

 | UNIVERSIDAD  
SEÑOR DE SIPÁN

**FACULTAD DE INGENIERÍA, ARQUITECTURA Y  
URBANISMO**

**ESCUELA ACADÉMICO 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**

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

**Autor:**

**Bach. Vilchez Guivar Daniel Steven**

**Asesor:**

**Ing. Heber Iván Mejía Cabrera**

**Línea de Investigación:**

**Ciencias de la computación**

**Pimentel, Perú**

**2019**

## INFORMACIÓN GENERAL

### 1.1 Título del Informe de Investigación:

“COMPARACIÓN DE TÉCNICAS DE ESTIMACIÓN DEL GRADO DE SALINIDAD EN SUELOS DE ESCASA VEGETACIÓN MEDIANTE EL PROCESAMIENTO DE IMÁGENES MULTIESPECTRALES.”

### 1.2 Línea de Investigación:

Ciencias de la computación

### 1.3 Autor:

Bach. Vilchez Guivar Daniel Steven

### 1.4 Asesor Metodólogo:

Ing. Heber Iván Mejía Cabrera

### 1.5 Tipo y diseño de investigación.

Tipo experimental, metodología cuantitativa.

### 1.6 Facultad y Escuela Académico Profesional:

Facultad de Ingeniería, Arquitectura y Urbanismo  
Escuela Profesional de Ingeniería de Sistemas

### 1.7 Periodo: 2019-II

### 1.8 Fecha de inicio y término del proyecto:

Enero – Diciembre 2019

### 1.9 Firma de los autores del proyecto:

---

Vilchez Guivar Daniel Steven  
**AUTOR**

### 1.10 Aprobado:

---

Ing. Heber Iván Mejía Cabrera  
**ASESOR METODOLÓGICO**

### 1.11 Fecha de Presentación:

**Diciembre del 2019**



## DEDICATORIA

A Dios.

Por haberme brindado la salud y la fortaleza de poder culminar el presente trabajo de investigación.

A mi familia.

Por el apoyo y la motivación constante que me ha permitido fortalecer mis valores y principios, convirtiéndome en una persona de bien.



## AGRADECIMIENTO

A Dios por encaminar mi trayecto, de la misma manera a mi familia por ser mi fortaleza para salir adelante, así mismo agradezco a mis docentes y todas las personas que cooperaron para que pueda lograr mis metas, y que me brindaron su apoyo, siendo constante y firmes.



## INDICE

DEDICATORIA .....	3
AGRADECIMIENTO.....	4
RESUMEN.....	12
ABSTRACT .....	13
I. INTRODUCCION.....	14
1.1 Realidad problemática .....	14
1.2 Antecedentes de Estudio .....	17
1.2.1 A nivel mundial .....	17
1.2.2 A nivel nacional .....	18
1.2.3 Estado del arte .....	19
1.3 Teorías relacionadas al tema .....	23
1.3.1 Conductividad Eléctrica (CE) .....	23
1.3.2 Radiación Electromagnética .....	24
1.3.3 Espectro Electromagnético .....	25
1.3.4 Radiancia Espectral.....	26
1.3.5 Reflectancia Espectral .....	27
1.3.6 Firma Espectral .....	27
1.3.7 Teledetección o Detección Remota .....	28
1.3.8 Técnicas de estimación .....	31
1.4 Formulación del problema.....	38
1.5 Justificación e importancia del estudio. ....	38
1.6 Hipótesis .....	38
1.7 Objetivos .....	39
1.7.1 Objetivo General.....	39



1.7.2	Objetivos Específicos .....	39
II.	MATERIAL Y MÉTODOS .....	39
2.1	Tipo y diseño de la investigación.....	39
2.1.1	Tipo de investigación .....	39
2.1.2	Diseño de la investigación .....	39
2.2	Población y muestra.....	40
2.3	VARIABLES Y OPERACIONALIZACIÓN .....	40
2.3.1	VARIABLES .....	40
2.3.2	Operacionalización de Variables .....	41
2.4	Técnicas e instrumentos de recolección de datos. ....	42
2.4.1	Abordaje Metodológico .....	42
2.4.2	Técnicas de Recolección de datos .....	42
2.4.3	Procedimiento de recolección de datos .....	42
2.5	Procedimientos de análisis de datos .....	42
2.6	Criterios éticos .....	43
2.7	Principios de rigor científico .....	43
III.	RESULTADOS .....	44
3.1	Resultados en Tablas y Figuras .....	44
3.2	Discusión de resultados .....	46
3.3	Aporte Práctico .....	49
3.3.1	Selección de las técnicas de estimación a utilizar.....	49
3.3.2	Adquisición de imágenes multiespectrales del área de estudio ...	50
3.3.3	Realización de calibraciones atmosféricas a las imágenes multiespectrales adquiridas. ....	55
3.3.4	Selección indicadores de salinidad y vegetación. ....	58



3.3.5	Recolección muestras de suelos de escasa vegetación y realizar la prueba de conductividad eléctrica. ....	63
3.3.6	Evaluación de las técnicas de estimación. ....	67
IV.	CONCLUSIONES Y RECOMENDACIONES.....	74
4.1	Conclusiones .....	74
4.2	Recomendaciones .....	75
	REFERENCIAS.....	76
	ANEXOS .....	79
	Anexo 1: Población de Técnicas de estimación y la muestra escogida.....	79
	Anexo 2: Pasos para recolectar las imágenes multiespectrales.....	81
	Anexo 3: Pasos para convertir los valores digitales a reflectancia manualmente.....	85
	Anexo 4: Programa para convertir los valores digitales a reflectancia automáticamente. ....	89
	Anexo 5: Pasos para calcular los indicadores manualmente.....	91
	Anexo 6: Programa para calcular los indicadores de una imagen automáticamente. ....	93
	Anexo 7: Pasos para ubicar una coordenada en una imagen con información geoespacial, mediante una interfaz gráfica. ....	95
	Anexo 8: Programa para ubicar coordenadas en una imagen con información geoespacial. ....	97
	Anexo 9: Análisis de muestras por conductividad eléctrica. ....	99





## INDICE DE TABLAS

Tabla 1: <i>Superficie Agropecuaria en el Perú</i> .....	14
Tabla 2: <i>Capacidad de uso de las tierras en el Perú</i> .....	14
Tabla 3: <i>Clasificación de salinidad de suelo según valores de conductividad eléctrica (EC)</i> .....	24
Tabla 4: <i>Límites del espectro electromagnético según su división</i> .....	25
Tabla 5: <i>Límites de la luz visible según los colores RGB</i> .....	26
Tabla 6: <i>Resolución espacial de satélites.</i> .....	29
Tabla 7: <i>Población de técnicas de estimación</i> .....	49
Tabla 8: <i>Bandas del sensor OLI del satélite Landsat 8</i> .....	54
Tabla 9: <i>Archivos adicionales incluidos del satélite Landsat 8</i> .....	54
Tabla 10: <i>Imágenes espectrales de reflectancia, banda y su identificador.</i> .....	58
Tabla 11: <i>Correlación entre los indicadores de salinidad, vegetación y bandas, y el valor de conductividad eléctrica.</i> .....	59
Tabla 12: <i>Correlación entre los indicadores de salinidad, vegetación, y el valor de conductividad eléctrica</i> .....	60
Tabla 13: <i>Imágenes multiespectrales procesadas que se usarán para la estimación</i> ....	62
Tabla 14: <i>Ubicación de las muestras tomadas y su respectivo valor de conductividad eléctrica.</i> .....	65
Tabla 15: <i>Relación entre los indicadores elegidos y el valor de conductividad eléctrica.</i> .....	66
Tabla 16: <i>Resultados de la estimación usando SVM, relación entre la conductividad eléctrica estimada y real</i> .....	68
Tabla 17: <i>Resultados de la estimación usando PLSR, relación entre la conductividad eléctrica estimada y real</i> .....	70
Tabla 18: <i>Cálculo del indicador <math>R^2</math> para SVM</i> .....	71
Tabla 19: <i>Cálculo del indicador RMSE para SVM</i> .....	72
Tabla 20: <i>Cálculo del indicador <math>R^2</math> para PLSR</i> .....	73
Tabla 21: <i>Cálculo del indicador RMSE para PLSR</i> .....	73
Tabla 22: <i>Población de técnicas de estimación</i> .....	79
Tabla 23: <i>Variables de cálculo de reflectancia por banda</i> .....	86





Tabla 24: <i>Variables de ángulo de elevación para el cálculo de reflectancia</i> .....	86
Tabla 25: <i>Ubicación y valor de un pixel en una imagen espectral</i> .....	87
Tabla 26: <i>Información de coordenada a buscar y el indicador donde se buscará</i> .....	96
Tabla 27: <i>Ubicación y valor de un pixel en una imagen espectral</i> .....	96



## INDICE DE FIGURAS

Figura 1: <i>Clasificación del espectro electromagnético</i> .....	25
Figura 2: <i>Espectro electromagnético característico del suelo desnudo, vegetación y agua</i> .....	28
Figura 3 : <i>Proceso de calibración radiométrica y geométrica de imágenes</i> .....	30
Figura 4 : <i>Múltiples hiperplanos encontrados para una solución en SVM</i> .....	34
Figura 5 : <i>Hiperplano de menor margen</i> .....	35
Figura 6 : <i>Estimación de SVR, con un margen de tolerancia épsilon</i> .....	36
Figura 7: <i>Gráfica de evaluación de <math>R^2</math> con 0.32 y RMSE de 41</i> .....	37
Figura 8: <i>Comparación de PLSR y SVM en el indicador de precisión <math>R^2</math></i> .....	44
Figura 9: <i>Comparación de PLSR y SVM en el indicador de precisión RMSE</i> .....	45
Figura 10: <i>Gráfica de porcentaje de correlación entre los indicadores usados y el valor de conductividad eléctrica</i> .....	45
Figura 11: <i>Indicador NDVI a color</i> .....	46
Figura 12: <i>Indicador NDVI, en la plataforma de Google Maps</i> .....	47
Figura 13: <i>Estimación del grado de salinidad usando SVM</i> .....	47
Figura 14: <i>Estimación del grado de salinidad usando SVM, en la plataforma de Google Maps</i> .....	48
Figura 15: <i>Paquetes de imágenes del repositorio de Landsat 8</i> .....	51
Figura 16: <i>Estructura de la carpeta de las imágenes adquiridas de Landsat 8</i> .....	52
Figura 17: <i>Imágenes multiespectrales adquiridas del satélite Landsat 8</i> .....	53
Figura 18: <i>Imágenes de reflectancia generadas al procesar las imágenes</i> .....	57
Figura 19: <i>Imágenes procesadas con los indicadores antes mencionados</i> .....	61
Figura 20: <i>Ubicación de las muestras recogidas en la ciudad de Mórrope</i> .....	63
Figura 14: <i>Procedimiento para la medición de Conductividad eléctrica</i> .....	64
Figura 22: <i>Variables independientes (indicadores)</i> .....	67
Figura 23: <i>Variable dependiente (CE)</i> .....	67
Figura 24: <i>Código de implementación de estimación SVM</i> .....	68
Figura 25: <i>Código de implementación de estimación PLSR</i> .....	69
Figura 26: <i>Búsqueda de imágenes multiespectrales por lugar</i> .....	81
Figura 27: <i>Búsqueda de imágenes multiespectrales por fecha</i> .....	82
Figura 28: <i>Búsqueda de imágenes multiespectrales por satélite</i> .....	83



Figura 29: <i>Búsqueda de imágenes multiespectrales por filtro de porcentaje de nubosidad.....</i>	83
Figura 30: <i>Vista de adquisición de imágenes multiespectrales. ....</i>	84
Figura 31: <i>Adquisición de paquete de imágenes multiespectrales. ....</i>	84
Figura 32: <i>Archivo de metadata de un paquete de imágenes multiespectrales.....</i>	85
Figura 33: <i>Variables de metadata a utilizar para calcular la reflectancia.....</i>	85
Figura 34: <i>Búsqueda de valor de un pixel en una imagen espectral.....</i>	87
Figura 35: <i>Código fuente en Python para leer el archivo de metadata .....</i>	89
Figura 36: <i>Código fuente en Python de la función para leer una imagen .....</i>	89
Figura 37: <i>Código fuente en Python para calcular la reflectancia.....</i>	90
Figura 38: <i>Código fuente en Python para calcular la reflectancia.....</i>	90
Figura 39: <i>Código fuente en Python de la ejecución de las funciones de leer de imagen y leer la metadata.....</i>	90
Figura 40: <i>Búsqueda del valor de reflectancia de una imagen procesada .....</i>	91
Figura 41: <i>Búsqueda del valor de reflectancia de una otra imagen procesada.....</i>	92
Figura 42: <i>Código fuente en Python de lectura de una imagen espectral procesada ....</i>	93
Figura 43: <i>Código fuente en Python implementación del indicador NDVI.....</i>	93
Figura 44: <i>Código fuente en Python de lectura de los puntos de origen y el tamaño de pixel de una imagen espectral .....</i>	93
Figura 45: <i>Código fuente en Python de la exportación de un indicador calculado.....</i>	94
Figura 46: <i>Código fuente en Python de lectura de un indicador.....</i>	97
Figura 47: <i>Código fuente en Python de lectura de coordenada, conversión y búsqueda de coordenada. ....</i>	97
Figura 48: <i>Código fuente en Python de ejecución de búsqueda de coordenadas y extracción de indicadores.....</i>	98
Figura 49: <i>Código fuente en Python de ejecución de búsqueda de coordenadas y extracción de indicadores.....</i>	98
Figura 50: <i>Análisis de muestras por conductividad eléctrica. ....</i>	99



## RESUMEN

En el año 2005, El Ministerio de Agricultura (MINAG) realizó un estudio de las zonas agrícolas en el Perú, donde estimó que el 0.24% de 128 521 500 hectáreas de producción agrícola, estaban afectadas por problemas de salinización, y la totalidad de estas se ubican en la Costa del Perú.

Asimismo, el plan estratégico del 2009 al 2015 del sector agrario de la región Lambayeque, reportó que el sector agrario de la región aportó 0.8% al Producto Bruto Interno (PBI) nacional, y 16,2% a la formación del PBI regional, además los causantes del bajo desarrollo agrario de la región son: el deterioro de suelos por efecto de la salinidad, la escasez de agua y la débil organización de productores.

La noticia publicada por el portal del Gobierno Regional de Agricultura Lambayeque el 15 agosto del 2018, en relación con el distrito de Mórrope, detalla que Mórrope sufre de una baja productividad agrícola debido al cultivo del arroz, puesto que dicho cultivo emplea demasiada agua, y sumado a la falta de agua, ocasionó que los terrenos tengan altos niveles de salinidad, y la solución que proponen es el cultivo de algodón en reemplazo al cultivo de arroz.

Basado en estos datos 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” para que mediante el procesamiento de imágenes multiespectrales podamos estimar la salinidad de los terrenos de escasa vegetación, y así aprovecharlos mediante la agricultura.

Basado en investigaciones relacionadas a estimación a esta se propuso utilizar las técnicas de estimación SVM y PLSR, y utilizar imágenes multiespectrales como medio de extracción de características a partir de indicadores de salinidad y vegetación, y utilizando como validación las pruebas de conductividad eléctrica para el grado de salinidad, la evaluación de estas técnicas demostró que SVM tiene un 98% de precisión y un error promedio de 0.22 dS/m, y PLSR 97% y 0.51 dS/m respectivamente.

## PALABRAS CLAVE

Salinidad en suelos, conductividad eléctrica, SVM, PLSR, imágenes multiespectrales



## ABSTRACT

In 2005, El Ministerio de Agricultura (MINAG) conducted a study of agricultural areas in Peru, where estimated that 0.24% of 128 521 500 hectares of agricultural production, was related to problems of salinization, and all of these They are located on the Coast of Peru.

The strategic plan from 2009 to 2015 of the agricultural sector of the Lambayeque region was also included, report of the agricultural sector of the region contributed 0.8% to the national Gross Domestic Product (GDP), and 16.2% to the formation of the regional GDP, in addition The limiting factors for the agrarian development of the region are: the deterioration of the soil due to the effect of salinity, the scarcity of water and the weak organization of the producers.

The news published on the website of the Regional Gobierno Regional de Agricultura Lambayeque, 2018, in a relationship with the district of Mórrope. Summary to the lack of water, the occasion of the land, the high levels of salinity, and the solution proposed by the cultivation of cotton in the replacement of rice cultivation. The results are explained in the research work. Vegetation, and thus take advantage of them through agriculture.

The investigations based on the relations of the characteristics of the results of the same and the use of the same ones of salinity, the evaluation of these techniques showed that SVM has 98% accuracy and an average error of 0.22 dS / m, and PLSR 97% and 0.51 dS / m respectively.

## KEYWORDS

Soil salinity, estimation, electrical conductivity, SVM, PLSR, multispectral images





## I. INTRODUCCION

### 1.1 Realidad problemática

Entre los principales efectos de la degradación de suelo se encuentran, la pérdida de elementos nutritivos, la desertificación, la salinización, deterioro de la estructura del suelo, la erosión eólica e hídrica y la contaminación (Lopez, 1994).

La salinidad del suelo es un peligro ambiental presente en todo el mundo, especialmente en zonas áridas y semiáridas. La salinización del suelo ocurre principalmente debido al riego y otras actividades agrícolas intensificadas (Akramkhanov et al. 2011)

El censo agropecuario de 2012, realizado por el INEI, mostró que la superficie agrícola del Perú es de 7.1 millones de hectáreas (ha), Ver *Tabla 1*.

**Tabla 1:** Superficie Agropecuaria en el Perú

Aptitud del Suelo	Superficie (millones de ha)	Porcentaje (%)
<b>Superficie Total</b>	<b>128.5</b>	<b>100%</b>
Superficie agrícola	7.1	5.5%
Superficie pastos naturales	18.0	14.0%
Superficie Montes y bosques	10.9	8.5%
Superficie otros usos	2.6	2.0%
Superficie no agrícola	89.9	70.0%

Fuente: Censo Agropecuario 2012, INEI.

