



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

---

**FACULTAD DE INGENIERÍA**

**Inversión de datos  
electromagnéticos mediante  
algoritmos genéticos en los  
condados de Butte/Glenn,  
California**

**TESIS**

Que para obtener el título de  
**Ingeniero Geofísico**

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

Alejandro Valentin Carballo Jardon y

Ricardo Malagón Sánchez

**DIRECTOR DE TESIS**

Dr. Jorge López Alvis



**Ciudad Universitaria, Cd. Mx., 2026**



**PROTESTA UNIVERSITARIA DE INTEGRIDAD Y  
HONESTIDAD ACADÉMICA Y PROFESIONAL  
(Titulación con trabajo escrito)**



De conformidad con lo dispuesto en los artículos 87, fracción V, del Estatuto General, 68, primer párrafo, del Reglamento General de Estudios Universitarios y 26, fracción I, y 35 del Reglamento General de Exámenes, me comprometo en todo tiempo a honrar a la institución y a cumplir con los principios establecidos en el Código de Ética de la Universidad Nacional Autónoma de México, especialmente con los de integridad y honestidad académica.

De acuerdo con lo anterior, manifiesto que el trabajo escrito titulado INVERSION DE DATOS ELECTROMAGNETICOS MEDIANTE ALGORITMOS GENETICOS EN LOS CONDADOS DE BUTTE/GLENN, CALIFORNIA, que presenté para obtener el título de INGENIERO GEOFÍSICO es original, de mi autoría y lo realicé con el rigor metodológico exigido por mi Entidad Académica, citando las fuentes de ideas, textos, imágenes, gráficos u otro tipo de obras empleadas para su desarrollo.

En consecuencia, acepto que la falta de cumplimiento de las disposiciones reglamentarias y normativas de la Universidad, en particular las ya referidas en el Código de Ética, llevará a la nulidad de los actos de carácter académico administrativo del proceso de titulación.

**ALEJANDRO VALENTIN CARBALLO JARDON**  
Número de cuenta: 318242804



**PROTESTA UNIVERSITARIA DE INTEGRIDAD Y  
HONESTIDAD ACADÉMICA Y PROFESIONAL  
(Titulación con trabajo escrito)**



De conformidad con lo dispuesto en los artículos 87, fracción V, del Estatuto General, 68, primer párrafo, del Reglamento General de Estudios Universitarios y 26, fracción I, y 35 del Reglamento General de Exámenes, me comprometo en todo tiempo a honrar a la institución y a cumplir con los principios establecidos en el Código de Ética de la Universidad Nacional Autónoma de México, especialmente con los de integridad y honestidad académica.

De acuerdo con lo anterior, manifiesto que el trabajo escrito titulado INVERSION DE DATOS ELECTROMAGNETICOS MEDIANTE ALGORITMOS GENETICOS EN LOS CONDADOS DE BUTTE/GLENN, CALIFORNIA, que presenté para obtener el título de INGENIERO GEOFÍSICO es original, de mi autoría y lo realicé con el rigor metodológico exigido por mi Entidad Académica, citando las fuentes de ideas, textos, imágenes, gráficos u otro tipo de obras empleadas para su desarrollo.

En consecuencia, acepto que la falta de cumplimiento de las disposiciones reglamentarias y normativas de la Universidad, en particular las ya referidas en el Código de Ética, llevará a la nulidad de los actos de carácter académico administrativo del proceso de titulación.

---

**RICARDO MALAGON SANCHEZ**  
Número de cuenta: 318305075

# Índice

<b>1. Introducción</b>	<b>7</b>
<b>2. Fundamentos teóricos del método</b>	<b>7</b>
2.1. El Sistema de Ecuaciones de Maxwell . . . . .	7
2.2. La aproximación cuasi-estática . . . . .	8
2.3. La ecuación de difusión electromagnética . . . . .	9
2.4. Solución de las ecuaciones de Maxwell para el TEM 1D . . . . .	9
2.5. Método electromagnético en el dominio del tiempo (TEM) . . . . .	10
2.6. Concepto de resistividad aparente. . . . .	13
2.7. Configuración para los sondeos de TEM . . . . .	17
2.8. Sistemas Aerotransportados (AEM) . . . . .	18
2.8.1. Sistemas en el dominio de la frecuencia (FDEM) . . . . .	19
2.8.2. Sistemas en el dominio del tiempo (TEM) . . . . .	20
2.9. Inversión geofísica con el método de Gauss-Newton . . . . .	20
2.10. Biblioteca SimPEG . . . . .	22
2.11. Algoritmos genéticos . . . . .	22
2.12. Algoritmos Genéticos de Codificación Real (RCGA) . . . . .	24
2.13. Librería para implementar algoritmos genéticos (DEAP) . . . . .	25
2.14. Paralelismo . . . . .	26
<b>3. Área de Estudio</b>	<b>27</b>
3.1. Geología . . . . .	28
3.2. Características hidrogeológicas . . . . .	30
3.3. Datos de sondeos . . . . .	31
3.4. Adquisición y procesamiento de datos . . . . .	32
3.5. Sondeos AEM seleccionados . . . . .	45
<b>4. Metodología</b>	<b>48</b>
4.1. Inversión con datos sintéticos . . . . .	48
4.2. Inversión con datos AEM reales . . . . .	50
4.3. Función Objetivo . . . . .	51
4.4. Parámetros del Algoritmo Genético (AG) . . . . .	51
<b>5. Resultados y discusión</b>	<b>53</b>
5.1. Selección del factor de regularización ( $\beta$ ) para la inversión con AG . . . . .	53
5.2. Modelos invertidos y comparación con el método de Gauss-Newton . . . . .	54
<b>6. Conclusiones</b>	<b>63</b>
<b>7. Anexo 1</b>	<b>64</b>

<b>8. Anexo 2. Paralelización de inversión con AG</b>	<b>71</b>
8.1. Problema y arquitectura del paralelismo . . . . .	72
8.2. Componentes técnicos en el código . . . . .	72
8.3. Ventajas y eficiencia . . . . .	72

## Índice de figuras

1. Se muestra la forma de onda de la corriente del transmisor. Adaptada de [18, p. 12]. . . . .	11
2. Configuración de sondeo de bobina central. Tomado de [22, p. 47]. . . . .	12
3. Comparación de la Forma de onda de salida del receptor y la forma de onda del transmisor. Adaptada de [18, p. 12]. . . . .	13
4. Ubicaciones de las puertas receptoras. Adaptada de [18, p. 12]. . . . .	13
5. Gráfico logarítmico. Voltaje de salida del receptor vs Tiempo. Adaptada de [18, p. 12]. . . . .	14
6. TEM: resistividad aparente de un semiespacio homogéneo. Adaptada de [18, p. 12]. . . . .	15
7. TEM: Respuesta de un medio de un estrato y un semiespacio, con dos contrastes de densidades diferentes. Adaptada de [18, p. 12]. . . . .	16
8. TEM: resistividad aparente. Adaptada de [18, p. 12]. . . . .	17
9. Configuraciones de TEM. Adaptada de [20, p. 455]. . . . .	18
10. Diagrama de flujo de la computación evolutiva . . . . .	24
11. Localidades de los condados de Butte/Glenn . . . . .	28
12. Extensión superficial y subsuperficial aproximada de las formaciones Tehama y Toscana, y extensión superficial únicamente de la formación Laguna. [9, p. 24]. . . . .	30
13. Ubicación de los sondeos cerca del área de prospección AEM de Butte. Los círculos verdes representan sondeos con información litológica y los círculos amarillos, sondeos con información geofísica. Las líneas de vuelo 312 están en azul y las 304 en morado. Tomada de [1, p. 24]. . . . .	32
14. Marco SkyTEM304M/312, incluyendo ubicaciones de instrumentación y los ejes X y Y. Las distancias de los ejes están en metros. Ubicaciones de la instrumentación listadas en Tabla 1. Tomada de [1, p. 29]. . . . .	34
15. Líneas de vuelo de Butte AEM agrupadas por fecha de adquisición. Tomada de [1, p. 32]. . . . .	36
16. Gráfica de la forma de onda LM de 210 Hz para el sistema SkyTEM312 registrada durante el estudio de AEM de Butte. El aumento de corriente está a la izquierda y la disminución de la corriente para apagar está a la derecha. La corriente está normalizada. Tomada de [1, p. 38]. . . . .	39
17. Gráfica de la forma de onda HM de 30 Hz para el sistema SkyTEM312 registrada durante el estudio de AEM de Butte. El aumento de corriente está a la izquierda y la disminución de la corriente para apagar está a la derecha. La corriente está normalizada. Tomada de [1, p. 39]. . . . .	40

18.	Gráfico de la forma de onda LM de 210 Hz para el sistema SkyTEM304M registrado durante el estudio de Butte AEM. El aumento de corriente está a la izquierda y la disminución de la corriente para apagar está a la derecha. La corriente está normalizada. Tomada de [1, p. 40]. . . . .	41
19.	Gráfico de la forma de onda HM de 30 Hz para el sistema SkyTEM304M registrado durante el estudio de Butte AEM. El aumento de corriente está a la izquierda y la disminución de la corriente para apagar está a la derecha. La corriente está normalizada. Tomada de [1, p. 41]. . . . .	42
20.	Un ejemplo de un perfil AEM que ilustra el aumento del espesor de las capas del modelo con la profundidad. Este es un modelo de 30 capas. Tomada de [1, p. 41]. . . . .	45
21.	Líneas correspondientes a los Pozos 21N01W11A001M y 22N01W29N .	46
22.	Líneas correspondientes al Pozo 20N02E4C . . . . .	47
23.	Líneas correspondientes al Pozo 21N02E18C . . . . .	47
24.	Modelo de 5 capas generado con datos sintéticos a la izquierda en azul con las siguientes conductividades; capa 1: 0-20m, $\sigma=0.10\text{S/m}$ , capa 2: 20-50m, $\sigma=1.00\text{S/m}$ , capa 3: 50-90m, $\sigma=0.05\text{S/m}$ , capa 4: 90-140m, $\sigma=0.50\text{S/m}$ , capa 5: 140+m, $\sigma=0.01\text{S/m}$ . En la derecha los datos sintéticos generados representados con puntos en negro . . . . .	48
25.	Modelo de datos sintéticos de 5 capas invertido a la izquierda y ajuste de los datos sintéticos generados a la derecha. . . . .	49
26.	Gráfica que visibiliza la manera en la que converge el algoritmo genético utilizado . . . . .	52
27.	Comparación de la estación 200801 (línea en azul) con el Pozo 21N01W11A001M (línea punteada en rojo) para las inversiones con $\beta$ : 50, 5000 y 50,000 respectivamente de izquierda a derecha. . . . .	53
28.	De izquierda a derecha. Modelo de AG invertido para la estación 200801. Modelo de Gauss Newton invertido las líneas de colores son los modelos invertidos con distintos valores de beta, la línea negra es el modelo elegido como más adecuado debido a los valores de $\phi$ y $\beta$ . . . . .	55
29.	De izquierda a derecha. Modelo de AG invertido para la estación 710301. Modelo de Gauss Newton invertido las líneas de colores son los modelos invertidos con distintos valores de beta, la línea negra es el modelo elegido como más adecuado debido a los valores de $\phi$ y $\beta$ . . . . .	56
30.	De izquierda a derecha. Modelo de AG invertido para la estación 730101. Modelo de Gauss Newton invertido las líneas de colores son los modelos invertidos con distintos valores de beta, la línea negra es el modelo elegido como más adecuado debido a los valores de $\phi$ y $\beta$ . . . . .	57
31.	De izquierda a derecha. Modelo de AG invertido para la estación 730201. Modelo de Gauss Newton invertido las líneas de colores son los modelos invertidos con distintos valores de beta, la línea negra es el modelo elegido como más adecuado debido a los valores de $\phi$ y $\beta$ . . . . .	58

32.	De izquierda a derecha. Modelo de AG invertido para la estación 200601. Modelo de Gauss Newton invertido las líneas de colores son los modelos invertidos con distintos valores de beta, la línea negra es el modelo elegido como más adecuado debido a los valores de $\phi$ y $\beta$ . . . . .	59
33.	De izquierda a derecha. Modelo de AG invertido para la estación 710201. Modelo de Gauss Newton invertido las líneas de colores son los modelos invertidos con distintos valores de beta, la línea negra es el modelo elegido como más adecuado debido a los valores de $\phi$ y $\beta$ . . . . .	60
34.	De izquierda a derecha. Modelo de AG invertido para la estación 300901. Modelo de Gauss Newton invertido las líneas de colores son los modelos invertidos con distintos valores de beta, la línea negra es el modelo elegido como más adecuado debido a los valores de $\phi$ y $\beta$ . . . . .	61
35.	De izquierda a derecha. Modelo de AG invertido para la estación 730401. Modelo de Gauss Newton invertido las líneas de colores son los modelos invertidos con distintos valores de beta, la línea negra es el modelo elegido como más adecuado debido a los valores de $\phi$ y $\beta$ . . . . .	62
36.	De izquierda a derecha. Modelo de AG invertido para la estación 200801 sin regularización. Modelo de Gauss Newton invertido las líneas de colores son los modelos invertidos con distintos valores de beta, la línea negra es el modelo elegido como más adecuado debido a los valores de $\phi$ y $\beta$ . . . . .	64
37.	De izquierda a derecha. Modelo de AG invertido para la estación 710301 sin regularización. Modelo de Gauss Newton invertido las líneas de colores son los modelos invertidos con distintos valores de beta, la línea negra es el modelo elegido como más adecuado debido a los valores de $\phi$ y $\beta$ . . . . .	65
38.	De izquierda a derecha. Modelo de AG invertido para la estación 730101 sin regularización. Modelo de Gauss Newton invertido las líneas de colores son los modelos invertidos con distintos valores de beta, la línea negra es el modelo elegido como más adecuado debido a los valores de $\phi$ y $\beta$ . . . . .	66
39.	De izquierda a derecha. Modelo de AG invertido para la estación 730201 sin regularización. Modelo de Gauss Newton invertido las líneas de colores son los modelos invertidos con distintos valores de beta, la línea negra es el modelo elegido como más adecuado debido a los valores de $\phi$ y $\beta$ . . . . .	67
40.	De izquierda a derecha. Modelo de AG invertido para la estación 200601 sin regularización. Modelo de Gauss Newton invertido las líneas de colores son los modelos invertidos con distintos valores de beta, la línea negra es el modelo elegido como más adecuado debido a los valores de $\phi$ y $\beta$ . . . . .	68

41.	De izquierda a derecha. Modelo de AG invertido para la estación 710201 sin regularización. Modelo de Gauss Newton invertido las líneas de colores son los modelos invertidos con distintos valores de beta, la línea negra es el modelo elegido como más adecuado debido a los valores de $\phi$ y $\beta$ . . . . .	69
42.	De izquierda a derecha. Modelo de AG invertido para la estación 300901 sin regularización. Modelo de Gauss Newton invertido las líneas de colores son los modelos invertidos con distintos valores de beta, la línea negra es el modelo elegido como más adecuado debido a los valores de $\phi$ y $\beta$ . . . . .	70
43.	De izquierda a derecha. Modelo de AG invertido para la estación 730401 sin regularización. Modelo de Gauss Newton invertido las líneas de colores son los modelos invertidos con distintos valores de beta, la línea negra es el modelo elegido como más adecuado debido a los valores de $\phi$ y $\beta$ . . . . .	71

# Resumen

El estudio se centra en las áreas este y oeste de la prospección electromagnética en los condados de Butte/Glenn, California. La geología del área está dominada por una secuencia compleja de depósitos volcánicos, aluviales y fluviales. El principal reservorio de agua es el sistema Acuífero del Valle Central (AVC), compuesto por arenas y gravas que exhiben condiciones no confinadas en la parte somera y confinadas a mayor profundidad debido a la heterogeneidad de lechos arcillosos lenticulares. La empresa Aqua Geo Frameworks (AGF) realizó la adquisición de datos electromagnéticos aerotransportados (AEM) para caracterizar la estratigrafía y la arquitectura del acuífero. Se recopilaron y analizaron aproximadamente 800 kilómetros de datos AEM utilizando los sistemas aerotransportados SkyTEM304M y SkyTEM312, ambos sistemas TEM de doble momento (LM y HM). En el presente trabajo la inversión fue llevada a cabo utilizando un Algoritmo Genético (AG) para una búsqueda paramétrica de las resistividades de las capas con una función objetivo definida como la suma del misfit de los datos ( $d$ ), normalizado por las incertidumbres y un término de regularización de suavidad ( $s$ ), ponderado por el factor  $\beta$ :  $C(m) = \phi_d(m) + \beta\phi_s(m)$ . Para la optimización del algoritmo se implementó el paralelismo a nivel de tarea dentro de la fase de evaluación del AG, aprovechando la arquitectura de multiprocesamiento del hardware. La comparación de resultados se realizó contra los modelos invertidos con Gauss-Newton y registros geofísicos de pozo. En un rango somero (0 - 100 m) el modelo AG mostró oscilaciones más abruptas entre capas consecutivas y en algunos casos sobrestimó la resistividad. El modelo AG estimó bien la resistividad en los primeros 20 metros en algunos pozos, mientras que Gauss-Newton la sobrestimó. En un rango intermedio (100 - 250 m) el AG a menudo siguió la tendencia del registro de una mejor manera que Gauss-Newton, sin embargo, en otros casos, ambos métodos tuvieron dificultades notándose la pérdida de bajos resistivos. En un rango profundo ( $>$  a 250 m) en la mayoría de las comparaciones, el modelo invertido con Gauss-Newton mostró una mejor aproximación o predicción de las tendencias de resistividad. El modelo de AG tendió a sobrestimar las resistividades en estas profundidades.

Palabras clave: Prospección electromagnética, TEM, Inversión de datos, Gauss-Newton, Algoritmos Genéticos.

This study focus on the east and west areas of the electromagnetic prospecting in the counties Butte/Glenn, California. The geology of the area is dominated by a complex sequence of volcanic, alluvial and fluvial deposits. The main water reservoir is the Central Valley Aquifer system, composed by sand and gravel that exhibit non confined conditions on the shallow part and confined on grater depths due to heterogeneity of the lenticular clay beds. Aqua Geo Frameworks (AGF) conducted a project in Butte and Glenn counties, California, to acquire data using airborne electromagnetic (AEM) surveys to characterize the aquifer's stratigraphy and architecture. The TDEM data was acquired via the airborne electromagnetic systems SkyTEM304M and SkyTEM312, both dual moment TDEM systems (LM and HM). The inversion was done with Genetic Algorithms for the parametric search of the resistivity of the

layers with a fitness function defined as the sum the misfit of the data (d), normalized by the uncertainties and a term of regularization of smoothness (s), wheighted by a  $\beta$  factor:  $C(m)=\phi_d(m) + \beta\phi_s(m)$ .

## 1. Introducción

El presente trabajo resalta la importancia del método electromagnético en el dominio del tiempo (TDEM por sus siglas en inglés) también conocido como transitorio electromagnético (TEM) en español. Este método se utiliza para la caracterización del subsuelo debido a su sensibilidad a la resistividad eléctrica, una propiedad clave para el mapeo de acuíferos y la delimitación de ciertas estructuras geológicas. No obstante, a pesar de la gran utilidad de este método, la inversión de datos electromagnéticos obtenidos es un problema no lineal, lo que lo hace susceptible a la no-unicidad y a la dependencia de condiciones iniciales.

Como alternativa para abordar la resolución de este problema de inversión, el presente trabajo emplea técnicas de inteligencia artificial, específicamente métodos heurísticos como los Algoritmos Genéticos (AG). La aplicación de estos algoritmos se lleva a cabo utilizando datos reales de sondeo electromagnético aerotransportado (AEM) adquiridos en el área de Butte/Glenn, ubicada en el norte del valle de Sacramento, California. La metodología implementada utiliza la librería SimPEG para la modelación directa del modelo y la librería DEAP para el Algoritmo Genético. El AG es ideal para este tipo de problema ya que no requiere de un punto de partida inicial y es robusto ante la presencia de múltiples mínimos en el espacio de parámetros. La función de costo utilizada combina el misfit de los datos y un término de regularización ponderado por un factor  $\beta$ .

El trabajo se enfoca en dos objetivos principales: el primero, determinar la resistividad de un modelo de subsuelo 1D de múltiples capas que minimice la discrepancia de los datos AEM observados y los datos predichos por la simulación; y el segundo, comparar la capacidad de resolución y las tendencias del modelo generado por el Algoritmo Genético frente a modelos generados por el método iterativo determinístico de Gauss-Newton, utilizado en el estudio original de Kang et al., 2021, o como base para métodos clásicos de SimPEG, tomando como referencia registros geofísicos de pozo.

## 2. Fundamentos teóricos del método

### 2.1. El Sistema de Ecuaciones de Maxwell

Todo método de prospección electromagnética se fundamenta en las Ecuaciones de Maxwell. Estas ecuaciones describen cómo los campos eléctricos y magnéticos interactúan entre sí y con la materia. En su forma diferencial, son:

$$\nabla \times \bar{H} = \bar{J} + \frac{\partial \bar{D}}{\partial t} \quad (\text{Ley de Ampere - Maxwell}) \quad (1)$$

$$\nabla \times \bar{E} = -\frac{\partial \bar{B}}{\partial t} \quad (\text{Ley de Faraday}) \quad (2)$$

$$\nabla \cdot \bar{D} = \rho \quad (\text{Ley de Gauss}) \quad (3)$$

$$\nabla \cdot \bar{B} = 0 \quad (\text{Ley de Gauss para el campo magnético}) \quad (4)$$

Donde  $\bar{E}$  es el campo eléctrico,  $\bar{H}$  es el campo magnético,  $\bar{D}$  es el desplazamiento eléctrico,  $\bar{B}$  es la inducción magnética,  $\bar{J}$  es la densidad de corriente eléctrica y  $\rho$  es la densidad de carga.

Para relacionar estos campos con las propiedades físicas del medio, utilizamos las relaciones constitutivas:

$$\bar{D} = \epsilon \bar{E} \quad (\text{donde } \epsilon \text{ es la permitividad eléctrica}) \quad (5)$$

$$\bar{B} = \mu \bar{H} \quad (\text{donde } \mu \text{ es la permeabilidad magnética}) \quad (6)$$

$$\bar{J} = \sigma \bar{E} \quad (\text{Ley de Ohm, donde } \sigma \text{ es la resistividad eléctrica}) \quad (7)$$

## 2.2. La aproximación cuasi-estática

En este régimen, la corriente de conducción  $J$  que fluye a través de los materiales conductores del subsuelo es mucho mayor que la corriente de desplazamiento  $\frac{\partial \bar{D}}{\partial t}$  que está relacionada con la propagación de ondas de alta frecuencia.

Por lo tanto, podemos aplicar la aproximación cuasi-estática que consiste en despreciar el término de la corriente de desplazamiento:

$$J \gg \frac{\partial \bar{D}}{\partial t} \quad (8)$$

Con esta simplificación, la Ley de Ampere-Maxwell se reduce a:

$$\nabla \times \bar{H} \approx \sigma \bar{E} \quad (9)$$

Esta simplificación es la que define la física de los métodos electromagnéticos inductivos.

### 2.3. La ecuación de difusión electromagnética

El comportamiento del TEM no es el de una onda que se refleja, sino el de un campo que se difunde en el medio. Comenzamos aplicando el rotacional a la Ley de Faraday:

$$\nabla \times (\nabla \times \bar{E}) = \nabla \left( -\frac{\partial \bar{B}}{\partial t} \right) \quad (10)$$

Aplicando la identidad vectorial  $\nabla \times (\nabla \times \bar{E}) = \nabla(\nabla \cdot \bar{E}) - \nabla^2 \bar{E}$  y asumiendo que no hay acumulación de carga ( $\rho = 0$ , por lo que  $\nabla \cdot \bar{E} = 0$ ), obtenemos:

$$-\nabla^2 \bar{E} = -\frac{\partial}{\partial t} (\nabla \times \bar{B}) \quad (11)$$

Ahora se sustituye la relación constitutiva  $\bar{B} = \mu \bar{H}$  y la ley de Ampere cuasi-estática ( $\nabla \times \bar{H} = \sigma \bar{E}$ )

$$-\nabla^2 \bar{E} = -\frac{\partial}{\partial t} (\mu(\sigma \bar{E})) \quad (12)$$

Si asumimos que las propiedades  $\mu$  y  $\sigma$  son homogéneas en el espacio, llegamos a:

$$\nabla^2 \bar{E} = \frac{\partial}{\partial t} \mu \sigma \frac{\partial \bar{E}}{\partial t} \quad (13)$$

Esta es formalmente una ecuación de difusión. Nos dice que la variación temporal del campo eléctrico  $\frac{\partial \bar{E}}{\partial t}$  es proporcional a sus variaciones espaciales en el tiempo ( $\nabla^2 \bar{E} = \frac{\partial}{\partial t} \mu \sigma \frac{\partial \bar{E}}{\partial t}$ ). El método TEM funciona midiendo la tasa de este decaimiento (difusión) en el tiempo. La cual está directamente relacionada con la resistividad  $\rho$  y la permeabilidad  $\mu$  del subsuelo.

### 2.4. Solución de las ecuaciones de Maxwell para el TEM 1D

En el dominio de la frecuencia:

$$\nabla \times E_m + i\omega\mu H_m = -J_m^s \quad (14)$$

$$\nabla \times H_m - (\sigma + i\omega\epsilon)E_m = 0 \quad (15)$$

Donde  $\omega$  es la frecuencia angular,  $F$  es el potencial vectorial eléctrico o potencial de Schelkunoff.

$$H_m = -(\sigma + i\omega\epsilon)F + \frac{1}{i\omega\mu} \nabla(\nabla \cdot F) \quad (16)$$

$$E_m = -\nabla \times F \quad (17)$$

donde  $F$  es un potencial vectorial que contiene toda la información respecto a los campos  $E$  y  $H$ . De esta forma, estas expresiones se pueden sustituir en las ecuaciones de Maxwell con lo que se obtiene una ecuación de Helmholtz:

$$\nabla^2 F + k^2 F = -J_m^s, \quad \text{donde} \quad k^2 = \omega^2 \mu \epsilon - i \omega \mu \sigma \quad (18)$$

Donde  $k$  es el número de onda.

Ahora podemos resolver para el potencial  $F$  y, finalmente, usar este potencial  $F$  para obtener las expresiones para  $E$  y  $H$ . La solución dependerá de la geometría de la fuente y del medio conductivo en donde se inducen las corrientes. De forma general, el potencial  $F$  debido a una fuente cualquiera está dado por:

$$F(r) = \int_{V'} \frac{e^{-ik|r-r'|}}{4\pi|r-r'|} J_m^s(r') dV' \quad (19)$$

Donde  $r$  es el vector de posición del punto de observación,  $r'$  es el vector de posición del punto de la fuente,  $|r-r'|$  es la distancia entre el punto de la fuente y el punto de observación,  $V'$  es el volumen de la fuente,  $dV'$  elemento diferencial de volumen en la posición de la fuente,  $e^{-ik|r-r'|}$  término de propagación describe cómo la fase de la señal cambia y cómo su amplitud se atenúa a medida que se aleja de la fuente,  $4\pi|r-r'|$  representa la atenuación geométrica o decaimiento por expansión esférica en tres dimensiones.

