

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

ESCUELA NACIONAL DE CIENCIAS DE LA TIERRA

INFORME SERVICIO SOCIAL

NOMBRE DEL ALUMNO(A): Malagon Nabor Jonathan

CARRERA: LCT ciencias de la Tierra

NÚMERO DE CUENTA UNAM: 318335696

NOMBRE DE LA ENTIDAD RECEPTORA: Instituto de Ciencias Atmosféricas y
Cambio Climático

NOMBRE DEL PROGRAMA: Investigación Atmosféricas y Ambientales

CLAVE DEL PROGRAMA: 2024-12/72-3414

PERIODO DEL SERVICIO SOCIAL: 15-octubre-2024 al 9-Mayo-2025

**Nombre y firma del(a)
alumno(a)**

**Nombre, firma y sello de la
entidad receptora**

Índice

1. introducción
 - 1.1 Análisis de datos de parámetros atmosféricos y ambientales medidos en el Observatorio Atmosférico Calakmul.
 - 1.2 Elaboración de Scripts para el filtrado, control de calidad y visualización de los datos de distintos sensores.
 - 1.2.1 PMT640 Mass Monitor.
 - 1.2.2 Picarro Gas Concentration Analyzer
 - 1.2.3 Pandoras
 - 1.3 Apoyo en la instalación y revisión de instrumentos atmosféricos en campo.
2. EXPERIENCIA EN TÉRMINOS DE LO SOCIAL, FORMATIVO Y RETRIBUTIVO
3. Conclusiones
4. ANEXO

1. Introducción

DESARROLLO DE ACTIVIDADES Y RESULTADOS OBTENIDOS

1.1 Análisis de datos de parámetros atmosféricos y ambientales medidos en el Observatorio Atmosférico Calakmul.

Primera revisión de las bases de datos generadas por los instrumentos instalados en el observatorio, Organización y actualización de los repositorios, discusión de las estrategias para el manejo de datos.

Se descargó software Google Colab, Visual Code Estudio, Python 3, y Github para la revisión de los datos.

Los datos utilizados para revisión de datos del Observatorio atmosférico de Calakmul, se tomaron de los servidores de la ROUA.

Los archivos ya contaban con un nivel L1 teniendo datos por minuto y hora. Se tomaron los datos de toda la serie de tiempo disponible desde noviembre de 2023 hasta Marzo de 2025.

Los datos contaban con inconsistencias y pérdida de datos producto de fallas en la alimentación u fenómenos ajenos que hicieron que los sensores no registraron datos.

Se corrigieron los datos que tenían un comportamiento anómalo, para la serie de tiempo de Temperatura, se hizo una regresión lineal basados en las observaciones de otro sensor PMT640 mass monitor, que cuenta con un sensor de temperatura, es decir se buscó correlacionar datos faltantes en la serie de tiempo de meteorología usando datos existentes del instrumento PMT640 y se agregaron flags (a las columnas) que indicaban donde los datos donde se le llevo a cabo la regresión lineal 'RL' o la interpolación (IP).

Finalmente se entrego un archivo de tipo L2 que contenía estas correcciones en tipo .CSV con una gráfica que contiene 4 graficas que contienen la información de Temperatura, Humedad Relativa, Dirección y velocidad del viento, Radiación, presión atmosférica y precipitación de toda la serie de tiempo y un climograma, mostrando el valor promedio, máximo y mínimo de temperatura, así como la suma total de precipitación por mes.

1.2 Elaboración de Scripts para el filtrado, control de calidad y visualización de los datos de distintos sensores.

La codificación de scripts de para leer la información capturada en los archivos con datos crudos.

Definición de criterios de filtrado y depuración de datos incorrectos para la elaboración de nivel 1 y nivel 2 de las bases de datos con controles de calidad. Creación de promedios y tablas estadísticas de los datos. Elaboración de scripts para la visualización de la información en formato de series de tiempo, histogramas, rosas de viento y de contaminantes, correlaciones, etc.

Este apartado se detallarán los diferentes métodos y algoritmos utilizados para el procesamiento de los distintos sensores.

1.2.1 PMT640 Mass Monitor.

Es un instrumento para la lectura en tiempo real de material particulado PM10 y PM2.5.