Así mismo, en el mismo censo, se estimó que solo 7.6 millones de hectáreas poseen potencial o vocación natural para cultivos agrícolas (cultivos en limpio y permanentes) y 18 millones con aptitud para pastos, Ver *Tabla 2*.

**Tabla 2:** Capacidad de uso de las tierras en el Perú

Aptitud del Suelo	Superficie (millones de ha)	Porcentaje (%)
<b>Superficie Total</b>	<b>128.5</b>	<b>100%</b>
Superficie agrícola	7.1	5.5%
Superficie pastos naturales	18.0	14.0%
Superficie Montes y bosques	10.9	8.5%
Superficie otros usos	2.6	2.0%
Superficie no agrícola	89.9	70.0%

Fuente: ONERN, 1981.



En el año 2005, El Ministerio de Agricultura (MINAG) realizó un estudio de las zonas agrícolas en el Perú, donde estimó que el 0.24% de 128 521 500 hectáreas de producción agrícola, estaban afectadas por problemas de salinización, y la totalidad de estas se ubican en la Costa del Perú.

Asimismo el plan estratégico del 2009 al 2015 del sector agrario de la región Lambayeque, reportó que el sector agrario de la región aportó 0.8% al Producto Bruto Interno (PBI) nacional, y 16,2% a la formación del PBI regional, además las causantes del bajo desarrollo agrario de la región son: el deterioro de suelos por efecto de la salinidad, la escasez de agua y la débil organización de productores.

Larsen y Strukova (2005) estiman que la pérdida de ganancias causada por la salinización para los agricultores es de 918 millones de soles anuales en el Perú.

La noticia publicada por el portal del Gobierno Regional de Agricultura Lambayeque el 15 agosto del 2018, en relación con el distrito de Mórrope, detalla que Mórrope sufre de una baja productividad agrícola debido al cultivo del arroz, puesto que dicho cultivo emplea demasiada agua, y sumado a la falta de agua, ocasionó que los terrenos tengan altos niveles de salinidad, y la solución que proponen es el cultivo de algodón en reemplazo al cultivo de arroz.

Por lo tanto, para evitar el riesgo de salinización, es importante controlar la salinidad del suelo y mantenerlo por debajo de los límites de tolerancia de salinidad de los cultivos.(Bouksila et al. 2013).

Un método tradicional usado para estimar la salinidad en suelos es mediante la prueba de Conductividad Eléctrica (CE), que consiste en la medición física de la resistencia eléctrica de extractos de suelos saturados; o una variación más usada actualmente, la Conductividad Eléctrica aparente (CEa), mediante sondas EMI (Inducción electromagnética). (Heil and Schmidhalter 2017).

En 1999, la Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO) estimó que para determinar la salinidad del suelo en Estados Unidos de un terreno de 64 hectáreas se necesitan 144 muestras, el costo promedio de la prueba de Conductividad Eléctrica (EC) es de \$20 por muestra, eso sumado a otros pagos ascendería a un total de \$9340.80, pero para la prueba de Conductividad Eléctrica Aparente (ECa)





usando el sensor EM-38, los costos se reducen significativamente a un total de \$97.32, pero eso sumado al costo de sensor de \$10970.

Por lo tanto, existe una necesidad de usar metodologías prácticas y rentables para evaluar, monitorear y mapear la extensión y distribución de los suelos afectados por la salinidad.(FAO 2007)

Una alternativa en la actualidad es la teledetección, usada para adquirir datos del medio ambiente, mediante las imágenes procedentes de sensores multiespectrales como los que poseen los satélites.(Romero 2006).

Por ello, la teledetección es una alternativa en el estudio de suelos afectados por sales, ya que proporciona información de grandes áreas de terreno (Mashimbye et al. 2012).

La combinación de técnicas de teledetección (imágenes e índices espectrales) y pruebas de conductividad eléctrica con sensores como el EM-38, provee un enfoque prometedor para producir formas rápidas, de bajo costo y razonablemente precisas para monitorear, evaluar y predecir la salinidad del suelo en grandes regiones. (Ding and Yu 2014).

Además, la relación entre las propiedades del suelo y los factores ambientales depende de cada región, estos factores están relacionados con la salinidad, actualmente se usa la estimación de la Conductividad eléctrica (CE) a partir de valores multiespectrales y características del terreno para estimar la salinidad, algunos de estos incluyen regresión lineal múltiple, redes neuronales artificiales y otros enfoques de modelos estadísticos (Shahabi et al. 2017)

Por ellos es necesario comparar la técnica óptima para estimar la salinidad en los terrenos de escasa vegetación de la región.

## 1.2 Antecedentes de Estudio

### 1.2.1 A nivel mundial

(Carlos Gonzaga Aguilar, 2014). En su trabajo de investigación **“Aplicación de Índices de Vegetación Derivados de Imágenes Satelitales LANDSAT 7 ETM+ y ASTER para la Caracterización de la Cobertura Vegetal en la Zona Centro de la Provincia De Loja, Ecuador”**, de la Universidad Nacional de la Plata en Argentina. El investigador afrontó el problema de estimar el tipo de cobertura vegetal mediante el uso de la información de imágenes multiespectrales satelitales.

Determino en las imágenes, el Índice de Vegetación de Diferencia Normalizada (NDVI), Índice de Vegetación Ajustado al Suelo (SAVI), y Índice de Estrés Hídrico (MSI), uso imágenes multiespectrales provenientes de los sensores LANDSAT 7 y ASTER.

Como resultado obtuvo que los 3 índices de vegetación (NDVI, MSI, SAVI) usados permitieron discriminar e identificar los diferentes tipos de cobertura vegetal, y pudo estimar lo diferentes tipos de vegetación en las zonas.

(Raquel Anaya, 2015). En su trabajo de investigación **“Modelo de la salinidad del suelo en los Negritos Villama, Michoacán”** en el Instituto Politécnico Nacional, México. El investigador afrontó el problema de predecir la salinidad en espacio y tiempo, utilizando imágenes multiespectrales LANDSAT y muestras de conductividad eléctrica, aplicó métodos multivariados para generar grupos similares de suelos y comparo los modelos lineales generalizados y los modelos de efecto mixto eligiendo el modelo de mayor coeficiente de determinación, utilizo 240 muestras de suelo entre las profundidades de 0-30 cm y 30-60cm para su trabajo de investigación.

El resultado de los modelos se empleó para generar semivariogramas experimentales, los semivariogramas se combinaron con la técnica de extrapolación kriging y se obtuvieron los modelos de predicción de la conductividad eléctrica.



En el trabajo de investigación se concluyó que es posible generar mapas de conductividad eléctrica en espacio y tiempo, utilizo los modelos lineales generalizado (GLM) y los modelos de Efecto Mixto (MEM) para estimar la salinidad, donde el modelo MEM tuvo una mayor correlación con las muestras, el porcentaje de precisión usando GLM fue de 76% y de MEM fue 99%.

### 1.2.2 A nivel nacional

(Rolando Badaracco, 2012). En su trabajo de investigación **“Detección de la Salinidad en los Suelos de Cultivos de Algodón usando Imágenes de Satélite en la Costa Norte del Perú”**, de la Universidad Nacional Mayor de San Marcos en Perú. El investigador afrontó el problema de detectar la salinidad en los cultivos de algodón, para ello propuso estimar la salinidad haciendo uso de la teledetección, usó imágenes multiespectrales provenientes de los sensores ASTER, LANDSAT 5 TM y LANDSAT 7 ETM+, y el aplicativo ENVI 4.5 para el procesamiento de las imágenes.

En las zonas de estudio del trabajo de investigación se ubicó puntos de muestreo georreferenciados con un GPS WGS 84, Se extrajo muestras de suelo de 0 - 30 cm de profundidad, aproximadamente 1 kilogramo por muestra, se tomó 54 muestras en total divididas en las zonas de Saltur, San Antero y Ucupe.

Determinó en las imágenes los valores de reflectancia, Índice de sal, y el índice de vegetación de diferencia normalizada (NDVI); correlacionó los valores obtenidos con muestras de CE de la zona; y comparó los análisis de regresión lineal simple y compuesta en la obtención de modelos de EC estimada.

Como resultado obtuvo que la estimación usando la regresión lineal compuesta es más precisa que la regresión lineal simple, los valores obtenidos de la regresión lineal compuesta es de 0.72 en el indicador  $R^2$  o 72% de precisión.

Este trabajo de investigación es importante porque ha estimado el nivel de salinidad con imágenes multiespectrales LANDSAT y regresión lineal simple y compuesto.



(Renato Soca, 2016). En su trabajo de investigación **“Identificación de tierras degradadas por salinidad del suelo en los cultivos de caña de azúcar en Pomalca usando imágenes de satélite”** en la Universidad Nacional Mayor de San Marcos, Perú. El investigador afrontó el problema de identificar las tierras degradadas por la salinidad en las áreas de cultivos de caña de azúcar de la Empresa Agroindustrial Pomalca S.A y uso la teledetección como herramienta para desarrollar su trabajo.

Utilizó imágenes de los sensores LANDSAT 5 TM, LANDSAT 7 ETM y SPOT-5 HRG-2, el aplicativo ENVI 4.5 para el procesamiento de las imágenes.

Los datos de conductividad eléctrica fueron entregados por la Empresa Agroindustrial Pomalca S.A.A en un Excel conteniendo información de 1024 muestras, pero solo se escogieron los datos de los años 2008 y 2009.

Determinó en las imágenes los valores de reflectancia, Índice de sal, y NDVI; correlacionó los valores obtenidos con muestras de EC de la zona; y comparó los análisis de regresión lineal simple y compuesta en la obtención de modelos de EC estimada.

Mediante la estimación entre los valores de reflectancia y las pruebas de conductividad eléctrica, Utilizando la técnica de regresión múltiple para estimar la salinidad, se obtuvo un 0.73 en el indicador  $R^2$  o 73 % de precisión.

### 1.2.3 Estado del arte

(Guan et al. 2013) En el trabajo de investigación **“Dynamic prediction of soil salinization in an irrigation district based on the support vector machine”**. Los investigadores afrontaron el problema de estimar la salinidad usando Support vector machine (SVM), en un distrito irrigado, con base en la profundidad del agua subterránea, el volumen de agua de riego y los datos de evaporación de 1991 a 2010, se elaboró el modelo de predicción.



El propósito de usar SVM suelo es lograr una alta precisión de predicción en lugar de una alta precisión de entrenamiento, se utilizó el indicador Error cuadrático medio (RMSE) para mostrar qué tan bien el modelo predice las nuevas muestras.

Los resultados mostraron que los predictores SVM tienen una mayor precisión que las redes neuronales artificiales en estimar los valores de conductividad eléctrica (EC) del suelo, el valor del error relativo de la estimación fue de 0.37 o 37% de error para SVM, y 0.49 o 49% error para SVM.

(Allbed, Kumar, and Sinha 2014). En el trabajo de investigación “**Mapping and Modelling Spatial Variation in Soil Salinity in the Al Hassa Oasis Based on Remote Sensing Indicators and Regression Techniques**”. Los investigadores afrontaron el problema de mapear y modelar la variación espacial de la salinidad del suelo en el oasis de Al Hassa , para ello se propuso desarrollar modelos de regresión estadística basados en indicadores de detección remota para predecir y mapear la variación espacial, se aplicó regresión lineal múltiple para estimar la salinidad, usaron la información procedente del sensor IKONOS, y pruebas de conductividad eléctrica de suelo (EC), el análisis de los datos procedentes del sensor IKONOS mostró que el índice de salinidad (SI) y la banda roja (banda 3) tenían la mayor correlación con la EC, se utilizaron ambas como variables para realizar el modelo de estimación.

Los resultados revelaron que el alto rendimiento de este modelo de estimación se atribuye a la resolución espacial de las imágenes, al potencial de las imágenes y la superioridad de la banda 3 en la recuperación de características y patrones de salinidad del suelo, que se explica por la alta reflectancia; los mapas de salinidad del suelo generados utilizando el modelo seleccionado mostraron que los suelos fuertemente salinos eran predominantes sobre el área de estudio.

Se utilizó los indicadores de Coeficiente de determinación ( $R^2$ ) y el error cuadrático medio (RMSE), los valores de  $R^2$  indican la fuerza de la relación lineal estadística entre los valores de salinidad del suelo medidos y predichos, y RMSE indica errores de estimación absoluta.





Como resultado el modelo de estimación utilizado obtuvo un 65% de correlación entre los valores medidos y predichos.

(Mousavi et al. 2017). En el trabajo de investigación **“Digital Mapping of Topsoil Salinity Using Remote Sensing Indices in Agh-Ghala Plain, Iran”**. Los investigadores afrontaron el problema de mapeo digital de la salinidad del suelo usando índices de detección remota (teledetección), para ello se propuso comparar los modelos de regresión lineal múltiple y redes neuronales multicapa, y el uso de la información procedente del sensor LANDSAT 8 OLI, y pruebas de conductividad eléctrica de suelo (EC).

Utilizaron los índices de la banda azul, verde y roja, índices de intensidad, índices de salinidad e índices de vegetación, Índice de Vegetación de Diferencia Normalizada (NDVI), Índice de vegetación ajustado al suelo (DVI), índice de diferencia de vegetación (RVI), Índice de vegetación con diferencia ponderada (SAVI) para estimar la salinidad.

Para comparar el rendimiento y la suficiencia de los modelos de predicción regresión lineal múltiple y redes neuronales para estimar la salinidad del suelo, se usó cuatro criterios diferentes, error cuadrático medio (RMSE), RSE (Error estándar relativo), R<sup>2</sup> (coeficiente de determinación) y ME (error medio) para los valores EC estimados.

Como resultado se obtuvo que redes de neuronales multicapas fue de 92 % y 52% la regresión lineal múltiple en coeficiente de determinación, dando como conclusión que las redes neuronales estiman mejor la salinidad del suelo.

(Shahabi et al. 2017) En el trabajo de investigación **“Spatial modeling of soil salinity using multiple linear regression, ordinary kriging and artificial neural network methods”**

Los investigadores afrontaron el problema del modelado espacial de salinidad del suelo en la provincia Azerbaijan al este de Irán, usaron parámetros de la superficie



terrestre, la elevación, pendiente, el ángulo de aspecto y el índice de humedad topográfica.

Se estimó los valores de conductividad eléctrica (EC) mediante los modelos de regresión lineal múltiple, kriging ordinario y redes neuronales artificiales, y los valores de superficie terrestre, con 150 muestras de campo (120 muestra utilizadas el entrenamiento y 30 muestras en la validación).

Como resultado se obtuvo que las redes neuronales fue el que tuvo mayor precisión de los valores estimados, utilizando el indicador  $R^2$ , que los modelos de regresión lineal múltiple y kriging ordinario, las redes neuronales obtuvieron un 69% de precisión y regresión lineal múltiple 39%.

(Elhag and Bahrawi 2017) En el trabajo de investigación **“Soil salinity mapping and hydrological drought indices assessment in arid environments based on remote sensing techniques”**

Los investigadores afrontaron el problema de mapeo de salinidad del suelo y relación los índices de salinidad con los índices de sequía hidrológica, usaron la información procedente del sensor LANDSAT 8 OLI, el análisis de componentes principales (PCA) y análisis de redes neuronales artificiales (ANN) fueron herramientas complementarias utilizadas para comprender el patrón de regresión de los índices de sequía hidrológica en el área de estudio designada.

Se examinaron cuatro índices de sequía hidrológica, el índice de vegetación de suministro de agua (WSVI), el índice de vegetación ajustado al suelo (SAVI), el índice de estrés por humedad (MSI) y la diferencia normalizada índice infrarrojo (NDII).

Como resultado se obtuvo que el indicador NDII tuvo un 73% de precisión en la estimación de (Índice de salinidad de diferencia normalizada) NDSI, demostrando una relación de la salinidad y la sequía hidrológica.





### 1.3 Teorías relacionadas al tema

#### 1.3.1 Conductividad Eléctrica (CE)

Según (Flores et al, 1996), la salinidad es la acumulación de sales en los suelos, debido al resultado de procesos naturales y/o antrópicos, y el grado de acumulación de sales puede afectar a la fertilidad del suelo.

