

JORGE ELIECER GALVIS DAZA

Universidad de Manizales
Facultad de Ciencias e Ingeniería
Maestría en Tecnologías de la Información Geográfica
Manizales, 2022

USO DE TELEDETECCIÓN EN LA MODELACION DEL CARBONO DE LA BIOMASA HERBÁCEA EN TERRENOS REHABILITADOS POR LA EMPRESA DRUMMOND LTD EN EL DEPARTAMENTO DEL CESAR-COLOMBIA

JORGE ELIECER GALVIS DAZA

Propuesta de trabajo de grado presentado como requisito parcial para optar al
título de Magíster en Tecnologías de la Información Geográfica

Director (a):
M.Sc., Dino Carmelo Manco Jaraba

Línea de Investigación:
Análisis y Modelamiento Espacial
Grupo de Investigación y Desarrollo en Informática y Telecomunicaciones

Universidad de Manizales
Facultad de Ciencias e Ingeniería
Maestría en Tecnologías de la Información Geográfica
Manizales, 2022

Resumen

La empresa Drummond LTD ubicada en el departamento del Cesar Colombia basa su actividad económica en la extracción de carbón utilizando métodos de explotación a cielo abierto, en zonas ya explotadas a realizado labores de rehabilitación mediante reforestación utilizando árboles y arbustos nativos. En la investigación se modelo el carbono contenido por la biomasa herbácea en terrenos rehabilitados por la empresa Drummond, mediante la correlación de datos de campo y sensores remotos utilizando el método de regresión lineal, maquinas vectoriales de soporte, random forest y k-vecinos más cercanos. Los resultados obtenidos muestran que la especie utilizada para la reforestación *Andropogon gayanus* (Pasto carimagua), presentan concentraciones con promedios de carbono de 49.83% y el modelo de regresión Random forest en combinación con el índice de vegetación MSAVI2 muestra la mejor correlación estadística, con un valor de $R^2= 0.935$, $RMSE=0.78$ y $MAE=0.585$, para una concentración máxima de contenido de carbono contenido en la biomasa herbácea de 13.04 kg C/m².

Palabras clave: Sentinel, biomasa herbácea, modelos de regresión, carbono, gases de efecto invernadero.

Contenido

1. PLANTEAMIENTO DEL PROBLEMA Y JUSTIFICACIÓN	16
1.1 Descripción del área problemática	16
1.2 Formulación del problema	19
1.3 Justificación	20
2. ANTECEDENTES	22
3. OBJETIVOS	32
3.1 Objetivo general	32
3.2 Objetivos específicos.....	32
4. REFERENTE CONTEXTUAL.....	33
5.REFERENTE NORMATIVO	34
6. REFERENTE TEORICO.....	36
6.1 Minería y gases de efecto invernadero.	36
6.2.1 Gas efecto de invernadero.....	37
6.2.2 Ciclo del carbono.	37
6.2.3 Sumideros de carbono.....	38
6.3 Cuantificación de carbono en biomasa herbácea	41
6.3.1Biomasa	41
6.3.2 Cuantificación de carbono en biomasa herbácea	41
6.4. Teledetección en el estudio de biomasa.	43
6.4.1 Sentitel-2.....	43
6.4.2 Índice de vegetación (IV).	44

VI USO DE TELEDETECCIÓN EN LA ESTIMACIÓN DEL CARBONO DE LA BIOMASA HERBÁCEA EN TERRENOS REHABILITADOS POR LA EMPRESA DRUMMOND LTD EN EL DEPARTAMENTO DEL CESAR-COLOMBIA

6.5 Modelos de ajuste para la estimación de la biomasa herbácea	45
6.5.1 Mínimos cuadrados ordinarios	45
6.5.2 <i>Random forest (RF)</i>	49
6.5.3 Regresión maquina vectorial de soporte (SVM)	50
6.5.4 K-vecinos más cercanos (K-NN).....	51
7. HIPÓTESIS DE INVESTIGACIÓN	53
8. METODOLOGÍA.....	54
8.1 Enfoque metodológico	54
8.2 Tipo de Estudio.....	54
8.3 Diseño de la Investigación.....	54
8.6 Operacionalización de la variable.....	57
8.7 Técnicas e instrumentos de recolección de información.	57
9. RESULTADOS Y DISCUSIÓN.....	58
9.1 Datos preliminares y trabajo de campo.	58
Fuente. Departamento ambiental Drummond.....	61
9.1.1. Muestreo.	61
9.2 Índices de vegetación	64
9.3 Estadística descriptiva de los datos de trabajo.	66
9.3. Modelos para cálculo de carbono en vegetación herbácea	67
9.3.1 Regresión por mínimos cuadrados ordinarios (OLS).....	67
9.3.2 K-vecinos más cercanos (K-NN).....	70
9.3.3 Maquina vectorial de soporte (SVM).....	73
9.3.4. <i>Random forest (RF)</i>	75

9.4 Discusión general	78
10. CONCLUSIONES	81
11. RECOMENDACIONES	82
11. BIBLIOGRAFÍA.....	83
ANEXOS.....	103

Lista de figuras

Figura 1. Producción de carbón en el departamento del Cesar 2016 y 2021	17
Figura 2. Localización Drummond	18
Figura 3. Almacenamiento y flujo de carbono en bosques	39
Figura 4. Formas de acumulación del carbono terrestre	40
Figura 5. Esquema Random forest	49
Figura 6. Esquema modelo de regresión vectorial de soporte.....	50
Figura 7. Esquema K - vecinos más cercanos.....	51
Figura 8. Esquema de metodología de trabajo	56
Figura 9. Análisis multitemporal	59
Figura 10. Análisis multitemporal (Continuación).....	60
Figura 11. Áreas en proceso de rehabilitación – Combinación 843.....	61
Figura 12. Pasto sembrado.....	61
Figura 13. Punto de muestreo.....	62
Figura 14. Procesamiento de muestras en laboratorio	63
Figura 15. Matriz de correlación.....	67
Figura 16. Diagramas Índice de Vegetación vs Contenido de Carbono	68
Figura 17. Modelación de carbono OLS	70
Figura 18. RMSE vs K.....	71
Figura 19. Grafica de residuos KNN	71
Figura 20. Modelación de carbono KNN	72
Figura 21. Grafica de residuos SVM	74
Figura 22. Modelación de carbono SVM.....	75

Figura 23. Factor de importancia RF	76
Figura 24. Grafica de residuos RF	77
Figura 25. Modelación de carbono RF	78

Lista de tablas

Tabla 1. Características de las bandas Sentinel 2A.	44
Tabla 2. Índice de vegetación	47
Tabla 3. Operacionalización de variable.	57
Tabla 4. Georreferenciación y contenido de biomasa por parcela.....	62
Tabla 5. Contenido de carbono en muestras	64
Tabla 6. Índice de vegetación extraído para cada punto de muestreo.....	64
Tabla 7. Evaluación de modelos OLS.....	69
Tabla 8. Evaluación de modelo SVM con parámetros de cálculo por defecto	73
Tabla 9. Parámetros de cálculo para regresión SVM	73
Tabla 10. Evaluación de modelo SVM con parámetros optimizados	74
Tabla 11. Parámetros de cálculo RF.....	76

Lista de anexos

Anexo 1. Algoritmo para regresión por Minimos cuadrados ordinarios	104
Anexo 2. Algoritmo para regresión por K-vecinos mas cercanos	107
Anexo 3. Algoritmo para regresión por Maquina vectorial de soporte	110
Anexo 4. Algoritmo para regresión por Random forest	114

Lista de Símbolos y abreviaturas

Símbolos con letras latinas

Símbolo	Término	Unidad SI	Definición
C_{muestra}	Carbono en la muestra	kg C m ²	Ec. 1
BS	Peso de la Biomasa	<i>Kg</i>	Ec. 1
CF	Fracción de carbono en biomasa	C	Ec. 1
C	Cantidad de carbono en la biomasa	$\frac{t\ C}{ha}$	Ec. 2
$T_{\text{subparcela}}$	Tamaño de subparcela	m ²	Ec. 2

Abreviaturas

Abreviatura	Término
NDVI	Índice de Vegetación de la Diferencia Normalizada
SAVI	Índice de Vegetación Ajustado al Suelo
SR	Índice de Proporción Simple
GEI	Gases de efecto invernadero
IPCC	Panel intergubernamental del cambio climático
NDMI	Índice de Humedad Normalizada
UNFCCC	Convención Marco de Naciones Unidas sobre el Cambio Climático
CONPES	Consejo Nacional de Política Económica y Social
SIG	Sistemas de información Geográfica
FAO	Organización de las Naciones Unidas para la Alimentación y la Agricultura
MSR	Proporción Modificada Simple
OLS	Regresión lineal
PLS	Regresión por mínimos cuadrados parciales
SVM	Máquinas de vectores de soporte

Símbolos químicos

Abreviatura	Término
H_2O	<i>Vapor de agua</i>
CO_2	<i>Dióxido de carbono</i>
N_2O	<i>óxido nitroso</i>
C	Carbono
CH_4	<i>Metano</i>
O_3	Ozono

Introducción

Alrededor del mundo la minería se ha convertido en la principal fuente de obtención de recursos minerales metálicos, no metálicos y fósiles utilizados en diversos procesos industriales. Durante la extracción de estos recursos se generan una serie de impactos a los componentes agua, suelo y atmosfera (Marove et al., 2022). Las afectaciones a la atmosfera son producto de la combustión de combustibles fósiles y emisiones fugitivas por la autocombustión del carbón, lo genera gases de efecto de invernadero como el CO₂ causantes cambio climático (Guimar et al., 2018).

Las áreas afectadas durante los procesos extractivos deben ser rehabilitadas y/o reforestadas. La vegetación sembrada en zonas reforestadas cumple un rol importante en la mitigación de los impactos a la atmosfera pues puede fungir como un potencial sumidero de carbono (Placek-Lapaj et al., 2019).

En Colombia la empresa Drummond LTD realiza labores de extracción de carbón a cielo abierto en dos minas diferentes: Pribbenow y El Descanso. En Pribbenow se han rehabilitado más de 600 hectáreas utilizando vegetación nativa de la zona. Esto la hace un escenario ideal para realizar una investigación para generar conocimiento científico tendiente a mitigar impactos ambientales producidos por gases de efecto invernadero.

El objetivo de la investigación es modelar el carbono contenido por la biomasa aérea en terrenos rehabilitados mediante la adquisición de datos campo, uso de imágenes satelitales de mediana resolución y el análisis de correlación de variables utilizando técnicas de regresión y algoritmos de machine learning.

1. PLANTEAMIENTO DEL PROBLEMA Y JUSTIFICACIÓN

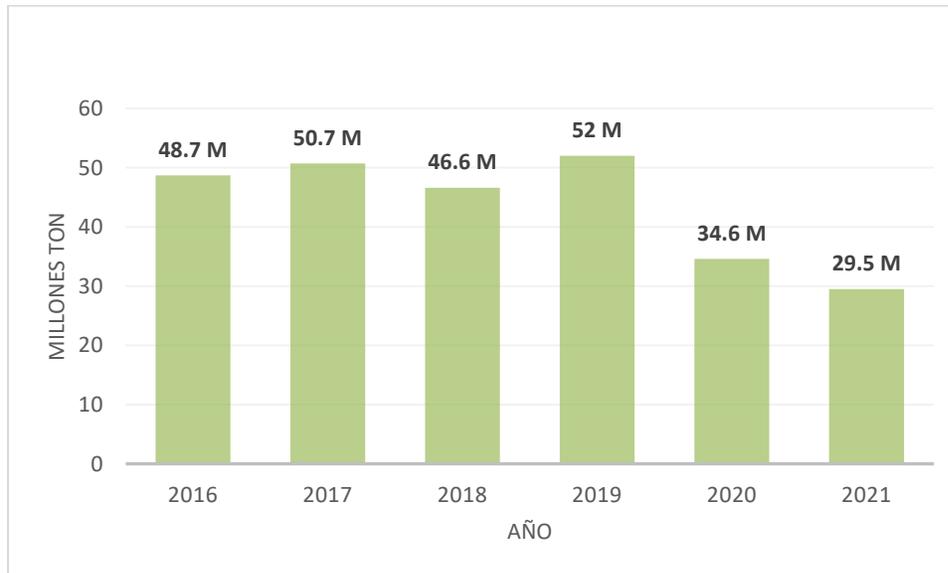
1.1 Descripción del área problemática

El departamento del Cesar se encuentra ubicado al noreste de Colombia, cuenta con una extensión de 22.905 km² y una población de 1.041.203 habitantes. Los principales sectores que mueven la economía del departamento son: la agronomía, agricultura y minería (Gobernación del Departamento del Cesar, 2020).

Entre los minerales extraídos en por empresas mineras son: 1) materiales de construcción entre los que predominan grava, arcilla, caliza y arena 2) metales principalmente hierro y cobre, 3) combustibles fósiles principalmente gas natural y Carbón térmico. Siendo la minería de carbón la de mayor producción. Las empresas Drummond, Carbones La Jagua, Prodeco, Colombian Natural Resources y el Consorcio Mineros Unidos son quienes explotan el carbón en el Cesar.

La producción de carbón en el departamento llegó a su punto más alto en el año 2016 con 52 millones de toneladas años, con la salida de operación de la multinacional Prodeco la en 2020 se registró un descenso en la producción, llegando a 29.5 millones de toneladas en el año 2021. Los datos de producción entre el año 2016 y 2020 se observan en la Figura 1.

Figura 1. Producción de carbón en el departamento del Cesar 2016 y 2021

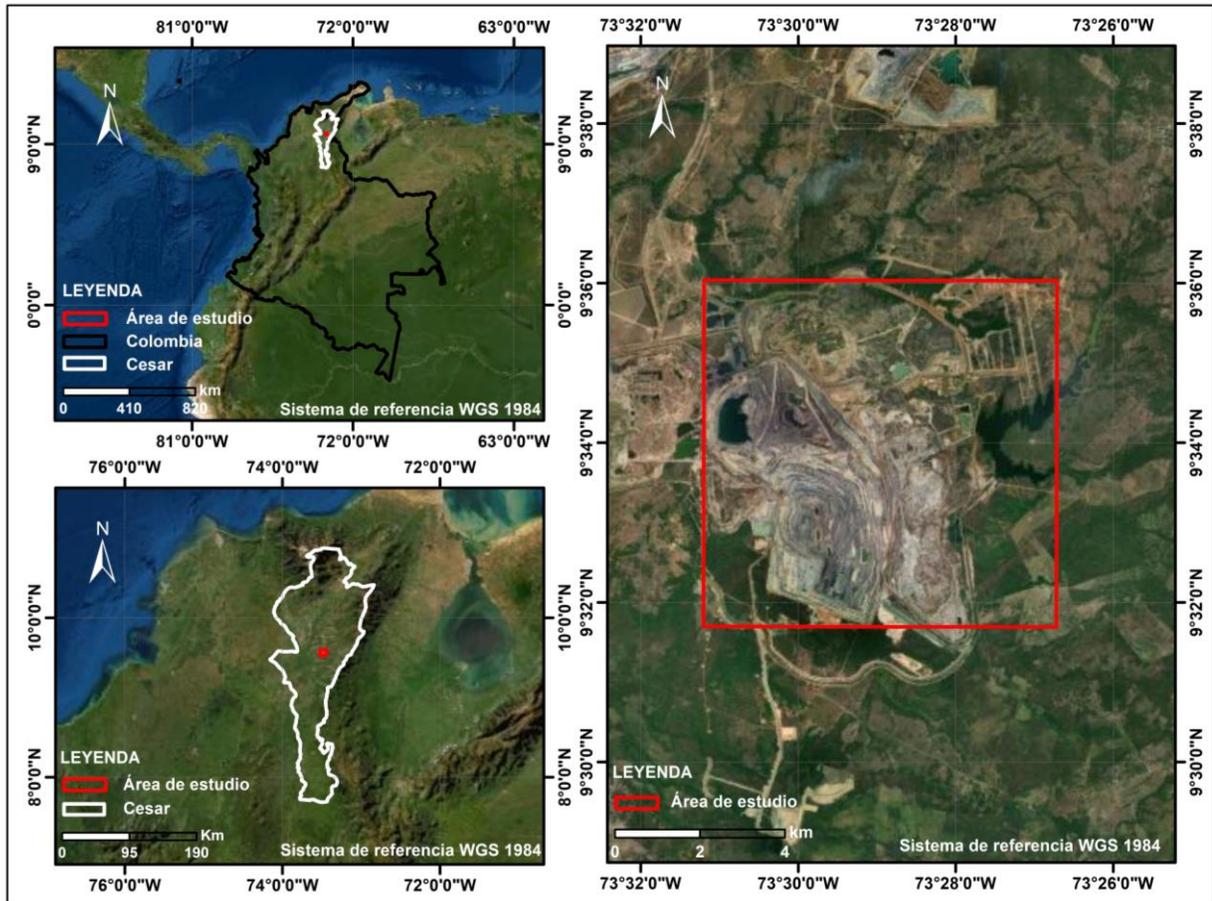


Fuente. (Gobernación del Departamento del Cesar, 2020).

Drummond inició sus trabajos exploratorios en 1990 y en 1995 comenzó la producción y exportación del mineral. Cuenta con dos minas: Pribbenow con una extensión de 6559.65 hectáreas y El Descanso con una extensión de 3157.7 hectáreas, la primera con influencia en los municipios de El Paso, La Jagua de Ibirico y Chiriguana y la segunda en El Paso, La Jagua de Ibirico, Chiriguana, Codazzi y Becerril.

18 USO DE TELEDETECCIÓN EN LA ESTIMACIÓN DEL CARBONO DE LA BIOMASA HERBÁCEA EN TERRENOS REHABILITADOS POR LA EMPRESA DRUMMOND LTD EN EL DEPARTAMENTO DEL CESAR-COLOMBIA

Figura 2. Localización Drummond



Fuente. Autor

Drummond aporta más de 7000 empleos directos e indirectos en la zona de influencia del proyecto minero, convirtiéndose en la mayor fuente generadora de empleo en el centro del Cesar. Adicionalmente, las regalías producidas ayudan al desarrollo de proyectos para el progreso de la región.

En la mina Pribbenow se han rehabilitado más de 600 hectáreas intervenidas por la minería utilizando árboles característicos de la zona como el *Pithecellobium*

saman (campano), *Mangifera indica* (mango), *Casearia sp* (roble), *Psidium guajava* (guayaba), *Ceiba pentandra* (ceiba) y *Pimenta dioica* (uvito) (Drummond Ltd., 2021).

1.2 Formulación del problema

La empresa Drummond LTD basa su actividad económica en la extracción de carbón térmico empleando métodos de explotación a cielo abierto, las diferentes etapas del proceso extractivo contribuyen a la emisión de gases de invernadero (GEI) causante de los cambios climáticos en sistemas naturales y humanos; la remoción de la cobertura vegetal en la fase inicial de la explotación, el consumo de combustible de los equipos mineros y la autocombustión de pilas y mantos de carbón son las principales fuentes de emisión de gases contaminantes.

Un porcentaje de los gases emitidos son absorbidos por la biomasa aérea presente en zonas rehabilitadas, mientras el resto de estos gases es liberado al ambiente generando un deterioro de la atmósfera.

Actualmente en la empresa no se conoce el potencial de secuestro de carbono de dichas áreas. Por tal motivo, esta investigación se proyecta con el fin de conocer el carbono almacenado en las áreas rehabilitadas y de esta manera contribuir al desarrollo sostenible del sector minero energético.

¿Cuál es el contenido de carbono contenido en la biomasa herbácea de las áreas rehabilitadas por la empresa por la empresa Drummond LTD?

1.3 Justificación

El motivo para el desarrollo de esta investigación se fundamenta en la importancia de generar conocimiento científico tendiente a mitigar los impactos ambientales producidos por la emisión de GEI durante la extracción de recursos minerales.

Así mismo, este estudio es importante debido a que apunta a modelar el almacenamiento de carbono de las áreas que las empresas mineras deben rehabilitar una vez culminada la actividad extractiva; lo que permitirá planificar la explotación mineral con miras a que la relación de dióxido de carbono producido versus la de carbono absorbido sea igual o menor.

El principio tenido en cuenta para el desarrollo de la investigación radica en que el dióxido de carbono producido por actividades antrópicas es absorbido de manera natural por las plantas durante el proceso de fotosíntesis, de esta forma la biomasa acumula carbono (C). De esta manera el proyecto busca modelar captura de C de las áreas rehabilitadas por la empresa Drummond LTD y así contribuir a la meta de mitigación de GEI por parte de Colombia; quien en el año 2015 se comprometió ante la secretaria de la Convención de Cambio Climático a reducir en un 20% las emisiones de GEI para el año 2030 (Ministerio de Ambiente y Desarrollo Sostenible, 2017).

Adicionalmente, el impacto de este proyecto gira en torno a la implementación de las acciones establecidas en la resolución 40807 correspondiente al Plan Integral

de Gestión de Cambio Climático para el Sector Minero Energético para mitigar el carbono expulsado a la atmosfera.

Drummond LTDA en sus dos minas tiene concesionadas un total de 38117.35 hectáreas y está obligada a recuperar todas las áreas impactadas negativamente por la explotación. Con el desarrollo de la investigación se establecerá una metodología para determinar la capacidad de almacenamiento de C de las áreas a recuperar.

Las herramientas de teledetección son indispensables para el desarrollo del proyecto. Permiten abarcar grandes extensiones de terreno disminuyendo tiempo y costos al momento de realizar un estudio.

Por último, la investigación busca la adaptación al cambio climático de las actividades desarrolladas por la empresa, lo que impacta el objetivo 13 (acción por el clima) de los objetivos de desarrollo sostenible (Programa de las Naciones Unidas para el Desarrollo, 2012).

2. ANTECEDENTES

Las actividades y operaciones de extracción de minerales en superficie generan impactos significativos en el suelo, agua y atmosfera. Castellanos Barliza, León Peláez, y Campo (2018) sostiene que las operaciones en minería de superficie degradan las características físicas y químicas de suelos limitando la instauración de vegetación nativa; en estudios realizados en áreas reforestadas por la multinacional Cerrejon en Colombia los investigadores observaron que las recuperaciones parciales de la química del suelo mejoran el secuestro de Carbón y ayuda mitigar los efectos del cambio climático. Tripathi, Singh, y Nathanail (2014) sostiene que la actividad minera genera impactos ambientales negativos sobre el suelo vegetación y atmosfera, pero cuando estas áreas son reforestadas pueden actuar como sumideros de carbono. Los Autores, en la investigación realizada concluyen que el restablecimiento de la vegetación en zonas ocupadas por la minería mejora la captura de carbono en: el suelo, biomasa aérea y microbiana. Shackelford, Miller, y Erickson (2017) sostienen que restaurar áreas afectadas por minería de superficie es una labor compleja que requiere *“logros complejos de cobertura, densidad, composición de la comunidad, riqueza de especies y elementos estructurales”*.

La minería aporta directa o indirectamente a la emisión de gases de efecto de invernadero contribuyendo así al deterioro de la atmosfera (Alvarez Davila y Perez, 2012). Muchos investigadores a nivel mundial han desarrollado investigaciones en torno a esta problemática (Xue, 2020; Wang et al., 2020). En su investigación Guimarães da Silva, Costa Muniz, Hoffmann, y Luz Lisbôa (2018) realizan un

inventario de los gases de efecto de invernadero producidos durante la producción de carbón en la mina Candiota en Brasil. Para la cuantificación de GEI producidos en la mina se consideró el consumo de combustible de todos los equipos que intervienen en el proceso minero, las emisiones fugitivas de metano de los mantos de carbón y los gases producidos por la auto combustión del carbón. La investigación fue realizada como una contribución al conocimiento de las emisiones de los GEI en la minería de superficie.