Se trabajaron con los datos crudos de toda la serie de tiempo, el primer paso fue revisar la serie de tiempo estuviese completa se agregaron fechas faltantes, los renombraron las columnas de PM 10 y 2.5, se creó un nuevo repositorio donde se guardaron todos los archivos de forma minutal y por hora los archivos fueron guardados en documentos .cvs.

Se llevó a cabo el ploteo de varias figuras combinando los archivos de meteorología con los de partículas para generar rosas de viento, para llevar a cabo el promedio de viento.

Se realizó scripts para poder graficar toda la serie de tiempo y también poder discriminar datos por fecha y valor de concentración, esto pensando a futuro en caso de querer estudiar eventos más concretos en una serie de tiempo o simplemente observar los eventos mayores de rosa de vientos.

1.2.2 Picarro Gas Concentration Analyzer

Es un instrumento que analiza simultáneamente gases de efecto invernadero, monóxido de Carbono (CO), dióxido de Carbono CO₂ y Metano CH₄

La ROUA procesa los datos crudos de este instrumento pero debido a algunos errores en su script, los datos presentaban inconsistencias, así que se partió de trabajar con los datos crudos para toda la serie de tiempo, debido a la incompatibilidad entre forma de trabajar, pues se consideraba conservar el script original, pero después de varias sesiones se decidió trabajar en uno nuevo, tomando de base la forma de filtrar y aplicarle controles de calidad, así mismo se trabajaron controles de calidad extra para la estación Calakmul. Al trabajar con los datos crudos fue necesario buscar otro tipo de separador de datos para hacer el procesamiento y debido a la cantidad y tamaño de los datos se hizo en varias partes una parte del procesamiento fue directamente realizada dentro del servidor 32 del instituto de ciencias atmosféricas, fue necesario filtrar a través flags los valores que correspondían la posición de la válvula teniendo esos cuatro valores para cada gas se ocupó una combinación para discriminar que datos tomar (tabla 1)

Tabla 1. Flags de filtrado	
GAS	Flag
CO	1,4
CO ₂	2,3
CH ₄	3

El picarro cuenta con una frecuencia de Hertz muy alta, es decir puede realizar mediciones en milisegundos, así que se una vez aplicados los primeros flags, se promediaron los valores restantes por minuto y así obtener la desviación estándar.

Segundo paso fue corregir la línea de tiempo utc -6h y una corrección de 170 segundos que toma llegar el aire de la torre de la estación al picarro dentro del contenedor de la estación.

A la serie de tiempo se le aplicó un umbral basado en los valores promedio de la desviación estándar obtenida previamente valores que no correspondieron a valores conocidos de concentración de los gases fueron eliminados

Debido a la pérdida de datos monóxido de carbono por fallas en las válvulas, los datos recabados para 2024 no son precisos, pero sí con datos de CO₂ y CH₄ los cuales se quitaron los valores menores a 300 ppm para CO₂ y menores a 1.6 ppm para CH₄ ya que estos ya son concentraciones menores al promedio mundial.

Se llevaron a cabo las series de tiempo anual para el año 2024 y se empezó a trabajar con perfiles de tiempo esto tomando como base en promedio por hora por cada mes para así obtener un perfil de cómo se comportan los gases anualmente

Que administraron los promedios horarios con distintos intervalos de tiempo en eso tuvieron promedios de las 19 horas en horas de las 9 a las 16 de todo el día, así como su promedio anual.

Se discutieron los datos y se llegó a la conclusión de que se debe llevar a cabo otro umbral para eliminar más valores por desviación estándar para producir un nuevo producto I1b y así limpiar un poco de ruido en la serie de tiempo se borró datos de desviación estándar mayor o igual a 0.2 para CO_2 y desviación estándar mayor igual a 0.002 para CH_4

Se compararon ambas series de tiempo I1 y I1 b y se decidió dejar el formato de I1b que muestra los valores más suavizados

1.2.3 Pandoras

Pandora es una red global lanzada por la NASA en colaboración con la ESA la estación espacial europea que establecen estaciones de medición de observaciones de concentración de gases estraza en columnas totales y verticales para apoyar a la validación y verificación de sensores y visible en satélites órbitas bajas y gestacionales.

Datos de nivel L2 se trabajó con un script previo, pero se creó un nuevo script a partir de Dataframes para modernizar el código y hacerlo más replicable, se revisó el número de columnas de interés la columna de tiempo columna de NO_2 y flags

El valor de tiempo tiene un formato ISO-86 aquí se muestra el formato de por lo que es necesario pasar al formato año mes día hora minuto segundo

Y aplicarle la corrección UTC - 6 horas para cada valor se decidió tomar los valores debido a que la corrección L2 deja pocos datos en un intervalo de aproximadamente las 7 a.m. a las 17 p.m.

