



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE INGENIERÍA

**ANÁLISIS DEL REGISTRO DE LA INTENSIDAD DE
CAMPO ELÉCTRICO ATMOSFÉRICO DE LA RED
UNIVERSITARIA DE OBSERVATORIOS
ATMOSFÉRICOS**

TESIS

Que para obtener el título de

INGENIERO GEOFÍSICO

P R E S E N T A

FIDEL ALEJANDRO ARGUETA HERNÁNDEZ

DIRECTOR DE TESIS

DR. LUIS XAVIER GONZÁLEZ MÉNDEZ



Ciudad Universitaria, Cd. Mx., 2020



Universidad Nacional
Autónoma de México



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.

“...Con aquella desaparición, comprendió Harlan, mientras Noys se acercaba lentamente hacia sus brazos, había llegado el fin de la Eternidad... Y el comienzo del Infinito...”

Isaac Asimov en el Fin de la Eternidad

A mi padre y a mi hermano

Agradecimientos

Quiero agradecer a la Universidad Nacional Autónoma de México y a la Facultad de Ingeniería, por haberme dado tanto a cambio de nada; los mejores amigos, profesores y académicos que pude haber encontrado; por abrirme las puertas de México y el mundo, por permitirme estudiar en el campus universitario más bonito que mis ojos han visto, el cual será mi lugar favorito en el mundo por siempre.

Muchas gracias a mi mamá que me dio la vida y a mi papá, que me hizo convertirme en la persona que soy ahora.

Muchas gracias a mi tutor Luis Xavier, por siempre tener las puertas de su cubículo abiertas y por todos sus consejos y aportes a mi formación académica. A mis sinodales: Michel, Víctor, Ernesto y Julián por sus valiosas contribuciones y recomendaciones para esta tesis.

A las personas que contribuyeron también a esta tesis, que me sacaron de más de algún aprieto, como Eugenia Gonzalez, Eduardo Acuña e Iza Canales.

A Rodrigo e Isabel, por siempre incluirme con ellos y ser mi familia todos estos años.

A mis amigos de la generación de geofísica, que siempre me motivaron a ser un mejor estudiante y con los cuales compartí tantas experiencias que llevaré conmigo siempre.

A mis roomies, sin los cuales esta tesis hubiera estado terminada en la mitad de tiempo, pero que definitivamente aligeran las cargas que se llevan en esta etapa de la vida.

A todos, muchas gracias.

...

ÍNDICE GENERAL

Agradecimientos	iii
Lista de Figuras	vi
Introducción	ix
1 Campos Eléctricos Atmosféricos	1
1.1 Introducción	1
1.2 Electricidad atmosférica	2
1.2.1 Circuito eléctrico atmosférico global	3
1.2.2 La curva de Carnegie	3
1.3 Formación de las tormentas eléctricas	5
1.3.1 Fase de desarrollo	5
1.3.2 Fase de madurez	6
1.3.3 Fase de disipación	6
1.3.4 Características eléctricas de las nubes de tormenta	7
2 Infraestructura	9
2.1 Introducción	9
2.2 La Red Universitaria de Observatorios Atmosféricos (RUOA)	9
2.2.1 Líneas temáticas	10
2.2.2 Estaciones	11
Observatorio atmosférico de Ciudad Universitaria, CDMX (UNAM)	12
Observatorio atmosférico de Campus UNAM en Juriquilla, Querétaro (JRQO)	13
Observatorio atmosférico de Campus UNAM en Hermosillo, Sonora (ERNO)	14
2.3 El detector de campos eléctricos Boltek EFM-100	15
2.3.1 Teoría de operación	15
2.3.2 Software del equipo	17
2.3.3 Interpretación de las gráficas	19
2.4 Instalación de un Boltek EFM-100 en Sierra Negra, Puebla	22
3 Base de datos	26
3.1 Introducción	26
3.2 Lenguaje de programación Python	26

3.2.1	Características de Python	27
3.3	Transformada de Fourier	27
3.4	Estado de la base de datos	28
4	Análisis de datos y resultados	31
4.1	Introducción	31
4.2	Niveles de calidad L0 y L1 aplicados a la base de datos GEOF	31
4.3	Correlación entre las bases de datos UNAM Y GEOF	35
4.4	Espectros de amplitudes	36
4.4.1	Espectro de amplitudes del CCA	36
4.4.2	Espectro de amplitudes del Instituto de Geofísica	39
4.4.3	Espectro de amplitudes de Juriquilla	41
4.4.4	Espectro de amplitudes de Hermosillo	46
4.5	Análisis L2	49
	Conclusiones	54
	Apéndice A: Códigos de Python	56
	Referencias	72

Lista de Figuras

Figure 1.1:	Experimento de Faraday, (Tomé, 2016)	1
Figure 1.2:	Representación esquemática del Circuito Eléctrico Atmosférico Global. Kelley and Holzworth, 2014	3
Figure 1.3:	Gráfica de la cantidad de área por continente bajo influencia de tormentas, donde la línea en negritas corresponde al promedio mundial y las línea punteada superior corresponde a la curva de Carnegie. (Harrison, 2013)	4
Figure 1.4:	Fases de la duración de una nube cumulonimbus, tipo de nube en la que se suele presentar actividad eléctrica: a) fase de desarrollo, cuando hay solo ascenso de aire caliente; b) fase madura, cuando coexiste ascenso y descenso; c) fase de disipación, cuando ya solo hay descenso de aire (Soula, 2012)	6
Figure 1.5:	Densidad de caída de rayos para diciembre de 2011. (Kelley and Holzworth, 2014)	8
Figure 1.6:	Densidad de caída de rayos para junio de 2011. (Kelley and Holzworth, 2014)	8
Figure 2.1:	Boltek EFM-100 instalado en el Instituto de Geofísica, desde marzo de 2017 a junio de 2019	10
Figure 2.2:	Página principal de la RUOA	11
Figure 2.3:	Estación en el Centro de Ciencias de la Atmósfera, donde se encuentra uno de nuestros monitores de campo eléctrico	12
Figure 2.4:	Estación en Juriquilla, Querétaro, donde se encuentra uno de nuestros monitores de campo eléctrico.	13
Figure 2.5:	Estación en Hermosillo, Sonora, donde se encuentra uno de nuestros monitores de campo eléctrico	14
Figure 2.6:	Diagrama de conexión del EFM-100, usando fibra óptica. Modificado de <i>EFM-100 Atmospheric Electric Field Monitor. Installation/Operators Guide for model EFM-100C 2016</i>	16
Figure 2.7:	Software que despliega la gráfica de intensidad del campo eléctrico en tiempo real, donde en la esquina inferior derecha se puede observar la lectura del campo en kV/m cada segundo	18
Figure 2.8:	Nube de tormenta aproximándose.	19
Figure 2.9:	Incremento del campo debido a la aproximación de la nube de tormenta	19
Figure 2.10:	Nube de tormenta justo sobre el monitor	20

Figure 2.11:	Inversión de la polaridad del campo eléctrico	20
Figure 2.12:	Nube de tormenta alejándose	21
Figure 2.13:	Oscilación del final de tormenta	21
Figure 2.14:	Salto en la intensidad del campo que indican rayos	22
Figure 2.15:	Cómo se observa el ruido por precipitación	22
Figure 2.16:	Boltek EFM-100 instalado en Sierra Negra	23
Figure 2.17:	Cima de Sierra Negra, donde quedó instalado el detector de campo eléctrico Boltek EFM-100, a un lado del GTM	24
Figure 2.18:	Computadora que recibe los datos del Boltek-EFM100	25
Figure 2.19:	Episodio de tormenta registrado en la cima de Sierra Negra el 11 de octubre del 2019	25
Figure 4.1:	Estructura que poseen los datos crudos de campo eléctrico atmosférico de los equipos de la RUOA	32
Figure 4.2:	a) Datos crudos del Boltek del Instituto de Geofísica, b) Datos transformados al formato crudo de la RUOA	33
Figure 4.3:	Código añadido al script master.py para que L0 y L1 pudieran correr	34
Figure 4.4:	Correlación entre los equipos del CCA y del Instituto de Geofísica para el día 15 de abril del 2018	36
Figure 4.5:	Campo eléctrico atmosférico para el año 2017 en el CCA	37
Figure 4.6:	Espectro de amplitudes para el CCA, mostrando una frecuencia dominante	38
Figure 4.7:	Campo eléctrico atmosférico diario promedio, en el año 2017 para el CCA	38
Figure 4.8:	Varianza por hora de los promedios diarios del campo eléctrico atmosférico, en el año 2017 para el CCA	39
Figure 4.9:	Campo eléctrico atmosférico para el año 2018 en el Instituto de Geofísica	40
Figure 4.10:	Espectro de amplitudes para el Instituto de Geofísica, mostrando una frecuencia dominante	41
Figure 4.11:	Campo eléctrico atmosférico diario promedio en el año 2018, para el Instituto de Geofísica	42
Figure 4.12:	Varianza por hora de los promedios diarios del campo eléctrico atmosférico, en el año 2018 para el Instituto de Geofísica	43
Figure 4.14:	Espectro de amplitudes para Juriquilla, mostrando dos frecuencias dominantes	43
Figure 4.13:	Campo eléctrico atmosférico para el año 2017 en Juriquilla	44
Figure 4.15:	Campo eléctrico atmosférico diario promedio en el año 2017, para Juriquilla	45
Figure 4.16:	Varianza por hora de los promedios diarios del campo eléctrico atmosférico, en el año 2017 para Juriquilla	45
Figure 4.17:	Campo eléctrico atmosférico para el año 2016 en Hermosillo	46
Figure 4.18:	Espectro de amplitudes para Hermosillo, mostrando dos frecuencias dominantes	47

Figure 4.19:	Campo eléctrico atmosférico diario promedio en el año 2016, para Hermosillo	47
Figure 4.20:	Varianza por hora de los promedios diarios del campo eléctrico atmosférico, en el año 2016 para Hermosillo	48
Figure 4.21:	Estructura de la hoja de cálculo realizada a partir de la inspección visual de los datos	50
Figure 4.22:	Ejemplo de cómo se aprecian los comentarios por día en la hoja de cálculo	51
Figure 4.23:	Gráfica de calidad para la estación del CCA	51
Figure 4.24:	Gráfica de calidad para la estación del Instituto de Geofísica	52
Figure 4.25:	Gráfica de calidad para la estación de Juriquilla	52
Figure 4.26:	Gráfica de calidad para la estación de Hermosillo	53

INTRODUCCIÓN

Actualmente el estudio de los campos eléctricos atmosféricos presentan varias áreas de oportunidad para la ciencia, porque es una variable meteorológica adicional que se puede ocupar para analizar fenómenos como las precipitaciones y las tormentas eléctricas. Además, se pueden utilizar para complementar estudios en otras áreas de la ciencia como el estudio de las afectaciones de los campos eléctricos atmosféricos en los rayos cósmicos secundarios detectados a nivel de tierra. También, el campo eléctrico se puede considerar como una fuente de ruido en las comunicaciones, como las que se establecen para GPS o se puede utilizar en la medición de riesgos para actividades en exteriores. Hemos observado en días recientes un acontecimiento tan importante como el primer despegue de un cohete de SpaceX con tripulantes, suspendido debido a fuertes campos eléctricos en la atmósfera circundante. De hecho, toda la instrumentación para cualquier experimento que esté expuesta a una tormenta eléctrica debe considerar relevadores, tierras físicas o diferentes equipos que puedan protegerlos de las descargas eléctricas.

En esta tesis se realiza un análisis de la base de datos de los detectores de campos eléctricos (Boltek EFM-100) de la Red Universitaria de Observatorios Atmosféricos (RUOA). Esto surge de la necesidad de conocer su estabilidad estadística y de determinar su viabilidad técnica para estudios científicos.

Los campos eléctricos atmosféricos se originan en las nubes de tormenta, por lo que en el capítulo 1 comenzamos repasando estos orígenes, además de como se forman las tormentas eléctricas. En el capítulo 2 exploramos en qué consiste la Red Universitaria de Observatorios Atmosféricos (RUOA), la cual es una plataforma de registro y difusión de datos meteorológicos de todo el territorio mexicano. Es la base de datos de intensidad de campo eléctrico atmosférico de 3 estaciones de la RUOA la que se utiliza en esta tesis, junto con una cuarta base de datos adicional registrada en el Instituto de Geofísica, todas estas grabadas con los equipos Boltek EFM-100 cuyo funcionamiento se describe detalladamente también el capítulo 2. Posteriormente en el capítulo 3 se presenta el lenguaje de programación Python, usado para manipular y tratar los datos, así como también mencionamos a la Transformada de Fourier. Además, hicimos una descripción del estado en el que se encontraba la base de datos. Finalmente, en el capítulo 4 presentamos los resultados y productos obtenidos, los cuales son:

1. El desarrollo y corrección de scripts de Python para aplicar los dos niveles de control de calidad (L0 y L1) a la base de datos exterior a la RUOA, que ya se realizan de forma continua sobre los registros de la RUOA y así poder normalizarla con los datos publicados

en el sitio web, y con cuya experiencia obtenida se podrán procesar los datos que se obtengan de nuevas estaciones.

2. Presentamos la correlación de dos días muestra, entre los datos que registraron las estaciones del Centro de Ciencias de la Atmósfera y el Instituto de Geofísica, ubicadas de forma contigua y comprobando al observar los datos, que los equipos tienen registros similares, lo que nos proporciona más confianza en estos datos.

3. Realizamos una transformada de Fourier sobre un año muestra para cada una de las 4 estaciones, para obtener sus espectros de frecuencia y confirmar la curva de Carnegie (de 24 horas) y otros ciclos de mayor frecuencia. Además obtuvimos gráficas de promedios diurnos y de varianza.

4. Para obtener un nivel de calidad L2, efectuamos una inspección visual de la base de datos completa, donde se generaron hojas de cálculo con un código de color según la calidad de los datos para cada día y se añadieron comentarios de los registros para la mayoría de los días. Cabe señalar que todos los scripts y productos de esta tesis se encuentran en un repositorio de Github cuya dirección de encuentra en la introducción del capítulo 4.

CAMPOS ELÉCTRICOS ATMOSFÉRICOS

1.1 Introducción

Desde tiempos remotos, a la humanidad le han sorprendido los fenómenos eléctricos y magnéticos, los cuales siempre se observaron como dos fenómenos separados, hasta que llegó el ingenio del físico experimental Michael Faraday en 1831. Faraday propuso su célebre experimento, donde descubrió que una corriente en un cable puede inducir una corriente en otro cable, mientras la corriente tenga variaciones. A este fenómeno se le denominó inducción magnética (Ver figura 1.1).

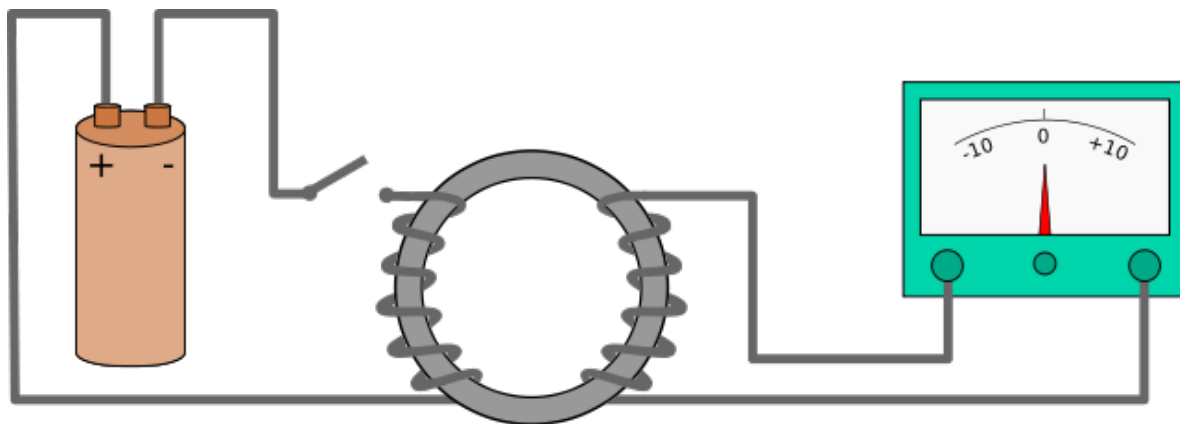


FIGURE 1.1: Experimento de Faraday, (Tomé, 2016)

Faraday en 1831 se encontraba experimentando con dos trozos de alambre que había enrollado en una pieza de hierro en forma de anillo. Se dio cuenta de que aparecía una corriente en uno de los trozos, llamado después el “secundario”, cuando se establecía una corriente en el otro trozo, el “primario”. Pero el fenómeno aparecía solo en el momento en el que se establecía la corriente, duraba solo un instante. Tan pronto como había una

corriente continua estable en el primario, la corriente en el secundario desaparecía. Cuando desaparecía la corriente en el primario, por un instante, volvía a aparecer la corriente en el secundario. Reflejó su hallazgo en su diario de laboratorio el 29 de agosto de 1831.

El descubrimiento de Faraday se puede describir en pocas palabras como que una corriente en un cable puede inducir una corriente en otro cable solo mientras la corriente está cambiando. Una corriente continua en un cable no puede inducir una corriente en otro cable.

En la actualidad, la generación de electricidad en el mundo se da con base en la Inducción Electromagnética. La ecuación que describe este fenómeno es una de las cuatro ecuaciones de Maxwell, conocida como Ley de Faraday. Esta ecuación en su forma diferencial, en el Sistema Internacional y aplicada para el vacío se ve de la siguiente forma:

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (1.1)$$

Donde:

\mathbf{B} es el campo magnético

\mathbf{E} es el campo eléctrico

Y es precisamente la ecuación 1.1 (que se puede leer como la divergencia del campo eléctrico es igual a menos la derivada del campo magnético con respecto al tiempo), la que describe cómo un campo magnético variable en el tiempo nos produce una corriente eléctrica.

1.2 Electricidad atmosférica

Por debajo de los 80 km, la atmósfera deja de ser plasma porque el movimiento de las partículas cargadas está controlado más por un arrastre neutro que por fuerzas electromagnéticas de gran escala. Sin embargo, la atmósfera es un débil conductor eléctrico con una gran variedad de fenómenos electromagnéticos ocurriendo a su alrededor. Uno de estos fenómenos se le conoce como el campo eléctrico de buen tiempo.

En 1911 Víctor Hess condujo un experimento para estudiar iones radiactivos en el aire. Usando un globo aerostático, encontró que la conductividad incrementaba con la altitud en vez de disminuir, llevándolo a descubrir la radiación cósmica. Con base en el estudio de estos experimentos, Wilson (Wilson, 1920) fue el primero en proponer que la Tierra era un conductor y que las tormentas eléctricas alrededor del mundo debían actuar poniendo carga negativa en el suelo y positiva en la atmósfera superior, funcionando como un capacitor. Así fue como nació el concepto de circuito eléctrico global.

1.2.1 Circuito eléctrico atmosférico global

La idea de un circuito eléctrico global (figura 1.2) fue rápidamente aceptada por la comunidad científica luego que Mauchly (Mauchly, 1923) usara datos de campo eléctrico del barco de investigación del Instituto Carnegie, que mostraban que el campo eléctrico de buen tiempo parecía tener un pico alrededor de las 19:00 horas tiempo universal (TU), independiente de donde fueran tomadas las mediciones.

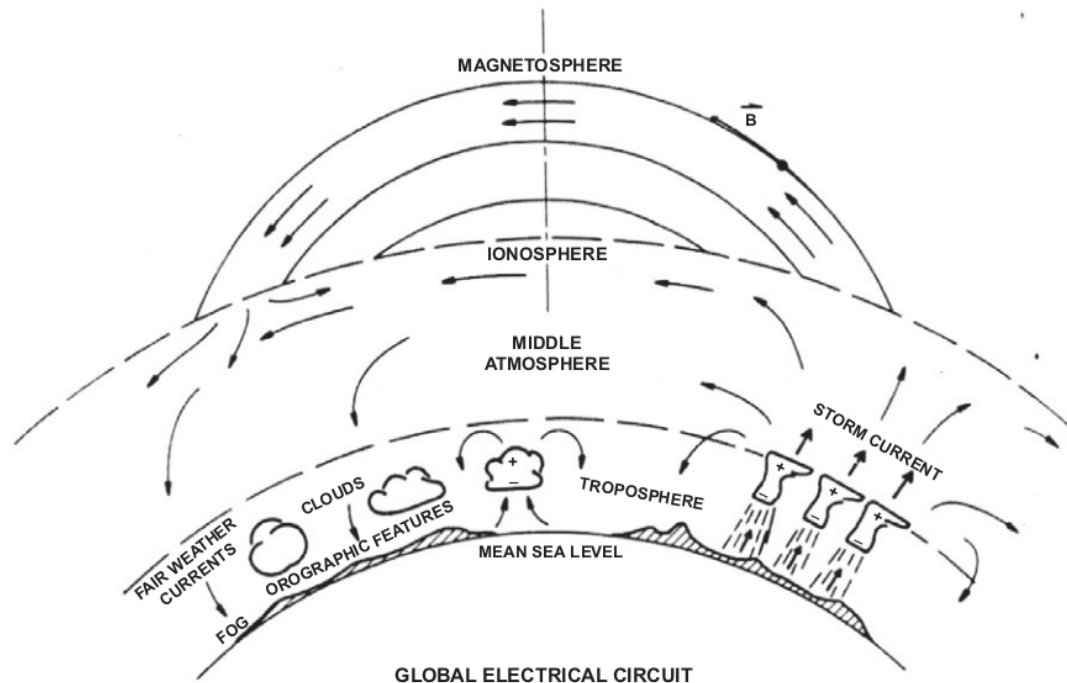


FIGURE 1.2: Representación esquemática del Circuito Eléctrico Atmosférico Global. Kelley and Holzworth, 2014

1.2.2 La curva de Carnegie

A principios del siglo XX, con base en una embarcación del instituto Carnegie, se desarrollaron experimentos que demostraron la existencia de una variación diurna en el campo eléctrico atmosférico, independiente del sitio donde se tomaran los datos, con un máximo alrededor de las 19:00 horas TU y un mínimo alrededor de las 04:00 horas TU (Harrison, 2013). Esta variación lleva por nombre curva de Carnegie.

La curva de Carnegie, nombrada en honor al Instituto Carnegie, es una de las oscilaciones fundamentales en electricidad atmosférica. Esta variación también cambia un poco con las estaciones. La hora del mínimo de la curva de Carnegie varía entre los equinoccios, mientras que la variación en el tiempo de los máximos, que es más grande, ocurre aproximadamente entre los solsticios. Una de las principales explicaciones para esta curva

es porque es alrededor de las 19:00 horas TU que se encuentra mayor área continental bajo influencia de tormentas eléctricas y se genera mayor actividad eléctrica atmosférica.

En la figura 1.3, se observa la comparación entre el área continental con tormentas (y sus contribuciones regionales) y la curva de Carnegie con valores relativos de promedios anuales (Harrison, 2013). En ella podemos observar que en promedio es alrededor de las 19:00 TU que se encuentra mayor área continental bajo tormentas eléctricas.

