



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

**PROGRAMA DE MAESTRÍA Y DOCTORADO EN
INGENIERÍA**

CENTRO DE INVESTIGACIÓN EN ENERGÍA

**MODELADO DE PROCESOS DE TRANSPORTE DE
FLUIDOS Y CALOR EN SISTEMAS GEOTÉRMICOS
USANDO REDES NEURONALES Y SIMULACIÓN MONTE
CARLO**

T E S I S

**QUE PARA OPTAR POR EL GRADO DE:
DOCTOR EN INGENIERÍA
ENERGÍA**

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

**TUTOR:
DR. EDGAR R. SANTOYO GUTIÉRREZ**



**TEMIXCO, MORELOS.
2012**



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

Jurado asignado

Presidente: DR. GARCÍA VALLADARES OCTAVIO

Secretario: DR. EDGAR R. SANTOYO GUTIÉRREZ

Vocal: DR. PANDARINATH KAILASA

1^{er} Suplente: DR. ANDAVERDE ARREDONDO JORGE ALBERTO

2^{do} Suplente: DR. HERNÁNDEZ PÉREZ JOSÉ ALFREDO

Lugar en donde se realizó la tesis:

CENTRO DE INVESTIGACIÓN EN ENERGÍA
UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Director de tesis:

Dr. EDGAR R. SANTOYO GUTIÉRREZ

FIRMA

La fe, el esfuerzo y optimismo dedicado a lo largo de los años de estudio, son el fruto de la gente que creyó en mi persona, apoyándome en todo sentido dándome la mano a través de la educación. Es por ello que este trabajo esta dedicado a las personas que a lo largo de mi vida me han dado la formación de ser persona.

Con mucho cariño a ti mamá, que siempre me diste el gusto en todo, me escuchaste, aconsejaste y creíste en mi hasta el último minuto, a ti papá, que pusiste la mano dura pero suave, siempre estuviste el ejemplo a seguir en todo, contigo padre aprendí aspirar la ciencia, a Mariam, Hassan, y Soundous Ali y sus familiares, por el amor y el cariño que tienen hacia mi, a ti Maya, que me motivaste y a Amina que estuvo a mi lado.

Bassam Ahmad Ali

AGRADECIMIENTOS

- * Al Dr. Edgar Santoyo, mi Director de Tesis, debo agradecerle de manera especial y sincera por su presencia incondicional, sus apreciados y relevantes aportes, críticas, comentarios y sugerencias durante el desarrollo de esta investigación. Debo agradecerle por su esfuerzo de entenderme en cada etapa de este trabajo. Su orientación hizo de este trabajo una tesis completa y terminada.
- * Al Dr. Jorge Andaverde, por su constante y paciente seguimiento y asistencia, compartiendo su tiempo y conocimiento de manera generosa durante el desarrollo del presente trabajo y por su amistad.
- * Al Dr. José Alfredo Hernández, por su participación y apoyo en el desarrollo de las herramientas computacionales, y por su sostenido apoyo moral.
- * Al Dr. Ignacio Torres (q.e.p.d), por su colaboración y orientación durante las evaluaciones de este proyecto como miembro del comité y su apoyo incondicional como jefe del posgrado.
- * A los Drs. Kailasa Pandarinath y García Valladares, miembros del Jurado de esta tesis, por sus valiosas sugerencias a la versión original del manuscrito, que contribuyeron al mejoramiento y ordenamiento del presente trabajo.
- * Al Dr. Luis Manuel Gagerro, por sus valiosos consejos durante estos años.
- * A todos mis amigos y colegas, (Alejandra, Cristian, Jimena, Daniel, Efraín, Michel, Sanjeet, Orlando, Wong, Gerardo, Osvaldo, Iván, Mauricio, German,... y la lista es larga) por los buenos momentos compartidos y su constante espíritu servicial.
- * A la dinámica del posgrado, (Cristina, Lulu y Sara), por su disposición, paciencia y apoyo en los tramites durante mi estancia doctoral.
- * A la Maestra Mirna, Alfredo y Margarita por el apoyo recibido en el equipamiento y computación.
- * Y en especial mi agradecimiento a mis compatriotas y compañeros de la comunidad Marroquí en México, (Amina, Outmane, Achir y Mia, Younes peque, Younes grande, Lailai, Kamal, Ferial, Abe, Jawad, Rachid Tadili, Benami, Abatal, Said y youssef El Wahabi...), por todo lo que hemos compartido, la amistad y el apoyo recibido.
- * A la Sra. Carmen y La Srta. Dama por apoyar en cuidar nuestra hija Maya durante la fase de redacción.
- * Mi agradecimiento al Posgrado de Ingeniería de la Universidad Nacional Autónoma de México, al Centro de Investigación en Energía de la misma universidad donde se realizó esta investigación, al Centro de Investigación en Ingeniería y Ciencias Aplicadas de la Universidad Autónoma del Estado de Morelos en donde se colaboró parte de este proyecto y a la Facultad de Química de la Universidad Veracruzana, Campos Coatzacoalcos en donde se realizó una estancia de investigación.
- La presente investigación se financió parcialmente con fondos provenientes del Consejo Nacional de Ciencias y Tecnología (México), y una beca del Programa Movilidad Estudiantil durante la estancia de investigación financiada por Banco Santander.

