



**FACULTAD DE INGENIERÍA, ARQUITECTURA Y
URBANISMO**

**ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS
TESIS**

**COMPARACIÓN DE TÉCNICAS DE ESTIMACIÓN
DEL GRADO DE SALINIDAD EN SUELOS DE
ESCASA VEGETACIÓN MEDIANTE EL
PROCESAMIENTO DE IMÁGENES
MULTIESPECTRALES DE SATÉLITE**

**PARA OPTAR EL TÍTULO PROFESIONAL DE INGENIERO
DE SISTEMAS**

Autor(a) (es):

Bach. Pardo Villegas Jose Diego

ORCID: <https://orcid.org/0000-0001-6170-3939>

Bach. Salinas Velez Bryan Jose Ricardo

ORCID: <https://orcid.org/0000-0002-3434-3330>

Asesor(a):

Mg. Mejia Cabrera Heber Ivan

ORCID: <https://orcid.org/0000-0002-0007-0928>

Línea de Investigación:

Infraestructura, Tecnología y Medio Ambiente

Pimentel – Perú 2022

APROBACIÓN DEL JURADO

**COMPARACIÓN DE TÉCNICAS DE ESTIMACIÓN DEL GRADO DE SALINIDAD
EN SUELOS DE ESCASA VEGETACIÓN MEDIANTE EL PROCESAMIENTO
DE IMÁGENES MULTIESPECTRALES DE SATÉLITE**

Bach. Pardo Villegas José Diego

Autor

Bach. Salinas Vélez Bryan José Ricardo

Autor

Mg. Mejía Cabrera Heber Ivan

Asesor

Mg. Mejia Cabrera Heber Ivan

Presidente de Jurado

Mg. Cachay Maco Junior Eugenio

Secretario de Jurado

Mg. Aguinaga Tello Juan

Vocal de Jurado

Dedicatorias

A nuestros padres

Por el esfuerzo, los valores inculcados y su constante apoyo.

A nuestros docentes

Por la constante asesoría y disposición para transmitir su conocimiento.

A nosotros

Por nuestro esfuerzo y constancia para alcanzar nuestra primera meta profesional.

Agradecimientos

A Dios

Por ser el inspirador y darnos fuerza para continuar en este proceso de obtener uno de los anhelos más deseados.

A nuestros padres

Por su amor, trabajo y sacrificio en todos estos años.

A todas las personas que nos han apoyado y han hecho que el trabajo se realice con éxito en especial a aquellos que nos abrieron las puertas y compartieron sus conocimientos.

Resumen

En 2005 el MINAG, concluyó que 0.24 % del total de suelo agricultor en Perú, es afectado por salinización y este se encuentra ubicado en suelos costeros. Del mismo modo se advirtió que la producción agrícola en el departamento de Lambayeque aportó 0.8% al PBI del país y 16,2% del PBI de la región.

Siendo la degradación del suelo a causa de la salinización, el escaso recurso hídrico y una deficiente planificación por parte de los productores, las principales causas de un bajo nivel de crecimiento agrícola en dicha región.

Por tal motivo y teniendo en cuenta los datos antes descritos, se propuso el trabajo de investigación “Comparación de técnicas de estimación del grado de salinidad en suelos de escasa vegetación, mediante el procesamiento de imágenes multiespectrales de satélite” y de esta forma usar el procesamiento de imágenes multiespectrales para estimar la salinidad de los terrenos de escasa vegetación, y así aprovecharlos mediante la agricultura dirigida.

Para lo cual se consideró utilizar las siguientes técnicas de estimación; SLR, MLR, RFR y DTR, para extraer características como los indicadores de salinidad y vegetación utilizar imágenes multiespectrales, para validar se realizaron pruebas de conductividad eléctrica para medir el grado de salinidad

La comparación de estas técnicas mostró resultado que dan a DTR como el de mejor con 95% de precisión y un error promedio de 0.11 dS/m, en segundo lugar, RFR con 88%, de precisión y un error promedio de 0.18 dS/m, mientras que SLR y MLR solo obtuvieron un 52% de precisión y un error promedio de 0.33 dS/m.

Palabras Clave:

Salinidad en suelos, conductividad eléctrica, Técnicas de Estimación, imágenes multiespectrales, grado de salinidad, escasa vegetación.

Abstract

In 2005 the MINAG concluded that 0.24% of the total agricultural land in Peru is affected by salinization and this is located in coastal soils. Similarly, it was noted that agricultural production in the department of Lambayeque contributed 0.8% to the country's GDP and 16.2% to the region's GDP.

Soil degradation due to salinization, scarce water resources and poor planning by producers are the main causes of a low level of agricultural growth in said region.

For this reason and taking into account the data described above, the research work "Comparison of techniques for estimating the degree of salinity in soils with little vegetation, through the processing of multispectral satellite images" was proposed and thus use the processing of multispectral images to estimate the salinity of sparsely vegetated lands, and thus take advantage of them through directed agriculture.

For which it was considered to use the following estimation techniques; SLR, MLR, RFR and DTR, to extract characteristics such as indicators of salinity and vegetation, use multispectral images, to validate electrical conductivity tests were carried out to measure the degree of salinity´

The comparison of these techniques showed results that give DTR as the best with 95% precision and an average error of 0.11 dS / m, secondly, RFR with 88%, precision and an average error of 0.18 dS / m , while SLR and MLR only obtained 52% precision and an average error of 0.33 dS / m.

Keywords:

Soil salinity, electrical conductivity, Estimation Techniques, multispectral imaging, Salinity Degree, Sparce Vegetation.

Índice

I. INTRODUCCIÓN	13
1.1. Realidad Problemática	13
1.2. Trabajos previos	16
1.3. Teorías relacionadas al tema	24
1.3.1. Conductividad Eléctrica (CE)	24
1.3.2. Radiación Electromagnética	25
1.3.3. Espectro Electromagnético	26
1.3.4. Radiancia Espectral (LT)	27
1.3.5. Reflectancia Espectral	28
1.3.6. Firma Espectral	29
1.3.7. Teledetección o Detección Remota	29
1.3.7.1. Imágenes multiespectrales de satélite	30
1.3.7.2. Pre procesamiento de Imágenes multiespectrales de satélite	31
1.3.7.3. Indicadores Espectrales	32
1.3.8. Técnicas de estimación	33
1.3.9. Suelos de Escasa Vegetación	37
1.4. Formulación del Problema	38
1.5. Justificación e importancia del estudio	38
1.6. Hipótesis	39
1.7. Objetivos	39
1.7.1. Objetivo general	39
1.7.2. Objetivos específicos	39
II. MATERIAL Y MÉTODO	40
2.1. Tipo y Diseño de Investigación	40
2.1.1 Tipo de la Investigación	40

2.1.2	Diseño de Investigación.	40
2.2.	Población y muestra.	40
2.3.	Variables, Operacionalización.	41
2.4.	Técnicas e instrumentos de recolección de datos, validez y confiabilidad.....	42
2.4.1	Abordaje Metodológico.....	42
2.4.2	Técnicas de Recolección de Datos.....	42
2.5.	Procedimiento de análisis de datos.	42
2.6.	Criterios éticos.	43
2.7.	Criterios de Rigor Científico.....	43
III.	RESULTADOS.	44
3.1.	Resultados en Tablas y Figuras.	44
3.2.	Discusión de resultados.....	45
3.3.	Aporte práctico.....	46
3.3.1.	Elegir las técnicas de estimación a utilizar.....	46
3.3.2.	Obtener imágenes multiespectrales del área de estudio	47
3.3.3.	Efectuar calibraciones atmosféricas a las imágenes multiespectrales adquiridas.....	51
3.3.4.	Elegir indicadores de salinidad y vegetación.....	53
3.3.5.	Recoger muestras de suelos de escasa vegetación y ejecutar la prueba de conductividad eléctrica.....	55
3.3.6.	Comparar las técnicas de estimación.	61
3.3.6.1.	Simple Lineal Regression (SLR).....	62
3.3.6.2.	Multiple Lineal Regression (MLR)	63
3.3.6.3.	Random Forest Regression (RFR).....	64
3.3.6.4.	Decision Tree Regression (DTR)	65
IV.	CONCLUSIONES Y RECOMENDACIONES	76

4.1. Conclusiones.....	76
4.2. Recomendaciones.....	77
REFERENCIAS.....	78
ANEXOS.	82

Índice de Figuras

Figura 1. Costos de Degradación Ambiental Anual.....	14
Figura 2. Clasificación del espectro electromagnético	26
Figura 3. Espectro electromagnético característico del suelo desnudo, vegetación y agua.	29
Figura 4. Proceso de calibración radiométrica y geométrica de imágenes.	32
Figura 5. Representación gráfica de la curva de regresión.	34
Figura 6. Representación de los mínimos cuadrados.	35
Figura 7. Valor esperado de Y dado x_1	35
Figura 8. Comparación de técnicas de estimación en el indicador de precisión R 2	44
Figura 9. Comparación de técnicas de estimación en el indicador de precisión RMSE	45
Figura 10. Comparación de técnicas de estimación en el indicador de precisión RMSE	48
Figura 11. Estructura de la carpeta de las imágenes adquiridas de Landsat 8	48
Figura 12. Imágenes multiespectrales adquiridas del satélite Landsat 8	49
Figura 13. Ubicación de las muestras recogidas en la ciudad de Mórrope	56
Figura 14. Procedimiento para la medición de Conductividad eléctrica	57
Figura 15. Variables independientes (indicadores)	61
Figura 16. Variables dependiente (CE)	61
Figura 17. Ingreso de coordenadas para la obtener las imágenes multiespectrales.	87
Figura 18. Ingreso de coordenadas para la obtener las imágenes multiespectrales.	88
Figura 19. Selección del satélite LANSAT 8.....	88
Figura 20. Selección Scene Cloud Cover.....	89
Figura 21. Descarga de paquetes de Imagen del LANSAT 8 OLI.....	90
Figura 22. Descarga de paquetes de Imagen del LANSAT 8 OLI.....	91
Figura 23. Archivo Metadata del conjunto de imágenes multiespectrales.....	91
Figura 24. Variables que se necesitan para calcular la reflectancia.....	92
Figura 25. Ubicación del valor de un pixel con ImageJ	93
Figura 26. Resultados del laboratorio de las pruebas.	100

Índice de Tablas

Tabla 1. Clasificación de Salinidad de suelo según valores de conductividad eléctrica (EC).....	25
Tabla 2. Límites del espectro electromagnético según su división.....	27
Tabla 3. Límites de la luz visible según los colores RGB.	27
Tabla 4. Resolución espacial de satélites.	31
Tabla 5. Variables, Operacionalización.	41
Tabla 6. Población de técnicas de estimación.....	46
Tabla 7. Bandas utilizadas del sensor OLI del satélite Landsat 8.	49
Tabla 8. Dataset del sensor OLI del satélite Landsat 8.....	50
Tabla 9. Identificador de Bandas de las Imágenes del satélite Landsat 8.....	53
Tabla 10. Correlación entre los indicadores de salinidad, vegetación y bandas, y el valor de conductividad eléctrica	54
Tabla 11. Correlación entre los indicadores de salinidad, vegetación y bandas, y el valor de conductividad eléctrica.	57
Tabla 12. Relación entre los indicadores elegidos y el valor de conductividad eléctrica.....	59
Tabla 13. Resultados de la estimación usando SLR, relación entre Ce. Estimada y Real.....	62
Tabla 14. Resultados de la estimación usando MLR, relación entre Ce. Estimada y Real.....	64
Tabla 15. Resultados de la estimación usando RFR, relación entre Ce. Estimada y Real.....	65
Tabla 16. Resultados de la estimación usando RFR, relación entre Ce. Estimada y Real.....	66
Tabla 17. Cálculo del indicador R^2 para SLR.	68
Tabla 18. Cálculo del indicador RMSE para SLR.	69
Tabla 19. Cálculo del indicador R^2 para MLR.	70
Tabla 20. Cálculo del indicador RMSE para MLR.	71
Tabla 21. Cálculo del indicador R^2 para RFR.	72
Tabla 22. Cálculo del indicador RMSE para RFR.	73
Tabla 23. Cálculo del indicador R^2 para DTR.....	74
Tabla 24. Cálculo del indicador RMSE para DTR.	75

Tabla 25. Población de técnicas de estimación.....	86
Tabla 26. Variables para el cálculo de reflectancia de cada banda.....	92
Tabla 27. Variables de ángulo de elevación solar para el cálculo de reflectancia.	93
Tabla 28. Ubicación y valor del pixel en una imagen multiespectral.	94
Tabla 29. Información de las coordenadas y la banda para buscar el valor del pixel.	98
Tabla 30. Ubicación y valor del pixel en una imagen espectral	98
Tabla 31. Cuadro de Verificación de la estimación de conductividad eléctrica.	101

INTRODUCCIÓN

1.1. Realidad Problemática.

La merma de los recursos nutritivos, a causa de la desertificación y salinidad, sumado al daño de los suelos, son las principales consecuencias generadas por un suelo degradado (López, 1944).

A nivel mundial los problemas de salinidad son considerados como un peligro ambiental latente. La intensificada actividad agrícola y el riego son las principales causas de la salinidad de suelo (Akramkhanov, Martius, Park, & Hendrickx, 2011)

En 2012, INEI realizó un censo agropecuario del cual se obtiene como resultado 7.1 millones de hectáreas de superficie agrícola en Perú. Del mismo modo se estima que tienen aptitud natural para el cultivo agrícola solo 7.6 millones.

En 2005 el MINAG, concluyó que 0.24 % del total de suelo agricultor en Perú, es afectado por salinización y este se encuentra ubicado en suelos costeros. Del mismo modo se advirtió que la producción agrícola en el departamento de Lambayeque, aportó 0.8% al PBI del país y 16,2% del PBI de la región. Siendo la degradación del suelo a causa de la salinización, el escaso recurso hídrico y una deficiente planificación por parte de los productores, las principales causas de un bajo nivel de crecimiento agrícola en dicha región.

En el estudio realizado por el AAP para identificar problemas ambientales, que la salinidad genera el costo más elevado en cuanto a degradación ambiental con respecto al PBI. (Larsen, 2005)

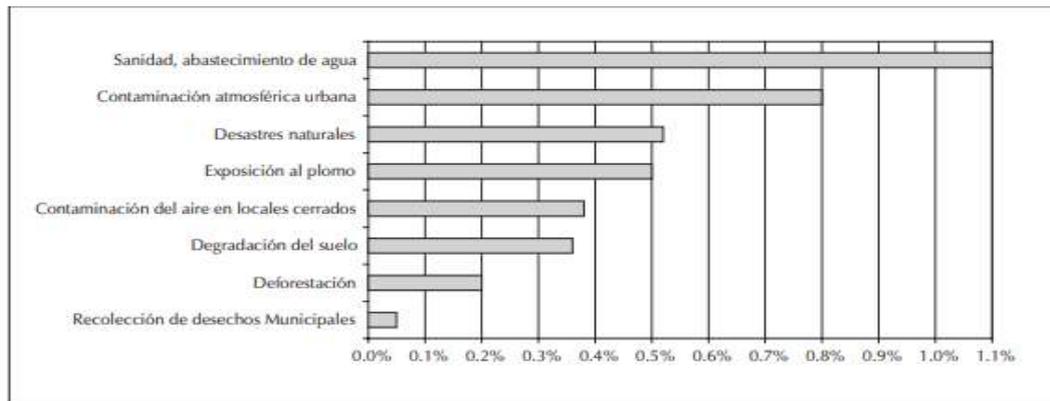


Figura 1

Figura 1. Costos de Degradación Ambiental Anual (% del PBI)

Fuente: (Larsen, 2005)

El antecedente publicado en la web de la Gerencia de Agricultura de Lambayeque en el mes de agosto de 2018 informó del bajo nivel de producción agrario en que se encuentra el distrito de Mórrope, a causa del cultivo de arroz, esto por la necesidad de abundante recursos hídricos, siendo este uno de los recursos que más escasea en la zona, lo cual deriva en un suelo con alto nivel de salinidad, a lo cual se plantea como posible solución el cambio de cultivo por uno que se ajuste más a las características de la zona, el cual sería el algodón.

Por lo cual, es de suma importancia realizar seguimiento a los niveles de salinidad en los suelos, para que de esta forma no superen los niveles tolerados por la actividad agrícola y de esa manera mitigar los riesgos a causa de la salinización. (Bouksila et al. 2013).

La conductividad eléctrica aparente del suelo (CEa), es de las más usadas por ser una variable correlacionada con propiedades físicas y químicas del suelo. Por consiguiente, esta característica sirve como parámetro de ayuda en la agricultura de exactitud, al permitir mejorar la estimación de otras características del suelo tales como textura, salinidad, niveles de agua. (Siqueira, 2009)

Para estimar el grado salinidad en los suelos se requieren 2.25 muestras por hectárea, lo cual supone un costo de \$20 por cada muestra con respecto a la prueba (EC). Por el contrario, si aplicamos el test Conductividad Eléctrica Aparente (ECa) usando el sensor EM-38, se ve una reducción significativa. (FAO, 2007)

Por lo cual, usar metodologías rentables y prácticas que nos permitan evaluar, monitorear, grandes extensiones de suelos mermados por salinidad, supone una necesidad. (FAO, 2007)

La teledetección, es utilizada para recoger datos por intermedio de sensores que se encuentran en satélites los cuales nos permiten capturar imágenes multiespectrales del medio ambiente, lo cual es una de las alternativas en la actualidad. (Romero 2006).

De esta manera, la teledetección se presenta como una de las mejores opciones, para el estudio de la degradación de suelo causada por la salinidad, puesto que facilita la recolección de información en grandes áreas de suelo de cultivo (Mashimbye, y otros, 2012)

Así mismo, las características de los suelos y el factor ambiental guardan una relación en la cual influye la región, dichos factores tienen relación directa con la el grado de sales en los suelos. En la actualidad estimar el grado de salinidad por intermedio de Conductividad eléctrica (CE) mediante imágenes multiespectrales y propiedades del terreno, incluye técnicas como MLR, RSL, RF, DT entre otras. (Shahabi et al. 2017).

Lo cual genera la necesidad de comparar las técnicas que nos permitan estimar de manera más eficiente el grado de salinidad en los suelos de escaza vegetación del departamento.

1.2. Trabajos previos.

(Yu, Wenbin, & Shuwen, 2017). Debido a la influencia significativa en la fertilización racional, la expansión de la producción y la reducción de la contaminación causada por el uso excesivo de fertilizantes nitrogenados, con el fin de adquirir rápidamente información sobre nutrientes del maíz en el campo, realizando un método no destructivo de la medición de contenido de nitrógeno del maíz basado en imágenes multiespectrales mediante un la plataforma de helicóptero no tripulados con la implementación de una cámara multisectorial ADC(producida por Tetracon Corporation) y se obtuvieron las imágenes de los canales Rojo(R),Verde(G) e infrarrojo cercano (NIR), obteniendo las muestras del dosel de maíz en la etapa de unión mediante , y para el contenido de nitrógeno de cada muestra se midió con un analizador de flujo continuo(CFA) AA3. Para ello se calcularon 11 índices de vegetación, A continuación, se llevó a cabo un análisis de correlación para reducir la dimensión de los datos a fin de adquirir 3 parámetros característicos espectrales sensibles, en la etapa de experimentar de donde se obtuvieron las muestras fue en la granja de Xiangfang, Harbin (125° 42' E,44° 04' N, Noreste de China). Dongnong 253, donde el periodo de crecimiento del maíz fue de 128 días con una temperatura de crecimiento de 10 ° C y temperatura acumulada activa de 2700 ° C, además se establecieron 4 niveles de fertilizante nitrogenado, 0(N1),65 (N2),109 (NE) y 163 (N4) kg/ha , donde cada una puede ser reconocida da como grave falta de nitrógeno, falta de nitrógeno, nitrógeno moderado y nitrógeno excesivo, respectivamente, utilizando el método de análisis clínico convencional que para determinar el contenido de nitrógeno en experimento, para la extracción de características se calculó el promedio de grises del canal rojo(AVSR), el promedio de grises del canal verde (AVSG) y el promedio de grises del canal del infrarrojo cercano (AVSNIR) se extraen mediante la ecuación, los otros ocho índices , NDVI, GNDVI, RVIR / G, RVIR / NIR, RVINIR / R, RNR, GNR y NIRNR que son calculados mediante su ecuación, siendo el índice de vegetación de diferencia normalizada (NDVI) el que se utiliza ampliamente, en la etapa de unión se extrajeron los parámetros característicos espectrales de 80 grupos, y los parámetros característicos parciales (N3 en la etapa de unión) donde los GNR,NIRNR,AVSR, AVSG son omitidos, el modelo de detección del contenido

de maíz , el estudio mostro que el coeficiente de correlación R puede alcanzar el 0,6 (60%) del modelo de recuperación del contenido de clorofila basados en los 7 índices de vegetación, Zhang llevo a cabo el análisis de los 18 índices de vegetación para la recuperación del contenido de clorofila y los resultados mostraron que había 5 coeficientes de correlación superiores a 0,75, Wang determino que la pendiente de del borde rojo , la reflectancia máxima del pico verde, la relación del índice de vegetación y el NDVI demostraron buena correlación de las hojas en los 19 parámetros característicos espectrales. Este experimento selecciono 3 de los índices espectrales sensibles NDVI, RVIR/NIR, RNR para realizar los modelos de parámetros de vegetación únicos y múltiples, con métodos de análisis en la plataforma SPSS2.0, obteniendo como resultado del Modelado basado en índice de vegetación, el análisis de regresión realizo utilizando los 3 indicadores espectrales RVIR/NIR, RNR y el contenido de nitrógeno de la planta y se estableció el modelado de regresión. Las curvas de regresión lineal y cuadrática del contenido de nitrógeno de las hojas y los 3 índices espectrales en la etapa de unión dieron mejores resultados, y el grado R2 fue de 0,854 y 0,848 respectivamente, para el modelo basado en el índice de vegetación múltiple y mejorar el grado de ajuste del modelo de regresión, se seleccionaron y analizaron los 3 índices espectrales sensible, por el análisis de regresión por pasos, el modelo de recuperación de tres parámetros fue, el modelo de recuperación de los tres parámetros fue $y=1.776 + 32.559RNR + 15.512$ y $RVIR/NIR + 1.085 NDVI$, el coeficiente y la ecuación del modelo de regresión son estadísticamente significativos. El grado de ajuste de R2 fue de 0.870 (87%) y el R2 revisado fue de 0.86(87%) que fue mejor que en el modelo de índice de vegetación simple, en conclusión, la tecnología multiespectral puede adquirir de forma rápida y no destructiva información sobre nutrición de los cultivos, para realizar un diagnóstico de la nutrición nitrogenada. El índice característico espectral NDI, RVIR/NIR/ RNR están significativamente correlacionados con el contenido de nitrógeno de las hojas, según se concluye en el análisis de correlación, Se estableció el modelo de regresión lineal y regresión polinomial con parámetros espectrales seleccionados y parámetros de nitrógeno, el modelo de diagnóstico de nutrición nitrogenada del maíz, que se estableció en el índice de vegetación único RVIR/NIR, encaja bien (R2=0.854), también se estableció

un modelo de regresión lineal múltiple para parámetros espectrales seleccionados. El modelo de diagnóstico de nutrición de nitrógeno del maíz se ajusta mejor que el modelo de índice único ($R^2 = 0.870$).