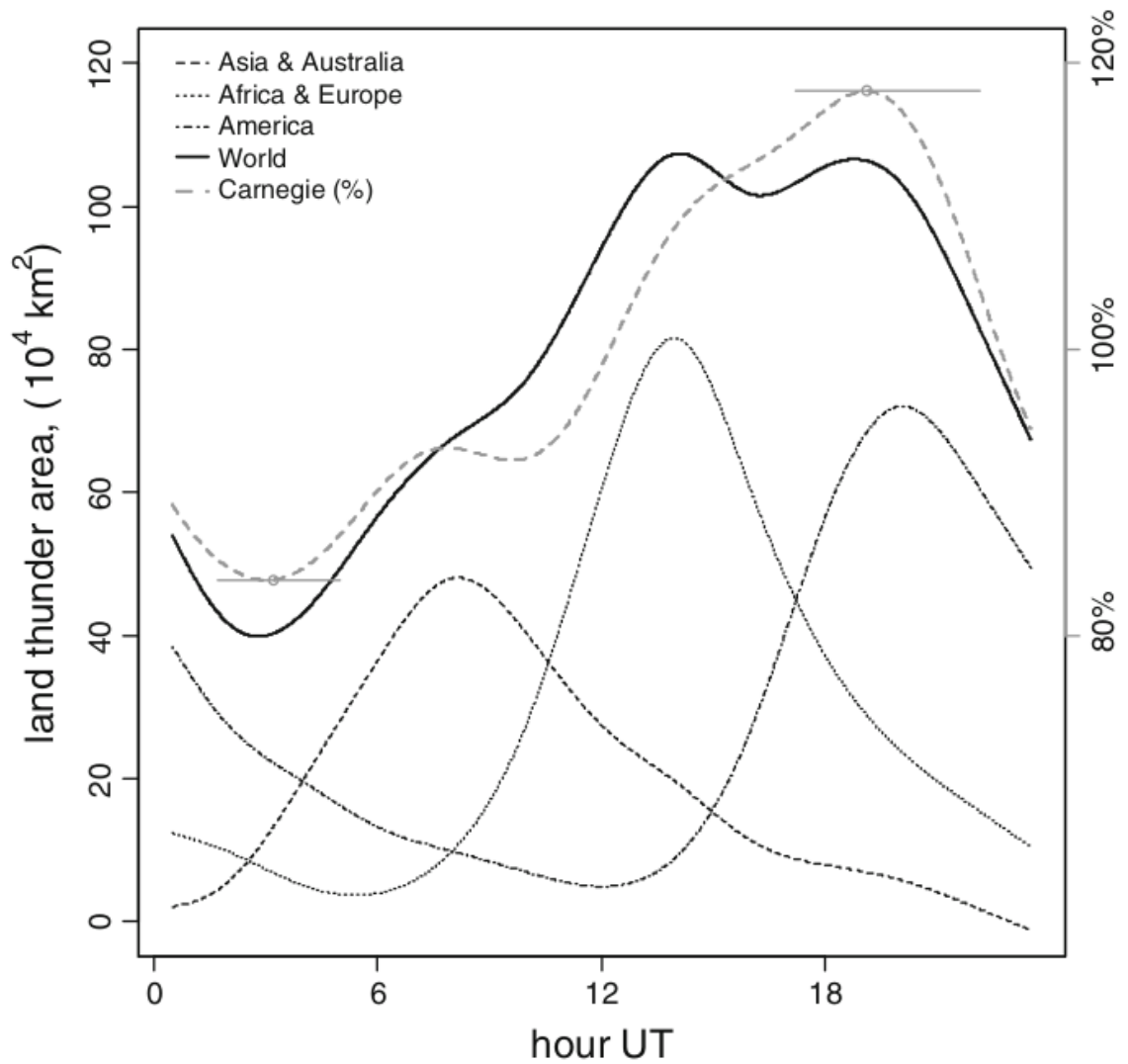


FIGURE 1.3: Gráfica de la cantidad de área por continente bajo influencia de tormentas, donde la línea en negritas corresponde al promedio mundial y la línea punteada superior corresponde a la curva de Carnegie. (Harrison, 2013)

1.3 Formación de las tormentas eléctricas

Para poder entender cómo se originan los campos eléctricos en la atmósfera debemos comprender cómo se generan las tormentas eléctricas, ya que se encuentran completamente relacionadas. Las nubes de tormenta son el resultado de la convección atmosférica. La convección puede iniciar cuando existe inestabilidad condicional. Para describir la inestabilidad condicional se usa la teoría de parcelas (gran número de moléculas que se mueven en forma más o menos coherente empleando un volumen de varios metros cúbicos). Cuando una parcela de aire se mueve en dirección vertical hacia arriba, lo hace bajo un proceso adiabático, lo que significa que no hay intercambio de energía entre la parcela de aire y el aire circundante, lo cual reduce su temperatura a una tasa cercana a $10\text{ }^{\circ}\text{C}$ cada kilómetro. Si la parcela resulta ser menos fría que el aire circundante a su nueva altitud, puede continuar ascendiendo por una fuerza de flotabilidad. Además, si la parcela de aire se satura en su ascenso, su temperatura decrece más lentamente debido al calor latente de condensación que es liberado. Bajo estas condiciones, el movimiento vertical del aire saturado es conocido como 'convección por humedad' y sucede a larga escala dentro de la nube de tormenta. Así, las probabilidades para convección por humedad son determinadas por la cantidad de humedad y alta temperatura en los primeros kilómetros de la atmósfera, junto con un fuerte decrecimiento de la temperatura con la altura (aire frío arriba) en la capa de 2 a 5 kilómetros. Adicionalmente, si existen circulaciones atmosféricas, estas pueden enfriar adiabáticamente (por levantamiento) capas de medio nivel y forzar a las parcelas de bajos niveles a ascender, así estas pueden volverse más cálidas que el aire circundante y continuarán ascendiendo (después de haber alcanzado el nivel de convección libre), con lo cual tenemos la formación de tormentas eléctricas (Soula, 2012).

1.3.1 Fase de desarrollo

En la fase de desarrollo (figura 1.4 a) se forman torres de nube debido a corrientes de aire ascendentes. La intensidad de la tormenta será determinada por la fuerza ejercida por estas corrientes, cuyas altas velocidades propician que exista una tasa de condensación elevada. Al elevarse, el aire cálido y húmedo se comienza a enfriar y el vapor de agua que contiene se condensa. En caso de que haya algún núcleo, es decir, una partícula sólida alrededor de la cual el agua se pueda cristalizar, se formarán copos de nieve.

Los bordes de la nube durante esta fase son definidos y distinguibles, indicando que la nube está compuesta principalmente de gotas de agua. La nube convectiva continúa creciendo verticalmente, hasta sobrepasar la altura de congelación donde la temperatura es de $0\text{ }^{\circ}\text{C}$. Mientras que algunas gotas de agua se congelan, otras moléculas de agua son depositadas sobre núcleos de condensación, formando de esta manera cristales de hielo, (Newton-Bosch, 2018).

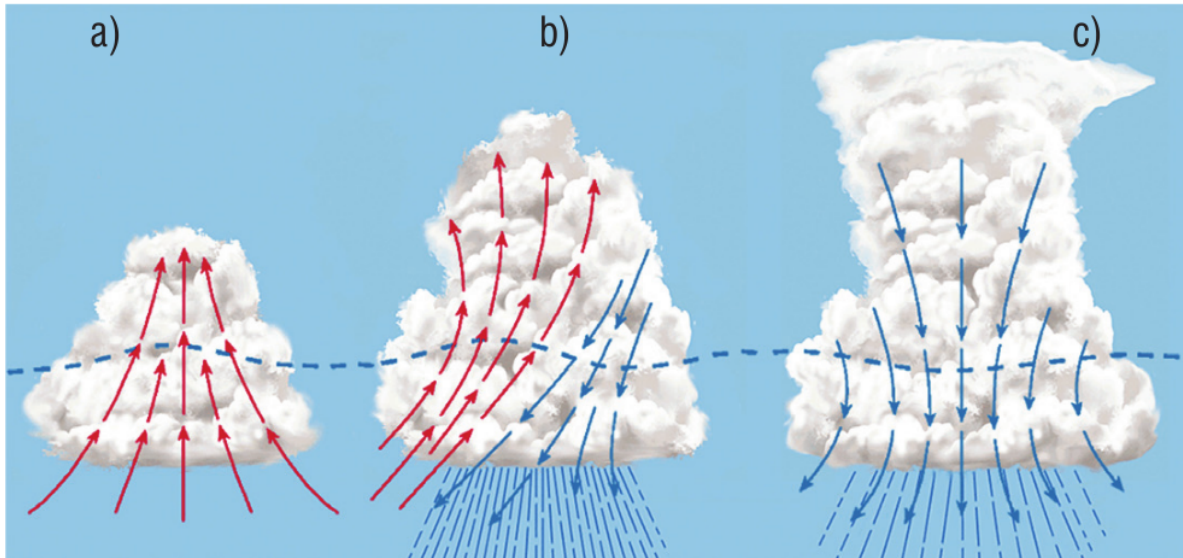


FIGURE 1.4: Fases de la duración de una nube cumulonimbus, tipo de nube en la que se suele presentar actividad eléctrica: a) fase de desarrollo, cuando hay solo ascenso de aire caliente; b) fase madura, cuando coexiste ascenso y descenso; c) fase de disipación, cuando ya solo hay descenso de aire (Soula, 2012)

1.3.2 Fase de madurez

Durante esta fase (figura 1.4 b), los copos de nieve formados en la fase de desarrollo seguirán siendo arrastrados por las corrientes de aire ascendentes. Es aquí donde la tormenta se considera madura. Los cristales en esta zona rápidamente se aglomerarán en grandes partículas de hielo. Una vez que esto sucede, las corrientes de aire ascendentes no serán capaces de soportar el peso del hielo, y éste caerá, arrastrando el aire a su alrededor hacia abajo. Es entonces que las corrientes de aire se vuelven descendentes y comienza la precipitación. La lluvia se genera por la fundición de algunas partículas de hielo, debida al incremento de temperatura que experimentan al caer. Las descargas eléctricas entre la nube y la tierra, comienzan a suceder en este momento. La nube cumulonimbus alcanzará su máxima altura, en la base de la estratosfera, donde la corriente ascendente se volverá inestable, extendiéndose lateralmente y formando así una forma característica de yunque.

1.3.3 Fase de disipación

Durante esta fase (figura 1.4 c) la corriente descendente principal se ve fortalecida por la penetración de aire frío desde la parte superior de la nube. La temperatura de la nube disminuye debido a este proceso, aunado con la gran cantidad de partículas de hielo y nieve que contiene. Esta disminución en la temperatura provoca que las gotas de agua que previamente alcanzaron un sobreenfriamiento sin congelarse, por fin se congelen y liberen de esa manera calor latente. Debido a eso, la corriente descendente se debilita y la nube comienza a disiparse desde la parte superior. La disipación de la cumulonimbus también puede ocurrir

en el caso de que la precipitación sea tan intensa que la corriente de aire ascendente no sea capaz de soportarla y por ende cese, provocando la evaporación de la nube.

1.3.4 Características eléctricas de las nubes de tormenta

Una de las preguntas remanentes por mayor tiempo ha sido cómo las nubes convectivas se electrifican para producir rayos. Para responder a esta pregunta, los científicos han desarrollado experimentos de laboratorio y observaciones de campo, pero es difícil obtener explicaciones completas y definitivas, debido al rango de distancia que hay entre la microescala de los procesos físicos que conciernen a los hidrometeoros (partículas de agua) de las nubes y el tamaño de la nube de tormenta para la estructura de carga. Igualmente, los múltiples procesos que toman parte en el fenómeno de carga y descarga dentro de la nube de tormenta cubren un gran rango de tiempo. Además, es difícil tomar medidas in situ, porque las condiciones de tormenta son hostiles a la instrumentación.

De todos los mecanismos de carga propuestos durante el siglo pasado, el que es considerado como el más capaz de desencadenar el proceso de electrificación dentro de la nube, reproducir las capas de carga en la vertical y la cantidad de carga observada, involucra colisiones por rebote entre los cristales de hielo y los granos de hielo. Los granos de hielo pueden ser descritos como pequeños granizos porosos y crecen como pequeñas gotículas superenfriadas hacia su superficie. Este mecanismo es no inductivo, es decir, este no necesita un campo eléctrico externo para crear carga en una partícula. Para estudiar este tipo de mecanismo, los científicos usan experimentos de laboratorio para analizar empíricamente los diferentes aspectos de las propiedades eléctricas de las partículas. Takashi mostró que de acuerdo a la temperatura ambiente y al contenido de agua líquida, los bordes de las partículas se cargan positivamente (para más altas temperaturas y para muy alto o bajo contenido de agua en la nube) o negativamente (para más bajas temperaturas y para un contenido de agua promedio en la nube), (Takashi, 1978).

En las siguientes figuras se graficó la densidad de rayos para junio y diciembre de 2011 a partir de los datos del World Wide Lightning Location Network (WWLLN), mostrando un agudo contraste entre los hemisferios norte y sur a través del año. En diciembre (Figura 1.5) los rayos sobre los continentes bajo el ecuador se ven muy marcados en América del sur, África, Indonesia y Australia, mientras que la localización de los rayos para junio es opuesta (figura 1.6), dándose la mayor actividad en el hemisferio norte, donde vemos una fuerte intensidad en México, que es nuestra área de interés y que será comprobado en los resultados en el capítulo 4

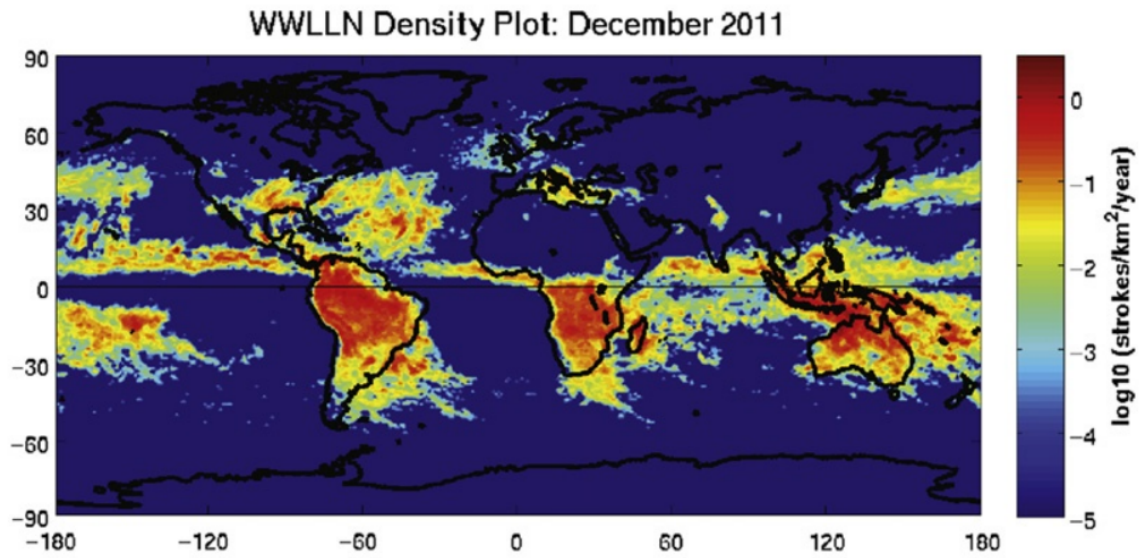


FIGURE 1.5: Densidad de caída de rayos para diciembre de 2011. (Kelley and Holzworth, 2014)

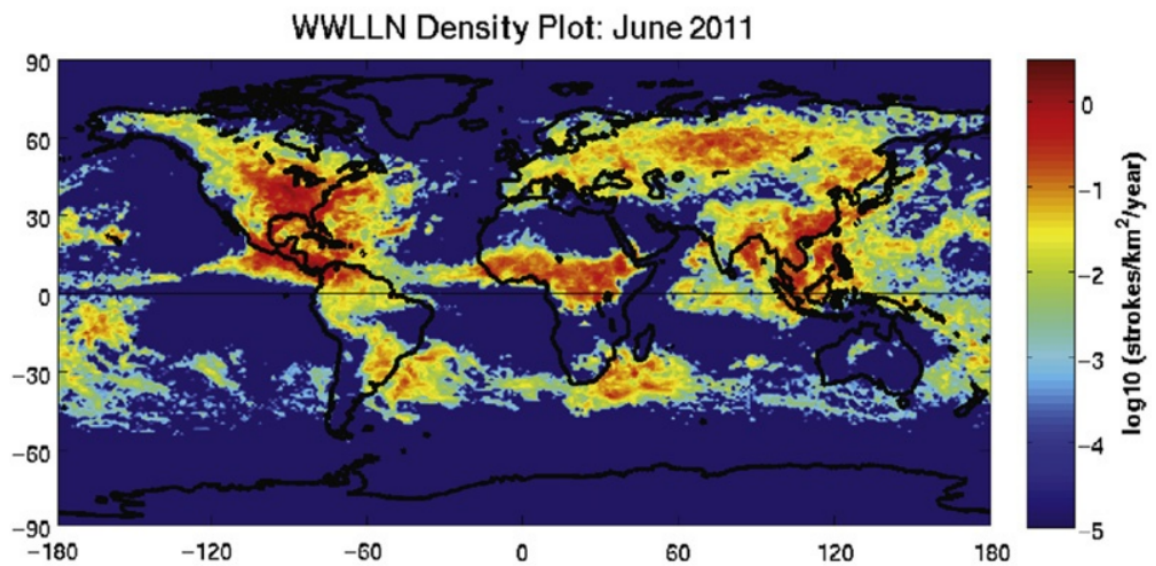


FIGURE 1.6: Densidad de caída de rayos para junio de 2011. (Kelley and Holzworth, 2014)

2

RED UNIVERSITARIA DE OBSERVATORIOS ATMOSFÉRICOS (RUOA)

2.1 Introducción

En este capítulo se presenta la Red Universitaria de Observatorios Atmosféricos (RUOA), sus líneas de investigación y su red de estaciones que cuentan con el registro de campo eléctrico atmosférico.

Además, veremos cómo funciona el monitor de campos eléctricos Boltek EFM-100 fabricado por la compañía canadiense Boltek. Este es el equipo que usa la RUOA y con el que contábamos en el Instituto de Geofísica para la base de datos de campo eléctrico (figura 2.1). Además, presentamos la instalación que llevamos a cabo de uno de estos equipos en Sierra Negra, Puebla, a 4,590 m.s.n.m. Posiblemente este es el Boltek-EFM 100 instalado a mayor altura en el mundo y sus datos serán clave para estudiar el comportamiento del campo eléctrico atmosférico a grandes alturas.

De este modo, se pueden llevar a cabo correlaciones con otras mediciones que se registran en esta alta montaña, como son las componentes secundarias de los rayos cósmicos, por lo que el tratamiento de los datos en esta tesis puede llegar a ser esencial para los estudios que se puedan hacer en la cima de Sierra Negra.

2.2 La Red Universitaria de Observatorios Atmosféricos (RUOA)

La Red Universitaria de Observatorios Atmosféricos (RUOA) es un proyecto de adquisición y difusión de datos atmosféricos y ambientales, utilizando una red de estaciones albergadas en instituciones académicas de todo México.



FIGURE 2.1: Boltek EFM-100 instalado en el Instituto de Geofísica, desde marzo de 2017 a junio de 2019

A través de una amplia cooperación entre las diversas instituciones participantes, se pone a la disposición de la comunidad científica y el público en general, esta información de consulta en tiempo real y para su libre descarga.

Los diversos parámetros y observables pueden ser empleados tanto para fines educativos como para llevar a cabo investigaciones en meteorología, contaminación, cambio climático y disciplinas afines.

2.2.1 Líneas temáticas

La RUOA maneja diferentes líneas temáticas de investigación con los datos que se adquieren con la red y estas son:

- Meteorología
- Gases de efecto invernadero
- Calidad del aire

- Caracterización de la precipitación
- Aerosoles
- Carbono negro
- Electricidad Atmosférica
- Química del depósito atmosférico
- Columna de agua precipitable
- Compuestos orgánicos volátiles
- Aerobiología



FIGURE 2.2: Página principal de la RUOA

2.2.2 Estaciones

La RUOA cuenta con estaciones distribuidas por todo el país en diferentes instalaciones universitarias, y es su base de datos de campo eléctrico atmosférico la que ha sido utilizada en esta tesis. También, muchas de las variables pueden consultarse con gráficas en tiempo real. Además, cabe mencionar que las variables de meteorología que se estarán mencionando a continuación incluyen temperatura, rapidez de viento, radiación, humedad, dirección de viento, visibilidad, precipitación total y presión atmosférica.

Observatorio atmosférico de Ciudad Universitaria, CDMX (UNAM)

Esta estación se encuentra en el Centro de Ciencias de la Atmósfera de la UNAM, en Ciudad Universitaria, en las coordenadas 19.3262 °N, 99.1761 °W, a una altitud de 2,280 m.s.n.m.

En esta estación se registran las variables de:

- Meteorología
- perfiles de viento
- Gases de efecto invernadero
- Calidad del aire
- Carbón negro
- Características de la precipitación
- Campo eléctrico
- Depositación húmeda
- Agua total precipitabe
- Perfil de retrodispersión



FIGURE 2.3: Estación en el Centro de Ciencias de la Atmósfera, donde se encuentra uno de nuestros monitores de campo eléctrico

Observatorio atmosférico de Campus UNAM en Juriquilla, Querétaro (JRQO)

Esta estación se encuentra en la Unidad Multidisciplinaria de Docencia e Investigación, perteneciente a la Facultad de Ciencias de la UNAM, en el campus Juriquilla, Querétaro. Sus coordenadas son 20.7030 °N, 100.4473 °W, a una altura de 1,945 m.s.n.m.

Las variables que registra esta estación son:

- ▶ Meteorología
- ▶ Gases de efecto invernadero
- ▶ Calidad del aire
- ▶ Carbón negro
- ▶ Características de la precipitación
- ▶ Campo eléctrico
- ▶ Depositación húmeda
- ▶ Perfil de retrodispersión



FIGURE 2.4: Estación en Juriquilla, Querétaro, donde se encuentra uno de nuestros monitores de campo eléctrico.

Observatorio atmosférico de Campus UNAM en Hermosillo, Sonora (ERNO)

Esta estación se encuentra en la Estación Regional del Noroeste (ERNO), perteneciente al Instituto de Geología de la UNAM, en el campus de la Universidad de Sonora, en Hermosillo. Sus coordenadas son 29.0814 °N, 110.9706 °W, a una altura de 154 m.s.n.m.

Las variables que esta estación registra son:

- ▶ Meteorología
- ▶ Gases de efecto invernadero
- ▶ Calidad del aire
- ▶ Campo eléctrico
- ▶ Perfil de retrodispersión



FIGURE 2.5: Estación en Hermosillo, Sonora, donde se encuentra uno de nuestros monitores de campo eléctrico

2.3 El detector de campos eléctricos Boltek EFM-100

El Boltek EFM-100 (ver figura 2.1) cuyo modo de detección está bajo el esquema de molino de campos eléctricos, posee la facultad de detectar las descargas eléctricas en un radio de hasta 38 kilómetros y mide la intensidad del campo eléctrico atmosférico. La información obtenida por el sensor del equipo, que se encuentra en el exterior, es guiada a través de un cable de fibra óptica, el cual llega a un convertidor de intensidad de luz a niveles analógico y digital, en este caso se emplea el sistema digital, la información es enviada a una PC que funciona como sistema de visualización y registro.

El Boltek EFM-100 tiene un rango de medición que va de -20 kV/m a 20 kV/m. El rango de temperaturas dentro del cual el equipo puede funcionar es de -40 a 60 °C.

El molino de campos eléctricos Boltek EFM-100 es del tipo obturador rotante, el cual consta de unas placas estacionarias con un disco de electrodos en rotación. Los molinos de campos de tipo obturador distorsionan el campo eléctrico y no son capaces de medir el gradiente del campo. Por esto mismo, no se deben considerar las mediciones de la intensidad de campo eléctrico como absolutas, sino debemos priorizar sus variaciones.

En la figura 2.6 encontramos el esquema de instalación del Boltek que debe ser siguiendo estrictas medidas de seguridad para evitar accidentes, además que debe ser apropiadamente aterrizado. Tenemos que el sensor que va en el exterior se conecta por medio de 3 cables, un cable de datos RJ-45 (como los usados para Internet) de color negro en la figura, que además es el que provee la energía para el sensor; un cable de fibra óptica (en naranja) y un cable de cobre que provee la conexión a tierra (en verde). El cable que provee nuestros datos es el de fibra óptica, que ya en el interior del edificio se conecta a un convertidor a USB para conectarse a la PC que registra y despliega los datos.

2.3.1 Teoría de operación

Los campos eléctricos se desarrollan donde quiera que haya una diferencia en el potencial eléctrico. Un campo eléctrico es lo que atrae nuestro cabello a un peine cargado o a un globo.