Para el caso de un medio estratificado, la solución analítica considera que los cambios de propiedades físicas (conductividad) coinciden con planos de  $z$  constante en donde parte de la energía se transmite a las capas más profundas. Esta geometría permite convertir la ecuación diferencial parcial de  $F$  en una ecuación diferencial ordinaria en  $z$  (Ward and Hohmann, 1987). Esta conversión toma la forma de una transformada de Hankel o una transformada de Fourier 2D debido a que los planos mencionados tienen extensión infinita en las direcciones  $x$  y  $y$ . Por lo tanto, la solución analítica para un medio 1D en el dominio de la frecuencia involucra integrales con funciones de Bessel que sólo se pueden resolver de forma numérica (por ejemplo, usando filtros digitales; Werthmüller, 2017), por esta razón, este tipo de solución se denomina semi-analítica. Finalmente, para obtener la solución en el dominio del tiempo se usa una transformada inversa de Fourier. Este tipo de solución semi-analítica se encuentra implementada en la biblioteca SimPEG usada en este trabajo como se describe más adelante.

## 2.5. Método electromagnético en el dominio del tiempo (TEM)

Las técnicas electromagnéticas que funcionan con transmisor se pueden dividir en dos grandes grupos: Método electromagnético en el dominio del tiempo TEM y Método electromagnético en el dominio de la frecuencia FDEM. En el dominio de frecuencia (FDEM), la corriente del transmisor varía sinusoidalmente con el tiempo a una frecuencia fija que se selecciona en función de la profundidad de exploración deseada. En el de dominio de tiempo (TEM), por otro lado, la corriente del transmisor, aunque sigue siendo periódica, pero es una onda cuadrada simétrica modificada, como

se muestra en la figura 1. Se observa que después de cada cuarto de período  $t/4$ , la corriente del transmisor se reduce abruptamente a cero durante un cuarto de período, tras lo cual fluye en la dirección opuesta.

Las aplicaciones más comunes del TEM abarcan desde la hidrogeología, para la búsqueda de agua subterránea, hasta la exploración de hidrocarburos y energía geotérmica. En el sector minero, destaca por su precisión en la localización de depósitos profundos de sulfuros masivos y uranio (Nabighian & Macnae, 1991). Por su capacidad para caracterizar la resistividad del terreno, se ha consolidado como una técnica estándar para el mapeo de estructuras geológicas complejas.

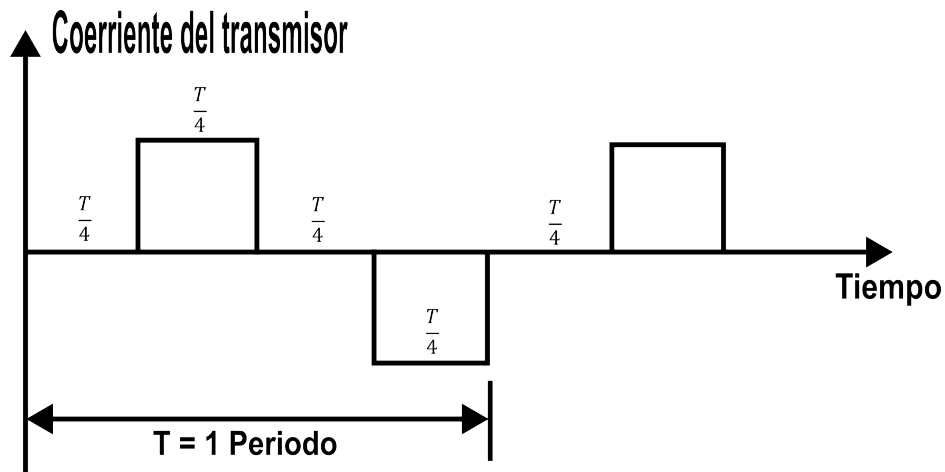


Figura 1: Se muestra la forma de onda de la corriente del transmisor. Adaptada de [18, p. 12].

En la figura 2 se muestra una configuración típica de un sondeo de resistividad TEM, donde se observa que el transmisor está conectado a un bucle cuadrado (generalmente de una sola vuelta) de alambre tendido en el suelo. Una bobina receptora de múltiples vueltas, ubicada en el centro del bucle del transmisor, está conectada al receptor a través de un cable corto.

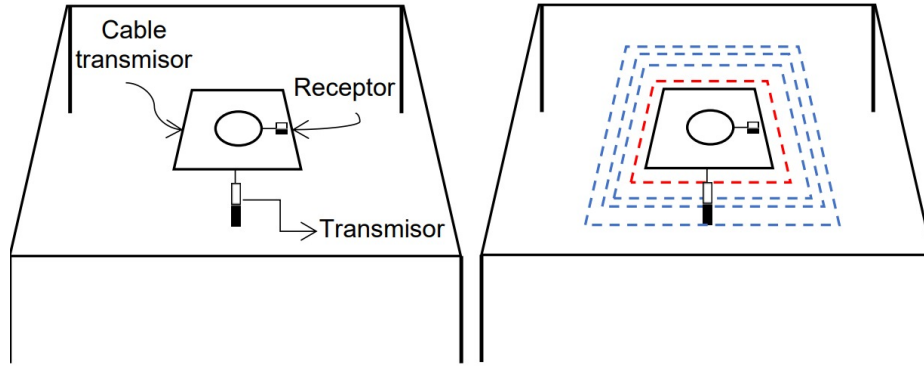


Figura 2: Configuración de sondeo de bobina central. Tomado de [22, p. 47].

El proceso de reducir abruptamente la corriente del transmisor a cero induce, de acuerdo con la ley de Faraday Ecuación 2, un campo electromagnético que induce corrientes eléctricas en el subsuelo de corta duración, lo que provoca que un bucle de corriente fluya en las inmediaciones del cable del transmisor. Sin embargo, debido a la resistividad finita del terreno, la amplitud de la corriente comienza a decaer inmediatamente. Esta corriente decreciente induce un pulso de voltaje que provoca un flujo de corriente, pero ahora a mayor distancia del bucle transmisor y a mayor profundidad. Este flujo de corriente a mayor profundidad también decae debido a la resistividad finita del terreno, induciendo un flujo de corriente aún más lejano, y así sucesivamente. Para estimar la distribución de las corrientes inducidas en el subsuelo a lo largo del tiempo, se mide el decaimiento transitorio del campo magnético secundario. Esto se logra empleando una pequeña bobina receptora de múltiples vueltas, posicionada habitualmente en el centro geométrico del bucle transmisor.

El voltaje en la bobina receptora se muestra esquemáticamente (junto con la corriente del transmisor) en la figura 3. Para medir con precisión las características de decaimiento de este voltaje, el sistema receptor contiene 20 ventanas temporales figura 4, cada una de las cuales se fija secuencialmente para medir y registrar la amplitud del voltaje decreciente en 20 instantes sucesivos. Para minimizar la distorsión al medir el voltaje transitorio, el ancho de las ventanas de muestreo se adapta a la tasa de decaimiento de la señal. Las primeras ventanas son estrechas para capturar la rápida caída inicial, mientras que las ventanas posteriores se diseñan progresivamente más anchas para ajustarse a la variación más lenta del campo electromagnético. Esta técnica es conveniente ya que las ventanas más anchas mejoran la relación señal/ruido, que disminuye a medida que la amplitud del voltaje decae en instantes posteriores. Como se observa en la figura 3, se generan cuatro señales transitorias de voltaje en el receptor durante cada período completo, compuesto por un pulso positivo más un pulso negativo del flujo de corriente del transmisor. Dado que el registro del transitorio electromagnético se efectúa en ausencia de la corriente del transmisor, la medición aísla la respuesta del subsuelo. Esta característica vuelve al método prácticamente insensible a pequeñas imprecisiones en la colocación de la bobina receptora, aumentando la fiabilidad de los datos.

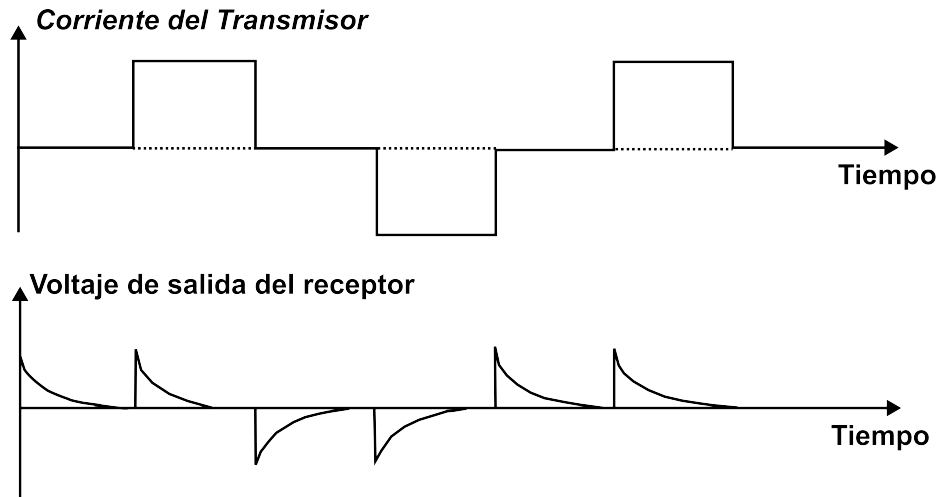


Figura 3: Comparación de la Forma de onda de salida del receptor y la forma de onda del transmisor. Adaptada de [18, p. 12].

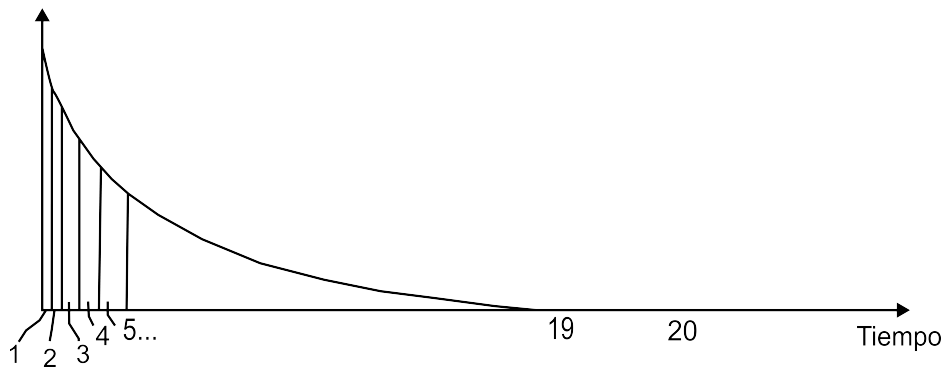


Figura 4: Ubicaciones de las puertas receptoras. Adaptada de [18, p. 12].

## 2.6. Concepto de resistividad aparente.

En la figura 5 se observa la gráfica lineal de la respuesta transitoria típica de un semi-espacio homogéneo, cuando esta respuesta se gráfica en un espacio logarítmico contra el tiempo. La respuesta se divide en 3 etapas:

-Etapa temprana: Al ser interrumpida la corriente de forma abrupta, la respuesta es constante en el tiempo debido a la aproximación cuasi-estática.

-Etapa intermedia: Debido a la pérdida de corriente la forma de la respuesta varía continuamente con el tiempo.

-Etapa tardía: Las corrientes ya no varían en el tiempo, la respuesta ahora es una variación lineal en la gráfica log-log.

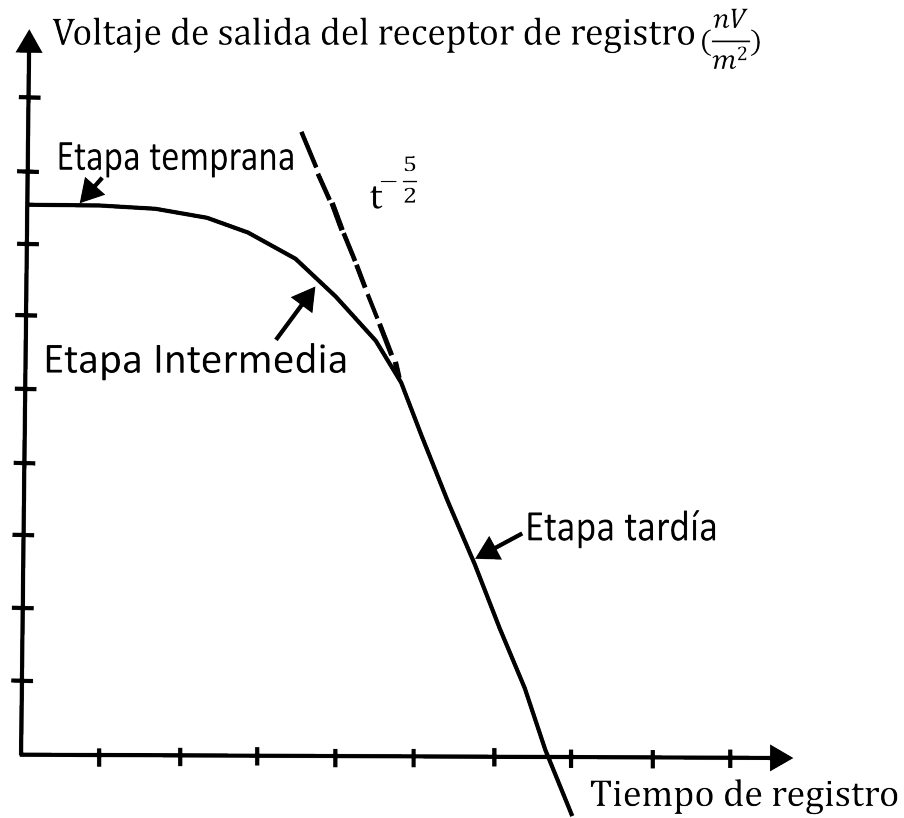


Figura 5: Gráfico logarítmico. Voltaje de salida del receptor vs Tiempo. Adaptada de [18, p. 12].

La respuesta es una función matemática compleja de la resistividad y el tiempo, sin embargo en esta última etapa tardía, la matemática se simplifica y se puede demostrar que durante ese tiempo la respuesta varía de manera bastante simple, esto como:

$$e(t) = \frac{k_1 M \sigma^{\frac{3}{2}}}{t^{\frac{5}{2}}} \quad (20)$$

donde  $k_1$  es una constante.

$M$  = Momento magnético ( $A$ ) x área ( $m^2$ )

$\sigma$  = resistividad de terreno ( $S/m$ )

$t$  = tiempo ( $s$ )

$e(t)$  = Voltaje de salida de una bobina receptora de una sola vuelta con un área de ( $m^2$ ).

Se observa que, a diferencia de la medición de resistividad convencional donde el voltaje varía linealmente con la resistividad del terreno, para TEM el voltaje medido,  $e(t)$ , varía como  $\sigma^{\frac{3}{2}}$  por lo que es más sensible a variaciones pequeñas en la resistividad, así mismo durante la etapa final el voltaje decae a una tasa de  $t^{\frac{5}{2}}$ , finalmente la señal desaparece en el ruido del mismo sistema alcanzando la profundidad máxima haciendo que medir sea irrelevante.

En los métodos de resistividad de corriente continua, para aumentar la profundidad de exploración debemos aumentar el espaciado entre electrodos. En el caso del sondeo TEM, por otro lado, observamos anteriormente que a medida que aumentaba el tiempo, la profundidad de los bucles de corriente aumentaba, por lo que este fenómeno permite realizar un sondeo de resistividad con la profundidad. Así, para obtener la resistividad aparente, podemos invertir la ecuación 20 para leer (ya que  $\rho = \frac{1}{\sigma}$ ).

$$\rho_a(t) = \frac{k_2 M^{\frac{2}{3}}}{e(t)^{\frac{2}{3}} t^{\frac{5}{3}}} \quad (21)$$

Supongamos una vez más que la resistividad del terreno no varía con la profundidad y es de resistividad  $\rho_1$ . Para este caso, una gráfica de  $\rho_a(t)$  en función del tiempo sería como se muestra en la Figura 6. Se observa que, en tiempos tardíos, la resistividad aparente  $\rho_a(t)$  es igual a  $\rho_1$ , pero en tiempos iniciales,  $\rho_a(t)$  es mucho mayor que  $\rho_1$ .

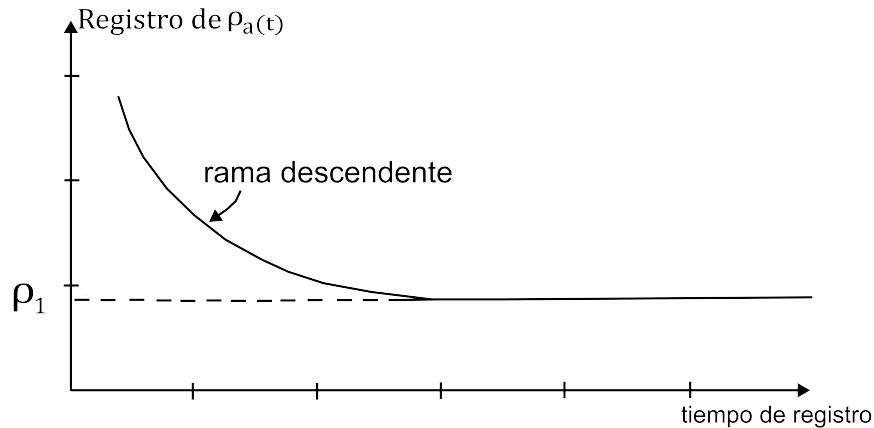


Figura 6: TEM: resistividad aparente de un semiespacio homogéneo. Adaptada de [18, p. 12].

La razón de esto es que la definición de resistividad aparente se basa (como se ve en la Figura 6) en el comportamiento temporal del voltaje de salida de la bobina recep-

tora en tiempos tardíos, cuando decae como  $t^{\frac{5}{2}}$ . En tiempos tempranos e intermedios, la figura 5 muestra que el voltaje del receptor es demasiado bajo (la línea discontinua indica el voltaje dado por la (“aproximación de etapa tardía”) y, por lo tanto, según la ecuación (21), la resistividad aparente será demasiado alta. Por esta razón, siempre habrá, como se muestra en la figura 6, una rama descendente en tiempos tempranos donde la resistividad aparente es mayor que la resistividad del semiespacio.

Supongamos, que el modelo de la tierra tiene dos capas, con una resistividad de capa superior  $\rho_1$  y espesor  $h$ , y una resistividad de base  $\rho_2 > \rho_1$ . Inicialmente, cuando las corrientes se encuentran completamente en la capa superior de resistividad  $\rho_1$ , la curva de decaimiento será similar a la de la figura 5, y la curva de resistividad aparente será similar a la de la figura 6. Sin embargo, posteriormente, las corrientes se distribuirán en ambas capas, y mucho más tarde se ubicarán completamente en la capa inferior, de resistividad  $\rho_2$ . Dado que  $\rho_2 > \rho_1$ , la ecuación (21) muestra que, como se indica en la figura 7a, el voltaje medido será menor del que debería ser para el semi-espacio homogéneo de resistividad  $\rho_1$ . El efecto sobre la curva de resistividad aparente se muestra en la figura 8a y consiste (ya que posteriormente todas las corrientes se encuentran en la capa inferior) en que la resistividad aparente  $\rho_a(t)$  se iguala a  $\rho_2$ . En el caso de que  $\rho_2 < \rho_1$ , el comportamiento inverso también es el esperado; es decir, posteriormente la respuesta de voltaje mostrada en la figura 7, es mayor que la de un semi-espacio homogéneo de resistividad  $\rho_1$ , y la curva de resistividad aparente se convierte en la de la Figura 8b, igualándose al nuevo valor de  $\rho_2$  en instantes posteriores.

En resumen, vemos que, a excepción de la rama descendente inicial y la región anómala intermedia, el comportamiento de sondeo del TEM es análogo al de la resistividad de corriente continua si dejamos que el paso del tiempo alcance mayor profundidad de exploración en lugar del aumento del espaciado entre electrodos.

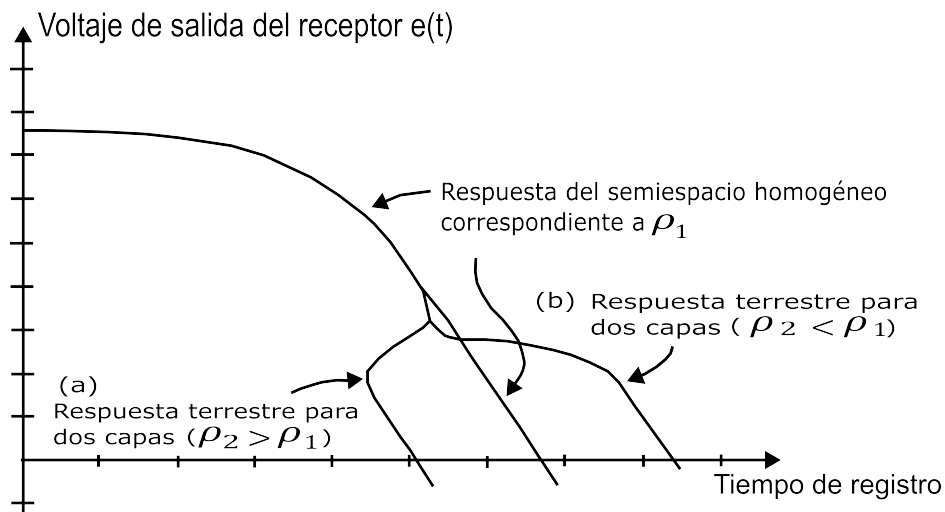


Figura 7: TEM: Respuesta de un medio de un estrato y un semiespacio, con dos contrastes de densidades diferentes. Adaptada de [18, p. 12].

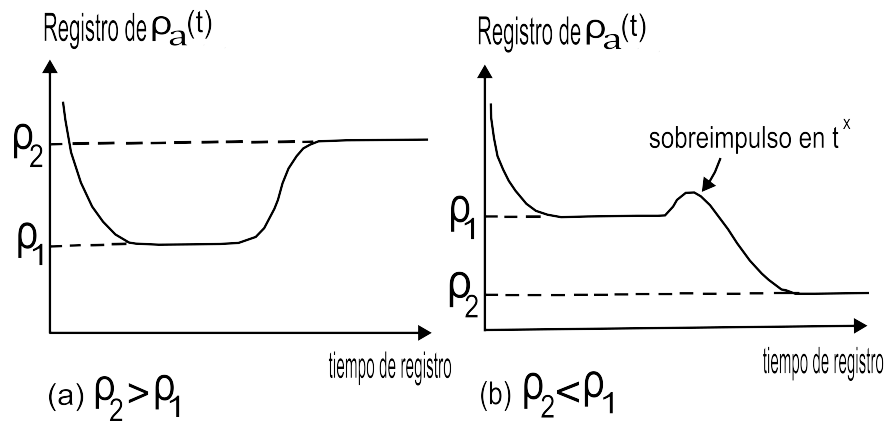


Figura 8: TEM: resistividad aparente. Adaptada de [18, p. 12].

## 2.7. Configuración para los sondeos de TEM

Un factor común en todas las técnicas inductivas de dominio del tiempo existentes es el hecho de que todas utilizan un bucle sin conexión a tierra más o menos rectangular como transmisor, y varían de acuerdo a la configuración del transmisor-receptor (Nabighian & Macnae, 1991).

Las configuraciones más comunes se explican a continuación:

-Transmisor y receptor de una sola bobina: En esta se utiliza una sola bobina tanto para el receptor como para el transmisor, cuando la corriente fluye igualmente actúa como transmisor (Figura 9a).

-De bobina coincidente: En este existe la misma geometría que el de una sola bobina con la diferencia que el receptor y el transmisor son bobinas separadas espacialmente coincidentes (Figura 9a).

-De bobina central: Se utiliza un receptor en la parte central de la bobina transmisora (Figura 9b).

-Transmisor y receptor separados: Las bobinas transmisor y receptor están separadas por una distancia fija conocida (Figura 9c-d).

-De doble bobina: Se utilizan dos arreglos adyacentes conectados en paralelo (Figura 9e).

-Transmisor fijo grande y receptor móvil: Consta de una gran bobina transmisora fija y un pequeño receptor móvil que se mide perpendicularmente al bucle (Figura 9f).

-TEM de perforación: El dipolo receptor puede bajar a un pozo de perforación, puede encontrarse como extensión del sistema transmisor fijo.

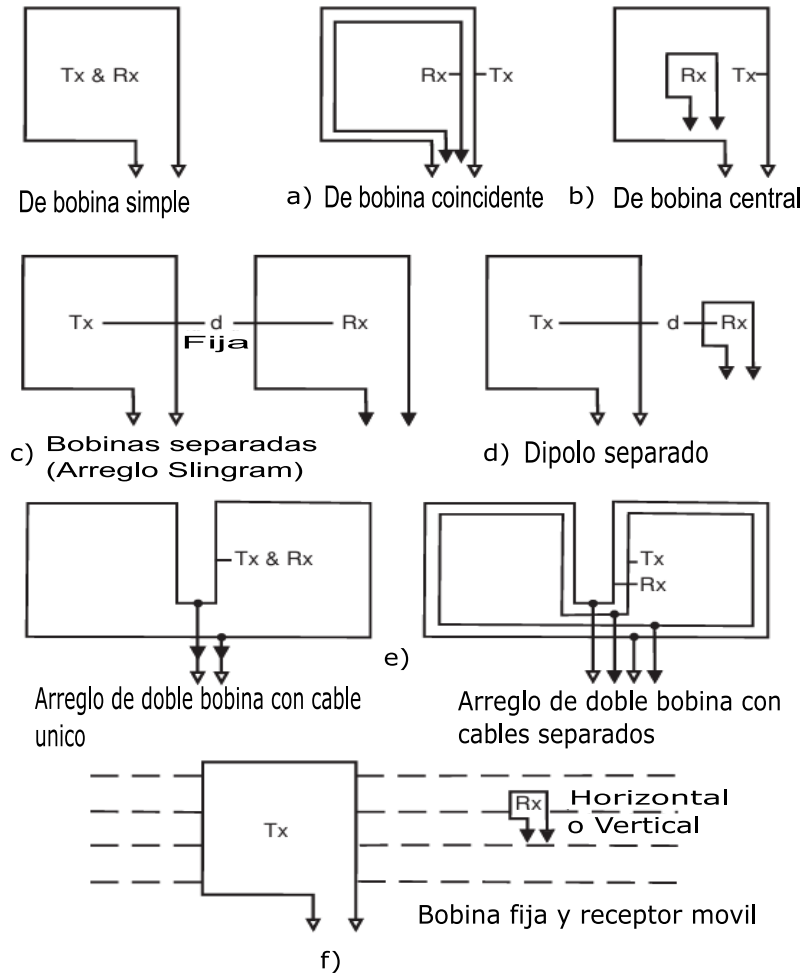


Figura 9: Configuraciones de TEM. Adaptada de [20, p. 455].

## 2.8. Sistemas Aerotransportados (AEM)

Una de las principales ventajas de los sistemas aereo, es que la adquisición de datos se puede realizar desde el aire. Solo el magnetismo y la radiactividad, entre otros

métodos, se pueden utilizar de esta manera. Se ha desarrollado una gran variedad de sistemas electromagnéticos aerotransportados para aprovechar este hecho (Telford et al., 1990). Las frecuencias empleadas en estos sistemas son las mismas que en los sistemas electromagnéticos terrestres; generalmente se utilizan frecuencias más bajas con mayor potencia.

La posición de la aeronave se determina normalmente a partir de GPS y otros sensores como altímetros. Las aeronaves más usadas son los helicópteros, ya que son más adecuados para trabajos detallados y terrenos accidentados.

El proceso se inicia con una bobina transmisora (Tx) que genera un campo electromagnético primario ( $H_p$ ), el cual se propaga hacia el subsuelo. Este campo variable en el tiempo induce corrientes de vórtice en los materiales conductores presentes. Estas corrientes, a su vez, generan un campo magnético secundario ( $H_s$ ) que es detectado por una bobina receptora (Rx) (Nabighian, 1991). Las características de este campo secundario son su amplitud, forma de onda y desfase con respecto al campo primario y contienen la información diagnóstica sobre la resistividad, geometría y profundidad de las estructuras geológicas.

Los sistemas aéreos AEM han sido clasificados típicamente dependiendo de la forma de onda del transmisor fuente y del cómo son medidos, es decir, en el dominio de la frecuencia (FDEM) o en el dominio del tiempo (TEM) (Telford et al., 1990).