Para medir la salinidad de en los suelos se usa la conductividad eléctrica, para ello se realiza una solución de una muestra del suelo y agua, como consecuencia de las sales solubles en agua, la solución del suelo tiene conductividad que se puede medir, esta conductividad está relacionada positivamente con el contenido de sal en el suelo; por lo tanto el método de conductividad eléctrica es un método rápido y preciso para medir el contenido de sal del suelo, y representa el contenido total de sal en el suelo, en la de unidad de medida siemens por metro (S / m), la conductividad eléctrica del extracto de suelo se utiliza para reflejar el contenido de sal del suelo (Bado, Forster, and Jankowicz-cieslak 2016)

En la siguiente ecuación:

$$CE_e = CE_w \times 1.5$$

Donde:

$CE_e$  = La conductividad eléctrica de extracto saturado de suelo

$CE_w$  = La conductividad eléctrica del agua



**Tabla 3: 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

“La radiación electromagnética se genera por la transformación de energía en otras formas de energía” (Ormeño-Villajos 2006)

Expresa que en la teoría ondulatoria, la radiación electromagnética es un fenómeno ondulatorio de propagación, que está compuesta por un campo eléctrico y uno magnético, y ambos son perpendiculares. (Pérez and Muñoz 2006)

En la cual la propagación de la radiación electromagnética es la velocidad de la luz, y está formada por la ecuación:

$$c = \lambda * F$$

Donde:

$\lambda$  = Longitud de onda

$F$  = Frecuencia

Y Según la teoría cuántica, se basa en propiedades energéticas:

$$Q = h * F$$



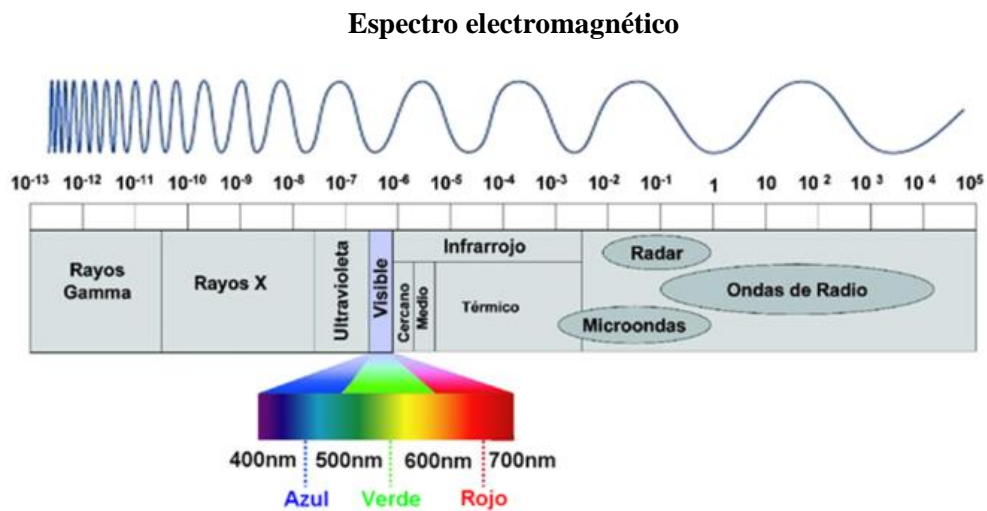
Donde:

$h$  = Es la constante de Planck

$F$  = Frecuencia

### 1.3.3 Espectro Electromagnético

Es la distribución continua de longitudes de onda y frecuencias, desde las longitudes de onda más cortas a las longitudes de onda más largas (Ormeño-Villajos 2006)



**Figura 1:** Clasificación del espectro electromagnético

**Fuente:** Radio Frecuencia España

**Tabla 4:** 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 $\mu$ m
Luz Visible	0.38–0.72 $\mu$ m
Infrarrojo cercano	0.72–1.30 $\mu$ m
Infrarrojo medio	7.0–1,000 $\mu$ m (1 mm)
Radiación microondas	1 mm–30 cm
Ondas de Radio	$\geq$ 30 cm

**Fuente:** Introduction to Remote Sensing Fifth Edition



**Tabla 5: 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 Fifth Edition

### 1.3.4 Radiancia Espectral

“La radiancia superficial es el total de energía radiada por una superficie, en una determinada dirección por unidad de área y por ángulo sólido de medida”, expresada en la siguiente ecuación.” (Pérez and Muñoz 2006)

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

Donde:

$L$  = Radiancia

$\partial I$  = Intensidad de la radiancia

$\partial A$  = Unidad de área

$\cos \theta$  = coseno de la dirección

Por lo tanto “la radiancia espectral ( $L_T$ ) es la energía radiada en una determinada longitud de onda” (Pérez and Muñoz 2006)

$$L_T = \frac{\partial L}{\partial \lambda}$$

Donde:

$L_T$  = Radiancia espectral

$\partial L$  = Energía radiada

$\partial \lambda$  = Longitud de onda



### 1.3.5 Reflectancia Espectral

La reflectancia espectral se puede definir como la relación entre la cantidad de energía reflejada por un objeto y la cantidad de energía incidente, en una determinada longitud de onda. (Pérez & Muñoz, 2006, p. 81)

$$p_{\lambda} = \frac{ER(\lambda)}{EI(\lambda)}$$

Donde:

$p_{\lambda}$  = Reflectancia espectral

$ER(\lambda)$  = Energía reflejada

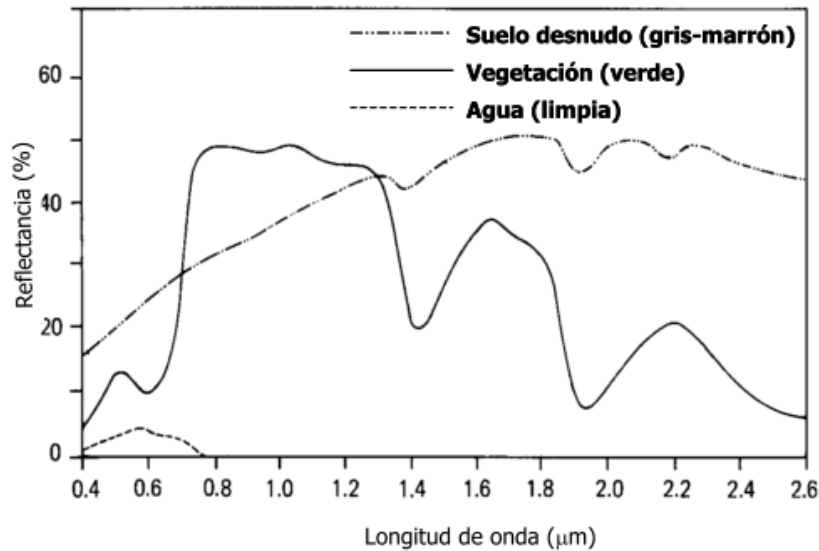
$EI(\lambda)$  = Energía incidente

### 1.3.6 Firma Espectral

Es una curva o trazo característico que identifica a un fenómeno en específico por reflectancia y longitud de onda, este comportamiento se determina en los laboratorios, y del comportamiento de estos se puede extraer información importante. (Pérez & Muñoz, 2006, p. 5)



Espectro electromagnético



**Figura 2:** Espectro electromagnético característico del suelo desnudo, vegetación y agua.  
 μm: micro-metros

**Fuente:** Teledetección Nociones y Aplicaciones

**1.3.7 Teledetección o Detección Remota**

(Campbell and Wynne. 2011) definen a la teledetección como la práctica de derivar información de las superficies terrestres de la tierra mediante imágenes adquiridas desde una perspectiva elevada, debido a la radiación electromagnética en una o más regiones del espectro electromagnético, reflejadas o emitidas desde la superficie de la Tierra.

**1.3.7.1 Imágenes multispectrales de satélite**

Es una imagen capturada mediante sensores fotoeléctricos que transforman los niveles de radiancia procedentes de la superficie terrestre, en valores numéricos que se denominan Niveles o Números Digitales para cada una de las bandas, generando una matriz de tres dimensiones, donde cada número digital pertenece a una fila, una columna, y una banda del espectro electromagnético.

Existen 5 tipos de resoluciones para las imágenes multispectrales, la resolución espacial, se refiere a la cantidad de m<sup>2</sup> que representa cada pixel; la resolución radiométrica, representa los valores de radiancia que puede el sensor representar en valores digitales;





la resolución espectral, es la cantidad y el rango de bandas espectrales obtenidas por el sensor; la resolución temporal, es el tiempo en el cual el satélite paso sobre una zona para tomar la imagen; y la resolución angular, la capacidad que tiene el sensor para tomar imágenes oblicuas. (Pérez & Muñoz, 2006, p. 8)

**Tabla 6:** Resolución espacial de satélites.

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

*m:* Metros

**Fuente:** Teledetección Nociones y Aplicaciones

### 1.3.7.2 Pre procesamiento de imágenes multiespectrales de satélite

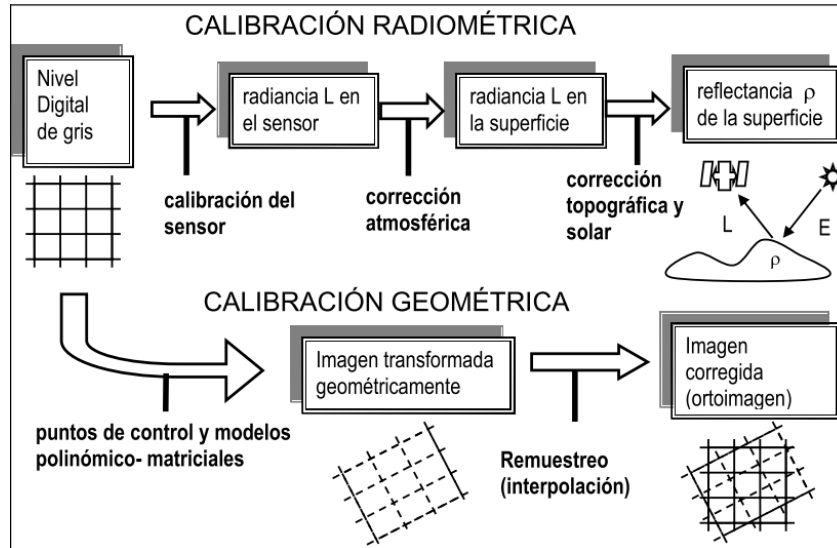
En general para usar la usar la información de las imágenes multiespectrales, si es que es necesario, se lleva a cabo dos tipos de correcciones, una corrección geométrica, para cambiar el sistema de coordenadas y la proyección cartográfica y una radiométrica para cambiar los valores de los niveles digitales de la imagen depurando cualquier defecto.

Además, aparte de las correcciones se debe realizar las calibraciones geométrica y radiométrica. La calibración geométrica corrige las distorsiones espaciales de la imagen y la calibración radiométrica convierte los niveles digitales de la imagen a valores físicos de radiancia y reflectancia. (Pérez and Muñoz 2006)





**Calibración radiométrica y geométrica**



**Figura 3** : *Proceso de calibración radiométrica y geométrica de imágenes.*  
**Fuente:** Teledetección Nociones y Aplicaciones

**1.3.7.3 Indicadores espectrales**

Los indicadores espectrales son ecuaciones, a partir de los datos de reflectancia de las bandas espectrales, que los científicos de han utilizado para predecir, modelar o inferir fenómenos en la superficie terrestre. (Campbell and Wynne. 2011)

**a) Indicadores de Vegetación**

Son indicadores utilizados con el objetivo de obtener información relacionada con la vegetación, y son aproximaciones que utilizan las bandas roja e infrarroja, debido a que en estas bandas contienen se presenta la mayor parte información de la cubierta vegetal.(Pérez and Muñoz 2006)

**Indicador de Vegetación de Diferencia Normalizada (NDVI)**

Rouse et al. (1974)

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

Donde:

$p_{NIR}$  = Reflectancia en la banda infrarrojo cercano

$p_R$  = Reflectancia en la banda roja



**b) Indicadores de salinidad**

Son indicadores aproximados utilizados con el objetivo de obtener información relacionada con la salinidad en suelos.(Pérez and Muñoz 2006)

$$SI = \sqrt{p_R \times p_{NIR}} \quad \text{Dehni and Lounis (2012)}$$

$$SI = \frac{p_{NIR} - p_{SWIR}}{p_{NIR} + p_{SWIR}} \quad \text{Asfaw et al. (2016)}$$

$$SI = \frac{p_G \times p_R}{p_B} \quad \text{Elhag (2016)}$$

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 infrarroja de onda corta

**1.3.8 Técnicas de estimación**

Las técnicas de estimación mediante un conjunto de características numéricas que se usa como datos de entrada, y un resultado numérico, como variable de salida, permite determinar una relación entre las características y el resultado que se desea lograr, mediante la creación de relaciones y asociaciones entre las características y el resultado, con el objetivo de poder predecir resultados con un conjunto de características diferentes. (Sarkar, Bali, and Sharma 2018)

**a) Partial Least Square Regression (PLSR)**

PLSR, es un método de regresión lineal que acepta un número muy grande de variables, realiza la predicción a partir un número pequeño de componentes, que pueden ser



optimizados para mejorar el rendimiento, transforma las variables de entrada en una matriz en variables latentes, maximizando el cuadrado de la covarianza. (Ng 2013)

Partiendo de la siguiente ecuación:

$$Y = X\beta + e$$

Donde:

y : Variable a predecir ( vector de salida)

x : Variables de entrada ( matriz entrada)

e : coeficiente de error

Se estima que  $\beta$  es :

$$\beta = | x_1 - \bar{x}_1 \dots x_m - \bar{x}_m |$$

Donde:

$x_1$ : Vector X en la posición 1

$\bar{x}_1$ : Mediana Vector X en la posición 1

La idea detrás de PLSR es descomponer la matriz de entrada X y el vector de salida Y, siguiendo las siguientes ecuaciones (Ng 2013):

$$X = TP^T$$

$$Y = UQ^T$$

Donde **T** y **U**, son matrices de coeficientes o pesos, **P** y **Q** son la traspuesta de **X** y **Y** respectivamente ,para luego realizar una regresion entre de **T** y **U** (Ng 2013)

$u = x_j$  para una posicion j

Procesar

$$p = \frac{Y^T u}{||Y^T u||}$$

Hasta u deje de cambiar.



$t = x_j$  para una posición  $j$

Procesar

$$p = \frac{X^T t}{\|X^T t\|}$$

Hasta  $t$  deje de cambiar.

Donde:

$\|X^T t\|$ : es la norma de del producto  $X^T t$

Con la finalidad de descomponer  $X$  y  $Y$  entre si para tener informacion de ambas variables en cuenta, esto intercambiando  $t$  y  $u$  como resultado lo siguiente. (Ng 2013)

$u = y_j$  para una posición  $j$

Procesar

$$p = \frac{X^T u}{\|X^T u\|}$$

$$t = Xp$$

$$q = \frac{Y^T t}{\|Y^T t\|}$$

$$u = Yt$$

Hasta que  $t$  deje de cambiar.

Pero al ser  $Y$  un vector ,  $q = 1$  (Ng 2013):

$u = y_j$  para una posición  $j$

Procesar

$$p = \frac{X^T u}{\|X^T u\|}$$

$$t = Xp$$

$$q = 1$$

$$u = X$$

$$X = X - tp^T$$

$$Y = Y - uq^T$$

Hasta que  $t$  deje de cambiar.



**b) Support Vector Machine (SVM)**

SVM, es un metodo clasificador de aprendizaje supervisado , en el cual se selecciona un hiperplano de separación que equidiste dos ejemplos mas cercanos de cada clase, para conseguir un márgen máximo.(Guareño 2016)

**SVM Lineal Separable**

Dado un conjunto  $C$ ,  $x_i \in R^d$  e  $y_i \in \{+1, -1\}$ , se define el hiperplano ( $D(x)$ ) de separación:

$$C = \{ (x_1, y_1), (x_2, y_1), \dots, (x_n, y_n) \}$$

$$D(x) = (w_1x_1 + \dots + w_dx_d) + b = \langle w, x \rangle + b$$

Dicho hiperplano debera cumplir con las siguientes desigualdades

$$\langle w, x_i \rangle + b \geq 0 \text{ si } y_i = +1$$

$$\langle w, x_i \rangle + b \leq 0 \text{ si } y_i = -1$$

$$y_i(\langle w, x_i \rangle + b) \geq 0 \text{ ó } y_i(D(x)) \geq 0$$

La selección del hiperplano se realiza por el menor margen ( $r$ ) existente, y la distancia minima ( $t$ ) entre el hiperplano y el ejemplo mas cercano de cada clase. (Guareño 2016)

**Múltiples hiperplanos en SVM**

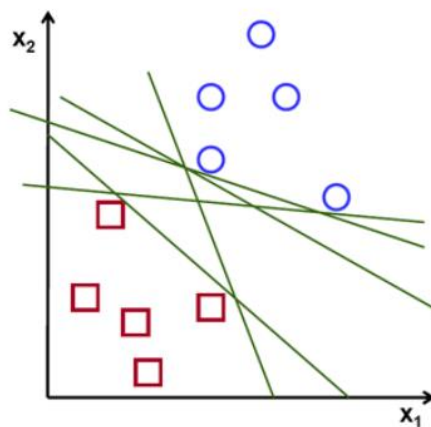
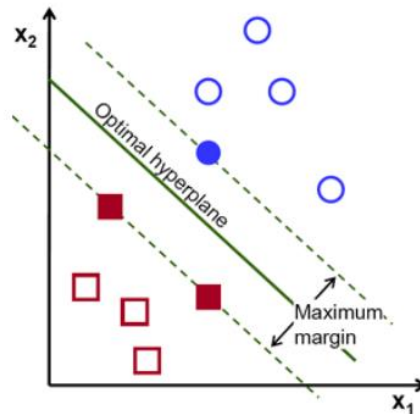


Figura 4 : Múltiples hiperplanos encontrados para una solución en SVM

**Fuente:** Support Vector Regression: propiedades y aplicaciones



**Hiperplano de menor margen**



**Figura 5 :** *Hiperplano de menor margen*

**Fuente:** Support Vector Regression: propiedades y aplicaciones

**Support Vector Regresion (SVR)**

SVM también se puede utilizar como método de estimación, se usa en minería de datos para modelar diferentes relaciones de entrada y salida de un sistema o proceso dado, con un margen de tolerancia (épsilon), el objetivo de este algoritmo es minimizar los errores de predicción y la complejidad de la función. (Guareño 2016)

Para una regresion

$$y = wx + b$$

Donde :

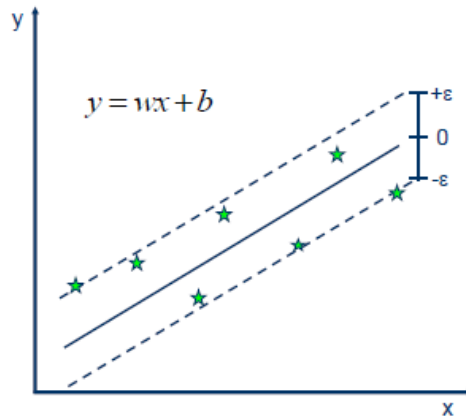
w : vectores de coeficientes

b: coeficiente independiente

x: variables independientes







**Figura 6 :** Estimación de SVR, con un margen de tolerancia epsilon

**Fuente:** Support Vector Regression: propiedades y aplicaciones

### 1.3.8.1 Evaluación de técnicas

La evaluación de las técnicas de estimación al ser de aprendizaje supervisado, es relativamente más fácil que los no supervisados, debido a la disponibilidad de las variables numéricas. (Sarkar et al. 2018).