El campo eléctrico se mide en volts por metro (V/m). Esto viene de la definición de campo eléctrico dada por:

$$\vec{E} = \frac{F}{Q} \quad (2.1)$$

Donde la fuerza (F) se mide en Newtons (N) y la carga eléctrica (Q) se mide en Coulombs (C). Realizando el análisis dimensional obtenemos:

$$\frac{N}{C} = \frac{N}{C} \left(\frac{m}{m} \right) = \frac{J}{C \cdot m} = \frac{V}{m} \quad (2.2)$$

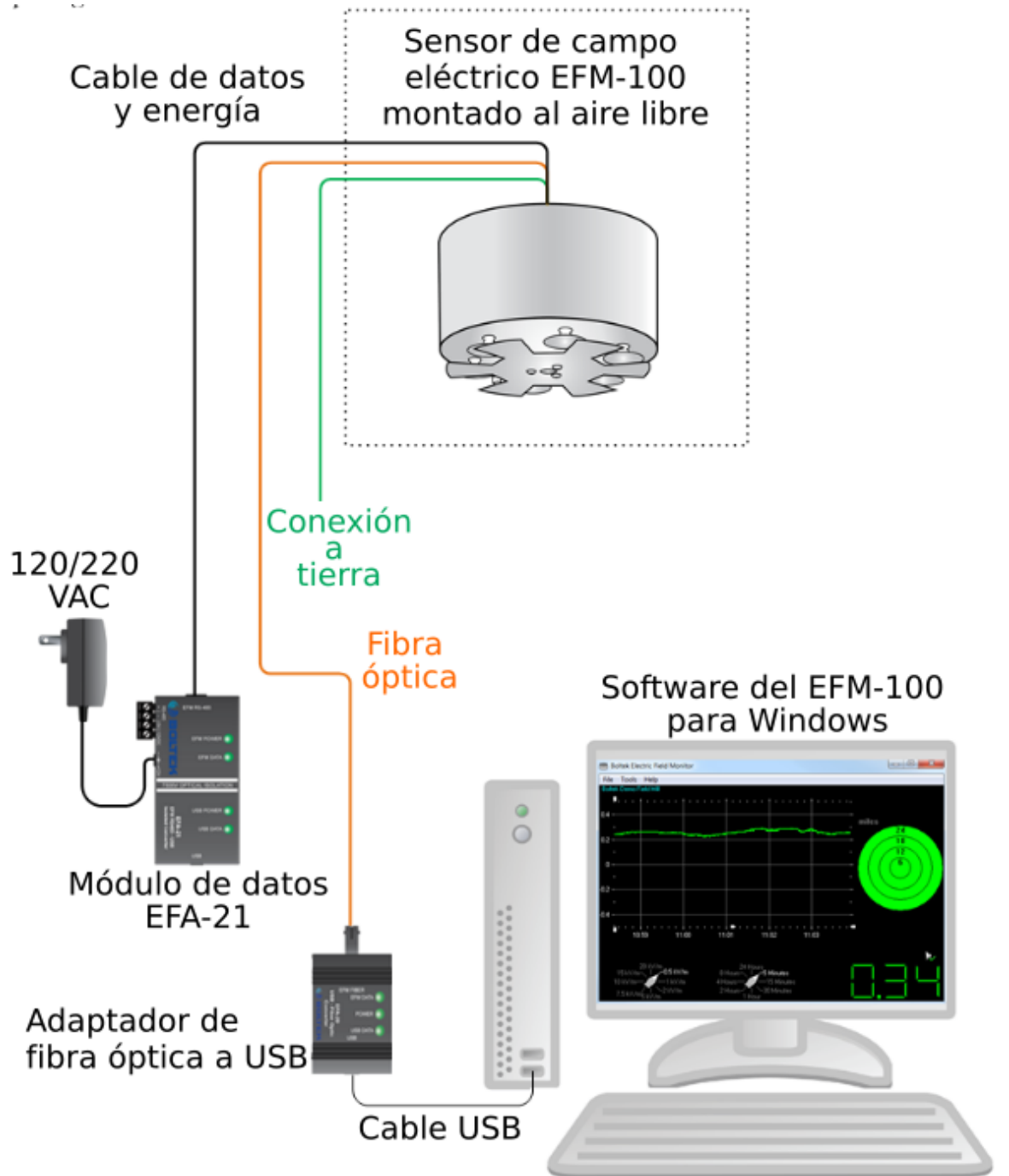


FIGURE 2.6: Diagrama de conexión del EFM-100, usando fibra óptica. Modificado de *EFM-100 Atmospheric Electric Field Monitor. Installation/Operators Guide for model EFM-100C 2016*

Los campos eléctricos que acompañan a las tormentas eléctricas normalmente los medimos en miles de volts por metro [kV/m]. Los rayos pueden ser detectados como un cambio

repentino en el campo eléctrico. La carga eléctrica contenida en una nube de tormenta también genera un campo eléctrico y este campo se puede medir en superficie. Idealmente se debe montar el molino a ras del suelo, pero esto no es práctico, porque el agua, la suciedad, los insectos, entre otras cosas, se acumularían alrededor de los electrodos y los contaminarían, disminuyendo considerablemente su vida útil.

Un molino de campo eléctrico utiliza un obturador mecánico para proteger y exponer alternativamente varias placas sensibles a un campo eléctrico. Cuando las placas de sensibilidad se exponen al campo eléctrico, una carga eléctrica es atraída desde la tierra a las placas a través de un resistor de sensibilidad. Cuando las placas de sensibilidad están protegidas del campo, la carga vuelve a fluir a la tierra de nuevo a través del resistor de sensibilidad. Esta carga en movimiento es una corriente eléctrica que se mide como una diferencia de potencial de corriente alterna (CA, corriente variable en la que las cargas eléctricas cambian el sentido del movimiento de manera periódica) a través del resistor de sensibilidad. El tamaño del voltaje es proporcional al tamaño del campo eléctrico aplicado a las placas.

La carga que fluye desde y hacia los electrodos desarrollará una diferencia de potencial a través de resistor de sensibilidad. Esta diferencia de potencial es amplificada y alimentada a un interruptor analógico junto con una versión desfasada de la señal (Monterde-Andrade, 2017).

2.3.2 Software del equipo

El software de visualización del EFM-100 interpreta y muestra las lecturas y el estado del molino de campos eléctricos. Se pueden establecer conexiones remotas con el software para recibir los datos, pero una técnica que nos ha funcionado bien es que desde la carpeta local donde se almacenan los datos se establezca una conexión a la nube usando cualquier software destinado para esto.

El software requerido para la operación del Boltek EFM-100 es admitido por los sistemas operativos pertenecientes a la plataforma de Microsoft; el equipo de registro utiliza la plataforma de Windows 10 con la versión V1.33 del software.

Un sólo sistema de visualización y registro puede recibir lecturas de hasta 4 fuentes de datos, es decir; molinos de campos eléctricos. Cada una de las fuentes de datos están conectadas a un puerto USB. Todos los datos registrados son almacenados en un archivo único que se crea automáticamente de manera continua; en otras palabras, un nuevo archivo es creado dentro del primer minuto de un día y se guardan todas las lecturas registradas hasta el último minuto de ese mismo día en un único archivo. Los datos se almacenan en archivos de extensión '.efm' y pueden ser administrados como archivos ASCII y visualizados en modo gráfico. En la pantalla del equipo de visualización, el software despliega una ventana donde todas las lecturas del día son visibles (ver figura 2.7), está disponible una opción para seleccionar los valores máximos en las gráficas. El selector del rango de tiempo permite cambiar la escala horizontal de los ejes, las opciones disponibles son 5, 15 y 30 minutos, y 1, 2, 4, 8

y 24 horas. El selector de rango de amplitud tiene las escalas de 0.5, 1, 2, 5, 7.5, 10, 15 y 20 [kV/m] para permitir un mejor análisis de los datos en tiempo real.

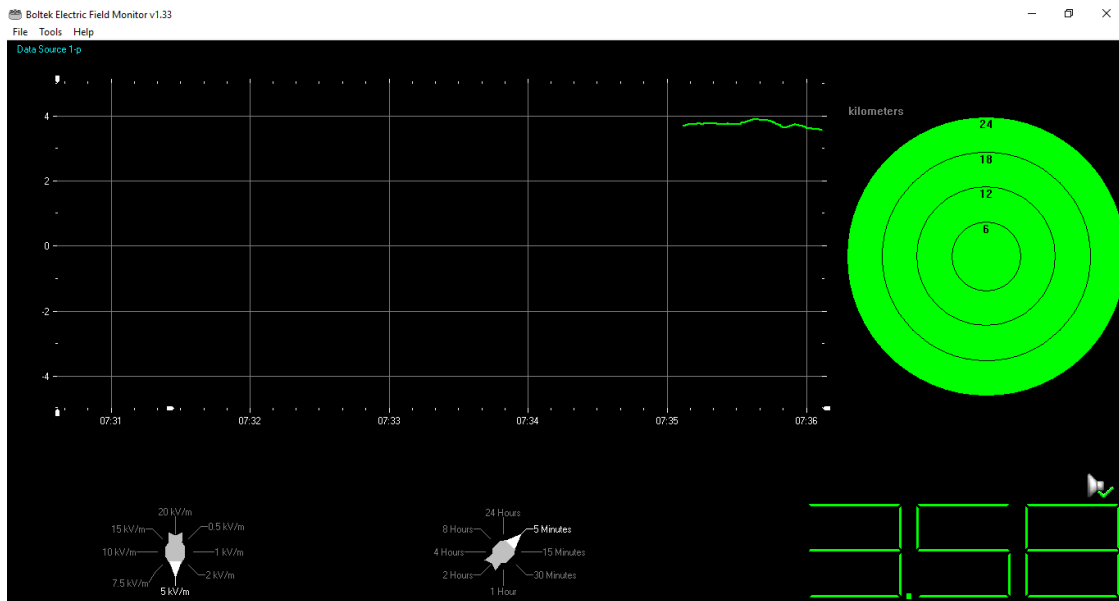


FIGURE 2.7: Software que despliega la gráfica de intensidad del campo eléctrico en tiempo real, donde en la esquina inferior derecha se puede observar la lectura del campo en kV/m cada segundo

Como una característica especial, el software del EFM-100 es capaz de emitir cuatro tipos de alarmas:

- ▶ Alarma de campo eléctrico alto: se activa cuando las lecturas exceden un valor fijado por el usuario y se mantienen un cierto intervalo de tiempo por encima del mismo. El intervalo de tiempo y la duración de la emisión de la alarma también es seleccionado por el usuario.
- ▶ Alarma de campo eléctrico muy alto: funciona de la misma manera que la alarma de campo eléctrico alto, pero con valores especificados como extremos.
- ▶ Alarma de rayos: se activa cuando el equipo detecta eventos de descargas eléctricas dentro de su radio de medición. Se mantiene activada hasta que los eventos de descargas cesan.
- ▶ Alarma Delta: funciona de manera similar a las dos primeras. Por ejemplo, si se desea que se active cuando las lecturas sobrepasen el valor de 1 [kV/m] o -1 [kV/m] y con un intervalo de 60 segundos, el campo eléctrico necesita incrementarse por encima de 1 kV/m y durante al menos 60 segundos para activar la alarma Delta.

Esto es importante en los casos donde se pretende utilizar el EFM-100 como un sistema de alerta para actividades al aire libre, como se promociona en el mismo sitio web de Boltek y se muestran ejemplos de escuelas de Estados Unidos usándolos.

2.3.3 Interpretación de las gráficas

A menudo, uno de los primeros indicadores de la aproximación de una nube de tormenta es una lectura positiva del campo, seguida por una inversión hacia una lectura negativa a medida la nube se mueve sobre el monitor. (ver figuras 2.8 y 2.9)

Las figuras que se muestran a continuación son esquemáticas y nos dan una idea de cómo se ven los registros a partir de lo que ocurre en el exterior. Son modificadas a partir del manual del Boltek EFM-100.

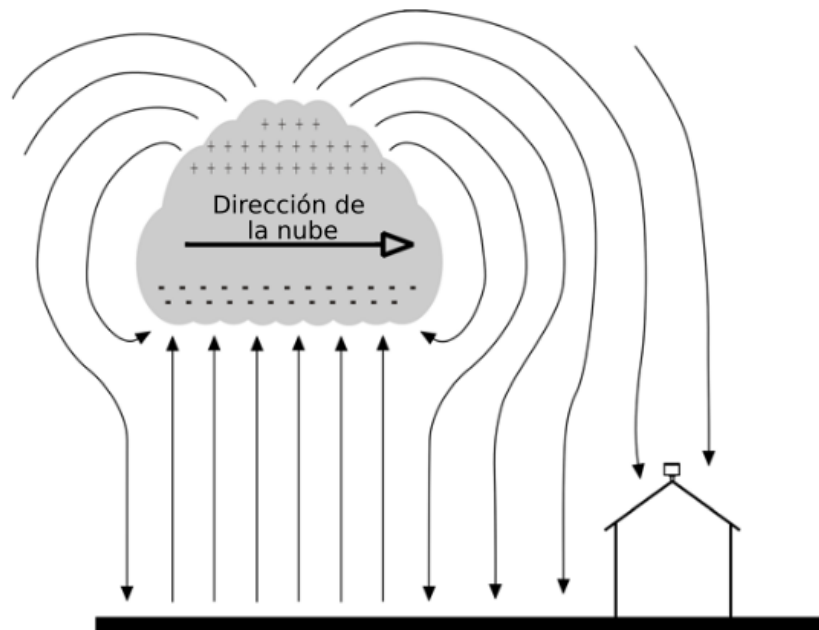


FIGURE 2.8: Nube de tormenta aproximándose.

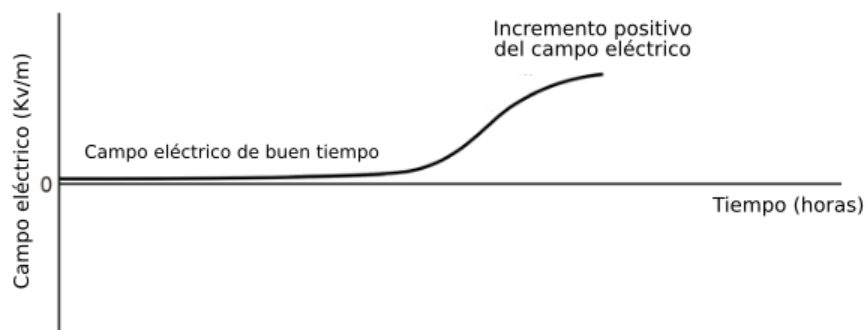


FIGURE 2.9: Incremento del campo debido a la aproximación de la nube de tormenta

Con la nube de tormenta directamente sobre el monitor de campo eléctrico la polaridad del campo se invierte hacia lecturas de campo eléctrico negativas (ver figuras 2.10 y 2.11).

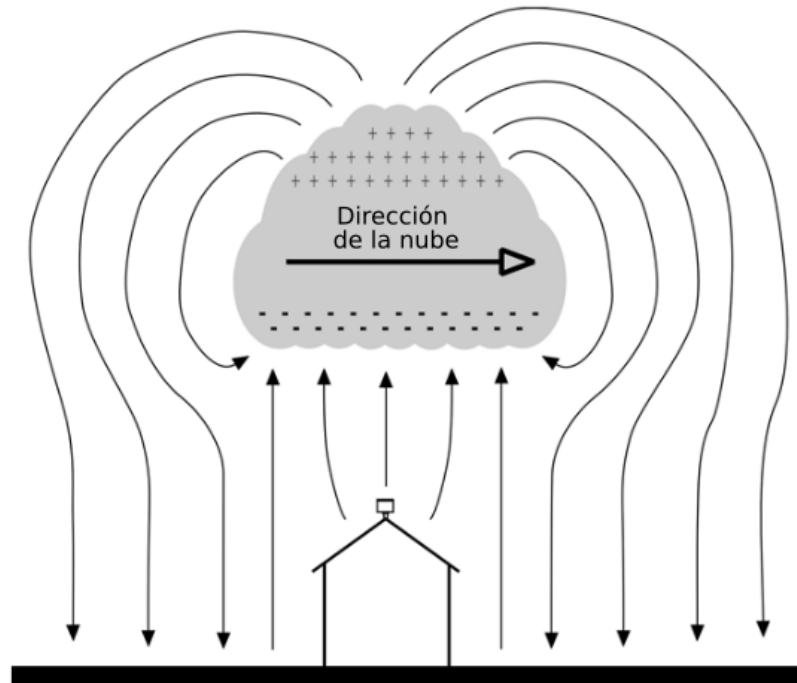


FIGURE 2.10: Nube de tormenta justo sobre el monitor

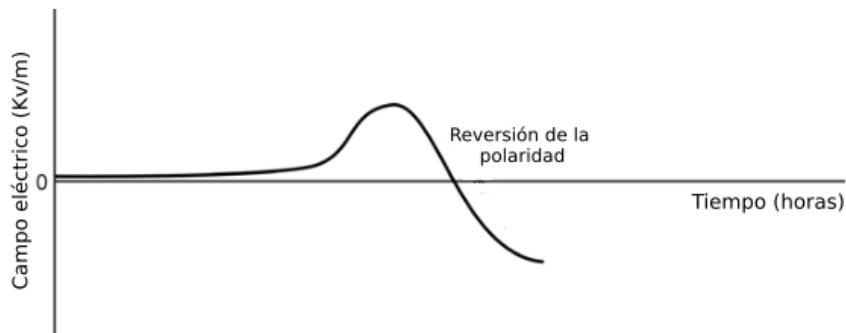


FIGURE 2.11: Inversión de la polaridad del campo eléctrico

Una vez que la nube ha pasado sobre el monitor, el campo va a regresar a positivo antes de decaer otra vez a una lectura de campo eléctrico de tiempo calmo, de aproximadamente 0.1 [kV/m]. Este cambio en el campo eléctrico hacia el fin de la tormenta se le conoce como Oscilación de Fin de Tormenta (OFT) (ver figuras 2.12 y 2.13)

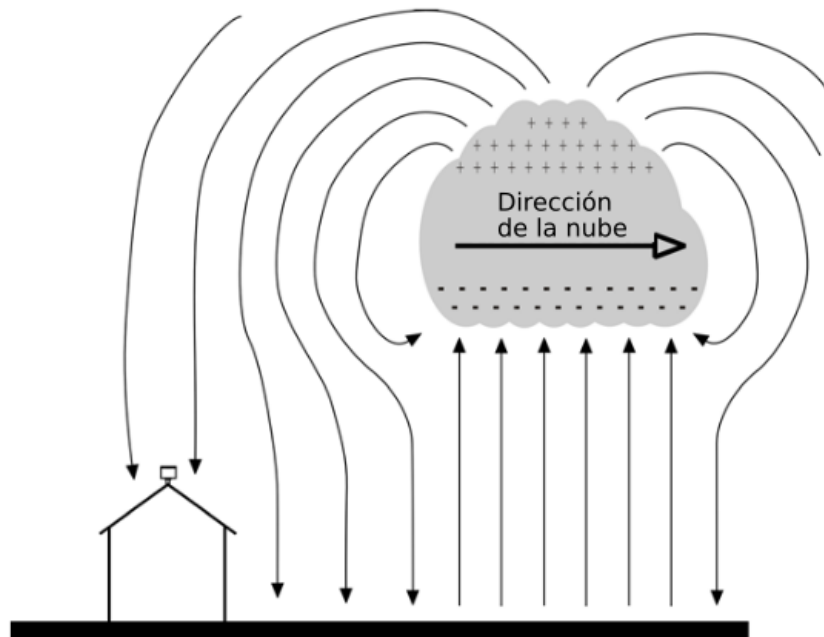


FIGURE 2.12: Nube de tormenta alejándose

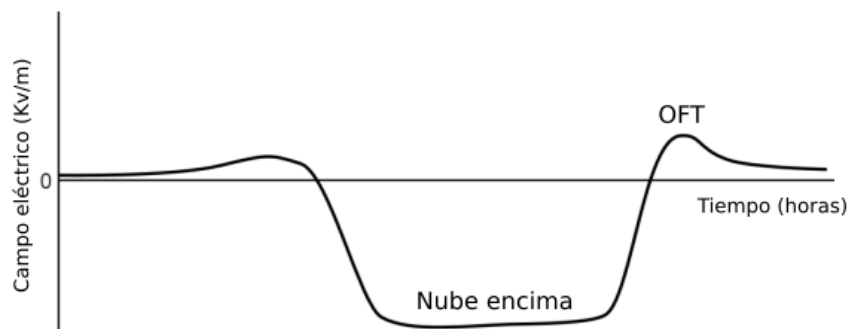


FIGURE 2.13: Oscilación del final de tormenta

Cambios bruscos en la magnitud del campo eléctrico indican rayos. Rayos cercanos producen mayores cambios del campo que rayos lejanos. El EFM-100 puede detectar rayos en un rango de 38 Km. (Ver figura 2.14)

La lluvia, nieve o el polvo pueden llevar una carga eléctrica. Si partículas cargadas tocan el plato del sensor, la carga eléctrica va a transferirse de la partícula al plato del sensor y se muestran como ruido en las lecturas. Durante períodos de lluvia intensa el ruido por precipitación puede volverse mayor (ver figura 2.15). Una vez que la tormenta pasa el ruido desaparece. El ruido por precipitación puede reducir la capacidad del software de detectar rayos lejanos, pero los cercanos aun pueden detectarse.

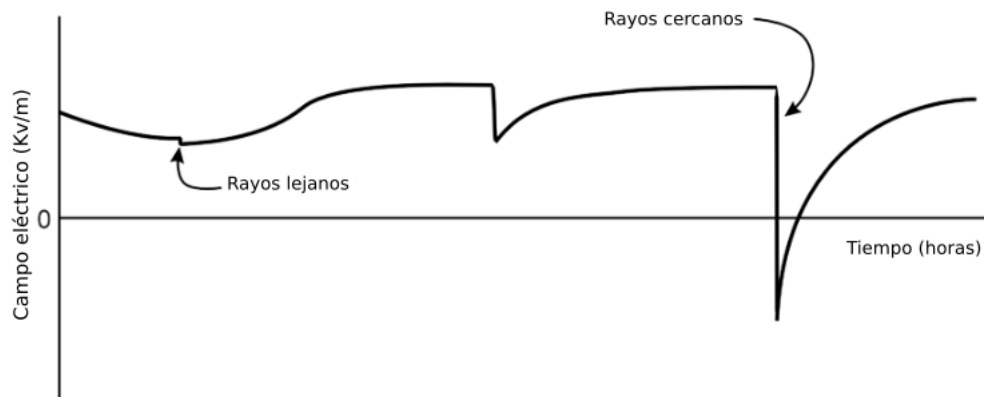


FIGURE 2.14: Saltos en la intensidad del campo que indican rayos

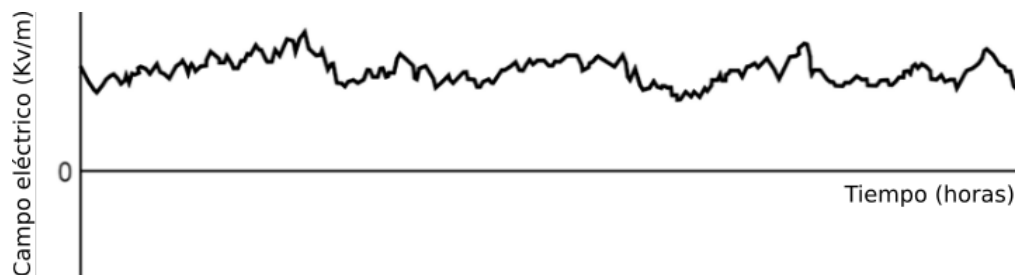


FIGURE 2.15: Cómo se observa el ruido por precipitación

2.4 Instalación de un Boltek EFM-100 en Sierra Negra, Puebla

La RUOA cuenta con 3 monitores de campo eléctrico EFM-100 en su red, mientras que en esta tesis ocupamos también la base de datos de un cuarto EFM-100 que estuvo registrando datos durante aproximadamente 2 años en la terraza del Instituto de Geofísica de la UNAM.