1.3 Apoyo en la instalación y revisión de instrumentos atmosféricos en campo.

Revisión de sensores para la medición de parámetros micrometeorológicos y de respiración de suelo. Configuración y adecuación de equipos para la adquisición de los datos en formatos requeridos.

Pruebas en laboratorio. Instalación y operación de equipos en campo

Se llevó a cabo la revisión de diversa instrumentación de fabricación propia por parte de la técnica académica María Eugenia y la limpieza de distintos muestreadores e instrumentos que se encontraban en la torre ubicada en Cu para prepararlos en la salida de campo hacia Calakmul la cual costó de cuatro fases la primera fase consistió en preparar en campo el equipo de muestreo y sincronizar los Datalogger con los sensores y asegurar que estos contarán con la suficiente batería, también se verificó si el sello era hermético para que el muestreador de gases pudiese dar buenas señales, en la segunda fase se llevó a cabo el muestreo en el sitio Oxpeul y se llevaron a cabo 16 muestreos siguiendo las direcciones norte sur este oeste cada muestreo duró aproximadamente 10 minutos donde la cámara de aire se colocaba en la superficie del suelo haciendo un sello entre la cámara de gas y la superficie dejando que el sensor captara la respiración asimismo sensores de temperatura

a 15 cm del suelo fueron colocados así como dos sensores de humedad fueron colocados uno a nivel de suelo y el otro enterrado aproximadamente a 10 cm dentro del suelo para poder conocer las condiciones del terreno.

La fase la fase 3 consistió en darle mantenimiento a la estación de Calakmul dentro del contenedor, se hizo un cambio en el motor que permite que el Domo que protege al espectrometro pudiese abrir y cerrar de nuevo, después una modificación al tablero de control agregando un switch que permite cambiar entre la alimentación de corriente alterna a corriente directa esto con el fin de facilitar el cambio entre alimentación cuando no haya electricidad en la estación y que los equipos puedan seguir operando, se hizo un respaldo y un cambio de un disco duro sólido tipo hdd a un disco de tipo sólido ssd para la computadora que lleva a cabo el procesamiento de los datos del espectrómetro, y se cambió el tanque de nitrógeno líquido que da refrigeración al espectrómetro y finalmente se intentó hacer un cambio en las válvulas del picarro para poder solucionar los problemas en la medición del monóxido de carbono los cuales habían causado la falta de datos previamente mencionados durante el trabajo el procesamiento de estos datos.

La fase 4 consistió en realizar los mismos procedimientos para la toma de datos de suelo ahora en el sitio denominado El palmar ubicado en sisal Yucatán se llevan a cabo el mismo muestreo y adicionalmente se tomaron tres puntos cercanos a la torre de El palmar para comparar esos datos con los que habrá en un futuro de forma fija alrededor de la estación.

2. EXPERIENCIA EN TÉRMINOS DE LO SOCIAL, FORMATIVO Y RETRIBUTIVO

Durante mi estancia en el grupo de espectroscopía tuve la oportunidad de conocer de manera más directa el funcionamiento de un grupo de trabajo académico. Aprendí a procesar datos utilizando herramientas y software de programación, y pude colaborar/participar en reuniones donde se discutían las distintas fases de los proyectos y los tópicos que aborda el grupo. El acercamiento que tuve con mi asesor, el doctor Michael Grutter, me permitió comprender mucho mejor cómo trabajar con estos datos, identificar patrones y solucionar problemas al tener retroalimentación directa a lo largo de mi servicio social, constantemente discutíamos los datos obtenidos y analizábamos cómo mejorar los procedimientos o implementar nuevos algoritmos que ayudaran en la interpretación de los resultados.

Esto fue fundamental, ya que no basta con aprender a manejar datos; es imprescindible saber interpretarlos y entender para qué pueden ser utilizados. Considero que el hecho de aprender a discriminar errores, identificar eventos y conocer cómo deben presentarse los datos que se quieren analizar, me permite seleccionar mejor la información que deseo obtener y definir claramente su propósito.