#### a) Coeficiente de Determinación o $R^2$

El coeficiente de determinación mide el porcentaje de relación entre los valores reales y los estimados, un valor de  $R^2$  de 1 denota un modelo perfecto, ya que indica que toda la varianza se explica por las variables independientes, y cuan probable es que prediga las muestras futuras. (Sarkar et al. 2018).

La fórmula matemática estándar para calcular el  $R^2$  es la siguiente:

$$R^2(y, \hat{y}) = 1 - \left( \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y}_i)^2} \right)$$

Donde:

$n$ : Es la cantidad de muestras

$\bar{Y}_i$ : Es la media de la variable dependiente

$Y_i$ : Es el valor real esperado

$\hat{Y}_i$ : Es la salida estimada del modelo



**b) Error cuadrático medio o RMSE**

El error cuadrático medio (RMSE), calcula la desviación estándar muestral de las diferencias entre los valores estimados y los valores esperados del modelo, en la unidad de medida de los valores estimados, muestra la diferencia entre los datos medidos y simulados, un menor valor de RMSE indica una mayor precisión en la estimación de valores. (Sarkar et al. 2018)

La fórmula matemática estándar para calcular el RMSE es la siguiente

$$RMSE(y, \hat{y}) = \sqrt{\frac{1}{n} * \left( \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \right)}$$

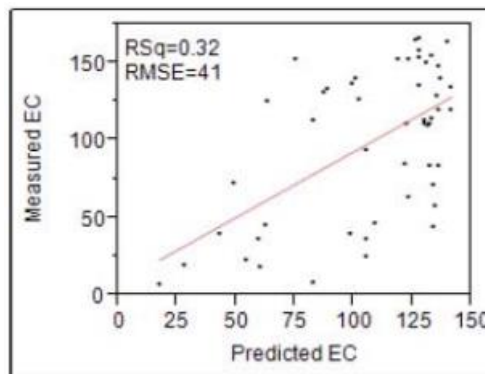
Donde:

$n$ : Es la cantidad de muestras

$Y_i$ : Es el valor real esperado

$\hat{Y}_i$ : Es la salida estimada del modelo

**Grafica de evaluación de  $R^2$  y RMSE**



**Figura 7:** Gráfica de evaluación de  $R^2$  con 0.32 y RMSE de 41

**Fuente:** Mapping and Modelling Spatial Variation in Soil Salinity in the Al Hassa Oasis Based on Remote Sensing Indicators and Regression Techniques



## 1.4 Formulación del problema

¿Qué técnica tendrá mayor precisión en la estimación de la salinidad en suelos de escasa vegetación?

## 1.5 Justificación e importancia del estudio.

La identificación de la salinidad mediante los métodos tradicionales es costosa, si se piensa usar en grandes extensiones de terrenos, por ello es necesario buscar otras alternativas con menor costo.

En el aspecto económico, la identificación de grado de salinidad usando técnicas de teledetección y pruebas tendrá menor costo que sólo usando los métodos tradicionales, y puede beneficiar a los agricultores, ya que podrían aprovechar los terrenos de escasa vegetación mediante la cosecha selectiva de cultivos de acuerdo a la tolerancia salina de estos y a la salinidad de los terrenos.

La investigación tendrá como objetivo la estimación del grado de salinidad en suelos de escasa vegetación en grandes extensiones de terreno, aplicando las técnicas de teledetección ya existentes en las imágenes multiespectrales obtenidas del satélite y al comparar las técnicas de estimación, se logrará encontrar el modelo más óptimo de estimación de conductividad eléctrica, y mediante la clasificación de este valor de estimación, el grado de salinidad.

Además, el proceso de procesamiento de imágenes multiespectrales y la comparación de técnicas de estimación, servirá como base a la línea de investigación de la escuela para futuros trabajos relacionados.

## 1.6 Hipótesis

Utilizando de teledetección en el procesamiento de imágenes multiespectrales, y la aplicación de técnicas de estimación de Partial Least Squares Regression (PLSR) y Support Vector Machine (SVM), se podrá estimar el grado de salinidad en suelos de escasa vegetación.

La estimación del grado de salinidad con PLSR, obtendrá mejores resultados en los indicadores usados ( $R^2$  y RMSE), que la técnica de SVM, por lo tanto, al evaluar ambas

técnicas en dichos indicadores, la precisión al utilizar PLSR es mayor a la precisión de SVM.

## **1.7 Objetivos**

### **1.7.1 Objetivo General**

Comparar técnicas de estimación del grado de salinidad en imágenes multiespectrales en suelos de escasa vegetación.

### **1.7.2 Objetivos Específicos**

1. Seleccionar las técnicas de estimación a utilizar.
2. Adquisición de imágenes multiespectrales del área de estudio
3. Realizar calibraciones atmosféricas a las imágenes multiespectrales adquiridas.
4. Seleccionar indicadores de salinidad y vegetación.
5. Recolectar muestras de suelos de escasa vegetación y realizar la prueba de conductividad eléctrica.
6. Evaluar las técnicas de estimación.

## **II. MATERIAL Y MÉTODOS**

### **2.1 Tipo y diseño de la investigación**

#### **2.1.1 Tipo de investigación**

De acuerdo a Domínguez, 2007 , la Investigación Cuantitativa estudia la relación entre las variables de la investigación que han sido cuantificada para luego procesarlas y analizarlas estadísticamente; tomando como referencia la cita anterior, el tipo de investigación es cuantitativo porque los indicadores están representados en cantidades numéricas.

#### **2.1.2 Diseño de la investigación**

De acuerdo Agudelo, Aigner, & Ruiz, 2008 la investigación cuasi experimental estudia las relaciones de causa-efecto entre la variable independiente y dependiente, por lo cual la manipulación de la variable independiente afecta a la dependiente; tomando como referencia la cita anterior, el diseño de investigación es cuasi



experimental porque las técnicas de estimación de grado de salinidad afectan en la estimación del grado de salinidad.

## **2.2 Población y muestra**

La población son las técnicas de estimación mediante información proveniente de imágenes multiespectrales (**Anexo 1**) y la muestra escogida son las técnicas de estimación de SVM y PLSR.

## **2.3 Variables y Operacionalización**

### **2.3.1 Variables**

#### **Variable Independiente**

Técnicas de estimación

#### **Variable Dependiente**

Grado de salinidad en suelos de escasa vegetación



### 2.3.2 Operacionalización de Variables

Variable Independiente	Dimensiones	Indicadores	Unidad de Medida	Formula	Descripción
<b>Técnicas de estimación</b>	Precisión de la técnica de estimación	Coficiente de determinación	Porcentaje	$R^2 = \left( \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2 + \sum_{i=1}^n (Y_i - \bar{Y})^2} \right)$	$X_i$ = 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
		Raíz del Error cuadrático medio	Porcentaje	$RMSE = \sqrt{\frac{\sum_{i=1}^n (X_i - Y_i)^2}{n}}$	$X_i$ = valor medido $Y_i$ = valor estimado $n$ = el número de muestras

Variable Dependiente	Dimensiones	Indicadores	Unidad de Medida	Formula	Descripción
<b>Grado de salinidad en suelos de escasa vegetación</b>	Grado de Salinidad	Conductividad Eléctrica (CE)	ds/m (deciSiemens / metro)	$CE_e = CE_w \times 1.5$	$CE_e$ = Conductividad eléctrica de extracto saturado de suelo $CE_w$ = Conductividad eléctrica del agua





## **2.4 Técnicas e instrumentos de recolección de datos.**

### **2.4.1 Abordaje Metodológico**

El abordaje metodológico fue cuasi experimental, de acuerdo a Agudelo, Aigner, & Ruiz, 2008, este tipo de diseño estudia las relaciones de causa y efecto entre la variable dependiente e independiente, por lo cual la manipulación de la variable independiente afecta a la dependiente.

### **2.4.2 Técnicas de Recolección de datos**

Se eligió aplicar la técnica la recopilación documental, debido a que se revisó y se usó como guía libros, artículos de investigación y papers.

Se eligió aplicar la técnica la observación científica, debido a que se utilizaran imágenes multiespectrales para realizar el proyecto de investigación.

### **2.4.3 Procedimiento de recolección de datos**

Una vez identificados los métodos de estimación a utilizar se realizó lo siguiente:

- a) En primera fase se seleccionó las técnicas de estimación a utilizar
- b) En segunda fase se debe adquirir las imágenes multiespectrales del área de estudio.
- c) En tercera fase se realizó las calibraciones atmosféricas para obtener la reflectancia por pixel de la imagen.
- d) En cuarta fase se seleccionó los índices espectrales de salinidad y vegetación.
- e) En quinta fase se tomó muestra de suelos de escasa vegetación y se realizara la prueba de conductividad eléctrica.
- f) En sexta fase se estimó la salinidad con PARTIAL LEAST SQUARE REGRESSION (PLSR) y SUPPORT VECTOR MACHINE (SVM).

## **2.5 Procedimientos de análisis de datos**

Para analizar e interpretar los resultados obtenidos de los indicadores en la operacionalización de variables, se realizará un procesamiento estadístico en lenguaje Python utilizando la librería scikit-learn.



## 2.6 Criterios éticos

En la investigación se utilizó el principio ético de honestidad e integridad, debido a que no se manipuló la información ni los resultados obtenidos, y el principio ético de propiedad intelectual, debido a que se citó y referencio los libros, trabajos de investigación y papers, que se utilizaron como base para desarrollar el presente trabajo.

## 2.7 Principios de rigor científico

**Fiabilidad:** La investigación cumplió con las expectativas fijadas en su contenido y su implementación.

**Validez:** Los datos obtenidos en la investigación fueron analizados y evaluados por personal especializado en el tema para decretar su autenticidad.

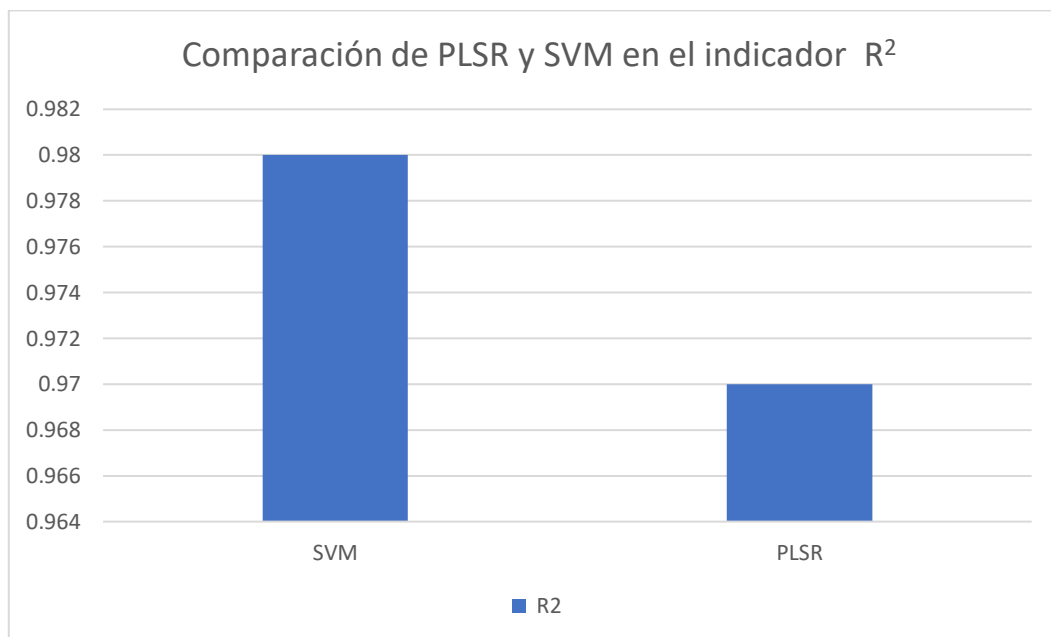


### III. RESULTADOS

#### 3.1 Resultados en Tablas y Figuras

La hipótesis inicial para este proyecto de investigación, fue que PLSR obtendría mejores resultados, con base a la investigación de la cual se extrajo la población (**Anexo 1**), pero los resultados fueron diferentes; SVM fue la técnica que 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 SVM obtuvo 0.98 y la técnica PLSR 0.97, lo que significó que la técnica SVM tuvo un 98 % de precisión y la técnica PLSR 97.

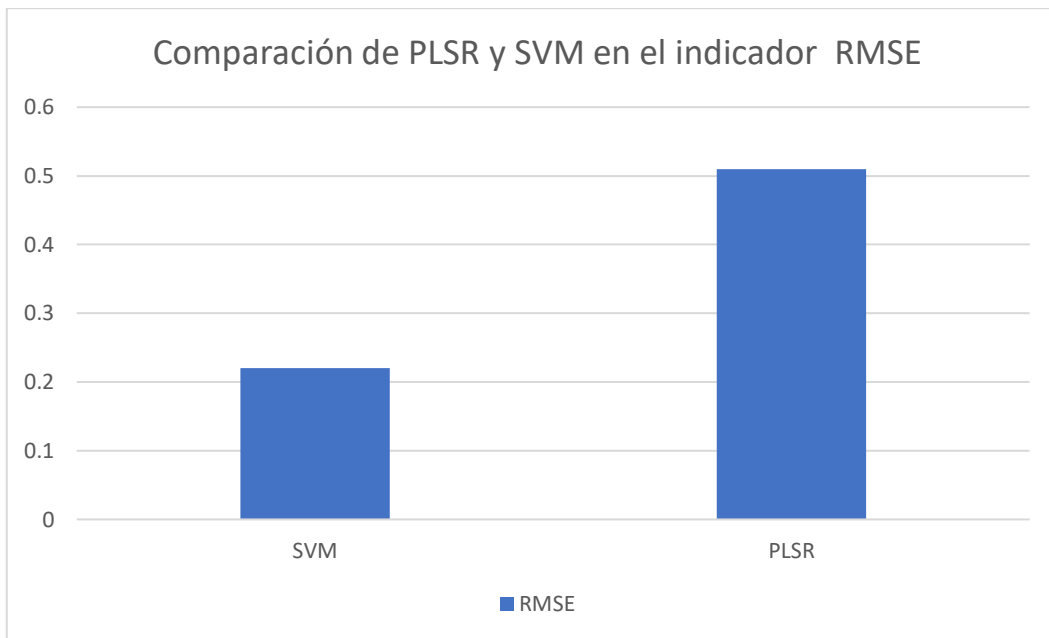


**Figura 8:** Comparación de PLSR y SVM en el indicador de precisión  $R^2$

**Fuente:** Elaboración propia

En el indicador de operacionalización RMSE la técnica SVM obtuvo un error de 0.22 dS/m y la técnica PLSR 0.51 dS/m, ambas en error en medida de Conductividad Eléctrica (EC).

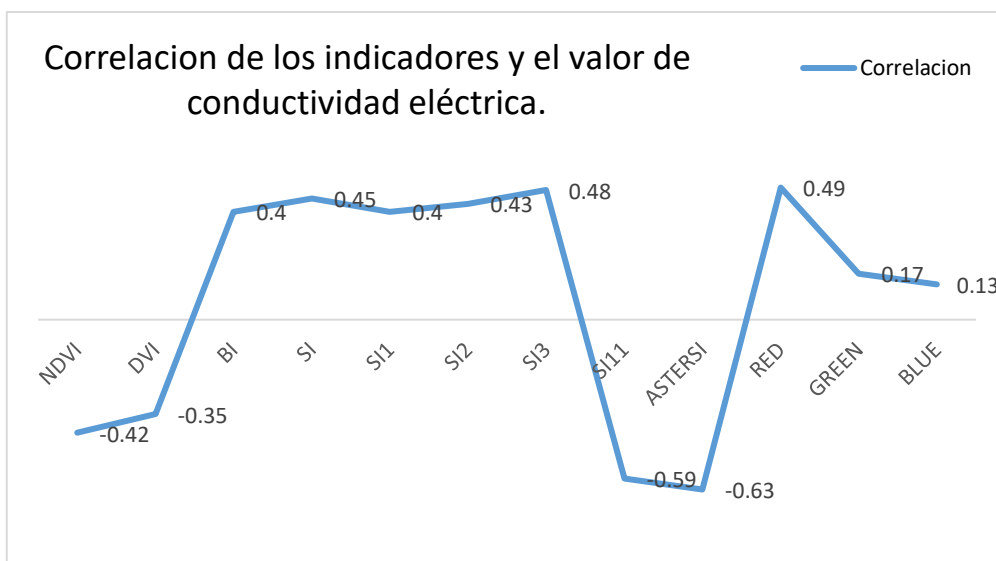




**Figura 9:** Comparación de PLSR y SVM en el indicador de precisión RMSE

**Fuente:** Elaboración propia

Aparte de la evaluación de ambas técnicas, también se aplicó el coeficiente de correlación de Pearson, para analizar la relación entre los indicadores y los valores de conductividad, se expresa la siguiente tabla:



**Figura 10:** Gráfica de porcentaje de correlación entre los indicadores usados y el valor de conductividad eléctrica.

**Fuente:** Elaboración Propia



### 3.2 Discusión de resultados

Al analizar los resultados de indicadores obtenidos de precisión  $R^2$  (*Figura 8*) RMSE (*Figura 9*), la técnica SVM tiene mayor porcentaje de precisión y tiene menor error promedio en la estimación de la salinidad.

La *Figura 10* demuestra que las variables SI, SI3, RED poseen una relación directa (sus valores son directamente proporcional al valor de conductividad eléctrica) con 0.45, 0.48, 0.49 de correlación y los indicadores NDVI, SI11 y ASTERSI poseen una relación inversa (sus valores son inversamente proporcional al valor de conductividad eléctrica) con 0.42, 0.59, 0.63 de correlación.

Se esperaba que los indicadores GREEN y BLUE tuvieran una mayor correlación, de acuerdo a los trabajos de investigación anteriores, pero el resultado fue de 0.17 y 0.13 respectivamente.

Se generó una imagen de colores a partir del indicador de NDVI, que es un índice de vegetación (*Figura 11*) y se correspondió las coordenadas con las de la plataforma de Google Maps (*Figura 12*).