En junio de 2019 este equipo fue desmontado de este sitio y se le hicieron pruebas al interior de un cubículo para verificar su correcto funcionamiento, ya que se tuvo la idea de trasladarlo a Sierra Negra, Puebla, en las instalaciones del Gran Telescopio Milimétrico Alfonso Serrano (GTM) perteneciente al Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE). En este sitio el Instituto de Geofísica ya cuenta con dos observatorios de Rayos cósmicos, uno llamado Telescopio de Neutrones Solares (TNS) y el otro Telescopio de Centelleo de Rayos Cósmicos (SciCRT). Esto es importante porque posiblemente sea el EFM-100 instalado a mayor altura en el mundo, a 4,590 m.s.n.n. y permitirá a los científicos obtener datos valiosos para estudiar las variaciones que existen entre el registro de rayos cósmicos secundarios y el campo eléctrico local.

Por lo tanto, El lunes 7 de octubre de 2019 nos dispusimos un equipo de 3 personas conformado por el Dr. Luis Xavier González Méndez (tutor de esta tesis), el estudiante de maestría Fernando Monterde Andrade y mi persona a llevar, instalar, calibrar y poner en marcha el cuarto equipo EFM-100 en la cima de Sierra Negra.

Primero llegamos al campamento base del INAOE ubicado en Ciudad Serdán, para descansar la noche y al siguiente día emprender viaje a Sierra Negra. Ya en el sitio, desplegamos el equipo necesario e instalamos, calibramos y pusimos en marcha el Boltek EFM-100 apoyándonos de un contenedor contiguo al TNS. La instalación quedó finalmente en las coordenadas $18.985459^{\circ}\text{N}$, $97.315341^{\circ}\text{W}$ y a una altura de 4,590 m.s.n.m. (figuras 2.16 y 2.17)



FIGURE 2.16: Boltek EFM-100 instalado en Sierra Negra

El sensor quedó instalado de la forma usual, con el rotor mirando hacia arriba, a diferencia de los otros tres equipos de la RUOA donde el rotor queda mirando hacia abajo y con



FIGURE 2.17: Cima de Sierra Negra, donde quedó instalado el detector de campo eléctrico Boltek EFM-100, a un lado del GTM

un vaso protector de intemperie. La principal ventaja de instalar el sensor de manera inversa se encuentra en un menor mantenimiento porque no se acumula suciedad o nieve en las aspas y además, protege del ruido por lluvia.

En las mismas instalaciones del TNS establecimos la computadora (figura 2.18) que recibe y almacena los datos del Boltek EFM-100, además que envía estos datos por medio de una conexión a Internet a una base de datos en la nube, que esperamos incluir próximamente en la RUOA. Cabe mencionar que los datos medidos en esta ubicación son muy prometedores y de buena calidad.

Finalmente, en la figura 2.19 observamos 3 días después de la instalación, la firma característica de un episodio de tormenta, donde identificamos aspectos como su duración, de aproximadamente una hora y también los constantes cambios de polaridad que alcanzan los niveles de saturación de ± 20 kV/m debido a la altura, ya que el equipo se encuentra muy cerca de las nubes de tormenta.



FIGURE 2.18: Computadora que recibe los datos del Boltek-EFM100

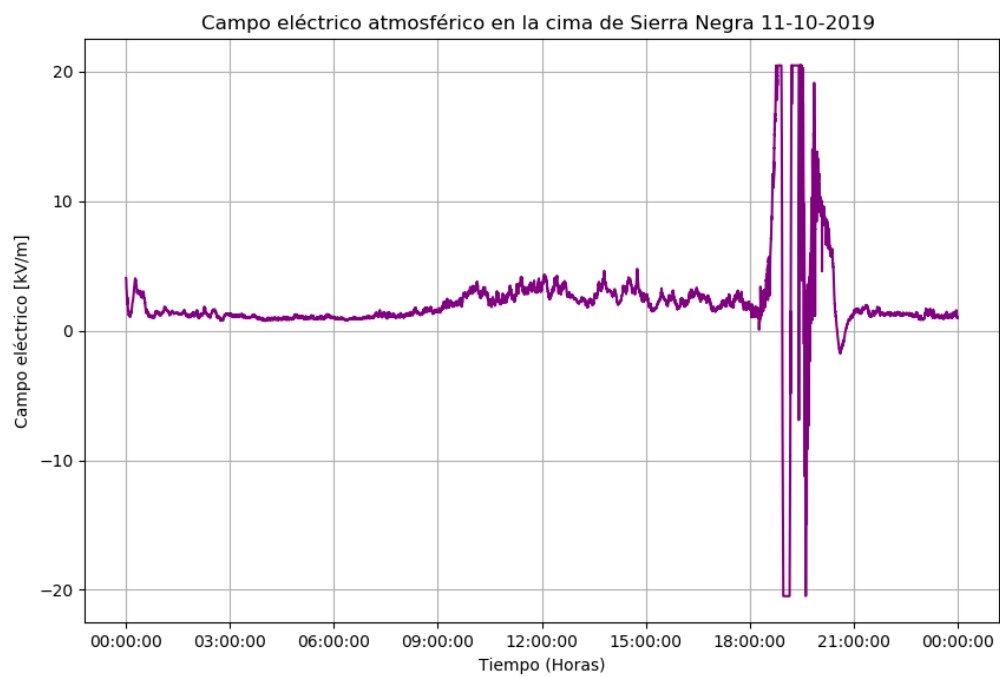


FIGURE 2.19: Episodio de tormenta registrado en la cima de Sierra Negra el 11 de octubre del 2019

3

BASE DE DATOS

3.1 Introducción

En el capítulo 3 hacemos una revisión del lenguaje de programación que se utilizó para analizar nuestra base de datos, así como la teoría detrás de la transformada de Fourier, que nos permite pasar una señal del dominio del tiempo al dominio de la frecuencia. Por último, expondremos el estado en el que se encontraba la base de datos, tanto para los 3 equipos Boltek EFM-100 activos con los que cuenta la RUOA como el cuarto equipo del Instituto de Geofísica, el cual se encuentra ahora registrando datos en Sierra Negra, Puebla.

3.2 Lenguaje de programación Python

Durante la carrera de Ingeniería Geofísica en la Facultad de Ingeniería recibimos clase de programación desde el inicio, muy básicas para el caso de C, pero poco a poco nos fuimos adentrando en el mundo del tratamiento y procesamiento de datos, para lo cual iniciamos con Matlab, lenguaje amigable y poderoso, para luego continuar con Python y Fortran, los reyes del procesamiento de datos en la actualidad. Decidí decantarme por Python para la presente tesis por ser un lenguaje de programación con una sintaxis sencilla, pero poderoso y altamente flexible, que además, tiene librerías muy especializadas para una amplia gama de análisis, incluyendo las visualizaciones gráficas. Otra ventaja reside en que es de código abierto.

Una aclaración importante que debe hacerse es la diferencia entre lenguajes compilados y lenguajes interpretados, porque una cosa es el lenguaje de programación llamado Python y otra cosa es el interprete de Python. El interprete es un programa que ejecuta otros programas. Lenguajes como Fortran o C necesitan ser compilados, en ellos un programa llamado compilador, traduce a lenguaje de máquina el código creado por el programador y produce un ejecutable. En Python por el contrario el interprete lee la primera línea el código escrito por el programador, lo traduce a lenguaje de máquina y la ejecuta, luego lee la siguiente y así sucesivamente. Esto hace que la ejecución sea mas lenta que en lenguajes como C pero se gana en flexibilidad y permite trabajar en forma interactiva (Suárez Ledo, 2014)

3.2.1 Características de Python

Python es un lenguaje de programación interpretado de alto nivel, entre sus principales características están:

Lenguaje de alto nivel: Este tipo de lenguajes se caracterizan principalmente por contar con una estructura sintáctica y semántica legible, acorde a las capacidades cognitivas humanas, (Bahit, 2012).

Es un lenguaje interpretado: A diferencia de los lenguajes compilados Python está diseñado para ser ejecutado por medio de su intérprete. Un intérprete funciona de manera similar a un compilador, con la diferencia que ejecuta el código de manera directa y no necesita ejecutar primeramente un ejecutable para el programa.

Multiplataforma: Python está diseñado para poder ser ejecutado en diferentes plataformas informáticas como por ejemplo: Microsoft Windows, GNU Linux, Mac OS X, etc.

Filosofía: la filosofía de Python hace especial énfasis en una sintaxis clara y limpia que favorezca un código legible.

Multiparadigma: Se dice que Python es multiparadigma ya que soporta diferentes estilos de programación como: la orientación a objetos, programación imperativa y, en menor medida, programación funcional.

Tipado dinámico: Esto quiere decir que Python nos da la facilidad de no definir el tipo al declarar una variable, por lo que ésta puede tomar valores de distinto tipo en diferentes momentos.

En esta tesis se usó como intérprete el IDE (Integrated Development Environment) llamado VS Code, perteneciente a Microsoft.

3.3 Transformada de Fourier

Las principales aportaciones de Jean-Baptiste Joseph Fourier (1768-1830) a la física y las matemáticas fueron el desarrollo de la teoría del calor, ecuaciones diferenciales, ecuaciones algebraicas, series trigonométricas, estadística matemática y teoría de las probabilidades; siendo su obra monumental la *Théorie analytique de la chaleur*; en la que desarrolla ecuaciones para explicar la conducción térmica en metales. La primera publicación sobre este tema fue una memoria presentada a la Academia en 1807 y completada con el texto *Théorie analytique de la chaleur* en 1822. Para explicar la conducción térmica, Fourier utiliza series matemáticas infinitas que permiten encontrar las soluciones del problema. El desarrollo de una función en forma de series infinitas de funciones trigonométricas se conoce como series de Fourier y tienen numerosas aplicaciones en prácticamente todas las áreas científico-técnicas con las que hay que modelar numerosos datos y procesos complejos. Tiene especial importancia en física, química, climatología, oceanografía, e ingeniería.

Para el caso de esta tesis, la aportación que nos interesa de Fourier es la Transformada de Fourier, que se define para una función $f(t)$ como:

$$\mathcal{F}[f(t)] = F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t} dt \quad (3.1)$$

y por lo tanto si

$$\mathcal{F}[f(t)] = F(\omega)$$

entonces

$$\mathcal{F}^{-1}[F(\omega)] = f(t)$$

Quedando la Transformada inversa de Fourier como:

$$\mathcal{F}^{-1}[F(\omega)] = f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)e^{i\omega t} d\omega \quad (3.2)$$

En el caso de esta tesis se uso la función de Python para la Transformada rápida de Fourier (FFT), que consiste en un algoritmo que permite calcular eficientemente la Transformada de Fourier discreta (TFD) y su inversa.

La Transformada Rápida de Fourier es de suma importancia en el análisis, diseño y realización de algoritmos y sistemas de procesamiento de señales dado que brinda mayor eficiencia tanto en tiempo como en recursos. La FFT se basa en la división del tiempo, eliminando así gran parte de los cálculos repetitivos que hay que llevar a cabo si se desea resolver la TFD de forma directa, (Schmidt, 2013).

3.4 Estado de la base de datos

Todo el procesamiento y análisis de esta tesis proviene de la base de datos de campo eléctrico atmosférico de la RUOA, la cual consiste en 3 estaciones que cuentan con un Boltek EFM-100, como se vio en el capítulo 2, las cuales son:

- ▶ El Observatorio atmosférico de Ciudad Universitaria, CDMX. (UNAM)
- ▶ El Observatorio atmosférico de Campus UNAM en Juriquilla, Querétaro (JQRO)
- ▶ El Observatorio atmosférico de Campus UNAM en Hermosillo, Sonora. (ERNO)

Estos datos se encuentran abiertos al público en el sitio web de la RUOA y para referirnos a las las bases de datos usamos los acrónimos de sus respectivas estaciones.

Además se cuenta con la base de datos de un cuarto equipo instalado en el Instituto de geofísica (acrónimo GEOF), de cuya instalación y calibración se habla en la tesis de Fernando Monterde (Monterde-Andrade, 2017). Su ubicación exacta (ver figura 2.1) se encontró en las coordenadas 19.3109545 °N, 99.162774 °W a una altura de 2,270 m.s.n.m. Además, este

equipo fue traslado después de 2 años y medio a Sierra Negra, Puebla; como se puede ver en el capítulo 3. Estos datos se han estado recopilando de forma continua y posteriormente pasarán a formar parte oficial de la base de datos de la RUOA.

Los datos se encuentran tanto en sus versiones crudas, como en los niveles de control de calidad L0 y L1. El proceso para llegar a esto consiste en que primero los registradores (dataloggers) de las tres estaciones donde existe un Boltek EFM-100 generan dos tablas de datos 'crudos' a partir de mediciones de intensidad del campo eléctrico (kV/m) que realiza el instrumento cada segundo:

- ▶ Promedios de intensidad de campo eléctrico cada minuto.
- ▶ Varianza de la intensidad de campo eléctrico cada hora.

Cabe mencionar que los datos crudos de varianza por hora no se utilizan más que para generar gráficas de desviación estándar de campo eléctrico que se muestran en el sitio web y las pantallas encontradas en las estaciones. No se utilizan en ninguna otra parte del control de calidad y no están disponibles para descarga.

Posteriormente, los datos crudos promediados por minuto los toma un script de Python que genera datos de un primer nivel de control de calidad (L0). Lo que hace este código es eliminar líneas repetidas, es decir, con la misma estampa de tiempo, y si hay periodos de minutos faltantes, se introducen las líneas con la estampa de tiempo faltante y se rellenan los datos con 'null'. Esencialmente, este script asegura la integridad temporal de cada archivo.

Después, los datos de nivel L0 son tomados por otro script, que acota los promedios de minuto, eliminando aquellos promedios que no cumplen con que el campo eléctrico se encuentre entre -20.0 y 20.0 [kV/m]. Estas líneas eliminadas son reemplazadas con 'null'. El archivo resultante después de pasar este filtro constituye el nivel L1 de minuto, y son los que se ponen a disposición de los usuarios para descarga. Además, los datos L1 de minuto se utilizan por este último script para generar archivos de desviación estándar nivel L1 de hora.

Para el caso de la estación UNAM los datos L1 van de febrero de 2015 a mayo de 2018. Los códigos de Python que usa la RUOA para los niveles de calidad L0 y L1 dejaron de correr adecuadamente para datos posteriores a mayo de 2018, pero en esta tesis se encontró y corrigió un error que impedía la correcta corrida de los códigos, posiblemente debido a una actualización de Python y se espera que esto solucione algunos de los problemas (como se ve en el capítulo 4).

En cuanto a la estación JQRO los datos L1 también van de febrero de 2015 a mayo de 2018, encontrándose el mismo error para correr los códigos de calidad que en UNAM, mientras que en la estación ERNO los datos L1 van de octubre de 2015 a noviembre de 2019, mes en el que se hizo el corte datos para esta tesis, ya que la estación ERNO sigue publicando datos de campo eléctrico de forma continua hasta la fecha.

Finalmente, tenemos la base de datos GEOF cuyos datos crudos van de marzo de 2017 a junio de 2019, fecha en la que el equipo fue bajado del techo del Instituto de Geofísica

para realizarle las pruebas correspondientes para trasladarlo a la cima de Sierra Negra. Cabe destacar que los datos crudos de este equipo son diferentes a los de los equipos de la RUOA, por lo que uno de los productos de esta tesis (capítulo 4) es un conjunto de scripts de Python que permiten llevar los datos crudos de GEOF a los datos crudos de los equipos de la RUOA y así poder ocupar los scripts destinados a los controles de calidad L0 y L1.

4

ANÁLISIS DE DATOS Y RESULTADOS

4.1 Introducción

En este capítulo final se presentan el procesamiento, análisis e interpretación que se realizó de la base de datos de campo eléctrico que se describe en el capítulo 4, y se exponen los productos obtenidos durante la realización de esta tesis.

Estos productos consisten básicamente en 4 apartados:

- ▶ Scripts de Python que permiten traducir los registros crudos de la base de datos GEOF a los registros crudos de la RUOA, con el objetivo de poder aplicarles los scripts para los niveles de control de calidad L0 y L1, los cuales también fueron corregidos.
- ▶ Correlaciones entre las bases de datos GEOF y UNAM, que registraron datos en el mismo sitio geográfico durante más de dos años, para corroborar si hay estabilidad estadística entre ellos.
- ▶ Transformada de Fourier para obtener los espectros de frecuencia de todas las estaciones y así identificar ciclos presentes en ellas.
- ▶ Inspección de toda la base de datos, obteniendo hojas de cálculo para cada estación donde se describe con un código de color la calidad de los datos y se añaden comentarios para cada día.

Además, todos los códigos usados para los productos de esta tesis y otros códigos necesarios para graficar los datos aquí expuestos, se podrán encontrar en los anexos de esta tesis y en el repositorio de Github con dirección:

<https://github.com/AlejandroArgueta/Tesis-Licenciatura>

4.2 Niveles de calidad L0 y L1 aplicados a la base de datos GEOF

En un inicio del procesamiento realizado en esta tesis nos dimos cuenta de un detalle importante: si queríamos añadir el boltek del Instituto de Geofísica a la RUOA (Antes de planear

moverlo a Sierra Negra), necesitábamos uniformizar su base de datos con las demás base de datos de la RUOA, es decir, aplicarle a los datos los niveles de calidad L0 y L1. Para llevar a cabo esta tarea podíamos diseñar los códigos L0 y L1 desde cero para la forma en que se encuentran los datos crudos del Boltek del Instituto de Geofísica, pero una mejor opción fue crear scripts que nos permitieran llevar los datos crudos del equipo del Instituto de Geofísica a los datos crudos de la RUOA, sobre los cuáles eran aplicados los scripts de L0 y L1. Esto será importante para estandarizar la base de datos que se está registrando en estos momentos en la cima de Sierra Negra.

En la imagen 4.1 podemos ver la estructura que tienen los datos crudos de los equipos de la RUOA. Consisten en 4 líneas de encabezados y ya en los datos tenemos 3 columnas que corresponde a la estampa de tiempo (entre comillas), a un contador y a un promedio por minuto de la intensidad de campo eléctrico, obtenido a partir de un muestreo de un dato por segundo.

```

1 "TOA5", "UNAM", "CR3000", "10004", "CR3000.Std.31", "CPU:UNAM.CR3", "58355", "VoltMin"
2 "TIMESTAMP", "RECORD", "E_Avg"
3 "TS", "RN", "kV/m"
4 "", "", "Avg"
5 "2017-10-16 15:58:00", 0, 1.373
6 "2017-10-16 15:59:00", 1, 1.379
7 "2017-10-16 16:00:00", 2, 1.402
8 "2017-10-16 16:01:00", 3, 1.283
9 "2017-10-16 16:02:00", 4, 1.34
10 "2017-10-16 16:03:00", 5, 1.12
11 "2017-10-16 16:04:00", 6, 1.187
12 "2017-10-16 16:05:00", 7, 0.924
13 "2017-10-16 16:06:00", 8, 0.98
14 "2017-10-16 16:07:00", 9, 0.939
15 "2017-10-16 16:08:00", 10, 0.882
16 "2017-10-16 16:09:00", 11, 0.965
17 "2017-10-16 16:10:00", 12, 0.877
18 "2017-10-16 16:11:00", 13, 1.041
19 "2017-10-16 16:12:00", 14, 1
20 "2017-10-16 16:13:00", 15, 0.751
21 "2017-10-16 16:14:00", 16, 1.018
22 "2017-10-16 16:15:00", 17, 0.963
23 "2017-10-16 16:16:00", 18, 1.154
24 "2017-10-16 16:17:00", 19, 1.181
25 "2017-10-16 16:18:00", 20, 1.015
26 "2017-10-16 16:19:00", 21, 1.097
27 "2017-10-16 16:20:00", 22, 1.09

```

FIGURE 4.1: Estructura que poseen los datos crudos de campo eléctrico atmosférico de los equipos de la RUOA

La misión de los scripts elaborados para esta parte de la tesis es transferir los datos crudos del equipo instalado en el Instituto de Geofísica lo más parecido posible a los datos crudos de la RUOA, con el fin de hacer los mínimos cambios necesarios en los scripts de L0 y L1. Como se puede observar en la figura 4.2, tenemos que de a) a b) los datos sufren una

transformación en su formato. Los pasos que se llevan a cabo aquí se describen a continuación:

- ▶ Un script llamado 'Average4rawdata.py' toma los datos crudos como se ven en a) y generan los promedios por minuto a partir de la tasa de muestreo de 2 datos por segundo, además de incluir en la estampa de tiempo año, mes y día, donde ya se encontraba hora, minuto y segundo. Cabe mencionar que realizar estos promedios permitió también verificar la integridad de cada archivo.
- ▶ Otro script llamado 'mergedata.py' une estos datos por día en un solo archivo mensual.
- ▶ Finalmente, un script llamado 'namecolumns.py' le asigna encabezados en blanco al archivo y una fila con los nombres de las columnas obteniendo lo que se ve en b).

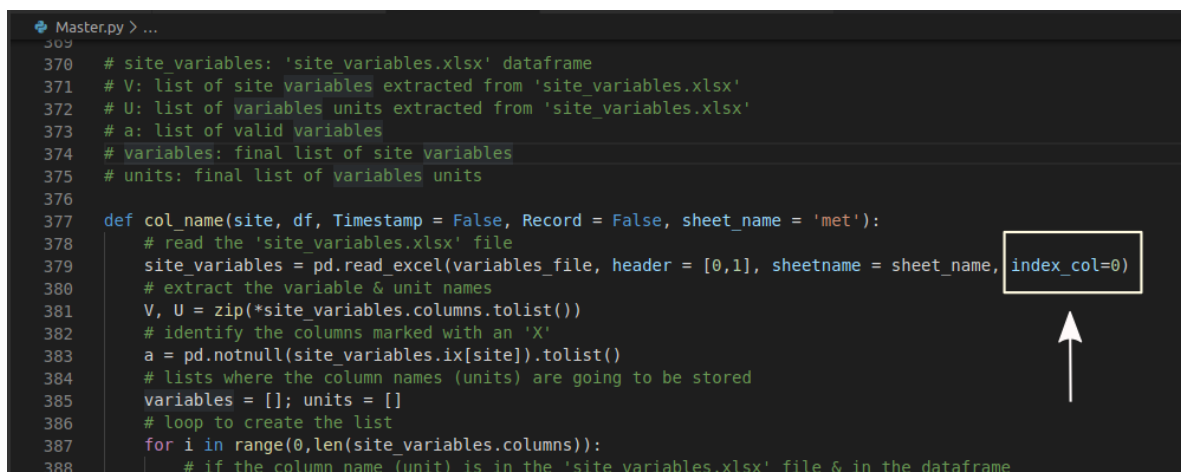
The image shows a side-by-side comparison of data before and after transformation. On the left, labeled 'a)', is a list of raw data points. Each line contains a timestamp in HH:MM:SS format followed by a comma and a numerical value. For example, '1 00:00:00,+1.03,0'. On the right, labeled 'b)', is the transformed data. The first line is a header: '2 TIMESTAMP,RECORD,E_Avg'. The subsequent lines have the same structure as in 'a)', but the timestamps are now in YYYY-MM-DD HH:MM:SS format, such as '5 2017-04-01 00:01:00,0,0.9039344262295088'. A large black arrow points from the raw data in 'a)' to the transformed data in 'b)'. The text editor interface includes a dark header with 'Abrir' buttons and a file path '~ /MEGA/'.

FIGURE 4.2: a) Datos crudos del Boltek del Instituto de Geofísica, b) Datos transformados al formato crudo de la RUOA

Después de realizado el proceso, procedemos a aplicar a estos archivos mensuales los códigos L0 y L1. Cabe mencionar que los códigos no corrían y aunque las modificaciones efectuadas para su correcto funcionamiento fueron mínimas, fue bastante extenuante encontrar estos errores. Los códigos L0 y L1 necesitan un tercer script master, que funciona para otros scripts usados para tratar otras variables meteorológicas de la RUOA, a parte de unas hojas de cálculo, donde se encuentran los umbrales para los filtros aplicados en el nivel L1, los encabezados de salida para L1, y los datos de la estación. Todos estos archivos se pueden encontrar en el repositorio de Github mencionado en la Introducción de este capítulo.

En un inicio, el problema consistió en comprender las tareas realizadas por los scripts L0, L1 y master, ya que los desarrolladores de estos códigos fueron otros estudiantes que tuvieron su paso por la RUOA. Se analizaron las líneas que describen el proceder del código, que no necesariamente coincidía siempre con los comentarios. Además, se cambiaron las rutas donde se adquirirían los datos y donde se guardaban los archivos de salida. Otro detalle es que se comentó toda la sección de 'rainfall check' en el archivo 'master.py', ya que son líneas de código que se utilizan para otras variables meteorológicas y no campo eléctrico.

