



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

---

**FACULTAD DE INGENIERÍA**

**Modelación de la concentración de  
monóxido de carbono en la  
atmósfera de la ZMVM mediante  
información obtenida de sensores  
remotos**

**TESIS**

Que para obtener el título de

**Ingeniero Civil**

**P R E S E N T A**

Kevin Nieto Pedraza

**DIRECTOR DE TESIS**

M.I. Rodrigo Takashi Sepúlveda Hirose



Ciudad Universitaria, Cd. Mx., 2018



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

## **AGRADECIMIENTOS:**

A mi papá, por haberme apoyado a lo largo de la carrera y ser siempre el eje de soporte familiar, a mi mamá, por alentarme a superarme día con día, por creerme en mí y ayudarme a terminar este proyecto, a José Luis, por creer en mí y darme fuerzas para no rendirme, a mi hermana por aguantarme y exhortarme a darlo todo para concluir este ciclo, a mi novia por ser mi soporte a lo largo de la carrera y apoyarme en todas las decisiones realizadas, a mi primo Erick que siempre creyó en mí, y a mi tutor Rodrigo Sepúlveda, por guiarme en la realización de este trabajo y aceptarme en la participación del proyecto SECITI, además de ser una excelente persona y profesor.

A la Universidad, por darme las herramientas necesarias para mi educación, a la Yeungnam University de Corea del Sur, por haberme formado en algún punto de mi carrera.

A Fundación Telmex por haberme otorgado apoyos económicos para que me enfocara en siempre dar lo mejor de mí.

A mis amigos, por el apoyo, aliento y siempre estar ahí.

## **AGRADECIMIENTO ESPECIAL:**

Agradecimiento especial a la Secretaría de Ciencia, Tecnología e Innovación de la Ciudad de México; por los fondos, asesorías y apoyos otorgados durante el desarrollo de esta investigación que formará parte del proyecto **SECITI/089/2017 “Desarrollo de modelos para la estimación de material particulado suspendido de la Zona Metropolitana de la Ciudad de México mediante el uso de imágenes satelitales e información geoespacial”**

## Resumen

La Zona Metropolitana del Valle de México se encuentra en un crecimiento continuo y acelerado, por lo que el número de industrias y vehículos también se ha incrementado en los últimos años ocasionando que los niveles de contaminación atmosférica incrementen a niveles nunca antes vistos, debido a esto los gobiernos se ven en la necesidad de tomar muestras de manera continua con equipos de alto costo y movilizándolo personal a las zonas de estudio, por lo que la percepción remota surge como una herramienta alternativa para la generación de modelos atmosféricos.

Se tomaron en un principio los datos de monóxido de carbono de la Red Automática de Monitoreo Atmosférico, este contaminante fue escogido debido al gran aumento del parque vehicular en los últimos años dentro de la Zona de estudio, por lo que las emisiones provenientes de vehículos han aumentado causando impacto en la salud de niños y adultos de la tercera edad, así como en la atmósfera oxidándose y transformándose en dióxido de carbono. Por otro lado, imágenes satelitales fueron adquiridas por medio de la plataforma Landsat 8 teniendo acceso al nivel 2 de información, con esto se evitó la realización de correcciones atmosféricas facilitando el procesamiento de las imágenes y la generación de modelos.

El objetivo de este proyecto es generar modelos atmosféricos que puedan predecir el comportamiento del monóxido de carbono en distintas fechas, utilizando como datos, la reflectancia de las bandas 1 a la 7 de la plataforma Landsat 8 y la concentración de monóxido de carbono obtenido por medio de la RAMA.

Durante el desarrollo de la investigación se utilizaron parámetros estadísticos distintos a la  $R^2$ , para la selección de modelos, como el Criterio de Información de Akaike que es la relación entre el sesgo y la varianza, así como el parámetro BIC entre otros. Además de que se analizó que en época de lluvias los modelos tienden a tener un comportamiento menos confiable en comparación al periodo de estiaje, donde los modelos tienen parámetros de mayor valor y confiabilidad para con ellos definir en qué época del año la percepción remota entrega mejores resultados, a su vez se utilizaron y compararon distintas regresiones lineales para tener una gama amplia de selección.

# Índice

<b>Introducción</b> .....	<b>7</b>
<b>Capítulo 1. Antecedentes de la Contaminación Atmosférica</b> .....	<b>9</b>
<b>Justificación</b> .....	<b>15</b>
<b>Objetivo general</b> .....	<b>16</b>
<b>Objetivos específicos</b> .....	<b>16</b>
<b>Capítulo 2. Calentamiento Global</b> .....	<b>17</b>
2.1. Efectos del calentamiento global.....	20
2.2. Contaminantes Atmosféricos .....	22
<b>Capítulo 3. Percepción Remota</b> .....	<b>26</b>
3.1 Elementos de la Percepción Remota.....	27
3.2 Sistemas de la Percepción Remota .....	29
3.2.1 Percepción Remota Aérea. ....	29
3.2.1 Percepción Remota Espacial.....	29
3.3 Plataformas de la Percepción Remota Espacial.....	30
3.3.1 Satélites Landsat TM.....	30
3.3.2 Sensor MODIS .....	33
3.3.3 Satélites Spot .....	33
3.3.4 Plataforma NOAA .....	34
3.3.5 Satélites IKONOS .....	35
3.3.6 Aplicaciones de los sistemas de percepción remota .....	36
<b>Capítulo 4. Regresiones y parámetros utilizados para elección de modelos</b> .....	<b>37</b>
4.1 Regresiones usadas comúnmente para la elección de modelos.....	38
4.2 Parámetros utilizados en la elección de modelos.....	44
<b>Capítulo 5. Metodología</b> .....	<b>46</b>
5.1 Estado del arte .....	46
5.2 Parámetros de la calidad del aire .....	47
5.3 Zona de estudio.....	49
5.4 Plataforma seleccionada.....	55
5.4 Procesamiento de las imágenes .....	58
<b>Capítulo 6. Resultados y Análisis</b> .....	<b>66</b>

<b>Capítulo 7. Conclusiones</b> .....	<b>78</b>
<b>Anexos</b> .....	<b>80</b>
<b>Referencias</b> .....	<b>96</b>

## Introducción

El aire, al ser uno de los recursos naturales esenciales para nuestro desarrollo debería de ser estudiado como lo es el agua, pero hoy en día la población sigue pensando que la calidad de la atmósfera es un término ajeno a nuestras actividades por lo que su cuidado pasa desapercibido, pero año con año médicos y científicos dan a conocer nuevas cifras alarmantes sobre los efectos ocasionados por aerosoles y emisiones tóxicas sobre la salud, asociado a esto, el monitoreo ambiental se ha convertido en un tema necesario en cada proyecto a realizar, dado las consecuencias que tienen las emisiones dentro de la atmósfera, por lo que cada día existen nuevas investigaciones y herramientas para mejorar el entendimiento de las causas, impacto y control de los parámetros contaminantes y evitar así repercusiones en nuestro medio ambiente y salud.

El proceso de emisiones a gran escala comenzó en el siglo XVIII, teniendo graves repercusiones en los países líderes de ese entonces, haciendo que estos mismos crearan leyes para poder controlar este problema y tener una calidad apropiada para realizar las labores diarias, pero como los efectos de este problema no son tan inmediatos como lo son al contaminar el agua, las leyes fueron perdiendo importancia, creando con esto que más países se sumaran a este nuevo modelo económico, emitiendo cada vez más a una mayor escala contaminantes hacia la atmósfera, y no fue hasta 1940 que se introdujeron leyes internacionales para el control y límites permisibles de los parámetros atmosféricos. En México no fue hasta 1966 que se creó una red de monitoreo ambiental en la Zona Metropolitana del Valle de México, cuando la entonces Secretaría de Salubridad integró la primera red para medir el dióxido de azufre y partículas suspendidas; en ese entonces se contaba con cuatro estaciones de monitoreo y hasta 1993 se comenzó a monitorear al monóxido de carbono por medio de la Red Automática de Monitoreo Atmosférico (RAMA) y comunicando los datos en tiempo real.

Debido a esto muestras de aire son tomadas continuamente, con la finalidad de evaluar parámetros para la determinación de su calidad, y aunque se cuente con tecnología avanzada para medir la calidad, esta es de acceso limitado y alto costo, por lo que el uso de la percepción remota surge como principal herramienta para su análisis y esto ha tomado una gran importancia en el contorno científico alrededor del mundo, sobre todo en países asiáticos donde la calidad del aire se ha tornado en un tema de suma importancia.



En el primer capítulo de este documento se habla de los antecedentes y contexto histórico de la contaminación atmosférica. En el segundo capítulo se trata de exponer el calentamiento global, tanto las causas y efectos que tienen en las actividades humanas, así como los principales contaminantes atmosféricos y efectos que tienen en la salud humana. Seguidamente el tercer capítulo, se aborda a la percepción remota de un enfoque general, exponiendo los diferentes sistemas, plataformas y usos, así como la justificación del uso del Landsat 8. En el cuarto capítulo se realiza una descripción del proceso realizado, así como de la zona de estudio, del parámetro seleccionado y la metodología, En el capítulo siguiente se exponen los resultados obtenidos, así como las diferentes regresiones que se aplicaron, el por qué la regresión Stepwise resulto la óptima y los diferentes parámetros de selección que existen dentro de la estadística para modelos y en el último capítulo se tienen las conclusiones.

## Capítulo 1. Antecedentes de la Contaminación Atmosférica

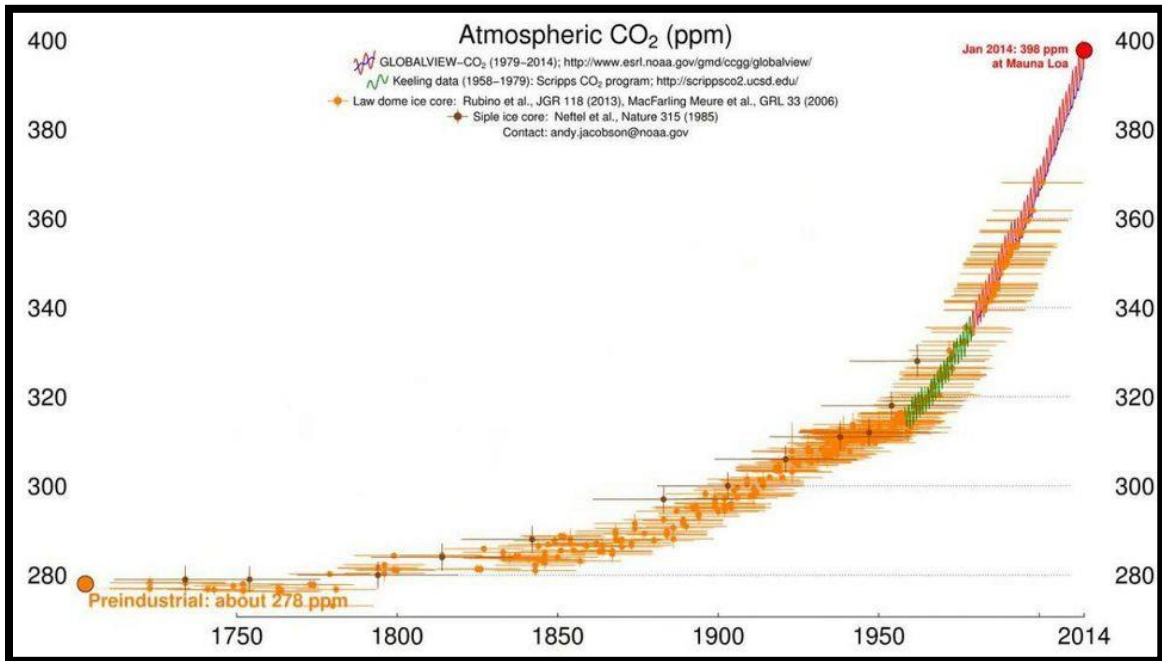


Imagen 1.1: Aumento del CO<sub>2</sub> desde la pre-industrialización (Cohen 2013)

La atmósfera terrestre comenzó a formarse hace unos 4,600 millones de años, junto con el origen de la Tierra. En sus primeros inicios la atmósfera comenzó su evolución; comprendida en un principio por Hidrógeno, Helio, vapor de agua, entre otros gases, que rodeaban una Tierra joven e inestable térmicamente. El planeta en un principio no podía retener gases ligeros como el Hidrógeno y el Helio que se escapaban al espacio, haciendo que la temperatura de terrestre incrementara haciendo que la formación de agua y vida como la conocemos se dificultarían. La presencia de compuestos más sofisticados como Dióxido de Carbono, Monóxido de Carbono, Nitrógeno y Moléculas de Agua, estos compuestos hicieron que la temperatura terrestre disminuyera logrando la formación de la hidrosfera hace unos 4000 millones de años y con ello la formación de grandes masas de agua.

Para organismos como los humanos, el cambio atmosférico más importante fue la formación del oxígeno ya que ningún otro proceso químico directo o procesos geológicos producen este elemento; Por eso la hidrósfera la formación de la hidrósfera, una atmósfera estable y la energía del Sol, fueron las condiciones ideales para que dentro del mar se formaran las primeras proteínas que darían inicio a las primeras formaciones de vida. Hace algunos mil millones de años, las algas empezaron a usar energía del Sol para dividir moléculas de agua y dióxido de carbón para obtener oxígeno libre, así empezó el proceso de fotosíntesis, lo cual constituyo un paso fundamental en la formación de la atmósfera terrestre como se conoce hoy en día, en ese momento empezó la formación de partículas de ozono que absorberían los rayos UV (Ultravioleta) provenientes del Sol, al paso de los años la cantidad de ozono fue suficiente para que la formación de vida terrestre fura posible.

Actualmente la atmósfera juega un papel no sólo de protección contra varios fenómenos físicos que suceden en el espacio, sino también como una extraordinaria reguladora de procesos termodinámicos, químicos y biológicos inherentes a la evolución y al acontecer terrestre, sin los cuales la vida no existiría como la conocemos. La interacción de diversos factores como la temperatura de los océanos, la ayuda del ozono al contener los rayos UV emitidos por el Sol, y un clima relativamente estable permitieron que la vida siguiera su evolución.

La atmósfera ha sufrido cambios en su composición a lo largo de los años, existen fuentes naturales que contribuyen a que existan cambios en la atmósfera; las erupciones volcánicas emiten una gran cantidad de Dióxido de Azufre, un claro ejemplo es el Monte Etna, se estima que cada día añade 16,000 toneladas de dióxido de carbono a la atmósfera, un aproximado de 5.8 millones

de toneladas al año y esto solo es un volcán de los muchos que existen alrededor del planeta, se estimó que en el 2013 que 60 millones de toneladas de dióxido de carbono fue emitido de 33 desgasificaciones de volcanes, que es el proceso en el cuál la magma y las cenizas ascienden juntos a la atmósfera haciendo que el magma pierda su volatilidad y reaccionando con los gases de la atmósfera, de los 30 volcanes que se tienen observados se emitió 6.4 millones de dióxido de carbono, extrapolando esto a los 550 volcanes históricamente activos se estima que se emitieron 117 millones de toneladas de dióxido de carbono al año. (Siegel, 2017).

De las erupciones volcánicas se emite también dióxido de azufre el cual queda atrapado en la estratosfera, un gas volátil que con el tiempo se oxida y produce ácido sulfúrico que queda en suspensión en la estratosfera por hasta dos años en forma de pequeñas gotas llamadas aerosoles, en 1983 el Programa Ambiental de las Naciones Unidas estimó que, entre 80 y 288 millones de toneladas de Óxidos de Azufre al año, comparado con los aproximadamente 69 millones de toneladas que son emitidos por actividades humanas. De igual manera entre 20 y 90 millones de toneladas de Óxidos de Nitrógeno son emitidas al año debido a actividades naturales; como las erupciones volcánicas, las tormentas eléctricas, entre otras, contaminantes secundarios como el ozono, que es formado mediante reacciones fotoquímicas a partir de emisiones procedentes de la vegetación, procesos de fermentación o volcanes, y se encuentra tanto en la troposfera como en la estratosfera, cuando el ozono troposférico aumenta en mayores cantidades, se convierte en un contaminante tóxico que contribuye al incremento de enfermedades cardiovasculares en el ser humano. (Pabón, 2005).

Aunque las actividades naturales contribuyen en gran medida a la contaminación atmosférica, no son de gran preocupación como la aportación debido a las actividades humanas, esto debido a que estos fenómenos aparecen en periodos intermitentes en el tiempo y no son prolongados además cuando comparamos los 0.645 mil millones de toneladas de dióxido de carbono emitido cada año de manera natural contra los 29 mil millones emitidos por el hombre es claro entender porque el hombre está modificando la atmósfera y causando daños al ambiente.

Para el ser humano el descubrimiento del fuego simbolizó un gran avance, ayudándonos a protegernos contra los animales, cocinar, elaborar herramientas y armas, pero también el descubrimiento del fuego el hombre empezó a contaminar el aire poco a poco, claro que en esos tiempos la contribución del hombre era de manera insignificante, pero gracias a la formación de comunidades sedentarias se la contaminación atmosférica comenzó a incrementar; se empezó a

tener un impacto ambiental mayor con más intensidad, ya que los grupos dejaron de ser nómadas que generaban impactos ambientales en un lugar y en el tiempo mínimos.

Durante los siglos XII y XIII, fue cada vez más difícil en Europa seguir usando madera como combustible principal para las actividades del día al día, ya que era un combustible de rápida combustión y que su obtención no era de manera tan inmediata, por eso fue necesario encontrar un combustible alternativo como el carbón que ofrece mantener una temperatura elevada por mayor tiempo y que transmite el calor de manera más eficiente, pero la combustión de carbón fue origen de crecientes problemas, cuando el carbón se convirtió en un combustible común para la generación de calor y energía, la contaminación del aire aumentó en forma exponencial como se observa en la imagen 1.1 al inicio del capítulo.

En Londres existen los primeros antecedentes de contaminación del aire en 1285, cuando fue aprobada una ley para prohibir la combustión en Londres, al principio el carbón se empleaba de manera industrial pero el problema empeoró cuando se empezó a emplear en las casas, la mejor información que se dispone acerca de la de la situación en esta época, cuando a principios de diciembre en 1952 una especie de smog (humo de carbón atrapado en la neblina) ocasionó la muerte prematura de 4000 personas debido principalmente de enfermedades crónicas de naturaleza respiratoria y circulatoria, esto provocó la creación de la Ley de Aire Puro de 1956 la cuál proporcione que las autoridades crearan zonas sin humo, mostradas en la imagen 1.2, gracias a esto algunas ciudades británicas tuvieron una notable mejora en la calidad del aire, lo cual sentó las bases para el cuidado de la atmósfera y la calidad del aire. (Enviropedia 2018)

En los principios de la revolución industrial, Gran Bretaña era famoso por sus amaneceres oscuros, todo porque el país era dependiente del carbón y cada día millones de emisiones de carbono eran descargadas al ambiente, se estima que a principios de 1850 debido a un incremento en el uso del carbón se incrementaron las tasas de mortandad en 1% en las zonas norte y sur de Inglaterra (Hatton 2010), donde se ubicaban las industrias en ese tiempo, de hecho podría decirse que la calidad de aire que se tiene hoy en día en China e India son comparables con la que tenía Gran Bretaña a finales del siglo XIX. Con el avance de la revolución industrial, las industrias empezaron a demandar cantidades mayores de energía, aumentando así las emisiones de gases de efecto invernadero como el dióxido de carbono y el metano, así como las emisiones de óxidos de nitrógeno que causan la lluvia ácida, pero dado que la producción en masa comenzaba y las necesidades poblacionales iban incrementando se fue dejando de lado el daño tanto a la salud

como al ambiente siendo la prioridad el poder producir más en el menor tiempo posible utilizando todos los recursos naturales.

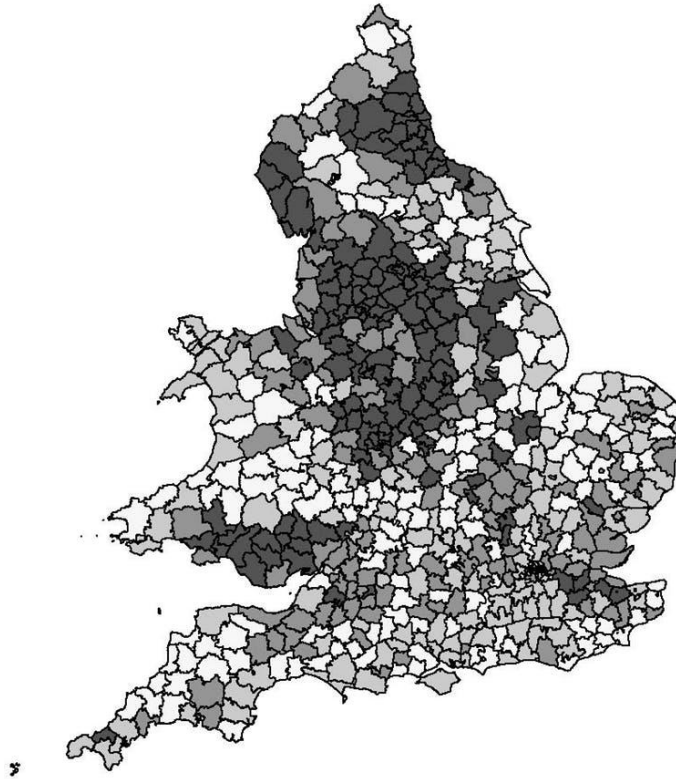


Imagen 1.2: Zonas libres de humo en Inglaterra y Gales, 1901. (Bailey et al. 2016)

La industria metalúrgica comenzó a prosperar, pero no fue hasta 1859, cuando en Pensilvania se inició la perforación de los pozos petroleros iniciando así lo que sería la principal fuente contaminante producida por el hombre. Tal como lo proclamó el científico sueco Svante Arrhenius en 1896, los combustibles fósiles podrán dar lugar a acelerar el calentamiento de la Tierra, Arrhenius estableció una relación entre concentraciones de dióxido de carbono atmosférico y temperatura. También determinó que la media de la temperatura superficial de la tierra es de 15°C debido a la capacidad de absorción de la radiación Infrarroja del vapor de agua y el Dióxido de Carbono, esto se denomina el efecto invernadero natural. Arrhenius sugirió que una concentración doble de gases de CO<sub>2</sub> provocaría un aumento de temperatura de 5°C. El junto con Thomas Chamberlin calculó que las actividades humanas podrían provocar el aumento de la temperatura mediante la adición de dióxido de carbono a la atmósfera, En 1940 se produjeron desarrollos en

las mediciones de radiaciones de onda larga mediante espectroscopia de Infrarrojo. En este momento se comprobó que el aumento del dióxido de carbono en la atmosfera provoca una mayor absorción de radiación Infrarrojo.

También se comprobó que el vapor de agua absorbe radiaciones diferentes que el dióxido de carbono. pero dichas investigaciones se fueron olvidando con el paso del tiempo ya que durante los años 1980 la curva de temperatura media anual comenzó a tener incrementos y las teorías de un calentamiento global fueron siendo más recurrentes, pero aun así en ese mismo periodo el número de vehículos motorizados fue incrementando en las zonas urbanas y los problemas de calidad del aire fueron más cotidianos, tanto el avance de la industria petroquímica y el desarrollo de la industria del transporte, con el uso de combustibles fósiles, permanece hasta hoy como la causa principal de emitir gases de efecto invernadero.

## Justificación

En los últimos años, el calentamiento global, sus causas y efectos han sido un tema de controversia a nivel mundial, por lo que cada vez más países se ven en la necesidad de monitorear parámetros de calidad, para evitar enfermedades y condiciones desagradables para el desarrollo humano, por lo que la Secretaría del Medio Ambiente de la CDMX (SEDEMA) lleva a cabo, por medio del Sistema de Monitoreo Atmosférico, un continuo muestreo en distintas estaciones de la Zona Metropolitana del Valle de México, realizando esto por medios convencionales como la absorción del infrarrojo para determinar el monóxido de carbono en una celda de correlación, esto hace que ciertos limitantes aparezcan como los mantenimientos a equipos costosos, desplazamiento de personal de operación y la realización de pruebas en laboratorios inaccesibles, por lo que se propone emplear la percepción remota como una herramienta novedosa para la obtención de información de las estaciones de interés por medio de imágenes satelitales, aunque esta técnica se encuentre en desarrollo, muchos países asiáticos (Sohrabinia, 2007, Elbir, 2010, Mohammed Hashim, 2010) han tomado la iniciativa para obtener mejor información y nuevas investigaciones utilizando los nuevos sensores que las plataformas recién lanzadas ofrecen, como la posibilidad de medir aerosoles a partir de la reflectancia obtenida, utilizando la información de las longitudes de onda de la luz visible, Además que la percepción remota ofrece mayores ventajas que el monitoreo tradicional, en áreas de difícil acceso y con un mayor rango de visión.



## Objetivo general

Aplicar distintas regresiones para la generación de modelos atmosféricos dentro de un rango de fechas utilizando la reflectancia obtenida por medio de la plataforma Landsat 8 y la concentración de monóxido de carbono obtenida de la Red Automática de Monitoreo Atmosférico de la Zona Metropolitana del Valle de México, identificando la mejor opción utilizando parámetros estadísticos.

## Objetivos específicos

- Analizar la importancia de monitorear contaminantes atmosféricos y las repercusiones que tienen.
- Explicar los distintos sistemas de percepción remota existentes y las plataformas en órbita.
- Comparar los distintos niveles de monóxido de carbono en el 2016, 2017 y 2018.
- Aplicar diferentes regresiones, así como parámetros estadísticos para la elección de modelos.
- Estudiar la relación que existe entre la reflectancia y los valores obtenidos de la RAMA.

## Capítulo 2. Calentamiento Global

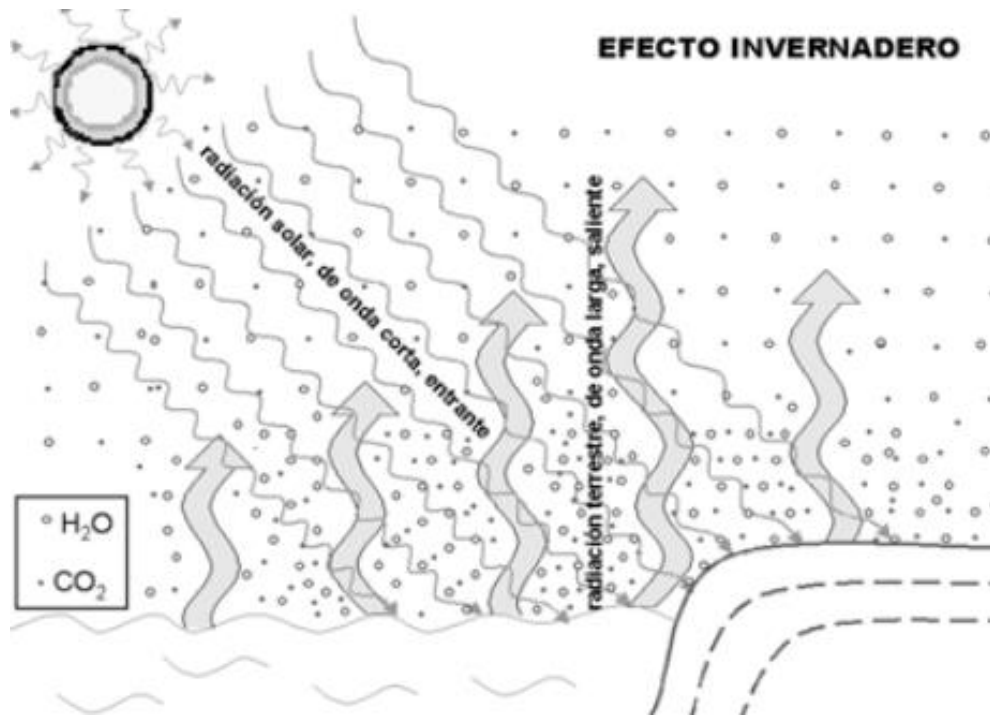


Imagen 2.1: Explicación del efecto invernadero (Garduño R 2018).

Todos los cuerpos emiten una radiación por medio de ondas electromagnéticas estas se caracterizan por su longitud de onda y su frecuencia las cuales son inversamente proporcionales, la radiación emitida depende de la temperatura del cuerpo en dos aspectos; en la cantidad de radiación que el cuerpo posee y la longitud de onda que va disminuyendo a medida que la temperatura va aumentando. En la atmósfera y en el clima intervienen dos tipos distintos de radiación, por un lado, la luz visible que emite el Sol a más de 6 mil grados Celsius y la radiación infrarroja emitida por la Tierra, en la imagen 2.1 se da una explicación gráfica de este fenómeno.

La Tierra por estar a una cierta distancia del Sol y tener una determinada capacidad de reflejar la radiación que recibe del mismo tiene una temperatura de equilibrio o efectiva, la cual es el resultado del balance entre la radiación solar de onda corta absorbida por la Tierra y la emitida por ella misma en onda larga, es por eso que se calcula que la temperatura efectiva de la Tierra es de -18 grados Celsius, un valor alejado a la temperatura promedio terrestre y que es posible gracias al efecto invernadero.

La atmósfera al ser un filtro que deja pasar los rayos solares que son absorbidos en su mayoría por la superficie y cuerpos terrestres que van aumentando su temperatura y con ello emitiendo radiación terrestre que es absorbida por la atmósfera que se va calentando al recibir más radiación, para después ser reemitido hacia el espacio y hacia la superficie terrestre de nuevo es decir que aunque el Sol es la fuente original de energía térmica que se encuentra en la atmósfera, ésta no se calienta por arriba sino por medio de intercambios de calor que suceden dentro de la Tierra, tampoco toda la radiación es retenida en la Tierra una parte escapa hacia el espacio manteniendo así el equilibrio térmico en nuestro planeta, el nombre efecto invernadero se debe a la similitud que existe con los invernaderos para cultivar plantas, se trata de crear un ambiente más cálido en el interior que en el exterior, se debe mencionar también que existen ciertos procesos que intervienen en el proceso, uno de ellos es la convección atmosférica, donde el aire inferior al calentarse poco a poco por el contacto directo con la superficie y por la radiación procedente de ella se va dilatando haciendo que su densidad disminuya y con ello subiendo hacia la atmósfera desplazando a las porciones frías que se encuentran en la parte superior creando así un intercambio calorífico entre las masas de aire. En segundo lugar, se encuentra el cambio de fase del agua del agua, cuando el aire se va enfriando debido a la a que la densidad y la presión atmosférica van disminuyendo a medida que la altura es menor provocando que el vapor de agua contenido en él sea condensado, cambiando de fase gaseosa a líquida liberando calor.

El efecto invernadero es un proceso natural que ha existido desde que la atmósfera se formó y es gracias a este efecto que la temperatura de la tierra se ha podido mantener estable; el problema surge con el calentamiento global, el cual se debe a un incremento en la temperatura de la atmósfera terrestre, océanos y superficie terrestre, siendo las causas principales los gases de efecto invernadero como el vapor de agua, el Dióxido de Carbono, el Metano, los óxidos de nitrógeno y clorofluorocarbonos (CFCs).

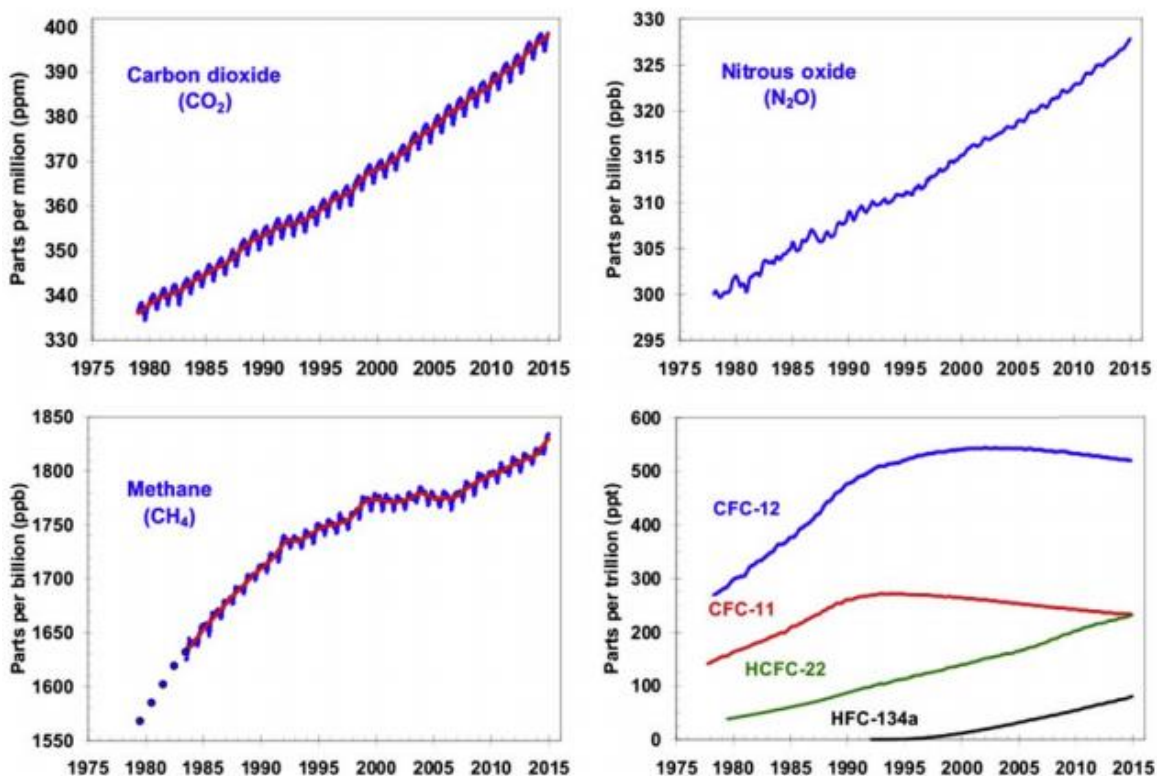


Imagen 2.2: Cambios en diferentes gases de efecto invernadero de 1980 al 2015 (Cohen 2013).