El problema ambiental ocasionado por la emisión de gases contaminantes ha captado la atención de gobiernos e investigadores a nivel mundial. Con el objetivo de reducir las emisiones GEI a la atmosfera en 1997 se adoptó el protocolo de Kyoto; en donde se compromete a los países industrializados a estabilizar las emisiones (Naciones Unidas, 1998). Así mismo, en el convenio firmado se reconoce el papel de la biomasa en la captura y almacenamiento de carbono.

Diversos estudios realizados en todo el mundo demuestran la capacidad de secuestro de carbono de los bosques (Khanal et al., 2017) ecosistemas forestales (Chuai et al., 2022), suelos (Post y Kwon, 2008) y biomasa herbácea (Li et al., 2013).

La biomasa aérea juega un papel fundamental en el ciclo global del carbono, por tal motivo, realizar un inventario y monitoreo preciso de la biomasa aérea es indispensable para comprender su rol como fuente de almacenamiento de carbono atmosférico (Dube et al., 2016).

La biomasa aérea puede calcularse utilizando métodos directos, los cuales requieren tomar muestras directas en campo. Esto implica derribar árboles para la medición de sapos (González Zárate, 2008), y métodos indirectos que estiman la biomasa a partir de medición de variables dasométricas tomadas en el área de estudio (López-Guerrero et al., 2011).

En los ecosistemas el carbono es almacenado en la biomasa aérea viva, biomasa subterránea viva, biomasa muerta en pie, biomasa muerta en piso y hojarasca o material descompuesto (Casiano-Domínguez et al., 2018). Además, describen que el contenido de carbono varía entre grupos de vegetación, se presenta un aumento en el contenido de C en plantaciones más antiguas y la captura varía en función del clima.

En Colombia (Yepes Quintero et al., 2011) en su investigación patrocinada por el Instituto de Hidrología, Meteorología y Estudios Ambientales-IDEAM establecieron la metodología para estimar el contenido de carbono en bosques, pastizales, raíces, detritos de madera en bosques naturales y en el suelo, a partir de muestras tomados en campo y su posterior análisis en laboratorio. Esta metodología ha sido ampliamente utilizada por diversos autores para la estimación de carbono capturado en los bosques. Zuluaga y Castro (2018) calcularon el carbono de la biomasa aérea hasta el fuste en arboles del bosque seco tropical en el Carmen de Bolívar, Colombia. Del mismo modo Salas, Alegre y Iglesias (2017), estiman la biomasa aérea y el carbono contenido en formaciones vegetales y en el suelo del bosque seco en la costa ecuatoriana.

Avtar et al., (2020) describen los sensores remotos como una tecnología ideal para el desarrollo de estudio en el marco del desarrollo sostenible con aplicaciones directas a la gestión de recursos naturales. En la revisión realizada por Ali et al., (2016), exponen que el monitoreo de biomasa herbácea puede realizarse aplicando diversas metodologías de medición terrestre como la evaluación visual, cosecha y secado de pasto seguido de análisis en laboratorio, medidor de placa ascendente y espectrometría de campo, para su posterior comprobación mediante teledetección óptica o por microondas. En 2006, Gao describe que mediante el uso de técnicas geoinformáticas especializadas aumenta la viabilidad de una cuantificación precisa de vegetación herbácea.

Rossini et al. (2014) estimaron la producción primaria de pasto en el cinturón subalpino de los Alpes italiano. Usando imágenes MODIS evaluaron diversos índices de vegetación, donde se obtuvo como resultado que el NDVI es el índice que mejor resultados ofrece a la hora de modelar la producción de pasto.

Cáceres, Martín, y Salas (2015) estiman la biomasa herbácea mediante el uso de índices espectrales de imágenes Landsat en la zona de Desea, provincia de Cáceres, España. La metodología planteada para el desarrollo de la investigación consta de 5 fases: selección y procesamiento de imágenes, muestreo en campo para medición de variables de la biomasa, generación de modelos para estimación de la biomasa y finalmente evaluación de la biomasa. Los resultados obtenidos muestran una buena correlación de los datos analizados.

En el trabajo realizado por Kolecka et al. (2018) realizan el mapeo de la frecuencia de corte de pasto en la region del canton Aargau, Suiza mediante el uso de una serie temporal de imágenes Sentinel-2. El mapeo de los pastos se realizó mediante la evaluación del índice de vegetación NDVI y el índice de corte de pasto GMI obteniendo resultados satisfactorios.

Los humedales corresponden a los ecosistemas terrestres con mayor capacidad de almacenamiento de carbono (Kayranli et al., 2009). En su investigación Naidoo et al. (2019) evalúan las capacidades de las imágenes Sentinel y WorldView en la estimación de la biomasa aérea herbácea en humedales en la cordillera de Amathole en la provincia de Eastern, Sur Africa. Obteniendo mejores resultados con los ensayos realizados con las imágenes WorldView.

Sousa et al. (2013) desarrollaron una metodología para cuantificar la biomasa forestal en la región Alto Alentejo al sur de Portugal basadas en el uso de imágenes de satélite QuickBird. Los investigadores aplicaron la técnica clasificación supervisada a la imagen de satélite para determinar las especies de árboles presente en el área de estudio, posteriormente elaboraron un inventario forestal tomando muestras en 26 parcelas, donde midieron variables dendrométricas de las especies, y finalmente, estimaron la biomasa a partir de ecuaciones alométricas. Los resultados obtenidos muestran una excelente correlación entre las variables analizadas con un error del 5%.

Para el desarrollo de estudios de modelamiento de Biomasa aérea y almacenamiento de carbono la teledetección es un aliado fundamental por su

capacidad para captar información en un rango amplio del espectro electromagnético (Marabel-García, 2016). Los detalles obtenidos por las imágenes dependen de las características del sensor utilizado (Chuvieco Salinero, 1995).

La estimación de biomasa es una labor compleja y la precisión de los resultados obtenidos dependen de la resolución espacial de la imagen utilizada. El uso de imágenes Ikonos proporcionan resultados precisos para trabajar a nivel local, imágenes de resolución media como Landsat son idóneas para la cuantificación de escala a nivel regional y las imágenes de resolución espacial gruesa como Modis pueden usarse cálculos a escala nacional o global (Lu, 2006).

En su investigación Banerjee y Raval (2017) resaltan la idoneidad del uso de sensores remotos (satélites, aviones, vehículos no tripulados) para el monitoreo de entornos mineros. En la misma línea, Petropoulos et al. (2013) aplican métodos Support Vector Machines en combinación con datos de imágenes LandsatTM en la gestión de impactos ocasionados por la actividad minera de superficie. Mohd Zaki y Abd Latif (2016) sostiene que la biomasa aérea y subterránea acumula grandes cantidades de carbono, y que la tecnología de la teledetección es un gran aliado para el mapeo de las estructuras de los bosques y de la de la biomasa aérea, proporcionando precisión espacial y temporal. En el mismo sentido, el Panel Intergubernamental de Expertos en Cambio Climático IPCC (2006) describe que los sensores remotos son adecuados para la cuantificación de carbono a partir de la biomasa.

La correlación de datos de campo y los sensores remotos se puede realizar con métodos de correlación estadísticos (Barrachina, Cristóbal, & Tulla, 2015) y técnicas de Machine learning (Pham et al., 2018; Vafaei et al., 2018).

Guerini et al. (2020) realizan un estudio experimental para determinar la biomasa en el pastizal ubicado en el Bioma Pampa Brasileño, sur de Brasil. Los investigadores usaron imágenes Sentinel-2, tomaron muestras de campo y aplicaron análisis de correlación lineal para la estimación de la biomasa; la evaluación de los resultados se realizó aplicando el coeficiente de correlación R^2 . El modelo resultante muestra que existe correlación significativa de los datos de campo y los valores extraídos de la imagen satelital para la estimación de biomasa herbácea.

Machine learning es una rama de la inteligencia artificial que permite a las máquinas mediante el uso de algoritmos aprender de los datos a analizar sin necesidad de estar programado para ello, son útiles en la descripción de patrones y correlación de datos (Michie, 1968). Entre las técnicas de machine learning más utilizadas para correlacionar datos de muestreo con índices de vegetación están el Random Forest (Vreugdenhil et al., 2018), Support Vector Regression (Navarro et al., 2019), Multi-Layer Perceptron Neural Networks (Brivio et al., 2003) y Support Vector Machine (Wang et al., 2019).

Múltiples investigadores a nivel mundial han empleado los datos de sensores remotos en investigaciones relacionadas con la biomasa aérea y la estimación de carbono capturado. Hou et al. (2015) analizaron los cambios en el almacenamiento de

carbono en el tiempo en la zona de Xuzhou, China empleando datos de sensores remotos entre los años 1987 y 2008. Las imágenes fueron clasificadas obteniendo como principales coberturas: tierras cultivadas, tierras forestales, agua, zona urbanizada y tierras abandonadas por minería. Posteriormente, estimaron el contenido de carbono de la biomasa en zonas mineras recuperadas y donde no se realizó actividad extractiva concluyendo que las biomásas en zonas antiguas de actividad minera contienen menos carbono.

La minería a cielo abierto ha sido objeto de estudio de muchos científicos, debido a que se debe planificar las labores de recuperación ambiental en zonas alteradas por las actividades extractivas. Raval et al. (2014) usan la teledetección como una alternativa de bajo costo a escalas espaciales variadas para estimar la producción de biomasa en minas de carbón reforestadas en los Apalaches centrales, en Norte América. Los autores hacen seguimiento a parcelas experimentales utilizando entre los años 2008 y 2010 utilizando imágenes landsat-5 TM proporcionando resultados aceptables en el monitoreo de biomasa en áreas recuperadas.

La investigación realizada por Pietrzykowski y Daniels (2014) estiman el secuestro de carbono en suelos recuperados en la mina de carbón Belchatow en Polonia. Se tomaron muestras en 32 parcelas y combinaron datos de los tipos de suelo y biomasa para determinar el potencial de captura de carbono después de la minería en la hojarasca, horizonte orgánico y biomasa leñosa. Un punto importante de

investigación muestra que en antiguas zonas de acopio de carbón el contenido de C retenido es mayor que en otras áreas.

Del mismo modo, Das y Maiti (2016) calculan la capacidad de almacenamiento de carbono de plantaciones de 4 años en zonas reforestadas de la mina de carbón Jharia en la India. Se recolectaron muestras del suelo, biomasa aérea y biomasa subterránea donde encontraron que el carbono secuestrado varía dependiendo de variedad de la vegetación, tipo de sustrato y condiciones climáticas. Adicionalmente, Ahirwal y Maiti (2017) evalúan la calidad del suelo y el potencial de secuestro de carbono de terrenos recuperados después de las faenas mineras en la mina de carbón Rohini en el estado de Jharkhand, India. Tomaron muestras en 20 parcelas de áreas reforestadas con edad joven, intermedia y avanzada. Los resultados obtenidos muestran que el mayor secuestro de carbono se presenta en zonas reforestadas mayores a 14 años.

La investigación realizada por Tao et al. (2016) cuantifican la precisión la biomasa forestal utilizando imágenes de satélite en escalas espaciales cambiantes en bosque de coníferas en las montañas de Sierra Nevada, California. En el estudio emplearon imágenes landsat las cuales se les realizó calibraron radiométrica y corrección atmosférica. Posteriormente las imágenes se volvieron a muestrear con resoluciones más gruesas empleando diversas técnicas de muestreo. Los resultados obtenidos muestran que la precisión del estudio es buena hasta una resolución espacial de 60 -90 m.

El fundamento científico para la cuantificación precisa de la biomasa utilizado por diversos investigadores describe que las mediciones de biomasa se interpolan, extrapolan o cartografían siguiendo la siguiente metodología: 1) clasificación de cobertura mediante muestreo en campo, 2) cuantificación de biomasa a partir de regresiones y 3) relación entre biomasa insitu y datos obtenidos por imágenes satelitales. Sin embargo, los resultados obtenidos pueden ser variables Houghton et al. (2009).

Karan et al. (2016) monitorearon la recuperación de áreas impactadas en la mina de carbón a cielo abierto Jharia en la India utilizando imágenes Landsat. Estudiaron la salud de la vegetación utilizando los Índice de Vegetación Normalizada NDVI y el Índice de Humedad Normalizada NDMI. El resultado obtenido utilizando el NDVI muestra que la salud de la vegetación aumenta en zonas más antiguas.

Xin et al. (2020) investigaron la variación de carbono contenido en la biomasa aérea y subterránea en sistemas agropastoriles del norte de China. Aplicaron modelos de regresión lineal combinando datos de: inventarios de la vegetación realizados en la pradera entre los años 2002 y 2009, NDVI extraídos de imágenes MODIS, muestras de 255 parcelas, datos climáticos, textura de suelo y capacidad de carga del ganado. Los resultados obtenidos revelaron que existe una mayor acumulación de carbono en la biomasa subterránea.

3. OBJETIVOS

3.1 Objetivo general

Modelar el carbono contenido por la biomasa herbácea en terrenos rehabilitados por la empresa Drummond LTDA en el departamento del Cesar-Colombia, a partir de teledetección con imágenes ópticas Sentinel-2.

3.2 Objetivos específicos

- Cuantificar carbono contenido en la biomasa herbácea a partir de observaciones de datos tomados en campo.
- Identificar los índices de vegetación asociados a la biomasa herbácea en las áreas reforestadas, mediante el uso de herramientas de la teledetección.
- Realizar análisis de correlación estadística a partir de los índices de vegetación y datos de campo.

4. REFERENTE CONTEXTUAL

El estudio se llevará a cabo en las zonas reforestadas por parte de la empresa Drummond LTDA la cual desarrolla su actividad extractiva en el área de influencia de los municipios El Paso, La Jagua de Ibirico, Chiriguana, Codazzi y Becerril en el centro del departamento del Cesar.

La ejecución de la investigación se realizará en una superficie de 700 hectáreas reforestadas en la mina Pribbenow.

En la zona de estudio predomina el clima cálido con temperaturas que oscilan entre los 23°C y 37°C y cuenta con una precipitación promedio de 6 mm entre los meses de diciembre – enero y de 92 mm en el mes de octubre.

En los programas de reforestación implementados por parte de la empresa encontramos utilizan vegetación nativa y otras especies resistentes a las características biofísicas de la zona. En la primera fase de reforestación utilizan el pasto *Andropogon gayanus* (carimagua), *Panicum máximum* (guinea) y *Dichantium aristatum* (angletón); siendo el pasto carimagua el que mejores resultados a presentado en suelos postmineria. También se simbran arboles como el *Coccoloba acuminata* (maíz tostao), *Samanea saman* (Campano, genizaro), *Pseudosamanea guachapele* (Igua amarillo), *Tabebuia chrysea* (cañahuate), *Crescentia cujete* (totumo), *Guayacum officinales* (guayacan), *Ceratonia siliqua* (algarrobo) (Guerra Peralta, 2017).

5.REFERENTE NORMATIVO

El primer gran acuerdo internacional respecto a las políticas relacionadas con el cambio climático fue “El convenio marco de las Naciones Unidas sobre el Cambio Climático UNFCCC” en 1992. En donde se reconoce que las actividades antrópicas han aumentado la concentración de GEI en la atmosfera lo que ha intensificado el efecto invernadero natural, suscitando un aumento de la temperatura superficial.

En Colombia por medio de La Ley 164 de 1994 el Congreso de Colombia aprueba la UNFCCC donde el estado colombiano se compromete a formular e implementar políticas que incluyan medidas tendientes a mitigar el Cambio Climático.

El Consejo Nacional de Política Económica y Social (2011) expide el documento COMPES 3700 en el que se adopta la “Estrategia Institucional para la articulación de políticas y acciones en materia de Cambio Climático”, la cual busca garantizar que las entidades de orden nacional y territorial incorporen medidas para mitigar emisiones de gases de efecto invernadero.

En el ámbito internacional Colombia en el año 2015 se comprometió ante la secretaria de la Convención de Cambio Climático a reducir las emisiones de GEI en un 20% para el año 2030. Ministerio de Minas y Energía (2017).

El Ministerio de Ambiente y Desarrollo Sostenible (2017) suscribió el Plan Nacional de Cambio Climático el cual busca integrar en las políticas que influyen en el desarrollo del país acciones de mitigación y adaptación. Entre las estrategias propuestas está el desarrollo minero-energético bajo en carbono.

El Ministerio de Minas y Energía (2018) expide la resolución 40807 correspondiente al Plan Integral de Gestión de Cambio Climático para el Sector Minero Energético, el cual tiene por objetivo promover el desarrollo minero con bajas emisiones de carbono, mediante la estructuración de tres componentes: Mitigación, Adaptación y Gobernanza.

El componente de mitigación promueve la reducción de gas de efecto de invernadero en el sector Minero-Energético. Dentro de las estrategias de mitigación se propone la contabilidad de emisiones/absorciones de GEI por compensaciones por el uso y cambio de uso del suelo.

6. REFERENTE TEORICO

6.1 Minería y gases de efecto invernadero.

La **minería** se define como las técnicas y actividades relacionadas con el descubrimiento y posterior explotación de acumulaciones anómalas de minerales o rocas que tiene propiedades físicas y químicas definidas, útiles en la industria y genera un rendimiento económico con su extracción. En la práctica, la minería se clasifica en **Minería a cielo abierto** cuando las actividades y operaciones son realizadas en superficie y **Minería subterránea** cuando las actividades y operaciones desarrolladas bajo tierra. (Ministerio de minas y energía, 2003).

Durante las actividades de extracción de recursos minerales se genera bióxido de carbono durante la quema de combustible por parte de camiones y equipos mineros (Dönicke, 2017), y de manera indirecta por la combustión espontánea de pilas de carbón (Mendoza Daza, 2009). Este fenómeno se atribuye "*a la reacción química entre el carbón y el oxígeno lo que genera energía en forma de calor*" (Gutiérrez et al., 2018).

Rehabilitación. Hace referencia al conjunto de tareas propuestas para devolver terrenos degradados a condiciones habitables para organismos originalmente presentes. (Ministerio de minas y energía, 2003).

Para el desarrollo de actividades extractivas toda empresa minera debe desarrollar un Plan de rehabilitación ambiental, de tal forma que se mitiguen o compensen las afectaciones ambientales generadas. Investigadores a nivel mundial

han desarrollado estudios en esta temática, Wang et al, (2020) evalúan los impactos ambientales generados por la actividad minera y su posterior rehabilitación de servicios ecosistémicos en la mina de carbón Currangh en Australia. En el mismo sentido, Lechner et al. (2016) evalúan la rehabilitación progresiva de las minas en la cuenca Fitzroy en Australia utilizando técnicas de mapeo SIG, realizan un análisis multitemporal.

6.2 Ciclo del carbono y sumideros de carbono

6.2.1 Gas efecto de invernadero. *“Componente gaseoso de la atmósfera, natural o antropógeno, que absorbe y emite radiación en determinadas longitudes de onda del espectro de radiación terrestre emitida por la superficie de la Tierra, por la propia atmósfera y por las nubes. Esta propiedad ocasiona el efecto invernadero. El vapor de agua (H_2O), el dióxido de carbono (CO_2), el óxido nitroso (N_2O), el metano (CH_4) y el ozono (O_3) son los gases de efecto invernadero primarios de la atmósfera terrestre”* (IPCC, Glosario, 2013).

6.2.2 Ciclo del carbono. El carbono es un elemento que se presenta como constituyente fundamental de todos los organismos vivos. El carbono puede ser emanado a la atmosfera por procesos naturales como erupciones volcánicas (Sanci et al., 2010) y por procesos antrópicos en la combustión de combustibles fósiles (Henao et al., 2010).

El CO_2 en la atmosfera es absorbido por las plantas en el proceso de fotosíntesis; en donde el bióxido de carbono es transformado a glucosa y se genera

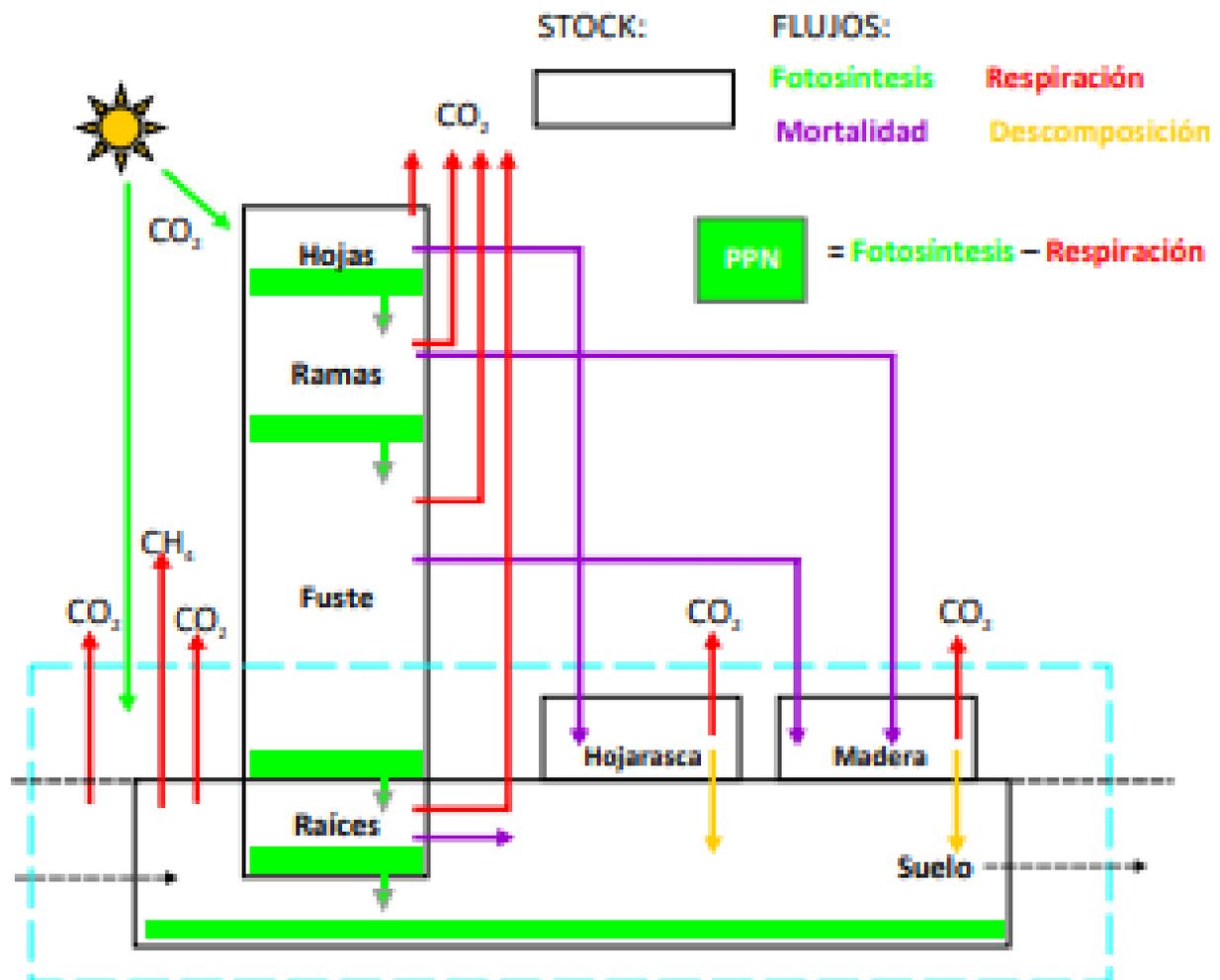
oxígeno, posteriormente se convierte en carbohidratos y tejidos que posteriormente van a convertirse biomasa; una porción de la glucosa es utilizada para el crecimiento de la planta (FAO, 2011).

En la figura 3 se observa un diagrama del almacenamiento (stock) y flujo de carbono en un bosque; los rectángulos de color negro representan los componentes donde se almacena y las flechas muestran el flujo del carbono. Cuando se desea estudiar el stock de carbono se debe estudiar la biomasa en superficie, bajo tierra y en la materia orgánica del suelo. Si se desea conocer el flujo de carbono se debe estudiar el proceso de fotosíntesis en hojas, la respiración de la planta, la mortalidad de la biomasa y la descomposición de la madera (Honorio C. y Baker, 2010).