Con absoluta sinceridad, mi agradecimiento a todos los que hice mención ya que con su aporte hicieron posible esta presentación.

Índice

Lista de Figuras	i
Lista de Tablas	ii
Resumen	1
Abstract	3

Introducción General **5**

Capítulo 1 Redes Neuronales Artificiales y Simulación Monte Carlo

1.1 Introducción	11
1.2 Redes Neuronales Artificiales (RNA)	12
1.2.1 <i>Definición de las RNA</i>	13
1.2.2 <i>Funciones de activación</i>	14
1.2.3 <i>Clasificación de las RNA</i>	15
1.2.4 <i>Métodos de aprendizaje</i>	17
1.2.5 <i>Modelo de Retro-propagacion 'Backpropagation'</i>	18
1.2.6 <i>Ventajas y desventajas de las RNA</i>	18
1.3 Simulación de Monte Carlo	19
1.3.1 <i>Propagación de incertidumbres con Monte Carlo</i>	19
1.3.2 <i>Números aleatorios</i>	21
1.3.3 <i>Función de distribución (Gaussiana)</i>	21
1.3.4 <i>Software utilizado</i>	22

Capítulo 2 Desarrollo de un nuevo método para estimación de las temperaturas estabilizadas de formación en sistemas geotérmicos usando redes neuronales artificiales

2.1 Introducción	23
2.2 Metodología	25
2.2.1 <i>Creación de base de datos para el desarrollo del método</i>	26
2.2.2 <i>Estimación inicial de las Temperaturas Estabilizadas de Formación normalizadas</i>	28
2.2.3 <i>Procedimiento general para la estimación de las TEF con RNA</i>	29
2.2.4 <i>Arquitectura propuesta de la RNA</i>	32
2.3 Modelo de la RNA	34
1.3.1 <i>Entrenamiento computacional del modelo</i>	35
1.3.2 <i>Validación numérica del modelo</i>	37

Capítulo 3 Determinación de gradientes de presión en pozos geotérmicos usando redes neuronales artificiales y un simulador numérico de pozos: Estudio comparativo

3.1 Introducción	39
3.2 Marco teórico	42
3.2.1 <i>Gradientes de presión</i>	42
3.2.2 <i>Simulador de pozos geotérmicos (GEOWELLS)</i>	43
3.2.3 <i>Redes neuronales artificiales</i>	45
3.3 Metodología Computacional	45
3.4 Resultados y discusión	51
3.4.1 <i>Evaluación de los modelos desarrollados</i>	51
3.4.2 <i>Análisis de sensibilidad</i>	55
3.4.3 <i>Validación de los gradientes de presión calculados por las RNA y GEOWELLS</i>	56
3.4.4 <i>Análisis estadístico comparativo entre datos de predicción y mediciones de campo</i>	58

Capítulo 4 Estimaciones de Flujo de calor y sus incertidumbres mediante un modelo de transferencia de calor simplificado y el método de Monte Carlo

4.1 Introducción	61
4.2 El Campo geotérmico Los Humeros (CGLH)	62
4.2.1 <i>Evolución volcánica</i>	64
4.2.1 <i>Termodinámica del yacimiento del CGLH</i>	64
4.2.3 <i>Geología del subsuelo del CGLH</i>	66
4.3 Metodología usada para la determinación de flujo de calor	68
4.3.1 <i>Temperaturas Estabilizadas de Formación</i>	69
4.3.2 <i>Conductividad térmica de rocas y su error</i>	71
4.3.3 <i>Estimación del flujo de calor mediante un modelo de transferencia de calor simplificado y el método Monte Carlo</i>	72
4.4 Análisis de los resultados y discusiones	76
4.4.1 <i>Validación de la metodología</i>	76
4.4.2 <i>Aplicación de la metodología al CGLH</i>	79