Como se observa en la imagen 2.2, existe un incremento de los gases de efecto invernadero ocasionando que la atmósfera retenga una mayor cantidad de calor y radiación solar, limitando la cantidad habitual de radiación que puede escapar hacia el espacio haciendo que quede atrapada en la atmósfera y con ello provocando que la superficie terrestre aumente su temperatura terrestre teniendo como consecuencias el cambio climático, extinción de especies, deshielo de las capas polares y aumento del nivel medio del mar, es por eso que cada vez es más común el uso de tecnologías e instrumentos que nos ayuden a detectar las fuentes y concentraciones de los contaminantes atmosféricos así como herramientas que modelen el posible comportamiento de los mismos gases y sus efectos en la salud, actividades tanto humanas como del medio ambiente,

para evaluar las posibles medidas a escoger para contrarrestar los efectos del calentamiento global.

## **2.1. Efectos del calentamiento global**

El efecto más obvio del calentamiento global es el cambio climático, el clima es el resultado de una compleja interacción entre varios elementos y factores, por ello, una pequeña modificación altera todo el sistema. Para adaptarnos a un clima diferente necesitamos saber cómo está cambiando el clima a nivel global, regional y local. ya que condiciones extremas en el clima afectarían la vida en el planeta. En México dado que gran parte del territorio nacional es de tipo semiárido, los cambios en el ciclo hidráulico y de temporada de lluvias provocarían sequías más prolongadas y con mayor frecuencia; los huracanes así como los fenómenos del niño y la niña han aparecido con mayor regularidad en las costas de México ocasionando que los patrones de lluvia se vean afectados repercutiendo en pérdidas humanas y económicas de mayor grado que años anteriores, se estima que habrá regiones donde la precipitación llegará a ser más intensa y frecuente aumentando el riesgo de inundación para alrededor de dos millones de personas en situación vulnerable. Se puede decir que mientras que la variabilidad natural del clima se ha presentado en ciclos desde su inicio, con el cambio climático los ciclos poco a poco han dejado de existir y las variaciones se han presentado con mayor frecuencia otros cambios que han surgido en México es que el país se ha vuelto más cálido desde la década de 1960, las temperaturas promedio a nivel nacional aumentaron 0.85 grados Celsius, la precipitación ha disminuido en el sureste del país y se han reducido la cantidad de días frescos desde los años setenta. En el sector agrícola se espera una disminución de la productividad del maíz para la década de 2050 y existe cierta evidencia de que la mayoría de los cultivos resultarán menos adecuados para la producción además de que un aumento de la temperatura en el océano llegara a ocasionar un colapso demográfico en poblaciones marinas. (SEMARNAT, 2018).

Otro efecto importante a mencionar es el aumento en el nivel medio del mar, una causa importante es la dilatación térmica, el agua al calentarse al haber un aumento en la temperatura tiende a dilatarse alrededor de la mitad del aumento del nivel del mar que se produjo a lo largo del siglo

pasado es atribuible al hecho de que los océanos, al calentarse, ocupan más espacio. El otro factor importante es el deshielo de los casquetes polares y glaciares estas grandes formaciones de hielo se derriten de forma natural en verano.

Pero en invierno, las precipitaciones en forma nieve, compuestas en su mayor parte de agua marina evaporada, bastan normalmente para compensar el deshielo. Sin embargo, las altas y persistentes temperaturas registradas recientemente a causa del calentamiento global, son las responsables de que la cantidad de hielo que se derrite en verano haya aumentado y de que las nevadas hayan disminuido debido a que los inviernos se retrasan y las primaveras se adelantan, provocando que el nivel del mar se eleve. Desde el año 1970 el promedio global del nivel del mar ha estado subiendo a un ritmo de  $2,0 \pm 0,3$  mm/año, de los cuales  $0,8 \pm 0,3$  mm/año (Garduño, 2004), Cuando el nivel del mar se eleva con rapidez, tal y como ha estado haciéndolo en los últimos tiempos, incluso un pequeño aumento puede tener consecuencias devastadoras en los hábitats costeros. El agua de mar penetra en zonas cada vez más alejadas de la costa, lo cual puede generar consecuencias catastróficas como la erosión, las inundaciones de humedales, la contaminación de acuíferos y la pérdida del hábitat debido al periodo de adaptación que cada especie sufrirá. Los ciclos de cada organismo se ven influenciado directamente por la temperatura y precipitación, en el futuro con el aumento de temperaturas y variaciones en el ciclo hídrico; tanto como vegetaciones y hasta insectos se verían obligados a cambiar sus ciclos biológicos creando con esto una extinción masiva de nuestra biodiversidad afectando con esto las actividades humanas y la producción de productos.

Si bien no se conoce bien que los efectos de la exposición a la contaminación atmosférica son múltiples y de diferente severidad, siendo los más afectados los sistemas respiratorio y cardiocirculatorio. En los últimos años se han llevado a cabo diversos proyectos multicéntricos utilizando técnicas de análisis de series temporales. En Europa el proyecto APHEA, (por sus siglas en inglés Agency for Public Health Education Accreditation) y en Estados Unidos el estudio NMMAPS (por sus siglas en inglés National Morbidity, Mortality, and Air Pollution Study se encuentran entre los que han aportado más al conocimiento del impacto agudo de la contaminación en la salud realizándose estudios multicéntricos que han valorado el impacto de la contaminación en las principales ciudades de cada uno de estos países. En general el contaminante más estudiado ha sido las partículas, encontrándose que un incremento de  $10 \mu\text{g}/\text{m}^3$  en los niveles atmosféricos de  $\text{PM}_{10}$  se asocia, según los estudios, con un aumento de un 0,2 a un 1% en la mortalidad por todas las causas y un 0,5 a un 2 % en la mortalidad

cardiorrespiratoria. En estados Unidos y como parte del Estudio para la Prevención del Cáncer. En total se recogieron desde 1982 datos sobre factores de riesgo y contaminación atmosférica para unos 500.000 adultos de 151 áreas metropolitanas de los Estados Unidos. (Health Effects Institute, 2018)

En marzo de 2002 se publicaron los resultados del seguimiento de dicha cohorte hasta el año 1998. Las partículas finas ( $PM_{2.5}$ ) y los óxidos de azufre mostraron una asociación con la mortalidad para todas las causas, para enfermedades del aparato circulatorio y por cáncer de pulmón. Cada aumento de  $10 \mu g/m^3$  en los niveles atmosféricos de partículas finas se asoció aproximadamente con un aumento de un 4%, 6%, y 8% respectivamente del riesgo de morir por alguna enfermedad cardiovascular o respiratoria. (Health Effects Institute, 2018) Todas las poblaciones se verán afectadas por el cambio climático, pero algunas son más vulnerables que otras. Los habitantes de los pequeños estados insulares en desarrollo y de otras regiones costeras, megalópolis y regiones montañosas y polares son especialmente vulnerables, según una investigación realizada por la Organización Mundial de la Salud (OMS) donde solo se toman en cuenta posibles repercusiones sanitarias se estimó que el cambio climático provocará anualmente unas 250.000 defunciones adicionales entre 2030 y 2050; 38.000 por exposición de personas ancianas al calor; 48.000 por diarrea; 60.000 por paludismo; y 95.000 por desnutrición infantil. (OMS, 2018)

## **2.2. Contaminantes Atmosféricos**

Se denomina gas de efecto invernadero al componente gaseoso de la atmósfera, natural o antropógeno, que absorbe eficazmente la radiación infrarroja emitida por la superficie de la Tierra, por la propia atmósfera y por las nubes, dando lugar al denominado efecto invernadero y estos pueden ser contaminantes primarios que son emitidos de manera directa por la atmósfera; en esta clase se describe al dióxido de azufre; el cual además de ser irritante para la vegetación de todos los ambientes, también suelen ser irritantes para los pulmones y todo su funcionamiento como los  $PM_{10}$  que proceden de emisiones vehiculares, procesos industriales y del humo del tabaco y que son partículas gruesas que constan de un elemento natural por lo cual se clasifica dentro de los contaminantes primarios. Estas sustancias se van generando posterior a procesos de evaporación resultando en minerales locales, aerosoles, restos vegetales; así como asfalto erosionado y restos de neumáticos formando así contaminantes secundarios los cuales son aquellas sustancias originarias de procedimientos químicos que actúan en conjunto con los contaminantes primarios y dejan múltiples secuelas en la atmósfera.

Estos son parte de los contaminantes que mayor actividad tienen en contra del contenido atmosférico como los  $PM_{2.5}$  que son partículas finas que tienen una composición tóxica de origen antropogénico; especialmente dadas por emisiones de vehículos cuando se han formado totalmente por contaminantes secundarios, El vapor de agua ( $H_2O$ ), el dióxido de carbono ( $CO_2$ ), el óxido nitroso ( $N_2O$ ), el metano ( $CH_4$ ) y el ozono ( $O_3$ ) son los gases de efecto invernadero primarios de la atmósfera terrestre. La atmósfera contiene, además, cierto número de gases de efecto invernadero enteramente antropogénicos, como los halocarbonos u otras sustancias que contienen cloro y bromo. De acuerdo con el protocolo de Kioto se enumeraron seis gases o grupos de gases de efecto invernadero (GEI) como los constituyentes principales del total de las emisiones procedentes de las actividades humanas; el  $CO_2$ , el  $CH_4$ , los óxidos de nitrógeno, los hexafluoruros de azufre, los hidrocarbonados y los perfluorocarbonados (PFCs).

La exposición a estos contaminantes tienen graves consecuencias para la salud humana, un reciente estudio realizado por el Comité de contaminación atmosférica y sus relaciones con la salud del Ministro de Salud de la India, reveló que la exposición a ambientes con grados altos de estos gases conlleva al aumento a tener mayores problemas crónicos respiratorios como el asma, un mayor riesgo mortal de enfermedades no contagiosas como obstrucciones crónicas pulmonares, enfermedades cardiovasculares así hasta cáncer de pulmón en adultos mayores y niños, además el estudio reveló que en el 2015 un aumento a la exposición de  $PM_{2.5}$  en la India trajo como resultado un millón de muertes que asociadas a este factor, esto simbolizó casi un cuarto de las muertes referidas a la contaminación atmosférica en el mundo siendo la cifra exacta 4.2 millones de personas, este número se ha ido incrementando en la India y en el mundo a lo largo de 25 años, si las pérdidas humanas se trasladan a términos económicos se considera que se ha perdido más de 225 billones de dólares en mano de obra y 5.11 trillones de dólares en pérdidas del bienestar según un análisis realizado por el Banco Mundial en conjunto con el Instituto de evaluaciones y mediciones para la salud. (Cohen, 2003)

Otro efecto es el causado por el Monóxido de Carbono (CO) que es un gas incoloro, inodoro e insípido, ligeramente menos denso que el aire, siendo su origen artificial las combustiones



incompletas, el cual al ser liberado a la atmósfera puede permanecer ahí aproximadamente dos meses, para eventualmente reaccionar con otros compuestos y formar dióxido de carbono.

El monóxido de carbono al ser inhalado se une con la hemoglobina en las células rojas formando carboxihemoglobina reduciendo la capacidad de traer oxígeno en la hemoglobina lo que interfiere con su liberación en los tejidos y produce hipoxia y alteraciones del funcionamiento celular (en las neuronas, en las células del corazón y de otros músculos). Esto conlleva efectos cardiovasculares como aterosclerosis y arritmias, todo esto dependiendo de la concentración en partes por millón que se encuentren en el aire como se muestra en la tabla 2.1:

Tabla 2.1: Efectos del monóxido de carbono en el humano dependiendo su concentración (OMS 2018).

<b>Concentración CO [ppm]</b>	<b>Concentración CO [%]</b>	<b>Síntomas</b>
<b>Menos de 35</b>	0.0035	Ningún efecto en adultos sanos.
<b>100</b>	0.01	Fatiga, ligeros dolores de cabeza.
<b>200</b>	0.02	Mareos, náuseas.
<b>400</b>	0.04	Migrañas.
<b>800</b>	0.08	Desmayos si la exposición es prolongada.
<b>1,500</b>	0.15	Convulsiones, si la exposición es por más de una hora puede llegar a ser mortal.
<b>3,000</b>	0.30	Mortal si se tiene una exposición de más de 30 minutos.
<b>6,000</b>	0.60	Mortal si la exposición es entre 15 y 30 minutos.
<b>12,000</b>	1.2	Muerte instantánea

Además de que estudios realizados por el Instituto de los Efectos de la Salud de los Estados Unidos (Health Effects Institute “HEI”, por sus siglas en inglés), en países desarrollados y en vías

de desarrollo, se ha demostrado que existe un incremento del 2% al 4% en enfermedades cardiovasculares y de la arteria coronaria debido principalmente al aumento de CO, además de que se demostró en California que la mayoría de los niños han tenido un aumento en enfermedades respiratorias debido al incremento a la exposición a gases contaminantes en horas donde la actividad vehicular ha aumentado. Por lo que cada vez más países se ven en la necesidad de realizar más estudios sobre como los gases de efecto invernadero afectan la salud pública. (Health Effects Institute, 2018)

## Capítulo 3. Percepción Remota

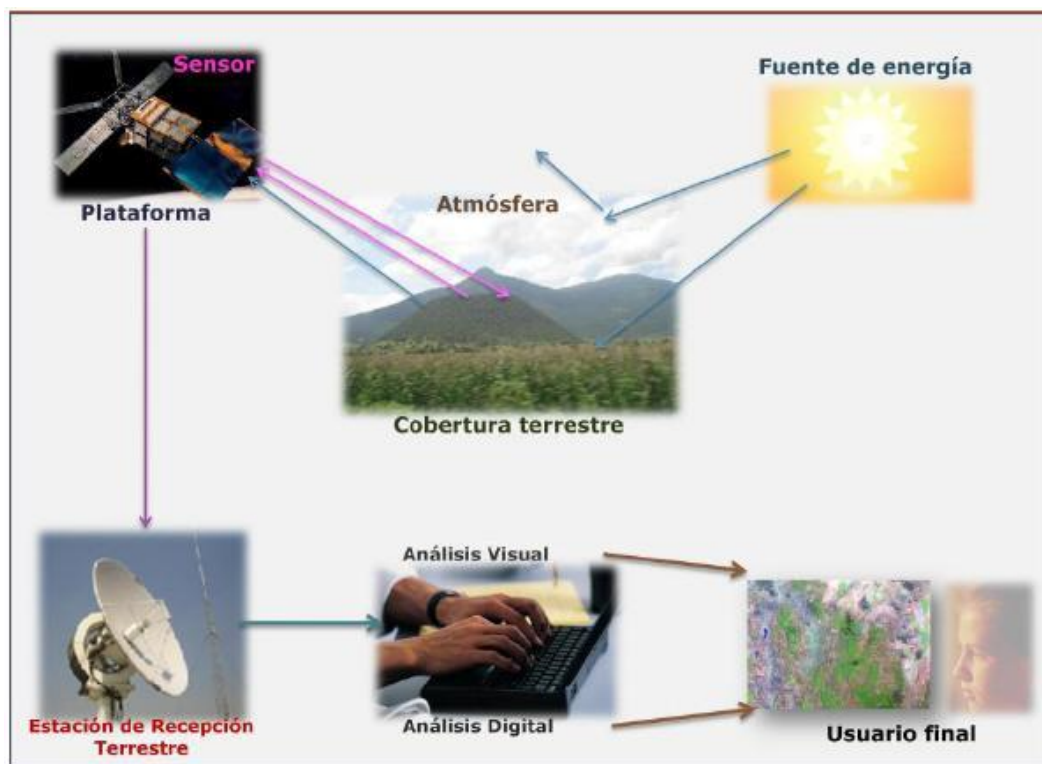


Imagen 3.1: Diagrama de los elementos del proceso de percepción remota o teledetección (INEGI 2018)

El Instituto Nacional de Estadística y Geografía (INEGI), define a la percepción remota como “la ciencia o técnica, que permite observar y obtener información del planeta, desde el espacio de manera indirecta, es decir sin tener contacto con la Tierra”, otra definición de la percepción remota puede ser:” Es la utilización de sensores para la adquisición de información sobre objetos o fenómenos sin que haya contacto directo entre el sensor y los objetos”. La historia de la Percepción Remota primitiva comenzó hace más de 600 millones de años cuando, un ente físico desarrollo células fotosensibles capaces de capturar la información de una fuente emisora como la luz y procesarla para así empezar a distinguir el medio que lo rodeaba, pero tuvo que pasar mucho tiempo para que dicho rudimento fotosensible evolucionará convirtiéndose así en un poderoso y sofisticado sensor como él ojo humano, con el paso del tiempo y siendo el ojo humano poderoso pero limitado el ser humano se vio en la necesidad de crear tecnologías capaces que le ayudarán en sus actividades, la cámara fotográfica fue un gran imitador mecánico del ojo humano que ayudo a que la investigación de la percepción remota incrementara de mayor manera. La segunda guerra mundial impulso la fotografía aérea, siendo de las primeras formas de percepción remota realizadas por el hombre, sin embargo, el salto fundamental de esta disciplina no fue hasta los años 60’s cuando las plataformas satelitales reemplazaron a las aéreas y los sensores electrónicos multiespectrales, acoplados a computadoras, fueron reemplazando poco a poco a las cámaras fotográficas.

### **3.1 Elementos de la Percepción Remota**

El INEGI y otros autores indican que la percepción remota consta de una serie de pasos, los cuales implican interacción entre el sensor, la superficie y el usuario, debiendo de tener en consideración que la percepción remota depende de la radiación que emiten los objetos y la reflectancia siendo los pasos sugeridos por INEGI y el proceso como se muestra en la imagen 3.1:

1. Fuente de energía o iluminación. Se denomina así al emisor de la energía lumínica o radiación electromagnética que utiliza el sistema de recepción remota La fuente puede ser externa al sensor remoto, caso que aplica para los sistemas pasivos, o puede ser propia del sensor remoto a través de un haz de energía, caso que aplica para los sistemas activos. Los sistemas pasivos suelen utilizar la radiación del Sol y los sistemas activos envían microondas para que choquen con el objeto y se devuelvan enseguida al sensor, brindando la posibilidad de obtener mejor información de la cobertura de manera espontánea y constante.

2. Radiación y medio de transmisión. Como la atmósfera es el medio por el cuál la energía viaja de la fuente al objeto y viceversa, está siempre modificada, al pasar por ella, la fuerza y la distribución espectral de la energía que recibe el sensor.
  
3. Interacción energía-objeto. La energía interactúa con el objeto dependiendo de las propiedades físicas del objeto y la radiación, con ello el objeto puede reflejar la energía que recibe, dispersarla y absorberla, dichas interacciones están siendo estudiadas.
  
4. Percepción de energía por el sensor. Este graba la radiación electromagnética reflejada o emitida por agentes externos, el sensor ideal debe ser altamente sensible a las distintas longitudes de onda que se emiten de diversos objetos.
  
5. Transmisión, recepción y procesamiento. La energía grabada por el sensor se transmite a una estación receptora, en donde los datos se procesan y son convertidos a imágenes digitales.
  
6. Interpretación, análisis y aplicación. Consiste en interpretar la imagen para extraer la información de los objetos captados y aplicar la información recolectada en el estudio de interés.

## **3.2 Sistemas de la Percepción Remota.**

De acuerdo a la plataforma donde se ubique el sensor, se distinguen dos grandes tipos de sistemas de percepción remota:

### **3.2.1 Percepción Remota Aérea.**

Se considera percepción remota aérea a aquella que se obtiene a partir de plataformas como un avión, dron, helicóptero, globo, incluyendo sensores como una cámara fotográfica, multibanda, digital, de video, barredor multiespectral o un barredor térmico. Las fotografías aéreas son las imágenes de percepción remota más comúnmente utilizadas, ya sea para análisis de recursos naturales, cartográficos y estudios de aproximación para sectores público y privado.

Las ventajas que esta técnica ofrece son varias como; el permitir tener una vista de los detalles de la superficie en su contexto espacial a menor escala que la imagen satelital, esto lo convierte en una herramienta muy útil cuando se busca trabajar en áreas pequeñas, por el hecho de utilizar plataformas aéreas para embarcar los sensores de imagen, es muy fácil rediseñar el plan de trabajo según las necesidades que van surgiendo, además de que la fotografía nos da una visión instantánea de los procesos bajo condiciones dinámicas, la fotografía permite desplazarse al lugar de interés sin demoras. El sensor usado en esta técnica es la película fotográfica y a las que estamos más acostumbrados son aquellas que son capaces de obtener la imagen de forma similar a como lo hace un ojo humano.

### **3.2.1 Percepción Remota Espacial**

Se considera percepción remota espacial aquella que se obtiene a partir de las plataformas como pueden ser plataformas espaciales tripuladas, estaciones orbitales o satélites autónomos, que giran alrededor de la Tierra en órbitas polares, oblicuas o ecuatoriales, las ventajas que tiene esta técnica con respecto a la otra es la visión sinóptica, la periodicidad en la toma de datos, el sincronismo que tiene con el sol, el bajo costo por hectárea de los datos obtenidos y la cantidad de información que se puede recopilar, de la imagen 3.2 a la 3.4, se pueden observar diferentes imágenes satelitales tomadas desde el Landsat 7 y comparando su resolución.

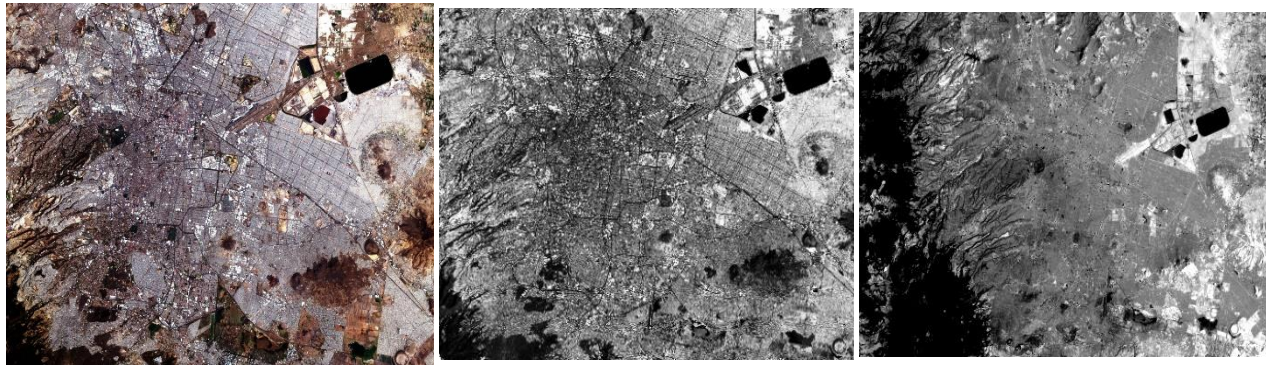


Imagen 3.2. Imagen LANDSAT7 formada por las bandas 1,2 y 3, con 30 m. de resolución espacial, que muestra la Ciudad de México (INEGI 2018)

Imagen 3.3. Imagen LANDSAT7 pancromática con 15 m. de resolución térmica con 60 m. de resolución espacial, que muestra la Ciudad de México (INEGI 2018)

Imagen 3.4. Imagen LANDSAT7 espacial, que muestra la Ciudad de México (INEGI 2018)

### 3.3 Plataformas de la Percepción Remota Espacial

#### 3.3.1 Satélites Landsat TM

El primer satélite puesto en órbita para la evaluación de los recursos terrestres por la agencia espacial norteamericana (por sus siglas en inglés NASA) ocurrió en el mes de julio de 1972, dando inicio con esto, a la nueva era en la forma de captar información sobre los recursos terrestres. Esta familia de satélites denominada Landsat han constituido uno de los proyectos más fructíferos de la teledetección espacial, la buena resolución de sus sensores, el carácter global, el periodo de las observaciones que realizan y su buena comercialización han ayudado a que estos satélites se conviertan en unos de los más utilizados en la actualidad.

De los cinco satélites Landsat enviados al espacio, los tres primeros tenían una forma similar a la de una mariposa, debido a su disposición de los paneles solares, su órbita era helio síncrono, polar, con una inclinación de 9.1 grados, una altura orbital de 917 kilómetros que circundaban la Tierra cada 103 minutos que equivalen a 14 órbitas diarias para volver a la misma posición de la superficie terrestre a los 18 días y a la misma hora local. El Landsat 4 y 5 modificaron su fisionomía y sus características orbitales. La altura de vuelo se reduce de 917 kilómetros a 705, se mejora el ciclo de recubrimiento a 16 días, debido a un menor periodo orbital, el cual es de 98.9 minutos, mientras que el Landsat 7 fue lanzado en abril de 1999 con un nuevo sensor denominado ETM+ (Enhanced Thematic Mapper Plus), una imagen de este satélite está compuesta por 8 bandas espectrales que pueden ser combinadas de distintas maneras para formar varias composiciones de color u opciones de procesamiento, una órbita del Landsat 7 se realiza en aproximadamente

99 minutos, permitiendo al satélite dar 14 vueltas a la Tierra por día y cubrir la totalidad del globo en 16 días, la imagen cubre áreas de 165x165 kilómetros lo que facilita que el proceso de captura de imágenes. El Landsat 8 es el satélite de observación más reciente, puesto en órbita en febrero del 2013, el satélite completa su órbita de 705 kilómetros de altura cada 99 minutos con un ciclo de recubrimiento de 16 días con un desfase de 8 días con respecto al Landsat 7 adquiriendo 650 imágenes diariamente, cuenta con dos sensores el Operational Land Imager (OLI) y el Thermal Infrared Sensor (TIRS) que completan las 11 bandas espectrales las demás características se observan en la tabla 3.1 y 3.2.

Tabla 3.1: Características Técnicas Landsat 8 (INEGI 2018)

<b>Atributo</b>	<b>Descripción</b>
Tipo de Producto	Imagen L1T (Precision, Terrain Corrected) - ortorrectificada
Espaciamiento de píxel	15 m Pancromático 30 m Multiespectral
Tamaño de Producto	185 x 185 km
Correcciones Geométricas	Proyectadas a un datum y proyección cartográfica, usando un Modelo Digital de Elevación (SRTM) y puntos de control terrestre (GCP)
Proyección Cartográfica	UTM (Universal Transversa de Mercator)
Bandas Espectrales	11 bandas: Aerosol costero (Coastal Aerosol)      430 – 450 nm Azul (Blue)                                      450 – 510 nm Verde (Green)                                   530 – 590 nm Rojo (Red)                                        640 – 670 nm Infrarrojo cercano (NIR)                    850 – 880 nm SWIR 1 (SWIR 1)                            1570 – 1650 nm SWIR 2 (SWIR 2)                            2110 – 2290 nm Pancromático (Pan)                           500 – 680 nm Cirrus (Cirrus)                                1360 – 1380 nm Infrarrojo térmico (TIRS ) 1              10600 – 11190 nm Infrarrojo térmico (TIRS ) 2              11500 – 12510 nm
Resolución Radiométrica	12 bits



Tabla 3.2: Bandas y aplicaciones de los satélites Landsat TM (CONAGUA)

Banda	Rango espectral	Resolución	Características y aplicaciones
TM1	0.45 – 0.52 Visible Azul-Verde	30 m	Batimetría en aguas de menor turbidez, diferencia suelo y vegetación, tipos de suelos.
TM2	0.52 – 0.60 Visible verde	30 m	Indicador de tasa de crecimiento y vigor de vegetación, estimación de la concentración de sedimentación.
TM3	0.63 – 0.69 Visible rojo	30 m	Absorción de clorofila y diferenciación de especies, aplicaciones geológicas.
TM4	0.76 – 0.90 Infrarrojo solar cercano	30 m	Delineación de cuerpos de agua, variaciones de biomasa.
TM5	1.55 – 1.75 Infrarrojo solar medio	30 m	Humedad de estrés de vegetación y minerales.
TM6	10.4 – 12.5 Infrarrojo térmico emitido	120 m	Temperatura de superficie aparente, separación urbano y uso de suelo, distinción de áreas quemadas.
TM7	2.08 – 2.35 Infrarrojo solar	30 m	Zonas hidrotermales alteradas, exploración minera, discriminación de tipos de suelo.
TM8	0.43 – 1.39 OLI 10.30 – 12.50 TIRS	30 m 100 m	Temperatura de superficie aparente Zonas hidrotermales alteradas, exploración minera.

### **3.3.2 Sensor MODIS**

El sensor MODIS (Moderate Resolution Imaging Spectroradiometer) es un sensor que recolecta datos sobre el verdor de la vegetación, pero también ofrece otras capacidades como la habilidad de recolectar datos a una resolución de 250 metros, comparada con los 1000 metros de resolución que ofrece el sensor AVHRR, el sensor MODIS ofrece la posibilidad de dar índices de vegetación como el NDVI que es la diferencia normalizada de índice de vegetación así como el EVI, el índice vegetativo perfeccionado que sirve también para medir el verdor de la vegetación.

### **3.3.3 Satélites Spot**

Spot es el Sistema de observaciones terrestres diseñado en Francia bajo el control de la CNES (Centre National D'Études Spatiales), desde 1986 fecha del lanzamiento del primer satélite Spot, el Spot 2 se le unió en órbita en enero de 1990, seguidos por el Spot 3 en 1993, Spot 4 en 1998, Spot 5 en 2002, Spot 6 en 2012 y el Spot 7 lanzado en 2014 creando así un servicio de imágenes para clientes el Spot Image que es el líder actual a nivel mundial en distribución de imágenes satelitales y productos derivados de la observación satelital terrestre con aplicaciones en campos como la Ingeniería Civil, Cartografía, Geología, Agricultura, Telecomunicaciones entre otras. Cada satélite Spot se coloca en una órbita polar, circular, heliocéntrica y en fases con relación a la Tierra, cuenta con una inclinación de 98 grados teniendo un ciclo de 26 días, los satélites Spot cuentan con el sensor multispectral HRV (XS) integrado tienen un rango de la longitud de onda verde hasta el infrarrojo cercano. La cobertura HRV (XS) en sus primeros satélites es en tres bandas espectrales, con una resolución espacial de 20 metros, llevan un sensor pancromático (HRV-P) que cubre desde el rango espectral Verde hasta el Rojo del espectro visible en una sola franja del terreno con 10 metros de resolución ambos sensores cubren una franja de 60 kilómetros a lo largo de la trayectoria orbital, los satélites actuales integran sensores de una alta resolución visible infrarroja (HRVIR) y una alta resolución geométrica (HRG), en la tabla 3.3 se muestran las demás características técnicas de los distintos Spot que existen.

Tabla 3.3: Características Técnicas Spot (CONAGUA)

Sensores	Espectro Electromagnético	Tamaño de Píxel	Rango Espectral
Spot 3 Spot 2 Spot 1	Pancromático	10 m	0.50-0.73 $\mu\text{m}$
	B1: Verde	20 m	0.50-0.59 $\mu\text{m}$
	B2: Rojo	20 m	0.61-0.68 $\mu\text{m}$
	B3: Infrarrojo Cercano	20 m	0.78-0.89 $\mu\text{m}$
Spot 4	Monespectral	10 m	0.61-0.68 $\mu\text{m}$
	B1: Verde	20 m	0.50-0.59 $\mu\text{m}$
	B2: Rojo	20 m	0.61-0.68 $\mu\text{m}$
	B3: Infrarrojo Cercano	20 m	0.78-0.89 $\mu\text{m}$
	B4: Infrarrojo Medio (MIR)	20 m	1.58-1.75 $\mu\text{m}$
Spot 5	Pancromático	2.5 y 5 m	0.48-0.71 $\mu\text{m}$
	B1: Verde	10 m	0.50-0.59 $\mu\text{m}$
	B2: Rojo	20 m	0.61-0.68 $\mu\text{m}$
	B3: Infrarrojo Cercano	20 m	0.78-0.89 $\mu\text{m}$
	B4: Infrarrojo Medio (MIR)	20 m	1.58-1.75 $\mu\text{m}$
Spot 6 Spot 7	Pancromático	1.5 m	0.45-0.74 $\mu\text{m}$
	B1: Verde	10 m	0.53-0.59 $\mu\text{m}$
	B2: Rojo	20 m	0.62-0.69 $\mu\text{m}$
	B3: Infrarrojo Cercano	20 m	0.76-0.89 $\mu\text{m}$
	B4: Azul	20 m	0.45-0.52 $\mu\text{m}$

### 3.3.4 Plataforma NOAA

La National Oceanic and Atmospheric Administration (NOAA) dispone de su programa de estudios de los fenómenos climáticos, para la predicción del tiempo a corto plazo, utiliza satélites en órbita estacionaria, como el GOES (Geostationary Operational Environmental Satellites), mientras que para predicciones a largo plazo utiliza satélites en órbita polar, los POES (Polar Orbiting Operational Environmental Satellites). Los satélites polares llevan en su carga útil el sensor AVHRR (Advanced Very High Resolution Radiometer) que puede proporcionar información radiométrica de las zonas visible e infrarroja del espectro, lo que permite complementar las funciones de observación meteorológica como se observa en la tabla 3.4.

Tabla 3.4: Resolución Espacial y Espectral del satélite NOAA (CONAGUA)

Sensores	Canales	Espectro electromagnético	Tamaño de Píxel	Rango Espectral
AVHRR	Canal 1	Visible	1.1 Km	0.58 - 0.68 $\mu\text{m}$
	Canal 2	Infrarrojo cercano	1.1 Km	0.72 – 1.00 $\mu\text{m}$
	Canal 3	Infrarrojo medio	1.1 Km	1.58 – 1.64 $\mu\text{m}$
	Canal 4	Infrarrojo medio	1.1 Km	3.55 – 3.93 $\mu\text{m}$
	Canal 5	Infrarrojo térmico	1.1 Km	10.3 – 11.3 $\mu\text{m}$
	Canal 6	Infrarrojo térmico	1.1 Km	11.5 – 12.5 $\mu\text{m}$

Tabla 3.5: Resolución Espacial y Espectral del satélite GOES (CONAGUA)

Canales	Espectro electromagnético	Tamaño de Píxel	Rango Espectral
Canal 1	Visible	1.1 Km	0.65 $\mu\text{m}$
Canal 2	Infrarrojo medio	4.0 Km	3.90 $\mu\text{m}$
Canal 3	Infrarrojo medio	8.0 Km	6.75 $\mu\text{m}$
Canal 4	Infrarrojo térmico	4.0 Km	10.70 $\mu\text{m}$
Canal 5	Infrarrojo térmico	4.0 Km	12.00 $\mu\text{m}$

### 3.3.5 Satélites IKONOS

El satélite IKONOS es de los primeros de tipo comercial que permite obtener imágenes con 1 metro de resolución espacial, recogiendo información de cualquier zona promedio dos veces al día, cubriendo áreas de 20 000 kilómetros en una misma pasada y produciendo como resultado imágenes de un metro de resolución cada tres días y de dos metros de resolución todos los días, el satélite IKONOS orbita la Tierra cada 98 minutos a una altitud máxima de 680 kilómetros, en forma sincronizada con el Sol, pasando sobre un determinado lugar a las 10:30 hora local las diferentes resoluciones de esta plataforma se muestran en la tabla 3.6.

