

ANALISIS TEMPORAL DE LOS METODOS DE REFORESTACION APLICADOS POR
JUSTDIGGIT EN AREAS DE KENIA Y TANZANIA MEDIANTE IMÁGENES SATELITALES

Yenny Paola Betancur Torres

Maestría en Ingeniería y Analítica de Datos
Facultad de Ciencias Naturales e Ingeniería
Universidad de Bogotá Jorge Tadeo Lozano
Bogotá D.C., Colombia

2024

ANALISIS TEMPORAL DE LOS METODOS DE REFORESTACION APLICADOS POR
JUSTDIGGIT EN AREAS DE KENIA Y TANZANIA MEDIANTE IMÁGENES SATELITALES

Yenny Paola Betancur Torres

Tesis de grado presentado para optar por el título de Magíster en Ingeniería y Analítica de Datos.

Director de Tesis:

Ixent Galpin

Maestría en Ingeniería y Analítica de Datos
Facultad de Ciencias Naturales e Ingeniería
Fundación Universidad de Bogotá Jorge Tadeo Lozano
Bogotá D.C., Colombia

2024

Resumen

El presente trabajo analiza los cambios en la vegetación de áreas de Kenia y Tanzania, generados por las actividades de reforestación de la ONG Justdiggít. Estas actividades consisten en cavar semicírculos en el suelo con y sin siembra de semillas de pasto, con el fin de retener el agua de lluvias y que esta penetre en el suelo y las plantas puedan tomarla.

En este trabajo se emplean técnicas de teledetección, con el fin de calcular la variación en los índices NDVI, SAVI y NDWI. Además, se entrenaron los modelos de aprendizaje automático supervisado SVM, Random Forest y Árboles de decisión, para cuantificar los cambios en la cobertura de vegetación en imágenes satelitales. Estas imágenes hacen parte de las colecciones del satélite LandSat 8 y fueron procesadas en Google Earth Engine. Lo cual contempla tres años antes de las actividades de reforestación, hasta el año 2022, con el fin de evidenciar los cambios.

En el análisis de resultados se tiene en cuenta la cobertura de suelos, índice SAVI y precipitaciones, además de tener en cuenta la estacionalidad climática de la zona. Se observa que en promedio el método de reforestación con siembra de semillas de pasto aumentó el índice SAVI en 0.06 y el porcentaje de cobertura de vegetación en 3.39%. Por otro lado, el método de reforestación sin siembra, aunque disminuye el índice SAVI en 0.04, aumenta el porcentaje de cobertura de vegetación en 9.04%. Se concluye que la técnica con siembra de semillas de pasto tiene mejores resultados en comparación con la técnica de semicírculos sin siembra. No obstante, la sequía prolongada en la zona tuvo un importante impacto en los resultados observados.

Tabla de Contenido

Tabla de Contenido	4
1. Introducción.....	12
1.1. Planteamiento del problema.....	13
1.2. Objetivos.....	13
1.2.1. Objetivo general	13
1.2.2. Objetivos específicos	14
1.3. Metodología.....	14
2. Marco Teórico	16
2.1. Teledetección	16
2.2. Índices espectrales de vegetación	18
2.2.1. NDVI	18
2.2.2. SAVI.....	18
2.2.3. NDWI	19
2.3. Calentamiento Global, Erosión y Reforestación	19
2.4. Aprendizaje automático.....	19
2.4.1. Evaluación	20
3. Estado del arte.....	22
3.1. Estudios realizados en Colombia	22
3.2. Estudios realizados a nivel internacional.....	24
3.3. Estudios relacionados con Justdiggitt	27
4. Entendimiento del negocio y preparación de los datos	28
4.1. Fuente de datos	28
4.2. Definición de áreas de estudio	28
4.3. Exploración y selección de imágenes satelitales.....	31
4.3.1. Exploración de imágenes satelitales	31
4.3.2. Selección de imágenes satelitales.....	33

4.4.	Preparación de las imágenes satelitales	34
5.	Entrenamiento de modelos y evaluación	38
5.1.	Iteración I	38
5.1.1.	Definición de polígono de entrenamiento	38
5.1.2.	Captura de puntos de entrenamiento y pruebas	39
5.1.3.	Entrenamiento de modelos de aprendizaje automático	43
5.1.4.	Evaluación de modelos	46
5.2.	Iteración II	47
5.2.1.	Definición de polígono de entrenamiento	47
5.2.2.	Captura de puntos de entrenamiento y pruebas	48
5.2.3.	Entrenamiento de modelos de aprendizaje automático	49
5.2.4.	Evaluación de modelos	52
6.	Aplicación del modelo clasificación	54
6.1.	Polígonos con semicírculos de suelo sin siembra	55
6.1.1.	Inkiito.....	55
6.1.2.	KukuB	57
6.1.3.	KukuD	59
6.1.4.	Nasipa.....	61
6.1.5.	KukuA	63
6.1.6.	Risa.....	65
6.2.	Polígonos con semicírculos de suelo con siembra	67
6.2.1.	Pembamoto.....	67
6.2.2.	Enkii	69
6.2.3.	Baraka.....	71
6.2.4.	Engaruka.....	73
6.2.5.	EsilaleiSite1.....	74
6.2.6.	EsilaleiSite2.....	76

6.2.7.	Inkisanjani	78
6.2.8.	Meserani	80
6.2.9.	Selela.....	81
7.	Resultados y Discusión	84
8.	Conclusiones.....	91
9.	Referencias Bibliográficas	92

Lista de Figuras

Fig. 1 Flujo de trabajo.....	15
Fig. 2 Componentes de un sistema de teledetección	16
Fig. 3 Espectro electromagnético	17
Fig. 4 Curvas de reflectancia espectral.....	17
Fig. 5 Semicírculos de suelo en Kuku, Kenia.....	29
Fig. 6 Imagen LandSat 7	32
Fig. 7 Imagen LandSat 8	32
Fig. 8 NDVI preliminar del polígono Enkii	33
Fig. 9 Climatología mensual promedio de Kenia.....	33
Fig. 10 Climatología mensual promedio de Tanzania	34
Fig. 11 Composición de imágenes	35
Fig. 12 Metodología CRISP-DM	38
Fig. 13 Polígono de entrenamiento - Iteración I	39
Fig. 14 Comparación visualización NDVI e Imagen RGB	40
Fig. 15 Visualización SAVI Iteración I.....	41
Fig. 16 Visualización NDWI Iteración I.....	41
Fig. 17 Visualización zonas urbanas	42
Fig. 18 Visualización zonas agrícolas.....	42
Fig. 19 Composición de bandas	43
Fig. 20 Clasificación SVM de Iteración I	44
Fig. 21 Clasificación Arboles de decisión de Iteración I	45
Fig. 22 Clasificación Random Forest de Iteración I.....	46
Fig. 23 Matriz de confusión modelo SVM Iteración I.....	46
Fig. 24 Matriz de confusión modelo Arboles de decisión Iteración I.....	46
Fig. 25 Matriz de confusión modelo Random Forest Iteración I	47
Fig. 26 Costas de cuerpos de agua – Visualización RGB.....	47
Fig. 27 Polígono de entrenamiento - Iteración II	48
Fig. 28 Visualización SAVI Iteración II.....	49
Fig. 29 Clasificación SVM de Iteración II	50
Fig. 30 Clasificación Arboles de decisión de Iteración II	51
Fig. 31 Clasificación Random Forest de Iteración II.....	52
Fig. 32 Lluvias e índice SAVI del polígono Inkiito	56
Fig. 33 Lluvias y % de cobertura vegetal del polígono Inkiito.....	56

Fig. 34 índice SAVI y Cobertura vegetal del polígono Inkiito	57
Fig. 35 Lluvias e índice SAVI del polígono KukuB	58
Fig. 36 Lluvias y % de cobertura vegetal del polígono KukuB	58
Fig. 37 índice SAVI y Cobertura vegetal del polígono KukuB	59
Fig. 38 Lluvias e índice SAVI del polígono KukuD	60
Fig. 39 Lluvias y % de cobertura vegetal del polígono KukuD	60
Fig. 40 índice SAVI y Cobertura vegetal del polígono KukuD	61
Fig. 41 Lluvias e índice SAVI del polígono Nasipa.....	62
Fig. 42 Lluvias y % de cobertura vegetal del polígono Nasipa	62
Fig. 43 índice SAVI y Cobertura vegetal del polígono Nasipa.....	63
Fig. 44 Lluvias e índice SAVI del polígono KukuA	64
Fig. 45 Lluvias y % de cobertura vegetal del polígono KukuA	64
Fig. 46 índice SAVI y Cobertura vegetal del polígono KukuA	65
Fig. 47 Lluvias e índice SAVI del polígono Risa	66
Fig. 48 Lluvias y % de cobertura vegetal del polígono Risa.....	66
Fig. 49 índice SAVI y Cobertura vegetal del polígono Risa.....	67
Fig. 50 Lluvias e índice SAVI del polígono Pembamoto.....	68
Fig. 51 Lluvias y % de cobertura vegetal del polígono Pembamoto.....	68
Fig. 52 índice SAVI y Cobertura vegetal del polígono Pembamoto.....	69
Fig. 53 Lluvias e índice SAVI del polígono Enkii.....	70
Fig. 54 Lluvias y % de cobertura vegetal del polígono Enkii	71
Fig. 55 índice SAVI y Cobertura vegetal del polígono Enkii	71
Fig. 56 Lluvias e índice SAVI del polígono Baraka	72
Fig. 57 Lluvias y % de cobertura vegetal del polígono Baraka.....	72
Fig. 58 índice SAVI y Cobertura vegetal del polígono Baraka	73
Fig. 59 índice SAVI y Cobertura vegetal del polígono Engaruka.....	74
Fig. 60 Lluvias e índice SAVI del polígono EsilaleiSite1	75
Fig. 61 Lluvias y % de cobertura vegetal del polígono EsilaleiSite1	75
Fig. 62 índice SAVI y Cobertura vegetal del polígono EsilaleiSite1	76
Fig. 63 Lluvias e índice SAVI del polígono EsilaleiSite2	77
Fig. 64 Lluvias y % de cobertura vegetal del polígono EsilaleiSite2.....	77
Fig. 65 índice SAVI y Cobertura vegetal del polígono EsilaleiSite2	78
Fig. 66 Lluvias e índice SAVI del polígono Inkisanjani.....	79
Fig. 67 Lluvias y % de cobertura vegetal del polígono Inkisanjani	79

Fig. 68 índice SAVI y Cobertura vegetal del polígono Inkisanjani	80
Fig. 69 índice SAVI y Cobertura vegetal del polígono Meserani	81
Fig. 70 Lluvias e índice SAVI del polígono Selela.....	82
Fig. 71 Lluvias y % de cobertura vegetal del polígono Selela	82
Fig. 72 índice SAVI y Cobertura vegetal del polígono Selela.....	83
Fig. 73 Promedio precipitaciones de todos los polígonos	86
Fig. 74 Comparación de precipitaciones de polígonos reforestados en 2018	86
Fig. 75 Comparación de % de vegetación de polígonos reforestados en 2018.....	87
Fig. 76 Comparación del índice SAVI de polígonos reforestados en 2018	87
Fig. 77 Comparación de precipitaciones de polígonos reforestados en 2019	89
Fig. 78 Comparación de % de vegetación de polígonos reforestados en 2019.....	89
Fig. 79 Comparación del índice SAVI de polígonos reforestados en 2019	90

Lista de Tablas

TABLA 1 ÁREAS OBJETO DE ESTUDIO	30
TABLA 2 SATÉLITES CON COLECCIÓN DE REFLECTANCIA DE LA SUPERFICIE TERRESTRE	31
TABLA 3 IMÁGENES CON PORCENTAJE DE NUBOSIDAD SUPERIOR A 20%	36
TABLA 4 PUNTOS DE ENTRENAMIENTO Y PRUEBAS POR CATEGORÍA DE LA ITERACIÓN I.....	43
TABLA 5 EVALUACIÓN DE MODELOS ITERACIÓN I	47
TABLA 6 PUNTOS DE ENTRENAMIENTO Y PRUEBAS POR CATEGORÍA DE LA ITERACIÓN II.....	49
TABLA 7 EVALUACIÓN DE MODELOS ITERACIÓN II	52
TABLA 8 VALIDACIÓN CRUZADA	53
TABLA 9 RESULTADOS INKIITO	55
TABLA 10 RESULTADOS KUKUB.....	57
TABLA 11 RESULTADOS KUKUD.....	59
TABLA 12 RESULTADOS NASIPA.....	61
TABLA 13 RESULTADOS KUKUA.....	63
TABLA 14 RESULTADOS RISA.....	65
TABLA 15 RESULTADOS PEMBAMOTO	67
TABLA 16 RESULTADOS ENKII.....	69
TABLA 17 RESULTADOS BARAKA.....	71
TABLA 18 RESULTADOS ENGARUKA	73
TABLA 19 RESULTADOS ESILALEISITE1	74
TABLA 20 RESULTADOS ESILALEISITE2.....	76
TABLA 21 RESULTADOS INKISANJANI	78
TABLA 22 RESULTADOS MESERANI	80
TABLA 23 RESULTADOS SELELA.....	81
TABLA 24 DIFERENCIA DE PROMEDIOS DE RESULTADOS DE REFORESTACIÓN CON SEMICÍRCULOS SIN SIEMBRA	85
TABLA 25 DIFERENCIA DE PROMEDIOS DE RESULTADOS DE REFORESTACIÓN CON SEMICÍRCULOS CON SIEMBRA	85
TABLA 26 DIFERENCIA DE PROMEDIOS DE RESULTADOS DE POLÍGONOS REFORESTADOS EN 2018.....	88

TABLA 27 DIFERENCIA DE PROMEDIOS DE RESULTADOS DE POLÍGONOS REFORESTADOS EN 2019.....90

1. Introducción

El calentamiento global y el cambio climático son un hecho que se ha demostrado por diferentes investigaciones [1], [2] y todos los seres humanos han aportado para que progresivamente aumente dicho fenómeno [3], [4], [5]. Si bien, el calentamiento global es causado por las concentraciones atmosféricas de gases de invernadero, también afecta de forma indirecta a la capa vegetal, pues en algunos lugares del mundo se ha evidenciado que el suelo se está degradando y perdiendo total o parcialmente su capacidad productiva a causa de la erosión [2].

Tanzania y Kenia son unos de los países africanos que se ven afectados por las sequías [6], [7]. El suelo en esta región es árido y compacto, por lo que el agua de lluvia corre por la superficie del suelo, pero no es absorbida por este, como consecuencia las cosechas y vegetaciones no tiene el suministro adecuado para poder vivir, lo que conlleva a la poca producción de alimentos y las altas temperaturas que afectan a la población [8].

Afortunadamente algunas organizaciones están trabajando para contrarrestar el calentamiento global, como es el caso de Justdiggit¹, una ONG que ha trabajado en los últimos años en reverdecer y restaurar algunas zonas degradadas de Kenia y Tanzania con el fin de que dicha vegetación ayude a disminuir la temperatura del país y del planeta [8].

Justdiggit usa técnicas basadas en la naturaleza para recuperar la naturaleza y el equilibrio del planeta. Una de ellas es cavar semi círculos en la tierra con y sin siembra de semillas de pasto. Dichos semi círculos cumplen la función de dique natural para retener el agua de lluvia, permitiendo de esta forma que el agua tenga tiempo de penetrar en la tierra y generar un ambiente propicio para que crezcan pastizales. Las raíces de dichos pastizales a su vez servirán de diques para retener más agua y ayudarán a des compactar el suelo, generando un ambiente propicio para que insectos puedan vivir ahí, los cuales polinizarán las plantas y darán paso a los ciclos de la naturaleza, devolviendo vida al suelo. Estas actividades de restauración impactan positivamente la biodiversidad, la seguridad hídrica y alimentaria de los pobladores [8].

Este trabajo está centrado en la medición de los resultados obtenidos con las actividades de Justdiggit en Kenia y Tanzania, cuyos inicios tienen lugar entre los años 2016 y 2021, donde se evalúan los cambios en la vegetación acorde a las técnicas de reforestación aplicadas, con el fin de identificar cual es más eficiente. Para lo cual se usan técnicas de teledetección y modelos de aprendizaje automático que permitan identificar y clasificar las áreas con vegetación desde el inicio de dichos proyectos hasta el 2022. Se tiene como insumo imágenes satelitales

¹ <https://justdiggit.org/>

multiespectrales capturadas por el Satélite LandSat 8 [9], en las cuales se identifican como ha cambiado el terreno de las zonas de estudio a partir de la reflectancia, la cual nos permite identificar con cada haz de luz los cambios en el área de estudio.

1.1. Planteamiento del problema

La deforestación ocasiona la pérdida del hábitat de miles de especies y afecta el ciclo del agua, lo que a corto y largo plazo pueden llegar a impactar en los recursos naturales con los que se realizan actividades de agrícolas. Por tanto, la deforestación puede llegar a influir en la economía de las regiones donde se presente, además de contribuir al calentamiento global, pues sin árboles el dióxido de carbono se queda en la atmósfera y produce el efecto invernadero [2].

La importancia de la preservación de los suelos alrededor del mundo repercute en el cambio climático del planeta, pues la vegetación actúa como una especie de aire acondicionado en el suelo, minimiza las temperaturas de la zona, disminuye la concentración de dióxido de carbono en el ambiente y ayuda a la retención del agua, lo que facilita las actividades de agricultura y pastoreo [2]. Algunos de los países más afectados por el calentamiento global son Kenia y Tanzania, pues tiene elevadas temperaturas y su suelo árido no permite que el agua penetre para alimentar a la poca vegetación que hay en el sector. Afortunadamente Justdiggit ha realizado actividades de restauración del suelo con técnicas naturales [8].

En dichas zonas de estudio se han cavado medias lunas o semicírculos que funcionan como diques naturales para la retención del agua de lluvia y en algunos casos de forma adicional se han esparcido semillas de pasto, para finales del 2022 se había cavado en total 314.038 semicírculos en el suelo de las zonas objeto de estudio de este proyecto [8]. Si bien ambas técnicas dependen de la cantidad de agua lluvia que se presente, no se ha determinado con cual de estas dos técnicas se obtienen mejores resultados.

1.2. Objetivos

1.2.1. Objetivo general

Cuantificar el cambio en la vegetación causado por las actividades de la ONG Justdiggit en Kenia y Tanzania usando imágenes satelitales, con la aplicación de técnicas de aprendizaje automático supervisado y teledetección, desde el inicio de los proyectos en dichas áreas hasta el 2022 y determinar cuál de las técnicas es más efectiva.

1.2.2. Objetivos específicos

- Recopilar conjunto de imágenes satelitales multiespectrales con nubosidad inferior a 20% en cada zona de estudio cuya reforestación tuvo lugar entre los años 2016 y 2021, desde el inicio de cada proyecto hasta el año 2022.
- Generar índices espectrales y etiquetar manualmente muestras de los tipos de áreas a nivel de píxeles, que se usarán para entrenar y evaluar los modelos de aprendizaje automático.
- Entrenar y evaluar modelos de clasificación supervisados (Árboles de decisión, Random Forest y SVM) para identificar el modelo con mejor rendimiento para clasificar la cobertura de suelos.
- Aplicar el modelo con mejor rendimiento para clasificar los suelos con vegetación en las áreas de estudio.
- Analizar la información obtenida de los modelos de clasificación e índices.
- Identificar cuál de las técnicas de reforestación aplicadas en las áreas de estudio es más efectiva.

1.3. Metodología

Este el desarrollo de este trabajo se elige la metodología CRISP-DM, la cual tiene seis etapas [10]. En este proyecto se agrupan en tres fases principales, las cuales se describen a continuación:

- Entendimiento del negocio y preparación de los datos: en esta fase se definen las áreas objeto de estudio, las fuentes de imágenes satelitales y contemplar los criterios mínimos que deben cumplir. Esto se presenta en el capítulo cuatro.
- Entrenamiento de modelos y evaluación: en esta fase se define el polígono de entrenamiento de los modelos, además de generar los índices necesarios para capturar los puntos de entrenamiento. Con las imágenes seleccionadas en la fase anterior se generan modelos de clasificación supervisados (Árboles de decisión, Random Forest y SVM) y se evalúan dichos modelos para identificar el que clasifique de la mejor manera las áreas vegetación/no vegetación. Esto se presenta en el capítulo cinco.
- Aplicación del modelo seleccionado y análisis de resultados: En esta fase se usa el modelo seleccionado para clasificar los píxeles de las imágenes y de estar forma

determinar anualmente el área con vegetación de cada una de las áreas objeto de estudio, lo cual se analiza junto con la información obtenida con los índices para determinar cuál de los dos métodos de reforestación es más efectivo. Esto se presenta en el capítulo seis.

En la figura 1 se muestra un diagrama que refleja el flujo de trabajo con el que se desarrolla este proyecto y las actividades a realizar para cumplir con los objetivos, las cuales están clasificadas en las fases definidas.

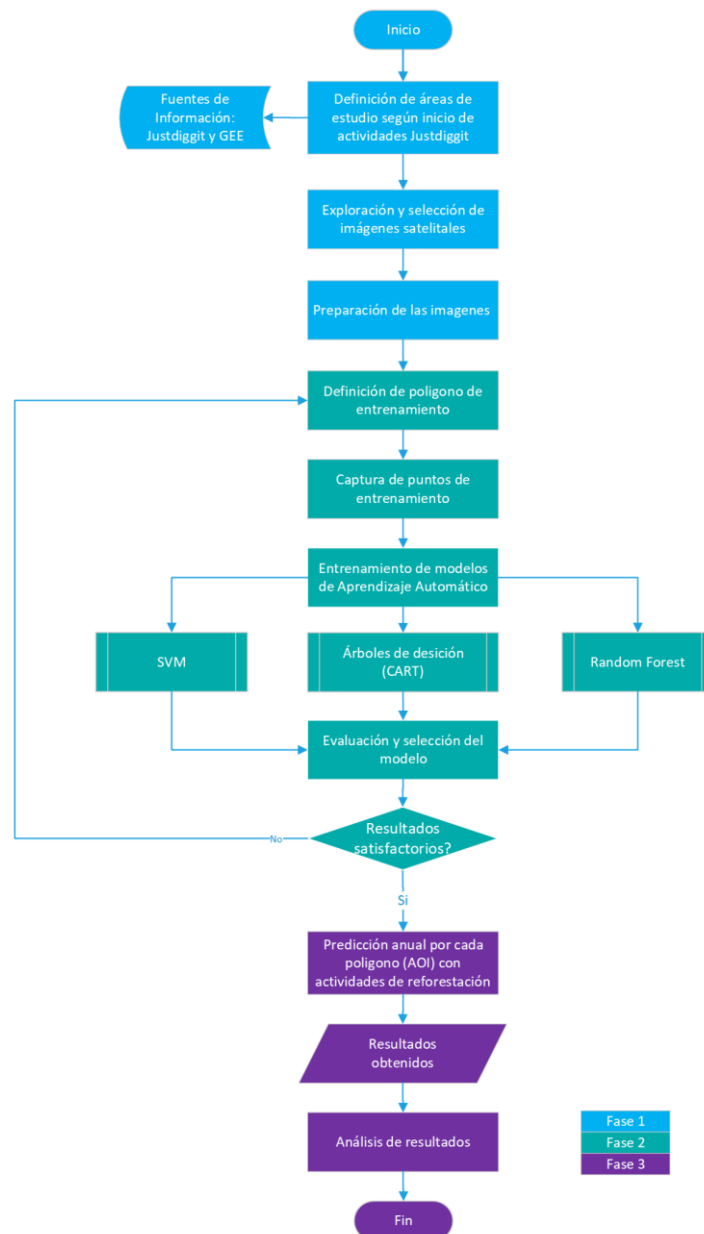


Fig. 1 Flujo de trabajo

2. Marco Teórico

En esta sección se mencionan conceptos, definiciones y técnicas que se usan en el desarrollo de este trabajo de grado.

2.1. Teledetección

Acorde a la definición de Emilio Chuvieco, la teledetección es “aquella técnica que nos permite obtener información a distancia de los objetos situados sobre la superficie terrestre” [11]. Partiendo de esto, se puede aplicar teledetección a todas las imágenes de la tierra capturadas a distancia por sensores, incluyendo imágenes capturadas por drones y satélites. Estas últimas se usan en este trabajo, específicamente las imágenes satelitales de las áreas objeto de estudio definidas.

Existen dos tipos de sensores [11]:

- Activos: tienen su propia fuente de energía, por ejemplo, radares y sonares [11].
- Pasivos: sensores cuya fuente de energía es obtenida del sol o de la atmósfera de la tierra, dentro de esta clasificación se encuentran los satélites LandSat y Sentinel [11].

Estos sensores pasivos pueden detectar diferentes longitudes de onda en el espectro electromagnético visible e infrarrojo y posteriormente el satélite transmite la información para su interpretación y análisis [11]. En la Figura 2 se representa el funcionamiento de los sensores pasivos.

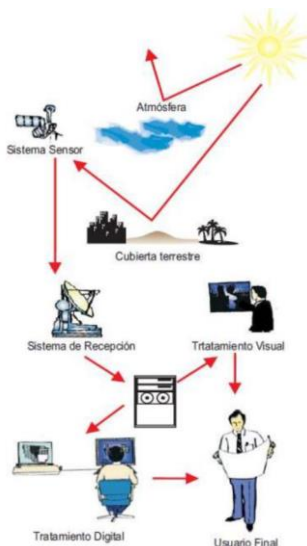


Fig. 2 Componentes de un sistema de teledetección

Nota: Imagen tomada del libro *Fundamentos de la Teledetección Espacial* [11].

Dicho espectro electromagnético es el rango de los diferentes tipos de radiación electromagnética u ondas, como se muestra en la Figura 3, las cuales se clasifican según la longitud y frecuencia de las ondas [11].

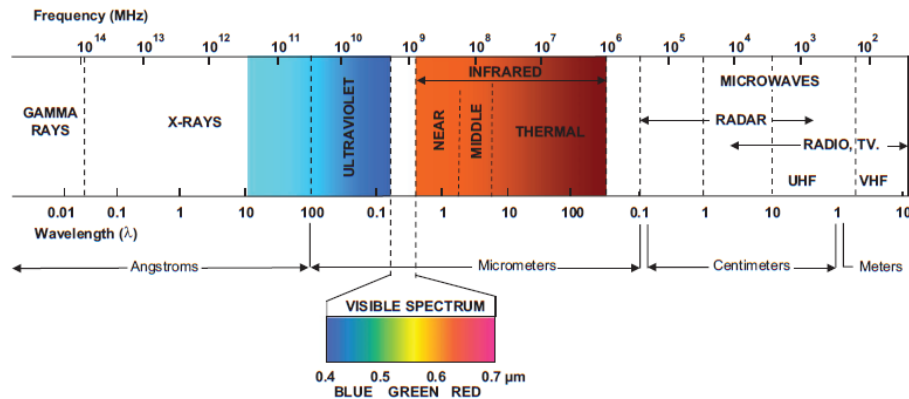


Fig. 3 Espectro electromagnético

Nota: Imagen tomada del libro *Fundamentals of satellite remote sensing* [12].

Acorde a la interacción de la superficie de la tierra con la radiación se obtendrán diferentes lecturas de longitudes de onda en los sensores [13], por ejemplo, la vegetación sana refleja las longitudes de onda verdes e infrarroja, pero absorbe las rojas y azules, a estas curvas de reflectancia espectral se llama firma espectral. Estas firmas varían acorde al tipo de terreno, es decir, los cuerpos de agua y suelos descubiertos tienen firmas diferentes a las de vegetación [14], como se muestra en la Figura IV.

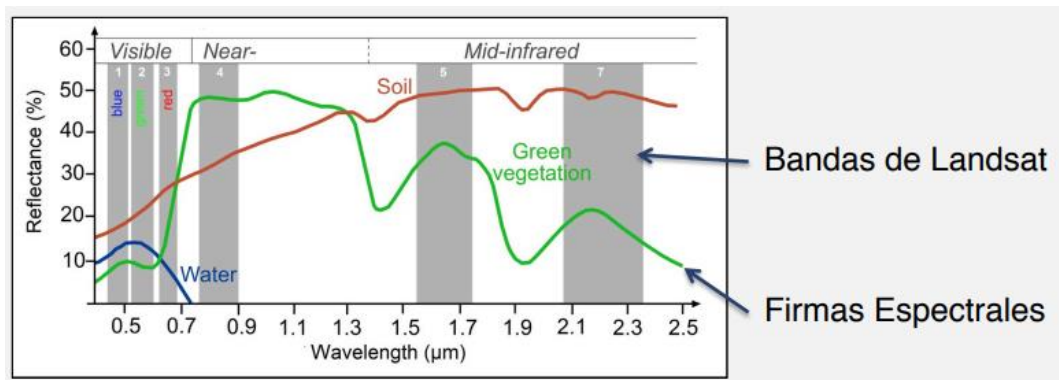


Fig. 4 Curvas de reflectancia espectral

Nota: Imagen tomada de presentación de *Fundamentos de la Teledetección para Aplicaciones de Salud Pública de la NASA* [15].

Cada sensor maneja bandas o rangos de longitud de ondas específicos, por ejemplo, los satélites LandSat y Sentinel tienen sensores diferentes y tiene variación en sus bandas, lo cual se debe

tener en cuenta en los análisis que se realicen a las imágenes capturadas en estos satélites [14]. En la Figura IV se muestran en color gris las bandas que manejan los satélites LandSat [15].

Las imágenes capturadas por estos sensores son multispectrales, ya que contiene la reflectancia de la superficie en diferentes bandas, es decir, las imágenes tienen valores numéricos a nivel de píxeles por cada banda, al combinarse dichas bandas nos pueden brindar información valiosa [14].

En LandSat 8 la combinación de bandas cuya composición se aproxima al color real o natural son 4, 3 y 2 (Red, Green, Blue RGB). Sin embargo, la combinación de las bandas 5, 4 y 3 resalta la vegetación sana en color rojo [16].

2.2. Índices espectrales de vegetación

Los índices espectrales de vegetación ayudan a resaltar los fenómenos que se estudian en este proyecto, como por ejemplo los cambios en la vegetación, para lo cual se combinan bandas específicas [17]. En este trabajo se usan los índices que se relacionan en los siguientes ítems.

