01
	S1A_IW_GRDH_1SDV_20191213T113041_20191213T113106_030329_037825_9240	2019-12-13
	S1A_IW_GRDH_1SDV_20191225T113041_20191225T113106_030504_037E2F_1147	2019-12-25
Sentinel-2	S2B_MSIL2A_20191130T160619_N0500_R097_T16PFS_20230526T173841	2019-11-30
	S2A_MSIL2A_20191225T160651_N0213_R097_T16PFS_20191225T201205	2019-12-25

4.2. Preprocesamiento

En la Figura 6, es posible observar el efecto de los distintos tamaños de ventana del filtro Gamma MAP. Donde en la imagen sin filtro se aprecia el efecto de “sal y pimienta” típico del ruido granular presente en las imágenes obtenidas mediante los radares de apertura sintética. Además, para el caso de los tamaños de ventana de 5x5 y 7x7, se contempla como a mayor tamaño de ventana, menor es la resolución obtenida.

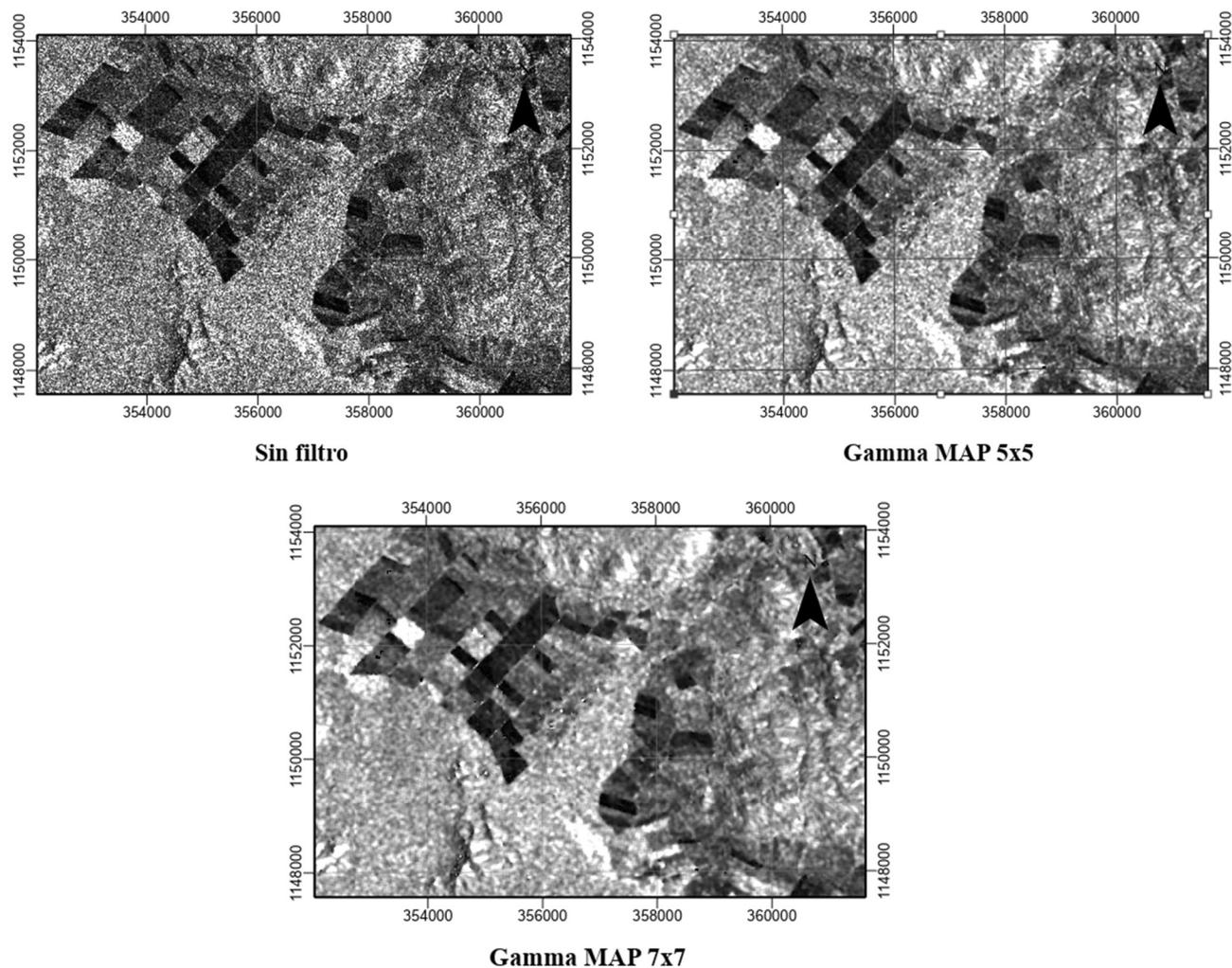


Figura 6. Preprocesamiento de imágenes Sentinel-1 (Polarización VH, filtro Gamma MAP y calibración radiométrica Sigma 0) del 25 de diciembre del 2019.

En el caso de desear consultar de forma visual los demás productos Sentinel-1 preprocesados, estos se encuentran disponibles en el apéndice A. De forma específica, se encuentran ilustrados desde la Figura A1 a la A27. Los valores de MSE, AD, SNR Y PNSR para los dos tamaños de ventana estudiados, se presentan en la Tabla 2, donde se puede observar que el filtro Gamma MAP 5x5 muestra un mejor desempeño en términos de preservación de información.

Tabla 2. Parámetros de evaluación para las imágenes preprocesadas con filtro Gamma MAP.

Filtro	Fecha de obtención del producto	MSE	SNR	AD	PNSR
Gamma MAP 5x5	25/12/2019	0,000175	8,960183	0,000146	47,514847
	13/12/2019	0,000184	9,002104	0,000321	47,455561
	1/12/2019	0,000190	8,989294	0,000349	47,406200
	19/11/2019	0,000192	8,995385	0,000363	50,936380
	7/11/2019	0,000194	8,994307	0,000385	49,257852
	26/10/2019	0,000264	9,047962	0,000488	48,565976
	14/10/2019	0,000288	9,037642	0,000502	48,104646
	2/10/2019	0,000214	8,795444	0,000423	53,123059
	20/9/2019	0,000223	9,062545	0,000458	48,214006
Gamma MAP 7x7	25/12/2019	0,000237	7,502420	0,000334	46,164883
	13/12/2019	0,000249	7,551623	0,000375	46,132434
	1/12/2019	0,000255	7,579758	0,000397	46,114030
	19/11/2019	0,000261	7,532420	0,000407	49,606014
	7/11/2019	0,000264	7,536976	0,000425	47,932185
	26/10/2019	0,000358	7,604957	0,000541	47,249599
	14/10/2019	0,000393	7,554517	0,000573	46,759349
	2/10/2019	0,000279	7,583108	0,000468	51,981264
	20/9/2019	0,000302	7,613732	0,000507	46,893098

Nota: AD: diferencia promedio; MSE: error cuadrático medio; SNR: relación señal-ruido; PSNR: relación señal-ruido máxima. Métricas utilizadas para cuantificar el error o la fidelidad entre imágenes, donde valores bajos de AD y MSE, y altos de SNR y PSNR, indican mejor calidad.

Es importante tomar en cuenta que valores bajos de MSE y AD evidencian que el filtro logró reducir el ruido sin perder información importante. De forma inversa, valores altos de SNR y PNSR demuestran un rendimiento óptimo del filtro de efecto granular. En las Figuras 7 y 8 se presenta de una forma más clara los valores obtenidos para cada uno de los casos estudiados. En primer lugar, se obtuvo un valor de MSE menor para las imágenes preprocesadas con un tamaño de ventana de 5x5 (Figura 8). Este es un comportamiento esperado, debido a que tamaños de ventana menores permiten que pérdida de información sea más pequeña, lo cual coincide con los resultados obtenidos por Rana & Suryanarayana (2019). Sucede de la misma forma para el parámetro de AD, el cual también refleja un mejor rendimiento para el preprocesamiento realizado con el tamaño de ventana menor.

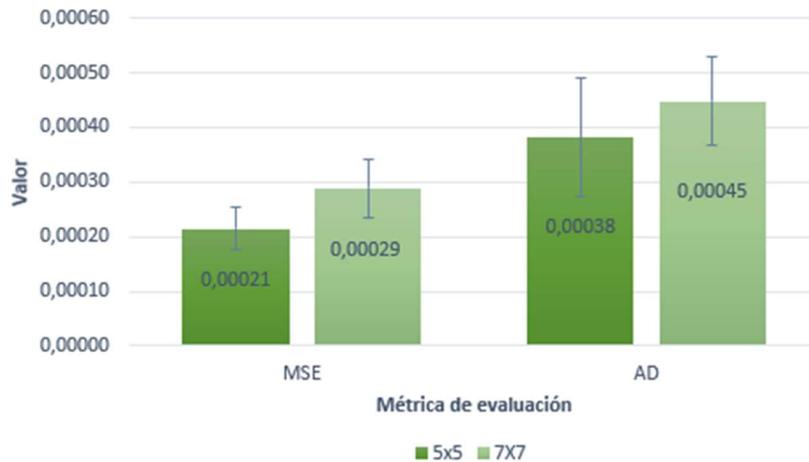


Figura 7. Promedio de los valores de MSE y AD obtenidos para los tamaños de ventana 5x5 y 7x7 del filtro Gamma Map.

En el caso de los parámetros de SNR y PNSR, como se mencionó anteriormente, un valor mayor demuestra un mejor rendimiento del filtro granular y, por ende, de la configuración de ventana utilizada. Por esto es posible apreciar como en el caso de la ventana 7x7 se presentan valores menores, lo cual concluye que para esta configuración se da una mayor pérdida de información al momento de eliminar el ruido presente en la retrodispersión de las imágenes de radar. Esto se alinea con los resultados obtenidos por Dingle et al. (2020), donde se concluye que entre mayor sea el tamaño de ventana, el rendimiento obtenido puede ser reducido.



Figura 8. Promedio de los valores de SNR y PNSR obtenidos para los tamaños de ventana 5x5 y 7x7 del filtro Gamma Map.

Considerando los valores obtenidos para las métricas utilizadas el filtro Gamma MAP presentó un buen funcionamiento en la eliminación del ruido de efecto granular, esto en las dos configuraciones empleadas. Sin embargo, la configuración de ventana de 5x5 presentó una pequeña mejora en los valores de MSE, AD, SNR y PNSR, logrando reducir el ruido sin perder información valiosa que puede ser utilizada en procesamientos posteriores. Es importante tener en cuenta también que menores tamaños de ventana pueden provocar aumentos en los tiempos de procesamiento de las imágenes en SNAP (Dingle et al., 2020).

Tomando esto a consideración, pese a presentar un mejor rendimiento el preprocesamiento de las imágenes con un tamaño de ventana de 5x5, se procederá a evaluar la clasificación de los dos grupos, para así conocer si se genera una diferencia significativa entre las dos configuraciones estudiadas.

4.3. Evaluación de la clasificación de cultivos

Se evaluaron múltiples alternativas con el fin de maximizar los rendimientos del modelo, buscando siempre mantener un equilibrio entre rendimiento y practicidad. En la Figura 9, se aprecian los polígonos empleados para la conformación del set de datos, los cuales fueron seleccionados asegurando que presentaran las características de interés a lo largo del periodo de estudio. Esto se debe a que algunos polígonos podrían presentar condiciones de nubosidad o presencia de suelo desnudo en alguna de las imágenes satelitales empleadas.

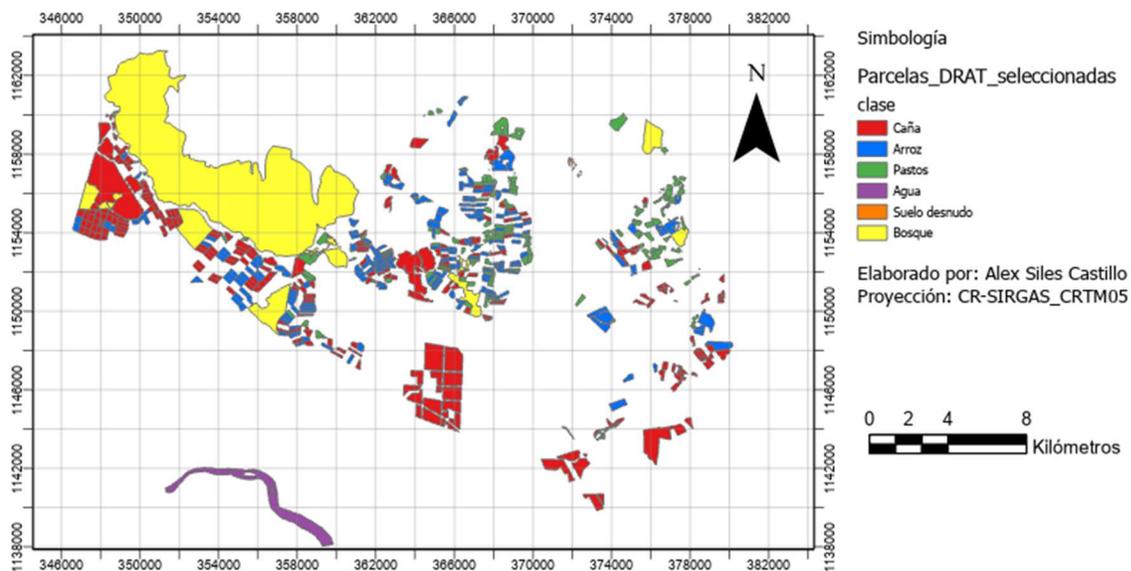


Figura 9. Polígonos de validación disponibles en la zona de estudio.

Esta decisión limitó la cantidad de puntos máximos que se podían presentar en los polígonos, ya que el buffer no lo permitía. Por este motivo, para la clase agua se obtuvieron un total de 470 puntos como máximo, por lo que, al buscar un balance de clases, a todas las demás se les aplicó la misma cantidad de muestreos. En el caso de la clase de suelo desnudo, al presentar polígonos con áreas muy pequeñas, no fue posible obtener esta cantidad de muestras, por lo cual fue eliminada del análisis.

Por último, al revisar de forma detallada los polígonos se observó que los que correspondían a arroz y pastos presentaban características similares lo cual podría provocar que, al momento de realizar la clasificación, estas clases presentaran problemas en su identificación. Tomando esto a consideración, se redujo la clasificación a cuatro clases finales “caña”, “arroz-pastos”, “bosque” y “agua”.

En la Figura 10 se observan los puntos aleatorios generados para cada una de las clases estudiadas con sus respectivas coordenadas. Como se mencionó anteriormente, estos puntos se mantuvieron para todos los escenarios e imágenes analizadas.

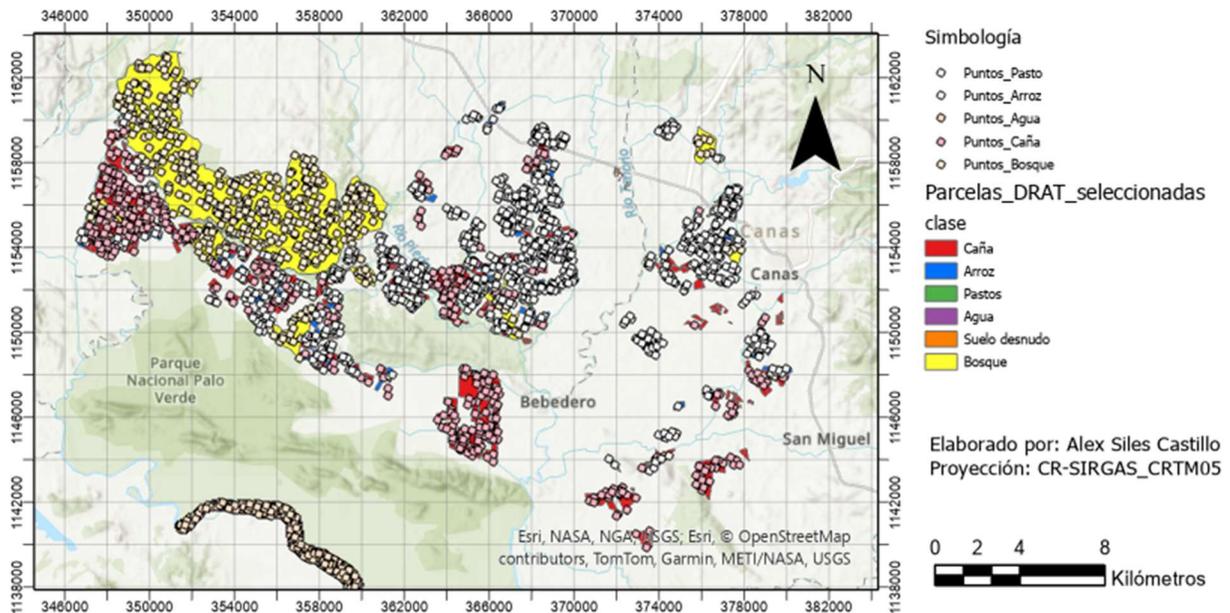


Figura 10. Puntos aleatorios generados para las clases “caña”, “arroz-pastos”, “bosque” y “agua”.

Como se aprecia en la Figura 11, el menor rendimiento con el enfoque multitemporal para la ventana 5x5 obtuvo un valor de índice kappa de 0,686, el cual corresponde a un kernel polinomial de grado 4.

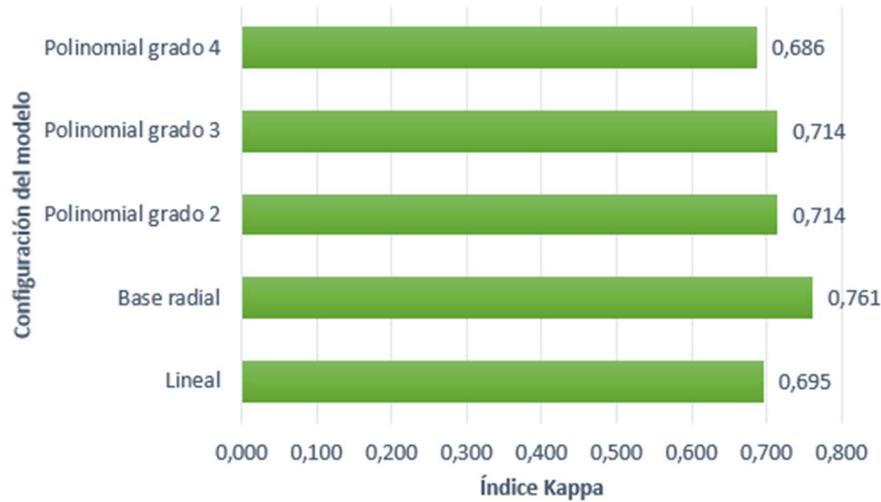


Figura 11. Índice kappa obtenido para cada una de las configuraciones de kernel estudiadas con el set de datos multitemporal y tamaño de ventana 5x5.

Por otra parte, el kernel de base radial presentó el mejor rendimiento, específicamente un índice kappa de 0,761. En relación con el tamaño de ventana 7x7, se dio el mismo comportamiento, el kernel polinomial de grado 4 generó un índice kappa de 0,688 mientras que el de base radial de 0,759. Estos resultados se pueden consultar en la Figura 12.

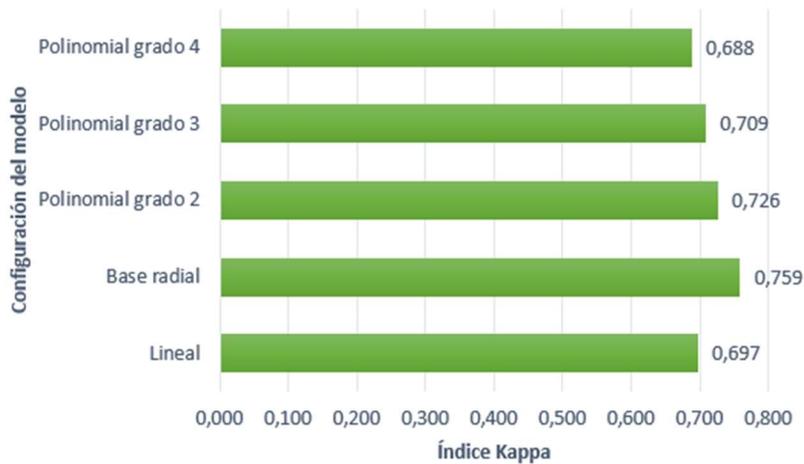


Figura 12. Índice kappa obtenido para cada una de las configuraciones de kernel estudiadas con el set de datos multitemporal y tamaño de ventana 7x7.

Este resultado permite conocer que las características dimensionales de los datos no pueden ser delimitadas fácilmente mediante funciones lineales y polinomiales. No obstante, ya que el kernel de base radial se adapta a infinitas formas de fronteras de decisión, genera un hiperplano óptimo para las características de los datos presentes en el modelo.

Con esta configuración, las diferencias presentadas en los rendimientos son mínimas teniendo un índice kappa máximo de 0,761 para 5x5 y de 0,759 para 7x7. No obstante, debido a que el kernel de base radial presentó los mejores resultados, se procede a realizar un análisis de sensibilidad para el valor de regularización utilizando este kernel como base, con el fin de conocer cuál es la mejor configuración para el modelo.

Las métricas asociadas a la medida F1 para cada una de las configuraciones analizadas previamente se presentan en las tablas B1 y B2 del apéndice B. Sin embargo, en el caso de la configuración con kernel de base radial, la medida F1 será examinada en la sección dedicada al análisis de sensibilidad que se presenta a continuación.

4.4. Análisis de sensibilidad de parámetros

En la Figura 13, se visualiza como, con un valor de C equivalente a 15, se obtuvo un rendimiento de 0,827 para el tamaño de ventana de 5x5, lo cual equivale a un rendimiento casi perfecto según la escala de Landis & Koch (1977).

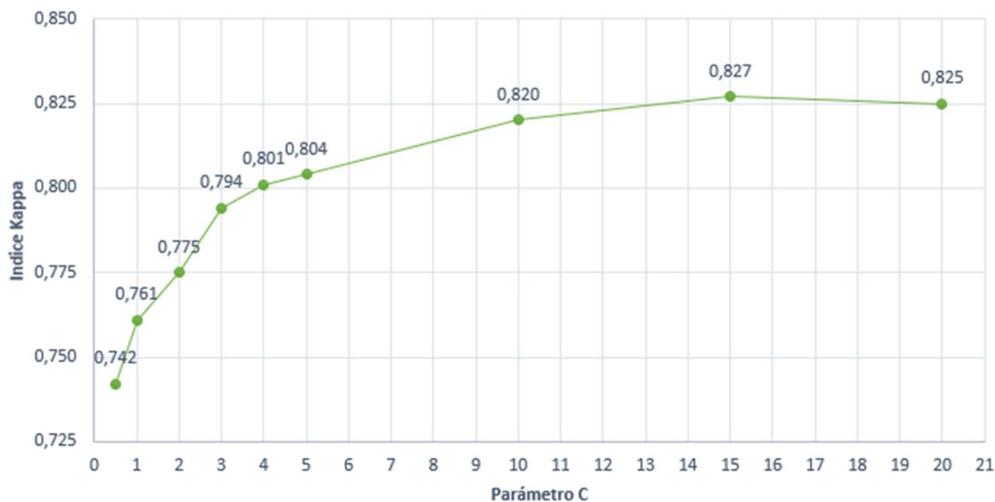


Figura 13. Análisis de sensibilidad del valor de regularización para el tamaño de ventana 5x5.

Por otra parte, en la Figura 14, se muestran los rendimientos obtenidos con la ventana de 7x7, este al igual que en el caso anterior presentó su valor más alto con un valor de C de 15. No obstante, el valor de 0,835 es mayor que el obtenido con la ventana más pequeña, específicamente un 0,97 % mayor a su contraparte. Si bien las diferencias no son abultadas, es importante tomar en cuenta que un menor tamaño de ventana requiere un mayor tiempo de preprocesamiento, por lo que, para este caso en específico, emplear una ventana de 7x7 genera mejores resultados con un menor tiempo de preprocesamiento.