### **2.8.1. Sistemas en el dominio de la frecuencia (FDEM)**

Estos sistemas operan con un transmisor que emite una onda continua a una o varias frecuencias específicas. La mitigación del campo primario se logra mediante una geometría de acoplamiento nulo o mediante compensación electrónica (Telford et al., 1990).

El receptor mide el campo secundario en términos de dos componentes:

**Componente En-Fase (In-Phase):** La porción del  $H_s$  que está en fase con el Campo primario  $H_p$ . Es altamente sensible a buenos conductores, pero también es susceptible al ruido generado por los movimientos relativos entre Tx y Rx.

**Componente en Cuadratura (Quadrature):** La porción desfasada  $90^\circ$  con respecto al Campo primario  $H_p$ . Es menos sensible al ruido de movimiento y responde bien a conductores de moderados a pobres.

Algunos sistemas como VLF utilizan transmisores militares de baja frecuencia en tierra y los sistemas de componentes en fase como Dighem y Aerodat usan un transmisor aéreo. Estos sistemas se caracterizan por proporcionar información detallada de alta resolución, aunque su profundidad de penetración es generalmente más limitada.

### 2.8.2. Sistemas en el dominio del tiempo (TEM)

Los sistemas transitorios (TEM o TDEM) no emiten una onda continua, sino pulsos intensos de corriente. El principio operativo clave es medir el campo secundario únicamente durante el tiempo de apagado ('off-time') del transmisor. En el instante en que el campo primario se corta, campo magnético primario  $H_p$ , pero las corrientes inducidas en los conductores del subsuelo no cesan instantáneamente sino que decaen lentamente. El receptor capta esta curva de decaimiento transitorio en múltiples ventanas de tiempo. Como se observa en la figura 5.

Esta metodología presenta dos ventajas:

1. Eliminación del campo primario  $H_p$ : El receptor mide en ausencia del campo primario, eliminando la principal fuente de interferencia.
2. Profundidad de penetración: El uso de pulsos de alta potencia y la medición del decaimiento en tiempos tardíos que corresponden a corrientes más profundas otorgan a los sistemas TEM una profundidad de penetración significativamente mayor. En comparación con los métodos electromagnéticos en el dominio de la frecuencia (FDEM)

La elección entre un sistema FDEM o TEM depende de los objetivos de la exploración. Los sistemas AEM se han convertido en herramientas indispensables no solo en la exploración mineral, sino también en aplicaciones hidrogeológicas para el mapeo de acuíferos (DWR, 2021) y en estudios ambientales para la caracterización de la intrusión salina o la delimitación de contaminantes.

## 2.9. Inversión geofísica con el método de Gauss-Newton

De acuerdo con Aster et al (2019), el método de Gauss-Newton constituye una estrategia iterativa esencial para la resolución de problemas de mínimos cuadrados no lineales. Su funcionamiento se basa en la linealización local de la función objetivo mediante una expansión de Taylor de primer orden en cada iteración, procediendo posteriormente a resolver las ecuaciones normales para actualizar el modelo. Este enfoque es la base de numerosos algoritmos de inversión geofísica debido a su eficiencia computacional, ya que simplifica el proceso al prescindir del cálculo de la matriz Hessiana de segundo orden.

Tiene como objetivo encontrar el vector de parámetros  $m$  que minimiza la función de objetivo  $\phi(m)$ , donde:

$$\phi = \phi_d + \beta\phi_m \quad (22)$$

Donde  $\phi_d = ||W_d(\mathcal{F}[m] - d_{obs})||^2$   
 $\phi_m = \alpha_s ||(m - m_{ref})||^2 + \alpha_x ||W_x m||^2$

- $\phi$ : Valor escalar a minimizar.
- $\beta$ : Parámetro de regularización.
- $m$ : El vector de parámetros del modelo.
- $F[m]$ : El operador del modelo directo. Proyecta el modelo al espacio de los datos.
- $d_{obs}$ : El vector de datos observados.
- $W_d$ : La matriz de pesos en los datos (matriz diagonal para errores independientes).
- $m_{ref}$ : El modelo de referencia.
- $\alpha_s$ : Es el coeficiente del término de pequeñez.
- $W_x$ : Matriz de suavidad.
- $\alpha_x$ : Es el coeficiente del término de suavizado.

El método no requiere calcular la matriz Hessiana completa. En su lugar, utiliza una aproximación lineal de la función no lineal  $f(m)$  en cada iteración. Esto simplifica el cálculo, ya que solo necesita el Jacobiano ( $J$ ), es decir, la matriz de las primeras derivadas de las predicciones del modelo con respecto a los parámetros.

El paso de actualización del modelo  $\delta m$  se encuentra resolviendo las ecuaciones normales de Gauss-Newton:

$$(J^T W_d^T W_d J + \beta \alpha_s I_m + \beta \alpha_x W_x^T W_x) \delta m = -(J^T W_d^T W_d \delta d + \beta \alpha_s (m^{k-1} - m_{ref}) + \beta \alpha_x W_x^T W_x m^{k-1}) \quad (23)$$

donde

$$\delta d = F[m] - d_{obs} \quad (24)$$

- $F[m]$  - El operador geofísico no lineal aplicado al modelo  $m$
- $W_x$  - es la matriz de derivadas (diferencias finitas) en la dirección  $x$
- $I_m$  - es una matriz identidad de tamaño  $m \times m$
- $m^{k-1}$  - es el modelo evaluado en la iteración anterior

## 2.10. Biblioteca SimPEG

SimPEG (Simulation and Parameter Estimation in Geophysics) es una librería de código abierto desarrollada por investigadores, estudiantes y profesores en geofísica asociados con la UBC Geophysical Inversion Facility de la University of British Columbia (Canadá), escrita en el lenguaje de programación Python, diseñada específicamente para la simulación y estimación de parámetros en aplicaciones geofísicas.

En el contexto del método electromagnético transitorio (TEM), SimPEG actúa como el solucionador del problema directo. Su función es predecir la respuesta electromagnética teórica que generaría un modelo de subsuelo específico, resolviendo las ecuaciones de difusión electromagnética para una configuración de transmisor y receptor dada.

Para el caso específico de un medio estratificado horizontalmente (modelado 1D), SimPEG utiliza el módulo **Simulation1DLayered**. Este módulo no requiere una discretización volumétrica compleja (como elementos finitos 3D), sino que emplea soluciones semianalíticas basadas en la transformada de Hankel, o filtros digitales lineales, para calcular la convolución de la respuesta del sistema con la estructura de resistividades del subsuelo.

SimPEG proporciona las herramientas matemáticas para definir la función objetivo que guía la inversión. La librería permite calcular el desajuste de datos comparando los datos observados ( $d_{obs}$ ) con los simulados ( $d_{pred}$ ), normalizados por sus incertidumbres ( $\sigma_i$ ):

$$\phi_d(m) = \sum_{i=1}^N \left( \frac{d_{obs,i} - F[m]_i}{\sigma_i} \right)^2 \quad (25)$$

Donde  $F[m]$  es el operador de simulación directa ejecutado por SimPEG sobre el modelo de resistividades  $m$ . Esta capacidad de generar respuestas sintéticas rápidas y precisas es lo que permite a algoritmos externos, como los Algoritmos Genéticos, evaluar iterativamente la aptitud de cientos de modelos candidatos en cada generación.

## 2.11. Algoritmos genéticos

Según Goldberg (1989), los algoritmos genéticos constituyen métodos de búsqueda inspirados en la genética y la selección natural. Estos sistemas operan mediante la combinación de la supervivencia de los individuos más aptos y un intercambio de información que, aunque posee componentes aleatorios, sigue una estructura lógica similar a la evolución humana. En este proceso, cada nueva generación de soluciones de construye a partir de fragmentos de las versiones más exitosas de la generación anterior, permitiendo que el algoritmo utilice la información histórica de manera eficiente

para proponer nuevos puntos de búsqueda con un rendimiento superior.

Los algoritmos genéticos son utilizados para la optimización; en este caso lo que buscamos es minimizar la función objetivo de la inversión geofísica. Nosotros utilizamos los algoritmos genéticos para una búsqueda paramétrica que refiere a la búsqueda de las distintas resistividades para cada capa dentro de nuestra inversión. Los algoritmos genéticos tienen una población inicial la cual se genera aleatoriamente y básicamente son diferentes propuestas que serán mejoradas mientras avanzan las generaciones. La población se compone de individuos que serán estas propuestas individuales, los individuos se generan con base a la cantidad de parámetros que se quieran estimar.

Para avanzar de una generación a otra, la población hace un proceso de elección. El proceso más simple es por torneo que implica escoger una cantidad de individuos dentro de la población y compararla para escoger al mejor. Una vez que el torneo escoge a los mejores individuos se hace la cruce, la cual determinará la población de la siguiente generación.

En este proceso, se ordena a los individuos de menor a mayor ajuste y se escoge un punto de cruce aleatorio que es el punto en el que cortaremos al individuo y lo combinaremos con otro para generar, a partir de los dos individuos padres, dos individuos hijos nuevos.

Padre 1	00100 — 101
Padre 2	01001 — 010

Se elige un punto de cruce aleatorio; en este ejemplo es la cuarta posición comenzando a contar desde el cero y de izquierda a derecha, es decir, en el cambio de celda después del 00100 para el padre 1.

Hijo 1	00100 — 010
Hijo 1	01001 — 101

Nótese que se tiene la misma cantidad de hijos que de padres para conservar siempre el número de individuos en la población.

Luego de la cruce se aplica una mutación en cada generación; el tipo de mutación que se hace varía con el problema que se está resolviendo y se utiliza una probabilidad de mutación generalmente baja, menor al diez por ciento.

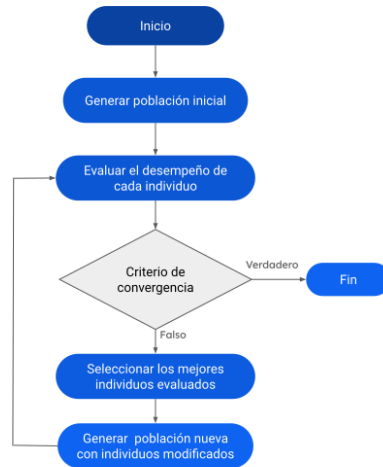


Figura 10: Diagrama de flujo de la computación evolutiva

También se puede aplicar elitismo, que significa conservar a los mejores individuos de cada generación. Se elige la cantidad de individuos dependiendo el problema pero normalmente se quiere solo conservar al mejor individuo para poder aprovechar la explotación; si conservas muchos individuos dejarían de lado la exploración.

Para entender la exploración y la explotación es importante hablar del espacio exploratorio, que es el lugar en donde nuestras posibles soluciones se encuentran; imaginemos un espacio en dos dimensiones finito al que le asignamos coordenadas  $(x,y)$  en el cual se encuentra nuestra solución.

En la búsqueda de una solución, el escenario ideal consistiría en evaluar exhaustivamente todo el espacio de parámetros para garantizar el hallazgo del óptimo global. Sin embargo, en la práctica, este espacio suele ser demasiado grande o presentar múltiples óptimos locales que pueden provocar un estancamiento. Ante la inviabilidad de una búsqueda exhaustiva, los algoritmos genéticos inician proponiendo un conjunto de soluciones aleatorias distribuidas en distintas posiciones; estas conforman los individuos de nuestra población inicial.

El algoritmo evoluciona dando mayor peso a los individuos con mejor ajuste para refinar la búsqueda en su entorno, proceso conocido como explotación. No obstante, para evitar el estancamiento en un óptimo local, esta fuerza debe equilibrarse simultáneamente con la exploración continua de nuevas regiones en el espacio de parámetros.

## 2.12. Algoritmos Genéticos de Codificación Real (RCGA)

La transición del espacio de búsqueda binario al real representa una evolución significativa en la capacidad de los algoritmos genéticos para resolver problemas de ingeniería. Mientras que el esquema binario clásico descompone el espacio en una

cuadrícula discreta, los RCGAs operan en el dominio continuo utilizando vectores de números de punto flotante como cromosomas. Esta representación permite al algoritmo manejar la topología del problema de forma directa, eliminando la necesidad de procesos de decodificación que a menudo introducen sesgos o pérdida de resolución (Michalewicz, 1992).

Una de las justificaciones principales para el uso de RCGAs en la optimización de parámetros geofísicos es la superación de los llamados acantilados de Hamming (Goldberg, 1989). En la codificación binaria, dos valores numéricamente cercanos pueden tener representaciones de bits radicalmente opuestas, lo que genera una discontinuidad artificial en el espacio de búsqueda. Los RCGAs, al emplear operadores numéricos, aseguran que pequeñas variaciones en el genotipo se traduzcan en cambios proporcionales en el fenotipo, facilitando una exploración más suave y coherente (Herrera et al., 2003).

En esta investigación se implementó un Algoritmo Genético de Codificación Real (RCGA, por sus siglas en inglés). La aproximación de codificación real es particularmente natural al abordar la optimización de parámetros con variables en dominios continuos. En un RCGA, el cromosoma que representa a un individuo abandona las cadenas de bits y se define directamente como un vector de números de punto flotante. Aunque el proceso de selección de individuos sigue basándose puramente en los valores de la función de aptitud (fitness), los operadores evolutivos tradicionales deben sustituirse por operaciones de naturaleza numérica.

Para la generación de nuevos individuos, se emplean operadores de cruce diseñados para el espacio continuo. El operador más destacado en la literatura actual es la Cruza Binaria Simulada (SBX), la cual utiliza un índice de distribución para controlar la varianza de la descendencia en relación con la posición de los padres, permitiendo un equilibrio dinámico entre la búsqueda global y el refinamiento local (Deb & Agrawal, 1995).

De manera análoga a la cruce, la mutación en un RCGA descarta la simple inversión lógica de estados discretos. Entre las metodologías destacan, la mutación de distribución normal gaussiana y la mutación no uniforme.

### **2.13. Librería para implementar algoritmos genéticos (DEAP)**

La librería Distributed Evolutionary Algorithms in Python (DEAP) es un marco de trabajo de código abierto diseñado para el desarrollo, prototipado y prueba rápida de algoritmos de computación evolutiva. A diferencia de otras librerías que encapsulan la lógica en estructuras rígidas, DEAP se distingue por una arquitectura desacoplada que prioriza la transparencia y la modularidad explícita. Este diseño facilita la implementación de Algoritmos Genéticos (AG), Programación Genética (GP), Estrategias Evolutivas (ES) y optimización multiobjetivo (como NSGA-II o SPEA2).

DEAP trabaja principalmente con una representación de vectores (listas o arreglos) para los individuos, y el tipo de datos dentro de este vector depende enteramente de la configuración dada por el usuario.

DEAP puede usar representaciones binarias (Algoritmo Genético Simple), esto requiere una configuración explícita y se utiliza principalmente cuando el problema de optimización es de naturaleza discreta.

La arquitectura de DEAP se fundamenta en el uso extensivo de la metaprogramación en Python para separar la estructura de los datos de los operadores genéticos. El funcionamiento de la librería se articula a través de dos componentes principales:

El módulo `creator`: Permite la generación dinámica de clases y tipos de datos en tiempo de ejecución. Esto posibilita que el programador defina individuos con atributos arbitrarios (listas, arreglos de NumPy, árboles binarios) y asocie métricas de aptitud personalizadas (minimizadas, maximizadas o multiobjetivo) sin alterar el código fuente de la librería.

La clase `Toolbox`: Actúa como un contenedor de operadores persistentes. Mediante mecanismos de registro (`register`), el usuario define los operadores de inicialización, selección, cruce (`crossover`) y mutación. Esta abstracción permite intercambiar algoritmos complejos modificando únicamente la configuración del `Toolbox`, manteniendo intacto el bucle principal de ejecución.

Una de las características técnicas más relevantes de DEAP en el contexto de la investigación es su soporte nativo para la paralelización. Dado que la evaluación de la función objetivo suele ser el paso más costoso computacionalmente en la optimización evolutiva, DEAP permite distribuir estas evaluaciones a través de múltiples núcleos de CPU o clústeres mediante la sustitución trivial de la función `map` estándar, función que permite aplicar una función específica a cada uno de los elementos de una estructura de datos (como una lista o un diccionario) sin necesidad de escribir un bucle `for` explícito, por implementaciones paralelas.

La librería también integra algunas herramientas estadísticas (`Logbook` y `Statistics`) que facilitan la recopilación estructurada de métricas generacionales (promedio, desviación estándar, mínimo, máximo).

## 2.14. Paralelismo

El paralelismo es un concepto fundamental en la informática y la ingeniería que se refiere a la ejecución simultánea de múltiples instrucciones o procesos.

El paralelismo se entiende como la aptitud de un sistema para abordar simultáneamente múltiples componentes de un mismo problema. Según Quinn (2004), esto im-

plica el uso de diversas unidades de procesamiento que operan de forma sincronizada para superar las limitaciones de velocidad de un solo elemento de cómputo. En el contexto de los algoritmos de búsqueda, esta simultaneidad a nivel de tarea permite que diferentes puntos de exploración sean evaluados al mismo tiempo, logrando una reducción significativa en los tiempos de ejecución globales.

Su objetivo principal es reducir el tiempo de ejecución de una tarea compleja dividiéndola en subtareas más pequeñas que se pueden resolver al mismo tiempo, típicamente utilizando múltiples unidades de procesamiento (como núcleos de CPU o GPU).

En el contexto del trabajo, el paralelismo se aplica como paralelismo de tareas donde cada individuo de la población del algoritmo genético (AG) representa una solución candidata y puede ser evaluado de forma independiente en un núcleo de CPU diferente, lo cual acelera drásticamente el AG, ya que el cálculo de la función de costo es la parte más intensiva y puede distribuirse.

### **3. Área de Estudio**

Las dos áreas de interés, la zona este y la zona oeste, en el área de estudio AEM se ubican en el norte del Valle de Sacramento, California, y abarcan aproximadamente 1700 kilómetros cuadrados. Estas áreas se encuentran en partes de dos condados: Butte y Glenn. Las principales localidades del área son Chico (86,187 hab.), Orland (7,291 hab.) y Durham (5,518 hab.) Según la Oficina del Censo de EE. UU (2018). También se encuentran cercanas Oroville, Corning y Willows. La precipitación y la escorrentía de riego dentro del área de estudio alimentan la cuenca del río Sacramento a través de numerosos afluentes, siendo los más importantes Pine Creek, Burch Creek, Rock Creek, Big Chico Creek, Little Chico Creek, Butte Creek y Little Dry Creek. El área cuenta con un suministro de agua subterránea localizada principalmente en los acuíferos formados por las arenas y gravas de las formaciones Tehama, Tuscan y Laguna, y los materiales aluviales no consolidados que la cubren. El flujo de agua subterránea se dirige hacia el río Sacramento desde ambos lados del valle. El uso del suelo combina agricultura de riego, pastoreo, forestación y uso municipal. El riego proviene de pozos de agua subterránea y de fuentes de agua superficial (Asch et al., 2019).

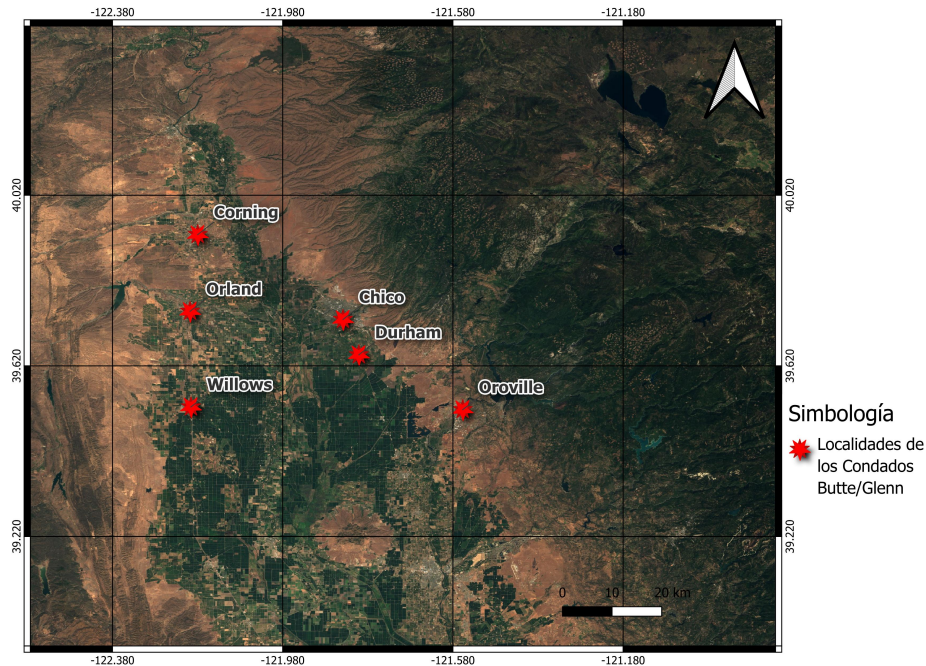


Figura 11: Localidades de los condados de Butte/Glenn

### 3.1. Geología

La geología del área es predominantemente volcánica, con una mezcla de unidades aluviales y fluviales que incluyen limo, arena, grava, y estratos de arcilla, limolita, arenisca y conglomerado (Page, 1986). Los valles están ocupados por depósitos de aluvión y coluviones no diferenciados ocupan gran parte de los valles incisos de las tierras altas.

Un gran paleovalle conforma gran parte del Valle Central (VC), compuesto por depósitos cuaternarios aluviales y marinos no consolidados. Las tierras altas y cordilleras circundantes están formadas por rocas volcánicas del Mioceno y Plioceno, así como rocas del Cretácico y Jurásico (Blake et al., 1992; Saucedo y Wagner, 1992). Entre las principales unidades cuaternarias destacan los depósitos aluviales, lacustres, de playas y terrazas (CA-DWR, 2014). El noreste de la cuenca representa el límite sur de los volcanes y mesetas de la Cordillera de las Cascadas (Planert y Williams, 1995). En el Valle Central se han cartografiado terrazas fluviales con edades entre 10,000 y 1.25 millones de años (Steele, 1980). La recarga del Acuífero del Valle Central proviene casi exclusivamente de la escorrentía de arroyos de montaña, ya que la mayor parte de la precipitación se evapora (Planert y Williams, 1995).

La geología superficial presenta una mezcla compleja de depósitos volcánicos, aluviales y coluviales indiferenciados. El sistema acuífero del Valle Central (AVC), principal reservorio de aguas subterráneas del área, es un sistema de relleno de cuenca con agua dulce hasta profundidades de 760 m (Planert y Williams, 1995). Dependiendo del nivel freático, puede ser confinado o no confinado, predominando este último en aguas someras. En la sección norte del Valle de Sacramento, los sedimentos tienen una

profundidad media entre 300 y 600 m (Planert y Williams, 1995).

Los depósitos cuaternarios están dominados por aluviones de la Formación Modesto (**Qm**), depósitos de cuenca (**Qb**) y de canal (**Qa**) (Burnett et al., 1969; Blake et al., 1992; Saucedo y Wagner, 1992). En las tierras altas orientales afloran la Formación Red Bluff del Pleistoceno (**Qrb**), la Formación Riverbank (**Qr**), la Formación Toscana del Plioceno (**Tt**), pequeñas cantidades de las Tobas de Oroville del Plioceno (**QPto**), y rocas volcánicas del Plioceno (**Pv**). También se encuentran depósitos de deslizamientos del Holoceno (**Qls**), rocas volcánicas del Mioceno-Plioceno (**MPv**), el Basalto Lovejoy del Mioceno (**Tl**), gravas auríferas del Eoceno (**Tg**) y la Formación Chico del Cretácico (**Kc**) (Blake et al., 1992; Saucedo y Wagner, 1992).

La geología terciaria está compuesta por una secuencia compleja de sedimentos marinos y no marinos, que incluyen areniscas, limolitas, lutitas, lutitas limosas, conglomerados, lahares volcánicos y basaltos. La Formación Tehama (**Tte**) es una unidad masiva y extensa que se adelgaza hacia el este, compuesta principalmente por depósitos aluviales y fluviales provenientes de las montañas costeras. Contiene areniscas y limolitas, además de lentes de conglomerados con guijarros y cantos rodados. Se intercala con la Formación Toscana (**Tt**) en el centro del Valle de Sacramento, alcanzando un espesor máximo de 610 m (CA-WDR, 2014).

La Formación Toscana (**Tt**), que se identifica como cuatro unidades litológicamente similares, incluye lahares, conglomerados volcánicos, areniscas volcánicas, limolitas y tobas pómez, también de gran extensión. Algunas de sus unidades han sido clasificadas por algunos autores como del Pleistoceno (CA-WDR, 2014). La Tt se intercala con la Tte y se adelgaza hacia el oeste.

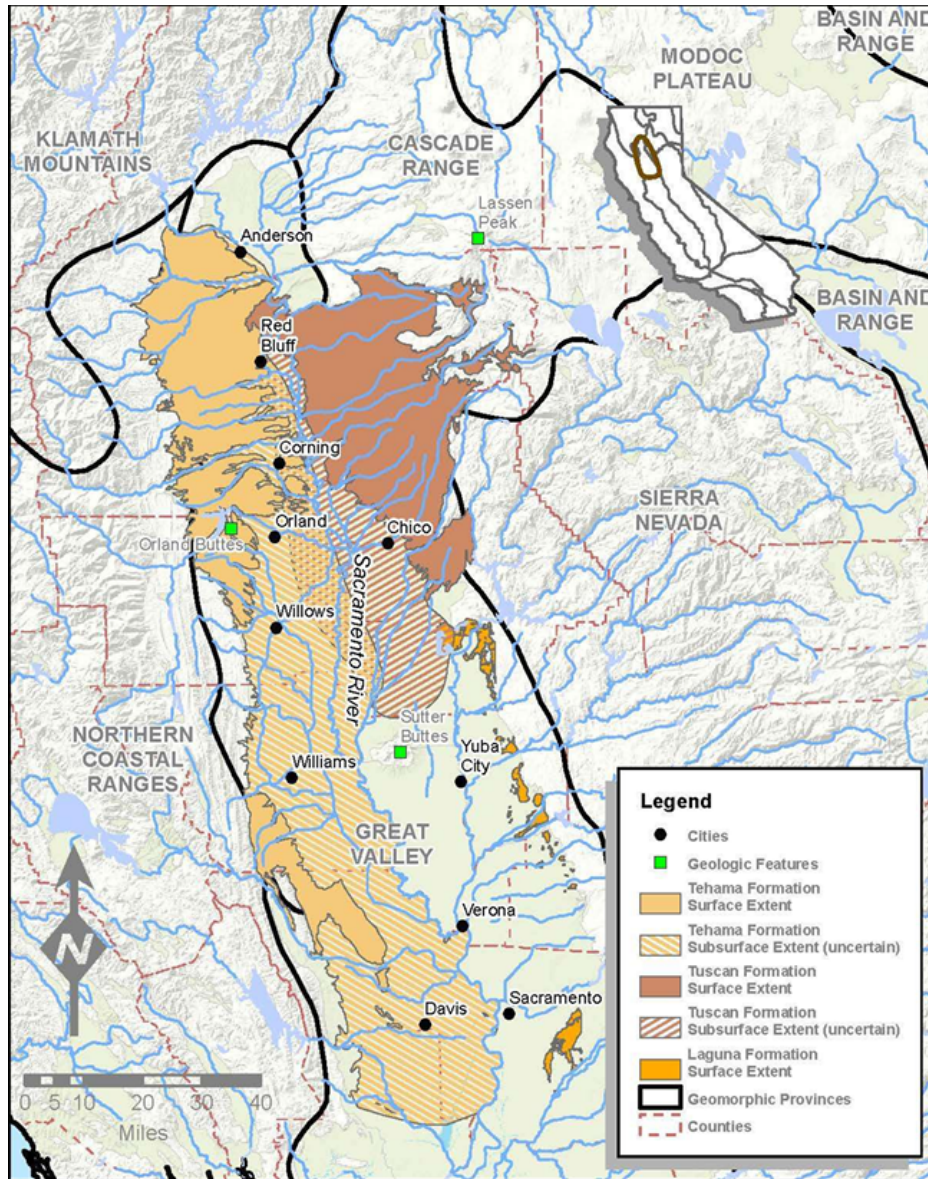


Figura 12: Extensión superficial y subsuperficial aproximada de las formaciones Tehama y Toscana, y extensión superficial únicamente de la formación Laguna. [9, p. 24].