2.2.1. NDVI

El índice diferencial de vegetación normalizado NDVI (Normalized Difference Vegetation Index), el cual nos permite identificar la densidad de masas vegetales o su estado de salud [18]. La fórmula para calcular este índice es la siguiente y usa las bandas infrarrojo cercano (NIR) y roja (Red) [13]:

$$NDVI = (NIR - Red)/(NIR + Red)$$

2.2.2. SAVI

El índice de vegetación ajustado al suelo SAVI (Soil Adjusted Vegetation Index) también permite identificar masas vegetales al igual que el NDVI, con la diferencia de que minimiza la influencia del brillo del suelo desnudo, para lo cual usa un factor de corrección L , por lo cual es ideal para identificar los cambios en la vegetación de las áreas objeto de estudio [19]. Su fórmula es la siguiente y usa las bandas infrarrojo cercano (NIR) y roja (Red) [19]:

$$SAVI = \frac{NIR - Red}{NIR + Red + L} * (1 + L)$$

2.2.3. NDWI

El índice normalizado diferencial de agua NDWI (Normalized Difference Water Index) es idóneo para identificar cuerpos de agua [20]. Su fórmula es la siguiente y usa las bandas infrarrojo cercano (NIR) y verde (Green) [21]:

$$NDWI = (Green - NIR)/(Green + NIR)$$

2.3. Calentamiento Global, Erosión y Reforestación

El *calentamiento global* o *cambio climático* es un fenómeno del que se ha hablado mucho y nos afecta a todos, este puede denominarse como el aumento medio de temperatura natural del planeta, el cual es causando por la acumulación de gases de invernadero en la atmósfera, cuyas consecuencias principales son aumento en el desplazamiento y extinción de especies animales y vegetales, mayor aridez y desertificación, entre otros [2].

Precisamente esto es lo que está afectando a Kenia y Tanzania, la *degradación del suelo*, esta puede definirse como la pérdida total o parcial de la calidad del suelo, afectando directamente su productividad, esta puede ser causada por erosión, desertificación entre otros [2].

Es por esto por lo que algunas organizaciones a nivel mundial le han apuntado a la recuperación de suelos para restaurar algunas zonas. La *reforestación* es definida por la real academia de la lengua española como “repoplar un terreno con plantas forestales” [22], sin embargo, a nivel agroecológico se le da una connotación más funcional, dando un objetivo asociado al territorio y a la producción [23], como es el caso de la restauración de las áreas objeto de estudio de este proyecto. Por otra parte, la *deforestación* es la disminución de la superficie boscosa de un área, lo que conlleva a la pérdida de flora y fauna [24]. Para cuantificar los cambios en la vegetación se aplica la clasificación de coberturas vegetación/no vegetación, se usan modelos de clasificación supervisada.

2.4. Aprendizaje automático

El aprendizaje automático es una rama de la inteligencia artificial. Russel Stuart lo define de la siguiente forma “En el aprendizaje de máquinas un computador observa datos, construye un modelo basado en esos datos y utiliza ese modelo a la vez como una hipótesis acerca del mundo y una pieza de software para resolver problemas” [25].

La clasificación de suelos se puede realizar de forma manual, sin embargo, es más precisa y menos dispendiosa si se realiza con técnicas de aprendizaje automático. Los principales tipos de aprendizaje automático son los siguientes:

- **Aprendizaje supervisado:** Este es tipo de aprendizaje automático que entrena modelos a partir de datos etiquetados, aprende identificando patrones en los atributos de dichos datos etiquetados [26]. En este caso los datos corresponden a los píxeles de las imágenes satelitales y deben estar categorizados según el tipo de cobertura de suelo (vegetación, suelo, agricultura, urbanización y cuerpos de agua) siendo esta la variable objetivo de los modelos, es decir, el modelo clasificará nuevos puntos en dichas categorías.
- **Aprendizaje no supervisado:** Este tipo de aprendizaje automático también encuentra patrones en los datos suministrados, con la diferencia de que estos no estarán etiquetados [26]. Además de esto agrupa la información acorde a las similitudes de los datos [26].

Partiendo de esta teoría, estos los algoritmos de aprendizaje no supervisado no son los más adecuados para suplir la necesidad de este proyecto, que es clasificar los tipos de cobertura de suelo. Ya que con aprendizaje supervisado se sabrá qué resultados esperar mientras que con el aprendizaje no supervisado se espera descubrir algo nuevo en los datos. Por tal razón en este proyecto solo se exploran técnicas de aprendizaje supervisado.

En este proyecto se usan los siguientes modelos de clasificación:

- **Arboles de decisión:** esta técnica usa un algoritmo que clasifica los datos en ramas acorde al valor del atributo de cada nodo y acorde a la cantidad de niveles definidos [27].
- **Random Forest:** este algoritmo consiste en un conjunto de árboles de decisión aleatorios, cada uno generado con un subconjunto de datos aleatorio [27].
- **Regresión logística:** este algoritmo separa de forma óptima los puntos de cada clase, que para este ejercicio corresponde al tipo de área (vegetación/no vegetación) y a partir de esta separación realiza la clasificación de elementos [28].

2.4.1. Evaluación

Después de contar con los modelos entrenados, estos se evalúan y afinan con el fin de encontrar el modelo que tiene más exactitud identificando áreas con vegetación y este será el que se aplique a todas las imágenes, con el fin de identificar en cada año cuantos metros cuadrados de vegetación había en las áreas de estudio y de esta forma determinar la eficiencia de los métodos de reforestación.

En este caso se usará la Matriz de confusión como técnica de evaluación de los modelos, en donde el objetivo es medir las predicciones correctas maximizando el número de verdaderos positivos y negativos [29].

También se usará el método de validación cruzada, el cual consiste en dividir los datos en k segmentos, de los cuales uno se usará para evaluar la métrica especificada en el modelo entrenado con los segmentos restantes. El mismo número de segmentos corresponde al número de iteraciones a realizar, en donde el segmento de evaluación y entrenamiento cambia en cada iteración. Este método permite evaluar la generalización de los modelos y comparar el desempeño de cada uno [30].

3. Estado del arte

Existen muchos trabajos de investigación en los cuales se han tocados temas relacionados con el proyecto descrito en este documento, como la deforestación y las técnicas de teledetección que se han aplicado con diferentes fines, a continuación, se relacionan algunos trabajos a nivel nacional, internacional y asociados con JustdiggIt, en orden cronológico para los tres casos.

3.1. Estudios realizados en Colombia

Una de las investigaciones nacionales fue la realizada por Jorge Polanco en el año 2012 donde aplicó técnicas de teledetección y aprendizaje no supervisado para identificar la cobertura de suelos del páramo de Belmira en Antioquia, utilizando imágenes satelitales de los años 1996, 2002 y 2003 [31]. Aunque este trabajo tiene factores en común con el proyecto descrito en este documento, están enfocados en áreas de estudio diferentes y en periodos diferentes.

Análogamente, Juan Rojas y Carmenza Castiblanco en el año 2018 indagaron sobre la variación de coberturas del municipio de San Carlos de Guaroa – Colombia, en los años 1985, 2000 y 2016, causada principalmente por el cultivo de palma de aceite, para lo cual se cuantificaron en dichos años la cobertura de bosques de galería, cultivo de palma de aceite y áreas de uso agropecuario, en donde se identificó una pérdida del 19.38% de bosques con una reducción media anual de 3.22% [32]. En contraste con el trabajo descrito en este documento, se encuentran diferencias significativas pues en este se busca identificar los cambios en la vegetación en un rango de tiempo más corto, pero con una periodicidad anual y en áreas de estudio diferentes.

Por otra parte, Jhon Guzmán y Andrés Flórez en el año 2019 investigaron la deforestación en el periodo de tiempo comprendido entre los años 1986 y 2013, causada por actividades mineras en el municipio de Montelíbano en Córdoba-Colombia, para lo cual también usaron técnicas de teledetección y correcciones radiométricas a imágenes satelitales, en donde se pudo evidenciar que pérdida de la vegetación nativa y fragmentación de hábitats [33]. Tanto en dicho proyecto como en el descrito en este documento, se busca identificar la deforestación en zonas específicas y diferentes, en las cuales las causas de esta tienen algunos orígenes distintos pero el factor común son actividades realizadas por la población.

Analógicamente, John Camacho analizó en el año 2020 las causas de la deforestación en el departamento del Guaviare con regresión logística del año 2005 al 2020 e hizo una proyección de escenarios para el 2030 [34]. Este difiere del presente proyecto ya que este no tiene como objetivo identificar las causas de la deforestación y las áreas de estudio son diferentes.

Por otra parte, Claudia Cardona y sus colegas en el año 2020 aplicaron técnicas de teledetección y cálculo de índices para estimar la capacidad de captura de dióxido de carbono de del Jardín Botánico Cedro Rosado en Quindío – Colombia [35], con lo que es notorio que el objetivo de dicho proyecto difiere de presente proyecto.

Jully Forero y Manuel Polanco en el año 2021 investigaron sobre la deforestación presentada en el municipio La Macarena antes y después de la firma de los acuerdos de paz en el año 2016, usando imágenes satelitales y tomando como base el protocolo de cuantificación de la tasa de deforestación definido por el IDEAM, en donde se encontró un aumento significativo de deforestación posterior a la firma de los acuerdos [36]. Esto difiere del trabajo que se describe en este documento, pues el área objeto de análisis es diferente y no está asociado a los acuerdos de paz.

Daissy Herrera y Edier Aristizábal en el año 2022 también aplicaron técnicas de aprendizaje automático para predecir eventos de sequía en el departamento del Magdalena para permitir a las entidades gubernamentales tomar medidas [37]. Este artículo y el proyecto actual en este documento usan técnicas similares pero el objetivo está enfocado a la sequía y no a la deforestación, además las áreas de estudio también difieren.

Por otra parte, Edier Avila y sus colegas en el año 2022 también aplico técnicas de teledetección orientadas a identificar plantas invasoras que afectan negativamente la biodiversidad y el desarrollo de los ecosistemas en Sibate-Cundinamarca [38]. Es notorio que el objetivo de dichos proyectos difiere de presente proyecto pues este está orientado a la detección de ciertas especies de plantas mas no a identificar los cambios en la cobertura de este.

De manera similar, Carlos Mejía y sus colegas en el año 2023 aplicaron técnicas de teledetección y calcularon índices para estimar la biomasa aérea y carbono en bosques secos tropicales del Centro universitario regional norte de la Universidad del Tolima [39]. Así mismo, Raúl Diaz y sus colegas en el año 2024 aplicaron técnicas de teledetección con imágenes obtenidas remotamente por un sensor empírico y otro semi empírico, con lo cual se calcularon índices y se generaron dos modelos para estimar la cantidad de materia seca acumulada por las plantas en la biomasa aérea, lo cual es información relevante para la toma decisiones en el manejo de sistemas de producción bovino en Villavicencio – Colombia [40]. Es notorio que el objetivo de dichos proyectos difiere de presente proyecto pues está orientado a la detección de biomasa aérea y de carbono mas no a identificar los cambios en la cobertura de este.

3.2. Estudios realizados a nivel internacional

A nivel internacional, Daniel Bezanilla, Joaquin Sosa y Abraham de Alba en el año 2008 publicaron un artículo sobre la degradación y recuperación de bosques en la Sierra Fría en Aguascalientes – México, teniendo en cuenta la extracción de leña, incendios forestales y pastoreo intensivo, para lo cual se compararon imágenes aéreas de 1956, 1970 y 1993, a partir de las cuales se tele detectaron áreas boscosas, en donde se reconocieron cambios en la superficie y grados de fragmentación, aunque la pérdida de boques fue nivelada con la recuperación [41], con lo que contrasta con el trabajo descrito en este documento, pues en este se busca identificar los cambios en la vegetación en un rango de tiempo más corto, pero con una periodicidad anual y en áreas de estudio diferentes.

Así mismo, Dolors Armenteras y Nelly Rodríguez en el año 2014 publicaron un artículo como resultado del análisis las dinámicas y causas de la deforestación en bosque de Latinoamérica, se consideraron 283 artículos con información histórica de este tema en los diferentes países del área objeto de estudio [42]. Esto contrasta con el trabajo descrito en este documento, debido a que el área de objeto de análisis es mucho más amplia y no se realiza el cálculo anual de los cambios en la vegetación, esto se toma de los artículos encontrados y con esto se haya la tasa de deforestación por países.

De modo similar, Gilberto Medina y Abel Mejía publicaron un artículo en el año 2014 en el cual analizaron la pérdida del volumen glaciar de los nevados de la cordillera Parón en los Andes del Perú usando imágenes capturadas por los satélites LandSat, para lo cual usaron el índice de nieve (NDSI) en el periodo de 1987 al 2011, con el cual se pudo identificar una disminución glacial promedio del 18% [43]. En comparación con el trabajo descrito en este documento, se puede identificar que tienen propósitos diferentes, pero tienen en común el uso de índices generados a partir de imágenes satelitales multiespectrales para obtener información que contribuya con el objetivo respectivo.

Pérez-Planells y sus colegas también publicaron un artículo en el año 2015 en el que exploraron técnicas de aprendizaje automático aplicadas en la teledetección, específicamente métodos lineales y no lineales, con el fin de identificar por medio de la validación cruzada cuales tenían resultados más confiables [44]. Si bien, tiene en común con este proyecto la aplicación de técnicas de aprendizaje automático, están enfocados a objetivos diferentes, pues en el proyecto descrito en este documento se enfoca en identificar los cambios en la vegetación de las áreas objeto de estudio y no la estabilidad de los resultados de los algoritmos.

Así mismo, Kozhikkodan Veetil y sus colegas también publicaron un artículo en el año 2016 en el que estuvieron analizando el comportamiento de los glaciares de los Andes Tropicales, a partir de técnicas de teledetección, con lo cual se pudo identificar que a nivel general dichos glaciares están retrocediendo aceleradamente desde 1970. Sin embargo, se detectó que los glaciares de Bolivia y Perú retroceden más rápido que los demás [45]. En comparación con el trabajo descrito en este documento, se puede identificar que tienen propósitos diferentes, pero tienen en común el uso de índices generados a partir de imágenes satelitales multiespectrales para obtener información que contribuya con el objetivo respectivo.

Por otro lado, el mexicano Baltazar Sánchez también publicó un artículo en el año 2018 de su investigación, en la cual aplicó técnicas de teledetección para identificar y monitorear a través de imágenes satelitales patrones de biodiversidad y la influencia de las actividades humanas en sus cambios a nivel local y mundial, en donde su objetivo es resaltar las ventajas de la teledetección como fuente de información [46], lo cual difiere del objetivo del presente documento, puesto que este último busca identificar los cambios en la vegetación causados por las actividades de reforestación realizadas por Justdiggit.

De modo similar, Alexander Ariza y sus colegas en el año 2019 publicaron un artículo resultado de la investigación en la que implementaron modelos de máxima verosimilitud y modelos de regresión con el fin de validar que índice permitía resaltar más los efectos de quemaduras en la vegetación causada por incendios en el centro de España [47]. Al igual que en el presente proyecto se aplican técnicas de aprendizaje automático en la teledetección de suelos, sin embargo, el objetivo es validar que índice resalta más los efectos de quemaduras en suelos mas no identificar los cambios en la vegetación causados por las actividades de reforestación.

Malvina Serra y sus colegas también publicaron un artículo en el año 2019 sobre su investigación en la que aplicaron técnicas de teledetección y cálculo de índices orientados a la caracterización de la geomorfología y pulsos de lava del volcán Aguas Calientes y sus alrededores en el departamento de Tinogasta en Argentina [48]. Al igual que en el presente proyecto se aplican técnicas de teledetección, sin embargo, el objetivo no es identificar los cambios en la vegetación causados por las actividades de reforestación, como si lo es el proyecto descrito en este documento.

En contraste, Ruben Valcarce en su tesis publicada en el año 2020 implementó técnicas de aprendizaje automático en teledetección y seguimiento de cultivos sin importar las condiciones climáticas, a partir de imágenes obtenidas por sensores de radar de apertura sintética (SAR) [49]. Si bien, tiene en común con este proyecto la aplicación de técnicas de aprendizaje automático,

están enfocados a objetivos diferentes, pues esta focalizado a identificar los cambios en las áreas objeto de estudio que son terrenos yermos mas no a cultivos.

Shujun Liang y sus colegas publicaron un artículo en el año 2020 en el que presentan su investigación en la que aplicaron técnicas de aprendizaje profundo con el algoritmo de máxima verosimilitud a los suelos en China, con el fin de compararlos con los métodos de clasificación de imágenes satelitales e identificar cual es más optimo [50]. Si bien, se aplicaron técnicas de aprendizaje automático en la teledetección de suelos, el objetivo es validar que método es más optimo por lo cual difiere del objetivo del presente proyecto.

De forma semejante, M. Ramírez y sus colegas en el año 2020 publicaron un artículo en donde obtuvieron los cambios de cobertura terrestre con el algoritmo de Random Forest aplicado a imágenes satelitales fusionadas con imágenes capturadas por drones [51]. Analógicamente al proyecto propuesto, el objetivo es identificar cambios en la cobertura del suelo, sin embargo, difiere del presente proyecto ya que usan un tipo de imágenes diferente en zonas de estudio distintas.

En cambio, Ana Alvarado y sus colegas en el año 2021 publicaron un artículo resultado de su estudio del cambio de la extensión agrícola en el distrito Yarada de los Palos en Perú en el periodo del 2000 al 2020, usando imágenes satelitales multiespectrales y técnicas de teledetección, en donde se encontró un aumento del 265.84% de dichas áreas [52]. Por otra parte, Jesús Mejía y sus colegas evaluaron la erosión hídrica en un área de Tacna en Perú, para lo cual entrenaron modelos y calcularon índices [53]. Si bien, las finalidades de estos proyectos son diferentes a la del proyecto que se describe en este documento, también se usan técnicas de teledetección para cumplir con los objetivos trazados.

Por otra parte, Fatima Rahal y sus colegas en el año 2022 publicaron un artículo resultado del mapeo de suelos arcillosos en el noroeste de Argelia con imágenes satelitales ASTER con el fin de establecer la mineralogía del suelo [54], con lo que es notorio que el objetivo de dicho proyecto difiere de presente proyecto pues está orientado a la composición del suelo mas no a identificar los cambios en la cobertura de este.

Así mismo, Andrés Estrada y sus colegas en el año 2022 publicaron un artículo derivado de su trabajo en el que aplicaron técnicas de aprendizaje automático a imágenes capturadas por vehículos no tripulados (drones) para identificar la biomasa de los pastizales en una comunidad vegetal altoandina en Perú [55]. Aunque este trabajo tiene factores en común con el proyecto descrito en este documento, difieren en que se busca identificar las áreas boscosas con imágenes

satelitales y no con imágenes capturadas con drones, además de estar enfocados en áreas de estudio diferentes.

Ali Ahman y sus colegas, abarcan en su artículo publicado en el año 2022 la importancia de la inteligencia artificial en la enseñanza de la fisiología vegetal, ya que esta complementa la teledetección y permite detectar enfermedades, predecir el rendimiento y generar simulaciones entre otros [56]. Evidentemente los objetivos del artículo y el presente proyecto son diferentes, sin embargo, tienen en común resaltar las bondades de la inteligencia artificial aplicada en diferentes ámbitos.

Por otra parte, Yoansy García y sus colegas en el año 2023 publicaron un artículo de su investigación que estaba enfocada en la identificación de los bosques ecuatorianos, los cuales han tenido una disminución de bosques de un 15% en los años 1990 y 2020, causados por la tala generados por el aumento de la superficie agrícola en las cuencas hidrográficas el Zapotal, para lo cual utilizaron técnicas de teledetección e índices de vegetación [57]. Aunque este trabajo tiene factores en común con el proyecto descrito en este documento, están enfocados en áreas de estudio diferentes y en periodos diferentes.

3.3. Estudios relacionados con Justdiggit

Las actividades de Justdiggit también han sido estudiadas por Susan Steele y Thom Bogaard, los cuales publicaron en el año 2018 su tesis sobre el aumento de la vegetación en el área de Kuku en Kenia a través de satélites remotos de VanderSat [58]. De forma similar, Lorenzo Villani y sus colegas publicaron un artículo en el año 2020 en el que identificaron los cambios de la vegetación en Dodoma Tanzania, en donde Justdiggit también ha realizado actividades de reforestación [59]. De forma similar, Mendy van der Vliet y sus colegas publicaron un artículo en el año 2023 sobre su investigación donde aplicaron técnicas de teledetección para identificar el agua retenida en el suelo y los cambios en la temperatura de dos áreas de Tanzania en las cuales Justdiggit ha realizado actividades [60]. A diferencia del proyecto descrito en este documento, ya que en este se estudian los cambios en la vegetación en áreas reforestadas en Kenia y Tanzania cuyas actividades iniciaron antes del 2021.

Para concluir este capítulo, se puede identificar que existen varias investigaciones relacionadas con el proyecto propuesto, sin embargo, todas varían en uno o varios aspectos.

4. Entendimiento del negocio y preparación de los datos

En esta fase se realizan los pasos que nos permiten identificar las áreas objeto de estudio, las fuentes de imágenes satelitales y contemplar los criterios mínimos que deben cumplir, además de generar los índices necesarios para capturar los puntos de entrenamiento.

4.1. Fuente de datos

Este proyecto tiene dos fuentes principales de información:

- Justdiggit: Esta ONG amablemente proveyó a través de correo electrónico el tipo de reforestación, año de inicio de actividades y cantidad de semicírculos de cada área, además de esto envió un archivo con formato KML (Keyhole Markup Language) con la ubicación y área de los polígonos en los cuales Justdiggit ha realizado actividades de reforestación en Tanzania y Kenia². Estas áreas no tenían información de la vegetación, puesto que solo corresponden a los polígonos de dichas áreas y su ubicación.
- Google Earth Engine: Esta herramienta tiene un amplio catálogo de imágenes satelitales y datos geoespaciales, además de permitir el procesamiento de datos en línea y en sus servidores, lo cual es gratuito para fines académicos y de investigación [61]. Todas estas cualidades de Google Earth Engine (GEE) optimizan tiempo, espacio y procesamiento en gran medida.

Las imágenes satelitales obtenidas de GEE son uno de los insumos principales pues a partir de estas se calculan índices que permiten resaltar características de la cobertura del suelo [13] y de estar forma ayudan a la captura de puntos de entrenamiento, los cuales conforman el conjunto de datos de entrenamiento de los modelos de aprendizaje automático supervisado. Así mismo, estas imágenes son usadas para la aplicación del modelo de aprendizaje automático supervisado seleccionado.

4.2. Definición de áreas de estudio

Actualmente Justdiggit tiene actividades en Tanzania, Kenia, Uganda y Ethiopia las cuales han tenido inicio entre los años 2016 y 2022. Sin embargo, este proyecto se centra solo en las

² T. Zaan [Comunicación personal]. 22 de septiembre del 2023.

actividades realizadas en Tanzania y Kenia y cuyo inicio de actividades tuvo lugar entre los años 2016 y 2021[8]. Se excluyen las áreas cuyo inicio de actividades se dió en el año 2022 debido a que ha pasado poco tiempo y no sería posible evidenciar los cambios en la vegetación en un periodo tan corto.

En Kenia y Tanzania, Justdiggít ha realizado actividades de reforestación con dos métodos en diferentes áreas, las cuales se describen a continuación:

- Semicírculos de suelo sin siembra: este método consiste en la cavar en el suelo en forma de semicírculos contra la pendiente de terreno, como se observa en la Figura V, de forma que estos hoyos cumplan la función de diques naturales y retengan el agua lluvia, para que esta tenga el tiempo suficiente para penetrar en el suelo, que de esta forma la vegetación crezca de forma natural [8].
- Semicírculos de suelo con siembra: este método al igual que en el método anterior consiste en cavar semicírculos en el suelo contra la pendiente y de forma adicional esparcir semillas de pasto [8], de las especies “African Foxtail” and “Maasai Lovegrass”³.

Dichos semicírculos tienen entre 3 y 6 metros de ancho y 0.5 metros de profundidad⁴, tienen la forma que se muestra en la Figura V.



Fig. 5 Semicírculos de suelo en Kuku, Kenia.

Nota. Imagen tomada de la página oficial de *Justdiggít* [8].

En la Tabla 1 se relacionan todas las áreas de Tanzania y Kenia que cumplen con las características descritas anteriormente y que son objeto de estudio en este trabajo, además de

³ T. Zaan [Comunicación personal]. 22 de septiembre del 2023.

⁴ T. Zaan [Comunicación personal]. 14 de marzo del 2024.

esto, se menciona el año en el que se iniciaron actividades, los métodos usados y la cantidad de semicírculos cavados.

TABLA 1
ÁREAS OBJETO DE ESTUDIO

País	Nombre del área	Técnica	Año de inicio de actividades	Cant. de semicírculos
Kenia	Inkiito	Semicírculos de suelo sin siembra	2016	1700
Kenia	KukuB	Semicírculos de suelo sin siembra	2016	52319
Kenia	KukuD	Semicírculos de suelo sin siembra	2016	25129
Kenia	Nasipa	Semicírculos de suelo sin siembra	2018	5432
Kenia	Enkii	Semicírculos de suelo con siembra	2019	32400
Kenia	KukuA	Semicírculos de suelo sin siembra	2019	6400
Kenia	Risa	Semicírculos de suelo sin siembra	2019	18000
Kenia	Inkisanjani	Semicírculos de suelo con siembra	2021	33800
Tanzania	Pembamoto	Semicírculos de suelo con siembra	2018	5143
Tanzania	Baraka	Semicírculos de suelo con siembra	2021	3200
Tanzania	Engaruka	Semicírculos de suelo con siembra	2021	3800
Tanzania	EsilaleiSite1	Semicírculos de suelo con siembra	2021	1352
Tanzania	EsilaleiSite2	Semicírculos de suelo con siembra	2021	3648
Tanzania	Meserani	Semicírculos de suelo con siembra	2021	5000
Tanzania	Selela	Semicírculos de suelo con siembra	2021	5006
Total				202329

Nota: Adaptado de los datos suministrados por Justdiggitt vía correo electrónico⁵.

⁵ T. Zaan [Comunicación personal]. 22 de septiembre del 2023.

4.3. Exploración y selección de imágenes satelitales

4.3.1. Exploración de imágenes satelitales

En el catálogo de *Google Earth Engine* existen colecciones de imágenes satelitales climáticas, termales, geofísicas, de reflectancia de superficie, entre otras [62]. Para seleccionar la colección de imágenes satelitales con las que se desarrolla este proyecto se tuvo en cuenta las siguientes consideraciones:

- Tiempo: teniendo en cuenta que para identificar los cambios en la vegetación se debe analizar información de años anteriores al inicio de las actividades de reforestación de Justdiggit hasta el año 2022, es necesario que la colección tenga disponibilidad de datos capturados por los satélites desde el 2013 al 2022.
- Bandas espectrales: se verifica que las imágenes capturadas por los satélites tengan las bandas con las cuales se calculan los índices necesarios para capturar los puntos de entrenamiento de los modelos de aprendizaje automático supervisado y para analizar los cambios en la vegetación de las áreas objeto de estudio, por tanto, es necesario que tengan la reflectancia de la superficie terrestre.
- Propiedades de la imagen: es importante que las imágenes de la colección estén rectificadas, para evitar inconsistencia en los datos.
- Resolución espacial: esta corresponde al tamaño de los píxeles que componen la imagen y la información incluida en cada píxel, esta se mide en metros, entre más pequeña sea mejor pues representa los metros a los que representa cada píxel de la imagen y la escala de la información.

En la Tabla 2 se relacionan los satélites cuyas colecciones de imágenes se encuentran rectificadas y tienen reflectancia de la superficie terrestre.

TABLA 2

SATÉLITES CON COLECCIÓN DE REFLECTANCIA DE LA SUPERFICIE TERRESTRE

Satélite	Resolución espacial (metros)	Tiempo: Desde	Tiempo: Hasta
LandSat 4	60 y 30	22/08/1982	24/06/1993
LandSat 5	60 y 30	16/03/1984	5/05/2012
LandSat 6		3/10/1993	Lanzamiento fallido
LandSat 7	30	28/05/1999	Presente
LandSat 8	30	18/03/2013	Presente

LandSat 9	30	31/10/2021	Presente
Sentinel 2	10, 20 y 60	28/03/2017	Presente

Nota: Información recopilada de *Google Earth Engine* [62].

Se observa que las colecciones de los satélites LandSat 7 y 8 cumplen con las condiciones definidas. Sin embargo, el sensor del satélite LandSat 7 tuvo una falla del corrector de línea del escaneo en mayo de 2003, por lo que las imágenes capturadas a partir de esa fecha tienen espacios, lo que corresponde a una pérdida de aproximadamente un 25% de los datos [63]. Como se puede observar en la Figura VI, la cual se tomó del catálogo de datos de Google Earth Engine, de la colección de imágenes del satélite LandSat 7 [64], capturada el 28/07/2019 en el sector Kuku en Kenia.

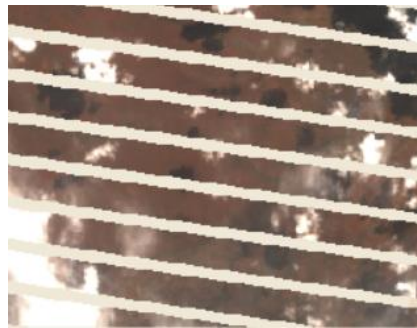


Fig. 6 Imagen LandSat 7

Nota. Imagen generada en Google Earth Engine.

Por tal razón, se selecciona la colección de imágenes capturadas por el satélite LandSat 8 para el desarrollo de este proyecto, las cuales no presentan el error de LandSat 7. Como se muestra en la Figura 7, la cual se tomó del catálogo de datos de Google Earth Engine [9], de la colección de imágenes del satélite LandSat 8, capturada el 04/07/2019 en el sector Kuku en Kenia, misma área que se muestra en la imagen VI.



Fig. 7 Imagen LandSat 8

Nota. Imagen generada en Google Earth Engine.

4.3.2. Selección de imágenes satelitales

Como parte de la exploración de los datos, se genera el Índice Diferenciado de Vegetación Normalizado (NDVI) a las imágenes de cada área de estudio⁶, este índice representa de la densidad y salud de la vegetación. Dichas imágenes fueron capturadas en el periodo de tiempo comprendido entre tres años anteriores al inicio de actividades de reforestación hasta el año 2022, con el fin de identificar la tendencia en los cambios en la vegetación en dicho rango de años. Por ejemplo, en la Figura 8 se representa el NDVI del área Enkii de los años 2016 al 2022.

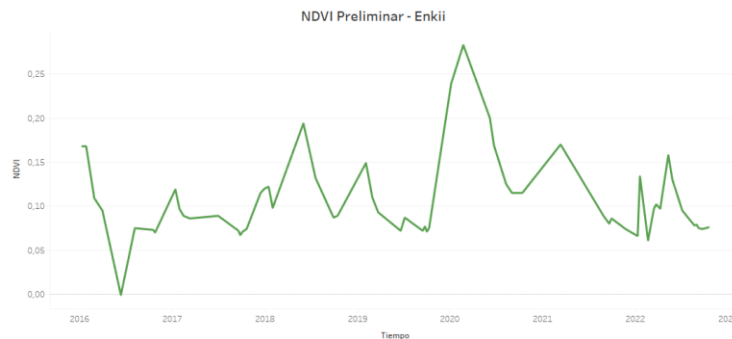


Fig. 8 NDVI preliminar del polígono Enkii

Acorde a lo informado por el Banco Mundial, en las últimas tres décadas tanto Kenia como Tanzania tienen meses en los cuales hay más precipitaciones y meses más secos [65], [66]. Mostrado en las Figuras 9 y 10.

