

UNIVERSIDAD NACIONAL DE CHIMBORAZO



FACULTAD DE INGENIERÍA

CARRERA DE INGENIERÍA AMBIENTAL

Proyecto de Investigación previo a la obtención del título de Ingeniera Ambiental

TRABAJO DE TITULACIÓN

Título del proyecto

**USO DE ALGORITMO DE INTELIGENCIA ARTIFICIAL PARA DESARROLLAR
UNA METODOLOGÍA PARA MEDIR LA TEXTURA DE LOS SUELOS**

Autor:

Estéfany Alejandra Erazo López

Tutor:

PhD. Víctor Julio García

Riobamba - Ecuador

Año 2019

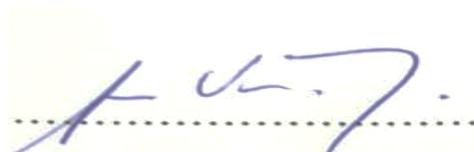
CERTIFICACION DEL TRIBUNAL

Los miembros del Tribunal de Graduación el proyecto de investigación de título: **“USO DE ALGORITMO DE INTELIGENCIA ARTIFICIAL PARA DESARROLLAR UNA METODOLOGÍA PARA MEDIR LA TEXTURA DE LOS SUELOS”**. Presentado por: Estéfany Alejandra Erazo López, dirigida por: PhD. Víctor García

Una vez escuchada la defensa oral y revisado el informe final del proyecto de investigación con fines de graduación escrito en la cual se ha constatado el cumplimiento de las observaciones realizadas, remite la presente para uso y custodia en la biblioteca de la Facultad de Ingeniería de la Universidad Nacional de Chimborazo.

Para constancia de lo expuesto firman:

PhD. Víctor García
Tutor del Proyecto de Investigación



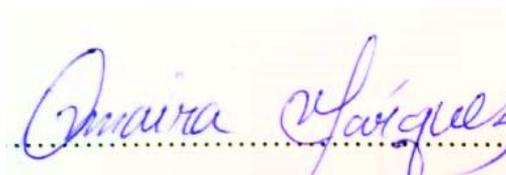
Firma

MsC. Patricia Andrade
Miembro del Tribunal



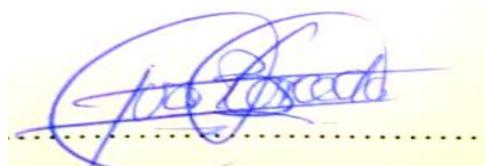
Firma

PhD. Omaira Márquez
Miembro del Tribunal



Firma

MsC. Juan Carlos Caicedo
Presidente del Tribunal

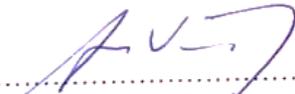


Firma

DECLARACIÓN EXPRESA DE TUTORÍA

Certifico que el presente trabajo de investigación previo a la obtención del Grado de INGENIERA AMBIENTAL, con el tema: **“USO DE ALGORITMO DE INTELIGENCIA ARTIFICIAL PARA DESARROLLAR UNA METODOLOGÍA PARA MEDIR LA TEXTURA DE LOS SUELOS”**, ha sido elaborado por ERAZO LOPEZ ESTÉFANY ALEJANDRA, el mismo que ha sido revisado y analizado en un cien por ciento con el asesoramiento permanente de mi persona en calidad de Tutor, por lo que se encuentra apta para su presentación y defensa respectiva.

Es todo cuanto puedo informar en honor a la verdad.


.....
PhD. Víctor García
C.I. 1757244122

AUTORÍA DE LA INVESTIGACIÓN

Yo, Estéfany Alejandra Erazo López con cédula de identidad 060488612-7, hago referencia como autora del presente trabajo de investigación, titulado: **“USO DE ALGORITMO DE INTELIGENCIA ARTIFICIAL PARA DESARROLLAR UNA METODOLOGÍA PARA MEDIR LA TEXTURA DE LOS SUELOS”**, dirigida por el tutor del proyecto, PhD. Víctor García.

Manifiesto la originalidad en la conceptualización de ideas, obtención de datos, interpretación de resultados, con el sustento de autores que se han referenciado debidamente en el documento.



.....
C.I. 060488612-7



AGRADECIMIENTO

Agradezco a Dios por siempre guiarme e iluminarme durante toda mi vida, y a mi familia que con su apoyo, consejos y valores me han ayudado a ser una mejor persona y a seguir adelante en todas las metas propuestas.

Al PhD. Víctor García, PhD. Omaira Márquez y a la MsC. Patricia Andrade por sus conocimientos, tiempo y guía durante el desarrollo del trabajo de investigación.

A la Universidad Nacional de Chimborazo, a la Carrera de Ingeniería ambiental, a los docentes por sus conocimientos y experiencias durante toda la formación profesional y a mis amigos por su amistad sincera y motivación.

DEDICATORIA

El presente trabajo dedico a mis padres Pedro y Rosario quienes con su apoyo incondicional, amor, me han inculcado que el respeto, responsabilidad, constancia, puntualidad, humildad y honestidad son indispensables para alcanzar el éxito en cada una de las metas de la vida, a mis abuelitas y tíos por siempre motivarme y a mi hermano Santiago por ayudarme en todo momento.

LISTA DE SIGLAS Y ABREVIATURAS

ERD: Espectro de reflectancia difusa

VIS: Visible (Luz)

NIR: Infrarrojo Cercano, por las siglas en inglés (Near Infrared)

MIR: Infrarrojo Medio, por las siglas en inglés (Mid Infrared)

A: Arcillosa

AL: Arcillo limosa

FAa: Franco arcillo arenosa

FA: Franco arcillosa

FAL: Franco arcillo limosa

Fa: Franco arenosa

F: Franca

FL: Franco limosa

aF: Areno francosa

L: Limosa

a: Arenosa

CONTENIDO

CERTIFICACION DEL TRIBUNAL.....	I
DECLARACIÓN EXPRESA DE TUTORÍA.....	II
AUTORÍA DE LA INVESTIGACIÓN	III
AGRADECIMIENTO	IV
DEDICATORIA.....	V
LISTA DE SIGLAS Y ABREVIATURAS.....	VI
RESUMEN	XI
ABSTRACT	XII
INTRODUCCIÓN.....	1
PLANTEAMIENTO DEL PROBLEMA.....	3
JUSTIFICACIÓN.....	4
OBJETIVOS.....	5
1. General	5
2. Específicos.....	5
ESTADO DEL ARTE	6
1. Textura.....	6
2. Clases texturales	7
3. El Método de Bouyoucos	8
4. Técnica tradicional	9
4.1. La estimación de la textura por el método del diagrama de flujo adaptado de Thien, 1979	10
5. Espectro de reflectancia difusa.....	11
5.1. Radiómetro.....	14
6. Definición de inteligencia artificial	14
6.1. Breve historia de la Inteligencia Artificial	15
6.2. Los procedimientos básicos de la inteligencia artificial	15

METODOLOGÍA	17
1. Preparación de muestras sintéticas y determinación de su textura en el laboratorio de suelos de la UNACH	17
2. Observables más relevantes y procesamiento	19
RESULTADOS Y DISCUSIÓN.....	21
1. Determinación de la textura con la técnica tradicional de diagrama de flujo propuesto por Thien.....	21
2. Caracterización de la textura usando el espectro de reflectancia difusa que permita el desarrollo del algoritmo.....	23
3. Estadística de los pares de árboles	25
4. Determinación de Atributos del árbol A1	29
CONCLUSIONES Y RECOMENDACIONES	35
1. Conclusiones	35
2. Recomendaciones	36
BIBLIOGRAFÍA.....	37
ANEXOS.....	39

ÍNDICE DE TABLAS

<i>Tabla 1.</i> Rango de los agregados arena, limo y arcilla.....	8
<i>Tabla 2.</i> Separados de suelos	9
<i>Tabla 3.</i> Determinación de la textura del suelo por propiedades organolépticas.....	10
<i>Tabla 4.</i> Porcentajes de cada uno de los agregados arena, limo y arcilla	17
<i>Tabla 5.</i> Comparación de textura con el Triángulo textural y la textura obtenida empleando el Diagrama de flujo de Thien.....	21
<i>Tabla 6.</i> Número de muestras erróneas de textura	23
<i>Tabla 7.</i> Características de los árboles	24
<i>Tabla 8.</i> Longitud de onda de las variables de importancia.....	24
<i>Tabla 9.</i> Tabla de tabulación cruzada.....	25
<i>Tabla 10.</i> Tabulación cruzada para comparar el desempeño del par de árboles A1 A2	25
<i>Tabla 11.</i> Tabulación cruzada para comparar el desempeño del par de árboles A1 A3	26
<i>Tabla 12.</i> Tabulación cruzada para comparar el desempeño del par de árboles A1 A4	26
<i>Tabla 13.</i> Tabulación cruzada para comparar el desempeño del par de árboles A1 A5	26
<i>Tabla 14.</i> Tabulación cruzada para comparar el desempeño del par de árboles A2 A3	26
<i>Tabla 15.</i> Tabulación cruzada para comparar el desempeño del par de árboles A2 A4	27
<i>Tabla 16.</i> Tabulación cruzada para comparar el desempeño del par de árboles A2 A5	27
<i>Tabla 17.</i> Tabulación cruzada para comparar el desempeño del par de árboles A3 A4	27
<i>Tabla 18.</i> Tabulación cruzada para comparar el desempeño del par de árboles A3 A5	27
<i>Tabla 19.</i> Tabulación cruzada para comparar el desempeño del par de árboles A4 A5	28
<i>Tabla 20.</i> Estadística de los pares de árboles en referencia al test estadístico de McNemar y precisión general.....	28
<i>Tabla 21.</i> Matriz de confusión para la etapa de prueba del árbol A1.....	29
<i>Tabla 22.</i> Matriz de confusión para la etapa de aprendizaje del árbol A1	30
<i>Tabla 23.</i> Tabulación de valores	30
<i>Tabla 24.</i> Fórmulas para atributos del árbol.....	31
<i>Tabla 25.</i> Atributos de los árboles de decisión	31
<i>Tabla 26.</i> Fase mineralógica identificada con la ayuda del árbol de decisión A1	32

ÍNDICE DE FIGURAS

<i>Figura 1.</i> Triángulo de textura.....	7
<i>Figura 2.</i> Método del diagrama de flujo adaptado de Thien, 1979.....	11
<i>Figura 3.</i> Detalle del espectro electromagnético visible e infrarrojo, regiones que han sido abordadas en el estudio de las propiedades del suelo.....	12
<i>Figura 4.</i> Diagrama de flujo del desarrollo de una metodología para medir la textura de los suelos	20
<i>Figura 5.</i> Árbol de decisión para la clasificación de la textura del suelo... ..	33
<i>Figura 6.</i> Curva de error del árbol A1 en función del número de nodos	34

RESUMEN

Esta investigación tiene como objetivo desarrollar una metodología para determinar la textura del suelo empleando el espectro de reflectancia difusa a través del uso del algoritmo de inteligencia artificial. Se utilizó el RadiometerFieldSPec 4 con el propósito de registrar el espectro de reflectancia difusa, el modelador estadístico y el Test de McNemar para posteriormente seleccionar el algoritmo con el mejor desempeño. Se generó 5 algoritmos (árboles de decisión) que tengan un menor número de nodos, a partir de estos se formaron 10 pares de árboles para elegir el árbol con menor porcentaje de error. El árbol A1 presentó un mejor desempeño con una precisión general de 87.50 %, el árbol A1 se obtuvo de segmentar el espectro de reflectancia difusa en un ancho de 10 nm, con esto se identificaron las variables de importancia. Los observables O 25, O 27, O 24, O 26 y O 29 están en un rango de 590 - 640 nm por la presencia de limo mientras que O 65, O 145 y O 149 están entre 990 – 1840 nm por el contenido de arena y O 177 y O 180 están entre 2110 – 2150 nm por la presencia de arcilla. Mediante la matriz de confusión para la etapa de prueba del árbol A1 se logró una precisión de 89.00 %, una sensibilidad de 92.00 %, especificidad de 71.00 % y una tasa de clasificación de incorrectas de 11.00 %. La metodología empleada para medir la textura del suelo con el uso del algoritmo de inteligencia artificial el error es de 13.60 % mientras que con la técnica tradicional de 33.33 %, logrando así obtener una menor incertidumbre y un porcentaje de error.