### 3.2. Características hidrogeológicas

Las principales características hidráulicas están asociadas al río Sacramento, los arroyos Big Chico y Butte, y estructuras de aguas superficiales como el canal Glenn-Colusa y el lago Oroville. El río Sacramento es el principal cauce de la zona, atravesando el área cubierta por los vuelos AEM. Entre las estaciones hidrométricas USGS 11377100 (cerca de Red Bluff) y USGS 11389500 (en Calusa), el caudal medio diario anual disminuyó de 560.1 a 515.1 m<sup>3</sup>/s durante el año hidrológico 2017 (USGS, 2018a). Aunque el informe no analiza directamente esta disminución, podría deberse a la recarga en el Área de Conservación de Aguas y a las desviaciones hacia canales.

La conexión entre aguas superficiales y subterráneas es compleja debido a las múltiples fuentes de recarga superficial. En la primavera de 2018, los niveles freáticos indicaron un flujo subterráneo general hacia el sur, con descargas al río Sacramento (CA-DWR, 2018). Datos del Sistema Nacional de Información Hídrica (USGS, 2018b) muestran que los niveles de agua subterránea más someros (<3 m) se encuentran generalmente al sur del área de estudio, cerca del río Sacramento o aguas abajo en gran medida adyacentes del cuerpo de Thermalito Afterbay. Los niveles moderadamente someros (3–15 m) son comunes a lo largo del río Sacramento y parecen reflejar una distribución más regional. Niveles entre 15 y 30 m se hallan cerca de las estribaciones de la Cordillera de las Cascadas/Sierra Nevada, y los niveles más profundos (>30 m) se localizan principalmente en zonas montañosas.

Según Planert y Williams (1995), las rocas volcánicas y metamórficas que rodean y subyacen al valle están consolidadas y son casi impermeables, por lo que permiten escaso flujo de agua. En cambio, el AVC está compuesto por arenas y gravas con limos y arcillas, todos provenientes de la erosión de las montañas circundantes.

El sistema acuífero es heterogéneo, con condiciones no confinadas en los primeros 100 metros, y confinadas en mayor profundidad, debido a la presencia de horizontes arcillosos.

### 3.3. Datos de sondeos

Se utilizaron datos provenientes de registros litológicos y geofísicos de fondo de pozo. El reporte de estos datos (Kang et al., 2021) indica lo siguiente: “Esta información fue recopilada por Todd Greene, Profesor Asociado de Ciencias Geológicas y Ambientales en la Universidad Estatal de California, Chico. Los registros geofísicos se entregaron el 19 de octubre de 2018 y los litológicos el 18 de octubre de 2018, con adiciones el 2 de noviembre del mismo año.”

En total, dentro del área de estudio AEM de Butte, se recopilaron 362 sondeos con información litológica y 183 con datos geofísicos, cuya ubicación se muestra en la Figura 13

Dado que los registros de pozo de resistividad provienen de diferentes épocas, operadores y equipos, es importante considerar posibles errores de calibración y funcionamiento del equipo al analizar los valores absolutos. Tradicionalmente, los registros de resistividad se usaban solo para observar deflexiones relativas y no requerían una calibración precisa. Esto ha llevado a la existencia de numerosos registros potencialmente no calibrados. Sin embargo, siguen teniendo valor científico por su capacidad de mostrar variaciones relativas en la resistividad de las formaciones. Los registros se utilizan para comparaciones cualitativas con los datos AEM.

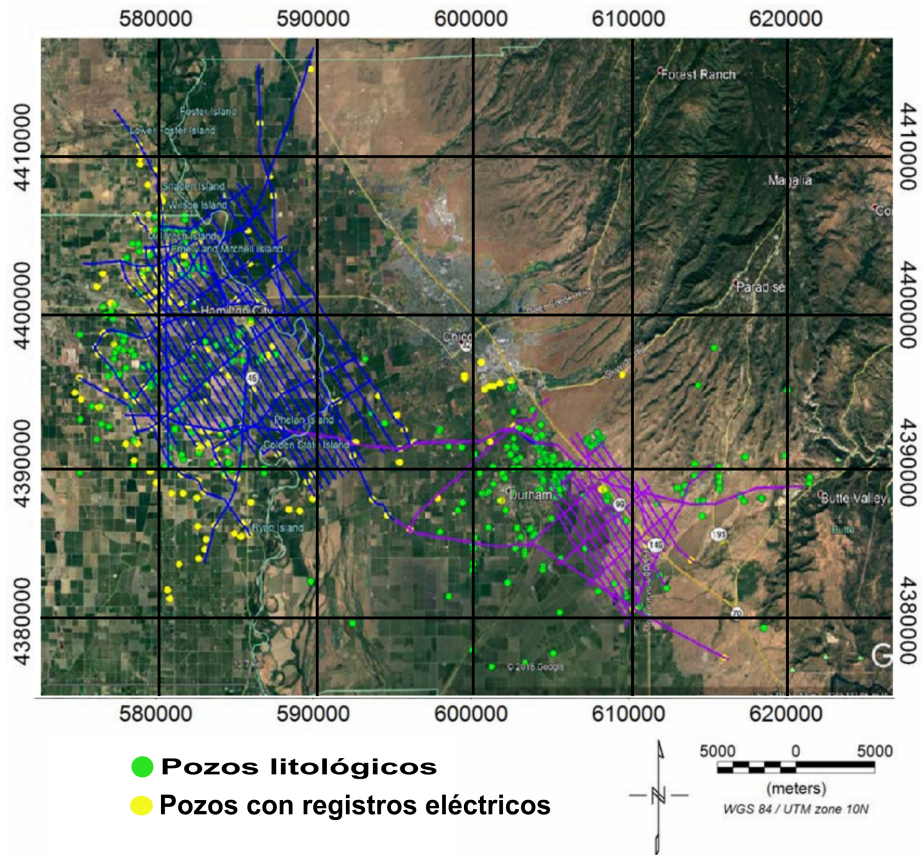


Figura 13: Ubicación de los sondeos cerca del área de prospección AEM de Butte. Los círculos verdes representan sondeos con información litológica y los círculos amarillos, sondeos con información geofísica. Las líneas de vuelo 312 están en azul y las 304 en morado. Tomada de [1, p. 24].

### 3.4. Adquisición y procesamiento de datos

La principal fuente de ruido en los sondeos electromagnéticos geofísicos son otros dispositivos electromagnéticos que forman parte de la infraestructura típica de servicios. Estos incluyen, por ejemplo, casas, líneas de alta tensión, ferrocarriles, oleoductos y bombas de agua. Antes de la adquisición de datos AEM en los condados de Butte y Glenn, se localizaron tres tipos de servicios (oleoductos, ferrocarriles y líneas de alta tensión).

Las ubicaciones de las líneas de vuelo se convirtieron de una cuadrícula regularmente espaciada a una con líneas de vuelo optimizadas para evitar el acoplamiento electromagnético, evitando en el camino los objetos anteriormente mencionados. También se modificaron las trayectorias de las líneas de vuelo para volar más cerca de las ubicaciones de pozos de sondeo conocidos.

Al concluir el proceso de diseño, las líneas de vuelo AEM de Butte se dividieron en

líneas de vuelo de bloque y de reconocimiento. Las líneas de vuelo de reconocimiento tenían aproximadamente 22 millas de longitud (36 km) en su punto más largo y alrededor de 3.1 millas (5 km) en su punto más corto. Los vuelos de bloque tenían líneas de vuelo de hasta 12 millas de longitud (aproximadamente 20 km) con líneas de unión de aproximadamente 8.6 millas de longitud (aproximadamente 14 km) para el área del bloque occidental. Las líneas de vuelo del bloque oriental eran de aproximadamente 5.5 millas de longitud, (aproximadamente 9 km) con líneas de amarre de 4.9 millas de largo (aproximadamente 8 km). Las líneas de vuelo en bloque estaban separadas por aproximadamente 1500 pies (aproximadamente 450 m a 500 m).

Los datos de AEM se adquirieron utilizando los sistemas electromagnéticos aéreos SkyTEM304M (304M) y SkyTEM312 (312) (SkyTem Airborne Surveys Worldwide, 2018).

El 304M es un sistema TEM de marco rígido y momento magnético dual (bajo y alto). El área de la bobina Tx del 304M es de 342 m<sup>2</sup> y la bobina contiene cuatro vueltas. Se hace pasar una corriente pico de 9 amperios a través de una vuelta en el Tx para mediciones de Bajo Momento y una corriente pico de 110 amperios se pasa a través de las cuatro vueltas para mediciones de Alto Momento. Esto resulta en momentos magnéticos Tx de Bajo y Alto de 3,000 Amperio-metro-cuadrado ( $Am^2$ ) y 150,000  $Am^2$ , respectivamente.

El SkyTEM312 utiliza el mismo marco que el 304M, pero con diferentes parámetros eléctricos y cableado del transmisor. Se pasa una corriente máxima de 6 amperios a través de dos vueltas en el Tx para mediciones de Bajo Momento y una corriente máxima de 110 amperios a través de las doce vueltas para mediciones de Alto Momento. Esto resulta en momentos magnéticos Tx máximo de Bajo y Alto de aproximadamente 4,100 amperios metro cuadrado ( $Am^2$ ) y 450,000  $Am^2$ , respectivamente.

Los sistemas SkyTEM304M y 312 utilizan un Rx desfasado colocado ligeramente detrás del Tx (Figura 14), lo que resulta en una posición de 'acoplamiento nulo', que es un lugar donde la intensidad del campo primario del transmisor del sistema se minimiza. Esto es deseable para minimizar la amplitud del campo primario en el Rx y maximizar la sensibilidad del Rx a los campos secundarios. La bobina vertical (Z) de Rx de múltiples vueltas de los modelos 304M y 312 tiene un área efectiva de 105 m<sup>2</sup>. Además del Tx y el Rx que constituyen el instrumento TEM, los modelos 304M y 312 también están equipados con un magnetómetro de campo total y sistemas de adquisición de datos para ambos instrumentos. Los 304M y 312 también incluyen dos altímetros láser, inclinómetros y receptores de sistema de posicionamiento global diferencial (DGPS). Los datos de posición de los receptores DGPS montados en el marco son registrados por el sistema de adquisición de datos AEM. El magnetómetro incluye un tercer receptor DGPS cuyos datos de posición son registrados por el sistema de adquisición de datos del magnetómetro.

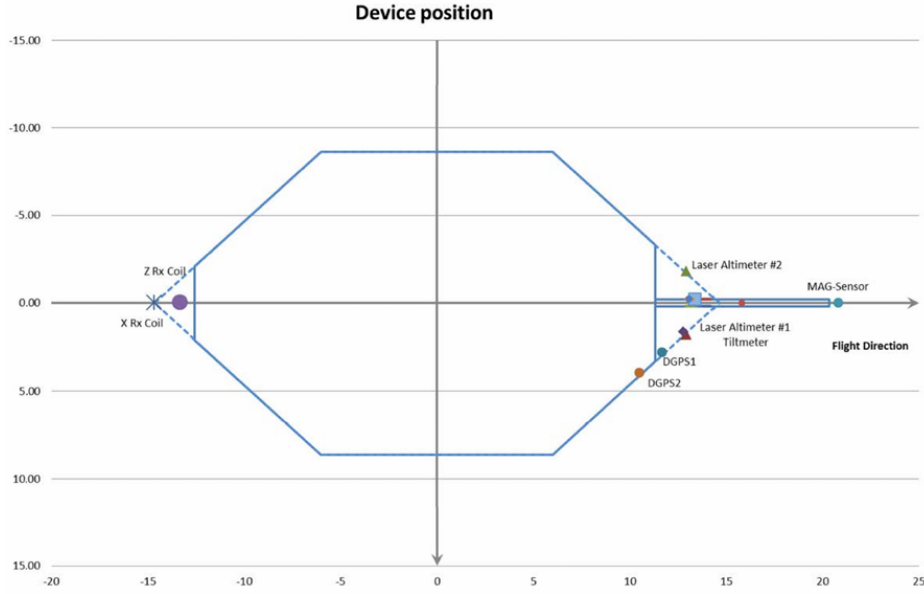


Figura 14: Marco SkyTEM304M/312, incluyendo ubicaciones de instrumentación y los ejes X y Y. Las distancias de los ejes están en metros. Ubicaciones de la instrumentación listadas en Tabla 1. Tomada de [1, p. 29].

La Figura 14 muestra la geometría del sistema vista en planta desde el eje z positivo. El magnetómetro está ubicado en un brazo en la parte delantera del marco (lado derecho de la imagen). La bobina Tx está ubicada alrededor del marco octogonal y la bobina Rx está ubicada en la parte trasera del marco (lado izquierdo de la Figura).

El sistema de coordenadas utilizado por el 304M y 312 define la dirección (+x) como la dirección del vuelo, la dirección (+y) se define a 90 grados a la derecha y la dirección (+z) es hacia abajo. El centro del bucle del transmisor, montado en el marco octogonal de SkyTEM, se utiliza como el origen en referencia a las posiciones de los instrumentos. La Tabla 1 lista las posiciones de los instrumentos y la Tabla 2 lista las esquinas del bucle del transmisor.

	DGSP 1	DGSP 2	Inclinómetro 1	Inclinómetro 2	Altímetro 1	Altímetro 2	Sensor magnético	Bobina Rx
X	11.68	10.51	12.79	12.79	12.94	12.94	20.50	-13.25
Y	2.79	3.95	1.64	1.64	1.79	-1.79	0.00	0.00
Z	-0.16	-0.16	-0.12	-0.12	-0.12	-0.12	-0.56	-2.00

Tabla 1: Posiciones de los instrumentos en el marco SkyTEM304M/312, utilizando el centro del marco como origen, en metros. Tomada de [1, p. 30].

Esquinas de Tx	1	2	3	4	5	6	7	8
X	-12.55	-6.03	6.03	11.34	11.34	6.03	-6.03	-12.55
Y	-2.10	-8.63	-8.63	-3.31	3.31	8.63	8.63	2.10

Tabla 2: Posiciones de las esquinas de la bobina del transmisor SkyTEM304M/312, usando el centro como el marco de origen. Tomada de [1, p. 30].

El DGPS y el magnetómetro montados en el marco del 304M y 312 requieren el uso de estaciones base, que se encuentran en el suelo y están posicionadas en un área con bajo ruido cultural. En este caso, estos instrumentos se ubicaron en el aeropuerto cerca de Orland, California. Los datos del magnetómetro y las estaciones base de DGPS se descargaron cada día después del final de los vuelos de AEM del día. Las estaciones base de DGPS y magnetómetros se colocaron en la ubicación del sistema de coordenadas Universal Transverse Mercator (UTM) que se indica en la Tabla 3. El referente geodésico horizontal utilizado es el Datum Norteamericano de 1983 (NAD83). Todas las elevaciones provienen del Conjunto de Datos de Elevación Nacional del USGS, referenciado al Datum Vertical Norteamericano de 1988; con pies como unidad de medida.

Ubicación de los instrumentos de la estación base DGPS y del campo magnético en el aeropuerto de Orland, CA:

Instrumento	Este (m)	Norte (m)	Zona UTM
Estación Base Magnetómetro	573399	4396864	10 N
Estación Base DGPS	573367	4396876	10 N

Tabla 3: Ubicación de los instrumentos de la estación base DGPS y del campo magnético en el aeropuerto de Orland, CA. Tomada de [1, p. 30].

Para este proyecto, el 304M fue volado a una velocidad promedio de 56 millas por hora (90.8 kilómetros por hora) a una altura de vuelo promedio de 44.5 m sobre la superficie terrestre, utilizando el sistema de carga por eslinga de un helicóptero Eurocopter AS350.

Todos los sistemas SkyTEM están calibrados en un sitio de prueba en tierra en Lyngby, Dinamarca, antes de ser utilizados para trabajos de producción (Grupo de Hidrogeofísica de la Universidad de Aarhus, 2010; Grupo de Hidrogeofísica de la Universidad de Aarhus, 2011; Foged et al., 2013). El proceso de calibración implica adquirir datos con el sistema suspendido a diferentes altitudes, desde 16 pies hasta 164 pies (5 m a 50 m), sobre el sitio de Lyngby. Los datos adquiridos se procesan y se aplica un factor de escala (tiempo y amplitud) para que el proceso de inversión produzca el modelo que aproxima la geología conocida en Lyngby.

Antes del proyecto en el Valle de San Joaquín, se instaló y ensambló el sistema SkyTEM312 en un helicóptero en el aeropuerto de Orland. Se realizaron vuelos de prueba y calibración, y se verificaron los instrumentos de posicionamiento y la forma

de onda del transmisor. Una prueba de alta altitud confirmó el rendimiento del sistema. Después de cada vuelo, se controló la calidad de los parámetros operativos con software del propietario. Tras 312 vuelos, el SkyTEM312 fue desmantelado y se instaló y configuró el sistema 304M, realizando controles similares antes de la adquisición de datos.

Aproximadamente 361 millas de línea (585 kilómetros) fueron adquiridas por el SkyTEM312 sobre el área de estudio en condados de Butte y Glenn al oeste de Chico del 30 de noviembre al 2 de diciembre de 2018. Luego, el 3 de diciembre de 2018, aproximadamente 138 millas de línea (224 kilómetros) fueron adquiridas por el SkyTEM304M al sureste de Chico. El campo Haigh en el aeropuerto de Orland se utilizó para el aterrizaje y el re-abastecimiento entre los vuelos de producción. Un mapa de adquisición de datos se presenta en la Figura 15 con las líneas de vuelo agrupadas por fecha de adquisición y la Tabla 4 enumera las fechas de adquisición, los vuelos y la cantidad adquirida en cada día.

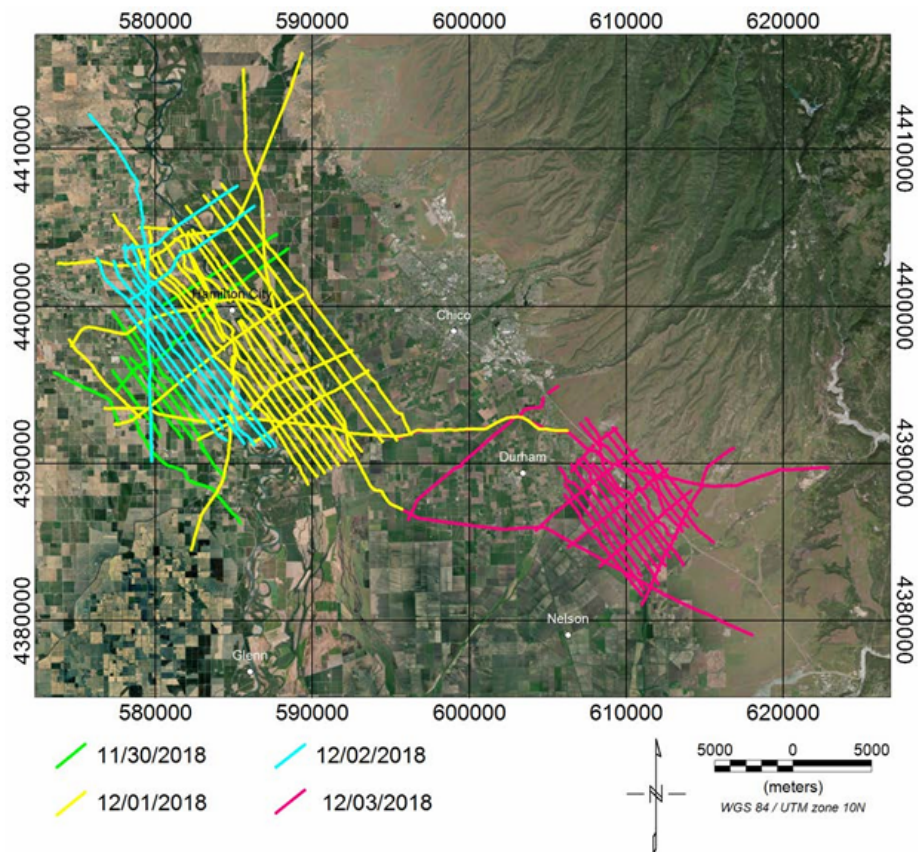


Figura 15: Líneas de vuelo de Butte AEM agrupadas por fecha de adquisición. Tomada de [1, p. 32].

Fecha	Vuelo	Sistema	Distancia (km)
30/Noviembre/2018	1	312	26.7
30/Noviembre/2018	4	312	53.2
1/Diciembre/2018	1	312	199.9
1/Diciembre/2018	2	312	179.2
2/Diciembre/2018	1	312	125.5
3/Diciembre/2018	2	304M	158.3
3/Diciembre/2018	3	304M	65.9
Total	7		808.7

Tabla 4: Producción de la línea de vuelo de Butte AEM por vuelo. Tomada de [1, p. 32].

La altura del sistema se especificó en 30-35 metros sobre el nivel del suelo; sin embargo, debido a juicios de seguridad y otros del piloto, las alturas de vuelo varían. El objetivo es mantener una altura lo más baja posible en la ventana de 25 a 50 m sobre el nivel del suelo. En el conjunto de datos AEM de Butte, la altura promedio fue de 44.5 m sobre el nivel del suelo, con un mínimo de 22.3 m sobre el nivel del suelo y un máximo de 236.4 m. Las alturas de vuelo máximas se encontraron sobre grandes líneas eléctricas. Los datos contaminados por las líneas eléctricas serán eliminados del conjunto de datos antes de la inversión debido al acoplamiento electromagnético, por lo que no afectarán al producto final.

La velocidad determina la distancia entre las muestras en el suelo. Sin embargo, hay un compromiso entre el costo del sondeo y la velocidad del sistema relacionada con la huella del sistema. En muchos sondeos, la velocidad especificada es de 100 km/h. El factor crítico en la velocidad de vuelo es mantener una velocidad en la que el sistema esté lo más nivelado posible. Esto puede requerir que el piloto acelere en la dirección de sotavento o reduzca la velocidad en la dirección de barlovento. El piloto utiliza la pantalla de visualización de los ángulos de inclinación del sistema para ayudar a mantener esta velocidad. La velocidad promedio en el suelo de la encuesta fue de 90.8 km/h, con una velocidad mínima en el suelo de 0.3 km/h y una velocidad máxima en el suelo de 115.3 km/h.

Los ángulos del sistema son críticos para asegurar que se envíen datos de calidad a la inversión. La corriente inicial del Tx del sistema en el tiempo de apagado de 0.0 segundos es la imagen del tamaño del lazo en la superficie. Si el sistema está inclinado, esa imagen será menor que el tamaño original del TX. Los algoritmos de inversión pueden contabilizar  $\pm 10$  grados de ángulo al calcular el tamaño efectivo del Tx. Con este fin, es importante mantener el marco del Tx dentro de  $\pm 10$  grados. La posición del Rx también se ve afectada por el ángulo del sistema y cualquier desviación de la perpendicular impacta al incluir componentes fuera de la perpendicular. Como se mencionó, los algoritmos pueden contabilizar  $\pm 10$  grados en el ángulo del Rx. Tanto el ángulo X (en la dirección de vuelo) como el ángulo Y (perpendicular a la dirección

de vuelo) se verificaron durante la adquisición de AEM de Butte. Cuando el sistema está volando sobre obstáculos o al girar al final de una línea, los ángulos pueden ser superiores a  $\pm 10$  grados. Estos bordes de la línea de vuelo generalmente se eliminan del conjunto de datos de la encuesta antes de la inversión. Durante la adquisición de AEM de Butte, ambos ángulos estuvieron dentro de los rangos aceptables. El ángulo X tuvo un promedio de aproximadamente -1.5 grados con un mínimo de -29.8 grados y un máximo de 29.40 grados. La inclinación del ángulo Y tuvo un promedio de aproximadamente 0.63 grados con un mínimo de -18.9 grados y un máximo de 25.9 grados. Las inclinaciones máximas y mínimas ocurrieron alrededor de la infraestructura y no afectarán mucho los datos ya que gran parte de esa área será eliminada durante el procesamiento de desacoplamiento.

El sistema SkyTEM utiliza un sistema de doble momento (alto (HM) y bajo (LM)) y dos corrientes y formas de onda Tx diferentes. Estas formas de onda se registran antes y después de la adquisición para asegurar que no hayan ocurrido cambios durante la misma. Las figuras y gráficos de las formas de onda Tx de bajo momento (LM) y alto momento (HM) registradas para el sistema SkyTEM312, respectivamente. Las figuras y gráficos de las formas de onda Tx de bajo momento (LM) y alto momento (HM) registradas para el sistema SkyTEM304M, respectivamente. La fuente Tx de LM se utiliza para resaltar la geología muy cerca de la superficie y la fuente de corriente HM se utiliza para obtener más energía electromagnética a profundidad para caracterizar las unidades geológicas más profundas. La corriente debe ser estable durante el estudio, pero los cambios en la temperatura pueden afectar la resistencia del cable Tx y del circuito al aumentar o disminuir la salida de la corriente máxima. La corriente máxima se registra durante la adquisición de cada sondeo y se utiliza para ajustar la forma de onda Tx en la inversión. Para el estudio AEM de Butte, la corriente media de LM 304M fue de 9.07 amp con una corriente mínima de 9.06 amp y una corriente máxima de 9.08 amp. Para el HM 304M, la corriente media fue de 111.79 amp con una corriente mínima de 111.19 amp y una corriente máxima de 112.86 amp. Para el estudio de AEM de Butte con el sistema 312, la corriente media de LM fue de 5.95 amp con una corriente mínima de 5.94 amp y una corriente máxima de 5.97 amp. Para el HM 304M, la corriente media fue de 111.01 amp con una corriente mínima de 108.38 amp y una corriente máxima de 113.52 amp. Todos los momentos del sistema muestran estabilidad en la corriente y no presentaron problemas en las inversiones.

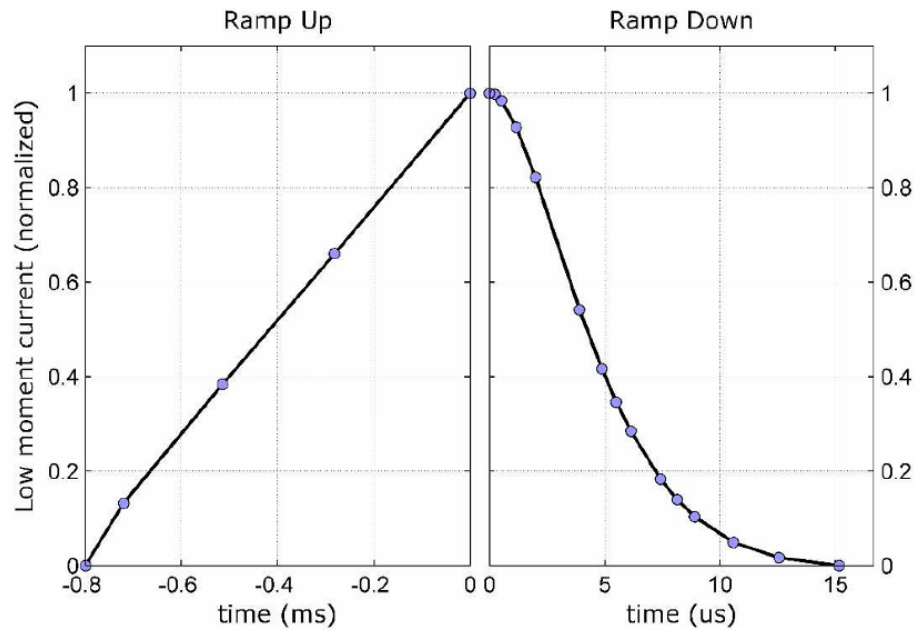


Figura 16: Gráfica de la forma de onda LM de 210 Hz para el sistema SkyTEM312 registrada durante el estudio de AEM de Butte. El aumento de corriente está a la izquierda y la disminución de la corriente para apagar está a la derecha. La corriente está normalizada. Tomada de [1, p. 38].