6.2.3 Sumideros de carbono. *“El sumidero de CO₂ se define como el proceso que implica la remoción de gases de invernadero de la atmósfera”* (FAO, 2020). En el mismo sentido, el Panel Intergubernamental en cambio climático define los sumideros como un *“Reservorio (de origen natural o producto de la actividad humana, en suelos, océanos y plantas) en el que un gas de efecto invernadero, un aerosol o un precursor de un gas de efecto invernadero se almacenan”* (IPCC, Glosario, 2013).

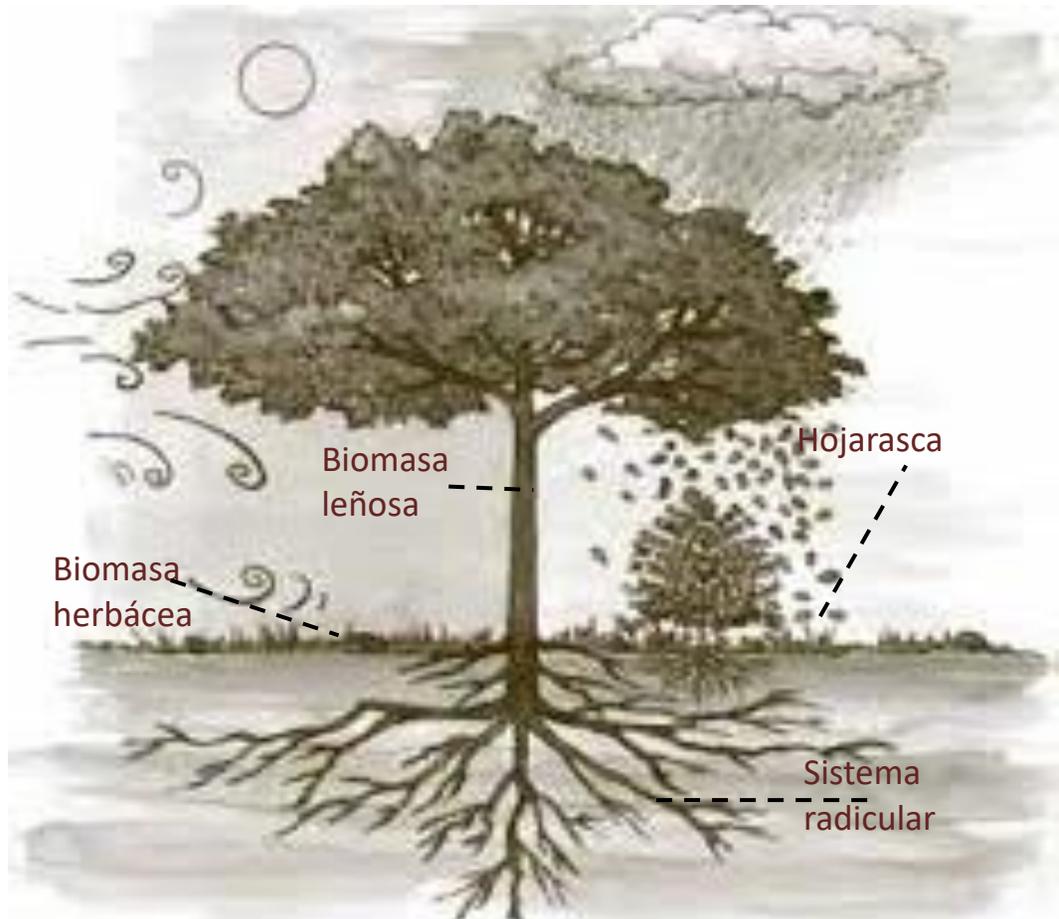
El carbono terrestre en la biomasa se puede acumular sobre el suelo (biomasa leñosa y herbácea), bajo el suelo (sistemas radiculares), hojarasca y el carbono orgánico del suelo (Snowdon et al., 2002).

Figura 3. Almacenamiento y flujo de carbono en bosques



Fuente. (Honorio C. y Baker, 2010)

Figura 4. Formas de acumulación del carbono terrestre



Fuente: Adaptado de (Beer et al., 2003)

Una vez los árboles son talados o cosechados se libera parte del carbono contenido; el resto aporta a la formación de la materia orgánica del suelo (Álvarez Sánchez y Harmon, 2003).

6.3 Cuantificación de carbono en biomasa herbácea

6.3.1 Biomasa. El termino hace referencia a toda “*material orgánico vivo o recientemente muerto*” (IPCC, 2013); es decir toda materia orgánica procedente de la agricultura, forestaría y restos animales que pueden transformarse en energía.

La biomasa asociada a la vegetación en términos generales es clasificada en: **Biomasa arbórea por encima suelo** la cual está relacionada a partes leñosas (fuste, corteza, ramas, tocones) y pastizales. **Y Biomasa arbórea por debajo del suelo** incluye raíces con diámetro mayor a 2 mm (FAO, 2006).

Para la reforestación y rehabilitación la empresa Drummond LTD emplea la especie herbácea *Andropogon gayanus* (Carimagua) por su capacidad de adaptación a las condiciones climáticas de la zona. Esta gramínea ha sido ampliamente utilizada en diversos estudios de rehabilitación de suelos degradados (Muller et al., 2001; Miranda et al., 2012)

6.3.2 Cuantificación de carbono en biomasa herbácea. Para determinar la cantidad de la biomasa y su carbono contenido se utilizan dos métodos directos o indirectos. En el primero es necesario cosechar muestras en subparcelas de muestreo en campo (Yaranga Cano y Custodio Villanueva, 2013) y los métodos indirectos se basan en estimaciones visuales, medición de altura del pasto, fotografías e imágenes de satélites (Mónaco et al., 2017).

En Colombia (Yepes et al., 2011) estipula los siguientes pasos para la estimación de la biomasa utilizando métodos directos:

Tipo, tamaño y número de parcelas. Las subparcelas para recolección de muestras son cuadradas, con tamaño de 1X1 m. (Fonseca González, 2017) utiliza como unidad de muestreo parcelas de 50X50 cm y (Connolly Wilson y Corea Siu, 2007) empleo subparcelas de 1X1 m.

Establecimiento de las parcelas. Los vértices de se deben fijar con pines d metal y delimitar el área con cuerda. Posteriormente recortar a ras de piso todo el material dentro de la subparcela. El material obtenido debe pesarse en campo, luego se debe tomar una muestra aproximada de 250 g (utilizar bolsas de plástico o papel) para procesarla en laboratorio. Una vez en el laboratorio se seca en el horno a 60 °C.

Contenido de carbono. se calcula a partir de la siguiente expresión:

$$C_{\text{muestra}} = BS * CF$$

Donde:

C_{muestra} (kg C m²) = Carbono en la muestra

BS = Peso de la biomasa seca

CF = fracción de carbono en biomasa 0.5 (IPCC, 2006),

Contenido de carbono por hectárea. Para su estimación es empleada la siguiente expresión:

$$C = \frac{10000 \text{ m}^2}{T_{\text{subparcela}}} * \frac{C_{\text{muestra}}}{1000}$$

C ($t\ C\ ha^{-1}$) = es la cantidad de carbono en la biomasa de vegetación no arbórea

$T_{subparcela}$ = Tamaño de subparcela en m^2

Otra forma de estimar el carbono fijo es realizar análisis próximos (ASTM D3172-89(02)) a cada una de las submuestras. La norma indicada establece estándares para determinar humedad, cenizas, materia volátil y carbono fijo (Vásquez Sierra y Herrera Builes, 2006).

6.4. Teledetección en el estudio de biomasa.

El objeto de estudio de la teledetección es la observación remota de la superficie terrestre, incluyendo la adquisición de imágenes y su posterior tratamiento (Chuvienco Salinero, 1996). Esta observación remota es debido a la interacción de un flujo electromagnético con las coberturas sobre la superficie (Pérez Gutierrez y Muñoz Nieto, 2006).

Las imágenes satelitales se caracterizan por la toma de datos en varias zonas del espectro electromagnético lo que lo hace ideal para el estudio de la biomasa (Song et al., 2015; Xiao et al., 2019).

Diferentes investigadores a nivel mundial han realizado investigaciones de estimación de biomasa utilizando datos de satélite Sentinel_2A obteniendo resultados precisos (Castillo et al., 2017; Punalekar et al, 2018).

6.4.1 Sentitel-2. El programa Sentinel forma de las misiones de la Agencia Espacial Europea (ESA). Fue lanzado al espacio al espacio el 23 de junio de 2015, a una altura

orbital de 786 km; este programa tiene por objetivo la observación de la tierra para el desarrollo de investigaciones asociadas al monitoreo terrestre como es el caso de la vegetación, suelo, zonas costeras (ESA, 2020).

Las características técnicas asociadas al proyecto Sentinel-2

Resolución temporal: 10 días

Resolución espacial: 10 – 60 metros

Resolución espectral: los detalles se observan en la Tabla 1.

Tabla 1. Características de las bandas Sentinel 2A.

	Longitud de onda central (nm)	Ancho de banda (nm)	Resolución espacial (m)
Banda 1 - Aerosol costero	442,7	21	60
Banda 2 – Azul	492,4	66	10
Banda 3 – Verde	559,8	36	10
Banda 4 – Rojo	664,6	31	10
Banda 5 - Borde rojo de vegetación	704,1	15	20
Banda 6 - Borde rojo de vegetación	740,5	15	20
Banda 7 - Borde rojo de vegetación	782,8	20	20
Banda 8 – NIR	832,8	106	10
Banda 8A - NIR estrecho	864,7	21	20
Banda 9 - vapor de agua	945,1	20	60
Banda 10 - SWIR - Cirrus	1373,5	31	60
Banda 11 – SWIR	1613,7	91	20

Fuente: (ESA, 2020)

6.4.2 Índice de vegetación (IV). Medida utilizada para extraer información de las coberturas vegetales mediante análisis de los valores de reflectancia a diferentes

longitudes de onda (Gilabert et al., 1997). Al determinar IV generalmente los valores inferiores revelan una vegetación poco vigorosa y los valores superiores muestran vegetación vigorosa (Muñoz, 2013).

No existe el Índice de Vegetación perfecto, los índices documentados son una aproximación al ideal (Pérez Gutiérrez et al., 2002). En la Tabla 2 se visualizan los Índices utilizados por diversos investigadores para el desarrollo de

6.5 Modelos de ajuste para la estimación de la biomasa herbácea

El análisis de datos del área de estudio se realizará a partir de la construcción de modelos de regresión, el cual tiene como *“objetivo predecir o aproximar el valor de una característica (variable dependiente) en función de los valores que pueden tomar en conjunto otra serie de características (variables independientes)”* (Peláez, 2016).

En el estudio evalúan el comportamiento de cuatro técnicas de modelación de datos para la estimación de carbono contenido en la biomasa herbácea en los terrenos rehabilitados por la empresa Drummond. Se compararon los siguientes modelos de regresión: Mínimos cuadrados ordinarios (OLS), Random Forest (RF), Maquinas vectoriales de soporte (SVM) y K-Vecinos más cercanos (K-NN).

6.5.1 Mínimos cuadrados ordinarios. Este método tiene por objetivo modelar una variable dependiente relacionada con una variable explicativa, generando una línea recta ajustada.

“La línea que mejor se ajusta a un conjunto de datos X-Y es aquella que minimiza la suma del cuadrado de las distancias desde los puntos hasta la línea, midiéndola en dirección vertical, o Y. Se conoce como recta de mínimos

cuadrados o regresión ajustada y su ecuación se llama Ecuación de regresión ajustada” (Hanke y Wichern, 2006, p. 213).

La línea ajustada es de forma:

$$\hat{Y} = \beta_0 + \beta_1 X$$

Donde:

Y es la variable dependiente (respuesta)

X es la variable independiente (explicativa)

β_0 constante del modelo (intersección)

β_1 pendiente de la recta

Este método ha sido utilizado con éxito para modelar biomasa utilizando datos de campo correlacionados con imágenes de satélite (Wnag et al.; 2020; Liu et al., 2021).

Tabla 2. Índice de vegetación

Índice	Ecuación	Características	Desventajas	Referencia
1. Índice de Vegetación de la Diferencia Normalizada (NDVI)	$\frac{NIR - Red}{NIR + Red}$	Utilizado para monitorear el verdor de la vegetación. Toma valores entre -1 a 1, siendo los valores cercanos a 1 indicador de vegetación saludable y abundante, a 0 presencia de agua y valores negativos suelo desnudo.	Sensible a la reflectividad del suelo; en zonas donde hay poca vegetación el suelo ejerce un efecto perturbador que varía dependiendo el tipo de suelo y porcentaje de cobertura vegetal.	(Rouse et al., 1974)
2. Índice de Vegetación Ajustado al Suelo (SAVI)	$\frac{NIR - Red}{(NIR + Red + L)} * (1 + L)$	Empleado para corregir el efecto de la reflectancia del suelo en zonas áridas con cobertura vegetal escasa, como parámetro de cálculo incluye la cantidad de cubierta de vegetación verde. Toma valores entre -1 a 1.	Condiciones de aplicación ideal para zonas con reducida cobertura vegetal. Una mala elección del parámetro L (cantidad de cobertura vegetal) altera significativamente los resultados.	(Huete, 1988)
3. Índice de vegetación ajustado por suelo modificado (MSAVI2)	$\frac{(2NIR + 1 - \sqrt{(2NIR + 1)^2 - 8(NIR - Red)})}{2}$	Este índice disminuye los efectos del suelo desnudo sobre la vegetación.		(Qi et al., 1994)
4. Índice de Vegetación Ajustado al Suelo Optimizado (OSAVI)	$\frac{NIR - Red}{NIR + Red + 0.16}$	Método empírico de la familia del índice savi, ayuda a disminuir los efectos de la	Recomendado para aplicaciones agrícolas	(Rondeau et al., 1996)

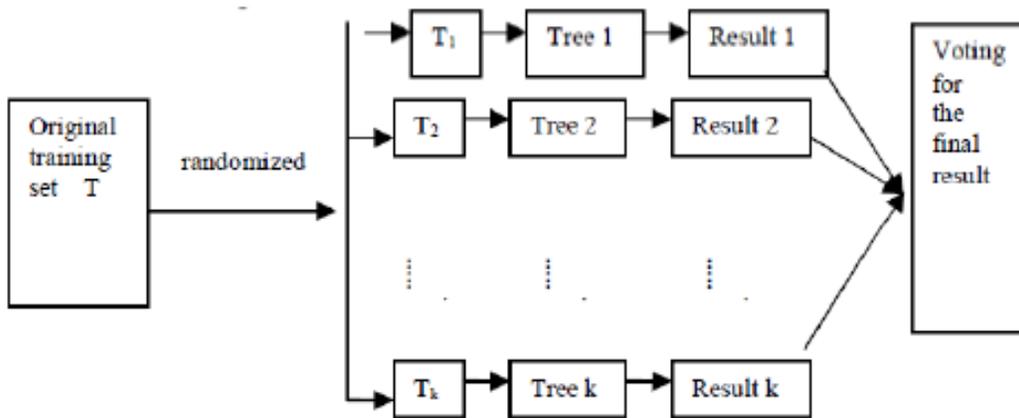
48 USO DE TELEDETECCIÓN EN LA ESTIMACIÓN DEL CARBONO DE LA BIOMASA HERBÁCEA EN TERRENOS REHABILITADOS POR LA EMPRESA DRUMMOND LTD EN EL DEPARTAMENTO DEL CESAR-COLOMBIA

		reflectancia del suelo gracias a la inclusión de un parámetro x con un valor de 0.16 en el numerador de la expresión matemática.	
5. Índice de vegetación mejorado (EVI)	$\frac{2.5(NIR - Red)}{NIR + 6Red - 7.5Blue + 1}$	Especialmente sensible en zonas de abundante vegetación, optimiza la señal del índice captado. Se caracteriza por tener menor sensibilidad al ruido generado por la atmosfera.	Pierde precisión en zonas con escasa vegetación (Huete et al., 1997)
6. Índice de Vegetación Atmosféricamente Resistente (ARVI)	$\frac{NIR - Red}{NIR + (Red - \gamma(Buel - red))}$	Índice mejorado ampliamente utilizado para atenuar la influencia de la atmosfera. Empleando en zonas con una alta presencia de aerosol atmosférico.	Para una mayor precisión del método debe conocerse previamente el modelo de aerosol γ . De lo contrario el valor a utilizar es 1. (Tanré et al., 1992)
7. Cociente simple (VARI)	$\frac{Green - red}{Green + Red - Blue}$	Este índice se caracteriza por la detección de la vegetación solo con las bandas del rango visible	Poco sensible para suelo desnudo, agua y hielo.

Fuente. Datos de estudio

6.5.2 Random forest (RF). Es un método de aprendizaje automático que permite crear modelos y generar predicciones. El algoritmo crea muchos árboles, cada uno de los cuales trabaja con un subconjunto de datos ligeramente distintos seleccionados aleatoriamente y obtiene una estimación (cada árbol es un modelo que genera su predicción). Luego, las respuestas de todos los árboles son analizadas tomando el dato más frecuente para dar la respuesta final (Rodrigo, 2020).

Figura 5. Esquema Random forest



Fuente. Tomado de (Liu et al., 2012)

La predicción del modelo RF está representada por la siguiente expresión:

$$\mu(x) = \sum_{i=1}^n w_i(x) Y_i$$

Donde:

X son las variables predictivas

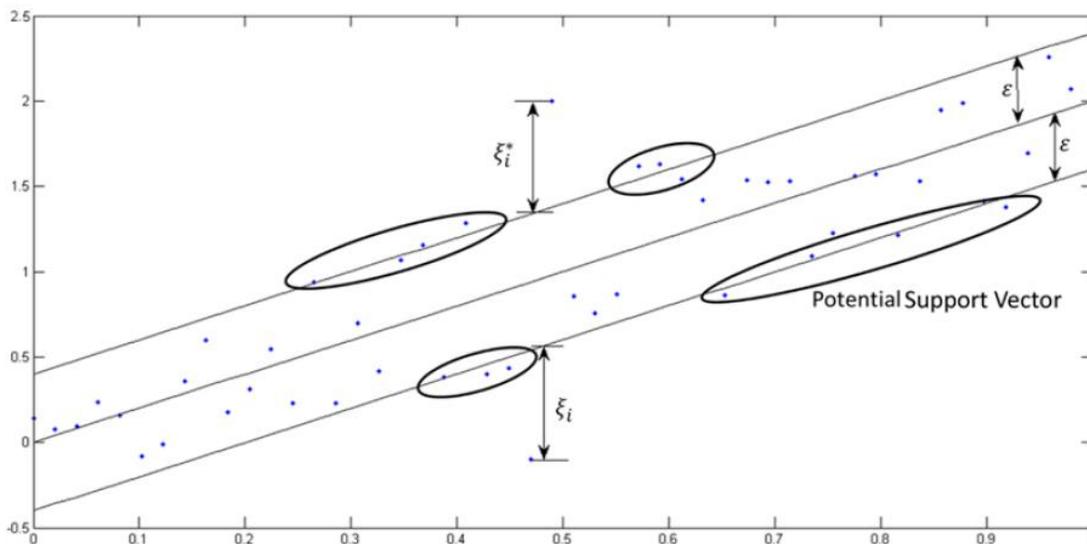
Y variable de respuesta

n número de arboles

Diversos investigadores a nivel mundial han experimentado resultados favorables al generar modelos predictores en el mapeo de biomasa utilizando el algoritmo Random forest (Wang et al., 2016; Fernandez-Habas et al., 2022).

6.5.3 Regresión maquina vectorial de soporte (SVM). Es un algoritmo de aprendizaje automático que utiliza el algoritmo de clasificación maquina vectorial de soporte, empleado para predecir una variable dependiente. Este algoritmo modela datos de entrenamiento en una región llamada hiperplano, el cual se encuentra limitado por una región tubular llamada ϵ – tube (Hawkins, 2015). Los valores por fuera de la función de predicción son penalizados son denominados vector de soporte (Vapnik , 1998).

Figura 6. Esquema modelo de regresión vectorial de soporte.

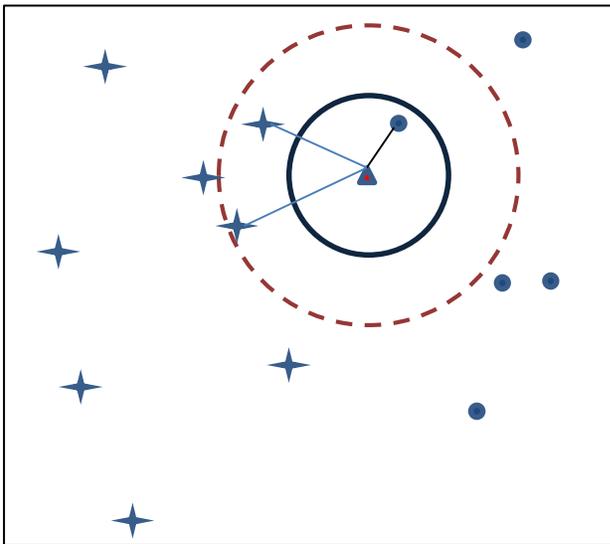


Fuente. Tomado de (Hawkins, 2015).

La calidad del modelo generado depende del parámetro ϵ el cual controla el ancho del tubo y el parámetro C controla el número de estimaciones (Durbha et al., 2007).

6.5.4 K-vecinos más cercanos (K-NN). Es un método supervisado no paramétrico empleado para predecir valores tanto en problemas de clasificación como de regresión; el algoritmo utiliza el promedio de los datos de entrenamiento más cercanos para predecir un valor (Rezaei y Jabbari, 2022).

Figura 7. Esquema K - vecinos más cercanos.



Nota. La figura muestra la aplicación del K-NN con un valor de $k=1$ (circunferencia con línea continua) y un $K=3$ (circunferencia línea discontinua).

Para su aplicación requiere el número de vecinos más cercanos (K) que se utilizará en el análisis. La definición de los vecinos más cercanos es realizada a partir del cálculo de la distancia euclidiana; entre más cercano se encuentre el vecino tiene mayor peso en el valor a calcular (Pérez Trujillo, 2021).

$$d = \sqrt{\sum_{i=1}^n (p_i - q_i)^2}$$

Donde:

p registro vacío

q datos de entrenamiento

n número de variables de cada registro

Esta técnica fue aplicada con resultados satisfactorios para la estimación de biomasa seca de arroz mediante la correlación de datos de campo con imágenes Sentinel-1A (Mansaray, Zhang, & Kanu, 2020).

7. HIPÓTESIS DE INVESTIGACIÓN

Las herramientas de los sistemas de información geográfica relacionadas con variables de crecimiento de pastos obtenidos de trabajo de campo permiten estimar el contenido de biomasa herbácea y carbono en áreas rehabilitadas por la empresa Drummond LTD.

8. METODOLOGÍA

8.1 Enfoque metodológico

Para alcanzar el objetivo general planteado se define un enfoque metodológico **Cuantitativo**, debido a que los datos utilizados son de carácter numérico, requieren un análisis estadístico y los resultados obtenidos son medibles.

8.2 Tipo de Estudio

Para determinar el contenido de carbono en la biomasa herbácea de las áreas rehabilitadas por la multinacional Drummond se requieren datos de campo, los cuales se relacionan con datos de índices de vegetación obtenidos de imágenes satelitales, lo que hace que este estudio sea tipo **Correlacional**.

8.3 Diseño de la Investigación.

El diseño de la investigación a desarrollar es de tipo transeccionales correlacionales-causales, debido a que se buscan datos en un momento determinado para su posterior correlación.

La investigación se desarrolla siguiendo tres fases. La primera fase corresponde a las labores de campo, la segunda fase comprende la adquisición y procesamiento de la imagen satelital y en la tercera fase se realiza la correlación los datos de campo y sensores remotos.