Palabras claves: espectro de reflectancia difusa, algoritmo de inteligencia artificial

Abstract

This research is intended to develop a methodology to determine soil texture using the spectrum of diffuse reflectance through the use of artificial intelligence algorithm. RadiometerFieldSpec 4 was used with the purpose of registering the diffuse reflectance spectrum, the statistical modeler and the McNemar test and then selecting the algorithm with the best performance. It generated 5 algorithms (decision trees) that have a less number of nodes, from these were formed 10 pairs of trees to choose the tree with the lowest percentage of error. The tree A1 presented a better performance with a general precision of 87.50 %, the tree A1 was obtained from segmenting the spectrum of diffuse reflectance in a width of 10 nm, with this, the variables of importance were identified . The observables 0 25, 0 27, 0 24, 0 26 and 0 29 are in a range of 590 - 640 nm by the presence silt while 0 65, 0 145 and 0 149 are between 990 - 1840 nm for the content of sand and 0 177 and 0 180 are between 2110 - 2150 nm by the presence of clay. Through the confusion matrix for the test stage of the tree A1 was achieved a precision of 89.00 %, a sensitivity of 92.00 %, specificity of 71.00 % and an incorrect classification rate of 11.00 %. The methodology used to measure the texture of the soil with the use of the artificial intelligence algorithm the error is 13.60 % while with the traditional technique of 33.33 %, achieving thus obtain a less uncertainty and a percentage of error.

Keywords: Spectrum of diffuse reflectance, artificial intelligence algorithm.



Reviewed by: Granizo, Sonia

Language Center Teacher

INTRODUCCIÓN

La textura es una de las propiedades físicas más importantes del suelo porque determina en gran parte la retención de agua y las propiedades de transmisión del suelo. La arena puede drenar rápidamente y en un suelo arcilloso los poros son demasiado pequeños para permitir un drenaje adecuado y donde las proporciones de arcilla y limo son elevadas, la penetración de la raíz resulta dificultosa. Los suelos con grandes áreas superficiales con abundancia de fracciones finas son: más activos químicamente, tienen más capacidad amortiguadora y también más capacidad de retención de agua (Argüello, 2008).

La textura del suelo se refiere a la proporción de componentes inorgánicos de diferentes formas y tamaños como arena, limo y arcilla, esta puede influir de manera importante en la funcionalidad del suelo, afectando la fertilidad, habilidad de retener agua, aireación, drenaje, contenido de materia orgánica, entre otras. Para su determinación se usan métodos oficiales de análisis, como es el caso del método del densímetro de Bouyoucos, aunque también se puede realizar de forma indirecta en campo, usando el diagrama de flujo basado en propiedades organolépticas y propuesto por Thien (1979).

La espectroscopia de reflectancia difusa, que se basa en la energía que refleja un material, es un método que se puede utilizar para reforzar o sustituir métodos convencionales de análisis de suelos (Xie, Pan, & Sun, 2012).

En el laboratorio de Suelos de la UNACH no se cuenta con un densímetro ni el equipo necesario para realizar la textura por el método de Bouyoucos, por lo cual, se recurre al método del diagrama de flujo propuesto por Thien, sin embargo, este método requiere que el investigador tenga experiencia en este tipo de pruebas organolépticas para determinar la textura del suelo de una manera asertiva.

En el laboratorio contamos con el equipo RadiometerFieldSPec 4 (Fabricante) que sirve para registrar el espectro de reflectancia difusa (ERD) desde cualquier superficie y en particular, cuidando las condiciones experimentales se puede decir que se registra la firma espectral del material cuya superficie interacciona con la fuente de luz. Cada longitud de onda porta información de moléculas en la superficie que interaccionan con la fuente luz; así el ERD es complejo de analizar usando técnicas convencionales, por tanto, en esta investigación se propone el desarrollo de una metodología que permita conocer la textura de una muestra de suelo a través del ERD.

Para lograr el propósito se usará muestras modelo, donde cada una se puede clasificar dentro de una de las doce clases texturales que conforman el triángulo textural. Esto mezclando en proporciones controladas la cantidad de arena, limo y arcilla. Posteriormente se registrará el ERD de cada muestra. Estos espectros se segmentarán en ventanas espectrales, cuyo valor medio se define como un observable que potencialmente puede ser un predictor de la textura de la muestra analizada.

La metodología empleada para identificar la textura es una opción frente a las técnicas convencionales como la de Bouyoucos, la tradicional (propiedades organolépticas), siendo así se identificará con la ayuda de un algoritmo de Inteligencia Artificial los predictores más relevantes que servirán para desarrollar un algoritmo que permita caracterizar la textura de una muestra de suelo. Con esta investigación se estaría determinando la textura a través de un método no destructivo, que puede ser usado en campo y que potencialmente elimina fuentes de incertidumbre, ahorra tiempo, dinero y puede conducir a resultados más ajustados a la realidad.

PLANTEAMIENTO DEL PROBLEMA

La textura del suelo es una propiedad física del suelo, de la cual dependen algunas características del suelo como la permeabilidad, la consistencia, la capacidad de intercambio de iones, de retención hídrica, distribución de poros, infiltración y estructura. Se establece tradicionalmente con la metodología de diagrama de flujo de textura propuesto por Thien, pero también se aplica la metodología de Bouyoucos o la de difracción laser, pero para estas técnicas se requiere experticia con el tacto, un mayor tiempo para la obtención de resultados e involucra una mayor inversión económica para la determinación de la textura, por lo cual este estudio se ha enfocado en el desarrollo de una metodología para medir la textura de los suelos mediante el uso de algoritmo de inteligencia artificial, a través de un método no destructivo, optimizando tiempo y dinero en la obtención de resultados confiables.

JUSTIFICACIÓN

La textura influye en la retención y capacidad de infiltración del agua o de un determinado contaminante, también en el riesgo de compactación, disponibilidad de nutrientes, en la aireación cuando se dificulta el intercambio de aire y permeabilidad que disminuye la productividad del suelo

Para la determinación de la textura se usan métodos de análisis, como es el caso del método del densímetro de Bouyoucos (fase de laboratorio) y el diagrama de flujo de Thien mediante la formación de una pequeña bola humedecida entre los dedos (con ayuda de una pequeña adición de agua), de este modo, cuanto más moldeable sea la bola, mayor proporción de arcilla tendrá mientras que cuanto menos moldeable sea la bola mayor fricción se nota entre las partículas, la proporción de arena será mayor. Así se pueden determinar las clases texturales y puede deducirse el contenido en las diversas fracciones (Gisbert, 2010).

La textura del suelo puede ser identificada tradicionalmente con la técnica que aplica propiedades organolépticas y hace uso del diagrama de Thien o la de Bouyoucos que involucran tiempo, recursos económicos, el uso de la materia prima, por tanto es importante emplear una metodología no destructiva, que pueda ser utilizada en campo y que tenga una mayor eficiencia, siendo así es importante desarrollar una metodología haciendo uso de un algoritmo de inteligencia artificial para medir la textura del suelo, empleando el espectro de reflectancia difusa, el modelador estadístico y el test estadístico de McNemar con el fin de obtener el algoritmo con mayor precisión.

OBJETIVOS

1. General

Desarrollar una metodología para medir la textura de los suelos mediante el uso de algoritmo de inteligencia artificial.

2. Específicos

1. Preparar muestras sintéticas y determinar la textura con la técnica tradicional de diagrama de flujo propuesto por Thien.
2. Caracterizar la textura usando el espectro de reflectancia difusa que permita el desarrollo del algoritmo.

ESTADO DEL ARTE

1. Textura

La textura del suelo hace referencia al tamaño de las partículas que lo componen. Las partículas se clasifican en varios grados de tamaño de partículas desde las más gruesas como las gravas, hasta la más fina la arcilla. La textura también influye en el movimiento de flujos de calor, agua y aire que acontecen tanto en el interior como en el exterior del perfil. En relación con la erosión hídrica, los suelos con alta proporción de limo tienden a ser muy erosionados (Argüello, 2008).

En los horizontes del suelo, la textura del suelo cambia, la textura del suelo depende del número de partículas inorgánicas que están presentes en el suelo. Estas han sido divididas en tres grupos en base a su tamaño: arcilla, limo y arena (FAO, 2017).

Los científicos del suelo pueden determinar la textura del suelo usando un triángulo de textura de suelos. Este es un ejercicio práctico donde se siente el suelo para averiguar su textura (FAO, 2017).

Los tres bloques de construcción de tierra (arena, limo y arcilla) se sienten muy diferentes y prestan diferentes propiedades a un suelo. Aunque los tres tipos de partículas de suelo se diferencian por su tamaño, que es una propiedad física, la cantidad relativa de cada uno de estos componentes tiene una gran influencia en las propiedades físicas, químicas y biológicas de un suelo. Los espacios de poros entre las partículas del suelo son en gran parte responsables de la cantidad de agua que puede contener un suelo. Las texturas más finas del suelo tienen una mayor superficie, poros de suelo más pequeños y una infiltración de agua más lenta en el perfil del suelo (Ritchey, McGrath, & Gehring, 2015).

La arena es la partícula más grande del suelo, mide de 0.05 a 2 mm de diámetro y puede verse a simple vista. La arena se siente arenosa al tacto y retiene muy poca agua. El limo es el componente de tamaño mediano del suelo, que mide 0.002 a 0.05 mm de diámetro. Las partículas de sedimento seco se sienten como la harina. Cuando está mojado, el limo se sentirá suave. El limo solo contiene una cantidad moderada de agua (Ritchey, McGrath, & Gehring, 2015).

La arcilla es la partícula más pequeña del suelo, mide menos de 0.002 mm de diámetro y solo se puede ver con microscopios potentes. La partícula de arcilla es 25 veces más pequeña que la partícula de limo y mil veces más pequeña que la partícula de arena. La arcilla se siente pegajosa cuando está húmeda y dura y quebradiza cuando está seca. La

arcilla puede contener mucha más agua que la arena o el cieno. En la mayoría de los suelos, el contenido de arcilla aumenta con la profundidad (Ritchey, McGrath, & Gehring, 2015).

2. Clases texturales

Las clases texturales son agrupamientos donde las propiedades dependientes de la textura presentan un comportamiento homogéneo (propiedades agronómicas homogéneas). Normalmente nos indica al grupo de partículas preponderantes. (Donald Gabriels, 2006).

Las clases texturales dentro del sistema del USDA son 12 y se disponen en un triángulo de textura como se puede apreciar en la *Figura 1*.