**Figura 11:** *Indicador NDVI a color*

**Fuente:** Elaboración propia

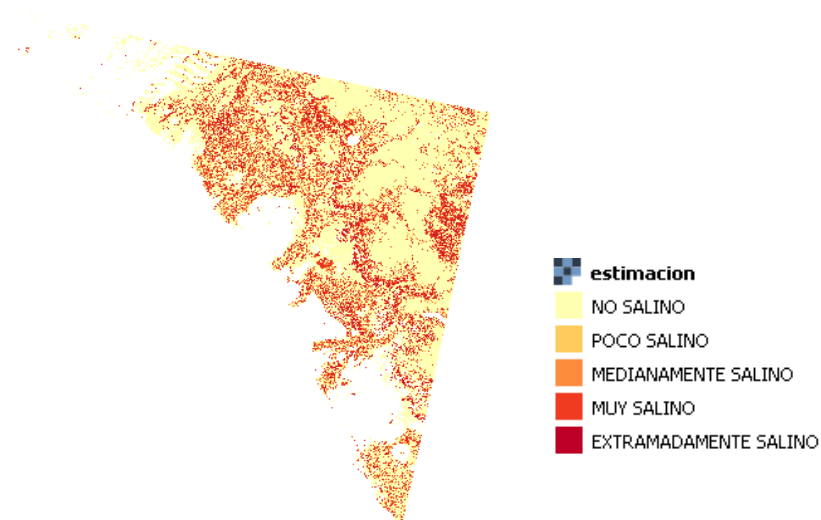




**Figura 12:** Indicador NDVI, en la plataforma de Google Maps.

**Fuente:** Elaboración propia

Además, se generó una imagen de colores a partir de la estimación de salinidad utilizando SVM (*Figura 13*), y se correspondió las coordenadas con las de la plataforma de Google Maps (*Figura 14*).

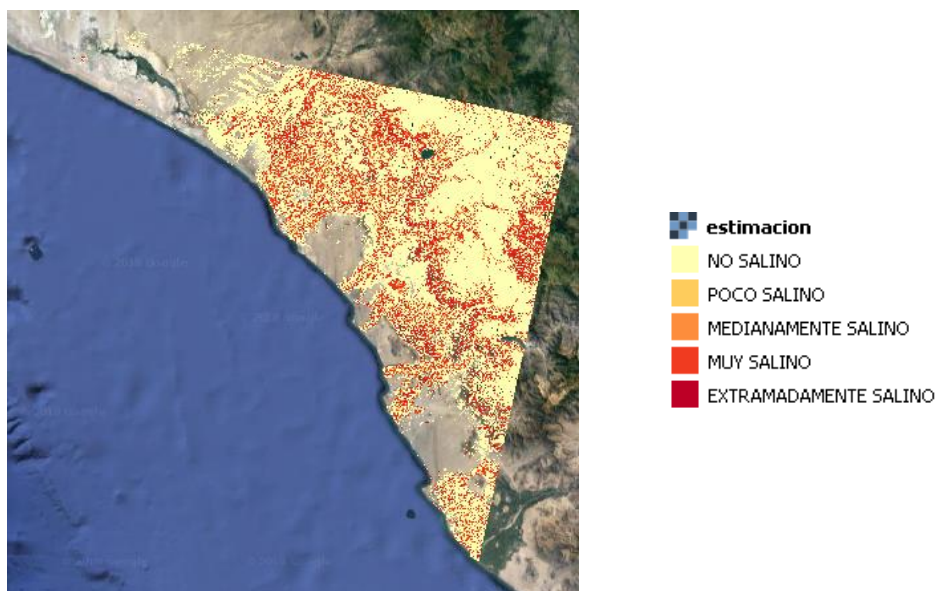


**Figura 13:** Estimación del grado de salinidad usando SVM

**Fuente:** Elaboración propia







**Figura 14:** *Estimación del grado de salinidad usando SVM, en la plataforma de Google Maps.*

**Fuente:** Elaboración propia



### 3.3 Aporte Práctico

#### 3.3.1 Selección de las técnicas de estimación a utilizar

La población de técnicas consideradas para esta investigación es la siguiente:

**Tabla 7:** Población de técnicas de estimación

TÉCNICAS DE ESTIMACIÓN	PRECISIÓN	
	R <sup>2</sup>	Error relativo
1. PARTIAL LEAST SQUARE REGRESSION (PLSR)	89.00%	
2. MULTIPLE LINEAR REGRESSION (MLR)	87.00%	
	52.00%	
	36.00%	
	96.30%	
	65.00%	
3. RANDOM FOREST REGRESSION (RF)	79.00%	
	98.30%	
	85.00%	
4. SUPPORT VECTOR MACHINE (SVM)	83.00%	
	61.70%	
		37.00%
5. BOOSTED BINARY REGRESSION (BBRT)	82.00%	
6. ARTIFICIAL NEURONAL NETWORK (ANN)	81.00%	
	93.00%	
	69.00%	
	99.70%	
	41.90%	
		49.00%
7. DECISION TREE REGRESSION (DT)	80.00%	
8. PRINCIPAL COMPONENT REGRESSION (PCR)	79.00%	

**Fuente:** Teledetección Nociones y Aplicaciones

Las fuentes con las que se elaboró la tabla anterior se encuentran en el **Anexo 1**, se eligió 2 técnicas de estimación siguiendo los siguientes criterios:

Según (Yue, Feng, Yang, & Li, 2018) la técnica de PLSR es mejor que la de SVM, la precisión de PLSR es de 89% y la de SVM es de 83% para estimar el vigor de las plantas, usando imágenes multiespectrales satelitales.



Según (Wang et al. 2016), la técnica SVM tuvo mayor precisión que la de ANN o redes neuronales en la estimación de la biomasa en trigos, usando imágenes multiespectrales satelitales, con un 61.70% y 41.90% respectivamente, pero SVM tiene menor precisión que RF con 79%.

Y Según (Guan et al. 2013), la técnica SVM tuvo mayor precisión que la de ANN en la estimación de salinidad en un distrito irrigado, mediante propiedades físicas del ambiente.

La técnica de ANN, tiene una muy precisión en la estimación de la salinidad, según (Mousavi et al. 2017) , ANN tuvo una precisión de 93% en comparación con el 52% de la técnica MLR o regresión lineal múltiple, y según (Shahabi et al. 2017) ANN tuvo una precisión de 69% en comparación con el 36% de la técnica MLR.

Con base a los trabajos anteriores se propuso comparar la técnica *SVM*, que obtuvo mejores resultados que la técnica de ANN, usada en los trabajos de investigación de estimación de salinidad usando imágenes multiespectrales, además posee un estudio previo en salinidad, pero mediante propiedades físicas, y la técnica *PLSR*, que fue la que obtuvo mejor resultado en la estimación del vigor de las plantas.

### **3.3.2 Adquisición de imágenes multiespectrales del área de estudio**

El objetivo principal de este trabajo de investigación es la comparación de técnicas de estimación para estimar el grado de salinidad en grandes extensiones de terrenos de escasa vegetación, para ello se necesitó los datos provenientes de las imágenes multiespectrales del de área de estudio, y se descargó dichas imágenes del proyecto espacial Landsat detallado a continuación.

El proyecto geoespacial Landsat lleva más de 40 años brindando datos calibrados de la superficie de la Tierra, en forma de imágenes multiespectrales, el programa Landsat 8 fue lanzado en 2013. Landsat 8 ofrece en su repositorio múltiples capturas de territorios a nivel mundial y datos gratuitos, siendo ofrecidos al público sin cargo alguno a través del United States Geological Survey (USGS) y Earth Resources Observation and Science (EROS). (Gonzaga Aguilar 2014)

El satélite Landsat 8 orbita la Tierra a una altitud de 705 kilómetros, en una franja de 185 kilómetros, moviéndose de norte a sur sobre el lado iluminado de la Tierra en una órbita



sincrónica solar, siguiendo el Worldwide Reference System (WRS-2). (Department of the Interior U.S. 2018)

Landsat 8 forma una órbita completa cada 99 minutos, unas 14 órbitas cada día y cruza cada punto de la Tierra una vez cada 16 días, la órbita del satélite permite cobertura repetida de 8 días de cualquier área en el globo terrestre. Además, cuenta con 2 sensores, el Operational Land Imager (OLI) y el Thermal Infrared Sensor (TIRS), los datos de ambos sensores son corregidos radiométricamente y co-registrados en una proyección cartográfica, con correcciones para el desplazamiento del terreno que resulta en una orto rectificación estándar de imagen digital. (Department of the Interior U.S. 2018)

Debido a los beneficios mencionados anteriormente, principalmente el de su fácil accesibilidad y descarga gratuita, se utilizó Landsat 8 como proveedor de imágenes multiespectrales.

La adquisición de imágenes del satélite Landsat 8, se obtuvo a través del sitio web de USGS de forma gratuita, se siguió los pasos del *anexo 2*, se filtró los paquetes de imágenes por cobertura de nubes, debido a que estos poseían mucha nubosidad, dificultando el trabajo de investigación.



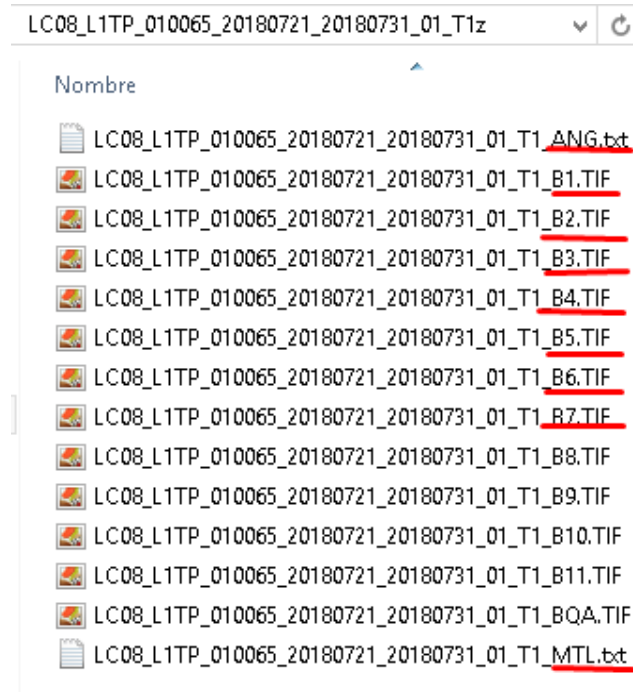
**Figura 15:** Paquetes de imágenes del repositorio de Landsat 8

**Fuente:** Elaboración propia





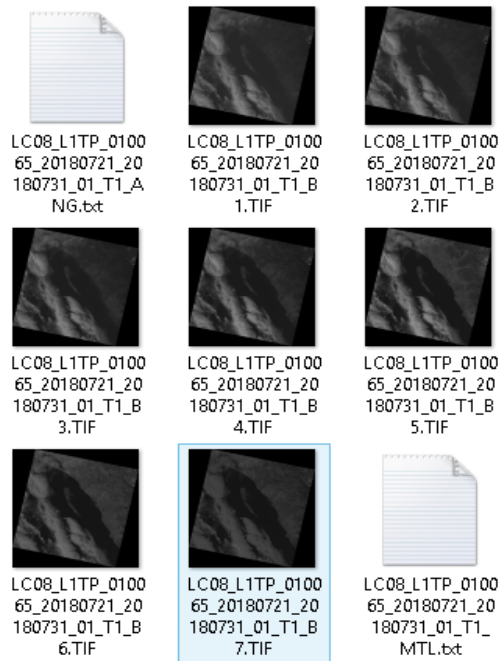
Por tal motivo se adquirió el paquete de imágenes del 21 julio del 2018, por ser la que presenta menor nubosidad, como resultado de la descarga del paquete de imágenes se obtuvo una carpeta con la siguiente estructura.



**Figura 16:** Estructura de la carpeta de las imágenes adquiridas de Landsat 8

**Fuente:** Elaboración propia





**Figura 17:** Imágenes multiespectrales adquiridas del satélite Landsat 8

**Fuente:** Elaboración propia

Donde las imágenes multiespectrales se encuentran en formato .TIF y las imágenes provenientes del sensor OLI a utilizar se especifica en el siguiente cuadro, cada pixel de la imagen tiene una resolución de 30m x 30m o 900m<sup>2</sup>, la imagen está compuesta de 7761 x 7611 pixeles, los pixeles se encuentran georreferenciados en el formato UTM 17N.





**Tabla 8:** *Bandas 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 biomasas
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:** Caracteres que identifican a una banda del espectro electromagnético en el conjunto de imágenes.

**Límites del espectro electromagnético:** Valores entre los que esta comprendidos una banda electromagnética.

**Fuente:** Landsat FAQs,USGS

En la tabla 8 se explica los archivos en formato de texto que también están incluidos en cada paquete de imágenes adquiridas.

**Tabla 9:** *Archivos adicionales incluidos 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:**

Las imágenes seleccionadas de la banda 1 a la banda 7 tienen una resolución espacial de 30 metros (m) x 30 m, es decir que cada pixel equivale a 900 m<sup>2</sup> de terreno, la imagen se compone de pixeles de 16 bits, cada pixel está entre el rango de 0 a 65536, y un ocupa un espacio de almacenamiento alrededor de 112MB(Megabytes). El tamaño aproximado de la escena (imagen capturada de Landsat) es de 170 km de norte a sur por 183 km de este a oeste, cada imagen está compuesta por 7761 y 7611. (Department of the Interior U.S. 2018)

**3.3.3 Realización de calibraciones atmosféricas a las imágenes multiespectrales adquiridas.**

Luego de haber adquirido las imágenes multiespectrales se realizó las calibraciones atmosféricas imagen, para obtener datos significativos que se utilizaran en la comparación de técnicas, las imágenes multiespectrales son entregadas en números digitales, con esta información no es posible realizar la comparación de técnicas de estimación, para obtener información relevante se necesita obtener la reflectancia por pixel de cada imagen.

Los productos de datos Landsat 8 constan de números o valores digitales escalados cuantificados y calibrados (ND) que representan los datos de imágenes multiespectrales, estos datos adquiridos por el sensor Operational Land Imager (OLI), y se entregan en un formato entero sin signo de 16 bits (Department of the Interior U.S. 2018).

Luego de la adquisición de las imágenes multiespectrales, se realizó la calibración atmosférica de las imágenes adquiridas, se cambió los números digitales a reflectancia, y luego de reflectancia a reflectancia superior atmosférica (TOA) utilizando variables proporcionadas en el archivo de metadatos (MTL), que se obtuvo junto a las imágenes.



Según (Department of the Interior U.S. 2018), en la documentación de Landsat 8, se expresan las siguientes ecuaciones de conversión de valor digital a radiancia y reflectancia para realizar la calibración atmosférica, que son las siguientes :

**Conversión de pixeles números digitales a radiancia.**

$$L_{\lambda} = M_L Q_{cal} + A_L$$

Donde:

$L_{\lambda}$  = Radiancia espectral al nivel de la superficie

$M_L$  = Factor de re escalado multiplicativo presente en la metadata **RADIANCE\_MULT\_BAND\_x**

$A_L$  = Factor de re escalado aditivo presente en la metadata **RADIANCE\_ADD\_BAND\_x**

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

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

$$p_{\lambda'} = M_p Q_{cal} + A_p$$

Donde:

$p_{\lambda'}$  = Radiancia espectral al nivel de la superficie

$M_p$  = Factor de re escalado multiplicativo presente en la metadata **REFLECTANCE\_MULT\_BAND\_x**

$A_p$  = 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 el ángulo solar**

$$p_{\lambda} = \frac{p_{\lambda'}}{\cos(\theta_{sz})} = \frac{p_{\lambda'}}{\sin(\theta_{SE})}$$



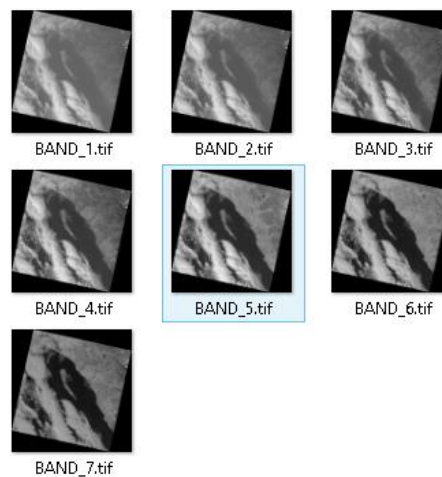
$\rho_{\lambda'}$  = Reflectancia con calibración atmosférica

$\theta_{SE}$  = Angulo solar de elevación

$\theta_{sz}$  = Angulo solar zenith

Para esta investigación se aplicó la conversión de valores digitales a reflectancia con corrección de ángulo solar, para ello se transformó cada pixel de cada imagen, como se muestra en el caso de ejemplo del *Anexo 3*, además se programó un pequeño algoritmo que realice dicho proceso automáticamente *Anexo 4*.

Como resultado de este proceso se obtuvo 7 imágenes nuevas, que contienen los valores de reflectancia de cada imagen espectral original.



**Figura 18:** *Imágenes de reflectancia generadas al procesar las imágenes*

**Fuente:** Elaboración Propia



Las imágenes procesadas se especifican en el siguiente cuadro:

**Tabla 10:** *Imágenes espectrales de reflectancia, banda y su identificador.*

IDENTIFICADOR	IMAGEN	BANDA
$p_B$	BAND_2	Banda Azul
$p_G$	BAND_3	Banda Verde
$p_R$	BAND_4	Banda Roja
$p_{NIR}$	BAND_5	Banda infrarroja cercano
$p_{SWIR}$	BAND_6	Banda infrarroja de onda corta

**Fuente:** Elaboración propia.

### 3.3.4 Selección indicadores de salinidad y vegetación.

Luego de convertir cada pixel de la imagen de la imagen a reflectancia superior atmosférica (TOA) se realizó la selección de indicadores de salinidad, vegetación y bandas, que están relacionados con los valores de salinidad, se recopiló indicadores de los siguientes trabajos de investigación:

En el trabajo de investigación de “**Digital Mapping of Topsoil Salinity Using Remote Sensing Indices in Agh-Ghala Plain, Iran**” (Mousavi et al. 2017), se comparó 17 indicadores de salinidad, vegetación, y bandas espectrales con la finalidad de hallar la correlación entre estos indicadores y el valor de conductividad eléctrica.