Fase I. Se realiza una visita de campo preliminar necesario para la identificación de la biomasa y delimitación del área a trabajar; para definir los límites del área a trabajar se tuvo

en cuenta la información sugerida por la empresa y una revisión multitemporal de imágenes que muestren las áreas reforestadas.

Durante la visita de campo se tomaron 10 muestras de la biomasa herbácea, las cuales fueron adquiridas en subparcelas de 1 m², posteriormente las muestras son llevadas al laboratorio para estimar el contenido de carbono. Las muestras tomadas tienen un peso de 250 g (Yepes et al., 2011). Adicionalmente, se tomarán 11 puntos de referencia en zonas donde no hay presencia de biomasa; en total 21 muestras.

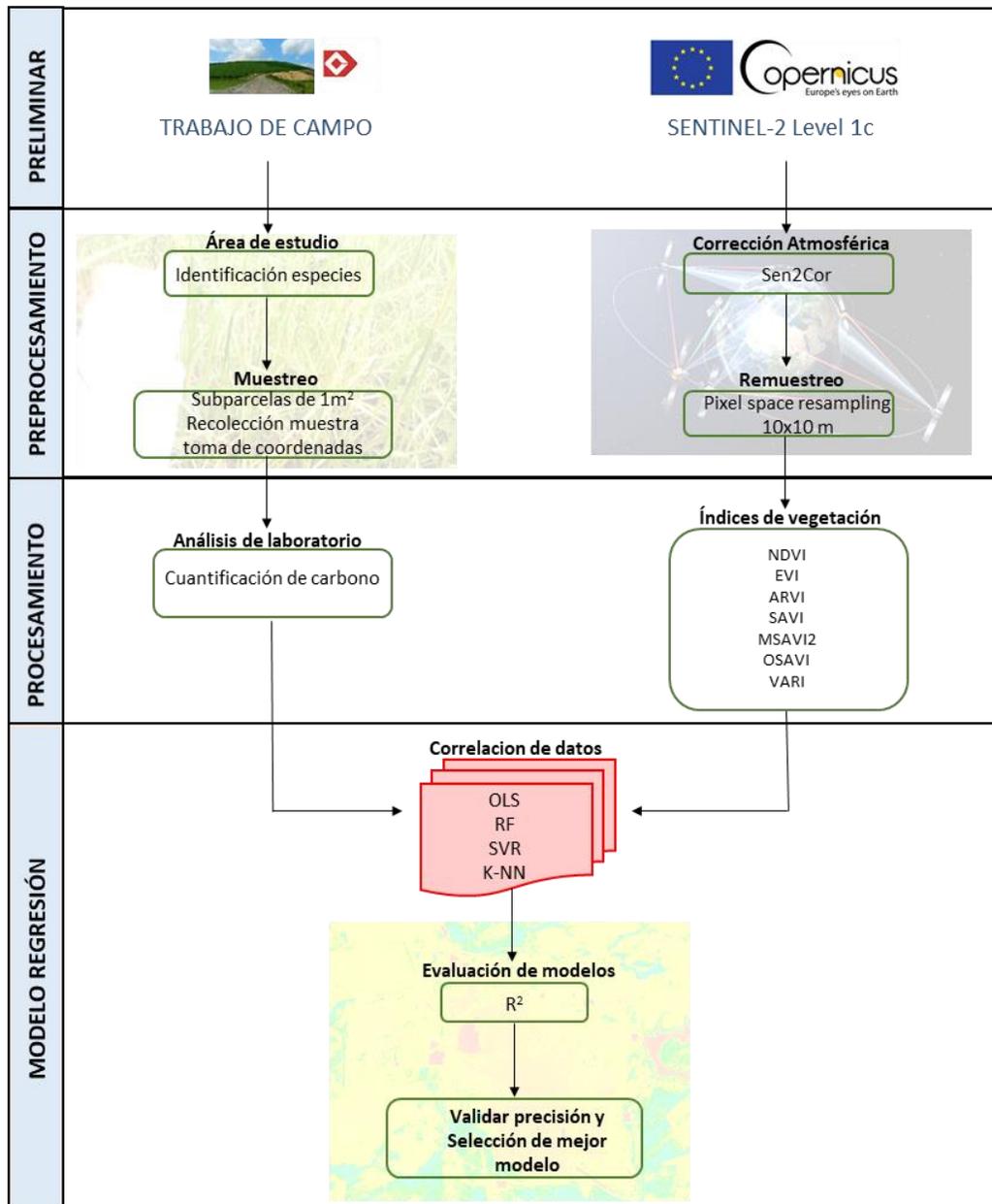
Fase II. Inicialmente se descargada una escena satelital Sentinel-2A con nivel de procesamiento TOA-1C desde la plataforma Copernicus (Jutz y Milagro-Pérez, 2018). Posteriormente se realiza la corrección atmosférica, corrección de terrenos y cirros de la imagen seleccionada utilizando el procesador SEN2COR (Gascón F. et al. , 2017).

Luego, se calculan múltiples índices de vegetación (NDVI, ARVI, EVI, SAVI, MSAVI2, OSAVI, VARI), que servirán como parámetro de entrada para el desarrollo de modelos predictivos.

Fase III. Se realiza la correlación de los datos tomados en campo con los índices de vegetación extraídos aplicando diversos métodos: Regresión por mínimos cuadrados ordinario (OLS), Random forest (RF), Maquinas de vectores de soporte (SVM) y K-Vecinos más cercanos (K-NN). La correlación de los datos se realiza en la plataforma Anaconda mediante el uso de la librería sklearn del lenguaje de programación Python. En la **¡Error! No se encuentra el origen de la referencia.** se observa un esquema con la metodología de trabajo.

Los resultados de cada uno de los análisis se evaluarán utilizando el coeficiente de correlación R² (Escribano Rodríguez et al., 2014).

Figura 8. Esquema de metodología de trabajo



Fuente. Datos de estudio

8.6 Operacionalización de la variable.

Las variables que componen el problema de investigación se descomponen en la siguiente tabla:

Tabla 3. Operacionalización de variable.

1. Variable	Áreas rehabilitadas por la empresa Drumond en la mina Pribbenow
2. Tipo de variable	Cuantitativa
3. Operacionalización	Plantas herbácea sembradas para la rehabilitación de áreas minadas
4. Categorías	Biomasa herbácea
5. Definición	La biomasa herbácea hace referencia al contenido de material orgánico presente en plantas herbáceas
6. Indicador	Toneladas de carbono en la biomasa herbácea por hectárea
7. Nivel de medición	De razón
8. Unidad de medida	t/ha C
9. Índice	Contenido de carbono por hectárea
10. Valor	100%

Fuente. Adaptado de (Betancur López, 1998)

8.7 Técnicas e instrumentos de recolección de información.

La recolección de información se realizará mediante trabajo de campo teniendo en cuenta los siguientes aspectos:

- Subparcelas. Tendrán un tamaño de 1m² y se distribuirán de manera aleatoria en el área de estudio
- Dentro de las subparcelas de muestreo se cortará todo el pasto dentro. El corte debe realizarse a nivel del suelo.
- La muestra se guarda en bolsas plásticas o de papel, se rotula y se envía a laboratorio.
- Determinación de contenido de biomasa y contenido de carbono.

9. RESULTADOS Y DISCUSIÓN

9.1 Datos preliminares y trabajo de campo.

Se descargó una secuencia de imágenes satelitales de los sensores Landsat 5, 7 y Sentinel_2A del área de estudio con intervalos de 5 años entre escenas. En la secuencia de imágenes de la Figura 9 se muestra:

1. Imagen Landsat-5 del año 1991, muestra la línea base antes de iniciar el proyecto minero.
2. Imagen Landsat-5 del año 1996, muestra la fase inicial de la operación minera, apreciándose al norte el área de acopio del material estéril (botaderos) y en la parte inferior el pit de explotación. La mina pribbenow inicio actividades extractivas el año de 1995 al norte del título minero (Drummond LTD, 2021).
3. Imagen Landsat-7 del año 2001, se observa la expansión del pit minero hacía el centro del título minero, la expansión de áreas de botadero hacia el este del título, construcción de vías de acarreo en dirección a los patios de acopio de carbón y se observa el inicio de la etapa de reforestación en los botaderos ubicados al norte.
4. Imagen Landsat-7 2006, se identifica la expansión de: los botaderos al norte y al este del título y del pit minero. Adicionalmente, se evidencia un aumento de las zonas reforestadas en los botaderos al norte y este del título.
5. Imagen Landsat-7 2011, se evidencia la expansión del pit minero, la construcción de un nuevo botadero al oeste y la expansión de botadero al este hasta el límite sur del Área de trabajo. De igual forma se aprecia un aumento en las zonas reforestadas.
6. Imagen Sentinel-2A 2016, al ser una imagen de mayor resolución se aprecia mayor detalle de la operación. Se evidencia el inicio de reforestación del botadero oeste.
7. Imagen Sentinel-2A 2021, esta imagen muestra el estado actual de la mina, observándose con mayor detalle aquellas áreas que han sido objeto de reforestación.

Figura 9. Análisis multitemporal

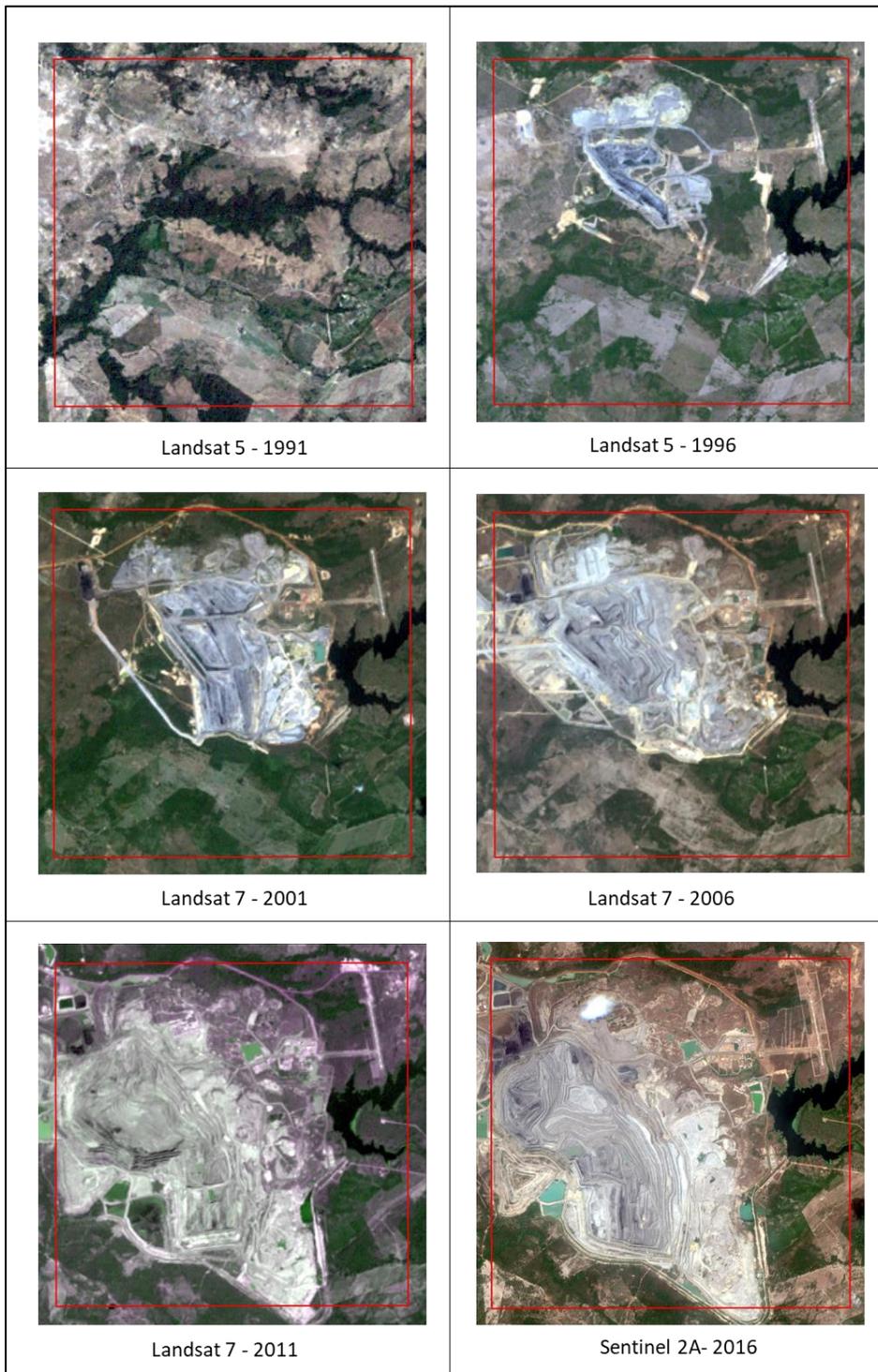
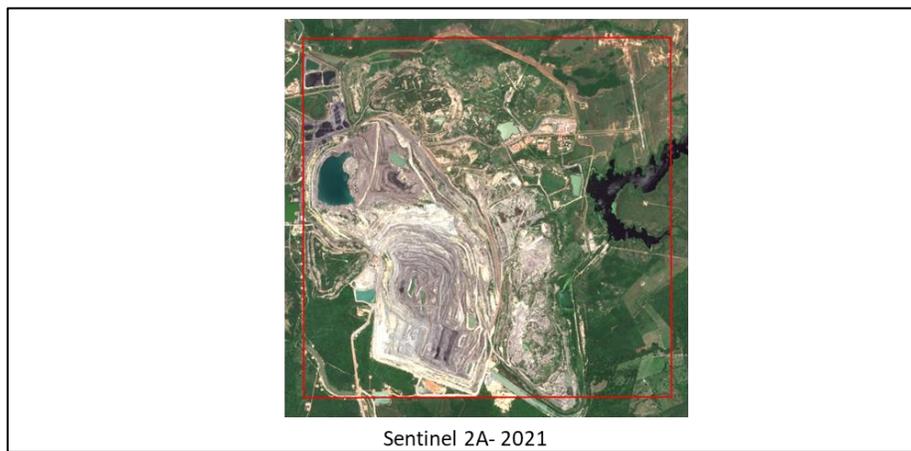


Figura 10. Análisis multitemporal (Continuación)

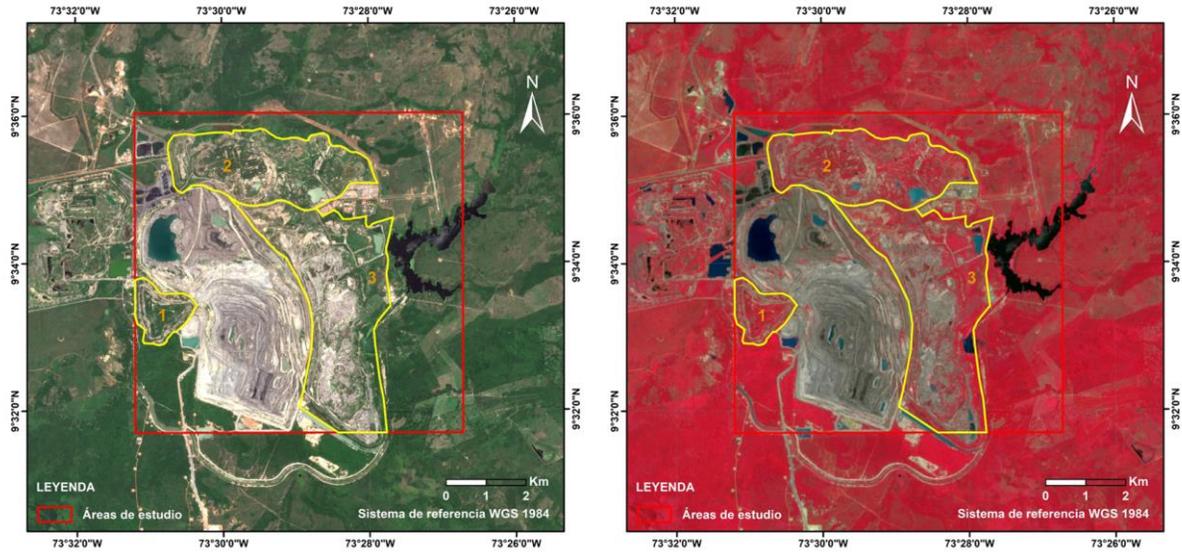


Fuente. Autor

A partir de la información recolectada en campo y los datos proporcionados por las imágenes de satélites se identificaron las áreas intervenidas para su rehabilitación. En la Figura 11. Áreas en proceso de rehabilitación – Combinación 843 se observa la delimitación de tres zonas: la 1 y 2 corresponden a zonas antiguas con pastos bien establecidos, mientras la zona 3 aún se encuentra en labores preparatorias.

El cálculo del carbono solo se realizará en las zonas 1 y 2, debido a que se encuentran en una etapa avanzada. La especie herbácea utilizadas en la etapa de rehabilitación es el *Andropogon gayanus* (Pasto carimagua); especie originaria de África occidental, se caracteriza por su resistencia a suelos de baja fertilidad, ácidos, bien drenados y aguante a sequias extensas (Laredo C. y Gomez s., 1982). Las características de esta especie lo hacen ideal para los procesos de rehabilitación de terrenos minados.

Figura 11. Áreas en proceso de rehabilitación – Combinación 843



Nota. La imagen de la derecha muestra la combinación 843 resaltando la vegetación.

Figura 12. Pasto sembrado



Fuente. Departamento ambiental Drummond

9.1.1. Muestreo. Se establecieron 21 puntos de muestreo de los cuales 10 corresponden a parcelas de 1m², donde se procedió a la toma de muestra de la biomasa herbácea y 11 puntos adicionales donde no existe presencia de vegetación (suelo desnudo y agua). Posteriormente, se georreferenciaron las parcelas utilizando los Módulos RTK GNSS M2 de la empresa emlid reach; que proporciona una precisión centimétrica en los procesos de medición (Fernández Álvarez et al., 2021). Los datos

de las coordenadas de las parcelas y la cantidad de biomasa herbácea cuantificada en el aforo de las parcelas se observan en la Tabla 1Tabla 4.

Figura 13. Punto de muestreo



Fuente. Autor

Las muestras de biomasa tomadas en campo fueron llevadas al laboratorio de química de la Facultad de ingeniería y ciencias básicas en la Fundación Universitaria del Área Andina ubicada en la ciudad de Valledupar, donde se determinó el contenido de carbono en la biomasa a partir de la materia orgánica contenida en las muestras procesadas. Los datos obtenidos se observan en la tabla 4.

Tabla 4. Georreferenciación y contenido de biomasa por parcela.

Muestra	Peso de biomasa por parcela (Kg)	Coordenadas	
		Latitud	Longitud
1	21.080	-73.5186438	9.55774462
2	5.780	-73.5115870	9.55634800
3	14.535	-73.5100234	9.58556140
4	14.280	-73.5095312	9.59340906

5	8.075	-73.5002874	9.58742947
6	12.070	-73.4954828	9.59281657
7	15.725	-73.4925590	9.58236598
8	22.355	-73.4837645	9.58727719
9	12.410	-73.4772657	9.59476773
10	8.755	-73.4708012	9.58790536
10	0	-73.5160359	9.56008041
11	0	-73.5157583	9.55906009
12	0	-73.4919463	9.58397406
13	0	-73.478262	9.58247888
14	0	-73.4826646	9.58436418
15	0	-73.4879326	9.58679147
16	0	-73.4893747	9.58773076
17	0	-73.4905418	9.58543923
18	0	-73.4756981	9.59258107
19	0	-73.4777094	9.58884237
20	0	-73.5089215	9.58728433

Fuente. Autor

Figura 14. Procesamiento de muestras en laboratorio



Fuente. Autor

Del análisis realizado se puede concluir que el contenido de carbono promedio en las parcelas de muestreo donde se obtuvo biomasa es de 49.83%, los resultados obtenidos se observan en la Tabla 5.

64 USO DE TELEDETECCIÓN EN LA ESTIMACIÓN DEL CARBONO DE LA BIOMASA HERBÁCEA EN TERRENOS REHABILITADOS POR LA EMPRESA DRUMMOND LTD EN EL DEPARTAMENTO DEL CESAR-COLOMBIA

Tabla 5. Contenido de carbono en muestras

Muestra	Peso muestra	Peso crisol	Peso muestra+cresol	% Materia seca	Peso Muestra seca	Peso material calcinado	% Carbono
1	2	20.08	22.08	19.92	1.602	0.792	50.12%
2	2	19.81	21.81	20.15	1.597	0.788	48.84%
3	2	22.16	24.16	20.14	1.597	0.803	49.82%
4	2	18.73	20.73	20.05	1.599	0.801	49.57%
5	2	21.94	23.94	20.21	1.596	0.795	48.98%
6	2	20.08	22.08	19.98	1.600	0.802	49.72%
7	2	19.81	21.81	20.11	1.598	0.792	50.09%
8	2	22.16	24.16	20.02	1.600	0.800	50.25%
9	2	18.73	20.73	20.03	1.599	0.798	49.89%
10	2	21.94	23.94	20.05	1.599	0.795	49.57%

Fuente. Autor

9.2 Índices de vegetación

De la imagen Sentinel se extraen los valores de los índices de vegetación los puntos de muestreo. Estos valores son la base para el desarrollo de los modelos de estimación de biomasa herbácea.

Tabla 6. índice de vegetación extraído para cada punto de muestreo.

ID	arvi	evi	msavi2	ndvi	osavi	Savi	vari
0	0.61262	0.000590	1.36675	0.639758	0.764832	1.14714	0.354798
1	0.33012	0.000403	1.19063	0.47224	0.527481	0.791134	0.111861
2	0.54033	0.000523	1.33193	0.627223	0.712221	1.06821	0.225313
3	0.52859	0.000472	1.32809	0.606342	0.706618	1.059820	0.242117
4	0.35524	0.000321	1.22476	0.55205	0.568348	0.85243	0.032967
5	0.52608	0.000483	1.32601	0.648332	0.7036	1.05529	0.147541
6	0.55994	0.000523	1.3426	0.640046	0.728013	1.09191	0.237838
7	0.70347	0.000632	1.4059	0.718613	0.827973	1.24185	0.393891
8	0.53594	0.000561	1.32779	0.595506	0.706184	1.05917	0.284483
9	0.40367	0.000437	1.24898	0.602108	0.598713	0.897989	-0.008282

10	-0.0067	0.000106	0.62005	0.087703	0.063867	0.095793	-0.041927
11	-0.0394	0.000059	0.68503	0.196392	0.101961	0.152931	-0.146154
12	-0.2185	-0.002888	-0.76527	-0.44763	-0.387831	-0.581664	0.118659
13	-0.2991	-0.002125	-2.15028	-0.62494	-0.570993	-0.856333	0.132867
14	-0.2574	-0.000675	-0.96263	-0.49260	-0.423011	-0.634382	0.13597
15	-0.0665	0.000043	0.66251	0.209155	0.088452	0.13267	-0.178867
16	-0.0414	0.000061	0.71955	0.242401	0.123322	0.184972	-0.178961
17	-0.1303	-0.000006	0.47525	0.104849	-0.012226	-0.018338	-0.170543
18	-0.0728	0.000036	0.64199	0.1994	0.076425	0.114632	-0.182734
19	-0.0091	0.000087	0.71465	0.187255	0.120248	0.180356	-0.108629
20	-0.0727	0.000039	0.65263	0.216216	0.082633	0.123938	-0.197647

Nota: Los ID del 0 al 9 corresponden a los valores espectrales en las parcelas de muestreo con presencia de biomasa mientras que los valores entre 10 y 21 corresponden a los puntos donde no hay presencia de biomasa

Fuente. Autor

Los valores espectrales asociados a los puntos de muestreo varían dependiendo el índice calculado. Para las zonas donde hay pastos sembrados el valor varía dependiendo la densidad de la biomasa, con valores de NDVI entre 0.3 y 0.5 para áreas con vegetación escasa y mayor a 0.5 para zonas con vegetación densa (Toosi et al., 2022). Lo que es consistente con los datos obtenidos de la imagen de satélite mostrando que la mayor parte de los puntos de muestreo tiene una cobertura vegetal densa.