Tabla 3.6: Resolución Espacial y Espectral del satélite IKONOS (CONAGUA)

Sensores	Espectro electromagnético	Tamaño de Píxel	Rango Espectral
IKONOS	Pancromático	1.0 m	0.45 - 0.90 $\mu\text{m}$
	Azul	4.0 m	0.45 – 0.52 $\mu\text{m}$
	Verde	4.0 m	0.52 – 0.60 $\mu\text{m}$
	Rojo	4.0 m	0.63 – 0.69 $\mu\text{m}$
	Infrarrojo próximo	4.0 m	0.76 – 0.90 $\mu\text{m}$

### 3.3.6 Aplicaciones de los sistemas de percepción remota

Las aplicaciones que se les puede dar a la información recaudada de las imágenes y sus recursos han ayudado a diferentes ramas de la ciencia para dar una nueva visión de análisis que permite comprender los fenómenos climáticos, la forma en que vamos afectando al medio ambiente y realizar proyecciones futuras, a continuación, se ejemplifican algunas aplicaciones en distintos campos.

- Geología; Reconocimiento de tipos de rocas, revisión de mapas geológicos, delineación de rocas y suelos no consolidados.
- Agricultura; Clasificación de usos del suelo, determinación de vegetación, medición de extensión de cultivos.
- Recursos Hídricos; Cartas de inundaciones, determinación de límites, áreas y volúmenes de superficies acuáticas, determinación de profundidades de cuerpos de agua.
- Planeamiento urbano y cartografía; Planificación regional, actualización de mapas, cartografía de redes de transporte.
- Medio ambiente; Detección de contaminantes atmosféricos, estimación de modelos de erosión, seguimiento de incendios forestales y sus efectos, determinación de efectos de desastres naturales.

## Capítulo 4. Regresiones y parámetros utilizados para elección de modelos

## 4.1 Regresiones usadas comúnmente para la elección de modelos

En estadística, cuando se detecta que dentro de una muestra se detecta que entre dos o más variables hay una relación significativa, por lo que se busca obtener una fórmula matemática o modelo que relacione varias variables, a esto se le conoce como regresión. El análisis de regresión es ampliamente utilizado para la predicción y previsión, donde su uso tiene superposición sustancial en el campo de aprendizaje automático. El análisis de regresión se utiliza también para comprender cuáles de las variables independientes están relacionadas con la variable dependiente, y explorar las formas de estas relaciones. En circunstancias limitadas, el análisis de regresión puede utilizarse para inferir relaciones causales entre las variables independientes y dependientes. Sin embargo, esto puede llevar a ilusiones o relaciones falsas.

Cuando se requiere elegir de entre un conjunto de modelos, se debe contar con una medida que permita compararlos, en este caso se optó por usar estimaciones por medio de la *prueba de error* utilizando la suma de cuadrados residuales para la elección; la suma de mínimos cuadrados mide el *training error*, que es el error que se tiene al utilizar un modelo con los datos que se tienen, y por lo solo se puede emplear para comparar modelos que tienen el mismo número de variables o predictores lo que realmente permite cuantificar cómo de útil es un modelo no es el *training error* si no la *prueba de error*, por lo que es esta última medida en la que hay que basarse para elegir entre modelos con diferente número de variables, generalmente cuando se habla de una *prueba de error*, se equipara con la prueba de la Suma Cuadrática de los Residuos (Test MSE, por sus siglas en inglés), dividido por el número de observaciones y existen la posibilidad de estimar el test error por medio de aproximaciones como lo es el Criterio de Información de Akaike (AIC, por sus siglas en inglés) que es una medida de la calidad relativa de un modelo estadístico, para un conjunto dado de datos.

$$AIC = 2k - 2\ln(L)$$

Donde  $k$  es el número de parámetros en el modelo estadístico y  $L$  es el máximo valor de la función de verosimilitud para el modelo estimado. Dado un conjunto de modelos candidatos para los datos, el modelo preferido es el que tiene el valor mínimo en el AIC. Por lo tanto, AIC no sólo recompensa la bondad de ajuste, sino también incluye una penalidad, que es una función creciente del número de parámetros estimados. Algunos de los principios del AIC indican que a menor valor el modelo se comportara de manera más parsimoniosa, solamente tiene validez el AIC si se compara con diferentes modelos e hipótesis, el AIC se puede utilizar para comparar modelos lineales con

modelos no lineales, las comparaciones solo son válidas si se espera el mismo resultado o variable.

Otra aproximación es el Criterio de Información Bayesiano (*BIC, por sus siglas en inglés*), que se basa en el método AIC solo que la penalización en este caso  $k=2$  será  $\log(n)$ , siendo  $n$  el número de observaciones. Esto implica que, cuando  $n > 7$ , el método BIC introduce mayores penalizaciones, tendiendo a seleccionar modelos con menos predictores que los seleccionados por AIC.

$$BIC = \ln(n) * k - 2\ln(\hat{L})$$

Siendo  $\hat{L}$  el máximo valor de la función de verosimilitud, que depende de los datos observados y los parámetros evaluados en la función de verosimilitud. Y por último la  $R^2$  ajustada, siendo  $R^2$  el porcentaje de varianza de la variable respuesta por el modelo respecto del total de varianza observada. En los modelos múltiples, cuantos más predictores se incluyan en el modelo, mayor es el valor de  $R^2$ , ya que, por poco que sea, cada predictor va a explicar una parte de la varianza observada. La idea detrás de  $R^2$  ajustada es que, una vez que los predictores correctos se han incluido en el modelo, la varianza extra que se consigue explicar añadiendo más predictores no compensa la penalización. Al contrario del AIC y BIC mientras mayor sea el valor del  $R^2$  ajustado mejor es el modelo.

Una vez conocido los parámetros para obtener la prueba MSE, es necesario conocer las diferentes regresiones que se aplicaron; El primero en aplicarse fue una regresión de Subset Selection o de Selección de Variables que tienen la finalidad de identificar y seleccionar, de entre todos los predictores disponibles, aquellos que están más relacionados con la variable respuesta y así crear el mejor modelo, dentro de este método se encuentra la manera de realizar la regresión por la Mejor Selección de Variable (*BSS, por sus siglas en inglés*) y una regresión por pasos (*Stepwise Selection, en inglés*), esta puede ser una eliminación progresiva (*en inglés, Backward Stepwise Selection*), una introducción progresiva (*en inglés, Forward Stepwise Selection*) o en ambas direcciones.



El BSS consiste en evaluar todos los posibles modelos que se pueden crear por combinación de los predictores disponibles, el algoritmo a seguir es el siguiente:

1. Se genera lo que se conoce como modelo nulo  $M_0$ , que es el modelo sin ningún predictor.
2. Se generan todos los posibles modelos que contienen un único predictor y se selecciona el que tiene la menor Suma de Residuos Cuadrados (*RSS, por sus siglas en inglés*) o mayor  $R^2$ , Al modelo seleccionado se denomina  $M_1$
3. Se repite el paso anterior para modelos con dos predictores y así sucesivamente hasta llegar al modelo con todos los predictores  $M_k$  y se selecciona de todos el mejor por medio de las aproximaciones mencionadas anteriormente.

La introducción progresiva es una alternativa computacionalmente más eficiente BSS en la que no se evalúan todas las posibles combinaciones de predictores sino un subconjunto de estas. El proceso se inicia generando el modelo nulo  $M_0$  sin predictores. A continuación, se generan todos los posibles modelos que se pueden crear introduciendo un predictor más al modelo nulo. De entre todos estos modelos con 1 predictor se selecciona el mejor basándose en la Suma de Residuos Cuadrados (RSS) o  $R^2$  al modelo elegido se denomina  $M_1$  se repite el paso anterior pero esta vez partiendo del último modelo seleccionado y así sucesivamente hasta llegar al modelo con todos los predictores. De entre los mejores modelos seleccionados para cada número de predictores  $M_k$ . La eliminación progresiva es equivalente a la introducción progresiva, pero en este caso iniciando el proceso a partir del modelo que contiene todos los posibles predictores, hasta el modelo que incluya todas las variables. En cada iteración se generan todos los modelos que se pueden crear eliminando un único predictor a la vez y se selecciona el que tiene menor RSS o mayor  $R^2$ . El proceso se repite hasta llegar al modelo nulo sin predictores. De entre los mejores modelos seleccionados para cada número de predictores  $M_k$ .

Los métodos mencionados anteriormente emplean mínimos cuadrados para ajustar al modelo lineal que contiene únicamente un subconjunto de variables, por ello existe otra alternativa conocida como regularización, que consiste en ajustar el modelo incluyendo todos los predictores pero empleando un método que fuerce a que las estimaciones de los coeficientes de regresión lineal tiendan a cero, es decir que se tienda a minimizar la influencia de los predictores menos importantes, dos de los métodos más ocupados son la Regresión Contraída (*en inglés, Ridge*

*Regression*) que aproxima a cero los coeficientes de los predictores pero sin excluir alguno y el Lasso que aproxima a cero coeficientes excluyendo variables.

La regresión contraída es similar al ajuste por mínimos cuadrados en cuanto que ambos tratan de minimizar el *RSS*. La diferencia reside en que la regresión contraída incorpora un término llamado *shrinkage penalty* o penalización por contracción que fuerza a que los coeficientes de los predictores tiendan a cero. El efecto de esta penalización está controlado por el parámetro Lambda. Cuando Lambda es igual a cero la penalización es nula y los resultados son equivalentes a los obtenidos por mínimos cuadrados, cuando lambda tiende a infinito todos los coeficientes son cero, lo que equivale al modelo sin ningún predictor.

La principal ventaja del ajuste por la regresión contraída frente al ajuste por mínimos cuadrados es la reducción de varianza. Por lo general, en situaciones en las que la relación entre la variable respuesta y los predictores es aproximadamente lineal, las estimaciones por mínimos cuadrados tienen poco Sesgo, pero aún pueden sufrir alta varianza. Este problema se acentúa conforme el número de predictores introducido en el modelo se aproxima al número de observaciones de entrenamiento, llegando al punto en que, si  $p > n$ , no es posible ajustar por mínimos cuadrados. Empleando un valor adecuado de Lambda, identificado mediante la validación cruzada, el método de regresión contraída es capaz de reducir varianza sin apenas aumentar el sesgo consiguiendo así un menor error total. La limitación del método de ajuste por regresión contraída en comparación a los métodos de selección de variable es que el modelo final va a incluir todos los predictores. Esto es así porque, si bien la penalización empleada fuerza a que los coeficientes tiendan a cero, nunca llegan a ser exactamente cero. Este método consigue minimizar la influencia sobre el modelo de los predictores menos relacionados con la variable buscada, pero en el modelo final van a seguir apareciendo. Aunque esto no supone un problema para la precisión del modelo, sí lo es para su interpretación. Por otro lado, se encuentra el método *Lasso* que es una variante del método anterior, que permite al igual que el método de regresión contraída, forzar a que las estimaciones de los coeficientes de los predicadores tiendan a cero con la variante de que *Lasso*, es capaz de fijar algunos de ellos exactamente a cero, lo que permite además de reducir la varianza, que se realice una selección de predictores, generando así modelos más fáciles de interpretar. La diferencia que existe de un método con el otro depende del escenario donde se apliquen y la precisión del modelo obtenido depende de la elección del valor Lambda empleado, ya que este determina el grado de penalización. Mediante el método

de validación cruzada, es posible identificar la Lambda más adecuada a utilizar en el modelo original.

Los métodos que se describieron con anterioridad controlan la varianza ya sea utilizando un parámetro penalizador como lambda haciendo que los coeficientes tiendan a cero, en ambos casos se emplean variables sin ser estandarizadas, por eso existe el método de reducción de dimensiones que crean un número reducido de nuevas variables a partir de las combinaciones lineales de las variables originales y con ellas ajustar el modelo. De esta manera se consigue un modelo con menor número de predictores pero que abarcan la misma información que la que aportan todas las variables y para esto existen dos métodos más comunes como el análisis de componentes principales (*PCA, por sus siglas en inglés*) y la Regresión de mínimos cuadrados parciales (*PLS, por sus siglas en inglés*), por lo general ambos métodos tienen buenos resultados en aquellos casos en que las variables están altamente correlacionadas. Es importante tener en cuenta que, aunque permiten generar modelos que contienen un número menor de predictores, no se trata de un método de selección de variables, ya que las nuevas variables son combinaciones lineales de todos los predictores originales. Dado que no dejan de ser un ajuste lineal por mínimos cuadrados que emplean componentes principales como predictores, para que sean válidos se tienen que cumplir las condiciones requeridas para la regresión por mínimos cuadrados. La regresión del componente principal (*PCR, por sus siglas en inglés*), consiste en ajustar un modelo de regresión lineal mediante mínimos cuadrados empleando como predictores las componentes generadas a partir de un *PCA*. De esta forma, con un número reducido de componentes se puede explicar la mayor parte de la varianza. Para método sigue el siguiente algoritmo:

- Los componentes principales se crean como combinación lineal de las variables originales.
- El vector de la primera componente principal (*first principal component*) define la línea más próxima a todos los puntos.
- La dirección del vector de la primera componente es aquella en la que las observaciones tienen mayor varianza.
- Proyectar un punto en una línea o un plano consiste en encontrar la localización de dicha línea o plano más próxima al punto.
- La segunda componente principal es la combinación lineal de las variables que explica mayor varianza después de la primera componente y que no está correlacionada con esta

última. La condición de no correlación entre componentes principales equivale a decir que las direcciones de las componentes son perpendiculares u ortogonales.

- Cuando el número de componentes es igual al número de predictores originales, el resultado de *PCR* es equivalente al de regresión por mínimos cuadrados.
- El número óptimo de componentes principales se puede elegir mediante la validación cruzada.
- Cuando se realiza *PCR* hay que estandarizar los predictores antes de calcular las *PCA*, de lo contrario, las variables que se miden en una escala mayor o las que presenten mayor varianza tendrán más peso. Si todos los predictores se miden con la misma escala y presentan la misma varianza, entonces no es necesaria la estandarización.

El último método ocupado en este proyecto fue el *PLS* es similar al *PCR* en cuanto que ambos emplean las componentes principales resultantes de un análisis *PCA* como predictores. La diferencia reside en que, mientras *PCR* ignora la variable respuesta *Y* para determinar las combinaciones lineales, *PLS* busca aquellas que, además de explicar la varianza observada, predigan un comportamiento futuro. Como se ha descrito a lo largo de este capítulo, no existe un método que por defecto supere a los otros ya que esto depende del escenario que se tenga ya sea con pocos predictores importantes, muchos predictores importantes, muchos predictores correlacionados, etc. Un método puede superar de forma sustancial a los otros, o puede que todos alcancen aproximadamente la misma precisión. La forma de identificar cuál es el mejor método para un determinado estudio consiste en dividir las observaciones disponibles en dos grupos (*training* y *test*). Siguiendo los pasos de cada uno de los métodos se ajusta el modelo empleando únicamente el *trainig set* y se calcula el  $MSE = (\text{predicción} - \text{valor real})^2$  utilizando el *test set*. Aquel método con el que se obtenga menor *MSE* es el que mejor precisión logra.

## 4.2 Parámetros utilizados en la elección de modelos

Uno de los parámetros más usados en la elección de modelos dentro de la estadística es el coeficiente de determinación o el  $R^2$ , que tiene como principal propósito el predecir futuros resultados, también este parámetro determina la calidad del modelo para replicar los resultados y la proporción de variación de los resultados que puede explicarse por el modelo, Siempre aumenta como el número de efectos que se incluyen en el modelo.

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}$$

$$SST = \sum (y - \bar{y})^2 \rightarrow \text{Sumatoria de cuadrados total}$$

$$SSR = \sum (\hat{y} - \bar{y})^2 \rightarrow \text{Sumatoria cuadrática de regresiones}$$

$$SSE = \sum (y - \hat{y})^2 \rightarrow \text{Sumatoria cuadrática de errores}$$

Por lo que  $R^2$  ajustado, que es otro parámetro, intenta corregir esta estimación excesiva. El  $R^2$  ajustado se calcula dividiendo el error cuadrático de la media residual por el error cuadrático promedio total que es la varianza muestral del campo objetivo y se resta 1 del resultado. El  $R^2$  ajustado siempre es menor o igual que  $R^2$ . Un valor de 1 indica un modelo que predice perfectamente valores del campo objetivo. Un valor que es menor o igual que 0 indica un modelo que no tiene ningún valor predictivo. En el mundo real, el  $R^2$  y el  $R^2$  ajustado se encuentra entre estos valores. También como se explica anteriormente otros parámetros son AIC y BIC que son medidas de la calidad relativa de un modelo estadístico, para un conjunto dado de datos.

Por otro lado, la desviación estándar o  $\sigma$ , que es una medida de dispersión para variables cuantitativas y de intervalo, que indica qué tan dispersos están los datos con respecto a la media. Mientras mayor sea la desviación estándar, mayor será la dispersión de los datos. La desviación estándar se puede utilizar para establecer un valor de referencia para estimar la variación general de un proceso y está se calcula como a raíz cuadrada de la varianza.

También se puede obtener el error estándar que es una estimación de cuánto varía el valor de una estadística de prueba de muestra a muestra. Es una medida de la incertidumbre de la estadística de prueba. El error estándar podría abreviarse como error est, el error estándar se calcula tomando la desviación estándar de la distribución de muestreo para la estadística de prueba, la distribución de muestreo es la distribución de todas las muestras posibles. Por último, el p.value o valor de p se define como la probabilidad correspondiente al estadístico de ser posible bajo la hipótesis nula, si cumple con la condición de ser menor al nivel de significancia impuesto arbitrariamente, entonces la hipótesis nula será, eventualmente, rechazada.

El p-valor es una medida directa de lo verosímil que resulta obtener una muestra como la actual si es cierta  $H_0$ . Los valores pequeños indican que es muy infrecuente obtener una muestra como la actual, en cambio, los valores altos que es frecuente. El p-valor se emplea para indicar cuánto (o cuán poco) contradice la muestra actual la hipótesis alternativa. Informar sobre cuál es el p-valor tiene la ventaja de permitir que cualquiera decida qué hipótesis acepta basándose en su propio nivel de riesgo  $\alpha$ . Esto no es posible cuando se informa, como ha sido tradicional, indicando sólo el resultado de la decisión, es decir, si se acepta o se rechaza  $H_0$  con un  $\alpha$  fijo. Al proporcionar el p-valor obtenido con la muestra actual, la decisión se hará de acuerdo con la regla siguiente:

si  $p_v \leq \alpha$ , aceptar  $H_1$

si  $p_v > \alpha$ , aceptar  $H_0$

## Capítulo 5. Metodología

En la actualidad existen diferentes maneras de medir el monóxido de carbono con equipos especializados como los sensores electroquímicos; que son sensores de estado sólido tienen una barrera de difusión capilar que controla la entrada del monóxido de carbono al interior de una cápsula en la cápsula se encuentran dos electrodos y un electrolito, generalmente ácido sulfúrico, otro manera de medirlo es con un detector de infrarrojo disperso o NDIR (*por sus siglas en inglés*) como lo hace la SEDEMA para la medición de monóxido de carbono en las estaciones de monitoreo en la Ciudad de México, los cuales determinan la concentración de monóxido midiendo la absorción de la longitud de onda del gas en el infrarrojo, pero debido a que estos equipos llegan a tener fallas y se debe de contar con personal para mantenimiento, se ha optado analizar la relación existente entre las mediciones de campo de monóxido reportadas por medio de la RAMA y la reflectancia obtenida empleando la percepción remota, esto aunado a la gran investigación que ha tenido esta ciencia en los últimos años y la relevancia del monitoreo de los parámetros de calidad atmosférica, las imágenes utilizadas fueron obtenidas gracias al sensor OLI perteneciente a la plataforma Landsat8.

### 5.1 Estado del arte

La plataforma Landsat8 fue puesta en órbita en 2013, debido al sensor TIRS, pero especialmente al sensor OLI ofreciendo más información en las imágenes que va capturando y teniendo dos bandas específicamente para objetivos de aerosoles en la atmósfera brindando así la posibilidad de que cada vez sea mayor el uso de la percepción remota para mediciones de contaminantes de manera indirecta. Por una parte, M Hamzelo et al. (2013) establecen un método para medir la concentración de monóxido de carbón en la ciudad de Terán en Irán por medio de imágenes obtenidas del Landsat8 utilizando el Spectral Angle Mapper (SAM) algoritmo que diferencia cada píxel usando una curva de reflectancia como parámetro de selección para finalmente presentar un mapeo elaborado en un software de información geográfica y correlacionando los datos de la red de monitoreo y lo obtenido de las imágenes. Otros investigadores han utilizado la plataforma TERRA de la NASA integrado con el sensor MODIS-MOPITT que mide específicamente contaminantes atmosféricos en especial el monóxido de carbono, pero en este proyecto se analizara la correlación con imágenes obtenidas del Landsat8. (Mohammed Hashim, 2010, Hung Nguyen, 2014, Elbir, 2010)

## 5.2 Parámetros de la calidad del aire

La contaminación atmosférica ha sido producto del proceso de la industrialización, así como de las grandes concentraciones urbanas, primordialmente por la emisión de humos, polvos y gases provenientes de fuentes móviles y fijas, por lo que fue necesario crear parámetros de calidad del aire que indiquen que tan contaminado se encuentra el aire, así como que tan perjudicial sería para la salud el respirar el aire en ese estado y realizar actividades en el exterior, debido a que se pueden contraer enfermedades en altas concentraciones el monóxido de carbono puede ser fatal. La intoxicación por este contaminante es uno de los tipos más comunes de envenenamiento, puede inhabilitar el transporte de oxígeno hacia las células y provocar mareos, dolor de cabeza, náuseas, estados de inconsciencia e inclusive la muerte. Se estima que cada año mueren alrededor de 18 000 mexicanos por contaminación del aire (OMS 2002) por lo que es necesario monitorear los contaminantes más dañinos a la salud. La peligrosidad del monóxido de carbono radica en que dicho gas presenta una gran afinidad por la hemoglobina dando lugar a la carboxihemoglobina, que al desplazar al oxígeno de la hemoglobina de la sangre provoca isquemia e hipoxia, situaciones a las que el corazón y el cerebro son especialmente sensibles, algunos estudios indican que más del 50% de los envenenamientos por monóxido de carbono son debidos a vehículos a motorizados y de estos más del 80% corresponde a vehículos estacionados (Martínez, 2015). Los estacionamientos subterráneos se han convertido en un sumidero de monóxido de carbono aun cuando dispongan de puertas o ventanas abiertas, puesto que la ventilación pasiva no es suficiente para reducir el riesgo en espacios semicerrados. La inhalación de humos de todo tipo de fuegos es la segunda causa de envenenamiento por monóxido de carbono.

Se estima que en la Ciudad de México la fuente principal de este contaminante son los automóviles dado que existe un parque vehicular de 5 millones, siendo el sector que contribuye en mayor medida a la emisión de este contaminante, de acuerdo con datos de la SEDEMA el consumo energético en la Zona Metropolitana del Valle de México es de 543 petajule, de los cuales 236.6 corresponden a la gasolina.

Los contaminantes atmosféricos se dividen de acuerdo con su origen en dos grupos; los primarios, que son emitidos de manera directa a la atmósfera y los secundarios, que son aquellos gases o químicos que reaccionan en la atmósfera formando otros gases tóxicos, a su vez los contaminantes pueden ser de dos clases; los aerosoles, formados por una mezcla entre partículas



que se encuentran suspendidas en un gas y los contaminantes gaseosos cuya principal fuente artificial es la quema de combustible.

En la Zona Metropolitana del Valle de México hay una gran variedad de contaminantes gaseosos según la SEDEMA:

- Compuestos Fotoquímicos.
- Compuestos de Azufre; siendo la principal fuente la combustión de sustancias que contienen azufre como el carbono y combustibles generando lluvia ácida.
- Compuestos de Nitrógeno; emitidos por los escapes de los automóviles, son formados debido a la reacción de nitrógeno con oxígeno, este contaminante llega a formar partículas suspendidas.
- Compuestos de Carbón; Se emiten principalmente durante la quema de combustibles o de cualquier material orgánico que se realice en la atmósfera con una cantidad de oxígeno limitada, propiciando así una combustión incompleta.

Este estudio se concentrará únicamente en el monóxido de carbono; El monóxido de carbono (CO), es un gas inodoro e incoloro que se produce por la combustión incompleta de compuestos de carbono, consecuentemente pueden verterlo al aire los vehículos automotores y la industria, aunque en menor escala; algunos procesos naturales son capaces de emitirlo, tales como los incendios forestales o su emisión de los procesos naturales que se llevan a cabo en los océanos. Mención especial debe hacerse de la acumulación intramuros por procesos domésticos y el hábito de fumar, la NORMA Oficial Mexicana NOM-021-SSA1-1993, Salud ambiental. Criterio para evaluar la calidad del aire ambiente con respecto al monóxido de carbono establece que La concentración de monóxido de carbono, como contaminante atmosférico, no debe rebasar el valor permisible de 11.00 ppm o lo que es equivalente a 12,595 µg/m<sup>3</sup> en promedio móvil de ocho horas una vez al año, como protección a la salud de la población susceptible.

### 5.3 Zona de estudio

La Ciudad de México se encuentra dentro de las ciudades más contaminadas del mundo de acuerdo con un estudio realizado por la Organización Mundial de la Salud (OMS 2018); La Ciudad de México se localiza en el Valle de México formando así la Zona Metropolitana del Valle de México (ZMVM), encontrándose a 2240 metros sobre el nivel medio del mar, en la latitud 19°25'42"N y en la longitud 99°07'40"O. Cuenta con una superficie de 9,560 km<sup>2</sup>, se forma por las 16 delegaciones de la Ciudad de México, 59 municipios del Estado de México y un municipio del Estado de Hidalgo, presenta valles montañosos, mesetas y cañadas, así como terrenos semiplanos, en lo que alguna vez fueron los lagos de Texcoco, Xochimilco y Chalco; está integrada por una parte del Estado de México, el sur del Estado de Hidalgo, el sureste de Tlaxcala y casi la totalidad de la Ciudad de México, que cuenta con prominencias topográficas aisladas como el Cerro de la Estrella, el Peñón y el Cerro de Chapultepec, entre otros; su ubicación geográfica y su entorno característico ejercen una influencia determinante sobre la calidad del aire existente en esta zona. El área se extiende a una cuenca semicerrada, en la porción sureste del Valle de México, la cual está sujeta de manera natural a condiciones que no favorecen una adecuada ventilación de la atmósfera.

Entre los principales factores fisiográficos y climáticos que afectan la calidad del aire destacan los siguientes:

- La altitud a la que se ubica el Valle de México (2,240 msnm), determina que el contenido de oxígeno sea 23% menor que a nivel del mar, lo cual tiende a hacer más contaminantes los procesos de combustión.
- Las frecuentes inversiones térmicas que ocurren en el valle, en más del 70% de los días del año, son un fenómeno natural que causa un estancamiento temporal de las masas de aire en la atmósfera. Ello inhibe la capacidad de autodepuración de ésta y favorece la acumulación de los contaminantes. El estancamiento perdura hasta que, al transcurrir el día y de manera gradual, la inversión térmica se rompe debido al calentamiento de la atmósfera, entonces los contaminantes se dispersan.
- El entorno montañoso que rodea la cuenca constituye una barrera natural que dificulta la libre circulación del viento y la dispersión de los contaminantes. Por ello, es un medio propicio para la acumulación de los contaminantes atmosféricos. Las montañas que delimitan la cuenca alcanzan una altitud promedio de 3,200 metros, con elevaciones que superan los 5,400 metros.

- La intensa y constante radiación solar que se registra en el Valle de México a lo largo de todo el año, favorece la formación del ozono. Ello es resultado de las complejas reacciones que la luz ultravioleta del sol desencadena entre los óxidos de nitrógeno y los hidrocarburos emitidos a la atmósfera, los cuales son precursores del ozono.



Imagen 5.1: Zona Metropolitana del Valle de México. (INEGI)

Desde la elección de la Ciudad de México (antes Distrito Federal), en 1824, como sede de los poderes federales y centro económico nacional se ha ido favoreciendo la concentración de población en este espacio geográfico y las zonas aledañas, consecuentemente, a partir de la década de los cuarenta e inicio de los cincuenta principalmente, el crecimiento urbano de la ciudad rebasó sus límites presentándose un proceso de conurbación con algunos municipios del Estado de México, surgiendo así la Zona Metropolitana del Valle de México. El crecimiento del área metropolitana ha obligado a incluir cada vez más municipios, por esta razón y datos del INEGI, en 1950 se consideraban como parte de la Ciudad de México a 12 delegaciones políticas del Distrito Federal, en este año la superficie urbana del Distrito Federal se incrementó en un 26% con relación

a 1941, y se inicia la conurbación con 2 municipios del Estado de México, sumando una población aproximada de 2.9 millones de habitantes, de los cuales aproximadamente el 90% se concentraba en el Distrito Federal.

En 1998 la población creció a 16.7 millones de habitantes y cubría una superficie de 3,489 kilómetros cuadrados, actualmente cuenta con una población total de 19.2 millones de habitantes y proyecciones futuras aseguran que la Zona Metropolitana del Valle de México podrá convertirse en la primera Megalópolis uniéndose en unos años con la Zona Metropolitana del Bajío y con la Zona Metropolitana de Toluca.

Por otro lado, debido a la posición latitudinal del territorio nacional, el país se ve afectado por masas de aire con características del tipo polar, durante la etapa invernal y de tipo tropical en la etapa del verano, las cuales, debido a su posición geográfica, afectan directamente al Valle de México. De esta manera, el clima de tipo frío está determinado por sistemas meteorológicos que provienen desde la parte norte del continente y el clima de tipo cálido es determinado por la afluencia de aire tropical, normalmente húmedo, proveniente del Pacífico, Mar Caribe y del Golfo de México. Como resultado de lo anterior, los sistemas meteorológicos predominantes definen claramente dos épocas climáticas con particularidades bien definidas: la época de “lluvias” de junio a octubre, caracterizada por aire marítimo tropical con alto contenido de humedad y la época de “secas” que se identifica con humedad relativa baja, debido a que la masa de aire correspondiente normalmente es de tipo polar continental.

La época de secas puede ser subdividida en dos: Seca-Caliente (marzo-mayo), en la cual predomina aire con características tropicales pero seco, y la época Seca-Fría (noviembre-febrero) cuyos rasgos meteorológicos la definen como aire de tipo polar con bajo contenido de humedad. Lo antes mencionado, como resultante de la posición geográfica particular donde se encuentra ubicado el país y el conjunto de montañas que conforman la rica variedad orográfica. La cuenca de México está rodeada de cadenas montañosas, al norte se encuentra la sierra de Guadalupe, al este la sierra de Santa Catarina y al suroeste las sierras de Las Cruces, Ajusco y Chichinautzin lo que limitan la dispersión horizontal de contaminantes. Hacia el Noroeste (NE). Todos estos factores intensifican los periodos de contaminación conduciendo con frecuencia a la mala dispersión de contaminantes atmosféricos y procesos internos de transporte de contaminantes y mayores tiempos de residencia y por tanto de exposición de la población.

## Sitios de Muestreo

El sistema de monitoreo de la Ciudad de México es operado por la Secretaría del Medio Ambiente (SEDEMA). Al sistema lo componen las siguientes redes que ayudan al monitoreo atmosférico:

- Red Automática de Monitoreo Atmosférico (RAMA), cuenta con 34 estaciones y mide de manera continua al ozono, óxidos de nitrógeno, dióxidos de azufre, monóxido de carbono, el  $PM_{10}$  y el  $PM_{2.5}$ .
- Red de Meteorología y Radiación Solar (REDMET), que cuenta con 16 estaciones midiendo continuamente la temperatura, humedad relativa, dirección del viento, velocidad del viento, presión atmosférica, radiación fotosintética activa, radiación global, radiación solar UV-A y UV-B.
- Red Manual de Monitoreo Atmosférico (REDMA), cuenta con 12 estaciones midiendo de manera manual las concentraciones de partículas suspendidas como el PST, Plomo,  $PM_{10}$  y el  $PM_{2.5}$ .
- Red de Depósito Atmosférico (REDDA), que cuenta con 16 estaciones, colectando muestras de lluvia para determinar volumen, acidez, conductividad y el contenido de nutrientes.

En la tabla 5.1 se pueden observar las estaciones de la RAMA que solamente monitorean monóxido de carbono el cual es el contaminante por modelar, así como su ubicación.