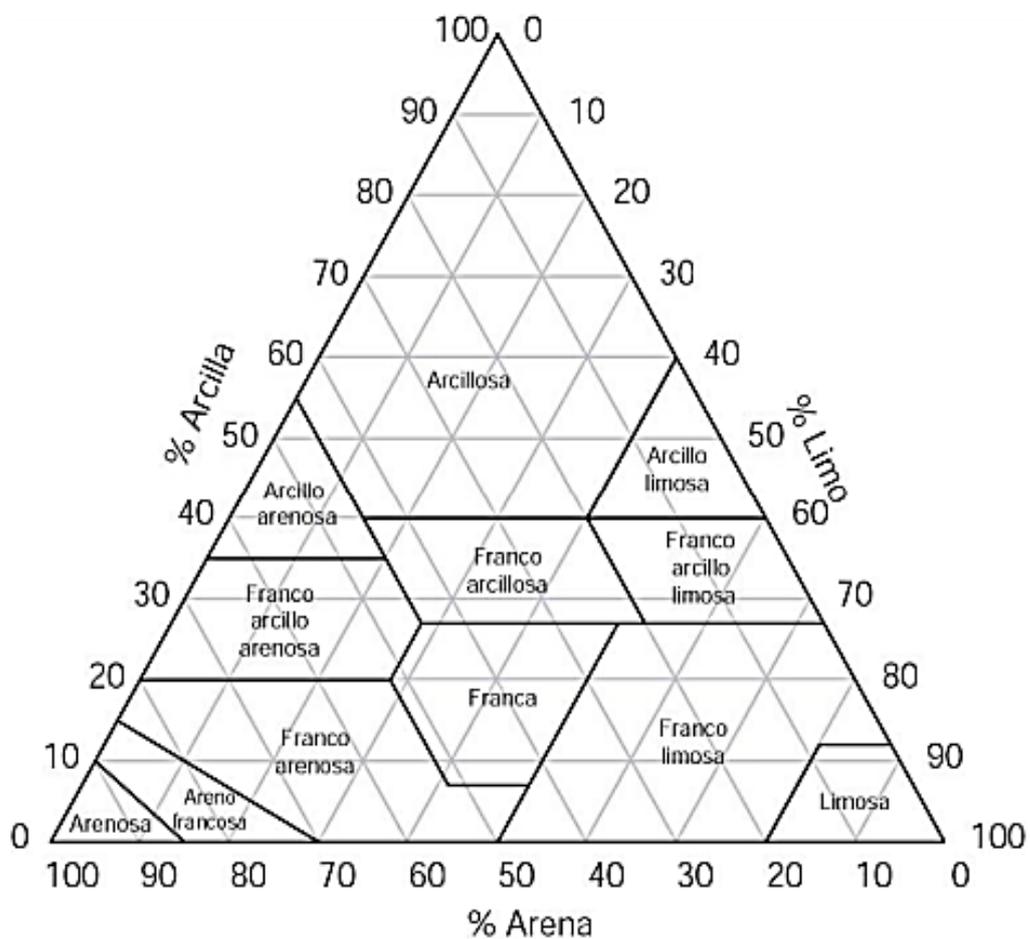


Figura 1. Triángulo de textura.

Fuente: (Donald Gabriels, 2006).

En la *Tabla 1* se puede observar el rango de cada agregado para la clase textural según el triángulo textural.

Tabla 1. Rango de los agregados arena, limo y arcilla.

Clase textural	Rango		
	Arcilla	Limo	Arena
	%	%	%
A	40-100	0-40	0-46
A a	35-55	0-20	46-100
AL	40-60	40-60	0-20
FAa	20-35	0-28	46-100
FA	27-40	15-56	20-46
FAL	27-40	60-73	0-20
Fa	15-20	0-50	44-100
F	8-27	28-50	24-52
FL	0-27	50-88	0-50
aF	10-15	0-30	70-85
L	0-12	80-100	0-20
a	0-10	0-12	86-100

Fuente: (Donald Gabriels, 2006).

3. El Método de Bouyoucos

El método del Bouyoucos emplea un hidrómetro. Este instrumento mide la disminución de la densidad de la suspensión debido al asentamiento de las partículas del suelo. Un pretratamiento es necesario para separar o dispersar los agregados del suelo en las partículas principales de arena, limo y arcilla (Maldonado, 2016).

En este procedimiento se usa una solución química como un dispersante. Este método, permitirá tener una adecuada información cuantitativa sobre los porcentajes de arena, limo y arcilla presentes en el suelo. Las partículas suspendidas en el agua se asientan diferencialmente dependiendo de la cantidad de superficie por unidad de volumen (Maldonado, 2016).

Las partículas de arcilla tienen una gran área superficial por unidad de volumen y se asientan lentamente, mientras que las partículas de arena se asientan rápidamente debido a su baja superficie específica. En el método de Bouyoucos es determinada usando el hidrómetro para medir la densidad de la suspensión (Cano, 2013).

4. Técnica tradicional

La técnica tradicional que se utiliza es el método del diagrama de flujo propuesto por Thien (1979), debido a que el método de Bouyoucos se requiere el equipo densímetro con el cual no se cuenta.

En la Tabla 2 se muestran los tamaños de partículas para separados de suelo Arcilla, Limo y Arena.

Tabla 2. Separados de suelos

Separados de suelo	Diámetro(mm)
Arcilla	<0.002
Limo	0.002 a 0.05
Arena	0.05 a 2.0
Arena muy fina	0.05 a 0.10
Arena fina	0.10 a 0.25
Arena media	0.25 a 0.5
Arena gruesa	0.5 a 1.0
Arena muy gruesa	1.0 a 2.0

Fuente: (Ritchey, McGrath, & Gehring, 2015).

4.1. La estimación de la textura por el método del diagrama de flujo adaptado de Thien, 1979

A continuación en la *Tabla 3* se detalla los pasos que se emplean en el diagrama de flujo de Thien, que permiten determinar la textura del suelo utilizando como herramienta el tacto.

Tabla 3. Determinación de la textura del suelo por propiedades organolépticas

Paso 1	Paso 2	Paso 3
		
<p>Comience con un puñado pequeño de tierra, aproximadamente del tamaño de una pelota de golf, y agregue lentamente una gota de agua a la vez, mezclándola a medida que avanza, hasta que tenga una bola de tierra que tenga la consistencia de una masilla. Aprieta suavemente la bola para determinar si se mantendrá junta en una bola o se romperá.</p>	<p>Si la bola de tierra permanece intacta, presiona suavemente la bola entre el pulgar y el índice, tratando de formar una cinta. Si puede formar una cinta, mida cuánto tiempo tiene la cinta antes de que se deshaga.</p>	<p>Después de completar la prueba de la cinta, agregue agua a una pizca de tierra en la palma de su mano hasta que tenga un charco fangoso. Frote el charco de lodo contra su palma y determine si se siente arenoso, suave o igualmente arenoso y suave.</p>

Fuente: (Ritchey, McGrath, & Gehring, 2015).

En la Figura 2 se observa el diagrama de flujo que ayuda a determinar a que clase textural pertenece la muestra de suelo.

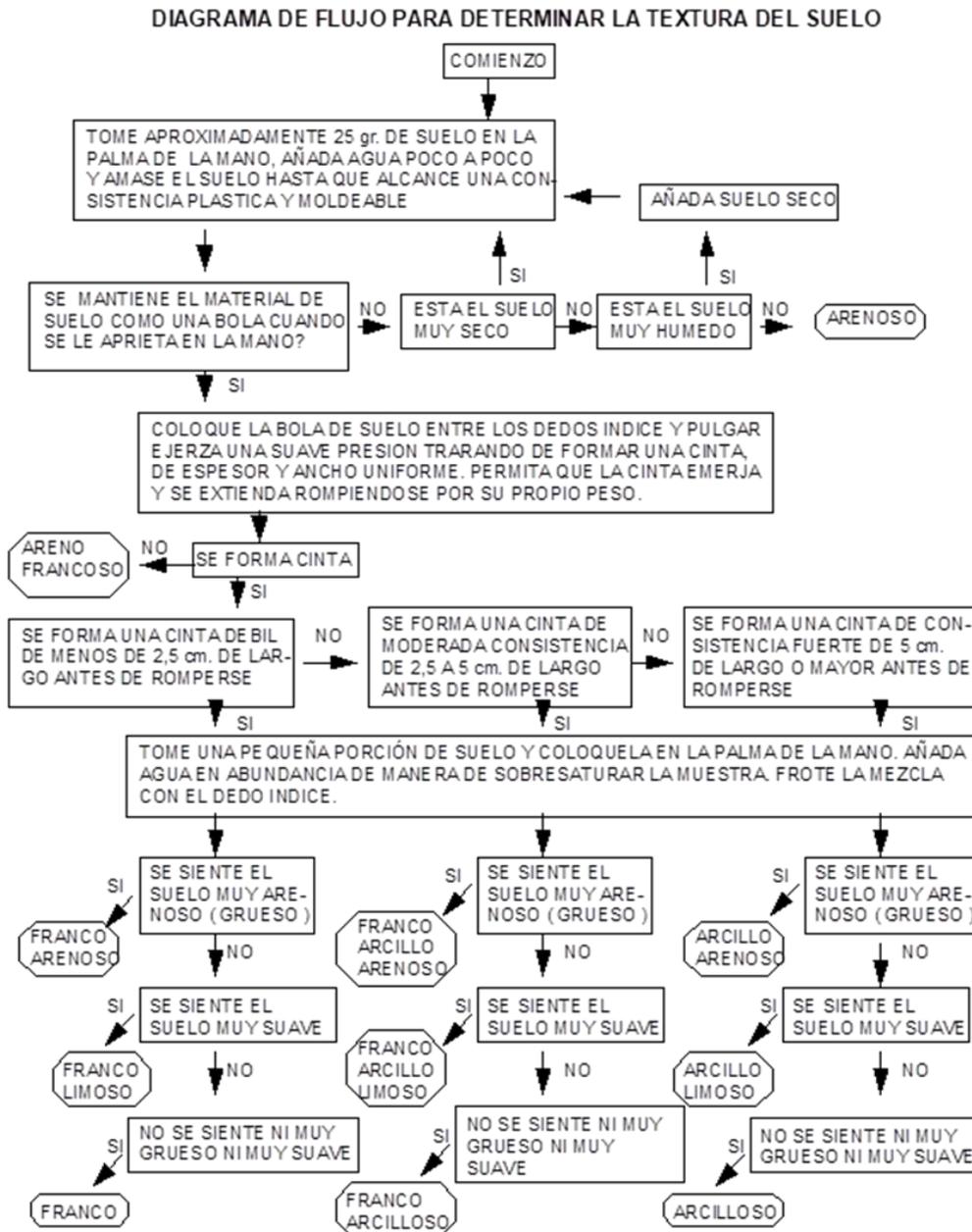


Figura 2. Método del diagrama de flujo adaptado de Thien, 1979.

Fuente: (Ritchey, McGrath, & Gehring, 2015).

5. Espectro de reflectancia difusa

La espectrometría de reflectancia es una técnica analítica utilizada desde principios del siglo XX por químicos y mineralogistas para la identificación de ciertos compuestos y minerales.

Todo cuerpo que sea sometido a efectos de radiación, como un haz de luz, experimenta un fenómeno de reflexión y absorción de energía (Páez, 2008).

El término espectro, y no es más que una gráfica continua con cambios de pendiente en dos dimensiones, donde los valores en el eje horizontal corresponden a la longitud de onda, que puede ser expresada en micrómetros(μm) o nanómetros(nm), y los valores del eje vertical al porcentaje de reflectancia (Páez, 2008).

En la espectroscopia aplicada al estudio de suelos, la percepción (próxima y remota) se refiere a la adquisición de información de un objeto, mediante pruebas no destructivas y sin entrar en contacto directo con él, a través de sensores de laboratorio, de campo, aéreos o ubicados en la órbita media o alta de la atmósfera terrestre. Se trata entonces de un proceso físico donde el suelo absorbe energía y refleja parte de la misma que incide sobre él. La absorción de la energía por el suelo es dada en función de los materiales que lo componen, los cuales reflejan energía a diferentes longitudes de onda. La energía reflejada por el suelo en las diferentes longitudes de onda es relacionada como una función de las propiedades del suelo químicas, físicas, mineralógicas y biológicas (Tamayo, 2013).

En este caso de detección espectral, la energía electromagnética funciona como un agente que interactúa con los componentes del suelo, de forma que cada atributo tiene una respuesta característica de reflectividad - absortividad en cada región espectral como se puede notar en la Figura 3.