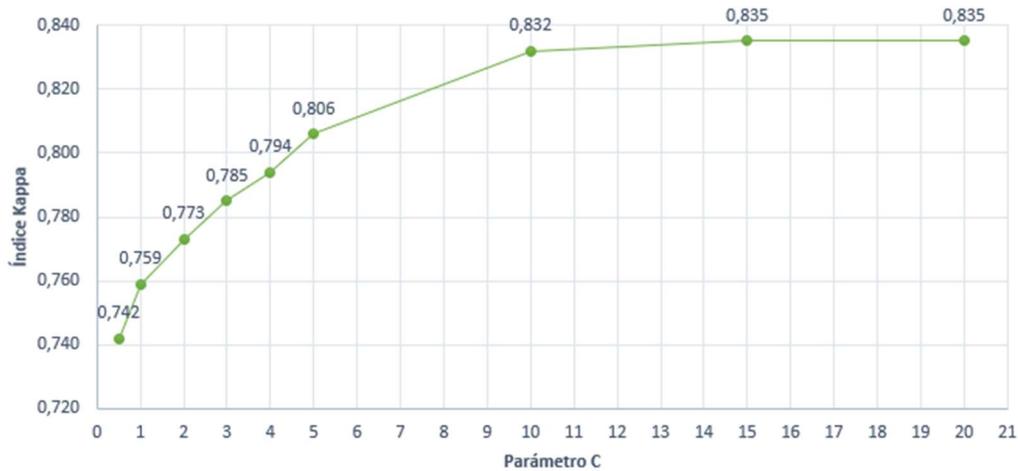


Figura 14. Análisis de sensibilidad para el valor de regularización para el tamaño de ventana 7x7.

Los resultados asociados a la medida F1 para cada una de las configuraciones de valor de regularización estudiadas, tanto para la ventana 5x5 como 7x7, se encuentran disponibles en el apéndice B. Específicamente lo componen las tablas B3 y B4.

Además de la evaluación del parámetro de regularización, se generó una matriz de confusión para cada una de las mejores configuraciones anteriormente descritas. Estas se pueden apreciar en la Figura 15 y 16, de donde se obtiene información importante con relación al rendimiento presentado en la clasificación.

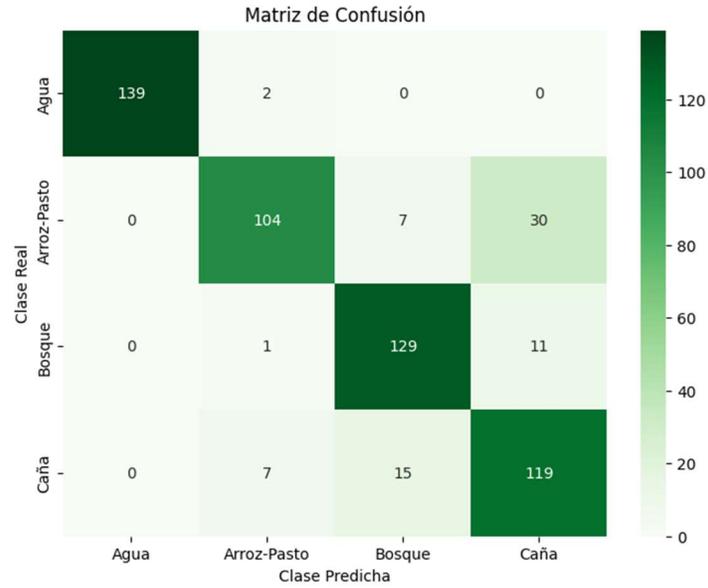


Figura 15. Matriz de confusión para tamaño de ventana 5x5.

La matriz de confusión permite visualizar el desempeño de un modelo de clasificación, donde las filas representan las clases reales y las columnas las clases predichas. Cada celda indica la cantidad de elementos clasificados en cada categoría. Los colores más oscuros reflejan una mayor cantidad de aciertos o errores, facilitando la identificación de patrones. Los valores en la diagonal (de arriba a la izquierda hacia abajo a la derecha) corresponden a las predicciones correctas, mientras más altos y oscuros sean, mejor es el desempeño del modelo.

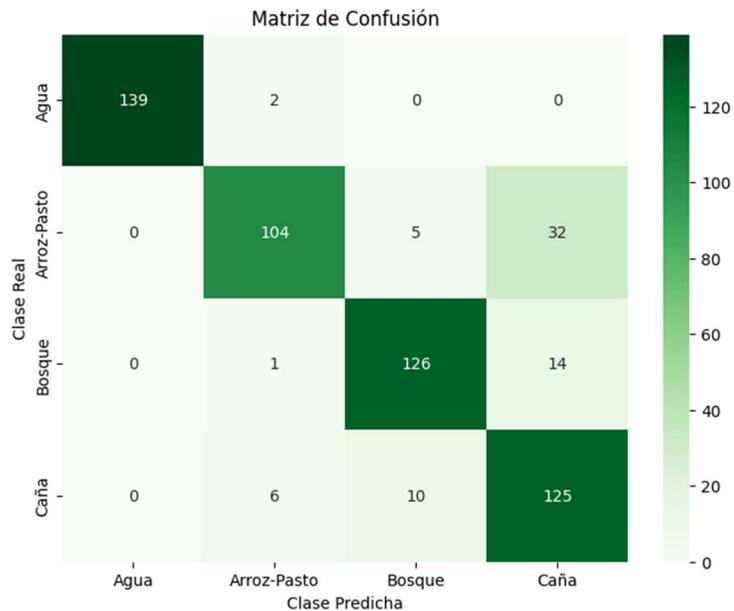


Figura 16. Matriz de confusión para tamaño de ventana 7x7.

Se observa que el mayor error de clasificación ocurre entre las clases “Arroz-Pasto” y “Caña”, lo que podría atribuirse a similitudes fenológicas en alguna de las etapas de crecimiento de estos cultivos. Durante este periodo, la vegetación puede presentar valores de retrodispersión y NDVI similares, llegando a dificultar su diferenciación.

El segundo caso con los valores más altos de error de clasificación se da entre el “Bosque” y “Caña”, los resultados se pueden deber a características propias de la zona de estudio. Ya que la mayor extensión de bosque presentada en los datos de entrenamiento y validación es la referente a la Reserva Biológica Lomas Barbudal, la cual se caracteriza por poseer un bosque tropical seco (SINAC, 2013). Esto es importante debido a que, como el análisis se realiza entre los meses de septiembre a diciembre, es posible que el bosque se encuentre en una fase de regeneración debido a la estación lluviosa, lo que puede provocar que el verdor del bosque se asemeje a la caña de azúcar. Por lo que la vegetación del bosque tropical seco y la caña de azúcar pueden presentar respuestas espectrales similares, sobre todo en los índices de vegetación como lo es el NDVI, el cual forma parte del conjunto de datos de entrenamiento.

Por otra parte, a diferencia de otros tipos de bosque que presentan doseles densos y cerrados, el bosque tropical seco generalmente posee una estructura más heterogénea, presentando en algunos casos espacios abiertos entre árboles, lo cual puede generar que firmas como la retrodispersión de las imágenes Sentinel-1 presenten similitudes, sobre todo en casos donde se emplea ventanas en su preprocesamiento.

En la Figura 17 se observan las diferencias en la clasificación al utilizar distintos tamaños de ventana. Específicamente, una ventana de 7x7 proporcionó una identificación más homogénea de las clases en la mayoría de los polígonos. Esto se debe a que ventanas de mayor tamaño suelen abarcar áreas más amplias de un mismo cultivo, lo que mejora la coherencia en la clasificación de cultivos extensivos como los presentes en el sitio de estudio.

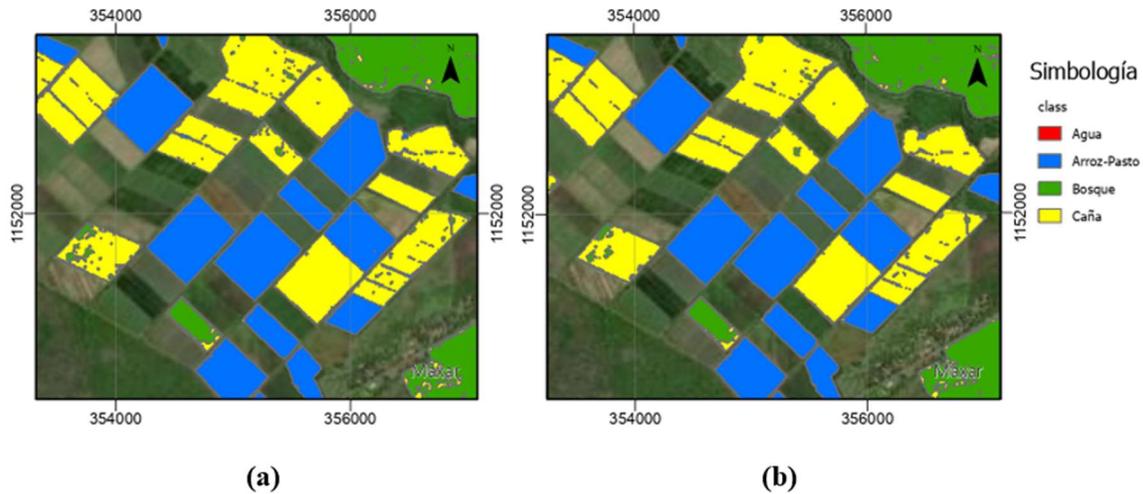


Figura 17. Clasificación de la zona de estudio para las ventanas de 5x5 (a) y 7x7 (b).

Estudios como el desarrollado por Dingle et al. (2018), han analizado las implicaciones del tamaño de ventana en la clasificación de cultivos con distintas dimensiones de parcelas agrícolas. Este trabajo destaca que el tamaño de las parcelas puede influir significativamente en la selección del filtro de moteado y el tamaño de ventana más adecuado, ya que una configuración apropiada podría suavizar en exceso o distorsionar los límites de las unidades agrícolas. Además, se observa cómo según las características del sitio de estudio en relación con el tamaño de sus parcelas varía el desempeño de las configuraciones de filtrado aplicadas. Esto evidencia la importancia de considerar particularidades espaciales del área de estudio al seleccionar los parámetros de preprocesamiento.

Para la extensión total de la zona de estudio, también se generaron mapas con el fin de visualizar el rendimiento del modelo. En el caso de la ventana de 7x7 la Figura 18 muestra la clasificación realizada, la cual presenta un comportamiento bastante homogéneo, por otra parte, la visualización de la ventana 5x5 se encuentra en la Figura B 1 disponible en el apéndice B.

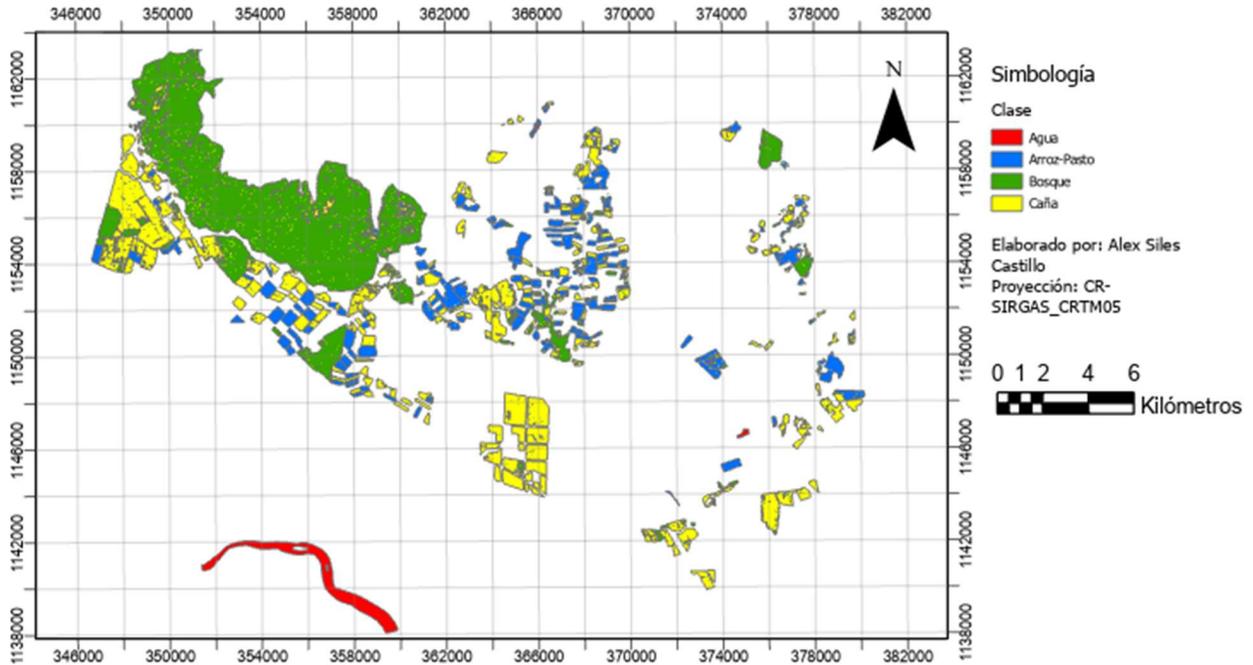


Figura 18. Clasificación de la zona de estudio para un tamaño de ventana de 7x7.

En relación con la medida F1, en la Figura 19 se identifica el rendimiento obtenido para la configuración de ventana 5x5 con un kernel de base radial y un valor C de 15, ya que esta fue la configuración con el mejor desempeño.

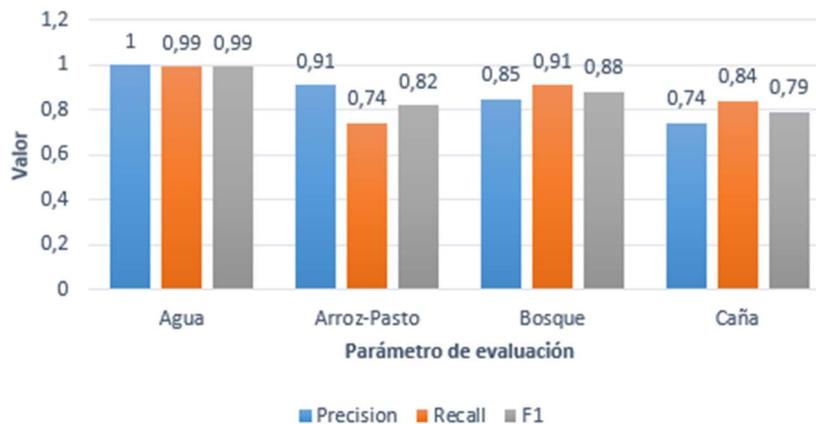


Figura 19. Desempeño de la medida F1 para la ventana de 5x5.

En relación con el escenario donde se empleó una ventana de 7x7, se obtuvieron valores altos de precision, recall y F1, los cuales fueron superiores a los generados en el caso donde se empleó una ventana menor. Esto se ilustra en la Figura 20.

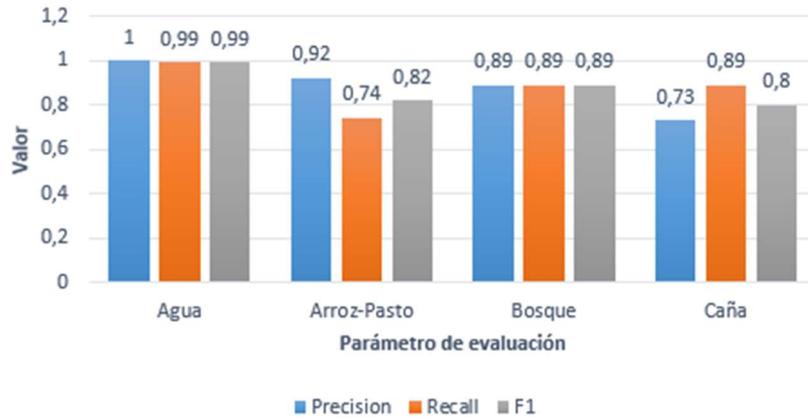


Figura 20. Desempeño de la medida F1 para la ventana de 7x7.

Estos resultados, complementan lo obtenido mediante el índice kappa en relación con el buen rendimiento que obtuvo el modelo. Además, permiten apreciar cuales fueron las clases en las que el algoritmo SVM tuvo un mejor desempeño para su clasificación. Donde la caña de azúcar pese a presentar un buen rendimiento, fue en promedio el más bajo de todas las clases estudiadas. No obstante, los valores del recall de 0,84 y 0,89 evidencian que un alto porcentaje de los puntos con caña de azúcar fueron identificados correctamente.

Seguido a la caña de azúcar, la clase Arroz-Pasto obtuvo el segundo rendimiento más bajo, con un F1 de 0,82, lo cual puede suceder debido a que se combinan dos coberturas de suelo que pueden presentar diferencias fenológicas las cuales afectan los datos que utiliza el modelo para realizar la clasificación.

Siguiendo esta línea, se genera la necesidad de conocer cuál fue el impacto que se obtuvo al unir las clases de “Arroz” y “Pasto”, motivo por el cual se desarrolló un análisis de sensibilidad en el cual se evaluó el rendimiento del modelo con estas clases separadas. Debido a que el tamaño de ventana 7x7 ha presentado los mejores rendimientos, el nuevo set de datos se generó con este tamaño de ventana.

Como se muestra en las Figuras 21 y 22, la separación de las clases de arroz y pasto no afecta el desempeño del modelo en la identificación de caña de azúcar, sin embargo, sí impacta su rendimiento general. Anteriormente, se indicó que el índice kappa para la configuración con las clases combinadas fue de 0,835, mientras que, al dividir las clases, este valor disminuyó a 0,784.

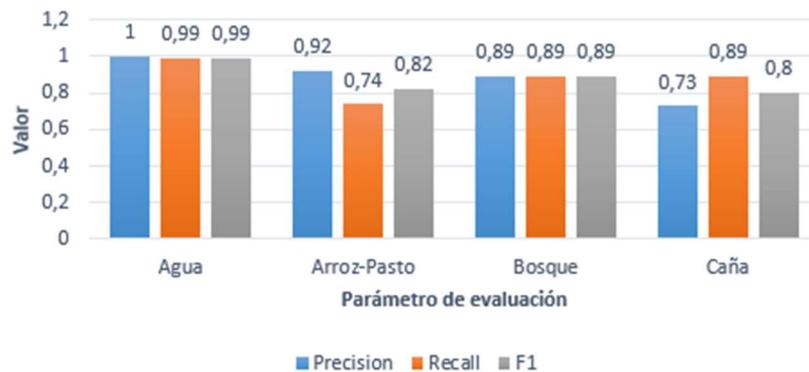


Figura 21. Desempeño del modelo SVM en los escenarios con las clases de arroz y pasto combinadas

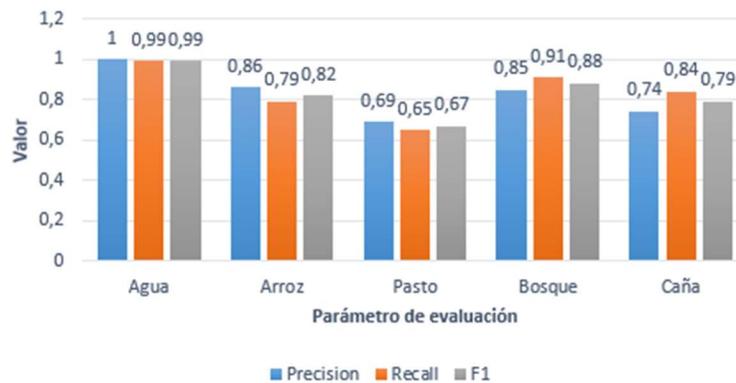


Figura 22. Desempeño del modelo SVM en los escenarios con las clases de arroz y pasto separadas.

Esta reducción del rendimiento general del modelo se debe a que la clase pasto presenta un desempeño significativamente inferior en comparación con las demás clases analizadas. Ya que presenta valores de 0,69, 0,65 y 0,67 para las métricas de precisión, recall y F1, respectivamente.

El bajo rendimiento presentado en la clase pasto se puede deber a diversos factores, tomando esto a consideración, se procedió a examinar el comportamiento de la retrodispersión de todas las clases a lo largo del periodo de estudio. Para esto, se generó un gráfico el cual muestra el promedio y la desviación estándar para las nueve imágenes Sentinel-1 que conforman el set de datos estudiado. Esto se observa en la Figura 23.

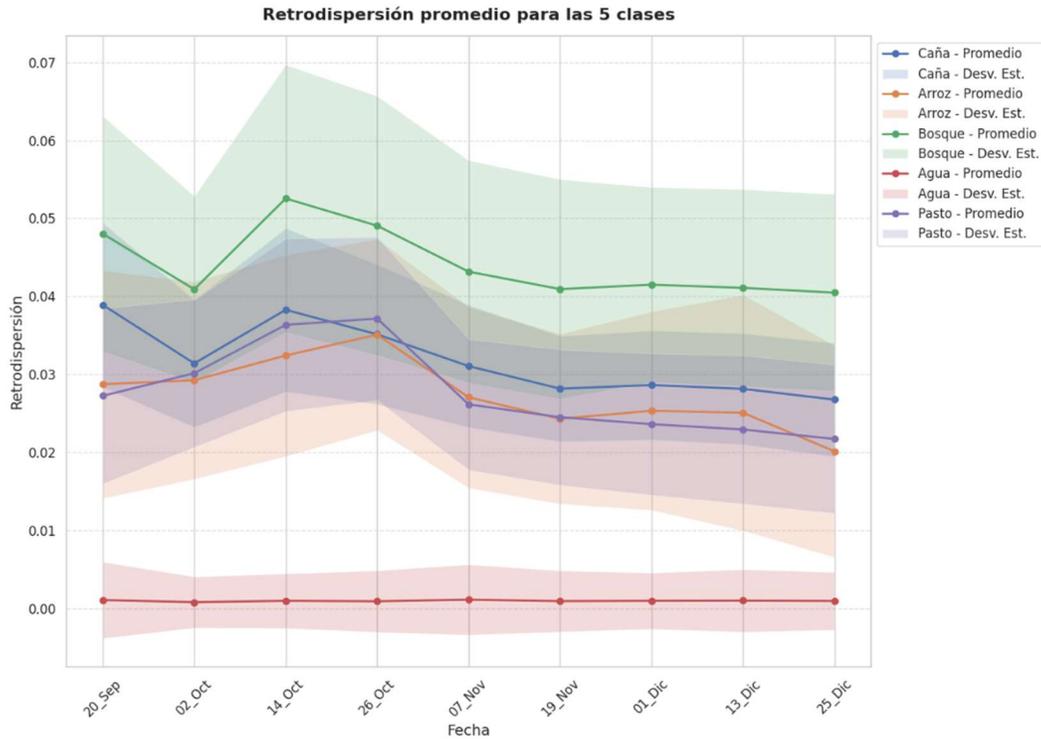


Figura 23. Comportamiento promedio de la retrodispersión para cada una de las clases estudiadas.

El comportamiento observado para las clases de arroz y pasto muestra una alta similitud, reflejada en valores muy cercanos en las imágenes obtenidas el 20 de septiembre, 02 de octubre, 07 y 19 de noviembre. Esta similitud podría atribuirse a las características fenológicas de ambos cultivos, las cuales no presentan diferencias significativas a nivel de los datos obtenidos mediante imágenes Sentinel-1.

En el DRAT, el cultivo de arroz se desarrolla en dos ciclos productivos: el primero, con siembras realizadas entre los meses de julio y agosto, y el segundo, entre diciembre y enero (Morales et al., 2018). Por su parte, los pastos en la región pueden presentar variabilidad en las fechas de siembra, sin embargo, se recomienda establecerlos durante la temporada lluviosa, siendo común su siembra en el mes de julio (Morales et al., 2003). Esta coincidencia en los períodos de establecimiento podría ser un factor que dificulte la diferenciación entre ambos cultivos en los análisis basados en imágenes Sentinel-1, afectando el desempeño del algoritmo de clasificación.

En relación con los datos de NDVI que se emplearon para el entrenamiento y validación del modelo, se realizaron gráficos de violín para conocer el comportamiento de este índice multispectral para las clases estudiadas.

En la Figura 24 se encuentra la distribución que presentó el NDVI el 30 de noviembre para todas las clases de interés. Como se puede observar las clases de agua y bosque presentan valores que permiten diferenciarlas fácilmente de las demás. Esto se debe a la propia naturaleza del índice multispectral, el cual se relaciona con la actividad fotosintética presente en el área estudiada. En el caso del agua, al no presentar vegetación, los valores de NDVI son muy cercanos a cero. Sucede lo contrario en los bosques, ya que al tener una estructura vegetal más robusta que las demás clases, presenta valores mayores.