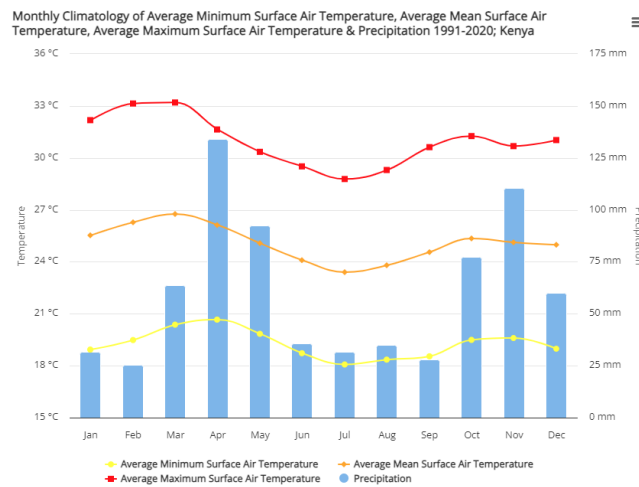


Fig. 9 Climatología mensual promedio de Kenia

Nota: Imagen tomada de la página web del Banco Mundial [65].

⁶ Este código se encuentra almacenado en Google Earth Engine.

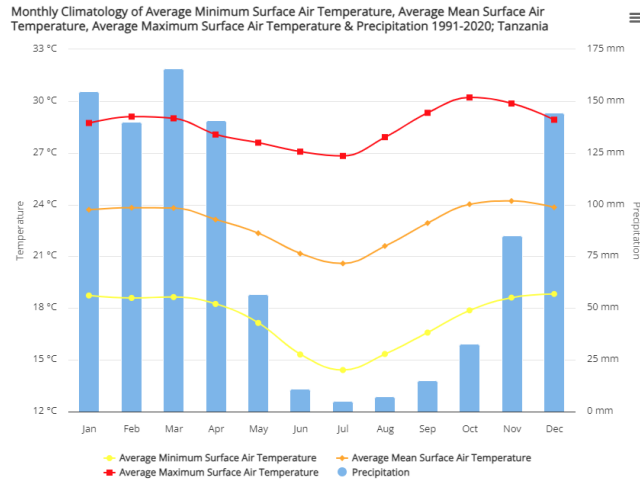


Fig. 10 Climatología mensual promedio de Tanzania

Nota: Imagen tomada de la página web del Banco Mundial [66].

Se evidencia que la vegetación cambia acorde a la estacionalidad del clima, ya que en los primeros meses de cada año se observa que el NDVI aumenta y estos corresponden a los meses en los cuales hay más precipitaciones. Esto es respaldado por lo que menciona Pedro Sánchez en su libro *Suelos del trópico: características y manejo*, en donde estudio los trópicos en diferentes países, entre los que se encuentra Kenia y concluye que las estaciones lluviosas y secas influyen en el crecimiento de las plantas de forma similar a la influencia que tiene el invierno y el verano en regiones templadas [67].

Con el fin de tener menos variabilidad en la información causada por la estacionalidad climática, se decide tomar las imágenes satelitales de los meses de junio y julio de cada año para poder comparar año a año los cambios en la vegetación en dichos meses. Estos corresponden a meses consecutivos y secos, lo cual ayuda a tener menos nubosidad y por tanto menos pixeles perdidos. Además de esto, para identificar los cambios causados por la reforestación, se toman las imágenes correspondientes al rango de tiempo comprendido entre tres años anteriores del inicio de actividades hasta el 2022 de cada área.

4.4. Preparación de las imágenes satelitales

Para el desarrollo de este proyecto usan las colecciones del satélite LandSat 8 que se encuentran en Google Earth Engine (GEE), específicamente la colección 2, por lo que se deben realizar los siguientes procesos para preparar las imágenes antes de empezar a trabajar con estas, acorde con las sugerencias de GEE [68]:

- Escalado de valor: el proveedor de las imágenes satelitales sugiere que las imágenes se escalen, debido a que estas vienen en números enteros. Cada satélite tiene escalas diferentes, en este caso LandSat 8 tiene $2.75e-05$ como factor de escala y -0.2 como desplazamiento [9], [68].
- Recorte: Las imágenes son recortadas, de forma que en estas quede solo la información correspondiente a cada área objeto de estudio.
- Enmascaramiento de nubes: las imágenes de esta colección se capturan cada diez y seis días y puede que parte de dichas imágenes tengan nubes y/o sombras de nubes las cuales pueden distorsionar la información y dificultan la identificación de la superficie del suelo. Las imágenes satelitales de la colección LandSat 8 tiene la banda QA_PIXEL, la cual permite enmascarar los pixeles con dicha problemática y quitarlos de las imágenes [9], [68], [69].
- Composición de imágenes: teniendo en cuenta que las imágenes pueden quedar con pixeles faltantes a causa del enmascaramiento de nubes, Google Earth Engine sugiere hacer una composición de imágenes de la misma estación climática, ya que la superficie del suelo no suele cambiar mucho en una misma estación y se puede hacer por medio de la reducción de rangos de fechas y la mediana de cada año [69]. En la Figura 11, a manera de ejemplo se muestran las imágenes A, B y C en donde cada celda representa un pixel, las celdas sin color corresponden a pixeles faltantes debido al enmascaramiento de nubes y las celdas con colores representan pixeles con información, también tenemos la imagen resultante de aplicar la reducción con la mediana de las imágenes A, B y C, la cual está mucho más completa y libre de nubes, aunque aún tiene algunos pixeles faltantes son menos que en las imágenes individuales A, B o C.

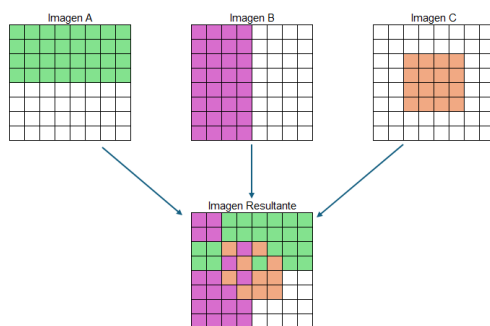


Fig. 11 Composición de imágenes

En todas las imágenes usadas en este proyecto se aplica escalamiento de valores, recorte y enmascaramiento de nubes, además de esto se hace uso de la función de reducción con la mediana para componer imágenes del 01 de junio al 31 de julio de cada año, generando de esta forma una sola imagen compuesta para cada año de cada área objeto de estudio.

Las imágenes satelitales completas corresponden a escenas superpuestas que cubren aproximadamente 170 km x 183 km y si bien estas imágenes tienen la propiedad de cobertura de nubes por la cual se pueden filtrar [9], este porcentaje corresponde a la nubosidad de la imagen completa mas no a la nubosidad del área objeto de estudio. Por tal razón, el porcentaje de nubosidad de cada imagen de las áreas objeto de estudio se determina con la cantidad de pixeles faltantes, resultado de aplicar las preparaciones mencionadas (escalamiento de valores, recorte, enmascaramiento de nubes y composición de imágenes).

De esta manera, se descartan algunas imágenes teniendo en cuenta lo mencionado y que a nivel general se considera que un valor superior al 20% de nubosidad interfiere con la investigación [70], pues representa perdida considerable de información. En la Tabla 3 se relacionan dichas imágenes compuestas descartadas, el año correspondiente, porcentaje de nubosidad y otros datos relevantes.

TABLA 3

IMÁGENES CON PORCENTAJE DE NUBOSIDAD SUPERIOR A 20%

Áreas	Año de imágenes	Cant. imágenes encontradas	Total pixeles por área	Cant. pixeles con datos	Cant. de pixeles faltantes	% nubosidad
Inkiito	2016	3	295	0	295	100.0%
Inkiito	2018	4	295	123	172	58.3%
Inkiito	2019	2	295	0	295	100.0%
KukuB	2014	3	3471	1209	2262	65.2%
KukuB	2015	4	3471	2711	760	21.9%
KukuD	2013	4	1441	954	487	33.8%
KukuD	2014	3	1441	922	519	36.0%
KukuD	2015	4	1441	383	1058	73.4%
KukuD	2017	4	1441	158	1283	89.0%
Nasipa	2016	3	135	0	135	100.0%
Nasipa	2021	5	135	27	108	80.0%
Enkii	2016	7	2377	1445	932	39.2%
KukuA	2017	4	1125	724	401	35.6%
Risa	2016	3	2578	0	2578	100.0%
Risa	2021	3	2578	1693	885	34.3%
Meserani	2020	8	41	0	41	100.0%
Meserani	2021	6	41	0	41	100.0%
EsilaleiSite1	2019	6	470	207	263	56.0%

EsilaleiSite1	2020	10	470	218	252	53.6%
EsilaleiSite1	2022	6	470	141	329	70.0%
Baraka	2018	12	339	114	225	66.4%
Baraka	2020	10	339	189	150	44.2%
Baraka	2022	6	339	1	338	99.7%
Selela	2018	8	547	132	415	75.9%
Selela	2020	6	547	402	145	26.5%
Selela	2022	3	547	70	477	87.2%

5. Entrenamiento de modelos y evaluación

En esta fase se realizan los pasos que nos permiten entrenar y evaluar los modelos de aprendizaje automático para la clasificación de los suelos de las áreas objeto de estudio. Como se menciona en la metodología CRISP-DM esta se puede realizar en iteraciones para ir mejorando los modelos [10], para este trabajo se ejecutan dos iteraciones de la metodología mostrada en la Figura 12.

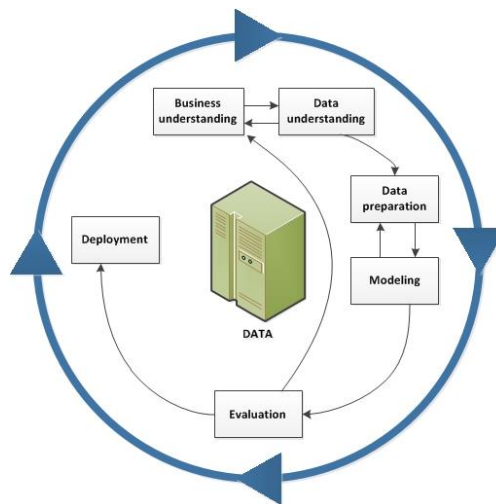


Fig. 12 Metodología CRISP-DM

Nota: Imagen tomada de página de IBM [71].

5.1. Iteración I

Acorde a lo definido en la metodología y como se ilustra en la Figura 1, en esta fase se define el polígono de entrenamiento, captura de puntos de entrenamiento, entrenamiento de los modelos y evaluación de esto.

5.1.1. Definición de polígono de entrenamiento

En esta primera iteración se definió un polígono para realizar el entrenamiento del modelo que abarcara la mayor parte del área del parque nacional Chyulu (Kuku) en el sur de Kenia, debido a que esta área tiene coberturas de suelo, vegetación, agricultura, urbanización y cuerpos de agua. Dicho polígono se muestra en la Figura 13.

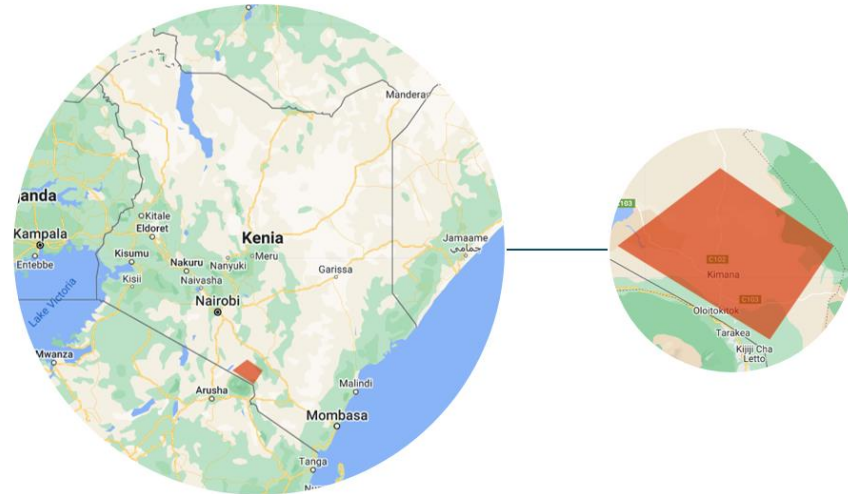


Fig. 13 Polígono de entrenamiento - Iteración I

Nota: Elaboración propia, graficado en Google Earth Engine [72].

5.1.2. Captura de puntos de entrenamiento y pruebas

Los puntos de entrenamiento y pruebas son capturados manualmente en la aplicación Google Earth Engine [72], en los cuales se obtiene solo la ubicación (latitud y longitud) y se les asigna una categoría, ya sea vegetación, suelo, agua, urbanización o agricultura. Para tal fin se selecciona el mes de enero del año 2018 para realizar dicha captura, esto con el fin de apoyarnos en las imágenes LandSat 8 [9] y Sentinel 2 [73], ya que estas últimas tienen menor resolución espacial y nos permite tener una visualización más nítida.

Los índices espectrales ayudan a resaltar los fenómenos que se estudian en este proyecto, como por ejemplo los cambios en la vegetación, suelo y agua, por tal razón se calculan los índices a partir de los cuales se capturan algunos de los puntos de entrenamiento, como se describe a continuación:

- Vegetación y suelo: para realizar la captura de esos puntos de estas dos categorías se resalta la vegetación inicialmente con el cálculo del índice diferencial de vegetación normalizado NDVI (Normalized Difference Vegetation Index) realizado con la imagen compuesta de LandSat 8. Este índice es ampliamente utilizado en teledetección debido a que permite resaltar las masas vegetales y su salud [18], lo cual genera como resultado la siguiente visualización en donde los valores se muestran en una gama de colores entre café y verde, donde el primero representa suelo y el segundo representa vegetación:

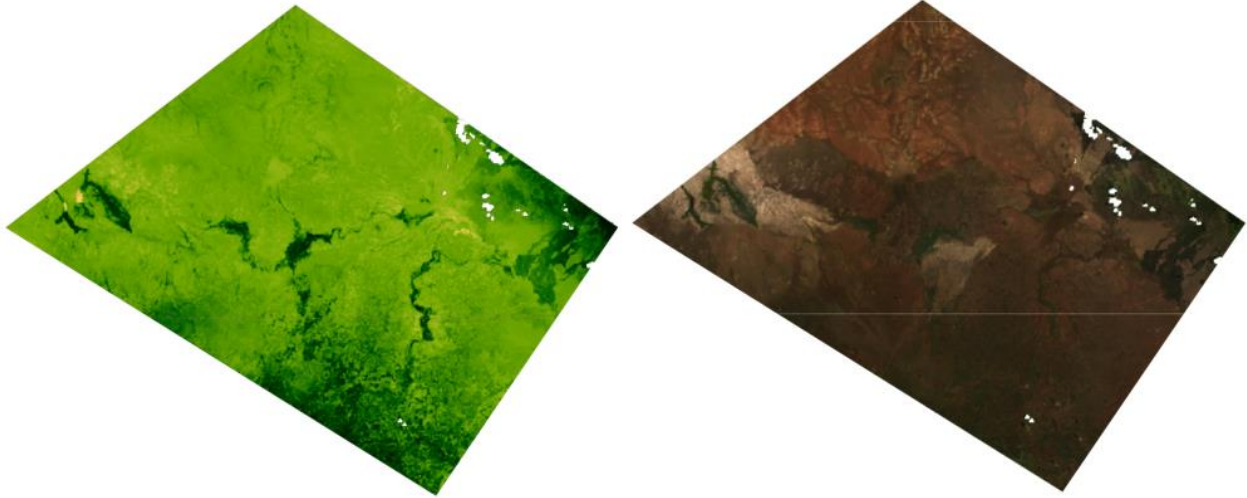


Fig. 14 Comparación visualización NDVI e Imagen RGB

Nota: Elaboración propia, graficado en Google Earth Engine [72].

En la Figura 14 se compara la visualización del NDVI vs la Imagen RGB (Red Green Blue) que es la visualización cercana al color natural o real, en donde se observa que el terreno esta mucho menos poblado de vegetación que lo que se muestra en el NDVI. Esto se debe a que el suelo desnudo ejerce una influencia sobre el espectro de dosel y por tanto en el resultado del NDVI [19]. Para este tipo de terreno en donde hay cobertura de vegetación y suelo desnudo es más útil el Índice de vegetación ajustado al suelo SAVI (Soil Adjusted Vegetation Index), el cual tiene un factor de ajuste, acorde a lo sugerido por Alfredo Huete para un terreno como este se sugiere usar un factor de corrección de 0.5 [19], con lo cual se obtiene la visualización generada a partir del cálculo de dicho índice de la imagen compuesta de LandSat 8, en donde los valores se muestran en una gama de colores entre café y verde, donde el primero representa suelo y el segundo representa vegetación. Esta se muestra en la Figura 15.

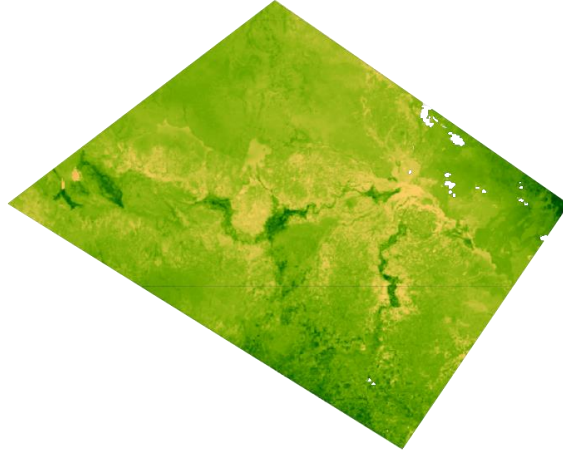


Fig. 15 Visualización SAVI Iteración I

Nota: Elaboración propia, graficado en Google Earth Engine [72].

Esta última visualización se ajusta más a la realidad del terreno, motivo por el cual a partir de esta visualización se capturan los puntos de entrenamiento y pruebas de las categorías vegetación y suelo.

- Agua: para realizar la captura de esos puntos de esta categoría se resaltan los cuerpos de agua con el cálculo del índice diferencial de agua normalizado NDWI (Normalized Difference Water Index) con la imagen compuesta de LandSat 8 [20], con lo cual se obtiene la siguiente visualización en donde los cuerpos de agua se muestran en gamas de color azul. Esto se muestra en la figura 16.

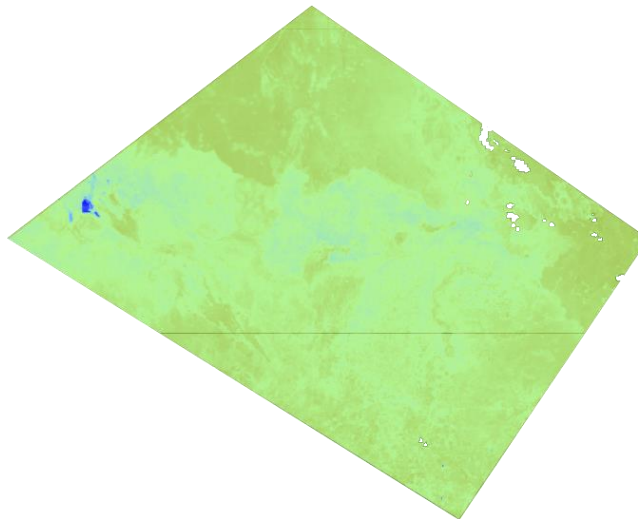


Fig. 16 Visualización NDWI Iteración I

Nota: Elaboración propia, graficado en Google Earth Engine [72].

- Agricultura y urbanización: En este caso se usa la visualización natural o real de la imagen Sentinel 2 para capturar para la captura de estos puntos de entrenamiento y pruebas, ya que dentro del polígono de entrenamiento hay áreas muy pequeñas correspondientes a estos tipos de cobertura de suelo, también es de resaltar que las carreteras y vías de las zonas urbanas no están pavimentadas en su gran mayoría acorde a la validación realizada en Google Earth Pro, como se muestra en las Figuras 17 y 18.



Fig. 17 Visualización zonas urbanas

Nota: Visualizaciones tomadas de Google Earth Pro [74] y Google Earth Engine [72].



Fig. 18 Visualización zonas agrícolas

Nota: Visualizaciones tomadas de Google Earth Pro [74] y Google Earth Engine [72].

Es de aclarar que las imágenes de Google Earth Pro [74] solo están disponibles para visualizar en dicha aplicación y solo corresponde a la visualización natural, estas imágenes no tienen las bandas con los datos que permiten entrenar los modelos de aprendizaje automático, a diferencia de Google Earth Engine [72] que, si permite acceder, descargar y trabajar las imágenes satelitales y los datos de sus bandas. Por lo que estas imágenes de Google Earth Pro solo se usan de guía para identificar zonas urbanas y de agricultura en las imágenes de Google Earth Engine.

En la Tabla 4 se relaciona la cantidad de puntos capturados en cada categoría.

TABLA 4

PUNTOS DE ENTRENAMIENTO Y PRUEBAS POR CATEGORÍA DE LA ITERACIÓN I

Categoría	Cantidad de puntos
Vegetación	340
Suelo	278
Agua	282
Urbano	140
Agricultura	90
Total	1130

Si bien es aconsejable hacer la verificación de la cobertura de algunos de los puntos de entrenamiento en sitio, esto se sale del alcance de este trabajo, pero se sugiere hacer dicha verificación para trabajos futuros.

5.1.3. Entrenamiento de modelos de aprendizaje automático

Después de realizar la captura de los puntos, lo que comprende categorización y ubicación (latitud y longitud), se realiza la extracción de los datos de la imagen de entrenamiento, es decir, los valores de las bandas correspondientes a los píxeles de dichos puntos capturados, esto teniendo en cuenta que cada píxel de las imágenes tiene un valor diferente en cada banda. Como se representa en la Figura 19, se podría decir que cada imagen esta compuestas por capas de datos sobrepuestas las cuales representan a las bandas y cada una de estas capas tiene valores diferentes para cada celda.

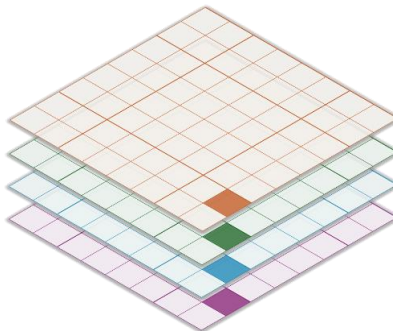


Fig. 19 Composición de bandas

Como se mencionó en el punto anterior, para el entrenamiento de los modelos se eligió la imagen de enero 2018 del satélite LandSat 8. Las imágenes de dicho satélite tienen diez y nueve bandas [9], sin embargo, solo se seleccionan las bandas de reflectancia para el entrenamiento y se descartan las bandas térmicas, aerosol y de QA, esto quiere decir, que el conjunto de datos comprende solo los valores de las bandas que se mencionan a continuación:

- SR_B2: Reflectancia de superficie de banda 2 (azul).
- SR_B3: Reflectancia de superficie de banda 3 (verde).
- SR_B4: Reflectancia de superficie de banda 4 (roja).
- SR_B5: Reflectancia de superficie de banda 5 (infrarrojo cercano).
- SR_B6: Reflectancia de superficie de banda 6 (infrarrojo de onda corta 1).
- SR_B7: Reflectancia de superficie de banda 7 (infrarrojo de onda corta 2).

Después de realizar esta extracción de datos se procede a realizar el particionamiento de los datos, dejando un 70% de los datos para entrenamiento y 30% para pruebas. Para realizar la clasificación de cobertura de las áreas de objeto de estudio se contemplan los siguientes modelos de aprendizaje automático supervisado:

- Máquinas de soporte de vectores SVM (Support Vector Machine): para el entrenamiento de este modelo se dejan los hiper parámetros por defecto⁷, con lo cual se observa la clasificación mostrada en la Figura 20.

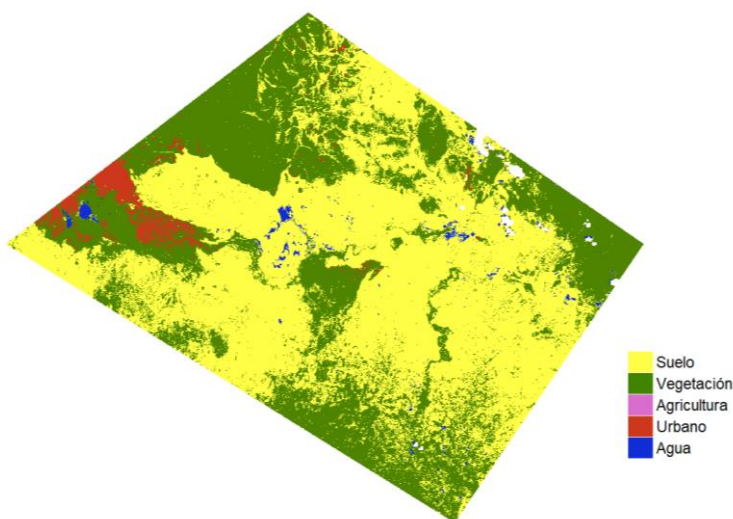


Fig. 20 Clasificación SVM de Iteración I

Visualmente se identifica que el modelo clasifica erradamente los píxeles de zonas urbanas pues dichas zonas realmente se encuentran en el centro del polígono; también se observa que algunos cuerpos de agua los clasifica correctamente, sin embargo, en algunas áreas los clasifica como vegetación; por otra parte, no clasifico ningún píxel con la categoría de Agricultura; en lo referente

⁷ Este código se encuentra almacenado en Google Earth Engine.

a las categorías de vegetación y suelo se percibe una clasificación cercana a lo observado con el índice SAVI.

- Árboles de decisión: para el entrenamiento de este modelo también se dejan los hiper parámetros por defecto⁸, con lo cual se obtiene la clasificación que se muestra en la Figura 21.

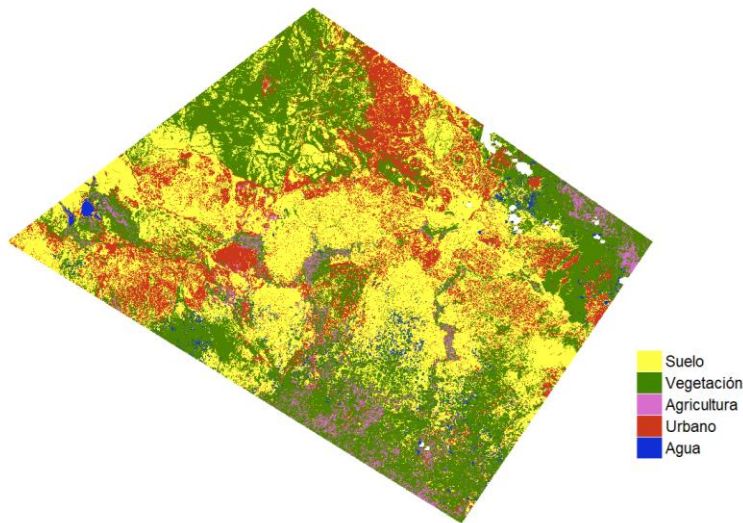


Fig. 21 Clasificación Árboles de decisión de Iteración I

Al igual que con el modelo anterior, se observa que el este modelo también clasifica erradamente los píxeles de zonas urbanas y algunos cuerpos de agua; por otra parte, si clasifico píxel con la categoría de Agricultura, pero estos realmente corresponden a zonas con vegetación frondosa; en lo referente a las categorías de vegetación y suelo se percibe una clasificación cercana a lo observado con el índice SAVI, pero menos acertada que en el modelo anterior.

- Random Forest: en este caso se realiza un entrenamiento con 100 árboles y los demás hiper parámetros se dejan por defecto⁹, con lo cual se obtiene la clasificación que se muestra en la Figura 22.

⁸ Este código se encuentra almacenado en Google Earth Engine.

⁹ Este código se encuentra almacenado en Google Earth Engine.

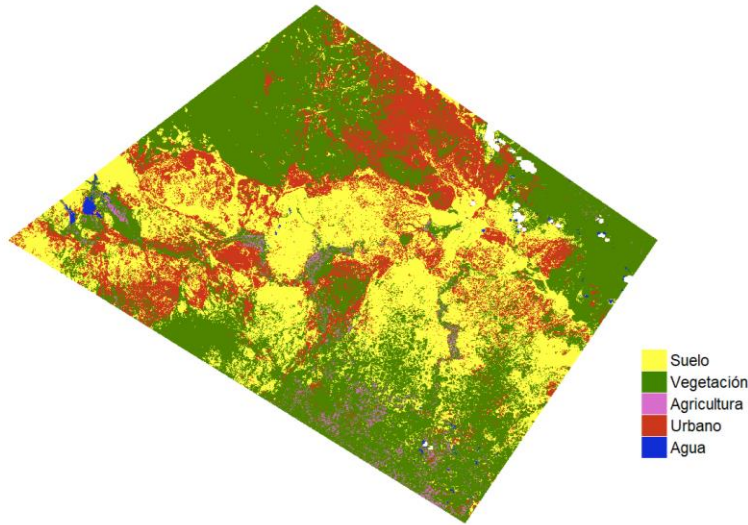


Fig. 22 Clasificación Random Forest de Iteración I

Este modelo también clasifica erradamente los píxeles de zonas urbanas y algunos cuerpos de agua; además de esto, cataloga algunos píxeles como Agricultura, pero estos corresponden a zonas con vegetación frondosa; en lo referente a las categorías de vegetación y suelo se percibe una clasificación más cercana a lo observado con el índice SAVI en comparación con los otros modelos.

5.1.4. Evaluación de modelos

Para evaluar los modelos de aprendizaje automático supervisado se elige la medición de la exactitud de la clasificación y el coeficiente Kappa para lo cual se usa la matriz de confusión. En las Figuras 23, 24 y 25 se muestran dichas matrices.

	Agua	Vegetación	Agricultura	Urbano	Suelo
Agua	78	2	0	0	2
Vegetación	0	83	0	1	13
Agricultura	0	18	0	0	0
Urbano	0	8	0	0	25
Suelo	13	2	0	0	70

Fig. 23 Matriz de confusión modelo SVM Iteración I

	Agua	Vegetación	Agricultura	Urbano	Suelo
Agua	79	0	1	0	2
Vegetación	2	80	11	1	3
Agricultura	0	3	15	0	0
Urbano	0	3	1	25	4
Suelo	2	1	0	11	71

Fig. 24 Matriz de confusión modelo Arboles de decisión Iteración I

	Agua	Vegetación	Agricultura	Urbano	Suelo
Agua	80	0	1	0	1
Vegetación	0	88	6	2	1
Agricultura	0	3	15	0	0
Urbano	0	4	1	24	4
Suelo	2	1	0	9	73

Fig. 25 Matriz de confusión modelo Random Forest Iteración I

En la tabla 5 se muestran los valores de exactitud y coeficiente Kappa obtenidos.