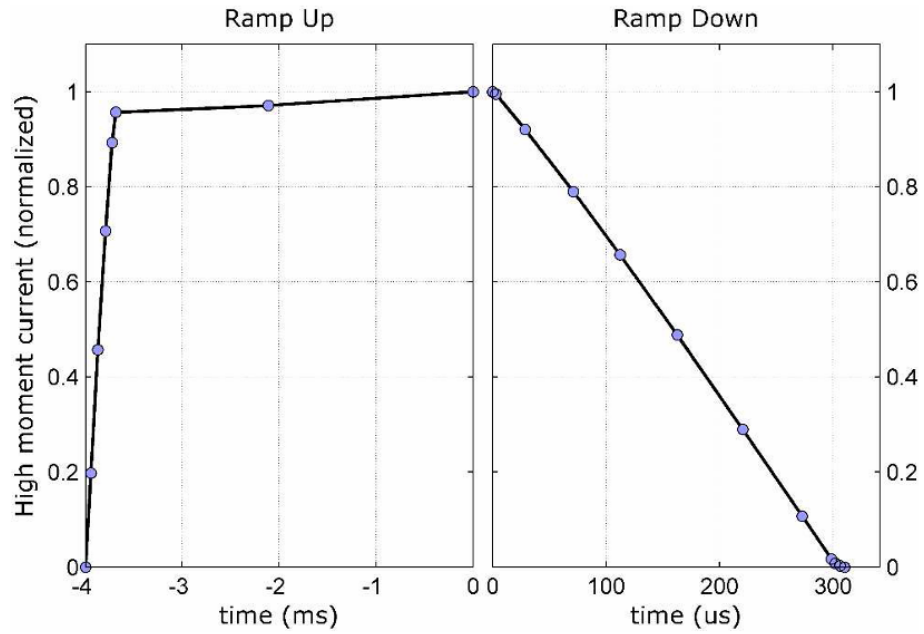


Figura 17: Gráfica de la forma de onda HM de 30 Hz para el sistema SkyTEM312 registrada durante el estudio de AEM de Butte. El aumento de corriente está a la izquierda y la disminución de la corriente para apagar está a la derecha. La corriente está normalizada. Tomada de [1, p. 39].

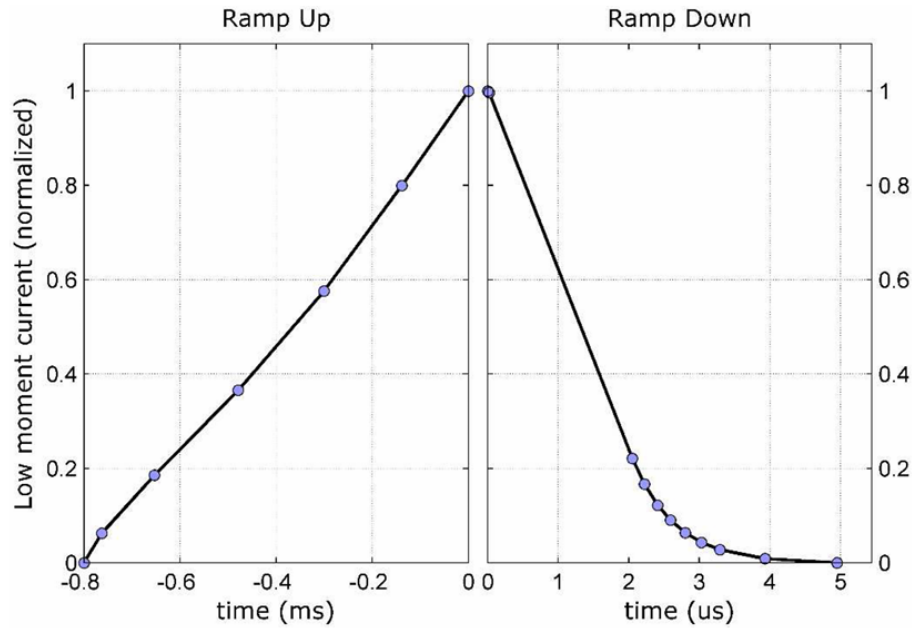


Figura 18: Gráfico de la forma de onda LM de 210 Hz para el sistema SkyTEM304M registrado durante el estudio de Butte AEM. El aumento de corriente está a la izquierda y la disminución de la corriente para apagar está a la derecha. La corriente está normalizada. Tomada de [1, p. 40].

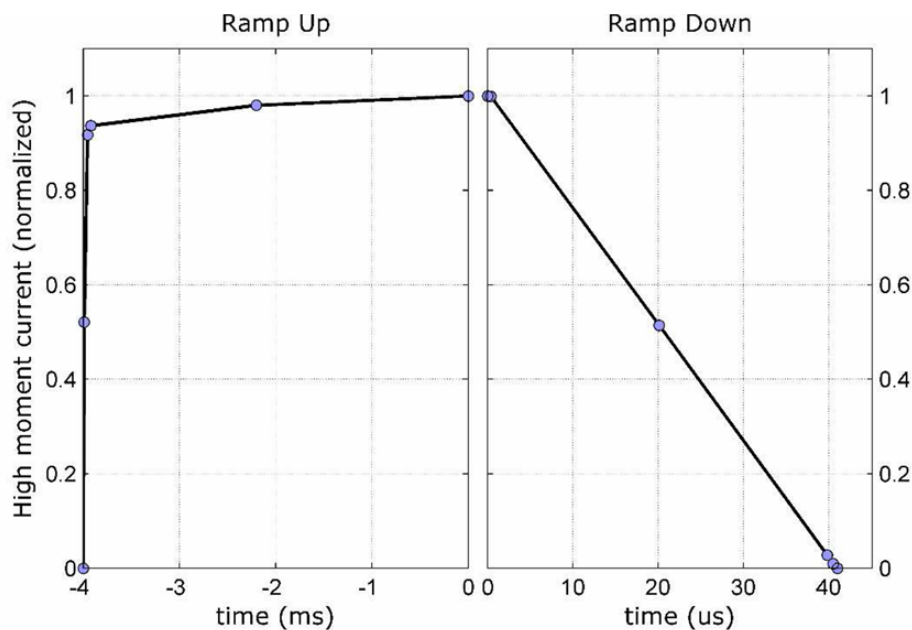


Figura 19: Gráfico de la forma de onda HM de 30 Hz para el sistema SkyTEM304M registrado durante el estudio de Butte AEM. El aumento de corriente está a la izquierda y la disminución de la corriente para apagar está a la derecha. La corriente está normalizada. Tomada de [1, p. 41].

Un procedimiento estándar de adquisición de datos de SkyTEM implica la revisión de los datos crudos adquiridos por SkyTEM en Dinamarca para la Compensación del Campo Primario (PFC) antes de continuar con el procesamiento de datos por parte de AGF (Aqua Geo Frameworks, LLC), (Schamper et al., 2014). El campo primario del transmisor afecta las ventanas de tiempo temprano registradas, que en el caso del momento bajo, son útiles para resolver la estructura de resistividad cerca de la superficie del suelo. El momento bajo utiliza una forma de onda de diente de sierra que se calcula y luego se utiliza en la corrección PFC para corregir las ventanas de tiempo temprano.

Los datos AEM recolectados por el 304M fueron procesados utilizando Aarhus Workbench versión 5.8.3 (en Aarhus Geosoftware (<https://www.aarhusgeosoftware.dk/workbench-overview>)) descrito por el Grupo de Hidrogeoquímica, Universidad de Aarhus (2011). Los algoritmos de procesamiento automático dentro del programa Workbench se aplican inicialmente a los datos AEM. Las ubicaciones DGPS fueron filtradas utilizando un filtro polinomial de segundo orden, paso a paso, de nueve segundos con un tiempo de ritmo de 0.5 segundos, basado en los parámetros de adquisición de vuelo. Los datos AEM fueron corregidos por desviaciones de inclinación desde el nivel y se aplicaron filtros también a ambas lecturas del medidor de inclinación con un filtro mediano de tres segundos y un filtro promedio de dos segundos. Los datos de altitud fueron corregidos utilizando una serie de dos filtros polinómiales. Las longitudes de ambos filtros polinómiales de octavo orden se establecieron en 15 segundos con longi-

tudes de desplazamiento de seis segundos. Los umbrales inferior y superior fueron de 1 y 100 metros, respectivamente.

Los filtros de promediado espacial trapezoidal fueron aplicados a los datos de AEM. Los tiempos utilizados para definir los filtros trapezoidales para el Bajo Momento fueron  $1.0 \times 10^{-5}$  seg,  $1.0 \times 10^{-4}$  seg, y  $1.0 \times 10^{-3}$  seg con anchos de 4, 7, y 18 segundos. Los tiempos utilizados para definir el trapezoide para el Alto Momento fueron  $1.0 \times 10^{-4}$  seg,  $1.0 \times 10^{-3}$  seg, y  $1.0 \times 10^{-2}$  seg con anchos de 10, 20, y 36 segundos. La distancia del sondeo trapezoidal se estableció en 1.0 segundos y la configuración izquierda/derecha, que requiere que el trapezoide esté completo en ambos lados, se activó. El factor de pico y el número mínimo de puertas se establecieron en un 25 por ciento para ambos sondeos. Por último, las ubicaciones de los sondeos promediados fueron sincronizadas entre los dos momentos.

Después de la implementación del filtrado automático, los datos AEM fueron examinados manualmente utilizando una ventana de tiempo móvil de dos minutos. Los datos fueron examinados en busca de posible acoplamiento electromagnético con infraestructura en la superficie y enterrada, así como por ruido en la ventana de tiempo tardía. Los datos afectados por estos factores fueron eliminados. Los datos AEM fueron luego invertidos utilizando un algoritmo de Inversión con Restricción Lateral (LCI) (Grupo de Hidrogeofísica de la Universidad de Aarhus, 2011). Se examinaron el perfil y los cortes de profundidad, y se enmascararon los acoplamientos electromagnéticos restantes del conjunto de datos. Después del procesamiento final, se retuvieron 109.3 millas de línea (177 km de línea) de datos 304M y 303 millas de línea (491 km de línea) de datos 312 para las inversiones finales del área de estudio AEM de Butte. Cada conjunto fue invertido por separado. Esto equivale a una retención de datos del 79 % para los datos 304 y del 84 % para los datos 312. Estas tasas de retención altas son el resultado de una planificación y diseño cuidadosos de las líneas de vuelo.

El canal de Intensidad de Ruido de Líneas de Distribución Eléctrica (PLNI) identifica posibles fuentes de ruido producidas por líneas de energía eléctrica. Las tuberías, a menos que tengan protección catódica, no son mapeadas por PLNI. El PLNI se genera mediante un análisis del contenido espectral de frecuencia de los datos brutos de SkyTEM (componentes Z). Se aplica una Transformada de Fourier a los datos de la ventana de tiempo del Bloque de Datos de Bajo Momento, evaluándose a 60 Hz para obtener la densidad espectral de amplitud del ruido.

Como se mencionó anteriormente, los sistemas SkyTEM 304M y 312 incluyen un magnetómetro de Campo Total cuya ubicación está listada en la Tabla 3. Los datos del Campo Total magnético pueden proporcionar información sobre la infraestructura así como sobre la geología. Estos datos también se utilizan en los esfuerzos de desacoplamiento.

Después del desacoplamiento inicial y el análisis LCI, se realizaron Inversiones con restricción espacial (SCI). La SCI utiliza datos EM a lo largo y a través de las líneas

de vuelo dentro de un criterio de distancia especificado por el usuario (Viezzoli et al., 2008). Los datos AEM de Butte se invirtieron utilizando modelos suaves de SCI con 40 capas, cada una con una resistividad inicial de 10 Ohm-m (equivalente a un medio semiespaciado de 10 ohm-m). Los grosores de las capas aumentan con la profundidad a medida que la resolución de la técnica disminuye. Las profundidades a los fondos de las 39 capas para el 304M se fijaron en 374.5992 metros, con grosores máximos de hasta aproximadamente 112 pies. Las profundidades a los fondos de las 39 capas para el 312 se fijaron en 1,804 pies, con grosores máximos de hasta aproximadamente 126 pies. La distancia de referencia espacial,  $s$ , para las restricciones se fijó en 328 pies (100 m) con una caída de ley de potencia de 0.75. Un ejemplo de un perfil obtenido con SCI se muestra en la Figura 20.

Además de los modelos de resistividad recuperados, los SCI también producen estimaciones de profundidad de investigación (DOI). El DOI proporciona una estimación general de la profundidad a la que los datos de AEM son sensibles a los cambios en la distribución de resistividad en profundidad (Christiansen y Auken, 2012). Se calculan dos DOI: un DOI superior con una sensibilidad acumulada de 1.2 y un DOI inferior establecido con una sensibilidad acumulada de 0.6.

La DOI calculada fue, en promedio, de 270 m para los sondeos del estudio SkyTEM 304M y de 300 m para los del estudio SkyTEM 312. Esta pequeña diferencia en la DOI entre ambos estudios se atribuyó a la relación señal-ruido relativamente alta en ambos estudios, debido a que el área de estudio es un entorno de resistividad relativamente baja.

Para el presente trabajo, no se consideraron las inversiones SCI sino inversiones 1D individuales por estación. El motivo es que inicialmente se quiere comparar el desempeño del método tradicional de inversión usando el método de Gauss-Newton con respecto al de algoritmos genéticos. Para trabajos futuros también se podría comparar el desempeño de ambos métodos para la inversión SCI pero esto requeriría mayor tiempo de cómputo.

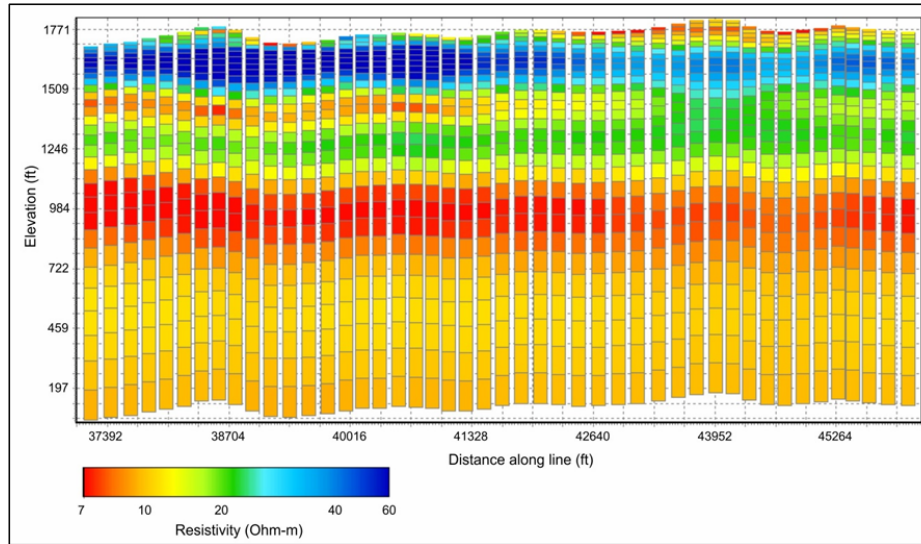


Figura 20: Un ejemplo de un perfil AEM que ilustra el aumento del espesor de las capas del modelo con la profundidad. Este es un modelo de 30 capas. Tomada de [1, p. 41].

### 3.5. Sondeos AEM seleccionados

Basándonos en el artículo de Kang et al., 2021 para este trabajo se utilizaron las 4 inversiones de los pozos que se presentaron en el mismo:

- Pozo 20N02E24C: Figura 15 de Kang et al., (2021).
- Pozo 22N01W29N: Figura 16 de Kang et al., (2021).
- Pozo 21N02E18C: Figura 17 de Kang et al., (2021).
- Pozo Pozo 21N01W11A001M: Figura A1 de Kang et al., (2021).

Los datos de los sondeos electromagnéticos de (AEM) y los pozos, se tomaron de Asch et al., (2019).

Para generar una comparativa entre las inversiones realizadas en el artículo y las generadas en este trabajo, Trabajamos con las líneas de datos más cercanas a estos pozos, basándonos en sus coordenadas UTM (WGS 84), esto en su lado norte y lado este para cada pozo reportado, las líneas más cercanas para cada pozo fueron las siguientes:

- Pozo 20N02E24C: Línea: 730201, Línea: 730101
- Pozo 22N01W29N: Línea: 710201, Línea: 200601

-Pozo 21N02E18C: Línea: 730401, Línea: 300901

-Pozo 21N01W11A001M: Línea: 710301, Línea: 200801

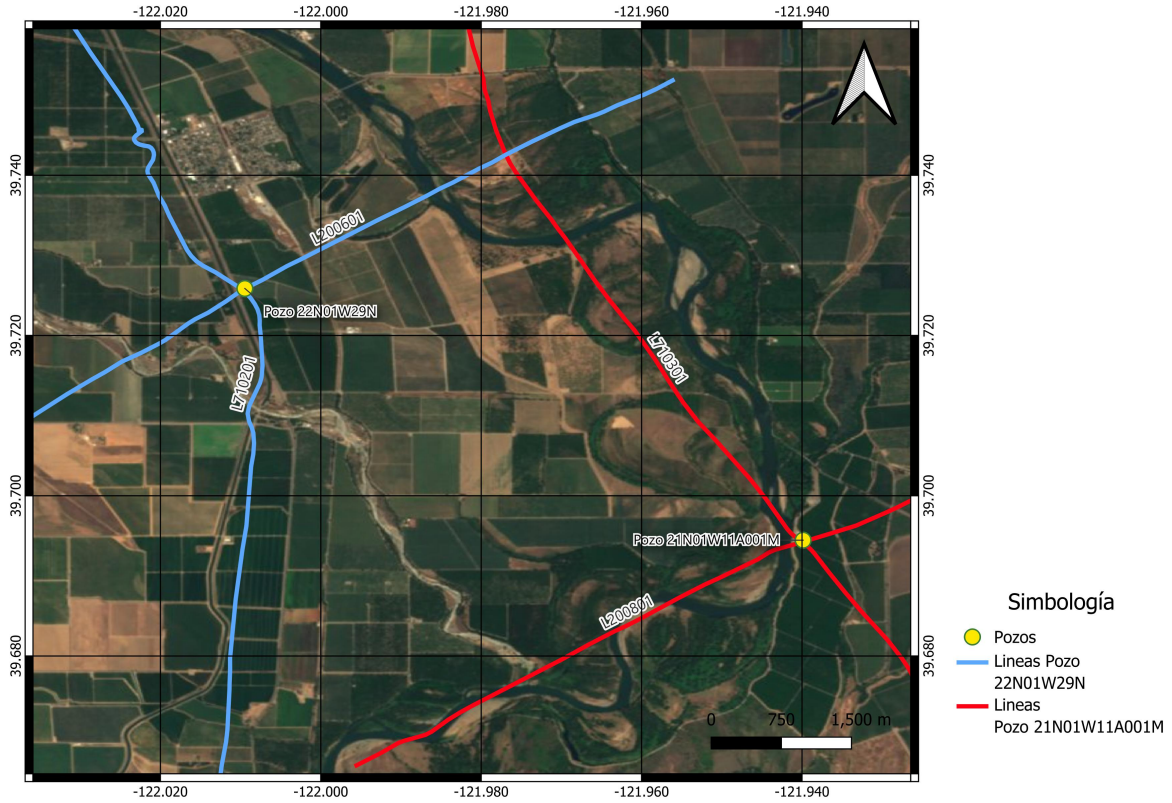


Figura 21: Líneas correspondientes a los Pozos 21N01W11A001M y 22N01W29N

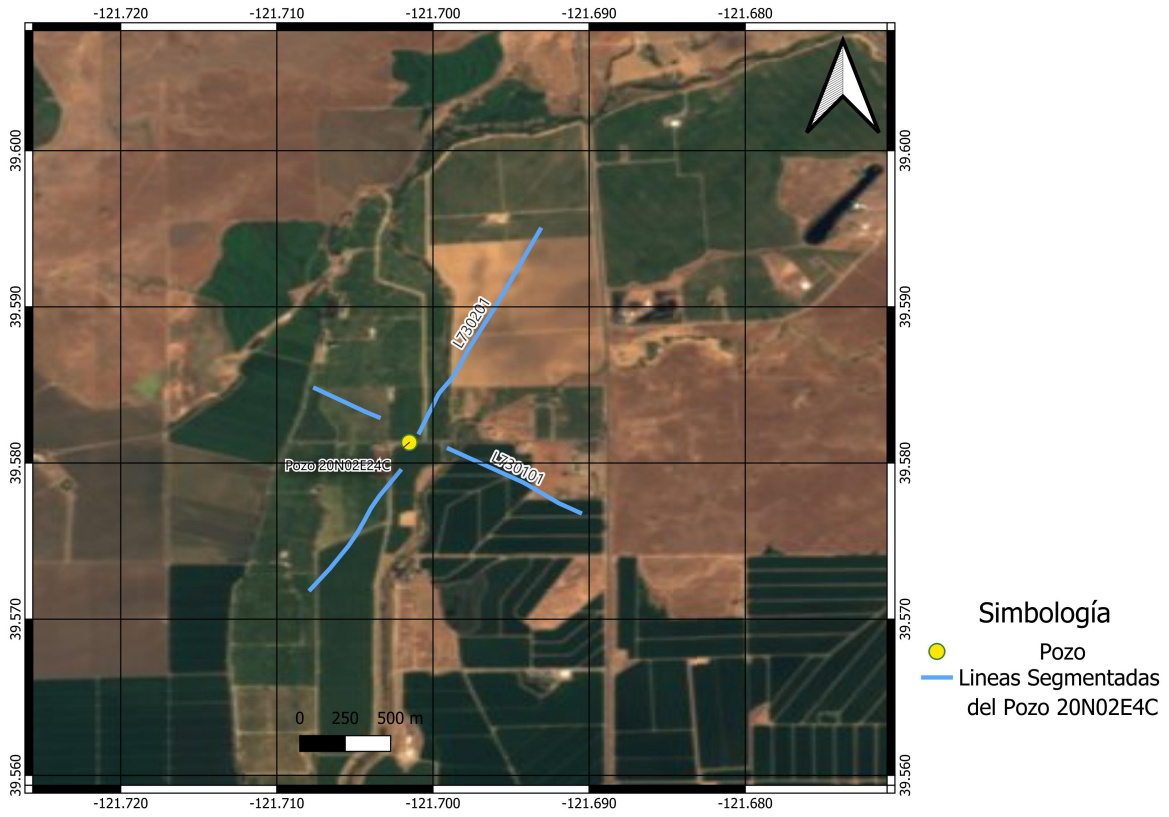


Figura 22: Líneas correspondientes al Pozo 20N02E4C

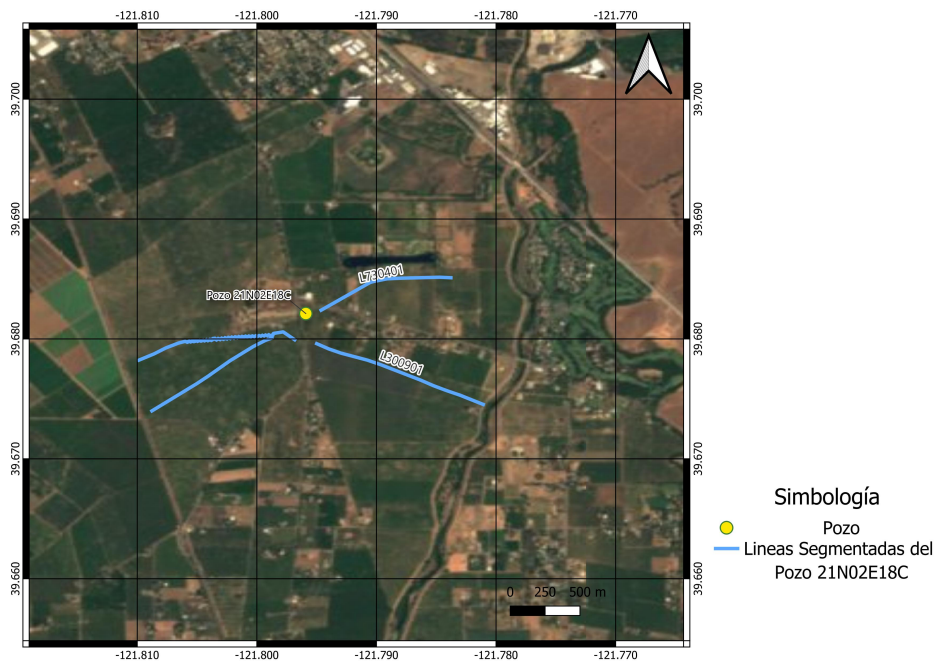


Figura 23: Líneas correspondientes al Pozo 21N02E18C

## 4. Metodología

Se implementa la inversión de datos electromagnéticos aéreos (AEM) utilizando un algoritmo genético para determinar la resistividad de un modelo de subsuelo 1D de múltiples capas. La metodología se basa en encontrar un modelo de resistividad del subsuelo que minimice la discrepancia entre los datos medidos y los datos predichos por la simulación así como el término de regularización.

### 4.1. Inversión con datos sintéticos

Esta sección describe la adaptación de la arquitectura de SimPEG para realizar la inversión 1D de datos reales de AEM. Asimismo, se detallan los parámetros preliminares establecidos para el Algoritmo Genético y se presenta una primera validación del método empleando datos sintéticos.

Este primer programa fue para asegurar que era viable trabajar con SimPEG y el algoritmo genético con datos electromagnéticos, confirmando que se ajustaron bien para este modelo sintético, esperando buenos resultados con datos reales, se comienza generando datos sintéticos a través de una simulación 1D de 5 capas horizontales añadiendo ruido sintético que añade un 2% de error aleatorio.