Participar en este grupo les dio mayor sentido a los años dedicados durante la carrera, ya que pude aplicar todos mis conocimientos y sentí una gran satisfacción al ser útil para el grupo, ayudando en el procesamiento de datos,

la obtención de nuevos productos y colaborando en nuevas ideas y mejoras. Me sentí arropado por todo el grupo de espectroscopía, ya que todos se comportaron de manera muy cálida y me ofrecieron su apoyo en todo momento. Asimismo, considero que la interacción con el doctor Michael Grutter fue excelente; la comunicación constante nos permitió resolver dudas rápidamente y estar al tanto de las actividades.

Estoy convencido de que estar en un grupo donde te sientes cómodo y en un ambiente favorable ayuda a desempeñar mejor tus actividades. La curva de aprendizaje que experimenté fue significativa, y considero que las actitudes y conocimientos adquiridos durante este servicio social serán de gran utilidad tanto en el ámbito profesional como académico. Además, me ayudó a perder el miedo de integrarme a un grupo, ya que muchas veces uno piensa que por la falta de experiencia o por ser reciente su formación no podrá aportar nada, pero esto es totalmente falso. Durante el servicio social descubrí que siempre se puede colaborar y contribuir en las actividades del grupo.

En resumen, esta experiencia me dejó un aprendizaje invaluable, tanto académico como profesional, y será de gran ayuda para mi trabajo de titulación, pues de toda la información recabada hay varios temas que me gustaría profundizar, como los datos de material particulado.

3. CONCLUSIONES

La creación de scripts para el procesamiento de datos me permitió involucrarme en técnicas avanzadas de análisis utilizando librerías del lenguaje de programación Python. Estas herramientas hacen posible trabajar con bases de datos que contienen series de tiempo con miles de mediciones, algo que sería prácticamente imposible de gestionar en hojas de cálculo como Excel, ya que el procesamiento sería mucho más lento y complicado.

Aplicar los conocimientos adquiridos durante la carrera sobre procesamiento de datos resultó fundamental. En primer lugar, pude extraer los datos de cada sensor y convertirlos a un formato más manejable. En segundo lugar, logré revisar y construir series de tiempo para obtener una primera aproximación a los datos en crudo, lo que permitió verificar su coherencia y calidad.

Posteriormente, empecé a implementar correcciones que permitieran obtener productos de mayor nivel de confianza, como los niveles L1 y L2, los cuales incluyen revisiones y validaciones en los datos. Los resultados se exportaron en archivos de texto o en hojas de cálculo formato .csv, lo que facilita el trabajo en distintas plataformas y la homogeneización de la información.

Esta homogeneización es especialmente útil porque permite cruzar y comparar datos entre diferentes productos. Por ejemplo, al trabajar con datos meteorológicos ya homogeneizados por fecha, es posible analizarlos junto con productos de gases o material particulado. Así, se logra un mejor análisis al

comparar variables como dirección y velocidad del viento con la cantidad de material particulado.

Por todo ello, una de las actividades más importantes fue homogeneizar los datos, trabajarlos en un formato sencillo y replicable, y dejar los scripts preparados para que otros puedan continuar con el procesamiento. Esta tarea es fundamental, ya que garantiza que los datos se puedan seguir utilizando y procesando, ya sea para trabajar con los productos ya terminados, revisarlos o incluso iniciar nuevos análisis desde los datos crudos.