(Lau Quan, Garea Llano, & Ruiz, 2005). Afrontaron el problema de estimar el porcentaje de salinidad en el suelo de cultivo de caña de azúcar usando la teledetección con ayuda de imágenes de satélite y la información de conductividad eléctrica, el lugar de la investigación se situó en Guantánamo donde se hace el sembrío de caña de azúcar en Paraguay, donde dichas tierras son de color Oscuro Plástico por motivo de sobre explotación de tierras y la utilización de aguas sobrecargadas de minerales, las imágenes fueron obtenidas por medio de captura electrónica a una resolución de 700 DPI que obtuvo como equivalencia una resolución espacial de 9,1 m/pixel en formato TIFF con una área equivalente a $80\,503\,232\text{ m}^2$, además de tener una base de datos de muestras de conductividad eléctrica con un total de 800 obtenidas de una profundidad de 20 cm con una distancia de 0 a 120 cm entre cada punto recaudadas por el instituto de suelos, la Estación de Suelos Salinos de Guantánamo y el INICA, como resultado de obtuvo que el raster de normalización de diferencia normalizada NDVI obtuvo una graduación de (-0,02 a 0.24) del nivel digital donde sus dichos valores son aceptable para la detección de la salinidad en sembrío de caña de azúcar, de donde se consideró 10 valores de interlineados de conductividad eléctrica prevista, siendo de menor valor a lo reportado -0,6 y 0,5 esto se debe al alto porcentaje de extensión de suelo lo que hace reducir el porcentaje de la reflectancia en la banda roja y a la poca visibilidad de elementos que no se aprecian bien en el infrarrojo, como la vegetación saludable, obteniendo como resultado final una correlación entre los datos estimados de salinidad y las medidas que se utilizaron con los valores de cada muestra, donde cuyo coeficiente de correlación resultó de 0,66, lo que se considera un valor aceptable teniendo en cuenta la desnivelación de los tiempos de recolección de imágenes referentes con los datos de conductividad eléctrica.

(Soca Flores, 2015). En su trabajo de investigación afronto el problema de estimar la salinidad del suelo con la ayuda de la teledetección en imágenes de satélite, validándolos con las muestras que se obtuvieron en el lugar a través de la CE, los lugares de donde se centró la investigación fue en la cuenca de Zaña provincia de Chiclayo, en donde se utilizaron imágenes del satélite utilizando los sensores que posteriormente fueron analizadas a través del software ENVI 4.5 y también del lenguaje de programación IDL 7.0, donde se logró diferenciar los suelos con alta y bajo nivel de salinidad además de calidad agrícola, se generó un array de tierras no cultivadas, para superponer a la imagen HRG que tiene una resolución de 10m x 10 m para mostrar a detalle la varianza entre el NVDI, tanto como en suelos normales y con problemas de salinidad, para ello se hizo un análisis de variación del NDVI de los periodos de 1999 – 2012, posteriormente se implementó el modelo de regresión lineal múltiple entre los las muestras de CE, frente a los valores d reflectancia generando el conjunto de datos que fueron analizados mediante el estimador de regresión lineal múltiple, usando las bandas B1, B2,B3 del sensor TM , alcanzando como resultado un coeficiente de correlación de 0.73, utilizando la ecuación de regresión se pudo hacer una estructuración desde el espacio de la salinidad del suelo, la cual fue dividida en 5 tipos de clases de suelos siendo estas, no salinos, ligeramente salinos, moderadamente salinos, fuertemente salinos y extremadamente salinos.

(Gamarra Agama, 2017). En su trabajo de investigación afronto el problema de la planificación del territorio seguido del ordenamiento, usando las imágenes del sensor Landsat 8 del año 2016 ubicado en Chanchamayo (Junín) con una superficie total de 591.67 km^2 , para ello se hizo uso de dos imágenes del Satélite descargadas del servidor The USGS Global Visualización Viewer (GloVis) de con una resolución media y una presencia de 50% de nubosidad, cuatro imágenes con una alta resolución, como también un Modelo Digital de Elevación (DEM) 30 metros, posteriormente se utilizó el programa ARCGIS 10.1 para poder procesar las imágenes y el Programa ENVIEX para poder hacer una visualización completa de las imágenes obteniendo así datos más completos de dichas imágenes, luego se procedió a usar ERDAS IMAGINE 9.1 para la extracción de información y construcción de imágenes de sistema a través de un Sistema de

Información Geográfica, para luego hacer uso de un clasificador semi automático para concluir con el mapeo de la cobertura y uso de la tierra, en el procesamiento de imágenes se utilizó el programa Erdas Imagine 9.1, con el cual se usó a unión de las bandas 1 a 7 y 9. Los resultados obtenidos en este estudio donde se utilizó el algoritmo Random Forest como clasificador obteniendo como resultado un 97,2 por ciento de precisión además de un valor de 0.922 del índice de Kappa. Dando como resultado final de la investigación que la mejor alternativa eficiente es el Random Forest para la clasificación de las imágenes satelitales.

(Mejia Cabrera, Vilchez, Tuesta Monteza, & Forero, 2020). En su investigación se afrontó el problema de estimar el grado de salinidad por medio de imágenes multiespectrales obtenidas del satélite, el área de estudio elegido fue en el distrito de Mórrope, Lambayeque, Perú(6°,37'37.9") con ayuda de un (GPS) Garmim Montana 650t que tiene una precisión de +/- 3.65m suficiente para su estudio porque cada pixel tiene una resolución de 30 x 30 m, de donde se obtuvieron 11 muestras del suelo para posteriormente ser evaluadas por método de conductividad eléctrica y correlacionarlas con las imágenes multiespectrales que fueron obtenidas de Earth Explorer del Servicio Geológico de los Estados Unidos de las cuales el conjunto contiene 11 imágenes de las cuales solo se utilizaron 7 debido a que para este estudio solo se necesitan las bandas 8,5,7 que permiten la observación de la salinidad en suelos. Para poder utilizar los modelos regresores fueron necesarios hacer una calibración atmosférica para cada una de las 7 bandas convirtiendo los valores en los valores de la imagen en reflectancia y luego en reflectancia de la parte superior de la atmósfera. Una vez que los valores del pixel se convirtieran en reflectancia se calculó los 11 índices de salinidad RED, GREEN, BLUE, NDVI, DVIBI, ASTERSI, SI, SI1, SI2, SI3, SI11, NIR, SWIR cada una con sus respectivas bandas, Los de estimadores que se utilizaron fueron Regresión de Mínimos Cuadrados Parciales (PLSR) y Maquinas de Vectores de Soporte(SVM), dando como resultado que el estimador SVM es mejor para estimar la salinidad alcanzando una precisión del 98 por ciento y 0.22 a diferencia de PLSR que tuvo una precisión menor de un 97 por ciento y 0.51.

(Guan, Shaoli, Zhanyi, & Ye, 2013). Se afrontaron el problema de un modelo para la predicción dinámica de la salinidad en el suelo ya que es fundamental para gestionar adecuadamente la irrigación de agua de acuerdo al grado de salinización, los datos recolectados para este estudio se tomaron del distrito de riego de Hetao en Mongolia Interior ya que es un distrito de riego más grande de Mongolia, La herramienta SVM se utilizó para la regresión para predecir la CE del suelo. Primero, se obtuvieron datos sobre la profundidad del agua subterránea, el volumen del agua de riego y la evaporación de 1991 a 2010 en el sub distrito de riego de Shahaoqu del distrito de riego de Hetao, se obtuvieron 40 muestras, las cuales se dividieron en 2 grupos. Un grupo constaba de 30 muestras designadas como conjunto de entrenamiento y el otro grupo constaba de 10 muestras designadas como conjunto de prueba. Luego se construyó el modelo SVM para predecir los valores de CE del suelo en el área de estudio. Para que el modelo se asemejara lo más posible a la evolución real de la salinidad del suelo y para mejorar la precisión de la predicción, se predijo el estado futuro de la salinidad del suelo de acuerdo con el modo entrenado. Los entrenamientos se basaron en los conjuntos de entrenamiento y los conjuntos de validación. Los núcleos y parámetros se seleccionaron de acuerdo con el entrenamiento RMSE. A través de las diferentes pruebas de parámetros ϵ , C y operaciones óptimas, se encontró que el sistema funciona mejor cuando $\epsilon = 0.0036$, $C = 26$. Con base en el modelo SVM construido, la Fig. 3 muestra el resultado de la relación entre lo observado y lo predicho datos de los valores de CE del suelo. Los valores probados y los valores predichos de este modelo tienen poca diferencia con sus valores reales. El valor absoluto del error relativo medio de ajuste es 2,14%, el del error relativo medio de prueba es 3,48% y el del error relativo medio de predicción es 6,37%. El modelo tiene alta velocidad de entrenamiento y prueba. La aplicación presentada en el documento actual se comparó con una herramienta de aprendizaje automático muy conocida, las ANN. El modelo ANN utiliza una función de pérdida por mínimos cuadrados a diferencia de las SVM que utilizan la función de pérdida ϵ -insensible como medida de aptitud. Las predicciones de ANN no son estables y dependen de los promedios de varias inicializaciones de red, que pueden producir un resultado diferente cada vez que se entrena un modelo. Por el contrario, los resultados de

SVM son estables y únicos. La Tabla 2 muestra los resultados de la comparación entre los métodos SVM y ANN basados en datos simulados y predichos. El error absoluto de los datos ajustados utilizando el método SVM fue de 0,18, pero el de ANN fue de 0,23, lo que indica que SVM predijo mejores resultados que ANN. El error máximo, el error mínimo y el error promedio de los resultados de SVM son menores que los de la ANN. Por lo tanto, el método SVM tiene una mejor capacidad para predecir el valor de CE del suelo.

(Amal, Lalit, & Priyakant, 2014). Para los investigadores debido a que mundialmente la salinidad en el suelo es un problema ambiental, tanto en regiones áridas como semiáridas, en este estudio se basó en desarrollar un modelo regresivo estadístico basándose en índices de detección remotos para predecir la salinidad del suelo. El área de estudio fue en oasis de Al Hassa está situado aproximadamente a 70 km tierra adentro de la costa del golfo entre una latitud $25^{\circ} 05'$ y $25^{\circ} 40'N$ y una longitud de $49^{\circ} 10'$ y $49^{\circ} 55' E$, El área de estudio cubre seis diferentes tipos de suelo, que son Torripsamments, Torriorthents, Calciorthids, Salorthids, Gypsiorthids y Haplaquepts. La distribución del tamaño de las partículas revela que los suelos son de textura franco arenosa. Se seleccionaron tres sitios de muestreo con base en la división del oasis y diferentes cantidades de vegetación. El primer sitio estaba ubicado en la parte norte del oasis en la ciudad de Al-Uyoun, que es caracterizado por una baja cobertura vegetal. El segundo sitio, en medio del oasis de Al-Bataliah pueblo, tenía una alta cobertura vegetal. El último conjunto de muestras se recogió bajo vegetación media. Las muestras se registraron utilizando un sistema de posicionamiento global (GPS) con una precisión de ± 5 m, cada muestra de suelo compuesta estaba compuesta por cuatro sub muestras centrales que se recolectaron a una distancia de 20 m al norte, sur, este y oeste del punto central de muestreo. Las sub muestras se recolectaron del horizonte superficial (0-20 cm) con una barrena manual (10 cm de diámetro) y se trituraron y mezclaron para formar una muestra. Se recolectaron un total de 149 muestras compuestas de suelo de los tres sitios definidos. La salinidad del suelo se puede medir directamente midiendo la CE en el campo y de forma remota, incluida la medición de laboratorio sin embargo, dado que el objetivo de este estudio es

establecer una relación entre la CE y la banda espectral del satélite y extrapolar la información del punto para generar un mapa de salinidad del suelo del área de estudio, la medición directa de la salinidad del suelo se realizó midiendo la CE, Para construir el modelo de regresión, las muestras se dividieron aleatoriamente en dos subconjuntos. Se utilizó un subconjunto para entrenamiento ($n = 98$), el otro para propósitos de prueba ($n = 51$). La eficiencia del modelo de regresión seleccionado para predecir y mapear la variación espacial en la salinidad del suelo se muestra por la buena relación ($R^2 = 0,65$) al nivel de probabilidad del 99%, RMSE de 39 dS/m y la normalidad de los residuos. El presente estudio demuestra que la combinación de la banda roja de IKONOS y el índice de salinidad en un modelo de regresión ofrecen un método potencialmente rápido y económico para mapear y modelar el espacio. Aunque este estudio demuestra que el mapeo y el modelado de la salinidad del suelo se pueden realizar con buena precisión sobre la base de imágenes multiespectrales de alta resolución espacial, se necesita más investigación para centrarse en investigar la posibilidad de datos híper espectrales en el mapeo y modelado de la salinidad del suelo en áreas dominadas por palmeras datileras e investigar si puede aumentar la precisión del modelado y el proceso de mapeo.

(Koulla et al. 2019). En su investigación los investigadores afrontaron el problema sobre la Estimación espacial y La predicción de la salinidad ya que es importante para predecir la evaluación de la tierra con el fin de desarrollar y determinar el factor de lixiviación y el manejo preciso para una máxima producción , dando su papel en la comprensión y estimación de la organización del espacio de la salinidad del suelo en la llanura del Bajo Cheliff en Argelia , de donde se extrajo 406 muestras , de las cuales el 80% fueron para el modelado y el 20 % para la validación, en esta investigación se utilizaron 5 modelos regresivos , el Modelo de Regresión Lineal Múltiple(MLP), Kriging Ordinario (OK), Red Neuronal Artificial(ANN), Red Perceptron Multicapa(MLP), Redes de Función de Base Radial(RFB), donde El desempeño de cada modelo se estudió bajo los criterios de rendimiento estadístico del coeficiente de determinación (R^2), el error cuadrático medio (RMSE), y el error absoluto medio (MAE), dando como

resultado la superioridad del modelo MLP estaba implícita con el valor de $R^2 = 0,994$, $RMSE = 0,63$ y $MAE = 0,33$ siendo este el mejor estimador.

(Mohamed & Bahrawi, 2017). En su trabajo de investigación actual se basa en evaluar una regresión correlación entre diferentes índices de vegetación y sus valores espaciales correspondientes de salinidad del suelo conducido en imágenes de satélite. El análisis de componentes principales (PCA) se utilizó para prever los impactos de la salinidad del suelo en el corriente alquiler de vegetación, utilizando cuatro índices el Índice de vegetación de suministro de agua (WSVI), el Índice de vegetación ajustado por suelo (SAVI), Índice de estrés por humedad (MSI), y Índice de infrarrojos de diferencia normalizada (NDII). El propósito de los análisis de regresión es prever los potenciales de regresión entre el índice de salinidad del suelo de un lado y el resto de los índices de sequía hidrológica del otro lado. Los análisis de componentes principales y los análisis de redes neuronales artificiales (ANN) fueron los enfoques implementados como complemento además de que El modelo ANN es imprescindible. Se utilizan cinco criterios diferentes: Coeficiente de correlación de Pearson (R), la raíz del error cuadrático medio (RMSE), la desviación media absoluta (MAD), la probabilidad logarítmica negativa y la suma de cuadrados del error (SSE). Básicamente, RMSE es el parámetro examinado para comparar, dando como resultado que el NDII logro alcanzar un 73 por ciento de precisión.

1.3. Teorías relacionadas al tema.

1.3.1. Conductividad Eléctrica (CE)

La conductividad eléctrica refleja la capacidad conductora de la solución del suelo, y dentro de una cierta concentración de sales solubles actuando como electrolitos fuertes obteniendo una solución que contiene conductividad que puede ser medida reflejando la relación positiva con la conductividad eléctrica, representando directamente el contenido total de salinidad y además siendo este un método preciso en la medición de sal en el suelo, su unidad para medición de conductividad eléctrica es siemens por metro (S / m) (Bado, y otros, 2016).

En la siguiente ecuación:

$$CEe = CEw \times 1.5$$

Donde:

CEe = La conductividad eléctrica de extracto saturado de suelo

CEw = La conductividad eléctrica del agua

Tabla 1.

Clasificación de Salinidad de suelo según valores de conductividad eléctrica (EC).

CE (dS/m)	Efecto sobre la cosecha	Clasificación del suelo
0-2	Comúnmente despreciable	No salino
2-4	Cultivos muy sensibles afectados	Poco Salino
4-8	Muchos cultivos afectados	Medianamente salino
8-16	Solamente cultivos tolerantes crecen bien	Muy salino
>16	Solamente crecen cultivos muy tolerantes	Extremadamente salino

Fuente. Soil Salinity Testing, Data Interpretation and Recommendations

1.3.2. Radiación Electromagnética

Es una onda que se auto propaga en el espacio o a través de la materia, y tiene tanto componentes eléctricos y magnéticos que oscilan en fase normalmente (en un ángulo de 90 °) entre sí y en la dirección de propagación de la onda o energía (Dervić, Sinik, & Despotovic, 2019).

La radiación electromagnética se propaga a la velocidad de la luz, y se denota por la siguiente expresión matemática.

$$c = \lambda * F$$

Donde:

λ = Longitud de onda

F = Frecuencia

Y de acuerdo a la teoría cuántica, se consiste en propiedades energéticas:

$$Q=h*F$$

Donde:

h = Es la constante de Planck

F = Frecuencia

1.3.3. Espectro Electromagnético

También puede expresarse con longitud de onda o frecuencia. Cada forma de pensar sobre el EM está relacionada con las otras de una manera matemática precisa (Butcher, 2010).

Su expresión matemática se denota de la siguiente manera:

$$\lambda = c / \nu$$

La ecuación para determinar la energía es la siguiente

$$E = h \times \nu$$

Donde:

λ = Es la longitud de onda.

ν = Es la frecuencia

E = Energía

c = Es la velocidad de la luz, c a 299,792,458 m/s (186,212 millas/segundo). h = Es la constante de Planck, $h = 6.626 \times 10^{-27}$ erg-segundos

Espectro electromagnético

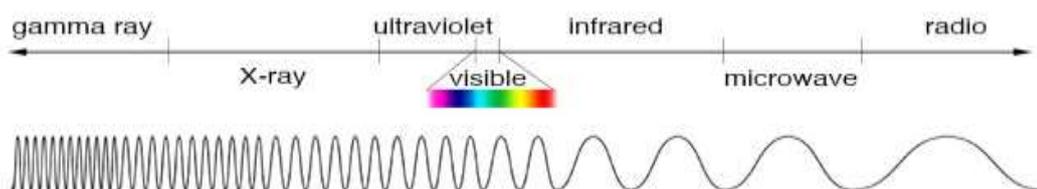


Figura 2. Clasificación del espectro electromagnético

Fuente: NASA's Imagine the Universe

Tabla 2.

Límites del espectro electromagnético según su división.

DIVISION	LIMITES
Rayos Gamma	<0.03nm
Rayos X	0.03–300 nm
Radiación ultravioleta	0.30-0.38 μm
Luz Visible	0.38–0.72 μm
Infrarrojo cercano	0.72–1.30 μm
Infrarrojo medio	7.0–1,000 μm (1 mm)
Radiación microondas	1 mm–30 cm
Ondas de Radio	≥ 30 cm

Fuente. Introduction to Remote Sensing Fift Edition

Tabla 3.

Límites de la luz visible según los colores RGB.

Colores	LIMITES
Azul	0.4 - 0.5 μm
Verde	0.5 - 0.6 μm
Rojo	0.6 - 0.7 μm

Fuente. Introduction to Remote Sensing Fift Edition

1.3.4. Radiancia Espectral (LT)

Toda la energía que irradia una superficie hacia una sola dirección, siguiendo un ángulo sólido, teniendo en cuenta a que es referido a un solo fragmento concreto. (Perez Gutierrez & Muñoz Nieto, 2006).

$$L = \frac{\partial I}{\partial A \cdot \cos \theta}$$

Donde:

L = Radiación

∂I = Magnitud de la radiancia

∂A = Unidad de área

$\cos \theta$ = coseno de la dirección

De modo que la LT se determina por la energía radiada y la extensión o longitud de onda denotada de la siguiente forma

$$LT = \partial L / \partial \lambda$$

Donde:

∂L = Energía radiada

$\partial \lambda$ = Extensión de onda.

1.3.5. Reflectancia Espectral

Es la reflectancia medida en un Ts y, dentro de un pequeño intervalo de longitud de onda (ILO). La reflectancia espectral se denomina típicamente reflectancia monocromática y puede ser especular, difusa o hemisférica.

La reflectancia espectral es la reflectancia medida en un intervalo dado Ts (transmitancia) y λ , dentro de un pequeño ILO, $\Delta\lambda$, centrado en λ . La reflectancia espectral se denomina típicamente reflectancia monocromática y puede ser especular, difusa o hemisférica. Este parámetro se indica con el subíndice “ λ ” después del símbolo ρ . En consecuencia, $\rho_\lambda(\lambda, \theta_i, \theta_r, \varphi, Ts)$ es la reflectancia especular espectral y $\rho_\lambda(\lambda, \theta_i, h, Ts)$ es la reflectancia hemisférica espectral. Un gráfico que representa la reflectancia especular en un rango de λ se llama espectro de reflectancia. (Fernández García, Sutter, Fernández Reche, & Lupfert, 2017).

$$\rho_\lambda = ER(\lambda) / EI(\lambda)$$

Donde:

p_{λ} = Reflectancia espectral

$E_r(\lambda)$ = Energía reflejada

$E_i(\lambda)$ = Energía incidente

1.3.6. Firma Espectral

Es un gráfico que determina el comportamiento que tiene la longitud de onda y la reflectancia, generando unas curvas o trazos que gracias a la examinación por laboratorio, se evalúa cada una de sus características obteniendo así información muy importante (Pérez Gutiérrez & Muñoz Nieto, 2006).