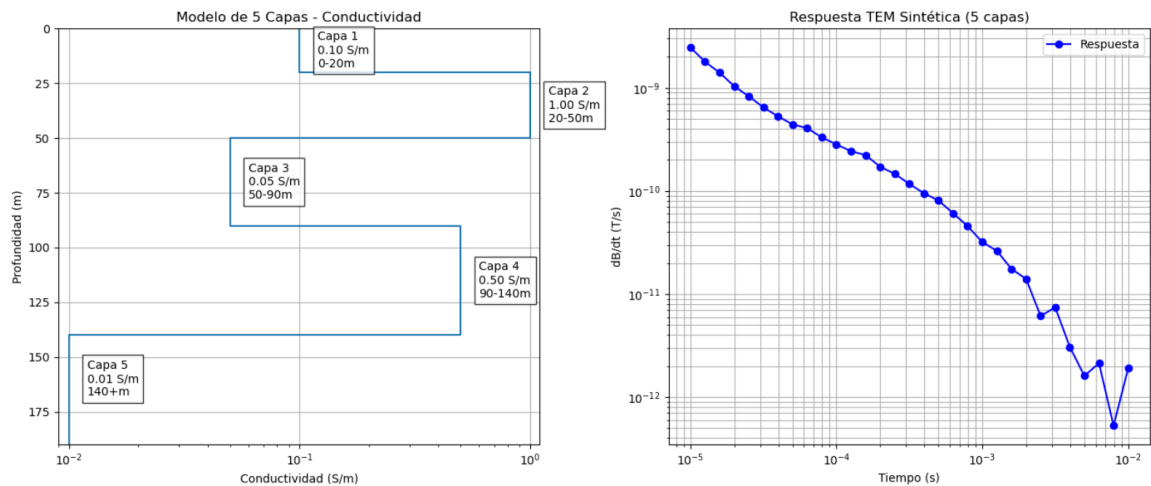


Figura 24: Modelo de 5 capas generado con datos sintéticos a la izquierda en azul con las siguientes conductividades; capa 1: 0-20m,  $\sigma=0.10\text{S/m}$ , capa 2: 20-50m,  $\sigma=1.00\text{S/m}$ , capa 3: 50-90m,  $\sigma=0.05\text{S/m}$ , capa 4: 90-140m,  $\sigma=0.50\text{S/m}$ , capa 5: 140+m,  $\sigma=0.01\text{S/m}$ . En la derecha los datos sintéticos generados representados con puntos en negro

Estos datos generados se guardaron en un archivo de texto con el que se trabajaron en la inversión con AG para hacer los ajustes pertinentes, dado que es mucho

más rápido que hacerlo con los datos reales por el tamaño de estos y por el número de capas con las que se trabajó.

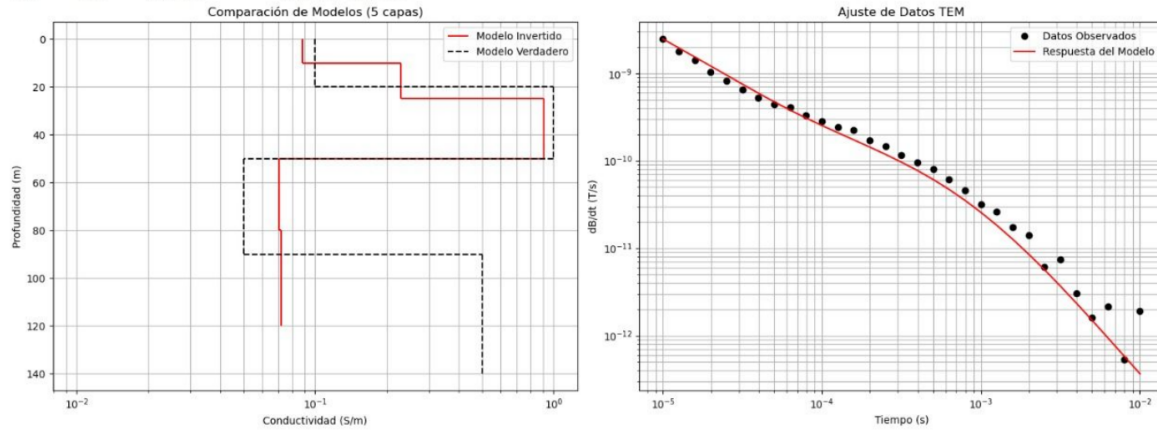


Figura 25: Modelo de datos sintéticos de 5 capas invertido a la izquierda y ajuste de los datos sintéticos generados a la derecha.

En la figura 25 podemos ver el modelo real de 5 capas generado con datos sintéticos en la parte izquierda con una línea punteada negra y como comparación el modelo invertido obtenido con el AG con una línea roja. En la sección de la parte derecha podemos ver los datos observados generados sintéticamente con puntos en negro y la respuesta del modelo, es decir, el ajuste de los datos con una línea en rojo.

Al tener un modelo real de datos generados sintéticamente, se pudieron hacer los primeros ajustes para saber que el código estuviera funcionando de manera correcta así mismo que las inversiones y el algoritmo genético estuvieran haciendo lo que se esperaba en las comparaciones del modelo invertido contra el verdadero. También funcionó como base para generar un buen AG que funcionara junto con SimPEG para el tema del problema directo, gracias a esto también se implementó DEAP para el ahorro de los tiempos de programación y sabiendo que funcionaba de una manera satisfactoria para los resultados esperados.

Para definir los parámetros del algoritmo, se realizaron diversas pruebas buscando un balance entre el tiempo de ejecución y la precisión de los resultados. Respecto al tamaño de la población, se evaluaron conjuntos de 50, 100, 150 y 200 individuos. Dado que al aumentar la cantidad de individuos el tiempo de cálculo se incrementó considerablemente, se seleccionó una población de 100 como la opción más eficiente.

De manera similar, para el número de generaciones, se observó que con 200 el algoritmo se estabiliza prematuramente, mientras que con 50 no siempre logra converger a una solución adecuada. Por lo tanto, se estableció un límite de 60 generaciones, valor con el cual se obtuvo convergencia en todos los casos de prueba.

Adicionalmente, se seleccionó la mutación polinomial acotada como operador de variación (detallada en la siguiente sección). Finalmente, se intentó implementar una estrategia de procesamiento simultáneo (paralelo) asignando una tarea por cada capa del modelo; sin embargo, este enfoque fue descartado debido a que se volvía ineficiente al aumentar el número de capas, por lo que se optó por el método definitivo que se describe más adelante.

## 4.2. Inversión con datos AEM reales

En esta sección se describe el proceso de inversión de los datos reales de AEM. Al Extraer la información más relevante de la estación de medición, como la geometría del transmisor (Tx) y receptor (Rx), las formas de onda, los tiempos de ventana y las lecturas de los canales de bajo momento (LM) y alto momento (HM), se pudo precisar la ubicación de las líneas trabajadas así mismo la respuesta que esta tiene, esto con la finalidad de poder plasmar la información recopilada en el código de AG implementado.

En el proyecto de Kang et al., 2021 se usó la librería SimPEG para configurar una simulación directa del modelo a los datos. Dicha simulación calcula la respuesta electromagnética de un modelo de subsuelo estratificado. SimPEG cumple la función de resolver el problema directo ( $f(m)$ ). Mientras que el Algoritmo Genético propone posibles modelos de subsuelo, SimPEG es la herramienta que predice qué datos mediría el instrumento SkyTEM si el subsuelo tuviera esas características específicas. Específicamente, se utiliza el módulo de simulación electromagnética para medios 1D, el cual asume un modelo de tierra estratificada horizontalmente, recibe un modelo de resistividades y espesores generado por la librería DEAP, resuelve las ecuaciones de Maxwell utilizando aproximación semianalítica para calcular la difusión de los campos electromagnéticos en las capas definidas y genera los datos predichos, es decir, las curvas de decaimiento de voltaje que se esperarían para ese modelo específico.

La inversión con el algoritmo genético es el principal objetivo; se implementó dicho algoritmo con ayuda de la librería DEAP para la búsqueda del mejor modelo de resistividades. DEAP no resuelve el problema directo; su única función es la búsqueda paramétrica para minimizar la función de costo.

El AG es ideal para este problema, ya que no requiere un punto de partida inicial es decir condiciones iniciales por la aleatoriedad del mismo y es robusto ante múltiples mínimos en el espacio de parámetros. El algoritmo genético evoluciona una población de modelos (individuos) a lo largo de generaciones para encontrar el que mejor se ajuste a los datos observados.

### 4.3. Función Objetivo

La función objetivo, también conocida como función de costo, tiene como propósito medir la calidad de un modelo de resistividad (cada individuo dentro de la población). Un valor más bajo de la función objetivo indica un mejor modelo. La función objetivo utilizada es una combinación del desajuste de datos y un término de regularización, lo que resulta en una función de costo regularizada:

$$C(m) = \phi_d(m) + \beta\phi_s(m) \quad (26)$$

Donde:

- $C(m)$  es la función de costo para un modelo  $m$
- $\Phi_d(m) = \sum_{i=1}^N \left( \frac{d_{obs,i} - d_{pred,i}}{\sigma_i} \right)^2$  es el misfit de los datos. Mide la diferencia entre los datos predichos ( $d_{pred,i}$ ) y los datos observados ( $d_{obs,i}$ ), normalizada por las incertidumbres ( $\sigma_i$ ). El objetivo es que este valor sea lo más cercano a cero posible.
- $\beta$  es un parámetro de regularización que controla el peso relativo del término de suavidad.

Esta función toma el modelo de resistividad, que es el individuo de DEAP, y se lo pasa a SimPEG para calcular la respuesta electromagnética, compara esa respuesta con los datos observados  $\phi_d(m)$ , añade la penalización por regularización  $\beta\phi_s(m)$ , y devuelve el valor de costo que DEAP busca minimizar.

### 4.4. Parámetros del Algoritmo Genético (AG)

Cada individuo (modelo) es una lista de valores de resistividad en escala logarítmica donde se define a cada individuo cómo una lista de números flotantes con el número de capas para un espacio base.

El tamaño inicial de la población es de 100 individuos, además de que se reguló a un mínimo de 39 capas, para que la inversión sea lo más parecida a los datos reales.

Se utilizó una selección por torneo, donde se eligen los mejores individuos de un subconjunto de tamaño 3. Esto garantiza que los individuos más aptos tengan más probabilidades de ser seleccionados para la cruce.

Se hizo uso de una cruce con un punto de cruce aleatorio. Esto combina las características de los dos padres para crear nuevos descendientes, fomentando la exploración del espacio de soluciones.

Para la mutación se usó una mutación polinomial acotada. Esta operación introduce pequeñas variaciones aleatorias en los individuos, asegurando que el algoritmo no quede atrapado en un mínimo local. Se aplican límites para mantener los valores de resistividad dentro de un rango realista. Se controla la intensidad de la mutación haciéndola un 10 % probable, es decir, que cada valor de resistividad dentro de un individuo tiene el 10 % de probabilidad de mutar.

El algoritmo corre durante 60 generaciones que determinamos que son las suficientes para que el modelo converja.

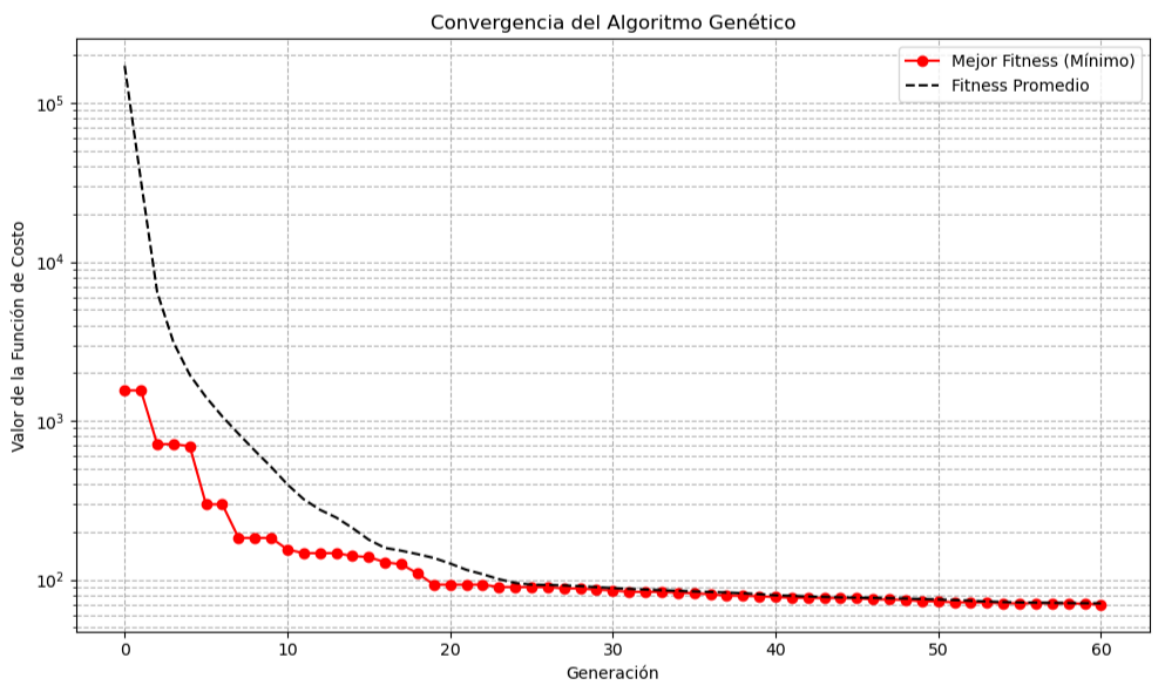


Figura 26: Gráfica que visibiliza la manera en la que converge el algoritmo genético utilizado

El procedimiento para la inversión se aceleró mediante la implementación de paralelismo a nivel de tarea en la fase de evaluación del Algoritmo Genético (AG), la cual es la parte computacionalmente más costosa del proceso.

La función de costo definida previamente, donde  $m$  es el modelo (resistividad). En el AG se calcula este costo para cada modelo propuesto (individuo). Dado que la evaluación de un modelo es independiente de la evaluación de cualquier otro modelo en la misma generación; se aprovecha la arquitectura de multiprocesamiento del hardware para ejecutar la evaluación de la población de los modelos de forma concurrente, en lugar de secuencial.

La aplicación de este paralelismo no afecta el resultado final de la inversión ( es decir, el modelo óptimo que minimiza la función de costo), su único efecto es la reducción del tiempo de cálculo. Esto permite utilizar poblaciones más grandes o un mayor número de generaciones en el AG, lo que incrementa las posibilidades de encontrar un óptimo global en un tiempo de cómputo razonable.

## 5. Resultados y discusión

### 5.1. Selección del factor de regularización ( $\beta$ ) para la inversión con AG

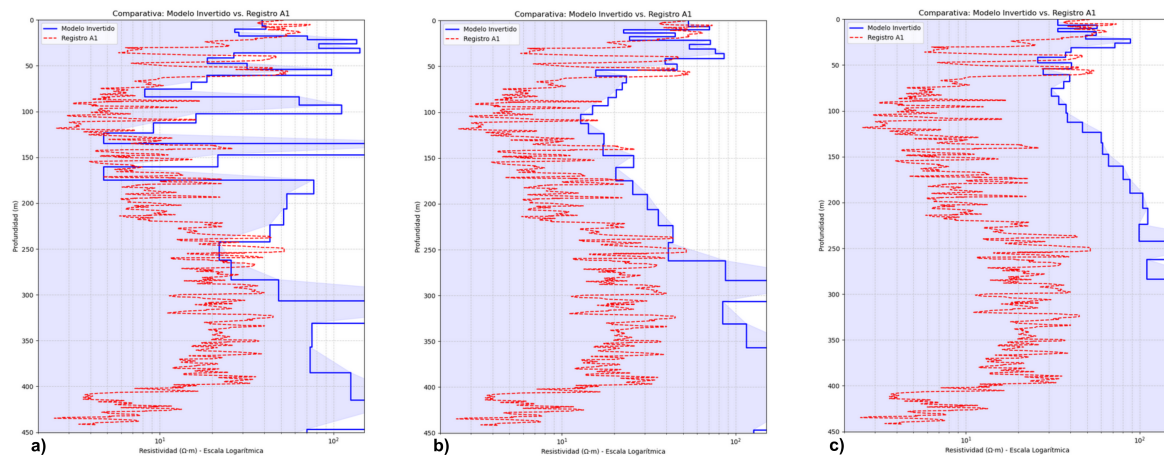


Figura 27: Comparación de la estación 200801 (línea en azul) con el Pozo 21N01W11A001M (línea punteada en rojo) para las inversiones con  $\beta$ : 50, 5000 y 50,000 respectivamente de izquierda a derecha.

En esta comparación utilizamos la estación (línea) 200801 correspondiente a el Pozo 21N01W11A001M (Figura A1 en el artículo de Kang et al., 2021). El análisis de sensibilidad realizado fundamenta la elección del parámetro  $\beta=5000$ , como parámetro para controlar el peso del término de suavidad después de como factor de ponderación para la restricción de suavidad (término de regularización). La determinación empírica de este coeficiente se basó en la búsqueda de un equilibrio entre el desajuste de los datos y la rugosidad del modelo recuperado. Dado que como se puede ver en el Anexo 1, al hacer falta el valor de  $\beta$  en las inversiones existen oscilaciones mayores, así como un desajuste mayor a los datos reales de AEM. Asimismo, se destaca que la sensibilidad del Algoritmo Genético frente al término de regularización difiere de la observada en métodos deterministas locales, como el de Gauss-Newton, debido a la naturaleza de la búsqueda global. Para seleccionar el valor de  $\beta$  se probaron 3 valores distintos como se muestra en la figura 27, iniciando con un valor de  $\beta=50$ , , en el cual se observan cambios de capa muy grandes generando una buena tendencia pero perdiendo

la suavidad entre capas, con un valor de  $\beta=5000$ , se observa cómo sigue la tendencia de los datos reales, así como una suavidad que corresponde con la misma, al no tener cambios tan abruptos entre capas y, finalmente, con el valor de  $\beta=50000$ , se puede observar que se aleja más de la tendencia original, generando mucha suavidad fuera de tendencia, así mismo perdemos resolución alrededor de los 250 metros de profundidad.

## 5.2. Modelos invertidos y comparación con el método de Gauss-Newton

En esta sección se presentan los modelos obtenidos con la inversión usando algoritmos genéticos y su respectiva comparación con los modelos en los que se usó Gauss-Newton junto con las especificaciones de la implementación de este último. Dentro de estas comparaciones podremos ver las distintas tendencias, diferencias y similitudes que tiene cada método, esto con la finalidad de poder validar los resultados obtenidos mediante los modelos invertidos de AG aplicados a las líneas y pozos reportados. Esta presentación de resultados es la parte más importante de la investigación ya que nos permite saber que tan útil es el AG para la inversión de datos electromagnéticos.

Para el método de Gauss-Newton se consideraron también modelos de capas. Los parámetros usados para la inversión de Gauss-Newton fueron los siguientes:

- Modelo inicial [rho background]=20  $\Omega m$
- Número de capas=39
- Número máximo de iteraciones [maxIter]=20
- $\beta = 100$
- Factor de enfriamiento [coolingFactor]=5
- Tasa de enfriamiento [coolingRate]=2
- $\alpha_x = 1.0$
- $\alpha_s = 50$
- Error relativo = 3

A continuación presentaremos las comparaciones más relevantes y daremos una descripción de ellas:

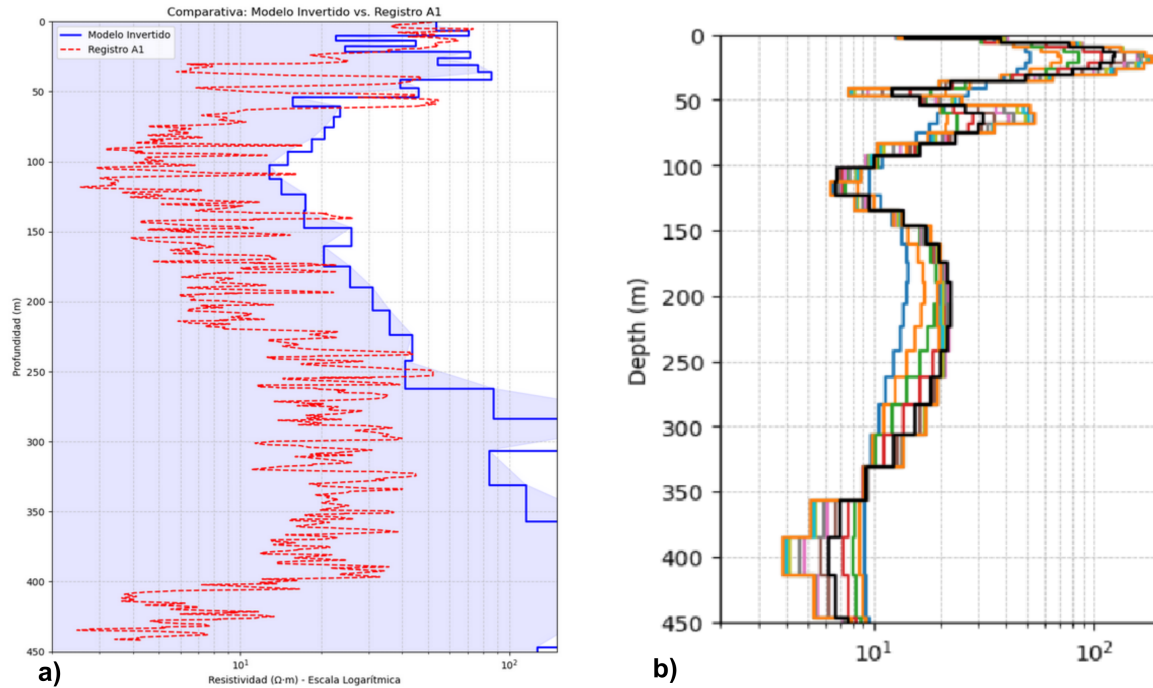


Figura 28: De izquierda a derecha. Modelo de AG invertido para la estación 200801. Modelo de Gauss Newton invertido las líneas de colores son los modelos invertidos con distintos valores de beta, la línea negra es el modelo elegido como más adecuado debido a los valores de  $\phi$  y  $\beta$ .

El análisis comparativo de la inversión para la estación 200801, validado con el registro del Pozo 21N01W11A001M (Figura A1 en Kang et al., 2021), revela diferencias significativas en la capacidad de resolución de ambos algoritmos.

En el intervalo somero (0 a 20 m), el modelo obtenido mediante Algoritmos Genéticos (AG) recupera fielmente los valores de resistividad, mientras que la solución de Gauss-Newton tiende a sobreestimar dicha propiedad. Sin embargo, en la zona inmediatamente posterior (25 a 50 m), el método de Gauss-Newton logra resolver una anomalía que el modelo de AG no consigue detectar, mostrando una limitación en la sensibilidad en esta área.

Posteriormente, en el rango de profundidad media (50 a 250 m), el modelo de (AG) reproduce con mayor consistencia la tendencia observada en la (Figura A1). En contraste, aunque el método Gauss-Newton identifica el descenso en la resistividad, subestima la profundidad.

Finalmente, a profundidades mayores a 250 m, el algoritmo genético presenta una tendencia a la sobreestimación de los valores resistivos, mientras que el modelo de

Gauss-Newton exhibe una mejor concordancia con los datos de pozo.

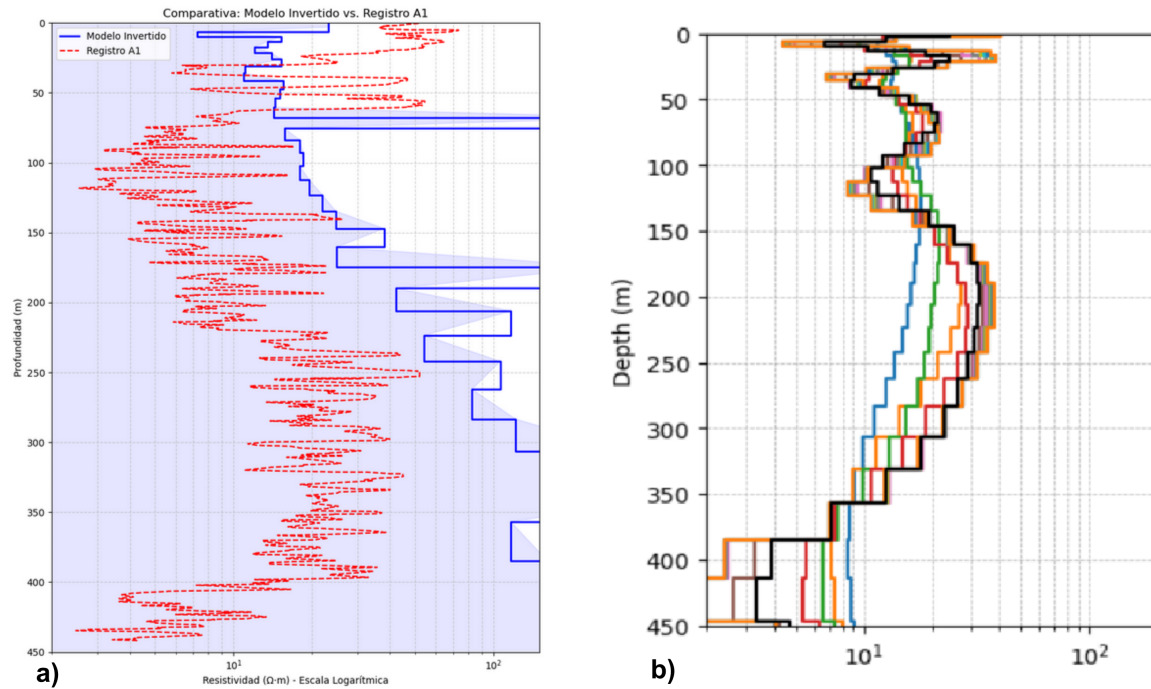


Figura 29: De izquierda a derecha. Modelo de AG invertido para la estación 710301. Modelo de Gauss Newton invertido las líneas de colores son los modelos invertidos con distintos valores de beta, la línea negra es el modelo elegido como más adecuado debido a los valores de  $\phi$  y  $\beta$ .

Este segundo punto invertido, comparado con referencia (Figura A1 en el artículo de Kang et al., 2021), estación-(línea) 710301, Se destaca la capa de los 80 a los 170 metros aproximadamente en el modelo de algoritmos genéticos que se observa una mejor resolución al espesor de esta capa que en el caso de Gauss-Newton que va de los 90 a los 120 metros aproximadamente, sin embargo, este método también tiene una mejor respuesta a esta otra capa de los 150 a los 400 metros cuando los Algoritmos genéticos tienden a sobrestimar las resistividades.

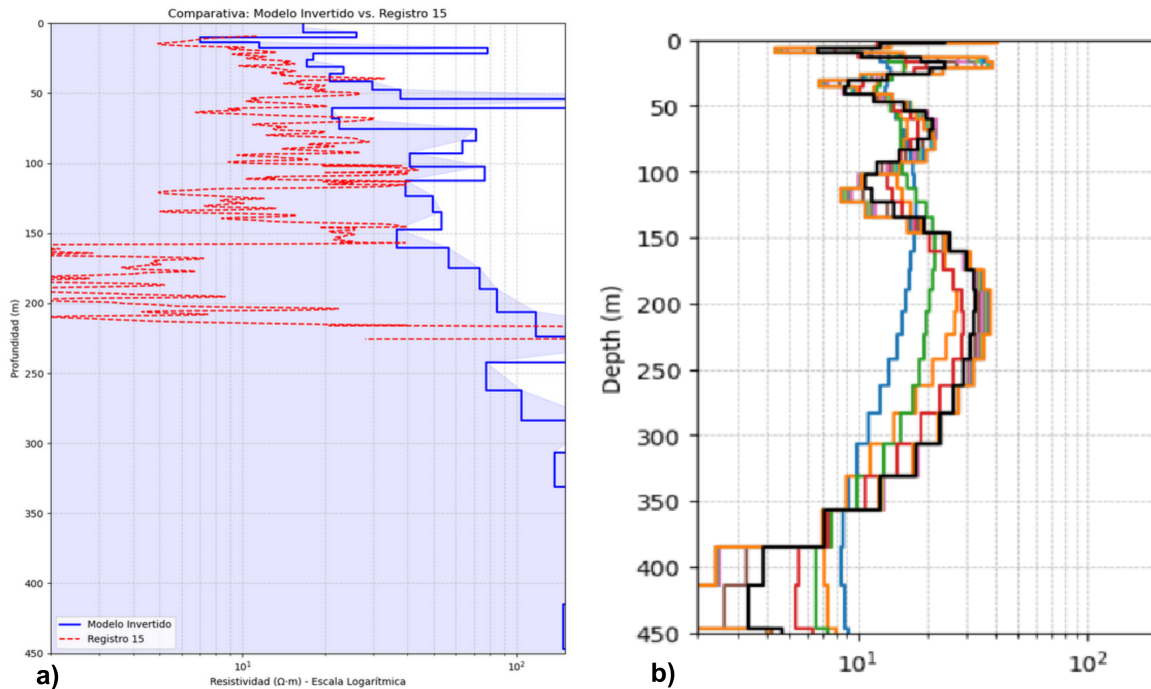


Figura 30: De izquierda a derecha. Modelo de AG invertido para la estación 730101. Modelo de Gauss Newton invertido las líneas de colores son los modelos invertidos con distintos valores de beta, la línea negra es el modelo elegido como más adecuado debido a los valores de  $\phi$  y  $\beta$ .