4. Anexos

Links de Repositorios

<https://github.com/mgrutter/met: datos de meteorología>

<https://github.com/mgrutter/T640> PM: Análisis de datos de partículas finas

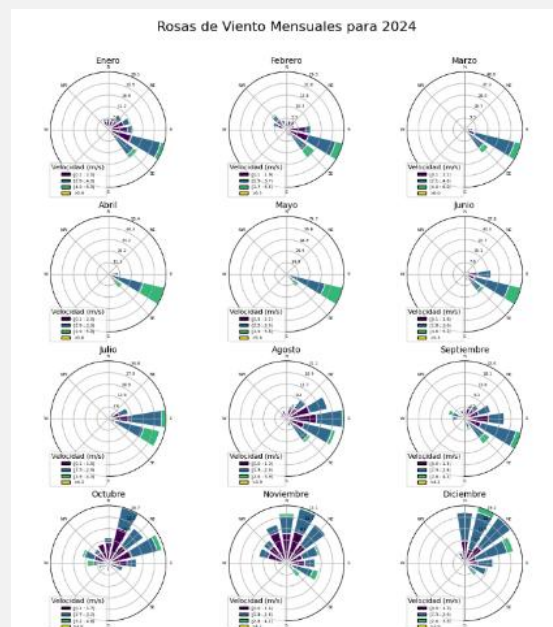
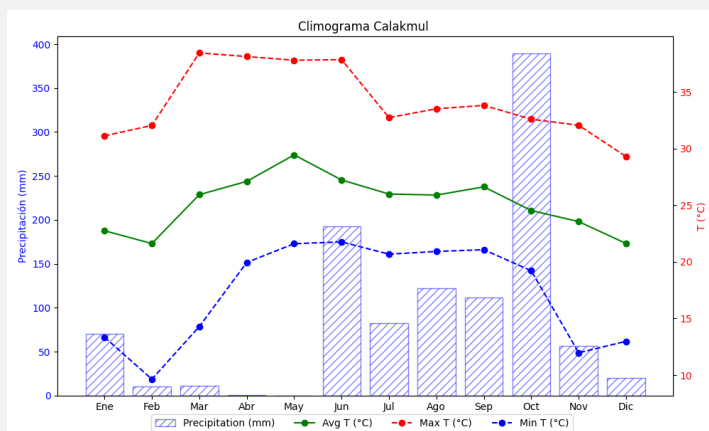
<https://github.com/mgrutter/pandora>: <https://www.pandonia-global-network.org/>

<https://github.com/Jonathanmn/scripts> ss/tree/33c49f704373c8e4e937bbbad92efa989cc9f84f/gei

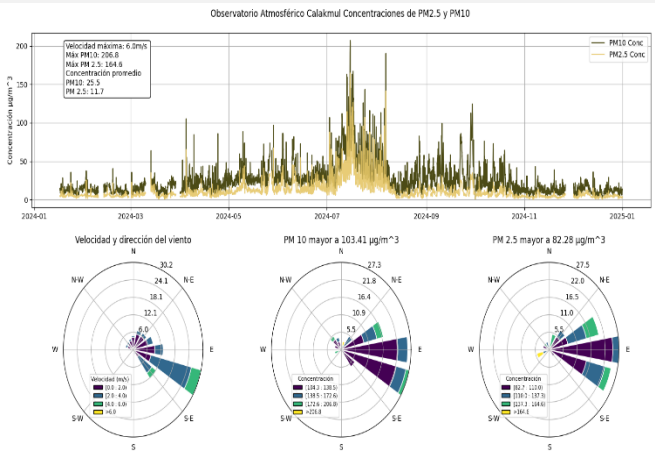
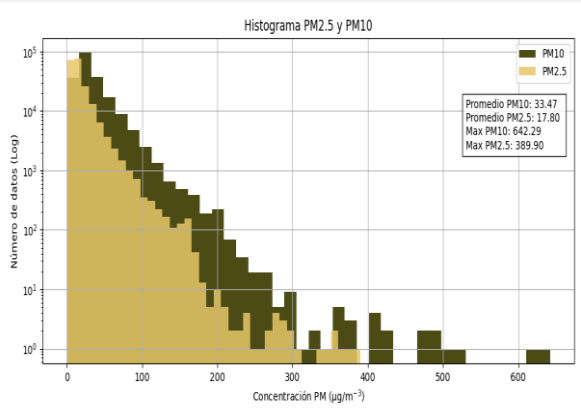
Página del grupo de EPR donde puedes descargar todos los productos de los sensores.