TABLA 5

EVALUACIÓN DE MODELOS ITERACIÓN I

Modelo	Exactitud	Kappa
SVM	0.7333	0.6295
Arboles de decisión	0.8571	0.8121
Random Forest	0.8888	0.8528

5.2. Iteración II

Teniendo en cuenta lo mencionado, se decide realizar otra iteración para entrenar nuevamente los modelos y mejorar su rendimiento para clasificar la cobertura de suelos.

5.2.1. Definición de polígono de entrenamiento

Con el fin de que el polígono de entrenamiento abarque más áreas en las cuales Justdiggit ha realizado actividades de reforestación, se decide cambiar dicho polígono, el cual también estará ubicado al sur de Kenia.

También se identifica que algunas áreas de los cuerpos acuáticos están siendo clasificadas como vegetación, esto posiblemente se deba al crecimiento de plantas subacuáticas cerca a las costas, como se muestra en la Figura XXVI y esto puede estar afectando la clasificación de los modelos.



Fig. 26 Costas de cuerpos de agua – Visualización RGB

Nota: Visualización tomada de Google Earth Engine [72], satélite Sentinel 2.

Además de esto, ya que las áreas en las cuales se realizan actividades de reforestación no comprenden cuerpos de agua se decide realizar el enmascaramiento de dichos cuerpos, de forma que estos no hagan parte del polígono de entrenamiento. Dicho polígono se muestra en la Figura 27.

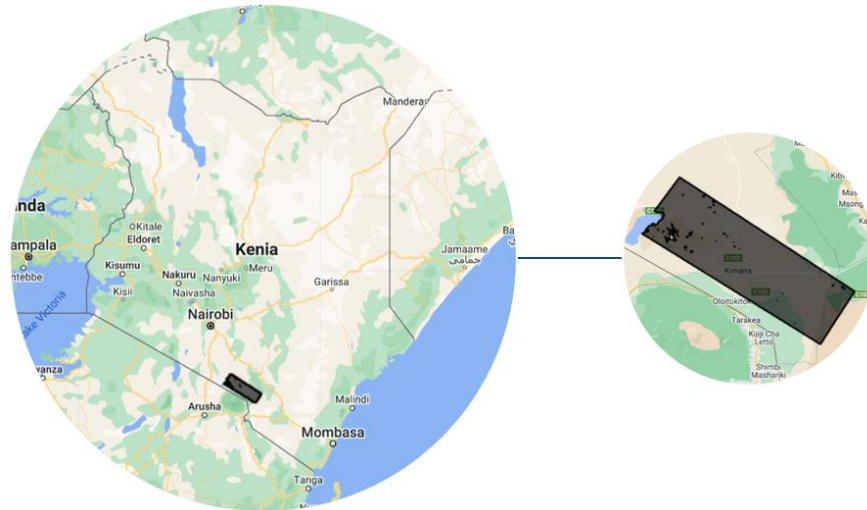


Fig. 27 Polígono de entrenamiento - Iteración II

Nota: Elaboración propia, graficado en Google Earth Engine [72].

5.2.2. Captura de puntos de entrenamiento y pruebas

Al igual que en la iteración anterior, los puntos de entrenamiento y pruebas son capturados manualmente en la aplicación Google Earth Engine [72], en los cuales se obtiene solo la ubicación (latitud y longitud) y se les asigna una categoría, sin embargo, para esta iteración se seleccionan los meses de junio y julio del año 2020 para realizar dicha captura, pues acorde a lo mencionado en la sección 4.3.2 se observa que la vegetación cambia acorde a la estacionalidad climática, por tal razón se selecciona estos dos meses que corresponden a temporada seca.

Se capturan solo puntos de coberturas de vegetación y suelo teniendo en cuenta que las áreas las cuales Justdiggit realizó actividades de reforestación no corresponden a zonas urbanas ni agrícolas y que se excluyeron los cuerpos de agua en el polígono de entrenamiento. Para la captura de estos puntos se usa el índice SAVI con factor de corrección de 0.5 con lo cual se obtiene la visualización generada a partir del cálculo de dicho índice de la imagen compuesta de LandSat 8, en donde los valores se muestran en una gama de colores entre café y verde, donde el primero representa suelo y el segundo vegetación¹⁰. Esto se muestra en la Figura 28.

¹⁰ Este código se encuentra almacenado en Google Earth Engine.

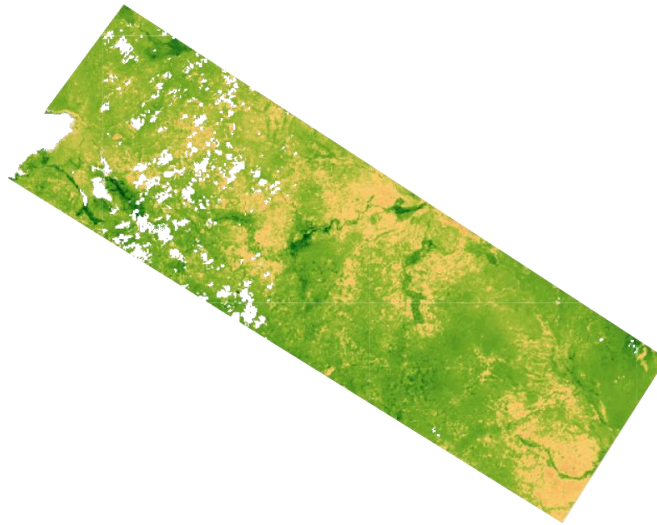


Fig. 28 Visualización SAVI Iteración II

Nota: Elaboración propia, graficado en Google Earth Engine [72].

Lastimosamente Google Earth Engine en la versión gratuita solo permite manejar archivos de hasta 512 KB, por tal razón solo se pudo realizar la captura de los puntos que se relacionan en la Tabla 6.

TABLA 6

PUNTOS DE ENTRENAMIENTO Y PRUEBAS POR CATEGORÍA DE LA ITERACIÓN II

Categoría	Cantidad de puntos
Vegetación	1124
Suelo	1129
Total	2253

5.2.3. Entrenamiento de modelos de aprendizaje automático

Al igual que en la iteración anterior, después de realizar la captura de los puntos, se realiza la extracción de los datos de la imagen de entrenamiento, es decir, se extraen los valores de las bandas correspondientes a los píxeles de dichos puntos capturados, para lo cual se usa la imagen compuesta de junio y julio del año 2020. Dichas bandas corresponden a las seleccionadas en la iteración anterior.

Después de realizar esta extracción de datos se procede a realizar el particionamiento de los datos, dejando un 70% de los datos para entrenamiento y 30% para pruebas. Para realizar la clasificación de cobertura de las áreas de objeto de estudio se contemplan los siguientes modelos de aprendizaje automático supervisado:

- Máquinas de soporte de vectores SVM (Support Vector Machine): para el entrenamiento de este modelo se usan los siguientes hiper parámetros¹¹:
 - Procedimiento de decisión: en este caso se usa la opción “Margin” con el fin de que se dejase un margen entre las clases objetivo.
 - Tipo de núcleo: se selecciona el “Kernel Linear” con el fin de que el modelo use una función lineal para clasificar los datos según los patrones que encuentre.

Los demás hiper parámetros se dejaron por defecto, con lo cual se obtiene la clasificación que se muestra en la Figura 29.

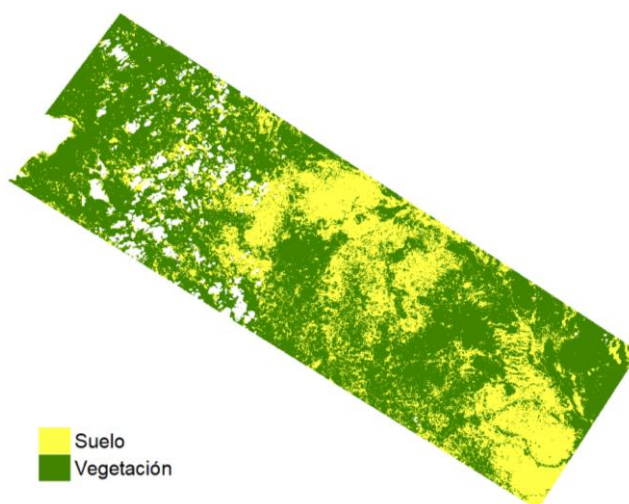


Fig. 29 Clasificación SVM de Iteración II

Visualmente se percibe que el modelo clasifica más acertadamente los píxeles de suelo y vegetación, pues la imagen se aproxima más a los resultados arrojados por el índice SAVI.

- Árboles de decisión: para el entrenamiento de este modelo se usan los siguientes hiper parámetros¹²:
 - Población por nodo: En este caso se usa 40, con el fin de que cada nodo tenga una población mínima de 40 puntos.
 - Nodos: en este caso, se definió una cantidad máxima de nodos de 80.

Con esta configuración se obtiene la clasificación que se muestra en la Figura 30.

¹¹ Este código se encuentra almacenado en Google Earth Engine.

¹² Este código se encuentra almacenado en Google Earth Engine.

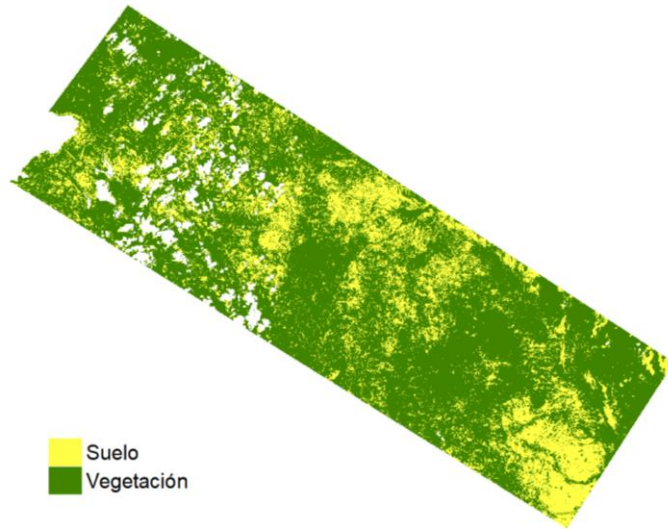


Fig. 30 Clasificación Arboles de decisión de Iteración II

Al igual que con el modelo anterior, se observa que este modelo también clasifica bien los píxeles en las categorías de vegetación y suelo ya que es cercana a lo observado con el índice SAVI.

- Random Forest: en este caso se realiza un entrenamiento del modelo de bosques aleatorios con los siguientes hiper parámetros¹³:
 - Numero de árboles: para este modelo se usan seis árboles para el entrenamiento.
 - Población por nodo: En este caso se usa 40, con el fin de que cada nodo tenga una población mínima de 40 puntos.
 - Nodos: en este caso, se definió una cantidad máxima de nodos de 80 por cada árbol.

Con esta configuración se obtiene la clasificación que se muestra en la Figura 31.

¹³ Este código se encuentra almacenado en Google Earth Engine.



Fig. 31 Clasificación Random Forest de Iteración II

Al igual que con los modelos anteriores, se percibe que este modelo también clasifica bien los píxeles en las categorías de vegetación y suelo ya que es cercana a lo observado con el índice SAVI.

5.2.4. Evaluación de modelos

Para evaluar los modelos de aprendizaje automático supervisado también se elige la medición de la exactitud de la clasificación y el coeficiente Kappa. Se encontraron los valores relacionados en la Tabla 7.

TABLA 7

EVALUACIÓN DE MODELOS ITERACIÓN II

Modelo	Exactitud	Kappa
SVM	0.9793	0.9584
Arboles de decisión	0.9630	0.9259
Random Forest	0.9615	0.9226

Como se muestra en la Tabla VII, los valores de exactitud son altos, con el fin de identificar si los modelos están generalizando bien, es decir, identificar si los modelos no están sobre ajustados o sub ajustados [30], se emplea el método de validación cruzada usando cinco segmentos e iteraciones con cada modelo, además de calcular el promedio y varianza de la exactitud de estos. Con lo cual se encontraron los valores relacionados en la Tabla 8.

TABLA 8
VALIDACIÓN CRUZADA

Iteración	SVM	Arboles de decisión	Random Forest
I	0.9869	0.9475	0.9540
II	0.9803	0.9694	0.9301
III	0.9868	0.9430	0.9605
IV	0.9861	0.9769	0.9446
V	0.9821	0.9576	0.9420
Promedio	0.9845	0.9589	0.9462
Varianza	0.000009147	0.000205135	0.000136407

Acorde los datos mostrados en la Tabla 8, el modelo que tiene mejor desempeño es SVM, esto teniendo en cuenta que es el que tiene la exactitud promedio más alta, la varianza más pequeña, ya que las exactitudes de las iteraciones no difieren significativamente, lo que indica que este modelo es el que está generalizando de mejor manera, en comparación con los otros dos modelos.

6. Aplicación del modelo clasificación

En esta fase se aplica el modelo con mejor desempeño con el fin de identificar la cobertura de suelos y vegetación de las áreas objeto de estudio, además de esto se analiza los resultados. Si bien se muestra la información de los años en los cuales se realizaron las actividades de reforestación, estas no se tienen en cuenta en la comparación de resultados, ya que no se tiene información de en qué mes se realizaron dichas actividades ni cuanto duraron realizándolas.

En dicho análisis se tienen en cuenta los siguientes factores:

- Áreas con vegetación: con la aplicación del modelo se determina el porcentaje de cobertura de vegetación y suelo en las áreas objeto de estudio.
- El índice SAVI: permite identificar la densidad y salud de la vegetación minimizando la influencia del brillo del suelo desnudo, cuya escala se presenta en valores entre -1 y 1, en donde los valores negativos o cercanos a cero representan superficies sin vegetación y los valores positivos representan cobertura vegetal, entre más cercano sea a 1 la vegetación será más densa [19].
- Las precipitaciones: este factor se muestra como parte del análisis teniendo en cuenta que la vegetación en este sector de África se ve afectada por la estacionalidad climática y la cantidad de precipitaciones que se hayan presentado en las áreas objeto de estudio [67]. Es importante recalcar que desde el 2019 África oriental está sufriendo de escasez de agua por falta de lluvias, ya que incluso en las temporadas del año en las que usualmente se dan más precipitaciones estas no se han presentado [75], [76]. Esta sequía ha sido catalogada por la Agencia de la ONU para los Refugiados ACNUR como la peor sequía de los últimos 40 años [76], y esto evidentemente afecta los resultados de la reforestación realizada por Justdiggitt y por tal razón hace parte de este análisis.

La información de las precipitaciones acumuladas mensuales es obtenida del conjunto de datos publicado por la Universidad de Idaho, las cuales se encuentran 'mm', es decir, litros de lluvia por metro cuadrado [77]. Ya que para este estudio se tienen en cuenta dos meses por cada año (junio y julio) se genera el promedio de las precipitaciones acumuladas de dichos meses.

6.1. Polígonos con semicírculos de suelo sin siembra

En esta sección se relacionan las áreas objeto de estudio en las cuales se realizó reforestación con semicírculos de suelo sin siembra de semillas de pasto.

6.1.1. Inkiito

En este polígono se iniciaron actividades de reforestación en el 2016 por lo cual la información se genera desde el 2013 con el fin de identificar si la vegetación varía con las labores de Justdiggit, es decir, con el objetivo de contrastar el antes y el después de la reforestación. En la Tabla 9 se muestra la información obtenida.

TABLA 9
RESULTADOS INKIITO

Año	% Suelo	% Vegetación	% Faltante	SAVI	Diferencia SAVI	Lluvia	Observación
2013	0.00	100.00	0.00	0.2137		3.5	
2014	24.75	75.25	0.00	0.1412	-0.0724	9	
2015	45.08	54.92	0.00	0.1458	0.0046	3	
2016						5	Imagen excluida
2017	36.95	63.05	0.00	0.1446	0.1446	8.5	
2018						7	Imagen excluida
2019						1.5	Imagen excluida
2020	0.00	100.00	0.00	0.2227	0.0781	2	
2021	0.00	100.00	0.00	0.2057	-0.0170	2	
2022	83.05	16.95	0.00	0.1382	-0.0675	2.5	

A pesar de que se tienen tres años excluidos (2016, 2018 y 2019), debido a que las imágenes compuestas tienen nubosidad superior al 20%, se tiene información del antes y después de las actividades de reforestación, en donde se destaca que en los años previos (2013, 2014 y 2015) se observa una cobertura de vegetación promedio de 76.72% y el índice promedio SAVI de 0.1669 con precipitación promedio de 5.17mm, mientras que en los años posteriores (2017, 2020, 2021 y 2022) se observa una cobertura de vegetación promedio de 70%, esto representa una disminución del área con vegetación, sin embargo, el índice promedio SAVI es de 0.1778 lo que representa un aumento en la densidad de las plantas, además de que esto se dio con una precipitación promedio de 3.75mm, inferior a las precipitaciones de los años previos a la reforestación.

En la Figura 32 se contrastan los resultados de lluvia y el índice SAVI, llama la atención la comparación de los años 2015 y 2020, años anterior y posterior a las actividades de reforestación,

en donde con precipitaciones de 3mm y 2mm se ve una gran diferencia en el índice SAVI. Esto indica que la técnica de reforestación tiene efectos positivos, a pesar de la sequía prolongada en el Cuerno de África.

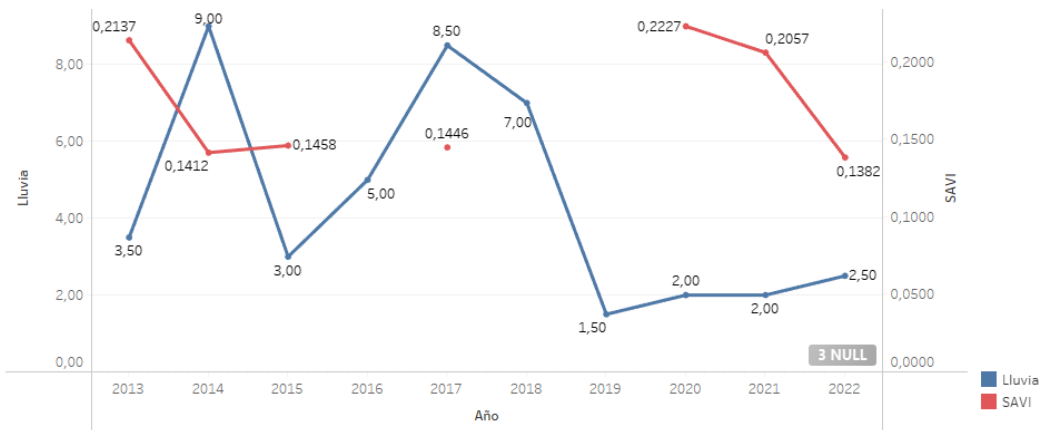


Fig. 32 Lluvias e índice SAVI del polígono Inkiito

En la Figura 33 se contrastan los resultados de lluvia y porcentaje de cobertura vegetal, al igual que en la gráfica anterior, llama la atención la comparación de los años 2015 y 2020, años anterior y posterior a las actividades de reforestación, en donde con precipitaciones de 3mm y 2mm se ve una gran diferencia en la cobertura de vegetación.

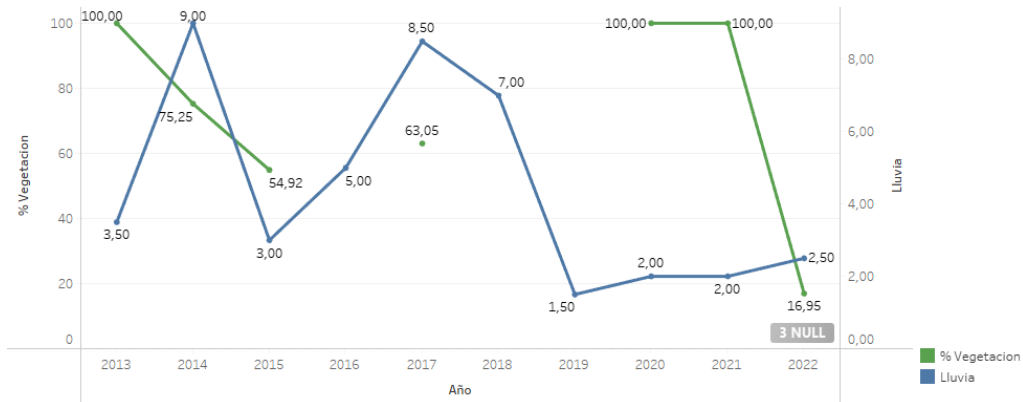


Fig. 33 Lluvias y % de cobertura vegetal del polígono Inkiito

En la Figura 34 se contrastan los resultados de cobertura de vegetación, con rango de cero a uno y el índice SAVI, cuyo rango máximo es 1. A pesar de que este último indica la densidad de la vegetación, este no necesariamente maneja las mismas curvas que la cobertura de vegetación, pues pueden existir áreas con vegetación más vigorosas que otras.

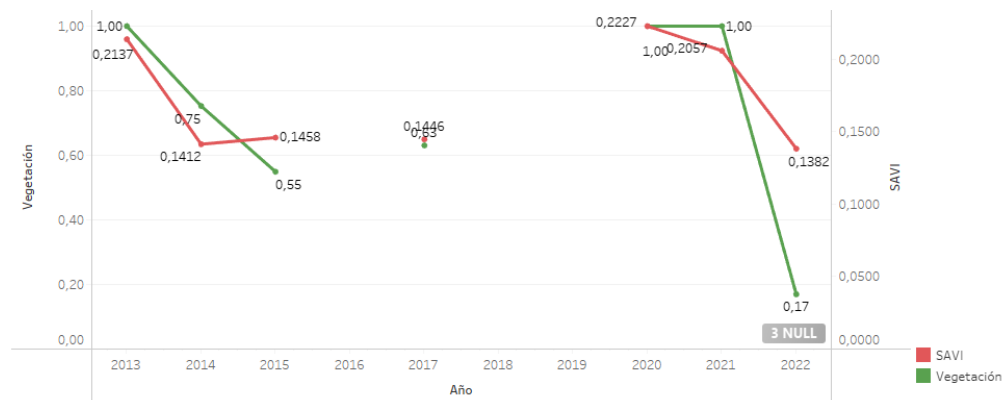


Fig. 34 índice SAVI y Cobertura vegetal del polígono Inkiito

También se destaca que en el año siguiente a la reforestación (2017) no hay variación significativa respecto al año anterior (2015).

6.1.2. KukuB

En este polígono también se iniciaron actividades de reforestación en el 2016 por lo cual la información se genera desde el 2013 con el fin de identificar si la vegetación varía con las labores de Justdiggit. En la Tabla 10 se muestra la información obtenida.

TABLA 10

RESULTADOS KUKUB

Año	% Suelo	% Vegetación	% Faltante	SAVI	Diferencia SAVI	Lluvia	Observación
2013	94.84	1.07	4.09	0.1121		3	
2014						7.5	Imagen excluida
2015						1	Imagen excluida
2016	100.00	0.00	0.00	0.0706	-0.015405	5.5	
2017	88.62	0.00	11.38	0.0657	-0.004896	9.5	
2018	46.07	44.37	9.56	0.1778	0.112112	6.5	
2019	99.91	0.09	0.00	0.1157	-0.062132	2	
2020	30.97	69.03	0.00	0.2029	0.087242	2	
2021	99.68	0.32	0.00	0.1156	-0.087288	1.5	
2022	98.91	1.09	0.00	0.1167	0.001084	1.5	

A pesar de que hay dos años excluidos (2014 y 2015), debido a que las imágenes compuestas tienen nubosidad superior al 20%, se tiene información de un año antes y varios años posteriores a las actividades de reforestación. En el año 2013 se observa una cobertura de vegetación de 1.07% y el índice SAVI de 0.1121 con precipitación de 3mm, mientras que en los años posteriores (2017 al 2022) se observa una cobertura de vegetación promedio de 19.15%, esto representa un

aumento del área con vegetación, sin embargo, en algunos años hubo muy poca o ninguna cobertura de vegetación. Por otra parte, se tiene un índice promedio SAVI de 0.1324 en los años posteriores a la intervención, lo que representa un aumento pequeño en la densidad de las plantas, además de que esto se dio con una precipitación promedio de 3.833mm.

En la Figura 35 se contrastan los resultados de lluvia y el índice SAVI, en donde se observa que después del 2017 hay aumento del índice SAVI a pesar de que las precipitaciones disminuyeron debido a la sequía del Cuerno de África, también es importante destacar que en el año siguiente a las actividades de reforestación se observa una disminución de la vegetación y el índice SAVI.

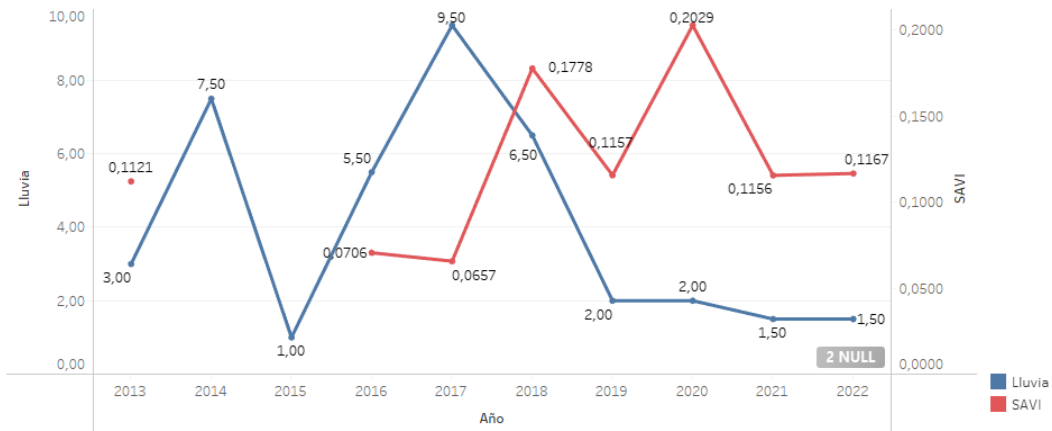


Fig. 35 Lluvias e índice SAVI del polígono KukuB

En la Figura 36 se contrastan los resultados de lluvia y porcentaje de cobertura vegetal, se observa fluctuación en la cobertura de vegetación y no se identifica un patrón acorde a las precipitaciones, sin embargo, se observa que en el año siguiente a las actividades de reforestación no se ve un incremento de la vegetación, pero en los años siguientes si se evidencia.

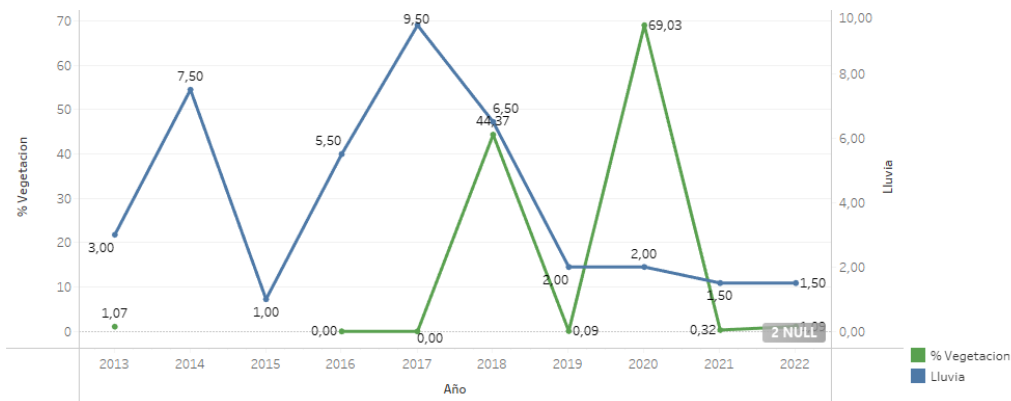


Fig. 36 Lluvias y % de cobertura vegetal del polígono KukuB

En la Figura 37 se contrastan los resultados de cobertura de vegetación y el índice SAVI. Al igual que en el polígono anterior se observa un patrón similar entre la cobertura de vegetación y el índice SAVI con algunas variaciones, las cuales se deben a la densidad de la vegetación, pues pueden existir áreas con vegetación más vigorosas que otras. También es importante destacar que en el año siguiente a las actividades de reforestación se observa una leve disminución del índice SAVI.

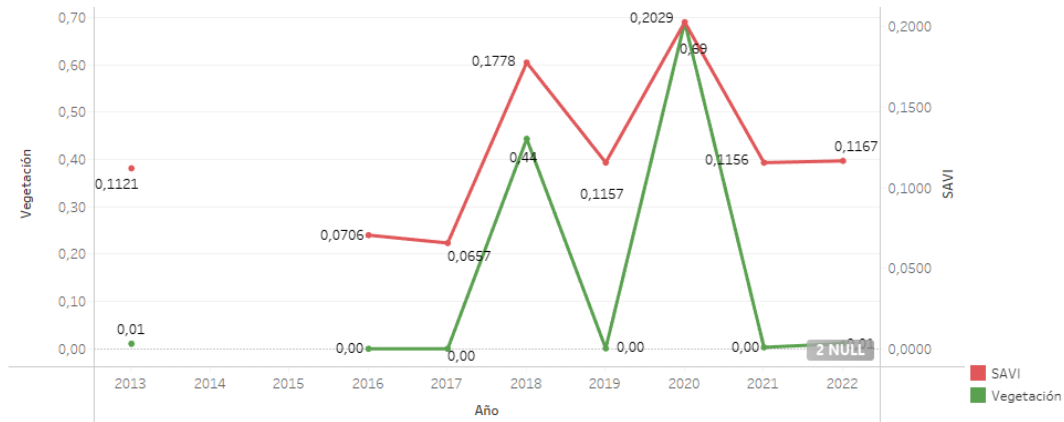


Fig. 37 índice SAVI y Cobertura vegetal del polígono KukuB

6.1.3. KukuD

En este polígono también se iniciaron actividades de reforestación en el 2016 por lo cual la información se genera desde el 2013 con el fin de identificar si la vegetación varía con las labores de Justdiggit. En la Tabla 11 se muestra la información obtenida.