El siguiente punto a analizar corresponde a la estación-(línea) 730101 correspondiente a el Pozo 20N02E24C (Figura 15 en el artículo de Kang et al., 2021). Aquí se hizo un seguimiento de tendencias similar, los primeros 60 metros en ambos casos, con algoritmos genéticos teniendo un par de oscilaciones a mayor resistividad de la esperada, también una capa que se pierde de los 100 a los 150 metros que sí detecta la inversión de Gauss-Newton, después el registro tiene una capa de bajos resistivos de los 150 a los 210 metros aproximadamente que se pierde totalmente en ambos métodos y luego a los 220 un alto de resistividad que los algoritmos genéticos sí parecen detectar mientras que Gauss-Newton tiene una resistividad menor del orden de los 10-20  $\Omega\text{m}$ .

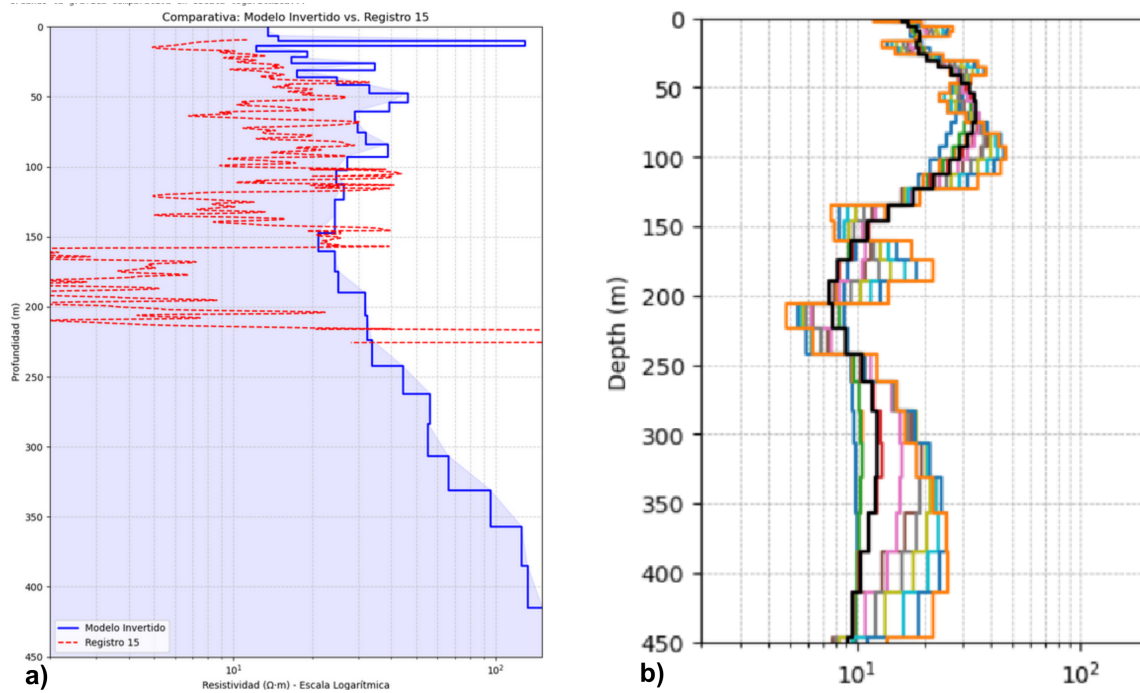


Figura 31: De izquierda a derecha. Modelo de AG invertido para la estación 730201. Modelo de Gauss Newton invertido las líneas de colores son los modelos invertidos con distintos valores de beta, la línea negra es el modelo elegido como más adecuado debido a los valores de  $\phi$  y  $\beta$ .

Para la estación-(línea) 730201 (Figura 15 en el artículo de Kang et al., 2021), Se tienen tendencias similares en ambos métodos con un alto resistivo detectado con algoritmos genéticos a los 15 metros, y estas tendencias comienzan a variar en los 100 metros donde en Gauss-Newton existe una capa de menor resistividad de un orden de  $7 \Omega\text{m}$  que en cambio en algoritmos genéticos no detecta y a los 200 metros comienza a aumentar la resistividad gradualmente de 20 a  $150 \Omega\text{m}$  hasta llegar a los 400 metros.

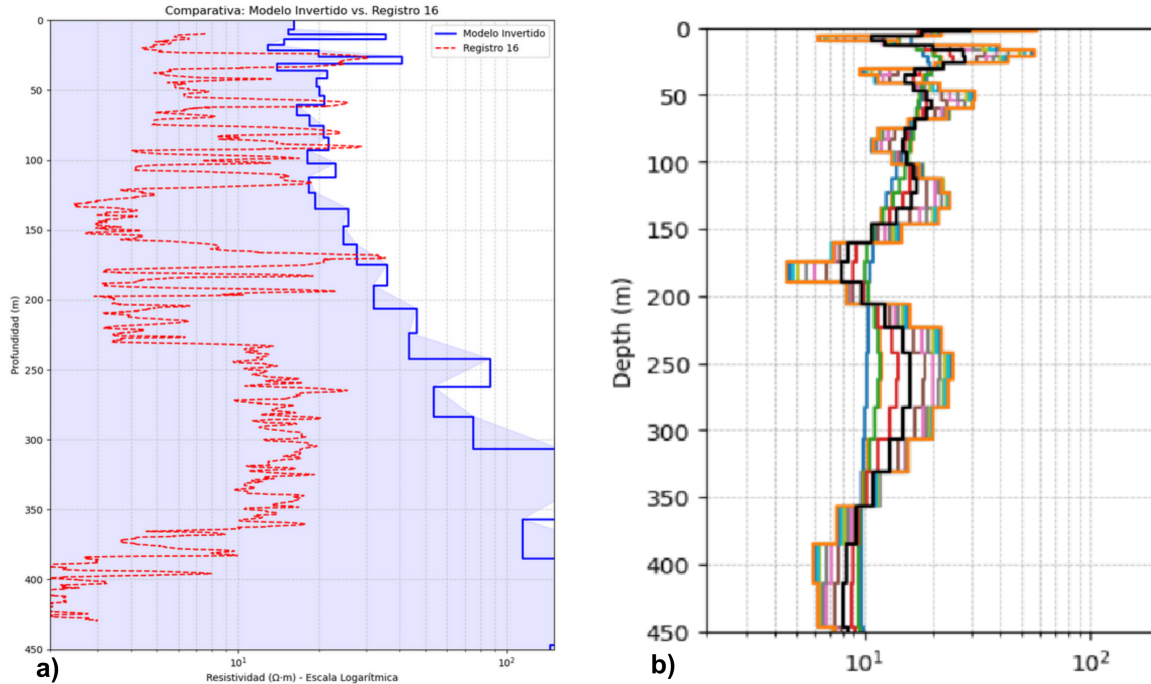


Figura 32: De izquierda a derecha. Modelo de AG invertido para la estación 200601. Modelo de Gauss Newton invertido las líneas de colores son los modelos invertidos con distintos valores de beta, la línea negra es el modelo elegido como más adecuado debido a los valores de  $\phi$  y  $\beta$ .

En la Figura 25 se muestran los resultados para la estación-(línea) 200601 Pozo 22N01W29N (Figura 16 en el artículo de Kang et al., 2021). Comparando los resultados de inversión usando algoritmos genéticos con la de Gauss-Newton se puede ver que que modelo a la izquierda sigue el rango de resistividades que el invertido con Gauss-Newton durante los primeros 100 metros. El registro a los 130-150 metros detecta un bajo resistivo que no llega a detectar ninguno de los dos modelos, después de eso el modelo de algoritmos genéticos tiende a sobrestimar las resistividades mientras que el modelo de inversión con Gauss-Newton si detecta de mejor manera las tendencias del registro en el intervalo de 250-350 metros.

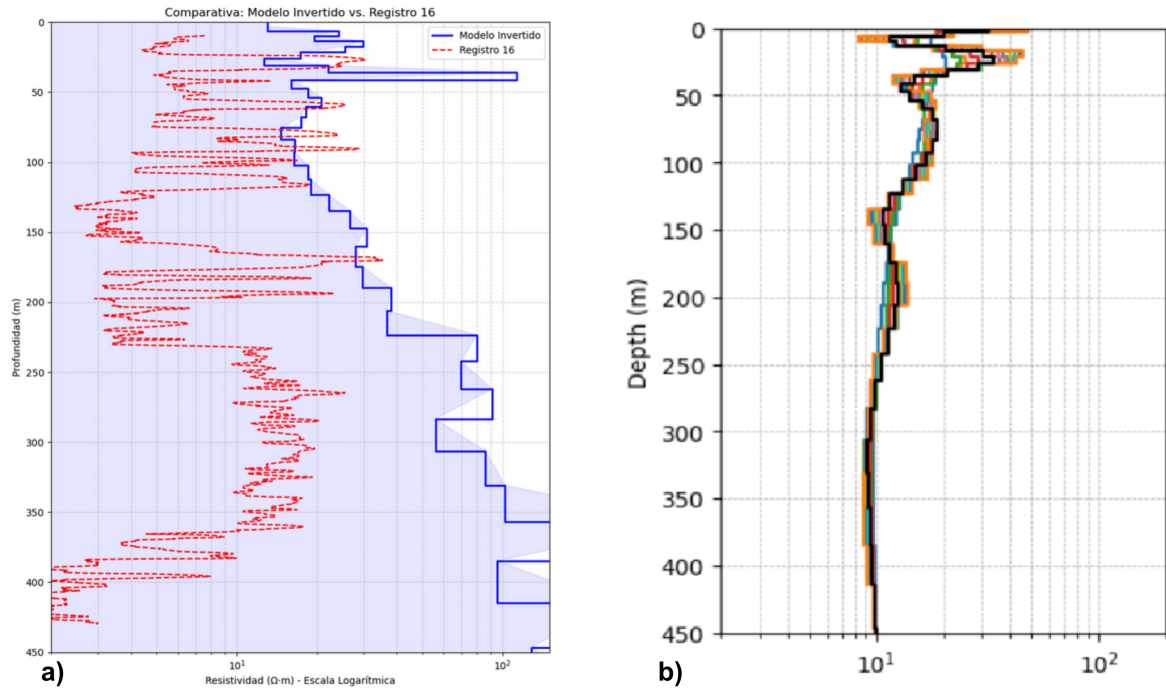


Figura 33: De izquierda a derecha. Modelo de AG invertido para la estación 710201. Modelo de Gauss Newton invertido las líneas de colores son los modelos invertidos con distintos valores de beta, la línea negra es el modelo elegido como más adecuado debido a los valores de  $\phi$  y  $\beta$

En esta comparación para la estación-(línea) 710201 (Figura 16 en el artículo de Kang et al., 2021), En donde los 50 metros en la inversión con el modelo de algoritmos genéticos detecta una elevación de resistividad con un orden de  $100 \Omega\text{m}$  mientras que el de Gauss-Newton se limita a un orden de  $20\text{-}40 \Omega\text{m}$ , en este, mantiene una resistividad de  $10\text{-}20 \Omega\text{m}$  de los 50 hasta los 450 metros mientras que con algoritmos genéticos hay una diferencia a partir de los 100 metros que el modelo tiende a predecir capas más resistivas por un orden de los 20 hasta los  $200 \Omega\text{m}$  que aumenta gradualmente mientras aumenta la profundidad. En este caso la aproximación del modelo invertido con Gauss-Newton tiene una mejor predicción en profundidades mayores a los 100 metros.

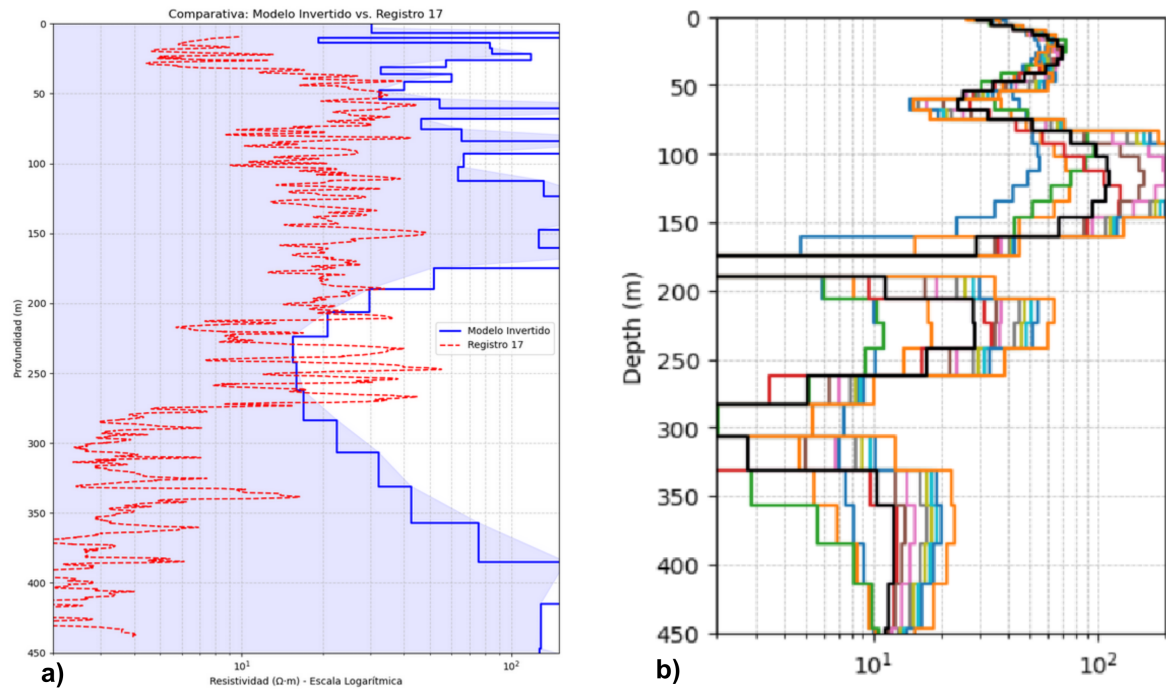


Figura 34: De izquierda a derecha. Modelo de AG invertido para la estación 300901. Modelo de Gauss Newton invertido las líneas de colores son los modelos invertidos con distintos valores de beta, la línea negra es el modelo elegido como más adecuado debido a los valores de  $\phi$  y  $\beta$

En esta inversión correspondiente a la estación-(línea) 300901 Pozo 21N02E18C (Figura 17 en el artículo de Kang et al., 2021), ambos modelos de inversión estiman de una manera similar los primeros 150 metros, pero en el modelo de inversión con algoritmos genéticos se observan oscilaciones más abruptas entre capas consecuentes. Después de los 150 a los 200 metros el modelo de Gauss-Newton detecta bajas resistividades menores a los  $0.2 \Omega\text{m}$  mientras que el modelo de AG si bien también detecta una baja de resistividades el orden es de  $10\text{-}20 \Omega\text{m}$ . En el rango de 200 a 250 metros ambos estiman las resistividades de manera similar, seguido del modelo de AG sobrestimando las resistividades después de los 275 metros aproximadamente y el modelo de Gauss-Newton haciendo un mejor trabajo de estimación hasta los 350 metros aproximadamente.

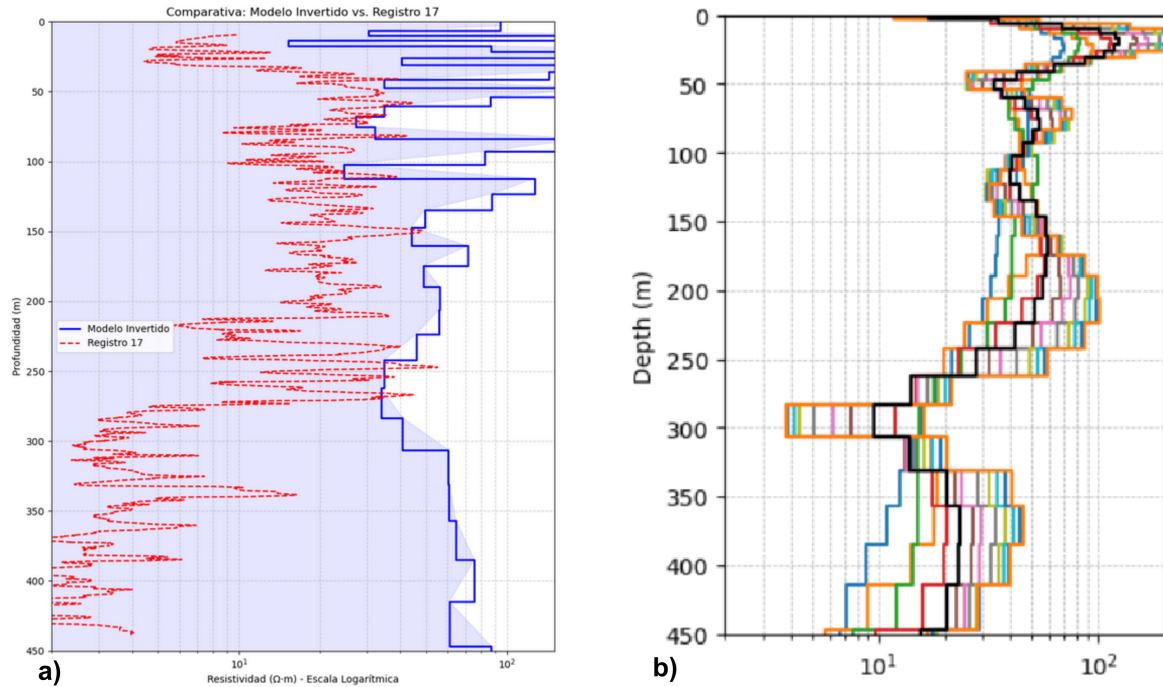


Figura 35: De izquierda a derecha. Modelo de AG invertido para la estación 730401. Modelo de Gauss Newton invertido las líneas de colores son los modelos invertidos con distintos valores de beta, la línea negra es el modelo elegido como más adecuado debido a los valores de  $\phi$  y  $\beta$

En esta inversión correspondiente a la estación-(línea) 730401 Pozo 21N02E18C (Figura 17 en el artículo de Kang et al., 2021) Sigue un seguimiento similar en ambos métodos de inversión donde existen tendencias similares pero se tienen mayores oscilaciones en el caso de algoritmos genéticos en las partes someras de los 10 a los 100 metros. En ambos se detecta una baja de resistividades en los 250 a los 325 metros y similares resistividades también en ambos a estas profundidades donde en el registro se detecta una baja de resistividades.

En todos los modelos obtenidos para las 8 estaciones se pudieron distinguir algunas tendencias generales, que ayudan a distinguir en distintas zonas el comportamiento de la inversión en cada una de las líneas, considerando que no se utilizó la técnica de inversión con restricción espacial (SCI), se optó por representar cada pozo con sus líneas respectivas, esto con la finalidad de observar los cambios respectivos, así como similitudes entre ambas, respecto al modelo de Gauss-Newton. El AG demostró capacidad para seguir las tendencias de los registros de pozo, especialmente en la parte somera e intermedia del subsuelo (hasta los 250 m). Sin embargo, el modelo de AG tuvo una tendencia a sobrestimar la resistividad y mostrar oscilaciones más abruptas entre capas a mayor profundidad ( $>250$  m), en comparación con el modelo de Gauss-Newton. Esto sugiere que, para profundidades donde la resolución del método TEM disminuye, la regularización aplicada al modelo AG puede no ser tan efectiva o su penalización

por  $\beta = 5000$  puede haber sido subóptima para estructuras más profundas.

La comparación de diferentes valores de  $\beta$  (50, 5000, 50000) fue crucial para la estabilidad del modelo. La selección del valor de 5000 fue determinante para evitar suavizar de más (visto en los valores más grandes) o ajustar de más (visto en valores bajos), aunque se reconoce el balance óptimo del peso de regularización sigue siendo un desafío en las partes más profundas de la inversión.

## 6. Conclusiones

La aplicación de Algoritmos Genéticos para la inversión de datos TEM demostró ser una alternativa viable a los métodos determinísticos tradicionales como el basado en Gauss-Newton, particularmente por su robustez ante la no linealidad del problema y la independencia de un modelo inicial.

Se insta a la posibilidad de generar programas de código abierto para la inversión de datos electromagnéticos con métodos de Algoritmos genéticos debido a la viabilidad de la utilización de este método en comparación a métodos determinísticos tradicionales.

El estudio documenta la factibilidad de utilizar AG con la arquitectura de SimPEG para la inversión TEM, proporcionando una herramienta poderosa para complementar los métodos determinísticos, especialmente la búsqueda de modelos geológicamente plausibles en un contexto de incertidumbre alta.

La implementación de paralelismo en la evaluación del Algoritmo Genético fue exitosa. Este procedimiento redujo significativamente el tiempo de cómputo, lo que permitió la ejecución de un mayor número de generaciones y un mayor tamaño de población, incrementando las posibilidades de encontrar un óptimo global en el espacio de soluciones.

Al implementarlo de manera temprana en el desarrollo del código nos dio mucho juego para ajustar parámetros ahorrándonos mucho tiempo. El tiempo de ejecución disminuye de forma casi lineal con el número de núcleos.

## 7. Anexo 1

En este anexo presentaremos algunos de los resultados previos a añadir el término  $\beta$  de regularización porque nos parece algo interesante para dejar reportado sin necesariamente meterlo dentro de la rama principal de esta tesis.

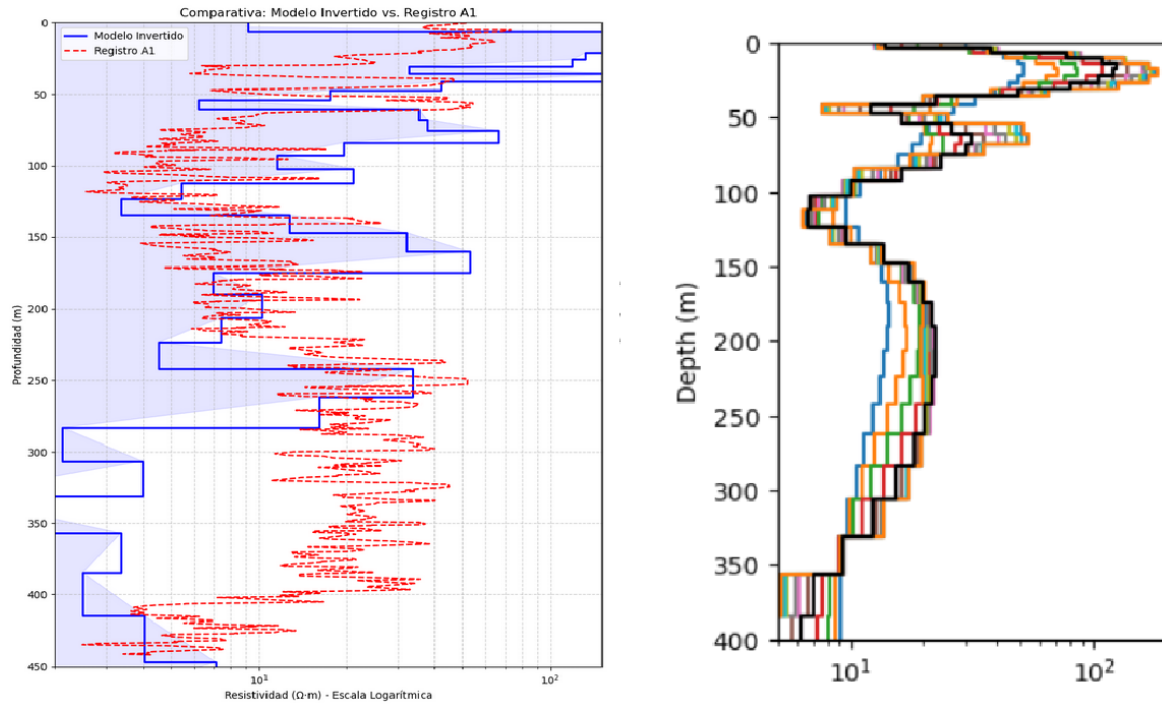


Figura 36: De izquierda a derecha. Modelo de AG invertido para la estación 200801 sin regularización. Modelo de Gauss Newton invertido las líneas de colores son los modelos invertidos con distintos valores de beta, la línea negra es el modelo elegido como más adecuado debido a los valores de  $\phi$  y  $\beta$

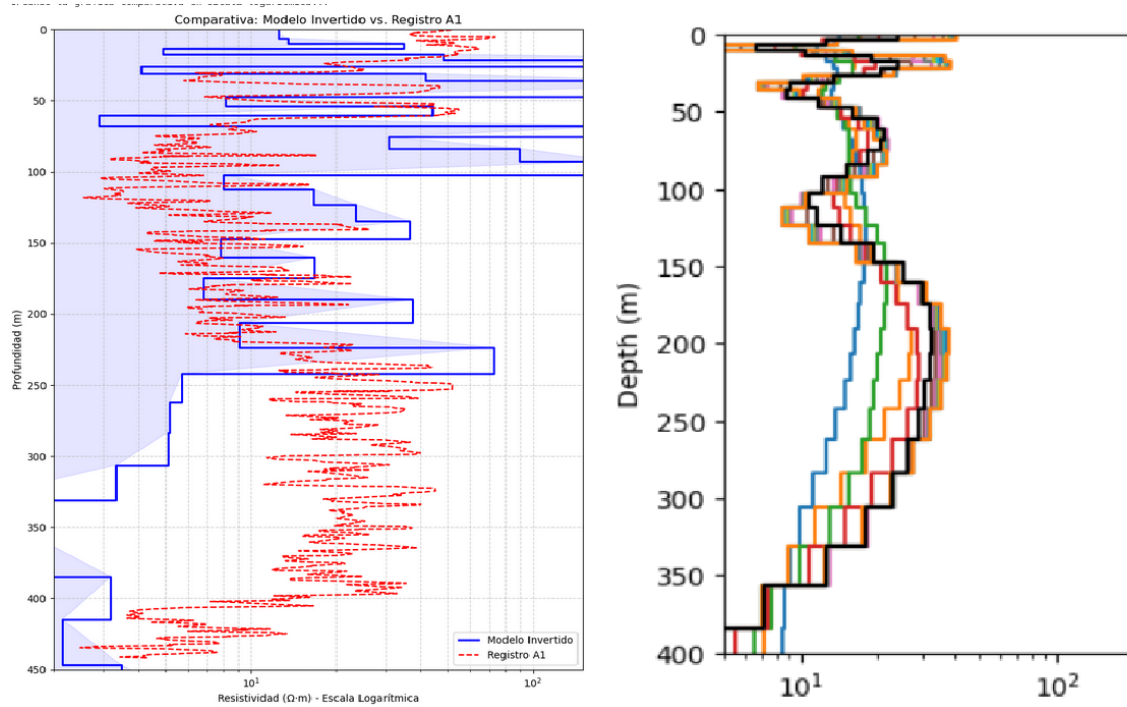


Figura 37: De izquierda a derecha. Modelo de AG invertido para la estación 710301 sin regularización. Modelo de Gauss Newton invertido las líneas de colores son los modelos invertidos con distintos valores de beta, la línea negra es el modelo elegido como más adecuado debido a los valores de  $\phi$  y  $\beta$