**Tabla 11:** *Correlación entre los indicadores de salinidad, vegetación y bandas, y el valor de conductividad eléctrica.*

INDICADOR	Coefficiente de correlación (R <sup>2</sup> )
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.

Por lo tanto con base a este trabajo de investigación se usara los indicadores *ASTER SI*, *SI1*, *SI11*, *SI3*, *B2*, *B3*, *B4* por tener una correlación mayor a 0.20.

En el trabajo de investigación “**Monitoring and evaluating spatial variability of soil salinity in dry and wet seasons in the Werigan–Kuqa Oasis, China, using remote sensing and electromagnetic induction instruments Jianli**” (Ding and Yu 2014), se comparó 5 indicadores de salinidad y vegetación, para hallar la correlación entre dichos indicadores y el valor de conductividad eléctrica.





**Tabla 12:** Correlación entre los indicadores de salinidad, vegetación, y el valor de conductividad eléctrica.

INDICADOR	Coefficiente de correlación
SI	0.310
NDSI	0.282
BI	0.338
NDVI	0.343
DVI	0.318

**Fuente:** Monitoring and evaluating spatial variability of soil salinity in dry and wet seasons in the Werigan–Kuqa Oasis, China, using remote sensing and electromagnetic induction instruments Jianli.

Por lo tanto, con base a este trabajo de investigación se usará los indicadores *NDVI*, *DVI*, *BI*, *SI* por tener una correlación mayor a 0.30 en relación a la conductividad eléctrica.

En el trabajo de investigación “**Mapping and Modelling Spatial Variation in Soil Salinity in the Al Hassa Oasis Based on Remote Sensing Indicators and Regression Techniques**” (Allbed, Kumar, and Sinha 2014), se utilize un indicador de salinidad (SI)

De los puntos anteriores, los indicadores seleccionados poseen las siguientes formulas:

$$NDVI = \frac{p_{NIR} - p_R}{p_{NIR} + p_R} \quad \text{Rouse et al. (1974)}$$

$$DVI = p_{NIR} - p_R \quad \text{Richardson y Wiegand (1977)}$$

$$SI_2 = \frac{p_R}{p_{NIR}} \quad \text{Tripathi et al. (1997)}$$

$$BI = \sqrt{p_G^2 + p_R^2} \quad \text{Inakwu and Odeh (2008)}$$

$$SI = \sqrt{p_B * p_R^2} \quad \text{Khan et al. (2005)}$$

$$SI_{11} = \frac{p_{NIR}}{p_{SWIR}} \quad \text{Khan et al. (2005)}$$



$$SI\ 1 = \sqrt{p_G * p_R} \quad \text{Dehni and Lounis (2012)}$$

$$ASTER\ SI = \frac{p_{NIR} - p_{SWIR}}{p_{NIR} + p_{SWIR}} \quad \text{Asfaw et al. (2016)}$$

$$SI\ 3 = \frac{p_G \times p_R}{p_B} \quad \text{Elhag (2016)}$$

Donde:

$p_B$  = Reflectancia en la banda azul

$p_G$  = Reflectancia en la banda verde

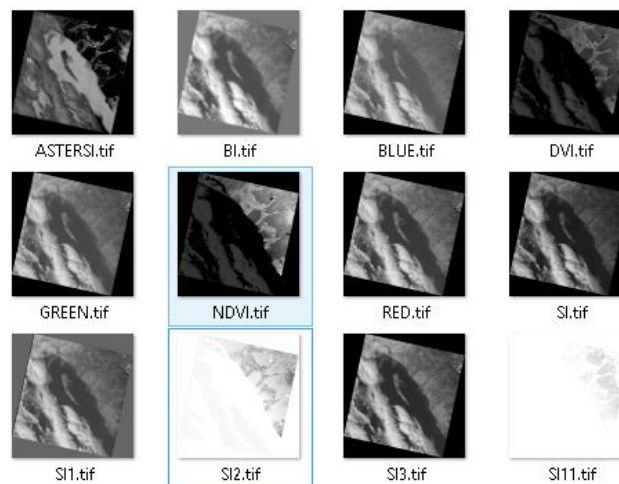
$p_R$  = Reflectancia en la banda roja

$p_{NIR}$  = Reflectancia en la banda infrarrojo cercano

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

Para calcular cada indicador, se tuvo que procesar las imágenes mediante operaciones de los píxeles de cada imagen, de acuerdo a su fórmula matemática como se muestra en el caso de ejemplo del *Anexo 5*, además se programó un pequeño algoritmo que realice este proceso automáticamente *Anexo 6*.

A partir de las fórmulas de los indicadores antes mencionados, se generaron las siguientes imágenes *Figura 19*, que fueron agrupadas en una carpeta, y a cada imagen se le asigno el nombre de su indicador correspondiente.



**Figura 19:** Imágenes procesadas con los indicadores antes mencionados

Fuente: Elaboración Propia



Al terminar el cálculo de los indicadores seleccionados, se detalla las imágenes generadas a partir de cada indicador.

**Tabla 13:** *Imágenes multiespectrales procesadas que se usarán para la estimación*

N°	Imagen	Descripción
1	NDVI	Imagen procesada con la fórmula del indicador de vegetación NDVI
2	DVI	Imagen procesada con la fórmula del indicador de vegetación DVI
3	BI	Imagen procesada con la fórmula del indicador de vegetación BI
4	SI	Imagen procesada con la fórmula del indicador de salinidad SI
5	SI1	Imagen procesada con la fórmula del indicador de salinidad SI1
6	SI2	Imagen procesada con la fórmula del indicador de salinidad SI2
7	SI3	Imagen procesada con la fórmula del indicador de salinidad SI3
8	SI11	Imagen procesada con la fórmula del indicador de salinidad SI11
9	ASTERSI	Imagen procesada con la fórmula del indicador de salinidad ASTER SI
10	RED	Imagen de la reflectancia banda 4
11	GREEN	Imagen de la reflectancia banda 3
12	BLUE	Imagen de la reflectancia banda 2

**Fuente:** Elaboración Propia

### 3.3.5 Recolección muestras de suelos de escasa vegetación y realizar la prueba de conductividad eléctrica.

Se escogió el distrito de Mórrope, Lambayeque para tomar las muestras de tierra, y realizar pruebas de Conductividad Eléctrica (CE) de estas, debido a que la prueba de CE sirve para medir la concentración de sales de una solución a partir un extracto de solución del suelo.

Se tomaron 10 puntos de muestreo a una profundidad de 30 cm del suelo, geo referenciados usando un Sistema de Posicionamiento Global (GPS), con un sistema de coordenadas UTM-WSG84, zona 17S, representado en la siguiente imagen de Google Earth.



**Figura 20:** *Ubicación de las muestras recogidas en la ciudad de Mórrope*

**Fuente:** Elaboración propia

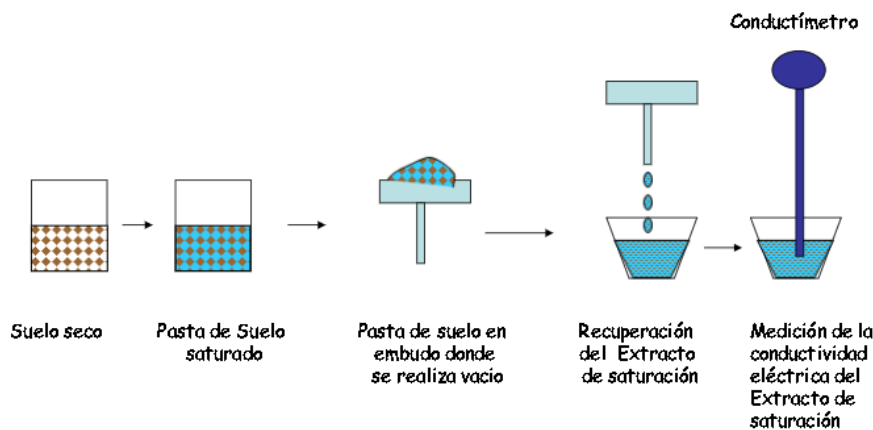
La distancia entre las muestras fue de mínimo 900 m<sup>2</sup>, el cual es el área de terreno que representa cada pixel en las imágenes multispectrales proporcionadas por el satélite Landsat 8, con el fin de obtener muestras de diferentes pixeles en suelos de escasa vegetación, la zona también tenía suelo de abundante vegetación a los alrededores por tal motivo se tomó otras 4 muestras más lejanas.



El GPS utilizado fue un Garmin Montana 650t el cual tiene una precisión entre +/- 3.65m, dicha precisión es suficiente debido al tamaño de pixel (30mx30m) antes mencionado.

El procedimiento que se realizó a las muestras fue el siguiente:

1. Se realizó el secado de las muestras de tierra, extendiéndolas al aire libre sobre un depósito.
2. Se procedió a moler la muestra con un rodillo para deshacer los agregados del suelo
3. Se tamizó la muestra de suelo con el tamiz N° 10, separando los elementos gruesos de la tierra fina, y se pesó y agregó 500 gramos de suelo tamizado en un recipiente de plástico.
4. Se agregó agua destilada al recipiente de plástico, y se bate con una espátula hasta llegar a una pasta de saturación.
5. En un embudo se colocó un papel filtro y se agregó la pasta saturada, y con una bomba al vacío se extrajo el extracto del suelo
6. Al final se calcula la conductividad eléctrica del extracto del suelo, con un conductímetro.



**Figura 21:** Procedimiento para la medición de Conductividad eléctrica

**Fuente:** Interpretación de Análisis de Suelos



**Tabla 14:** *Ubicación de las muestras tomadas y su respectivo valor de conductividad eléctrica.*

N° PRUEBA	UBICACIÓN UTM 17S	VALOR DE CONDUCTIVIDAD ELECTRICA (CE)	Grado de salinidad
D1	617226 , 9265494	1.86	No salino
D2	617206 , 9265509	1.26	No salino
D3	617203 , 9265512	1.44	No salino
D4	617196 , 9265493	1.50	No salino
D5	617175 , 9265478	0.90	No salino
D6	617160 , 9265457	1.14	No salino
D7	617183 , 9265434	1.26	No salino
D8	617259 , 9265455	9.15	Muy salino
D9	617291 , 9265427	8.80	Muy salino
D10	617241 , 9265490	1.32	No salino

**Fuente:** Elaboración Propia

Luego de acuerdo a las coordenadas de las muestras tomadas por GPS, se ubicó cada muestra en el formato UTM 17S en las imágenes multiespectrales (**Tabla 14**), para ello se convirtió las coordenadas en UTM 17S a coordenadas UTM 17N, debido a que el formato de coordenadas UTM 17N es el estándar del proyecto Landsat 8.

Para extraer los indicadores antes seleccionados, se toma como guía para dicho proceso los pasos del **Anexo 7**, y se desarrolló un programa que realiza este proceso automáticamente con ayuda de la librería GDAL (**Anexo 8**)

A partir de la extracción de los indicadores, se obtuvo la siguiente tabla en la cual se relacionan los indicadores y los valores de conductividad eléctrica.





**Tabla 15:** *Relación entre los indicadores elegidos y el valor de conductividad eléctrica.*

MUESTRA	RED	GREEN	BLUE	NDVI	DVI	BI	ASTERSI	SI	SI1	SI2	SI3	SI11	CE
D1	0.100982	0.110124	0.122933	0.366371	0.116777	0.149414	0.173738	0.035406	0.105454	0.463731	0.090460	1.420540	1.86
D2	0.099646	0.111538	0.124164	0.357052	0.110674	0.149566	0.247902	0.035112	0.105424	0.473783	0.089513	1.659227	1.26
D3	0.117773	0.120785	0.131316	0.247027	0.077275	0.168699	0.089632	0.042678	0.119269	0.603814	0.108328	1.196914	1.44
D4	0.109024	0.116201	0.129325	0.306622	0.096424	0.159339	0.187164	0.039207	0.112555	0.530664	0.097960	1.460521	1.50
D5	0.115389	0.115520	0.131656	0.223652	0.066483	0.163277	0.065531	0.041868	0.115454	0.634452	0.101247	1.140253	0.90
D6	0.114839	0.118166	0.132101	0.244854	0.074472	0.164776	0.129483	0.041739	0.116490	0.606614	0.102724	1.297487	1.14
D7	0.097629	0.109076	0.122959	0.348199	0.104308	0.146386	0.193991	0.034234	0.103194	0.483461	0.086605	1.481360	1.26
D8	0.111931	0.110071	0.125684	0.169404	0.045658	0.156985	-0.009223	0.039682	0.110997	0.710273	0.098027	0.981723	9.15
D9	0.128120	0.122514	0.132809	0.250651	0.085710	0.177269	-0.005725	0.046691	0.125285	0.599167	0.118189	0.988616	8.80
D10	0.110726	0.111093	0.126443	0.189143	0.051657	0.156850	0.000403	0.039373	0.110909	0.681884	0.097284	1.000807	1.32

**Fuente:** Elaboración Propia



### 3.3.6 Evaluación de las técnicas de estimación.

Luego, se utilizó la información de la relaciones entre los indicadores y bandas (*Figura 22*) y el valor de conductividad eléctrica (*Figura 23*) para estimar la salinidad, además se implementó la estimación usando la técnica de Partial least squares regression (PLSR), y la técnica de Support Vector Machine (SVM), los indicadores que se usaron para medir la precisión de ambas técnicas fueron, el coeficiente de determinación ( $R^2$ ) y la raíz del error cuadrático medio (RMSE).

X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
0.10098	0.11012	0.12293	0.36637	0.11678	0.14941	0.17374	0.03541	0.10545	0.46373	0.09046	1.42054
0.09965	0.11154	0.12416	0.35705	0.11067	0.14957	0.24790	0.03511	0.10542	0.47378	0.08951	1.65923
0.11777	0.12079	0.13132	0.24703	0.07728	0.16870	0.08963	0.04268	0.11927	0.60381	0.10833	1.19691
0.10902	0.11620	0.12933	0.30662	0.09642	0.15934	0.18716	0.03921	0.11256	0.53066	0.09796	1.46052
0.11539	0.11552	0.13166	0.22365	0.06648	0.16328	0.06553	0.04187	0.11545	0.63445	0.10125	1.14025
0.11484	0.11817	0.13210	0.24485	0.07447	0.16478	0.12948	0.04174	0.11649	0.60661	0.10272	1.29749
0.09763	0.10908	0.12296	0.34820	0.10431	0.14639	0.19399	0.03423	0.10319	0.48346	0.08661	1.48136
0.11193	0.11007	0.12568	0.16940	0.04566	0.15699	-0.00922	0.03968	0.11100	0.71027	0.09803	0.98172
0.12812	0.12251	0.13281	0.25065	0.08571	0.17727	-0.00573	0.04669	0.12529	0.59917	0.11819	0.98862
0.11073	0.11109	0.12644	0.18914	0.05166	0.15685	0.00040	0.03937	0.11091	0.68188	0.09728	1.00081

**Figura 22:** Variables independientes (indicadores)

Fuente: Elaboración propia

Y  
1.86  
1.26  
1.44  
1.5  
0.9  
1.14  
1.26  
9.15  
8.8  
1.32

**Figura 23:** Variable dependiente (CE)

Fuente: Elaboración propia



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

La técnica de estimación de SVM tuvo los siguientes parámetros:

Kernel	rbf
Coficiente de penalidad (C)	0.00001
Coficiente del kernel	0.8

Y el siguiente código de implementación:

```
dataset = pd.read_csv('G:/caracteristicas2.csv',",")

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

X = dataset[indicadores]
y = dataset['EC']
svr = SVR(kernel='rbf', C=1e5, gamma=0.8)
svr.fit(X, y)
y_pred = svr.predict(X)
```

**Figura 24:** Código de implementación de estimación SVM

**Fuente:** Elaboración propia

Y se obtuvo los siguientes resultados:

**Tabla 16:** Resultados de la estimación usando SVM, relación entre la conductividad eléctrica estimada y real.

Nº	CE Estimada	CE Real
D1	1.96	1.86
D2	1.36	1.26
D3	1.54	1.44
D4	1.40	1.50
D5	0.80	0.90
D6	1.24	1.14
D7	1.16	1.26
D8	9.05	9.15
D9	8.70	8.80
D10	1.94	1.32
Media	2.92	2.86

**Fuente:** Elaboración Propia



Algoritmo de PLSR:

$u = y_j$  para una posición  $j$

Procesar

$$p = \frac{X^T u}{\|X^T u\|}$$

$$t = Xp$$

$$q = 1$$

$$u = Y$$

$$X = X - tp^T$$

$$Y = Y - uq^T$$

Para  $n$  componentes .

Para el algoritmo de PLSR se realiza iteraciones consecutivas, donde cada iteración se le llama componente hasta que la variable  $t$  deje de cambiar.

Para la estimación dado las variables independientes (**Figura 15**) se desea estimar la variable dependiente (**Figura 16**).