TABLA 11

RESULTADOS KUKUD

Año	% Suelo	% Vegetación	% Faltante	SAVI	Diferencia SAVI	Lluvia	Observación
2013						3.115	Imagen excluida
2014						7.731	Imagen excluida
2015						1	Imagen excluida
2016	99.58	0.42	0.00	0.1099	-0.013	5.423	
2017						8.654	Imagen excluida
2018	24.22	75.78	0.00	0.2000	0.0956	6.231	
2019	99.93	0.07	0.00	0.1277	-0.0723	1.808	
2020	58.64	41.36	0.00	0.1847	0.0570	1.808	
2021	90.42	9.58	0.00	0.1483	-0.0364	1.308	
2022	98.68	1.32	0.00	0.1353	-0.0130	1.5	

Lastimosamente para este polígono se tienen cuatro imágenes excluidas (2013, 2014, 2015 y 2017), debido a que las imágenes compuestas tienen nubosidad superior al 20%, de las cuales tres corresponden a los años anteriores a las actividades de reforestación, por lo cual no se tienen valores de referencia para realizar la comparación de los cambios en la cobertura del área. En los años posteriores a la reforestación se observa una cobertura de vegetación promedio de 25.62%, el índice promedio SAVI es de 0.1592 y precipitación promedio de 2.5mm.

En la Figura 38 se contrastan los resultados de lluvia y el índice SAVI.

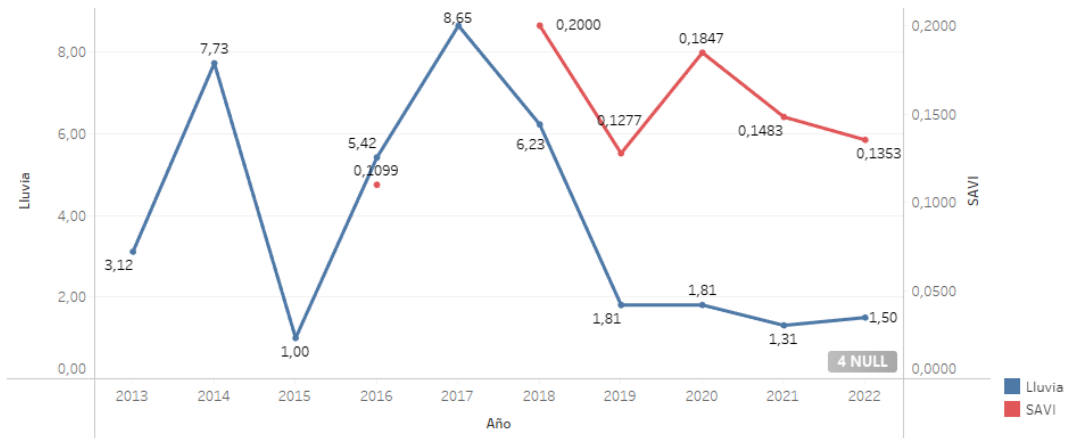


Fig. 38 Lluvias e índice SAVI del polígono KukuD

En la Figura 39 se contrastan los resultados de lluvia y porcentaje de cobertura vegetal.

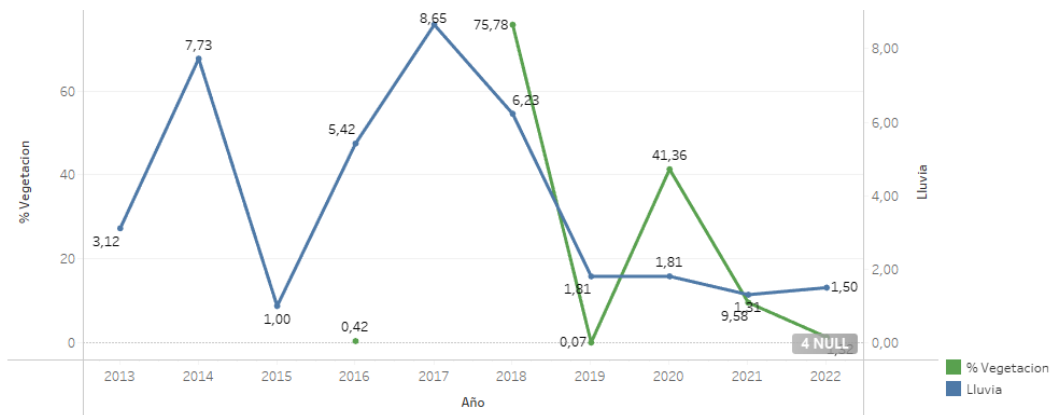


Fig. 39 Lluvias y % de cobertura vegetal del polígono KukuD

En la Figura 40 se contrastan los resultados de cobertura de vegetación y el índice SAVI. Al igual que en el polígono anterior se observa un patrón similar entre la cobertura de vegetación y el índice SAVI con algunas variaciones, las cuales se deben a la densidad de la vegetación, pues pueden existir áreas con vegetación más vigorosas que otras.

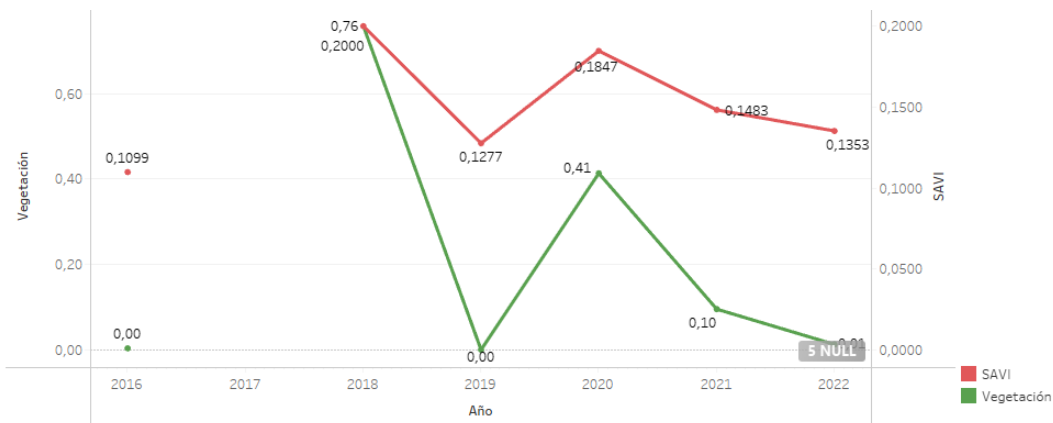


Fig. 40 índice SAVI y Cobertura vegetal del polígono KukuD

6.1.4. Nasipa

En este polígono se iniciaron actividades de reforestación en el 2018 por lo cual la información se genera desde el 2015 con el fin de identificar si la vegetación varía con las labores de Justdiggit. En la Tabla 12 se muestra la información obtenida.

TABLA 12
RESULTADOS NASIPA

Año	% Suelo	% Vegetación	% Faltante	SAVI	Diferencia SAVI	Lluvia	Observación
2015	80.74	19.26	0.00	0.1288	-0.0153	2.5	
2016						7.5	Imagen excluida
2017	98.52	1.48	0.00	0.1029	0.1029	11.5	
2018	0.00	100.00	0.00	0.2520	0.1491	8.5	
2019	100.00	0.00	0.00	0.0781	-0.1739	5	
2020	0.00	100.00	0.00	0.1830	0.1048	6.5	
2021						4	Imagen excluida
2022	100.00	0.00	0.00	0.1120	-0.0322	4	

A pesar de que se tienen dos años excluidos (2016 y 2021), debido a que las imágenes compuestas tienen nubosidad superior al 20%, se tiene información del antes y después de las actividades de reforestación, en donde se destaca que en los años previos (2015 y 2017) se observa una cobertura de vegetación promedio de 10.37% y el índice promedio SAVI de 0.1159 con precipitación promedio de 7mm, mientras que en los años posteriores (2019, 2020 y 2022) se observa una cobertura de vegetación promedio de 33.33%, esto representa aumento del área con vegetación, además el índice promedio SAVI es de 0.1244 lo que representa un aumento

pequeño en la densidad de las plantas; es de resaltar que esto se presentó una precipitación promedio de 5.167mm, inferior a los años anteriores a la reforestación.

En la Figura 41 se contrastan los resultados de lluvia y el índice SAVI, en donde se observa que después del 2017 hay aumento del índice SAVI a pesar de que las precipitaciones disminuyeron debido a la sequía del Cuerno de África.

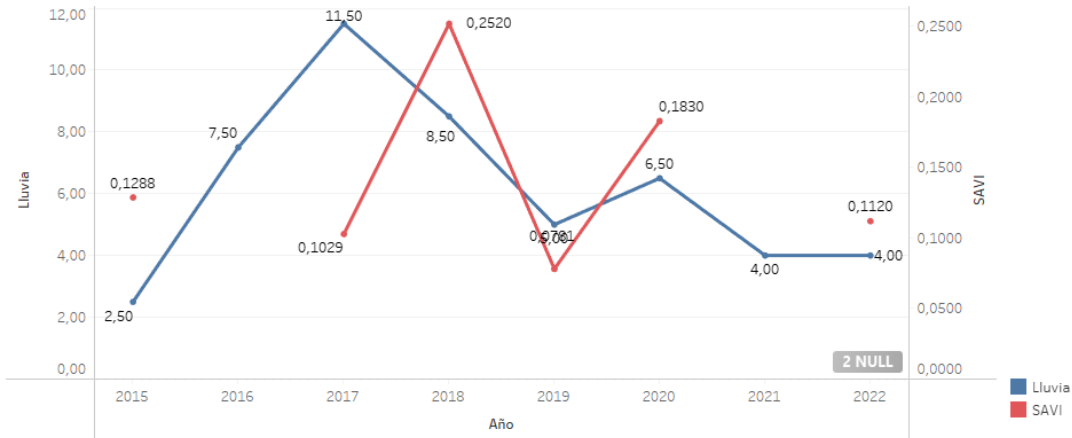


Fig. 41 Lluvias e índice SAVI del polígono Nasipa

En la Figura 42 se contrastan los resultados de lluvia y porcentaje de cobertura vegetal, se observa fluctuación en la cobertura de vegetación y no se identifica un patrón acorde a las precipitaciones.

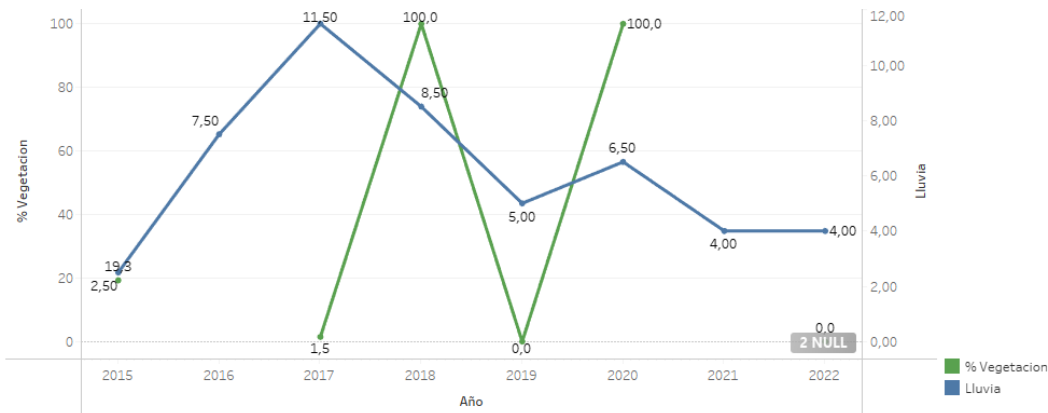


Fig. 42 Lluvias y % de cobertura vegetal del polígono Nasipa

En la Figura 43 se contrastan los resultados de cobertura de vegetación y el índice SAVI. Al igual que en el polígono anterior se observa un patrón similar entre la cobertura de vegetación y el índice SAVI con algunas variaciones, las cuales se deben a la densidad de la vegetación, pues pueden existir áreas con vegetación más vigorosas que otras. También es importante destacar

que en el año siguiente (2019) a las actividades de reforestación se observa una disminución de la vegetación y el índice SAVI respecto al año anterior (2017).

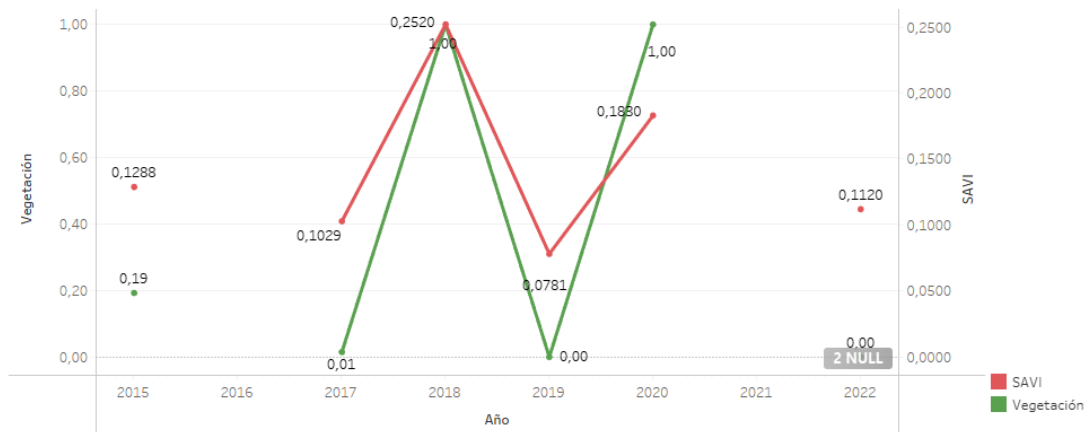


Fig. 43 índice SAVI y Cobertura vegetal del polígono Nasipa

6.1.5. KukuA

En este polígono también se iniciaron actividades de reforestación en el 2019 por lo cual la información se genera desde el 2016 con el fin de identificar si la vegetación varía con las labores de Justdiggit. En la Tabla 13 se muestra la información obtenida.

TABLA 13

RESULTADOS KUKUA

Año	% Suelo	% Vegetación	% Faltante	SAVI	Diferencia SAVI	Lluvia	Observación
2016	100.00	0.00	0.00	0.0868	-0.0087	5	
2017						7.5	Imagen excluida
2018	98.67	1.33	0.00	0.1280	0.0653	5.5	
2019	100.00	0.00	0.00	0.0956	-0.0324	1.5	
2020	92.18	7.82	0.00	0.1574	0.0618	2	
2021	96.71	3.29	0.00	0.1214	-0.0360	1.5	
2022	100.00	0.00	0.00	0.0795	-0.0419	1.5	

A pesar de que hay un año excluido (2017), debido a que la imagen compuesta tiene nubosidad superior al 20%, se tiene información de dos años antes y varios años posteriores a las actividades de reforestación, en donde se destaca que en los años previos (2016 y 2018) se observa una cobertura de vegetación promedio de 0.67% y el índice promedio SAVI de 0.1074 con precipitación promedio de 5.25mm, mientras que en los años posteriores (2020, 2021 y 2022) se observa una cobertura de vegetación promedio de 3.7%, esto representa aumento del área con

vegetación, el índice promedio SAVI es de 0.1194 lo que también representa aumento pequeño en la densidad de las plantas, a pesar de que esto se presentó con una precipitación promedio de 1.67mm, inferior a las precipitaciones de los años anteriores a la reforestación.

En la Figura 44 se contrastan los resultados de lluvia y el índice SAVI, en donde se observa que después del 2019 hay aumento del índice SAVI a pesar de que las precipitaciones disminuyeron debido a la sequía del Cuerno de África.

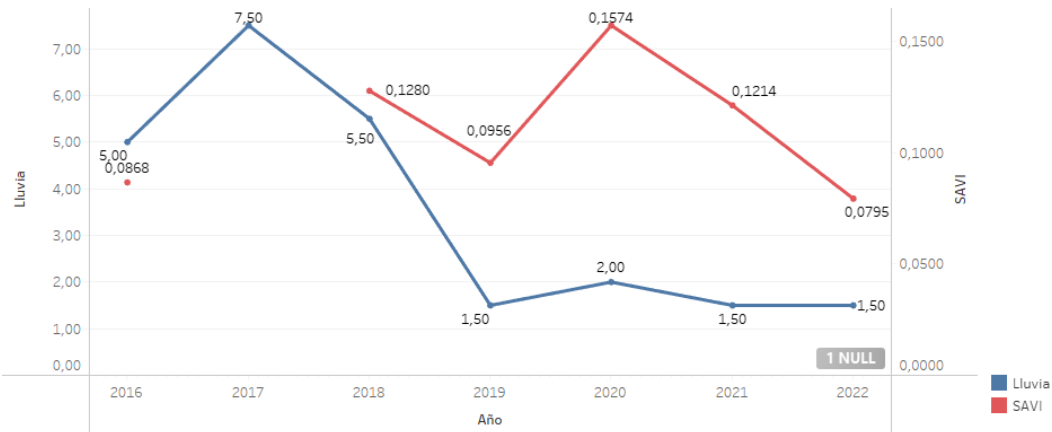


Fig. 44 Lluvias e índice SAVI del polígono KukuA

En la Figura 45 se contrastan los resultados de lluvia y porcentaje de cobertura vegetal, se observa un aumento y disminución en la cobertura de vegetación posiblemente causado por la disminución de las precipitaciones que se dieron a partir del 2019.

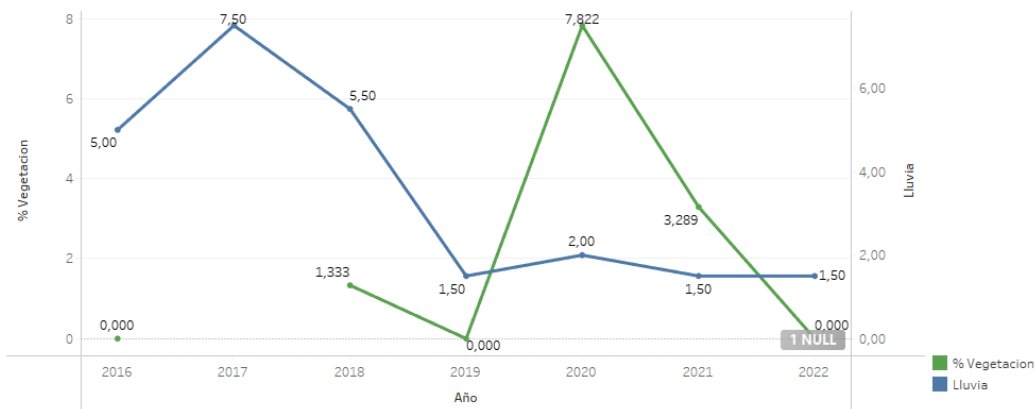


Fig. 45 Lluvias y % de cobertura vegetal del polígono KukuA

En la Figura 46 se contrastan los resultados de cobertura de vegetación y el índice SAVI. Al igual que en el polígono anterior se observa un patrón similar entre la cobertura de vegetación y el índice SAVI con algunas variaciones, las cuales se deben a la densidad de la vegetación, pues pueden existir áreas con vegetación más vigorosas que otras.



Fig. 46 índice SAVI y Cobertura vegetal del polígono KukuA

6.1.6. Risa

En este polígono también se iniciaron actividades de reforestación en el 2019 por lo cual la información se genera desde el 2016 con el fin de identificar si la vegetación varía con las labores de Justdiggit. En la Tabla 14 se muestra la información obtenida.

TABLA 14

RESULTADOS RISA

Año	% Suelo	% Vegetación	% Faltante	SAVI	Diferencia SAVI	Lluvia	Observación
2016						6.5	Imagen excluida
2017	43.99	56.01	0.00	0.127	0.1270	10.5	
2018	0.27	85.65	-14.08	0.485	0.3582	8.5	
2019	8.22	91.78	0.00	0.145	-0.3404	2.5	
2020	2.95	90.69	-6.36	0.207	0.0619	3	
2021						2.5	Imagen excluida
2022	46.82	53.18	0.00	0.122	-0.0697	3	

A pesar de que hay dos años excluidos (2016 y 2021), debido a que las imágenes compuestas tienen nubosidad superior al 20%, se tiene información de dos años antes y dos años posteriores a las actividades de reforestación, en donde se destaca que en los años previos (2017 y 2018) se observa una cobertura de vegetación promedio de 70.83% y el índice promedio SAVI de 0.3061 con precipitación promedio de 9.5mm, mientras que en los años posteriores (2020 y 2022) se observa una cobertura de vegetación promedio de 71.94%, esto representa un pequeño aumento del área con vegetación, el índice promedio SAVI es de 0.1642 lo que representa una disminución en la densidad de las plantas, a pesar de que esto se presentó con una precipitación promedio de 3mm.

En la Figura 47 se contrastan los resultados de lluvia y el índice SAVI, en donde se observa que a partir del 2019 hay disminución del índice SAVI posiblemente causado por la disminución de las precipitaciones.

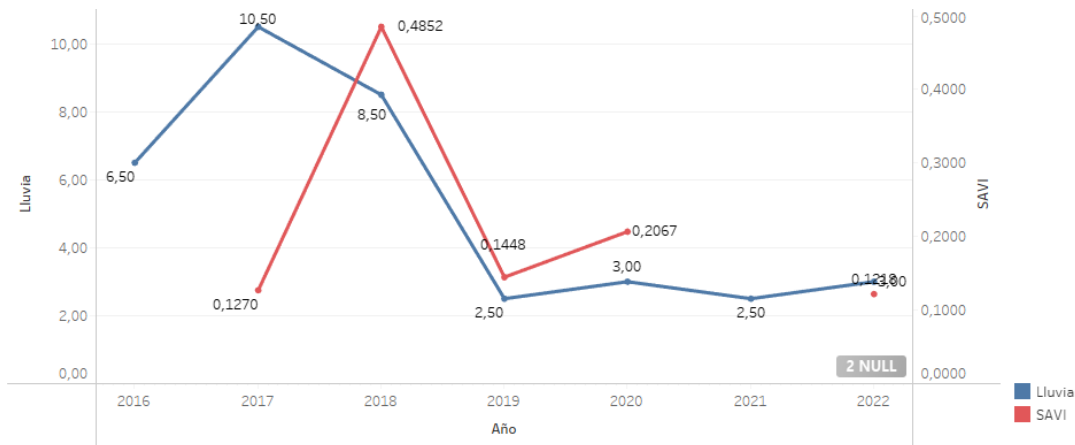


Fig. 47 Lluvias e índice SAVI del polígono Risa

En la Figura 48 se contrastan los resultados de lluvia y porcentaje de cobertura vegetal, se observa un aumento en la cobertura de vegetación en el 2018, el cual se mantuvo por dos años, a pesar de la disminución de las precipitaciones.

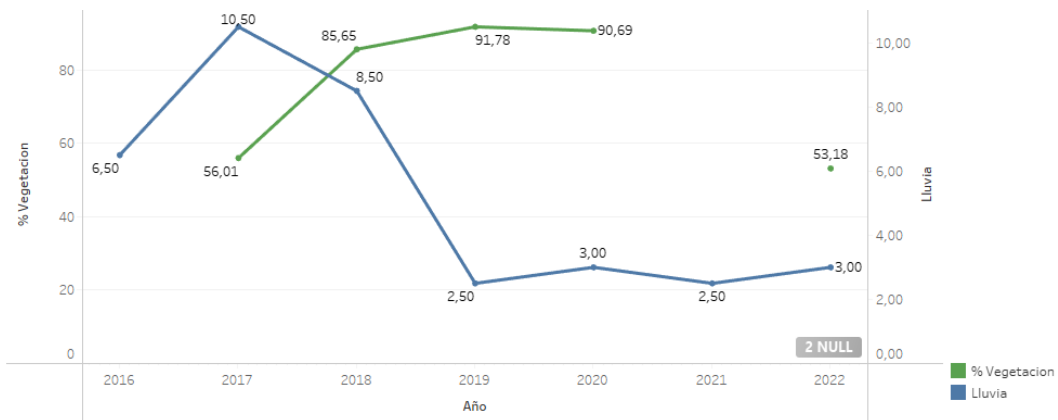


Fig. 48 Lluvias y % de cobertura vegetal del polígono Risa

En la Figura 49 se contrastan los resultados de cobertura de vegetación y el índice SAVI. No se observa un patrón similar entre la cobertura de vegetación y el índice SAVI.

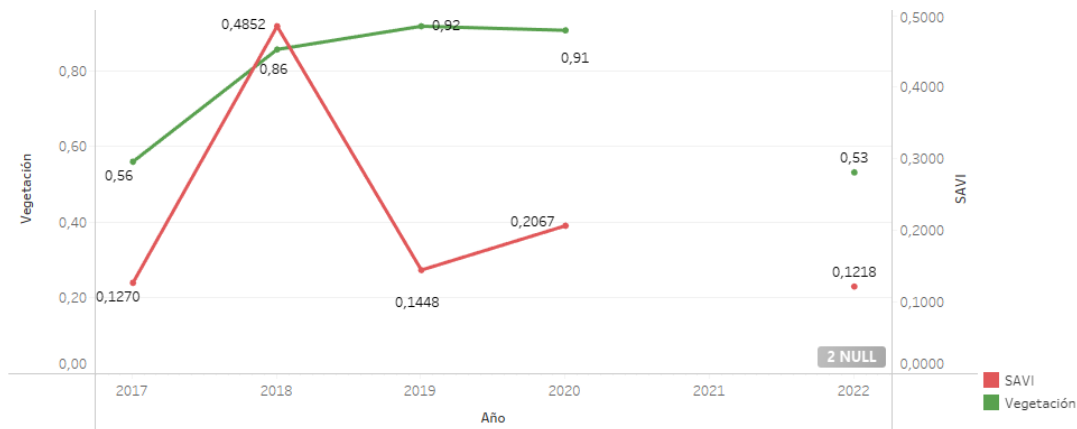


Fig. 49 índice SAVI y Cobertura vegetal del polígono Risa

6.2. Polígonos con semicírculos de suelo con siembra

En esta sección se relacionan las áreas objeto de estudio en las cuales se realizó reforestación con semicírculos de suelo con siembra.

6.2.1. Pembamoto

En este polígono se iniciaron actividades de reforestación en el 2018 por lo cual la información se genera desde el 2015 con el fin de identificar si la vegetación varía con las labores de Justdiggit. En la Tabla 15 se muestra la información obtenida.

TABLA 15

RESULTADOS PEMBAMOTO

Año	% Suelo	% Vegetación	% Faltante	SAVI	Diferencia SAVI	Lluvia	Observación
2015	0.00	100.00	0.00	0.13889	-0.03589		Sin información de precipitaciones
2016	0.00	100.00	0.00	0.17364	0.034752	2.5	
2017	0.00	100.00	0.00	0.14089	-0.032749	10	
2018	1.78	98.22	0.00	0.17941	0.038523	1	
2019	1.00	99.00	0.00	0.22017	0.040763	0.5	
2020	0.00	100.00	0.00	0.31158	0.091403	1	
2021	0.11	99.89	0.00	0.23447	-0.077105	1.5	
2022	2.78	90.88	-6.34	0.26097	0.026496	1.5	

Para este polígono no se tienen imágenes faltantes. Se destaca que en los años previos (2015, 2016 y 2017) se observa una cobertura de vegetación promedio de 100% y el índice promedio SAVI de 0.1511 con precipitación promedio de 6.25mm, lo que indica que en toda el área hay

vegetación pero es poco densa, mientras que en los años posteriores (2019, 2020, 2021 y 2022) se observa una cobertura de vegetación promedio de 97.44%, esto representa una pequeña disminución del área con vegetación, sin embargo, el índice promedio SAVI es de 0.2568 lo que significa que la densidad de la vegetación aumento, además de que esto se dio con una precipitación promedio inferior de 1.125mm.

En la Figura 50 se contrastan los resultados de lluvia y el índice SAVI, llama la atención que después de la intervención de Justdiggit se ve un aumento significativo del índice SAVI a pesar de la baja cantidad de precipitaciones. Esto indica que la técnica de reforestación tiene efectos positivos, a pesar de la sequía en el Cuerno de África.

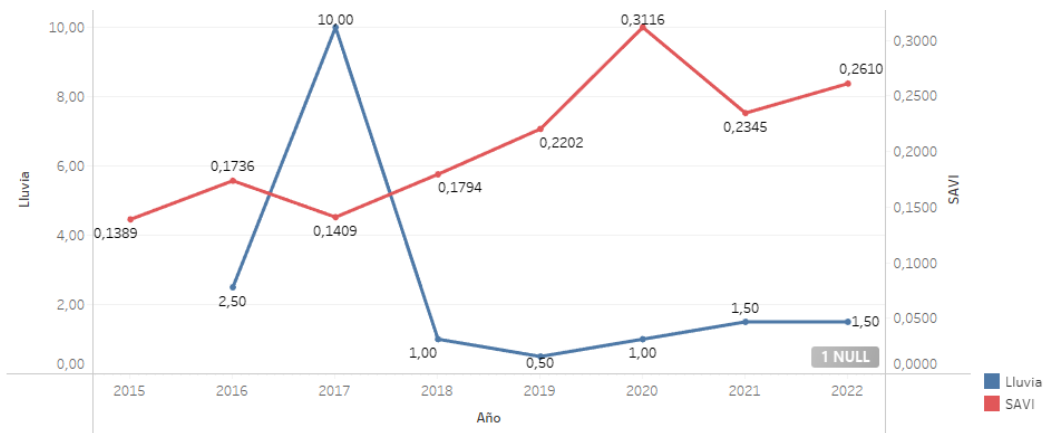


Fig. 50 Lluvias e índice SAVI del polígono Pembamoto

En la Figura 51 se contrastan los resultados de lluvia y porcentaje de cobertura vegetal, al igual que en la gráfica anterior, se ve un aumento significativo de la cobertura de vegetación con una cantidad baja de precipitaciones, sin embargo, esta se reduce la disminución prolongada de las lluvias, resalta la diferencia en la cobertura de vegetación.

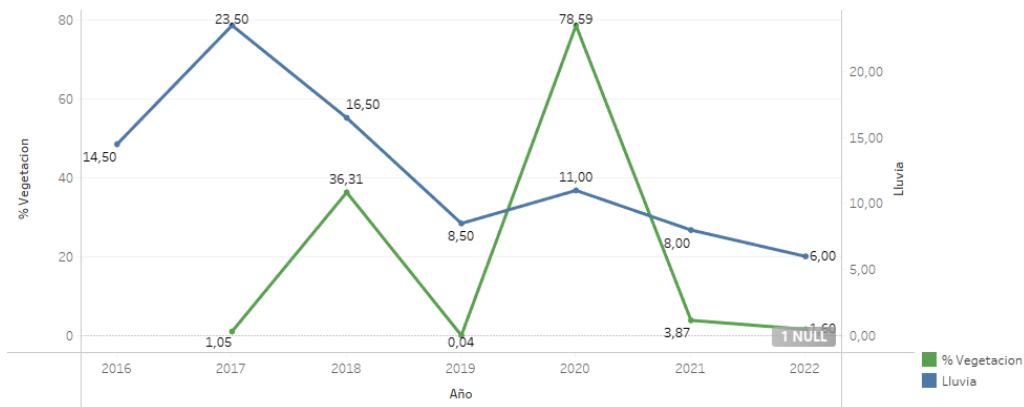


Fig. 51 Lluvias y % de cobertura vegetal del polígono Pembamoto

En la Figura 52 se contrastan los resultados de cobertura de vegetación y el índice SAVI. Se observa un patrón similar entre las dos variables con algunas variaciones, ya que esta última indica la densidad de la vegetación, este no necesariamente maneja las mismas curvas que la cobertura de vegetación, pues pueden existir áreas con vegetación más vigorosas que otras.