Tabla 5.1: Estaciones del Sistema de Monitoreo de la Ciudad de México. (Elaboración Propia)

<b>Clave</b>	<b>Nombre</b>	<b>Latitud</b>	<b>Longitud</b>
ACO	Acolman	19.635501	-98.912003
AJM	Ajusco Medio	19.272100	-99.207658
ATI	Atizapán	19.576963	-99.254133
BJU	Benito Juárez	19.371612	-99.158969
CAM	Camarones	19.468404	-99.169794
CCA	Centro de Ciencias de la Atmósfera	19.3262	-99.1761
CHO	Chalco	19.266948	-98.886088
CUA	Cuajimalpa	19.365313	-99.291705
FAC	FES Acatlán	19.482473	-99.243524
HGM	Hospital General de México	19.411617	-99.152207
INN	Investigaciones Nucleares	19.291968	-99.38052
IZT	Iztacalco	19.384413	-99.117641
LLA	Los Laureles	19.578792	-99.039644
LPR	La Presa	19.534727	-99.11772
MER	Merced	19.42461	-99.119594
MGH	Miguel Hidalgo	19.404050	-99.202603
MON	Montecillo	19.460415	-98.902853
MPA	Milpa Alta	19.176900	-98.990189
NEZ	Nezahualcóyotl	19.393734	-99.028212
PED	Pedregal	19.325146	-99.204136
SAG	San Agustín	19.532968	-99.030324
SFE	Santa Fe	19.357357	-99.262865
SJA	San Juan de Aragón	19.446203	-99.207868
TAH	Tláhuac	19.246459	-99.010564
TLA	Tlalnepantla	19.529077	-99.204597
TLI	Tultitlán	19.602542	-99.177173
UAX	UAM Xochimilco	19.304441	-99.103629
UIZ	UAM Iztapalapa	19.360794	-99.07388
VIF	Villa de las Flores	19.658223	-99.09659
XAL	Xalostoc	19.525995	-99.0824

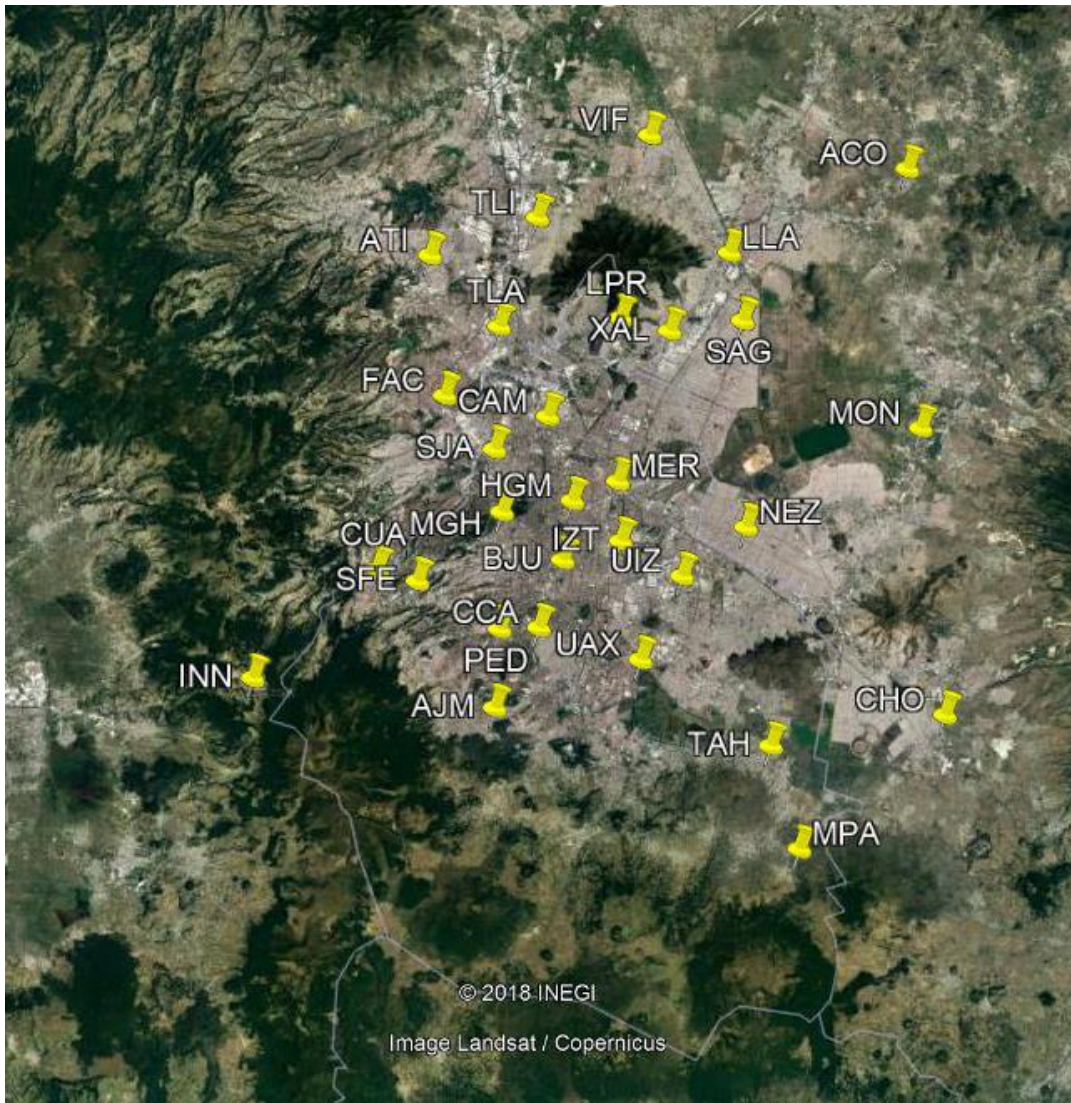


Imagen 5.2: Ubicación de las estaciones de monitoreo de CO en la Zona Metropolitana del Valle de México.  
(Elaboración Propia)

## 5.4 Plataforma seleccionada

Como fue posible obtener información del Landsat 8, la cual fue lanzada en el 2013, que cuenta con el sensor TIRS (*en inglés, Thermal Infrared Sensor*), que mide la temperatura superficial de la tierra en dos bandas, esto es posible ya que el sensor usa fotodetectores infrarrojos de pozos cuánticos “QWIPs” (*en inglés, Quantum Well Infrared Photodetectors*), para detectar longitudes de ondas largas que son emitidas por la Tierra, cuya intensidad depende de la temperatura superficial, dichas longitudes de onda son las conocidas como ondas de infrarrojo térmico, las cuales no son posibles de observar debido a que se encuentran fuera del rango de visión del ojo humano, el sensor TIRS colecta imágenes de dos bandas espectrales con una resolución de 120 m en 185 kilómetros de escena, las dos bandas se utilizan para garantizar la corrección atmosférica de la información térmica utilizando un algoritmo de división de ventana (Caselles et al, 1998), el cual representa una ventaja sobre la única banda térmica que poseen los satélites Landsat anteriores ya que sirve para distinguir entre la temperatura superficial terrestre y la atmosférica. Además del sensor TIRS, la plataforma del Landsat 8 cuenta con el sensor OLI (*en inglés, Operational Land Imager*), con un 90 por ciento de confiabilidad y relación radiométrica de 16 bits, que provee de dos nuevas bandas con respecto a la plataforma anterior Landsat7, una de ellas con objetivo de detectar nubes cirrus que son nubes formadas por cristales de hielo; en la banda 9, utilizando el infrarrojo próximo y la banda 1 utilizando un azul profundo en el rango visible para detección y estudio de zonas costeras y aerosoles en la atmósfera, ofreciendo 15 metros de resolución pancromática y 30 metros de resolución multi-espacial a lo largo de 185 kilómetros de escena en general, operando en la parte más baja del espectro visible entre 0.43  $\mu\text{m}$  y 0.45  $\mu\text{m}$ . Integrando así tres nuevas bandas, la banda 1, la banda 9 y banda QA (*en inglés, Quality Assesment*) donde cada píxel contiene combinaciones de bits que representan combinaciones de superficie, atmósfera y sensores que modifican la utilidad general del píxel como se muestra en la tabla siguiente. Por lo que el Landsat 8 cuenta con un total de 11 bandas; 8 bandas espectrales, una banda pancromática y dos bandas del sensor TIRS como se indica en la tabla 3.1 de este documento.



## Fechas de muestreo

Las fechas y horarios de muestreo, se fueron determinando en conjunto con la órbita del Landsat 8 dado que tiene un ciclo de 16 días, y el Servicio Meteorológico de los Estados Unidos (*USGS por sus siglas en inglés*) cuenta con un calendario donde se identifican las fechas en que el satélite se encuentra en posición de capturar imágenes en donde se encuentran las estaciones de monitoreo de la RAMA, siendo un total de 24 imágenes adquiridas y utilizables dado que se necesitaba que la nubosidad fuera nula o muy baja para poder realizar los modelos más precisos sin errores y que se pudiera descargar también con el nivel 2 SR que nos entrega reflectancia superficial de las bandas 1 a la 7. Las imágenes descargadas comprenden un periodo desde el 06 de febrero del 2016 hasta el 5 de julio del 2018, el Sistema Meteorológico de los Estados Unidos, por medio de los dominios <https://glovis.usgs.gov> y <https://earthdata.nasa.gov>, pone a disposición y de manera gratuita la descarga de los paquetes e imágenes que las distintas plataformas van capturando, para cualquier región de interés que se seleccione, las imágenes se van actualizando cada 24 horas sucesivas a su obtención. Además, se seleccionó este periodo debido a promedios anuales obtenidos de los años 2016, 2017 y hasta junio del 2018 tomado a las 12 de la tarde dado que a esa hora pasa el satélite Landsat 8 como se muestra en las imágenes 5.3 a 5.5.

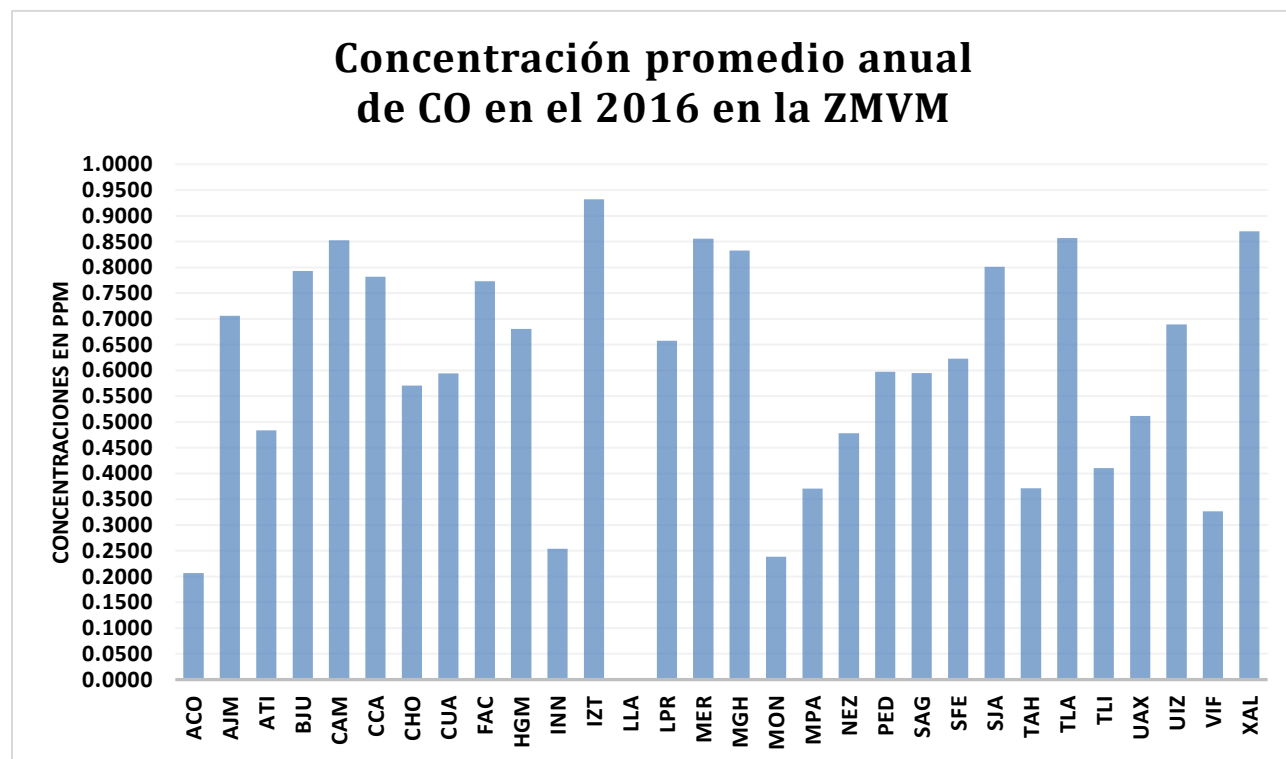


Imagen 5.3 Concentración promedio anual de monóxido de carbono en el 2016 en la ZMVM a las 12 pm.

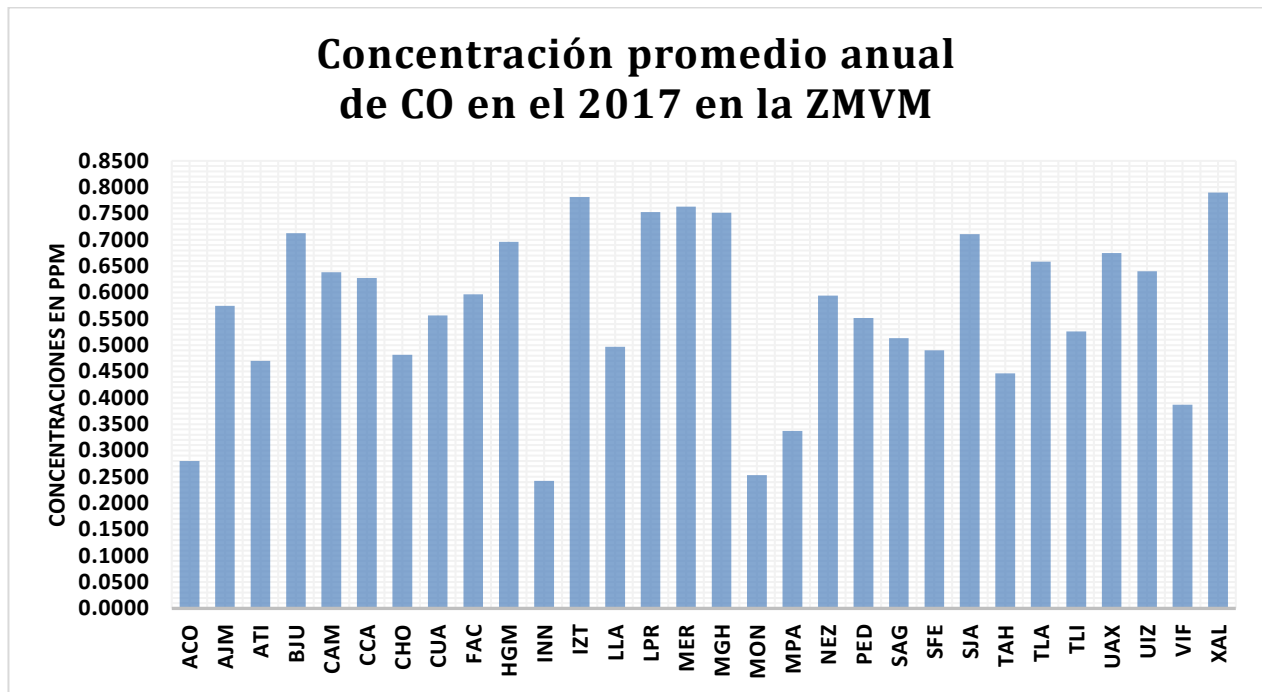


Imagen 5.4 Concentración promedio anual de monóxido de carbono en el 2017 en la ZMVM a las 12 pm.

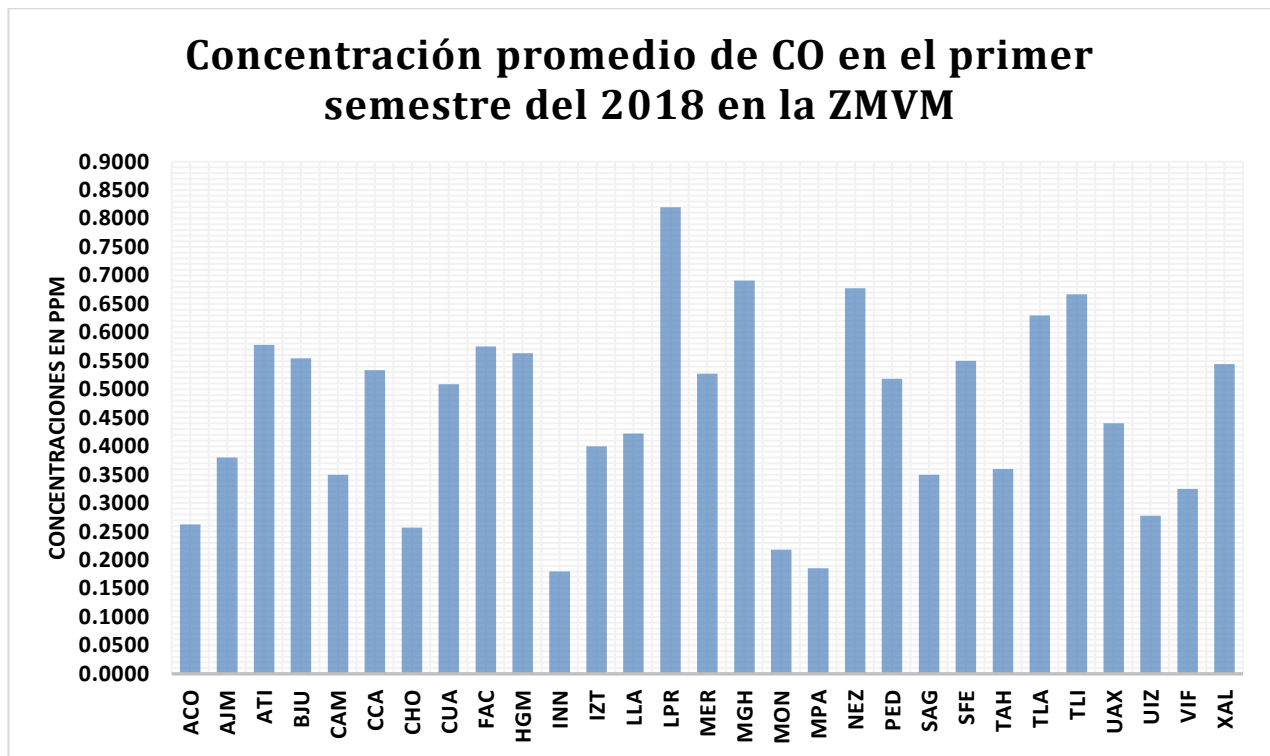


Imagen 5.5 Concentración promedio anual de monóxido de carbono en primer semestre del 2018 en la ZMVM a las 12 pm.

Como se puede observar en el 2016, Iztapalapa fue la estación que más concentración promedio anual captó con 0.93 ppm esto se puede deber a que en Iztapalapa se concentra una gran zona industrial además de un parque vehicular extenso que en la mayoría de las ocasiones se encuentra atascado y sin flujo lo cual aumenta las emisiones de monóxido de carbono, en el 2017 de nuevo Iztapalapa y Xalostoc, un pequeño pueblo de Ecatepec, la explicación de este año se debe a que existe un aumento en el parque vehicular año con año y zonas aledañas a la Ciudad de México también van incrementando su infraestructura y población como es el caso de Ecatepec que ha sido de las zonas conurbadas que más ha visto un incremento vehicular y poblacional en los últimos tiempos, mientras que en el primer semestre de este año la estación que más monóxido de carbono ha detectado se encuentra en La Presa ubicada en el municipio de Tlanepantla con 0.82 ppm, de nuevo un municipio del Estado de México que se encuentra en crecimiento poblacional y que mes con mes emite más contaminantes hacia la atmósfera, esto no quiere decir que la Ciudad de México este exenta de problemas, ya que esto simboliza promedios anuales, pero si se evalúa mes con mes, ha habido zonas de la Ciudad de México que exceden en más de una parte por millón la concentración promedio solamente a las 12 de la tarde, y dada la situación actual de la zona metropolitana, mucha gente pasa el tiempo atascado en el tráfico se estima que el mexiquense promedio pasa entre cinco y seis horas en transportarse a lo largo del día de su casa al trabajo y viceversa, esto simboliza una gran exposición a este contaminante el cual ha reflejado en un aumento de niños con problemas respiratorios y en adultos mayores con problemas cardiovasculares y aunque se cuente con el plan No circula, este no ha dado los resultados esperados ya que las emisiones han aumentando en estaciones donde antes no se presentaban altos niveles de monóxido de carbono.

#### **5.4 Procesamiento de las imágenes**

Hoy en día existen diversos softwares de información geográfica que tienen dentro de sus funciones el poder procesar imágenes satelitales, como el ArcGIS, ENVI, Qgis y aunque estos sean herramientas computacionales poderosas, se optó por usar al software IDRISI Selva debido a su amplia gama de funciones que se ajustan al proyecto, facilidad de uso y acceso. Para el proyecto se tomaron en cuenta seis imágenes dentro del periodo mencionado; el 6 de febrero del 2016, 19 de octubre del 2016, 24 de febrero del 2017, 7 de noviembre del 2017, 27 de febrero del 2018 y 18 de mayo del 2018, siendo así tres imágenes en periodo de lluvias y 3 imágenes en periodo de estiaje para diferentes años. Como fue posible obtener el nivel 2 del Landsat 8 por medio de los datos recolectados del EarthExplorer, el cual ya nos proporciona bandas con

reflectancia superficial (Bandas SR) y bandas TOA que ya tienen las correcciones como la geométrica que se debe realizar ya que existen alteraciones en el movimiento del satélite y el mecanismo de captación de los sensores que generan distorsiones en la imagen global, la corrección radiométrica que son los fallos en los sensores que generan pixeles incorrectos y la corrección atmosférica que de no realizarse existiría interferencia de la atmósfera, alterando la forma sistemática los valores de los pixeles. Ocupando solamente se ocuparon las bandas de la 1 a la 7 tanto las proporcionadas por la reflectancia superficial como las TOA, por lo que el procesamiento se facilitó dentro del software IDRISI Selva, el proceso inicial es acotar la imagen debido a que la imagen que se obtiene contiene zonas e información que no son necesarias para el proyecto, cortando imágenes dentro del mismo punto, todas las imágenes a procesar deben ser cortadas en el mismo píxel para que al agruparlas los datos correspondan unos con otros. Una vez que se han ajustado las imágenes, se debe componer una capa vectorial de puntos, la cual contiene las coordenadas de las estaciones de la RAMA, para dicho proceso las coordenadas deben estar en UTM ya que IDRISI opera en este sistema, como se mencionó anteriormente, fue posible obtener las imágenes con el nivel 2 agilizando así el procesamiento de las imágenes, seguidamente se realizó la resta de las bandas TOA menos la banda SR para así obtener la reflectancia que se utilizará para obtener los modelos.

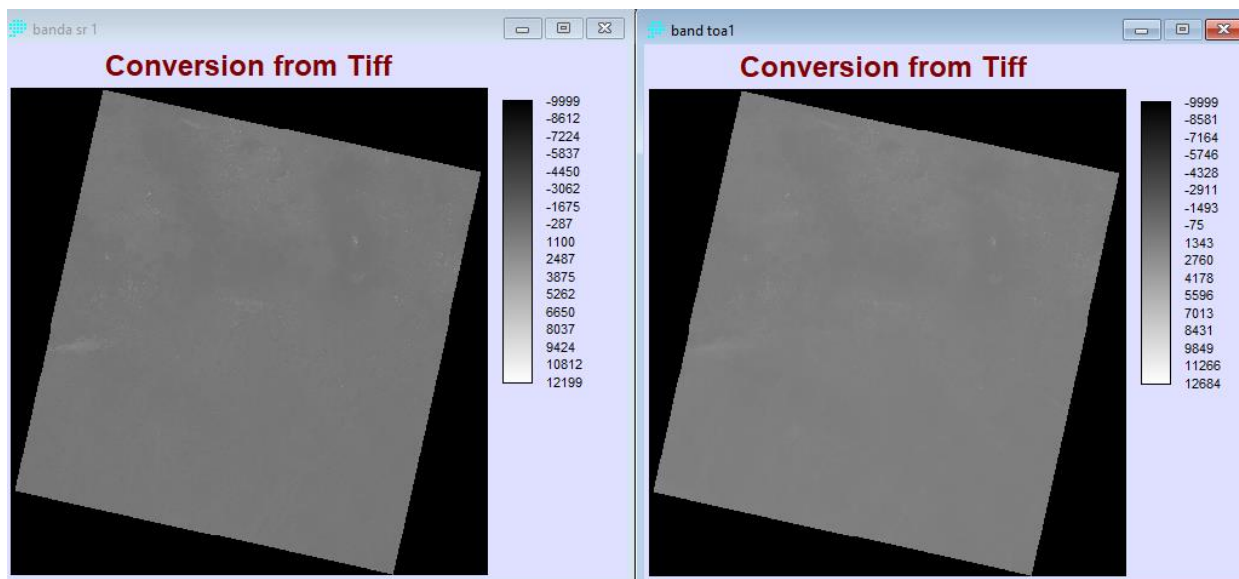


Imagen 5.6: Imagen de Bandas SR (Izquierda) y TOA (Derecha)

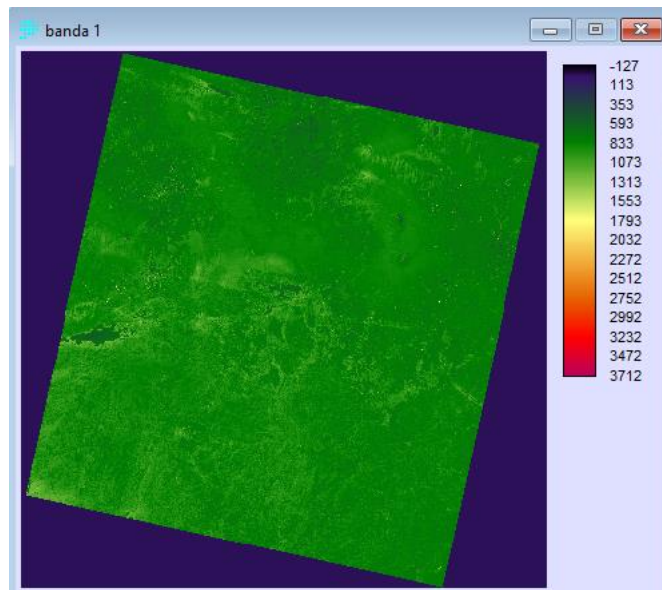


Imagen 5.7: Resta Banda 1

Así sucesivamente para las bandas de la 1 a la 7, una vez obtenidas las reflectancias resultantes de la resta con ayuda de la banda QA fue posible filtrar los datos para obtener imágenes claras como se muestra en la tabla 5.2.

Tabla 5.2: Landsat 8 Pixel Quality Attributes Values (USGS)

<b>Atributo</b>	<b>Valor del Píxel</b>
Fill	1
<b>Clear</b>	<b>322, 386, 834, 898, 1346</b>
Water	324, 388, 836, 900, 1348
Cloud shadow	328, 392, 840, 904, 1350
Snow/Ice	336, 368, 400, 432, 848, 880, 912, 944, 1352
Cloud	352, 368, 416, 432, 480, 864, 880, 928, 944, 992
Low confidence cloud	322, 324, 328, 336, 352, 368, 834, 836, 840, 848, 864, 880
Medium confidence cloud	386, 388, 392, 400, 416, 432, 898, 900, 904, 928, 944
High confidence cloud	480, 992
Low confidence cirrus	322, 324, 328, 336, 352, 368, 386, 388, 392, 400, 416, 432, 480
Medium confidence cirrus	834, 836, 840, 848, 864, 880, 898, 900, 904, 912, 928, 944, 992
High confidence cirrus	1346, 1348, 1350, 1352

Para que las imágenes obtenidas tengan menos errores es necesario que se capture una imagen clara (**clear**), para que la reflectancia sea de mayor calidad y los datos obtenidos que se manejen correspondan a imágenes sin obstrucciones como nubes, para comprobar los datos obtenidos es preciso hacer una composición de color, eso permite resaltar las características de una imagen, como diferenciar un cuerpo de agua, suelo y vegetación y en percepción remota esto se refiere al nivel de brillo en las distintas regiones del espectro electromagnético, este nivel de brillo depende de la asignación de los colores. La imagen que se logra por la composición que está formada por tres imágenes, puede ser color real o en falso color, esto se logra al formar la combinación RGB (Red, Green and Blue), la cual compone una imagen de contraste con puntos de saturación en cada píxel.

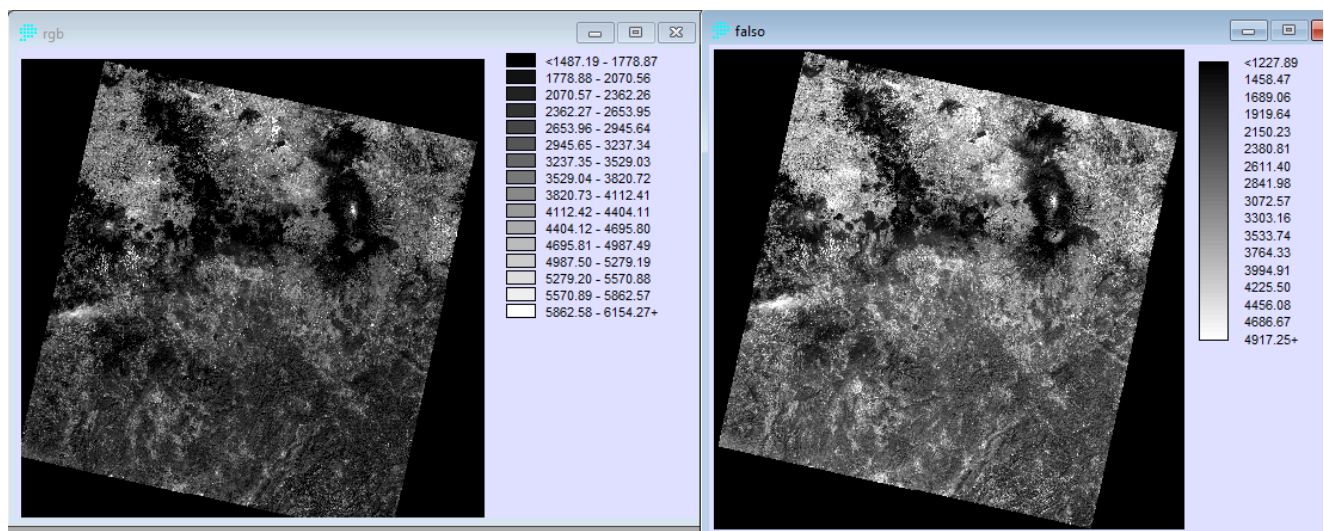


Imagen 5.8: Color verdadero (2,3,4) y Color Falso (4,3,2) del RGB

Comprobados los valores, se procedió a realizar gráficos con los valores obtenidos, con la idea de analizar de mejor manera las relaciones entre los valores obtenidos mediante la percepción remota y las mediciones de campo, así como analizar los datos obtenidos y posteriormente obtener un modelo de las imágenes que prediga el comportamiento del monóxido de carbono por medio de diferentes regresiones lineales, a continuación se muestran las gráficas por orden alfabético de las 30 estaciones evaluadas utilizando la bandas TOA para que el comportamiento se explicara de mejor manera.

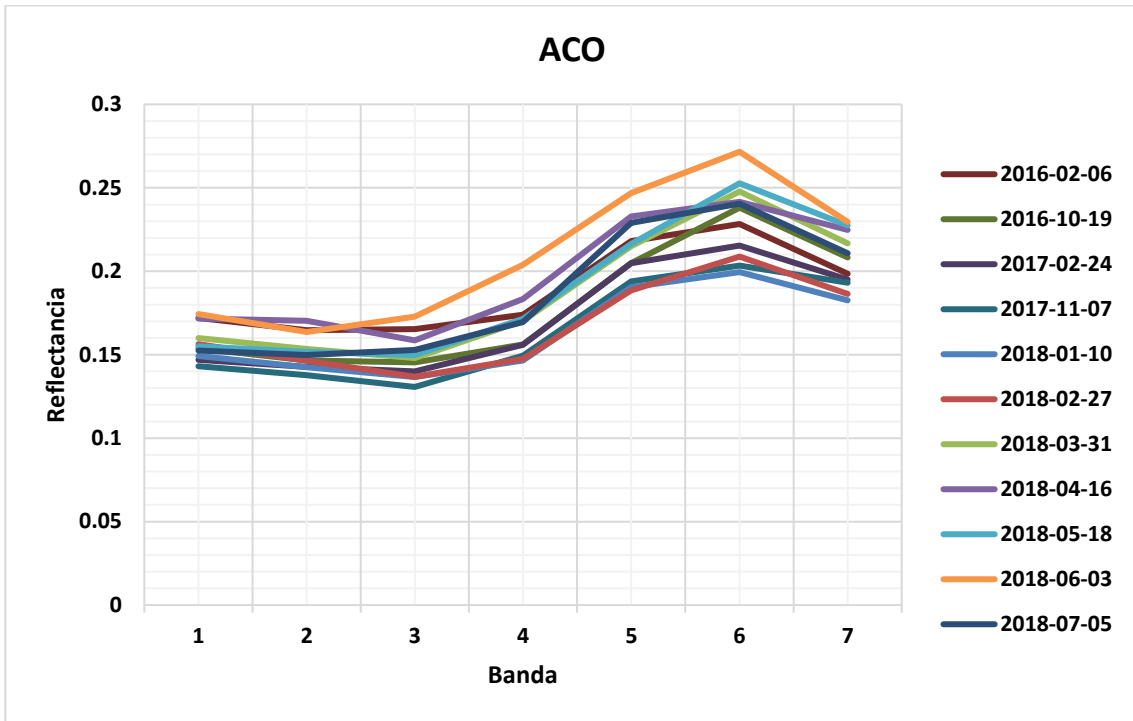


Imagen 5.9: Gráfica reflectancia Acolman.

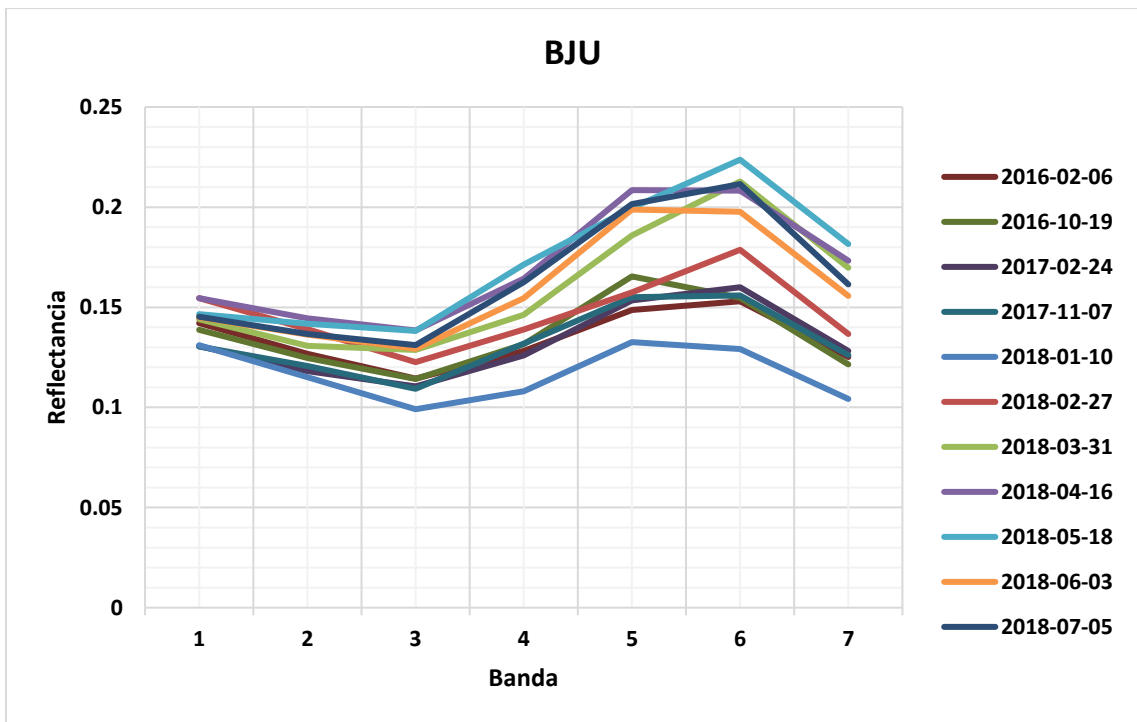


Imagen 5.10: Gráfica reflectancia Benito Juárez.