Los puntos de muestreo 10, 11, 15, 16, 17, 18, 19 y 20 se ubican sobre suelo desnudo, de los cuales los punto 10, 11, 18 y 19 se encuentran sobre las vías o es zonas aledañas a ellas, mientras que 15, 16, 17 y 20 se encuentran sobre los taludes. Adicionalmente, mientras puntos 12, 13, 14 se encuentran sobre el agua.

Los valores espectrales de los puntos que se encuentran sobre las vías muestran los valores más bajos; suelos estériles sin presencia de materia orgánica. Mientras que los puntos

ubicados sobre los taludes presentan valores más altos; suelos con mayor presencia de materia orgánica. En 1992, Kaufman y Tanré describen que los valores espectrales asociados al suelo desnudo oscilan entre -0.08 a 0.08. Mientras, Toosi et al. 2022 describen en su investigación que los valores de NDVI menores a 0.1 corresponden a áreas áridas

Los valores asociados a los cuerpos de agua adoptan valores negativos (Santiago et al., 2009), de los datos extraídos de la imagen Sentinel se observan valores negativos variables dependiendo el índice de vegetación evaluado, solo se exceptúa el VARI que toma valores positivos.

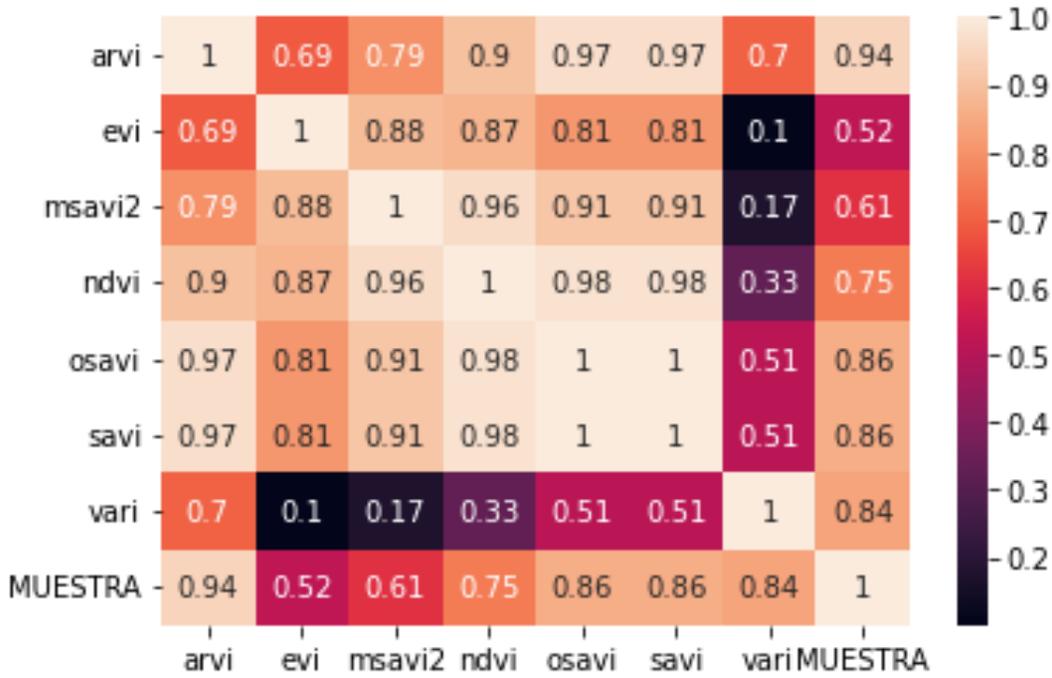
9.3 Estadística descriptiva de los datos de trabajo.

Inicialmente se analiza la relación entre los índices de vegetación extraídos de la imagen de satélite con los valores de contenido de carbono a partir del coeficiente de correlación de Pearson.

El mapa de calor de la matriz de correlación en La Figura 15 **Error! No se encuentra el origen de la referencia.** muestra una alta correlación positiva entre los índices de vegetación ARVI, OSAVI y el SAVI con el contenido de carbono de las muestras con un valor de 0.94, 0.86 Y 0.86 respectivamente, mientras el EVI es el índice con menor correlación con el contenido de carbono 0.52. Adicionalmente existe una alta correlación entre algunos predictores; entre el SAVI - OSAVI muestra una relación del 1 y entre el NDVI - SAVI la relación es de 0.98.

Estos valores calculados muestran un panorama general de los modelos a generar.

Figura 15. Matriz de correlación



Fuente. Autor

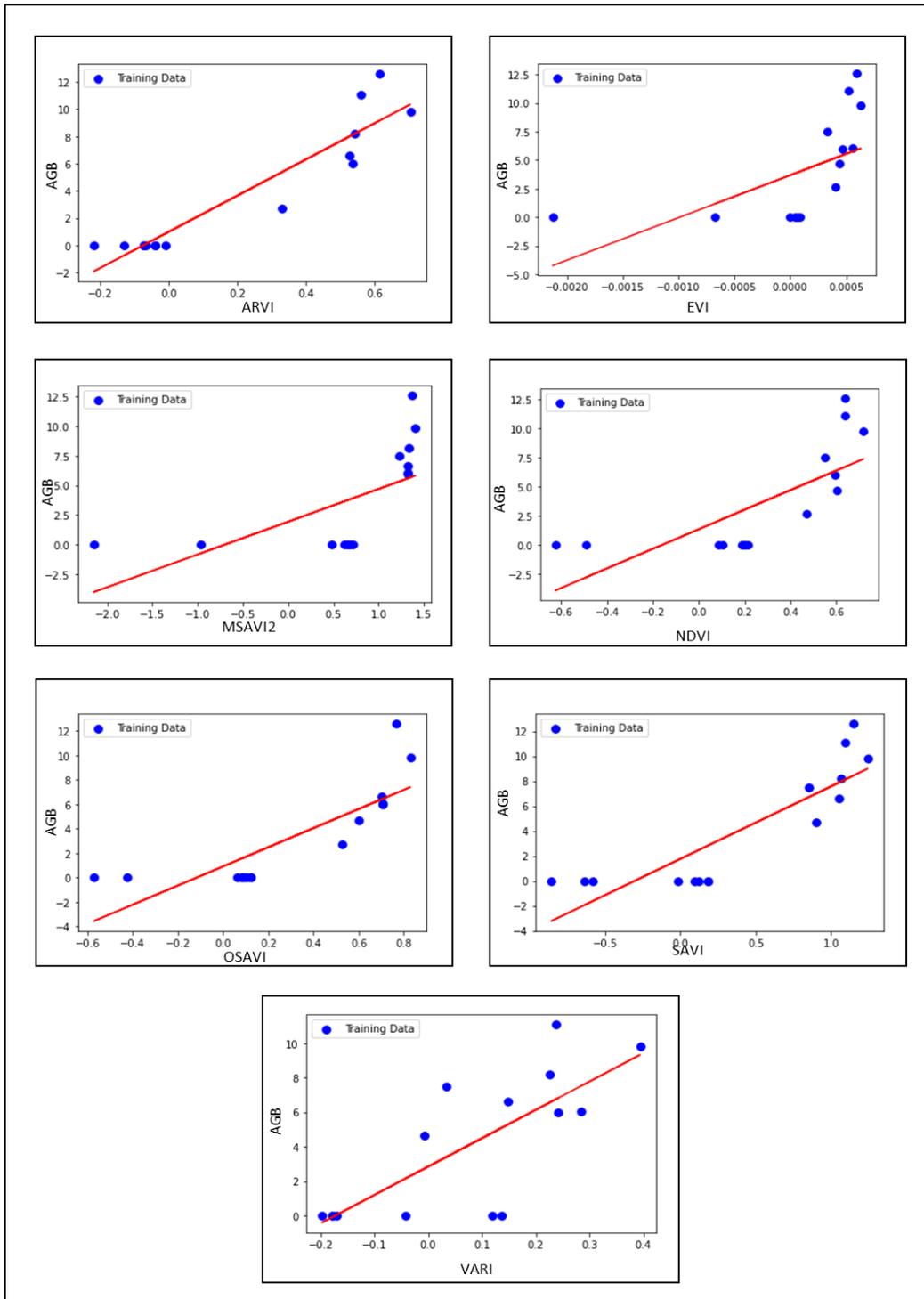
9.3. Modelos para cálculo de carbono en vegetación herbácea

Los modelos utilizaron como variable dependiente (respuesta) el contenido de carbono en la biomasa herbácea y como variable independiente (predictora) los índices de vegetación ARVI, EVI, MSAVI2, NDVI, OSAVI, SAVI y VARI.

9.3.1 Regresión por mínimos cuadrados ordinarios (OLS). Las ilustraciones de la muestra la relación entre los índices de vegetación extraído de la imagen de satélite y el contenido de carbono en la biomasa herbácea. Se puede apreciar que la mejor relación entre puntos y la línea de tendencia se presenta con los índices ARVI, OSAVI y SAVI.

68 USO DE TELEDETECCIÓN EN LA ESTIMACIÓN DEL CARBONO DE LA BIOMASA HERBÁCEA EN TERRENOS REHABILITADOS POR LA EMPRESA DRUMMOND LTD EN EL DEPARTAMENTO DEL CESAR-COLOMBIA

Figura 16. Diagramas Índice de Vegetación vs Contenido de Carbono



Fuente. Autor

Al evaluar los modelos realizados con la regresión OLS, los resultados más exactos se obtuvieron para los índices ARVI ($R^2 = 0.88$, RMSE=1.41, MAE=1.19) y OSAVI ($R^2 = 0.72$, RMSE=2.43, MAE=2.15). Mientras que los índices que presentaron menor exactitud en los cálculos fueron el EVI ($R^2 = 0.31$, RMSE=3.41, MAE=3.50) y el MSAVI2 ($R^2 = 0.35$, RMSE=3.57, MAE=3.07). Los resultados de la evaluación son concordatos con los resultados arrojados por la matriz de correlación; que muestra los índices ARVI, SAVI y OSAVI como los de mayor relevancia para la estimación del carbono y el EVI como el de menor importancia. Es notable resaltar que, aunque el MSAVI2 pertenezca a la misma familia del SAVI y el OSAVI no muestra un buen rendimiento para la estimación.

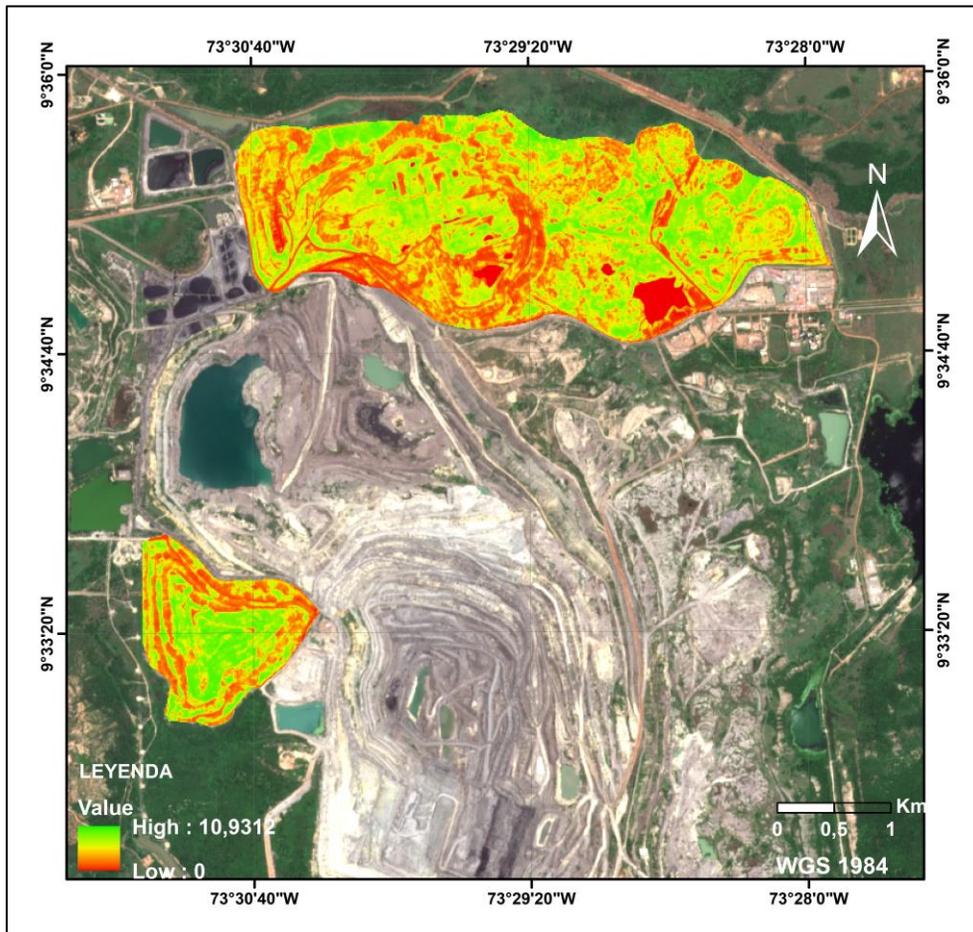
Tabla 7. Evaluación de modelos OLS

	ARVI	EVI	MSAVI2	NDVI	OSAVI	SAVI	VARI
MAE	1.19.	3.50	3.07	1.86	2.15	2.08	1.41
RMSE	1.41	3.49	3.57	2.12	2.43	2.35	2.00
R^2	0.88	0.31	0.35	0.55	0.72	0.70	0.68

Fuente. Autor

En la Figura 17 se observa la distribución del contenido de carbono en el área estudiada, en rojo las áreas con ausencia de biomasa (suelo desnudo, depósitos de agua) y en color verde con presencia de biomasa. Se observa que en las áreas de máxima concentración encontramos un valor de 10.93 Kg C/m², la concentración de carbono será mayor en aquellas áreas de mayor densidad de biomasa; es decir en las zonas de reforestación más antiguas.

Figura 17. Modelación de carbono OLS

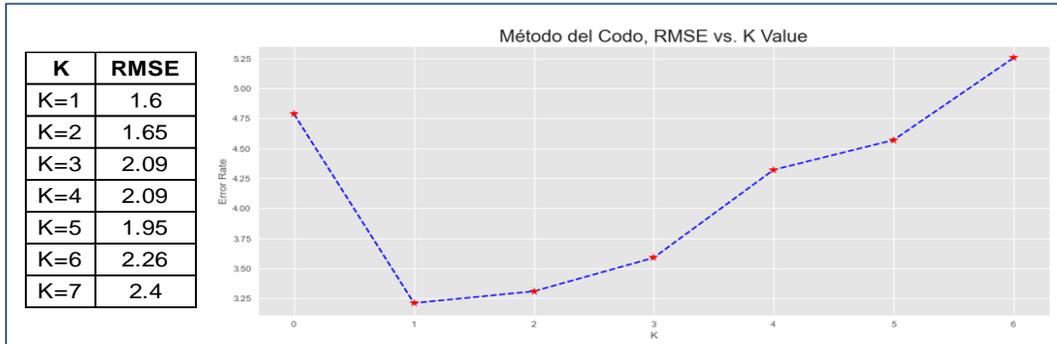


Fuente. Autor

9.3.2 K-vecinos más cercanos (K-NN). Inicialmente se evaluó un modelo general aplicando la regresión KNN, el modelo desarrollado empleando todas índices de vegetación como variable independiente, con un coeficiente $K = 3$ presento un $R^2=0.84$, $RMSE=1.51$ un $MAE=1.14$

A partir de los resultados iniciales, se aplicó el método del codo para determinar el Coeficiente K más idóneo para la estimación del carbono en la biomasa, siendo el valor de $K=1$ y $K=2$ con valores de error cuadrático medio de 1.60 y 1.65 respectivamente.

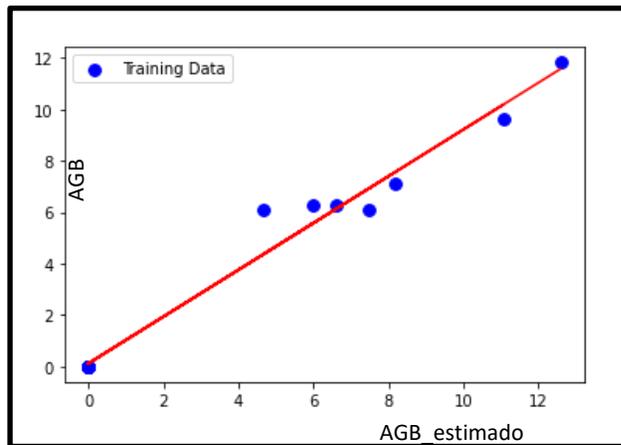
Figura 18. RMSE vs K



Fuente. Autor

Finalmente, el modelo predictivo para estimar el carbono en la biomasa se generó en función de los datos arrojados por la matriz de correlación, utilizando el índice ARVI.

Figura 19. Grafica de residuos KNN



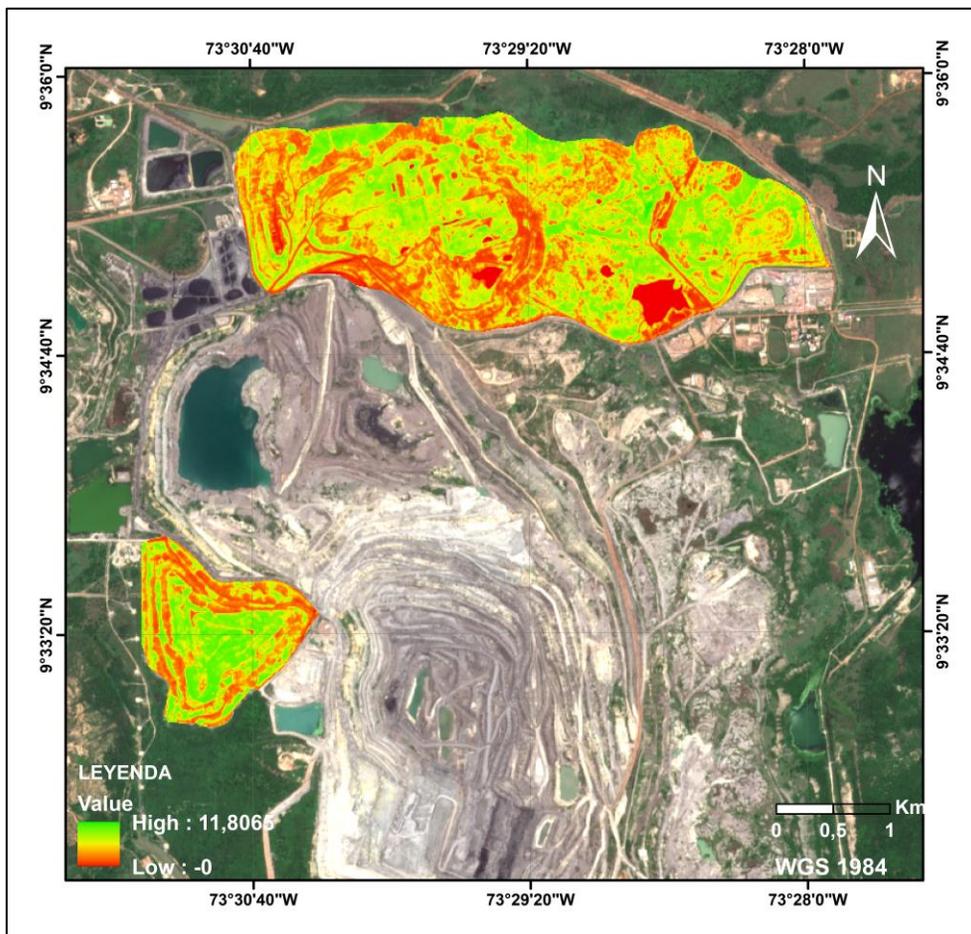
Fuente. Autor

La grafica de residuos muestra una dispersión adecuada de puntos sobre la recta, lo que nos permite inferir que los modelos generados pueden estimar de manera adecuada el contenido de carbono.

72 USO DE TELEDETECCIÓN EN LA ESTIMACIÓN DEL CARBONO DE LA BIOMASA HERBÁCEA EN TERRENOS REHABILITADOS POR LA EMPRESA DRUMMOND LTD EN EL DEPARTAMENTO DEL CESAR-COLOMBIA

La distribución de carbono es similar al modelo realizado empleando la regresión OLS, la diferencia radica en la cantidad de carbono estimada por metro cuadrado que pasa de 10.93 kg C/m² a 11.80 kg C/m² en las zonas de mayor concentración.

Figura 20. Modelación de carbono KNN



Fuente. Autor

9.3.3 Maquina vectorial de soporte (SVM). La aplicación de este método de regresión se aplicó de manera separada a cada uno de los índices de vegetación. El algoritmo inicialmente se corrió con los parámetros por defecto, mostrando que el mejor índice para estimar el contenido de carbono era el ARVI mostrando valores de $R^2=0.77$, $RMSE=2.16$ y $MAE=1.27$. La evaluación para los modelos realizados se observa en la Tabla 8.

Tabla 8. Evaluación de modelo SVM con parámetros de cálculo por defecto

	ARVI	EVI	MSAVI2	NDVI	OSAVI	SAVI	VARI
MAE	1.27	2.34	1.48	1.40	1.36	1.36	2.33
RMSE	2.16	3.03	2.42	2.35	2.44	2.34	3.11
R ²	0.77	0.55	0.71	0.73	0.73	0.73	0.52

Fuente Autor

A partir de los resultados iniciales se optimizo el modelo, mostrando los mejores C, kernel y gamma para estimar el contenido de carbono. Ver Tabla 9.

Tabla 9. Parámetros de cálculo para regresión SVM

	ARVI	EVI	MSAVI2	NDVI	OSAVI	SAVI	VARI
C	10	100	100	100	10	100	100
gamma	1	1	1	1	1	0.1	0.1
Kernel	Rbf	rbf	rbf	rbf	Rbf	rbf	rbf

Fuente. Autor.

Al evaluar los modelos utilizando los parámetros optimizados C, gamma y kernel se observa que los índices MSAVI2 Y OSAVI presentaron un mejor ajuste estadístico. El MSAVI2 presento valores de $R^2=0.90$, $RMSE=1.43$ y $MAE=0.88$; el índice OSAVI exhibió valores de $R^2=0.90$, $RMSE=1.46$ y $MAE=0.94$. Para modelar el contenido de carbono en la biomasa se emplea el índice con valores de más bajos, es decir el MSAVI2.