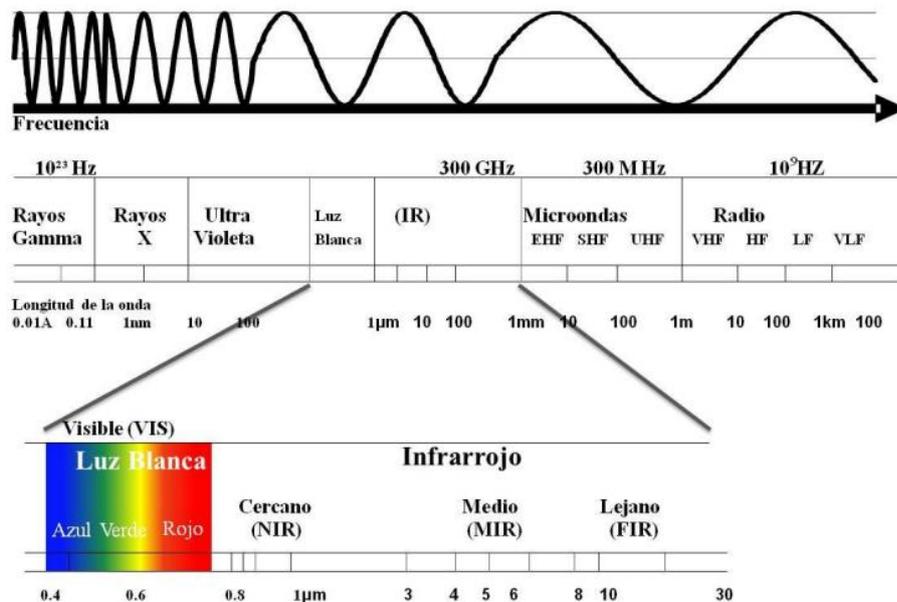


Figura 3. Detalle del espectro electromagnético visible e infrarrojo, regiones que han sido abordadas en el estudio de las propiedades del suelo.

Fuente: (Tamayo, 2013).

La espectroscopía de reflectancia estudia la energía que refleja un material, como parte del fraccionamiento de la energía incidente, en función de la longitud de onda. (Tamayo, 2013)

Aunque las bandas fundamentales de vibración se encuentran principalmente en regiones del infrarrojo medio y lejano, los procesos vibratorios producen características en la NIR debido a la excitación de armónicos y la combinación de los tonos de los periodos fundamentales de grupos aniónicos (Tamayo, 2013).

Para el caso específico de la determinación de textura del suelo, la influencia en el VIS-NIR se produce como la intensidad de la reflectancia sobre la firma espectral (albedo), pues la mayoría de los suelos con contenido de arcilla presentan una menor reflexión de la energía electromagnética incidente (Bellinaso, 2010).

Además, los minerales de arcilla tales como filosilicatos y óxidos, hidróxidos y oxihidróxidos de hierro y aluminio presentes en el suelo son los responsables de las características de absorción del espectro de reflectancia por VIS-NIR (Stenberg, 2010).

La respuesta espectral de cuarzo, mineral abundante en la fracción arena en el VIS-NIR, se caracteriza por un albedo mayor y ausencia de características de absorción, que se observa por lo general en el espectro de los suelos arenosos (Bellinaso, 2010).

La espectroscopía por reflexión difusa es una técnica de análisis directa y no destructiva y la reflexión difusa, es aquella temporalmente absorbida y luego re-emitida por la muestra y entrega información útil acerca de la naturaleza o composición de la muestra (Falcó, 2012).

Esta técnica presenta un alto grado de desarrollo en las últimas dos décadas, al superar algunas limitaciones y el perfeccionamiento de metodologías estadísticas robustas que consiguen relacionar de una forma más precisa la variabilidad del espectro con la variabilidad de las características del suelo (Camacho, 2013).

Para el caso específico de la determinación de textura del suelo, la influencia en el VIS-NIR se produce como la intensidad de la reflectancia sobre la firma espectral (albedo), pues la mayoría de los suelos con contenido de arcilla presentan una menor reflexión de la energía electromagnética incidente (Bellinaso, 2010).

La espectroscopia en las franjas del visible (VIS), infrarrojo cercano (NIR) e infrarrojo medio (MIR) permiten una rápida adquisición de la información de distintas variables suelo (arena, limo, arcilla) (Ben-Dor, Heller, & Chudnovsky, 2008).

5.1. Radiómetro

Un Radiómetro es un equipo usado para medir la intensidad de energía radiante. La mayoría de los radiómetros utilizan simplemente fotocélulas como sensores. Para poder medir la radiación emitida de un espectro específico o para incorporar el radiómetro dentro de cierta respuesta espectral, se utiliza normalmente un filtro óptico. Este filtro óptico ofrece una solución simple y de bajo costo (Minolta, 2018).

Las aplicaciones industriales del radiómetro principalmente son para medir irradiación y radiación. Para poder cuantificar la emisión de radiación de una fuente, se realizan normalmente mediciones de radiación. Por otro lado, cuando el nivel de exposición es el problema, se puede medir la irradiación o irradiación integrada (Minolta, 2018).

Aplicaciones de un Radiómetro

El Radiómetro es comúnmente usado en la industria para cuantificar la luz que está fuera del espectro visible, por ejemplo, ultravioleta e infrarroja. La luz ultravioleta (UV) es ampliamente utilizada en la industria para varias aplicaciones, por ejemplo (Minolta, 2018):

- Curado o fotorresistencia en la fabricación de semiconductores
- Curado de emulsiones para impresión o elaboración de placas
- Ensayos de solidez del color

Para realizar mediciones UV con un radiómetro, sea mediciones de radiación o irradiación, se debe especificar la respuesta espectral (longitud de onda y pico de longitud de onda), según la aplicación específica (Minolta, 2018).

6. Definición de inteligencia artificial

En el ámbito de las llamadas ciencias computacionales o ciencias informáticas, son múltiples y muy variadas las definiciones de la inteligencia artificial. A título de ejemplo mostramos varias de ellas (Martínez, 2012):

- La creación de máquinas que pueden pensar
- La construcción de máquinas que puedan ejecutar tareas que requieren inteligencia humana

- El arte de desarrollar máquinas con capacidad para realizar funciones que, cuando son realizadas por personas, requieren de inteligencia.
- La creación de herramientas de hardware y software capaces de desarrollar comportamientos inteligentes.

6.1. Breve historia de la Inteligencia Artificial

El nacimiento de la IA como disciplina de investigación se remonta a 1956, durante una conferencia sobre informática teórica que tuvo lugar en el Dartmouth College. A esa conferencia asistieron algunos de los científicos que posteriormente se encargaron de desarrollar la disciplina en diferentes ámbitos y de dotarla de una estructura teórica y computacional apropiada. Entre los asistentes estaban John McCarthy, Marvin Minsky, Allen Newll y Hebert Simon (Benitez, Escudero, Kanaan, & Rodó, 2014).

La inteligencia artificial es una disciplina académica relacionada con la teoría de la computación cuyo objetivo es emular algunas de las facultades humanas en sistemas artificiales. Con la inteligencia humana nos referimos típicamente a procesos de percepción sensorial (visión, audición, etc.) y a sus consiguientes procesos de reconocimiento de patrones (Benitez, Escudero, Kanaan, & Rodó, 2014).

6.2. Los procedimientos básicos de la inteligencia artificial

Estrictamente hablando, los objetivos que se proponen, a largo plazo, la Inteligencia Artificial tienden a desarrollar la tecnología adecuada para proporcionar a los ordenadores una capacidad de razonar o de discernir similar a las de la inteligencia humana. A pesar de la expectación que despertó desde sus comienzos esa disciplina, cuyo potencial de desarrollo y eficacia parecía irresistible en muchos campos (Martínez, 2012).

Si analizamos el abordaje utilizado desde los inicios para el desarrollo de la Inteligencia Artificial, vemos que se ha basado en intentar que los ordenadores lleguen a razonar. Con el fin de dotar a las computadoras de una capacidad de razonar, y resolver problemas, que fuese cada vez más eficaz, se han intentado dos posibles vías o procedimientos generales: algoritmos y las soluciones heurísticas (Martínez, 2012).

El diseño de un sistema de inteligencia artificial normalmente requiere la utilización de herramientas de disciplinas muy diferentes como el cálculo numérico, la estadística, la informática, el procesado de señales, el control automático, la robótica o la neurociencia. Por

este motivo, pese a que la inteligencia artificial se considera una rama de la informática teórica, es una disciplina en la que contribuyen de forma activa numerosos científicos, técnicos y matemáticos (Benitez, Escudero, Kanaan, & Rodó, 2014).

Un sistema de inteligencia artificial requiere de una secuencia finita e instrucciones que especifique las diferentes acciones que ejecuta la computadora para resolver un determinado problema esta secuencia de instrucciones constituye la estructura algorítmica del sistema de inteligencia artificial. Se conoce como método efectivo o algoritmo al procedimiento para encontrar la solución a un problema mediante la reducción del mismo a un conjunto de reglas (Benitez, Escudero, Kanaan, & Rodó, 2014).

En ocasiones, los sistemas de IA resuelven problemas de forma heurística mediante un procedimiento de ensayo y error que incorpora información relevante basada en conocimientos previos. En un sistema IA, en cambio, las instrucciones básicas son las propias de una computadora, es decir operaciones aritmético-lógicas, de lectura/escritura de registros y de control de flujo secuencial. En muchas ocasiones, los sistemas de IA constan de dos fases, una primera fase de aprendizaje y una segunda de predicción (Benitez, Escudero, Kanaan, & Rodó, 2014).

En la fase de aprendizaje es un conjunto de datos representativo de aquellas situaciones que se desea analizar, de forma que el sistema IA aprende las características fundamentales de los datos y es capaz de generalizar su estructura. El sistema de IA modularía la ejecución de las diferentes acciones, incrementando la probabilidad de aquellas que presenten un mejor rendimiento y restringiendo las que no comporten una mejora de la función objetivo (Benitez, Escudero, Kanaan, & Rodó, 2014).

METODOLOGÍA

1. Preparación de muestras sintéticas y determinación de su textura en el laboratorio de suelos de la UNACH

- a. Se recolectó la materia prima para la preparación de muestras sintéticas: arena, limo y arcilla.
 1. Se recolectó 5 Kilogramos de cada uno de los agregados arena, limo y arcilla
 2. Se secó la materia prima para evitar el efecto del agua en la medición del espectro de reflectancia difusa.
 3. Se etiquetó cada uno de los agregados (arena, limo y arcilla)
- b. Se preparó las muestras representativas de cada grupo textural
 1. Se determinó los porcentajes de cada uno de los agregados arena, limo y arcilla y se obtuvo 36 muestras sintéticas, empleado el triángulo textural *Figura.1*
 2. En cada una de las 36 muestras se pesó el porcentaje establecido de cada agregado dando un total de 100 gramos.
 3. Se etiquetó cada una de las 36 muestras.
 4. Se homogenizó cada una de las muestras de suelo previo al registro del ERD.

En la *Tabla 4* se encuentran los porcentajes de los agregados arena, arcilla y limo para cada una de las muestras sintéticas de suelo de cada grupo textural.

Tabla 4. Porcentajes de cada uno de los agregados arena, limo y arcilla

MUESTRAS		% ARENA	% ARCILLA	% LIMO	
1	Arcillosa (A)	m1	20	60	20
		m2	20	70	10
		m3	10	70	20
2	Arcillo arenosa (Aa)	m4	54	40	6
		m5	50	40	10
		m6	50	44	6
3	Arcillo limosa (AL)	m7	6	50	44
		m8	10	44	46
		m9	4	46	50
4	Franco arcillo arenosa (FAa)	m10	64	26	10
		m11	50	30	20
		m12	60	30	10
5	Franco arcillosa (FA)	m13	36	34	30
		m14	30	36	34
		m15	24	36	40

Tabla 4. Continuación

6	Franco arcillo limosa (FAL)	m16	10	34	56
		m17	4	32	64
		m18	14	36	50
7	Franco arenosa (Fa)	m19	66	10	24
		m20	70	10	20
		m21	60	10	30
8	Franca (F)	m22	40	20	40
		m23	46	20	34
		m24	44	16	40
9	Franco limosa (FL)	m25	20	14	66
		m26	20	20	60
		m27	20	10	70
10	Areno francosa (aF)	m28	84	6	10
		m29	76	4	20
		m30	80	6	14
11	Limosa (L)	m31	8	6	86
		m32	10	8	82
		m33	6	4	90
12	Arenosa (a)	m34	92	4	4
		m35	92	4	6
		m36	92	2	6

Fuente: El autor

c. Se determinó la textura usando las propiedades organolépticas.