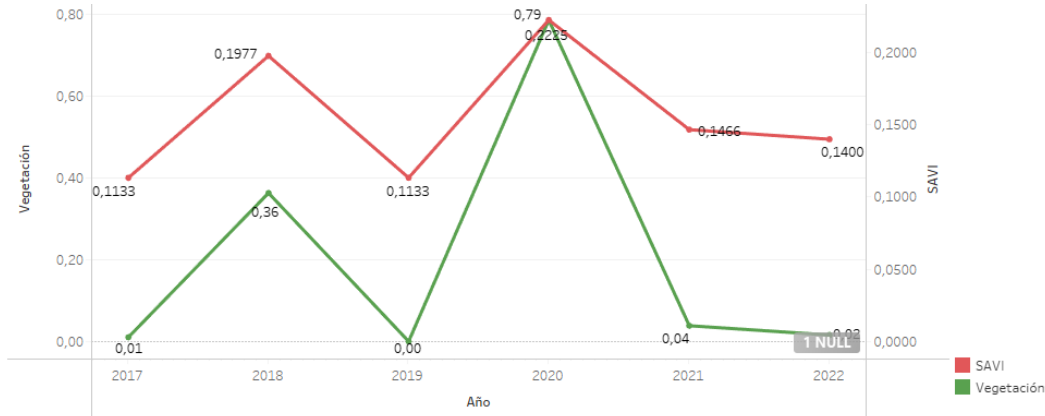


Fig. 52 índice SAVI y Cobertura vegetal del polígono Pembamoto

6.2.2. Enkii

En este polígono se iniciaron actividades de reforestación en el 2019 por lo cual la información se genera desde el 2016 con el fin de identificar si la vegetación varía con las labores de Justdiggit. En la Tabla 16 se muestra la información obtenida.

TABLA 16
RESULTADOS ENKII

Año	% Suelo	% Vegetación	% Faltante	SAVI	Diferencia SAVI	Lluvia	Observación
2016						14.5	Imagen excluida
2017	94.36	1.05	-4.59	0.1133	-0.0052	23.5	
2018	63.69	36.31	0.00	0.1977	0.0844	16.5	
2019	99.96	0.04	0.00	0.1133	-0.0843	8.5	
2020	21.41	78.59	0.00	0.2225	0.1092	11	
2021	96.13	3.87	0.00	0.1466	-0.0759	8	
2022	98.40	1.60	0.00	0.1400	-0.0066	6	

A pesar de que se tiene un año excluido (2016), debido a que la imagen compuesta tiene nubosidad superior al 20%, se tiene información del antes y después de las actividades de reforestación, en donde se destaca que en los años previos (2017 y 2018) se observa una cobertura de vegetación promedio de 18.68% y el índice promedio SAVI de 0.1555 con

precipitación promedio de 20mm, mientras que en los años posteriores (2020, 2021 y 2022) se observa una cobertura de vegetación promedio de 28.02%, esto representa un aumento del área con vegetación, el índice promedio SAVI es de 0.1697 también representa un aumento en la densidad de las plantas, además de que esto se dio con una precipitación promedio inferior de 8.33mm.

En la Figura 53 se contrastan los resultados de lluvia y el índice SAVI, llama la atención que después de la intervención de Justdiggít se ve un aumento significativo del índice SAVI con una cantidad baja de precipitaciones, sin embargo, este se reduce la disminución prolongada de las lluvias. Esto indica que la técnica de reforestación tiene efectos positivos, a pesar de la sequía en el Cuerno de África.

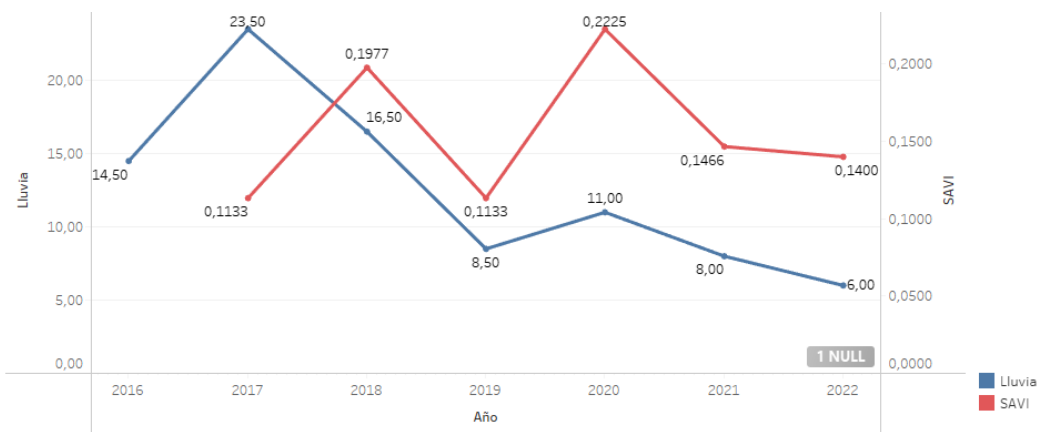


Fig. 53 Lluvias e índice SAVI del polígono Enkii

En la Figura 54 se contrastan los resultados de lluvia y porcentaje de cobertura vegetal, al igual que en la gráfica anterior, se ve un aumento significativo de la cobertura de vegetación con una cantidad baja de precipitaciones, sin embargo, esta se reduce la disminución prolongada de las lluvias, resalta la diferencia en la cobertura de vegetación.



Fig. 54 Lluvias y % de cobertura vegetal del polígono Enkii

En la Figura 55 se contrastan los resultados de cobertura de vegetación y el índice SAVI. Se observa un patrón similar entre las dos variables con algunas variaciones, ya que esta última indica la densidad de la vegetación y pues pueden existir áreas con vegetación más vigorosas que otras.

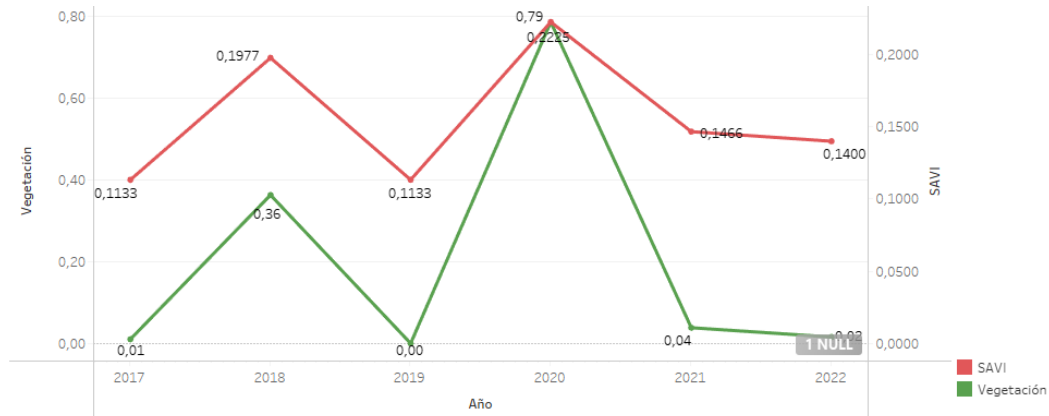


Fig. 55 índice SAVI y Cobertura vegetal del polígono Enkii

6.2.3. Baraka

En este polígono se iniciaron actividades de reforestación en el 2021 por lo cual la información se genera desde el 2018 con el fin de identificar si la vegetación varía con las labores de Justdiggit. En la Tabla 17 se muestra la información obtenida.

TABLA 17

RESULTADOS BARAKA

Año	% Suelo	% Vegetación	% Faltante	SAVI	Diferencia SAVI	Lluvia	Observación
2018						3	Imagen excluida
2019	99.71	0.29	0.00	0.1442	-0.0326	1.5	
2020						1.5	Imagen excluida
2021	100.00	0.00	0.00	0.1320	-0.0875	1.5	
2022						2.5	Imagen excluida

Lastimosamente para este polígono se tienen tres imágenes excluidas (2018, 2020 y 2022), debido a que las imágenes compuestas tienen nubosidad superior al 20%, de las cuales dos corresponden a los años anteriores a las actividades de reforestación y la otra correspondiente

al único año posterior, por lo cual no se tienen valores para realizar la comparación de los cambios en la cobertura del área contra los valores de referencia. Se observa una cobertura de vegetación de 0.29%, el índice SAVI es de 0.1442 y precipitación de 1.5mm en el año anterior a las actividades de Justdiggít.

En la Figura 56 se contrastan los resultados de lluvia y el índice SAVI, en donde no se puede observar ningún patrón debido a la falta de información.

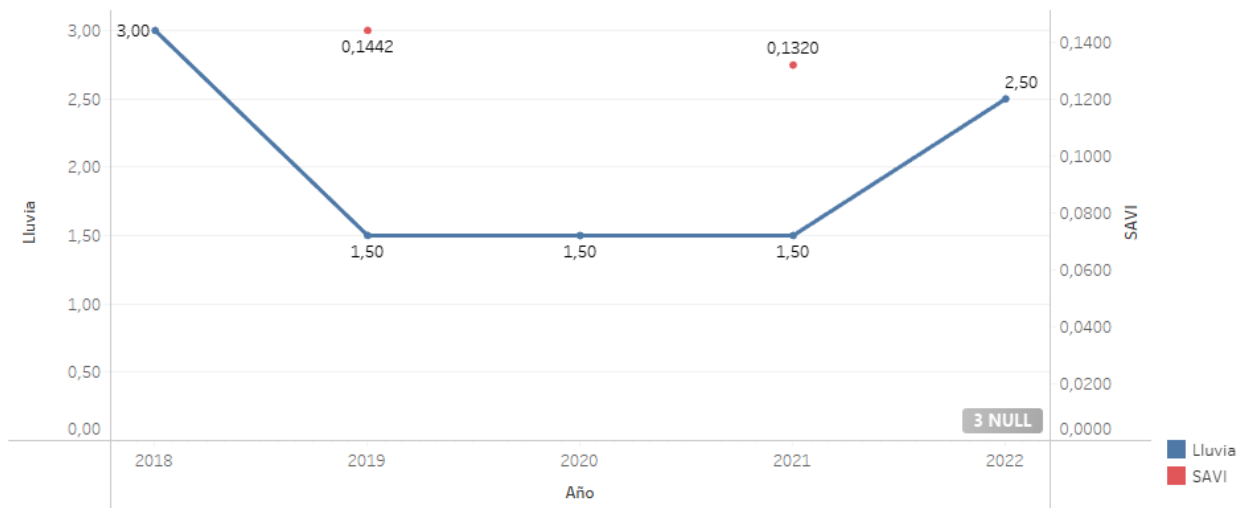


Fig. 56 Lluvias e índice SAVI del polígono Baraka

En la Figura 57 se contrastan los resultados de lluvia y porcentaje de cobertura vegetal, al igual que en la gráfica anterior, no se puede observar ningún patrón por falta de información.

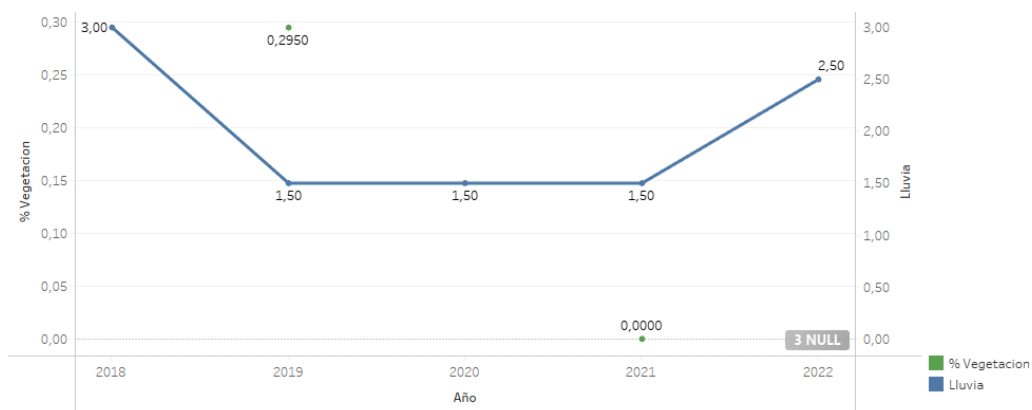


Fig. 57 Lluvias y % de cobertura vegetal del polígono Baraka

En la Figura 58 se contrastan los resultados de cobertura de vegetación y el índice SAVI. No se observa ningún patrón por falta de información.

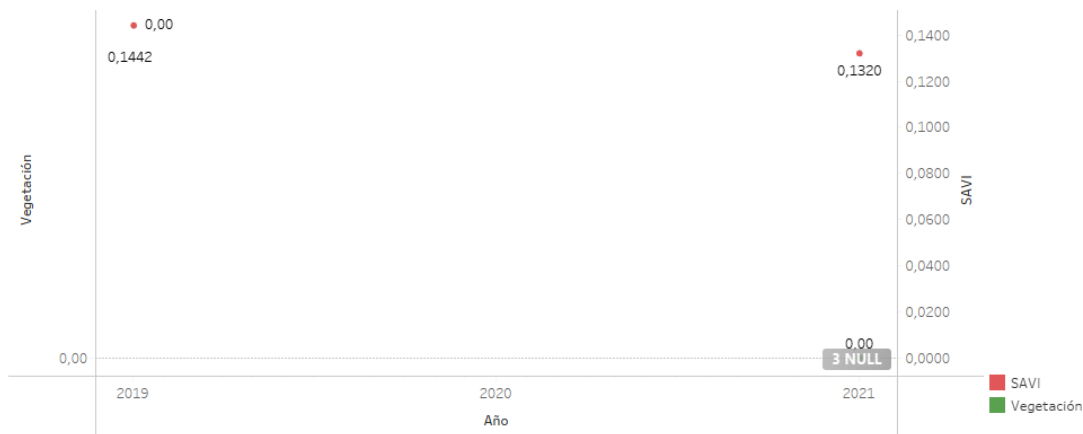


Fig. 58 índice SAVI y Cobertura vegetal del polígono Baraka

6.2.4. Engaruka

En este polígono se iniciaron actividades de reforestación en el 2021 por lo cual la información se genera desde el 2018 con el fin de identificar si la vegetación varía con las labores de Justdiggit. En la Tabla 18 se muestra la información obtenida.

TABLA 18

RESULTADOS ENGARUKA

Año	% Suelo	% Vegetación	% Faltante	SAVI	Diferencia SAVI	Lluvia	Observación
2018	13.78	86.22	0.00	0.1872	0.0749	0.5	
2019	96.79	3.21	0.00	0.1206	-0.0666		Sin información de precipitaciones
2020	4.81	95.19	0.00	0.2062	0.0856		Sin información de precipitaciones
2021	3.85	96.15	0.00	0.2093	0.0031		Sin información de precipitaciones
2022	88.14	11.86	0.00	0.1333	-0.0760	0.5	

Para este polígono no se tienen imágenes faltantes. Se destaca que en los años previos (2018, 2019 y 2020) se observa una cobertura de vegetación promedio de 61.54% y el índice promedio SAVI de 0.1713, mientras que en el año posterior (2019, 2020, 2021 y 2022) se observa una cobertura de vegetación de 11.86% e índice SAVI de 0.1333 lo que significa que tanto la cobertura como la densidad de la vegetación disminuyeron significativamente.

Debido a que no se cuenta con información de las precipitaciones de la mayoría de los años, no se contrastan las demás variables. En la Figura 59 se contrastan los resultados de cobertura de vegetación y el índice SAVI. Se observa un patrón similar entre las dos variables con algunas

variaciones, ya que esta última indica la densidad de la vegetación, este no necesariamente maneja las mismas curvas que la cobertura de vegetación, pues pueden existir áreas con vegetación más vigorosas que otras.

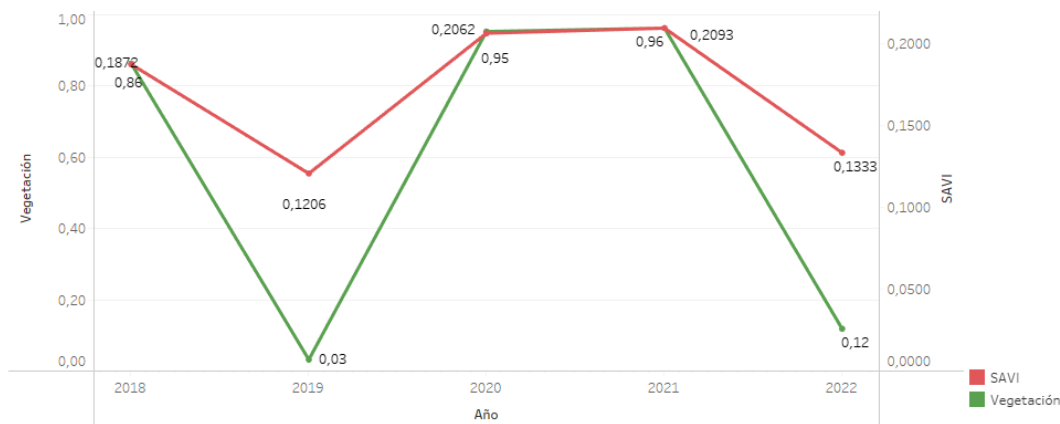


Fig. 59 índice SAVI y Cobertura vegetal del polígono Engaruka

6.2.5. EsilaleiSite1

En este polígono se iniciaron actividades de reforestación en el 2021 por lo cual la información se genera desde el 2018 con el fin de identificar si la vegetación varía con las labores de Justdiggit. En la Tabla 19 se muestra la información obtenida.

TABLA 19

RESULTADOS ESILALEISITE1

Año	% Suelo	% Vegetación	% Faltante	SAVI	Diferencia SAVI	Lluvia	Observación
2018	53.40	46.60	0.00	0.191	0.003418	3	
2019						1.5	Imagen excluida
2020						1	Imagen excluida
2021	75.11	24.89	0.00	0.143	-0.05467	1.5	
2022						3	Imagen excluida

Lastimosamente para este polígono se tienen tres imágenes excluidas (2019, 2020 y 2022), debido a que las imágenes compuestas tienen nubosidad superior al 20%, de las cuales dos corresponden a los años anteriores a las actividades de reforestación y la otra correspondiente al único año posterior, por lo cual no se tienen valores para realizar la comparación de los cambios en la cobertura del área contra los valores de referencia. Se observa una cobertura de vegetación de 46.6%, el índice SAVI es de 0.1913 y precipitación de 3mm en el año anterior a las actividades de Justdiggit.

En la Figura 60 se contrastan los resultados de lluvia y el índice SAVI, en donde no se puede observar ningún patrón debido a la falta de información.

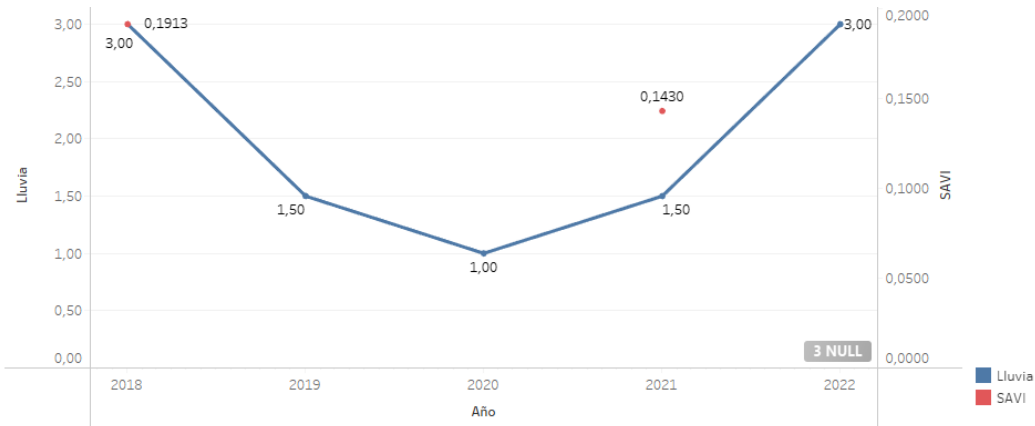


Fig. 60 Lluvias e índice SAVI del polígono EsilaleiSite1

En la Figura 61 se contrastan los resultados de lluvia y porcentaje de cobertura vegetal, al igual que en la gráfica anterior, no se puede observar ningún patrón por falta de información.

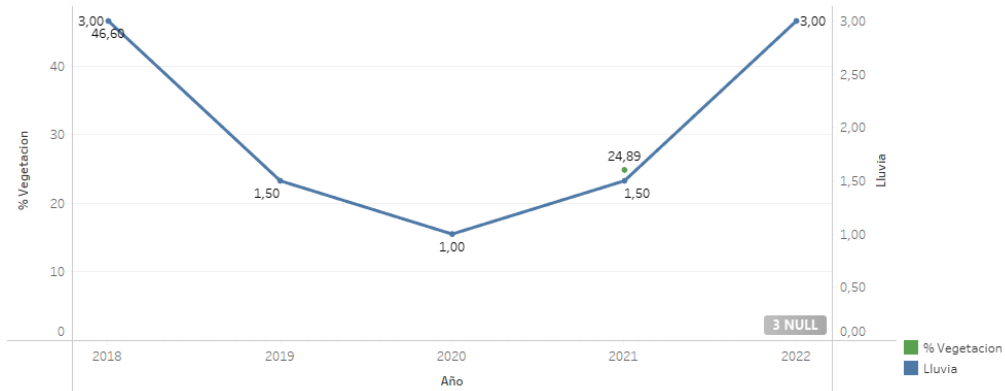


Fig. 61 Lluvias y % de cobertura vegetal del polígono EsilaleiSite1

En la Figura 62 se contrastan los resultados de cobertura de vegetación y el índice SAVI. No se observa ningún patrón por falta de información.

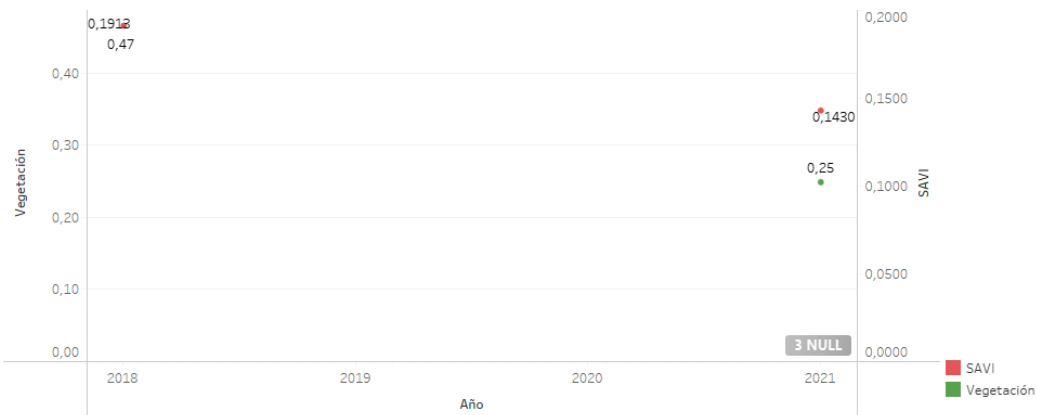


Fig. 62 índice SAVI y Cobertura vegetal del polígono EsilaleiSite1

6.2.6. EsilaleiSite2

En este polígono se iniciaron actividades de reforestación en el 2021 por lo cual la información se genera desde el 2018 con el fin de identificar si la vegetación varía con las labores de Justdiggit. En la Tabla 20 se muestra la información obtenida.

TABLA 20

RESULTADOS ESILALEISITE2

Año	% Suelo	% Vegetación	% Faltante	SAVI	Diferencia SAVI	Lluvia
2018	89.00	11.00	0.00	0.1358	0.0079	2
2019	97.51	2.49	0.00	0.1048	-0.0310	1
2020	20.38	74.63	-4.99	0.2086	0.1038	1
2021	75.95	24.05	0.00	0.1516	-0.0570	1
2022	81.82	1.47	-16.72	0.0762	-0.0754	2

Para este polígono no se tienen imágenes faltantes. Se destaca que en los años previos (2018, 2019 y 2020) se observa una cobertura de vegetación promedio de 29.37% y el índice promedio SAVI de 0.1497 con precipitación promedio de 1.333mm, mientras que en el año posterior (2022) se observa una cobertura de vegetación de 1.47% y el índice SAVI es de 0.0762 con una precipitación promedio inferior de 2mm, esto representa una disminución del área con vegetación posiblemente causada por la sequía prolongada en el cuerno de África.

En la Figura 63 se contrastan los resultados de lluvia y el índice SAVI, en donde no se observa un resultado positivo posterior a la reforestación.

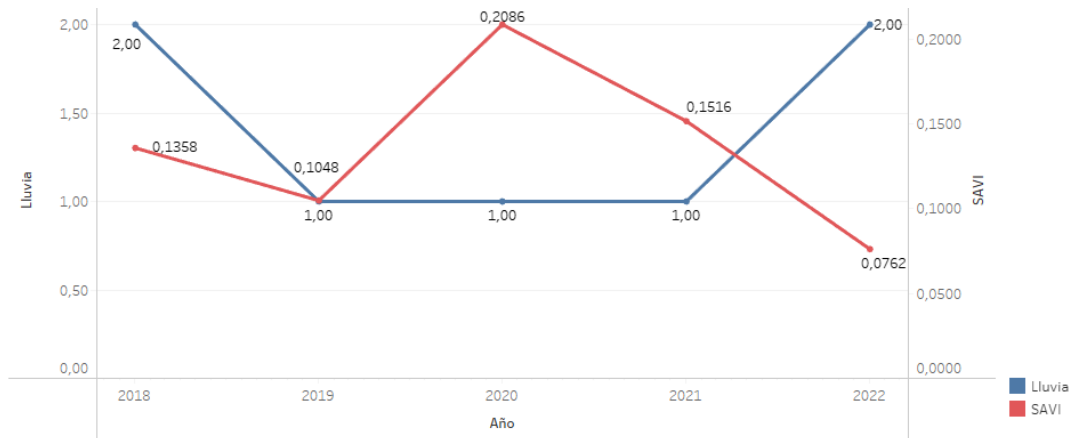


Fig. 63 Lluvias e índice SAVI del polígono EsilaleiSite2

En la Figura 64 se contrastan los resultados de lluvia y porcentaje de cobertura vegetal, al igual que en la gráfica anterior, no se observa un resultado positivo posterior a las actividades de reforestación.

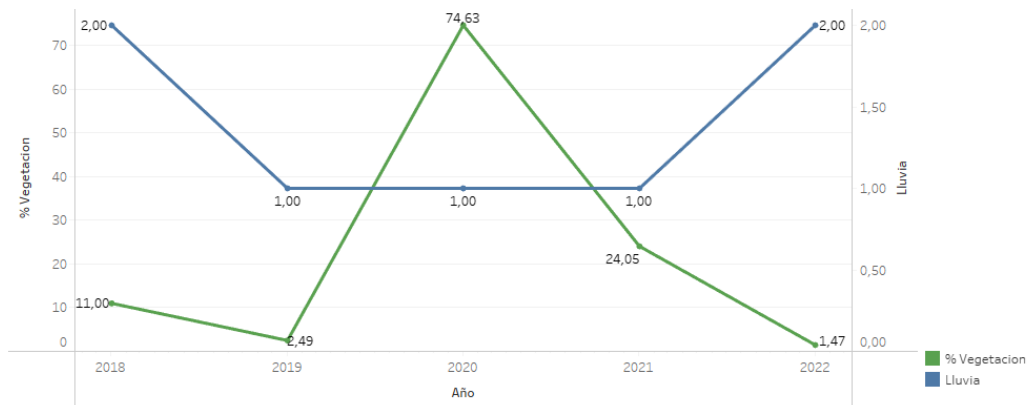


Fig. 64 Lluvias y % de cobertura vegetal del polígono EsilaleiSite2

En la Figura 65 se contrastan los resultados de cobertura de vegetación y el índice SAVI. Se observa un patrón similar entre las dos variables con algunas variaciones, ya que esta última indica la densidad de la vegetación, este no necesariamente maneja las mismas curvas que la cobertura de vegetación, pues pueden existir áreas con vegetación más vigorosas que otras.

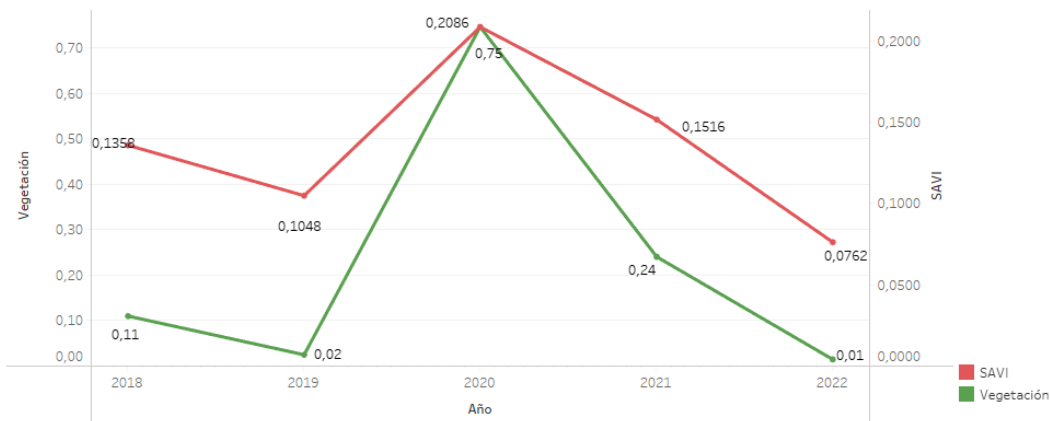


Fig. 65 índice SAVI y Cobertura vegetal del polígono EsilaleiSite2

6.2.7. Inkisanjani

En este polígono se iniciaron actividades de reforestación en el 2021 por lo cual la información se genera desde el 2018 con el fin de identificar si la vegetación varía con las labores de Justdiggit. En la Tabla 21 se muestra la información obtenida.

TABLA 21

RESULTADOS INKISANJANI

Año	% Suelo	% Vegetación	% Faltante	SAVI	Diferencia SAVI	Lluvia
2018	51.10	48.90	0.00	0.2213839	0.098572	17.9
2019	95.50	4.50	0.00	0.1418904	-0.079493	8.968
2020	32.84	67.16	0.00	0.2321792	0.090289	11.47
2021	86.65	13.35	0.00	0.1725078	-0.059671	8.468
2022	97.77	2.23	0.00	0.1400875	-0.03242	6.468

Para este polígono no se tienen imágenes faltantes. Se destaca que en los años previos (2018, 2019 y 2020) se observa una cobertura de vegetación promedio de 40.18% y el índice promedio SAVI de 0.1985 con precipitación promedio de 12.78mm, mientras que en el año posterior (2022) se observa una cobertura de vegetación promedio de 2.23%, esto representa una pequeña disminución del área con vegetación, el índice promedio SAVI es de 0.14, esto se dio con una precipitación promedio inferior de 6.468mm.