Conclusiones 87

Apéndice A 91

Apéndice B 94

Referencias 96

Anexo A 105

Anexo B 116

Anexo C 162

Lista de Figuras

Figura 1 Esquema de un sistema de roca seca caliente a escala comercial

Figura 2 Objetivos y justificaciones del proyecto doctoral

Figura 1.1 Presentación de una simple neurona

Figura 1.2 Diferentes tipos de funciones de activación

Figura 1.3 Modelo RNA monocapa

Figura 1.4 Modelo RNA multicapas

Figura 1.5 Modelos de aprendizaje de la RNA

Figura 1.6 Propagación de incertidumbres mediante Monte Carlo de un modelo estocástico

Figura 2.1 Esquema general de la metodología desarrollada

Figura 2.2 Arquitectura general de la RNA

Figura 2.3 Algoritmo utilizado para el proceso de la optimización y aprendizaje de RNA

Figura 2.4 Comparación estadística entre las TEF estimadas por los métodos analíticos y las simuladas por RNA

Figura 2.5 Evaluación de la precisión de la RNA utilizando una comparación estadística entre datos de TEF simulados por la red y las TEF verdaderas

Figura 3.1 Algoritmo de entrenamiento de las RNA

Figura 3.2 Esquema general de la metodología desarrollada

Figura 3.3 Arquitectura de modelos de las RNA

Figura 3.4 Comparación entre valores de presión medidos y los simulados por la RNA₁

Figura 3.5 Comparación entre valores de presión medidos y los simulados por la RNA₂

Figura 3.6 Perfiles de presión predichos mediante GEOWELLS y RNA vs. medidos para el pozo Okoy-7

Figura 3.7 Perfiles de presión predichos mediante GEOWELLS y RNA vs. medidos para el pozo Az-42

Figura 3.8 Perfiles de presión predichos mediante GEOWELLS y RNA vs. medidos para el pozo M-39

Figura 4.1 Localización y escenario tectónico del Cinturón Volcánico Mexicano

Figura 4.2 (a, b y c) Gradientes de TEF calculadas por el método de Horner (Lineal y Cuadrático) para los pozos H20, H28 y H31

Figura 4.3 Metodología desarrollada para la estimación de flujo de calor conductivo en pozos

Figura 4.4 Comparación de flujos de calor reportados en la literatura por Pfister et al. (1998) vs. los calculados por el método de Monte Carlo

Figura 4.5 Propagación de errores del FC con el método de MC en el campo geotérmico de Los Humeros

Figura 4.6 Distribuciones de FC calculado con MC usando las TEF estimadas con el método de Horner con regresiones (lineales y cuadráticas)

Lista de Tablas

- Tabla 2.1 Base de datos utilizada para el entrenamiento y validación de la RNA
- Tabla 2.2 Soluciones numéricas simplificadas de diferentes métodos analíticos (HM, BM, KEM, LM, MM, SRM y CSM) utilizados para la determinación de las TEF en pozos geotérmicos
- Tabla 2.3 Valores de TEF estimadas con los siete métodos analíticos utilizados
- Tabla 2.4 Parámetros óptimos de la RNA
-
- Tabla 3.1 Distribución estadística de la BDG utilizada en el proceso de entrenamiento y la evaluación de los modelos de la RNA
- Tabla 3.2 Diferentes pruebas estadísticas utilizadas para la evaluación de los modelos desarrollados
- Tabla 3.3 Parámetros óptimos resultado del entrenamiento de los modelos RNA₁ y RNA₂
- Tabla 3.4 Análisis de sensibilidad de los variables de entrada en los modelos de la RNA
- Tabla 3.5 Comparación estadística de los gradientes de presión calculados por (RNA₁, RNA₂ y GEOWELLS)
-
- Tabla 4.1 Características termodinámicas de las zonas productoras del CGLH
- Tabla 4.2 Geología del subsuelo del CGLH
- Tabla 4.3 Conductividad térmica de roca seca y saturada
- Tabla 4.4 Validación de la metodología valores de flujo de calor conductivo medidos en el campo vs. Flujo de calor calculado por el método de MC
- Tabla 4.5 Valores de TEF y flujo de calor conductivo en diferentes profundidades de los pozos estudiados del colapso central
- Tabla 4.6 Análisis de FC conductivo para el pozo H20

Resumen

La evaluación de sistemas geotérmicos hidrotermales tradicionales y mejorados requiere se desarrollen metodologías para evaluar de manera confiable las variables físicas que evidencian las reservas disponibles para la explotación de estos recursos.

En el presente trabajo se han planteado tres objetivos específicos relacionados con la evaluación de tres variables fundamentales para conceptualizar un sistema geotérmico: a) Las temperaturas estabilizadas de formación; b) Los perfiles de presión dentro de los pozos geotérmicos productivos; y c) la estimación de flujo de calor en estos sistemas.