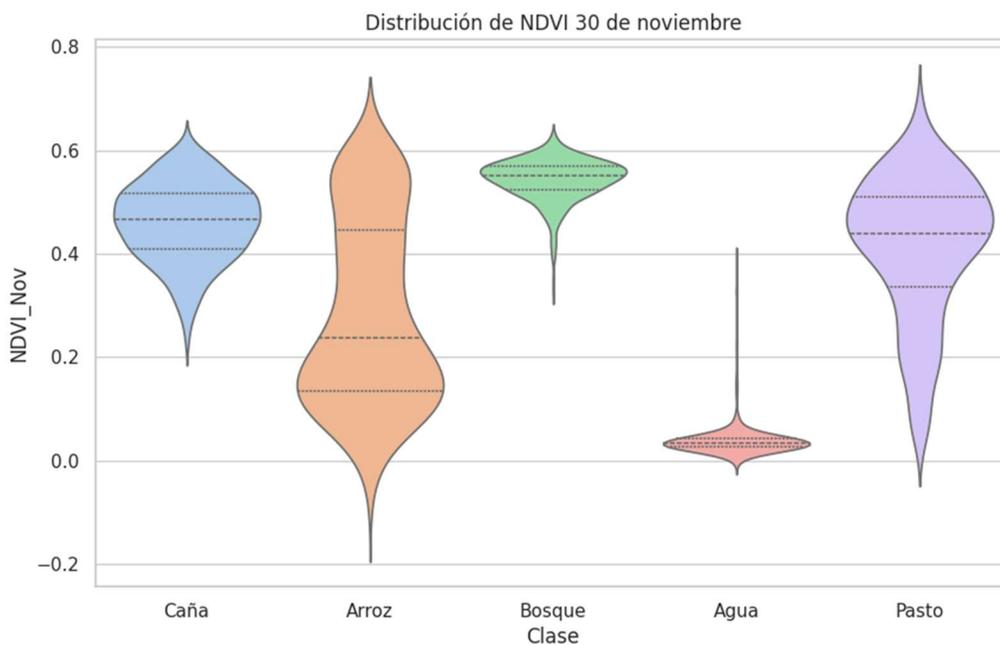


Figura 24. Distribución del NDVI para el producto Sentinel-2 del 30 de noviembre

No obstante, para las clases de arroz, pasto y caña los valores obtenidos presentan cierta similitud. De estas clases destaca el arroz el cual posee un rango muy amplio de NDVI, esto principalmente se puede atribuir a el tipo de riego del cultivo, ya que, en el caso de presentar un riego por inundación, es posible que el espejo de agua cause cierta interferencia en los datos presentados.

Además, en los casos donde se presente arroz de secano, la estructura vegetal puede ser menos robusta, lo cual genere una disminución en el NDVI presentado.

En el caso de los pastos también se observa un rango amplio de NDVI, en este caso se puede deber a los periodos de cosecha. El pasto posee diversas fechas de cosecha según el desarrollo fenológico que este disponga, generalmente se realiza cada 45 días, por lo que es posible que en las parcelas estudiadas algunas zonas se encontraran cosechadas lo que redujo el NDVI y amplió el rango de valores (Morales et al., 2003).

En relación con la caña de azúcar, el conjunto de datos obtenidos no presenta alta variabilidad en comparación con las clases de arroz y pasto, esto se debe a que el cultivo se da en condiciones más uniformes, donde la variación entre parcelas de los periodos de siembra y cosecha es mínima, esto supone una ventaja al momento de realizar un proceso de clasificación.

La Figura 25 permite apreciar el comportamiento para el 25 de diciembre, donde al igual que en el caso anterior se presenta una similitud en el rango de NDVI presentado para las clases de arroz y pasto. Como se mencionó anteriormente, aunque el DRAT cuenta con disponibilidad de riego durante todo el año, algunos productores optan por el cultivo de arroz de secano. La ausencia de riego en estos sistemas productivos puede generar un comportamiento espectral más similar al de los pastos, lo que incrementa la dificultad para su correcta clasificación.

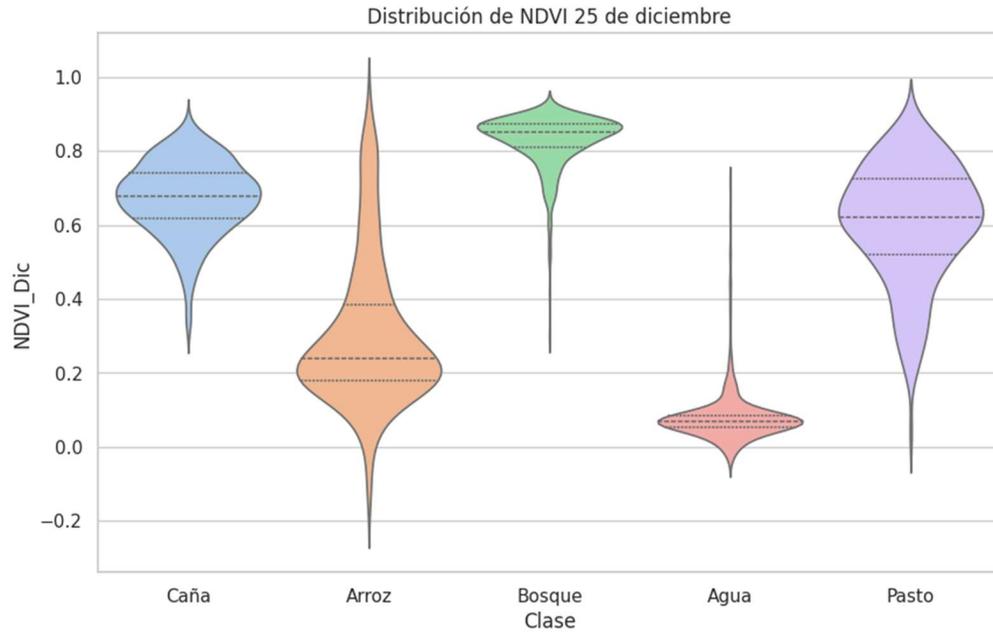


Figura 25. Distribución del NDVI para el producto Sentinel-2 del 25 de diciembre.

4.5. Desempeño del modelo en diferentes fechas

Los resultados presentados en la sección anterior muestran que existen diferencias en los valores de retrodispersión y NDVI a lo largo del tiempo, lo cual puede llegar a repercutir en el rendimiento del algoritmo. Por este motivo se procedió a seleccionar tres imágenes Sentinel-1 y los dos NDVI obtenidos, para realizarles un análisis monotemporal y observar el comportamiento del modelo SVM, lo que permitirá analizar de forma más detallada la interacción entre las clases de arroz, pasto y caña.

Para esta prueba, se empleó un kernel de base radial y un valor C de 15, además de emplear un análisis monotemporal para imágenes SAR con una ventana de 7x7, los resultados obtenidos se encuentran en las Figuras 26, 27 y 28. En un primer análisis, la aplicación de un enfoque monotemporal resultó en una reducción del rendimiento del modelo.

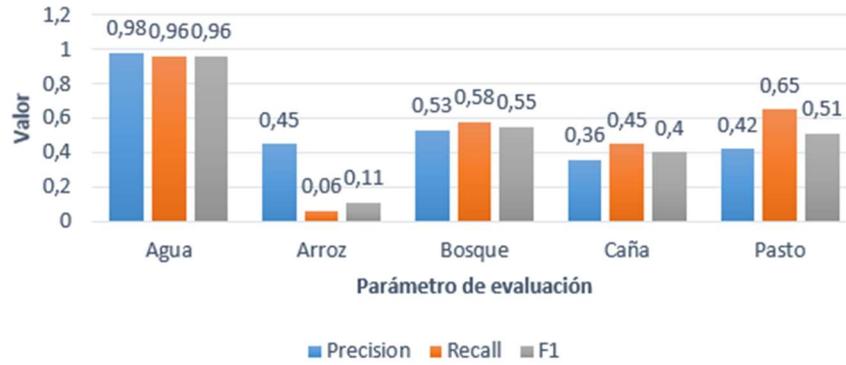


Figura 26. Rendimiento del algoritmo para el 20 de septiembre del 2019 mediante análisis monotemporal de imágenes Sentinel-1 y ventana 7x7

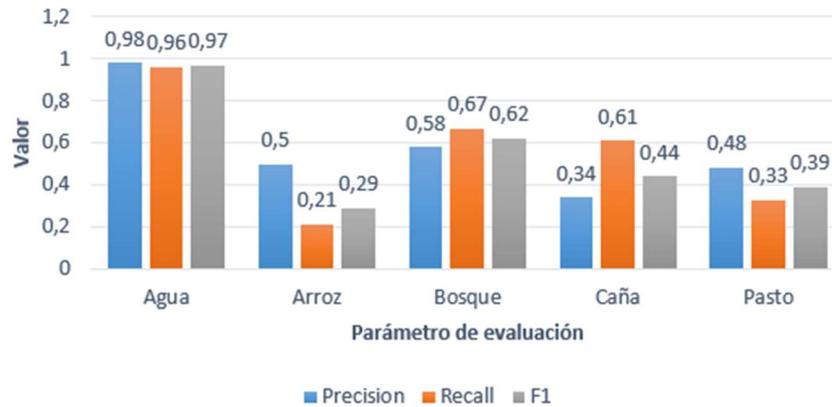


Figura 27. Rendimiento del algoritmo para el 19 de noviembre del 2019 mediante análisis monotemporal de imágenes Sentinel-1 y ventana 7x7

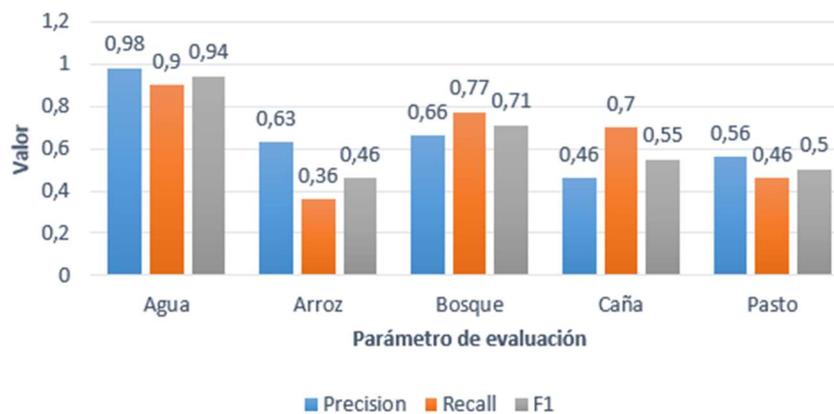


Figura 28. Rendimiento del algoritmo para el 25 de diciembre del 2019 mediante análisis monotemporal de imágenes Sentinel-1 y ventana 7x7

Esta disminución fue más pronunciada en la clase de arroz, donde la medida F1 alcanzó un valor de 0,11 con una precisión de 0,45 y una sensibilidad de 0,06 para la imagen obtenida el 20 de septiembre. En cuanto al índice kappa, los valores obtenidos fueron 0,504, 0,570 y 0,660 para las fechas del 20 de septiembre, 19 de noviembre y 25 de diciembre, respectivamente.

En relación con la caña de azúcar, también se presentó una reducción en su rendimiento de identificación, no obstante, no fue tan drástico como en el caso del arroz. En este escenario, la caña de azúcar obtuvo un F1 de 0,55 mientras que en el análisis multitemporal el valor reportado fue de 0,79.

En la Figura 29, se encuentra de forma visual la distribución presentada por las diversas clases para el análisis realizado con la imagen del 25 de diciembre del año de estudio. Es importante recalcar que a este producto se le aplicó el filtro Gamma MAP con un tamaño de ventana de 7x7. En caso de querer consultar la distribución de las otras fechas analizadas, estas se encuentran en el apéndice B en las Figuras B4 y B5. Se empleó este producto como ejemplo debido a que en la misma fecha se obtuvo una imagen Sentinel-2, lo que permitirá realizar comparaciones futuras, además fue donde se presentó un mayor rendimiento en el análisis monotemporal. Dicha distribución presentada posee un comportamiento bastante heterogéneo, indicando que el rendimiento de la clasificación fue bajo, concordando con las métricas evaluadas.

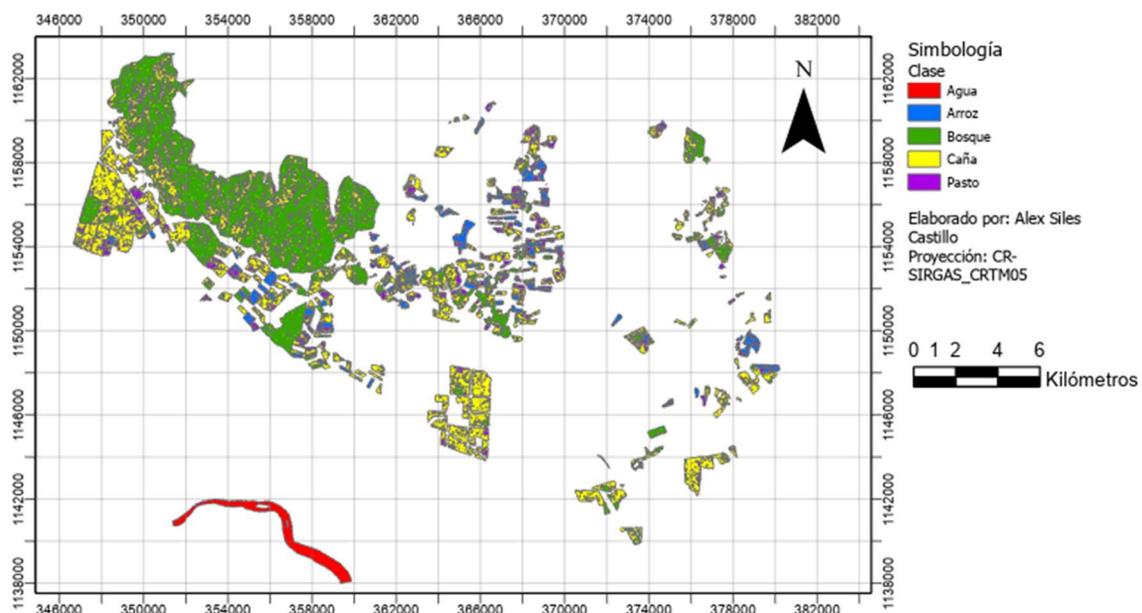


Figura 29. Clasificación de la zona de estudio para el producto Sentinel-1 del 25 de diciembre del 2019.

En el caso del NDVI, al igual que en caso anterior, el análisis monotemporal provoca una disminución del rendimiento de clasificación, generando un índice kappa de 0,505 y 0,580 para el NDVI del 30 de noviembre y 25 de diciembre, respectivamente. En relación con la medida F1 y parámetros relacionados, los resultados se encuentran en las Figuras 30 y 31.

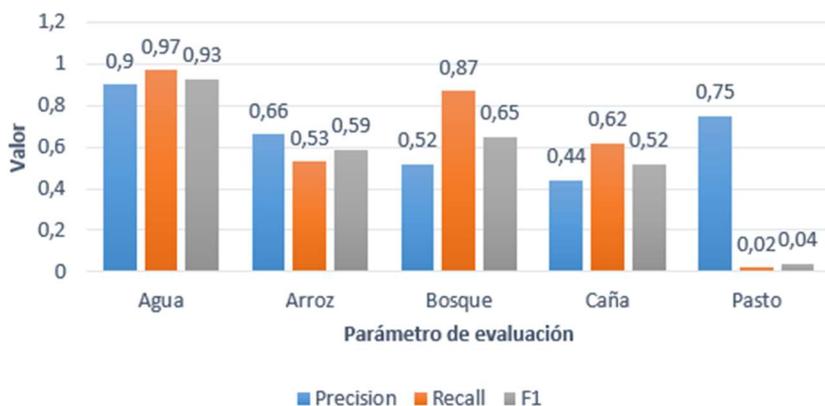


Figura 30. Rendimiento del algoritmo para el 30 de noviembre del 2019 mediante análisis monotemporal de imágenes de NDVI.

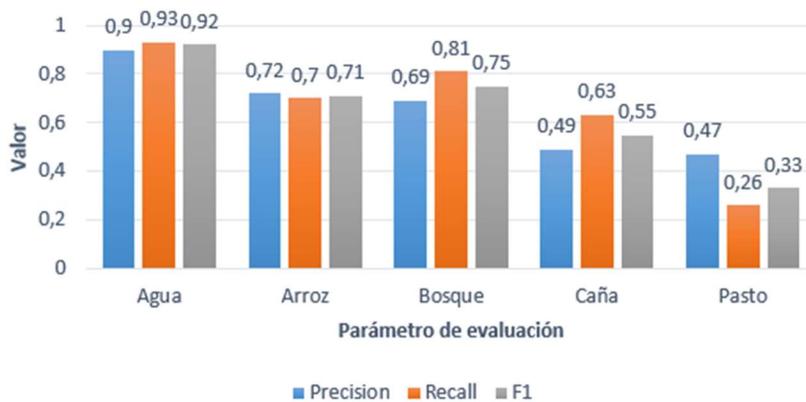


Figura 31. Rendimiento del algoritmo para el 25 de diciembre del 2019 mediante análisis monotemporal de imágenes de NDVI.

Para el NDVI obtenido el 30 de noviembre del año de estudio, se destaca el comportamiento de la clase pasto, donde pese a tener un precisión de 0,75, el recall fue de solo 0,02, lo que generó un F1 de 0,04. Esto evidencia que el modelo tiene dificultades para identificar esta clase tomando en cuenta las características suministradas para la fecha en la cual se realiza el análisis.

La caña de azúcar presentó un rendimiento similar en comparación con el análisis monotemporal empleando imágenes Sentinel-1, donde para los dos escenarios evaluados el 25 de diciembre del

2019 se presentó el mismo valor de F1. En la Figura 32 se encuentra la distribución espacial para el producto de NDVI del 25 de diciembre, por otra parte, la distribución del 30 de noviembre se encuentra en el apéndice B Figura B6.

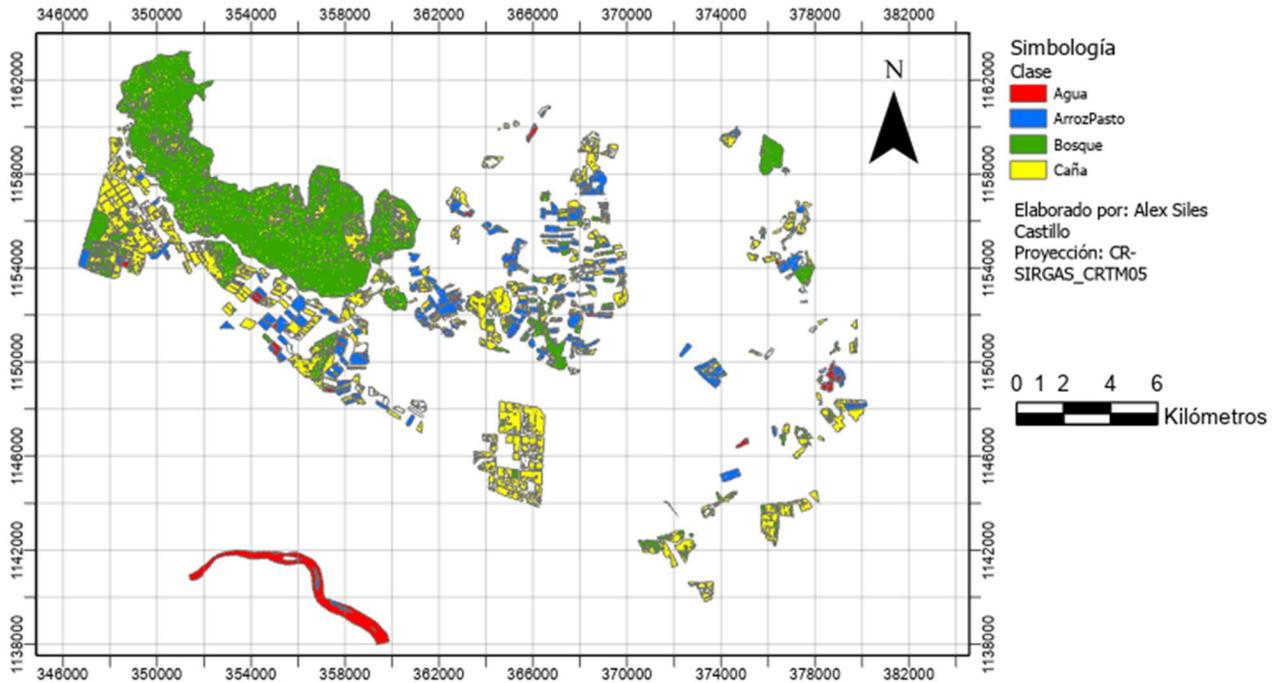


Figura 32. Clasificación de la zona de estudio para el análisis de NDVI del 25 de diciembre del 2019.

El análisis monotemporal de las imágenes Sentinel-1 y el NDVI, evidencia una tendencia en la cual, a medida que avanza el tiempo, los rendimientos del modelo presentan una mejora considerable. Esto se complementa con lo presentado por Jiang et al. (2019), donde se realizó un estudio sobre la determinación del periodo en el cual realizar la clasificación de caña de azúcar, lo cual generó como conclusión que el rendimiento del modelo en los últimos tres meses previo a la cosecha es mayor. Para esto el autor utilizó imágenes SAR, NDVI y una polarización VH.

Como se especificó anteriormente, los cultivos de arroz y pasto tienen fechas de siembra similares, generalmente alrededor del mes de julio, dependiendo de la región. Considerando que el ciclo fenológico del arroz desde la germinación hasta la cosecha dura aproximadamente entre 105 y 125 días, dependiendo de la variedad, es posible afirmar que su desarrollo fenológico alcanza su máximo alrededor del mes de diciembre (Morales et al., 2018).

En el caso de la caña de azúcar, este cultivo requiere altas temperaturas durante su período de crecimiento y temperaturas más bajas durante la maduración. En Costa Rica, la cosecha de caña de azúcar, conocida como zafra, suele realizarse en la época seca, que abarca desde diciembre hasta abril (Flores, 2017). Esto indica que, para diciembre, la caña también ha alcanzado una etapa avanzada en su ciclo de crecimiento.

Este desarrollo fenológico óptimo, de los cultivos analizados, en el mes de diciembre repercute directamente en el rendimiento del modelo de clasificación. Durante este período, las diferencias en los valores de NDVI y la retrodispersión SAR entre las distintas clases de cultivos son más pronunciadas, lo que facilita la discriminación precisa de los cultivos de interés. Estudios han demostrado que la combinación de datos SAR y ópticos mejora la discriminación de cultivos debido a la captura de características específicas durante sus fases fenológicas clave (Marini, 2021).

Este análisis del desempeño del modelo bajo un escenario monotemporal, también permite obtener nuevas conclusiones clave sobre las limitaciones de este enfoque. Los resultados obtenidos evidencian que una estrategia multitemporal ofrece un desempeño superior en la clasificación de cultivos, reflejando en valores más altos de índice kappa (0,835) en comparación con el análisis monotemporal (0,676).

Esta diferencia se debe a que la capacidad de identificación de los cultivos puede verse afectada en diversos puntos del ciclo de producción, dependiendo de las condiciones específicas de la zona de estudio. En un análisis monotemporal, se captura solo un instante de tiempo, lo que puede no ser suficiente para distinguir cultivos con patrones de crecimiento similares en esa fecha en particular. En contraste, un enfoque multitemporal permite monitorear la evolución de la retrodispersión y el NDVI a lo largo del tiempo, proporcionando una representación más completa del desarrollo fenológico de los cultivos facilitando su identificación con una mayor precisión.

Por lo tanto, la incorporación de múltiples imágenes a lo largo del ciclo de crecimiento no solo mejora la capacidad de clasificación que posee el modelo, sino que también permite comprender mejor las variaciones espectrales y estructurales de los cultivos en respuesta a factores ambientales y agronómicos.

Pese a que en diversos estudios un enfoque monotemporal ha presentado rendimientos más altos como los obtenidos por Prashanthika et al. (2019) y Saadat et al. (2011), donde se presentaron valores de 0,75 y 0,95, respectivamente. Esto dependerá de las características específicas de la zona de estudio y las bandas o índices empleados para realizar la clasificación.

Tomando esto a consideración, se reitera la importancia de desarrollar este tipo de investigaciones con el fin de profundizar en mayor medida sobre los alcances que pueden tener todo este tipo de herramientas a nivel nacional. Para así, poder validarlas y emplearlas en diversas aplicaciones tanto agrícolas como para la toma de decisiones en términos de seguridad alimentaria.

4.6. Comparación de métodos con y sin NDVI

Los resultados presentados hasta el momento demuestran la efectividad de la metodología aplicada para la clasificación de la caña de azúcar dentro del DRAT. Sin embargo, presenta la oportunidad de profundizar más en las posibles variaciones que puede tener la metodología con el fin de conocer su impacto bajo diferentes escenarios.

Como se mencionó anteriormente, para el desarrollo del conjunto de datos solo se utilizaron dos productos Sentinel-2, las cuales corresponden al 30 de noviembre y 25 de diciembre del 2019. Por otra parte, imágenes Sentinel-1 se emplearon nueve, lo cual puede llegar a generar intriga sobre si la incorporación de solo dos capturas ópticas genera un impacto en el rendimiento del modelo.

Tomando en consideración lo anterior, se generó un set de datos con los mismos puntos empleados en los análisis anteriores, no obstante, no se incorporó el valor de NDVI para los datos de entrenamiento y validación. El rendimiento de esta configuración se puede observar en las Figuras 33 y 34 para el tamaño de ventana 5x5.