La técnica de estimación de PLSR tuvo los siguientes parámetros:

Número de componentes	7
-----------------------	---

Y el siguiente código de implementación:

```
dataset = pd.read_csv('G:/caracteristicas2.csv',",")
indicadores = ['RED', 'GREEN', 'BLUE', 'NDVI', 'DVI',
               'BI', 'ASTERSI', 'SI', 'SI1', 'SI2', 'SI3', 'SI11']
X = dataset[indicadores]
y = dataset['EC']

pls2 = PLSRegression(n_components=7)
pls2.fit(X, y)
print pls2.coef_
y_pred = pls2.predict(X)
```

**Figura 25:** Código de implementación de estimación PLSR

Fuente: Elaboración propia



Y se obtuvo los siguientes resultados:

**Tabla 17:** Resultados de la estimación usando PLSR, relación entre la conductividad eléctrica estimada y real.

N°	CE Estimada	CE Real
D1	2.19	1.86
D2	1.75	1.26
D3	1.50	1.44
D4	0.25	1.50
D5	1.04	0.90
D6	1.84	1.14
D7	0.99	1.26
D8	8.88	9.15
D9	8.76	8.80
D10	1.42	1.32
MEDIA	2.86	2.86

Fuente: Elaboración Propia

Luego de desarrollar ambas técnicas de estimación, se procedió a calcular la operación de variables, utilizando los resultados 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

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

$X_i$  = valor medido

$Y_i$  = valor estimado

$n$  = el número de muestras

**DESARROLLO DE LA OPERACIONALIZACION DE VARIABLES PARA SVM**

- Coeficiente de correlación  $R^2$

$$n = 10$$

$$\bar{x} = 2.86$$

$$\bar{y} = 2.92$$

**Tabla 18:** *Cálculo del indicador  $R^2$  para SVM*

N°	$x$	$x - \bar{x}$	$y$	$y - \bar{y}$
D1	1.86	-1.00	1.96	-0.96
D2	1.26	-1.60	1.36	-1.56
D3	1.44	-1.42	1.54	-1.38
D4	1.50	-1.36	1.40	-1.52
D5	0.90	-1.96	0.80	-2.12
D6	1.14	-1.72	1.24	-1.68
D7	1.26	-1.60	1.16	-1.76
D8	9.15	6.29	9.05	6.13
D9	8.80	5.94	8.70	5.78
D10	1.32	-1.54	1.94	-0.98

Fuente: Elaboración Propia

$$R^2 = 0.98$$





- Raíz del error Cuadrático medio RMSE

$$n = 10$$

**Tabla 19:** Cálculo del indicador RMSE para SVM

N°	x	y	x - y
D1	1.86	1.96	-0.10
D2	1.26	1.36	-0.10
D3	1.44	1.54	-0.10
D4	1.50	1.40	0.10
D5	0.90	0.80	0.10
D6	1.14	1.24	-0.10
D7	1.26	1.16	0.10
D8	9.15	9.05	0.10
D9	8.80	8.70	0.10
D10	1.32	1.94	-0.62
Media	2.86	2.92	-0.10

Fuente: Elaboración Propia

$$RMSE = 0.22$$

**DESARROLLO DE LA OPERACIONALIZACION DE VARIABLES PARA PLSR**

- Coeficiente de correlación  $R^2$

$$n = 10$$

$$\bar{x} = 2.86$$

$$\bar{y} = 2.92$$



**Tabla 20:** Cálculo del indicador  $R^2$  para PLSR

N°	$x$	$x - \bar{x}$	$y$	$y - \bar{y}$
D1	1.86	-1.00	2.19	-0.67
D2	1.26	-1.60	1.75	-1.11
D3	1.44	-1.42	1.5	-1.36
D4	1.50	-1.36	0.25	-2.61
D5	0.90	-1.96	1.04	-1.82
D6	1.14	-1.72	1.84	-1.02
D7	1.26	-1.60	0.99	-1.87
D8	9.15	6.29	8.88	6.02
D9	8.80	5.94	8.76	5.9
D10	1.32	-1.54	1.42	-1.44

**Fuente:** Elaboración Propia

$$R^2 = 0.97$$

- Raíz del error Cuadrático medio RMSE

**Tabla 21:** Cálculo del indicador RMSE para PLSR

N°	$x$	$y$	$x - y$
D1	1.86	2.19	-0.33
D2	1.26	1.75	-0.49
D3	1.44	1.5	-0.06
D4	1.50	0.25	1.25
D5	0.90	1.04	-0.14
D6	1.14	1.84	-0.70
D7	1.26	0.99	0.27
D8	9.15	8.88	0.27
D9	8.80	8.76	0.04
D10	1.32	1.42	-0.10
Media	2.86	2.86	-0.33

**Fuente:** Elaboración Propia

$$RMSE = 0.51$$



## IV. CONCLUSIONES Y RECOMENDACIONES

### 4.1 Conclusiones

1. La técnica PLSR tuvo buenos resultados en los trabajos de investigación relacionados con imágenes multiespectrales, siendo esta la más precisa al estimar el vigor de plantas, y la técnica SVM también tuvo buenos resultados al estimar la salinidad mediante propiedades físicas del ambiente, por tal motivo se escogió ambas técnicas.
2. Las imágenes multiespectrales satelitales adquiridas de la región de Chiclayo fueron gratuitas y de fácil acceso, el único problema fue la nubosidad en las imágenes, por tal motivo solo se utilizó aquellas que tenían poca nubosidad para un mejor estudio.
3. La realización de calibraciones atmosféricas a las imágenes multiespectrales, permite poder extraer indicadores de salinidad, y vegetación de acuerdo a las reflectancias de las bandas espectrales.
4. El criterio de selección de los indicadores de salinidad y vegetación fue el porcentaje de correlación con el valor de conductividad eléctrica, una mayor relación permite análisis más precisos, y estos indicadores fueron extraídos de trabajos de investigación relacionados.
5. Para estimar la salinidad se debe tomar muestras de terreno de escasa vegetación, para la evaluación de la salinidad de las muestras se realizó la prueba de conductividad eléctrica.
6. Para evaluar las técnicas de estimación se utilizó los criterios de  $R^2$  y RMSE, el  $R^2$  es el porcentaje de precisión en la estimación y el RMSE es el error cuadrático en la estimación, la técnica de SVM obtuvo un 98% y 0.22 en cada uno respectivamente, y la técnica PLSR 97% Y 0.51, los resultados demuestran que SVM es más preciso y tiene menos error al estimar la salinidad que PLSR.



## 4.2 Recomendaciones

1. La precisión de las técnicas de estimación varía dependiendo al trabajo de investigación en el que se aplica, por ello se recomienda no asumir que una técnica de estimación es mejor que otra, si es que no se ha probado en el mismo contexto de investigación, se debería tomar dichos resultados como referencia para cuantificar las técnicas, más no para concluir que una es mejor que otra.
2. El presente trabajo de investigación se realizó con imágenes multiespectrales satelitales, un pixel de esta imagen representa 30mx30m o 900m<sup>2</sup> de terreno, se espera que en trabajos futuros las imágenes multiespectrales que se utilicen puedan ser con un vehículo aéreo no tripulado (dron) y utilizando cámaras multiespectrales para tener una mejor estimación de la salinidad en suelos de escasa vegetación.
3. Si se adquiere imágenes multiespectrales satelitales se recomienda realizar las calibraciones atmosféricas al procesar las imágenes adquiridas, debido a que las imágenes multiespectrales satélites son entregadas en valores numéricos.
4. El porcentaje de correlación entre los indicadores seleccionados y la conductividad eléctrica varió, relacionado a los trabajos de investigación que se tomaron como base, puesto que este trabajo de investigación se realizó en otro continente y utilizando imágenes multiespectrales procedentes de otro satélite.
5. Se recomienda tomar muestras de conductividad eléctrica, para compararlas con valores de indicadores de salinidad y vegetación ya existentes, y lograr estimar la salinidad.



## REFERENCIAS

- Agudelo, Gabriel, Miguel Aigner, and Jaime Ruiz. 2008. "Diseños De Investigación Experimental Y No-Experimental." *Centro de Estudios de Opinión* 1–46.
- Akramkhanov, A., C. Martius, S. J. Park, and J. M. H. Hendrickx. 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. "Mapping and Modelling Spatial Variation in Soil Salinity in the Al Hassa Oasis Based on Remote Sensing Indicators and Regression Techniques." *Remote Sensing* 6(2):1137–57.
- Bado, Souleymane, Brian P. Forster, and Joanna Jankowicz-cieslak. 2016. *Protocols for Pre- Field Screening of Mutants for Salt Tolerance in Rice , Wheat and Barley*.
- Bouksila, Fethi, Akissa Bahri, Ronny Berndtsson, Magnus Persson, Jelte Rozema, and Sjoerd E. A. T. M. Van der Zee. 2013. "Assessment of Soil Salinization Risks under Irrigation with Brackish Water in Semiarid Tunisia." *Environmental and Experimental Botany* 92:176–85.
- Campbell, James B. and Randolph H. Wynne. 2011. *Introduction to Remote Sensing FIFTH EDITION*. Vol. XXXIII.
- Department of the Interior U.S. 2018. "LANDSAT 8 DATA USERS HANDBOOK." 8(October):2–4.
- Ding, Jianli and Danlin Yu. 2014. "Monitoring and Evaluating Spatial Variability of Soil Salinity in Dry and Wet Seasons in the Werigan-Kuqa Oasis, China, Using Remote Sensing and Electromagnetic Induction Instruments." *Geoderma* 235–236:316–22.
- Domínguez, Yanetsys Sarduy. 2007. "Revista Cubana de Salud Pública El Análisis de Información y Las Investigaciones Cuantitativa y Cualitativa The Information Analysis and the Qualitative and Quantitative." 33(3):1–8.
- FAO. 2007. *Advances in the Assessment and Monitoring of Salinization and Status of Biosaline Agriculture*.



- Gonzaga Aguilar, Carlos. 2014. “Aplicación de Índices de Vegetación Derivados de Imágenes Satelitales Landsat 7 ETM + y ASTER Para La Caracterización de La Cobertura Vegetal En La Zona Centro de La Provincia De Loja, Ecuador.” 103.
- Guan, Xiaoyan, Shaoli Wang, Zhanyi Gao, and Ye Lv. 2013. “Dynamic Prediction of Soil Salinization in an Irrigation District Based on the Support Vector Machine.” *Mathematical and Computer Modelling* 58(3–4):719–24.
- Guareño, Juan José Martín. 2016. “SUPPORT VECTOR REGRESSION: PROPIEDADES Y APLICACIONES.”
- Heil, Kurt and Urs Schmidhalter. 2017. “The Application of EM38: Determination of Soil Parameters, Selection of Soil Sampling Points and Use in Agriculture and Archaeology.” *Sensors (Switzerland)* 17(11).
- Mashimbye, Z. E., M. A. Cho, J. P. Nell, W. P. De clerq, A. Van niekerk, and D. P. Turner. 2012. “Model-Based Integrated Methods for Quantitative Estimation of Soil Salinity from Hyperspectral Remote Sensing Data: A Case Study of Selected South African Soils.” *Pedosphere* 22(5):640–49.
- Mousavi, Seyedeh Zohreh, Mahmood Habibnejad, Ataollah Kavian, Karim Solaimani, and Farhad Khormali. 2017. “Digital Mapping of Topsoil Salinity Using Remote Sensing Indices In.” 5(2):1771–86.
- Ng, Kee Siong. 2013. “A Simple Explanation of Partial Least Squares.” 1–10.
- Ormeño-Villajos, S. 2006. “Teledeteccion Fundamental.” 344.
- Peréz, Carlos and Ángel Luis Muñoz. 2006. *Teledetección: Nociones y Aplicaciones*. Vol. 53.
- Romero, Francisco Sacristán. 2006. “La Teledetección Satelital y Los Sistemas de Protección Ambiental.” *Quivera* 8(1):315–56.
- Sarkar, Dipanjan, Raghav Bali, and Tushar Sharma. 2018. *Practical Machine Learning with Python*.
- Shahabi, Mahmoud, Ali Asghar Jafarzadeh, Mohammad Reza Neyshabouri, Mohammad Ali Ghorbani, and Khalil Valizadeh Kamran. 2017. “Spatial Modeling



of Soil Salinity Using Multiple Linear Regression, Ordinary Kriging and Artificial Neural Network Methods.” *Archives of Agronomy and Soil Science* 63(2):151–60.



ANEXOS

Anexo 1: Población de Técnicas de estimación y la muestra escogida

Tabla 22: Población de técnicas de estimación

TÉCNICAS DE ESTIMACIÓN	PRECISIÓN	
	R <sup>2</sup>	Error relativo
1. PARTIAL LEAST SQUARE REGRESSION (PLSR)	89.00%	
	87.00%	
2. MULTIPLE LINEAR REGRESSION (MLR)	52.00%	
	36.00%	
	96.30%	
	65.00%	
3. RANDOM FOREST REGRESSION (RF)	79.00%	
	98.30%	
	85.00%	
4. SUPPORT VECTOR MACHINE (SVM)	83.00%	
	61.70%	
		37.00%
5. BOOSTED BINARY REGRESSION (BBRT)	82.00%	
6. ARTIFICIAL NEURONAL NETWORK (ANN)	81.00%	
	93.00%	
	69.00%	
	99.70%	
	41.90%	
		49.00%
7. DECISION TREE REGRESSION (DT)	80.00%	
8. PRINCIPAL COMPONENT REGRESSION (PCR)	79.00%	

Fuente: Elaboración propia

Recopilado de:

A. “ A Comparison of Regression Techniques for Estimation of Above-Ground Winter Wheat Biomass Using Near-Surface Spectroscopy “ (Yue, Feng, Yang, & Li, 2018)



- B. “Mapping and Modelling Spatial Variation in Soil Salinity in the Al Hassa Oasis Based on Remote Sensing Indicators and Regression Techniques” (Allbed, Kumar, and Sinha 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”(Shahabi et al., 2017)
- 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" (Kayri, Kayri, and Gencoglu 2017)

El orden de las técnicas de estimación se extrajo de la fuente **A**, en dicha fuente se compara todas las técnicas de la población

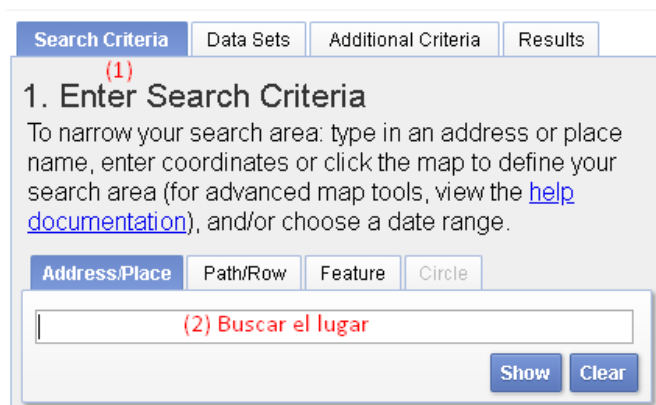


## Anexo 2: Pasos para recolectar las imágenes multiespectrales

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

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

2. En el primer menú se insertó el lugar **Chiclayo** a buscar las imágenes



**Figura 26:** *Búsqueda de imágenes multiespectrales por lugar.*

**Fuente:** Elaboración propia.

3. Luego se buscó las imágenes desde **01 de enero de 2018 hasta 30 de setiembre de 2018.**



**Search Criteria** | Data Sets | Additional Criteria | Results

### 1. Enter Search Criteria

To narrow your search area: type in an address or place name, enter coordinates or click the map to define your search area (for advanced map tools, view the [help documentation](#)), and/or choose a date range.

**Address/Place** | Path/Row | Feature | Circle

Show Clear

**Coordinates** | Predefined Area | Shapefile | KML

Degree/Minute/Second | Decimal

1. Lat: 06° 46' 35" S, Lon: 079° 50' 39" W

Use Map Add Coordinate Clear Coordinates

**Date Range** | Result Options

Search from: 01/01/2018 to: 09/30/2018

Search months: (all) (3) Fechas

Data Sets » Additional Criteria » Results »

**Figura 27:** *Búsqueda de imágenes multiespectrales por fecha.*

**Fuente:** Elaboración propia.

- En el segundo menú (DataSets) se seleccionó los orígenes de datos **Landsat 8 OLI**.



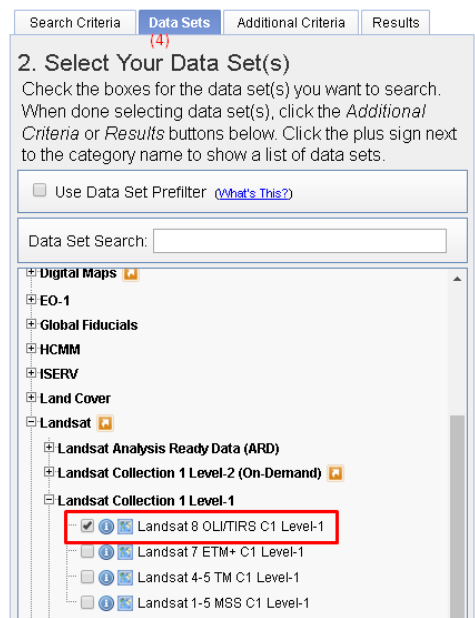


Figura 28: Búsqueda de imágenes multiespectrales por satélite.

Fuente: Elaboración propia.

- En el tercer menú (Additional Criteria), se seleccionó **less than 30%** en **Scene Cloud Cover** para obtener las imágenes que tengan cobertura de nubes menor al 30%.

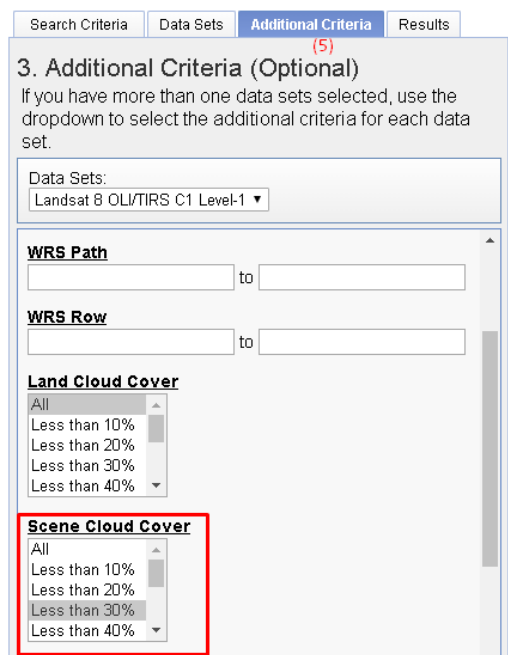


Figura 29: Búsqueda de imágenes multiespectrales por filtro de porcentaje de nubosidad.