En la Figura 66 se contrastan los resultados de lluvia y el índice SAVI, en donde no se observa un resultado positivo posterior a la reforestación.

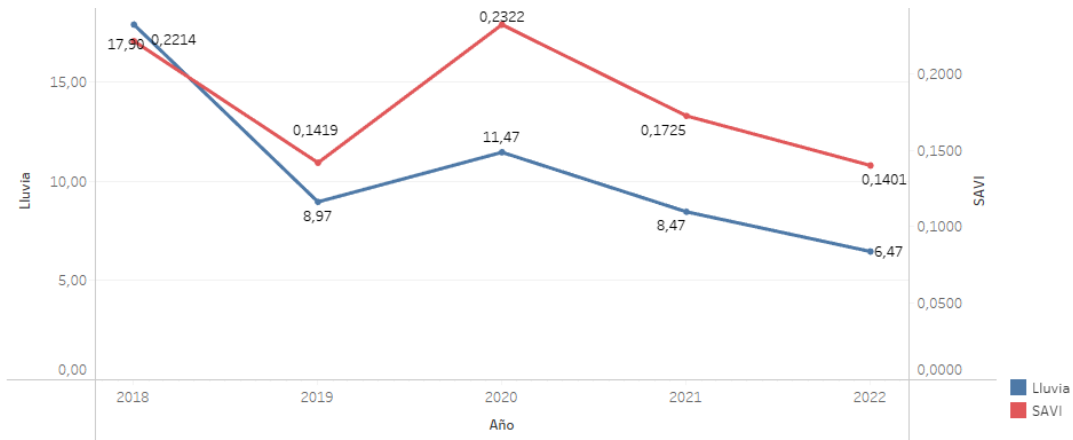


Fig. 66 Lluvias e índice SAVI del polígono Inkisanjani

En la Figura 67 se contrastan los resultados de lluvia y porcentaje de cobertura vegetal, al igual que en la gráfica anterior, no se observa un resultado positivo posterior a las actividades de reforestación.

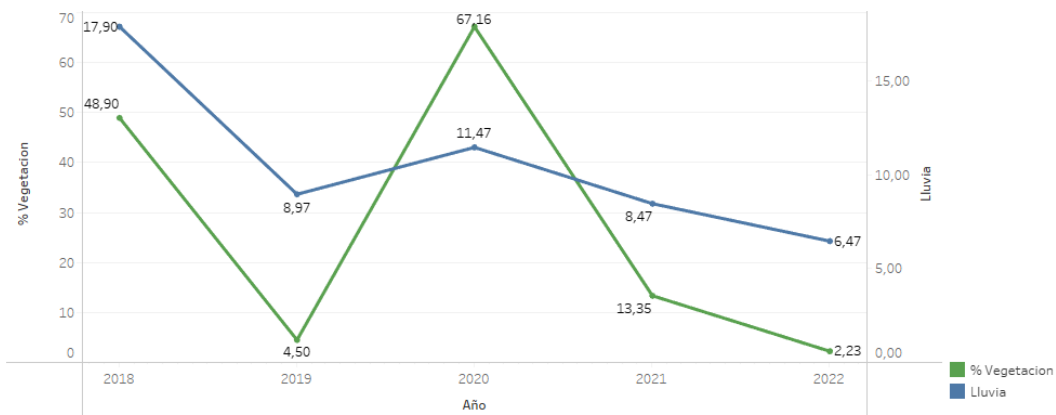


Fig. 67 Lluvias y % de cobertura vegetal del polígono Inkisanjani

En la Figura 68 se contrastan los resultados de cobertura de vegetación y el índice SAVI. Se observa un patrón similar entre las dos variables con algunas variaciones, ya que esta última indica la densidad de la vegetación, este no necesariamente maneja las mismas curvas que la cobertura de vegetación, pues pueden existir áreas con vegetación más vigorosas que otras.

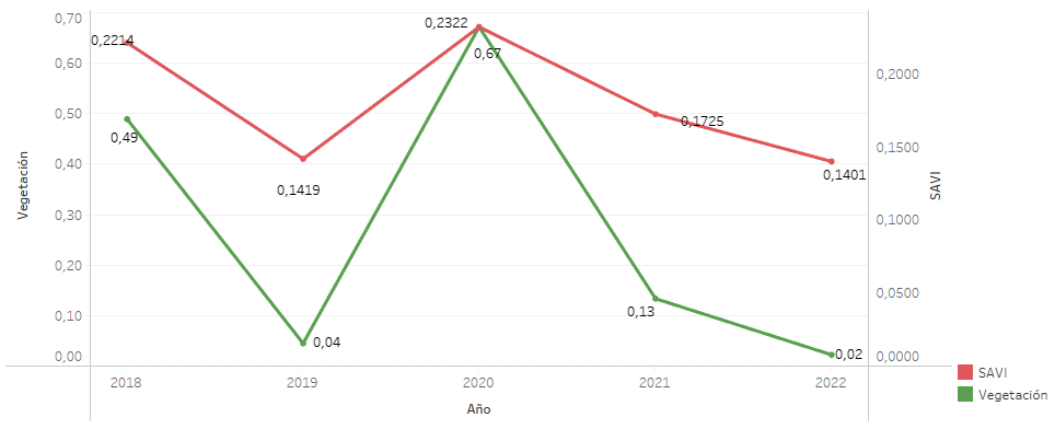


Fig. 68 índice SAVI y Cobertura vegetal del polígono Inkisanjani

6.2.8. Meserani

En este polígono se iniciaron actividades de reforestación en el 2021 por lo cual la información se genera desde el 2018 con el fin de identificar si la vegetación varía con las labores de Justdiggit. En la Tabla 22 se muestra la información obtenida.

TABLA 22

RESULTADOS MESERANI

Año	% Suelo	% Vegetación	% Faltante	SAVI	Diferencia SAVI	Lluvia	Observación
2018	95.12	4.88	0.00	0.1274	0.0030		
2019	100.00	0.00	0.00	0.1149	-0.0125		
2020							Imagen excluida
2021							Imagen excluida
2022	100.00	0.00	0.00	0.1144	0.1144		

A pesar de que se tiene dos años excluidos (2020 y 2021), debido a que la imagen compuesta tiene nubosidad superior al 20%, se tiene información del antes y después de las actividades de reforestación, además de esto no se cuenta con la información de las precipitaciones. Se destaca que en los años previos (2018 y 2019) se observa una cobertura de vegetación promedio de 2.44% y el índice promedio SAVI de 0.1212, mientras que en el año posterior (2022) se observa una cobertura de vegetación de 0% y el índice SAVI es de 0.1144 también representa una disminución.

Debido a que no se cuenta con información de las precipitaciones, esta no se contrastan las demás variables. En la Figura 69 se contrastan los resultados de cobertura de vegetación y el índice SAVI. Se observa un patrón similar entre las dos variables con algunas variaciones, ya

que esta última indica la densidad de la vegetación, este no necesariamente maneja las mismas curvas que la cobertura de vegetación, pues pueden existir áreas con vegetación más vigorosas que otras.

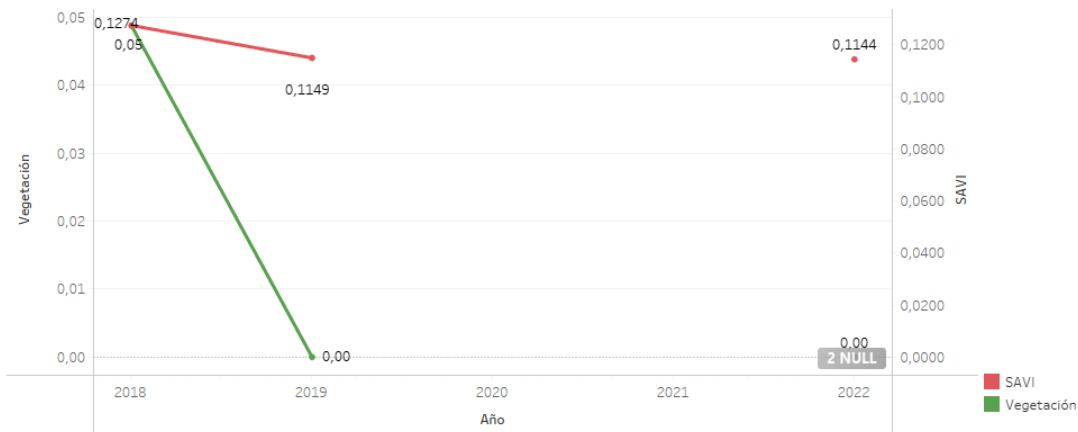


Fig. 69 índice SAVI y Cobertura vegetal del polígono Meserani

6.2.9. Selela

En este polígono se iniciaron actividades de reforestación en el 2021 por lo cual la información se genera desde el 2018 con el fin de identificar si la vegetación varía con las labores de Justdiggit. En la Tabla 23 se muestra la información obtenida.

TABLA 23

RESULTADOS SELELA

Año	% Suelo	% Vegetación	% Faltante	SAVI	Diferencia SAVI	Lluvia	Observación
2018						2	Imagen excluida
2019	86.47	13.53	0.00	0.1465	-0.0947	0.5	
2020						0.5	Imagen excluida
2021	26.33	73.67	0.00	0.1777	-0.0681	1	
2022						2	Imagen excluida

Lastimosamente para este polígono se tienen tres imágenes excluidas (2018, 2020 y 2022), debido a que las imágenes compuestas tienen nubosidad superior al 20%, de las cuales dos corresponden a los años anteriores a las actividades de reforestación y la otra correspondiente al único año posterior, por lo cual no se tienen valores para realizar la comparación de los cambios en la cobertura del área contra los valores de referencia. Se observa una cobertura de vegetación de 13.53%, el índice SAVI es de 0.1465 y precipitación de 0.5mm en el año anterior a las actividades de Justdiggit.

En la Figura 70 se contrastan los resultados de lluvia y el índice SAVI, en donde no se puede observar ningún patrón debido a la falta de información.

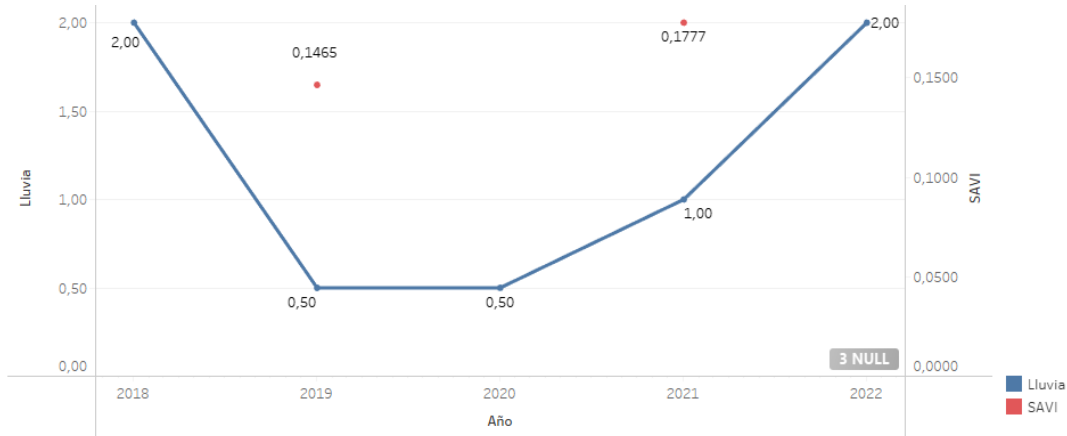


Fig. 70 Lluvias e índice SAVI del polígono Selela

En la Figura 71 se contrastan los resultados de lluvia y porcentaje de cobertura vegetal, al igual que en la gráfica anterior, no se puede observar ningún patrón por falta de información.

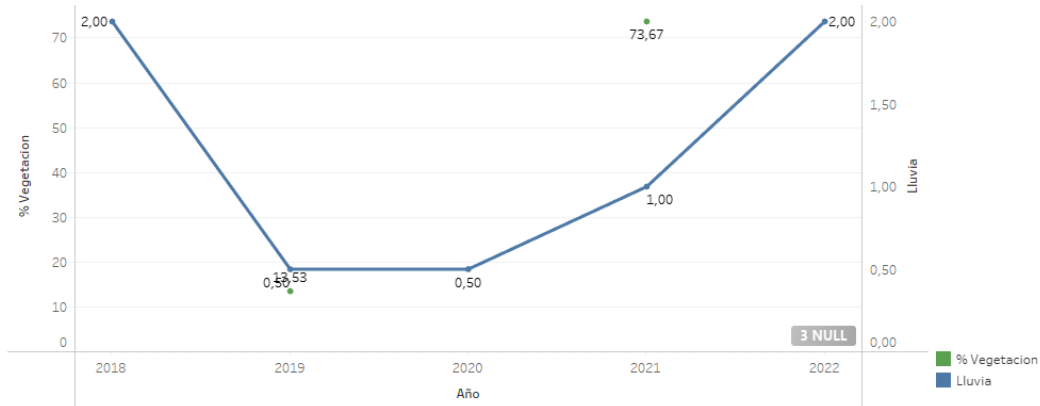


Fig. 71 Lluvias y % de cobertura vegetal del polígono Selela

En la Figura 72 se contrastan los resultados de cobertura de vegetación y el índice SAVI. No se observa ningún patrón por falta de información.

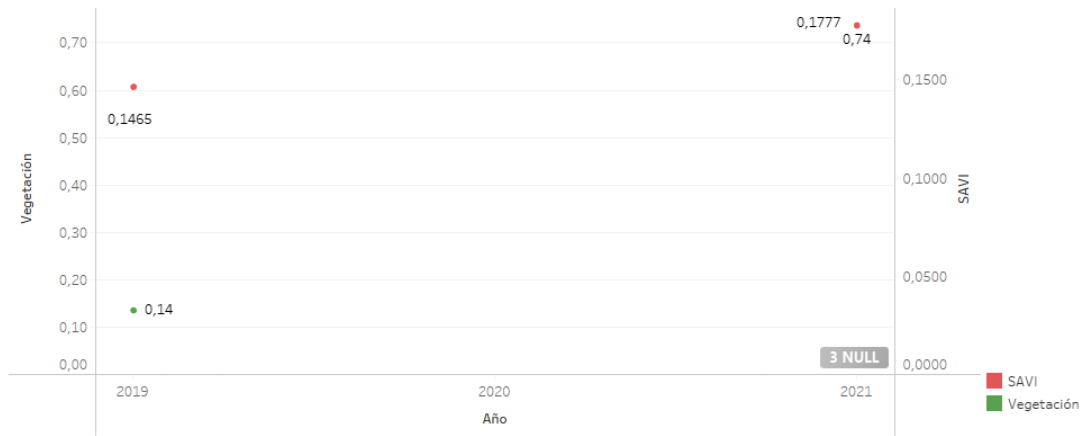


Fig. 72 índice SAVI y Cobertura vegetal del polígono Selea

7. Resultados y Discusión

Acorde a lo descrito en el capítulo seis, se evidencia que en algunos los polígonos no se puede realizar la comparación de la vegetación antes y después de las actividades de reforestación de Justdiggit. Por tal razón, se excluirán los siguientes polígonos en la evaluación de los resultados:

- KukuD: se excluyeron todas las imágenes de los años anteriores a las actividades de reforestación, debido a que tenían nubosidad superior a 20%. Por lo cual no se tienen valores de referencia para realizar la comparación de los cambios en la cobertura vegetal del área.
- Polígonos cuyas actividades de reforestación iniciaron en el año 2021: ya que se evalúa la cobertura vegetal de todos los polígonos hasta el año 2022 y que para los polígonos cuyas actividades de reforestación iniciaron en el año 2021, solo se tendría un año para evaluar los resultados. En otros polígonos se evidencia que en el año siguiente a las actividades de reforestación la vegetación y el índice SAVI disminuyen, pero en los años siguientes aumentan. Por tal razón se considera que para evaluar los resultados se requiere contar con mínimo tres años posteriores a la reforestación. Los polígonos que se excluyen por dicha razón son los siguientes:
 - Baraka
 - Engaruka
 - EsilaleiSite1
 - EsilaleiSite2
 - Inkisanjani
 - Meserani
 - Selela

A continuación, se relaciona la diferencia entre los promedios de los resultados de los años posteriores y anteriores a la reforestación de cada polígono, con el fin de identificar si hubo cambios y si estos fueron positivos o negativos.

TABLA 24

DIFERENCIA DE PROMEDIOS DE RESULTADOS DE REFORESTACIÓN CON
SEMICÍRCULOS SIN SIEMBRA

Polígono	% vegetación	SAVI	Lluvia
Inkiito	-6.72	0.0109	-1.417
KukuB	18.08	0.0203	0.833
Nasipa	22.96	0.0085	-1.833
KukuA	3.04	0.0120	-3.583
Risa	1.11	-0.1418	-6.5
Promedio	7.69	-0.0180	-2.5

TABLA 25

DIFERENCIA DE PROMEDIOS DE RESULTADOS DE REFORESTACIÓN CON
SEMICÍRCULOS CON SIEMBRA

Polígono	% vegetación	SAVI	Lluvia
Pembamoto	-2.56	0.1057	-5.13
Enkii	9.34	0.0142	-11.70
Promedio	3.39	0.0599	-8.41

En el grupo de polígonos en los cuales se realizaron semicírculos sin siembra, se identifica un aumento en porcentaje de áreas con vegetación. Sin embargo, se observa una disminución del índice SAVI. También se identifica que los resultados en general fueron positivos, aunque se presentaron menos precipitaciones que en los años previos a la reforestación.

En el grupo de polígonos en los cuales se realizaron semicírculos con siembra, se identifica un aumento en porcentaje de áreas con vegetación, aunque menor que en el primer grupo. Sin embargo, se observa aumento del índice SAVI. También se identifica que los resultados en general fueron positivos, aunque se presentaron menos precipitaciones que en los años previos a la reforestación y menores a las del primer grupo.

Es de resaltar que en todos los polígonos se evidencia una disminución significativa y prolongada de las precipitaciones a partir del año 2019. Como se muestra en la Figura 73.

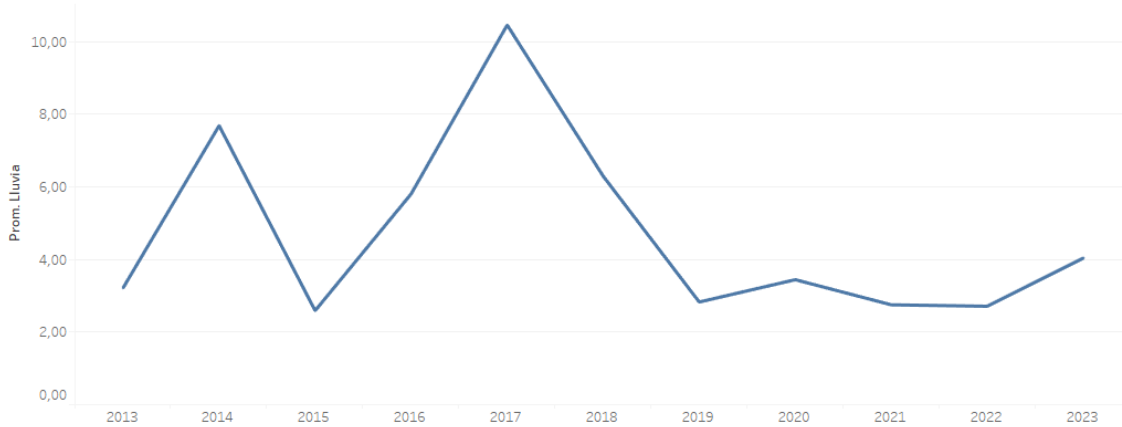


Fig. 73 Promedio precipitaciones de todos los polígonos

Teniendo en cuenta lo anterior y que los polígonos tuvieron diferentes años de reforestación lo cual puede hacer variar los resultados. A continuación, se comparan los resultados de polígonos reforestados el mismo año, pero con diferente técnica:

- 2018: En este año se realizaron las actividades de reforestación de los polígonos Pembamoto y Nasipa, el primero de ellos con siembra y el segundo sin siembra de semillas de pasto.

En la Figura 74 se contrastan los resultados de lluvia. A pesar de que hay información faltante de dos años, se observa que a nivel general Nasipa tuvo más precipitaciones que Pembamoto.

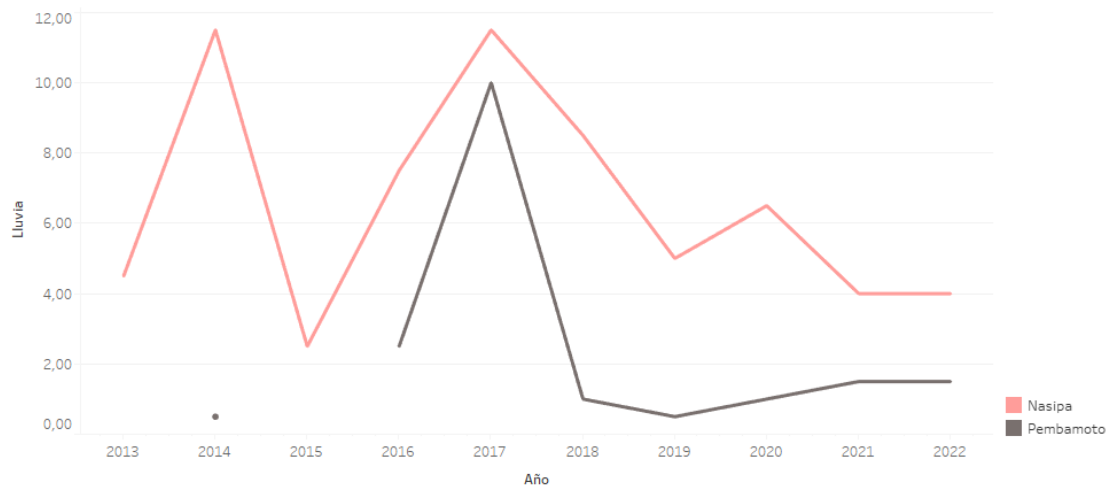


Fig. 74 Comparación de precipitaciones de polígonos reforestados en 2018

En la Figura 75 se contrastan los resultados de porcentaje de cobertura de vegetación. A pesar de que hay información faltante de un año posterior a la reforestación, se observa que a nivel general Pembamoto tuvo mayor porcentaje de cobertura de vegetación que Nasipa a partir del 2018.

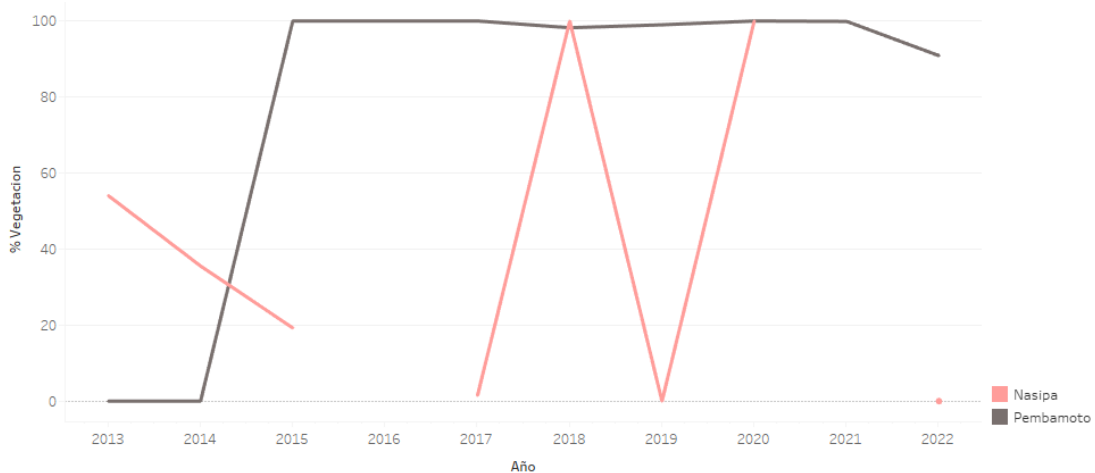


Fig. 75 Comparación de % de vegetación de polígonos reforestados en 2018

En la Figura 76 se contrastan los resultados del índice SAVI. A pesar de que hay información faltante de un año posterior a la reforestación, se observa que a nivel general Pembamoto tuvo mayor índice SAVI que Nasipa a partir del 2018.

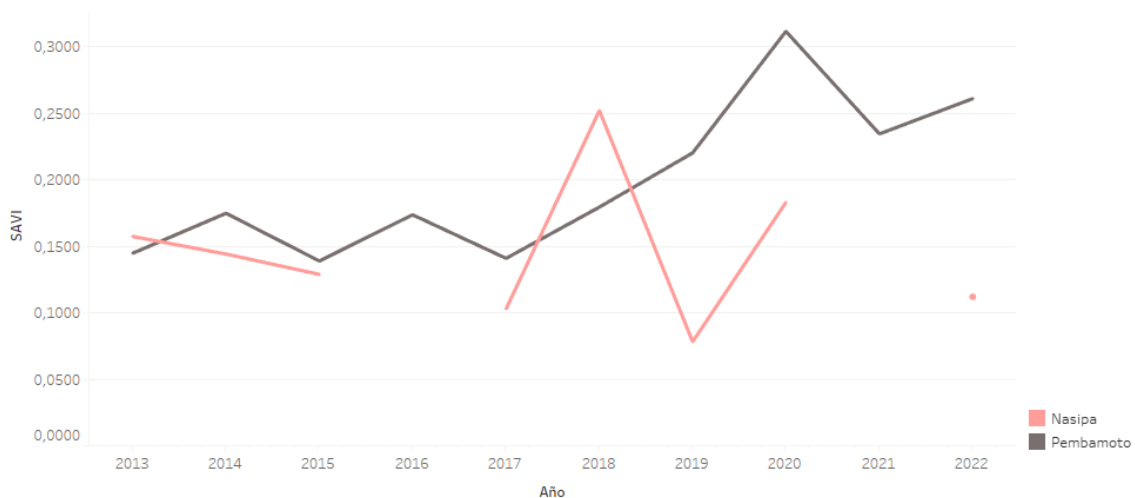


Fig. 76 Comparación del índice SAVI de polígonos reforestados en 2018

En la Tabla 26 se relaciona la diferencia entre los promedios de los resultados de los años posteriores y anteriores a la reforestación de cada polígono, con el fin de identificar si hubo cambios y si estos fueron positivos o negativos.

TABLA 26

DIFERENCIA DE PROMEDIOS DE RESULTADOS DE POLÍGONOS REFORESTADOS EN
2018

Polígono	Técnica	% vegetación	SAVI	Lluvia
Pembamoto	Semicírculos de suelo con siembra	-2.56	0.1057	-5.13
Nasipa	Semicírculos de suelo sin siembra	22.96	0.0085	-1.83

Acorde a lo anterior, el polígono que tuvo reforestación con siembra de semillas de pasto tuvo mejores resultados en la comparación anual de las variables % de vegetación y SAVI, con menos precipitaciones. Sin embargo, al validar la diferencia de promedio de resultados de los años anteriores y posteriores a la reforestación de cada polígono, se puede evidenciar que a pesar de que Pembamoto tuvo disminución de porcentaje de cobertura de vegetación, tuvo un aumento mayor el índice SAVI con menos precipitaciones en comparación con Nasipa.

Partiendo de lo anterior se concluye que, en el 2018 la técnica de reforestar con Semicírculos de suelo con siembra tiene mejores resultados.

- 2019: En este año se realizaron las actividades de reforestación de los polígonos Enkii, KukuA y Risa, el primero de ellos con siembra y los otros dos sin siembra de semillas de pasto.

En la Figura 77 se contrastan los resultados de lluvia. A pesar de que hay información faltante de un año, se observa que a nivel general Enkii tuvo más precipitaciones que Risa y KukuA.

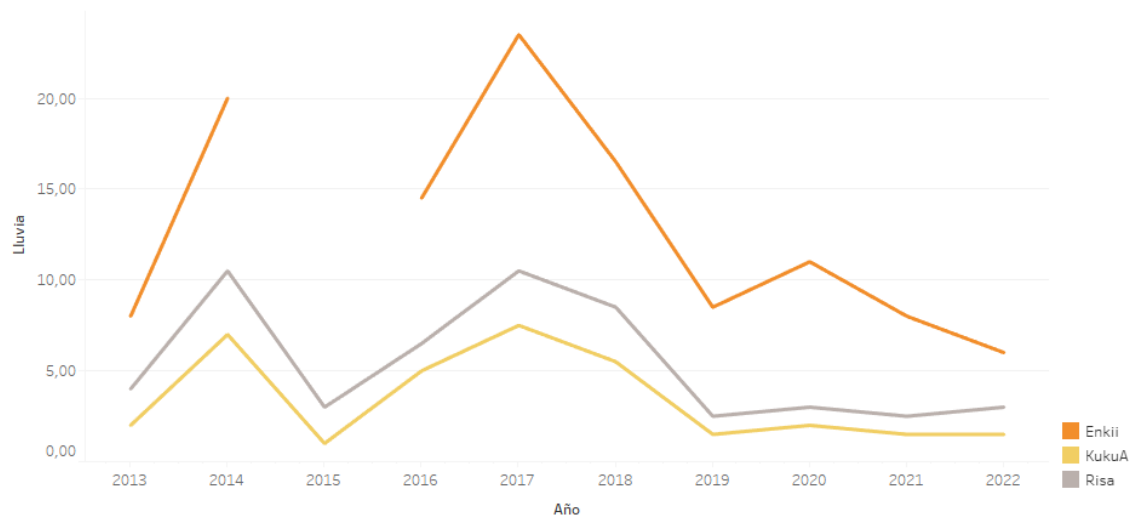


Fig. 77 Comparación de precipitaciones de polígonos reforestados en 2019

En la Figura 78 se contrastan los resultados de porcentaje de cobertura de vegetación. A pesar de que hay un dato faltante en los años posteriores al 2019, se observa que a nivel general Risa tuvo mayor porcentaje de cobertura de vegetación que Enkii y KukuA después de la reforestación. A su vez, Enkii tuvo mayor cobertura que KukuA.