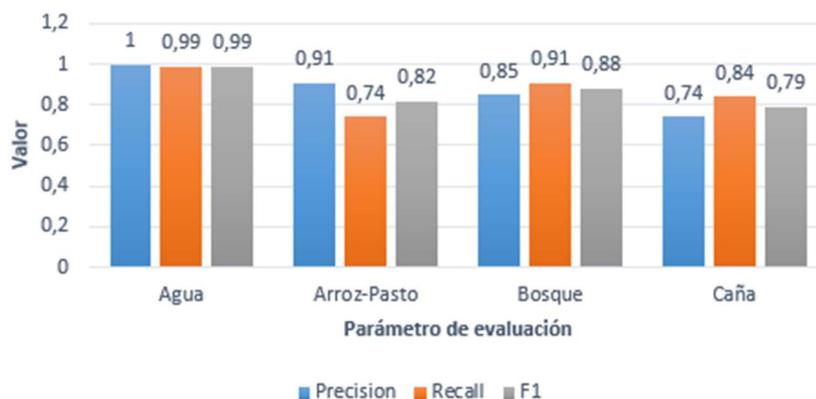


Figura 33. Rendimiento con NDVI 5x5



Figura 34. Rendimiento sin NDVI 5x5.

Como se observa, en el caso de la ventana 5x5, el desempeño bajó para todas las clases exceptuando para la clase “Agua”. En el caso de la caña de azúcar, el rendimiento bajó alrededor de 16,45 %. La estabilidad en relación con la clase agua se debe a que la retrodispersión generada por el agua presenta valores muy cercanos a cero, ya que el agua no permite reflejar la señal enviada por los satélites SAR. Esta característica hace que las imágenes de radar sean comúnmente utilizadas al momento de identificar inundaciones o cuerpos de agua (Jiménez & Guerrero, 2023).

En cambio, en el caso de las demás clases los valores de retrodispersión pueden llegar a ser similares entre sí. Por este motivo, el NDVI busca respaldar al modelo al incorporar más características que pueden provocar una correcta diferenciación entre clases. En el caso de la ventana 7x7, el rendimiento fue similar, donde el rendimiento al momento de identificar la caña se redujo en un 10 %, esto se aprecia en las Figuras 35 y 36.

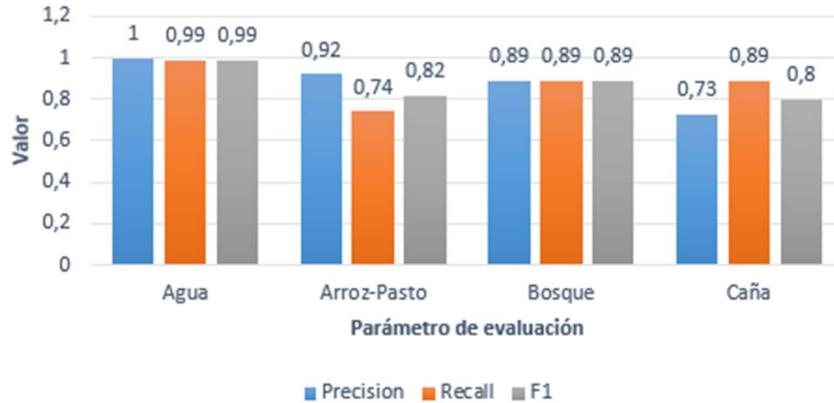


Figura 35. Rendimiento con NDVI 7x7.

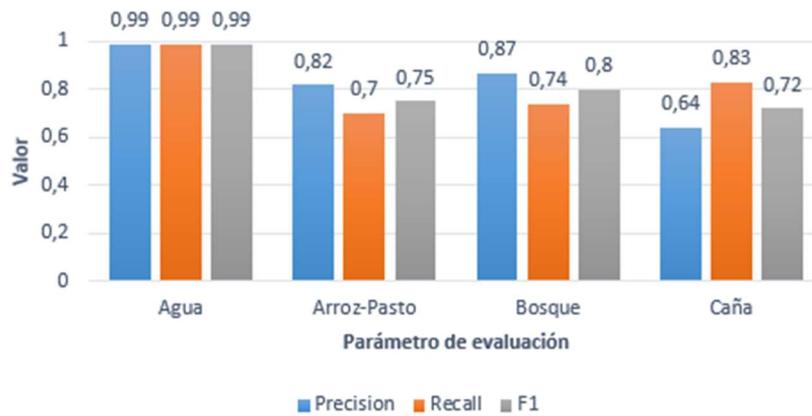


Figura 36. Rendimiento sin NDVI 7x7

En cuanto al índice kappa, los valores obtenidos para los conjuntos de datos sin NDVI fueron inferiores a aquellos que sí incluían este índice multispectral. Esto indica que, a pesar de emplear un número reducido de imágenes ópticas, su incorporación mejora el rendimiento en comparación con su ausencia. La Figura 37 muestra cómo, en ambas configuraciones de ventana, el desempeño del modelo disminuyó al excluir el NDVI y la Figura 38 se logra apreciar la distribución espacial para el análisis multitemporal sin NDVI con tamaño de ventana 7x7.

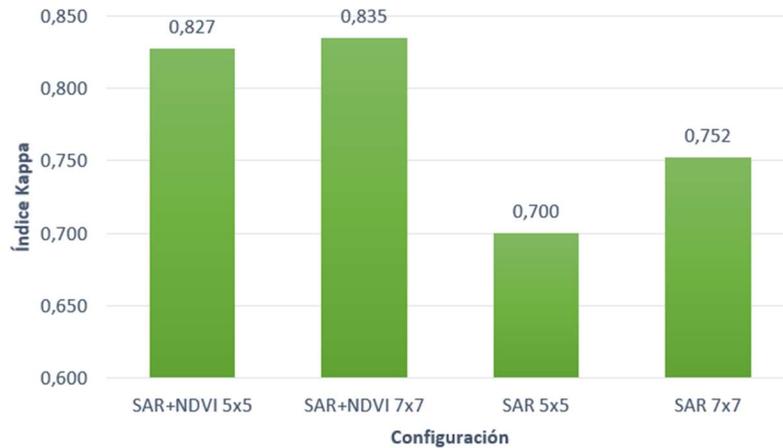


Figura 37. Índice Kappa del algoritmo SVM al excluir el NDVI del conjunto de datos.

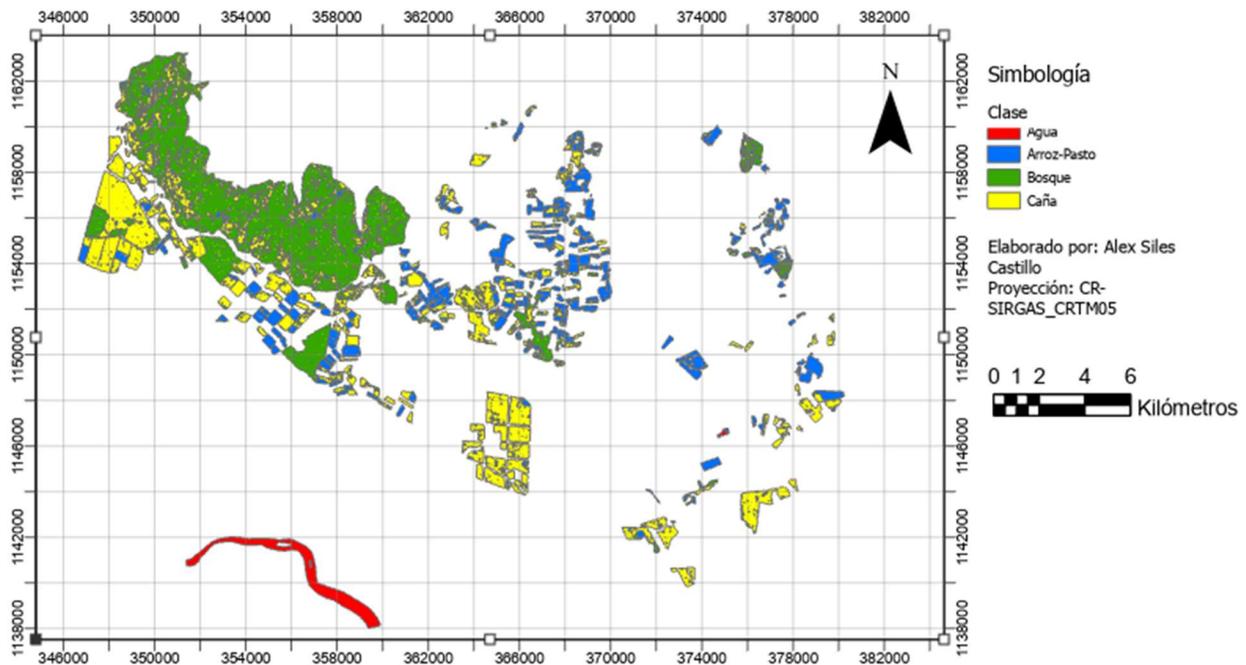


Figura 38. Clasificación de la zona de estudio para el análisis multitemporal sin NDVI con una ventana de 7x7.

Como se muestra en la Figura 38, la distribución de las clases identificadas es más heterogénea que en los casos donde el conjunto de datos se componía de productos Sentinel-1 y NDVI proveniente de Sentinel-2 (Figura 18). Esto era de esperar debido a los rendimientos de las métricas de índice kappa y medida F1 presentados eran menores que en las configuraciones que disponían

de los productos Sentinel-1 y Sentinel-2. En el caso de la ventana 5x5, esta distribución espacial del análisis multitemporal sin NDVI, se encuentra en el apéndice B Figura B7.

Con los resultados presentados en relación con la eliminación del NDVI del set de datos, se aprecia que estos aportan en gran medida al rendimiento del modelo, no obstante, se presenta la oportunidad de analizar qué pasaría en el caso de que el conjunto de datos empleado solo se componga de las dos imágenes Sentinel-2 disponibles y no se consideren las imágenes Sentinel-1.

Por este motivo, se realizó un análisis multitemporal compuesto solamente por las imágenes que contienen el NDVI para las fechas del 30 de noviembre y el 25 de diciembre, al cual se le aplicó un kernel de base radial y un valor C de 15. En relación con el Índice kappa, esta configuración arrojó un valor de 0,742, el cual es alto si se compara con los casos anteriores en los que no se aplicaba este índice multiespectral.

Si se observa la Figura 39, se evidencia de una forma más detallada el rendimiento del modelo donde para todas las clases se da un desempeño positivo. El cual, pese a ser menor que en el caso de incorporar imágenes Sentinel-1 más el NDVI de Sentinel-2, sigue estando en un rango donde se puede concluir que la clasificación realizada es sustancial. En el caso de la caña de azúcar la reducción presentada en relación con el desempeño del algoritmo bajo esta configuración se redujo en un 10 %. Estas observaciones evidencian el potencial de la aplicación de índices multiespectrales como lo es el NDVI para la clasificación de cultivos debido a que emplean numerosas bandas espectrales para su obtención.

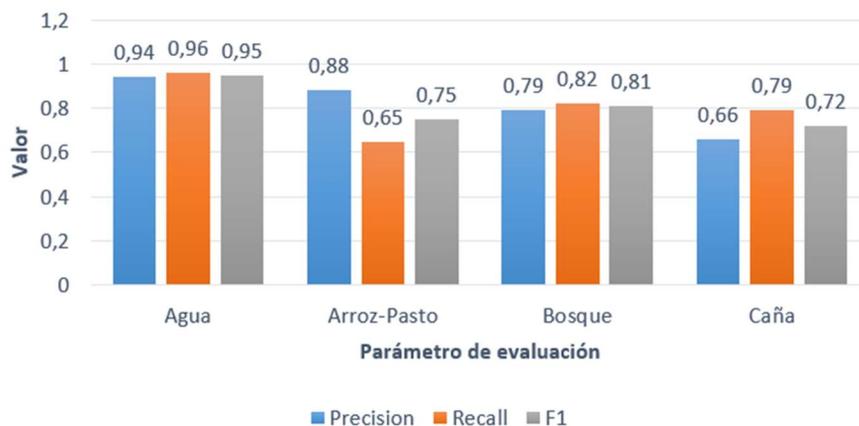


Figura 39. Métricas de desempeño del modelo utilizando únicamente datos de NDVI.

En la Figura 40, se encuentra de forma gráfica el rendimiento del modelo al clasificar empleando solo los dos productos que contenían información del NDVI. Como se puede apreciar, la clasificación es un poco más heterogénea que en los casos donde se combinó los dos tipos de productos (Figura 18), no obstante, el rendimiento es exitoso. Además, se observa la presencia de más píxeles de la clase agua, esto se puede deber al cultivo de arroz. Anteriormente se mencionó que, en el caso de presentar un riego por inundación en este cultivo, el espejo de agua podría generar una deficiencia en la clasificación del algoritmo empleado.

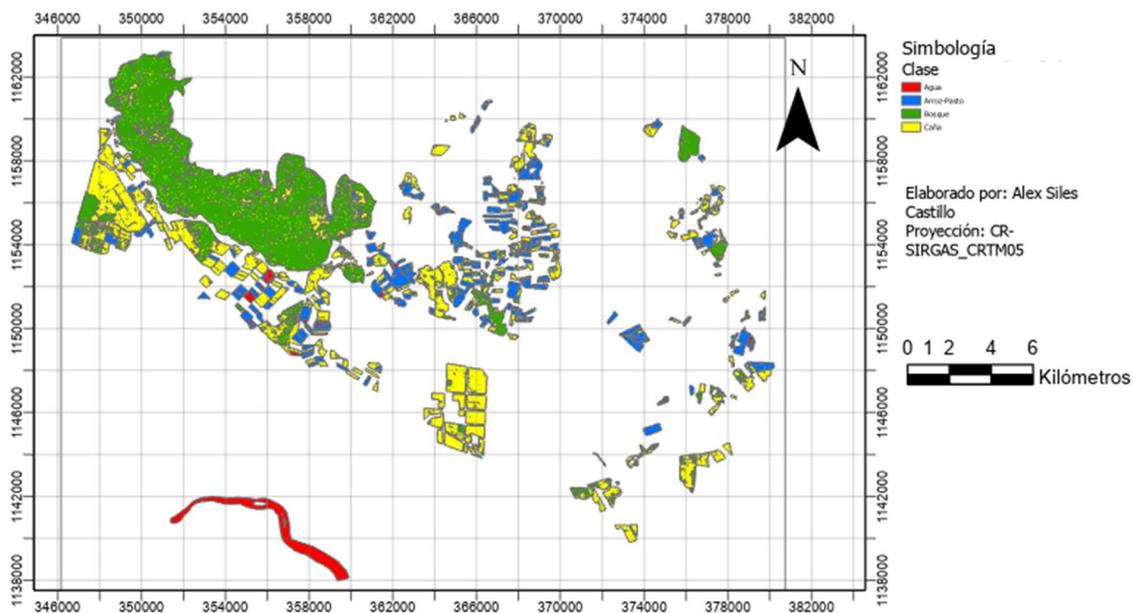


Figura 40. Clasificación de la zona de estudio para el análisis multitemporal solo empleando NDVI.

Prashanthika et al. (2019) obtuvo valores de índice kappa de hasta 0,75 lo que refuerza los rendimientos presentados en esta investigación para el análisis de NDVI son óptimos para la clasificación de cultivos bajo estas condiciones.

La tabla 3 muestra las configuraciones con los valores más altos reportados para los modelos. En el caso de desear consultar todos los valores obtenidos en cuanto a las métricas de la medida F1 e Índice kappa para todos los escenarios evaluados, estos se encuentran resumidos en el apéndice B Tabla B 14.

Tabla 3. Mejores rendimientos presentados para las configuraciones analizadas.

Producto	Tamaño de ventana	Clases	Metodología	Valor de regularización	Kernel	Medida F1	Índice kappa
SAR + NDVI	5x5	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	15	Base Radial	0,87	0,827
SAR + NDVI	7x7	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	15	Base Radial	0,88	0,835
SAR + NDVI	7x7	Caña, Arroz, Pasto, Bosque y Agua	Multitemporal	15	Base Radial	0,83	0,784
SAR	7x7	Caña, Arroz-Pasto, Bosque y Agua	Monotemporal 25 Dic	15	Base Radial	0,74	0,660
NDVI	-	Caña, Arroz-Pasto, Bosque y Agua	Monotemporal 25 Dic	15	Base Radial	0,76	0,676
SAR	7x7	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	15	Base Radial	0,81	0,752
NDVI	-	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	15	Base Radial	0,81	0,742

El escenario con mejor rendimiento fue el planteado inicialmente, el cual se componían de los nueve productos Sentinel-1 y los dos Sentinel-2, con un kernel de base radial y un valor de regularización de 15. El índice kappa de esta configuración obtuvo un valor de 0,835 el cual según la escala de Landis & Koch (1977) corresponde a un arreglo casi perfecto.

No obstante, los demás escenarios estudiados presentan rendimientos sustanciales con valores de índice kappa mayores a 0,660. Esto permite conocer los alcances de esta tecnología bajo diferentes condiciones, las cuales se pueden llegar a presentar según la zona de estudio. Otra observación importante es que en las evaluaciones donde se realizaron análisis monotemporales, el rendimiento fue mayor empleando productos Sentinel-2 que su contraparte Sentinel-1, donde se obtuvieron valores de índice kappa de 0,676 y 0,660, respectivamente.

Por otra parte, en la Tabla 3 también se observa como la unión de las clases de arroz y pasto aumenta el rendimiento general de modelo, obteniendo un índice kappa de 0,835, el cual comparado con el 0,784 de su contraparte con dichas clases separadas, es en un 6,11 % superior.

5. CONCLUSIONES

- El filtro Gamma MAP demostró ser una herramienta efectiva para la reducción de ruido granular en imágenes SAR, donde la configuración de ventana de 5x5 obtuvo mejores resultados en términos de reducción de ruido sin pérdida significativa de información.
- La combinación de imágenes SAR de Sentinel-1 y ópticas de Sentinel-2 permitió mejorar la clasificación del cultivo de caña de azúcar alrededor de un 18 %, destacando la importancia de emplear información multiespectral y de radar en estudios clasificación de cultivos.
- El uso de un enfoque multitemporal incrementó significativamente el rendimiento del modelo de clasificación, evidenciado por un aumento de alrededor de 19 % del Índice kappa en comparación con el enfoque monotemporal, lo que resalta la relevancia de considerar la variabilidad temporal en la identificación de cultivos.
- El clasificador Support Vector Machine con un kernel de base radial presentó el mejor rendimiento en la identificación de caña de azúcar, logrando un índice kappa de 0,835 para una ventana de 7x7, lo que indica que este algoritmo es adecuado para la clasificación de cultivos con características espectrales y de retrodispersión similares.
- El estudio demostró la viabilidad de utilizar metodologías basadas en imágenes de libre acceso para la identificación de cultivos, lo que sugiere su aplicabilidad en estudios a mayor escala o en diferentes tipos de cultivos dentro del DRAT y otras regiones agrícolas.

6. RECOMENDACIONES

- Ampliar el análisis multitemporal incluyendo un mayor número de imágenes Sentinel-1 y Sentinel-2 para evaluar el impacto de la variabilidad espectral en la precisión de la clasificación y reducir posibles sesgos en la selección de fechas óptimas.
- Explorar otros algoritmos de clasificación, como Random Forest o Redes Neuronales Artificiales, para determinar si estos modelos pueden mejorar la precisión en la identificación de caña de azúcar y reducir la confusión con otras clases.
- Realizar un análisis de sensibilidad más detallado sobre el efecto de diferentes valores de regularización en el SVM, con el fin de identificar configuraciones óptimas para estudios similares en otras regiones.
- Considerar la incorporación de índices espectrales adicionales además del NDVI, como el EVI o SAVI, para evaluar si pueden mejorar la discriminación entre cultivos con características fenológicas similares.
- Optimizar el proceso de preprocesamiento de imágenes Sentinel-1 para reducir tiempos computacionales, evaluando herramientas alternativas o automatización de procesos mediante scripts en Python.
- Explorar la posibilidad de generar mapas de clasificación a nivel nacional, utilizando la metodología aplicada en este estudio como base para identificar la caña de azúcar en otras regiones productoras del país.
- Ampliar la evaluación de otras técnicas de filtrado de ruido granular en imágenes SAR, como los filtros Lee o Frost, con el objetivo de determinar si pueden proporcionar mejores resultados en términos de balance entre reducción de ruido y preservación de detalles.

7. REFERENCIAS BIBLIOGRÁFICAS

- Adeli, S., Salehi, B., Mahdianpari, M., Quackenbush, L. J., Brisco, B., Tamiminia, H., & Shaw, S. (2020). Wetland monitoring using SAR Data: A meta-analysis and comprehensive review. *Remote Sensing*, *12*(14). <https://doi.org/10.3390/rs12142190>
- Agencia Espacial Europea. (n.d.). *Copernicus Programme*. <https://sentiwiki.copernicus.eu/web/copernicus-programme>
- Aguilar, A. (2019). Ubicación de áreas agrícolas por medio de imágenes satelitales de alta resolución en distintas zonas de Costa Rica. *Revista Tecnología En Marcha*, *32*, 43–50. <https://doi.org/10.18845/tm.v32i7.4258>
- Anand, N. S., Marappan, R., & Sethumadhavan, G. (2018). Performance Analysis of SAR Image Speckle Filters and its Recent Challenges. *2018 IEEE International Conference on Computational Intelligence and Computing Research, ICCIC 2018*, *5*, 1–4. <https://doi.org/10.1109/ICCIC.2018.8782425>
- Andrade, O., de Assunção, A., da Silva, M., de Barros, L., Brito, T., Mendes, J., Alves, A., da Silva, M., Costa, V., Ferreira, R., Freire, T., & Pinto, B. (2024). UAV-Based Classification of Intercropped Forage Cactus: A Comparison of RGB and Multispectral Sample Spaces Using Machine Learning in an Irrigated Area. *AgriEngineering*, *6*(1), 509–525. <https://doi.org/10.3390/agriengineering6010031>
- Arias, M., Campo, M. A., & Álvarez, J. (2020). Crop classification based on temporal signatures of Sentinel-1 observations over Navarre province, Spain. *Remote Sensing*. <https://doi.org/https://doi.org/10.3390/rs12020278>
- Ávila-Pérez, I. D., Ortiz-Malavassi, E., Soto-Montoya, C., Vargas-Solano, Y., Aguilar-Arias, H., & Miller-Granados, C. (2020). Evaluación de cuatro algoritmos de clasificación de imágenes satelitales Landsat-8 y Sentinel-2 para la identificación de cobertura boscosa en paisajes altamente fragmentados en Costa Rica. *Revista de Teledetección*, *57*, 37. <https://doi.org/10.4995/raet.2020.13340>
- Bargiel, D. (2017). A new method for crop classification combining time series of radar images and crop phenology information. *Remote Sensing of Environment*, *198*, 369–383. <https://doi.org/10.1016/j.rse.2017.06.022>
- Barni, M., & Ratto, E. (2022). “*Teledetección aplicada al Agro . Clasificación de cultivos a partir de imágenes satelitales*” [Universidad Nacional de La Plata]. https://sedici.unlp.edu.ar/bitstream/handle/10915/141770/Documento_completo.Clasificación_de_cultivos_a_partir_de_imagenes_satelitales.pdf-PDFA.pdf?isAllowed=y&sequence=1
- Belmonte, I., & Caamaño, S. (2009). *Una aplicación basada en imágenes SAR para el monitoreo y prevención de inundaciones en Argentina*.