Fuente: Elaboración propia.





6. En el cuarto menú (Results), se descargó las imágenes.

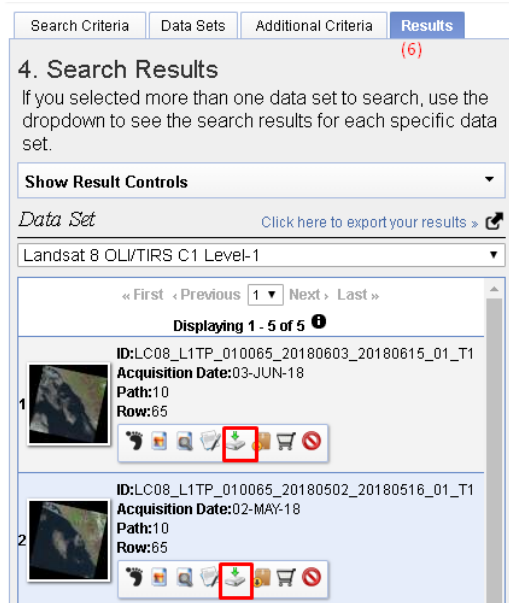


Figura 30: Vista de adquisición de imágenes multiespectrales.

Fuente: Elaboración propia.

7. Por último se seleccionó el paquete **Level 1- GeoTIFF Data Product** para su descarga.

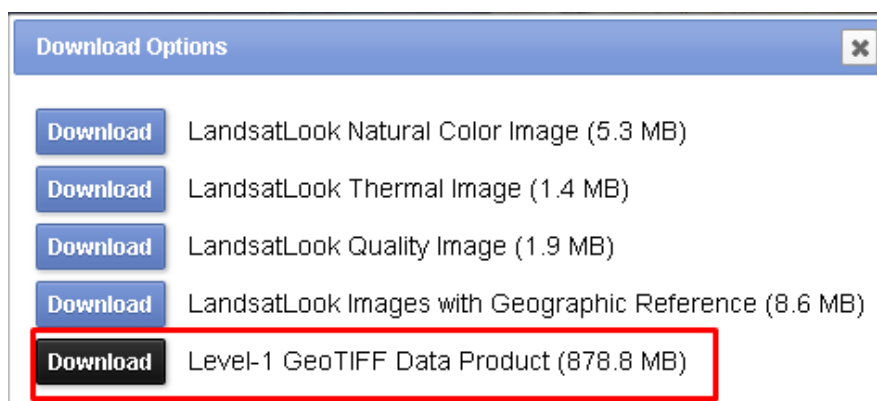


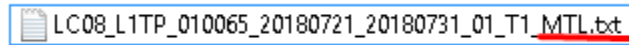
Figura 31: Adquisición de paquete de imágenes multiespectrales.

Fuente: Elaboración propia.



**Anexo 3: Pasos para convertir los valores digitales a reflectancia manualmente.**

1. Ubicar y leer el archivo de metadata, es un archivo que tiene el sufijo **\_MTL** al final y con extensión **.TXT**.



**Figura 32:** Archivo de metadata de un paquete 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 = 49.77414932
```

**Figura 33:** Variables de metadata a utilizar para calcular la reflectancia

**Fuente:** Elaboración propia.

3. Crear tablas con base a las variables anteriores de la banda 1-7.

$$p_{\lambda'} = M_p Q_{cal} + A_p$$

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



**Tabla 23:** Variables de cálculo de reflectancia por banda

BANDA	REFLECTANCE_MULT_BAND ( $M_p$ )	REFLECTANCE_ADD_BAND ( $A_p$ )
BANDA 1	$2 \times 10^{-5}$	-0.10
BANDA 2	$2 \times 10^{-5}$	-0.10
BANDA 3	$2 \times 10^{-5}$	-0.10
BANDA 4	$2 \times 10^{-5}$	-0.10
BANDA 5	$2 \times 10^{-5}$	-0.10
BANDA 6	$2 \times 10^{-5}$	-0.10
BANDA 7	$2 \times 10^{-5}$	-0.10

**Fuente:** Elaboración propia

**Tabla 24:** Variables de ángulo de elevación para el cálculo de reflectancia

Angulo elevación (SE) (grados)	Angulo elevación (SE) (radianes)
49.77414932	0.86872278801114

**Fuente:** Elaboración propia



4. Ubicar un pixel en la imagen de una banda, y desarrollar la ecuación.

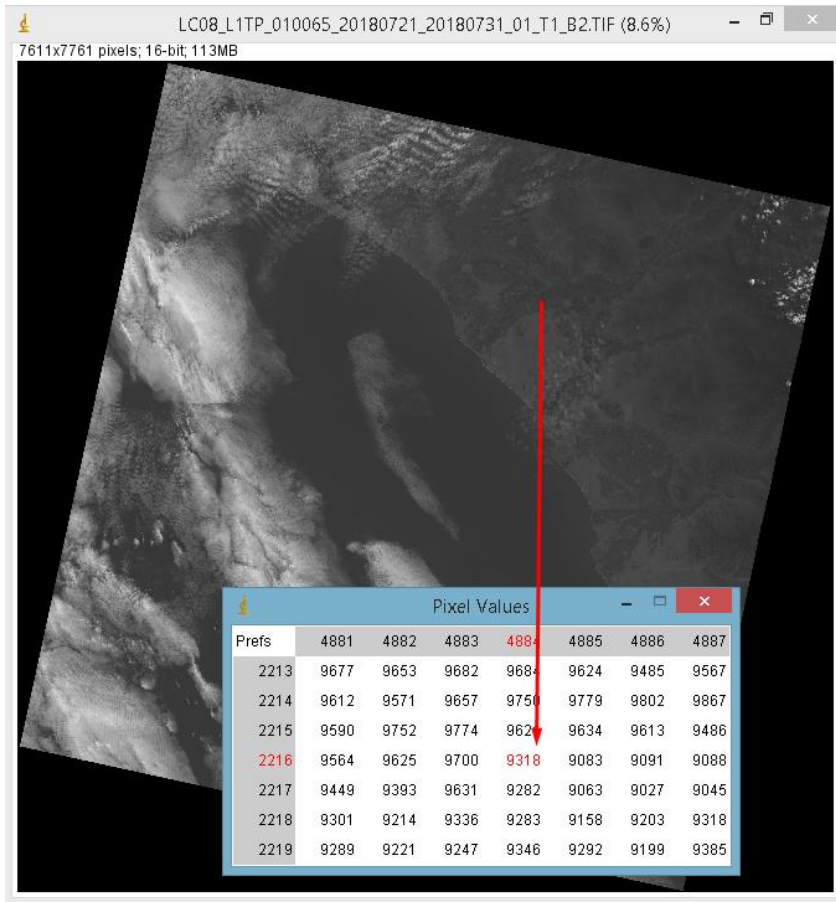


Figura 34: Búsqueda de valor de un pixel en una imagen espectral

Fuente: Elaboración propia

Tabla 25: Ubicación y valor de un pixel en una imagen espectral

BANDA	VALOR DEL PIXEL ( $Q_{cal}$ )	FILA EN LA IMAGEN	COLUMNA EN LA IMAGEN
BANDA 2	9318	4884	2216

Fuente: Elaboración propia

$$p_{\lambda'} = M_p Q_{cal} + A_p$$

$$p_{\lambda'} = 2 * 10^{-5} * 9318 + -0.10$$

$$p_{\lambda'} = 0.00002 * 9318 + -0.10$$

$$p_{\lambda'} = 0.08636$$



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

$$p_{\lambda} = \frac{0.08636}{\sin(0.86872)}$$

$$p_{\lambda} = 0.11311024$$

La reflectancia del pixel en la posición (4884,2216) de la imagen de la banda 2 es

$$0.11311024$$



## Anexo 4: Programa para convertir los valores digitales a reflectancia automáticamente.

- Métodos para leer el archivo de metadata

```
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,c,p):
    if filter(lambda x: c in x, p):
        return p

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

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

    metadata.close
```

Figura 35: Código fuente en Python para leer el archivo de metadata

Fuente: Elaboración propia

- Métodos para leer los pixeles de la imagen en una matriz numérica, y calcular la reflectancia.

```
def leerImagen(self):
    dataset = gdal.Open(self.ruta)
    self.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]
    matriz = band.ReadAsArray(0, 0, dataset.RasterXSize, dataset.RasterYSize).astype(numpy.float)
    self.calcularReflectancia(matriz)
```

Figura 36: Código fuente en Python de la función para leer una imagen

Fuente: Elaboración propia





```
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
```

**Figura 37:** Código fuente en Python para calcular la reflectancia

**Fuente:** Elaboración propia

- Extracción de valores en metadata

```
from metadata import Metadata
from imagen import Imagenes,Imagen,Banda
from osgeo import gdal
import matplotlib.pyplot as plt
import sys
if __name__ == '__main__':
    directorio = "C:\Users\Dante\Downloads\Compressed\LC08_L1TP_010065_20180721_20180731_01_T1"
    ruta = directorio+"\LC08_L1TP_010065_20180721_20180731_01_T1_MTL.txt"
    metadata = Metadata()
    metadata.leerMetadata(ruta)
    imagenes = Imagenes()
```

**Figura 38:** Código fuente en Python para calcular la reflectancia

**Fuente:** Elaboración propia

- Cálculo de la reflectancia de las imágenes espectrales.

```
for ban in metadata.bandas[:7]:
    ruta = metadata.buscarMetadata("FILE_NAME_"+ban)
    if ruta is not None:
        banda = Banda()
        banda.nombre= ban
        banda.REFLECTANCE_MULT_BAND = float(metadata.buscarMetadata("REFLECTANCE_MULT_"+ban)[0][1])
        banda.REFLECTANCE_ADD_BAND = float(metadata.buscarMetadata("REFLECTANCE_ADD_"+ban)[0][1])
        banda.SUN_ELEVATION = float(metadata.buscarMetadata("SUN_ELEVATION")[0][1])
        img = Imagen()
        img.banda = banda
        img.ruta = directorio + '\\' + ruta[0][1]
        imagenes.lista.append(img)

for img in imagenes.lista:
    print "Leyendo imagen"+ img.banda.nombre
    img.leerImagen()
```

**Figura 39:** Código fuente en Python de la ejecución de las funciones de leer de imagen y leer la metadata

**Fuente:** Elaboración propia



**Anexo 5: Pasos para calcular los indicadores manualmente.**

1. Identificar el indicador a utilizar, y las imágenes que se utilizaran.

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

2. Las bandas a utilizar para este indicador son las  $p_{NIR}$  (Imagen de la reflectancia en la banda infrarrojo cercano o Banda 6) y  $p_R$  (Imagen de la reflectancia en la banda roja o Banda 5).

Se toma como referencia el pixel en la fila 4255 y columna 2113, el valor de dicho pixel representa la reflectancia en dicha zona.

El valor de reflectancia para la **Banda 5** ( $p_{NIR}$ ) es **0.1752**



**Figura 40:** Búsqueda del valor de reflectancia de una imagen procesada

**Fuente:** Elaboración propia



El valor de reflectancia para la **Banda 4** ( $p_R$ ) es **0.1618**



**Figura 41:** Búsqueda del valor de reflectancia de una otra imagen procesada

**Fuente:** Elaboración propia

3. Reemplazar las reflectancias obtenidas en la ecuación del indicador.

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

$$NDVI = \frac{0.1752 - 0.1618}{0.1752 + 0.1618}$$

$$NDVI = \frac{0.0134}{0.337}$$

$$NDVI = 0.04$$



## Anexo 6: Programa para calcular los indicadores de una imagen automáticamente.

- Referenciación de las imágenes y el identificador de reflectancia.

```
pblue = imagenes.buscarImagen("BAND_2")
pgreen = imagenes.buscarImagen("BAND_3")
pred = imagenes.buscarImagen("BAND_4")
pnir = imagenes.buscarImagen("BAND_5")
pswir = imagenes.buscarImagen("BAND_6")
```

**Figura 42:** Código fuente en Python de lectura de una imagen espectral procesada

**Fuente:** Elaboración propia

- Implementación de los indicadores.

```
pnarvi = (pnir-pred)/( pnir + pred)
```

**Figura 43:** Código fuente en Python implementación del indicador NDVI

**Fuente:** Elaboración propia

- Copia de la metadata (georreferenciación) de la imagen.

```
output_image = Imagen()
imt = imagenes.lista[0]
output_image.xOrigin = imt.xOrigin
output_image.yOrigin = imt.yOrigin
output_image.pixelWidth = imt.pixelWidth
output_image.pixelHeight = imt.pixelHeight
output_image.metadata = imt.metadata
```

**Figura 44:** Código fuente en Python de lectura de los puntos de origen y el tamaño de pixel de una imagen espectral

**Fuente:** Elaboración propia

- Cálculo del indicador.

```

directorio = directorio + "\\INDICADORES\\"
print "Exportando ndvi"
output_image.guardarImagen(directorio + '\\NDVI.tif',pndvi)
    
```

**Figura 45:** Código fuente en Python de la exportación de un indicador calculado.

**Fuente:** Elaboración propia



## Anexo 7: Pasos para ubicar una coordenada en una imagen con información geoespacial, mediante una interfaz gráfica.

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.

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

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

Donde:

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 la coordenada a ubicar y la imagen donde se ubicará, para extraer el valor de dicho 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 26:** Información de coordenada a buscar y el indicador donde se buscará

COORDENADA UTM 17S	IMAGEN
617226 , 9265494	NDVI

Fuente: Elaboración Propia

**Conversión de la coordenada**

$$617226, 9265494 \text{ (UTM 17S)} = 617226, -734506 \text{ (UTM 17N)}$$

**Tabla 27:** Ubicación y valor de un pixel en una imagen espectral

A	B	$X_0$	$Y_0$	m	n
617226	-734506	495000	-683400	30	30

Fuente: Elaboración Propia

3. Reemplazar la coordenada en la fórmula.

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

$$fila = \frac{617226 - 495000}{7611}$$

$$fila = 4074$$

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

$$columna = \frac{-683400 - -734506}{30}$$

$$columna = 1703$$



## Anexo 8: Programa para ubicar coordenadas en una imagen con información geoespacial.

- Función que lee la información de una imagen, valor de pixel en valores numéricos, y la metadata (coordenadas asociadas a la imagen),

```
def leerIndice(self):
    dataset = gdal.Open(self.ruta)
    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(numpy.float)
```

**Figura 46:** Código fuente en Python de lectura de un indicador

**Fuente:** Elaboración propia

- Función que convierte la coordenada de formato UTM 17S (variable epsg:32717) a UTM 17N (variable epsg:32617), luego busca la posición en la imagen y extrae el valor del pixel.

```
def buscarPixel(self, coordenadas):
    valores = []

    for point in coordenadas:
        x, y = proj.transform(proj.Proj(init='epsg:32717'), proj.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 reflec
    return valores
```

**Figura 47:** Código fuente en Python de lectura de coordenada, conversión y búsqueda de coordenada.

**Fuente:** Elaboración propia



- Ejecución de las funciones anteriores.

```

directorio = "C:\Users\Dante\Downloads\Compressed\LC08_L1TP_010065_20180721_20180731_01_T1\INDICADORES"
array = ('RED','GREEN','BLUE','NDVI','DVI','BI','ASTERSI','SI','SI1','SI2','SI3','SI11')
imagenes = Imagenes()

df = pd.DataFrame()

cont = 0
for elmnt in array:

    path = directorio + '\\' +elmnt+ ".tif"
    img = Imagen()
    print path
    img.ruta = path
    img.leerIndice()
    points_list = [
        ( 617226 , 9265494 ),
        ( 617206 , 9265509 ),
        ( 617203 , 9265512 ),
        ( 617196 , 9265493 ),
        ( 617175 , 9265478 ),
        ( 617160 , 9265457 ),
        ( 617183 , 9265434 ),
        ( 617259 , 9265455 ),
        ( 617291 , 9265427 ),
        ( 617241 , 9265490 )
    ]
    valores = img.buscarPixel(points list)
    df_temp = pd.DataFrame({ elmnt : valores})
    df.insert( loc=cont, column=elmnt , value=valores)
    cont = cont +1
    
```

**Figura 48:** Código fuente en Python de ejecución de búsqueda de coordenadas y extracción de indicadores.

**Fuente:** Elaboración propia

- Exportar la información obtenida a un archivo csv

```

df.insert( loc=cont, column='EC' , value=[1.86,1.26,1.44,1.50,0.90,1.14,1.26,9.15,8.80,1.32])
print df
df.to_csv( 'E:\\' +"caracteristicas.csv", sep=',', encoding='utf-8')
    
```

**Figura 49:** Código fuente en Python de ejecución de búsqueda de coordenadas y extracción de indicadores.

**Fuente:** Elaboración propia



**Anexo 9: Análisis de muestras por conductividad eléctrica.**



**ANÁLISIS QUÍMICO DE SUELOS**

**SOLICITANTE** : STEVEN VILCHEZ GUIVAR

**TESIS** : COMPARACION DE TECNICAS DE ESTIMACION DEL GRADO DE SALINIDAD EN SUELOS DE ESCASA VEGETACION MEDIANTE EL PROCESAMIENTO DE IMAGENES MULTIESPECTRALES

**PROCEDENCIA** : DISTRITO MORROPE, PROVINCIA LAMBAYEQUE, REGION LAMBAYEQUE

**ASUNTO** : ANÁLISIS QUÍMICO DE DIEZ (10) MUESTRAS TRAÍDOS POR EL INTERESADO AL LABORATORIO DE SUELOS.

Muestra N°	C.E. dS/m
D1	1.86
D2	1.26
D3	1.44
D4	1.50
D5	0.90
D6	1.14
D7	1.26
D8	9.15
D9	8.80
D10	1.32

Lambayeque, 19 de octubre del 2018



**JEFATURA DE LABORATORIO DE SUELOS**

**Figura 50:** Análisis de muestras por conductividad eléctrica.

**Fuente:** Elaboración propia