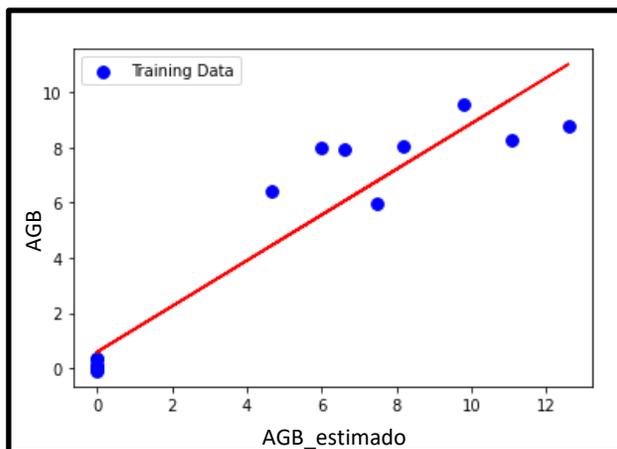
Tabla 10. Evaluación de modelo SVM con parámetros optimizados

	ARVI	EVI	MSAVI2	NDVI	OSAVI	SAVI	VARI
MAE	0.96	4.14	0.88	1.09	0.94	0.94	1.95
RMSE	1.60	4.86	1.43	1.65	1.46	1.49	2.65
R^2	0.87	0.16	0.90	0.87	0.90	0.89	0.66

Fuente. Autor

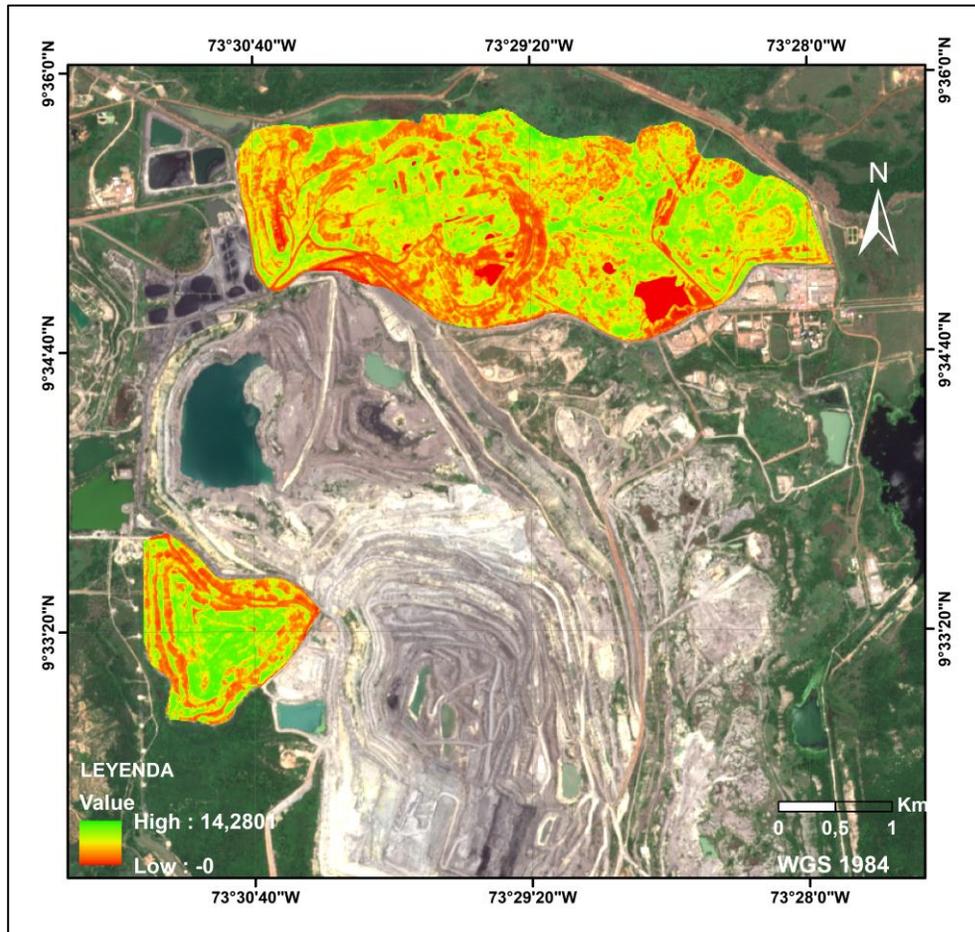
Al diagramar los datos de entrenamiento del modelo con el índice MSAVI2 se observa poca dispersión de los residuos, indicando una buena modelación de los datos.

Figura 21. Grafica de residuos SVM



En la Figura 22 se observa la distribución del contenido de carbono estimado, con un valor máximo de 14.28 kg C/m². Se aprecia la estimación de carbono en áreas donde hay presencia de suelo desnudo, lo que lleva a una sobreestimación del contenido de carbono.

Figura 22. Modelación de carbono SVM



Fuente. Autor

9.3.4. Random forest (RF). La estimación de carbono utilizando el modelo random forest se realiza empleando un modelo general; donde se utilizan todos los índices de vegetación. Se corrió el modelo con los parámetros `max_depth`, `max_features`, `n_estimato` y

random_state por defecto optimizados, mostrando valores de $R^2=0.93$, $RMSE=0.71$ y $MAE=0.59$.

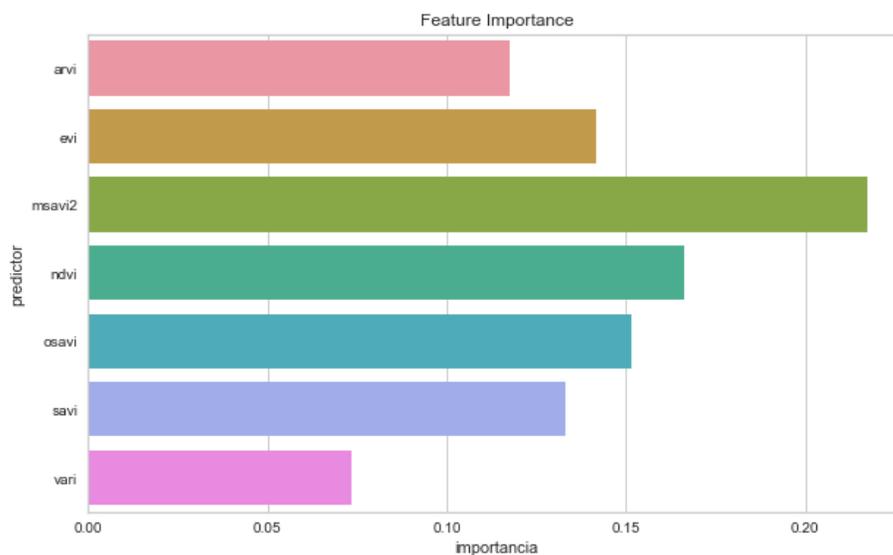
Tabla 11. Parámetros de cálculo RF

Max Depth	Max features	n estimators	Random state
5	log2	10	0

Fuente. Autor

Se estimó el factor de importancia para determinar cuál de los índices de vegetación tiene más incidencia en el modelo, donde se pudo determinar que el MSAVI2 tiene un mayor peso en la estimación de carbono.

Figura 23. Factor de importancia RF

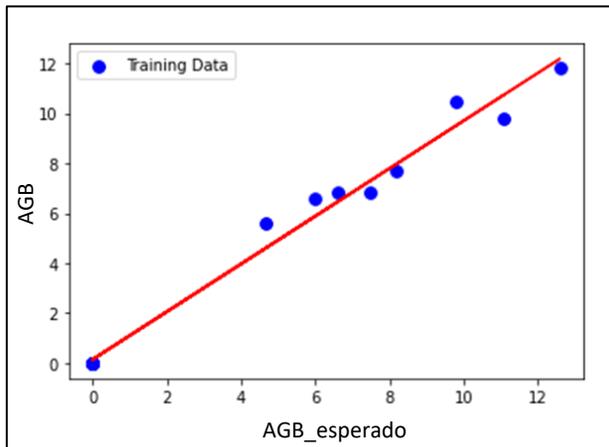


Fuente. Autor.

Una vez determinado el factor de importancia se aplica la regresión Random forest para determinar el contenido de carbono empleando como variable independiente el índice MSAVI2, al evaluar el modelo se observan valores de $R^2=0.935$, $RMSE=0.78$ y $MAE=0.585$.

La Figura 24 muestra la distribución de los residuos, donde se aprecian poca dispersión de los datos.

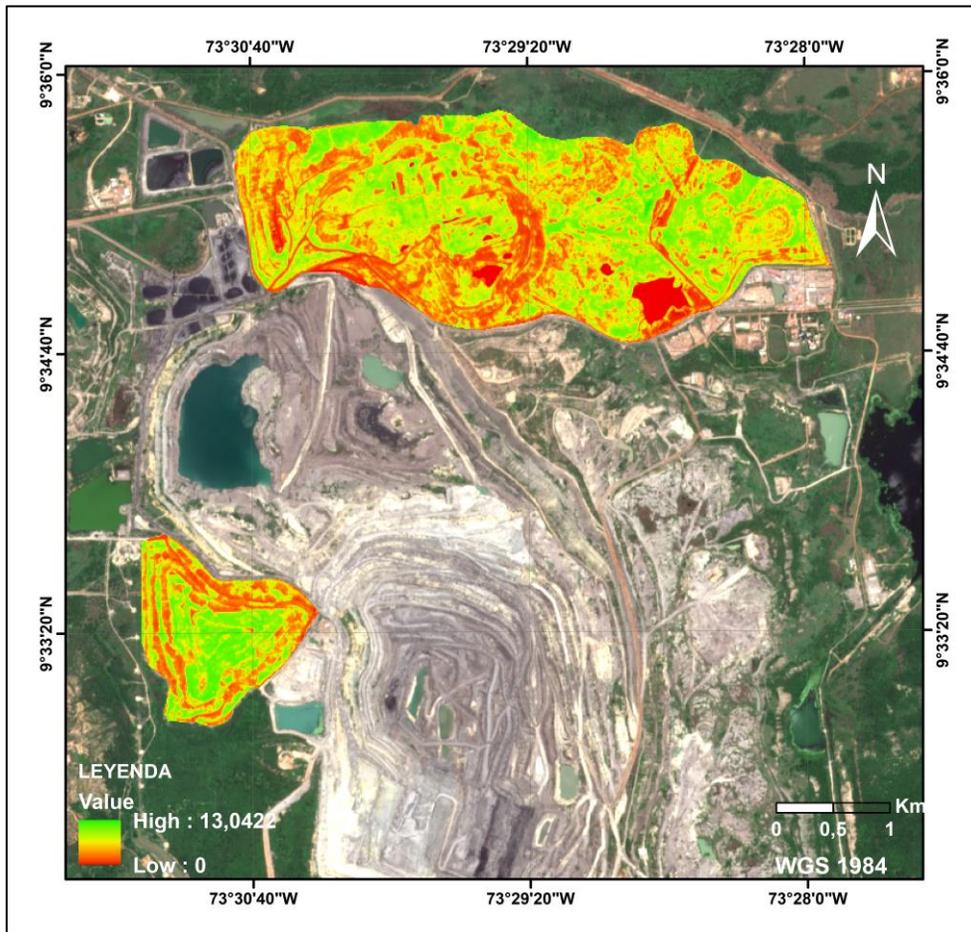
Figura 24. Grafica de residuos RF



Fuente. Autor

En la **¡Error! No se encuentra el origen de la referencia.** se observa la distribución del contenido de carbono estimado, con un valor máximo de 13.04 kg C/m². Se diferencian las áreas de suelo desnudo y agua de aquellas que contienen biomasa.

Figura 25. Modelación de carbono RF



Fuente. Autor

9.4 Discusión general

La presente investigación tubo por finalidad encontrar una relación entre los valores espectrales extraídos de una imagen satelital Sentinel y la biomasa herbácea de los terrenos rehabilitados por la empresa, con la finalidad de modelar el carbono contenido.

Las tecnologías de la percepción remota, posibilitan la observación de grandes extensiones de la superficie terrestre y permiten identificar los cambios que sobre esta se

presentan. Mediante el análisis multitemporal desarrollado en la investigación, se identificó como ha cambiado la cobertura vegetal en el área del proyecto minero, desde una etapa previa al desarrollo de las actividades extractivas en 1991 hasta el año 2021, donde se aprecian las áreas que han sido rehabilitadas. Sin embargo, aunque la empresa destine recursos y personal para la recuperación de los terrenos minados aun la cobertura vegetal se observa fraccionada, son escasas las áreas donde la vegetación cubre en su totalidad la superficie del terreno.

En cuanto a los métodos de regresión utilizados, se demostró que mediante el uso de modelos matemáticos como el Maquinas vectoriales de soporte, Random forest, K vecinos más cercanos y Mínimos cuadrados ordinarios se puede estimar de manera eficiente el contenido de carbono en la biomasa herbácea.

El modelo con mayor exactitud se obtuvo utilizando el método de regresión Random forest (RF) en combinación con el índice de vegetación MSAVI2 ($R^2=0.935$, $RMSE=0.78$ y $MAE=0.585$), con un valor máximo de trece kilogramos de carbono capturado por cada metro cuadrado de terrenos sembrados con biomasa herbácea (13.04 kg C/m^2).

Los resultados obtenidos muestran que la biomasa herbácea tiene un alto potencial para la captura de carbono (13.04 kg C/m^2) y eventualmente cuando la empresa aumente las superficies rehabilitadas, se aumentara el secuestro de carbono. La generación de conocimiento en torno a esta temática permitirá a la empresa proyectar un crecimiento económico sostenible y bajo en carbono.

Para tener una mayor comprensión de la cantidad de carbono almacenado en los terrenos recuperados por parte de Drummond, se deben elaborar estudios más detallados que incluyan modelos de captura de carbono en las arboles, raíces y el suelo. De esta manera la

empresa puede aportar al compromiso adquirido por Colombia en el Convenio Marco de las Naciones Unidas para el Cambio Climático donde se comprometió a elaborar, actualizar y publicar datos de emisiones y absorción de gases de efecto invernadero.

En Colombia no se han definido metas que obligue a las empresas del sector minero a medir y reducir las emisiones de gases de efecto invernadero producidos durante la extracción de minerales. Por tal motivo, los datos obtenidos en la presente investigación servirán de referencia para estudios tendientes a la mitigación de emisiones de carbono por parte del sector minero.

10. CONCLUSIONES

La minería a cielo abierto se caracteriza por dejar huellas visibles sobre la superficie terrestre, las cuales pueden ser mapeadas utilizando imágenes de satélite. Las características espectrales de las imágenes facilitan la diferenciación del suelo desnudo, cuerpos de agua y áreas que han sido objeto de reforestación. Mediante un análisis multitemporal, es factible identificar la evolución de las áreas objeto de explotación versus las zonas objeto de reforestación.

A partir de una correcta delimitación de la superficie que contiene la biomasa herbácea mediante el uso de siete índices de vegetación, el contenido de carbono en la biomasa y el uso de modelos de regresión OLS, SVM, K-NN y RF, fue posible modelar el contenido de carbono en la biomasa herbácea en los terrenos rehabilitados por la empresa Drummond. Se evidencia, que la adecuada configuración de los parámetros en los modelos de regresión mejora significativamente la estimación del carbono contenido en la biomasa; en el modelo Maquinas vectoriales de soporte los parámetros costo, kernel y gamma, en K-vecinos más cercanos el valor de K y en el modelo Random forest los parámetros max_depth, max_features, n_estimato y random_state.

En la investigación se demuestra que el modelo de regresión Random forest, puede estimar el contenido de carbono contenido en la biomasa herbácea con fidelidad a partir del índice de vegetación MSAVI2 extraído de imágenes de satélite Sentinel-2A y el contenido de carbono asociado a muestras de biomasa obtenidas en campo.

El uso combinado de las herramientas de la teledetección y los Sistemas de Información Geográfica facilitan la generación de cartografía para el seguimiento y control de áreas que cumplen la función de captura de carbono. Lo que permitirá a la empresa aportar con las medidas de reducción de emisiones de gases de efecto invernadero.

11. RECOMENDACIONES

De acuerdo con lo plasmado en la presente investigación, algunas medidas a tener en cuenta para ahondar en la temática asociada con el objeto del mismo son:

- Realizar modelos para la estimación del carbono contenido en la biomasa leñosa, raíces y suelo de los terrenos rehabilitados por la empresa Drummond y compararlos con los resultados obtenidos en terrenos aledaños donde no se ha desarrollado minería.
- Desarrollar un análisis económico para evaluar el potencial de captura de carbono de los terrenos rehabilitados para un eventual ingreso a los mercados de carbono.
- Hacer estudio de huella de carbono para determinar la totalidad de carbono emitido empresa y comparar dichos datos con la cantidad de carbono que está siendo almacenado en terrenos rehabilitadas e inalterados.
- Realizar investigación para modelar como ha sido la dinámica en el cambio de uso del suelo y evaluar cuando carbono se ha dejado de captar con la remoción de la cobertura vegetal desde el inicio de las actividades extractivas hasta el día año 2021.

11. BIBLIOGRAFÍA

- Ahirwal, J., & Maiti, S. K. (2017). Assessment of carbon sequestration potential of revegetated coal mine overburden dumps: A chronosequence study from dry tropical climate. *Journal of Environmental Management*, 201, 369–377. <https://doi.org/10.1016/j.jenvman.2017.07.003>
- Ali, I., Cawkwell, F., Dwyer, E., Barrett, B., & Green, S. (2016). Satellite remote sensing of grasslands: from observation to management. *Journal of Plant Ecology*, 9(6), 649–671. <https://doi.org/10.1093/JPE/RTW005>
- Álvarez Sánchez, J., & Harmon, M. E. (2003). Descomposición de hojarasca: Hojas y madera. In J. Álvarez Sánchez & E. Naranjo García (Eds.), *Ecología del suelo en la selva tropical húmeda de México* (pp. 108–122). Instituto de Ecología, A.C. https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0065-17372004000200018
- Avtar, R., Komolafe, A. A., Kouser, A., Singh, D., Yunus, A. P., Dou, J., Kumar, P., das Gupta, R., Johnson, B. A., Vuong, H., Minh, T., Kumar Aggarwal, A., & Kurniawan, A. (2020). Assessing sustainable development prospects through remote sensing: A review. *Remote Sensing Applications: Society and Environment*, 20, 100402. <https://doi.org/10.1016/j.rsase.2020.100402>
- Awad, M., & Khanna, R. (2015). Support Vector Regression. In: *Efficient Learning Machines*. Apress, Berkeley, 67–80. https://doi.org/10.1007/978-1-4302-5990-9_4
- Banerjee, B. P., & Raval, S. (2017). REMOTE SENSING TO ADVANCE SUSTAINABLE MINING PRACTICES. *Conference: Asian Conference on Remote Sensing*. https://a-a-r-s.org/proceeding/ACRS2017/ID_753_1652/1564.pdf

- Barliza, J. C., Peláez, J. D. L., & Campo, J. (2018). Recovery of biogeochemical processes in restored tropical dry forest on a coal mine spoil in La Guajira, Colombia. *Land Degradation & Development*, 29(9), 3174–3183. <https://onlinelibrary.wiley.com/doi/epdf/10.1002/ldr.3069>
- Barrachina, M., Cristóbal, J., & Tulla, A. F. (2015). Estimating above-ground biomass on mountain meadows and pastures through remote sensing. *International Journal of Applied Earth Observation and Geoinformation*, 38, 184–192. <https://doi.org/10.1016/J.JAG.2014.12.002>
- Beer, J., Ibrahim, M., Somarriba, E., Barrance, A., & Leakey, A. (2003). Establecimiento y manejo de árboles en sistemas agroforestales. In J. Cordero & D. H. Boshier (Eds.), *Árboles de centroamerica: un manual para extensionistas* (pp. 197–242). <https://repositorio.catie.ac.cr/handle/11554/9730>
- Brivio, P. A., Maggi, M., Binaghi, E., & Gallo, I. (2010). Mapping burned surfaces in Sub-Saharan Africa based on multi-temporal neural classification. *International Journal of Remote Sensing*, 24(20), 4003–4016. <https://doi.org/10.1080/0143116031000103835>
- Cáceres, J., Martín, M. P., & Salas, J. (2015). Análisis temporal de biomasa y stocks de carbono en un ecosistema de dehesa mediante imágenes Landsat, y su relación con factores climáticos. *Ciencias Espaciales*, 8(1), 190–211. <https://doi.org/10.5377/CE.V8I1.2049>
- Casiano-Domínguez, M., Paz-Pellat, F., Rojo-Martínez, M., Covalada-Ocon, S., & Aryal, D. R. (2018). Anatomía de la madera de tres especies de Mimosa (Leguminosae-Mimosoideae) distribuidas en México. *Madera y Bosques*, 24. <https://doi.org/10.21829/myb.2018.2401894>
- Castillo, J. A. A., Apan, A. A., Maraseni, T. N., & Salmo, S. G. (2017). Estimation and mapping of above-ground biomass of mangrove forests and their replacement land uses in the Philippines

using Sentinel imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 134, 70–85.
<https://doi.org/10.1016/J.ISPRSJPRS.2017.10.016>

Chu, X., Zhan, J., Li, Z., Zhang, F., & Qi, W. (2019). Assessment on forest carbon sequestration in the Three-North Shelterbelt Program region, China. *Journal of Cleaner Production*, 215, 382–389. <https://doi.org/10.1016/J.JCLEPRO.2018.12.296>

Chuai, X., Xia, M., Xiang, A., Miao, L., Zhao, R., & Zuo, T. (2022). Vegetation coverage and carbon sequestration changes in China's forest projects area. *Global Ecology and Conservation*, 38. <https://doi.org/10.1016/j.gecco.2022.e02257>

Chuvieco, E. (1996). Fundamentos de teledetección. *Ediciones RIALP S.A.*

LEY 164 DE 1994, (1994) (testimony of Congreso de Colombia).
<https://www.minambiente.gov.co/wp-content/uploads/2022/01/1.-Ley-160-de-1994.pdf>

Connolly Wilson, R. Y., & Corea Siu, C. A. (2007). *Cuantificación de la captura y almacenamiento de carbono en sistema agroforestal y forestal en seis sitios de cuatro municipios de Nicaragua*.
<https://repositorio.una.edu.ni/1103/>

Estrategia institucional para la articulación de políticas y acciones en materia de cambio climático en Colombia, (2011) (testimony of Consejo Nacional de Política Económica y Social & Departamento Nacional de Planeación).
<https://colaboracion.dnp.gov.co/cdt/conpes/econ%C3%B3micos/3700.pdf>

Das, R., & Maiti, S. K. (2016). Estimation of carbon sequestration in reclaimed coalmine degraded land dominated by *Albizia lebbeck*, *Dalbergia sissoo* and *Bambusa arundinacea* plantation: a case study from Jharia Coalfields, India. *International Journal of Coal Science and Technology*, 3(2), 246–266. <https://doi.org/10.1007/S40789-016-0131-4/FIGURES/12>

- Dávila, E. Á., & Pérez, A. (2012). Bosques y minería responsable en Antioquia. *Revista Ambiental ÉOLO*, 12(17), 1–14. <http://revistaeolo.fconvida.org/index.php/eolo/article/view/818>
- Dönicke, D. B. (2017). Cálculo del consumo de combustible y emisiones De co2 de camiones mineros, mediante simulación Discreta. *Revista Ingeniería Industrial*, 16(2), 151–168. <https://revistas.ubiobio.cl/index.php/RI/article/view/3316>
- Drummond Ltd. (2021). *Informes de Sostenibilidad 2021*. https://issuu.com/drummondLtd/docs/is_drummond_2021_20220727_830_pm_sencillas?fr=sMDdjNDM5NDk1ODc
- Dube, T., Mutanga, O., Shoko, C., Samuel, A., & Bangira, T. (2016). Remote sensing of aboveground forest biomass: A review ESA-Alcantra View project Remote sensing of water use and water stress in African savanna ecosystem from local to regional scale: Implications for land productivity View project. *International Society for Tropical Ecology*, 57(2), 125–132. www.tropecol.com
- Durbha, S. S., King, R. L., & Younan, N. H. (2007). Support vector machines regression for retrieval of leaf area index from multiangle imaging spectroradiometer. *Remote Sensing of Environment*, 107(1–2), 348–361. <https://doi.org/10.1016/J.RSE.2006.09.031>
- ESA. (2020). *Sentinel-2 MSI User Guide*. <https://sentinel.esa.int/web/Sentinel/user-guides/Sentinel-2-msi>
- Escribano Rodríguez, J. A., Díaz-Ambrona, C. G. H., & Tarquis Alfonso, A. M. (2014). Selección de índices de vegetación para la estimación de la producción herbácea en dehesas. *Pastos*, 44(2), 6–18. <http://polired.upm.es/index.php/pastos/article/view/3108>

-
- FAO. (2006). *Evaluación de los recursos forestales mundiales 2005*.
<https://www.fao.org/forestry/32145/es/>
- FAO. (2011). *La gestión de los bosques ante el cambio climático*.
<https://www.fao.org/3/i1960s/i1960s00.pdf>
- FAO. (2020). *Portal de Suelos de la FAO*. <https://www.fao.org/soils-portal/soil-management/secuestro-de-carbono-en-el-suelo/es/>
- Fernández Álvarez, H., Álvarez-Narciandi, G., García-Fernández, M., Laviada, J., Álvarez López, Y., & Las-Heras Andrés, F. (2021). A Portable Electromagnetic System Based on mm-Wave Radars and GNSS-RTK Solutions for 3D Scanning of Large Material Piles. *Sensors*, 21(3), 757. <https://doi.org/10.3390/S21030757>
- Fernández-Habas, J., Carriere Cañada, M., García Moreno, A. M., Leal-Murillo, J. R., González-Dugo, M. P., Abellanas Oar, B., Gómez-Giráldez, P. J., & Fernández-Rebollo, P. (2022). Estimating pasture quality of Mediterranean grasslands using hyperspectral narrow bands from field spectroscopy by Random Forest and PLS regressions. *Computers and Electronics in Agriculture*, 192, 106614. <https://doi.org/10.1016/J.COMPAG.2021.106614>
- Fonseca-González, W. (2017). Revisión de métodos para el monitoreo de biomasa y carbono vegetal en ecosistemas forestales tropicales. *Revista de Ciencias Ambientales*, 51(2), 91–109. <https://doi.org/10.15359/RCA.51-2.5>
- Gao, J. (2007). Quantification of grassland properties: how it can benefit from geoinformatic technologies? *International Journal of Remote Sensing*, 27(7), 1351–1365. <https://doi.org/10.1080/01431160500474357>

Gascon, F., Bouzinac, C., Thépaut, O., Jung, M., Francesconi, B., Louis, J., Lonjou, V., Lafrance, B., Massera, S., Gaudel-Vacaresse, A., Languille, F., Alhammoud, B., Viallefont, F., Pflug, B., Bieniarz, J., Clerc, S., Pessiot, L., Trémas, T., Cadau, E., ... Fernandez, V. (2017). Copernicus Sentinel-2A Calibration and Products Validation Status. *Remote Sensing*, 9(6), 584. <https://doi.org/10.3390/RS9060584>

Gilabert, M. A., González-Piqueras, J., & García-Haro, J. (1997). Acerca de los Indices de Vegetación. *Revista de Teledetección*, 1–10.