Se tomó 25 gramos de cada una de las 36 muestras sintéticas y mediante el uso de las propiedades organolépticas y el Diagrama de flujo de Thien *Figura 2* se identificó la textura de cada una de las 36 muestras sintéticas.

d. Se extrajo 3 submuestras de cada grupo textural dando un total de 36 muestras.

e. Se registró el ERD y se realizó el pretratamiento del ERD.

Para el registro del ERD se realizó rotando 90°, hasta obtener 4 ángulos de cada una de las 36 muestras sintéticas dando un total de 1440 datos en cada una de las matrices de 10, 20, 30, 40 y 50 nm.

2. Observables más relevantes y procesamiento

- a) Se segmentó el espectro en bandas espectrales de 10, 20, 30, 40 y 50 nm.

Los valores obtenidos del registro de ERD se segmentó para posteriormente la construcción de la matriz observable.

- b) Se calculó los valores medios de cada banda y ensambló la matriz de observables.

Para realizar los cálculos y el establecimiento de la matriz observable se empleó Microsoft Excel como herramienta.

- c) Se entrenó, validó e identificó las observables más relevantes.

Para la identificación de observables más relevantes o variables de importancia se empleó un modelador estadístico.

- d) Se identificó el árbol de decisión que conduce al menor error de predicción y que posee el menor número de nodos.

Se optimizó el algoritmo de inteligencia artificial en términos de número de nodos y el menor error de predicción tanto de entrenamiento como de validación.

- e) Se estableció el algoritmo con mejor desempeño aplicando el test estadístico de McNemar e identificando la mayor precisión general de los pares de árboles que no poseen diferencia significativa.

En el siguiente diagrama de la Figura 4 se puede observar la metodología empleada para medir la textura del suelo empleando el diagrama de flujo de Thien y el espectro de reflectancia difusa.

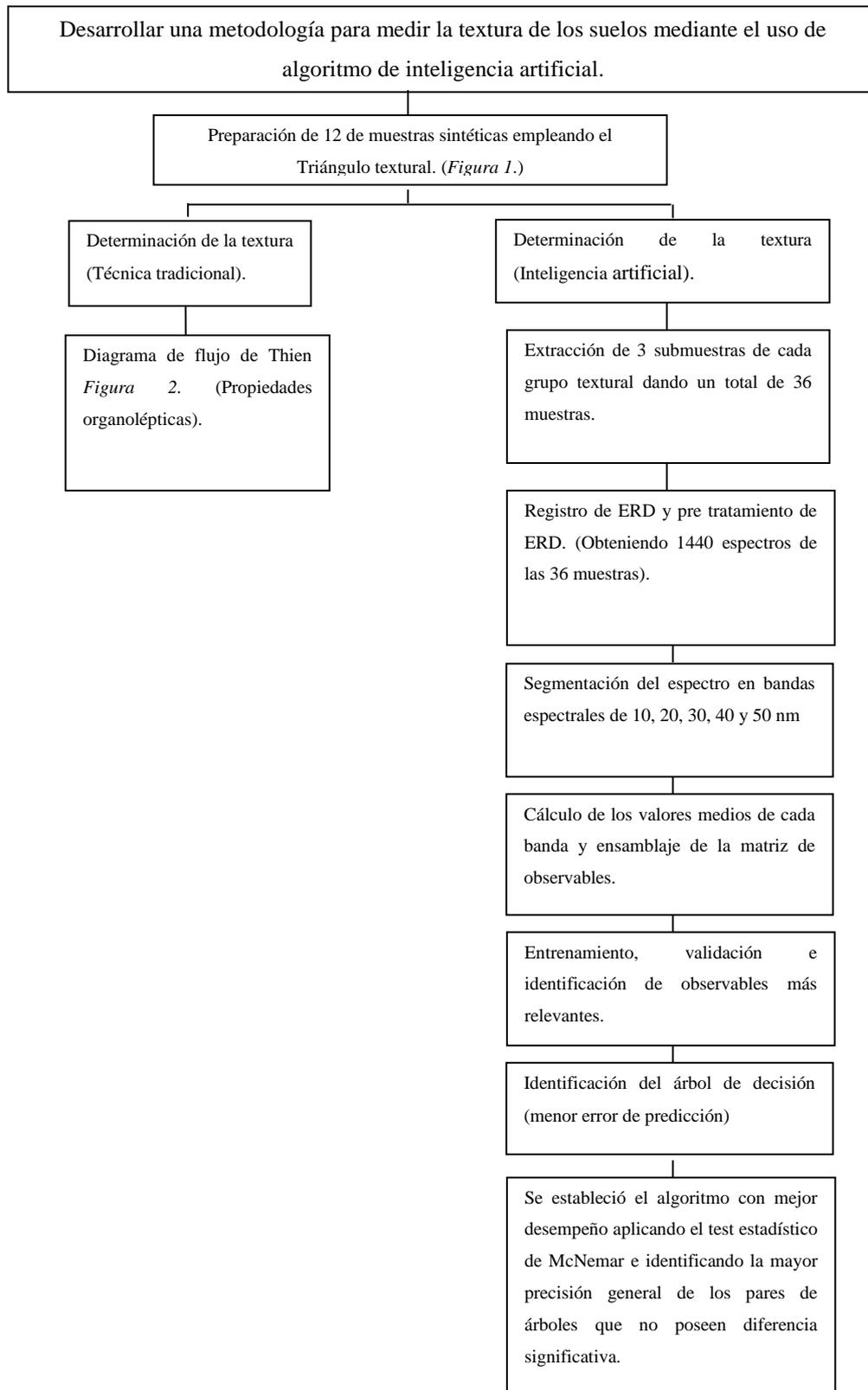


Figura 4. Diagrama de flujo del desarrollo de una metodología para medir la textura de los suelos.

Fuente: El autor

RESULTADOS Y DISCUSIÓN

1. Determinación de la textura con la técnica tradicional de diagrama de flujo propuesto por Thien.

Mediante el empleo del diagrama de flujo propuesto por Thien se identificó la textura de cada una de las 36 muestras sintéticas de las 12 clases texturales: Arcillosa, Arcillo arenosa, Arcillo limosa, Franco arcillo arenosa, Franco arcillosa, Franco arcillo limosa, Franco arenosa, Franca, Franco limosa, Areno francosa, Limosa y Arenosa. Los resultados de la textura con el uso de las propiedades organolépticas y el diagrama de flujo de Thien se detallan en la Tabla 5, en donde según el triángulo textural la muestra 4 pertenece a Franco arcillo arenosa, la 5 a Franco arcillosa, la 11 a limosa y la 12 a arenosa son las erróneas, obteniendo una precisión para medir la textura del suelo de 66.67 %.

En la *Tabla 5* se compara la textura según la clasificación del Triángulo textural y la textura obtenida con el análisis de las muestras de suelo, utilizando el Diagrama de flujo de Thien.

Tabla 5. Comparación de textura con el Triángulo textural y la textura obtenida empleando el Diagrama de flujo de Thien

MUESTRAS		Textura (Diagrama de flujo de Thien)	Textura (Triangulo textural)
1	m1	Arcilloso	Arcillosa (A)
	m2	Arcilloso	
	m3	Arcilloso	
2	m4	Arcillo arenosa	Arcillo arenosa (Aa)
	m5	Arcillo arenosa	
	m6	Arcillo arenosa	
3	m7	Arcillo limoso	Arcillo limosa (AL)
	m8	Arcillo limoso	
	m9	Arcillo limoso	
4	m10	Franco	Franco arcillo arenosa (FAa)
	m11	Franco	
	m12	Franco	
5	m13	Franco arcillo Limoso	Franco arcillosa (FA)
	m14	Franco arcillo Limoso	
	m15	Franco arcillo Limoso	

Tabla 5. Continuación

6	m16	Franco arcillo limoso	Franco arcillo limosa (FAL)
	m17	Franco arcillo limoso	
	m18	Franco arcillo limoso	
7	m19	Franco arenoso	Franco arenosa (Fa)
	m20	Franco arenoso	
	m21	Franco arenoso	
8	m22	Franco	Franca (F)
	m23	Franco	
	m24	Franco	
9	m25	Franco limoso	Franco limosa (FL)
	m26	Franco limoso	
	m27	Franco limoso	
10	m28	Areno francoso	Areno francosa (aF)
	m29	Areno francoso	
	m30	Areno francoso	
11	m31	Franco limoso	Limosa (L)
	m32	Franco limoso	
	m33	Franco limoso	
12	m34	Areno francoso	Arenosa (a)
	m35	Areno francoso	
	m36	Areno francoso	

Fuente: El autor

En la *Tabla 6* se presenta las muestras obtenidas con el diagrama de flujo de Thien y las propiedades organolépticas, las cuales no coinciden con el triángulo textural.

Tabla 6. Número de muestras erróneas de textura

MUESTRAS		Textura (Diagrama de flujo de Thien)	Textura (Triangulo textural)
4	m10	Franco	Franco arcillo arenosa (FAa)
	m11	Franco	
	m12	Franco	
5	m13	Franco arcillo Limoso	Franco arcillosa (FA)
	m14	Franco arcillo Limoso	
	m15	Franco arcillo Limoso	
11	m31	Franco limoso	Limosa (L)
	m32	Franco limoso	
	m33	Franco limoso	
12	m34	Areno francoso	Arenosa (a)
	m35	Areno francoso	
	m36	Areno francoso	

Fuente: El autor

Como se puede observar en la tabla se compara la textura según el triángulo textural y el diagrama de flujo de las 36 muestras sintéticas. El porcentaje de error empleando esta técnica tradicional del diagrama de flujo de Thien y las propiedades organolépticas para obtener la textura es del 33.33 %, el cual representa 12 muestras que pertenecen a los siguientes grupos texturales: 3 a Franco arcillo arenosa, 3 a Franco arcillosa, 3 a Limosa y 3 a Arenosa de las 36 analizadas

2. Caracterización de la textura usando el espectro de reflectancia difusa que permita el desarrollo del algoritmo.

Con el espectro de reflectancia difusa se calculó los valores medios de cada banda espectral de 10, 20, 30, 40 y 50 nm y se ensambló la matriz de observables de cada banda con 1440 espectros y variables a partir de los cuales con la utilización del modelador estadístico se obtuvo una matriz de confusión tanto de prueba como de aprendizaje y 5 árboles de decisión con un determinado porcentaje de precisión y la menor cantidad de nodos como se puede observar en la *Tabla 7*.

Tabla 7. Características de los árboles

Árbol	Banda espectral	Número de nodos	Precisión general %
1	10	23	87.50
2	20	20	86.60
3	30	23	85.56
4	40	23	86.74
5	50	23	84.72

Fuente: El autor

En la *Tabla 8* se puede observar la longitud de onda para cada uno de los árboles seleccionados.

Tabla 8. Longitud de onda de las variables de importancia

A1	Longitud de onda/nm	A2	Longitud de onda/nm	A3	Longitud de onda/nm	A4	Longitud de onda/nm	A5	Longitud de onda/nm
O 25	590	O 14	480	O 10	440	O 8	420	O 10	440
O 27	610	O 12	460	O 9	430	O 9	430	O 6	400
O 24	580	O 17	510	O 11	450	O 7	410	O 5	390
O 26	600	O 13	470	O 12	460	O 10	440	O 7	410
O 29	630	O 15	490	O 8	420	O 6	400	O 8	420
O 65	990	O 11	450	O 13	470	O 17	510	O 28	620
O 177	2110	O 75	1090	O 46	800	O 37	710	O 11	450
O 145	1790	O 73	1070	O 49	830	O 13	470	O 30	640
O 180	2140	O 31	650	O 61	950	O 29	630	O 14	480
O 149	1830	O 32	660	O 60	940	O 18	520	O 21	550

Fuente: El autor

3. Estadística de los pares de árboles

Test estadístico (McNemar)

En la *Tabla 9* se hallan las variables para comparar el desempeño de los árboles.