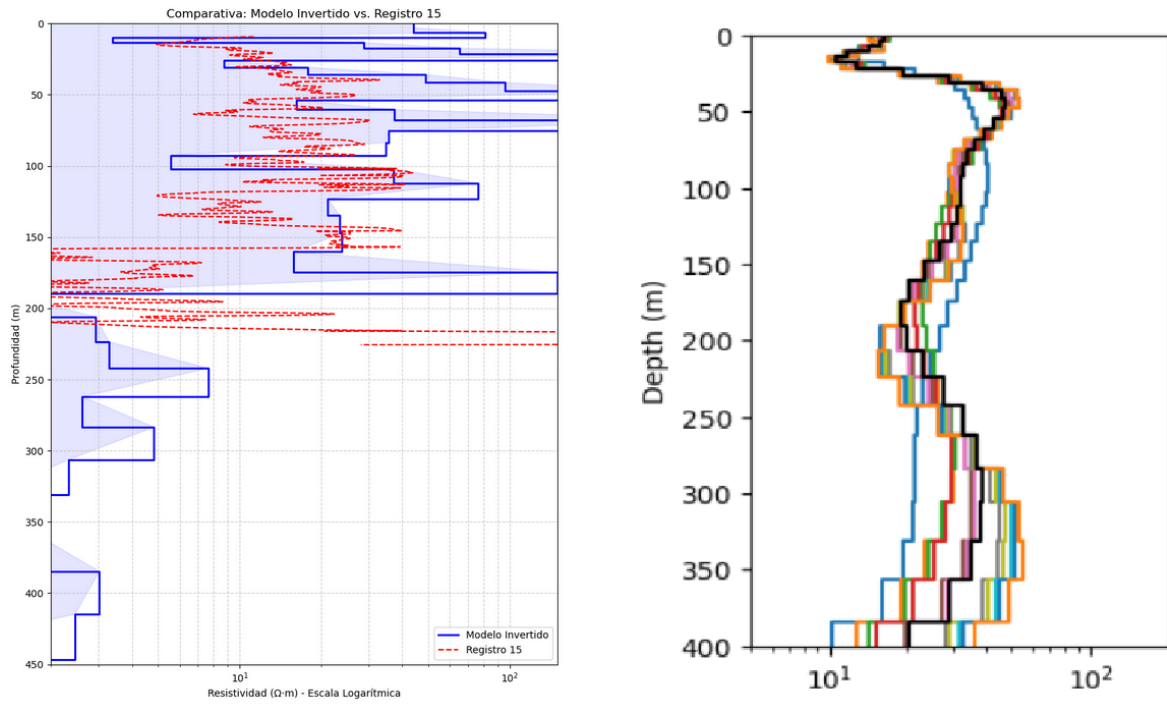


Figura 38: De izquierda a derecha. Modelo de AG invertido para la estación 730101 sin regularización. Modelo de Gauss Newton invertido las líneas de colores son los modelos invertidos con distintos valores de beta, la línea negra es el modelo elegido como más adecuado debido a los valores de  $\phi$  y  $\beta$

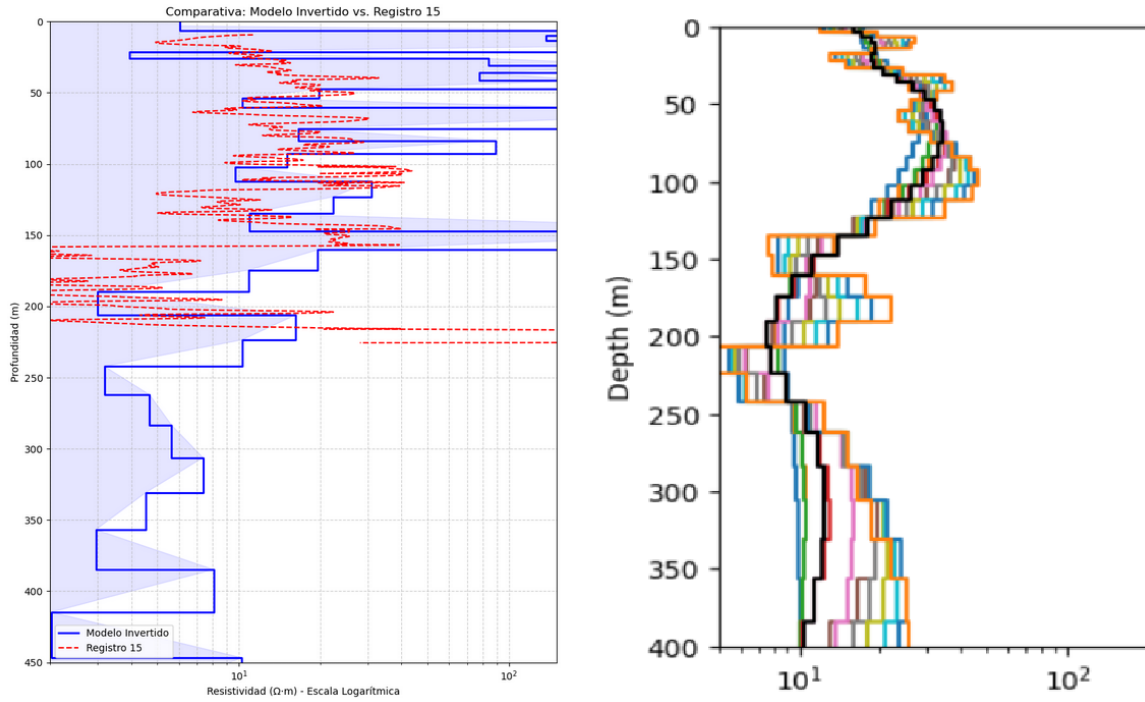


Figura 39: De izquierda a derecha. Modelo de AG invertido para la estación 730201 sin regularización. Modelo de Gauss Newton invertido las líneas de colores son los modelos invertidos con distintos valores de beta, la línea negra es el modelo elegido como más adecuado debido a los valores de  $\phi$  y  $\beta$

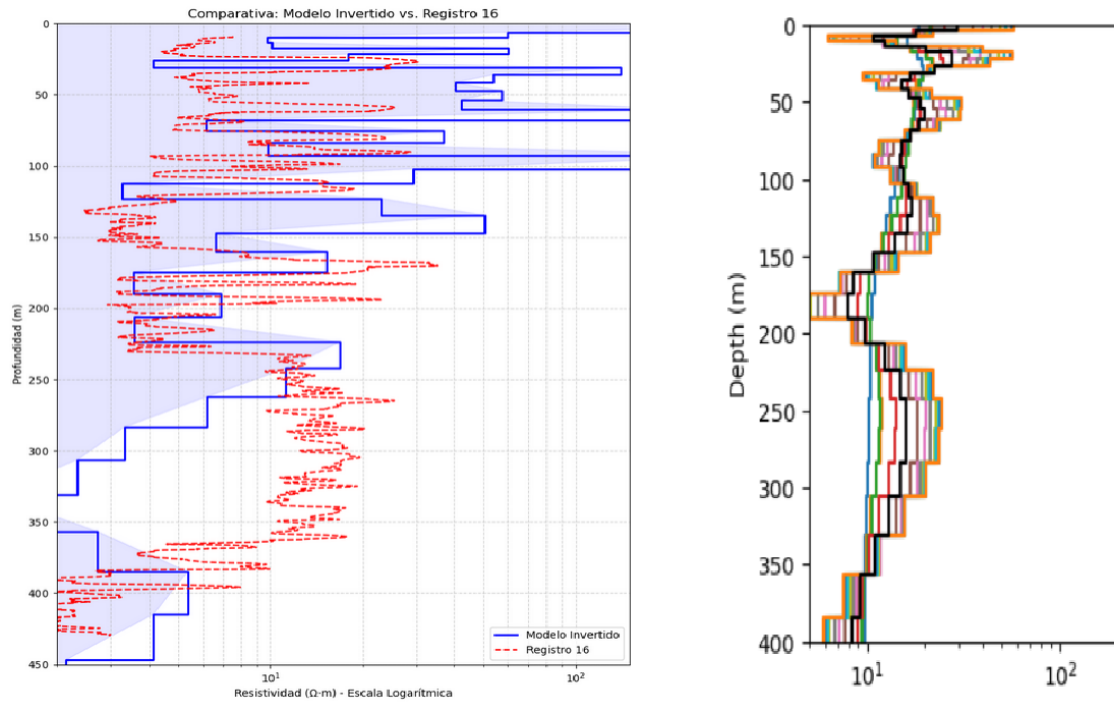


Figura 40: De izquierda a derecha. Modelo de AG invertido para la estación 200601 sin regularización. Modelo de Gauss Newton invertido las líneas de colores son los modelos invertidos con distintos valores de beta, la línea negra es el modelo elegido como más adecuado debido a los valores de  $\phi$  y  $\beta$

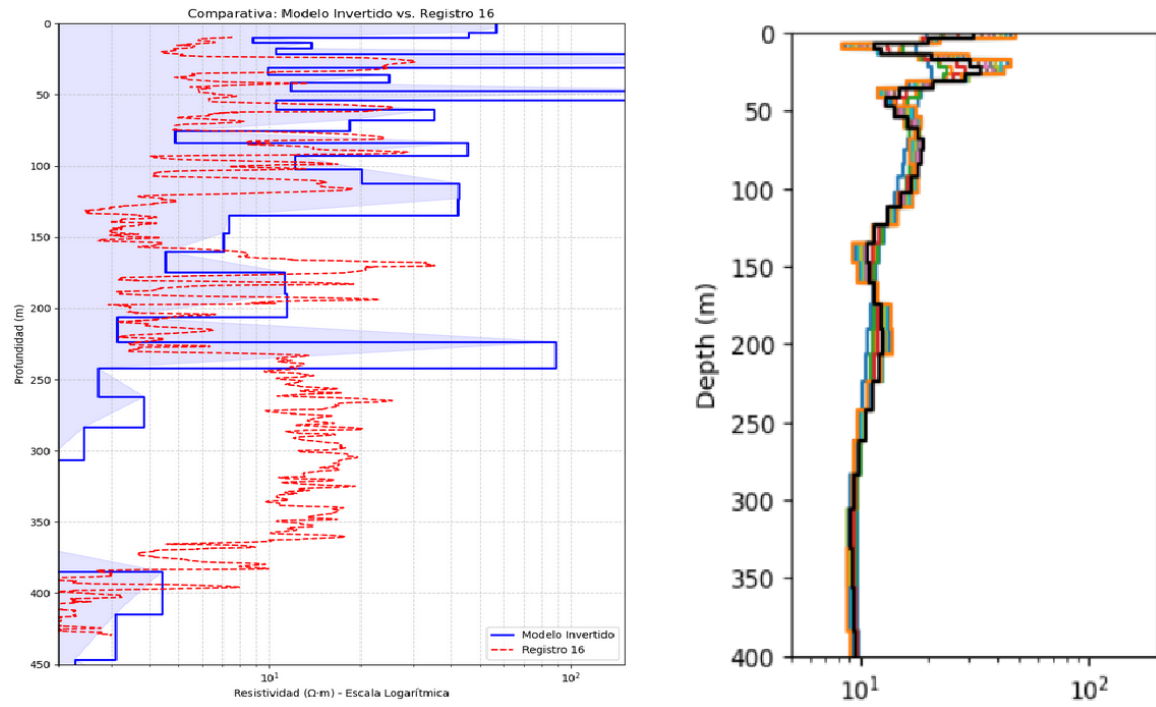


Figura 41: De izquierda a derecha. Modelo de AG invertido para la estación 710201 sin regularización. Modelo de Gauss Newton invertido las líneas de colores son los modelos invertidos con distintos valores de beta, la línea negra es el modelo elegido como más adecuado debido a los valores de  $\phi$  y  $\beta$

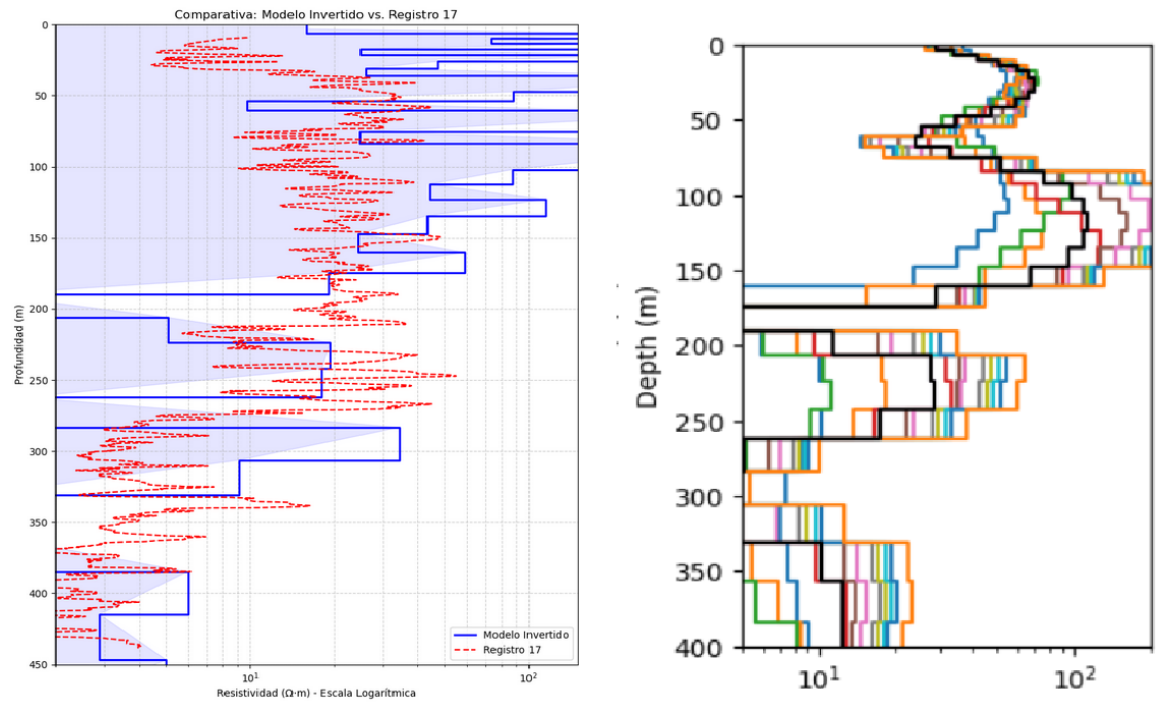


Figura 42: De izquierda a derecha. Modelo de AG invertido para la estación 300901 sin regularización. Modelo de Gauss Newton invertido las líneas de colores son los modelos invertidos con distintos valores de beta, la línea negra es el modelo elegido como más adecuado debido a los valores de  $\phi$  y  $\beta$

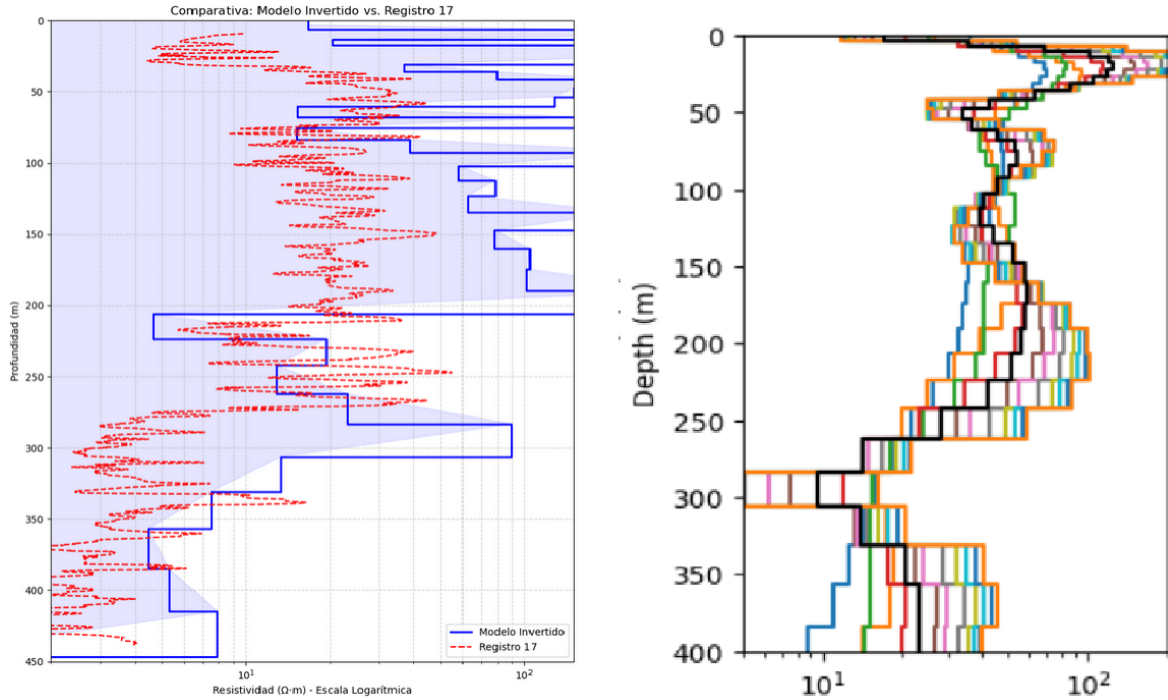


Figura 43: De izquierda a derecha. Modelo de AG invertido para la estación 730401 sin regularización. Modelo de Gauss Newton invertido las líneas de colores son los modelos invertidos con distintos valores de beta, la línea negra es el modelo elegido como más adecuado debido a los valores de  $\phi$  y  $\beta$

Tratando con los datos reales de TEM, en estos resultados sin el término de regularización  $\beta$ , se observan oscilaciones más grandes entre cada cambio de capa, esto porque el factor  $\beta$  controla la suavidad de las mismas, al tener solo el parámetro  $\phi$  la inversión tiende a salirse de la oscilación normal vista en las inversiones de SimPEG, que cuenta con el parámetro  $\beta$  en su función objetivo, es interesante observar cómo los datos sin este parámetro, siguen una tendencia hacia los datos reales esto por el ajuste del modelo de AG solo con  $\phi$  integrado a la función objetivo.

Como se puede ver en todas las imágenes anteriores la aplicación del término  $\beta$  fue fundamental para llegar a resultados satisfactorios.

## 8. Anexo 2. Paralelización de inversión con AG

En este anexo ahondaremos más en la implementación del paralelismo para la optimización del código, su arquitectura y sus ventajas.

## 8.1. Problema y arquitectura del paralelismo

En una inversión mediante Algoritmos Genéticos, el cuello de botella se encuentra en la etapa de evaluación de la función objetivo para cada individuo de la población en donde el código debe tomar las conductividades del modelo, ejecutar una simulación directa por medio de SimPEG, y resolver las ecuaciones de Maxwell para obtener la respuesta electromagnética y comparar el resultado con los datos observados.

Si una simulación toma, por ejemplo, 0.5 segundos y tenemos una población de 100 individuos durante 60 generaciones, una ejecución secuencial tardaría al menos 50 minutos por cada número de capas probado. El paralelismo reduce drásticamente este tiempo.

El código utiliza una arquitectura de paralelismo de datos llamada “Maestro-Trabajador” basada en el módulo multiprocessing de Python.

El proceso *maestro* gestiona el ciclo de vida del AG (selección, cruza y mutación). Estos operadores son rápidos ya que solo manipulan vectores numéricos. Los procesos *trabajadores* son en los cuales el *maestro* distribuye los individuos de la población entre los núcleos disponibles del CPU. Cada núcleo recibe un subconjunto de modelos, ejecuta la simulación de SimPEG de forma independiente y devuelve el valor objetivo al maestro.

## 8.2. Componentes técnicos en el código

La implementación se apoya en tres elementos clave:

1. Crear un grupo de procesos *trabajadores*, el código detecta automáticamente cuántos hilos de ejecución puede manejar el procesador.
2. Sobrescribir la función map estándar de Python (que es secuencial) por la versión paralela Pool. Así cuando el algoritmo intenta evaluar la población, lo hace enviando datos a todos los núcleos simultáneamente.
3. Empaquetar los parámetros fijos (la simulación, los datos observados y las incertidumbres). Lo que permite que la función sea serializable y enviada a través de los diferentes procesos sin errores de memoria.

## 8.3. Ventajas y eficiencia

Cada proceso *trabajador* tiene su propio espacio de memoria. Esto evita conflicto de recursos, aunque requiere que el equipo tenga suficiente RAM para cargar las instancias de la simulación en cada núcleo.

Y lo más importante y la razón por la que decidimos utilizar la paralelización, la optimización del tiempo, en sistemas modernos, este enfoque permite realizar una tarea que de otro modo tardaría horas en menos de media hora, facilitando la exploración de múltiples configuraciones de capas o parámetros de regularización.

## Referencias

- [1] Asch, T. H., Abraham, J. D., & Cannia, J. C. (2019). *Hydrogeologic framework of selected areas in Butte and Glenn Counties, California*. Aqua Geo Frameworks, LLC. <https://data.ca.gov/dataset/airborne-electromagnetic-aem-pilot-studies>
- [2] Aster, R. C., Borchers, B., & Thurber, C. (2019). *Parameter estimation and inverse problems* (3<sup>a</sup> ed.). Elsevier.
- [3] Aster, R., Borchers, B., & Thurber, C. (2019). *Parameter Estimation and Inverse Problems* (3rd ed., p. 219). Elsevier.
- [4] Blair, T. C., Baker, F. G., & Turner, J. B. (1991). Cenozoic fluvial-facies architecture and aquifer heterogeneity, Oroville, California, Superfund site and vicinity. En A. D. Miall & N. Tyler (Eds.), *The three-dimensional facies architecture of terrigenous clastic sediments and its implications for hydrocarbon discovery and recovery* (Concepts in Sedimentology and Paleontology, Vol. 3). SEPM.
- [5] Blake, M. S., Helley, E. J., Jayko, A. S., Jones, D. L., & Ohlin, H. N. (1992). *Geologic map of the Willows 1:100,000 quadrangle, California* (Open-File Report 92-271). U.S. Geological Survey.
- [6] Burnett, J. L., Ford, R. S., & Scott, R. G. (1969). *Geology of the Richardson Springs quadrangle, California* (Map Sheet 13). California Division of Mines and Geology.
- [7] California Department of Water Resources. (2014). *Geology of the Northern Sacramento Valley, California*. Northern Region Office. <https://cawaterlibrary.net/wp-content/uploads/2017/05/Geology-of-the-Northern-Sacramento-Valley.pdf>
- [8] California Department of Water Resources. (2018). *Groundwater information center groundwater interactive map application*. <https://gis.water.ca.gov/app/gicima/>
- [9] California Division of Mines and Geology. (1962). *Geologic Guide to the Gas and Oil Fields of Northern California*. California Division of Mines and Geology Bulletin 181. 412 pp.
- [10] Deb, K., & Agrawal, R. B. (1995). Simulated binary crossover for continuous search space. *Complex Systems*, 9(2), 115–148.

- [11] Fortin, F.-A., De Rainville, F.-M., Gardner, M.-A., Parizeau, M., & Gagné, C. (2012). DEAP: Evolutionary Algorithms Made Easy. *Journal of Machine Learning Research*, 13, 2171–2175.
- [12] Goldberg, D. E. (1989). *Genetic algorithms in search, optimization, and machine learning*. Addison-Wesley.
- [13] Herrera, F., Lozano, M., & Sánchez, A. M. (2003). A taxonomy for the crossover operator for real-coded genetic algorithms: An experimental study. *International Journal of Intelligent Systems*, 18(3), 309–338.
- [14] Ingersoll, R. V., & Dickinson, W. R. (1981). Great Valley Group (Sequence), Sacramento Valley, California. En V. Frizzell (Ed.), *Upper Mesozoic Franciscan rocks and Great Valley Sequence, Central Coast ranges, California* (pp. 1–33). Society of Economic Paleontologists and Mineralogists, Pacific Section.
- [15] Kang, S., Knight, R., Greene, T., Buck, C., & Fogg, G. (2021). Exploring the model space of airborne electromagnetic data to delineate large-scale structure and heterogeneity within an aquifer system. *Water Resources Research*, 57(10), e2021WR029699. <https://doi.org/10.1029/2021WR029699>
- [16] Knight, R., Smith, R., Asch, T., Abraham, J., Cannia, J., Viezzoli, A., & Fogg, G. (2018). Mapping aquifer systems with airborne electromagnetics in the Central Valley of California. *Groundwater*, 55(6), 893–908.
- [17] Ley-Cooper, Y., & Davis, A. (2010). Can a borehole conductivity log discredit a whole AEM survey? En *Extended abstracts of the Australian Society of Exploration Geophysicists Annual Meeting*. Sydney, Australia.
- [18] McNeill, J. D. (1994). *Principles and application of time domain electromagnetic techniques for resistivity sounding* (Technical Note TN-27). Geonics Limited.
- [19] Michalewicz, Z. (1992). *Genetic algorithms + data structures = evolutionary programs*. Springer-Verlag.
- [20] Nabighian, M. N., & Macnae, J. C. (1991). Time domain electromagnetic prospecting methods. En M. N. Nabighian & J. D. Corbett (Eds.), *Electromagnetic methods in applied geophysics: Volume 2, Application, Parts A and B* (pp. 453–456). Society of Exploration Geophysicists.
- [21] Page, R. W. (1986). *Geology of the fresh ground-water basin of the Central Valley, California, with texture maps and sections* (Professional Paper 1401-C). U.S. Geological Survey.
- [22] Pinto Herrera, I. M. (2022). *Caracterización estructural e hidrogeológica de la zona cárstica Joya de Luna, San Luis Potosí, mediante métodos electromagnéticos* [Tesis de Maestría, Universidad Nacional Autónoma de México]. Tesis Digitales UNAM.

- [23] Planert, M., & Williams, J. S. (1995). *Ground Water Atlas of the United States, Segment 1, California, Nevada* (Hydrologic Investigations Atlas 730-B). U.S. Geological Survey.
- [24] Quinn, M. J. (2004). *Parallel programming in C with MPI and OpenMP*. McGraw Hill.
- [25] Saucedo, G. J., & Wagner, D. L. (1992). *Geologic map of the Chico quadrangle* (Regional Geologic Map 7A). California Division of Mines and Geology.
- [26] Steele, W. C. (1980). *Quaternary stream terraces in the northwestern Sacramento Valley, Glenn, Tehama, and Shasta Counties, California* (Open-File Report 80-472). U.S. Geological Survey.
- [27] Telford, W. M., Geldart, L. P., & Sheriff, R. E. (1990). *Applied geophysics* (2<sup>a</sup> ed.). Cambridge University Press.
- [28] U.S. Census Bureau. (2018). *Census 2010 total population*. <https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml#none>
- [29] U.S. Geological Survey. (2018a). *3DEP products and services*. The National Map. [http://nationalmap.gov/3DEP/3dep\\_prodserv.html](http://nationalmap.gov/3DEP/3dep_prodserv.html)
- [30] U.S. Geological Survey. (2018b). *USGS stream stage water watch*. National Water Information System. [https://waterwatch.usgs.gov/index.php?id=real&sid=w\\_\\_gmap&r=ca](https://waterwatch.usgs.gov/index.php?id=real&sid=w__gmap&r=ca)
- [31] Ward, S. H., & Hohmann, G. W. (1987). Electromagnetic Theory for Geophysical Applications. En M. N. Nabighian & J. D. Corbett (Eds.), *Electromagnetic Methods in Applied Geophysics: Volume 1, Theory* (Vol. 3, p. 0). Society of Exploration Geophysicists. <https://doi.org/10.1190/1.9781560802631.ch4>
- [32] Werthmüller, D. (2017). An open-source full 3D electromagnetic modeler for 1D VTI media in Python: Empymod. *Geophysics*, 82(6), WB9–WB19. <https://doi.org/10.1190/geo2016-0626.1>
- [33] Zhdanov, M. S. (2009). *Geophysical electromagnetic theory and methods*. Elsevier.