Como un ejemplo del trabajo del desarrollo de los códigos, dentro de los problemas de programación remanente, el problema principal residió en la línea 380 del script original 'master.py', donde se lee el Excel 'site_variables', donde se tuvo que agregar un 'index_col=0' para que pudiera correr, esto se puede apreciar en la figura 4.3. Creemos que este error se debe a la versión anterior de la librería Pandas, cuando fue creado originalmente el script. Es posible que antes 'index_col' por default fuera cero, mientras que en la nueva versión es obligatorio especificar que este es cero, es decir, que no se cuenta con una columna índice, porque desde la primera columna de la hoja de cálculo se cuenta con datos.



```

370 # site_variables: 'site_variables.xlsx' dataframe
371 # V: list of site variables extracted from 'site_variables.xlsx'
372 # U: list of variables units extracted from 'site_variables.xlsx'
373 # a: list of valid variables
374 # variables: final list of site variables
375 # units: final list of variables units
376
377 def col_name(site, df, Timestamp = False, Record = False, sheet_name = 'met'):
378     # read the 'site_variables.xlsx' file
379     site_variables = pd.read_excel(variables_file, header = [0,1], sheetname = sheet_name, index_col=0)
380     # extract the variable & unit names
381     V, U = zip(*site_variables.columns.tolist())
382     # identify the columns marked with an 'X'
383     a = pd.notnull(site_variables.ix[site]).tolist()
384     # lists where the column names (units) are going to be stored
385     variables = []; units = []
386     # loop to create the list
387     for i in range(0, len(site_variables.columns)):
388         # if the column name (unit) is in the 'site_variables.xlsx' file & in the dataframe

```

FIGURE 4.3: Código añadido al script master.py para que L0 y L1 pudieran correr

Posteriormente, pudimos proceder a aplicar el script para L0, como se explicó en el capítulo 4, para rellenar las estampas de tiempo faltante y eliminar líneas repetidas, con el objetivo de garantizar la integridad temporal del archivo, además de crear un archivo de estadísticas que menciona los datos faltantes por día.

El script L1 toma los datos nivel L0 y les aplica un filtro donde elimina las lecturas fuera del rango de medición del equipo que es de -20 a 20 [kV/m], además de también crear un archivo de estadísticas de datos faltantes por día. Esto último es importante porque a partir de estos datos faltantes comenzamos a elaborar nuestros archivos de calidad L2, como se ve en la sección 5.5 de este capítulo.

Una característica de estos códigos es que están hechos para correrse de forma automática, por lo que se aplica solo para los datos nuevos que entran en tiempo real. Por lo que si queremos correrlos para una base de datos de muchos meses, como el trabajo desarrollado en esta tesis, no deben haber carpetas con los nombres de mes y año previamente creadas, porque el código asume que esas ya están procesadas.

De este modo, aplicando los scripts desarrollados, obtuvimos una base de datos uniforme en el nivel L1. Con base en la experiencia obtenida, se pueden tratar los datos que se registren en la cima de Sierra Negra.

4.3 Correlación entre las bases de datos UNAM Y GEOF

Una pregunta muy importante que surgió durante el desarrollo de esta tesis fue si los equipos Boltek EFM-100 medían con buena repetibilidad, es decir, si al encontrarse en el mismo lugar geográfico, estos podían medir los mismos datos de campo eléctrico atmosférico. Esto pudo corroborarse, ya que en nuestra base de datos actual tenemos más de un año de datos para comparar entre el Boltek del CCA y el del Instituto de Geofísica. Por lo que procedimos a tomar un día como ejemplo en los que las mediciones fueron casi completas para ambos equipos (los datos faltantes fueron rellenados con la tendencia). Este día fue 15 de abril del 2018, y sí, efectivamente ambos equipos tuvieron mediciones estadísticamente coherentes, como se puede comprobar en la figura 4.4. Esto se hizo con el script 'CorrelationsL1.py'.

El parecido en esta gráfica es particularmente gratificante, dado que ambos equipos fueron adquiridos en diferentes años, instalados en edificios con posiblemente diferentes condiciones de ruido, y manejados por equipos de investigación independientes; por un lado el equipo de investigación de la RUOA y por otro el grupo de Rayos Cósmicos del Instituto de Geofísica.

Es de notar que existía un desfase en tiempo, donde el Boltek del Instituto de Geofísica se encontraba exactamente una hora adelantado respecto al del CCA, afectando al índice de correlación obtenido. Esto se debe a que los equipos de la RUOA miden en hora local fija, es decir, siempre UTC -6, por lo que se procedió a hacer la corrección en tiempo.

Además, en la figura 4.4 puede verse en el campo eléctrico negativo que se da pasadas las 16:00 horas, que para el Boltek del Instituto de geofísica existe un mayor valor en la medida. Esto se debe a que ambos equipos pueden poseer un diferente plug atenuador, el cual es un resistor que reduce el valor del campo eléctrico obtenido de forma cruda por el equipo y puede llegar a funcionar de forma distinta cuando existen altos valores del campo, lo cual es corregido a la hora de calibrar el detector, tal como se realizó en la instalación en Sierra Negra.

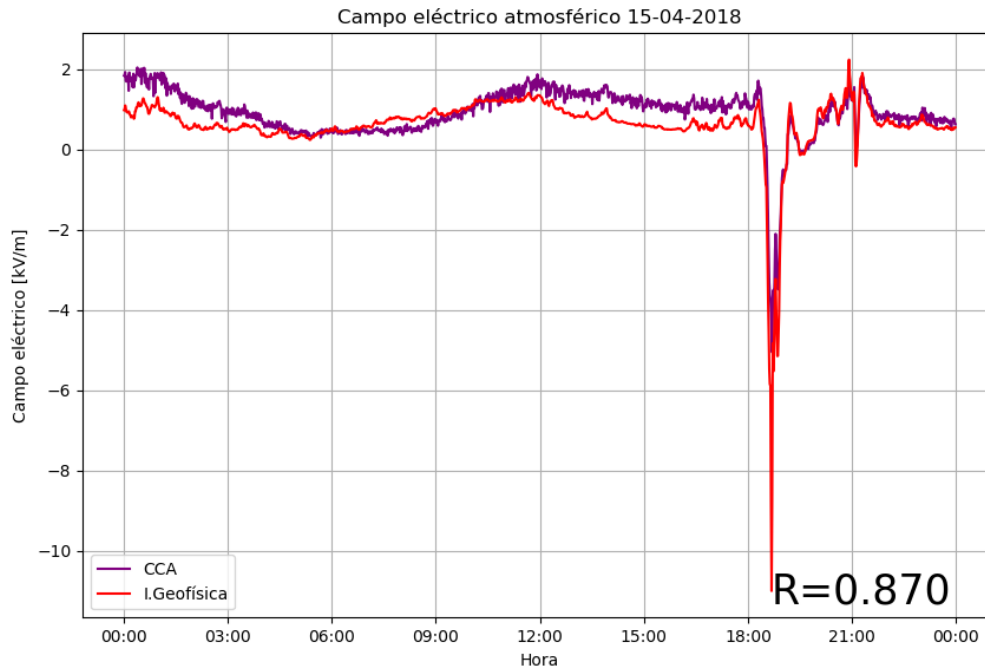


FIGURE 4.4: Correlación entre los equipos del CCA y del Instituto de Geofísica para el día 15 de abril del 2018

4.4 Espectros de amplitudes

Lo siguiente que se realizó en el análisis de esta base de datos fue una Transformada Rápida de Fourier (FFT), de la cual se expuso en el capítulo 4, para obtener el espectro de amplitudes de las señales por año y así poder confirmar de forma matemática los ciclos que se encuentran en el campo eléctrico local de cada estación. En el capítulo 1 se habló de la Curva de Carnegie, un aumento en el valor del campo, que normalmente debe observarse alrededor de las 19:00 horas UTC, debido al Circuito Eléctrico Global. Cabe mencionar que para poder correr el script 'fft.py' primero se tuvieron que rellenar los datos faltantes usando interpolación lineal con el script 'filldatatrend.py'.

A continuación se presentan las gráficas del campo eléctrico y sus respectivos espectros de amplitudes para cada estación, eligiendo un año modelo para cada una.

4.4.1 Espectro de amplitudes del CCA

En la figura 4.5 observamos la intensidad de campo eléctrico atmosférico para el año 2017 en el Centro de Ciencias de la Atmósfera de la UNAM. Apreciamos que los registros de episodios de tormenta se empiezan a ver desde marzo y continúan hasta mediados de octubre.

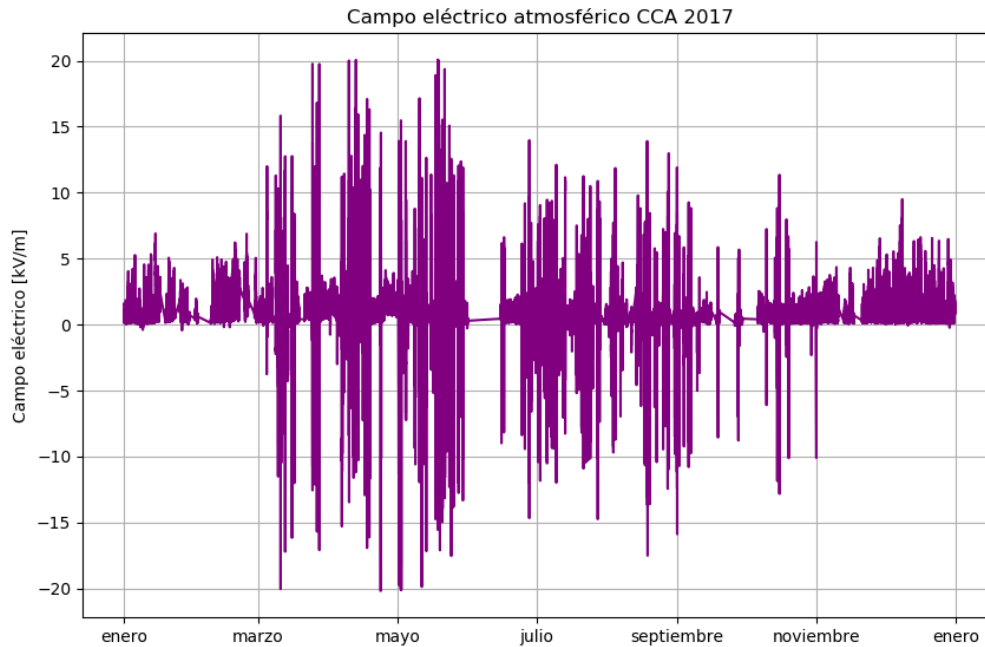


FIGURE 4.5: Campo eléctrico atmosférico para el año 2017 en el CCA

En la figura 4.6 vemos el espectro de amplitudes para el CCA, el cual se presenta bastante limpio y con una sola frecuencia dominante que corresponde a $x = f = 0.0006944[Hz]$. Como sabemos de nuestras relaciones de física, el inverso de la frecuencia $[f]$ es el período $[T]$, el cual nos da el dato de tiempo del ciclo que estamos observando. Para este caso:

$$T = 1/f = 1/0.0006944 = 1440.09[min] = 24.001[h]$$

Quedándonos así un ciclo de 24 horas, el cual corresponde a la curva de Carnegie.

Además, realizamos una gráfica del campo eléctrico atmosférico durante un día promediando los valores del mismo minuto para cada día a lo largo de un año, obteniendo la figura 4.7, donde se observa que la menor actividad se da durante las primeras horas del día para ir aumentando gradualmente y de forma suave hasta alcanzar un máximo a las 12:00 horas tiempo local, siendo esto coherente con la Curva de Carnegie que se observa en todo el mundo a las 18:00 TU. Continuando durante el día observamos como la curva deja de ser suave y pasa a ser más caótica debido que es durante la tarde donde se registran los episodios de tormenta y por lo tanto los valores más altos del campo eléctrico atmosférico, para comenzar a decaer de nuevo un poco antes de las 21:00 horas tiempo local.

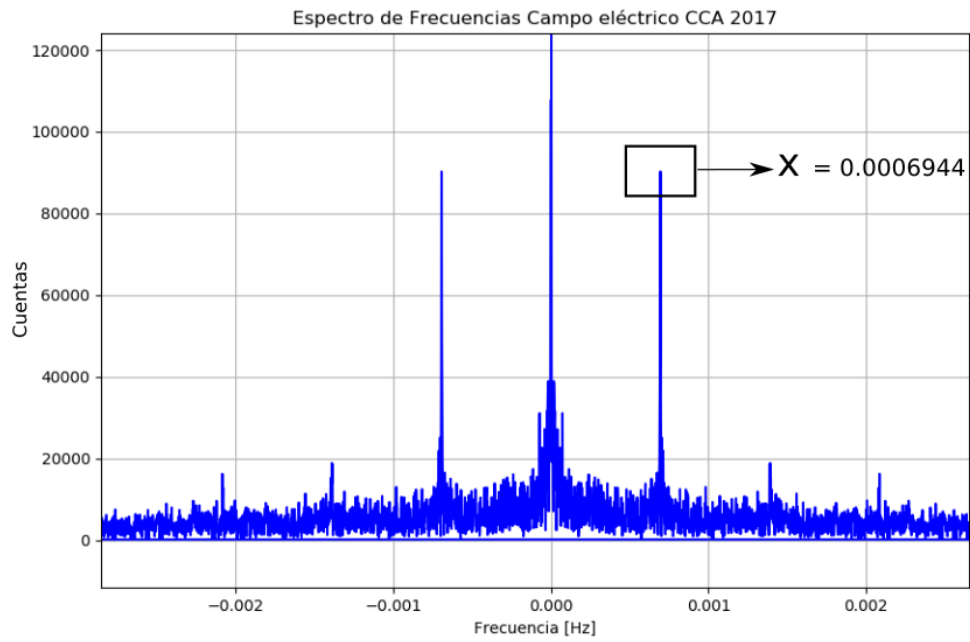


FIGURE 4.6: Espectro de amplitudes para el CCA, mostrando una frecuencia dominante

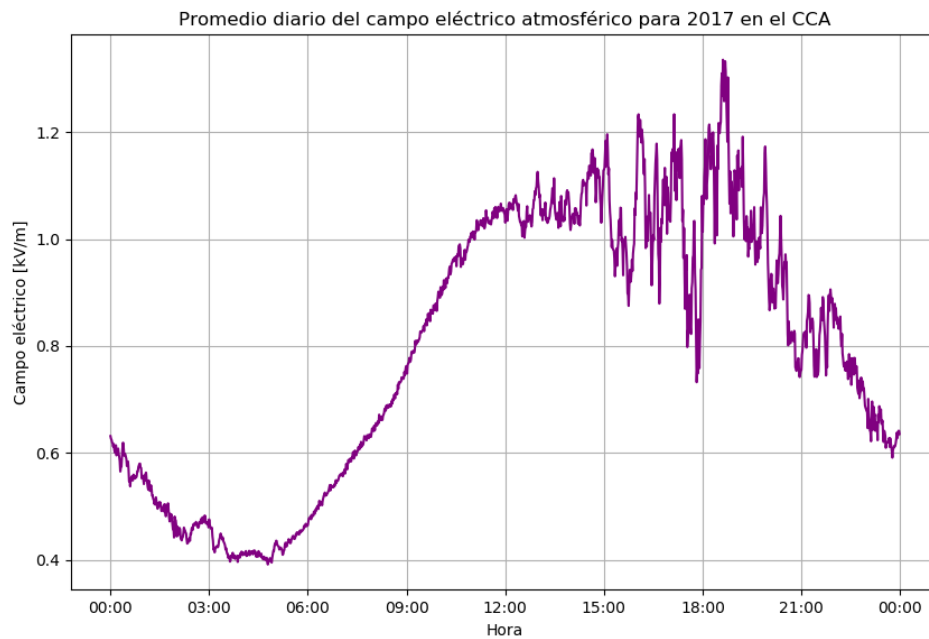


FIGURE 4.7: Campo eléctrico atmosférico diario promedio, en el año 2017 para el CCA

También tenemos una gráfica de varianza por hora (figura 4.8), que se elaboró con los datos promedio de cada hora del día en el año 2017. En ella apreciamos los cambios que se dan antes de llegar a la cima de la curva de Carnegie antes de las 12:00 horas tiempo local y además que las mayores variaciones se dan durante la tarde debido a los episodios de tormenta.

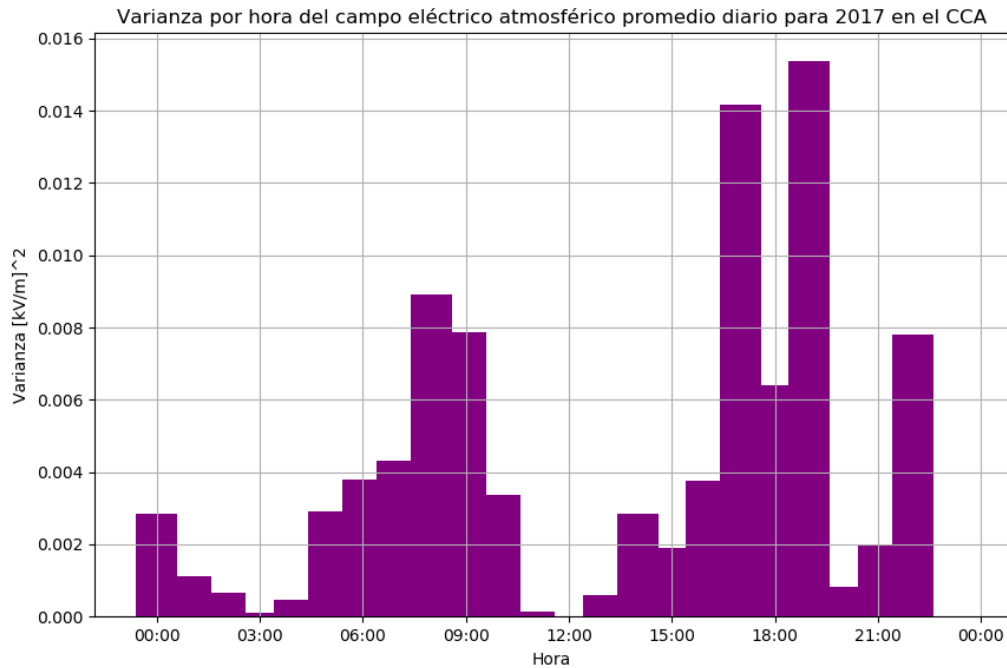


FIGURE 4.8: Varianza por hora de los promedios diarios del campo eléctrico atmosférico, en el año 2017 para el CCA

4.4.2 Espectro de amplitudes del Instituto de Geofísica

En la figura 4.9 observamos la intensidad de campo eléctrico atmosférico para el año 2018 en el Instituto de Geofísica de la UNAM. Vemos que los registros de episodios de tormenta aparecen durante casi todo el año. Esto puede deberse a que no siempre eventos que el Boltek registra como una tormenta corresponden a una tormenta real, tal vez solo nubes electrificadas que pasan muy cerca del equipo. A pesar que se observan dos brechas grandes en los registros para los meses de abril y julio, se decidió tomar este año de datos porque es el único que se encuentra completo de enero a diciembre para esta estación.

En la figura 4.10 vemos el espectro de amplitudes para el Instituto de Geofísica, que aunque contiene un poco más de ruido que el del CCA, sigue apreciándose una sola frecuencia dominante que corresponde a $x = f = 0.0006944[Hz]$. Para esta estación el período corresponde a exactamente el mismo que el CCA, dándonos una validación muy buena del dato de la anterior estación:

$$T = 1/f = 1/0.0006944 = 1440.09[\text{min}] = 24.001[\text{h}]$$

Confirmando así un ciclo de 24 horas de forma muy exacta. Es de notar lo sorprendente que fue obtener el mismo dato incluso usando 7 cifras significativas, ya que el Boltek EFM-100 no es un equipo de alta precisión y estos dos equipos instalados en edificios contiguos presentan registros con ligeras variaciones incluso apreciables a simple vista comparando las gráficas 4.5 y 4.9.

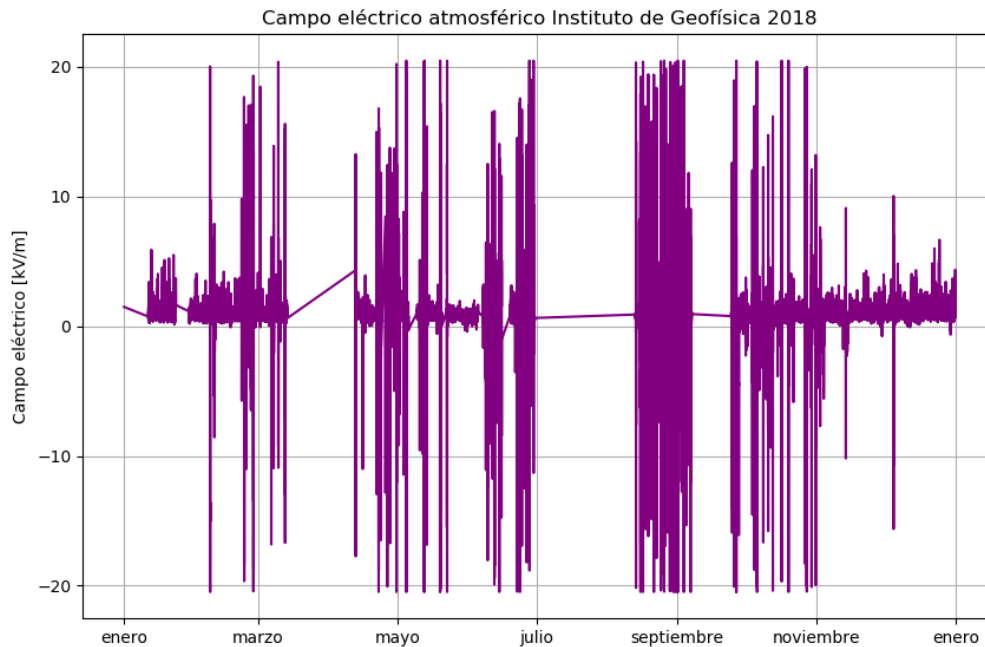


FIGURE 4.9: Campo eléctrico atmosférico para el año 2018 en el Instituto de Geofísica

También en esta estación se procedió a calcular los promedios diarios de un año y mostrarlos en una gráfica de 24 horas (figura 4.11), obteniendo un resultado muy similar al del CCA (figura 4.7), pero con ligeras variaciones debido a que en esta estación se registra en tiempo local que varía entre los horarios de invierno y de verano.

En la figura 4.12 observamos un comportamiento muy parecido de la varianza en los datos con el CCA (figura 4.8), obteniendo cambios durante el incremento del campo cuando se desarrolla la curva de Carnegie y una mayor variación durante los episodios de tormenta de la tarde.