- Brinkhoff, J., Vardanega, J., & Robson, A. J. (2020). Land cover classification of nine perennial crops using sentinel-1 and -2 data. *Remote Sensing*, 12(1), 1–26. <https://doi.org/10.3390/rs12010096>
- Campos, C. (2010). *Análisis de los cambios de cobertura de la cuenca alta y media del Río Reventazón, Costa Rica, periodo 2000-2010*. Tecnológico de Costa Rica.
- Cerezo, D. (2023). *Identificación de cultivos mediante teledetección y técnicas de machine learning en el ámbito territorial de la comarca de la Ribera Baja, Valencia*. <https://riunet.upv.es:443/handle/10251/197929>
- Copernicus. (2018). Fundamentos de teledetección aplicada. *Ministerio de Transportes, Movilidad y Agenda Urbana*, 1–27. https://www.ign.es/web/resources/docs/IGNCnig/actividades/OBS/Programa_Marco_Copernicus_User_Uptake/2_Fundamentos_teledeteccion_aplicada.pdf
- Cruz-Sanabria, H., Sanchez, M. G., Rivera-Caicedo, J. P., & Avila-George, H. (2020). Identification of phenological stages of sugarcane cultivation using Sentinel-2 images. *Applications in Software Engineering - Proceedings of the 9th International Conference on Software Process Improvement, CIMPS 2020*, 110–116. <https://doi.org/10.1109/CIMPS52057.2020.9390095>
- Dingle, L., Davidson, A., McNairn, H., Hosseini, M., Mitchell, S., De Abelleira, D., Verón, S., & Cosh, M. H. (2018). SAR speckle filtering and agriculture field size: Development of sar data processing best practices for the JECAM SAR intercomparison experiment. *International Geoscience and Remote Sensing Symposium (IGARSS), 2018-July*, 3828–3831. <https://doi.org/10.1109/IGARSS.2018.8519299>
- Dingle, L., Davidson, A., McNairn, H., Hosseini, M., Mitchell, S., De Abelleira, D., Verón, S., & Cosh, M. H. (2020). Synthetic Aperture Radar (SAR) image processing for operational space-based agriculture mapping. *International Journal of Remote Sensing*, 41(18), 7112–7144. <https://doi.org/10.1080/01431161.2020.1754494>
- ESA. (n.d.). *Overview of Sentinel-2 Mission*. <https://sentiwiki.copernicus.eu/web/s2-mission>
- Filipponi, F. (2019). *Sentinel-1 GRD Preprocessing Workflow*. 11. <https://doi.org/10.3390/ecrs-3-06201>
- Flores, J. (2017). Agrocadena de la Caña De Azúcar. *Mag*, 1–80. <http://www.mag.go.cr/bibliotecavirtual/E70-10273.pdf>
- Hidalgo, Natalia. (2022). *Análisis del rendimiento del cultivo de caña de azúcar mediante índices de vegetación y monitores de rendimiento durante el periodo de zafra 2021 – 2022 en la empresa Central Azucarera Tempisque S.A. (CATSA). Guanacaste, Costa Rica*.
- Instituto Nacional de Técnica Aeroespacial (INTA). (2018). *¡Productos L2A de Sentinel-2 en todo el mundo!* https://www.inta.es/INTA/en/blogs/copernicus/BlogEntry_1521119787301

- Jiang, H., Li, D., Jing, W., Xu, J., Huang, J., Yang, J., & Chen, S. (2019). Early season mapping of sugarcane by applying machine learning algorithms to sentinel-1A/2 time series data: A case study in Zhanjiang City, China. *Remote Sensing*, *11*(7). <https://doi.org/10.3390/RS11070861>
- Jiménez, A. V., & Guerrero, M. J. (2023). *Evaluación de dos clasificadores para la identificación de áreas inundables con diferentes metodologías de procesamiento mediante imágenes de Radar de Apertura Sintética en la planicie inundable de la margen derecha del Río Tempisque, Costa Rica.*
- Kuma, B., Ranjan, R. K., & Husain, A. (2022). A New Synthetic Aperture Radar (SAR) Image Classification Framework Model Using SAR Despeckling. *International Journal of Computing and Digital Systems*, *12*(4), 1129–1144. <https://doi.org/10.12785/ijcds/120190>
- Kumar, D. (2021). Urban objects detection from C-band synthetic aperture radar (SAR) satellite images through simulating filter properties. *Scientific Reports*, *11*(1), 1–24. <https://doi.org/10.1038/s41598-021-85121-9>
- Landis, J. R., & Koch, G. G. (1977). Landis and Koch agreement of categorical data. *Biometrics*, *33*(1), 159–174.
- Lipton, Z. C., Elkan, C., & Narayanaswamy, B. (2014). *Thresholding Classifiers to Maximize F1 Score*. <http://arxiv.org/abs/1402.1892>
- López Cabrera, J. D., & Pereira-Toledo, A. (2018). Análisis Del Comportamiento Del Algoritmo SVM Para Diferentes Kernel En Ambientes Controlados. *Holos*, *5*, 101–115. <https://doi.org/10.15628/holos.2018.5563>
- Lopez-Martinez, D. (2017). *Regularization approaches for support vector machines with applications to biomedical data*. *1*(4), 1–7. <http://arxiv.org/abs/1710.10600>
- Lozano-Garzon, C., Bravo-Cordoba, G., Castro, H., Gonzalez-Rodriguez, G., Nino, D., Nunez, H., Pardo, C., Vivas, A., Castro, Y., Medina, J., Motta, L. C., Rojas, J. R., & Suarez, L. I. (2022). Remote Sensing and Machine Learning Modeling to Support the Identification of Sugarcane Crops. *IEEE Access*, *10*, 17542–17555. <https://doi.org/10.1109/ACCESS.2022.3148691>
- Maity, A., Pattanaik, A., Sagnika, S., & Pani, S. (2015). A comparative study on approaches to speckle noise reduction in images. *Proceedings - 1st International Conference on Computational Intelligence and Networks, CINE 2015*, *1*, 148–155. <https://doi.org/10.1109/CINE.2015.36>
- Mandal, D., Kumar, V., & Rao, Y. S. (2022). An assessment of temporal RADARSAT-2 SAR data for crop classification using KPCA based support vector machine. *Geocarto International*, *37*(6), 1547–1559. <https://doi.org/10.1080/10106049.2020.1783577>
- Mandal, D., Vaka, D. S., Bhogapurapu, N. R., Vanama, K. S. K., Kumar, V., Rao, Y. S., & Bhattacharya, A. (2019). Sentinel-1 SLC preprocessing workflow for polarimetric applications: A generic practice for generating dual-pol covariance matrix elements in SNAP

- Marini, F. (2021). Discriminación de trigo y cebada empleando imágenes satelitales ópticas y radar. Estudio de caso: partido de Coronel Rosales (Argentina). *Investigaciones Geográficas*, 104, 1–17.
- McNairn, H., Shang, J., Jiao, X., & Champagne, C. (2009). The contribution of ALOS PALSAR multipolarization and polarimetric data to crop classification. *IEEE Transactions on Geoscience and Remote Sensing*, 47(12), 3981–3992. <https://doi.org/10.1109/TGRS.2009.2026052>
- Meenakshi, A., & Punitham, V. (2011). Performance of speckle noise reduction filters on active radar and SAR images. *International Journal of Technology And Engineering System*, 2(March), 1–4.
- Mishra, D., Pathak, G., Singh, B. P., Mohit, Sihag, P., Rajeev, Singh, K., & Singh, S. (2023). Crop classification by using dual-pol SAR vegetation indices derived from Sentinel-1 SAR-C data. *Environmental Monitoring and Assessment*, 195(1). <https://doi.org/10.1007/s10661-022-10591-x>
- Morales, J., Acuña, V., & Cruz, A. (2003). *Industrialización del Heno de Calidad en Sistemas Bajo Riego en Costa Rica*. 10, N.D.
- Morales, M., Vargas, A., Vignola, R., & Watler, W. (2018). Prácticas efectivas para la reducción de impactos por eventos climáticos en costa rica (cultivo de arroz). *Instituto Nacional de Tecnología Agropecuaria*, 1–85. <http://www.mag.go.cr/bibliotecavirtual/reduccion-impacto-por-eventos-climaticos/Informe-final-Arroz.pdf>
- Moskoliai, W. R., Abdou, W., Dipanda, A., & Kolyang. (2022). A Workflow for Collecting and Preprocessing Sentinel-1 Images for Time Series Prediction Suitable for Deep Learning Algorithms. *Geomatics*, 2(4), 435–456. <https://doi.org/10.3390/geomatics2040024>
- Nezry, E., Lopés, A., Ducrot-Gambart, D., Nezry, C., & Lee, J. Sen. (1996). Supervised classification of Jf-distributed SAR images of natural targets and probability of error estimation. *IEEE Transactions on Geoscience and Remote Sensing*, 34(5), 1233–1242. <https://doi.org/10.1109/36.536539>
- ONU. (2015). *Objetivos de Desarrollo Sostenible*. Organización de Las Naciones Unidas. <https://www.un.org/sustainabledevelopment/es/objetivos-de-desarrollo-sostenible/>
- Organización de las Naciones Unidas (ONU). (2019). *La población mundial sigue en aumento, aunque sea cada vez más vieja*. <https://news.un.org/es/story/2019/06/1457891>
- Pereira-Sandoval, M. (2023). *Desafíos y oportunidades de Sentinel-2 en la monitorización de la calidad de las aguas continentales*.

- Pisner, D. A., & Schnyer, D. M. (2019). Support vector machine. In *Machine Learning: Methods and Applications to Brain Disorders*. Elsevier Inc. <https://doi.org/10.1016/B978-0-12-815739-8.00006-7>
- Prashanthika, B., Choudhary, K., Pandey, V., & Murthy, C. (2019). *Mapping Crops in Heterogenous Landscape : Selection of Optimum Bands Combination and Period for Single-Date Classification*. 19(1), 134–141.
- Rana, V. K., & Suryanarayana, T. M. V. (2019). Evaluation of SAR speckle filter technique for inundation mapping. *Remote Sensing Applications: Society and Environment*, 16(October), 100271. <https://doi.org/10.1016/j.rsase.2019.100271>
- Rojas, S., & Medina, J. (2020). Supervised Markov contextual classification using LandSat 8 OLI and Sentinel 2A multispectral images. *Revista Ud Y La Geomática*, 15, 84–90.
- Saadat, H., Adamowski, J., Bonnell, R., Sharifi, F., Namdar, M., & Ale-Ebrahim, S. (2011). Land use and land cover classification over a large area in Iran based on single date analysis of satellite imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(5), 608–619. <https://doi.org/10.1016/j.isprsjprs.2011.04.001>
- Santoso, A. W., Pebrianti, D., Bayuaji, L., & Zain, J. M. (2015). Performance of various speckle reduction filters on Synthetic Aperture Radar image. *2015 4th International Conference on Software Engineering and Computer Systems, ICSECS 2015: Virtuuous Software Solutions for Big Data*, 11–14. <https://doi.org/10.1109/ICSECS.2015.7333103>
- Sara, U., Akter, M., & Uddin, M. S. (2019). Image Quality Assessment through FSIM, SSIM, MSE and PSNR—A Comparative Study. *Journal of Computer and Communications*, 07(03), 8–18. <https://doi.org/10.4236/jcc.2019.73002>
- Servicio Nacional de Aguas Subterráneas Riego y Avenamiento (SENARA). (2021). *Distrito de Riego Arenal Tempisque*. <http://www.senara.or.cr/proyectos/drat/index.aspx>
- SINAC. (2013). *Plan General de Manejo Reserva Biológica Lomas Barbudal 2014-2024*.
- Vignola, R., Poveda, K., Watler, W., Vargas, A., & Berrocal, Á. (2018). Cultivo de Caña de Azúcar en Costa Rica. *Programa de Cambio Climático y Cuencas, CATIE*, 1–127. <https://www.mag.go.cr/bibliotecavirtual/F01-8327.pdf>
- Wang, M., Liu, Z., Ali Baig, M. H., Wang, Y., Li, Y., & Chen, Y. (2019). Mapping sugarcane in complex landscapes by integrating multi-temporal Sentinel-2 images and machine learning algorithms. *Land Use Policy*, 88(July), 104190. <https://doi.org/10.1016/j.landusepol.2019.104190>
- Wood, D., McNairn, H., Brown, R. J., & Dixon, R. (2002). The effect of dew on the use of RADARSAT-1 for crop monitoring Choosing between ascending and descending orbits. *Remote Sensing of Environment*, 81(2–3), 456. [https://doi.org/10.1016/S0034-4257\(02\)00032-9](https://doi.org/10.1016/S0034-4257(02)00032-9)

- Yeasin, M., Haldar, D., Kumar, S., Paul, R. K., & Ghosh, S. (2022). Machine Learning Techniques for Phenology Assessment of Sugarcane Using Conjunctive SAR and Optical Data. *Remote Sensing*, *14*(14). <https://doi.org/10.3390/rs14143249>
- Zhang, X., Gao, H., Shi, L., Hu, X., Zhong, L., & Bian, J. (2024). Mapping Crop Evapotranspiration by Combining the Unmixing and Weight Image Fusion Methods. *Remote Sensing*, *16*(13). <https://doi.org/10.3390/rs16132414>

8. APÉNDICE

APÉNDICE A:

Objetivo específico 1

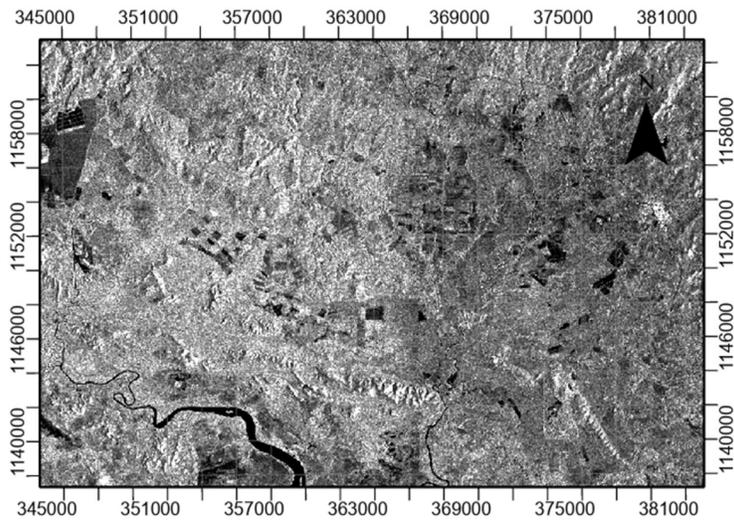


Figura A. 1. Imagen preprocesada filtro Gamma MAP 5x5 del 20 de septiembre del 2019.

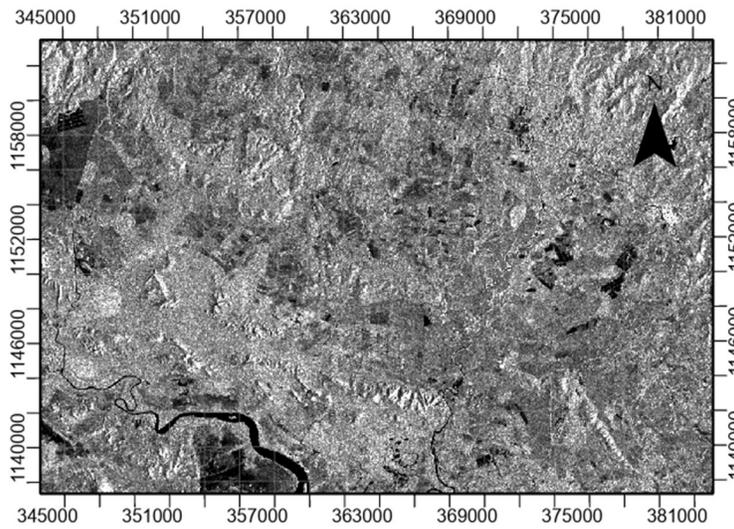


Figura A. 2. Imagen preprocesada filtro Gamma MAP 5x5 del 02 de octubre del 2019.

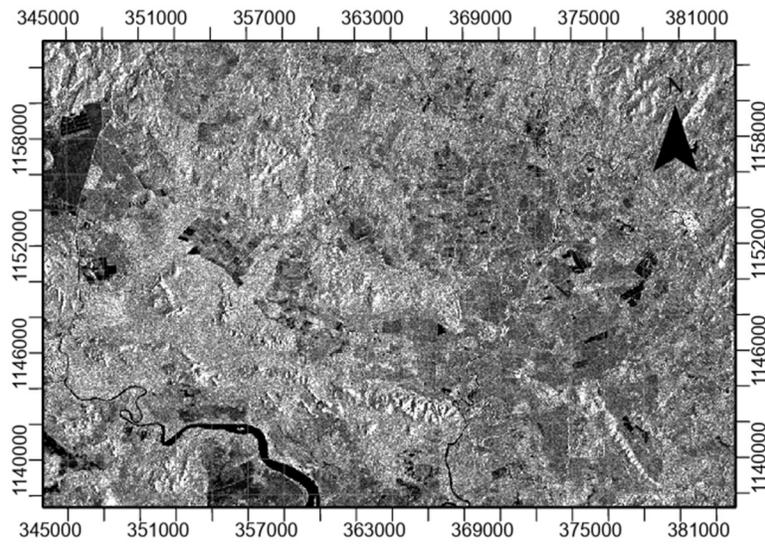


Figura A. 3. Imagen preprocesada filtro Gamma MAP 5x5 del 14 de octubre del 2019.

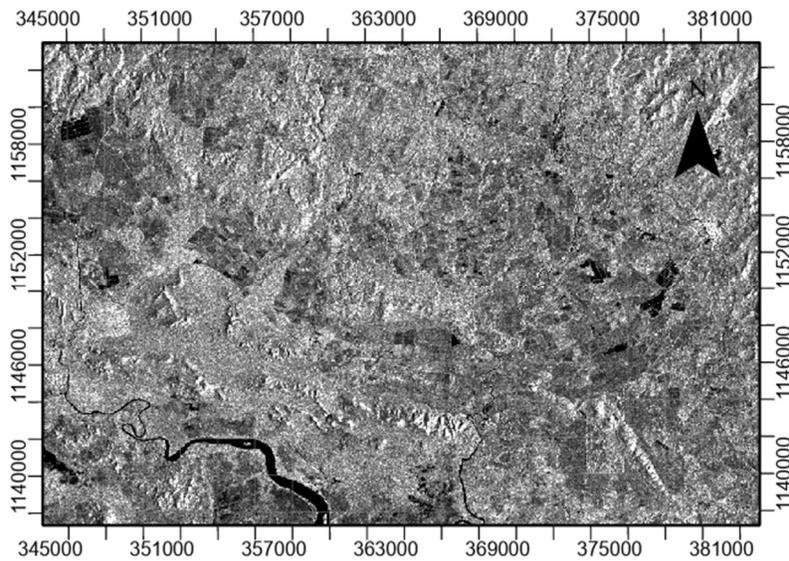


Figura A. 4. Imagen preprocesada filtro Gamma MAP 5x5 del 26 de octubre del 2019.

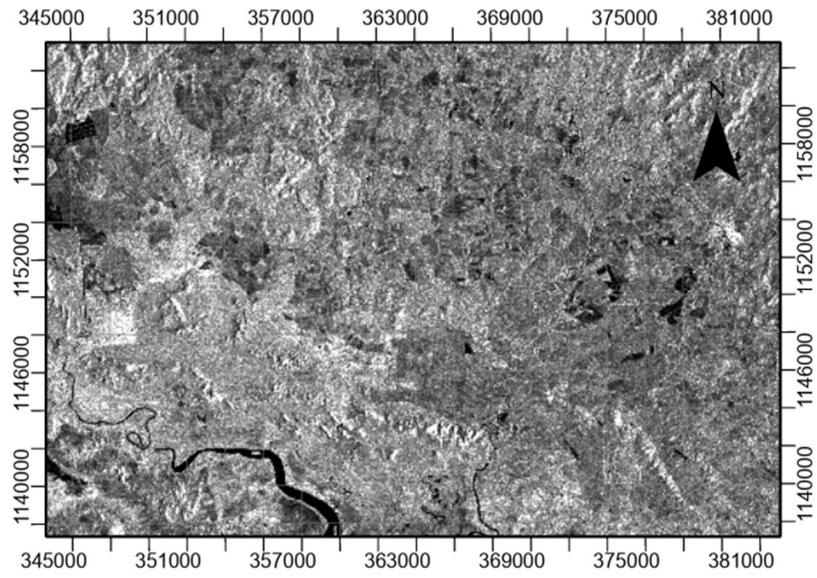


Figura A. 5. Imagen preprocesada filtro Gamma MAP 5x5 del 7 de noviembre del 2019.

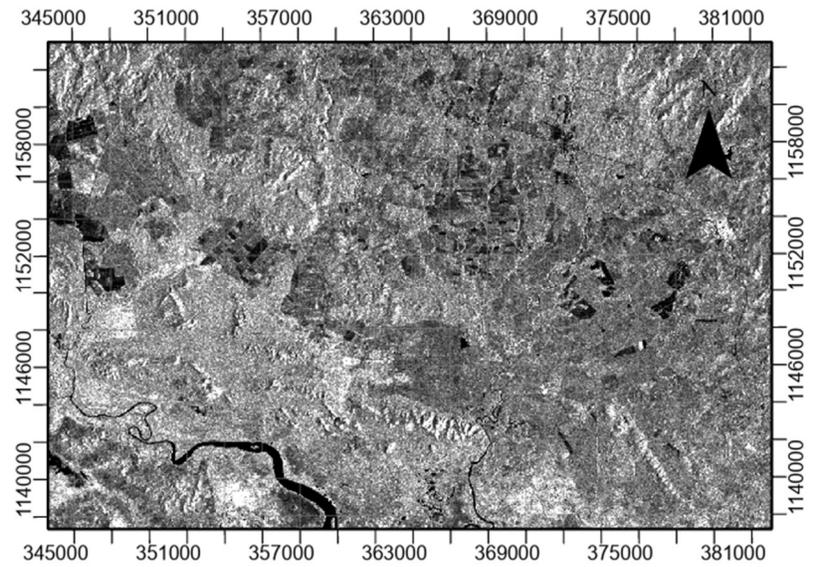


Figura A. 6. Imagen preprocesada filtro Gamma MAP 5x5 del 19 de noviembre del 2019.

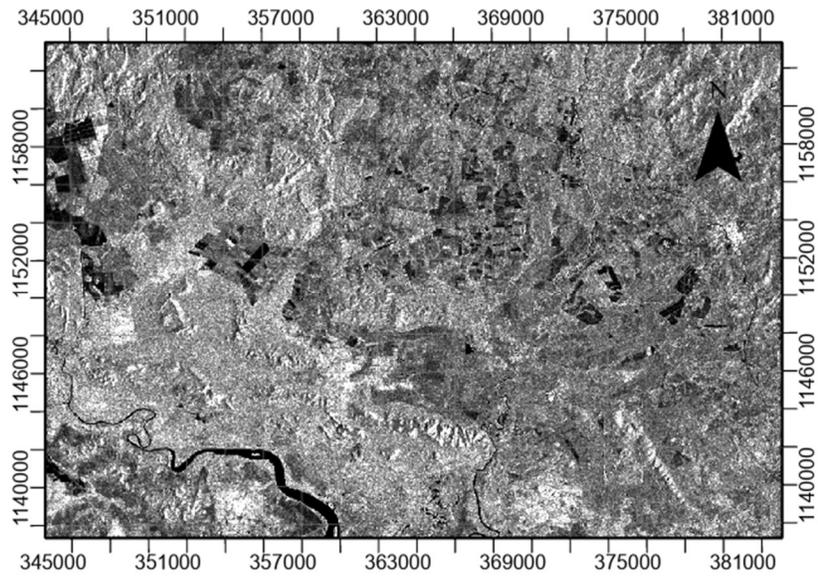


Figura A. 7. Imagen preprocesada filtro Gamma MAP 5x5 del 01 de diciembre del 2019.

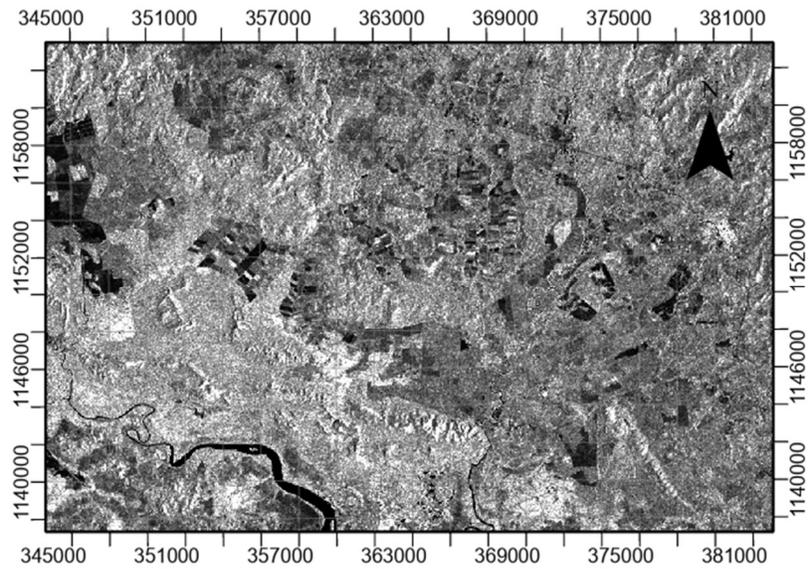


Figura A. 8. Imagen preprocesada filtro Gamma MAP 5x5 del 13 de diciembre del 2019.