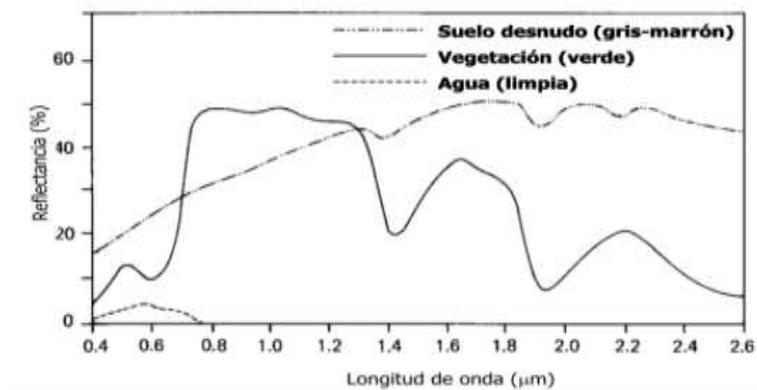


Figura 3. Espectro electromagnético característico del suelo desnudo, vegetación y agua.

μm : micro-metros

Fuente: Teledetección Nociones y Aplicaciones

1.3.7. Teledetección o Detección Remota

Definen a la teledetección es la instrumentación, técnicas y métodos para observar la superficie de la Tierra a una distancia y para interpretar las

imágenes o números valores obtenidos para adquirir información significativa de particular objetos en la Tierra (Buiten & Cleveres,1993).

1.3.7.1. Imágenes multiespectrales de satélite

Son imágenes capturadas mediante un conjunto de detectores fotoeléctricos que miden los valores irradiados provenientes de la superficie de la tierra, que posteriormente son transformados en valores numéricos denominados también como niveles digitales en sus respectivas bandas, obteniendo como resultado final un matriz tridimensional en la que sus valores son interpretados siguiendo el esquema de una matriz denotada por sus filas y columnas además agregando un índice que es el valor de banda del espectro electromagnético.

Las resoluciones que existen son 5 tipos gracias a ellas se evalúan los rangos de precisión con referencia a definiciones extremadamente importantes para el correcto estudio de las imágenes multiespectrales, la resolución espacial que especifica la capacidad que tiene el sistema para diferenciar los objetos de acuerdo a su tamaño interpretando de otro modo como la dimensión del pixel en el área del el terreno, la resolución radiométrica la cual refleja la dimensión del sensor para especificar niveles o intensidades de radiancia, la resolución espectral tiende a indicar el número y la anchura de las bandas espectrales que puede excluir el sensor, la resolución temporal la cual mide el tiempo del satélite al pasar por un punto sobre su línea vertical determinando así el periodo que toma para adquirir de una imagen de la misma zona, la resolución angular es la facultad que tiene un sensor para capturar imágenes trasversales (Pérez Gutiérrez & Muñoz Nieto, 2006).

Tabla 4.

Resolución espacial de satélites.

SATELITE	SENSOR	RESOLUCION ESPACIAL
IKONOS		1m
ORBVIEW		1m
SPOT 5		2.5m
SPOT 3	HRV	10m
LANDSAT 7	ETM+	15m
LANDSAT 8	OLI	30m

Nota. *m*: Metros

Fuente: Teledetección Nociones y Aplicaciones

1.3.7.2. Pre procesamiento de Imágenes multiespectrales de satélite

Particularmente para usar los datos de las imágenes multiespectrales(IM), se hacen dos tipos de correcciones una de ellas se utiliza para rectificar el sistema de coordenadas y la proyección cartográfica, además de otra que perfecciona los valores para hacer una correcta transformación de la imagen eliminando las fallas perfeccionándola de una manera eficaz.

Además, por otro lado, también se debe de hacer una calibración de tipo geométrica y radiométrica, donde una repara cualquier tipo de distorsión y la otra que cambia a valores físicos la imagen (Perez Gutierrez & Muñoz Nieto, 2006).

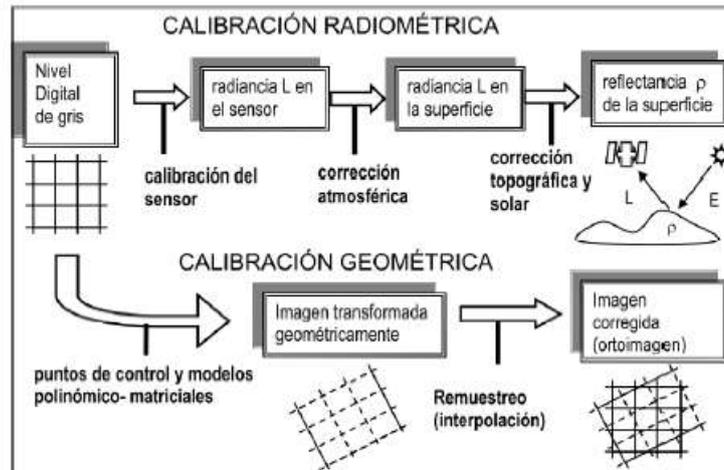


Figura 4. Proceso de calibración radiométrica y geométrica de imágenes.

Fuente: Teledetección Nociones y Aplicaciones

1.3.7.3. Indicadores Espectrales

Se define como ecuaciones, que a partir de información que se recluta de las bandas espectrales, que los investigadores utilizan para pronosticar, modelar y deducir sucesos en la superficie de la tierra.

a. Indicadores de Vegetación

Son variables que se calculan a partir de los datos de reflectividad en diferentes longitudes de onda con la finalidad de conocer información que se relacione con la vegetación. (Pérez Gutiérrez & Muñoz Nieto, 2006)

Indicador de Vegetación de Diferencia Normalizada (NDVI)

Rouse et al. (1974)

$$NDVI = \frac{pNIR - pR}{pNIR + pR}$$

Donde:

$pNIR$ = Reflectancia en la banda infrarrojo cercano

pR = Reflectancia en la banda roja

b. Indicadores de Vegetación

A través de índices aproximados que son utilizados con el fin de adquirir datos informáticos respecto al grado de salinidad que se encuentran en los suelos (Perez Gutierrez & Muñoz Nieto, 2006).

$$SI = \sqrt{pR \times pNIR} \quad (\text{Abdellatif \& Mourad, 2012})$$

$$SI = \frac{pNIR - pSWIR}{pNIR + pSWIR} \quad \text{Asfaw et al. (2016)}$$

$$SI = \frac{pG \times pR}{pB} \quad \text{Elhag (2016)}$$

Donde:

pB = Reflectancia en la banda azul

pG = Reflectancia en la banda verde

pR = Reflectancia en la banda roja

$pNIR$ = Reflectancia en la banda infrarrojo cercano

$pSWIR$ = Reflectancia en la banda infrarroja de onda corta

1.3.8. Técnicas de estimación

Por medio de un determinado grupo de cualidades que se representan por un número, el cual es tomado como valor de entrada y un resultado, salida que también es representado por un valor numérico, por el cual podemos determinar la relación entre variable dependiente e independiente y obtener el resultado deseado, a fin de poder predecir resultados mediante un grupo de distintas cualidades. (Sarkar, 2018)

a) Regresión Lineal Simple

Es un modelo que es muy utilizado a fin de explicar las relaciones entre dos variables, una explicativa y una de respuesta, por medio de la técnica de regresión de variables; Y sobre X, se busca una aproximación de una nube de puntos (x,y). (Limeres, 2011-2012)

Consideramos el modelado entre la variable dependiente y una independiente. Cuando solo hay una variable independiente, el modelo generalmente se denomina como un modelo lineal simple de regresión. (Shalabh)

Mediante una curva del tipo: $Y = f(X)$

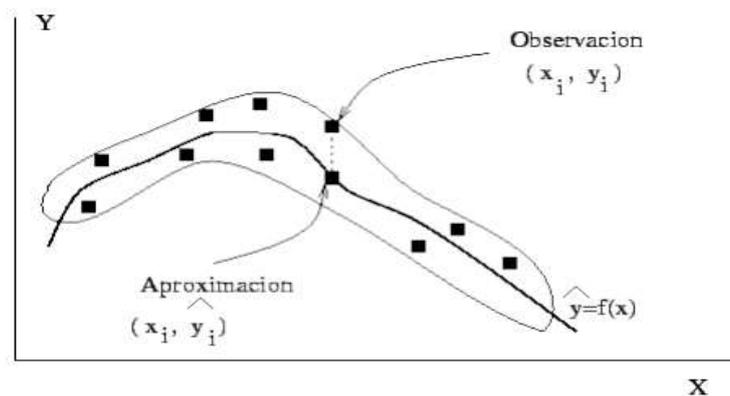


Figura 5. Representación gráfica de la curva de regresión.

Fuente: Regresión Lineal Simple

El modelo RSL tiene la siguiente expresión:

$$Y = \alpha + \beta X + \varepsilon$$

Donde

α = ordenada del origen

β = pendiente de la recta

ε = variable que incluye conjunto de factores

Este modelo consta de una sola variable predictora, la cual nos permite realizar predicciones de posibles respuestas basadas en valores de la variable predictora (Carrasquilla-Batista, 2016)

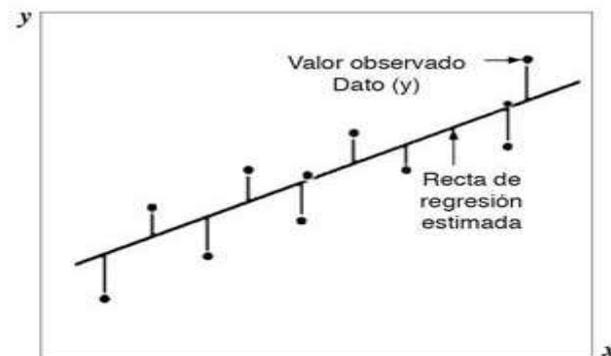


Figura 6. Representación de los mínimos cuadrados.

Fuente: Regresión Lineal Simple

Como mínimo deben existir al menos dos de los factores son numéricos, es posible preparar un base de respuesta a fin de predecir Y con múltiples combinaciones de los factores de diseño. Este modelo es generalmente usado con el fin de ajustar los modelos a los valores experimentales (Carrasquilla-Batista, 2016)

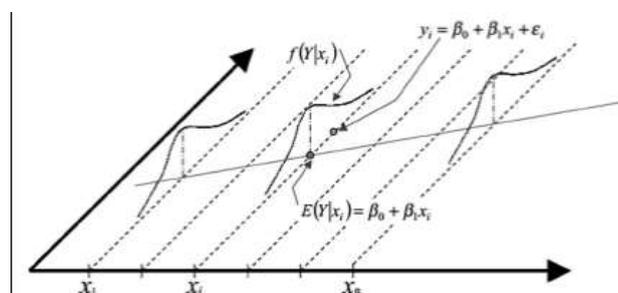


Figura 7. Valor esperado de Y dado x_1

Fuente: Regresión Lineal Simple

b) Regresión Lineal Múltiple

Es un modelo lineal donde la respuesta (Y) es determinada mediante un conjunto de predictores (X1, X2, X3...)" (Rodrigo, 2016)

Es igual al modelo de regresión lineal simple, con la variación que en este modelo intervienen más de una variable explicativa (Rojo, 2007)

Y los cuales se denotan en la ecuación siguiente:

$$Y_i = (\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_n X_{ni}) + e_i$$

β_0 : el valor de la variable.

β_i : es el promedio que incrementado en uno a la variable predictora x sobre y. Se conocen como coeficientes parciales de regresión.

e_i : es el error, la diferencia entre el valor observado y el estimado por el modelo.

c) Random Forest Regression (RF)

Es la combinación de árboles predictores de modo que cada uno de ellos es afectado por los valores muestrales de un vector al azar de manera independiente y con igual distribución para cada uno de los árboles que pertenecen al bosque. (Breiman, 2001)

Para regresión están formados por árboles en crecimiento dependiendo de un vector aleatorio Θ tal que el predictor de árbol $h(x, \Theta)$ toma valores numéricos a diferencia de las etiquetas de clase. (Breiman, 2001)

Los valores de salida son numéricos y asumimos que el conjunto de entrenamiento se extrae independientemente de la distribución de la distribución aleatoria del vector Y, X.

El error de generalización cuadrático medio para cualquier predictor numérico $h(x)$ es:

$$E_{X,Y}(Y - h(X))^2$$

El predictor de bosque aleatorio se forma tomando el promedio sobre k de los árboles

$\{h(x, \Theta_k)\}$. De manera similar al caso de clasificación, se cumple lo siguiente:

$$E_{X,Y}(Y - \text{avg}_k h(X, \Theta_k))^2 \rightarrow E_{X,Y}(Y - E_{\Theta} h(X, \Theta))^2$$

d) Decision Tree Regression (Dt)

Se construye mediante particiones recursivas: a partir del nodo raíz (primer padre), cada nodo se puede dividir en nodos secundarios izquierdo y derecho. Estos nodos pueden luego dividirse aún más y ellos mismos se convierten en nodos padres de sus nodos secundarios resultantes. (Li, 2019)

Es el algoritmo más simple, pero el más poderoso en el aprendizaje automático. El árbol de decisión utiliza un diagrama de flujo como una estructura de árbol para predecir la salida sobre la base de la entrada. (Chugh, 2019)

Se llega a una estimación al hacer una serie de preguntas a los datos, cada pregunta reduce nuestros valores posibles hasta que el modelo adquiere la confianza suficiente para hacer una sola predicción. En regresión normalmente usa el error cuadrático medio (MSE) para decidir dividir un nodo en dos o más sub nodos. (DRAKOS, 2019)

1.3.9. Suelos de Escasa Vegetación.

La usabilidad del suelo se determina por distintos tipos de vegetación, a tal medida que al mencionar sobre escasez se denota que para clasificado como tal; la cubierta del suelo tiene que tener un 90 % del área sin vegetación

o suelo desnudo, además se tiene que constar de distintas características del suelo. (Hernandez et al., 2009)

1. Suelo contextura arenosa.
2. Suelo contextura limosa o polvorosa.
3. Suelo contextura Arcillosa usualmente de color rojiza también conocidos como Luvisoles.

1.4. Formulación del Problema.

¿Qué técnica será más efectiva para estimar el grado de salinidad en imágenes multiespectrales en suelos de escasa vegetación?

1.5. Justificación e importancia del estudio.

Identificar el grado de salinidad en terrenos bien extensión tiene un costo elevado si se trabaja con los métodos tradicionales. Lo cual conlleva a buscar alternativas más económicas.

En lo económico, detectar el grado de salinidad utilizando técnicas de teledetección y pruebas supondría un menor presupuesto que usando métodos clásicos, por consiguiente, esto beneficiaría a los agricultores, los cuales pueden aplicar el cultivo selectivo utilizando la tolerancia salina y salinidad de cada terreno.

Esta investigación persigue el objetivo de estimar el grado de salinidad en grandes extensiones de terreno con suelos de escasa vegetación, utilizando técnicas de teledetección en imágenes multiespectrales tomadas por un satélite, las cuales servirán para comparar las técnicas de estimación y de esa forma encontrar el mejor modelo para estimar la conductividad eléctrica, la cual nos permitirá medir el grado de salinidad del suelo.

1.6. Hipótesis.

La técnica MLR, es más efectiva para estimar el grado de salinidad en imágenes multiespectrales en suelos de escasa vegetación

1.7. Objetivos.

1.7.1. Objetivo general.

Comparar técnicas de estimación del grado de salinidad en suelos de escasa vegetación mediante el procesamiento de imágenes multiespectrales de satélite.

1.7.2. Objetivos específicos.

1. Elegir las técnicas de estimación a utilizar.
2. Obtener imágenes multiespectrales del área de estudio
3. Efectuar calibraciones atmosféricas a las imágenes multiespectrales adquiridas.
4. Elegir indicadores de salinidad y vegetación.
5. Recoger muestras de suelos de escasa vegetación y ejecutar la prueba de conductividad eléctrica.
6. Comparar las técnicas de estimación.

II. MATERIAL Y MÉTODO

2.1. Tipo y Diseño de Investigación.

2.1.1 Tipo de la Investigación.

Tipo de investigación es cuantitativo, puesto que la índole se deriva de criterio científico basado en temas de ciencias de la computación. Cabe mencionar, que se utilizarán datos numéricos para su posterior procesamiento y análisis estadístico, los cuales estarán relacionados con las variables descritas en el apartado de operacionalización.

2.1.2 Diseño de Investigación.

Diseño cuasi experimental, puesto que se estudia las relaciones causa-efecto entre variables independiente y dependiente respectivamente, puesto que cualquier alteración en la variable independiente tiene reacción en la variable dependiente, por lo expuesto anteriormente, la investigación es cuasi experimental dado que las técnicas de estimación de grado de salinidad afectan directamente a la estimación del grado de salinidad.

2.2. Población y muestra.

Población: la población está determinada por 9 técnicas de estimación a través de información obtenida de imágenes multiespectrales. (Anexo 2)

Muestra:

La muestra fue determinada de manera no probabilística por conveniencia, y está conformada por cuatro técnicas de estimación que son: Multiple Linear Regression, Simple Linear Regression, Random Forest Regression, Decision Tree Regression

2.3. Variables, Operacionalización.

Tabla 5.

Variables, Operacionalización.

Variable Independiente	Dimensión	Indicador	Formula	Descripción	Técnica e instrumento de recolección de datos
Técnicas de estimación	Precisión de la técnica de estimación	Coeficiente de determinación	$R^2 = \left(\frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2 + \sum_{i=1}^n (Y_i - \bar{Y})^2} \right)^2$	X_i = valor medido Y_i = valor estimado \bar{X} = la media de valores medidos \bar{Y} = la media de valores estimados n = el número de muestras	Registro electrónico Ficha de observación
		Raíz del Error cuadrático medio	$RMSE = \sqrt{\frac{\sum_{i=1}^n (X_i - Y_i)^2}{n}}$	X_i = valor medido Y_i = valor estimado n = el número de muestras	
Grado de salinidad en suelos de escasa vegetación	Grado de Salinidad	Conductividad Eléctrica (CE)	$CE_e = CE_w \times 1.5$	CE_e = Conductividad eléctrica de extracto saturado de suelo CE_w = Conductividad eléctrica del agua	

2.4. Técnicas e instrumentos de recolección de datos, validez y confiabilidad.

2.4.1 Abordaje Metodológico

Sera cuasi experimental, de acuerdo a Hodrick, (1993), puesto que se busca probar la existencia de una relación entre dos o más variables.

2.4.2 Técnicas de Recolección de Datos

Se aplicó recopilación documental, puesto que se realizó revisiones de artículos científicos, libros y papers.

Se aplicará observación científica, puesto que se hará uso de las imágenes de satélite multiespectrales a fin de satisfacer nuestra investigación.

2.5. Procedimiento de análisis de datos.

De acuerdo al tema de investigación, se desarrolló un método para comparar las técnicas de estimación mediante la medición de los siguientes indicadores:

El coeficiente de correlación (R^2)

$$R^2 = \left(\frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2 + \sum_{i=1}^n (Y_i - \bar{Y})^2} \right)$$

Donde:

X_i = valor medido

Y_i = valor estimado

\bar{X}_i = la media de valores medidos

\bar{Y}_i = la media de valores estimados

n = el número de muestras

RMSE

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (X_i - Y_i)^2}{n}}$$

Donde:

X_i = valor medido

Y_i = valor estimado

n = el número de muestras

Conductividad Eléctrica (CE)

$$CE_e = CE_w \times 1.5$$

Donde:

CE_e = Conductividad eléctrica de extracto saturado de suelo

CE_w = Conductividad eléctrica del agua

2.6. Criterios éticos.

Se utilizó el principio ético de propiedad intelectual, dado que se referencio y cito trabajos de investigación, los libros y papers, que se nos permitieron el conocimiento base para el desarrollo del presente proyecto de investigación.

2.7. Criterios de Rigor Científico.

Fiabilidad: se cumplió con las expectativas plasmadas en su contenido y su implementación.

Validez: La información utilizada para la investigación fue analizada y evaluada por personas capacitadas y especializadas en el tema y de esa forma confirmar su veracidad.

III. RESULTADOS.

3.1. Resultados en Tablas y Figuras.

En principio se tuvo como hipótesis para este proyecto de investigación, que MLR sería la técnica con los mejores resultados, basándonos en los resultados de la investigación de la cual se extrajo la población (Anexo 2), los resultados obtenidos fueron distintos; DTR obtuvo mejores resultados en la estimación de salinidad, verificado por los indicadores de R^2 y RMSE.

En el indicador de operacionalización R^2 la técnica DTR obtuvo 0.95, la segunda técnica con mejor rendimiento fue RFR con 0.88, mientras que SLR y MLR solo obtuvieron un 0,52.

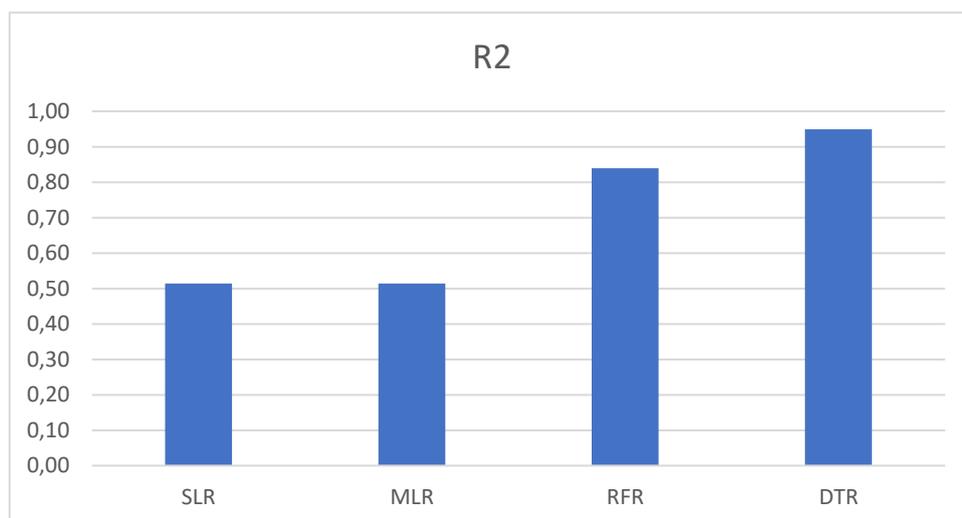


Figura 8.