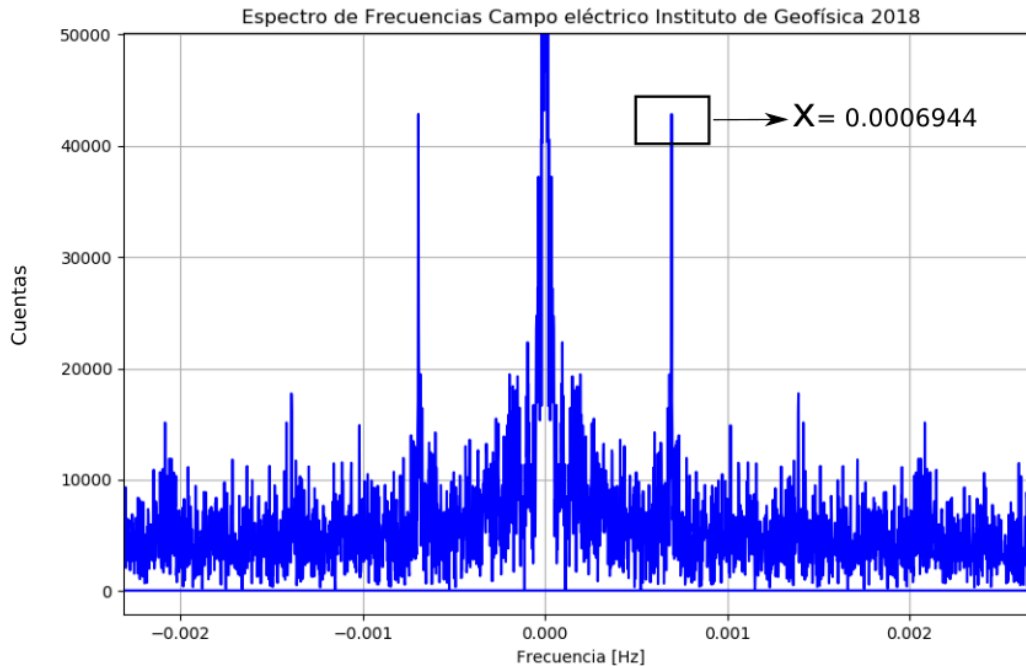


FIGURE 4.10: Espectro de amplitudes para el Instituto de Geofísica, mostrando una frecuencia dominante

4.4.3 Espectro de amplitudes de Juriquilla

En la figura 4.13 observamos la intensidad de campo eléctrico atmosférico para el año 2017 en la estación de la RUOA de Juriquilla. Vemos que los registros de episodios de tormenta aparecen desde casi el inicio del año y van disminuyendo a partir de septiembre. Curiosamente, para los últimos 3 meses del año se observan eventos de campo eléctrico alto pero solo en la parte positiva, por lo que interpretamos que estos no corresponden a episodios de tormenta, ya que las tormentas eléctricas tienen una clara firma, donde los registros del campo varían de polaridad y pueden alcanzar hasta de ± 20 kV/m. (Ver figura 2.19)

En la figura 4.14 vemos el espectro de amplitudes para la estación de Juriquilla, el cual está bastante libre de ruido y presenta dos frecuencias dominantes, aunque la primera con más del doble de cuentas que la segunda. Estas dos frecuencias dominantes corresponden a los valores de $x_1 = f_1 = 0.0006925$ [Hz] y $x_2 = f_2 = 0.0013900$ [Hz]. A la hora de sacar sus inversos, obtenemos:

$$T_1 = 1/f_1 = 1/0.0006925 = 1444.04[\text{min}] = 24.06[\text{h}]$$

$$T_2 = 1/f_2 = 1/0.0013900 = 719.42[\text{min}] = 11.99[\text{h}]$$

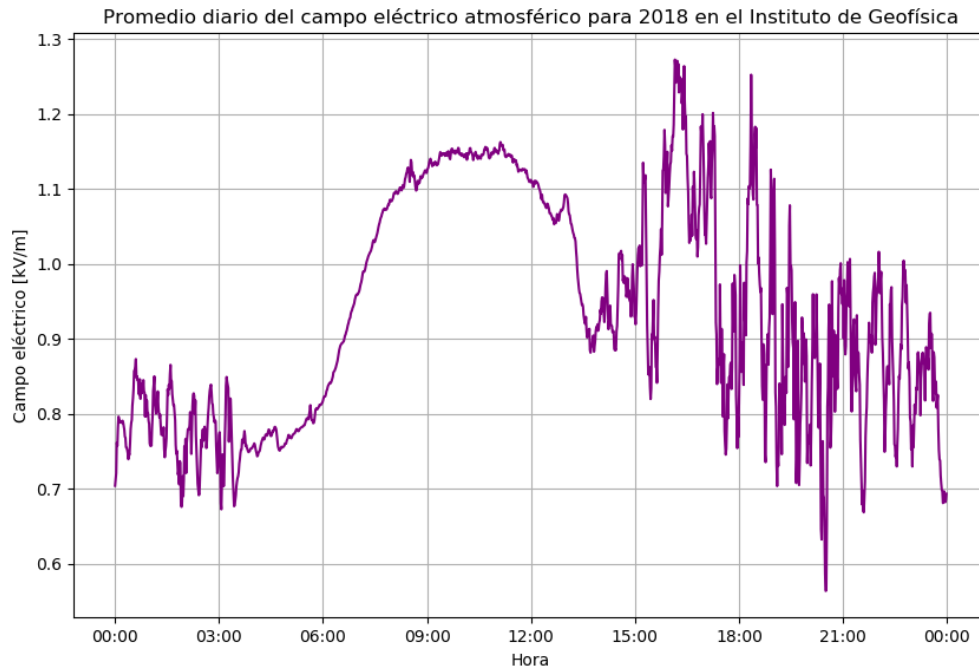


FIGURE 4.11: Campo eléctrico atmosférico diario promedio en el año 2018, para el Instituto de Geofísica

De estos datos observamos que para el caso de la estación de Juriquilla se presenta la esperada curva de Carnegie cada 24 horas, pero observamos otro ciclo, más tenue, el cual se presenta cada 12 horas.

En la figura 4.15 que corresponde al promedio diario de campo eléctrico atmosférico para el año 2017 en Juriquilla, observamos como en las demás estaciones la curva de Carnegie con su cima a las 12:00 horas tiempo local, además de un comportamiento más suave para la parte de las tormentas de la tarde que poco a poco va decayendo hasta llegar a un mínimo unos minutos después de las 04:00 horas tiempo local.

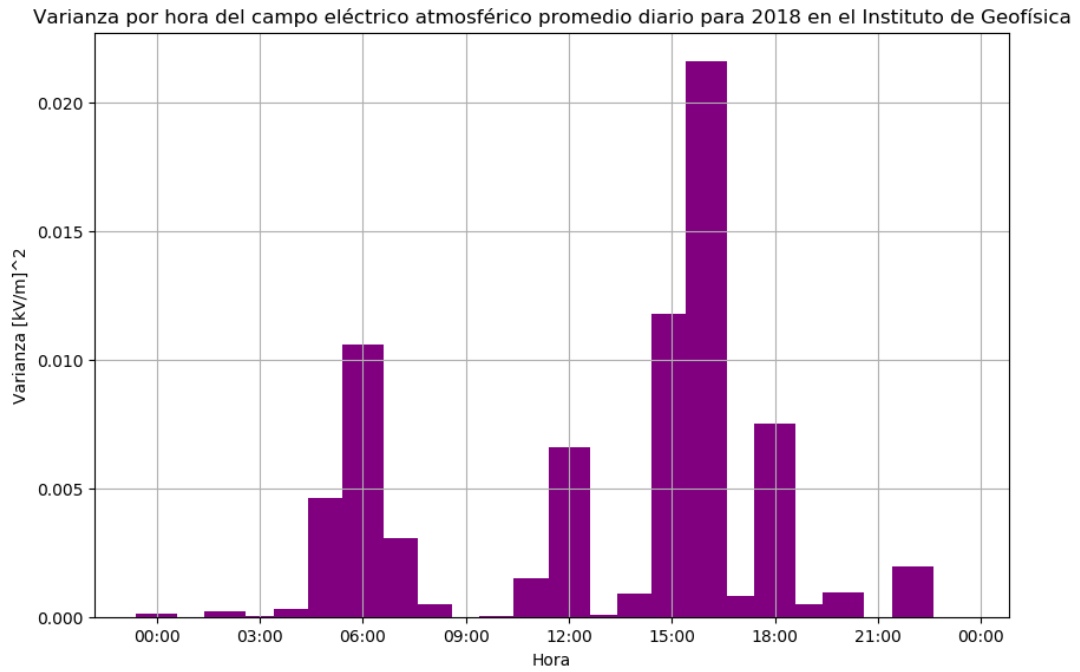


FIGURE 4.12: Varianza por hora de los promedios diarios del campo eléctrico atmosférico, en el año 2018 para el Instituto de Geofísica

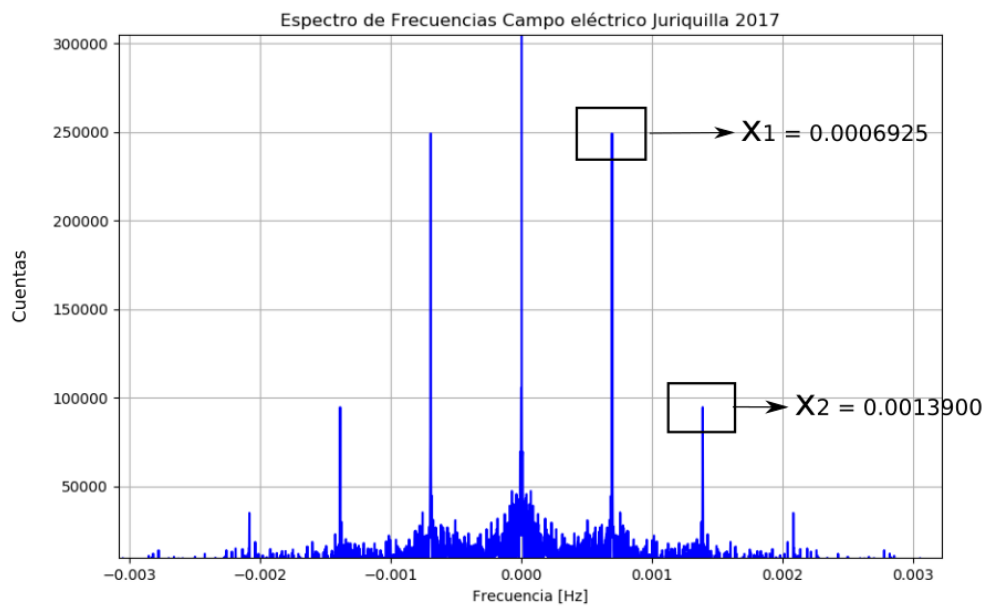


FIGURE 4.14: Espectro de amplitudes para Juriquilla, mostrando dos frecuencias dominantes

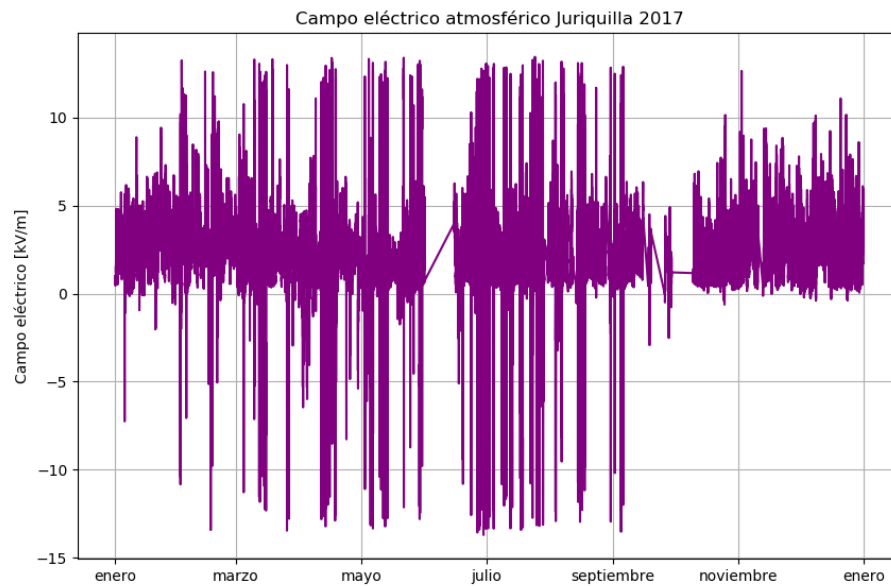


FIGURE 4.13: Campo eléctrico atmosférico para el año 2017 en Juriquilla

En la gráfica de varianza para la estación Juriquilla (figura 4.16) observamos que las mayores variaciones del campo se dan durante el desarrollo de la curva de Carnegie y ya no durante los episodios de tormenta como en las dos estaciones previas.

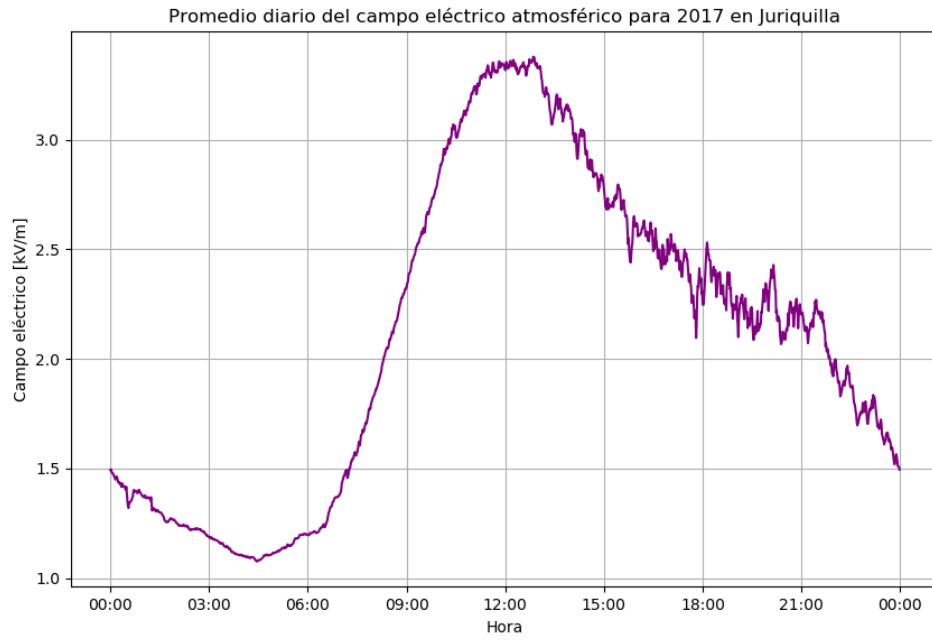


FIGURE 4.15: Campo eléctrico atmosférico diario promedio en el año 2017, para Juriquilla

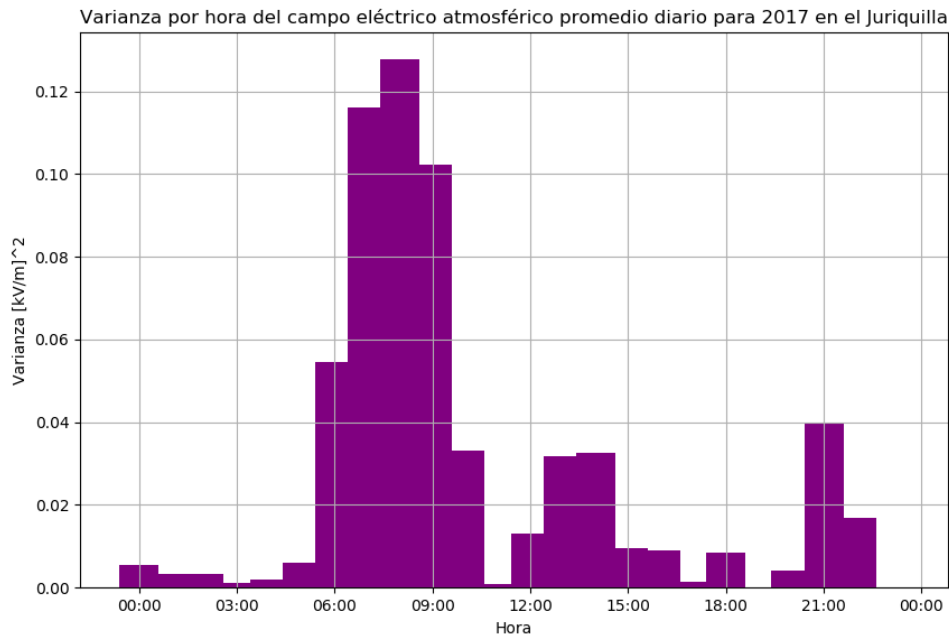


FIGURE 4.16: Varianza por hora de los promedios diarios del campo eléctrico atmosférico, en el año 2017 para Juriquilla

4.4.4 Espectro de amplitudes de Hermosillo

En la figura 4.17 observamos la intensidad de campo eléctrico atmosférico para el año 2016 en la estación de la RUOA de Hermosillo. A simple vista podemos apreciar que a diferencia de las demás estaciones, esta cuenta con menos episodios de tormenta. De hecho, la densidad de episodios de tormenta más considerable solo se encuentra de julio a septiembre, salvo algunas excepciones. Esto debido a que en Hermosillo estos 3 meses son los que acumulan la mayor parte de precipitación de todo el año.

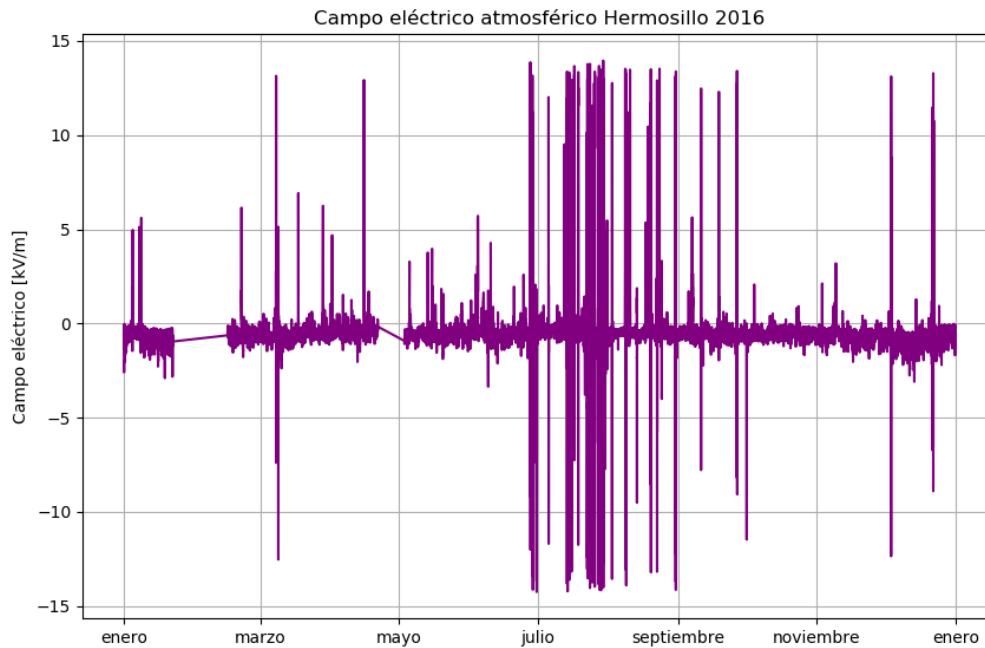


FIGURE 4.17: Campo eléctrico atmosférico para el año 2016 en Hermosillo

En la figura 4.18 vemos el espectro de amplitudes para la estación de Hermosillo, el cual contiene más ruido comparado al de la estación juriquilla, pero presenta los dos mismos ciclos al ver sus frecuencias dominantes. Estas dos frecuencias corresponden a los valores de $x1 = f1 = 0.0006956[Hz]$ y $x2 = f2 = 0.0013889[Hz]$ que al sacar sus inversos obtenemos:

$$T1 = 1/f1 = 1/0.0006945 = 1439.88[min] = 23.998[h]$$

$$T2 = 1/f2 = 1/0.0013889 = 719.99[min] = 11.999[h]$$

De estos datos observamos que al igual que para todas las demás estaciones se presenta la curva de Carnegie cada 24 horas aunque no es apreciable a simple vista en los promedios diarios (figura 4.19), y también podemos confirmar el ciclo de 12 horas descubierto en Juriquilla.

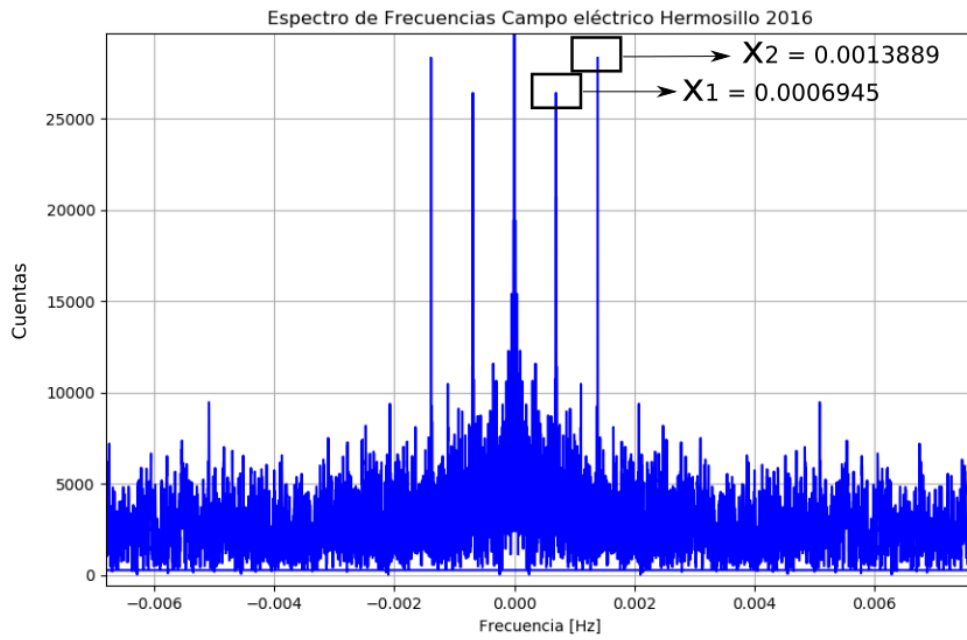


FIGURE 4.18: Espectro de amplitudes para Hermosillo, mostrando dos frecuencias dominantes

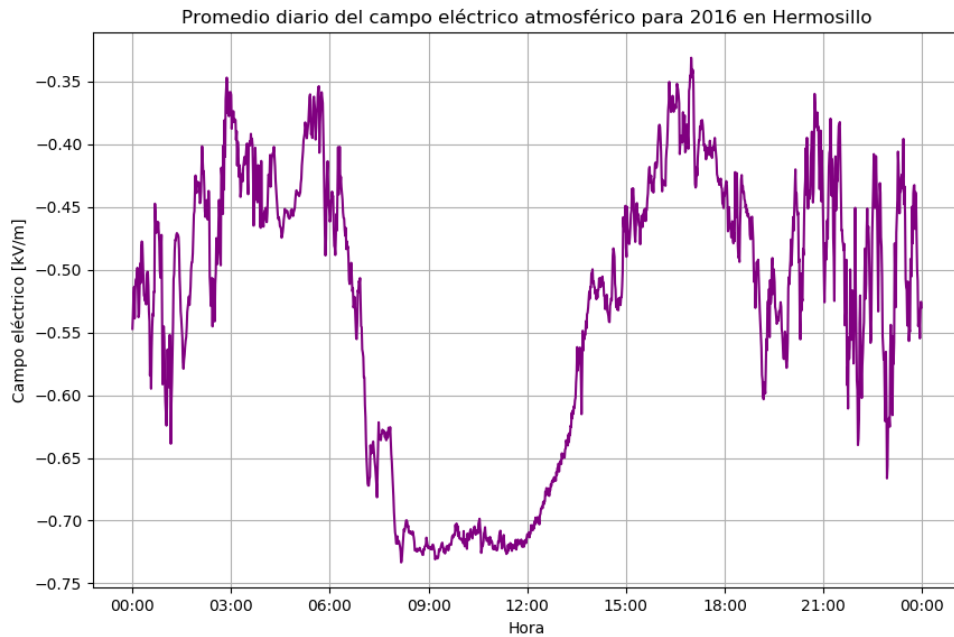


FIGURE 4.19: Campo eléctrico atmosférico diario promedio en el año 2016, para Hermosillo

Una anotación importante es que si una persona estuviera interesada en descartar la curva de Carnegie de los datos para concentrarse en otros ciclos, podría eliminarla en el espectro de frecuencias utilizando un filtro Notch.

En la figura 4.19 observamos la gráfica de promedio diario del campo en la estación Hermosillo durante el año 2016. Se puede observar en el eje Y que los valores del campo son muy pequeños porque en esta estación casi no se registran valores altos de campo debido a la escasez de tormentas en el noroeste de México. Además, cabe mencionar que en esta estación los pocos episodios de tormenta que se registran son mayormente durante las primeras horas del día y algunos después de las 15:00 horas tiempo local, pero sin una clara tendencia como en las demás estaciones.

Además, es muy útil ver la figura 4.20 de la varianza de los datos en Hermosillo para darnos cuenta que es en las primeras horas del día donde hay mayores cambio del campo eléctrico, correspondiendo con los episodios de tormenta en esa zona del país.