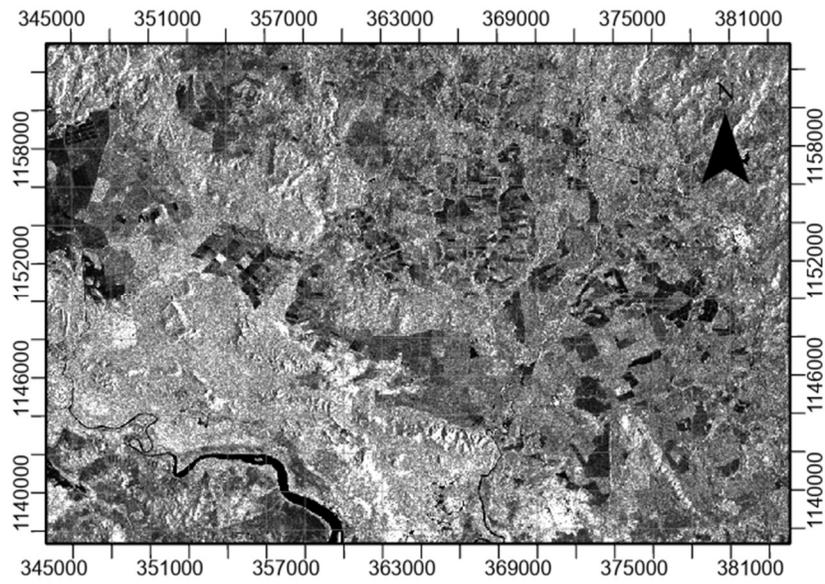


Figura A. 9. Imagen preprocesada filtro Gamma MAP 5x5 del 25 de octubre del 2019.

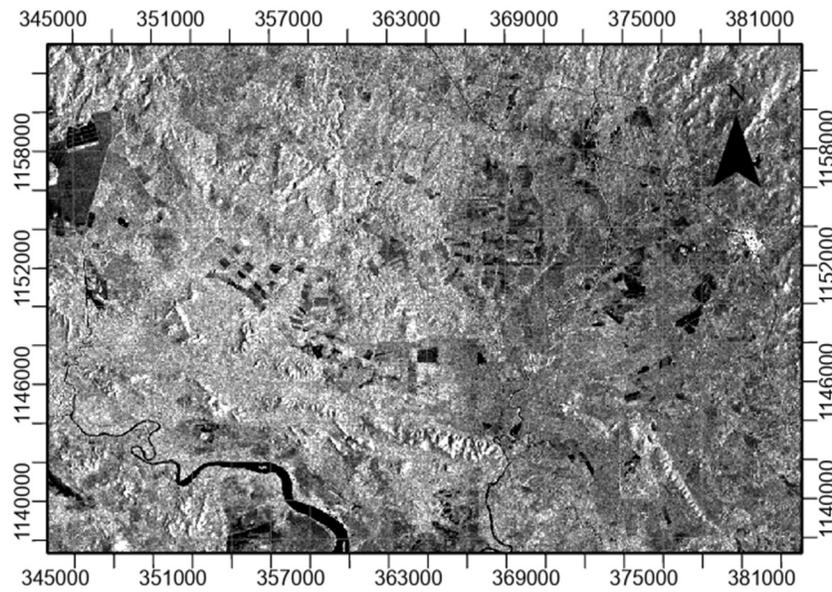


Figura A. 10. Imagen preprocesada filtro Gamma MAP 7x7 del 20 de septiembre del 2019.

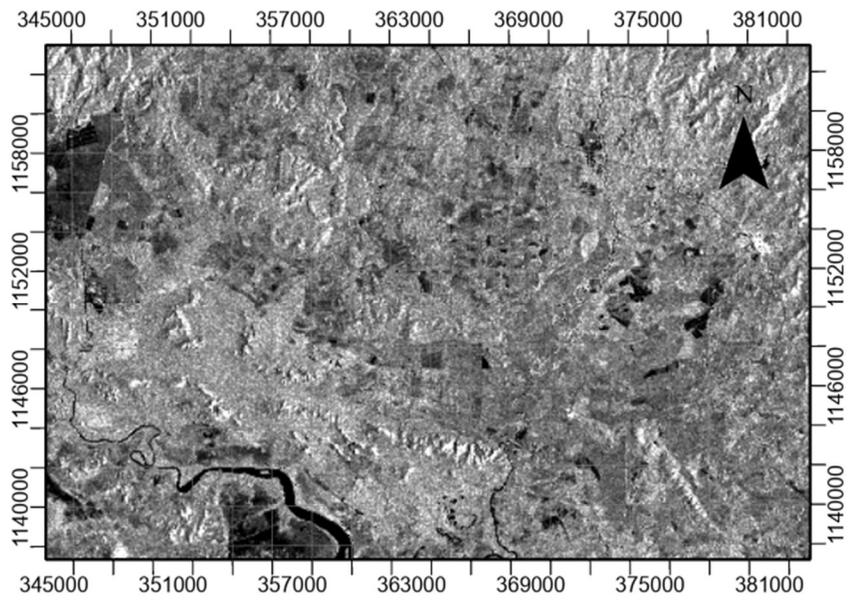


Figura A. 11. Imagen preprocesada filtro Gamma MAP 7x7 del 02 de octubre del 2019.

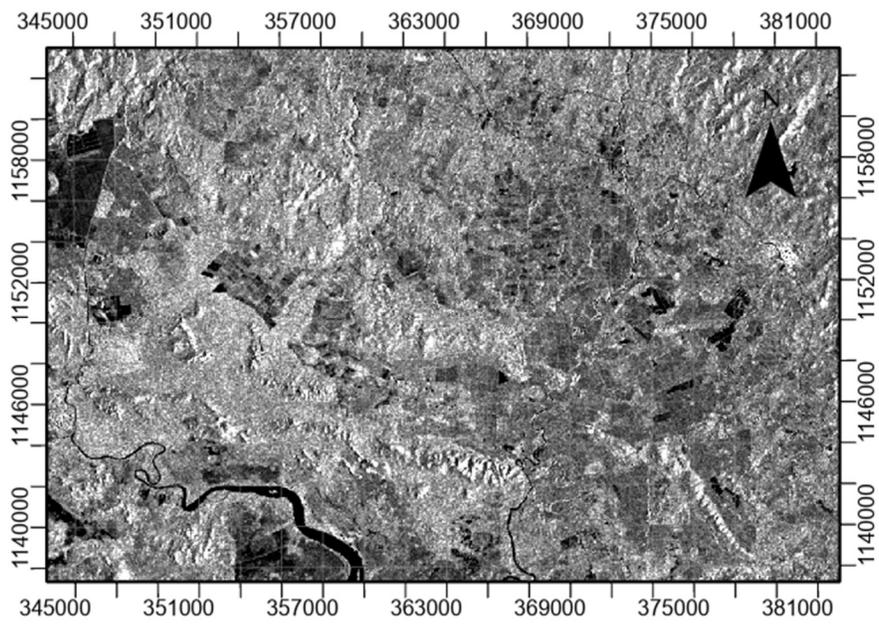


Figura A. 12. Imagen preprocesada filtro Gamma MAP 7x7 del 14 de octubre del 2019.

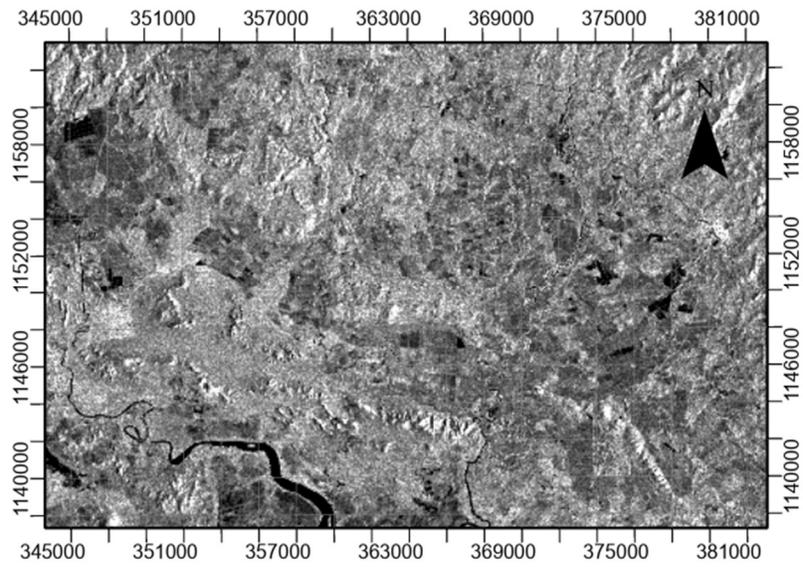


Figura A. 13. Imagen preprocesada filtro Gamma MAP 7x7 del 26 de octubre del 2019.

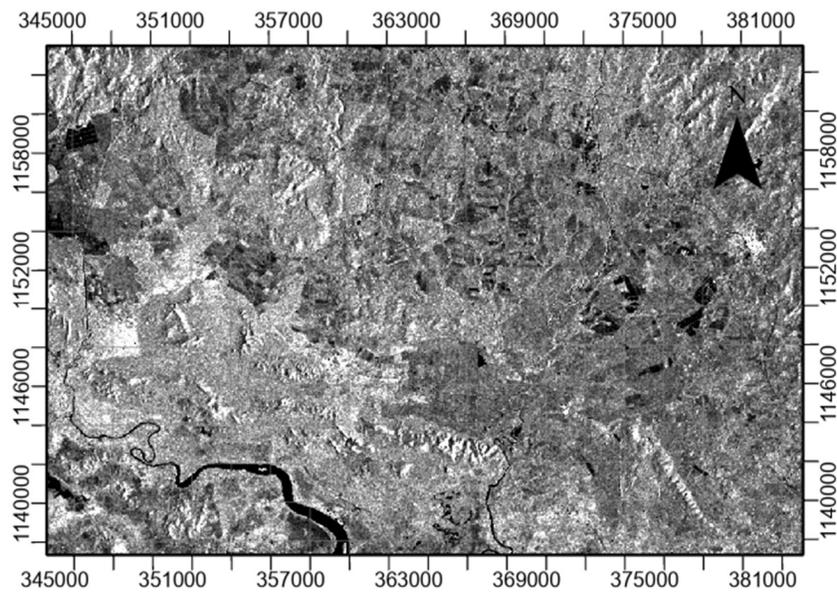


Figura A. 14. Imagen preprocesada filtro Gamma MAP 7x7 del 07 de noviembre del 2019.

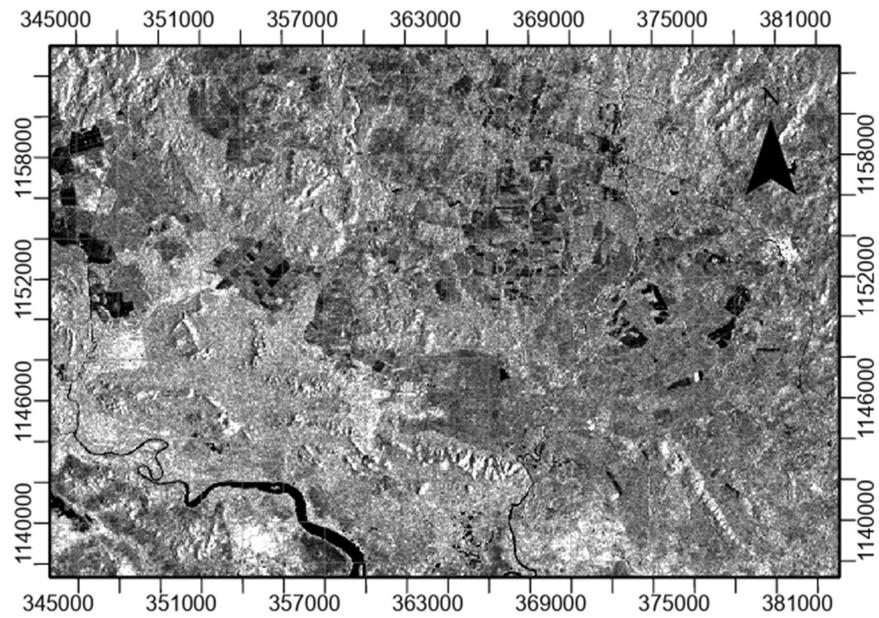


Figura A. 15. Imagen preprocesada filtro Gamma MAP 7x7 del 19 de noviembre del 2019.

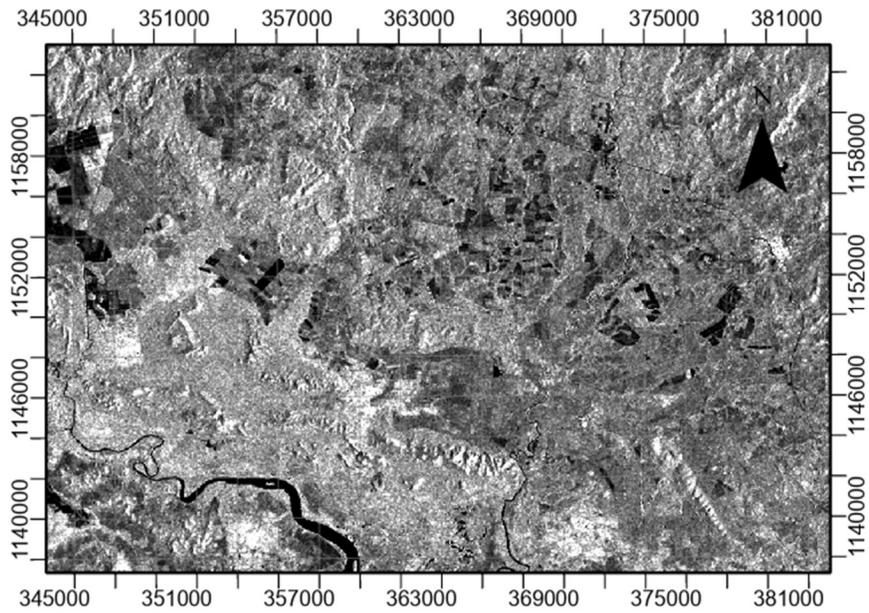


Figura A. 16. Imagen preprocesada filtro Gamma MAP 7x7 del 01 de diciembre del 2019.

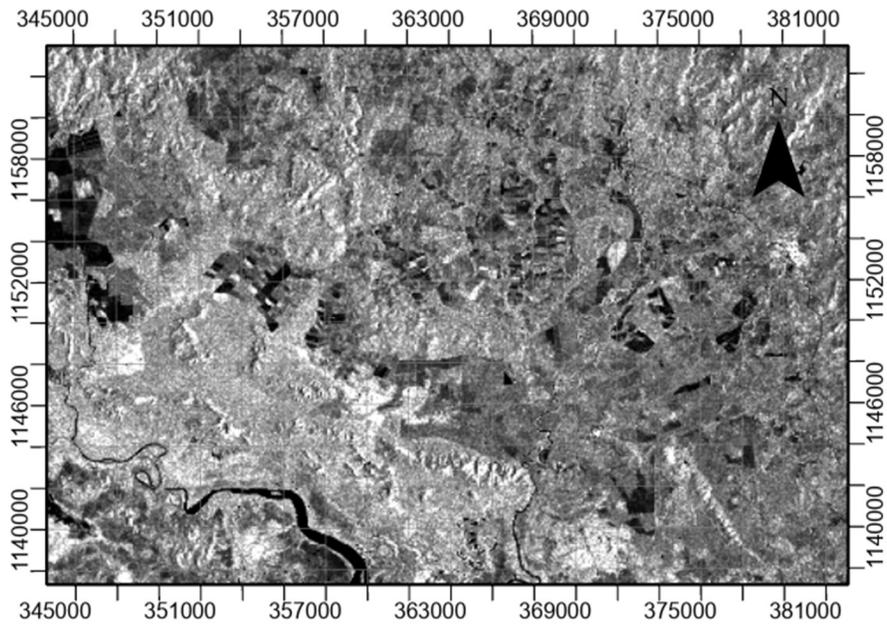


Figura A. 17. Imagen preprocesada filtro Gamma MAP 7x7 del 13 de diciembre del 2019.

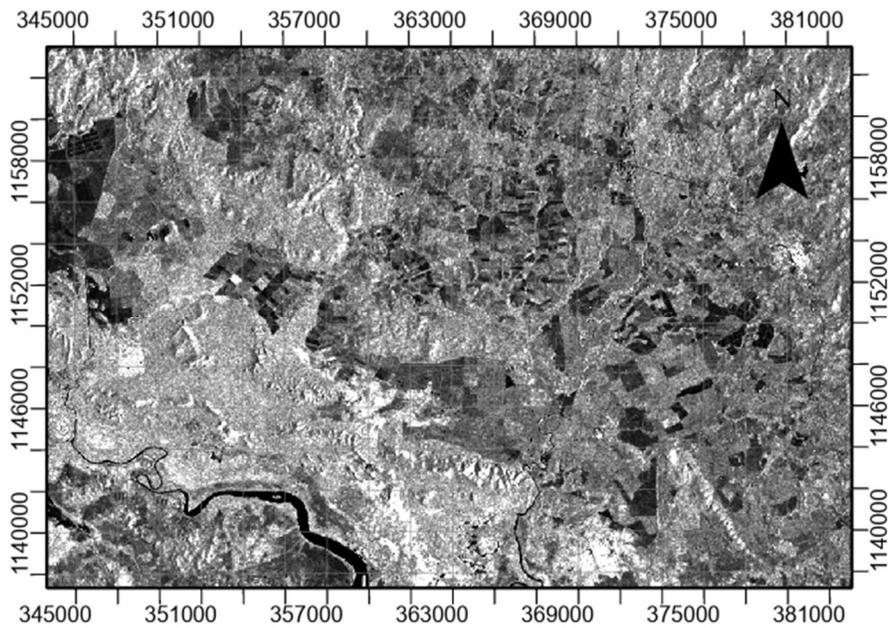


Figura A. 18. Imagen preprocesada filtro Gamma MAP 7x7 del 25 de diciembre del 2019.

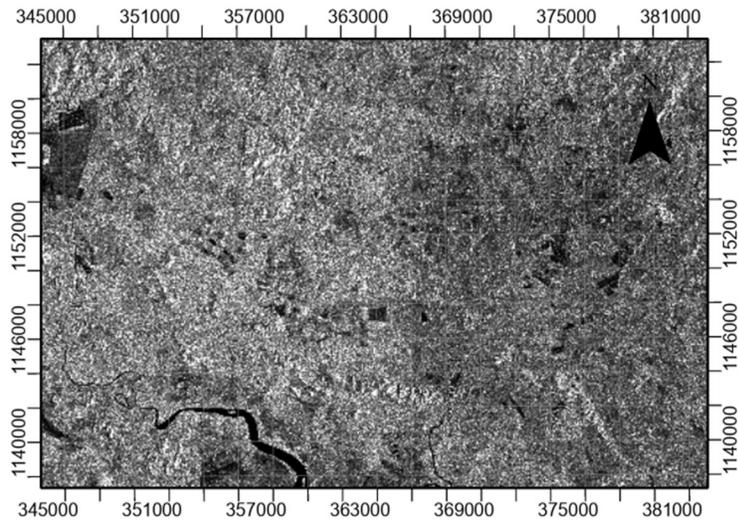


Figura A. 19. Imagen preprocesada sin filtro del 20 de septiembre del 2019.

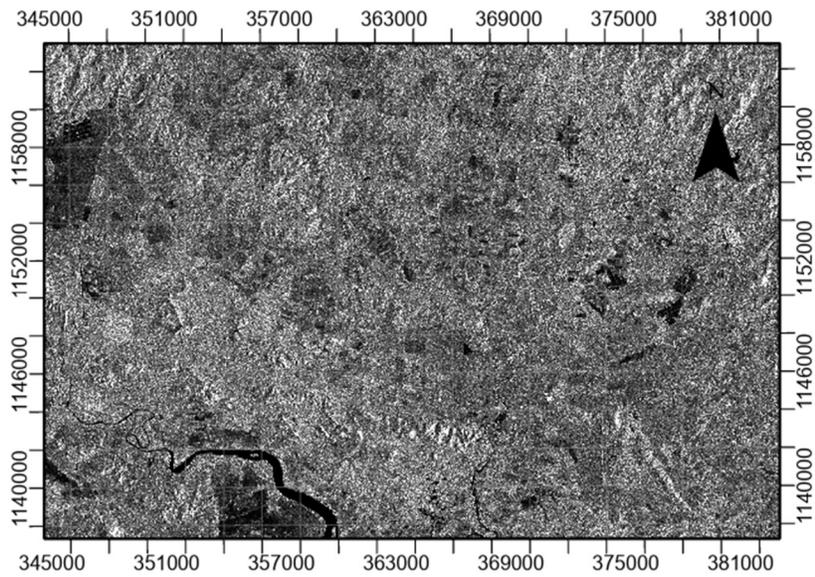


Figura A. 20. Imagen preprocesada sin filtro del 02 de octubre del 2019.

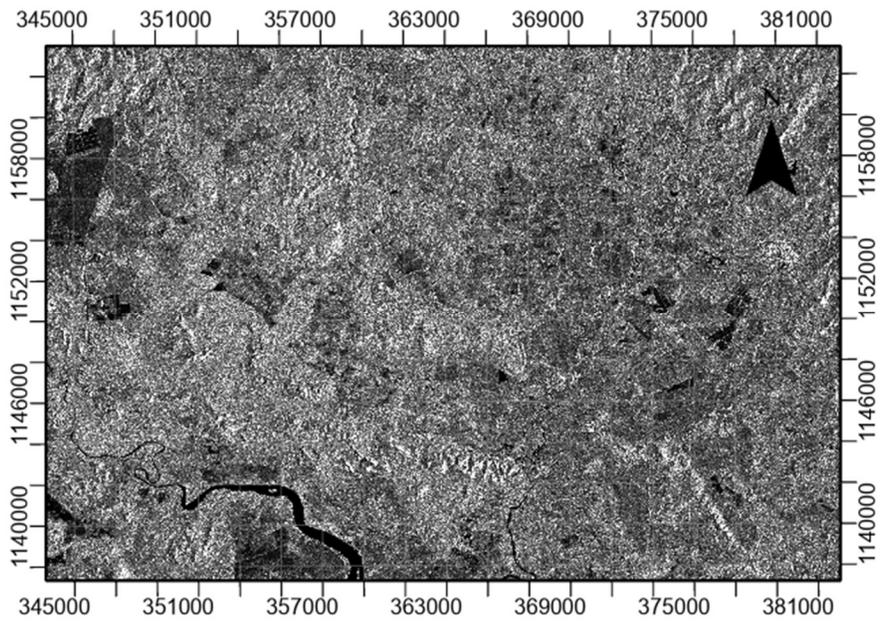


Figura A. 21. Imagen preprocesada sin filtro del 14 de octubre del 2019.

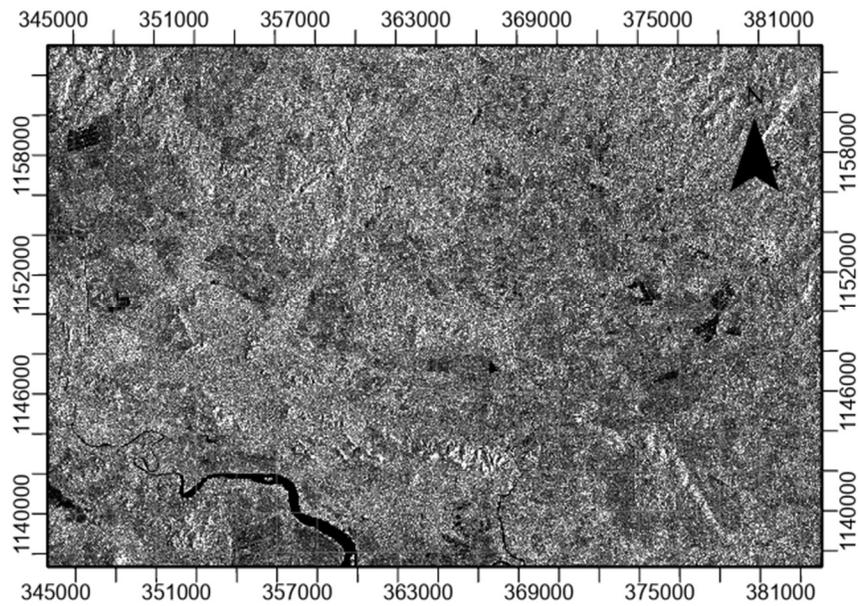


Figura A. 22. Imagen preprocesada sin filtro del 26 de octubre del 2019.