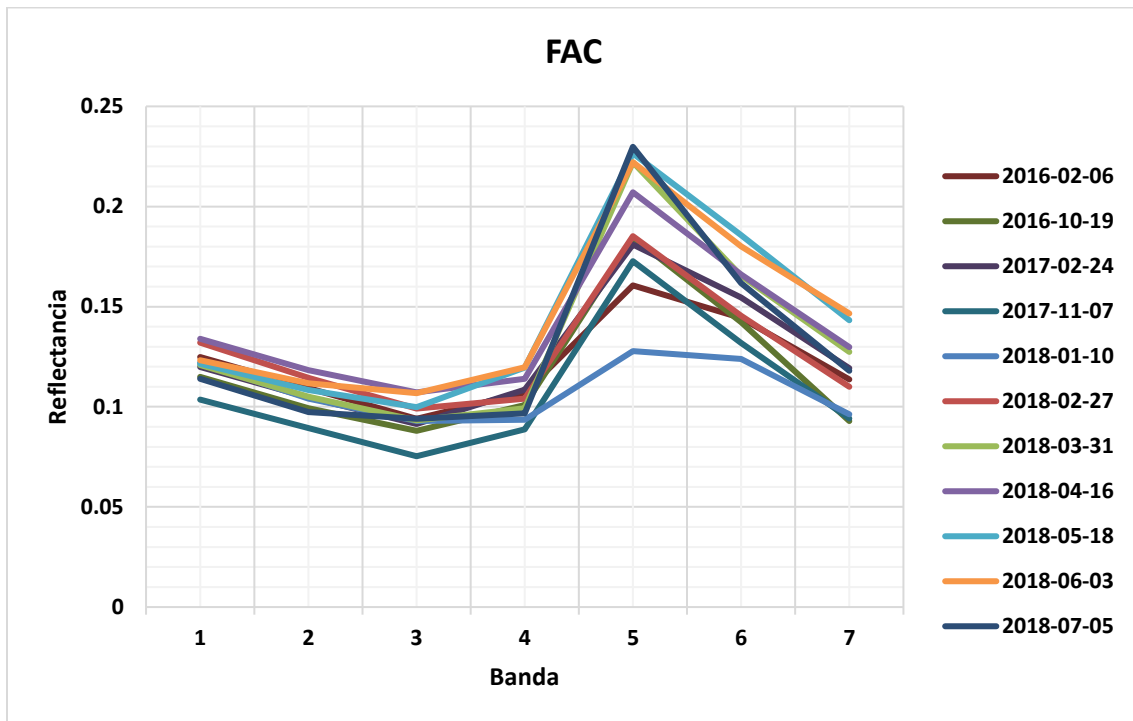


Imagen 5.11: Gráfica reflectancia FES Acatlán.

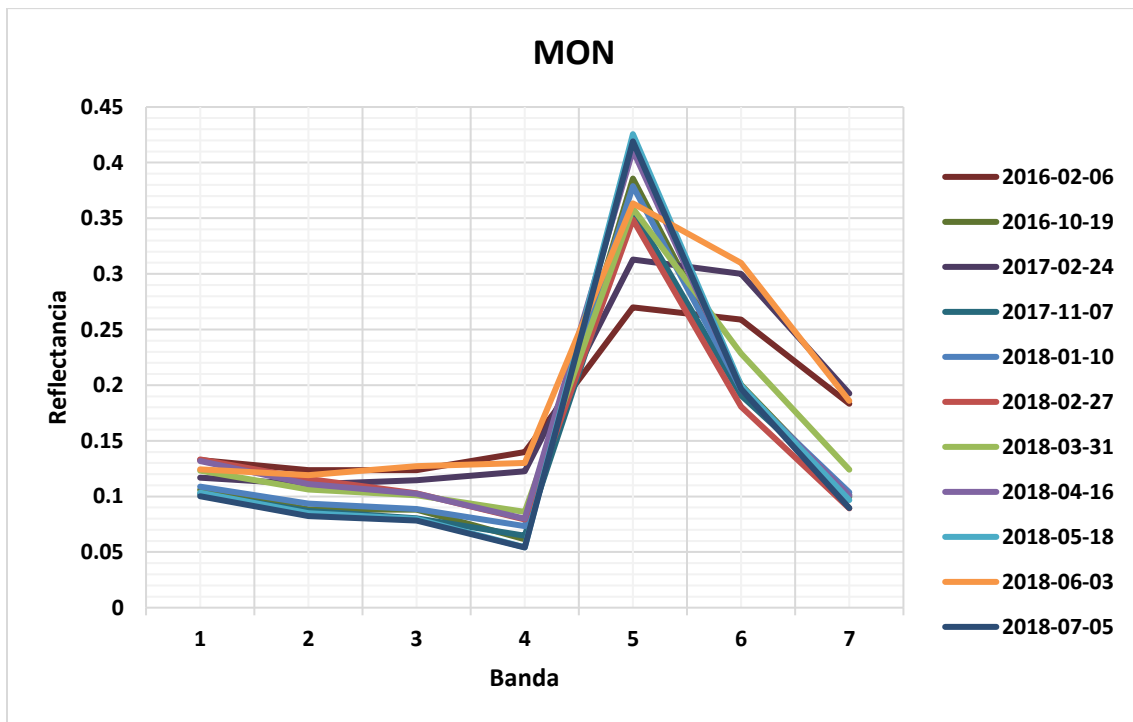


Imagen 5.12: Gráfica reflectancia Montecillo.



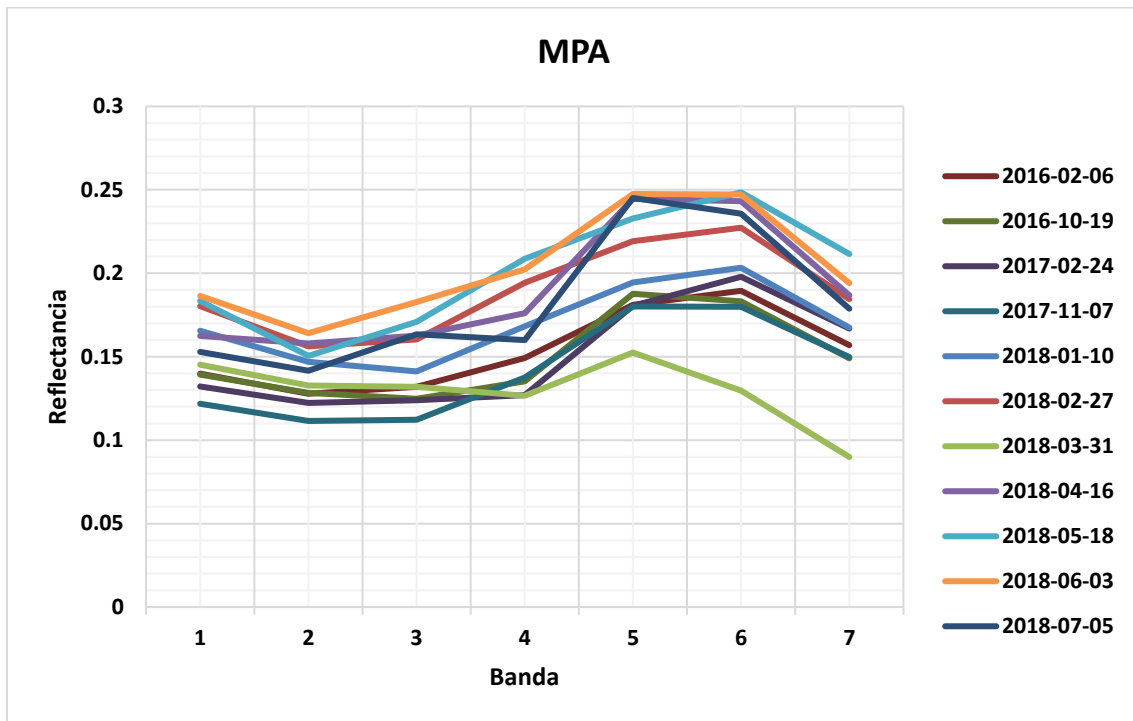


Imagen 5.13: Gráfica reflectancia Milpa Alta.

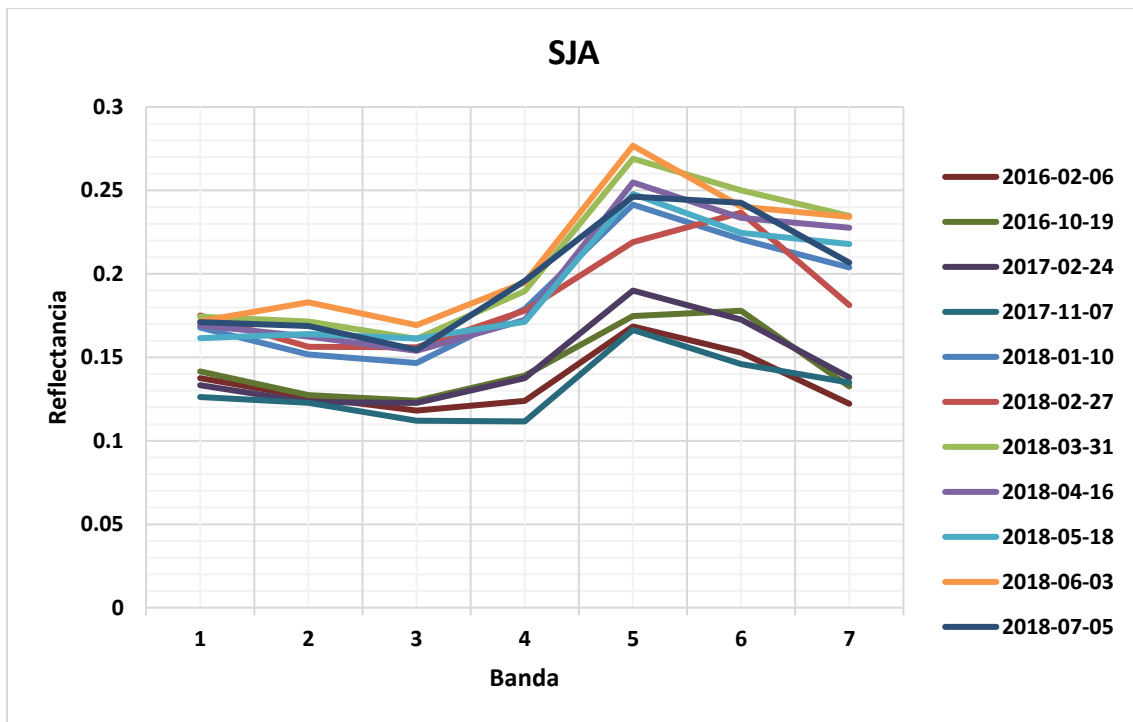


Imagen 5.14: Gráfica reflectancia San Juan de Aragón.

Como se puede observar en las imágenes de 5.9 a 5.14 y en las restantes encontradas dentro del anexo, se ha graficado las fechas indicadas, además de que todas las imágenes se capturaron en lo que va del 2018, en la mayoría de los casos, se empieza teniendo una tendencia negativa con valores de reflectancia oscilando entre 0.2 y 0.1, si solamente se contempla de la banda 1 a la 4 se observa que los valores más altos son los que se obtienen por medio de la banda 1 dado que ha este se le conoce como la banda para costas y aerosoles, además de que está integrado con el sensor OLI, lo cual estos valores son lo que se utilizan para el seguimiento de partículas suspendidas y aerosoles. Posteriormente podemos observar que existe un decremento de las bandas 2, 3 y 4, siendo éstas las bandas visibles del rojo = Banda 4, verde = Banda 3 y azul = Banda 2, la posible explicación a este fenómeno es que la zona urbana donde se toman las imágenes favorece la reflectancia del azul sobre el verde y del verde sobre el rojo, debido a que como la banda 1 refleja la presencia de aerosoles van difractando la energía de la banda 1, haciendo que la longitud de onda crezca a medida que se aleja la banda 1, por lo tanto existe una disminución en la reflectancia de las bandas 2 a la 4 ya que miden los últimos restos de aerosoles. El repunte que existe en la banda 5 que es el infrarrojo próximo, es debido a la vegetación que existe y la mancha urbana ya que las partículas de agua que se absorbe en la vegetación ayudan a que exista una mayor reflectancia. En la banda 6 en algunos de los casos se puede observar un incremento, pero no mayor a la banda 5, pero mayor que las bandas 2, 3 y 4, ya que la banda 6 es capaz de penetrar nubes finas, pero, aunque exista vegetación esto no es suficiente para elevar su reflectancia, pero debido al asfalto y concreto la reflectancia no decae tanto como en las bandas 2, 3 y 4. Finalmente la banda 7 muestra un claro decremento en todos los casos y aunque esta banda mide lo mismo que la banda 6, la banda 7 lo hace de una manera más profunda, haciendo que se ignore las partículas más grandes por lo que la reflectancia disminuye pero sin llegar a los valores del rango visible. Las fechas muestran un mismo comportamiento con excepción del 3 de junio del 2018 en la estación Ajusto Medio que muestra un comportamiento inusual a los demás comportamientos, esto puede deberse a la fuertes lluvias presentadas esa semana contribuyendo a una alta reflectancia del nivel azul, las mayores reflectancias se presentan en las estaciones de UAM Xochimilco y Villa Flores esto a que existe una gran vegetación en la mancha urbana de estas estaciones las cuales ayudan al incremento de la reflectancia.

## Capítulo 6. Resultados y Análisis

Como uno de los objetivos de este proyecto es encontrar un modelo que explique de mejor manera una relación entre las bandas teniendo como variable independiente al monóxido de carbono obtenido como dato de la RAMA y las reflectancias obtenidas de las siete primeras bandas del TOA y las siete primeras bandas resultantes de las restas, esto provenientes del Landsat 8, para esto fue necesario realizar varias regresiones para así encontrar la regresión a ocupar en cada fecha mencionada con anterioridad. Por lo que para este proyecto se analizaron las fechas mencionadas; 2016-02-06 (Estiague-NO verano), 2016-10-19 (LLuvias-NO verano), 2017-02-24 (Estiague-NO verano), 2017-11-07 (LLuvias-NO verano), 2018-02-27 (Estiague-NO verano) y 2018-05-18 (LLuvias-NO verano), la idea principal al escoger estas fechas era tener un grupo variado de observaciones, pero dado que algunas imágenes no fueron posible obtenerse en verano y temporada de lluvias porque las imágenes no eran claras o tenían nubosidad se fueron excluyendo, por eso al final se optó por estas seis imágenes y observar si la lluvia tiene algún efecto en el comportamiento de los modelos, por medio del software estadístico Rstudio, fue posible programar todas las regresiones y obtener el test MSE y graficar el error para elegir la mejor regresión para cada modelo, como cada archivo de reflectancia contaba con la resta obtenida de las bandas de la 1 a la 7 y la reflectancia de las bandas con TOA de la 1 a la 7, por eso fue necesario crear tres programas uno que solo tomara en cuenta las bandas de la 1 a la 7 otro que tomara las bandas con el TOA y el último que contara las 14 variables, dentro de todas las regresiones mencionadas se realizaron todas las regresiones para las fechas mencionadas para poder comparar que regresión se ajusta mejor a las variables, por lo que se realiza una comparación gráfica para facilitar el porqué de la elección de la regresión paso por paso, como se muestra en las imágenes 6.1 a 6.3.

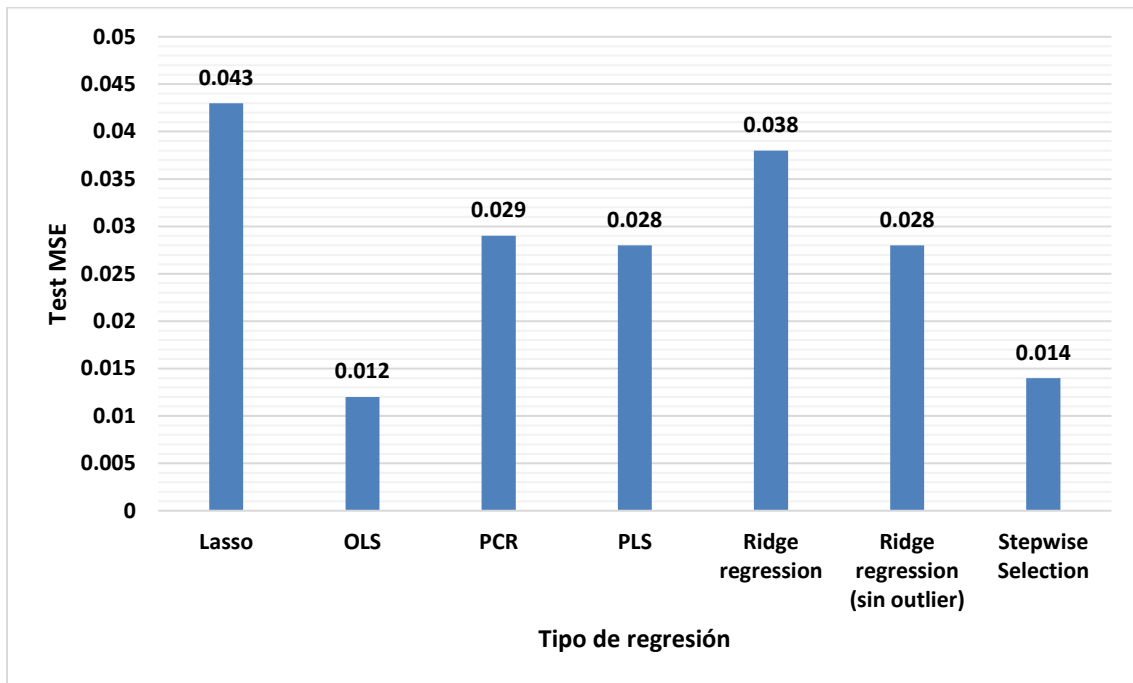


Imagen 6.1: Comparación de métodos de regresión para el 2016-02-06 utilizando solo las bandas.

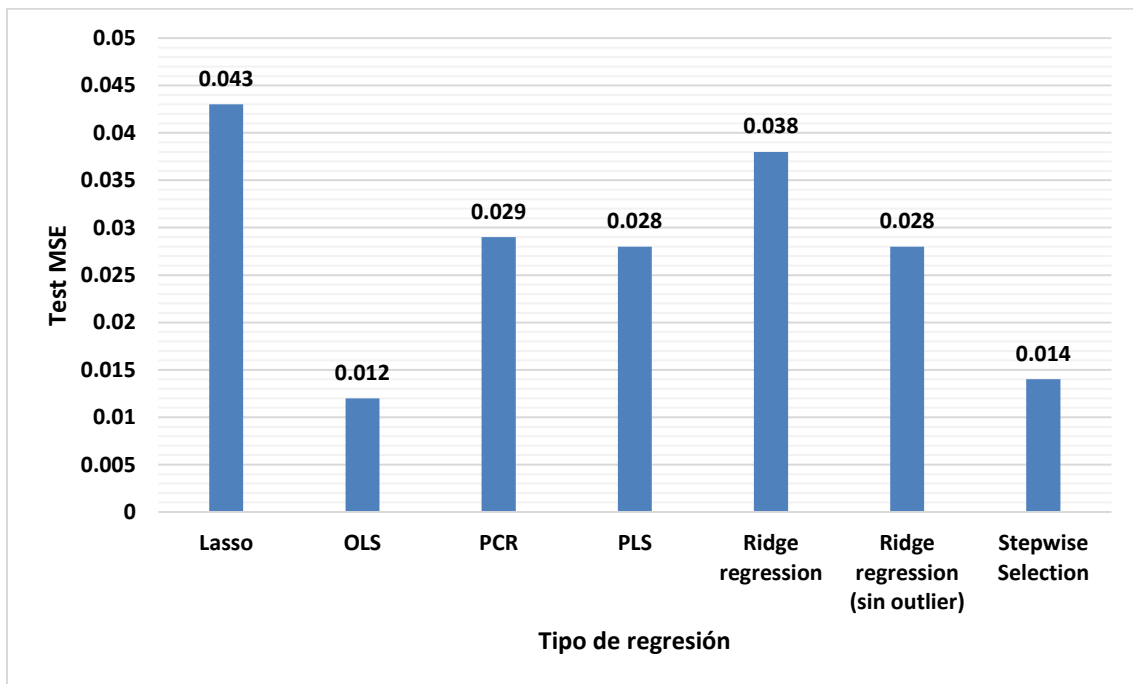


Imagen 6.2: Comparación de métodos de regresión para el 2016-02-06 utilizando solo las bandas con TOA.

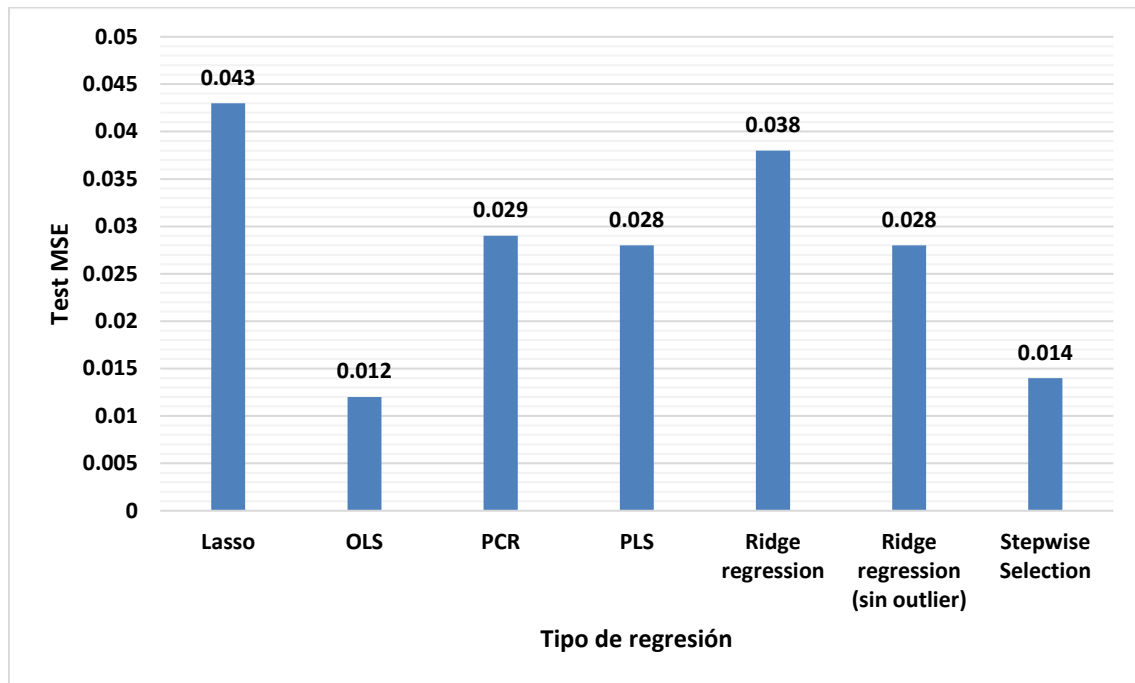


Imagen 6.3: Comparación de métodos de regresión para el 2016-02-06 utilizando todas las bandas.

Las gráficas anteriores muestran el resultado del test-MSE, cuando el valor obtenido es cero significa que hay una perfecta exactitud con los parámetros de observación, pero como esta condición es ideal y en la realidad no se puede aplicar se desechan, el valor con test menor será el modelo que mejor se comporte y tenga menor variación en la variancia. Para la fecha seleccionada se tiene que en las tres combinaciones posibles el valor de test menor es el OLS o la regresión por mínimos cuadrados, seguido por el PCR, PLS y el Stepwise Selection, por lo que las demás regresiones se descartan por tener un valor alto, dentro de las cuatro posibilidades la que mejor explicaría el comportamiento sin tanto error y mejor probabilidad sería la opción de OLS, pero en estadística existe la ley de parsimonia o Navaja de Ockham que tiene aplicaciones de importancia en el análisis exploratorio de modelos de regresión lineal múltiple de un conjunto de variables explicativas que forman parte del modelo a estudiar, debe seleccionarse la combinación más reducida y simple posible, teniendo en cuenta la varianza residual, la capacidad de predicción y la multicolinealidad, es decir se debe seleccionar el modelo que tenga menos variables y explique el mismo comportamiento. Al encontrar los coeficientes de las cuatro posibles regresiones, se descartaron las opciones de OLS, PLS y PCR dado que como se explicó anteriormente estos modelos incluyen la mayoría de las variables, por eso su error es bajo, ya que al incluir todas las variables la varianza no varía, es por eso que la regresión Stepwise surge como

la opción ideal, debido a que esta si va eliminando variables hasta encontrar el modelo que mejor explique el comportamiento sin tener que contar con todas las variables. Posteriormente después de escoger la regresión que mejor se adecua a las imágenes seleccionadas se comienza a aplicar el método Stepwise en dirección Backward utilizando nuevamente el software Rstudio y para las tres combinaciones posibles mencionadas, primeramente, se mostrara el modelo utilizando solo las restas, después solamente las bandas con TOA y por último las 14 variables, por lo que se obtuvo lo siguiente:

1) 06 de febrero del 2016.

i)  $CO_{12} \sim \text{Banda.2} + \text{Banda.7}$

$$CO = -22.939 * \text{Banda } 2 + 84.852 \text{ Banda } 7 + 2.627$$

Con un error residual estándar de 0.1963 con 16 grados de libertad, una  $R^2$  de 0.4965, una  $R^2$  ajustada de 0.4336,  $\sigma$  de 0.196, AIC de -3.22, BIC de 0.558 y un p.value de 0.00413.

ii)  $CO_{12} \sim \text{TOA.Banda.2} + \text{TOA.Banda.7}$

$$CO = 13.5902 * \text{TOA. Banda } 2 - 8.3153 \text{ TOA. Banda } 7 + 0.1737$$

Con un error residual estándar de 0.2027 con 16 grados de libertad, una  $R^2$  de 0.4631, una  $R^2$  ajustada de 0.396,  $\sigma$  de 0.203, AIC de -2.00, BIC de 1.78 y un p.value de 0.00690.

iii)  $CO_{12} \sim \text{Banda.1} + \text{Banda.2} + \text{Banda.3} + \text{Banda.5} + \text{Banda.6} + \text{Banda.7} + \text{TOA.Banda.1} + \text{TOA.Banda.2} + \text{TOA.Banda.3} + \text{TOA.Banda.6}$

$$CO = -583.263 * \text{Banda } 1 + 990.452 * \text{Banda } 2 - 272.586 * \text{Banda } 3 - 246.677 * \text{Banda } 5 - 1070.944 \text{ Banda } * 6 + 130.844 * \text{Banda } 7 - 168.489 * \text{TOA Banda } 1 + 236.774 * \text{TOA Banda } 2 - 62.777 * \text{TOA Banda } 3 - 42.110 * \text{TOA Banda } 6 + 2.885$$

Con un error residual estándar de 0.182 con 8 grados de libertad, una  $R^2$  de 0.4631, una  $R^2$  ajustada de 0.396,  $\sigma$  de 0.203, AIC de -2.00, BIC de 1.78 y un p.value de 0.00690.

De las tres opciones posibles evaluaremos parámetro por parámetro para elegir la mejor ecuación de regresión, el primer parámetro es el error residual este indica el porcentaje promedio de desviación que se tendrá en predicciones y que tanto se aleja de la media, es decir en el modelo (i) se tiene una concentración de 2.627 de CO cuando se iguala a cero los valores de las bandas y un error residual de 0.1963, por lo que el porcentaje de error es 7.47%, al igual manera el modelo (ii) tiene un porcentaje de error de 116.70% de error y el modelo (iii) con un porcentaje de error de 6.31%.

El siguiente parámetro es la  $R^2$ , este parámetro es de los más decisivos al momento de escoger entre distintas regresiones ya que indica que tan bien un modelo se ajusta a la información y tiene como propósito el predecir futuros resultados o probar hipótesis, este coeficiente determina la calidad del modelo para replicar resultados y que tanta varianza existe, entre mayor sea el valor se espera que el modelo sea mejor y este debe estar comprendido entre cero y uno, en el modelo (i) se tiene una  $R^2$  de 0.4965, en el modelo (ii) de 0.4631 y en el modelo (iii) de 0.7836.

La  $\sigma$  es la desviación estándar que es una medida de dispersión y puede ser interpretada como una medida de incertidumbre es decir que tan dispersos están los datos con respecto a la media, mientras más grande la incertidumbre aumenta, en el primer modelo se tiene una sigma de 0.196, en el segundo de 0.203 y en el tercero de 0.182, otro parámetro para medir la calidad relativa de un modelo estadístico para un conjunto de datos.

Es el AIC y el BIC, como ya se explicó anteriormente, entre menor sea el valor, indicara que se necesitan menos variables para representar mejor el fenómeno y que tan bien se predecirá el comportamiento, en el modelo (i) se tiene un AIC de -3.22 y un BIC de 0.558, en el (ii) AIC de -2.00 y BIC de 1.78 y en el (iii) un AIC de -3.26 y un BIC de 8.07.

Por último, el parámetro p.value que indica que tan buena es la relación entre el predictor y las variables de respuesta, entre más alejado este el valor de uno se considera que es mejor y que se rechaza la hipótesis nula de la prueba, en el primero modelo se tiene un valor de 0.00413, en el segundo un valor de 0.00690 y en el tercero un valor de 0.0726

Para poder elegir la mejor combinación de regresión Stepwise, es necesario analizar parámetro por parámetro y comprarlos, es necesario decir que ningún parámetro tiene mayor peso sobre otro, por eso es necesario calcular varios para escoger la mejor opción. Aunque el tercer modelo tenga el mayor valor de  $R^2$  y menor desviación estándar, se debe de descartar porque este modelo

contiene demasiadas variables, por esto la  $R^2$  es la mayor, pero si observamos el AIC y el p.value aunque este modelo nos explique en un 78.36 de las veces el comportamiento, esto no indica que sea el modelo que mejor correlación tiene con las variables y en realidad necesita tener la mayoría de ellas para explicar el fenómeno, por lo que el mejor modelo para esta fecha sería el modelo (i), dado que el BIC es menor que el segundo modelo, y siendo el parámetro BIC el que más penaliza las variables y su comportamiento, además que el p.value del modelo (i) es menor que el (ii) y el porcentaje de error del modelo dos es de 116.70%, esto sugiere mucha incertidumbre y si se cuenta un modelo con dos variables, que además es resultado de una resta de bandas (Bandas TOA-Bandas SR), con menos incertidumbre y mayor predicción. Por eso el modelo ideal para esta fecha es  $CO = -22.939 * Banda 2 + 84.852 Banda 7 + 2.627$

2) 19 de octubre del 2016.

i)  $CO_{12} \sim Banda.3 + Banda.5 + Banda.7$

$$CO = -66.458 * Banda 3 + 151.883 * Banda 5 + 90.900 Banda 7 + 2.561$$

Con un error residual estándar de 0.2187 con 20 grados de libertad, una  $R^2$  de 0.5609, una  $R^2$  ajustada de 0.495,  $\sigma$  de 0.219, AIC de 0.771, BIC de 6.66 y un p.value de 0.000765.

ii)  $CO_{12} \sim TOA.Banda.3 + TOA.Banda.5 + TOA.Banda.7$

$$CO = 11.9526 * TOA. Banda 3 - 2.2761 * TOA. Banda 5 - 9.0131 * TOA. Banda 7 + 0.9966$$

Con un error residual estándar de 0.2281 con 20 grados de libertad, una  $R^2$  de 0.5223, una  $R^2$  ajustada de 0.4507,  $\sigma$  de 0.228, AIC de 2.79, BIC de 8.68 y un p.value de 0.00172.

iii)  $CO_{12} \sim Banda.1 + Banda.2 + Banda.5 + Banda.6 + Banda.7 + TOA.Banda.1 + TOA.Banda.2 + TOA.Banda.5 + TOA.Banda.6 + TOA.Banda.7$

$$CO = -500.15 * Banda 1 + 718.75 * Banda 2 + 998.994 * Banda 5 - 2267.191 * Banda 6 - 180.882 * TOA Banda 1 + 199.873 * TOA Banda 2 + 10.270 * TOA Banda 5 - 73.843 * TOA Banda 6 - 19.281 * TOA Banda 7 - 1.377$$



Con un error residual estándar de 0.2165 con 14 grados de libertad, una  $R^2$  de 0.6987, una  $R^2$  ajustada de 0.505,  $\sigma$  de 0.203, AIC de 3.73, BIC de 16.7 y un p.value de 0.0158.