Gitelson, A. A., Kaufman, Y. J., & Merzlyak, M. N. (1996). Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sensing of Environment*, 58(3), 289–298. [https://doi.org/10.1016/S0034-4257\(96\)00072-7](https://doi.org/10.1016/S0034-4257(96)00072-7)

Gitelson, A. A., Stark, R., Grits, U., Rundquist, D., Kaufman, Y., & Derry, D. (2002). Vegetation and soil lines in visible spectral space: a concept and technique for remote estimation of vegetation fraction. *Int. j. Remote Sensing*, 23(13), 2537–2562. <https://doi.org/10.1080/01431160110107806>

Gobernación del Departamento del Cesar. (2020). *Departamento del Cesar*. <http://cesar.gov.co/d/index.php/es/>

González Zárate, M. (2008). Estimación de la biomasa aérea y la captura de carbono en regeneración natural de *Pinus masiminoi* H. E. Moore, *Pinus oocarpa* var. *ochoterenai* Mtz, y *Quercus* sp. en el norte del estado de Chiapas, México. *Revista Mexicana de Ciencias Pecuarias*, 2(2). <https://buscador.una.edu.ni/Record/RepoCATIE1699>

-
- Guerini Filho, M., Kuplich, T. M., & Quadros, F. L. F. D. (2019). Estimating natural grassland biomass by vegetation indices using Sentinel 2 remote sensing data. *International Journal of Remote Sensing*, 41(8), 2861–2876. <https://doi.org/10.1080/01431161.2019.1697004>
- Guimar, M., Aes Da Silva, ~, Rosa Costa Muniz, A., Hoffmann, R., Carlos, A., Lisb, L., & Oa, ^ (2018). *Impact of greenhouse gases on surface coal mining in Brazil*. <https://doi.org/10.1016/j.jclepro.2018.05.076>
- Gutiérrez, G. A., Cotes, D. A., & Bastidas, M. J. (2018). Estudio de la Combustión Espontánea del Carbón durante el Acopio en la Mina El Hatillo, ubicada en Cesar, Colombia. *Información Tecnológica*, 29(6), 287–294. <https://doi.org/10.4067/S0718-07642018000600287>
- Hanke, J. E. (2006). Pronósticos en los negocios. *Pearson Educación de México, SA de CV*, 556P. https://books.google.com/books/about/Pron%C3%B3sticos_en_los_negocios_8a_ed.html?hl=es&id=JqqwtAEACAAJ
- Henao, S. A. F., A., J. D. M., & M., J. C. M. (2010). Análisis de emisiones de CO2 para diferentes combustibles en la población de taxis en Pereira y Dosquebradas. *Scientia et Technica*, 2(45). <https://doi.org/10.22517/23447214.385>
- Honorio, E. N., Timothy, C., & Baker, R. (2010). Manual para el monitoreo del ciclo del carbono en bosques amazónicos. *Universidad de Leeds. Lima*, 54p.
- Hou, H., Zhang, S., Ding, Z., Huang, A., & Tian, Y. (2015). Spatiotemporal dynamics of carbon storage in terrestrial ecosystem vegetation in the Xuzhou coal mining area, China. *Environmental Earth Sciences*, 74, 1657–1669. <https://doi.org/10.1007/S12665-015-4171-7/FIGURES/6>

- Houghton, R. A., Hall, F., & Goetz, S. J. (2009). Importance of biomass in the global carbon cycle. *Journal of Geophysical Research: Biogeosciences*, 114(G2), 1–13. <https://doi.org/10.1029/2009JG000935>
- Huete, A. R. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, 25(3), 295–309. [https://doi.org/10.1016/0034-4257\(88\)90106-X](https://doi.org/10.1016/0034-4257(88)90106-X)
- Huete, A. R., Liu, H. Q., Batchily, K., & van Leeuwen, W. (1997). A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *Remote Sensing of Environment*, 59(3), 440–451. [https://doi.org/10.1016/S0034-4257\(96\)00112-5](https://doi.org/10.1016/S0034-4257(96)00112-5)
- IPCC. (2006). 2006 IPCC Guidelines for National Greenhouse Gas Inventories, Prepared by the National Greenhouse Gas Inventories Programme. *Eggleston H.S., Buendia L., Miwa K., Ngara T. y Tanabe K. (Eds) Publicado Por: IGES, Japón*. . https://www.ipcc-nggip.iges.or.jp/public/2006gl/spanish/pdf/0_Overview/V0_0_Cover.pdf
- IPCC. (2013). *Glosario [Planton, S. (ed.)]. En: Cambio Climático 2013. Bases físicas. Contribución del Grupo de trabajo I al Quinto Informe de Evaluación del Grupo Intergubernamental de Expertos sobre el Cambio Climático [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex y P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, Reino Unido y Nueva York, NY, Estados Unidos de América.* https://www.ipcc.ch/site/assets/uploads/2018/08/WGI_AR5_glossary_ES.pdf
- Irene Moral Peláez. (2016). Modelos de regresión: lineal simple y regresión logística. *SEDEN*, 195–214.

-
- Jutz, S., & Milagro-Pérez, M. P. (2018). Copernicus Program. In L. Shunlin (Ed.), *1.06 - Comprehensive Remote Sensing* (pp. 150–191). Elsevier. <https://doi.org/10.1016/B978-0-12-409548-9.10317-3>
- Karan, S. K., Samadder, S. R., & Maiti, S. K. (2016). Assessment of the capability of remote sensing and GIS techniques for monitoring reclamation success in coal mine degraded lands. *Journal of Environmental Management*, *182*, 272–283. <https://doi.org/10.1016/J.JENVMAN.2016.07.070>
- Kaufman, Y. J., & Tanré, D. (1992). Atmospherically Resistant Vegetation Index (ARVI) for EOS-MODIS. *IEEE Transactions on Geoscience and Remote Sensing*, *30*(2), 261–270. <https://doi.org/10.1109/36.134076>
- Kayranli, B., Scholz, M., Mustafa, A., & Hedmark, Å. (2009). Carbon Storage and Fluxes within Freshwater Wetlands: a Critical Review. *Wetlands*, *30*(1), 111–124. <https://doi.org/10.1007/S13157-009-0003-4>
- Khanal, P. N., Grebner, D. L., Munn, I. A., Grado, S. C., Grala, R. K., & Henderson, J. E. (2017). Evaluating non-industrial private forest landowner willingness to manage for forest carbon sequestration in the southern United States. *Forest Policy and Economics*, *75*, 112–119. <https://doi.org/10.1016/J.FORPOL.2016.07.004>
- Kolecka, N., Ginzler, C., Pazur, R., Price, B., & Verburg, P. H. (2018). Regional Scale Mapping of Grassland Mowing Frequency with Sentinel-2 Time Series. *Remote Sensing 2018, Vol. 10, Page 1221*, *10*(8), 1221. <https://doi.org/10.3390/RS10081221>

- Laredo C., M. A., & Gómez S., J. (1982). Valor nutritivo de pastos tropicales, 4: Pasto Carimagua-1 (*Andropogon gayanus* Kunth) anual y estacional. *Revista ICA*, 17(1), 29–36. <https://repository.agrosavia.co/handle/20.500.12324/35352>
- Lechner, A. M., Kassulke, O., & Unger, C. (2016). Spatial assessment of open cut coal mining progressive rehabilitation to support the monitoring of rehabilitation liabilities. *Resources Policy*, 50, 234–243. <https://doi.org/10.1016/J.RESOURPOL.2016.10.009>
- Li, W. J., Wang, Z., Han, Q. F., Ren, C. H., Yan, M. K., Zhang, P., Jia, Z. K., & Yang, B. P. (2013). Evaluation on carbon sequestration effects of artificial alfalfa pastures in the Loess Plateau area. *Acta Ecologica Sinica*, 33(23), 7467–7477. <https://doi.org/10.5846/STXB201209101276>
- Liu, Y., Wang, Y., & Zhang, J. (2012). New machine learning algorithm: Random forest. *ICICA 2012: Information Computing and Applications*, 7473, 246–252. https://doi.org/10.1007/978-3-642-34062-8_32/COVER
- López Betancur, S. I. (1998). *Operacionalización de Variables* (p. 8p). Universidad de Caldas . http://fcaenlinea.unam.mx/anexos/1349/1349_u2_Act2.pdf
- López-Guerrero, I., Fontenot, J. P., & Beatriz García-Peniche, T. (2011). Comparaciones entre cuatro métodos de estimación de biomasa en praderas de festuca alta Comparison of four biomass estimation methods in Tall Fescue pastures. *Rev Mex Cienc Pecu*, 2(2), 209–220.
- Lu, D. (2007). The potential and challenge of remote sensing-based biomass estimation. *International Journal of Remote Sensing* , 27(7), 1297–1328. <https://doi.org/10.1080/01431160500486732>

-
- Mansaray, L. R., Zhang, K., & Kanu, A. S. (2020). Dry biomass estimation of paddy rice with Sentinel-1A satellite data using machine learning regression algorithms. *Computers and Electronics in Agriculture*, 176, 105674. <https://doi.org/10.1016/J.COMPAG.2020.105674>
- Marabel-García, M. (2016). *DEPARTAMENTO DE INGENIERÍA Y CIENCIAS AGRARIAS ESCUELA SUPERIOR Y TÉCNICA DE INGENIERÍA AGRARIA ESTIMACIÓN DE LA BIOMASA AÉREA DE HERBÁCEAS CON SENSORES AEROTRANSPORTADOS DE ALTA RESOLUCIÓN*.
<https://buleria.unileon.es/bitstream/handle/10612/5735/Tesis%20Miguel%20Marabel.pdf?sequence=1&isAllowed=y>
- Marove, C. A., Sotozono, R., Tangviroon, P., Tabelin, C. B., & Igarashi, T. (2022). Assessment of soil, sediment and water contaminations around open-pit coal mines in Moatize, Tete province, Mozambique. *Environmental Advances*, 8. <https://doi.org/10.1016/J.ENVADV.2022.100215>
- Mendoza Daza, A. D. (2009). *Prevención y control de autocombustión del carbón in situ en la mina Carbones del Cerrejón Limited* [Universidad Nacional de Colombia, sede medellin.].
<https://repositorio.unal.edu.co/handle/unal/2584>
- Michie, D. (1968). "Memo" Functions and Machine Learning. *Nature*, 218, 19–22.
<https://doi.org/10.1038/218019a0>
- Ministerio de ambiente y desarrollo sostenible. (2017). *Política nacional de cambio climático* (Á. Barragan, Ed.). www.minambiente.gov.co
- Ministerio de Ambiente y Desarrollo Sostenible. (2017). *POLÍTICA NACIONAL DE CAMBIO CLIMÁTICO*. www.minambiente.gov.co

Ministerio de minas y energía. (2003). *Glosario técnico minero* (p. 168p).
<https://www.anm.gov.co/sites/default/files/DocumentosAnm/glosariominero.pdf>

Resolucion 40807, (2018) (testimony of Ministerio de minas y energía).

Ministerio de minas y energia. (2018, August 2). *Resolución 40807 de 2018*.
https://oab.ambientebogota.gov.co/?post_type=dlm_download&p=21766

Mohd Zaki, N. A., & Abd Latif, Z. (2016). Carbon sinks and tropical forest biomass estimation: a review on role of remote sensing in aboveground-biomass modelling. *Geocarto International*, 32(7), 701–716. <https://doi.org/10.1080/10106049.2016.1178814>

Mónaco, N., V., S., M.J., R., & V., A. (2017). Evaluación De Métodos Indirectos Para Estimar Biomasa En Un Pastizal Natural Del Sur De Córdoba (Argentina Central). *European Scientific Journal, ESJ*, 13(36), 59. <https://doi.org/10.19044/esj.2017.v13n36p59>

Muñoz Aguayo, P. (2013). Apuntes de teledetección: índices de vegetación. In *Centro de Información de Recursos Naturales* (p. 15p). CIREN.
<https://bibliotecadigital.ciren.cl/handle/20.500.13082/26389>

Naciones Unidas. (1998). *PROTOCOLO DE KYOTO DE LA CONVENCIÓN MARCO DE LAS NACIONES UNIDAS SOBRE EL CAMBIO CLIMÁTICO*.
<https://unfccc.int/resource/docs/convkp/kpspan.pdf>

Naidoo, L., van Deventer, H., Ramoelo, A., Mathieu, R., Nondlazi, B., & Gangat, R. (2019). Estimating above ground biomass as an indicator of carbon storage in vegetated wetlands of the grassland biome of South Africa. *International Journal of Applied Earth Observation and Geoinformation*, 78, 118–129. <https://doi.org/10.1016/J.JAG.2019.01.021>

-
- Navarro, J. A., Algeet, N., Fernández-Landa, A., Esteban, J., Rodríguez-Noriega, P., & Guillén-Climent, M. L. (2019). Integration of UAV, Sentinel-1, and Sentinel-2 Data for Mangrove Plantation Aboveground Biomass Monitoring in Senegal. *Remote Sensing*, 11(1), 77. <https://doi.org/10.3390/RS11010077>
- Pérez Gutiérrez, C., & Muñoz Nieto, Á. L. (coord.). (2002). *Teledetección : nociones y aplicaciones*. Universidad de Salamanca. <https://dialnet.unirioja.es/servlet/libro?codigo=5256>
- Pérez Gutiérrez, C., Pérez Gutiérrez, C. (coord.), & Muñoz Nieto, Á. L. (coord.). (2002). Introducción a la Teledetección Espacial. *Teledetección : Nociones y Aplicaciones*, 1–15. <https://dialnet.unirioja.es/servlet/articulo?codigo=1010787>
- Pérez Trujillo, M. A. (2021). *Técnicas de Minería de Datos Supervisadas en el Espacio de Vida de Adultos Mayores de la Ciudad de Manizales* [Universidad Nacional de Colombia]. <https://repositorio.unal.edu.co/bitstream/handle/unal/80826/1075282588.2021.pdf?sequence=2&isAllowed=y>
- Petropoulos, G. P., Partsinevelos, P., & Mitraka, Z. (2012). Change detection of surface mining activity and reclamation based on a machine learning approach of multi-temporal Landsat TM imagery. *Geocarto International*, 28(4), 323–342. <https://doi.org/10.1080/10106049.2012.706648>
- Pham, T. D., Yoshino, K., Le, N. N., & Bui, D. T. (2018). Estimating aboveground biomass of a mangrove plantation on the Northern coast of Vietnam using machine learning techniques with an integration of ALOS-2 PALSAR-2 and Sentinel-2A data. *International Journal of Remote Sensing*, 39(22), 7761–7788. <https://doi.org/10.1080/01431161.2018.1471544>

- Pietrzykowski, M., & Daniels, W. L. (2014). Estimation of carbon sequestration by pine (*Pinus sylvestris* L.) ecosystems developed on reforested post-mining sites in Poland on differing mine soil substrates. *Ecological Engineering*, 73, 209–218. <https://doi.org/10.1016/J.ECOLENG.2014.09.058>
- Placek-Lapaj, A., Grobelak, A., Fijalkowski, K., Singh, B. R., & Almás, Á. R. (2019). *Post – Mining soil as carbon storehouse under polish conditions*. *Journal of Environmental Management*. <https://doi.org/doi.org/10.1016/j.jenvman.2019.03.005>
- Post, W. M., & Kwon, K. C. (2008). Soil carbon sequestration and land-use change: processes and potential. *Global Change Biology*, 6(3), 317–327. <https://doi.org/10.1046/J.1365-2486.2000.00308.X>
- Programa de las Naciones Unidas para el Desarrollo. (2012). *Objetivos de Desarrollo Sostenible | Programa De Las Naciones Unidas Para El Desarrollo*. <https://www.undp.org/es/sustainable-development-goals>
- Punalekar, S. M., Verhoef, A., Quaife, T. L., Humphries, D., Bermingham, L., & Reynolds, C. K. (2018). Application of Sentinel-2A data for pasture biomass monitoring using a physically based radiative transfer model. *Remote Sensing of Environment*, 218, 207–220. <https://doi.org/10.1016/J.RSE.2018.09.028>
- Qi, J., Chehbouni, A., Huete, A. R., Kerr, Y. H., & Sorooshian, S. (1994). A modified soil adjusted vegetation index. *Remote Sensing of Environment*, 48(2), 119–126. [https://doi.org/10.1016/0034-4257\(94\)90134-1](https://doi.org/10.1016/0034-4257(94)90134-1)

-
- Raval, S., Sarver, E., Shamsoddini, A., & Zipper, C. E. (2018). Satellite remote sensing-based: Estimates of biomass production on reclaimed coal mines. *Mining Engineering*, 66(4), 76–82. www.miningengineeringmagazine.com
- Rezaei, N., & Jabbari, P. (2022). K-nearest neighbors in R. *Immunoinformatics of Cancers*, 181–190. <https://doi.org/10.1016/B978-0-12-822400-7.00006-3>
- Rodrigo, J. A. (2020, October). *Random Forest con Python*. Association for Computing Machinery. <https://doi.org/10.1145/2939672.2939785>
- Rondeaux, G., Steven, M., & Baret, F. (1996). Optimization of soil-adjusted vegetation indices. *Remote Sensing of Environment*, 55(2), 95–107. [https://doi.org/10.1016/0034-4257\(95\)00186-7](https://doi.org/10.1016/0034-4257(95)00186-7)
- Rossini, M., Migliavacca, M., Galvagno, M., Meroni, M., Cogliati, S., Cremonese, E., Fava, F., Gitelson, A., Julitta, T., di Cella, U. M., Siniscalco, C., & Colombo, R. (2014). Remote estimation of grassland gross primary production during extreme meteorological seasons. *International Journal of Applied Earth Observation and Geoinformation*, 29(1), 1–10. <https://doi.org/10.1016/J.JAG.2013.12.008>
- Rouse, J. W. , Jr., Haas, R. H., Schell, J. A., Deering, D. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. NASA. *Goddard Space Flight Center 3d ERTS-1 Symp., Vol. 1, Sect. A*, 309–317. <https://ntrs.nasa.gov/citations/19740022614>
- Salas Macías, C. A., Alegre Orihuela, J. C., & Iglesias Abad, S. (2017). Estimation of above-ground live biomass and carbon stocks in different plant formations and in the soil of dry forests of the Ecuadorian coast. *Food and Energy Security*, 6(4). <https://doi.org/10.1002/FES3.115>

- Sanci, R., Panarello, H. O., & Ostera, y H. A. (2010). Flujo de dióxido de carbono en el flanco oriental del volcán Peteroa, Andes del Sur. *Revista Mexicana de Ciencias Geológicas*, 27(2), 225–237.
- Santiago, M. M., LastNameAlves Da Silva, H., Domiciano, J., Tiago, G., & LastNameLastNameDe Oliveira, H. (2009). Análise da Cobertura Vegetal Através dos Índices de Vegetação (NDVI, SAVI e IAF) no Entorno da Barragem do Botafogo-PE. *XIV Simpósio Brasileiro de Sensoriamento Remoto*, 3003–3009.
- Shackelford, N., Miller, B. P., & Erickson, T. E. (2018). Restoration of Open-Cut Mining in Semi-Arid Systems: A Synthesis of Long-Term Monitoring Data and Implications for Management. *Land Degradation & Development*, 29(4), 994–1004. <https://doi.org/10.1002/LDR.2746>
- Snowdon, P., Raison, R. J., Keith, H., Ritson, P., Grierson, P., Adams, M., Montagu, K., Bi, H.-Q., Burrows, W., & Eamus, D. (2002). Protocol for sampling tree and stand biomass. In *Environment Australia Catalogación en la publicación* (p. P66). Canberra, Australian Greenhouse Office. https://www.researchgate.net/publication/298252249_Protocol_for_Sampling_Tree_and_Stand_Biomass
- Soares Miranda, V., Guimarães Ribeiro, K., Christófaro Silva, A., Cristina Pereira, R., Gomes Pereira, O., Vidal Torrado, P., Sebastião Cunha Fernandes, J., Coura Oliveira, M., & Brasileira de Zootecnia, R. (2012). Rehabilitation with forage grasses of an area degraded by urban solid waste deposits. *Revista Brasileira de Zootecnia*, 21(1), 18–23. www.sbz.org.br
- Song, C., Chen, J. M., Hwang, T., Hakkenberg, C., & Li, J. (2015). Ecological characterization of vegetation using multisensor remote sensing in the solar reflective spectrum. En Prasad

Thenkabail. In *Land Resources Monitoring, Modeling, and Mapping with Remote Sensing* (pp. 533–575). United States: Geological Survey. <https://doi.org/10.1201/b19322>

Sousa, A. M. O., Marques Da Silva, J. R., Gonçalves, A. C., Mesquita, P. A., Silva, L. L., & Batista, F. (2013). ESTIMACIÓN DE LA BIOMASA FORESTAL PARA LAS ESPECIES QUERCUS ROTUNDIFOLIA Y QUERCUS SUBER BASADA EN IMÁGENES DEL SATÉLITE QUICKBIRD. In T. de A. INTA (Ed.), *XV Congreso de la Asociación Española de Teledetección*.