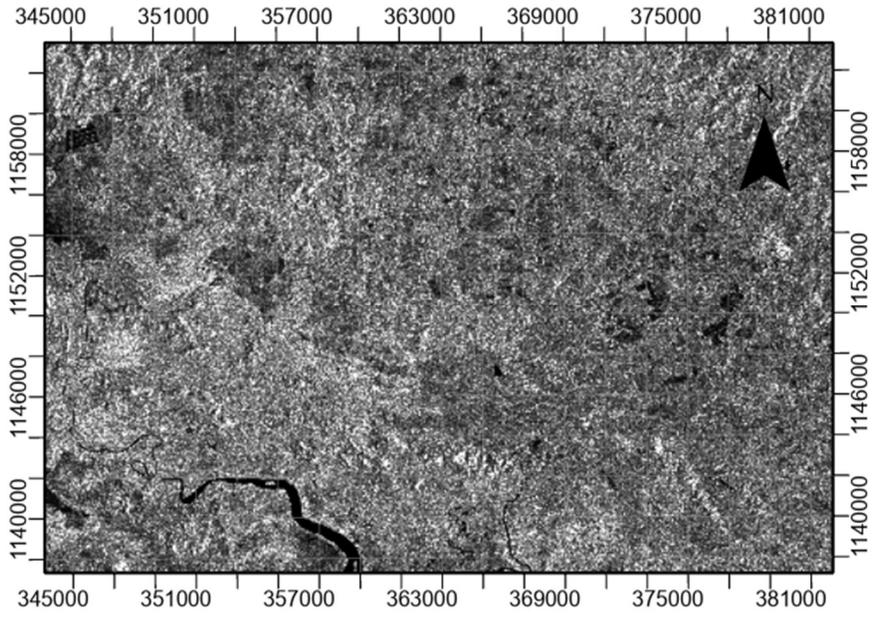


Figura A. 23. Imagen preprocesada sin filtro del 07 de noviembre del 2019.

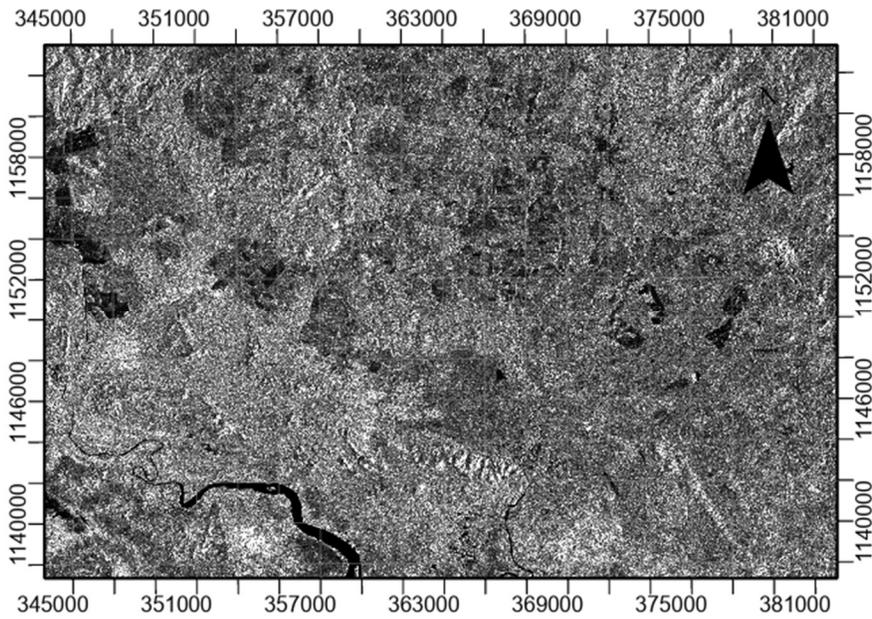


Figura A. 24. Imagen preprocesada sin filtro del 19 de noviembre del 2019.

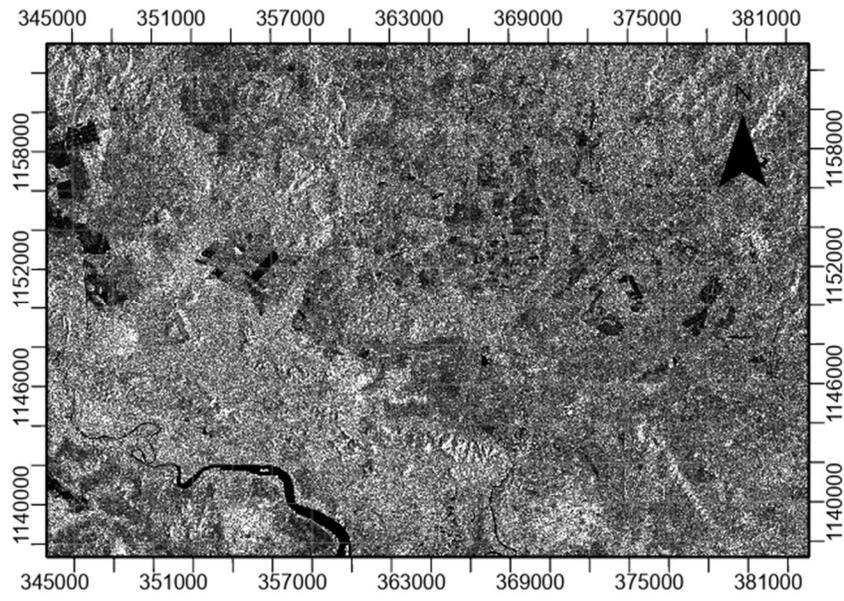


Figura A. 25. Imagen preprocesada sin filtro del 01 de diciembre del 2019.

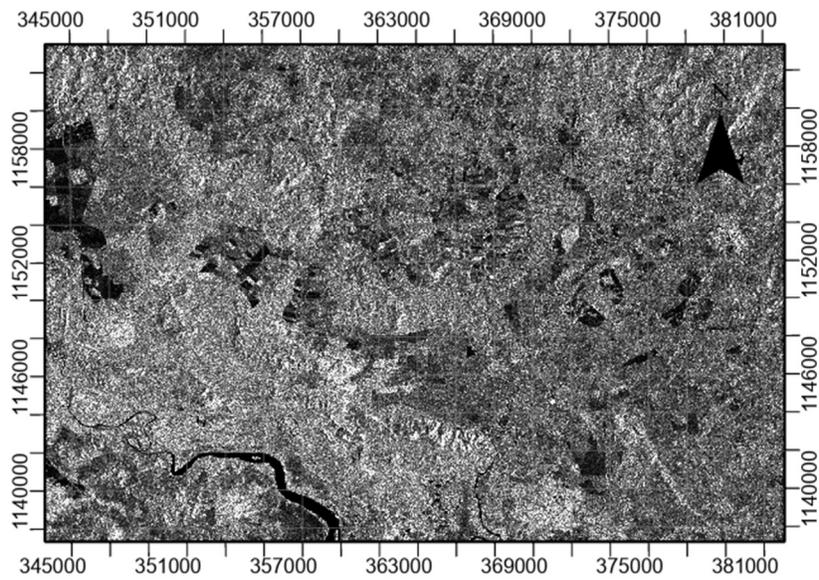


Figura A. 26. Imagen preprocesada sin filtro del 13 de diciembre del 2019.

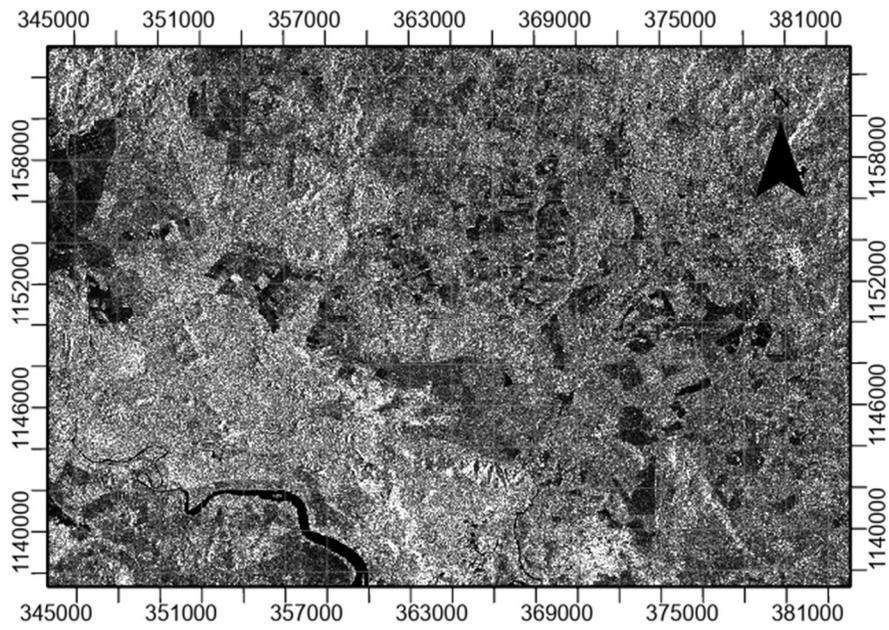


Figura A. 27. Imagen preprocesada sin filtro del 25 de diciembre del 2019.

APÉNDICE B:

Objetivos específicos 2 y 3

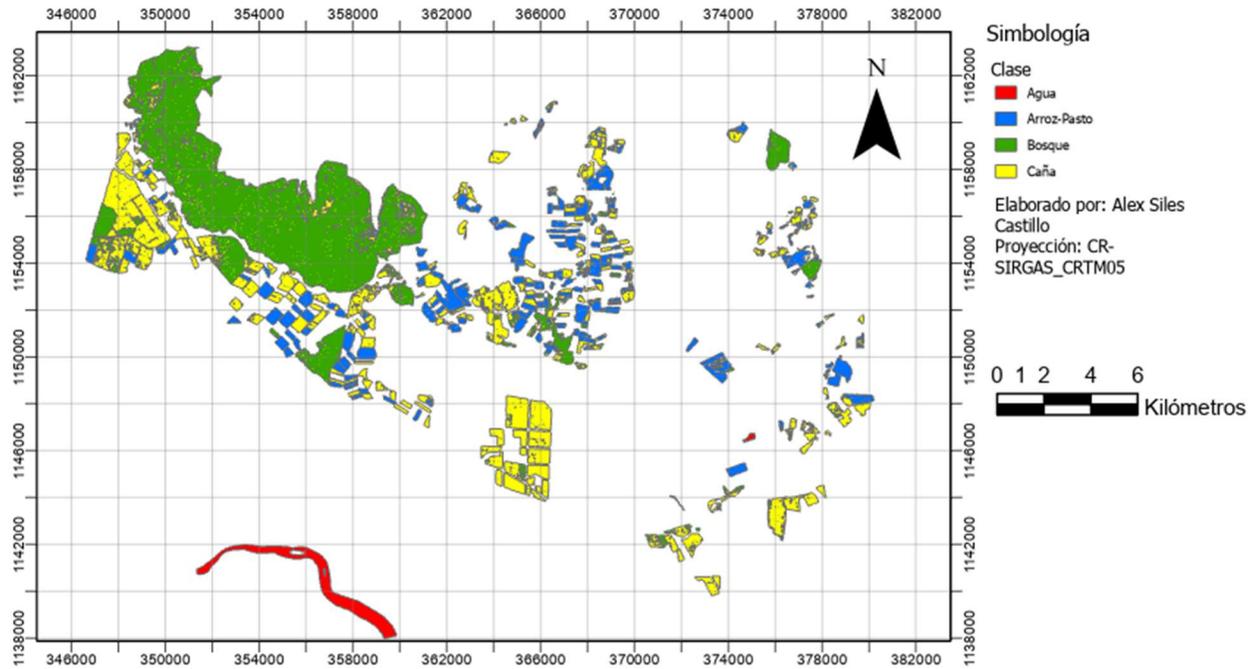


Figura B 1. Visualización de la clasificación multitemporal con las clases de arroz y pasto unidas para un tamaño de ventana de 5x5.

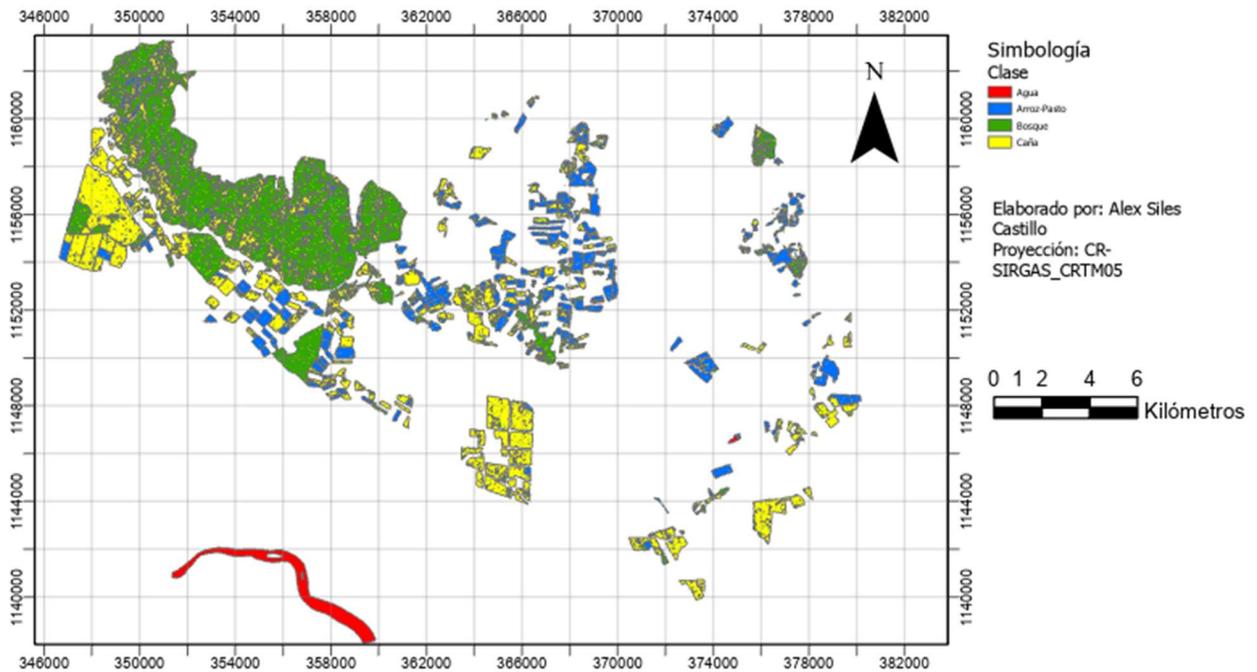


Figura B 2. Visualización de la clasificación multitemporal con las clases de arroz y pasto unidas para un tamaño de ventana de 5x5 sin NDVI.

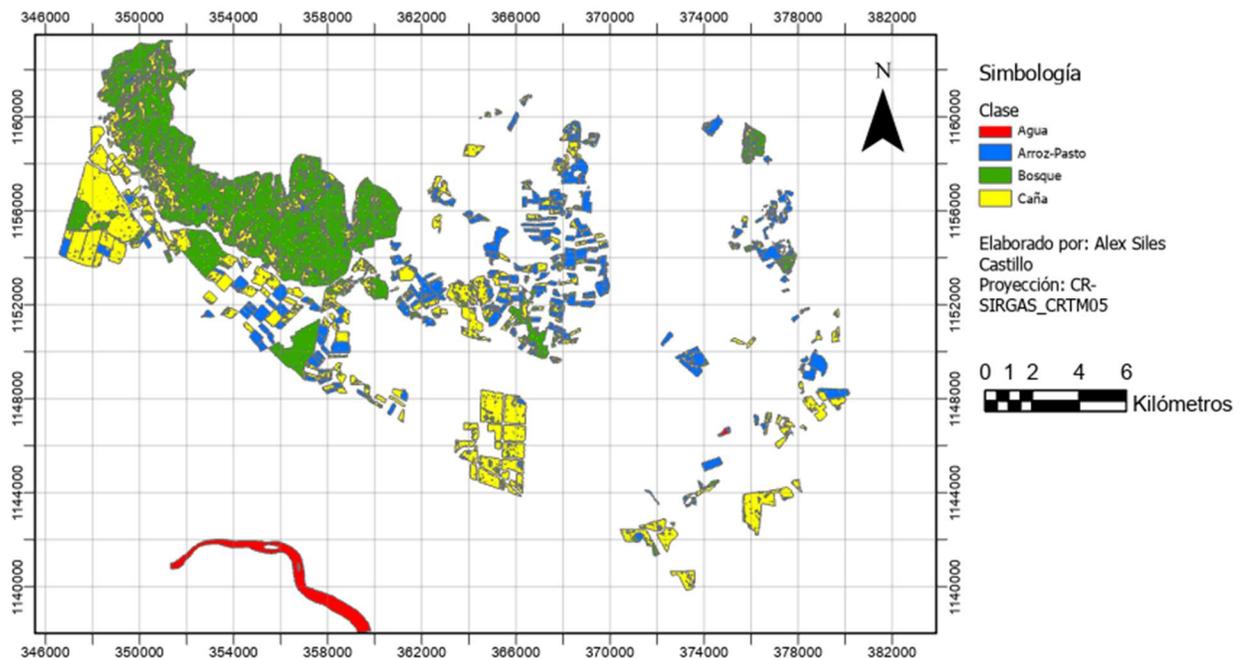


Figura B 3. Visualización de la clasificación multitemporal con las clases de arroz y pasto unidas para un tamaño de ventana de 7x7 sin NDVI.

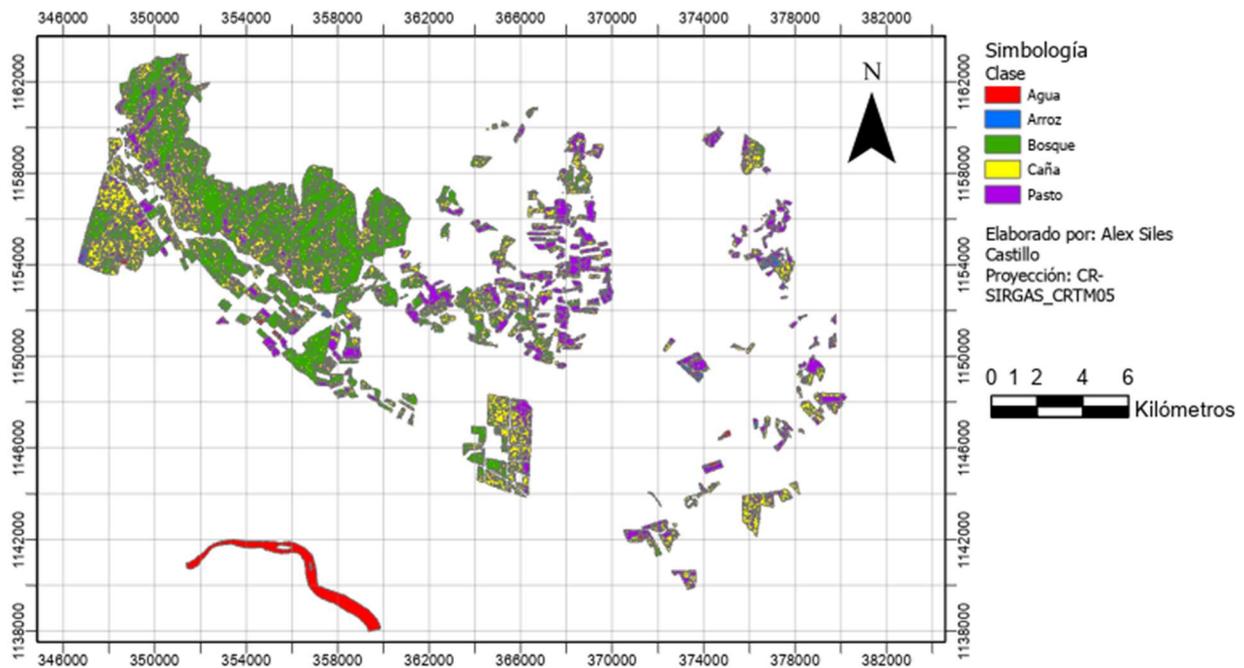


Figura B 4. Clasificación de la zona de estudio para el producto Sentinel-1 del 20 de septiembre del 2019.

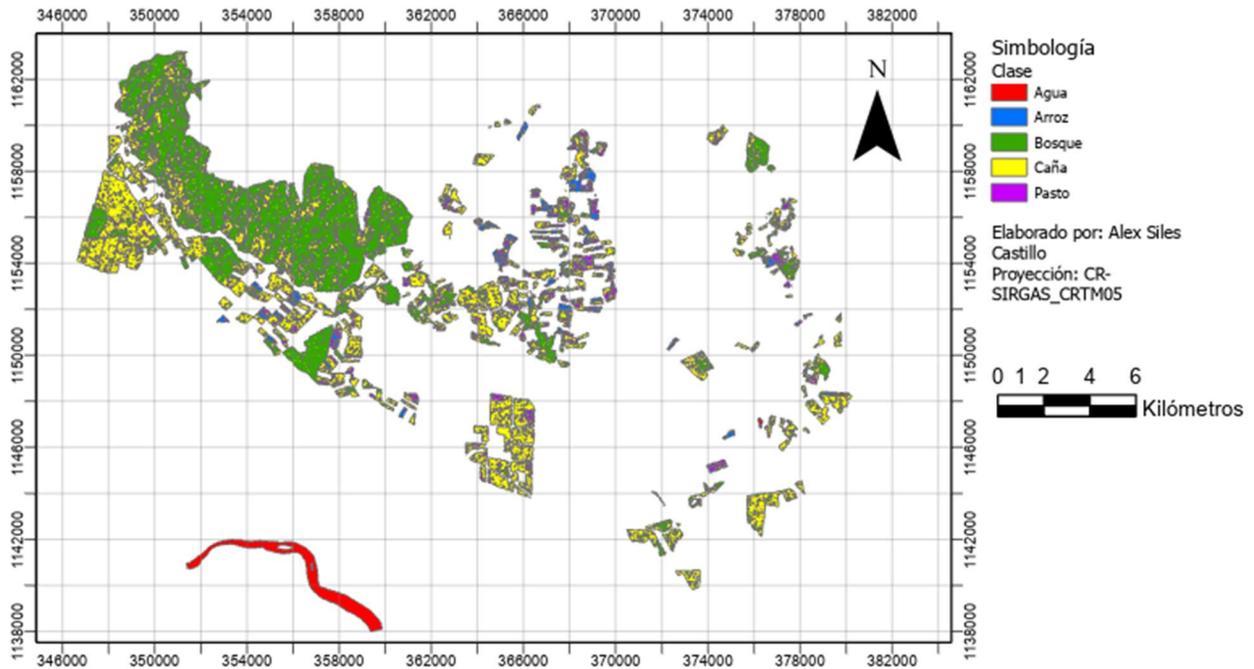


Figura B 5. Clasificación de la zona de estudio para el producto Sentinel-1 del 19 de noviembre del 2019.

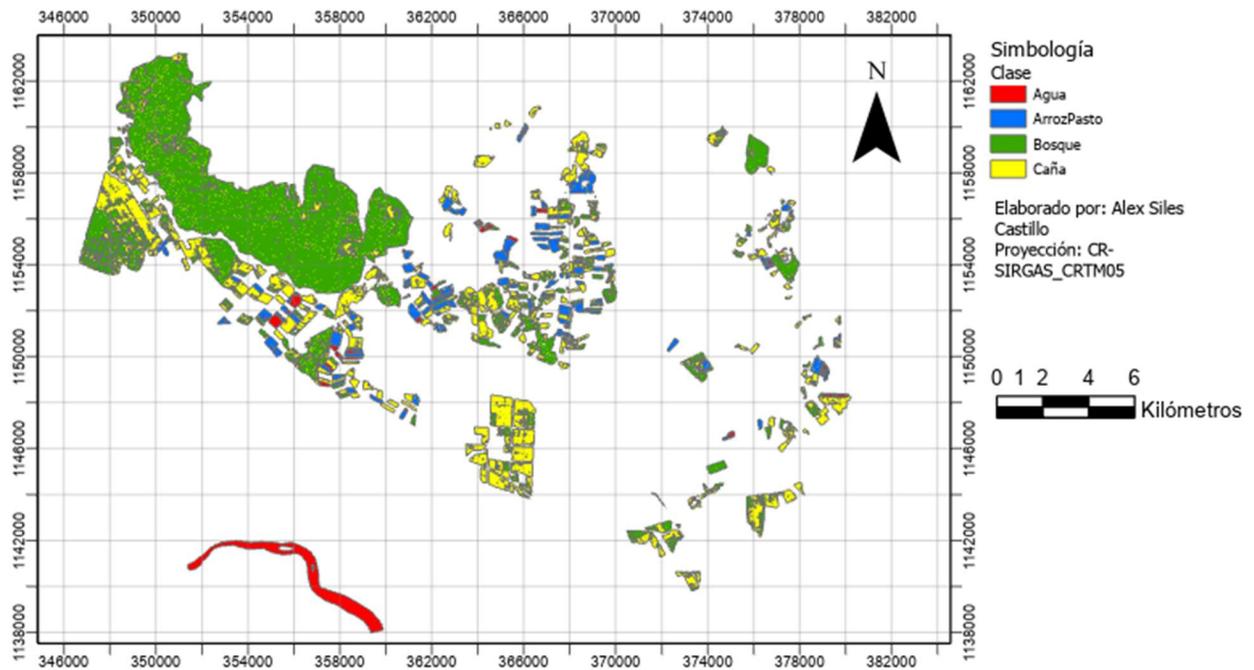


Figura B 6. Clasificación de la zona de estudio para el producto NDVI del 30 de noviembre del 2019.