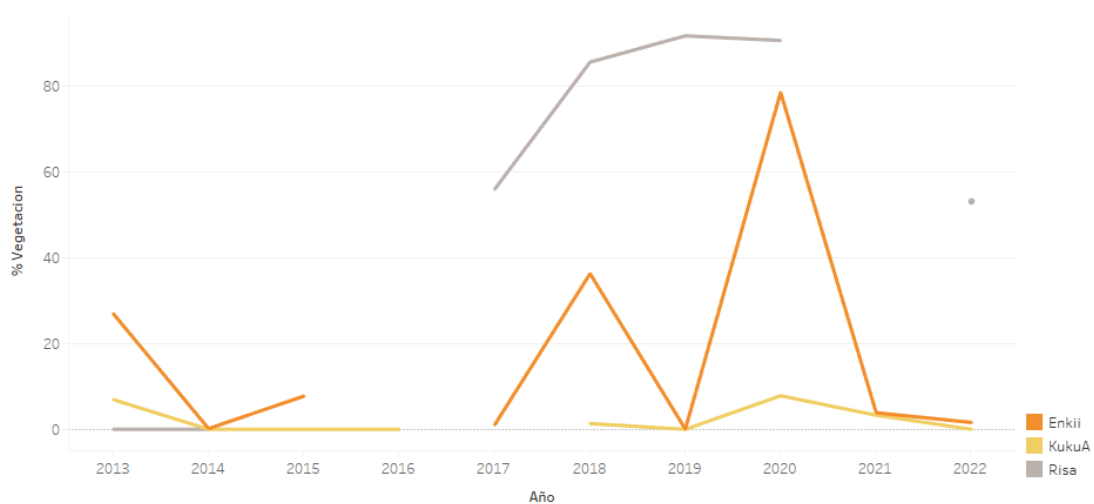


Fig. 78 Comparación de % de vegetación de polígonos reforestados en 2019

En la Figura 79 se contrastan los resultados del índice SAVI. A pesar de que hay un dato faltante en los años posteriores al 2019, se observa que a nivel general Enkii tuvo índice SAVI mayor que KukuA y Risa después de la reforestación.

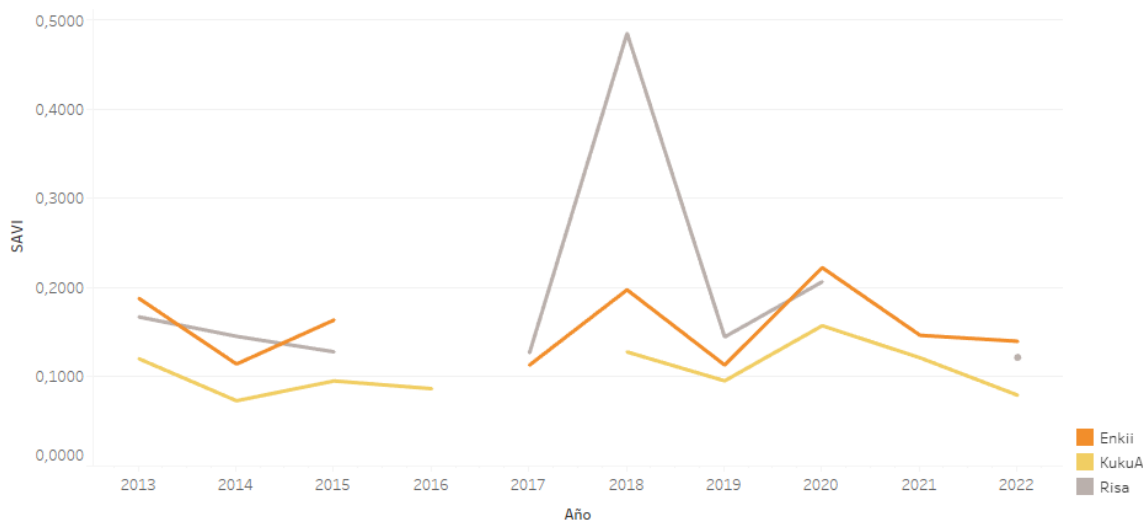


Fig. 79 Comparación del índice SAVI de polígonos reforestados en 2019

En la Tabla 27 se relaciona la diferencia entre los promedios de los resultados de los años posteriores y anteriores a la reforestación de cada polígono, con el fin de identificar si hubo cambios y si estos fueron positivos o negativos.

TABLA 27

DIFERENCIA DE PROMEDIOS DE RESULTADOS DE POLÍGONOS REFORESTADOS EN
2019

Polígono	Técnica	% vegetación	SAVI	Lluvia
Enkii	Semicírculos de suelo con siembra	9.34	0.0142	-11.70
Risa	Semicírculos de suelo sin siembra	1.11	-0.1418	-6.5
KukuA	Semicírculos de suelo sin siembra	3.04	0.0120	-3.583

Acorde a lo anterior, el polígono que tuvo reforestación con siembra de semillas de pasto tuvo mejores resultados en la comparación anual de las tres variables % de vegetación y SAVI, aunque con mayores precipitaciones. Sin embargo, al validar la diferencia de promedio de resultados de los años anteriores y posteriores a la reforestación de cada polígono, se puede evidenciar que a pesar de que tuvo aumento de porcentaje de cobertura de vegetación e índice SAVI con menos precipitaciones en comparación con Risa y KukuA.

Partiendo de lo anterior se concluye que, en el 2019 la técnica de reforestar con Semicírculos de suelo con siembra tiene mejores resultados.

8. Conclusiones

El objetivo de este trabajo es determinar cuál de las técnicas de reforestación implementadas por Justdiggít en áreas de Kenia y Tanzania es más efectiva. Dichas técnicas hacen referencia a cavar semicírculos en el suelo con y sin siembra de semillas en las áreas objeto de estudio, cuya reforestación tuvo inicio entre los años 2016 y 2021.

Para llevar a cabo este análisis se aplican técnicas de teledetección y aprendizaje automático supervisado, con el fin de cuantificar el cambio en la vegetación causado por las actividades de la ONG Justdiggít, desde el inicio de los proyectos de reforestación en dichas áreas hasta el 2022.

Se determina que para los polígonos cuya reforestación tuvo lugar en 2018 y 2019, en promedio el método de reforestación con siembra de semillas de pasto aumentó el índice SAVI en 0.06 y el porcentaje de cobertura de vegetación en 3.39%. Mientras que en promedio el método de reforestación sin siembra disminuye el índice SAVI en 0.04, este método aumenta el porcentaje de cobertura de vegetación en 9.04%. Lo cual permite comprobar que ambos métodos son efectivos, pero la técnica con siembra de semillas de pasto tiene mejores resultados en comparación con la técnica de semicírculos sin siembra. También se determina que las precipitaciones son determinantes en los resultados de la reforestación, ya que en ambos escenarios tuvieron resultados alterados debido a la sequía prolongada del Cuerno de África.

Si bien este proyecto se centra en obtener información de imágenes satelitales, la calidad de los datos es una importante consideración, debido a que se encuentra información faltante. Por ende, sería fructífero trabajar en proyectos futuros con imágenes con mayor resolución espacial y menor nubosidad, además de realizar la verificación en sitio de la cobertura de suelo de los puntos de entrenamiento. También sería ideal que en los trabajos futuros se contemplen más años en el estudio para evidenciar la evolución de la vegetación en las áreas reforestadas.

9. Referencias Bibliográficas

- [1] Lester. R. Brown, *El Mundo Al Borde Del Abismo, Cómo Evitar el Declive Ecológico y el Colapso de la Economía: Ensayo Ecológico y Económico.*, 1st ed. Cork: Centro de Estudios para el Desarrollo Sostenible, 2015.
- [2] J. Amestoy Alonso, *El Planeta Tierra en peligro: calentamiento global, cambio climático, soluciones.* Alicante: Club Universitario, 2013.
- [3] Jaime Martínez Urtaza, “Cambio climático y patógenos en el agua: el fenómeno de El Niño y su impacto en la salud”, *Rev Salud Ambient*, vol. 11, núm. 1–2, 2011.
- [4] E. Fuenzalida Labarca, *El lado oculto del cambio climático y el calentamiento global.* Santiago de Chile: RIL editores, 2020.
- [5] I. Gough, *Calentamiento Global, Codicia y Necesidades Humanas: Cambio Climático, Capitalismo y Bienestar Sostenible*, 1a ed. Montevideo: Miño y Dávila Editores, 2023.
- [6] EFE News Service, “Más de 9 millones de personas en riesgo de hambre por sequía en sur de África: ÁFRICA SEQUÍA”, *EFE News Service*. EFE News Services, Inc, Madrid, 2019. Consultado: el 5 de abril de 2024. [En línea]. Disponible en: <https://www.proquest.com/docview/2299489875?pq-origsite=primo&sourcetype=Wire%20Feeds>
- [7] EFE News Service, “ONU pide a países donantes más ayudas para combatir sequía en Cuerno África: ONU ÁFRICA”, *EFE News Service*. EFE News Services, Inc, Madrid, 2011. Consultado: el 5 de abril de 2024. [En línea]. Disponible en: ONU pide a países donantes más ayudas para combatir sequía en Cuerno África: ONU ÁFRICA
- [8] Justdigg, “Justdigg | Cooling Down The Planet | Global Warming Charity”. Consultado: el 6 de agosto de 2023. [En línea]. Disponible en: <https://justdigg.org/>
- [9] Google LLC y Servicio Geológico de EE. UU, “USGS Landsat 8 Level 2, Collection 2, Tier 1 | Earth Engine Data Catalog | Google for Developers”. Consultado: el 17 de marzo de 2024. [En línea]. Disponible en: https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LC08_C02_T1_L2
- [10] P. Chapman, *CRISP-DM 1.0: Step-by-step Data Mining Guide*. SPSS, 2000. [En línea]. Disponible en: <https://books.google.com.co/books?id=po7FtgAACAAJ>

- [11] E. Chuvieco, *Fundamentos de Teledetección Espacial*, Segunda edición. Madrid: Ediciones RIALP S.A., 1995.
- [12] E. Chuvieco y A. Huete, *Fundamentals of satellite remote sensing / Emilio Chuvieco and Alfredo Huete*. 2010. [En línea]. Disponible en: <https://search.ebscohost.com/login.aspx?direct=true&db=cat09062a&AN=cbj.oai.edge.tad.eo.folio.ebsco.com.fs00001093.a35417d3.6f2b.5249.80f1.dcc99390649d&site=eds-live>
- [13] P. M. Mather y M. Koch, *Computer Processing of Remotely-Sensed Images: An Introduction*, Fourth edition. Somerset: John Wiley & Sons, Incorporated, 2011. doi: 10.1002/9780470666517.
- [14] J. A. (John A. Richards y X. Jia, *Remote sensing digital image analysis: an introduction*, 2a ed. New York: Springer, 1993. doi: 10.1007/978-3-662-03978-6.
- [15] NASA, “Fundamentos de la Teledetección para Aplicaciones de Salud Pública”, jun. 2016. Consultado: el 8 de mayo de 2023. [En línea]. Disponible en: <https://appliedsciences.nasa.gov/sites/default/files/2020-11/healthmonitoringpart1spanish.pdf>
- [16] M. Gonzalez y E. Mendoza, “Introducción a la teledetección SESIÓN 2”. Consultado: el 8 de mayo de 2023. [En línea]. Disponible en: https://appliedsciences.nasa.gov/sites/default/files/EO4IM_Session_2_Espanol.pdf
- [17] C. Sáiz Escrivá de Romaní, “Estudio de índices de vegetación a partir de imágenes aéreas para la detección de déficit hídrico y aplicaciones en la agricultura de precisión y en la evaluación de daños en seguro agrario del almendro, en parcela de 20 ha en Chiloeches (Guadalajara)”, UNIVERSIDAD POLITÉCNICA DE MADRID, 2022.
- [18] R. P. Gupta, *Remote Sensing Geology*, 3rd ed. 2018. Berlin, Heidelberg: Springer Berlin / Heidelberg, 2017. doi: 10.1007/978-3-662-55876-8.
- [19] A. R. Huete, “A soil-adjusted vegetation index (SAVI)”, *Remote Sens Environ*, vol. 25, núm. 3, pp. 295–309, 1988, doi: 10.1016/0034-4257(88)90106-X.
- [20] B. Gao, “NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space”, *Remote Sens Environ*, vol. 58, núm. 3, pp. 257–266, 1996, doi: 10.1016/S0034-4257(96)00067-3.

- [21] F. Cigna y H. Xie, *Imaging Floods and Glacier Geohazards with Remote Sensing*. Basel, Switzerland: MDPI - Multidisciplinary Digital Publishing Institute, 2021. doi: 10.3390/books978-3-0365-0067-6.
- [22] Real Academia Española, “reforestar | Definición | Diccionario de la lengua española | RAE - ASALE”. Consultado: el 6 de agosto de 2023. [En línea]. Disponible en: <https://dle.rae.es/reforestar>
- [23] C. J. González Martínez, *Reforestación agroecológica : una alternativa para la protección del recurso hídrico*. Bogotá: UNIMINUTO Corporación Universitaria Minuto de Dios, 2019.
- [24] WWF, “Glosario ambiental: ¿Deforestación o degradación? ¿Reforestación o restauración? Resolvamos dudas | WWF”. Consultado: el 7 de mayo de 2023. [En línea]. Disponible en: <https://www.wwf.org.co/?uNewsID=323725>
- [25] S. Russell y P. Norving, *Inteligencia artificial : un enfoque moderno / Stuart J. Stuart, Peter Norving ; traducción Juan Manuel Corchado Rodríguez ... [et al.]*. 2003. [En línea]. Disponible en: <https://search.ebscohost.com/login.aspx?direct=true&db=cat09062a&AN=cj.oai.edge.tad.eo.folio.ebsco.com.fs00001093.a8e70af5.4094.5316.b044.4cb913998d91&site=eds-live>
- [26] R. Chopra, A. England, y M. N. Alaudeen, *Data Science with Python : Combine Python with Machine Learning Principles to Discover Hidden Patterns in Raw Data*. Birmingham: Packt Publishing, 2019. [En línea]. Disponible en: <https://search.ebscohost.com/login.aspx?direct=true&AuthType=sso&db=e000xww&AN=2204654&lang=es&site=eds-live&scope=site&custid=s6026984>
- [27] D. Natingga, *Data Science Algorithms in a Week : Top 7 Algorithms for Scientific Computing, Data Analysis, and Machine Learning, 2nd Edition*, vol. Second edition. Birmingham, UK: Packt Publishing, 2018. [En línea]. Disponible en: <https://search.ebscohost.com/login.aspx?direct=true&AuthType=sso&db=e000xww&AN=1925349&lang=es&site=eds-live&scope=site&custid=s6026984>
- [28] Bernhard. Schölkopf, *Learning with kernels : support vector machines, regularization, optimization, and beyond*. en Adaptive computation and machine learning. Cambridge, Mass: MIT Press, 2002.
- [29] K. Jolly, *Machine Learning with Scikit-Learn Quick Start Guide: Classification, Regression, and Clustering Techniques in Python*, 1a ed. Birmingham: Packt Publishing, Limited, 2018.

- [30] P. Refaeilzadeh, L. Tang, y H. Liu, "Cross-Validation", en *Encyclopedia of Database Systems*, New York, NY: Springer New York, 2018, pp. 677–684. doi: 10.1007/978-1-4614-8265-9_565.
- [31] Jorge Andres Polanco López de Mesa, "Teledetección de la vegetación del Paramo de Belmira con imágenes LANDSAT", *Dyna (Medellín, Colombia)*, vol. 79, núm. 171, pp. 222–231, 2012.
- [32] J.-C. Rojas y C. Castiblanco-Rozo, "Variabilidad espacial y análisis multitemporal sobre los cambios en las coberturas de la tierra ocasionados por la expansión del cultivo de la palma de aceite en los años 1985, 2000 y 2016 en el municipio de San Carlos de Guaroa (Colombia).", *Gestión y Ambiente*, vol. 21, núm. 1, pp. 59–68, ene. 2018, doi: 10.15446/ga.v21n1.69096.
- [33] J. Guzmán-Manrique y A. Camilo Flórez-García, "Fragmentación del paisaje empleando análisis multitemporal de imágenes de satélite Landsat TM y ETM+ en el municipio de Montelíbano, Córdoba-Colombia.", *Gestión y Ambiente*, vol. 22, núm. 1, pp. 31–41, ene. 2019, doi: 10.15446/ga.v22n1.76116.
- [34] J. F. Camacho Peña, "Análisis de regresión logística aplicado al modelamiento espacial de las causas de deforestación en el departamento del Guaviare durante el periodo 2005-2020 y proyección de escenarios de deforestación a 2030", *Cuadernos de geografía (Bogotá)*, vol. 31, núm. 2, pp. 255–280, 2022, doi: 10.15446/rcdg.v31n2.98012.
- [35] C. M. C. Lindo, J. G. Barrero, y G. J. Cleves, "Estimacion de la captura de carbono en un relicto de bosque urbano mediante tecnicas de teledeteccion/Estimation of Carbon Capture in an Urban Forest Relict through Teledetection Techniques/Estimativa da captura de carbono num relicto de floresta urbana mediante tecnicas de sensoriamento remoto", *Revista ingenierías (Medellín, Colombia)*, vol. 19, núm. 37, Universidad de Medellin, p. 13, 2020. doi: 10.22395/rium.v19n37a1.
- [36] J. A. F. Riano y M. F. P. Puerta, "Análisis de la deforestacion en La Macarena, antes y despues de los acuerdos de paz/Analysis of deforestation in La Macarena before and after the peace agreements", *Colombia forestal*, vol. 24, núm. 2, p. 9, 2021, doi: 10.14483/2256201X.16479.
- [37] D. M. Herrera Posada y E. Aristizábal, "Artificial Intelligence and Machine Learning Model for Spatial and Temporal Prediction of Drought Events in the Department of Magdalena,

Colombia.”, *Inge-Cuc*, vol. 18, núm. 2, pp. 249–265, jul. 2022, [En línea]. Disponible en: <https://search.ebscohost.com/login.aspx?direct=true&AuthType=sso&db=asn&AN=161704750&lang=es&site=eds-live&scope=site&custid=s6026984>

- [38] E. F. Avila Velez, L. A. Herrera Martinez, y A. Celis Forero, “Identificación de zonas con presencia de retamo espinoso (*ulex europaeus*), utilizando teledetección espacial, en el municipio de Sibaté [Cundinamarca]”, *Revista de Investigación Agraria y Ambiental*, vol. 14, núm. 1, 2022, doi: 10.22490/21456453.5840.
- [39] Carlos E. Mejía, Hernán J. Andrade, y Milena Segura, “Estimación de biomasa y carbono con herramientas de teledetección en bosques secos tropicales del Tolima, Colombia”, *Revista de teledetección*, núm. 62, pp. 57–70, 2023, doi: 10.4995/raet.2023.19242.
- [40] R. A. Díaz Giraldo, M. Álvarez de León, O. Pérez López, S. L. Gutiérrez Parrado, y M. A. Arango Argoti, “Aplicación de dos modelos basados en teledetección para estimar la producción primaria neta aérea en un sistema pastoril en la Altillanura colombiana: Estimación de productividad primaria de pasturas por teledetección”, *Tropical grasslands-Forrajeros tropicales*, vol. 12, núm. 1, pp. 24–35, 2024, doi: 10.17138/tgft(12)24-35.
- [41] D. C. Bezanilla, J. S. Ramírez, y A. de A. Ávila, “Estudio multitemporal de fragmentación de los bosques en la Sierra Fría, Aguascalientes, México.”, *Madera y Bosques*, vol. 14, núm. 1, pp. 37–51, mar. 2008, [En línea]. Disponible en: <https://search.ebscohost.com/login.aspx?direct=true&AuthType=sso&db=fua&AN=35903269&lang=es&site=eds-live&scope=site&custid=s6026984>
- [42] D. Armenteras y N. Rodríguez, “DINÁMICAS Y CAUSAS DE DEFORESTACIÓN EN BOSQUES DE LATINO AMÉRICA: UNA REVISIÓN DESDE 1990”, *Colombia Forestal*, vol. 17, núm. 2, p. 233, dic. 2014, doi: 10.14483/udistrital.jour.colomb.for.2014.2.a07.
- [43] G. Medina y A. Mejía, “Análisis Multitemporal y Multifractal de la Deglaciación de la Cordillera Parón en los Andes de Perú.”, *Ecología Aplicada*, vol. 13, núm. 1, pp. 35–42, ene. 2014, doi: 10.21704/rea.v13i1-2.452.
- [44] LI. Pérez-Planells, J. Delegido, J. P. Rivera-Caicedo, y J. Verrelst, “Análisis de métodos de validación cruzada para la obtención robusta de parámetros biofísicos”, *Revista de Teledetección*, núm. 44, pp. 55–65, dic. 2015, doi: 10.4995/raet.2015.4153.

- [45] B. Kozhikkodan Veetil *et al.*, “Un análisis comparativo del comportamiento diferencial de los glaciares en los Andes Tropicales usando teledetección”, *Investigaciones Geográficas*, núm. 51, p. 3, jul. 2016, doi: 10.5354/0719-5370.2016.41215.
- [46] B. Sánchez-Díaz, “La teledetección en investigaciones ecológicas como apoyo a la conservación de la biodiversidad: una revisión”, *Revista científica*, vol. 3, núm. 33, pp. 243–253, sep. 2018, doi: 10.14483/23448350.13370.
- [47] A. Ariza, J. Salas Rey, y S. M. de Miguel, “Comparison of maximum likelihood estimators and regression models for burn severity mapping in Mediterranean forests using Landsat TM and ETM+ data.”, *Revista Cartográfica*, núm. 98, pp. 145–177, ene. 2019, [En línea]. Disponible en: <https://search.ebscohost.com/login.aspx?direct=true&AuthType=sso&db=asn&AN=136455800&lang=es&site=eds-live&scope=site&custid=s6026984>
- [48] S. Malvina, C. G. Herrera, y A. E. Niz, “Teledetección aplicada al mapeo geomorfológico de los volcanes de la cuenca alta del río Chaschuil, provincia de Catamarca, Argentina”, *Tecnura*, vol. 23, núm. 60, pp. 13–26, 2019, doi: 10.14483/22487638.14642.
- [49] R. Valcarce Diñeiro, “Seguimiento y clasificación de parámetros biofísicos de superficies agrícolas a partir de sensores remotos radar”, 2020. [En línea]. Disponible en: <https://search.ebscohost.com/login.aspx?direct=true&AuthType=sso&db=edsred&AN=edsred.10366.145429&lang=es&site=eds-live&scope=site&custid=s6026984>
- [50] S. Liang, J. Cheng, y J. Zhang, “Maximum Likelihood Classification of Soil Remote Sensing Image Based on Deep Learning”, *Earth sciences research journal*, vol. 24, núm. 3, pp. 357–365, 2020, doi: 10.15446/esrj.v24n3.89750.
- [51] M. Ramírez, L. Martínez, M. Montilla, O. Sarmiento, J. Lasso, y S. Diaz, “Obtención de coberturas del suelo agropecuarias en imágenes satelitales sentinel-2 con la inyección de imágenes de dron usando random forest en google earth engine”, *Revista de teledetección*, vol. 2020, núm. 56, pp. 49–68, 2020, doi: 10.4995/raet.2020.14102.
- [52] A. I. Alvarado Huapaya, P. Capristán Sotelo, C. Corahua Benites, C. Ruiz Philipps, y G. Velásquez Bejarano, “VARIACIÓN DEL ÁREA AGRÍCOLA EN EL DISTRITO LA YARADA LOS PALOS, TACNA, PERÚ.”, *Espacio y Desarrollo*, núm. 35, pp. 99–120, ene. 2020, [En línea]. Disponible en:

<https://search.ebscohost.com/login.aspx?direct=true&AuthType=sso&db=fua&AN=151638307&lang=es&site=eds-live&scope=site&custid=s6026984>

- [53] J. Mejía-Marcacuzco, E. Pino-Vargas, E. Guevara-Pérez, V. Olivos-Alvites, y M. Condori-Ventura, “Predicción espacial de la erosión del suelo en zonas áridas mediante teledetección. Estudio de caso: Quebrada del Diablo, Tacna, Perú”, *Revista Ingeniería UC*, vol. 28, núm. 2, pp. 252–264, sep. 2021, doi: 10.54139/revinguc.v28i2.24.
- [54] F. Rahal y F.-Z. Baba-Hamed, “Mapping of clay soils by remote sensing in the area of Mers El Kébir, Algeria”, *Revista Facultad de Ingeniería Universidad de Antioquia*, núm. 109, p. 9, oct. 2022, doi: 10.17533/udea.redin.20221099.
- [55] A. C. E. Zúñiga, J. C. Rodríguez, J. V. B. Saya, y J. A. Ñ. Vázquez, “Estimación de la biomasa de una comunidad vegetal altoandina utilizando imágenes multiespectrales adquiridas con sensores remotos UAV y modelos de Regresión Lineal Múltiple, Máquina de Vectores Soporte y Bosques Aleatorios”, *Scientia Agropecuaria*, vol. 13, núm. 3, oct. 2022, doi: 10.17268/sci.agropecu.2022.027.
- [56] A. Ahmad, S. E. Noor, P. Cartujo Cassinello, y V. Martos Núñez, “Artificial Intelligence (AI) as a complementary technology for agricultural Remote Sensing (RS) in plant physiology teaching.”, *REiDoCrea: Revista Electrónica de Investigación y Docencia Creativa*, vol. 11, pp. 695–701, ene. 2022, [En línea]. Disponible en: <https://search.ebscohost.com/login.aspx?direct=true&AuthType=sso&db=fua&AN=161802827&lang=es&site=eds-live&scope=site&custid=s6026984>
- [57] Y. García Ortega, D. Valdez Rivera, y D. Mancero Castillo, “Dinámica de la frontera agrícola del sistema de cuencas hidrográficas del zapotal mediante herramientas de teledetección”, *Ciencia y Tecnología*, vol. 16, núm. 1, pp. 12–23, jun. 2023, doi: 10.18779/cyt.v16i1.637.
- [58] M. Mulder, “Monitoring land restoration projects of Justdiggit in Kenya, using downscaled passive microwave remote sensing products of VanderSat”, 2018. Consultado: el 6 de agosto de 2023. [En línea]. Disponible en: <https://repository.tudelft.nl/islandora/object/uuid%3A86a24122-e3b5-4cc7-b580-f70fdecb78f8>
- [59] L. Villani, G. Castelli, F. Sambalino, L. A. Almeida Oliveira, y E. Bresci, “Integrating UAV and satellite data to assess the effects of agroforestry on microclimate in Dodoma region, Tanzania”, *2020 IEEE International Workshop on Metrology for Agriculture and Forestry*,

MetroAgriFor 2020 - Proceedings, pp. 338–342, nov. 2020, doi: 10.1109/METROAGRIFOR50201.2020.9277643.

- [60] M. Van der Vliet *et al.*, “Quantifiable impact: monitoring landscape restoration from space”, mar. 2023, doi: 10.21203/RS.3.RS-2669521/V1.
- [61] Google LLC, “Google Earth Engine”. Consultado: el 29 de octubre de 2023. [En línea]. Disponible en: <https://earthengine.google.com/>
- [62] Google LLC, “Earth Engine Data Catalog | Google for Developers”. Consultado: el 11 de marzo de 2024. [En línea]. Disponible en: <https://developers.google.com/earth-engine/datasets/>
- [63] La Administración Nacional de Aeronáutica y del Espacio - NASA, “Landsat 7 | Landsat Science”. Consultado: el 12 de marzo de 2024. [En línea]. Disponible en: <https://landsat.gsfc.nasa.gov/satellites/landsat-7/>
- [64] Google LLC y Servicio Geológico de EE. UU, “USGS Landsat 7 Level 2, Collection 2, Tier 1 | Earth Engine Data Catalog | Google for Developers”. Consultado: el 17 de marzo de 2024. [En línea]. Disponible en: https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LE07_C02_T1_L2#description
- [65] The World Bank Group, “Kenya - Climatology | Climate Change Knowledge Portal”. Consultado: el 13 de marzo de 2024. [En línea]. Disponible en: <https://climateknowledgeportal.worldbank.org/country/kenya/climate-data-historical>
- [66] The World Bank Group, “Tanzania - Climatology | Climate Change Knowledge Portal”. Consultado: el 13 de marzo de 2024. [En línea]. Disponible en: <https://climateknowledgeportal.worldbank.org/country/tanzania/climate-data-historical>
- [67] P. Sanchez y John Wiley, *Suelos del trópico: características y manejo*. Universidad Estatal de Carolina del Norte Raleigh, 1981.
- [68] Google LLC, “Landsat Collection 1 to Collection 2 migration | Google Earth Engine | Google for Developers”. Consultado: el 17 de marzo de 2024. [En línea]. Disponible en: https://developers.google.com/earth-engine/landsat_c1_to_c2
- [69] Google LLC, “ImageCollection Visualization | Google Earth Engine | Google for Developers”. Consultado: el 17 de marzo de 2024. [En línea]. Disponible en: https://developers.google.com/earth-engine/guides/ic_visualization

- [70] IDEAM, “Formato Común Hoja Metodológica”, Consultado: el 3 de abril de 2024. [En línea]. Disponible en: www.ideam.gov.co
- [71] International Business Machines Corporation (IBM), “Conceptos básicos de ayuda de CRISP-DM - Documentación de IBM”. Consultado: el 20 de marzo de 2024. [En línea]. Disponible en: <https://www.ibm.com/docs/es/spss-modeler/saas?topic=dm-crisp-help-overview>
- [72] Google LLC, “Platform – Google Earth Engine”. Consultado: el 20 de marzo de 2024. [En línea]. Disponible en: <https://earthengine.google.com/platform/>
- [73] Google LLC y Agencia Espacial Europea, “Harmonized Sentinel-2 MSI: MultiSpectral Instrument, Level-2A | Earth Engine Data Catalog | Google for Developers”. Consultado: el 20 de marzo de 2024. [En línea]. Disponible en: https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_SR_HARMONIZED
- [74] Google LLC, “Google Earth Pro”. Consultado: el 22 de marzo de 2024. [En línea]. Disponible en: <https://www.google.com/intl/es-419/earth/about/>
- [75] Agencia de la ONU para refugiados - ACNUT, “Emergencia por la sequía en el Cuerno de África | ACNUR”. Consultado: el 31 de marzo de 2024. [En línea]. Disponible en: <https://www.acnur.org/emergencias/cuerno-de-africa>
- [76] Agencia de la ONU para refugiados - ACNUR, “El Cuerno de África sufre la peor sequía de los últimos años”. Consultado: el 31 de marzo de 2024. [En línea]. Disponible en: <https://eacnur.org/es/actualidad/noticias/desplazados/el-cuerno-de-africa-sufre-la-peor-sequia-de-los-ultimos-anos>
- [77] J. T. Abatzoglou, S. Z. Dobrowski, S. A. Parks, y K. C. Hegewisch, “TerraClimate: Monthly Climate and Climatic Water Balance for Global Terrestrial Surfaces, University of Idaho | Earth Engine Data Catalog | Google for Developers”. Universidad de Idaho, 2018. Consultado: el 30 de marzo de 2024. [En línea]. Disponible en: https://developers.google.com/earth-engine/datasets/catalog/IDAHO_EPSCOR_TERRACLIMATE#citations