[https://dspace.uevora.pt/rdpc/bitstream/10174/8936/1/XV_CongressoAET_out2013_BIOMAS SA.pdf](https://dspace.uevora.pt/rdpc/bitstream/10174/8936/1/XV_CongressoAET_out2013_BIOMAS_SA.pdf)

Tanré, D., Holben, B. N., & Kaufman, Y. J. (1992). Atmospheric Correction Algorithm for NOAA-AVHRR Products: Theory and Application. *IEEE Transactions on Geoscience and Remote Sensing*, *30*(2), 231–248. <https://doi.org/10.1109/36.134074>

Tao, S., Guo, Q., Wu, F., Li, L., Wang, S., Tang, Z., Xue, B., Liu, J., & Fang, J. (2016). Spatial scale and pattern dependences of aboveground biomass estimation from satellite images: a case study of the Sierra National Forest, California. *Landscape Ecology*, *31*(8), 1711–1723. <https://doi.org/10.1007/S10980-016-0357-Y/FIGURES/5>

Toosi, A., Javan, F. D., Samadzadegan, F., Mehravar, S., Kurban, A., & Azadi, H. (2022). Citrus orchard mapping in Juybar, Iran: Analysis of NDVI time series and feature fusion of multi-source satellite imageries. *Ecological Informatics*, *70*, 101733. <https://doi.org/10.1016/J.ECOINF.2022.101733>

Tripathi, N., Singh, R. S., & Nathanail, P. (2013). Mine spoil acts as a sink of carbon dioxide in Indian dry tropical environment. *Science of the Total Environment*, *468–469*, 1162–1171. <https://doi.org/10.1016/j.scitotenv.2013.09.024>

Unidad de Planeación Minero Energética. (2022). *Carbon*. <https://www1.upme.gov.co/simco/Cifras-Sectoriales/Paginas/carbon.aspx>

Vafaei, S., Soosani, J., Adeli, K., Fadaei, H., Naghavi, H., Pham, T. D., & Bui, D. T. (2018). Improving Accuracy Estimation of Forest Aboveground Biomass Based on Incorporation of ALOS-2 PALSAR-2 and Sentinel-2A Imagery and Machine Learning: A Case Study of the Hyrcanian Forest Area (Iran). *Remote Sensing* 2018, Vol. 10, Page 172, 10(2), 172. <https://doi.org/10.3390/RS10020172>

Vapnik, V. (1998). The Support Vector Method of Function Estimation. In: Suykens, J.A.K., Vandewalle, J. (Eds) *Nonlinear Modeling*. Springer, Boston, MA., 55–85. https://doi.org/10.1007/978-1-4615-5703-6_3

Vásquez Sierra, E. B., & Herrera Builes, J. F. (2006). METODOLOGÍA PARA LA CARACTERIZACIÓN DE COMBUSTIBLES SÓLIDOS MADERABLES DEL ÁREA METROPOLITANA DEL VALLE DE ABURRÁ “AMVA”, COLOMBIA. *Revista Facultad Nacional de Agronomía Medellín*, 59(2), 3557–3564. http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0304-28472006000200011&lng=en&nrm=iso&tlng=es

Vreugdenhil, M., Wagner, W., Bauer-Marschallinger, B., Pfeil, I., Teubner, I., Rüdiger, C., & Strauss, P. (2018). Sensitivity of Sentinel-1 Backscatter to Vegetation Dynamics: An Austrian Case Study. *Remote Sensing* 2018, Vol. 10, Page 1396, 10(9), 1396. <https://doi.org/10.3390/RS10091396>

Wang, J., Xiao, X., Bajgain, R., Starks, P., Steiner, J., Doughty, R. B., & Chang, Q. (2019). Estimating leaf area index and aboveground biomass of grazing pastures using Sentinel-1,

Sentinel-2 and Landsat images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 154, 189–201. <https://doi.org/10.1016/J.ISPRSJPRS.2019.06.007>

Wang, L., Zhou, X., Zhu, X., Dong, Z., & Guo, W. (2016). Estimation of biomass in wheat using random forest regression algorithm and remote sensing data. *The Crop Journal*, 4(3), 212–219. <https://doi.org/10.1016/J.CJ.2016.01.008>

Wang, Y., Zhang, Z., Feng, L., Du, Q., & Runge, T. (2020). Combining Multi-Source Data and Machine Learning Approaches to Predict Winter Wheat Yield in the Conterminous United States. *Remote Sensing*, 12(8), 1232. <https://doi.org/10.3390/RS12081232>

Wang, Z., Lechner, A. M., Yang, Y., Baumgartl, T., & Wu, J. (2020). Mapping the cumulative impacts of long-term mining disturbance and progressive rehabilitation on ecosystem services. *Science of The Total Environment*, 717, 137214. <https://doi.org/10.1016/J.SCITOTENV.2020.137214>

Xiao, J., Chevallier, F., Gomez, C., Guanter, L., Hicke, J. A., Huete, A. R., Ichii, K., Ni, W., Pang, Y., Rahman, A. F., Sun, G., Yuan, W., Zhang, L., & Zhang, X. (2019). Remote sensing of the terrestrial carbon cycle: A review of advances over 50 years. *Remote Sensing of Environment*, 233, 111383. <https://doi.org/10.1016/J.RSE.2019.111383>

XiaoPing, X., Lei, D., Wei, C., XiaoYu, Z., BaoRui, C., ZhongLing, L., GuangLi, H., GeLe, Q., GuiXia, Y., & HuaJun, T. (2020). Biomass Carbon Storage and Its Effect Factors in Steppe and Agro-Pastoral Ecotones in Northern China. *Scientia Agricultura Sinica*, 53(13), 2757–2768. <https://doi.org/10.3864/J.ISSN.0578-1752.2020.13.022>

Xue, Y. (2020). Empirical research on household carbon emissions characteristics and key impact factors in mining areas. *Journal of Cleaner Production*, 256, 120470. <https://doi.org/10.1016/j.jclepro.2020.120470>

Yaranga Cano, R. M., & Custodio Villanueva, M. (2013). *Storage of carbon in natural grasses high andean*. 4, 313–319. www.sci-agropecu.unitru.edu.pe

Yepes Quintero, A. Patricia., Duque Montoya, A. Javier., Navarrete Encinales, D. Alejandro., Phillips Bernal, J. Fernando., & Instituto de Hidrología, M. y E. A. (Colombia). (2011). Protocolo para la estimación nacional y subnacional de biomasa-carbono en Colombia. *Instituto de Hidrología, Meteorología, y Estudios Ambientales-IDEAM-. Bogotá D.C.*, 162 p. http://www.ideam.gov.co/documents/13257/13548/Protocolo+para+la+estimaci%C3%B3n+nacional+y+subnacional_1.pdf/11c9d26b-5a03-4d13-957e-0bcc1af8f108

Zhu, L., Song, R., Sun, S., Li, Y., & Hu, K. (2022). Land use/land cover change and its impact on ecosystem carbon storage in coastal areas of China from 1980 to 2050. *Ecological Indicators*, 142. <https://doi.org/10.1016/j.ecolind.2022.109178>

Zuluaga, L. Z., Stiven, E., & Escobar, C. (2018). ASSESSMENT OF ENVIRONMENTAL SERVICES BY CO₂ CAPTURE IN A TROPICAL DRY FOREST ECOSYSTEM IN THE MUNICIPALITY OF EL CARMEN DE BOLIVAR, COLOMBIA. *Luna Azul*, 47, 1–20. <https://doi.org/10.17151/luaz.2019.47.1>

Zuluaga Zuluaga, L., & Castro Escobar, E. S. (2018). Valoración de servicios ambientales por captura de CO₂ en un ecosistema de Bosque seco Tropical en el municipio de El Carmen de Bolívar, Colombia. *Luna Azul*, 47, 01–20. <https://doi.org/10.17151/LUAZ.2019.47.1>

ANEXOS

Anexo 1. Algoritmo para regresión por Mínimos cuadrados ordinarios

```
# Importing the libraries

import numpy as np
import pandas as pd
import requests, io
import matplotlib.pyplot as plt
import seaborn as sns

%matplotlib inline

# scikit-learn modules

from sklearn.model_selection import train_test_split
from sklearn.metrics import mean_squared_error
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error, r2_score
from sklearn import metrics

#Load data

df = pd.read_table('C:/Users/USER/Documents/jupyter/BDD20/20.txt')

df.head()

#Load info data

(df.info())

#define the target(y) and learning variable(x)

ARVI = df.iloc[:, 0]
AGB = df.iloc[:, -1]

#Splitting the dataset

ARVI_train, ARVI_test, AGB_train, AGB_test = train_test_split(ARVI, AGB)
```

```
#OLS model
olsreg = LinearRegression()

#Fitting the model
olsreg.fit(ARVI_train.values.reshape(-1,1), AGB_train.values)
ols = olsreg.fit(ARVI_train.values.reshape(-1,1), AGB_train.values)

#Prediction on the testing data
y_pred = olsreg.predict(ARVI_train.values.reshape(-1,1))
y_pred

#Create first simple correlation matrix
corrMatrix = df.corr()
sns.heatmap(corrMatrix, annot=True)
plt.show()

#Draw training, testing and prediction data
plt.scatter(ARVI_train, AGB_train, label='Training Data', s=60, c='blue')
plt.scatter(ARVI_test, AGB_test, label='Testing Data', s=60, c='green')
plt.legend()
plt.plot(ARVI_train.values.reshape(-1,1), y_pred, color='red')
plt.show()

#Regression model evaluation
print ("Mean Squared Error:" ,
metrics.mean_squared_error(AGB_train.values.reshape(-1,1), y_pred))

print ("Root Mean Squared Error:" ,
np.sqrt(metrics.mean_squared_error(AGB_train.values.reshape(-1,1), y_pred)))

print('Coefficient of determination: %.2f' % r2_score(AGB_train.values.reshape(-1,1),
y_pred))

#Raster generation of the model
inpRaster = 'C:/Users/USER/Documents/jupyter/BDD20/ARVI.tif' #define path to the
input raster (on which we will make predic
```

```
outraster = 'C:/Users/USER/Documents/jupyter/BDD20/AGB_OLS_ARVI.tif' #spath to
the output raster - where it will be saved

df = pd.read_table('C:/Users/USER/Documents/jupyter/BDD20/20.txt') #apath to the
dataset with the training data

data = df.iloc[:, 0].values #List of predictors

label = df.iloc[:, 1].values

ds = gdal.Open(inpRaster, gdal.GA_ReadOnly) #with gdal module we will open the
input raster

rows = ds.RasterYSize

cols = ds.RasterXSize

bands = ds.RasterCount

geo_transform = ds.GetGeoTransform()

projection = ds.GetProjectionRef()

array = ds.ReadAsArray()

ds = None

array = np.stack(array,axis=2)

array = np.reshape(array, [rows*cols,bands])

test = pd.DataFrame(array, dtype='float32')

outRaster = 'C:/Users/USER/Documents/jupyter/BDD20/AGB_OLS_ARVI.tif'
```

Anexo 2. Algoritmo para regresión por K-vecinos más cercanos

```
# Importing the libraries
import numpy as np
import pandas as pd
import requests, io
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
# scikit-learn modules
from sklearn.model_selection import train_test_split
from sklearn.metrics import mean_squared_error
from sklearn import preprocessing
from sklearn.metrics import mean_squared_error, r2_score
from sklearn import metrics
from sklearn.neighbors import KNeighborsRegressor
from sklearn import neighbors
from sklearn.metrics import mean_squared_error
from yellowbrick.regressor import ResidualsPlot
#Load data
df = pd.read_table('C:/Users/USER/Documents/jupyter/BDD20/20.txt')
df.head()
#Load info data
(df.info())
#define the target(y) and learning ARVIable(x)
ARVI = df.iloc[:, 0]
AGB = df.iloc[:, -1]
#Splitting the dataset
```

10 USO DE TELEDETECCIÓN EN LA ESTIMACIÓN DEL CARBONO DE LA BIOMASA HERBÁCEA EN TERRENOS REHABILITADOS POR LA EMPRESA DRUMMOND LTD EN EL DEPARTAMENTO DEL CESAR-COLOMBIA

```
ARVI_train, ARVI_test, AGB_train, AGB_test = train_test_split(ARVI, AGB, test_size = 0.20,
random_state = 0)

#KNN model

knnreg = KNeighborsRegressor(n_neighbors=2)

#Fitting the model

knnreg.fit(ARVI_train.values.reshape(-1,1), AGB_train)

mknn = knnreg.fit(ARVI_train.values.reshape(-1,1), AGB_train)

#Prediction on the testing data

AGB_pred = knnreg.predict((ARVI_train.values.reshape(-1,1)))

AGB_pred

#Elbow method RMSE vs K Value

rmse_val = []

for K in range(7):

    K = K+1

    modelknn = neighbors.KNeighborsRegressor(n_neighbors = K)

    modelknn.fit(ARVI_train.values.reshape(-1,1), AGB_train)

    pred=modelknn.predict(ARVI_test.values.reshape(-1,1))

    error = sqrt(mean_squared_error(AGB_test,pred))

    rmse_val.append(error) #store rmse values

    print('Valor de RMSE para k= ', K , 'es:', error)

#Draw training, testing and prediction data

plt.scatter(AGB_train, AGB_PRED_train, label='Training Data', s=60, c='blue')

plt.legend()

plt.plot(AGB_train.values.reshape(-1,1), y_pred, color ='red')

plt.show()

#Regression model evaluation

print ("Mean Squared Error:" , metrics.mean_squared_error(AGB_train.values.reshape(-1,1),
AGB_pred))
```

```
print("Root Mean Squared Error:",
      np.sqrt(metrics.mean_squared_error(AGB_train.values.reshape(-1,1), AGB_pred)))

print('Coefficient of determination: %.2f%' % r2_score(AGB_train.values.reshape(-1,1),
      AGB_pred))

#Raster generation of the model

from osgeo import gdal

inpRaster = 'C:/Users/USER/Documents/jupyter/BDD20/ARVI.tif'

outraster = 'C:/Users/USER/Documents/jupyter/BDD20/AGB_KNN_ARVI.tif'

df = pd.read_table('C:/Users/USER/Documents/jupyter/BDD20/20.txt')

data = df.iloc[:, 0].values

label = df.iloc[:, 1].values

ds = gdal.Open(inpRaster, gdal.GA_ReadOnly)

rows = ds.RasterYSize

cols = ds.RasterXSize

bands = ds.RasterCount

geo_transform = ds.GetGeoTransform()

projection = ds.GetProjectionRef()

array = ds.ReadAsArray()

ds = None

array = np.stack(array,axis=2)

array = np.reshape(array, [rows*cols,bands])

test = pd.DataFrame(array, dtype='float32')

outRaster = 'C:/Users/USER/Documents/jupyter/BDD20/AGB_KNN_ARVI.tif'
```

Anexo 3. Algoritmo para regresión por Máquina vectorial de soporte

```
# Importing the libraries

import numpy as np
import pandas as pd
import requests, io
import matplotlib.pyplot as plt
import seaborn as sns

%matplotlib inline

# scikit-learn modules

from sklearn.model_selection import train_test_split
from sklearn.metrics import mean_squared_error
from sklearn.svm import SVR
from sklearn.metrics import mean_squared_error, r2_score
from sklearn import metrics
from yellowbrick.regressor import ResidualsPlot
from sklearn.model_selection import cross_val_score
from sklearn.model_selection import GridSearchCV

#Load data

df = pd.read_table('C:/Users/USER/Documents/jupyter/BDD20/20.txt')

df.head()

#Load info data

(df.info())

#define the target(y) and learning MSAVI2able(x)

MSAVI2 = df.iloc[:, 2]
```

```
AGB = df.iloc[:, -1]
#Splitting the dataset
rand_state=0
MSAVI2_train, MSAVI2_test, AGB_train, AGB_test = train_test_split(MSAVI2, AGB,
test_size = 0.20, random_state = rand_state)
#SVR model
svrreg = SVR()
#Fitting the model
svrreg.fit(MSAVI2_train.values.reshape(-1,1), AGB_train)
msvr = svrreg.fit(MSAVI2_train.values.reshape(-1,1), AGB_train)
#Prediction on the testing data
AGB_train_pred = svrreg.predict((MSAVI2_train.values.reshape(-1,1)))
AGB_train_pred
#Regression model evaluation
print          ("Mean          Squared          Error:"
metrics.mean_squared_error(AGB_train.values.reshape(-1,1), AGB_pred))
print          ("Root          Mean          Squared          Error:"
np.sqrt(metrics.mean_squared_error(AGB_train.values.reshape(-1,1), AGB_pred)))
print('Coefficient of determination: %.2f' % r2_score(AGB_train.values.reshape(-1,1),
AGB_pred))
#Tuning hyperparameters
my_param_grid = {'C': [1,10,100], 'gamma': [1,0.1,0.01], 'kernel': ['rbf']}
#Import gridsearchCV
GridSearchCV(estimator=SVR(),param_grid= my_param_grid, refit = True,
verbose=3, cv=5 )
#Define grid
grid = GridSearchCV(estimator=SVR(),param_grid= my_param_grid, refit = True,
verbose=2, cv=5 )
```

```
#Iteration for parameter adjustment
grid.fit(MSAVI2_train.values.reshape(-1,1), AGB_train)
#the best parameter
grid.best_params_
#the best estimator
SVR(C=100, gamma=1)
#Optimized prediction data
AGB_train_pred_optimized = grid.predict(MSAVI2_train.values.reshape(-1,1))
AGB_train_pred_optimized
#Regression model evaluation
print          ("Mean          Squared          Error:"          ,
metrics.mean_squared_error(AGB_train.values.reshape(-1,1),
AGB_train_pred_optimized))
print          ("Root          Mean          Squared          Error:"          ,
np.sqrt(metrics.mean_squared_error(AGB_train.values.reshape(-1,1),
AGB_train_pred_optimized)))
print('Coefficient of determination: %.2f'% r2_score(AGB_train.values.reshape(-1,1),
AGB_train_pred_optimized))
#Draw training, testing and prediction data
plt.scatter(AGB_train, AGB_train_pred_optimized_train, label='Training Data', s=60,
c='blue')
plt.legend()
plt.plot(AGB_train.values.reshape(-1,1), AGB_train_pred_optimized, color='red')
plt.show()
#Raster generation of the model
from osgeo import gdal
inpRaster = 'C:/Users/USER/Documents/jupyter/BDD20/MSAVI2.tif'
outraster = 'C:/Users/USER/Documents/jupyter/BDD20/AGB_SVM_MSAVI2.tif'
```

```
df = pd.read_table('C:/Users/USER/Documents/jupyter/BDD20/20.txt')
data = df.iloc[:, 0].values
label = df.iloc[:, 1].values
ds = gdal.Open(inpRaster, gdal.GA_ReadOnly)
rows = ds.RasterYSize
cols = ds.RasterXSize
bands = ds.RasterCount
geo_transform = ds.GetGeoTransform()
projection = ds.GetProjectionRef()
array = ds.ReadAsArray()
ds = None
array = np.stack(array,axis=2)
array = np.reshape(array, [rows*cols,bands])
test = pd.DataFrame(array, dtype='float32')
outRaster = 'C:/Users/USER/Documents/jupyter/BDD20/AGB_SVM_MSAVI2.tif'
```

Anexo 4. Algoritmo para regresión por Random forest

```
# Importing the libraries

import numpy as np
import pandas as pd
import requests, io
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline

# scikit-learn modules
from sklearn.model_selection import train_test_split
from sklearn.metrics import mean_squared_error
from sklearn.metrics import mean_squared_error, r2_score
from sklearn import metrics
from sklearn.ensemble import RandomForestRegressor
from sklearn.model_selection import train_test_split
from yellowbrick.regressor import ResidualsPlot
from sklearn.model_selection import cross_val_score
from sklearn.model_selection import GridSearchCV

#Load data
df = pd.read_table('C:/Users/USER/Documents/jupyter/BDD20/20.txt')
df.head()

#Load info data
(df.info())

#define the target(y) and learning IVable(x)
IV = df.drop(columns = "MUESTRA")
```

```
AGB = df['MUESTRA']

#Splitting the dataset

rand_state=0

IV_train, IV_test, AGB_train, AGB_test = train_test_split(IV, AGB, test_size = 0.20,
random_state = rand_state)

#Random forest model

rfreg = RandomForestRegressor(n_estimators = 100, max_depth= None, n_jobs= -1,
random_state = 0)

#Fitting the model

rfreg.fit(IV_train, AGB_train)

rf = rfreg.fit(IV_train, AGB_train)

#Prediction on the testing data

AGB_train_pred = rfreg.predict((IV_train))

AGB_train_pred

#Regression model evaluation

print          ("Mean          Squared          Error:"
metrics.mean_squared_error(AGB_train.values.reshape(-1,1), AGB_pred))

print          ("Root          Mean          Squared          Error:"
np.sqrt(metrics.mean_squared_error(AGB_train.values.reshape(-1,1), AGB_pred)))

print('Coefficient of determination: %.2f%' r2_score(AGB_train.values.reshape(-1,1),
AGB_pred))

#Importance of predictors in the model

importancia_predictores = pd.DataFrame(

        {'predictor': df.drop('MUESTRA', axis = 1).columns,

        'importancia': rfreg.feature_importances_}

)

print("Importancia de los predictores en el modelo")

print("-----")
```

```
importancia_predictores.sort_values('importancia', ascending=False)

#Feature importance

plt.figure(figsize=(10,6))

plt.title('Feature Importance')

sns.barplot(y='predictor', x='importancia', data=importancia_predictores)

plt.show()

#Tuning hyperparameters

my_param_grid = {'n_estimators': [10,100,500], 'max_features':['sqrt','log2'],
'max_depth':[5,10,20]}

#Impor gridserchCV

GridSearchCV(estimator=RandomForestRegressor(random_state=rand_state),param
_grid= my_param_grid, refit = True, verbose=2, cv=5 )

grid =
GridSearchCV(estimator=RandomForestRegressor(random_state=rand_state),param
_grid= my_param_grid, refit = True, verbose=2, cv=5 )

#Iteration for parameter adjustment

grid.fit(IV_train,AGB_train)

#The best parameter

grid.best_params_

#The best estimator

grid.best_estimator_

#Optimized prediction data

AGB_train_pred_optimized = grid.predict((IV_train))

AGB_train_pred_optimized

#Iteration for parameter adjustment

grid.fit(IV_train.values.reshape(-1,1), AGB_train)

#the best parameter
```

```
grid.best_params_  
#the best estimator  
SVR(C=100, gamma=1)  
#Optimized prediction data  
AGB_train_pred_optimized = grid.predict(IV_train.values.reshape(-1,1))  
AGB_train_pred_optimized  
#Regression model evaluation  
print ("Mean Squared Error:" , metrics.mean_squared_error(AGB_train,  
AGB_train_pred_optimized))  
print ("Root Mean Squared Error:" , np.sqrt(metrics.mean_squared_error(AGB_train,  
AGB_train_pred_optimized)))  
print('Coefficient of determination: %.2f%' r2_score(AGB_train,  
AGB_train_pred_optimized))  
#Regression model evaluation  
print ("Mean Squared Error:" ,  
metrics.mean_squared_error(AGB_train.values.reshape(-1,1),  
AGB_train_pred_optimized))  
print ("Root Mean Squared Error:" ,  
np.sqrt(metrics.mean_squared_error(AGB_train.values.reshape(-1,1),  
AGB_train_pred_optimized)))  
print('Coefficient of determination: %.2f%' r2_score(AGB_train.values.reshape(-1,1),  
AGB_train_pred_optimized))  
#Draw training, testing and prediction data  
plt.scatter(AGB_train, AGB_train_pred_optimized_train, label='Training Data', s=60,  
c='blue')  
plt.legend()  
plt.plot(AGB_train.values.reshape(-1,1), AGB_train_pred_optimized, color='red')  
plt.show()  
#Raster generation of the model  
from osgeo import gdal
```

```
inpRaster = 'C:/Users/USER/Documents/jupyter/BDD20/MSAVI2.tif'
outraster = 'C:/Users/USER/Documents/jupyter/BDD20/AGB_RF_MSAVI2.tif'
df = pd.read_table('C:/Users/USER/Documents/jupyter/BDD20/20.txt')
data = df.iloc[:, 0].values
label = df.iloc[:, 1].values
ds = gdal.Open(inpRaster, gdal.GA_ReadOnly)
rows = ds.RasterYSize
cols = ds.RasterXSize
bands = ds.RasterCount
geo_transform = ds.GetGeoTransform()
projection = ds.GetProjectionRef()
array = ds.ReadAsArray()
ds = None
array = np.stack(array,axis=2)
array = np.reshape(array, [rows*cols,bands])
test = pd.DataFrame(array, dtype='float32')
outRaster = 'C:/Users/USER/Documents/jupyter/BDD20/AGB_RF_MSAVI2.tif'
```