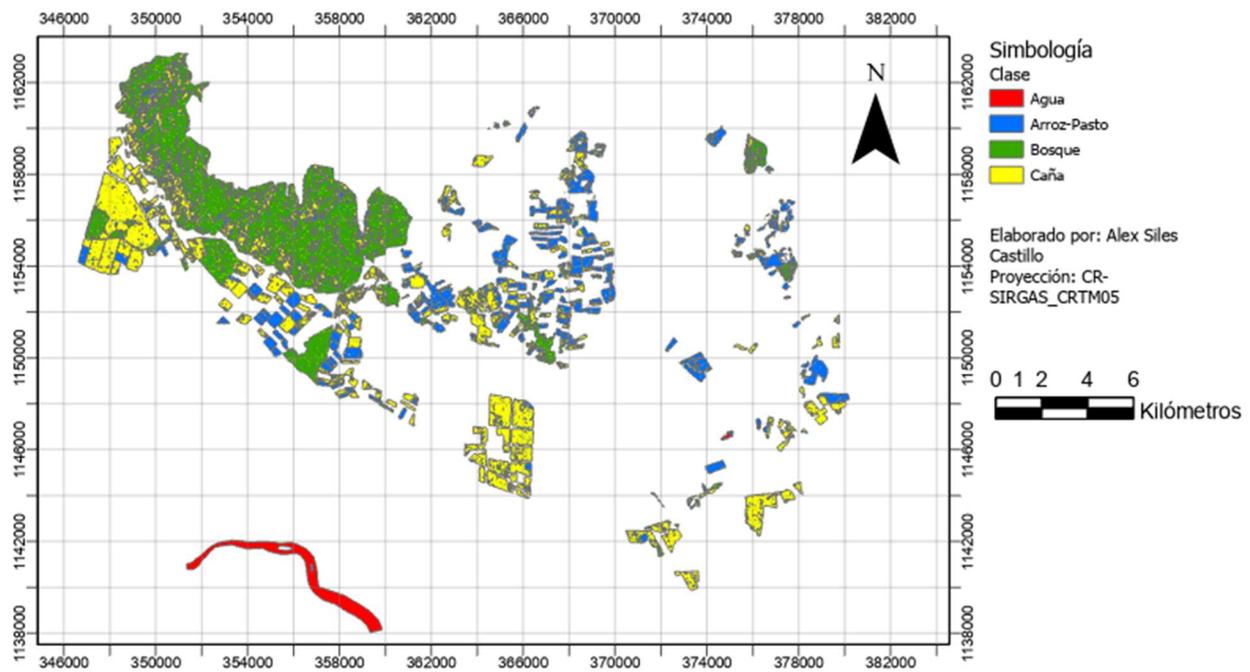


Figura B 7. Clasificación de la zona de estudio para el análisis multitemporal sin NDVI con una ventana de 5x5.

Tabla B 1. Métricas de rendimiento para el análisis multitemporal para ventana 5x5.

Kernel	C	Clase	Precision	Recall	F1	Support
Linear	1	Agua	0,89	0,99	0,94	141
		Arroz-Pasto	0,87	0,49	0,63	141
		Bosque	0,75	0,89	0,81	141
		Caña	0,63	0,72	0,67	141
		Accuracy			0,77	564
		Macro avg	0,79	0,77	0,76	564
		Weighted avg	0,79	0,77	0,76	564
		Índice Kappa	0,695			
Radial basis	1	Agua	0,97	0,99	0,98	141
		Arroz-Pasto	0,93	0,62	0,74	141
		Bosque	0,78	0,9	0,84	141
		Caña	0,67	0,72	0,72	141
		Accuracy			0,82	564
		Macro avg	0,84	0,82	0,82	564
		Weighted avg	0,84	0,82	0,82	564
		Índice Kappa	0,761			
Poly 2	1	Agua	0,84	0,99	0,91	141
		Arroz-Pasto	0,85	0,5	0,63	141
		Bosque	0,8	0,89	0,85	141
		Caña	0,68	0,76	0,72	141
		Accuracy			0,79	564
		Macro avg	0,79	0,79	0,78	564
		Weighted avg	0,79	0,79	0,78	564
		Índice Kappa	0,714			
Poly 3	1	Agua	0,76	0,99	0,86	141
		Arroz-Pasto	0,79	0,46	0,58	141
		Bosque	0,85	0,89	0,87	141
		Caña	0,75	0,81	0,78	141
		Accuracy			0,79	564
		Macro avg	0,79	0,79	0,77	564
		Weighted avg	0,79	0,79	0,77	564
		Índice Kappa	0,714			
Poly 4	1	Agua	0,72	0,99	0,83	141
		Arroz-Pasto	0,71	0,42	0,53	141
		Bosque	0,86	0,89	0,87	141
		Caña	0,76	0,77	0,76	141
		Accuracy			0,76	564
		Macro avg	0,76	0,76	0,75	564
		Weighted avg	0,76	0,76	0,75	564
		Índice Kappa	0,686			

Tabla B 2. Métricas de rendimiento para el análisis multitemporal para ventana 7x7.

Kernel	C	Clase	Precision	Recall	F1	Support
Linear	1	Agua	0,91	0,99	0,94	141
		Arroz-Pasto	0,9	0,49	0,63	141
		Bosque	0,75	0,89	0,81	141
		Caña	0,61	0,72	0,66	141
		Accuracy			0,77	564
		Macro avg	0,79	0,77	0,76	564
		Weighted avg	0,79	0,77	0,76	564
		Índice Kappa	0,697			
Radial basis	1	Agua	0,99	0,99	0,99	141
		Arroz-Pasto	0,92	0,6	0,73	141
		Bosque	0,8	0,89	0,84	141
		Caña	0,65	0,8	0,72	141
		Accuracy			0,82	564
		Macro avg	0,84	0,82	0,82	564
		Weighted avg	0,84	0,82	0,82	564
		Índice Kappa	0,759			
Poly 2	1	Agua	0,83	0,99	0,9	141
		Arroz-Pasto	0,9	0,49	0,63	141
		Bosque	0,82	0,88	0,85	141
		Caña	0,69	0,82	0,75	141
		Accuracy			0,79	564
		Macro avg	0,81	0,79	0,78	564
		Weighted avg	0,81	0,79	0,78	564
		Índice Kappa	0,726			
Poly 3	1	Agua	0,77	0,99	0,86	141
		Arroz-Pasto	0,79	0,48	0,59	141
		Bosque	0,85	0,87	0,86	141
		Caña	0,73	0,79	0,76	141
		Accuracy			0,78	564
		Macro avg	0,78	0,78	0,77	564
		Weighted avg	0,78	0,78	0,77	564
		Índice Kappa	0,709			
Poly 4	1	Agua	0,73	0,99	0,84	141
		Arroz-Pasto	0,72	0,43	0,54	141
		Bosque	0,88	0,86	0,87	141
		Caña	0,74	0,79	0,76	141
		Accuracy			0,77	564
		Macro avg	0,77	0,77	0,75	564
		Weighted avg	0,77	0,77	0,75	564
		Índice Kappa	0,688			

Tabla B 3. Métricas de rendimiento para el análisis de sensibilidad del valor de regularización para la ventana 5x5.

Kernel	C	Clase	Precision	Recall	F1	Support
Radial basis	0,5	Agua	0,97	0,98	0,98	141
		Arroz-Pasto	0,91	0,6	0,72	141
		Bosque	0,77	0,89	0,83	141
		Caña	0,64	0,77	0,7	141
		Accuracy			0,81	564
		Macro avg	0,82	0,81	0,81	564
		Weighted avg	0,82	0,81	0,81	564
		Índice Kappa	0,742			
Radial basis	1	Agua	0,97	0,99	0,98	141
		Arroz-Pasto	0,93	0,62	0,74	141
		Bosque	0,78	0,9	0,84	141
		Caña	0,67	0,72	0,72	141
		Accuracy			0,82	564
		Macro avg	0,84	0,82	0,82	564
		Weighted avg	0,84	0,82	0,82	564
		Índice Kappa	0,761			
Radial basis	2	Agua	0,99	0,99	0,99	141
		Arroz-Pasto	0,92	0,65	0,76	141
		Bosque	0,8	0,89	0,85	141
		Caña	0,67	0,8	0,73	141
		Accuracy			0,83	564
		Macro avg	0,85	0,83	0,83	564
		Weighted avg	0,85	0,83	0,83	564
		Índice Kappa	0,775			
Radial basis	3	Agua	0,99	0,99	0,99	141
		Arroz-Pasto	0,92	0,69	0,79	141
		Bosque	0,82	0,9	0,86	141
		Caña	0,7	0,81	0,75	141
		Accuracy			0,85	564
		Macro avg	0,86	0,85	0,85	564
		Weighted avg	0,86	0,85	0,85	564
		Índice Kappa	0,794			
Radial basis	4	Agua	0,99	0,99	0,99	141
		Arroz-Pasto	0,93	0,7	0,8	141
		Bosque	0,82	0,9	0,86	141
		Caña	0,71	0,82	0,76	141
		Accuracy			0,85	564
		Macro avg	0,86	0,85	0,85	564
		Weighted avg	0,86	0,85	0,85	564
		Índice Kappa	0,801			

Tabla B 3. (Continuación)

Kernel	C	Clase	Precision	Recall	F1	Support
Radial basis	5	Agua	0,99	0,99	0,99	141
		Arroz-Pasto	0,93	0,71	0,8	141
		Bosque	0,82	0,9	0,86	141
		Caña	0,71	0,82	0,76	141
		Accuracy			0,85	564
		Macro avg	0,86	0,85	0,85	564
		Weighted avg	0,86	0,85	0,85	564
		Índice Kappa	0,804			
Radial basis	10	Agua	0,99	0,99	0,99	141
		Arroz-Pasto	0,92	0,72	0,81	141
		Bosque	0,85	0,9	0,88	141
		Caña	0,73	0,85	0,79	141
		Accuracy			0,87	564
		Macro avg	0,87	0,87	0,87	564
		Weighted avg	0,87	0,87	0,87	564
		Índice Kappa	0,820			
Radial basis	15	Agua	1	0,99	0,99	141
		Arroz-Pasto	0,91	0,74	0,82	141
		Bosque	0,85	0,91	0,88	141
		Caña	0,74	0,84	0,79	141
		Accuracy			0,87	564
		Macro avg	0,88	0,87	0,87	564
		Weighted avg	0,88	0,87	0,87	564
		Índice Kappa	0,827			
Radial basis	20	Agua	1	0,99	0,99	141
		Arroz-Pasto	0,9	0,74	0,81	141
		Bosque	0,85	0,91	0,88	141
		Caña	0,74	0,84	0,79	141
		Accuracy			0,87	564
		Macro avg	0,88	0,87	0,87	564
		Weighted avg	0,88	0,87	0,87	564
		Índice Kappa	0,825			

Tabla B 4. Métricas de rendimiento para el análisis de sensibilidad del valor de regularización para la ventana 7x7.

Kernel	C	Clase	Precision	Recall	F1	Support
Radial basis	0,5	Agua	0,98	0,99	0,98	141
		Arroz-Pasto	0,92	0,59	0,72	141
		Bosque	0,78	0,88	0,82	141
		Caña	0,63	0,77	0,7	141
		Accuracy			0,81	564
		Macro avg	0,83	0,81	0,81	564
		Weighted avg	0,83	0,81	0,81	564
		Índice Kappa	0,742			
Radial basis	1	Agua	0,99	0,99	0,99	141
		Arroz-Pasto	0,92	0,6	0,73	141
		Bosque	0,8	0,89	0,84	141
		Caña	0,65	0,8	0,72	141
		Accuracy			0,82	564
		Macro avg	0,84	0,82	0,82	564
		Weighted avg	0,84	0,82	0,82	564
		Índice Kappa	0,759			
Radial basis	2	Agua	0,99	0,99	0,99	141
		Arroz-Pasto	0,94	0,62	0,74	141
		Bosque	0,83	0,87	0,85	141
		Caña	0,65	0,84	0,74	141
		Accuracy			0,83	564
		Macro avg	0,85	0,83	0,83	564
		Weighted avg	0,85	0,83	0,83	564
		Índice Kappa	0,773			
Radial basis	3	Agua	0,99	0,99	0,99	141
		Arroz-Pasto	0,93	0,66	0,77	141
		Bosque	0,83	0,87	0,85	141
		Caña	0,67	0,84	0,75	141
		Accuracy			0,84	564
		Macro avg	0,86	0,84	0,84	564
		Weighted avg	0,86	0,84	0,84	564
		Índice Kappa	0,785			
Radial basis	4	Agua	0,99	0,99	0,99	141
		Arroz-Pasto	0,93	0,67	0,78	141
		Bosque	0,83	0,88	0,86	141
		Caña	0,69	0,84	0,76	141
		Accuracy			0,85	564
		Macro avg	0,86	0,85	0,85	564
		Weighted avg	0,86	0,85	0,85	564
		Índice Kappa	0,794			

Tabla B 4. (Continuación).

Kernel	C	Clase	Precision	Recall	F1	Support
Radial basis	5	Agua	0,99	0,99	0,99	141
		Arroz-Pasto	0,93	0,69	0,79	141
		Bosque	0,85	0,89	0,87	141
		Caña	0,7	0,86	0,77	141
		Accuracy			0,85	564
		Macro avg	0,77	0,85	0,85	564
		Weighted avg	0,77	0,85	0,85	564
		Índice Kappa	0,806			
Radial basis	10	Agua	1	0,99	0,99	141
		Arroz-Pasto	0,94	0,72	0,82	141
		Bosque	0,89	0,89	0,89	141
		Caña	0,73	0,9	0,8	141
		Accuracy			0,87	564
		Macro avg	0,89	0,87	0,87	564
		Weighted avg	0,89	0,87	0,87	564
		Índice Kappa	0,832			
Radial basis	15	Agua	1	0,99	0,99	141
		Arroz-Pasto	0,92	0,74	0,82	141
		Bosque	0,89	0,89	0,89	141
		Caña	0,73	0,89	0,8	141
		Accuracy			0,88	564
		Macro avg	0,89	0,88	0,88	564
		Weighted avg	0,89	0,88	0,88	564
		Índice Kappa	0,835			
Radial basis	20	Agua	0,99	0,99	0,99	141
		Arroz-Pasto	0,91	0,74	0,82	141
		Bosque	0,9	0,89	0,9	141
		Caña	0,74	0,89	0,8	141
		Accuracy			0,88	564
		Macro avg	0,89	0,88	0,88	564
		Weighted avg	0,89	0,88	0,88	564
		Índice Kappa	0,835			

Tabla B 5. Métricas de rendimiento en el análisis de clases unidas.

Ventana 7x7 Clases de arroz y pasto unidas							
Cantidad de puntos por clase	Kernel	C	Clase	Precision	Recall	F1	Support
470	Radial basis	15	Agua	1	0,99	0,99	141
			Arroz-Pasto	0,92	0,74	0,82	141
			Bosque	0,89	0,89	0,89	141
			Caña	0,73	0,89	0,8	141
			Accuracy			0,88	564
			Macro avg	0,89	0,88	0,88	564
			Weighted avg	0,89	0,88	0,88	564
			Índice kappa	0,835			

Tabla B 6. Métricas de rendimiento en el análisis de clases separadas.

Ventana 7x7 Clases de arroz y pasto separadas							
Cantidad de puntos por clase	Kernel	C	Clase	Precision	Recall	F1	Support
470	Radial basis	15	Agua	1	0,99	0,99	141
			Arroz	0,86	0,79	0,82	141
			Bosque	0,85	0,91	0,88	141
			Caña	0,74	0,84	0,79	141
			Pasto	0,69	0,65	0,67	141
			Accuracy			0,83	705
			Macro avg	0,83	0,83	0,83	705
			Weighted avg	0,83	0,83	0,83	705
Índice kappa	0,784						

Tabla B 7. Métricas de rendimiento del análisis monotemporal para la imagen SAR del 20 de septiembre.

Ventana 7x7 single 20 sept							
Cantidad de puntos por clase	Kernel	C	Clase	Precision	Recall	F1	Support
470	Radial basis	15	Agua	0,98	0,96	0,96	141
			Arroz	0,45	0,06	0,11	141
			Bosque	0,53	0,58	0,55	141
			Caña	0,36	0,45	0,4	141
			Pasto	0,42	0,65	0,51	141
			Accuracy			0,54	705
			Macro avg	0,55	0,54	0,51	705
			Weighted av	0,55	0,54	0,51	705
Índice kappa	0,422						

Tabla B 8. Métricas de rendimiento del análisis monotemporal para la imagen SAR del 19 de noviembre.

Ventana 7x7 single 19 nov							
Cantidad de puntos por clase	Kernel	C	Clase	Precision	Recall	F1	Support
470	Radial basis	15	Agua	0,98	0,96	0,97	141
			Arroz	0,5	0,21	0,29	141
			Bosque	0,58	0,67	0,62	141
			Caña	0,34	0,61	0,44	141
			Pasto	0,48	0,33	0,39	141
			Accuracy			0,55	705
			Macro avg	0,58	0,55	0,54	705
			Weighted av	0,58	0,55	0,54	705
Indice kappa				0,443			

Tabla B 9. Métricas de rendimiento del análisis monotemporal para la imagen SAR del 25 de diciembre.

Ventana 7x7 single 25 dic							
Cantidad de puntos por clase	Kernel	C	Clase	Precision	Recall	F1	Support
470	Radial basis	15	Agua	0,98	0,9	0,94	141
			Arroz	0,63	0,36	0,46	141
			Bosque	0,66	0,77	0,71	141
			Caña	0,46	0,7	0,55	141
			Pasto	0,56	0,46	0,5	141
			Accuracy			0,64	705
			Macro avg	0,66	0,64	0,63	705
			Weighted av	0,66	0,64	0,63	705
Indice kappa				0,548			

Tabla B 10. Métricas de rendimiento del análisis monotemporal para la imagen NDVI del 30 de noviembre.

NDVI 30 Nov							
Cantidad de puntos por clase	Kernel	C	Clase	Precision	Recall	F1	Support
470	Radial basis	15	Agua	0,9	0,97	0,93	141
			Arroz	0,66	0,53	0,59	141
			Bosque	0,52	0,87	0,65	141
			Caña	0,44	0,62	0,52	141
			Pasto	0,75	0,02	0,04	141
			Accuracy			0,60	705
			Macro avg	0,65	0,60	0,55	705
			Weighted av	0,65	0,60	0,55	705
			Indice kappa	0,505			

Tabla B 11. Métricas de rendimiento del análisis monotemporal para la imagen NDVI del 25 de diciembre.

NDVI 25 dic							
Cantidad de puntos por clase	Kernel	C	Clase	Precision	Recall	F1	Support
470	Radial basis	15	Agua	0,9	0,93	0,92	141
			Arroz	0,72	0,7	0,71	141
			Bosque	0,69	0,81	0,75	141
			Caña	0,49	0,63	0,55	141
			Pasto	0,47	0,26	0,33	141
			Accuracy			0,66	705
			Macro avg	0,65	0,66	0,65	705
			Weighted av	0,65	0,66	0,65	705
			Indice kappa	0,580			

Tabla B 12. Métricas de rendimiento del análisis multitemporal sin NDVI para la ventana de 5x5.

Ventana 5x5 Sin NDVI								
Cantidad de	Kernel	C	Clase	Precision	Recall	F1	Support	
470	Radial basis	15	Agua	1	0,99	0,99	141	
			Arroz-Pasto	0,79	0,67	0,72	141	
			Bosque	0,82	0,66	0,73	141	
			Caña	0,58	0,79	0,66	141	
			Accuracy			0,77	564	
			Macro avg	0,8	0,77	0,78	564	
			Weighted av	0,8	0,77	0,78	564	
			Indice kappa	0,700				

Tabla B 13. Métricas de rendimiento del análisis multitemporal sin NDVI para la ventana de 7x7.

Ventana 7x7 Sin NDVI							
Cantidad de	Kernel	C	Clase	Precision	Recall	F1	Support
470	Radial basis	15	Agua	0,99	0,99	0,99	141
			Arroz-Pasto	0,82	0,7	0,75	141
			Bosque	0,87	0,74	0,8	141
			Caña	0,64	0,83	0,72	141
			Accuracy			0,81	564
			Macro avg	0,83	0,81	0,82	564
			Weighted av	0,83	0,81	0,82	564
			Indice kappa	0,752			

Tabla B 14. Resumen de las principales métricas de rendimiento obtenidas para todos los escenarios evaluados.

Producto	Tamaño de ventana	Clases	Metodología	Valor de regularización	Kernel	Medida F1	Índice Kappa
SAR + NDVI	5x5	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	1	Polinomial grado 4	0,76	0,686
SAR + NDVI	5x5	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	1	Polinomial grado 3	0,79	0,714
SAR + NDVI	5x5	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	1	Polinomial grado 2	0,79	0,714
SAR + NDVI	5x5	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	1	Lineal	0,77	0,695
SAR + NDVI	5x5	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	1	Base Radial	0,82	0,761
SAR + NDVI	5x5	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	0,5	Base Radial	0,81	0,742
SAR + NDVI	5x5	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	2	Base Radial	0,83	0,775
SAR + NDVI	5x5	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	3	Base Radial	0,85	0,794
SAR + NDVI	5x5	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	4	Base Radial	0,85	0,801
SAR + NDVI	5x5	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	5	Base Radial	0,85	0,804
SAR + NDVI	5x5	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	10	Base Radial	0,87	0,820
SAR + NDVI	5x5	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	15	Base Radial	0,87	0,827
SAR + NDVI	5x5	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	20	Base Radial	0,87	0,825

Tabla B 14. (Continuación).

Producto	Tamaño de ventana	Clases	Metodología	Valor de regularización	Kernel	Medida F1	Índice Kappa
SAR + NDVI	7x7	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	1	Polinomial grado 4	0,76	0,688
SAR + NDVI	7x7	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	1	Polinomial grado 3	0,79	0,709
SAR + NDVI	7x7	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	1	Polinomial grado 2	0,79	0,726
SAR + NDVI	7x7	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	1	Lineal	0,77	0,697
SAR + NDVI	7x7	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	1	Base Radial	0,82	0,759
SAR + NDVI	7x7	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	0,5	Base Radial	0,81	0,742
SAR + NDVI	7x7	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	2	Base Radial	0,83	0,773
SAR + NDVI	7x7	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	3	Base Radial	0,84	0,785
SAR + NDVI	7x7	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	4	Base Radial	0,85	0,794
SAR + NDVI	7x7	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	5	Base Radial	0,85	0,806
SAR + NDVI	7x7	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	10	Base Radial	0,87	0,832
SAR + NDVI	7x7	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	15	Base Radial	0,88	0,835
SAR + NDVI	7x7	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	20	Base Radial	0,88	0,835
SAR + NDVI	7x7	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	15	Base Radial	0,88	0,835
SAR + NDVI	7x7	Caña, Arroz, Pasto, Bosque y Agua	Multitemporal	15	Base Radial	0,83	0,784

Tabla B 14. (Continuación).

Producto	Tamaño de ventana	Clases	Metodología	Valor de regularización	Kernel	Medida F1	Índice Kappa
SAR	7x7	Caña, Arroz, Pasto, Bosque y Agua	Single 20 Sep	15	Base Radial	0,54	0,422
SAR	7x7	Caña, Arroz, Pasto, Bosque y Agua	Single 19 Nov	15	Base Radial	0,55	0,443
SAR	7x7	Caña, Arroz, Pasto, Bosque y Agua	Single 25 Dic	15	Base Radial	0,64	0,548
SAR	7x7	Caña, Arroz-Pasto, Bosque y Agua	Single 20 Sep	15	Base Radial	0,63	0,504
SAR	7x7	Caña, Arroz-Pasto, Bosque y Agua	Single 19 Nov	15	Base Radial	0,68	0,570
SAR	7x7	Caña, Arroz-Pasto, Bosque y Agua	Single 25 Dic	15	Base Radial	0,74	0,660
NDVI	-	Caña, Arroz-Pasto, Bosque y Agua	Single 30 Nov	15	Base Radial	0,71	0,619
NDVI	-	Caña, Arroz-Pasto, Bosque y Agua	Single 25 Dic	15	Base Radial	0,76	0,676
NDVI	-	Caña, Arroz, Pasto, Bosque y Agua	Single 30 Nov	15	Base Radial	0,60	0,505
NDVI	-	Caña, Arroz, Pasto, Bosque y Agua	Single 25 Dic	15	Base Radial	0,66	0,580
SAR	5x5	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	15	Base Radial	0,77	0,700
SAR	7x7	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	15	Base Radial	0,81	0,752
NDVI	-	Caña, Arroz-Pasto, Bosque y Agua	Multitemporal	15	Base Radial	0,81	0,742