Figura 8. Comparación de técnicas de estimación en el indicador de precisión R^2

Fuente: Elaboración propia

En el indicador de operacionalización RMSE la técnica DTR obtuvo un error de 0.11, la técnica RFR 0.18 mientras que SLR y MLR obtuvieron 0.33, en error en medida de Conductividad Eléctrica (EC).

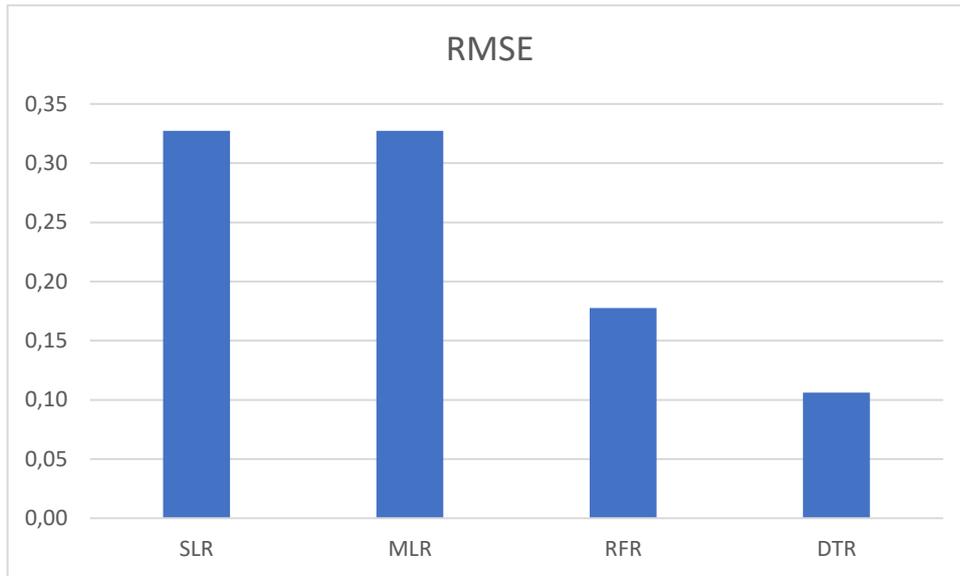


Figura 9. Comparación de técnicas de estimación en el indicador de precisión RMSE

Fuente: Elaboración propia

3.2. Discusión de resultados.

Al analizar los resultados obtenidos con respecto a la precisión R^2 (**Figura 8**), la técnica DTR tiene el mayor porcentaje de precisión seguido de RFR y finalmente SLR y MLR obtuvieron el mismo porcentaje de precisión.

Para el indicador RMSE, analizamos la **Figura 9**, la cual demuestra que las técnicas SLR y MLR obtuvieron el error promedio más alto en la estimación, así mismo el RFR y DTR obtuvieron un menor error promedio.

Se esperaba que la técnica MLR tuviera un mejor desempeño, de acuerdo a los trabajos de investigación previos.

Las **Figura 8 y Figura 9**, demuestran que a mayor porcentaje de precisión menor es el error promedio en la estimación de la salinidad.

3.3. Aporte práctico.

3.3.1. Elegir las técnicas de estimación a utilizar.

La siguiente población se determinó para esta investigación:

Tabla 6.

Población de técnicas de estimación.

TÉCNICAS DE ESTIMACIÓN	PRECISIÓN	
	R^2	Error relativo
PARTIAL LEAST SQUARE REGRESSION (PLSR)	89.00%	
	87.00%	
MULTIPLE LINEAR REGRESSION (MLR)	52.00%	
	36.00%	
	65.00%	
RANDOM FOREST REGRESSION (RFR)	79.00%	
	98.30%	
	85.00%	
	83.00%	
SUPPORT VECTOR MACHINE (SVM)	61.70%	37.00%
BOOSTED BINARY REGRESSION (BBRT)	82.00%	
	81.00%	
	93.00%	
ARTIFICIAL NEURONAL NETWORK (ANN)	69.00%	
	99.70%	
	41.90%	49.00%
DECISION TREE REGRESSION (DTR)	80.00%	
PRINCIPAL COMPONENT REGRESSION (PCR)	79.00%	
SIMPLE LINEAR REGRESSION (SLR)	87.00%	

Fuente. Teledetección Nociones y Aplicaciones

Con base en los resultados obtenidos en las investigaciones anteriores que sirvieron para concluir la tabla anterior, se propuso compara las técnicas de estimación; **SLR**, **MLR**, **RFR**, **DTR**. Los cuales obtuvieron resultado por encima del 80% en la estimación de salinidad de suelos mediante el procesamiento de imágenes multiespectrales. (Anexo 2)

3.3.2. Obtener imágenes multiespectrales del área de estudio

El principal objetivo que persigue esta investigación es comparara las técnicas de estimación a fin de lograr estimar el grado de salinidad en grandes extensiones de suelos de escasa vegetación, para la cual fue necesario utilizar imágenes multiespectrales del área de interés, dichas imágenes pudimos obtenerlas del satélite Landsat, mediante su descarga.

El satélite Landsat brinda datos calibrados de la superficie de la tierra desde hace poco más de 40 años, los cuales se dan en forma de imágenes multiespectrales, en 2013 fue su lanzamiento. Landsat 8 consta de múltiples capturas de territorios alrededor de todo el mundo y el de acceso gratuito, en su repositorio, a los cuales podemos acceder a través del United States Geological Survey (USGS) y Earth Resources Observation and Science (EROS). (Gonzaga Aguilar, 2014)

El satélite Landsat 8 hace una trayectoria alrededor de la tierra con una elevación de 705 kilómetros, en una franja de 185 kilómetros, con un movimiento de norte a sur sobre el lado de iluminación terrestre en una órbita sincrónica solar, siguiendo el Worldwide Reference System (WRS-2). (U.S D. o., 2018)

El satélite Landsat 8 genera cada 99 minutos una órbita completa, 14 órbitas diarias y cruza el mismo punto cada 16 días, esta órbita permite que la cobertura del satélite sea repetida por 8 días en cualquier área del planeta. Asimismo, tiene dos sensores, el Operational Land Imager (OLI) y el Thermal Infrared Sensor (TIRS), teniendo una corrección de datos radiométricamente y co-registrados en una proyección cartográfica, permitiendo de esta manera hacer correcciones en la translación de terreno obteniendo como resultado una orto rectificación estándar de imagen digital (U.S D. o., 2018).

Por las facilidades para el acceso a la información y principalmente uso gratuito de descarga, se usó Landsat 8 para la creación del dataset de imágenes multiespectrales.

El proceso que seguimos para la descarga de imágenes del satélite Landsat 8, se siguió los pasos del **anexo 3**, se realizó el filtró de paquetes de imágenes por porcentaje de cobertura de nubes.



Figura 10. Comparación de técnicas de estimación en el indicador de precisión RMSE

Fuente: Elaboración propia

Se de el paquete de imágenes del 03 de febrero del 2021, por ser la que más cercana a la fecha de recogida de muestras de suelo, se obtuvo una carpeta con el paquete de imágenes con la siguiente estructura.



Figura 11. Estructura de la carpeta de las imágenes adquiridas de Landsat

Fuente: Elaboración propia

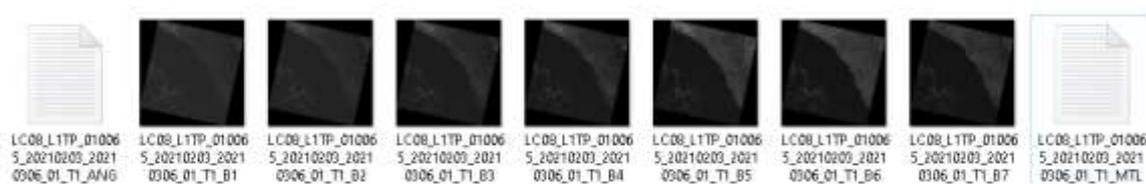


Figura 12. Imágenes multiespectrales adquiridas del satélite Landsat 8

Fuente: Elaboración propia

El formato .TIF muestra las imágenes multiespectrales y las imágenes provenientes del sensor OLI a utilizar se detallan en el siguiente cuadro, la resolución de cada pixel de la imagen es de 30m x 30m o 900m², la imagen tiene una composición de 7761 x 7611 pixeles, el formato UTM 17N, es utilizado para georreferenciar los pixeles.

Tabla 7.

Bandas utilizadas del sensor OLI del satélite Landsat 8.

Identificador de la banda	Banda del espectro electromagnético	Límites de espectro electromagnético (um) micras	Aplicación y uso
B1	Coastal	0.435 – 0.451	Estudios costeros
B2	Blue – Azul	0.452 – 0.512	Distinguir suelos de vegetación
B3	Green – Verde	0.533 – 0.590	Vigor de las plantas
B4	Red – Roja	0.636 – 0.673	Pendientes de la vegetación
B5	NIR – Infrarrojo cercano	0.851 – 0.879	Contenido de biomasa
B6	SWIR – 1	1.566 – 1.651	Humedad del suelo y vegetación
B7	SWIR – 2	2.107 – 2.294	Humedad del suelo y Vegetación.

Identificador de la banda: Identificadores de una banda del espectro electromagnético en la data de imágenes.

Límites del espectro electromagnético: Valores comprendidos de las bandas electromagnéticas

Fuente: Landsat FAQs, USGS

En la tabla 8 se explica el contenido de los dataset que están en extensión .txt y que vienen dentro del paquete de imágenes.

Tabla 8.

Dataset del sensor OLI del satélite Landsat 8.

Archivo	Banda del espectro electromagnético
MTL.TXT	- Ángulo de inclinación del satélite al capturar la zona. -Factores de cálculo de reflectancia y puntos de ubicación geográfica.
ANG.TXT	-El tamaño del pixel por imagen, -Coeficientes de cálculos de reflectancia.

Fuente: Landsat FAQs, USGS

Características de las imágenes:

Se seleccionó como características a las imágenes de la banda 1 a la banda 7 cuya resolución espacial de 30 x 30 (metros), que equivale a 900 m² de terreno por pixel, los cuales están compuestos por 16 bits, donde el pixel tiene como rango un valor de 0 a 65536, y con 112MB (Megabytes) de peso aproximado en almacenamiento. Con tamaño alrededor de 170 km de norte a sur por 183 km de este a oeste, de escena (imagen capturada de Landsat). Con una composición de imagen de 7761 y 7611. (U.S D. o., 2018).

Una vez obtenidas las imágenes se procedió con la calibración atmosférica, a fin de obtener los datos que nos permiten la comparación entre técnicas de estimación, puesto que las imágenes descargadas contienen números digitales, las cuales no son de utilidad para realizar las comparaciones de las técnicas, por lo cual es necesario calcular la reflectancia por pixel de cada imagen.

Los datos que se obtienen de Landsat 8 se representan por valores digitales escalados cuantificados y calibrados (ND), obteniendo los datos de imágenes multiespectrales, recolectados por el sensor Operational Land Imager (OLI), entregando un formato entero de 16 bits sin signo. (U.S D. o., 2018).

3.3.3. Efectuar calibraciones atmosféricas a las imágenes multiespectrales adquiridas.

Según (U.S D. o., 2018) en la documentación del satélite Landsat 8, muestra ecuaciones para convertir valores digitales a radiancia y reflectancia que posteriormente permiten hacer las calibraciones atmosféricas, las cuales son las siguientes:

Conversión de pixeles números digitales a radiancia.

$$L\lambda = MLQcal + AL$$

Donde:

$L\lambda$ = Radiancia espectral al nivel de la superficie

ML = Factor de re escalado multiplicativo presente en la Metadata
RADIANCE_MULT_BAND_x

AL = Factor de re escalado aditivo presente en la Metadata
RADIANCE_ADD_BAND_x

$Qcal$ = El valor digital (ND) del pixel

Conversión de pixeles de valores digitales a reflectancia a nivel de la superficie.

$$p\lambda' = MpQcal + Ap$$

Donde:

$p\lambda'$ = Radiancia espectral al nivel de la superficie

MP = Factor de re escalado multiplicativo presente en la Metadata
REFLECTANCE_MULT_BAND_x

AP = Factor de re escalado aditivo presente en la Metadata
REFLECTANCE_ADD_BAND_x

Q_{cal} = El valor digital (ND) del pixel

Corrección de reflectancia respecto al ángulo solar

$$p\lambda = p\lambda' / \cos(\theta_{sz}) = p\lambda' / \sin(\theta_{SE})$$

Donde:

$p\lambda'$ = Reflectancia con calibración atmosférica

θ_{SE} = Angulo solar de elevación

θ_{sz} = Angulo solar zenith

Se realizo una aplicación de conversión de valores digitales a reflectancia, aplicando una corrección de Angulo solar, transformando los valores de cada pixel de banda, como se observa en el ejemplo del **Anexo 4**, para hacer más fácil el proceso, se desarrolló un pequeño algoritmo que realiza el proceso de forma automática **Anexo 5**.

Codigo donde se calculo la reflectancia a travez de los valores de las bandas que se extrajeron de la metadata, se infresa la matriz de donde contiene los valores de los pixeles.

```
def calcularReflectancia(self, matriz):  
    sun_correction = math.sin(math.radians(self.banda.SUN_ELEVATION))  
    self.reflectancia = ((matriz * self.banda.REFLECTANCE_MULT_BAND) +  
        self.banda.REFLECTANCE_ADD_BAND) / sun_correction
```

Las imágenes procesadas se especifican en el siguiente cuadro:

Tabla 9.

Identificador de Bandas de las Imágenes del satélite Landsat 8.

IDENTIFICADOR	IMAGEN	BANDA
<i>pB</i>	BAND_2	Banda Azul
<i>pG</i>	BAND_3	Banda Verde
<i>pR</i>	BAND_4	Banda Roja
<i>pNIR</i>	BAND_5	Banda infrarrojo cercano
<i>pSWIR</i>	BAND_6	Banda infrarroja de onda corta

Fuente: Elaboración propia.

3.3.4. Elegir indicadores de salinidad y vegetación.

Posteriormente de convertir los valores de los pixeles de cada imagen a reflectancia superior atmosférica (TOA), se seleccionó los indicadores de salinidad, bandas y vegetación, que guardan relación con los valores de salinidad, los indicadores utilizados se obtuvieron de las siguientes investigaciones:

En el trabajo de investigación (Mousavi et al, 2017), comparó 17 indicadores de vegetación, salinidad, y bandas espectrales con el fin de encontrar una relación entre estos indicadores y los datos de conductividad eléctrica.

Tabla 10

Correlación entre los indicadores de salinidad, vegetación y bandas, y el valor de conductividad eléctrica.

INDICADOR	Coefficiente de correlación (R²)
ASTER-SI	0.20 - 0.40
SI1	0.20 - 0.40
SI2	0.00 - 0.20
SI3	0.20 - 0.40
SI11	0.20 - 0.40
INT1	0.00 - 0.20
B2(Banda azul)	0.40 - 0.50
B3(Banda verde)	0.30 - 0.40
B4(Banda Roja)	0.20 - 0.30

Fuente. Digital Mapping of Topsoil Salinity Using Remote Sensing

Indices in Agh-Ghala Plain, Iran.

Llegando a la conclusión, que en el trabajo de investigación se utilizaran los indicadores NDVI, DVI, BI, SI porque presenta una correlación que sobrepasa a 0.30 con relación a los valores de conductividad eléctrica.

(Allbed, Amal, Lalit Kumar, and Priyakant Sinha, 2014). En su investigación se utilizó un indicador de salinidad (SI).

Los indicadores estudiados en los puntos anteriores presentan las siguientes formulas:

$$NDVI = \frac{p_{NIR} - p_R}{p_{NIR} + p_R}$$

$$SI = \sqrt{p_B * p_R^2}$$

$$DVI = p_{NIR} - p_R$$

$$SI_{11} = \frac{p_{NIR}}{p_{SWIR}}$$

$$SI2 = \frac{p_R}{p_{NIR}}$$

$$SI_{1} = \sqrt{p_B * p_R}$$

$$BI = \sqrt{p_{NIR}^2 + p_R^2}$$

$$ASTER SI = \frac{p_{NIR} - p_{SWIR}}{p_{NIR} + p_{SWIR}}$$

$$SI 3 = \frac{p_G * p_R}{p_B}$$

Donde:

p_B = Reflectancia en la banda azul

p_G = Reflectancia en la banda verde

p_R = Reflectancia en la banda roja

p_{NIR} = Reflectancia en la banda infrarrojo cercano

p_{SWIR} = Reflectancia en la banda infrarrojo de onda corta

3.3.5. Recoger muestras de suelos de escaza vegetación y ejecutar la prueba de conductividad eléctrica.

Se decidió tomar como zona de interés el distrito de Mórrope, el cual pertenece a la provincia de Lambayeque. Se tomaron muestra de la tierra y luego fueron llevadas al laboratorio de suelos para realizar las pruebas de Conductividad Eléctrica (CE), puesto que la prueba de CE nos permite medir la concentración de sales de una solución con respecto a un extracto de solución del suelo.

Se recogieron 30 muestras en 30 puntos distintos, los cuales fueron recogidos a 30 cm de profundidad del suelo, las cuales se geo referenciaron mediante un Sistema de Posicionamiento Global (GPS), a través del sistema de coordenadas UTM-WGS84, zona 17S, el cual se representa en la siguiente imagen de Google Map.



Figura 13. Ubicación de las muestras recogidas en la ciudad de Mórrope

Fuente: Elaboración propia

La distancia aproximada entre cada una de las muestras es de 30m x 30m, que es la dimensión que representa cada pixel de las imágenes multiespectrales del satélite Landsat 8, a fin de que las muestras obtenidas representen en distintos pixeles en suelos de escasa vegetación, se consideró tomar 6 muestras de un área distinta al de las 24 primeras por recomendación de los trabajadores de la zona.

Se utilizó un GPS modelo Garmin Montana 650t dado que nos ofrece una precisión de +/- 3.65m, la cual es muy buena con respecto al tamaño de pixel (30mx30m) antes descrito.

El proceso que se aplicó a cada muestra fue la siguiente:

1. Se procedió a secar las muestras recogidas, exponiéndolas en un recipiente a la intemperie.
2. Se realizó un procedimiento para deshacer los agregados del suelo, para el cual hicimos uso de un rodillo.
3. Se procedió a tamizar cada una de las muestras, para la cual se utilizó el tamiz N° 10, para separar la tierra fina de elementos gruesos, posteriormente

se pesó y finalmente en un depósito plástico se agregó 500 gramos de suelo tamizado.

4. Se añadió agua destilada al depósito de plástico, y se realizó un proceso de batido, usando una espátula hasta lograr una pasta de saturación.
5. Con ayuda de un embudo se depositó en un papel filtro y se añadió la pasta saturada, y se extrajo una porción de suelo mediante una bomba al vacío.
6. Finalmente haciendo uso conductímetro, se procede a calcular la conductividad eléctrica de suelo.

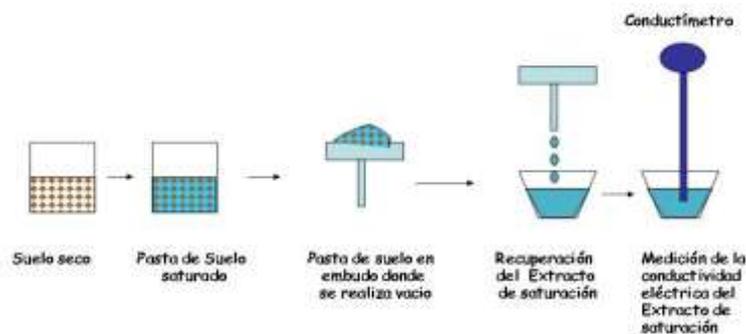


Figura 14. Procedimiento para la medición de Conductividad eléctrica

Fuente: Elaboración propia

Tabla 11.

Correlación entre los indicadores de salinidad, vegetación y bandas, y el valor de conductividad eléctrica

N	Ubicación UTM 17 S		Valor de Conductividad Eléctrica (CE)	Grado de Salinidad
1	617655,	9266050	1.574	No salino
2	617634,	9266062	1.368	No salino
3	617603,	9266069	2.482	Poco Salino
4	617573,	9266077	0.713	No salino
5	617564,	9266047	1.621	No salino
6	617553,	9266011	0.601	No salino
7	617548,	9265978	0.746	No salino
8	617543,	9265941	0.994	No salino
9	617634,	9266024	0.947	No salino
10	617633,	9265993	0.311	No salino
11	617627,	9265958	14.47	Muy salino
12	617612,	9265929	38.9	Extremadamente salino

13	617599,	9265897	84.5	Extremadamente salino
14	617659,	9266019	0.917	No salino
15	617653,	9265984	7.21	Medianamente salino
16	617646,	9265943	45.3	Extremadamente salino
17	617633,	9265908	104.4	Extremadamente salino
18	617550,	9265897	4.2	Medianamente salino
19	617614,	9265886	166.4	Extremadamente salino
20	617612,	9265842	134.6	Extremadamente salino
21	617549,	9265809	42.4	Extremadamente salino
22	617549,	9265815	46.0	Extremadamente salino
23	617564,	9265840	43.6	Extremadamente salino
24	617550,	9265868	28.0	Extremadamente salino
25	617757,	9266011	83.6	Extremadamente salino
26	617768,	9266039	48.4	Extremadamente salino
27	617778,	9266065	71.6	Extremadamente salino
28	617806,	9266057	110.4	Extremadamente salino
29	617797,	9266026	77.7	Extremadamente salino
30	617784,	9265998	117.9	Extremadamente salino

Fuente: Elaboración propia

Se ubicó las muestras en formato UTM 17S en las imágenes multiespectrales, usando las coordenadas tomadas del GPS (**Tabla 11**), para lo cual se precedió con la conversión de las coordenadas en UTM 17S a coordenadas UTM 17N, puesto que el Satélite Landsat 8 tiene como formato las coordenadas UTM 17N.