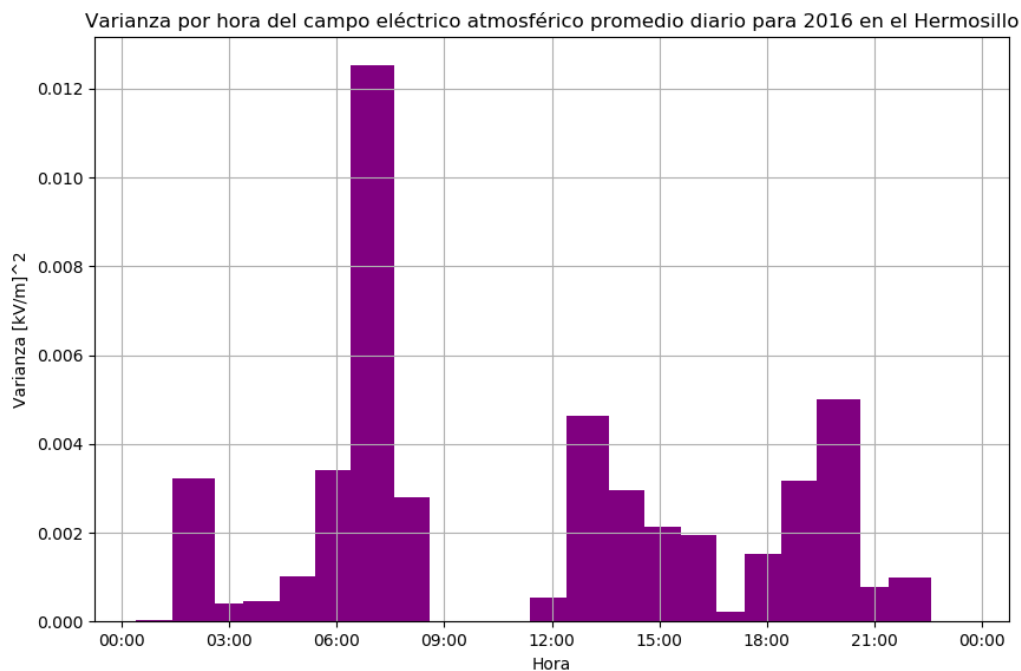


FIGURE 4.20: Varianza por hora de los promedios diarios del campo eléctrico atmosférico, en el año 2016 para Hermosillo

4.5 Análisis L2

Finalmente, lo que se realizó con la base de datos fue una inspección por día, para tener esta base de datos en un nivel de calidad L2, abarcando alrededor de 4,000 gráficas. Lo que buscamos en las gráficas al hacer esto fue identificar los días con pocos datos o completamente vacíos, lecturas fuera de rango (aunque ya sabemos que el nivel L1 se encarga de esto, no se podían descartar errores en su aplicación), o anomalías en estos datos, como saturaciones exageradas.

Por lo que el producto de esta parte fue una hoja de cálculo para cada estación, donde por día se establece un código de color para la calidad de los datos, además de comentarios para la mayor parte de los días. Por otro lado, también se elaboró una gráfica de calidad para cada estación. Estas hojas de cálculo y los scripts usados para sus respectivas gráficas de calidad se encuentran en el repositorio de Github mencionado en la introducción de este capítulo, para que puedan ser actualizados por otros estudiantes.

En la figura 4.21 vemos la estructura de la hoja de cálculo obtenida a partir de la inspección visual de toda la base de datos. Esta contiene el título de la hoja de cálculo, mi nombre y contacto para futuras referencias, así como su fecha de actualización. Para los datos tenemos fecha, número de datos faltantes por día, una escala arbitraria de calidad donde el 0, 2, 5 y 10 corresponden a los colores gris, rojo, amarillo y verde, respectivamente. Estos se basan principalmente en la cantidad de datos para el día, ya que por lo general, si tenemos datos completos se considera como un día con datos de calidad alta.

Los umbrales de cantidad de datos para esta escala son:

- ▶ $0 \leq \text{datos} < 300$, corresponde a un día sin datos, es decir color gris.
- ▶ $300 \leq \text{datos} < 600$, color rojo.
- ▶ $600 \leq \text{datos} < 1000$, color amarillo.
- ▶ $1000 \leq \text{datos} \leq 1440$, color verde.

Para los días con saturación exagerada se consideró una calidad media, y no se encontraron días donde se excediera el rango de campo eléctrico de -20 a 20 [kV/m].

En la figura 4.22 se muestra un ejemplo de los comentarios que se pueden encontrar en el documento, entre ellos tenemos:

- ▶ Tiempo calmo
- ▶ Traza de rayo (también se puede especificar si pequeña o grande)
- ▶ Episodio de tormenta (a veces con su duración, o a veces se especifica si pequeño o grande)
- ▶ Evento de campo eléctrico alto (positivo o negativo)
- ▶ Ruido por lluvia

- Casi sin datos
- Gran saturación
- La cantidad de horas sin datos
- Si no hay comentario, es porque no hay mucho que destacar en ese día en particular

	A	B	C	D	E	F	G	H	I
1	Tabla de calidad para datos de intensidad de campo eléctrico atmosférico en el Centro de Ciencias de la Atmósfera, UNAM.								
2	Elaborado por Alejandro Argueta, F-UNAM alejo.tigres@yahoo.com Actualizado: 13 de diciembre 2019								
3	Fecha	No. datos faltantes	Calidad	Escala	Comentarios				
4	2015-02-01	1440		0	Sin datos			0	
5	2015-02-02	1440		0	Calidad alta			10	
6	2015-02-03	1440		0	Calidad media			5	
7	2015-02-04	1440		0	Calidad baja			2	
8	2015-02-05	1440		0					
9	2015-02-06	1440		0					
10	2015-02-07	1440		0					
11	2015-02-08	1440		0					
12	2015-02-09	1440		0					
13	2015-02-10	1440		0					
14	2015-02-11	1440		0					
15	2015-02-12	1440		0					
16	2015-02-13	1440		0					
17	2015-02-14	1440		0					
18	2015-02-15	1440		0					
19	2015-02-16	1440		0					
20	2015-02-17	1096		10	Tiempo calmo				
21	2015-02-18	12		10	Pequeña traza de rayo				
22	2015-02-19	0		10					
23	2015-02-20	7		10					
24	2015-02-21	11		10					
25	2015-02-22	9		10					
26	2015-02-23	2		10					
27	2015-02-24	9		10					
28	2015-02-25	8		10					
29	2015-02-26	15		10					

FIGURE 4.21: Estructura de la hoja de cálculo realizada a partir de la inspección visual de los datos

La diferencia entre un episodio de tormenta y un evento de campo eléctrico alto, reside en que en este último no hay inversiones de polaridad. Estos comentarios pueden ser importantes para una persona que solo esté interesada en días con tormenta, para lo cual ya no tendría que revisar la base de datos completa buscando estos días, sino solo buscar las casillas que correspondan al comentario.

Otro aspecto importante que se identificó en la inspección es que para la estación de Hermosillo el tiempo calmo que normalmente estaría por debajo de 0.1 kV/m se encuentra en negativo, esto puede deberse a un pequeño error de calibración del equipo.

Además, para las estaciones del CCA y el Instituto de Geofísica, las tormentas suelen ocurrir alrededor de las 6:00 pm hora local, a diferencia de Hermosillo donde las tormentas no tienen una regularidad, aunque algunas suceden durante la madrugada.

	A	B	C	D	E	F	G	H	I	J	K
46	2015-03-15	3		10							
47	2015-03-16	0		10							
48	2015-03-17	17		5	Falta la segunda mitad del día, a pesar que el código arroja solo "17 datos faltantes"						
49	2015-03-18	7		5	Faltan las primeras 9 horas de datos						
50	2015-03-19	0		10	Episodio de tormenta						
51	2015-03-20	4		10	Trazas de rayos						
52	2015-03-21	0		10	Ruido por lluvia						
53	2015-03-22	0		10	Episodio de tormenta						
54	2015-03-23	0		10	Episodio de tormenta (12 horas)						
55	2015-03-24	1		10	Episodio de tormenta						
56	2015-03-25	0		10	Episodio de tormenta						
57	2015-03-26	0		10	Episodio de tormenta						
58	2015-03-27	2		10	Episodio de tormenta						
59	2015-03-28	2		10	Episodio de tormenta						
60	2015-03-29	9		10	Episodio de tormenta						
61	2015-03-30	0		10							
62	2015-03-31	0		10	Traza de rayo						
63	2015-04-01	0		10	Episodio de tormenta						
64	2015-04-02	1		10	Episodio de tormenta, posible nube justo arriba del detector						
65	2015-04-03	0		10	Tiempo calmo						
66	2015-04-04	0		10	Tiempo calmo						
67	2015-04-05	0		10							
68	2015-04-06	0		10	Subida positiva del campo						
69	2015-04-07	0		10	Tiempo calmo						
70	2015-04-08	0		10	Episodio de tormenta						
71	2015-04-09	0		10	Traza de rayo						
72	2015-04-10	0		10	Trazas de rayos						
73	2015-04-11	0		10	Episodio de tormenta						

FIGURE 4.22: Ejemplo de cómo se aprecian los comentarios por día en la hoja de cálculo

En Hermosillo, nunca se rebasa la barrera de los 13 [kV/m], posiblemente debido al plug atenuador con el que cuenta este equipo. Por otro lado, en Juriquilla casi no se aprecian días con tiempo calmo.

Por último, en las figuras 4.23, 4.24, 4.25 y 4.26 tenemos las gráficas de la calidad de los datos para cada estación. Abarcando la totalidad de datos de cada estación, y haciendo uso de la misma escala arbitraria de colores para la calidad, donde un punto corresponde a un día.

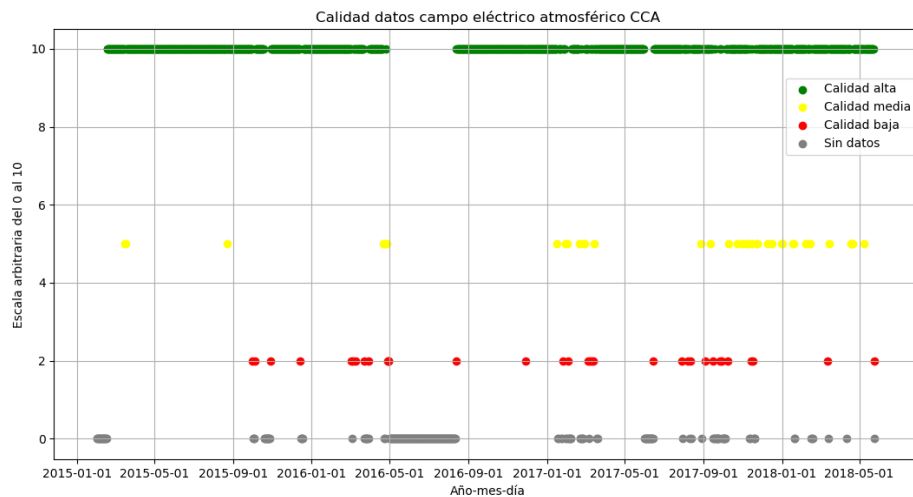


FIGURE 4.23: Gráfica de calidad para la estación del CCA

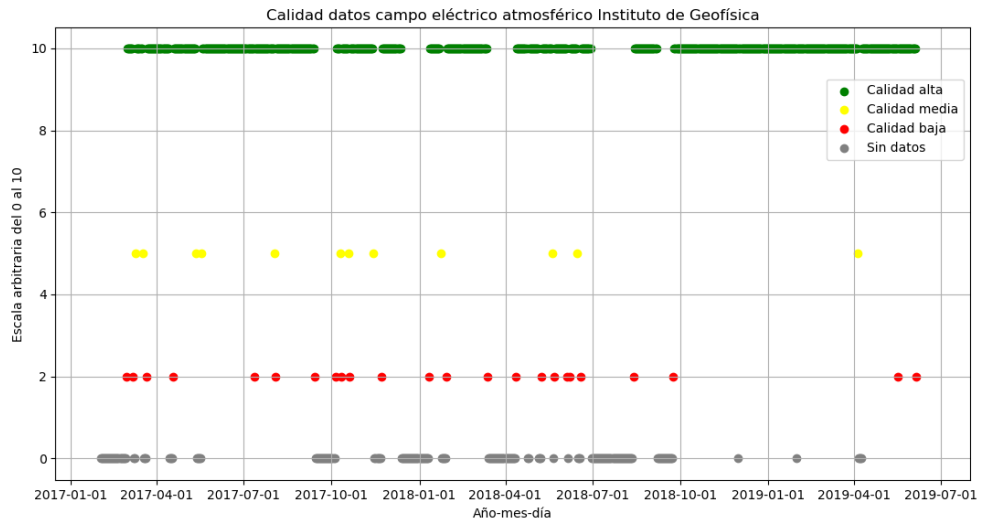


FIGURE 4.24: Gráfica de calidad para la estación del Instituto de Geofísica

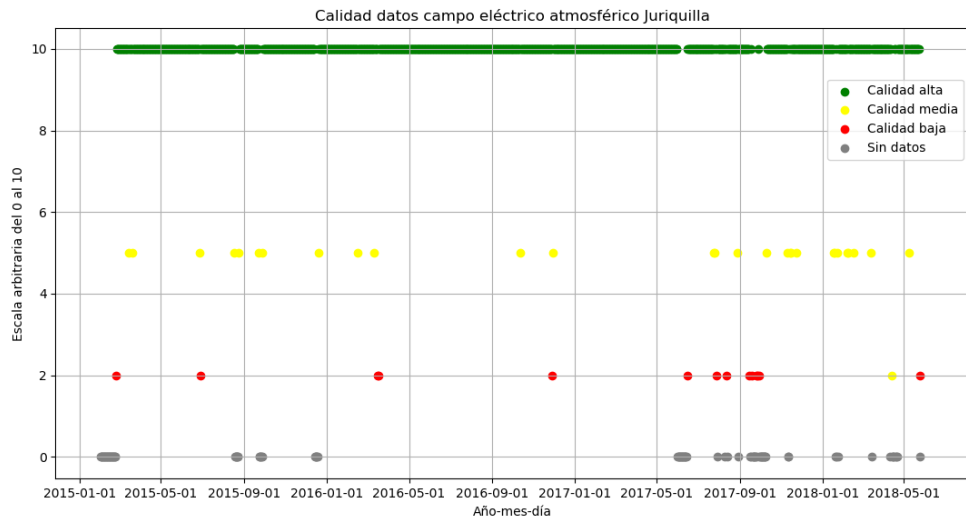


FIGURE 4.25: Gráfica de calidad para la estación de Juriquilla

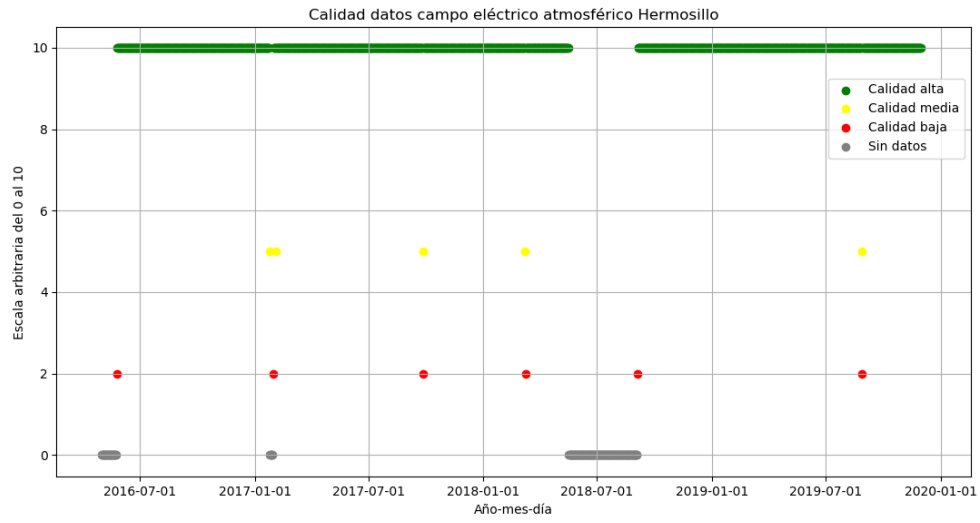


FIGURE 4.26: Gráfica de calidad para la estación de Hermosillo

CONCLUSIONES

Es de mucha importancia el conocer cómo se tratan los datos de intensidad de campo eléctrico atmosférico, para tener una base de datos con calidad óptima para la realización de investigaciones científicas donde éstos sean una de las variables a considerar.

En esta tesis, mencionamos las teorías actuales de los orígenes de los campos eléctricos en la atmósfera, que son de origen no inductivo y vienen junto con las nubes de tormenta. Conocimos en qué consiste la Red Universitaria de Observatorios Atmosféricos, la cual es un proyecto de recolección, tratamiento y análisis de diversas variables atmosféricas, a lo largo de todo México.

También, describimos a profundidad el funcionamiento de los detectores de campos eléctricos Boltek EFM-100, que resultan ser bastante confiables para la medición de la intensidad de campo eléctrico. Además, vimos cómo se efectuó la instalación de uno de estos equipos en el volcán Sierra Negra en Puebla, el cual está mediendo datos con un valor científico muy alto, dado que es posiblemente el Boltek EFM-100 instalado a mayor altura en el mundo, a 4,590 m.s.n.m., y se encuentra instalado en el mismo sitio que otros dos observatorios de rayos cósmicos.

Estudiamos el estado en el que se encontraba la base de datos obtenida con estos detectores Boltek EFM-100 y se desarrollaron scripts de python para su visualización y tratamiento. Se logró aplicar los niveles de control de calidad L0 y L1 a una base de datos en estado crudo, que no pertenece a la RUOA y se llevó toda la base de datos a un nivel de control de calidad L2, el cual implica una inspección de todos los datos.

Se comprobó que los equipos miden de forma precisa, ya que hicimos correlaciones entre los datos obtenidos por dos equipos instalados en el mismo sitio geográfico y confirmamos en el espectro de frecuencias de los datos que la Curva de Carnegie se ve de forma exacta cada 24 horas, independientemente de la ubicación o condiciones de estos equipos. Además, se observó otro ciclo de 12 horas, del cual aun se desconoce su origen, pero puede ser estudiado para discernir entre ruido o algún proceso físico externo. También comprobamos la curva de Carnegie alrededor de las 12:00 tiempo local, por medio de gráficas de promedios diurnos y visualizamos la varianza de los datos de forma promedio durante el día.

Es por todo lo anterior que podemos afirmar que la base de datos de intensidad de campo eléctrico atmosférico de la RUOA se encuentra en óptimas condiciones y lista para su uso en investigaciones meteorológicas. Por otro lado, con la experiencia y productos

obtenidos, se puede continuar con el tratamiento de los nuevos datos que se adquieran en el tiempo o en nuevas estaciones.

Por último, este trabajo facilita a cualquier usuario acceder de forma fácil a un sitio web libre, donde se encuentran toda la base de datos, los productos obtenidos y los códigos utilizados debidamente comentados.

APÉNDICE A: CÓDIGOS DE PYTHON

rawdata_graph.py

```
#UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
#Argueta Hernández, Fidel Alejandro alejo_tigres@yahoo.com
#Asesor: Dr. Luis Xavier González Méndez
#Programa que grafica los datos crudos de campo eléctrico atmosférico del EFM-100
import pandas as pd
import matplotlib.pyplot as plt
import matplotlib.dates as md
#Leemos el archivo efm
data=pd.read_csv('Graphs raw data IGEOF/Data Source 1-04162018.efm')
#Damos nombre a las columnas
data.columns = ['date', 'field', 'v1']
y = data['field']
#Convertimos a datetime la columna de tiempo (que inicialmente solo es object)
data['date'] = pd.to_datetime(data['date'], format='%H : %M : %S ')

fig, ax = plt.subplots(figsize=(10,10))
ax.plot(data['date'], y, color = 'purple')
#Establecemos el formato H:M:S para el eje x
xfmt = md.DateFormatter('%H : %M : %S ')
ax.xaxis.set_major_formatter(xfmt)
ax.set_title('Campo eléctrico atmosférico 16-04-2018')
ax.set_xlabel('Hora')
ax.set_ylabel('Campo eléctrico [kV/m]')
plt.grid()
plt.show()
```

Average4rawdata.py

```
#UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
```



```

#Argueta Hernández, Fidel Alejandro alejo_tigres@yahoo.com
#Asesor: Dr. Luis Xavier González Méndez
#Programa que saca promedio por minuto de los datos crudos de campo
eléctrico
import pandas as pd
data = pd.read_csv('Graphs raw data IGEOF/Data Source 1-06052019.
efm')
#Leemos el archivo efm. If wrong won't run
data.columns = ['time', 'field', 'v1'] #Damos nombre a las columnas
data['dates'] = "2019-06-05" #Creamos columna con fecha del día. If
wrong we'll see in terminal
data['datetime'] = data['dates'] + ' ' + data['time'] #Sumamos
columnas de fecha y de hora
data=data[['datetime', 'field']] #Nos quedamos con las columnas
que nos interesan
data['datetime'] = pd.to_datetime(data['datetime'],
format='%Y-%m-%d %H:%M:%S')
#Covertimos a formato datetime
data = data.set_index(data['datetime']) #Trasladamos a Index
columna datetime
data=data['field'].resample('1Min').mean() #Promediamos por minuto
print(data.head()) #Imprimimos para verificar correcto
funcionamiento
#Creamos nuevo archvivo. If wrong, merge won't work
data.to_csv('Data E Geofisica avg min/June2019/EFavgmin06052019.csv
', encoding='utf-8', header=False)

```

GraphDayRawAvgMin.py

```

#UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
#Argueta Hernández, Fidel Alejandro alejo_tigres@yahoo.co
#Asesor: Dr. Luis Xavier González Méndez
#Programa que grafica datos crudos de un día promediados por minuto
import pandas as pd
import matplotlib.pyplot as plt
import matplotlib.dates as md
#Leemos el archivo efm
data = pd.read_csv('Data E Geofisica avg min/May2019/
EFavgmin05202019.csv')
#Damos nombre a las columnas
data.columns = ['datetime', 'field']
y = data['field']

```

```
#Convertimos a datetime la columna de tiempo (que inicialmente solo
    es object)
data['datetime'] = pd.to_datetime(data['datetime'], format='%Y -
    %m - %d %H : %M : %S ')

fig, ax = plt.subplots(figsize=(10,10))
ax.plot(data['datetime'], y, color = 'purple')
#Establecemos el formato H:M:S para el eje x
xfmt = md.DateFormatter('%H : %M ')
ax.xaxis.set_major_formatter(xfmt)
ax.set_title('Campo eléctrico atmosférico 19-05-2018')
ax.set_xlabel('Hora')
ax.set_ylabel('Campo eléctrico [kV/m]')
plt.grid()
plt.show()
```