<http://www.epr.atmosfera.unam.mx/>

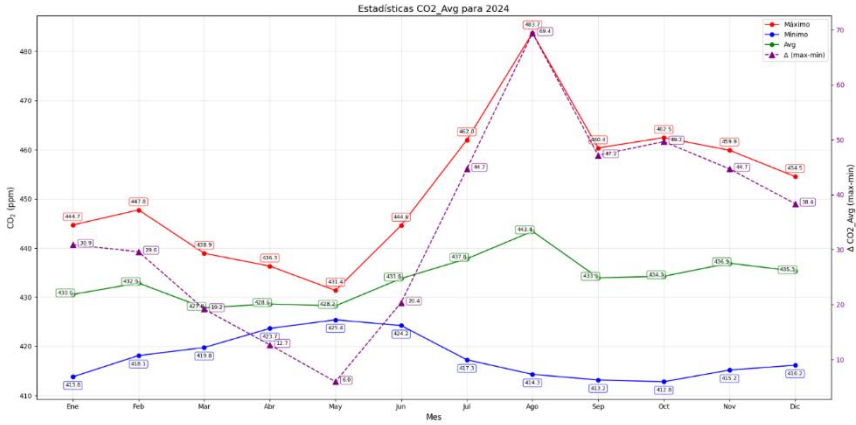
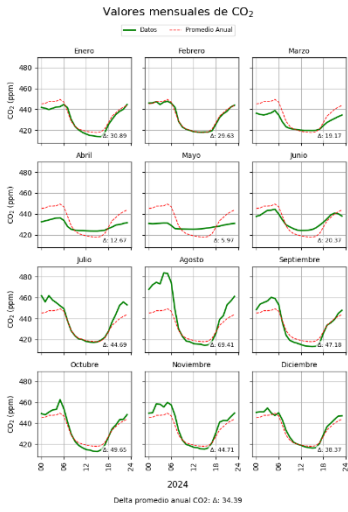
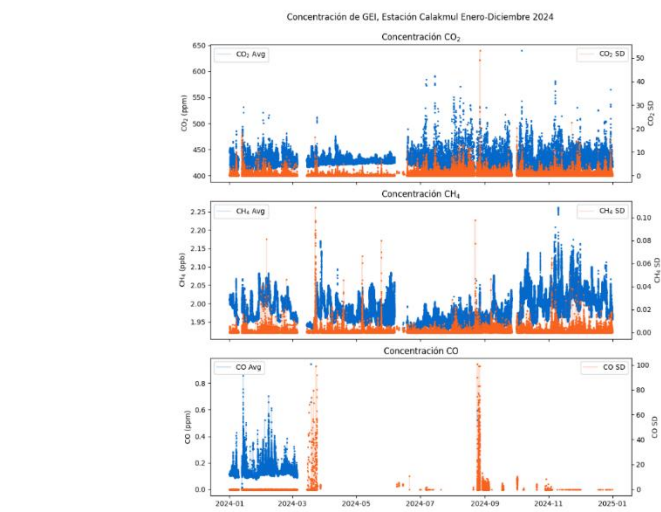
Productos meteorología



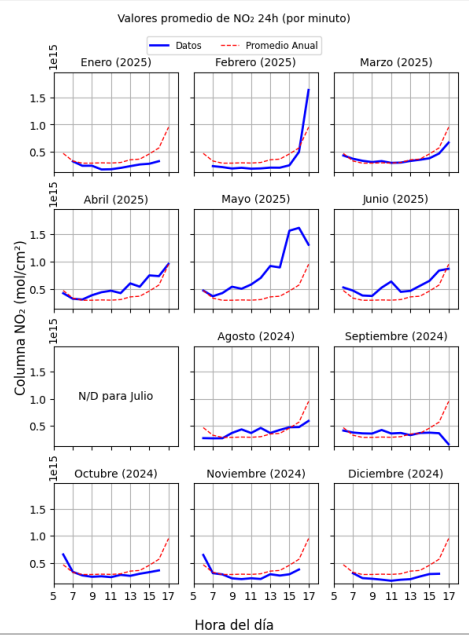
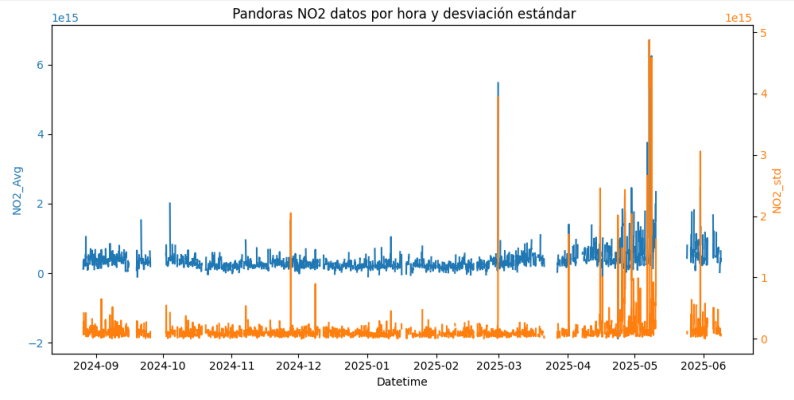
Productos PM T640



Productos Picarro



Productos Pandoras



Evidencia de campo