Para lograr la extracción de los indicadores, se utilizó de guía los pasos del (**Anexo 6**). Y se implementó un algoritmo que permite realizar el proceso de forma automática, mediante el uso de GDAL, librería de Python (**Anexo 7**)

La extracción de los indicadores, permitió obtener la siguiente tabla, mediante la cual se relacionan el valor de conductividad eléctrica y los indicadores.

Tabla 12.

Relación entre los indicadores elegidos y el valor de conductividad eléctrica.

MUESTRAS	RED	GREEN	BLUE	NDVI	DVI	BI	ASTERSI	SI	SI1	SI2	SI3	SI11	CE
m1	0,158590	0,144398	0,147015	0,132112	0,048282	0,214480	-0,045002	0,414928	0,152693	0,766610	0,155767	0,913872	1.574
m2	0,158590	0,144398	0,147015	0,132112	0,048282	0,214480	-0,045002	0,414928	0,152693	0,766610	0,155767	0,913872	1.368
m3	0,152696	0,139824	0,142370	0,146415	0,052384	0,207044	-0,049497	0,407046	0,147443	0,744568	0,149966	0,905674	2.482
m4	0,133105	0,131549	0,134402	0,257496	0,092320	0,187142	0,073176	0,390024	0,133752	0,590462	0,130280	1,157907	0,713
m5	0,093546	0,116438	0,117899	0,479845	0,172594	0,149361	0,298706	0,355880	0,105019	0,351493	0,092387	1,851870	1.621
m6	0,097460	0,118041	0,119785	0,469966	0,172830	0,153075	0,223912	0,359561	0,108047	0,360576	0,096040	1,577029	0,601
m7	0,114740	0,126363	0,127966	0,401427	0,153899	0,170684	0,181319	0,375675	0,121173	0,427117	0,113303	1,442953	0,746
m8	0,148359	0,138905	0,140602	0,192998	0,070961	0,203236	0,027275	0,403252	0,144428	0,676448	0,146567	1,056079	0,994
m9	0,162834	0,146614	0,147840	0,145173	0,055307	0,219113	-0,032619	0,417558	0,155156	0,746461	0,161483	0,936823	0,947
m10	0,176507	0,158402	0,153734	0,160556	0,067519	0,237162	-0,022845	0,429987	0,164727	0,723312	0,181867	0,955330	0,311
m11	0,183155	0,163635	0,158048	0,183886	0,082537	0,245606	0,001244	0,437714	0,170139	0,689352	0,189630	1,002491	14,47
m12	0,179973	0,161419	0,156539	0,214932	0,098544	0,241757	0,014861	0,434660	0,167848	0,646182	0,185583	1,030171	38,9
m13	0,175470	0,157718	0,152461	0,225575	0,102222	0,235934	0,019253	0,428077	0,163561	0,631887	0,181521	1,039262	84,5
m14	0,165946	0,150551	0,149797	0,168949	0,067472	0,224061	-0,004625	0,421111	0,157664	0,710938	0,166781	0,990794	0,917
m15	0,176507	0,158402	0,153734	0,160556	0,067519	0,237162	-0,022845	0,429987	0,164727	0,723312	0,181867	0,955330	7,21
m16	0,191006	0,169340	0,160735	0,181285	0,084588	0,255264	-0,010245	0,444093	0,175218	0,693071	0,201231	0,979718	45,3
m17	0,178346	0,158708	0,153592	0,190823	0,084116	0,238737	-0,004070	0,430580	0,165507	0,679511	0,184286	0,991892	104,4
m18	0,158684	0,148193	0,145223	0,204468	0,081570	0,217122	-0,000882	0,412800	0,151805	0,660485	0,161930	0,998237	4,2
m19	0,175470	0,157718	0,152461	0,225575	0,102222	0,235934	0,019253	0,428077	0,163561	0,631887	0,181521	1,039262	166,4
m20	0,174881	0,156728	0,152248	0,169317	0,071291	0,234834	-0,034132	0,427588	0,163173	0,710400	0,180026	0,933989	134,6
m21	0,210644	0,185136	0,174008	0,158900	0,079590	0,280439	-0,008217	0,467311	0,191452	0,725774	0,224114	0,983700	42,4
m22	0,210644	0,185136	0,174008	0,158900	0,079590	0,280439	-0,008217	0,467311	0,191452	0,725774	0,224114	0,983700	46
m23	0,172877	0,154747	0,152131	0,202761	0,087935	0,232020	-0,022617	0,426634	0,162172	0,662840	0,175850	0,955767	43,6
m24	0,162103	0,148547	0,146567	0,225894	0,094607	0,219872	0,008054	0,415745	0,154139	0,631463	0,164293	1,016239	28
m25	0,151140	0,145482	0,144539	0,261958	0,107291	0,209782	0,044099	0,409124	0,147803	0,584839	0,152126	1,092268	83,6

m26	0,160500	0,151305	0,147180	0,247402	0,105522	0,220575	0,029891	0,415860	0,153696	0,603332	0,164999	1,061624	48,4
m27	0,212742	0,190181	0,177993	0,163554	0,083197	0,285356	0,002956	0,472495	0,194593	0,718872	0,227310	1,005930	71,6
m28	0,216114	0,190629	0,175753	0,122649	0,060423	0,288174	-0,027726	0,471654	0,194891	0,781500	0,234406	0,946044	110,4
m29	0,203100	0,181859	0,170071	0,174571	0,085908	0,272621	0,006073	0,459697	0,185853	0,702749	0,217177	1,012220	77,7
m30	0,179431	0,160594	0,155195	0,175317	0,076289	0,240802	-0,003583	0,432886	0,166874	0,701669	0,185672	0,992860	117,9

Fuente: Elaboración Propia

3.3.6. Comparar las técnicas de estimación.

Se elaboró un dataset a partir de la información proporcionada de las relaciones entre bandas e indicadores, el cual se exporto en archivo .csv, el cual permite estimar la salinidad mediante las técnicas de estimación como son el SLR, MLR, RFR, DTR. Se usaron dos indicadores; coeficiente de determinación (R^2) y la raíz del error cuadrático medio (RMSE), a fin de medir la precisión de cada técnica de estimación.

X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
0,15859	0,144398	0,147015	0,132112	0,048282	0,21448	-0,045002	0,414928	0,152693	0,76661	0,155767	0,913872
0,15859	0,144398	0,147015	0,132112	0,048282	0,21448	-0,045002	0,414928	0,152693	0,76661	0,155767	0,913872
0,152696	0,139824	0,14237	0,146415	0,052384	0,307044	-0,049497	0,407046	0,147443	0,744568	0,149966	0,905674
0,133105	0,131549	0,134402	0,257496	0,09232	0,187142	0,073176	0,390024	0,133752	0,590462	0,13028	1,157907
0,093546	0,116438	0,117899	0,479845	0,172594	0,149361	0,298706	0,35588	0,105019	0,351493	0,092387	1,85187
0,09746	0,118041	0,119785	0,469966	0,17283	0,153075	0,223912	0,359561	0,108047	0,360376	0,09604	1,577029
0,11474	0,126363	0,127966	0,401427	0,153899	0,170684	0,181319	0,375675	0,121173	0,427117	0,113303	1,442953
0,148359	0,138905	0,140602	0,192998	0,070961	0,203236	0,027275	0,403252	0,144428	0,676448	0,146567	1,056079
0,162834	0,146614	0,14784	0,145173	0,055307	0,219113	-0,032619	0,417558	0,155156	0,746461	0,161483	0,936823
0,176507	0,158403	0,153734	0,160556	0,067519	0,237162	-0,022845	0,429987	0,164737	0,723312	0,181867	0,95533
0,183155	0,163635	0,158048	0,183886	0,082537	0,245606	0,001244	0,437714	0,170139	0,689352	0,18963	1,002491
0,179973	0,161419	0,156539	0,214932	0,098544	0,241757	0,014861	0,43466	0,167846	0,646182	0,185583	1,030171
0,17547	0,157718	0,152461	0,225575	0,102222	0,235934	0,019253	0,428077	0,163561	0,631887	0,181521	1,039262
0,165946	0,150551	0,149797	0,168949	0,067472	0,224061	-0,004625	0,421111	0,157664	0,710938	0,166781	0,990794
0,176507	0,158402	0,153734	0,160556	0,067519	0,237162	-0,022845	0,429987	0,164737	0,723312	0,181867	0,95533
0,191006	0,16934	0,160735	0,181285	0,084588	0,255264	-0,010245	0,444093	0,175218	0,693071	0,201231	0,979718
0,178346	0,158708	0,153592	0,190823	0,084116	0,238737	-0,00407	0,43058	0,165507	0,679511	0,184286	0,991892
0,158684	0,148193	0,145223	0,204468	0,08157	0,217122	-0,000882	0,4128	0,151805	0,660485	0,16193	0,998237
0,17547	0,157718	0,152461	0,225575	0,102222	0,235934	0,019253	0,428077	0,163561	0,631887	0,181521	1,039262
0,174881	0,156728	0,152248	0,169317	0,071291	0,234834	-0,034132	0,427588	0,163173	0,7104	0,180026	0,933989
0,210644	0,185136	0,174008	0,1589	0,07959	0,280439	-0,008217	0,467311	0,191452	0,725774	0,224114	0,9837
0,210644	0,185136	0,174008	0,1589	0,07959	0,280439	-0,008217	0,467311	0,191452	0,725774	0,224114	0,9837
0,172877	0,154747	0,152131	0,202761	0,087935	0,23202	-0,022617	0,426634	0,162172	0,66284	0,17585	0,955767
0,162103	0,148547	0,146567	0,225894	0,094607	0,219872	0,000054	0,415745	0,154139	0,631463	0,164293	1,016239
0,15114	0,145482	0,144539	0,261958	0,107291	0,209782	0,044099	0,409124	0,147803	0,384839	0,152126	1,092268
0,1605	0,151305	0,14718	0,247402	0,105322	0,220575	0,029891	0,41586	0,153696	0,603332	0,164999	1,061624
0,212742	0,190181	0,177993	0,163554	0,083197	0,285356	0,002956	0,472495	0,194593	0,718872	0,22731	1,00593
0,216114	0,190629	0,175753	0,122649	0,060423	0,288174	-0,027726	0,471654	0,194891	0,7815	0,234406	0,946044
0,2031	0,181859	0,170071	0,174571	0,085908	0,272621	0,006073	0,459697	0,185853	0,702749	0,217177	1,01222
0,179431	0,160594	0,155195	0,175317	0,076289	0,240802	-0,003583	0,432886	0,166874	0,701669	0,185672	0,99286

Figura 15. Variables independientes (indicadores)

Fuente: Elaboración propia

Y
1,374
1,388
2,482
0,713
1,831
0,601
0,746
0,994
0,947
0,311
14,47
38,9
84,5
0,817
7,21
45,3
104,4
4,2
166,4
134,6
42,4
46
43,6
28
83,6
48,4
71,6
110,4
77,7
117,8

Figura 16. Variables dependiente (CE)

Fuente: Elaboración propia

Las técnicas se implementaron en Python y usando la librería sklearn.

3.3.6.1. Simple Lineal Regression (SLR)

Código en Python

```

indicadores=['RED', 'GREEN', 'BLUE',
'NDVI','DVI','BI','ASTERSI','SI','SI1','SI2','SI3','SI11']
x=dataset[indicadores]
y=dataset['CE']

regr = linear_model.LinearRegression()

regr.fit(x,y)

```

Se obtuvo los resultados siguientes:

Tabla 13.

Resultados de la estimación usando SLR, relación entre Ce. Estimada y Real.

MUESTA	CE Estimada	CE Real
1	-9,15	1,574
2	-9,15	1,368
3	16,933	2,482
4	-8,052	0,713
5	-0,381	1,621
6	-3,852	0,601
7	19,472	0,746
8	8,688	0,994
9	9,944	0,947
10	35,936	0,311
11	44,929	14,470
12	66,148	38,900
13	98,547	84,500
14	7,923	0,917
15	35,936	7,210
16	85,753	45,300
17	72,065	104,400
18	41,579	4,200
19	98,547	166,400
20	51,458	134,600
21	57,322	42,400
22	57,322	46,000
23	58,814	43,600
24	66,089	28,000
25	24,418	83,600
26	70,334	48,400
27	49,383	71,600
28	101,857	110,400
29	84,905	77,700
30	48,135	117,900
Media	42,7284	42,72846667

Fuente: Elaboración propia

3.3.6.2. Multiple Lineal Regression (MLR)

Código en Python donde se ingresan los indicadores para el entrenamiento además de seleccionar los valores de conductividad eléctrica, se instancia el Objeto que contiene los métodos para el entrenamiento con el algoritmo de MLR.

```
indicadores=['RED', 'GREEN', 'BLUE',  
'NDVI', 'DVI', 'BI', 'ASTERSI', 'SI', 'SI1', 'SI2', 'SI3', 'SI11']  
x=dataset[indicadores]  
y=dataset['CE']  
  
regresor = LinearRegression()  
regresor.fit(x,y)  
  
LinearRegression(copy_X=True,fit_intercept=True,n_jobs=1000,nor  
malize=False)  
y_pred=regresor.predict(x)
```

Tabla 14.

Resultados de la estimación usando MLR, relación entre Ce. Estimada y Real.

Muestra	CE Estimada	CE Real
1	-9,15	1,574
2	-9,15	1,368
3	16,933	2,482
4	-8,052	0,713
5	-0,381	1,621
6	-3,852	0,601
7	19,472	0,746
8	8,688	0,994
9	9,944	0,947
10	35,936	0,311
11	44,929	14,470
12	66,148	38,900
13	98,547	84,500
14	7,923	0,917
15	35,936	7,210
16	85,753	45,300
17	72,065	104,400
18	41,579	4,200
19	98,547	166,400
20	51,458	134,600
21	57,322	42,400
22	57,322	46,000
23	58,814	43,600
24	66,089	28,000
25	24,418	83,600
26	70,334	48,400
27	49,383	71,600
28	101,857	110,400
29	84,905	77,700
30	48,135	117,900
Media	42,7284	42,7284667

Fuente: Elaboración propia

3.3.6.3. Random Forest Regression (RFR)

Código en Python donde se ingresan los indicadores para el entrenamiento además de seleccionar los valores de conductividad eléctrica, se instancia el Objeto que contiene los métodos para el entrenamiento con el algoritmo de RFR.

```

indicadores=['RED', 'GREEN', 'BLUE',
'NDVI', 'DVI', 'BI', 'ASTERSI', 'SI', 'SI1', 'SI2', 'SI3', 'SI11']
x=dataset[indicadores]
y=dataset['CE']
rfgm = RandomForestRegressor()
rfgm.fit(x,y)
ypred = rfgm.predict(x)

```

Tabla 15.

Resultados de la estimación usando RFR, relación entre Ce. Estimada y Real.

Muestra	CE Estimada	CE Real
1	2,860	1,574
2	2,860	1,368
3	3,314	2,482
4	7,178	0,713
5	3,799	1,621
6	3,374	0,601
7	3,367	0,746
8	2,290	0,994
9	10,103	0,947
10	20,675	0,311
11	31,178	14,470
12	59,461	38,900
13	121,242	84,500
14	11,132	0,917
15	20,675	7,210
16	44,026	45,300
17	86,390	104,400
18	9,401	4,200
19	121,242	166,400
20	102,478	134,600
21	47,408	42,400
22	47,408	46,000
23	65,916	43,600
24	32,534	28,000
25	57,512	83,600
26	48,656	48,400
27	65,059	71,600
28	85,563	110,400
29	64,335	77,700
30	85,314	117,900
Media	42,225	42,7284667

Fuente: Elaboración propia

3.3.6.4. Decision Tree Regression (DTR)

Código en Python donde se ingresan los indicadores para el entrenamiento además de seleccionar los valores de conductividad eléctrica, se instancia el Objeto que contiene los métodos para el entrenamiento con el algoritmo de DTR.

```
indicadores=['RED', 'GREEN', 'BLUE',
```

```

'NDVI', 'DVI', 'BI', 'ASTERSI', 'SI', 'SI1', 'SI2', 'SI3', 'SI11']
x=dataset[indicadores]
y=dataset['CE']

regressor = DecisionTreeRegressor(random_state=0)
regressor.fit(x,y)
y_pred=regressor.predict(x)
for a in y_pred :
    print(round(a,3))
r2=regressor.score(x,y) # R2
rmse=np.sqrt(r2)

```

Tabla 16.

Resultados de la estimación usando RFR, relación entre Ce. Estimada y Real.

Muestra	CE Estimada	CE Real
1	1,471	1,574
2	1,471	1,368
3	2,482	2,482
4	0,713	0,713
5	1,621	1,621
6	0,601	0,601
7	0,746	0,746
8	0,994	0,994
9	0,947	0,947
10	3,760	0,311
11	14,470	14,470
12	38,900	38,900
13	125,450	84,500
14	0,917	0,917
15	3,760	7,210
16	45,300	45,300
17	104,400	104,400
18	4,200	4,200
19	125,450	166,400
20	134,600	134,600
21	44,200	42,400
22	44,200	46,000
23	43,600	43,600
24	28,000	28,000
25	83,600	83,600
26	48,400	48,400
27	71,600	71,600
28	110,400	110,400
29	77,700	77,700
30	117,900	117,900
Media	42,72843333	42,7284667

Fuente: Elaboración propia

Posterior al desarrollo de las técnicas de estimación seleccionadas, se procedió a calcular la Operacionalización de las variables, con los resultados obtenidos de cada técnica.

1. El coeficiente de correlación (R^2)

$$R^2 = \left(\frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 + \sum_{i=1}^n (Y_i - \bar{Y})^2}} \right)^2$$

Donde:

X_i = valor medido

Y_i = valor estimado

\bar{X}_i = la media de valores medidos

\bar{Y}_i = la media de valores estimados

n = el número de muestras

2. RMSE

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (X_i - Y_i)^2}{n}}$$

Donde:

X_i = valor medido

Y_i = valor estimado

n = el número de muestras

DESARROLLO DE LA OPERACIONALIZACION DE VARIABLES PARA SLR

- El coeficiente de correlación (R^2)

$$\bar{X} = 42,73$$

$$\bar{Y}_i = 42,73$$

$$n = 30$$

Tabla 17.

Cálculo del indicador R^2 para SLR.

MUESTA	X	$x - \bar{X}$	Y	$y - \bar{Y}$
1	-9,15	-51,88	1,57	-41,15
2	-9,15	-51,88	1,37	-41,36
3	16,93	-25,80	2,48	-40,25
4	-8,05	-50,78	0,71	-42,02
5	-0,38	-43,11	1,62	-41,11
6	-3,85	-46,58	0,60	-42,13
7	19,47	-23,26	0,75	-41,98
8	8,69	-34,04	0,99	-41,73
9	9,94	-32,78	0,95	-41,78
10	35,94	-6,79	0,31	-42,42
11	44,93	2,20	14,47	-28,26
12	66,15	23,42	38,90	-3,83
13	98,55	55,82	84,50	41,77
14	7,92	-34,81	0,92	-41,81
15	35,94	-6,79	7,21	-35,52
16	85,75	43,02	45,30	2,57
17	72,07	29,34	104,40	61,67
18	41,58	-1,15	4,20	-38,53
19	98,55	55,82	166,40	123,67
20	51,46	8,73	134,60	91,87
21	57,32	14,59	42,40	-0,33
22	57,32	14,59	46,00	3,27
23	58,81	16,09	43,60	0,87
24	66,09	23,36	28,00	-14,73
25	24,42	-18,31	83,60	40,87
26	70,33	27,61	48,40	5,67
27	49,38	6,65	71,60	28,87
28	101,86	59,13	110,40	67,67
29	84,91	42,18	77,70	34,97
30	48,14	5,41	117,90	75,17

Fuente: Elaboración propia

$$R^2 = 0,51$$

- RMSE

$$n = 30$$

Tabla 18.

Cálculo del indicador RMSE para SLR

MUESTA	x	y	x-y
1	-9,15	1,57	-10,72
2	-9,15	1,37	-10,52
3	16,93	2,48	14,45
4	-8,05	0,71	-8,77
5	-0,38	1,62	-2,00
6	-3,85	0,60	-4,45
7	19,47	0,75	18,73
8	8,69	0,99	7,69
9	9,94	0,95	9,00
10	35,94	0,31	35,63
11	44,93	14,47	30,46
12	66,15	38,90	27,25
13	98,55	84,50	14,05
14	7,92	0,92	7,01
15	35,94	7,21	28,73
16	85,75	45,30	40,45
17	72,07	104,40	-32,34
18	41,58	4,20	37,38
19	98,55	166,40	-67,85
20	51,46	134,60	-83,14
21	57,32	42,40	14,92
22	57,32	46,00	11,32
23	58,81	43,60	15,21
24	66,09	28,00	38,09
25	24,42	83,60	-59,18
26	70,33	48,40	21,93
27	49,38	71,60	-22,22
28	101,86	110,40	-8,54
29	84,91	77,70	7,21
30	48,14	117,90	-69,77

Fuente: Elaboración propia

$$RMSE = 0,33$$

DESARROLLO DE LA OPERACIONALIZACION DE VARIABLES PARA MLR

- El coeficiente de correlación (R^2)

$$\bar{X} = 42,73$$

$$\bar{Y} = 42,73$$

$$n = 30$$

Tabla 19.

Cálculo del indicador R^2 para MLR.