GraphMonthRawAvgMin.py

```
#UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
#Argueta Hernández, Fidel Alejandro alejo_tigres@yahoo.co
#Asesor: Dr. Luis Xavier González Méndez
#Programa que grafica datos crudos de un mes promediados por minuto
import pandas as pd
import matplotlib.pyplot as plt
import matplotlib.dates as md
data = pd.read_csv('Data E Geofisica avg min/Geof/raw/
    EFRawAvgMinJuly2017.csv', skiprows=[0,2,3])
y = data['E_Avg']
data['TIMESTAMP'] = pd.to_datetime(data['TIMESTAMP'], format='%Y
    - %m - %d %H : %M : %S ')
print(data.head())
fig, ax = plt.subplots(figsize=(10,10))
ax.plot(data['TIMESTAMP'], y, color = 'purple')
xfmt = md.DateFormatter('%d %H : %M ')
ax.xaxis.set_major_formatter(xfmt)
ax.set_title('Campo eléctrico atmosférico Julio 2017')
ax.set_xlabel('Día Hora:Minuto')
ax.set_ylabel('Campo eléctrico [kV/m]')
plt.grid()
plt.show()
```

mergedata.py

```
#UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
#Argueta Hernández, Fidel Alejandro alejo_tigres@yahoo.com
#Asesor: Dr. Luis Xavier González Méndez
#Programa que une datos diarios o mensuales de campo eléctrico
    promediados
#Y les da el formato crudo de los dataloggers de la RUOA
import pandas as pd
import shutil
#Creamos nuevo archivo mensual a partir de datos diarios
#with open('Data E Geofisica avg min/EFRawAvgMw.csv', 'wb') as wfd:
with open('/home/alejandro/MEGA/MEGAsync/Python codes/Fourier
    Analysis/2017-jqro_minuto_L1-cael.csv', 'wb') as wfd:
    for f in ([
        # 'Data E Geofisica avg miw/EFavgmin06012019.csv',
        # 'Data E Geofisica avg miw/EFavgmin06022019.csv',
        # 'Data E Geofisica avg miw/EFavgmin06032019.csv',
        # 'Data E Geofisica avg miw/EFavgmin06042019.csv',
        # 'Data E Geofisica avg miw/EFavgmin06052019.csv',
        # 'Data E Geofisica avg miw/EFavgmin05062019.csv',
        # 'Data E Geofisica avg miw/EFavgmin05072019.csv',
        # 'Data E Geofisica avg miw/EFavgmin05082019.csv',
        # 'Data E Geofisica avg miw/EFavgmin05092019.csv',
        # 'Data E Geofisica avg miw/EFavgmin05102019.csv',
        # 'Data E Geofisica avg miw/EFavgmin05112019.csv',
        # 'Data E Geofisica avg miw/EFavgmin05122019.csv',
        # 'Data E Geofisica avg miw/EFavgmin05132019.csv',
        # 'Data E Geofisica avg miw/EFavgmin05142019.csv',
        # 'Data E Geofisica avg miw/EFavgmin05152019.csv',
        # 'Data E Geofisica avg miw/EFavgmin05162019.csv',
        # 'Data E Geofisica avg miw/EFavgmin05172019.csv',
        # 'Data E Geofisica avg miw/EFavgmin05182019.csv',
        # 'Data E Geofisica avg miw/EFavgmin05192019.csv',
        # 'Data E Geofisica avg miw/EFavgmin05202019.csv',
        # 'Data E Geofisica avg miw/EFavgmin05212019.csv',
        # 'Data E Geofisica avg miw/EFavgmin05222019.csv',
        # 'Data E Geofisica avg miw/EFavgmin05232019.csv',
        # 'Data E Geofisica avg miw/EFavgmin05242019.csv',
        # 'Data E Geofisica avg miw/EFavgmin05252019.csv',
        # 'Data E Geofisica avg miw/EFavgmin05262019.csv',
        # 'Data E Geofisica avg miw/EFavgmin05272019.csv',
        # 'Data E Geofisica avg miw/EFavgmin05282019.csv',
```

```

# 'Data E Geofisica avg miw/EFavgmin05292019.csv ',
# 'Data E Geofisica avg miw/EFavgmin05302019.csv ',
# 'Data E Geofisica avg miw/EFavgmin05312019.csv ',
#Con lo siguiente vamos a unir los archivos de 2018
'Data L1/Juriquilla/2017-01-jqro_minuto_L1_cael.csv ',
'Data L1/Juriquilla/2017-02-jqro_minuto_L1_cael.csv ',
'Data L1/Juriquilla/2017-03-jqro_minuto_L1_cael.csv ',
'Data L1/Juriquilla/2017-04-jqro_minuto_L1_cael.csv ',
'Data L1/Juriquilla/2017-05-jqro_minuto_L1_cael.csv ',
'Data L1/Juriquilla/2017-06-jqro_minuto_L1_cael.csv ',
'Data L1/Juriquilla/2017-07-jqro_minuto_L1_cael.csv ',
'Data L1/Juriquilla/2017-08-jqro_minuto_L1_cael.csv ',
'Data L1/Juriquilla/2017-09-jqro_minuto_L1_cael.csv ',
'Data L1/Juriquilla/2017-10-jqro_minuto_L1_cael.csv ',
'Data L1/Juriquilla/2017-11-jqro_minuto_L1_cael.csv ',
'Data L1/Juriquilla/2017-12-jqro_minuto_L1_cael.csv ']) :
    with open(f, 'rb') as fd:
        shutil.copyfileobj(fd, wfd)
        #Abrimos archivo recién creado para darle formato
        #data = pd.read_csv('/Fourier Analysis/2018-
            geof_minuto_l1_cael.csv ')
# data.columns = ['datetime', 'field'] #Damos nombres a las
    columnas
# #data['datetime'] = '' + data['datetime'] + '' #Ponemos
    comillas al datetime
# data['index'] = data.index #Creamos índice
# data = data[['datetime', 'index', 'field']] #Reorganizamos
    columnas
# print(data) #Imprimimos datos para visualizar correcto
    funcionamiento
# #Sobreescribimos archivo
# data.to_csv('Data E Geofisica avg min/EFRawAvgMw.csv', encoding='
    utf-8', index=False)

```

namecolumns.py

```

#UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
#Argueta Hernández, Fidel Alejandro alejo_tigres@yahoo.com
#Asesor: Dr. Luis Xavier González Méndez
#Programa que les da nombre a las columnas de los datos crudos del
    Boltek de Geofísica
#Y reemplaza las comillas

```

```

import pandas as pd
import csv
data = pd.read_csv('Data E Geofisica avg min/Geof/raw/
    EFRawAvgMinFebruary2017.csv',)
data.columns = ['TIMESTAMP', 'RECORD', 'E_Avg']
data['TIMESTAMP'] = data['TIMESTAMP'].str.replace('\"', '')
print(data.head())
#Sobreescribimos archivo
data.to_csv('Data E Geofisica avg min/Geof/raw/
    EFRawAvgMinFebruary2017.csv', encoding='utf-8', index=False)

```

L1datamonth_graph.py

```

#UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
#Argueta Hernández, Fidel Alejandro alejo_tigres@yahoo.com
#Asesor: Dr. Luis Xavier González Méndez
#Programa que grafica los datos procesados L1 de campo eléctrico
    atmosférico del EFM-100
import pandas as pd
import matplotlib.pyplot as plt
import matplotlib.dates as md
#Leemos el archivo csv
#data = pd.read_csv('L2 analysis/CCA/2017-01-unam_minuto_L1_cael.
    csv', skiprows=(0,1,2,3,4,5,6))
#Para graficar ano
data = pd.read_csv('Fourier Analysis/2017-
    unam_minuto_L1_cael_filltrend.csv', skiprows=(0,1,2,3,4,5,6))
#Nombramos las columnas
data.columns = ['datetime', 'field']
y= data['field']
#Convertimos a datetime la columna de tiempo (que inicialmente solo
    es object)
x = pd.to_datetime(data['datetime'], format='%Y-%m-%d %H :
    %M : %S ')

fig, ax = plt.subplots(figsize=(10,10))
ax.plot(x, y, color = 'purple')
#Establecemos el formato de días y luego H:M:S para el eje x
#xfmt = md.DateFormatter('%d %H : %M ')
#Establecemos el formato de mes-día para cuando graficamos ano
xfmt = md.DateFormatter('%m-%d ')
ax.xaxis.set_major_formatter(xfmt)

```

```

ax.set_title('Campo eléctrico atmosférico CCA 2017')
#Título eje x cuando graficamos por mes
#ax.set_xlabel('Día Hora:Minuto')
#Título eje x cuando graficamos ano
ax.set_xlabel('Mes-Día')
ax.set_ylabel('Campo eléctrico [kV/m]')
plt.grid()
plt.show()

```

correlationsL1.py

```

#UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
#Argueta Hernández, Fidel Alejandro alejo_tigres@yahoo.com
#Asesor: Dr. Luis Xavier González Méndez
#Programa que grafica en una misma ventana los datos diarios de los
    Boltek del Instituto de Geofísica
#Y del Centro de Ciencias de la Atmósfera para buscar la correlación
    n entre ellos
import pandas as pd
import matplotlib.pyplot as plt
import matplotlib.dates as md
import numpy as np
#Leemos el archivo csv
data = pd.read_csv('/home/alejandro/MEGA/MEGAsync/Python codes/
    Correlations/CorrL1/ECCA2018-04-14-L1.csv', skiprows
    =(0,1,2,3,4,5,6,7))
data2 = pd.read_csv('/home/alejandro/MEGA/MEGAsync/Python codes/
    Correlations/CorrL1/Egeof2018-04-14-L1.csv', skiprows
    =(0,1,2,3,4,5,6,7))
#Damos nombre a las columnas
data.columns = ['TIMESTAMP', 'field']
data2.columns = ['TIMESTAMP', 'field']
#print (data2.head())
y = data['field']
z = data2['field']

#Calculamos el Coeficiente de correlación
c = np.corrcoef(y, z)[0, 1]
print(c)

#Convertimos a datetime la columna de tiempo (que inicialmente solo
    es object)

```

```

data['TIMESTAMP'] = pd.to_datetime(data['TIMESTAMP'], format='%Y
    - %m - %d %H : %M : %S ')
fig, ax = plt.subplots(figsize=(10,10))
ax.plot(data['TIMESTAMP'], y, color = 'purple', label="CCA")
ax.plot(data['TIMESTAMP'], z, color = 'red', label="I.Geofísica")
#Establecemos el formato H:M:S para el eje x
xfmt = md.DateFormatter('%H : %M ')
ax.xaxis.set_major_formatter(xfmt)
ax.text(0.95, 0.01, 'r=0.465',
        verticalalignment='bottom', horizontalalignment='right',
        transform=ax.transAxes,
        color='black', fontsize=25)
ax.set_title('Campo eléctrico atmosférico 14-04-2018')
ax.set_xlabel('Hora')
ax.set_ylabel('Campo eléctrico [kV/m]')
plt.legend()
plt.grid()
plt.show()

```

filldatatrend.py

```

#UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
#Argueta Hernández, Fidel Alejandro alejo_tigres@yahoo.com
#Asesor: Dr. Luis Xavier González Méndez
#Programa que realiza interpolación lineal para datos faltantes
import pandas as pd
import matplotlib.pyplot as plt
import matplotlib.dates as md
import numpy as np
#Leemos el archivo csv
data = pd.read_csv('/home/alex/Escritorio/MEGAsync/Python codes/
    Fourier Analysis/2018-geof_minuto_L1-cael_filltrend.csv',
skiprows=(0,1,2,3,4,5,6,7))
data.columns = ['TIMESTAMP', 'field']
data['field'].interpolate(method='linear', inplace=True)
data.to_csv('/home/alex/Escritorio/MEGAsync/Python codes/Fourier
    Analysis/2018-geof_minuto_L1-cael_filltrend.csv', encoding='utf
    -8', index=False)

```

fft.py

```

#UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

```

```

#Argueta Hernández, Fidel Alejandro alejo_tigres@yahoo.com
#Asesor: Dr. Luis Xavier González Méndez
#Programa que realiza la Transformada rápida de Fourier a los datos
  de campo eléctrico atmosférico
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
from scipy import fftpack
from scipy.stats import pearsonr
#Leemos el archivo csv
data = pd.read_csv('Fourier Analysis/2017-
  jgro_minuto_L1_cael_filltrend.csv', skiprows=(0,1,2,3,4,5,6))
#Nombramos las columnas
data.columns = ['datetime', 'field']
y= data['field']
Y    = np.fft.fft(y)
freq = np.fft.fftfreq(len(y))
fig, ax = plt.subplots(figsize=(10,10))
ax.plot(freq, np.abs(Y), color = 'blue')
ax.set_title('Espectro de Frecuencias Campo eléctrico CCA 2017')
ax.set_xlabel('Frecuencia [Hz]')
plt.grid()
plt.show()

```

Graphqualitytable.py

```

#UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
#Argueta Hernández, Fidel Alejandro alejo_tigres@yahoo.com
#Asesor: Dr. Luis Xavier González Méndez
#Programa que crea una gráfica de calidad L2 de los datos por
  estación a partir de un CSV
import pandas as pd
import matplotlib.pyplot as plt
import matplotlib.dates as md
#Leemos el archivo csv
data = pd.read_csv('L2 analysis/Juriquilla/L2/graficacalidadjgro.
  csv')
#Nombramos las columnas
data.columns = ['datetime', 'escala1', 'escala2', 'escala3', '
  escala4']
v= data['escala1']
w= data['escala2']

```



```

y= data['escala3']
z= data['escala4']
#Convertimos a datetime la columna de tiempo (que inicialmente solo
  es object)
x = pd.to_datetime(data['datetime'], format='%Y-%m-%d')

fig, ax = plt.subplots(figsize=(10,10))
ax.scatter(x, w, color = 'green', label='Calidad alta')
ax.scatter(x, y, color = 'yellow', label='Calidad media')
ax.scatter(x, z, color = 'red', label='Calidad baja')
ax.scatter(x, v, color = 'gray', label='Sin datos')
#Establecemos el formato de días y luego H:M:S para el eje x
xfmt = md.DateFormatter('%Y-%m-%d')
ax.xaxis.set_major_formatter(xfmt)
ax.set_title('Calidad datos campo eléctrico atmosférico Juriquilla
  ')
ax.set_xlabel('Año-mes-día')
ax.set_ylabel('Escala arbitraria del 0 al 10')
plt.legend(loc='upper right', bbox_to_anchor=(1, 0.9))
plt.grid()
plt.show()

```

minutemean_peryear.py

```

#UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
#Argueta Hernández, Fidel Alejandro alejo_tigres@yahoo.com
#Asesor: Dr. Luis Xavier González Méndez
#Programa que calcula el promedio por minuto diurno de todo un año
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
#Leemos el archivo csv
data = pd.read_csv('Fourier Analysis/2017-
  unam_minuto_L1_cael_filltrend.csv', skiprows=(0,1,2,3,4,5,6))
#Nombramos las columnas
data.columns = ['dates', 'field']
data['dates'] = pd.to_datetime(data['dates'])
data = data.set_index(['dates'])
newdata = data.groupby([data.index.hour, data.index.minute]).mean()
newdata.to_csv('/home/alejandro/MEGA/MEGAsync/Python codes/Fourier
  Analysis/Meansperminute/meansperminutoUNAM.csv', encoding='utf
  -8')

```

graphmeanperminute.py

```

#UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
#Argueta Hernández, Fidel Alejandro alejo_tigres@yahoo.com
#Asesor: Dr. Luis Xavier González Méndez
#Programa que grafica el promedio por minuto diurno de todo un ano
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
import matplotlib.pyplot as plt
import matplotlib.dates as md
import matplotlib.ticker as plticker
data = pd.read_csv('/home/alejandro/MEGA/MEGAsync/Python codes/
    Fourier Analysis/Meansperminute/meansperminuteUNAM.csv',
    skiprows=0)
data.columns = ['hour', 'minute', 'field', 'points', 'datetime']
x = data['datetime']
z = data['field']
x = pd.to_datetime(data['datetime'], format='%H:%M')
fig, ax = plt.subplots(figsize=(10,10))
ax.plot(x,z, color = 'purple')
xfmt = md.DateFormatter('%H:%M')
ax.xaxis.set_major_formatter(xfmt)
ax.set_title('Promedio diurno del campo eléctrico atmosférico para
    2017 en el CCA')
ax.set_xlabel('Hora')
ax.set_ylabel('Campo eléctrico [kV/m]')
plt.grid()
plt.show()

```

varianza.py

```

#UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
#Argueta Hernández, Fidel Alejandro alejo_tigres@yahoo.com
#Asesor: Dr. Luis Xavier González Méndez
#Programa que grafica la varianza horaria de campo eléctrico atmosférico
#a partir del promedio diurno de todo un ano
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
import matplotlib.pyplot as plt
import matplotlib.dates as md
import matplotlib.ticker as plticker

```

```

data = pd.read_csv('/home/alejandro/MEGA/MEGAsync/Python codes/
    Fourier Analysis/varianza/varianzaJQROgraph.csv', skiprows=0)
data.columns = ['hora', 'promedio', 'varianza']
x = data['hora']
y = data['varianza']
x = pd.to_datetime(data['hora'], format='%H:%M')
fig, ax = plt.subplots(figsize=(10,10))
ax.bar(x,y, color = 'purple', width=0.05)
xfmt = md.DateFormatter('%H:%M')
ax.xaxis.set_major_formatter(xfmt)
ax.set_title('Varianza diurna del campo eléctrico atmosférico
    promedio para 2017 en Juriquilla')
ax.set_xlabel('Hora')
ax.set_ylabel('Varianza [kV/m]^2')
plt.grid()
plt.show()

```

DatosL0_ce.py (modificado)

```

import time
import os
#os.chdir('/home/met/Documents/Scripts')
import Master
start = time.time()

#-----
# define the path where the sites folders are stored
path = '/home/alex/Escritorio/MEGAsync/Python codes/Data E
    Geofisica avg min/'
# define the path where the month data folders are going to be
    stored
path1 = '/home/alex/Escritorio/MEGAsync/Python codes/Data E
    Geofisica avg min/'
#-----
# list with the name of the sites
#sites = ['erno', 'jqro', 'unam']
#sites = ['erno', 'unam']
sites = ['geof']
# loop for each site
for site in sites:
    print ('#####')
    print (site)

```

```

# create the folder where the month site data are going to be
  stored
L0 = os.path.join(path1,site,'L0')
if not os.path.exists(L0):
    os.makedirs(L0)
    os.makedirs(L0 + '/stat')
# the route of each site folder
route0 = os.path.join(path,site,'raw')
# merge the data
appended_data = Master.merge_files(route0, skip_rows =[0,2,3])

# name of the columns to be extracted
cols = ['TIMESTAMP','RECORD','E_Avg']
# print (appended_data.head())
# appended_data['E_Avg'].update(appended_data['E']) ; del
  appended_data['E']
# re-arrange the columns order

appended_data = appended_data[cols]
# Avg_frequencies
appended_data = Master.Avg_frequencies(appended_data)
# fill_missing_measurements
appended_data, oldest, youngest = Master.
  fill_missing_measurements(appended_data.reset_index())
# get the last data month uploaded
# 'site+RECORD.dat' is file with the record of the upload
  dates
Rrd = os.path.join(path,site,'L0',site+'_record.txt')
# check the record
F, R1, FileDate = Master.record(site, Rrd, oldest, youngest)
# 'date2' is the string of the first date to be taken
# 'date2' is the string of the last date to be taken
date1 = str(FileDate.strftime('%Y')) + '-' + str(FileDate.
  strftime('%m'))
date2 = youngest.strftime('%Y') + '-' + youngest.strftime('%m')
  + '-' + youngest.strftime('%d')
# the new 'appended_data' is the range of days contained
  between data1 & data2
appended_data = appended_data[date1:date2]
# split & save the data by month
Master.save_month(L0, site, appended_data, extension = '.dat',
  version = 'L0-cael', MDR = True, sheet_name = 'cael', area =
  'Electric field intensity')

```

```

    # write the record
    Master.write_record(site , youngest , Rrd , F , R1)
end = time.time()
print ('#####')
print ('___ Elapsed Time: ' + '%d' % (end - start) + ' s ___')
```

DatosL1_ce.py (modificado)

```

import pandas as pd
import numpy as np
from dateutil.relativedelta import relativedelta
import time
import os
#os.chdir('/home/met/Documents/Scripts')
import Master
start_time = time.time()
#-----
# define the path where the site folders are stored
path = '/home/alex/Escritorio/MEGAsync/Python codes/Data E
    Geofisica avg min/'
#-----
# list with the name of the sites
#sites = ['erno', 'jqro', 'unam']
sites = ['geof']
# loop for each site
for site in sites:
    print('#####')
    print(site)
    # find where the L0 data is stored
    L0 = os.path.join(path, site, 'L0')
    # specify where the L1 data is going to be stored
    # create the folder where the L1 data are going to be stored
    L1 = os.path.join(path, site, 'L1')
    if not os.path.exists(L1):
        os.makedirs(L1)
        os.makedirs(os.path.join(L1, 'Minuto'))
        os.makedirs(os.path.join(L1, 'Hora'))
        os.makedirs(os.path.join(L1, 'stat'))
        #os.makedirs(os.path.join(L1, 'stat'))
    # Rrd_L0 is the path of the file that indicates what was the
    last L0 date generated
    Rrd_L0 = os.path.join(L0, site + '_record.txt')
```

```

# Rrd_L1 is the path of the file that indicates what was the
  last L1 date generated
Rrd_L1 = os.path.join(L1, site+'_record_L1.txt')
# obtain the youngest date of the site
start_year, start_month = Master.youngest_date(site, sf = path)
start = pd.to_datetime(start_year+'-'+start_month)
# obtain the last date
_, _, end = Master.record(site, Rrd_L0, None, None)
# obtain the first date
F, R1, start = Master.record(site, Rrd_L1, start, end)
# create an array with dates to work (start - end)
dates_array = pd.date_range(start = start, end = end +
  relativedelta(months = 1), freq = 'M').strftime('%Y-%m')
# loop for each month in 'dates_array'
for date in dates_array:
  print(date)
  # 'file_name' is the name of the file (L0 file) to be read
  file_name = os.path.join(L0, date + '-' + site + '_L0_cael.
    dat')
  print(file_name)
  # read the data
  df = pd.read_csv(file_name, sep=',', skiprows
    =[0,1,2,3,4,5,7], encoding='latin1', na_values=['null', '
    NAN'])
  # indicate the date format
  df['TIMESTAMP'] = pd.to_datetime(df['TIMESTAMP'], dayfirst
    = False)
  # set the date as the index
  df = df.set_index(['TIMESTAMP'])
  # delete consecutive data
  df['E_Avg'][((df.E_Avg - df.E_Avg.shift(1)) == 0.0) | ((df.
    E_Avg - df.E_Avg.shift(-1)) == 0.0)] = np.nan
  # count the number of observations per day
  CC_min_i = df.resample('D').count()
  # apply the L1 filters
  df = Master.L1_filter(site, df, extra_filters = False,
    sheet_name = 'cael')
  # recount the number of observations per day
  CC_min_f = df.resample('D').count()
  # obtain the number of observations filtered
  CC_min = CC_min_i - CC_min_f
  # define where the filtered data report is going to be
    stored

```


Referencias

- Bahit, Eugenia (2012). *Curso: Python para Principiantes*.
- EFM-100 Atmospheric Electric Field Monitor. *Installation/Operators Guide for model EFM-100C* (2016). Boltek Corporation.
- Harrison, R. Giles (2013). "The Carnegie Curve." In: *Surv. Geophys.* 2013.34, pp. 209–232.
- Kelley, Michael C. and Robert H. Holzworth (2014). *The Earth's Electric Field Sources from Sun to Mud*. Elsevier, San Diego.
- Mauchly, S.J. (1923). "On the diurnal variation of the potential gradient of atmospheric electricity." In: *Terr. Mag. Atmos. Elect.* 28.3, pp. 61–81.
- Monterde-Andrade, Fernando (2017). *Puesta en marcha y calibración de equipo de monitoreo de campos eléctricos atmosféricos y primeros análisis de datos*.
- Newton-Bosch, Jania (2018). *Variaciones en la componente secundaria de los rayos cósmicos galácticos, detectadas por el Observatorio de Rayos Cósmicos de la Ciudad de México, atribuidas a campos eléctricos atmosféricos*.
- Schmidt, Ana Lucía (2013). *FFT: Transformada Rápida de Fourier*. accedido en 28-05-2020 a url <http://lcr.uns.edu.ar/fvc/NotasDeAplicacion/FVC-Schmidt20Ana20Luc3ADa.pdf>.
- Soula, Serge (2012). "Electrical Environment in a Storm Cloud." In: *Lightning Hazards to Aircraft and Launchers*.
- Suárez Ledo, Miguel (Jan. 2014). *Ciencias e Ingeniería para estudiantes de Python*. ISBN: 978-980-12-9190-9.
- Takashi, Tsutomu (1978). "Riming Electrification as a Charge Generation Mechanism in Thunderstorms." In: *J. Atmos. Sci.* 35, pp. 1536–1548.
- Tomé, César (2016). *El descubrimiento de la inducción electro-magnética*. accedido en 10-03-2020 a url <https://culturacientifica.com/2016/06/14/descubrimiento-la-induccion-electro-magnetica-1/>.
- Wilson, C.T.R. (1920). "Investigations on lightning discharges and on the electric field of thunderstorms." In: *Philos. Trans. R. Soc. A.* 221, pp. 73–115.