Para poder comparar los tres modelos de la segunda fecha, primero debemos obtener el porcentaje de error, del primer modelo se tiene que el porcentaje de error es de 8.54%, del segundo modelo, se tiene que el error es de 22.89% y para el último modelo se tiene que el porcentaje de error es de 15.72%, este primer parámetro nos indica que mejor modelo es la primera opción dado que sus predicciones futuras solo tendrán un 8.54% de desviación o de error, si además queremos incluir los demás parámetros se observa que el p.value nos indica que una vez más el primer modelo tiene el valor de probabilidad más alejado de uno que indica el mejor comportamiento de las variables, pero no solamente eso tanto los valores de AIC y de BIC resultantes son menores en el primer modelo, lo que explica que las variables se comportan de mejor manera en el primer modelo y aunque el valor de la  $R^2$  sea mayor en el tercer modelo y la mayor desviación estándar sea del segundo modelo, el primer modelo se acepta como el mejor debido a que tiene más parámetros estadísticos de elección.

$$CO = -66.458 * Banda\ 3 + 151.883 * Banda\ 5 + 90.900\ Banda\ 7 + 2.561$$

3) 24 de febrero del 2017.

i)  $CO_{12} \sim Banda.2 + Banda.3$

$$CO = -85.986 * Banda\ 2 + 84.279 * Banda\ 3 + 3.348$$

Con un error residual estándar de 0.276 con 25 grados de libertad, una  $R^2$  de 0.2492, una  $R^2$  ajustada de 0.1892,  $\sigma$  de 0.276, AIC de 12.2, BIC de 17.5 y un p.value de 0.0278.

ii)  $CO_{12} \sim TOA.Banda.2 + TOA.Banda.3 + TOA.Banda.4 + TOA.Banda.5 + TOA.Banda.6 + TOA.Banda.7$

$$CO = 20.9847 * TOA.Banda\ 2 - 22.0970 * TOA.Banda\ 3 + 9.2994 * TOA.Banda\ 4 - 6.0306 TOA.Banda\ 5 + 9.7285 TOA.Banda\ 6 - 10.9630 TOA.Banda\ 7 + 0.3833$$

Con un error residual estándar de 0.2825 con 21 grados de libertad, una  $R^2$  de 0.3392, una  $R^2$  ajustada de 0.1505,  $\sigma$  de 0.283, AIC de 16.6, BIC de 27.3 y un p.value de 0.1484.

iii)  $CO_{12} \sim Banda.2 + Banda.4 + Banda.6 + Banda.7 + TOA.Banda.3 + TOA.Banda.4 + TOA.Banda.6 + TOA.Banda.7$

$$CO = -96.992 * Banda 2 + 399.784 * Banda 4 - 2280.577 * Banda 6 \\ + 3408.390 Banda 7 - 17.965 * TOA Banda 3 + 45.282 \\ * TOA Banda 4 - 80.232 * TOA Banda 6 + 201.399 \\ * TOA Banda 7 - 0.136$$

Con un error residual estándar de 0.2796 con 19 grados de libertad, una  $R^2$  de 0.4144, una  $R^2$  ajustada de 0.1678,  $\sigma$  de 0.168, AIC de 17.2, BIC de 30.6 y un p.value de 0.168.

Nuevamente para poder tener un parámetro más que nos ayude a la elección de los modelos, obtenemos el porcentaje de error y del primero modelo el resultado es de 7.44%, del segundo modelo es de 28.35% y del tercer modelo de 205.59% un porcentaje de error muy alto y aunque existen otros parámetros se puede rechazar el último por este resultado y además de que es el modelo que mayor número de variables tiene. Comparando los dos modelos restantes, el segundo modelo tiene mayor  $R^2$  que el primero y aunque este sea el parámetro que genera mayor confianza dentro de la estadística, no es suficiente para explicar el comportamiento, dentro del primer modelo se obtiene una  $R^2$  cercana a la del segundo modelo valores tanto de AIC, BIC y p.value de mejor calidad. Explicaría la mejor correlación de las variables con los datos y aunque existan razones para elegir el segundo modelo como el más confiable, se tendrá que aplicar la ley de la navaja de Ockham para elegir el primer modelo ya que con menor parámetros se obtienen resultados favorables y confiables.

$$CO = -85.986 * Banda 2 + 84.279 * Banda 3 + 3.348$$

4) 7 de noviembre del 2017.

i)  $CO_{12} \sim Banda.1 + Banda.3$

$$CO = -55.133 * Banda 1 + 56.420 * Banda 3 + 3.534$$

Con un error residual estándar de 0.2792 con 23 grados de libertad, una  $R^2$  de 0.2684, una  $R^2$  ajustada de 0.2047,  $\sigma$  de 0.279, AIC de 12.3, BIC de 17.3 y un p.value de 0.0275.

ii)  $CO_{12} \sim TOA.Banda.1 + TOA.Banda.3$

$$CO = 17.7018 * TOA.Banda 1 - 11.5065 * TOA.Banda 3 - 0.4983$$

Con un error residual estándar de 0.2989 con 23 grados de libertad, una  $R^2$  de 0.1616, una  $R^2$  ajustada de 0.08866,  $\sigma$  de 0.299, AIC de 15.8, BIC de 20.8 y un p.value de 0.132.

iii) CO<sub>12</sub>~Banda.1+Banda.3+Banda.4+Banda.5+TOA.Banda.3+  
TOA.Banda.4+TOA.Banda.5+TOA.Banda.7

$$CO = -63.139 * Banda 1 - 1394.083 * Banda 3 + 2148.471 * Banda 4 \\ + 536.584 Banda 5 - 238.466 * TOA Banda 3 + 200.912 \\ * TOA Banda 4 + 7.108 * TOA Banda 5 + 6.770 * TOA Banda 7 \\ + 12.592$$

Con un error residual estándar de 0.2623 con 17 grados de libertad, una R<sup>2</sup> de 0.5229, una R<sup>2</sup> ajustada de 0.2984, σ de 0.262, AIC de 13.1, BIC de 25.7 y un p.value de 0.0680.

Si se elige la R<sup>2</sup> como el parámetro con mayor peso de decisión, el modelo (iii) sería el adecuado ya que en 52.29% de las veces explicaría el comportamiento en predicciones futuras y además se analiza el porcentaje de error que es de 2.08% indicando que no habría mucha variación en las predicciones futuras, pero observado los parámetros de p.value AIC, BIC y desviación estándar, el primer modelo tiene más parámetros que ajustan los datos a la regresión realizada y solamente utilizando dos variables, como se explicó anteriormente entre más variables tenga los modelos el valor de la R<sup>2</sup> aumentara, por eso en este caso el alto valor que se obtiene en el tercer modelo va perdiendo certidumbre y confiabilidad, contrario al caso del primer modelo y segundo, pero lo que diferencia a estos dos modelos es que el error porcentual del primero es de 7.9% y del segundo 59.98%. por lo que descartando el tercer modelo por la cantidad de variables que tiene y el segundo por el alto error porcentual, AIC, BIC y p.value que tiene en comparación con el primer modelo, por lo que este es el indicado para esta fecha.

$$CO = -55.133 * Banda 1 + 56.420 * Banda 3 + 3.534$$

5) 27 de febrero del 2018

i)  $CO_{12} \sim \text{Banda.2} + \text{Banda.6} + \text{Banda.7}$

$$CO = -20.653 * \text{Banda } 2 + 152.973 * \text{Banda } 6 - 68.689 * \text{Banda } 7 + 2.211$$

Con un error residual estándar de 0.2355 con 17 grados de libertad, una  $R^2$  de 0.53, una  $R^2$  ajustada de 0.4471,  $\sigma$  de 0.235, AIC de 4.42, BIC de 9.64 y un p.value de 0.00426.

ii)  $CO_{12} \sim \text{TOA.Banda.6} + \text{TOA.Banda.7}$

$$CO = -9.502 * \text{TOA.Banda } 6 + 10.911 * \text{TOA.Banda } 7 + 1.047$$

Con un error residual estándar de 0.258 con 18 grados de libertad, una  $R^2$  de 0.4026, una  $R^2$  ajustada de 0.3362,  $\sigma$  de 0.258, AIC de 7.46, BIC de 11.6 y un p.value de 0.00970.

iii)  $CO_{12} \sim \text{Banda.1} + \text{Banda.3} + \text{Banda.6} + \text{Banda.7} + \text{TOA.Banda.1} +$

$\text{TOA.Banda.2} + \text{TOA.Banda.3} + \text{TOA.Banda.4} + \text{TOA.Banda } 6$

$$\begin{aligned} CO = & -214.572 * \text{Banda } 1 + 752.288 * \text{Banda } 3 - 3118.079 * \text{Banda } 6 \\ & - 182.423 \text{Banda } 7 - 92.127 * \text{TOA Banda } 1 + 36.849 \\ & * \text{TOA Banda } 2 + 131.872 * \text{TOA Banda } 3 - 9.871 * \text{TOA Banda } 4 \\ & - 131.557 * \text{TOA Banda } 6 - 0.953 \end{aligned}$$

Con un error residual estándar de 0.2291 con 11 grados de libertad, una  $R^2$  de 0.7122, una  $R^2$  ajustada de 0.4766,  $\sigma$  de 0.229, AIC de 6.12, BIC de 17.6 y un p.value de 0.0437.

Obteniendo los porcentajes de error del primero modelo el resultado es de 10.65%, del segundo modelo el error es de 24.72% y del tercer modelo de 24.04%, este parámetro abre el panorama para escoger tanto al modelo uno como al último como posibles opciones y descartando la segunda regresión y aunque este tenga solamente dos variables dándole mayor peso a esta opción si se analiza con mayor detalle los parámetros arrojados, el valor de AIC es el mayor, esto indica que aunque solo tenga dos parámetros estos no tienen tan buena correlación con los datos y las predicciones como los otros modelos, por lo que este modelo por menor número de variables y mayor  $\sigma$  que arroja, este modelo se descarta por el AIC, p.value y BIC arrojados, mientras que por el otro lado el mejor valor de  $R^2$  que nos arrojan los modelos restantes es el del primer modelo de 0.53 además de que el p.value del tercer modelo no es tan significativo como el del primero indicando que la probabilidad de predicción de esta opción no será tan significativa conjuntamente

el porcentaje de error del primer modelo explicaría la relación con menos varianza, por lo que por eliminación, la mejor regresión es la del primer modelo.

$$CO = -20.653 * Banda 2 + 152.973 * Banda 6 - 68.689 * Banda 7 + 2.211$$

6) 18 de mayo del 2018

i)  $CO_{12} \sim Banda.2 + Banda.3 + Banda.7$

$$CO = 10.7817 * Banda 2 - 35.4572 * Banda 3 + 33.2662 * Banda 7 + 0.8466$$

Con un error residual estándar de 0.123 con 23 grados de libertad, una  $R^2$  de 0.2729, una  $R^2$  ajustada de 0.1781,  $\sigma$  de 0.123, AIC de -30.9, BIC de -24.4 y un p.value de 0.0580.

ii)  $CO_{12} \sim TOA.Banda.3 + TOA.Banda.7$

$$CO = 2.9640 * TOA.Banda 3 - 2.4993 * TOA.Banda 7 + 0.3579$$

Con un error residual estándar de 0.1271 con 24 grados de libertad, una  $R^2$  de 0.19, una  $R^2$  ajustada de 0.1225,  $\sigma$  de 0.127, AIC de -30.0, BIC de -24.8 y un p.value de 0.0798.

iii)  $CO_{12} \sim Banda.1 + Banda.2 + Banda.3 + Banda.5 + Banda.6 + Banda.7 + TOA.Banda.3 + TOA.Banda.4 + TOA.Banda.6$

$$CO = 31.279 * Banda 1 + 39.112 * Banda 2 - 287.794 * Banda 3 + 60.776 * Banda 5 + 984.881 * Banda 6 + 43.586 * Banda 7 - 33.146 * TOA Banda 3 + 3.794 * TOA Banda 4 + 36.521 * TOA Banda 6 + 4.158$$

Con un error residual estándar de 0.1002 con 17 grados de libertad, una  $R^2$  de 0.6432, una  $R^2$  ajustada de 0.4544,  $\sigma$  de 0.100, AIC de -38.1, BIC de -23.8 y un p.value de 0.0142.

Analizando los parámetros, se puede observar que el AIC del segundo modelo es el de menor valor comparándolo con los otros dos, indicando que con solamente con dos variables correlaciona de mejor manera los datos que los otros modelos, además de que el valor de la desviación estándar es el mayor indicando que la muestra estará comportándose dentro de la media la mayoría de las veces, esto serían parámetros suficientes para escoger este modelo como el ideal, pero si analizamos las variables se trata de bandas que contienen reflectancia de la parte superior de la atmósfera y aunque aerosoles y otras sustancias se encuentren ahí, por

lo que lo ideal es utilizar las bandas de las restas para obtener los resultados operados, y aunque este modelo sea el segundo modelo tenga parámetros estadísticos que lo eligen como el mejor, si se elige se debe tomar en cuenta que se está trabajando con la parte superior de la atmósfera, por lo que analizando los parámetros del modelo uno, se puede notar que el valor de AIC no varía mucho, el p.value resulta de mayor confianza y la  $R^2$  es de mayor valor, asimismo el error porcentual del segundo modelo es de 35.51% y del primer modelo 14.53%, con este último parámetro se opta por elegir el primer modelo, aclarando que también el segundo modelo puede resultar útil y de confianza.

$$CO = 10.7817 * \text{Banda 2} - 35.4572 * \text{Banda 3} + 33.2662 * \text{Banda 7} + 0.8466$$

## Capítulo 7. Conclusiones

Durante el desarrollo de este proyecto de investigación, se analizó una herramienta diferente para ayudar al área de la Ingeniería Sanitaria y Ambiental para llevar a cabo revisiones para el monóxido de carbono y así a tener un método diferente para realizar el monitoreo de los parámetros atmosféricos. Debido a que se tuvo la posibilidad de obtener el nivel 2 de información del Landsat 8, el procesamiento de las imágenes se realizó de manera más sencilla y eficiente, evitándose de esa manera el realizar las correcciones atmosféricas, confiando que las imágenes obtenidas habían sido revisadas por la NASA con un mínimo margen de error.

Aunque dentro de esta investigación decidió estudiar al monóxido de carbono, existen otros indicadores como las partículas suspendidas, partículas de plomo,  $PM_{2.5}$ ,  $PM_{10}$ , etcétera, se puede utilizar la percepción remota para la obtención de modelos que expliquen el comportamiento de estos parámetros.

En un inicio se tenía la incertidumbre de que los modelos obtenidos no tuvieran correlación algunas y tanto los datos como las variables tuvieran un gran grado de incertidumbre haciendo con esto que los modelos fueran poco precisos y confiables, por lo que fue necesario la obtención de parámetros estadísticos nuevos para darle un grado de certidumbre mayor y no solamente utilizar la  $R^2$ , como parámetro decisor, se debe de tomar en cuenta que dentro de las opciones de modelos se seleccionaron por medio de los valores y confiabilidad de los parámetros, pero estos no son los únicos que existen y si en un futuro se llegan a aplicar nuevos parámetros o distintos, como la Cross Validation, Boot Validation u otros, los modelos finales podrían ser otros, pero en este proyecto se utilizaron los que a criterio de varios autores son los parámetros con mayor confiabilidad y uso en la actualidad, pero dentro de lo obtenido se puede concluir que la reflectancia de las bandas del Landsat 8 presentan una buena correlación con los datos obtenidos de la RAMA, por lo que los modelos seleccionados tendrán una buena respuesta al momento de comparar los datos de campo con los datos de reflectancia.

Se observó que dentro de los modelos cuando se tiene temporada de lluvias el porcentaje de error es mayor, al contrario de la  $R^2$ , cuyo valor es más significativo, en periodo de estiaje, esto puede deberse a que la lluvia, como se sospechaba en un principio, afecta la reflectancia por lo que los modelos obtenidos en periodos de estiaje resultan con menor incertidumbre, haciendo con eso que los datos se comporten de mejor manera y el modelo se ocupe para predicciones

futuras con mayor grado de exactitud. Se espera que las políticas ambientales y el estudio de parámetros de calidad tomen mayor importancia dentro de los próximos años, como se ha visto en estudios realizados principalmente en China, Irán y Turquía, por lo que las investigaciones futuras tendrán parámetros estadísticos dedicados especialmente a la elección de modelos obtenidos por medio de la percepción remota y que las normas ambientales nacionales se revisen periódicamente ya que para el monóxido de carbono es de 1996, y en países asiáticos, de la Unión Europea y Estados Unidos se revisen cada año, para que se eviten condiciones extremas como las que se ven en Pekín, Shanghái y Terán. Por lo que con la investigación realizada contribuye al análisis y que, dentro del proyecto a entregar en la SECITI, se entregue un parámetro distinto al  $PM_{2.5}$  y  $PM_{10}$ , y existan modelos que expliquen los otros parámetros ambientales a monitorear con mayor grado de confianza para la generación de nuevas normas ambientales y crear un impacto dentro de la sociedad como muestra de ello fue la comparación realizada a los distintos niveles de monóxido de carbono en los últimos tres años para generar conciencia, además de que se explicaron los diferentes sistemas de percepción remota que existen mencionando que puede utilizarse otra plataforma o sistema para el mismo análisis pudiendo obtener resultados distintos.



# Anexos

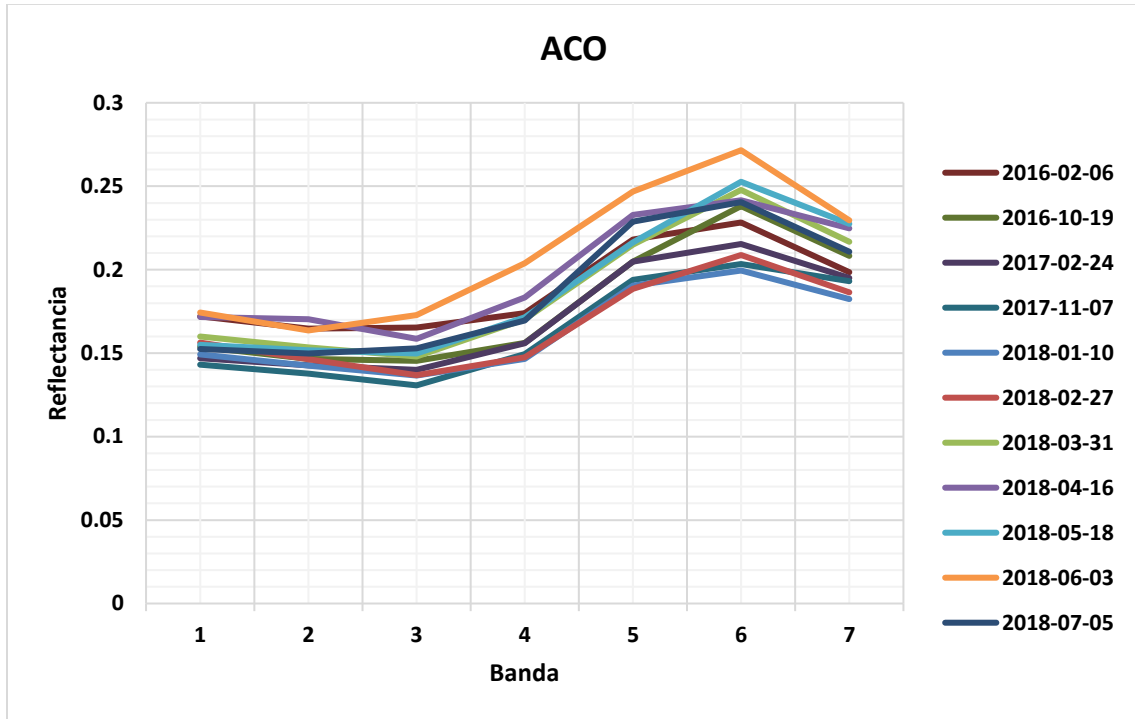


Imagen A.1: Gráfica reflectancia Acolman.

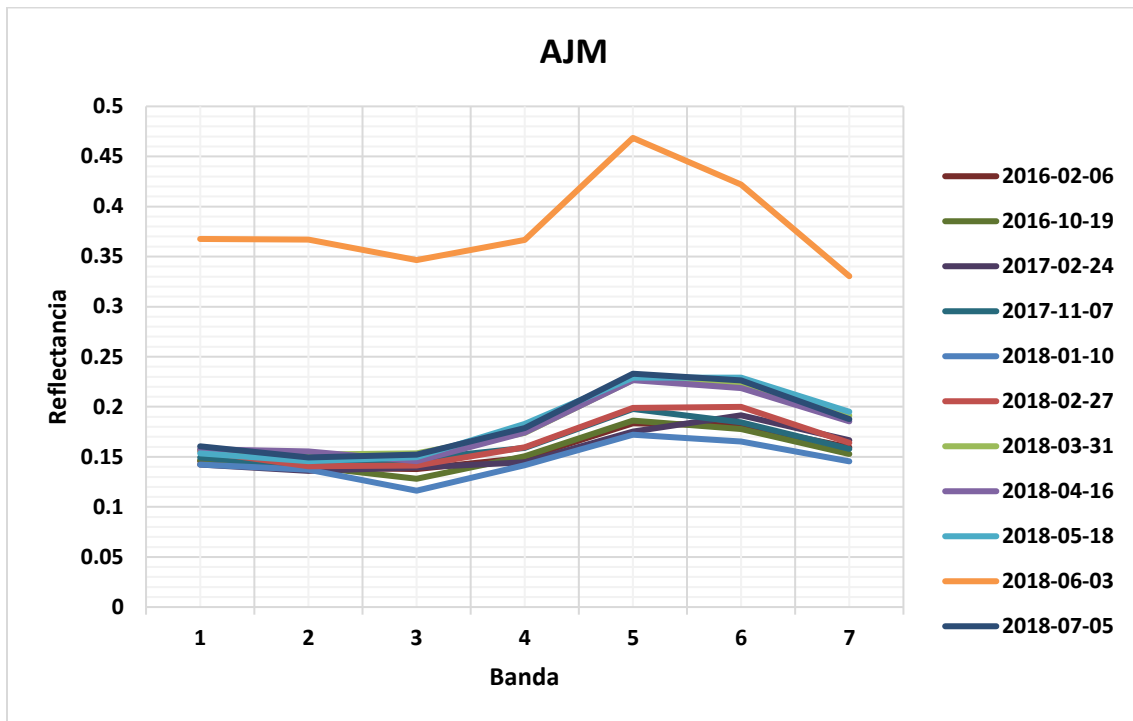


Imagen A.2: Gráfica reflectancia Ajusto Medio.

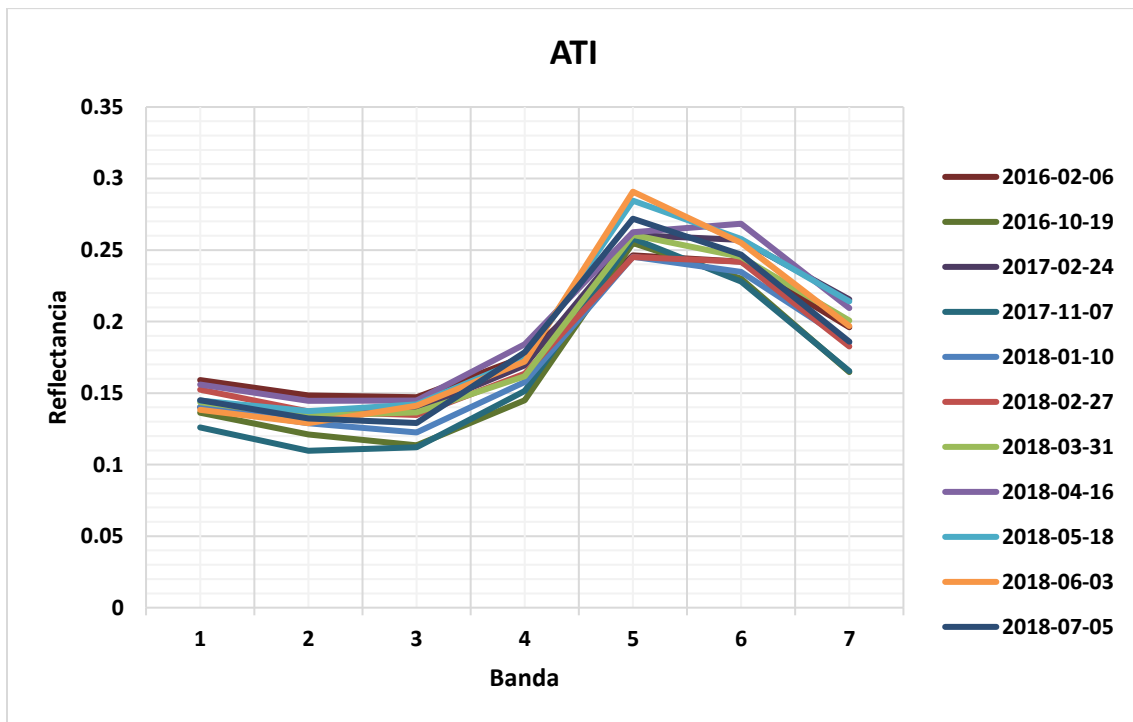


Imagen A.3: Gráfica reflectancia Atizapán.

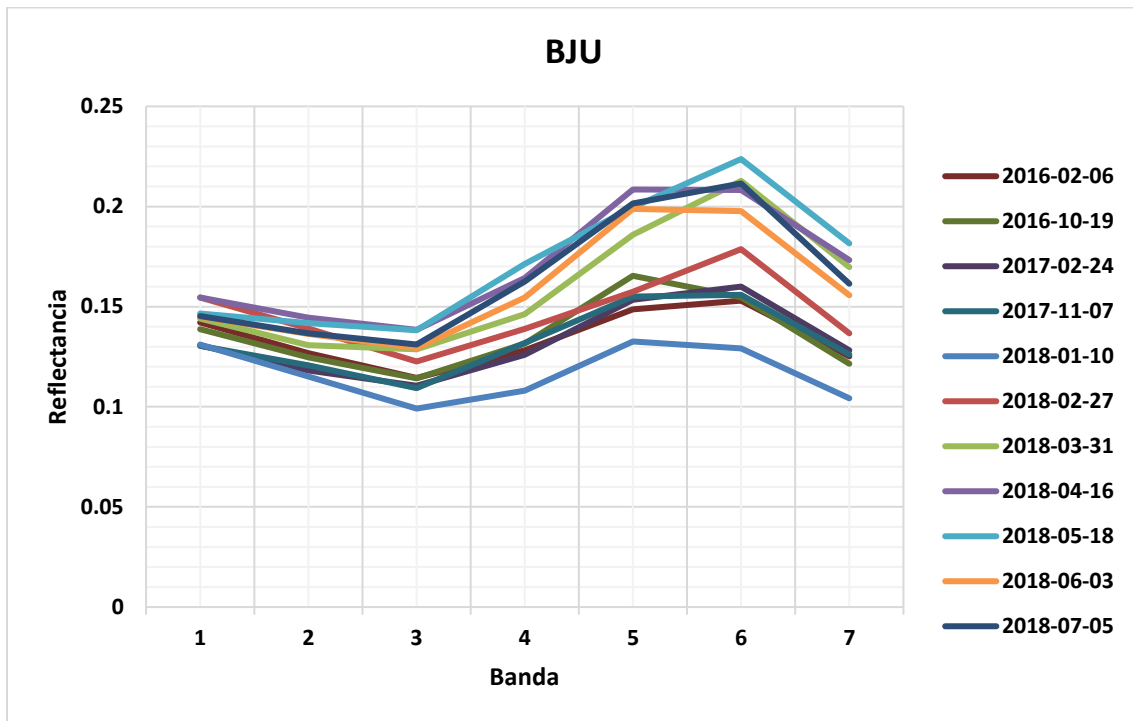


Imagen A.4: Gráfica reflectancia Benito Juárez.

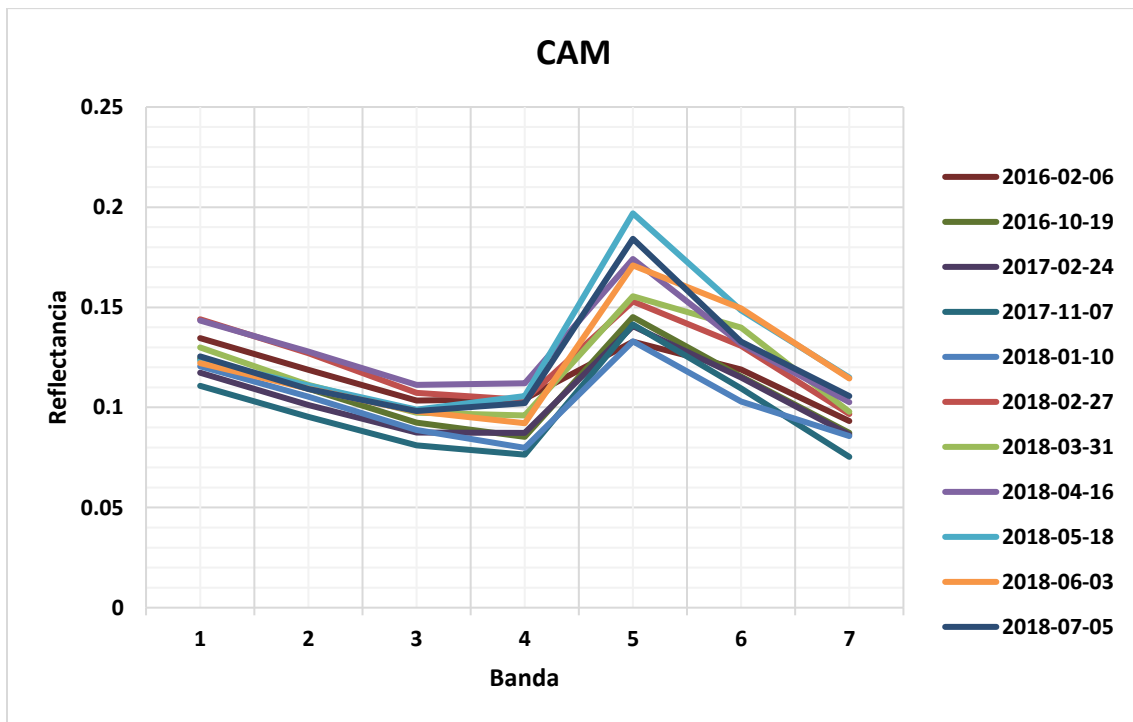


Imagen A.5: Gráfica reflectancia Camarones.

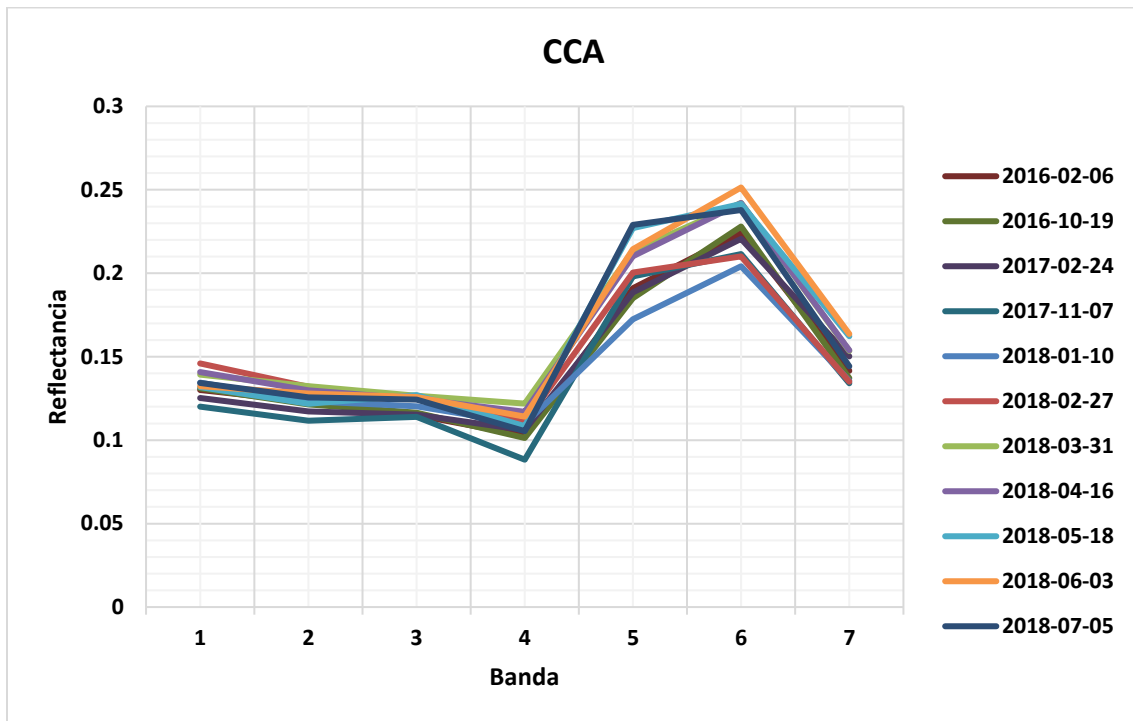


Imagen A.6: Gráfica reflectancia Centro de Ciencias de la Atmósfera.

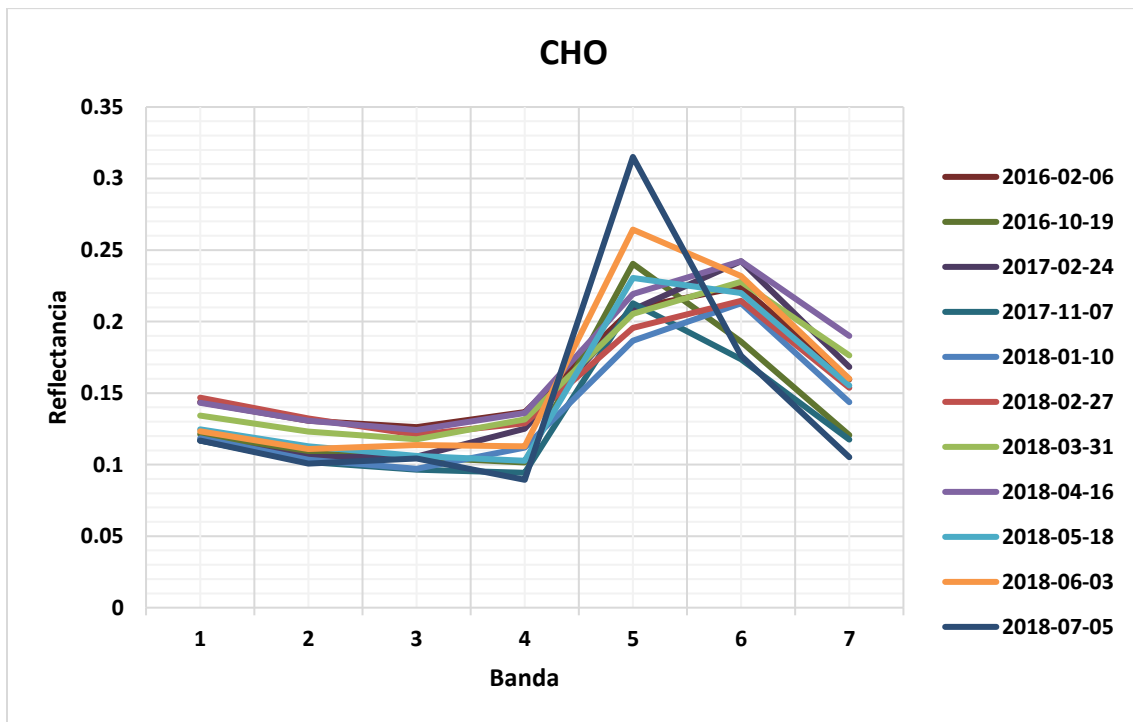


Imagen A.7: Gráfica reflectancia Chalco.