MUESTA	X	X- \bar{X}	Y	Y- \bar{Y}
1	-9,15	-51,88	1,57	-41,15
2	-9,15	-51,88	1,37	-41,36
3	16,93	-25,80	2,48	-40,25
4	-8,05	-50,78	0,71	-42,02
5	-0,38	-43,11	1,62	-41,11
6	-3,85	-46,58	0,60	-42,13
7	19,47	-23,26	0,75	-41,98
8	8,69	-34,04	0,99	-41,73
9	9,94	-32,78	0,95	-41,78
10	35,94	-6,79	0,31	-42,42
11	44,93	2,20	14,47	-28,26
12	66,15	23,42	38,90	-3,83
13	98,55	55,82	84,50	41,77
14	7,92	-34,81	0,92	-41,81
15	35,94	-6,79	7,21	-35,52
16	85,75	43,02	45,30	2,57
17	72,07	29,34	104,40	61,67
18	41,58	-1,15	4,20	-38,53
19	98,55	55,82	166,40	123,67
20	51,46	8,73	134,60	91,87
21	57,32	14,59	42,40	-0,33
22	57,32	14,59	46,00	3,27
23	58,81	16,09	43,60	0,87
24	66,09	23,36	28,00	-14,73
25	24,42	-18,31	83,60	40,87
26	70,33	27,61	48,40	5,67
27	49,38	6,65	71,60	28,87
28	101,86	59,13	110,40	67,67
29	84,91	42,18	77,70	34,97
30	48,14	5,41	117,90	75,17

Fuente: Elaboración propia

$$R^2 = 0,51$$

- RMSE

$$n = 30$$

Tabla 20.

Cálculo del indicador RMSE para MLR.

MUESTA	x	y	x-y
1	-9,15	1,57	-10,72
2	-9,15	1,37	-10,52
3	16,93	2,48	14,45
4	-8,05	0,71	-8,77
5	-0,38	1,62	-2,00
6	-3,85	0,60	-4,45
7	19,47	0,75	18,73
8	8,69	0,99	7,69
9	9,94	0,95	9,00
10	35,94	0,31	35,63
11	44,93	14,47	30,46
12	66,15	38,90	27,25
13	98,55	84,50	14,05
14	7,92	0,92	7,01
15	35,94	7,21	28,73
16	85,75	45,30	40,45
17	72,07	104,40	-32,34
18	41,58	4,20	37,38
19	98,55	166,40	-67,85
20	51,46	134,60	-83,14
21	57,32	42,40	14,92
22	57,32	46,00	11,32
23	58,81	43,60	15,21
24	66,09	28,00	38,09
25	24,42	83,60	-59,18
26	70,33	48,40	21,93
27	49,38	71,60	-22,22
28	101,86	110,40	-8,54
29	84,91	77,70	7,21
30	48,14	117,90	-69,77

Fuente: Elaboración propia

$$RMSE = 0,33$$

DESARROLLO DE LA OPERACIONALIZACION DE VARIABLES PARA RFR

- El coeficiente de correlación (R^2)

$$\bar{X} = 42,23$$

$$\bar{Y} = 42,73$$

$$n = 30$$

Tabla 21.

Cálculo del indicador R^2 para RFR.

MUESTA	X	X- \bar{X}	Y	Y- \bar{Y}
1	2,86	-39,37	1,57	-41,15
2	2,86	-39,37	1,37	-41,36
3	3,31	-38,91	2,48	-40,25
4	7,18	-35,05	0,71	-42,02
5	3,80	-38,43	1,62	-41,11
6	3,37	-38,85	0,60	-42,13
7	3,37	-38,86	0,75	-41,98
8	2,29	-39,94	0,99	-41,73
9	10,10	-32,12	0,95	-41,78
10	20,68	-21,55	0,31	-42,42
11	31,18	-11,05	14,47	-28,26
12	59,46	17,24	38,90	-3,83
13	121,24	79,02	84,50	41,77
14	11,13	-31,09	0,92	-41,81
15	20,68	-21,55	7,21	-35,52
16	44,03	1,80	45,30	2,57
17	86,39	44,17	104,40	61,67
18	9,40	-32,82	4,20	-38,53
19	121,24	79,02	166,40	123,67
20	102,48	60,25	134,60	91,87
21	47,41	5,18	42,40	-0,33
22	47,41	5,18	46,00	3,27
23	65,92	23,69	43,60	0,87
24	32,53	-9,69	28,00	-14,73
25	57,51	15,29	83,60	40,87
26	48,66	6,43	48,40	5,67
27	65,06	22,83	71,60	28,87
28	85,56	43,34	110,40	67,67
29	64,34	22,11	77,70	34,97
30	85,31	43,09	117,90	75,17

Fuente: Elaboración propia

$$R^2 = 0,88$$

- RMSE

$$n = 30$$

Tabla 22.

Cálculo del indicador RMSE para RFR.

MUESTA	x	y	x-y
1	2,86	1,57	1,29
2	2,86	1,37	1,49
3	3,31	2,48	0,83
4	7,18	0,71	6,47
5	3,80	1,62	2,18
6	3,37	0,60	2,77
7	3,37	0,75	2,62
8	2,29	0,99	1,30
9	10,10	0,95	9,16
10	20,68	0,31	20,36
11	31,18	14,47	16,71
12	59,46	38,90	20,56
13	121,24	84,50	36,74
14	11,13	0,92	10,22
15	20,68	7,21	13,47
16	44,03	45,30	-1,27
17	86,39	104,40	-18,01
18	9,40	4,20	5,20
19	121,24	166,40	-45,16
20	102,48	134,60	-32,12
21	47,41	42,40	5,01
22	47,41	46,00	1,41
23	65,92	43,60	22,32
24	32,53	28,00	4,53
25	57,51	83,60	-26,09
26	48,66	48,40	0,26
27	65,06	71,60	-6,54
28	85,56	110,40	-24,84
29	64,34	77,70	-13,37
30	85,31	117,90	-32,59

Fuente: Elaboración propia

$$RMSE = 0,18$$

DESARROLLO DE LA OPERACIONALIZACION DE VARIABLES PARA DTR

- El coeficiente de correlación (R^2)

$$\bar{X} = 42,73$$

$$\bar{Y} = 42,73$$

$$n = 30$$

Tabla 23.

Cálculo del indicador R^2 para DTR.

MUESTA	X	X- \bar{X}	Y	Y- \bar{Y}
1	1,47	-41,26	1,57	-41,15
2	1,47	-41,26	1,37	-41,36
3	2,48	-40,25	2,48	-40,25
4	0,71	-42,02	0,71	-42,02
5	1,62	-41,11	1,62	-41,11
6	0,60	-42,13	0,60	-42,13
7	0,75	-41,98	0,75	-41,98
8	0,99	-41,73	0,99	-41,73
9	0,95	-41,78	0,95	-41,78
10	3,76	-38,97	0,31	-42,42
11	14,47	-28,26	14,47	-28,26
12	38,90	-3,83	38,90	-3,83
13	125,45	82,72	84,50	41,77
14	0,92	-41,81	0,92	-41,81
15	3,76	-38,97	7,21	-35,52
16	45,30	2,57	45,30	2,57
17	104,40	61,67	104,40	61,67
18	4,20	-38,53	4,20	-38,53
19	125,45	82,72	166,40	123,67
20	134,60	91,87	134,60	91,87
21	44,20	1,47	42,40	-0,33
22	44,20	1,47	46,00	3,27
23	43,60	0,87	43,60	0,87
24	28,00	-14,73	28,00	-14,73
25	83,60	40,87	83,60	40,87
26	48,40	5,67	48,40	5,67
27	71,60	28,87	71,60	28,87
28	110,40	67,67	110,40	67,67
29	77,70	34,97	77,70	34,97
30	117,90	75,17	117,90	75,17

Fuente: Elaboración propia

$$R^2 = 0,95$$

- RMSE

$n = 30$

Tabla 24.

Cálculo del indicador RMSE para DTR.

MUESTA	x	y	x-y
1	1,47	1,57	-0,10
2	1,47	1,37	0,10
3	2,48	2,48	0,00
4	0,71	0,71	0,00
5	1,62	1,62	0,00
6	0,60	0,60	0,00
7	0,75	0,75	0,00
8	0,99	0,99	0,00
9	0,95	0,95	0,00
10	3,76	0,31	3,45
11	14,47	14,47	0,00
12	38,90	38,90	0,00
13	125,45	84,50	40,95
14	0,92	0,92	0,00
15	3,76	7,21	-3,45
16	45,30	45,30	0,00
17	104,40	104,40	0,00
18	4,20	4,20	0,00
19	125,45	166,40	-40,95
20	134,60	134,60	0,00
21	44,20	42,40	1,80
22	44,20	46,00	-1,80
23	43,60	43,60	0,00
24	28,00	28,00	0,00
25	83,60	83,60	0,00
26	48,40	48,40	0,00
27	71,60	71,60	0,00
28	110,40	110,40	0,00
29	77,70	77,70	0,00
30	117,90	117,90	0,00

Fuente: Elaboración propia

RMSE =0,11

IV. CONCLUSIONES Y RECOMENDACIONES

4.1. Conclusiones.

1. Se consideró utilizar cuatro técnicas de estimación, las de mejor resultado en trabajos previos relacionados a la estimación de la salinidad.
2. La adquisición de imágenes multiespectrales de satélite, se realizó de manera sencilla y gratuita, tomándose solo aquellas que representan el área de interés y que presentaron bajo nivel de nubosidad.
3. La calibración atmosférica de las imágenes multiespectrales de satélite, nos arrojaron valores que representan los indicadores de salinidad y vegetación que representan la reflectancia en cada una de las bandas espectrales
4. Los indicadores de salinidad y vegetación, se tomaron de investigaciones previas relacionadas, se realizó la selección teniendo en cuenta la conductividad eléctrica y su correlación con ellos.
5. Se recolecto muestras de suelo de la zona de interés, las cuales fueron sometidas a prueba de conductividad eléctrica, para poder estimar la salinidad.
6. La técnica DTR obtuvo el mejor rendimiento al estimar la salinidad con una precisión de 95% y un error promedio de 0.11, como criterios para comparar las técnicas se utilizó el R^2 y el **RMSE**.

4.2.Recomendaciones.

1. La técnica de estimación varía su precisión de acuerdo al modelo de datos y tamaño de las muestras, por esta razón es recomendable evaluar el modelo y tamaño de muestras con respecto a los resultados obtenidos por las técnicas de interés en trabajos previos.
2. El nivel de nubosidad es importante al momento de elegir las imágenes multiespectrales por tal motivo, se recomienda siempre tomar aquellas que tengan el nivel de nubosidad por debajo del 30 % y que además representa la zona de interés.
3. Se recomienda tomar las muestras de suelo con la mayor precisión posible ayudados por un GPS para poder identificar las coordenadas con la mayor precisión posible, puesto que es con ellas que se va trabajar para extraer los valores en las imágenes multiespectrales, adquiridas del satélite.

REFERENCIAS.

- Abdellatif, D., & Mourad, L. (2012). Remote Sensing Techniques for Salt Affected Soil Mapping: *Application to the Oran Region of Algeria*, 188-198.
- Akramkhanov, A., Martius, C., Park, S. J., & Hendrickx, J. (2011). Environmental factors of spatial distribution of soil salinity on flat irrigated terrain. *Geoderma*, 163(1–2):55–62.
- Allbed, Amal, Lalit Kumar, and Priyakant Sinha. (2014). investigación “Mapping and Modelling Spatial Variation in Soil Salinity in the Al Hassa Oasis Based on Remote Sensing Indicators and Regression Techniques”.
- Al-Shafie, W., Nangia, V., Mhaimed, A., Ziadat, F., & Payne, W. B. (2014). Soil Salinity Mapping by Multiscale Remote Sensing in Mesopotamia, Iraq. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote*, 4442-4451.
- Amal, A., Lalit, K., & Priyakant, S. (2014). Mapping and Modelling Spatial Variation in Soil Salinity in the Al Hassa Oasis Based on Remote Sensing Indicators and Regression Techniques. *Remote Sens*, 1137-1157.
- Bado, S., Ghanim, A. M., Berthold, G., Forster, B., Jankowicz-Cieslak, J., & Luxiang, L. (2016). *Protocols for Pre-Field Screening of Mutants for Salt Tolerance in Rice, Wheat and Barley*. Nueva York: Springer Open.
- Breiman, L. (2001). Random Forests. *Machine Learning*, 5–32.
- Butcher, G. (2010). *Tour of the Electromagnetic Spectrum*. Washington, DC.: National Aeronautics and Space.
- Carrasquilla-Batista, A. (2016). Regresión lineal simple y múltiple: aplicación en la predicción de variables naturales relacionadas con el crecimiento microalgal. *Scielo*.
- Chugh, K. (2019). Decision Tree Regression in Machine learning.
- Dervić, K., Sinik, V., & Despotovic, Z. V. (2019). BASICS OF ELECTROMAGNETIC RADIATION. *IX International Conference Industrial Engineering and Environmental*, 512-520.
- DRAKOS, G. (23 de 5 de 2019). *Decision Tree Regressor explained in depth*. Obtenido de GDCoder: <https://gdcoder.com/decision-tree-regressor-explained-in-depth/>

- FAO. (2007). *Advances in the assessment and monitoring of salinization and status of biosaline agriculture*.
- Fernández García, A., Sutter, F., Fernández Reche, J., & Lupfert, E. (2017). The Performance of Concentrated Solar Power (CSP) Systems. *Woodhead*, 67-98.
- Gamarra Agama, S. (2017). Analisis de la cobertura y uso de la tierra utilizando imágenes de resolución espacial media para el distrito de San Ramón-Chanchamayo-Junín-Perú. *Universidad Nacional Agraria de La Molina*.
- Gonzaga Aguilar, C. (2014). "Aplicación de Índices de Vegetación Derivados de Imágenes Satelitales Landsat 7 ETM + y ASTER Para La Caracterización de La.
- Guan, X., Shaoli, W., Zhanyi, G., & Ye, L. (2013). Dynamic prediction of soil salinization in an irrigation district based on the support vector machine. *Mathematical and Computer Modelling*, 719-724.
- Hernandez, M. M., Muñoz Iniestra, D. K., López, G. J., Soler, A. A., & López, G. F. (2009). IMPACTO DE LA PÉRDIDA DE LA VEGETACIÓN SOBRE LAS PROPIEDADES DE UN SUELO ALUVIAL. *Terra Latinoamericana*, 237-246.
- Hsiao, E., Lin, C., & Pham, N. (2018). APPLYING AN OBJECT-BASED SVM CLASSIFIER TO EXPLORE CANOPY CLOSURE OF MANGROVE FOREST IN THE MEKONG DELTA USING SENTINEL-2 MULTISPECTRAL IMAGE. *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, 5402-5404.
- Jimenez, H., Rodriguez, R., & Tiparra, J. (1978). *Diagnóstico de TEA*. Madrid: Latinoamérica SA.
- Larsen, S. (2005).
- Lau Quan, A., Garea Llano, E., & Ruiz, M. E. (2005). Estimación de la salinidad de los suelos utilizando una imagen espectrozonal y el sistema de información geográfica Telemap. *Revista Ciencias Técnicas Agropecuarias*, 47-54.
- Li, L. (2019). Classification and Regression Analysis with Decision Trees. *Towards Data Science*.
- Limeres, M. C. (2011-2012). Regresión lineal simple. *Estadística. FBA I*.

- López, F. (1944). Degradación del suelo ¿fatalidad climática o mala gestión humana?: hacia una gestión sostenible del recurso en el contexto mediterráneo. *University of Murcia*, 16.
- Mashimbye, Z., Nell, J. P., Cho, M. P., DE CLERCQ, W. P., VAN NIEKERK, A., & TURNER, D. P. (2012). *Model-Based Integrated Methods for Quantitative Estimation of Soil Salinity from Hyperspectral Remote Sensing Data: A Case Study of Selected South African Soils*. China: Pedosphere.
- Mejia Cabrera, H. I., Vilchez, D., Tuesta Monteza, V., & Forero, M. (2020). Soil salinity estimation of sparse vegetation based on multispectral image processing and machine learning. *Proc. SPIE 11510, Applications of Digital Image Processing XLIII*.
- Mejia, I., Ramirez, R., Jimenez, H., & Rosas, J. (2019). A new method a architecture entreprise. *Conference IEEE bussines*, 200-215.
- Mejia, I., Tuesta, M., & Forero, M. (2020). A new method of enterprise archicture small organizations. *Computer Science Technology*, 150-170.
- Mohamed, E., & Bahrawi, J. A. (2017). Soil salinity mapping and hydrological drought indices assessment in arid environments based on remote sensing techniques. *Copernicus Publications on behalf of the European Geosciences Union*.
- Mousavi et al. (2017). "Digital Mapping of Topsoil Salinity Using Remote".
- Palma, J., & Marín, R. (2008). *Inteligencia Artificial*. Madrid: McGrawHill. doi:978-84-481-5618-3
- Pérez Gutiérrez, C., & Muñoz Nieto, Á. L. (2006). *Teledetección o Detección Remota*. Salamanca: Universidad de Salamanca.
- Perez Gutierrez, C., & Muñoz Nieto, A. L. (2006). *Teledeteccion:Nociones y Aplicaciones*. España: Universidad de Salamanca.
- Rodrigo, J. A. (2016). Introducción a la Regresión Lineal Múltiple.
- Rojas, K. (2018). Identificación de efectos negativos de la TEA en el aprendizaje. *IEEE conference Technology children especial*, 200-215.
- Rojo, J. M. (2007). *Regresión lineal múltiple*. Madrid: Instituto de Economía y Geografía.
- Sarkar, D. R. (2018). Practical Machine Learning.
- Shalabh, I. K. (s.f.). Simple Linear Regression Analysis. *Regression Analysis*.

- Siqueira, G. M. (2009). Medida de la conductividad eléctrica aparente del suelo por inducción electromagnética y variabilidad espacial de propiedades físicas y químicas del suelo.
- Soca Flores, R. (2015). Identificación de tierras degradadas por salinidad del suelo en los cultivos de caña de azúcar en Pomalca usando imágenes de satélite. *UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS*, 139.
- SZNAJDLEDER, P. (2012). *Java a fondo - estudio del lenguaje y desarrollo de aplicaciones - 2a ed.* México: Alfaomega.
- U.S, D. o. (2018). "LANDSAT 8 DATA USERS HANDBOOK." 8(October):2–4.
- Upadhyay, M., Upadhyay, M., & Maurya, M. (2017). Regression and Artificial Neural Network based Improved Classification of LISS-III Satellite Image. *International Conference on Current Trends in Computer, Electrical, Electronics and Communication (ICCTCEEC-2017)*, 917-920.
- Yu, Z., Wenbin, Z., & Shuwen, W. (2017). Diagnosis Method for Corn Nitrogen Status Measurement based on Multispectral Imaging. *IEEE Instrumentation and Measurement Technology Conference*, 5.

ANEXOS.

Anexo 1: Resolución de aprobación de tesis



UNIVERSIDAD
SEÑOR DE SIPÁN

FACULTAD DE INGENIERÍA, ARQUITECTURA Y URBANISMO

RESOLUCIÓN N° 2363-2020/FIAU-USS

Pimentel, 16 de diciembre de 2020

VISTO:

El Acta de reunión N°1012 - 2020/FIAU-ISS del Comité de investigación de la Escuela profesional de INGENIERÍA DE SISTEMAS remitida mediante mensaje de correo electrónico del 15 de diciembre de 2020, y;

CONSIDERANDO:

Que, de conformidad con la Ley Universitaria N° 30220 en su artículo 48° que a letra dice: "La investigación constituye una función esencial y obligatoria de la universidad, que la fomenta y realiza, respondiendo a través de la producción de conocimiento y desarrollo de tecnologías a las necesidades de la sociedad, con especial énfasis en la realidad nacional. Los docentes, estudiantes y graduados participan en la actividad investigadora en su propia institución o en redes de investigación nacional o internacional, creadas por las instituciones universitarias públicas o privadas.";

Que, de conformidad con el Reglamento de grados y títulos en su artículo 21° señala: "Los temas de trabajo de investigación, trabajo académico y tesis son aprobados por el Comité de Investigación y derivados a la Facultad o Escuela de Posgrado, según corresponda, para la emisión de la resolución respectiva. El periodo de vigencia de los mismos será de dos años, a partir de su aprobación. En caso un tema perdiera vigencia, el Comité de Investigación evaluará la ampliación de la misma.

Que, de conformidad con el Reglamento de grados y títulos en su artículo 24° señala: La tesis es un estudio que debe denotar rigurosidad metodológica, originalidad, relevancia social, utilidad teórica y/o práctica en el ámbito de la escuela profesional. Para el grado de doctor se requiere una tesis de máxima rigurosidad académica y de carácter original. Es individual para la obtención de un grado; es individual o en pares para obtener un título profesional. Asimismo, en su artículo 25° señala: "El tema debe responder a alguna de las líneas de investigación institucionales de la USS S.A.C."

Que, según documentos de Visto el Comité de investigación de la Escuela profesional de INGENIERÍA DE SISTEMAS acuerda aprobar los temas de las Tesis a cargo de los egresados que se detallan en el anexo de la presente Resolución.



FACULTAD DE INGENIERÍA, ARQUITECTURA Y URBANISMO
RESOLUCIÓN N° 2363-2020/FIAU-USS

Pimentel, 16 de diciembre de 2020

Estando a lo expuesto, y en uso de las atribuciones conferidas y de conformidad con las normas y reglamentos vigentes;

SE RESUELVE:

ARTÍCULO 1°: APROBAR, el tema de la Tesis perteneciente a la línea de investigación de INFRAESTRUCTURA, TECNOLOGÍA Y MEDIO AMBIENTE, a cargo de los egresados del Programa de estudios de INGENIERÍA DE SISTEMAS según se detalla en el anexo de la presente Resolución.

ARTÍCULO 2°: ESTABLECER, que la inscripción del Tema de la Tesis se realice a partir de emitida la presente resolución y tendrá una vigencia de dos (02) años.

ARTÍCULO 3°: DEJAR SIN EFECTO, toda Resolución emitida por la Facultad que se oponga a la presente Resolución.