Tabla 9. Tabla de tabulación cruzada

	I	C	
I	n_{00}	n_{01}	n_{00} : Número de datos incorrectos para los algoritmos A_x y A_y
C	n_{10}	n_{11}	n_{01} : Número de datos incorrectos para el algoritmo A_x pero no para A_y
			n_{10} : Número de datos incorrectos para el algoritmo A_y pero no para A_x
			n_{11} : Número de datos incorrectos para ningún algoritmo A_x ni A_y

Fuente: (Dietterich, 1997)

La siguiente estadística se distribuye con 1 grado de libertad, esta incorpora un término de corrección de continuidad en el numerador (-1)

$$x^2 = \frac{(|n_{01} - n_{10}| - 1)^2}{n_{01} + n_{10}} \quad (1)$$

Fuente: (Dietterich, 1997)

Hipótesis nula: H_0 : No hay diferencia significativa entre el desempeño de los dos algoritmos (Árboles de decisión)

Si $x^2_{1,0.95}$ es mayor que 3.84 la hipótesis nula se rechaza y por lo tanto los algoritmos tienen una diferencia significativa mientras que si $x^2_{1,0.95}$ es menor que 3.84 la hipótesis nula se acepta puesto que no hay una diferencia significativa.

En las siguientes Tablas 10, 11, 12, 13, 14, 15, 16, 17, 18 y 19 muestran el desempeño de pares de árboles en base a la validación de los datos correctos e incorrectos, para identificar el árbol que posee el mejor desempeño con la utilización de la Ecuación (1) de Chi cuadrado.

Tabla 10. Tabulación cruzada para comparar el desempeño del par de árboles A1 A2

A1 A2	I	C	TOTAL
I	114	41	155
C	149	1136	1285
TOTAL	263	1177	1440
Test de McNemar: $x^2 = 60.26$			

Fuente: El autor

Tabla 11. Tabulación cruzada para comparar el desempeño del par de árboles A1 A3

A1 A3	I	C	TOTAL
I	132	23	155
C	36	1249	1285
TOTAL	168	1272	1440
Test de McNemar: $x^2 = 2.44$			

Fuente: El autor

Tabla 12. Tabulación cruzada para comparar el desempeño del par de árboles A1 A4

A1 A4	I	C	TOTAL
I	141	14	155
C	44	1241	1285
TOTAL	185	1255	1440
Test de McNemar: $x^2 = 14.50$			

Fuente: El autor

Tabla 13. Tabulación cruzada para comparar el desempeño del par de árboles A1 A5

A1 A5	I	C	TOTAL
I	141	14	155
C	46	1239	1285
TOTAL	187	1253	1440
Test de McNemar: $x^2 = 16.02$			

Fuente: El autor

Tabla 14. Tabulación cruzada para comparar el desempeño del par de árboles A2 A3

A2 A3	I	C	TOTAL
I	131	132	263
C	37	1140	1177
TOTAL	168	1272	1440
Test de McNemar: $x^2 = 52.28$			

Fuente: El autor

Tabla 15. Tabulación cruzada para comparar el desempeño del par de árboles A2 A4

A2 A4	I	C	TOTAL
I	126	137	263
C	59	1118	1177
TOTAL	185	1255	1440
Test de McNemar: $\chi^2 = 30.25$			

Fuente: El autor

Tabla 16. Tabulación cruzada para comparar el desempeño del par de árboles A2 A5

A2 A5	I	C	TOTAL
I	133	130	263
C	54	1123	1177
TOTAL	187	1253	1440
Test de McNemar: $\chi^2 = 30.57$			

Fuente: El autor

Tabla 17. Tabulación cruzada para comparar el desempeño del par de árboles A3 A4

A3 A4	I	C	TOTAL
I	145	23	168
C	40	1232	1272
TOTAL	185	1255	1440
Test de McNemar: $\chi^2 = 4.06$			

Fuente: El autor

Tabla 18. Tabulación cruzada para comparar el desempeño del par de árboles A3 A5

A3 A5	I	C	TOTAL
I	145	23	168
C	42	1230	1272
TOTAL	187	1253	1440
Test de McNemar: $\chi^2 = 4.98$			

Fuente: El autor

Tabla 19. Tabulación cruzada para comparar el desempeño del par de árboles A4 A5

A4 A5	I	C	TOTAL
I	173	12	185
C	14	1241	1255
TOTAL	187	1253	1440
Test de McNemar: $\chi^2 = 0.038$			

Fuente: El autor

En la Tabla 20 se identifica cual árbol tiene el mejor desempeño según el test estadístico de McNemar y la precisión general de los pares de árboles que no poseen diferencia significativa.

Tabla 20. Estadística de los pares de árboles en referencia al test estadístico de McNemar y precisión general

Algoritmos (Árboles)	Precisión general %		Mejor desempeño	ρ Value	χ^2	Decisión estadística	Nivel de significancia	Significancia estadística
A1 A2	A1	A2	A1	0.05	60.26	Rechaza Ho	95.00 %	Significante
	87.50	86.60						
A1 A3	A1	A3	A1	0.05	2.44	Acepta Ho	95.00 %	No Significante
	87.50	85.56						
A1 A4	A1	A4	A1	0.05	14.50	Rechaza Ho	95.00 %	Significante
	87.50	86.74						
A1 A5	A1	A5	A1	0.05	16.02	Rechaza Ho	95.00 %	Significante
	87.50	84.72						
A2 A3	A2	A3	A2	0.05	52.28	Rechaza Ho	95.00 %	Significante
	86.60	85.56						
A2 A4	A2	A4	A4	0.05	30.25	Rechaza Ho	95.00 %	Significante
	86.60	86.74						
A2 A5	A2	A5	A2	0.05	30.57	Rechaza Ho	95.00 %	Significante
	86.60	84.72						
A3 A4	A3	A4	A4	0.05	4.06	Rechaza Ho	95.00 %	Significante
	85.6	86.74						
A3 A5	A3	A5	A3	0.05	4.98	Rechaza Ho	95.00 %	Significante
	85.56	84.72						
A4 A5	A4	A5	A4	0.05	0.038	Acepta Ho	95.00 %	No Significante
	86.74	84.72						

Fuente: El autor

Con el test estadístico de McNemar se identifica cuales pares de árboles no poseen una significancia estadística, para posteriormente comparar la precisión general de cada uno de los árboles y determinar el árbol que tenga una mejor precisión.

Por lo tanto en esta tabla se puede observar 10 pares de árboles, siendo los pares A1 A3 y A4 A5 los que tienen un mejor desempeño, y se acepta la hipótesis nula con un nivel de significancia del 95.00 %, ya que no hay una diferencia significativa entre los dos algoritmos y el Chi cuadrado es menor a 3.84 mientras que los demás algoritmos se rechazan al ser estadísticamente significativos y por obtener un valor de Chi cuadrado mayor a 3.84.

El árbol con un mejor desempeño es el A1 ya que posee una precisión general de 87.50 % siendo el algoritmo escogido para medir la textura del suelo.

4. Determinación de Atributos del árbol A1

A continuación se muestran las Tablas 21 y 22 la matriz de confusión del árbol con mejor desempeño A1.

Tabla 21. Matriz de confusión para la etapa de prueba del árbol A1

Clase actual	N de datos	% Correcto	Clase a predecir											
			Aa	AL	A	aF	a	F	FAa	FAL	FA	Fa	FL	L
Aa	120	74.17	89	0	11	0	0	0	13	0	7	0	0	0
AL	120	99.17	0	119	0	0	0	0	0	0	1	0	0	0
A	120	100.00	0	0	120	0	0	0	0	0	0	0	0	0
aF	120	90.83	0	0	0	109	0	0	0	0	0	11	0	0
a	120	81.67	0	0	0	6	98	0	0	0	0	16	0	0
F	120	98.33	0	0	0	0	0	118	0	2	0	0	0	0
FAa	120	90.83	2	0	9	0	0	0	109	0	0	0	0	0
FAL	120	91.67	0	0	0	0	0	0	0	110	10	0	0	0
FA	120	89.17	0	2	0	0	0	0	11	0	107	0	0	0
Fa	120	62.50	0	0	0	12	10	13	0	10	0	75	0	0
FL	120	88.33	0	0	0	0	0	11	0	1	0	0	106	2
L	120	83.33	0	0	0	0	0	0	0	0	0	1	19	100
Total	1.440													
Promedio del % correcto		87.50												

Fuente: El autor

Tabla 22. Matriz de confusión para la etapa de aprendizaje del árbol A1

Clase actual	N de datos	% Correcto	Clase a predecir												
			Aa	AL	A	aF	a	F	FAa	FAL	FA	Fa	FL	L	
Aa	120	75.00	90	0	10	0	0	0	10	0	10	0	0	0	
AL	120	100.00	0	120	0	0	0	0	0	0	0	0	0	0	
A	120	100.00	0	0	120	0	0	0	0	0	0	0	0	0	
aF	120	91.67	0	0	0	110	0	0	0	0	0	10	0	0	
a	120	82.50	0	0	0	1	99	0	0	0	0	20	0	0	
F	120	100.00	0	0	0	0	0	120	0	0	0	0	0	0	
FAa	120	91.67	0	1	9	0	0	0	110	0	0	0	0	0	
FAL	120	91.67	0	0	0	0	0	0	0	110	10	0	0	0	
FA	120	90.83	0	1	0	0	0	0	10	0	109	0	0	0	
Fa	120	73.33	0	0	0	0	10	12	0	10	0	88	0	0	
FL	120	90.83	0	0	0	0	0	11	0	0	0	0	109	0	
L	120	91.67	0	0	0	0	0	0	0	0	0	0	10	110	
Total	1.440														
Promedio del % correcto		87.50													

Fuente: El autor

El porcentaje de precisión general del árbol A1 tanto para la matriz de confusión de prueba y aprendizaje es de 87.50 %.

A continuación en la Tabla 23, se muestra las variables para los valores positivos y negativos de la matriz de confusión en la etapa de prueba.

Tabla 23. Tabulación de valores

	PN	PP
NE	<i>a</i>	<i>b</i>
PE	<i>c</i>	<i>d</i>

Fuente: (Zelada, 2017)

- a:** Verdadero negativo (todos los valores menos los de la diagonal)
- b:** Falso positivo (valores arriba de la diagonal)
- c:** Falso negativo (valores abajo de la diagonal)
- d:** Verdadero positivo (los valores de la diagonal)
- PN:** Positivo negativo
- PP:** Positivo positivo
- NE:** Negativo Experimental
- PE:** Positivo Experimental

En la *Tabla 24* se encuentran las fórmulas que permitirán establecer que árbol tiene una mejor precisión, sensibilidad, especificidad y tasa de clasificación de incorrectas.

Tabla 24. Fórmulas para atributos del árbol

Fórmula	
Precisión	$\frac{a + d}{a + b + c + d}$
Tasa de clasificación de incorrectas	$\frac{b + c}{a + b + c + d}$
Sensibilidad	$\frac{d}{c + d}$
Especificidad	$\frac{a}{a + b}$

Fuente: (Zelada, 2017)

Se identificó y se calculó los valores de precisión, sensibilidad, especificidad y tasa de clasificación de incorrectas empleando la matriz de confusión para la etapa de prueba de los árboles A1, A2, A3, A4, y A5 como se muestra en la *Tabla 25*.

Tabla 25. Atributos de los árboles de decisión

Árbol	Precisión	Tasa de clasificación de incorrectas	Sensibilidad	Especificidad
1	0.89	0.11	0.92	0.71
2	0.88	0.12	0.92	0.69
3	0.87	0.13	0.91	0.72
4	0.88	0.12	0.92	0.69
5	0.87	0.13	0.90	0.74

Fuente: El autor

Esta tabla muestra los atributos de los árboles de decisión a partir de la clasificación de los datos de la *Tabla 21* que pertenece a la matriz de confusión en la etapa de prueba, por tanto el mejor árbol según los atributos que presentan cada uno de los cinco arboles es el árbol uno debido a que tiene una mejor precisión de 0.89, una sensibilidad de 0.92, especificidad de 0.71 y posee una tasa de clasificación de incorrectas menor que los arboles A2, A3, A4 y A5 que está representado por 0.11.