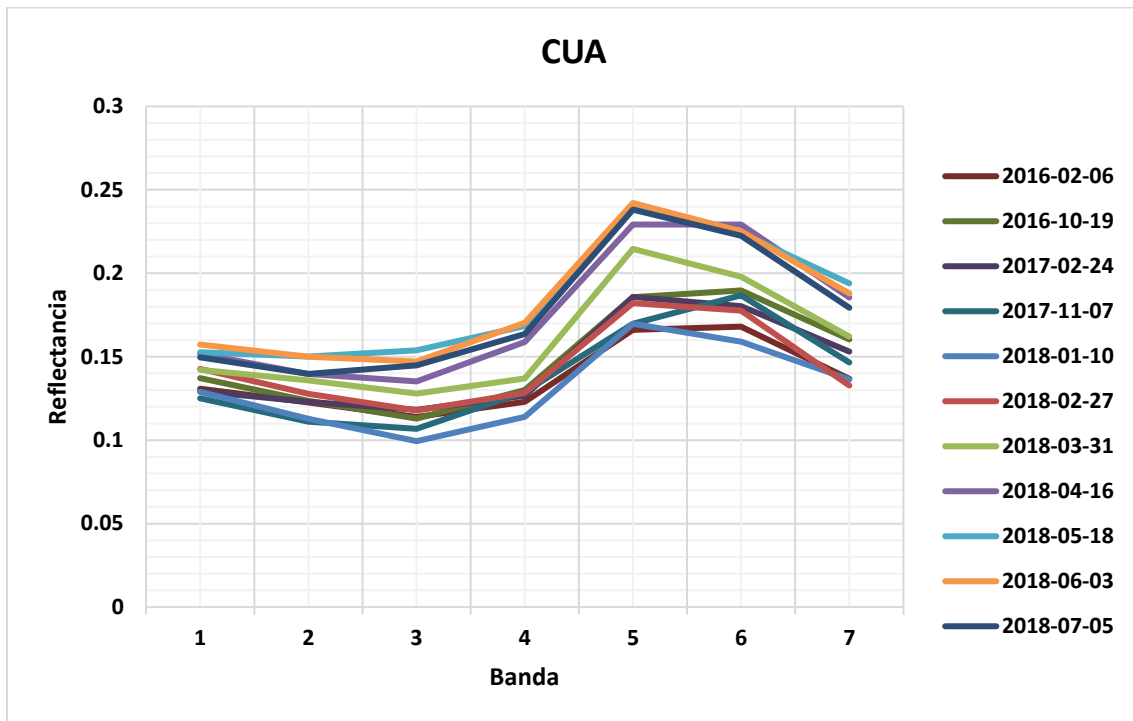


Imagen A.8: Gráfica reflectancia Cuajimalpa.

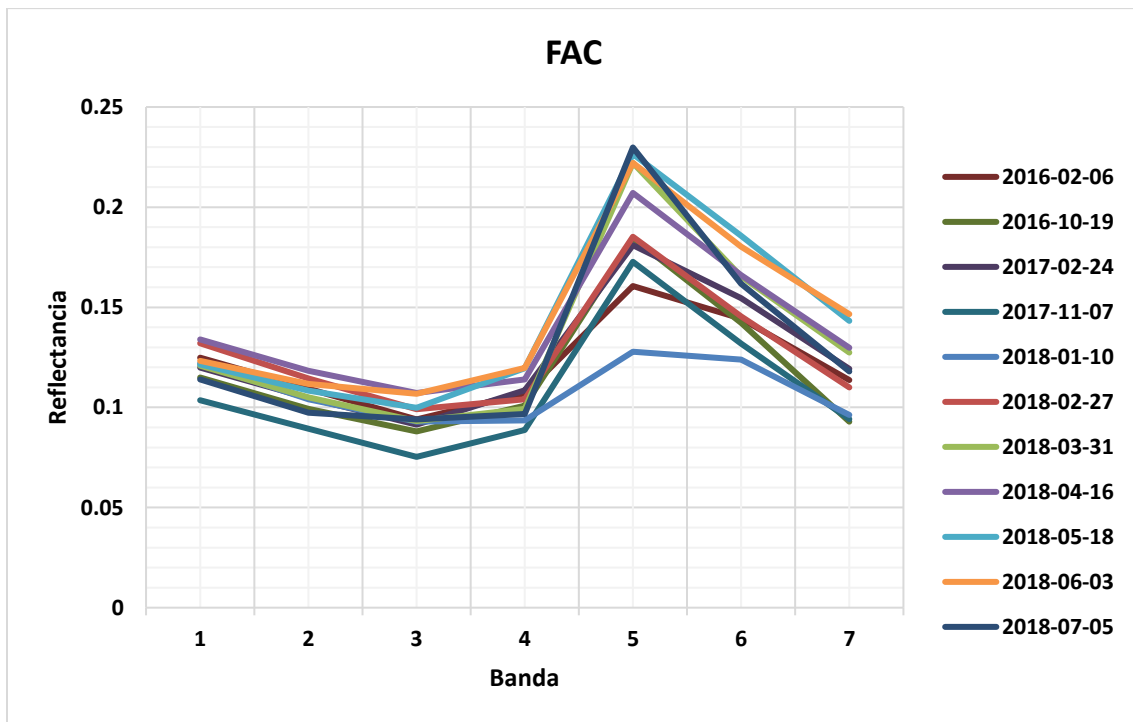


Imagen A.9: Gráfica reflectancia FES Acatlán.

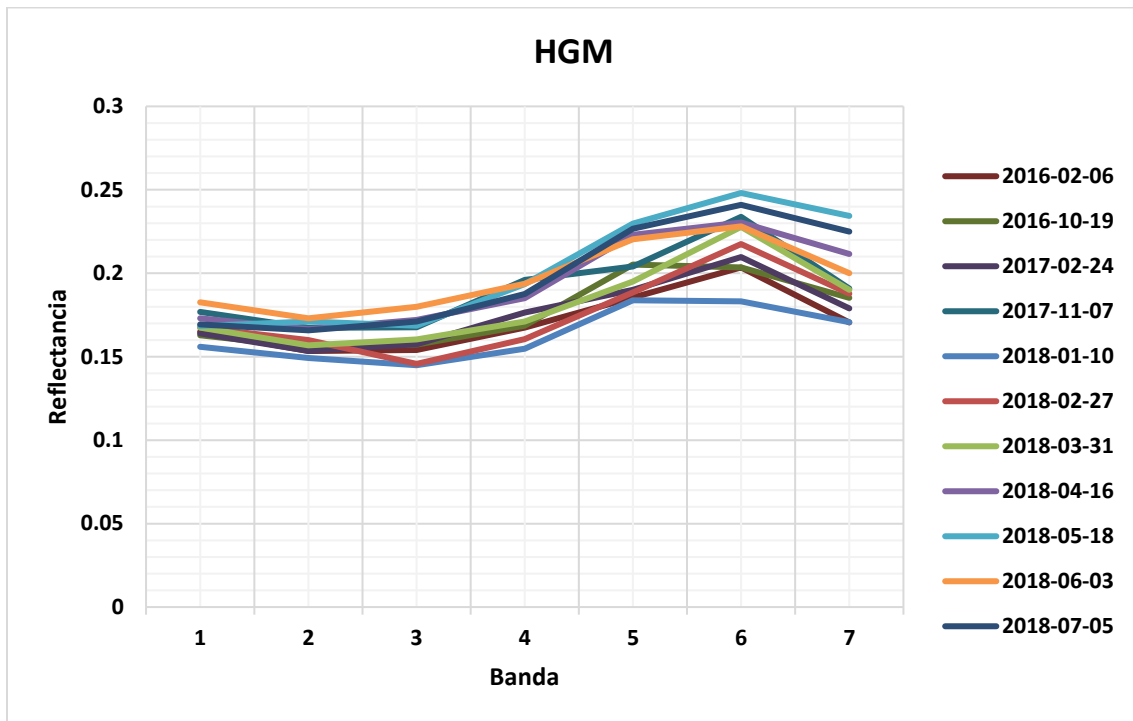


Imagen A.10: Gráfica reflectancia Hospital General de México.

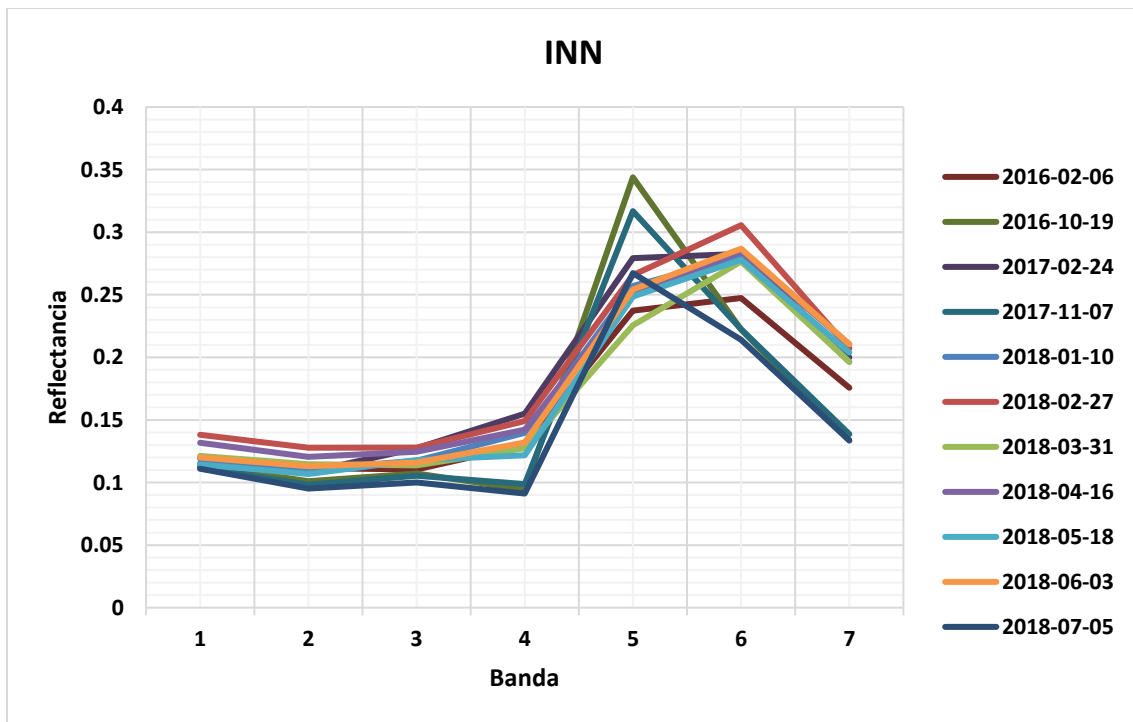


Imagen A.11: Gráfica reflectancia Investigaciones Nucleares.

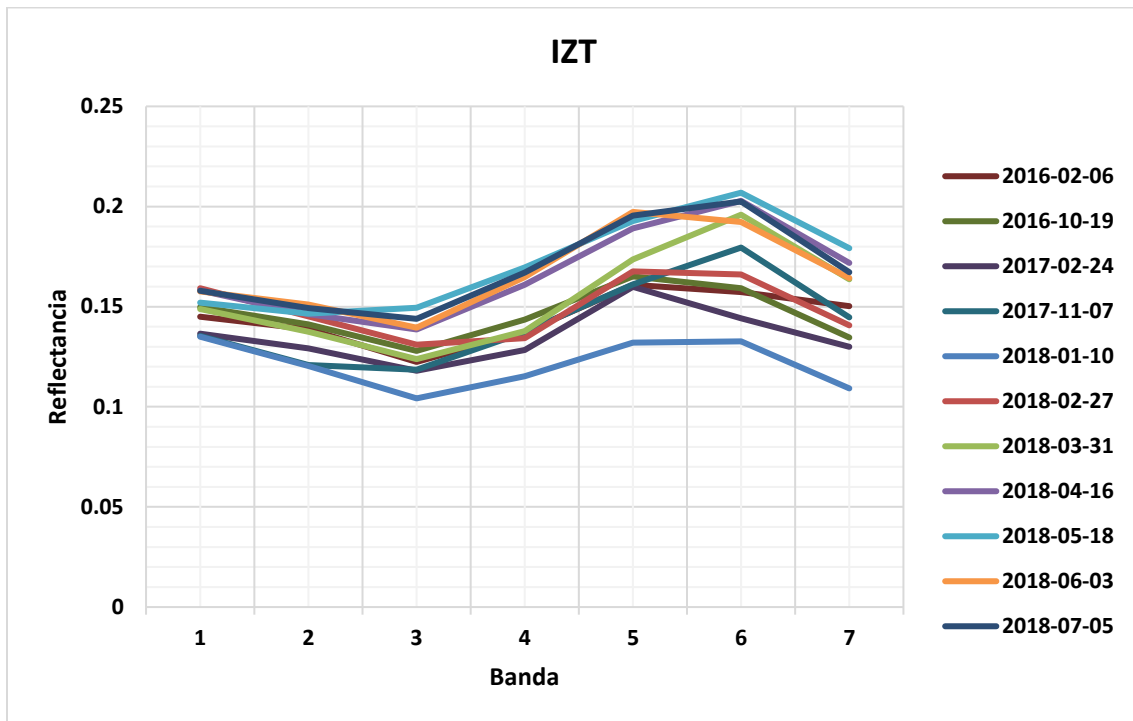


Imagen A.12: Gráfica reflectancia Iztacalco.

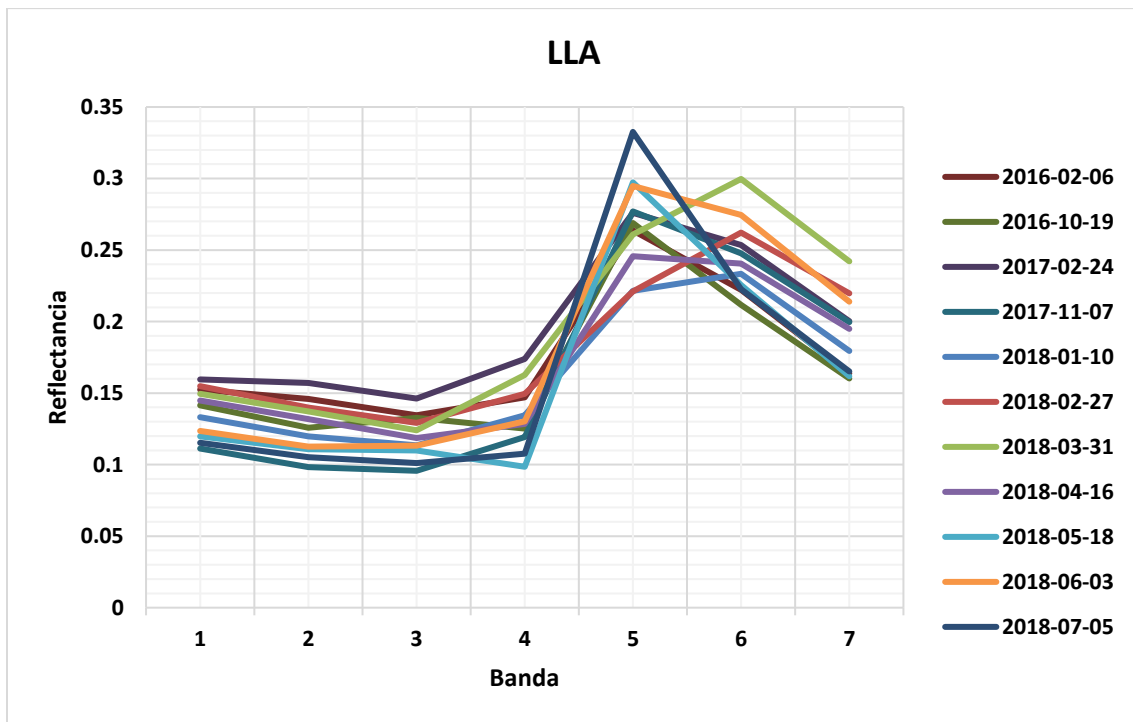


Imagen A.13: Gráfica reflectancia Los Laureles.

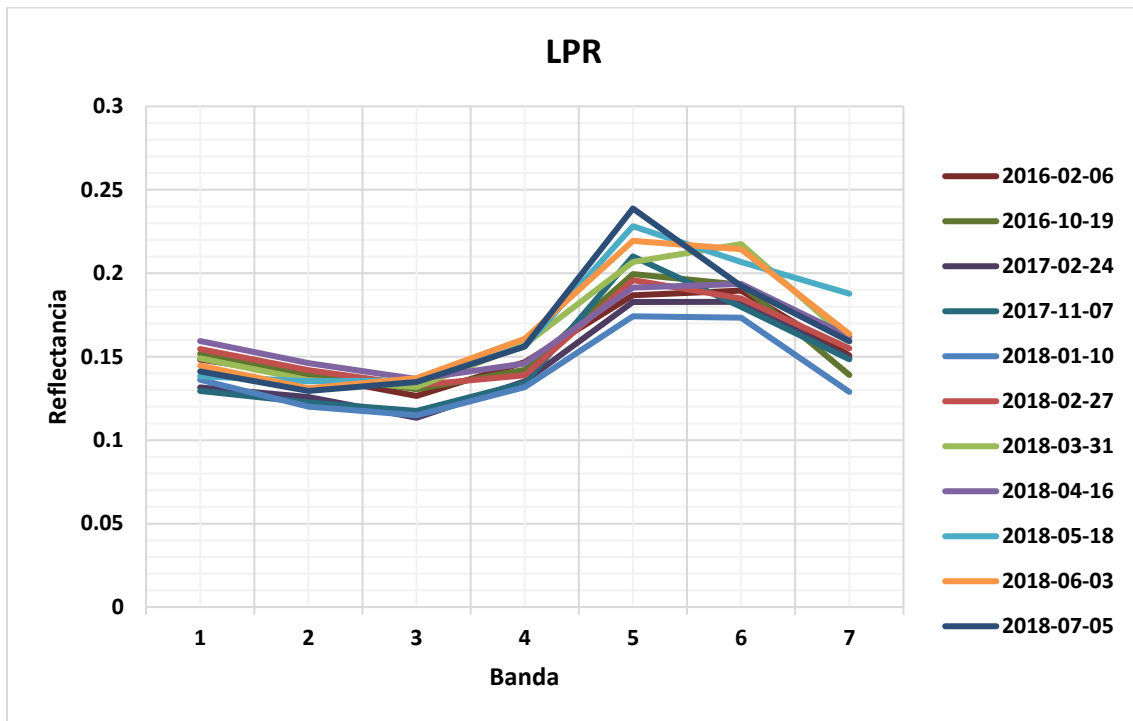


Imagen A.14: Gráfica reflectancia La Presa.

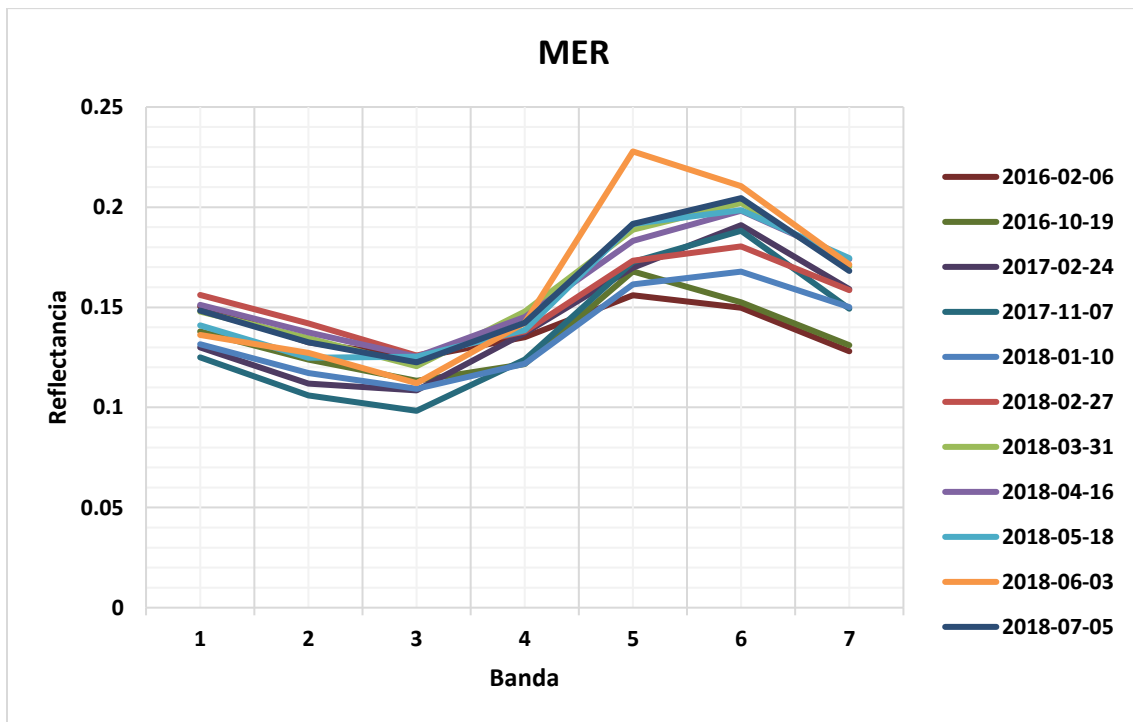


Imagen A.15: Gráfica reflectancia Merced.



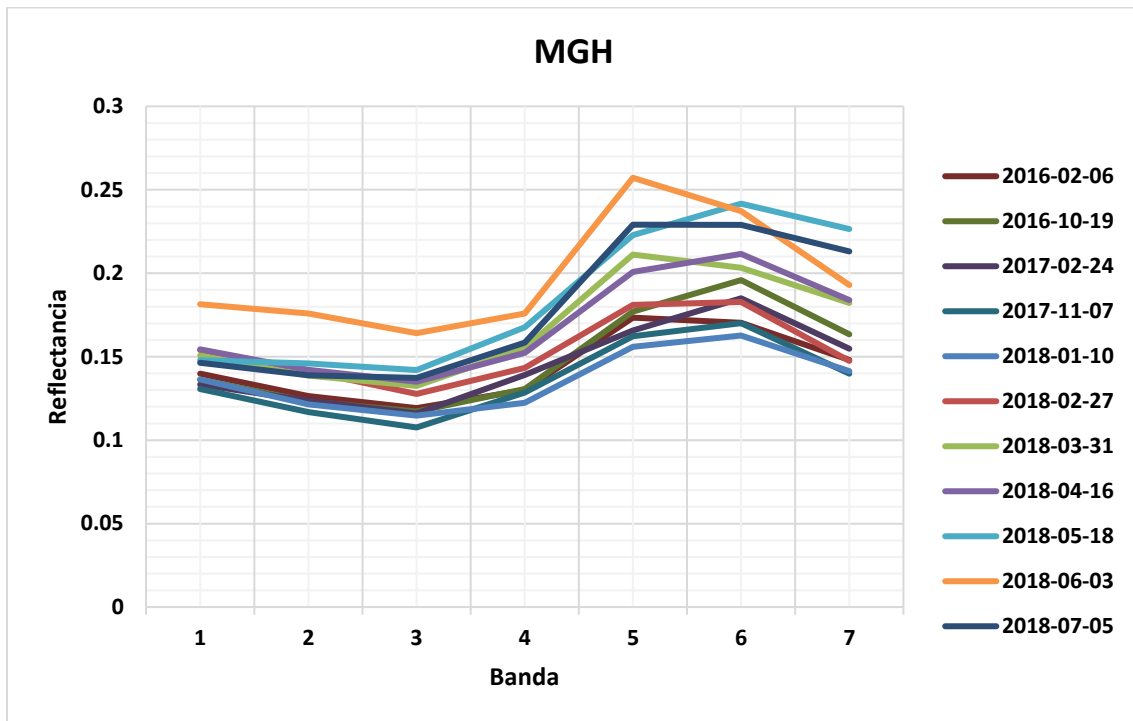


Imagen A.16: Gráfica reflectancia Miguel Hidalgo.

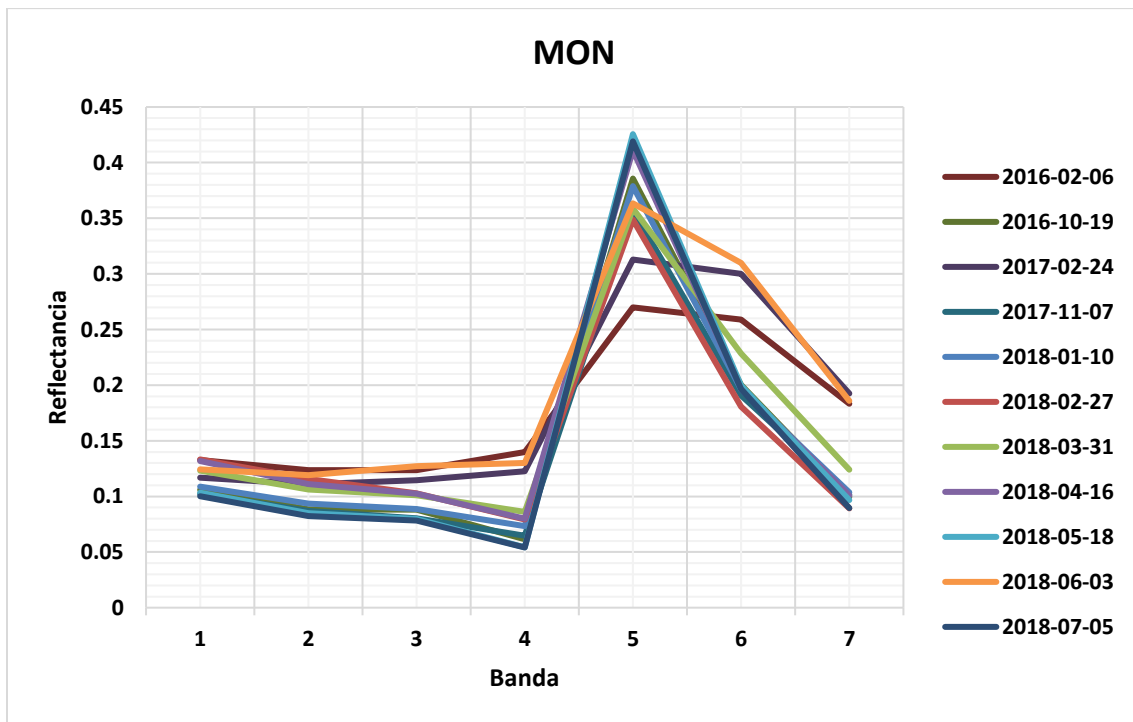


Imagen A.17: Gráfica reflectancia Montecillo.

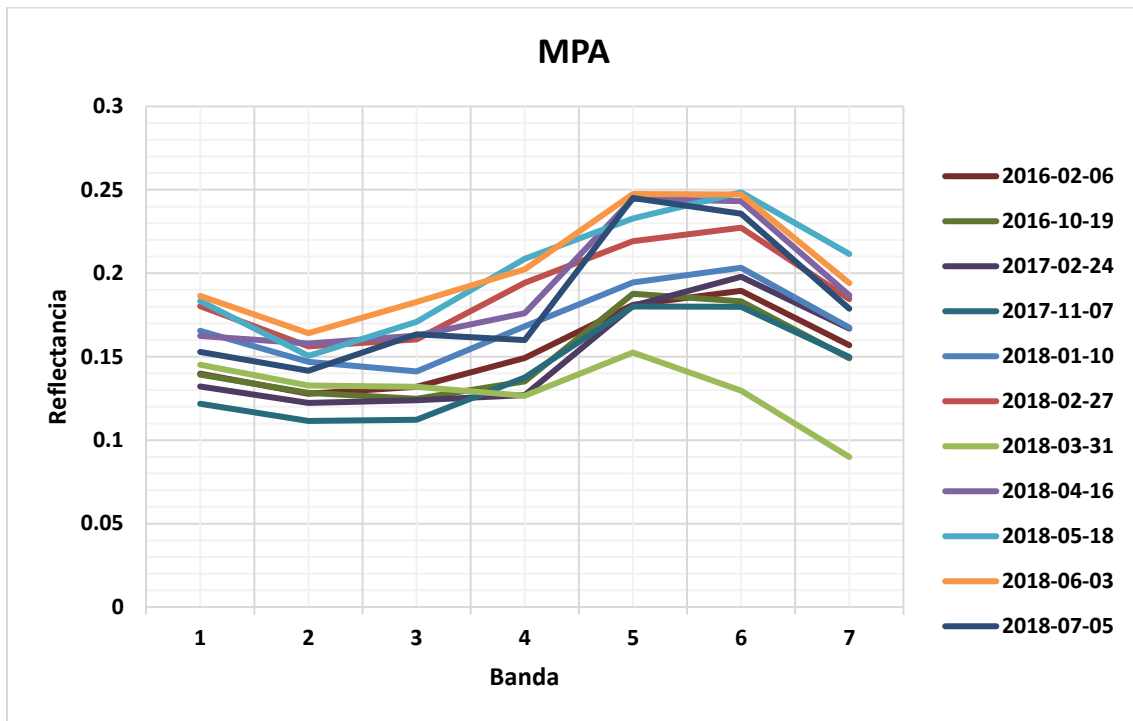


Imagen A.18: Gráfica reflectancia Milpa Alta.

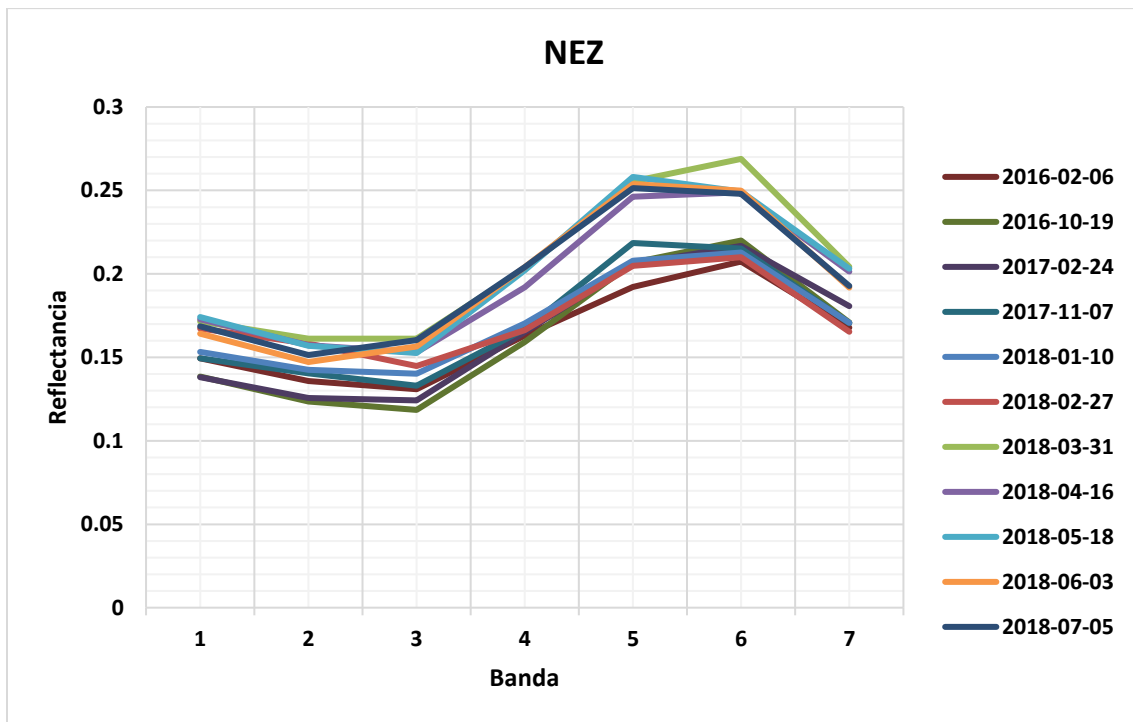


Imagen A.19: Gráfica reflectancia Nezahualcóyotl.

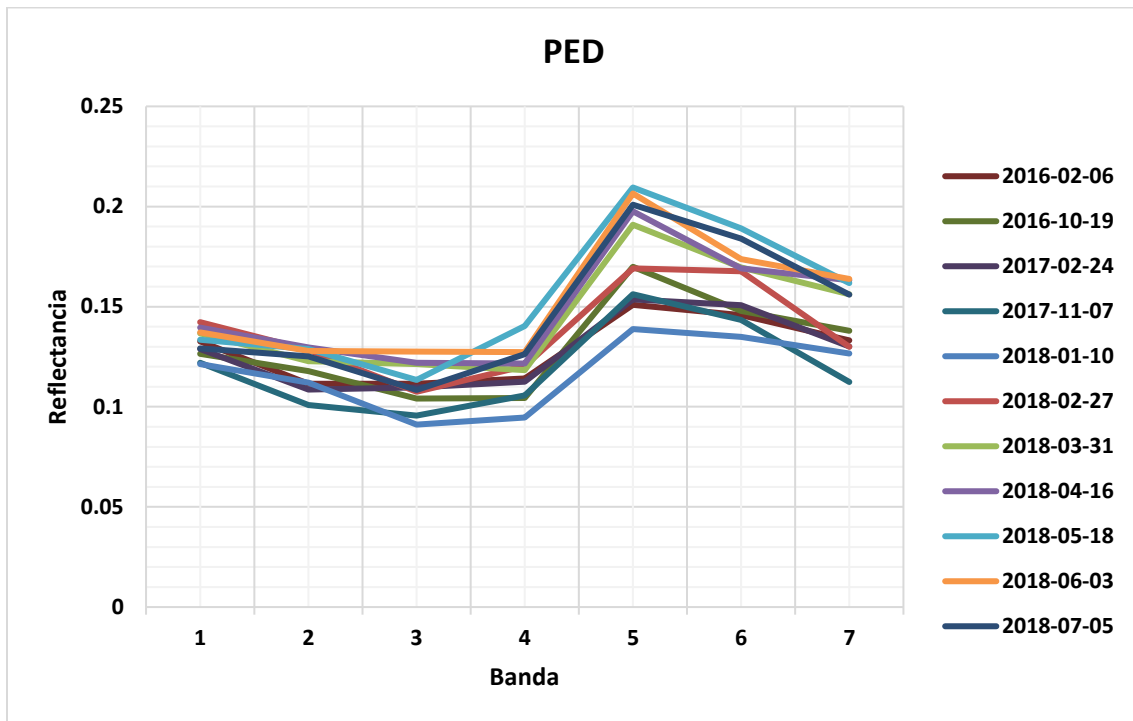


Imagen A.20: Gráfica reflectancia Pedregal.