REGÍSTRESE, COMUNÍQUESE Y ARCHÍVESE


 Dr. Mario Fernández Ramos Mesquí
Decano - Facultad de Ingeniería,
Arquitectura y Urbanismo
UNIVERSIDAD SEÑOR DE SIPÁN SAC


 FIAU, María Rosalva Siles Rivera
Secretaría Académica / Facultad de Ingeniería,
Arquitectura y Urbanismo
UNIVERSIDAD SEÑOR DE SIPÁN SAC

Cc: Interesado, Archivo

FACULTAD DE INGENIERÍA, ARQUITECTURA Y URBANISMO

RESOLUCIÓN N° 2363-2020/FIAU-USS

Pimentel, 16 de diciembre de 2020

ANEXO

Orden	Apellidos y Nombres	TITULO DEFINITIVO PROYECTO DE INVESTIGACIÓN
1	ALMENDRAS FLORES RONALD ELADIO	COMPARACIÓN DEL RENDIMIENTO DE PROTOCOLOS DE COMUNICACIÓN DE DISPOSITIVOS IOT PARA ASEGURAR LOS DATOS EN EL ÁMBITO DE LA SALUD
2	BOCANEGRA PINCHI YAN CARLOS	DESARROLLO DE UNA METODOLOGÍA PARA LA ADQUISICIÓN DE TECNOLOGÍAS DE LA INFORMACIÓN ORIENTADA A PEQUEÑAS ORGANIZACIONES BASADA EN ESTÁNDARES
3	BRENIS LLAGENTO JULIO ANTONIO	COMPARACIÓN DE ALGORITMOS DE AUTENTICACIÓN EN CONEXIONES DE REDES PRIVADAS VIRTUALES PARA USO EN TRABAJO REMOTO. CASO DE ESTUDIO: MUNICIPALIDAD PROVINCIAL DE FERREÑAFE
4	BUSTAMANTE GUERRERO FERNANDO LUCERO CHUCAS LUIS ALBERTO	EVALUACIÓN DE LA CALIDAD DE SOFTWARE DE LAS APLICACIONES EMBEBIDAS EN DISPOSITIVOS PUNTOS DE VENTA BASADOS EN ESTÁNDARES
5	CABANILLAS TORRES ALVARO PAUL HARBERT FARRO VARGAS LUIS ADEMAR	DESARROLLO DE UN SISTEMA INTELIGENTE PARA LA EVALUACIÓN DE LOS PERFILES POR COMPETENCIA LABORAL DE UN PUESTO GERENCIAL
6	CABRERA GUEVARA EDSON OMAR ROQUE IZQUIERDO ADELMO	DESARROLLO DE UNA METODOLOGÍA ÁGIL AD HOC PARA LA CREACIÓN DE APLICACIONES MÓVILES EN PEQUEÑAS EMPRESAS. CASO DE ESTUDIO: AMÉRICA PERÚ S.A.C, LIMA.
7	CAMPOS BARRERA SANDRO PAUL PASTOR OLIVA CESAR AUGUSTO	DETECCIÓN DE PATRONES DE COMPORTAMIENTO EN LA DESERCIÓN DE ESTUDIANTES DE LA ESCUELA PROFESIONAL DE INGENIERÍA DE INDUSTRIAS ALIMENTARIAS DE LA UNIVERSIDAD NACIONAL DE JAÉN UTILIZANDO MINERÍA DE DATOS
8	CARRASCO MARTINEZ LUIS MARLON	EVALUACIÓN DE MÉTODOS DE DETECCIÓN DE INTRUSOS PARA BLOQUEAR Y NOTIFICAR ATAQUES EN LA ZONA DESMILITARIZADA DE LA RED
9	CARREÑO CASTILLO JORGE LUIS SALAZAR AGUILAR LUIS	EVALUACIÓN DE ALGORITMOS PARA MEDIR EFICIENCIA EN EL TRÁFICO OCULTO DE VOZ IP
10	CARRERA PEÑA LUIS ARMANDO	COMPARACIÓN DE TÉCNICAS DE CLASIFICACIÓN EN EL RECONOCIMIENTO DE CARACTERES DE DÍGITOS DE IMAGEN DIGITAL DE MEDIDORES ELÉCTRICOS
11	CARRERA SANCHEZ JOSE ROSVEL	DESARROLLO DE LA ARQUITECTURA EMPRESARIAL BASADO EN MARCOS DE TRABAJO PARA EL ALINEAMIENTO ESTRATÉGICO DE TECNOLOGÍAS DE INFORMACIÓN. CASO DE ESTUDIO: COLEGIO RAMÓN CASTILLA – BAGUA GRANDE
12	CARRILLO CAJO CYNTHIA EFIGENIA PISCOYA INOÑAN SANDRA PAOLA	DESARROLLO DE UNA PLATAFORMA MÓVIL COMO APOYO A LA ATENCIÓN DE PACIENTES EN HOSPITALES Y CLÍNICAS UTILIZANDO CLOUD COMPUTING Y GEOREFERENCIACIÓN
13	CELIS SANCHEZ SUJEILY PEREZ ROJAS FRANKLIN ALEXI	MODELO DE EVALUACIÓN DE PROCESOS TI CON BASE EN EL MARCO DE REFERENCIA COBIT 5 PAM. CASO DE ESTUDIO SIEMPRESOFT E.I.R.L. 2020 - CHICLAYO
14	CONDOR ROJAS JHAIR	DESARROLLO DE UNA METODOLOGÍA DE GESTIÓN DE RIESGOS DE CIBERSEGURIDAD PARA UNA MUNICIPALIDAD PERUANA. CASO DE ESTUDIO: MUNICIPALIDAD DE CHICLAYO
15	LURUNEL LAJAN ERIK ARTURO	RECONOCIMIENTO DE EXPRESIONES FACIALES DE TRISTEZA UTILIZANDO APRENDIZAJE PROFUNDO
16	CORTEZ BURGOS JOHANDER ENRIQUE MEDRANO MORI JOSE LUIS	DESARROLLO DE UN MODELO DE PROCESOS PARA UNA FÁBRICA PERUANA DE SOFTWARE BASADA EN METODOLOGÍAS ÁGILES CASO DE ESTUDIO CONASTEC S.R.L.
17	GASTELO FERNANDEZ EDIN JONNY RODRÍGUEZ FLORES ALFREDO HIPÓLITO	DESARROLLO DE UN MODELO DE GESTIÓN DE RIESGOS BASADO EN LA METODOLOGÍA MAGERIT PARA MINIMIZAR LOS RIESGOS DE ADQUISICIÓN Y USO DE TI EN UNA MUNICIPALIDAD DEL PERÚ. CASO DE ESTUDIO: MUNICIPALIDAD DISTRITAL DE CUPISNIQUE - CAJAMARCA
18	GUERRERO SERNAQUE LUIGGI EMILIO LOZA VARGAS FERNANDO AUGUSTO	DESARROLLO DE UNA APLICACIÓN MIDDLEWARE PARA EL ENVÍO Y VALIDACIÓN DE COMPROBANTES ELECTRÓNICOS A SUNAT
19	HUAMAN CASAS JUNIOR ALDAIR SERRATO VILCHERRES FERNANDO JOSE	DESARROLLO DE UN MÉTODO PARA DETECCIÓN DE FRAUDES DE PAGOS EN LÍNEA UTILIZANDO APRENDIZAJE AUTOMÁTICO
20	IDROGO CORNEJO LEONCIO FERNANDEZ SALDAÑA CARLOS	DESARROLLO DE UN MÉTODO DE CLASIFICACIÓN AUTOMÁTICA DE CITRUS AURANTIFOLIA USANDO PROCESAMIENTO DIGITAL DE IMÁGENES
21	LOPEZ BARRETO ALDO HASSAN	COMPARACIÓN DE PROTOCOLOS DE REDES PRIVADAS VIRTUALES PARA AUMENTAR LA SEGURIDAD EN EL CONTEXTO DE TELETRABAJO

FACULTAD DE INGENIERÍA, ARQUITECTURA Y URBANISMO

RESOLUCIÓN N° 2363-2020/FIAU-USS

Pimentel, 16 de diciembre de 2020

22	MENDOZA RENGIFO GENARO	DESARROLLO DE UNA METODOLOGÍA ÁGIL AD HOC PARA LA CREACIÓN DE APLICACIONES WEB EN PEQUEÑAS EMPRESAS. CASO DE ESTUDIO: SOLTI S.A.C.
23	MONTALVO SANDOVAL JOSE LUIS RUBIO OTERO DANIEL	DESARROLLO DE UN MÉTODO DE IDENTIFICACIÓN DE DEFECTOS EXTERNOS DEL MANGIFERA INDICA L USANDO PROCESAMIENTO DE IMÁGENES DIGITALES Y APRENDIZAJE DE MÁQUINA.
24	ÑAÑEZ PEREDO ALEXIS JOEL	COMPARACIÓN DE ALGORITMOS DE DETECCIÓN DE BORDES PARA UNA EFECTIVA VECTORIZACIÓN DE IMÁGENES DIGITALES DE MOLDES TEXTILES.
25	ORTEGA PUENTE OSCAR FERNANDO	DESARROLLO DE UNA APLICACIÓN MÓVIL DE REPORTE CIUDADANO UTILIZANDO IMÁGENES Y GEOLOCALIZACIÓN
26	PARDO VILLEGAS JOSE DIEGO SALINAS VELEZ BRYAN JOSE RICARDO	COMPARACIÓN DE TÉCNICAS DE ESTIMACIÓN DEL GRADO DE SALINIDAD EN SUELOS DE ESCASA VEGETACIÓN MEDIANTE EL PROCESAMIENTO DE IMÁGENES MULTIESPECTRALES DE SATÉLITE
27	RAVELO RUIZ ALLEN MARCEL	EVALUACIÓN DE ALGORITMOS DE PREDICCIÓN APLICADOS AL MANTENIMIENTO PREDICTIVO DE MÁQUINAS ELÉCTRICAS ROTATIVAS
28	REQUEJO NAVARRO JERSONS EXFRANSHER	EVALUACIÓN DE ALGORITMOS CRIPTOGRÁFICOS PARA MEJORAR SEGURIDAD EN UNA RED PRIVADA VIRTUAL
29	SALAZAR DÁVILA GIANFRANCO STEVEN	COMPARACIÓN DE TÉCNICAS DE VALIDACIÓN DE REQUERIMIENTOS DE SOFTWARE EN PEQUEÑAS EMPRESAS
30	SANTISTEBAN VALDERA ELVIS ALBERTO VILLANUEVA CASTRO JOSE SEBASTIAN	DESARROLLO DE UN MODELO DE GESTIÓN DE INCIDENCIAS BASADO EN MARCOS DE REFERENCIA DE BUENAS PRÁCTICAS PARA MEJORAR EL SERVICIO DE TI. CASO DE ESTUDIO MUNICIPALIDAD DISTRITAL DE MÓRROPE – LAMBAYEQUE
31	THEOLOGITIS SANCHEZ DIMITRIS IOANNIS	CONTROL EN EL ABASTECIMIENTO DEL INVENTARIO BASADO EN REDES NEURONALES PARA PEQUEÑAS Y MEDIANAS EMPRESAS

Anexo 2: Población de Técnicas de estimación y la muestra escogida

Tabla 25.

Población de técnicas de estimación.

TÉCNICAS DE ESTIMACIÓN	PRECISIÓN	
	R^2	Error relativo
A. PARTIAL LEAST SQUARE REGRESSION (PLSR)	89.00%	
	87.00%	
B. MULTIPLE LINEAR REGRESSION (MLR)	52.00%	
	36.00%	
	65.00%	
C. RANDOM FOREST REGRESSION (RFR)	79.00%	
	98.30%	
	85.00%	
D. SUPPORT VECTOR MACHINE (SVM)	83.00%	
	61.70%	37.00%
E. BOOSTED BINARY REGRESSION (BBRT)	82.00%	
	81.00%	
	93.00%	
F. ARTIFICIAL NEURONAL NETWORK (ANN)	69.00%	
	99.70%	
	41.90%	49.00%
G. DECISION TREE REGRESSION (DTR)	80.00%	
H. PRINCIPAL COMPONENT REGRESSION (PCR)	79.00%	
I. SIMPLE LINEAR REGRESSION (SLR)	87.00%	

Fuente: Elaboración propia

Recopilado de:

A. "A Comparison of Regression Techniques for Estimation of Above-Ground Winter Wheat Biomass Using Near-Surface Spectroscopy"(Yue et al., 2018)
B. "Mapping and Modelling Spatial Variation in Soil Salinity in the Al Hassa Oasis Based on Remote Sensing Indicators and Regression Techniques"(Allbed et al., 2014)
C. "Dynamic prediction of soil salinization in an irrigation district based on the support vector machine"(Guan et al., 2013)
D. "Digital Mapping of Topsoil Salinity Using Remote Sensing Indices in Agh-Ghala Plain, Iran"(Mousavi et al., 2017)
E. "Spatial modeling of soil salinity using multiple linear regression, ordinary kriging and artificial neural network methods"(Noureddine et al., 2019)
F. "Estimation of biomass in wheat using random forest regression algorithm and remote sensing data"(Wang et al., 2016)
G. "The performance comparison of Multiple Linear Regression, Random Forest and Artificial Neural Network by using photovoltaic and atmospheric data"
H. "A Comparison of Regression Techniques for Estimation of Above-Ground Winter Wheat Biomass Using Near-Surface Spectroscopy"(Yue et al., 2018)
I. "Identificación de las tierras degradadas por la salinidad del suelo en los cultivos de caña de azúcar mediante imágenes de satélite"(Soca et al., 2017)

Anexo 3: Pasos para recolectar las imágenes multiespectrales

1. Para descargar las imágenes se accede al siguiente sitio web:

<https://earthexplorer.usgs.gov/>

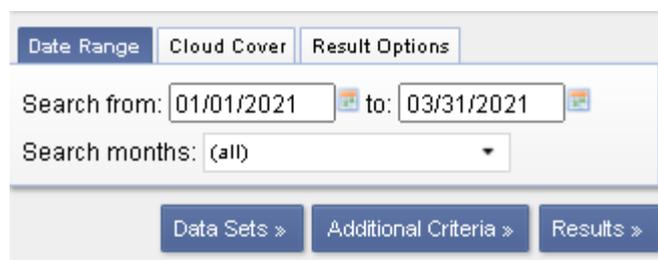
2. Se ingresarán las coordenadas de la zona de interes

Polygon	Circle	Predefined Area
Degree/Minute/Second		Decimal
1. Lat: 06° 38' 15" S, Lon: 079° 56' 09" W		
2. Lat: 06° 38' 16" S, Lon: 079° 56' 00" W		
3. Lat: 06° 38' 30" S, Lon: 079° 56' 03" W		
4. Lat: 06° 38' 30" S, Lon: 079° 56' 16" W		
Use Map	Add Coordinate	Clear Coordinates

Figura 17. Ingreso de coordenadas para la obtener las imágenes multiespectrales.

Fuente: Elaboración propia

3. Luego se buscó las imágenes desde **el 01 de enero del 2021 hasta el 03 de marzo del 2021**.

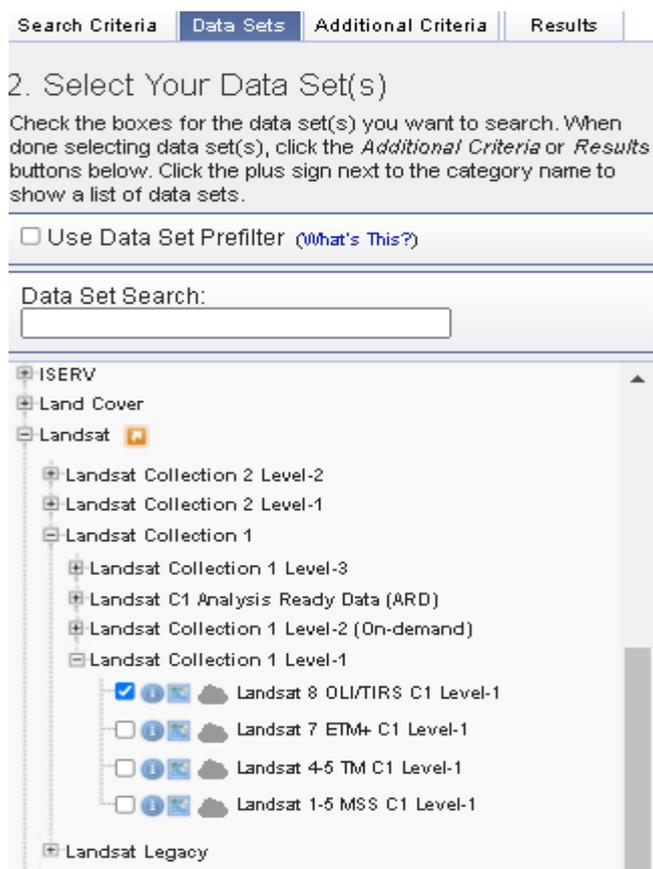


The screenshot shows a search interface with three tabs: "Date Range", "Cloud Cover", and "Result Options". The "Date Range" tab is active. It contains a "Search from:" field with the date "01/01/2021" and a "to:" field with the date "03/31/2021". Below these is a "Search months:" dropdown menu set to "(all)". At the bottom, there are three buttons: "Data Sets >", "Additional Criteria >", and "Results >".

Figura 18. Ingreso de coordenadas para la obtener las imágenes multiespectrales.

Fuente: Elaboración propia

4. En la segunda opción del Menú (DataSets) se procedió a seleccionar el satélite **LANSAT 8 OLI** con el cual se realizar la recolección de imágenes.



The screenshot shows the "Data Sets" tab in the search interface. It has a heading "2. Select Your Data Set(s)" and instructions: "Check the boxes for the data set(s) you want to search. When done selecting data set(s), click the Additional Criteria or Results buttons below. Click the plus sign next to the category name to show a list of data sets." There is a checkbox for "Use Data Set Prefilter (What's This?)". Below is a "Data Set Search:" text input field. A tree view shows the following categories and sub-items:

- ISERV
- Land Cover
- Landsat
 - Landsat Collection 2 Level-2
 - Landsat Collection 2 Level-1
 - Landsat Collection 1
 - Landsat Collection 1 Level-3
 - Landsat C1 Analysis Ready Data (ARD)
 - Landsat Collection 1 Level-2 (On-demand)
 - Landsat Collection 1 Level-1
 - Landsat 8 OLI/TIRS C1 Level-1
 - Landsat 7 ETM+ C1 Level-1
 - Landsat 4-5 TM C1 Level-1
 - Landsat 1-5 MSS C1 Level-1
 - Landsat Legacy

Figura 19. Selección del satélite LANSAT 8

Fuente: Elaboración propia

5. En el tercer Menú (Adicional Criteria) se eligio las propiedades de la búsqueda, en este caso se seleccionó el Scene Cloud Cover a un 30% para obtener las imágenes libres de nubosidad.

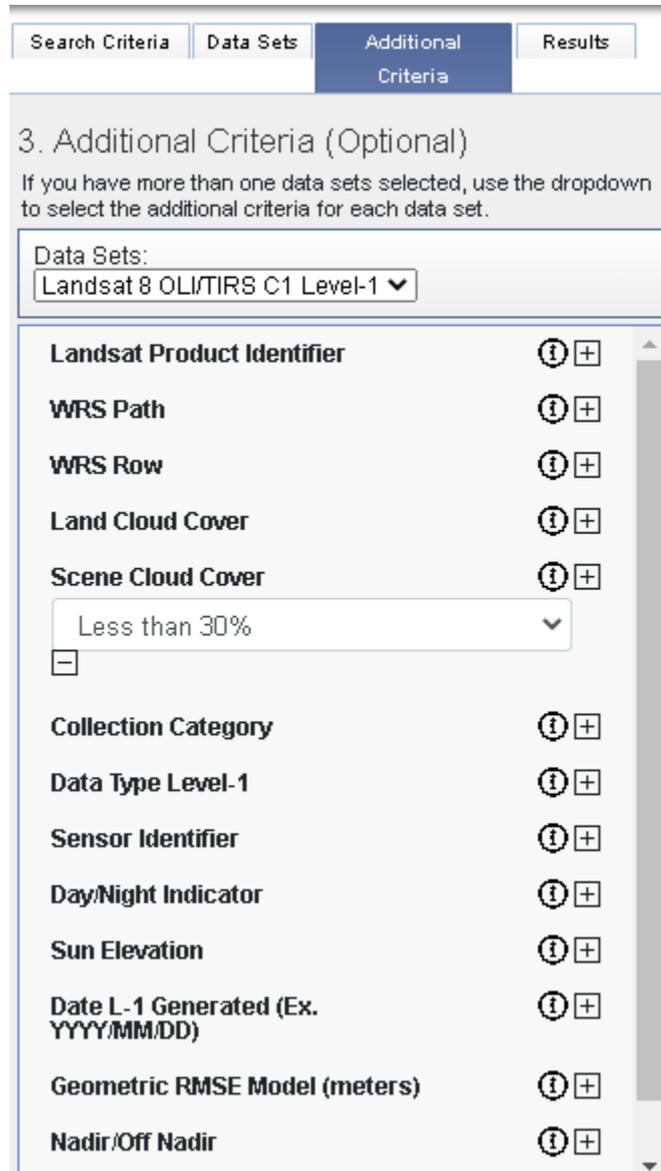


Figura 20. Selección Scene Cloud Cover.

Fuente: Elaboración propia

6. En el Menú (Results) se mostró los resultados de la búsqueda, de donde se seleccionó la imagen más óptima para continuar con el desarrollo de la investigación.



Figura 21. Descarga de paquetes de Imagen del LANSAT 8 OLI.

Fuente: Elaboración propia

7. Por último, se procedió a descargar la Colección Completa que contiene los datos de GeoTIFF

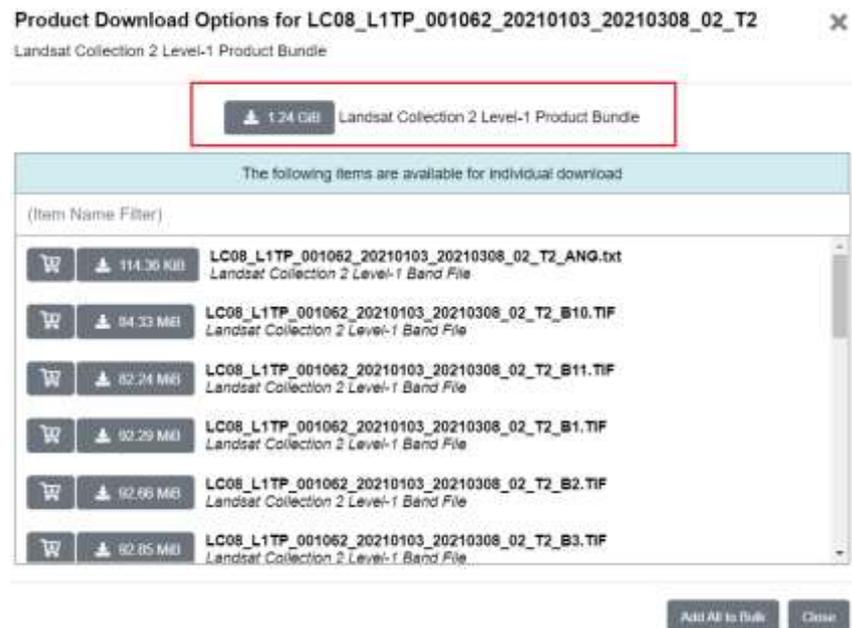


Figura 22. Descarga de paquetes de Imagen del LANSAT 8 OLI.