En la *Tabla 26* se identifica la fase mineralógica del árbol A1 el cual fue segmentado en ventanas espectrales con un ancho de 10 nm.

Tabla 26. Fase mineralógica identificada con la ayuda del árbol de decisión A1

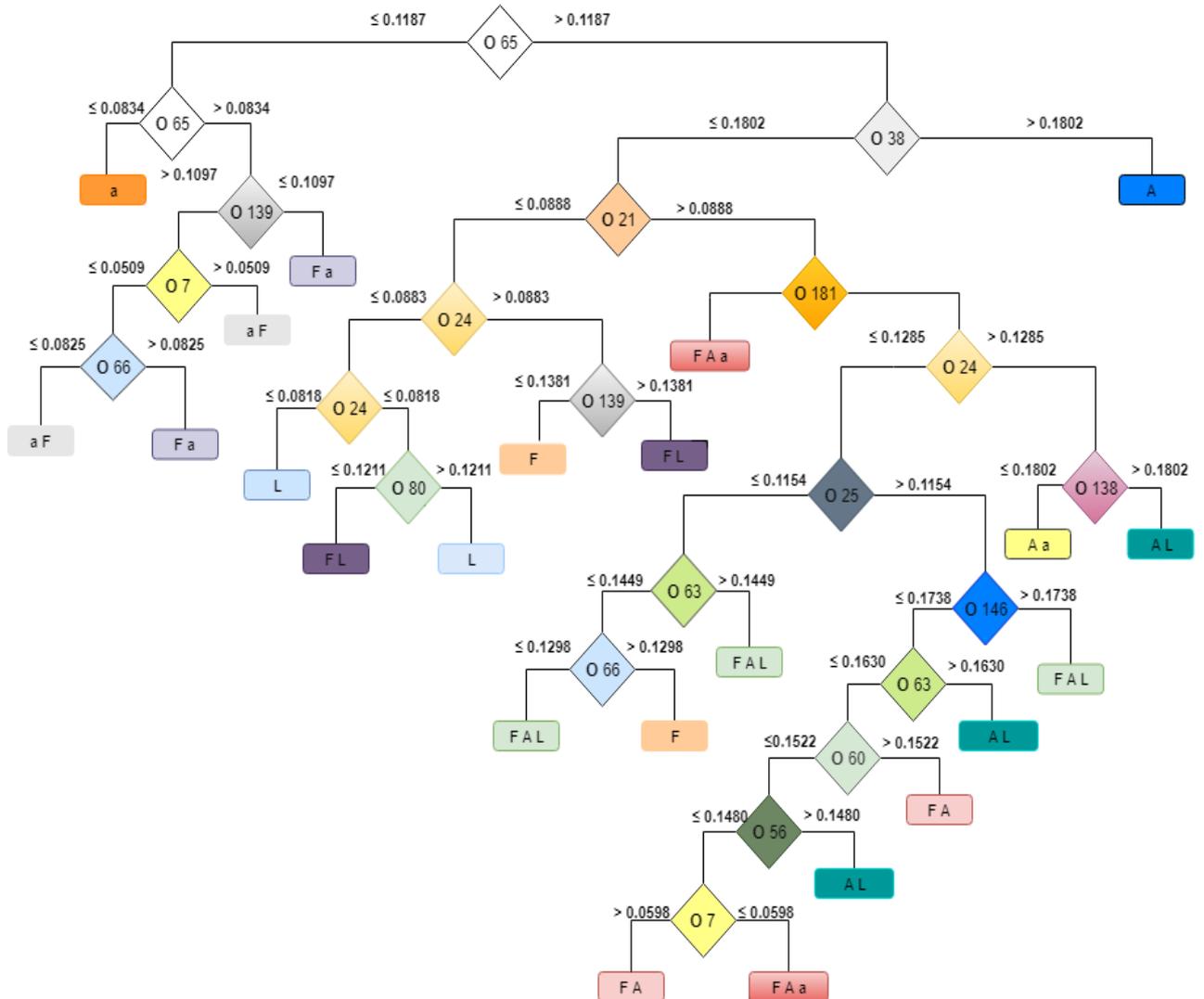
Observable	Porcentaje de importancia %	Longitud de onda (nm)		Fase mineralógica
		Obtenida	Referencia *	
O 25	100	590 - 600	585	Limo
O 27	99.83	610 - 620	620	Limo
O 24	86.20	580 - 590	585	Limo
O 26	74.28	600 - 610	585	Limo
O 29	74.28	630 - 640	620	Limo
O 65	65.25	990 - 1000	1015	Arena
O 177	64.40	2110 - 2120	2080	Arcilla
O 145	63.36	1790 - 1800	1855	Arena
O 180	52.29	2140 - 2150	2200	Arcilla
O 149	51.05	1830 - 1840	1855	Arena

Fuente: El autor

* Curcioa, (2013)

Para la fase mineralógica del árbol A1 con un mejor desempeño se ha identificado 10 variables de importancia (O 25, O 27, O 24, O 26, O 29, O 65, O 145, O 177, O 180, O 149), las mismas que se caracterizan por tener un porcentaje mayor al 50 % y están relacionadas con la banda espectral de 10 nm. Obteniéndose que las observables O 25, O 27, O 24, O 26 y O 29 se encuentran en un rango de 590 - 640 nm por la presencia de limo mientras que para O 65, O 145 y O 149 se hallan entre 990 – 1840 nm por su contenido de arena y finalmente O 177 y O 180, que están en la longitud de onda de 2110 - 2150 nm por la presencia de arcilla.

En la Figura 5 se identifica el árbol de decisión con mejor desempeño para la clasificación de textura del suelo.



Textura	
	Arcillosa (A)
	Arcillo limosa (AL)
	Franco arcillo arenosa (FAa)
	Franco arcillosa (FA)
	Franco arcillo limosa (FAL)
	Franco arenosa (Fa)
	Franca (F)
	Franco limosa (FL)
	Areno francosa (aF)
	Limosa (L)
	Arenosa (a)

Figura 5. Árbol de decisión para la clasificación de la textura del suelo

Fuente: El autor

A continuación en la Figura 6 se puede observar la curva de error del algoritmo (Árbol de decisión) A1.

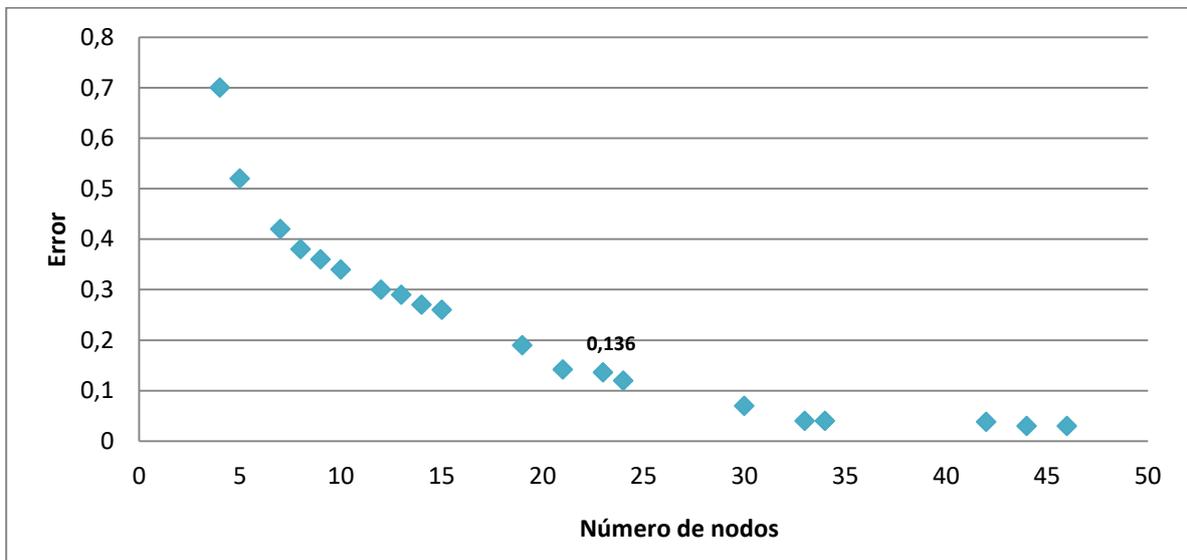


Figura 6. Curva de error del árbol A1 en función del número de nodos

Fuente: El autor

CONCLUSIONES Y RECOMENDACIONES

1. Conclusiones

Con el empleo del RadiometerFieldSPec 4 se ha caracterizado la textura del suelo usando el espectro de reflectancia difusa que ha permitido establecer la matriz observable en cada una de las bandas espectrales de 10, 20, 30, 40, 50 nm para el desarrollo del algoritmo empleando el modelador estadístico y el test estadístico de McNemar.

El árbol utilizado para medir la textura del suelo es el A1 puesto que tiene una precisión general de 87.50 % siendo un porcentaje mayor que los árboles A3 con 85.56 %, A4 con 86.74% y A5 con 84.72 %, los cuales fueron seleccionados aplicando el test estadístico de McNemar donde los pares de árboles A1 A3 y A4 A5 no presentan diferencia significativa al obtener un $\chi^2_{1,0.95}$ menor a 3.84.

Se ha identificado que el porcentaje en la obtención de la textura con el Diagrama de Thien da un error del 33.33% siendo un porcentaje mayor a 13.60 % que es el obtenido con el árbol uno, que tiene un mejor desempeño, por tanto la metodología para determinar la textura es más eficiente que la técnica que implica el uso de las propiedades organolépticas.

Mediante la matriz de confusión en la etapa de prueba del árbol uno se logró obtener una precisión de 89.00 %, una sensibilidad de 92.00 %, especificidad de 71.00 % y una tasa de clasificación de incorrectas de 11.00 %.

Las fase mineralógica del algoritmo A1 el cual fue segmentado en la ventana de banda espectral de 10 nm, con esto se determinaron los observables O 25, O 27, O 24, O 26 y O 29 debido a la presencia de limo y que se encuentra en el rango de 590 - 640 nm mientras que para O 65, O 145 y O 149 por su contenido en arena está entre 990 - 1840 nm y O 177 y O 180 por contener arcilla se halla entre 2110 – 2150 nm.

2. Recomendaciones

- Seguir con la investigación del ERD para determinar la textura del suelo aplicando metodologías que permitan mejorar el algoritmo de clasificación textural.
- Realizar estudios similares empleando el espectro de reflectancia difusa con diferentes muestras de suelo.
- Investigar estrategias para conseguir un algoritmo con mayor eficiencia y menor porcentaje de error tomando como referencia las variables de importancia.