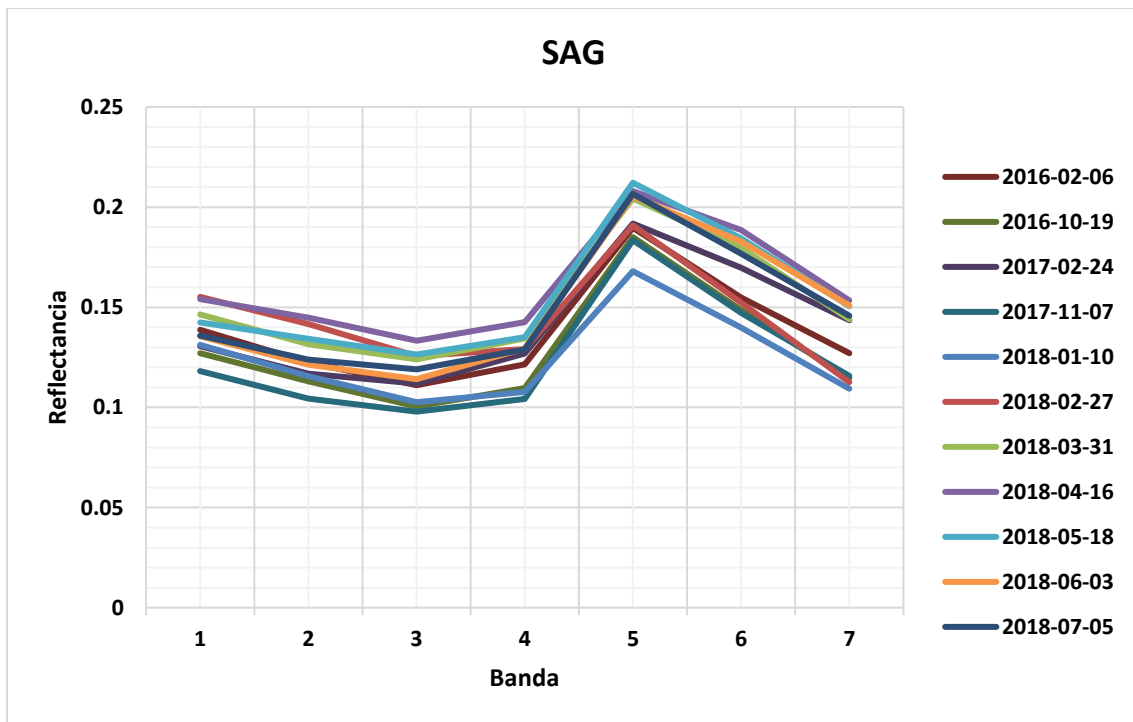


Imagen A.21: Gráfica reflectancia San Agustín.

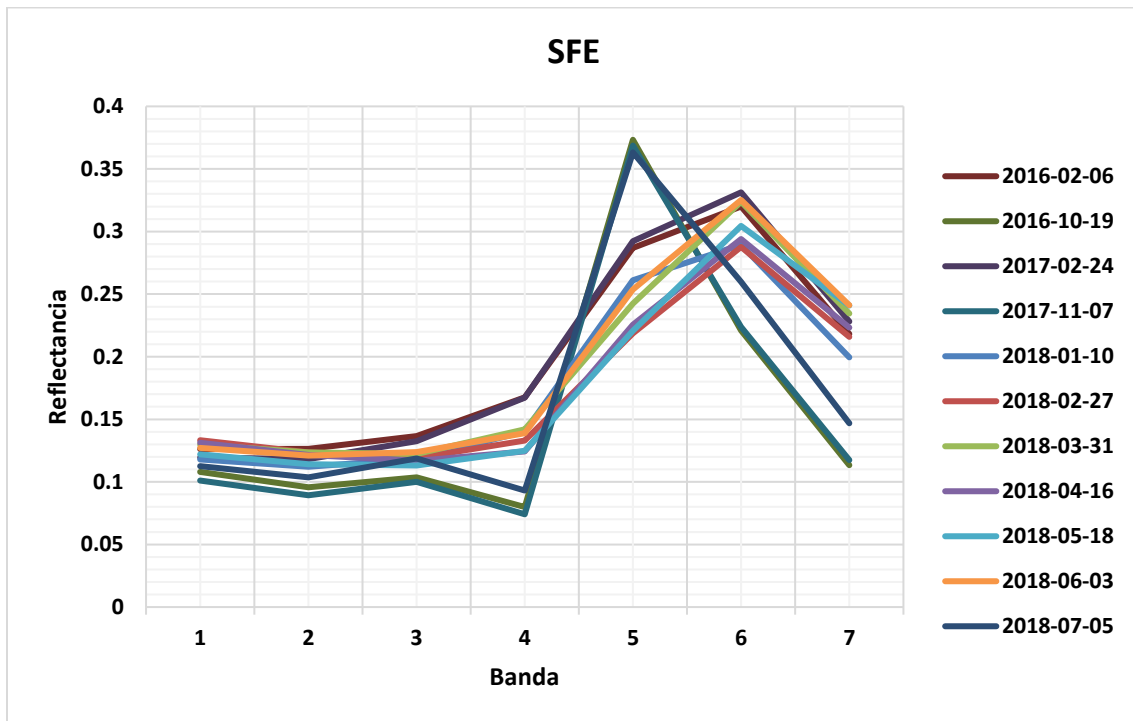


Imagen A.22: Gráfica reflectancia Santa Fe.

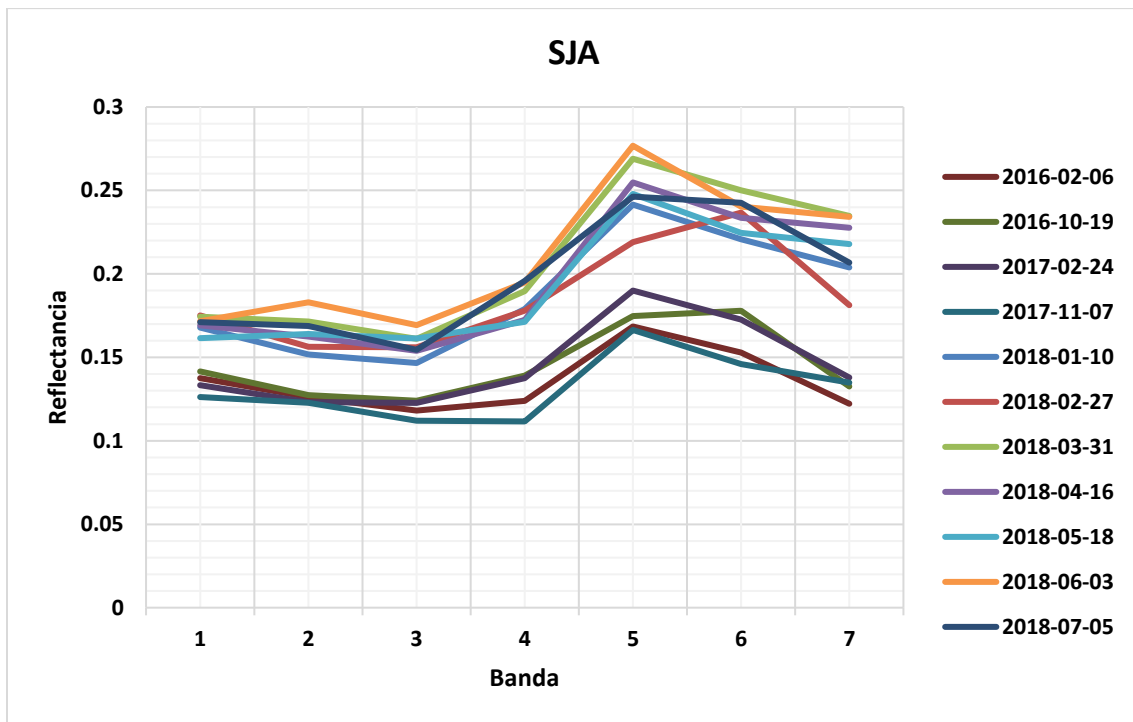


Imagen A.23: Gráfica reflectancia San Juan de Aragón.

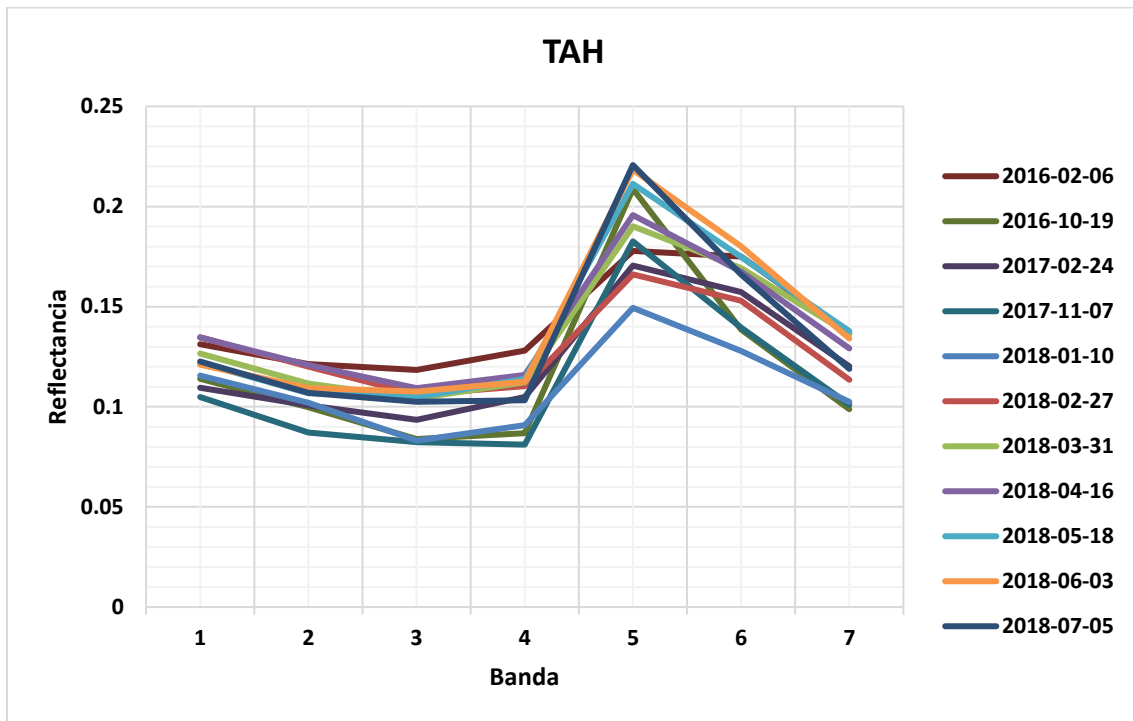


Imagen A.24: Gráfica reflectancia Tláhuac.

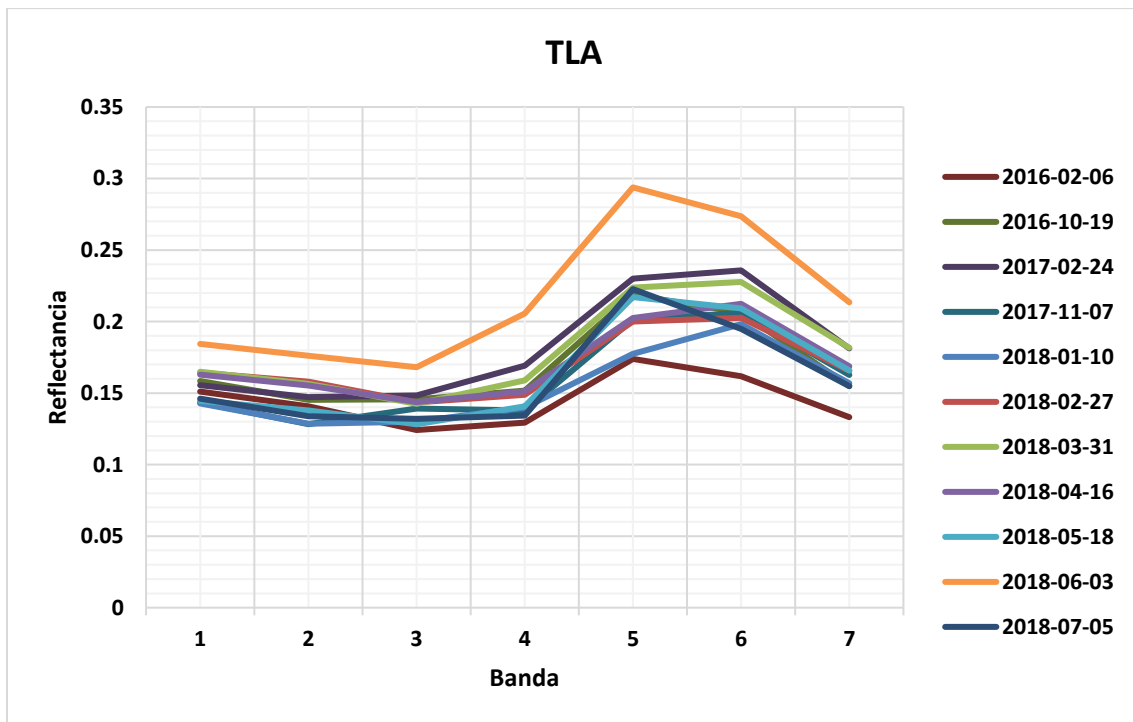


Imagen A.25: Gráfica reflectancia Tlanepantla.

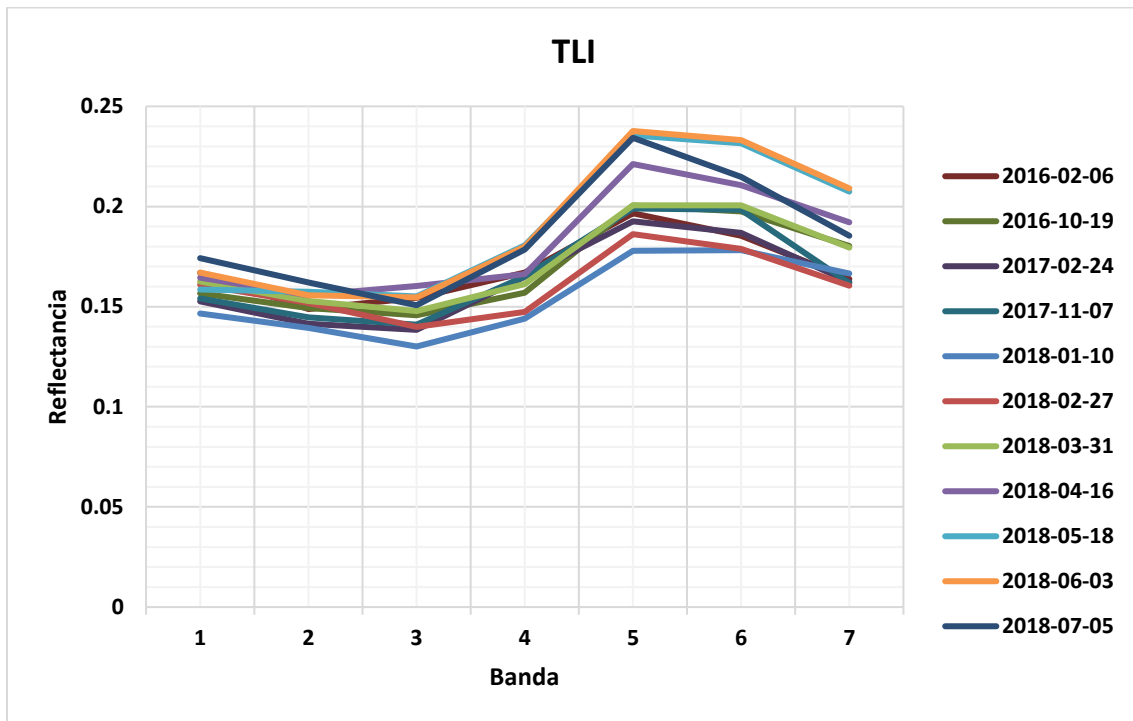


Imagen A.26: Gráfica reflectancia Tultitlán.

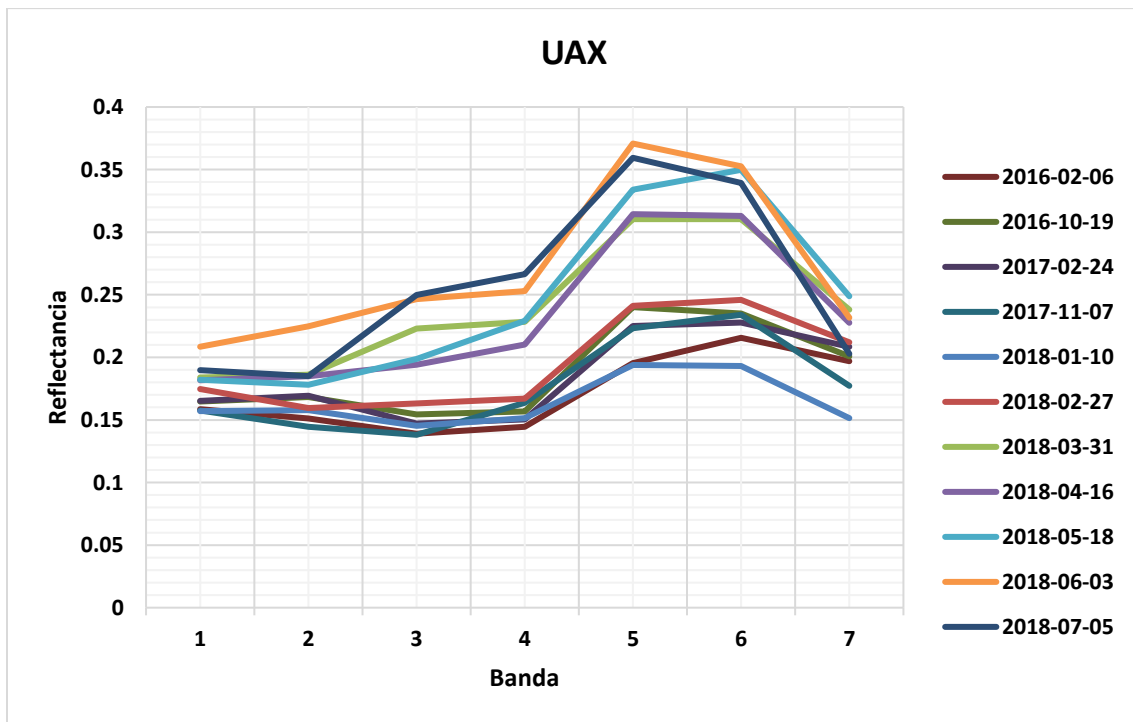


Imagen A.27: Gráfica reflectancia UAM Xochimilco.

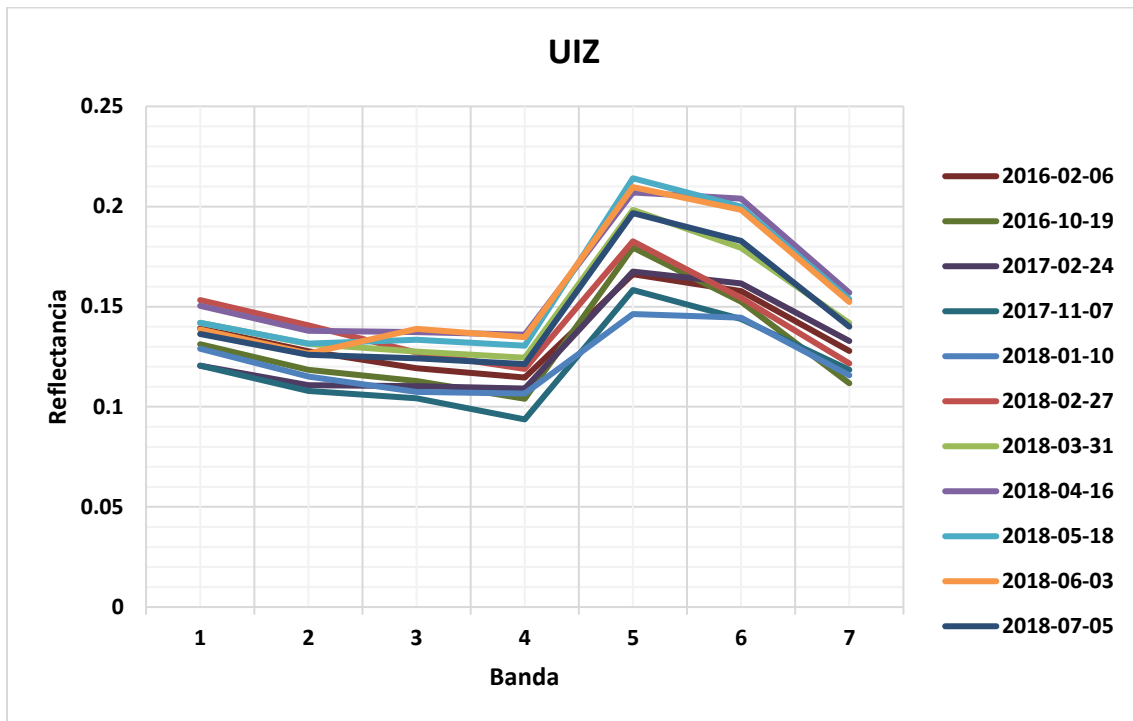


Imagen A.28: Gráfica reflectancia UAM Iztapalapa.

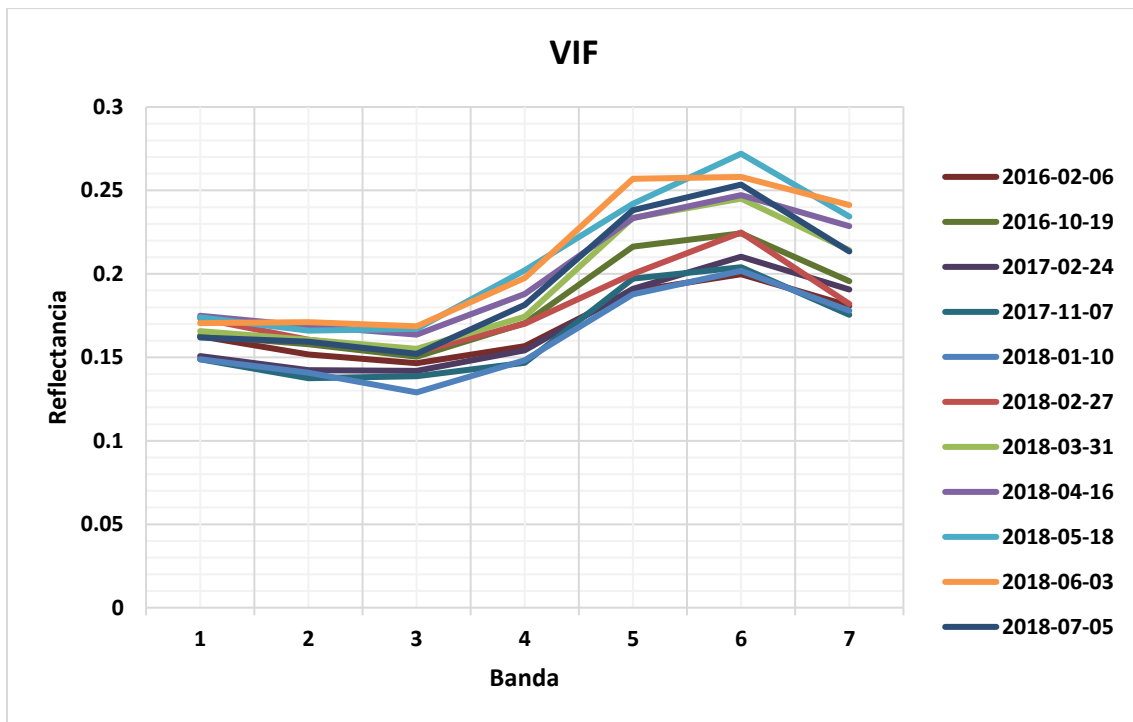


Imagen A.29: Gráfica reflectancia Villa de las Flores.

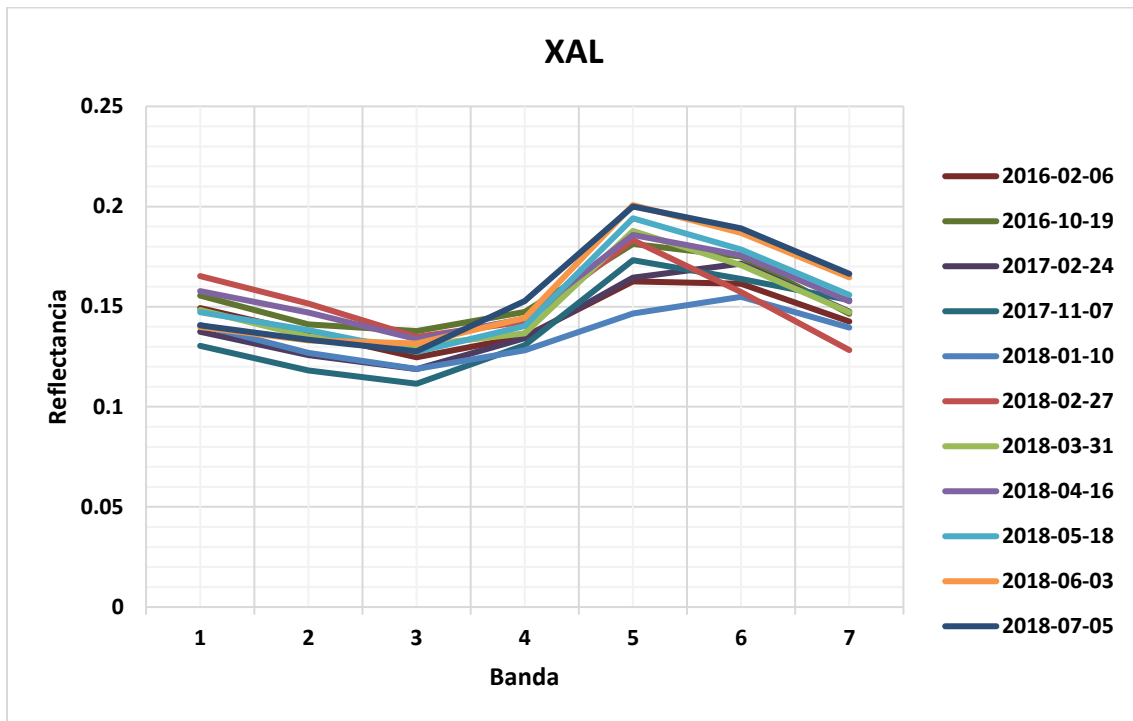


Imagen A.30: Gráfica reflectancia Xalostoc.



## Referencias

- Chuvieco, Emilio. (1996) *“Fundamentos y teledetección espacial”* 3° edición, ed. RIALP S.A. Madrid.
- Gónzales, J. (2005). *Formación y evolución de la atmósfera terrestre*. marzo 27, 2018, de UNAM, CIENCIORAMA Sitio web: [http://www.cienciorama.unam.mx/a/pdf/66\\_cienciorama.pdf](http://www.cienciorama.unam.mx/a/pdf/66_cienciorama.pdf)
- Buchdahl J., Hare S. (2010). *Natural Air Pollution*. marzo 15, 2018, de Manchester Metropolitan University Sitio web: [http://www.enviropedia.org.uk/Air\\_Quality/Natural\\_Air\\_Pollution.php](http://www.enviropedia.org.uk/Air_Quality/Natural_Air_Pollution.php)
- Portillo G. (2017). *¿Qué es la atmósfera y por qué es importante?* marzo 17, 2018, de Meteorología en Red Sitio web: <https://www.meteorologiaenred.com/la-atmosfera.html>
- Siegel E. (2017). *How Much CO2 Does A Single Volcano Emit?*. marzo 17, 2018, de Forbes Sitio web: <https://www.forbes.com/sites/startswithabang/2017/06/06/how-much-co2-does-a-single-volcano-emit/#5b0fded85cbf>
- Buchdahl J., Hare S. (2010). *Air Quality*. marzo 15, 2018, de Manchester Metropolitan University Sitio web: [http://www.enviropedia.org.uk/Air\\_Quality/History.php](http://www.enviropedia.org.uk/Air_Quality/History.php)
- Conde, C. (2006). *México y el cambio climático global*. México D.F., Centro de Ciencias de la Atmósfera, UNAM. Pp.28.
- Instituto Nacional de Ecología y Cambio Climático. (2018). *Efectos del cambio climático*. junio 17, 2018, de Instituto Nacional de Ecología y Cambio Climático Sitio web: <https://www.gob.mx/inecc/acciones-y-programas/efectos-del-cambio-climatico>
- National Geographic. (2010). *El aumento del nivel del mar*. junio 17, 2018, de National Geographic Sitio web: <https://www.nationalgeographic.es/medio-ambiente/el-aumento-del-nivel-del-mar>
- Ballester F. (2005). *Contaminación atmosférica, cambio climático y salud*. Revista Española de Salud Pública, 79, pp 1-5.
- Organización Mundial de la Salud (2002). *Informe de Estado de la salud en el mundo* junio 17, 2018, de Organización Mundial de la Salud. Sitio web: <http://www.who.int/whr/2002/chapter4/en/index7.html>.
- Cohen AJ, et al. (2003) *Mortality impacts of Urban Air Pollution*. Comparative Quantification of Health Risks: Global and Regional Burden of Disease Attributable to Selected Major Risk Factors. Ginebra: Organización Mundial de la Salud.
- Organización Mundial de la Salud. (2018). *Cambio climático y salud*. junio 20, 2018, de Organización Mundial de la Salud Sitio web: <http://www.who.int/es/news-room/fact-sheets/detail/cambio-clim%C3%A1tico-y-salud>

- Garduño R., Martínez J., Fernández A. (2004). *¿Qué es el efecto invernadero?* En Cambio, climático: una visión desde México (pp. 29-38). México D.F.: Instituto Nacional de Ecología. Secretaría del Medio Ambiente y Recursos Naturales.
- Pabón J., Nicholls R. (2005). *El cambio climático y la salud humana*. Biomédica Instituto Nacional de Salud, 25, p. 2.
- Red Ambiental de Asturias. (2015). *Gases de efecto invernadero*. 21 junio, 2018, de Red Ambiental de Asturias Sitio web: <https://www.asturias.es/portal/site/medioambiente/menuitem.1340904a2df84e62fe47421ca6108a0c/?vgnextoid=2a0e1e861647d210VgnVCM10000097030a0aRCRD&vgnextchannel=1e1c1cc03aa1a110VgnVCM100006a01a8c0RCRD&i18n.http.lang=es>
- Arellano E. (2011). *Evaluación de aerosoles troposféricos del AMCM por medio de imágenes de satélite GOES*. México D.F.: Instituto Politécnico Nacional.
- Martínez W. (2015). *Efecto de la calidad del aire en la zona metropolitana del valle de México sobre la respuesta espectral registrada por un sensor remoto*. México D.F.: Universidad Nacional Autónoma de México.
- Hamzelo M., et al. (2015). *MODELLING OF CARBON MONOXIDE AIR POLLUTION IN LARG CITIES BY EVALUETION OF SPECTRAL LANDSAT8 IMAGES*. Kish Island, Iran: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences.
- Elbir, T., (2004). *A GIS based decision support system for estimation, visualization and analysis of air pollution for large Turkish cities*, Elsevier, Atmospheric Environment, pp. 4509–4517.
- Sohrabinia, M., Khorshiddoust, AM., (2007). *Application of satellite data and GIS in studying air pollutants in Tehran*. Habitat International, pp. 268-275.
- Elbir, T., Mangir, N., (2010). *Development of a GIS-based decision support system for urban air quality management in the city of Istanbul*. Elsevier, Atmospheric Environment, pp. 441-454.
- Mohammed Hashim, B., Abdullah Sultan M., 2010. Using remote sensing data and GIS to evaluate air pollution and their relationship with land cover and land use in Baghdad City, Iranian Journal of Earth Sciences, pp.20-24.
- Hung Nguyen N. & Anh Tran V. (2014). *ESTIMATION OF PM10 FROM AOT OF SATELLITE LANDSAT 8 IMAGE OVER HANOI CITY A GIS*. Hanoi, Vietnam: Hanoi University of Mining and Geology.
- Health Effects Institute. (2018). *Burden of Disease Attributable to Major Air Pollution Sources in India*. India: Health Effects Institute.

- CONAGUA (2018). *Comisión Nacional del Agua*. 21 junio, 2018, CONAGUA. Sitio web: <http://www.conagua.gob.mx>
- Glovis (2018). *USGS Global Visualization Viewer*. 21 junio, 2018. Sitio web: <https://glovis.usgs.gov>
- INEGI (2018). *Instituto Nacional de Estadística y Geografía*. 23 junio, 2018. Sitio web: <http://www.inegi.org.mx/>
- RAMA (2018). *Red Automática de Monitoreo Atmosférico*, 5 julio, 2018. Sitio web: <http://www.aire.cdmx.gob.mx/default.php?opc=%27aKBh%27>
- Camara O. (2005). *Método contextual para fusión de datos*. 10 julio, 2018. México, D.F.: Universidad Nacional Autónoma de México.
- SEMARNAT (2018). *Secretaría del Medio Ambiente y Recursos Naturales*. 20 marzo, 2018. Sitio web: <https://www.gob.mx/semarnat>
- Daniela W., et al., (2013), *Introduction to Statistical Learning: With Application in R*, Springer. Estados Unidos.
- Julian J. Faraway, *Linear Models with R*, (2004), Chapman & Hall/CRC, Estados Unidos.
- Hatton Tim. *Air pollution in Victorian-era Britain – its effects on health now revealed*, (2016), University of Essex, Reino Unido.
- Camara, Alejandro. *Método contextual para fusión de datos*, (2005), UNAM, CDMX.