Fuente: Elaboración propia

Anexo 4: Pasos para convertir los valores digitales a reluctancia manualmente

1. Leer el archivo .TXT que viene junto con el paquete de descargas de las imágenes del satélite tiene un sufijo _MTL.

 LC08_L1TP_010065_20210203_20210306_01_T1_MTL.txt

Figura 23. Archivo Metadata del conjunto de imágenes multiespectrales.

Fuente: Elaboración propia

2. Buscar las variables de REFLECTANCE_MULT_BAND_* y REFLECTANCE_ADD_BAND_*

```

REFLECTANCE_MAXIMUM_BAND_1 = 1.210700
REFLECTANCE_MINIMUM_BAND_1 = -0.099980
REFLECTANCE_MAXIMUM_BAND_2 = 1.210700
REFLECTANCE_MINIMUM_BAND_2 = -0.099980
REFLECTANCE_MAXIMUM_BAND_3 = 1.210700
REFLECTANCE_MINIMUM_BAND_3 = -0.099980
REFLECTANCE_MAXIMUM_BAND_4 = 1.210700
REFLECTANCE_MINIMUM_BAND_4 = -0.099980
REFLECTANCE_MAXIMUM_BAND_5 = 1.210700
REFLECTANCE_MINIMUM_BAND_5 = -0.099980
REFLECTANCE_MAXIMUM_BAND_6 = 1.210700
REFLECTANCE_MINIMUM_BAND_6 = -0.099980
REFLECTANCE_MAXIMUM_BAND_7 = 1.210700
REFLECTANCE_MINIMUM_BAND_7 = -0.099980
REFLECTANCE_MAXIMUM_BAND_8 = 1.210700
REFLECTANCE_MINIMUM_BAND_8 = -0.099980
REFLECTANCE_MAXIMUM_BAND_9 = 1.210700
REFLECTANCE_MINIMUM_BAND_9 = -0.099980
SUN_ELEVATION = 58.03269324

```

Figura 24. Variables que se necesitan para calcular la reflectancia.

Fuente: Elaboración propia

3. En base a las variables de reflectancia, se procede a crear la tabla utilizando solo las bandas 1-7.

$$p\lambda' = MpQcal + Ap$$

$$p\lambda = \frac{p\lambda'}{\sin(\theta_{SE})}$$

Tabla 26.

Variables para el cálculo de reflectancia de cada banda.

BANDA	REFLECTANCE_MULT_BAND (M_p)	REFLECTANCE_MULTI_BAND (A_p)
BANDA 1	2×10^{-5}	-0.10
BANDA 2	2×10^{-5}	-0.10
BANDA 3	2×10^{-5}	-0.10
BANDA 4	2×10^{-5}	-0.10
BANDA 5	2×10^{-5}	-0.10
BANDA 6	2×10^{-5}	-0.10
BANDA 7	2×10^{-5}	-0.10

Fuente: Elaboración propia

Tabla 27.

Variables de ángulo de elevación solar para el cálculo de reflectancia.

ANGULO ELEVACION SE (grados)	ANGULO ELEVACION SE (radianes)
58.03269324	1.0128615708353

Fuente: Elaboración Propia

- Ubicar un pixel en una imagen multiespectral con ImageJ

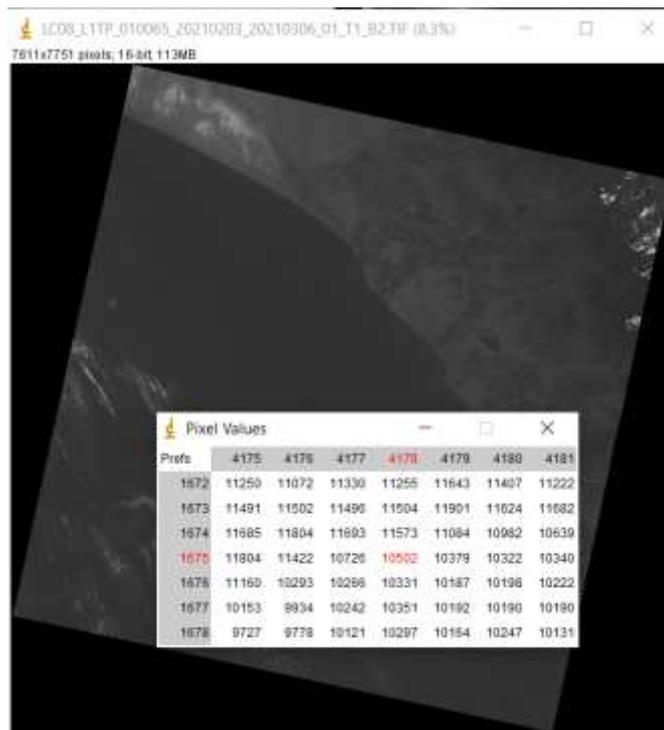


Figura 25. Ubicación del valor de un pixel con ImageJ

Fuente: Elaboración Propia

Tabla 28.

Ubicación y valor del pixel en una imagen multiespectral.

BANDA	VALOR DEL PIXEL	FILA EN LA IMAGEN	COLUMNA EN LA IMAGEN
BANDA 2	10502	4178	1675

$$p\lambda' = MpQcal + Ap$$

$$p\lambda' = 2 * 10^{-5} * 10502 + -0.10$$

$$p\lambda' = 0.00002 * 10502 + -0.10$$

$$p\lambda' = 0,11004$$

$$p\lambda = \frac{p\lambda'}{\sin(\theta_{SE})}$$

$$p\lambda = \frac{0,11004}{\sin(0,11004)}$$

$$p\lambda = 0,895716$$

El resultado de reflectancia en la Banda 2 de la posición (4178 x 1675) es de 0,895716.

Anexo 5: Programa en Python que convierte los valores de reflectancia automáticamente

1. Métodos para leer archivos de la metadata, cada método está relacionados para efectuar una determinada orden, buscar metadata, es donde busca el archivo TXT donde contiene las características de la imagen, buscar palabra la cual busca los nombres de variables, luego extraer bandas para remplazar los nombres de variables por datos vacíos así obtener solo el valor requerido y por último extraer bandas , donde obtenemos el valor para efectuar los cálculos.

```
def buscarMetadata(self, cadena):
    nlist = []
    for el in self.lista:
        c = self.buscarPalabra(cadena, el)
        if c is not None:
            nlist.append(self.buscarPalabra(cadena, el))

    return nlist

def buscarPalabra(self, cadena, p):
    if list(filter(lambda x: cadena in x, p)):
        return p
```

```

def extraerBandas(self):
    busq = self.buscarMetadatos("REFLECTANCE_ADD_")
    for el in busq:
        temp = el[0].replace("REFLECTANCE_ADD_", "")
        self.bandas.append(temp)

def leerMetadatos(self, ruta):
    f = open(ruta, "r")
    lines = f.readlines()
    for line in lines:
        line = line.replace(' ', '')
        line = line.replace('\n', '')
        line = line.replace("'", '')
        tline = line.split("=")
        self.lista.append(tline)
    self.extraerBandas()
    f.close

```

2. Método para obtener los píxeles de una imagen y guardarlas en una matriz numérico y calcular la reflectancia.

```

def leerImagen(self, rutaimg):
    dataset = gdal.Open(rutaimg)
    metadata = dataset.GetGeoTransform()
    band = dataset.GetRasterBand(1)
    transform = dataset.GetGeoTransform()
    self.xOrigin = transform[0]
    self.yOrigin = transform[3]
    self.pixelWidth = transform[1]
    self.pixelHeight = -transform[5]
    self.matriz = band.ReadAsArray(0, 0, dataset.RasterXSize,
dataset.RasterYSize).astype(float)
    self.calcularReflectancia(self.matriz)

def calcularReflectancia(self, matriz):
    sun_correction =
math.sin(math.radians(self.banda.SUN_ELEVATION))
    self.reflectancia = ((matriz *
self.banda.REFLECTANCE_MULT_BAND) +
self.banda.REFLECTANCE_ADD_BAND) / sun_correction

```

3. Clase donde se implementa los métodos para extraer los valores de la metadata.

```

from metadata import Metadata
from imagen import Imagenes, Imagen, Banda
from osgeo import gdal
import matplotlib.pyplot as plt

if __name__ == '__main__':
    metadata=Metadata()
    metadata.leerMetadatos(metadata.ruta)

```

4. Aplicando el algoritmo y utilizando el método para el cálculo de la reflectancia de las distintas bandas.

```
for ban in metadata.bandas[:7]:

    ruta = metadata.buscarMetadada("FILE_NAME_" + ban)
    print(ruta)
    if ruta is not None:
        banda = Banda()
        banda.nombre = ban
        banda.REFLECTANCE_MULT_BAND =
float(metadata.buscarMetadada("REFLECTANCE_MULT_" + ban)[0][1])
        banda.REFLECTANCE_ADD_BAND =
float(metadata.buscarMetadada("REFLECTANCE_ADD_" + ban)[0][1])
        banda.SUN_ELEVATION =
float(metadata.buscarMetadada("SUN_ELEVATION")[0][1])
        img = Imagen()
        img.banda = banda
        img.ruta = img.directorio + "\\\" + ruta[0][1]
        rutas.append(img.ruta)
        imagenes.lista.append(img)

for i in imagenes.lista:
    #valor=valor+1

    print("leyendo una img " + i.banda.nombre)
    #print(valor)

    i.leerIndice(i.ruta)
    i.leerImagen(i.ruta)
    print(i.buscarPixel(lista_cordernadas))
```

Anexo 6: Calcular los valores de los indicadores a través de código Python usando los valores de reflectancia obtenidos de distintas bandas.

1. Para el cálculo se utiliza los valores de reflectancia de distintas bandas, aplicando las distintas fórmulas para cada indicador

```
listadvni=[]
listadvi=[]
lista12=[]
listabl=[]
listasi=[]
listasi11=[]
listasi1=[]
listaAstersi=[]
listasi3=[]
for x in range(30):
    ndvi=(banda5[x]-banda4[x])/(banda5[x]+banda4[x])
    dvi=banda5[x]-banda4[x]
    s12=banda4[x]/banda5[x]
    b1=math.sqrt(pow(banda3[x],2)+pow(banda4[x],2))
    si=math.sqrt((banda2[x])+pow(banda4[x],2))
```

```

si11=banda5[x]/banda6[x]
si1=math.sqrt(banda2[x]*banda4[x])
asterSi=(banda5[x]-banda6[x])/(banda5[x]+banda6[x])
si3=(banda3[x]*banda4[x])/(banda2[x])
listadvni.append(ndvi)
listadvi.append(dvi)
lista12.append(s12)
listab1.append(b1)
listasi.append(si)
listasi11.append(si11)
listasi1.append(si1)
listaAstersi.append(asterSi)
listasi3.append(si3)

```

Anexo 7: Pasos para ubicar ubicación geoespacial y obtener valores de los pixeles para hacer el cálculo de una manera automática a través de un algoritmo en Python.

1. La ecuación para obtener la ubicación (fila y columna) de un pixel en una imagen (m filas y n columnas), mediante la coordenada de este es la siguiente.

$$fila = \frac{A - X_0}{m}$$

$$columna = \frac{Y_0 - B}{m}$$

Dónde:

A: es la coordenada horizontal a buscar

X_0 : La coordenada inicial horizontal

m: Cantidad de pixeles horizontales

B: es la coordenada vertical a buscar

Y_0 : La coordenada inicial horizontal n: Cantidad de pixeles verticales

2. Identificar las coordenadas a buscar para obtener el valor del pixel
La ecuación para convertir una coordenada UTM 17S A UTM 17N es la siguiente:

(X,Y)

$$X_N = X_S$$

$$Y_N = Y_S - 1000000$$

Donde:

X: Es la coordenada horizontal

Y: Es la coordenada vertical

X_N, Y_N : Coordenadas UTM 17N

X_S, Y_S : Coordenadas UTM 17S

Tabla 29.

Información de las coordenadas y la banda para buscar el valor del pixel.

COORDENADA UTM 17S	IMAGEN
617655,9266050	BANDA 1

Fuente: Elaboración propia

Conversión de la coordenada:

$$617655, 9266050 \text{ (UTM 17S)} = 617655, -733950 \text{ (UTM 17N)}$$

Tabla 30.

Ubicación y valor del pixel en una imagen espectral.

A	B	X_0	Y_0	m	n
617655	-733950	492285	-683685	30	30

Fuente: Elaboración Propia

3. Reemplazar valores en la formula

$$fila = \frac{A - X_0}{m}$$
$$fila = \frac{617655 - 492285}{30}$$

$$fila = 4179$$

$$columna = \frac{Y_0 - B}{n}$$
$$columna = \frac{-683685 - (-733950)}{30}$$
$$columna = 1675$$

ANEXO 8: Programa en Python para ubicar coordenadas en una imagen con información geoespacial.

1. Función que lee la imagen para obtener las variables necesarias para efectuar la fórmula de búsqueda del pixel, como la coordenada de origen, altura y ancho de la imagen así poder utilizarlas en el próximo método.

```
def leerIndice(self, rutaI):
    dataset = gdal.Open(rutaI)
    band = dataset.GetRasterBand(1)
    transform = dataset.GetGeoTransform()
    self.xOrigin = transform[0]
    self.yOrigin = transform[3]
    self.pixelWidth = transform[1]
    self.pixelHeight = -transform[5]
    self.reflectancia = band.ReadAsArray(0, 0, dataset.RasterXSize,
dataset.RasterYSize).astype(float)
```

2. Función en para convertir UTM 17S (variable epsg:32717) a UTM 17N (variable epsg:32617), utilizando las variables obtenidas en la función anterior.

```
def buscarPixel(self, coordenadas):
    valores = []
    for point in coordenadas:
        x, y = pyproj.transform(pyproj.Proj(init='epsg:32717'),
pyproj.Proj(init='epsg:32617'), point[0], point[1])

        row = int((x - self.xOrigin) / self.pixelWidth)
        col = int((self.yOrigin - y) / self.pixelHeight)
        reflec = self.reflectancia[col, row]
        valores.append(reflec)
        print(row,col,reflec)
    return valores
```

3. Ejecución de las Funciones anteriores para obtener los datos respectivas de las imágenes , Fila, Columna y Reflectancia.

```
for i in imagenes.lista:
    print("leyendo una img " + i.banda.nombre)
    i.leerIndice(i.ruta)
    i.leerImagen(i.ruta)
    print(i.buscarPixel(lista_cordenadas))
```

4. Exportar la información obtenida en un archivo CSV

```
df = pd.DataFrame(data, columns = ['MUESTRAS', 'RED', 'GREEN',
'BLUE', 'NDVI', 'DVI', 'BI', 'ASTERSI', 'SI', 'SI1', 'SI2', 'SI3', 'SI11', 'C
E'])
df.to_csv('datos.csv')
dataset = pd.read_csv('datos.csv')
```

Anexo 9: Análisis de muestras por conductividad eléctrica.

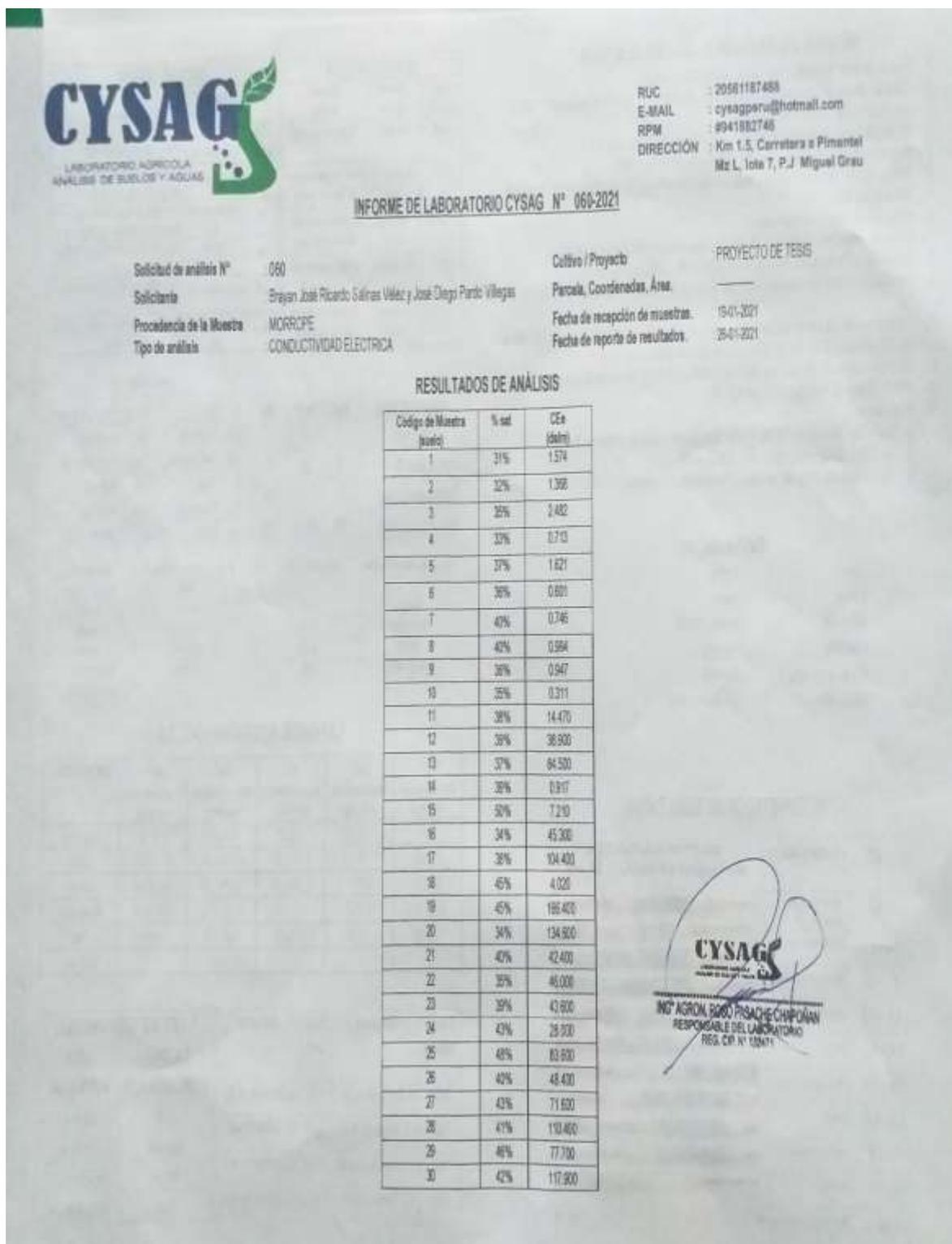


Figura 26. Resultados del laboratorio de las pruebas CE.

Fuente: Elaboración propia

Anexo 10: Cuadro de Verificación de la estimación de conductividad eléctrica.

Tabla 31.

Cuadro de Verificación de la estimación de conductividad eléctrica.

MUESTA	CE Real(x)	SLR		MLR		RFR		DTR	
		CE Estimada	Diferencia						
1	1,574	-9,15	10,724	-9,15	10,724	2,86	-1,286	1,471	0,103
2	1,368	-9,15	10,518	-9,15	10,518	2,86	-1,492	1,471	-0,103
3	2,482	16,933	-14,451	16,933	-14,451	3,314	-0,832	2,482	0
4	0,713	-8,052	8,765	-8,052	8,765	7,178	-6,465	0,713	0
5	1,621	-0,381	2,002	-0,381	2,002	3,799	-2,178	1,621	0
6	0,601	-3,852	4,453	-3,852	4,453	3,374	-2,773	0,601	0
7	0,746	19,472	-18,726	19,472	-18,726	3,367	-2,621	0,746	0
8	0,994	8,688	-7,694	8,688	-7,694	2,29	-1,296	0,994	0
9	0,947	9,944	-8,997	9,944	-8,997	10,103	-9,156	0,947	0
10	0,311	35,936	-35,625	35,936	-35,625	20,675	-20,364	3,76	-3,449
11	14,47	44,929	-30,459	44,929	-30,459	31,178	-16,708	14,47	0
12	38,9	66,148	-27,248	66,148	-27,248	59,461	-20,561	38,9	0
13	84,5	98,547	-14,047	98,547	-14,047	121,242	-36,742	125,45	-40,95
14	0,917	7,923	-7,006	7,923	-7,006	11,132	-10,215	0,917	0
15	7,21	35,936	-28,726	35,936	-28,726	20,675	-13,465	3,76	3,45
16	45,3	85,753	-40,453	85,753	-40,453	44,026	1,274	45,3	0
17	104,4	72,065	32,335	72,065	32,335	86,39	18,01	104,4	0
18	4,2	41,579	-37,379	41,579	-37,379	9,401	-5,201	4,2	0
19	166,4	98,547	67,853	98,547	67,853	121,242	45,158	125,45	40,95
20	134,6	51,458	83,142	51,458	83,142	102,478	32,122	134,6	0
21	42,4	57,322	-14,922	57,322	-14,922	47,408	-5,008	44,2	-1,8

22	46	57,322	-11,322	57,322	-11,322	47,408	-1,408	44,2	1,8
23	43,6	58,814	-15,214	58,814	-15,214	65,916	-22,316	43,6	0
24	28	66,089	-38,089	66,089	-38,089	32,534	-4,534	28	0
25	83,6	24,418	59,182	24,418	59,182	57,512	26,088	83,6	0
26	48,4	70,334	-21,934	70,334	-21,934	48,656	-0,256	48,4	0
27	71,6	49,383	22,217	49,383	22,217	65,059	6,541	71,6	0
28	110,4	101,857	8,543	101,857	8,543	85,563	24,837	110,4	0
29	77,7	84,905	-7,205	84,905	-7,205	64,335	13,365	77,7	0
30	117,9	48,135	69,765	48,135	69,765	85,314	32,586	117,9	0
MEDIA	X=42,72846667	Y=42,7284		Y=42,7284		Y=42,225		Y=42,72843333	