BIBLIOGRAFÍA

- Argüello, A. (2008). *Gestión y conservación del suelo*. Sasovi.
- Bellinaso, H. D. (2010). Soil spectral library and its use in soil. *Ciência do Solo*.
- Ben-Dor, E., Heller, D., & Chudnovsky, A. (2008). A novel method of classifying soil profiles. *Soil Science Society of American Journal*, 1113-1123.
- Benitez, R., Escudero, G., Kanaan, S., & Rodó, D. M. (2014). *Inteligencia artificial avanzada*. Barcelona: UOC.
- Camacho, J. (2013). *bdigital (unal)*. Recuperado el 27 de Noviembre de 2018, de <http://bdigital.unal.edu.co/11558/1/07797091.2013.pdf>
- Cano, A. (2013). *Manual de Practicas de la Materia de Edafologia*. Chiapas: Gobierno del Estado de Chiapas.
- Curcioa, D. (2013). Prediction of soil texture distributions using VNIR-SWIR. *ELSEVIER*.
- Dietterich, T. (30 de Diciembre de 1997). Obtenido de <https://sci2s.ugr.es/keel/pdf/algorithm/articulo/dietterich1998.pdf>
- Donald Gabriels, D. L. (2006). Métodos para determinar granulometría y densidad aparente del suelo . *Academia.edu*.
- Falcó, C. (Diciembre de 2012). *APLICACIONES DEL USO DE LA TÉCNICA DE ESPECTROSCOPIA DE REFLECTANCIA DIFUSA EN*. Obtenido de https://www.uv.es/GICF/4Ar6_Campins_GICF_05.pdf
- FAO. (2017). *Insignia de los suelos*. Obtenido de: <http://www.fao.org/3/a-i3855s.pdf>
- Gisbert, S. I. (2010). LA TEXTURA DE UN SUELO. *ResearchGate*. Obtenido de: https://www.researchgate.net/publication/50839531_La_textura_del_suelo

- Maldonado, A. (2016). *Evaluacion de diferentes dosis de hexametafosfato de sodio, en la determinacion de tres tipos texturales de suelo, mediante el metodo de bouyoucos*. Quito: Universidad Central del Ecuador.
- Martínez, D. (2012). *El yo y la máquina, Cerebro, mente, e inteligencia artificial*. España: Palabra.
- Minolta, K. (2018). Recuperado el 06 de Diciembre de 2018, de <http://sensing.konicaminolta.com.mx/learning-center/light-measurement/radiometers-photometers/>
- Páez, C. E. (2008). *Espectrometría de reflectancia (SWIR), aplicada para mapeo de alteración, Proyecto La India. Distrito minero Mulatos* . Obtenido de <http://tesis.uson.mx/digital/tesis/docs/18179/Capitulo6.pdf>
- Ritchey, E. L., McGrath, J. M., & Gehring, D. (2015). Determining Soil Texture by Feel. *UKnowledge*.
- Stenberg, B. V. (2010). Visible and near infrared spectroscopy in soil science. *Advances in Agronomy*.
- Tamayo, J. H. (2013). *Uso de la reflectancia difusa - NIR en a determinación de características físicas y químicas de un Oxisol* . Obtenido de <http://bdigital.unal.edu.co/11558/1/07797091.2013.pdf>
- Viscarra Rossel, R. W. (2006). Visible near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous. *Geoderma*.
- Xie, X., Pan, X., & Sun, B. (2012). Visible and near-infrared diffuse reflectance. *Pedosphere*, 351-366.
- Zelada, C. (10 de Mayo de 2017). Matriz de confusión. Obtenido de <https://rpubs.com/chzelada/275494>

ANEXOS



Preparación de la materia prima

Fuente: Autor



Preparación de la materia prima

Fuente: Autor



Toma del peso de cada uno de los agregados

Fuente: Autor



Pesando los agregados para las muestras sintéticas

Fuente: Autor



Muestras sintéticas etiquetadas

Fuente: Autor



La muestra de suelo mas agua(textura)

Fuente: Autor



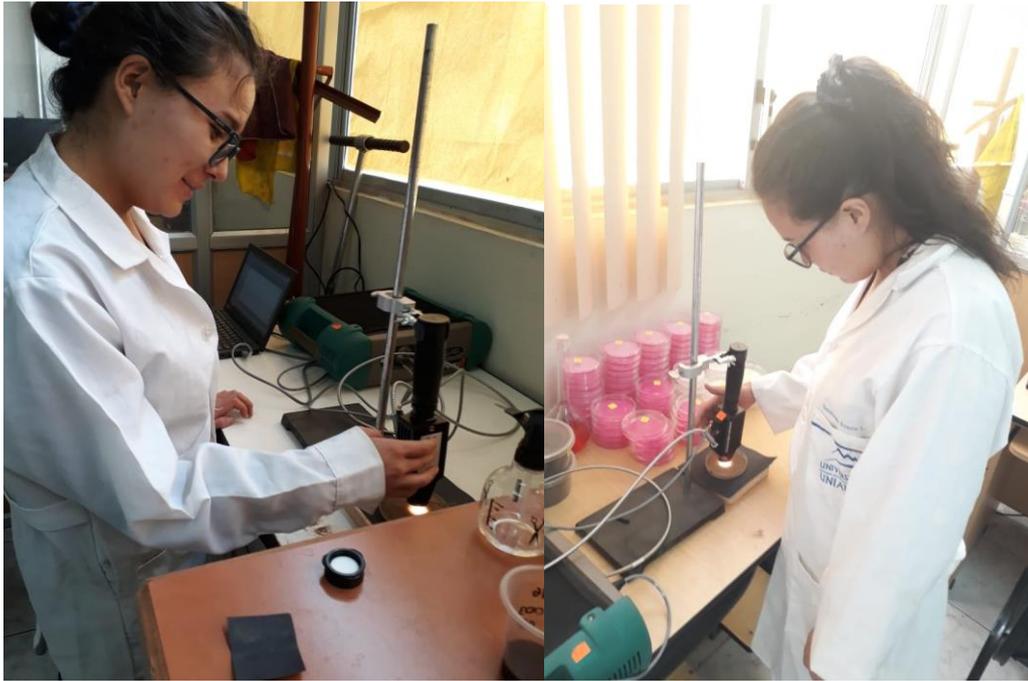
Formación de la cinta de las muestras sintéticas

Fuente: Autor



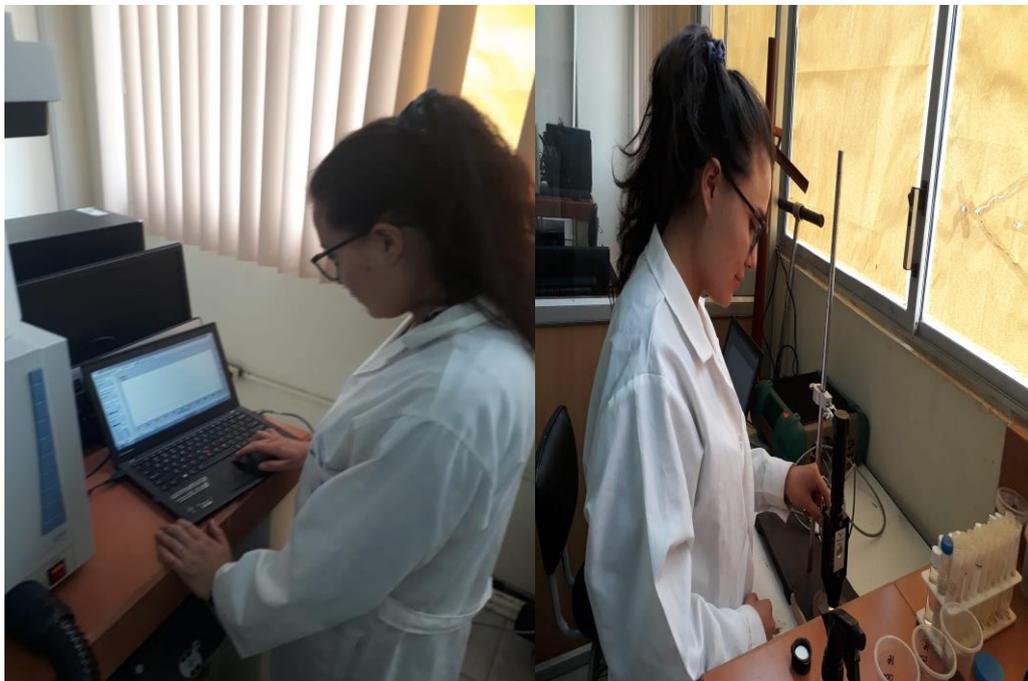
Equipo RadiometerFieldSpec 4

Fuente: Autor



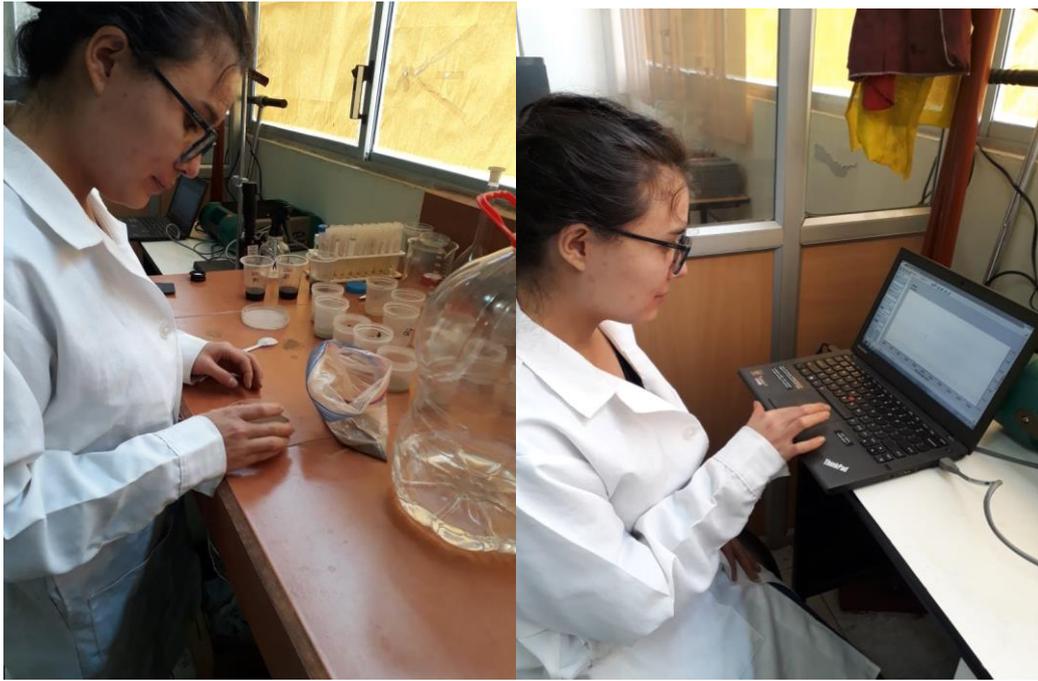
Instalación del Equipo RadiometerFieldSPec 4

Fuente: Autor



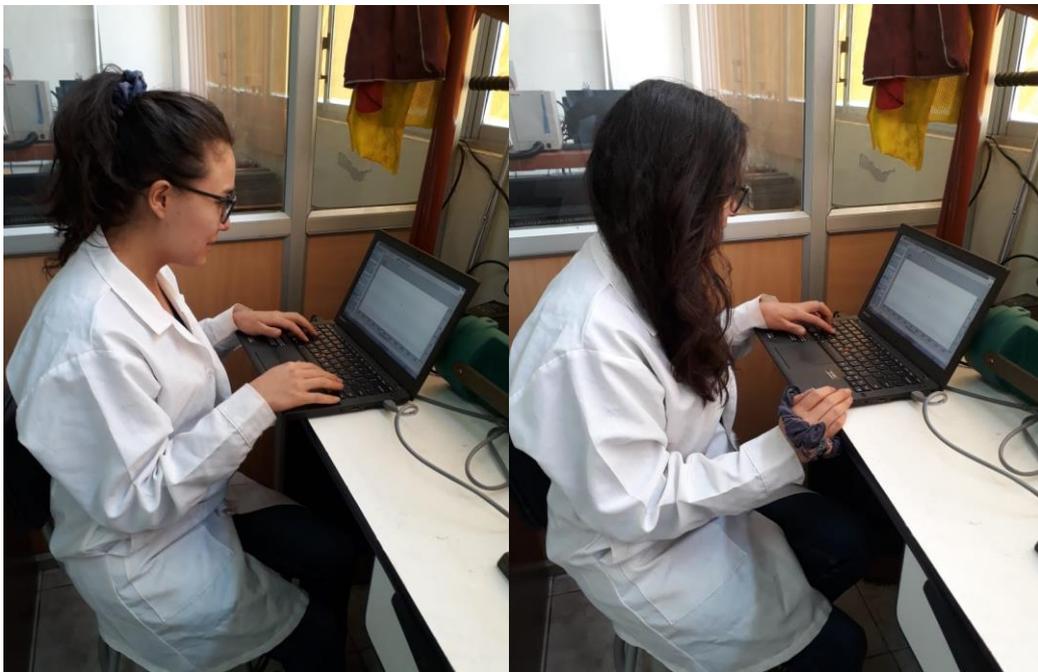
Estableciendo el blanco en el Equipo RadiometerFieldSPec 4

Fuente: Autor



Registro del ERD en 4 ángulos

Fuente: Autor



Registro del ERD

Fuente: Autor