Una nueva metodología, aplicando Redes neuronales artificiales (RNA) fue desarrollada para la estimación de las temperaturas estabilizadas de formación, así como sus incertidumbres. De esta metodología se obtuvo una nueva ecuación basada en RNA la cual fue validada estadísticamente con datos medidos en campo. Esta ecuación puede ser aplicada en el rango de temperatura de 45 °C a 263 °C. Aplicando también la técnica computacional de RNA se obtuvieron perfiles de presión de pozos y los resultados fueron validados con datos de 10 pozos productores de campos geotérmicos en explotación. Así mismo se realizó un análisis estadístico comparativo entre los valores simulados, tanto por modelos de RNA obtenidos en el presente trabajo, como los inferidos por un simulador de pozos geotérmicos, y las mediciones de campo. Se encontró que no existen diferencias significativas entre los valores predichos por los modelos RNA con los datos de campo. Asimismo, se desarrolló una nueva metodología estadística, en donde se aplica la simulación Monte Carlo para determinar valores de flujo de calor confiables en el campo geotérmico los Humeros, Puebla, incluyendo sus incertidumbres. Finalmente se concluye que todas las herramientas desarrolladas en este proyecto doctoral pueden ser aplicadas de manera confiable para ayudar a definir un modelo conceptual de flujo de calor, tanto en sistemas geotérmicos hidrotermales tradicionales, como en los sistemas de roca seca mejorados. Las herramientas desarrolladas en el presente proyecto doctoral pueden ser utilizadas en la industria geotérmica para la estimación y optimización de los recursos geotérmicos, además de contribuir al ahorro en los programas de explotación actuales y los que se pudieran desarrollar en sistemas geotérmicos mejorados.

Durante el desarrollo de este trabajo se ha observado que en la actual industria geotérmica no se tiene cuidado en la calidad de mediciones en cuanto a sus incertidumbres, por lo que se sugiere que en las mediciones físicas de campo se incluya el registro de ellas y se reporten a la comunidad científica y tecnológica, esto con el fin de tener una evaluación más fiable de estos sistemas geotérmicos.

Abstract

The evaluation of hydrothermal and enhanced geothermal systems requires the development of new computational tools to evaluate their thermophysical properties and heat reserves. Such parameters are fundamental for the exploitation of these geothermal resources in terms of confidence estimates of power potential and the exploitable energy base. These parameters can be also used for decision making tasks in newly explored geothermal prospects or fields.

The present PhD work has been carried out to develop new computation tools to be used for the reliable estimation of heat flows and some other fundamental variables. Such parameters are proposed to be used to conceptualize a geothermal system under either exploration or exploitation conditions. The following specific objectives were defined to reach the main goal of the research project: (a) the estimation of static formation temperatures and wellbore pressure profiles of drilled wells using artificial neural network (ANN) techniques; and (b) the determination of heat flows in geothermal systems using a simple heat transfer model, together with some error propagation and Monte Carlo methods to estimate the associated uncertainties.

In relation with the first objective, a new methodology based on several ANN architectures was successfully developed for the estimation of the static formation temperatures with uncertainties. As a result of this study, a new practical equation was developed and successfully validated statistically with field data (i.e., bottomhole temperatures and shut-in times logged during the wellbore drilling operations). This equation has been effectively applied for bottomhole temperatures ranging from 45 °C to 263 °C. For the second part of this objective, wellbore pressure profiles of producer geothermal wells were also predicted using ANN techniques and wellbore numerical simulations (using the GEOWELLS simulator). Such results were efficiently validated and statistically well compared with field data reported for some producer wells drilled in geothermal fields under exploitation conditions. As a result of the statistical comparison, the pressure profiles predicted with ANN tools showed an acceptable agreement with those results inferred from numerical simulations and field data.

Finally and as part of the last objective, a new numerical and statistical methodology based on a simple version of the Fourier heat transfer model, together with a coupling of Monte Carlo simulations was developed for the determination of heat flows in geothermal

systems. This methodology was effectively validated with data reported in the literature, and applied for the estimation of heat flows in Los Humeros geothermal field.

Introducción general

*“La Tierra proporciona lo suficiente para satisfacer las necesidades de cada hombre,
pero no de la codicia del ser humano”*

Partiendo de esta reflexión del gran líder y pensador hindú Gandhi (1869 - 1948), se puede explicar la situación energética y el cambio climático que vive el mundo actual.

Hoy día estamos asistiendo a los primeros estadios del agotamiento del petróleo. Las subidas reiteradas del precio del barril desde el año 2000 solo es un indicativo del principio del fin, pero aunque los expertos cifren su final a partir de 20 años, es el cenit de su producción su verdadera preocupación. Un fenómeno que ya han experimentado de forma local la mayoría de los países productores de crudo fuera de la Organización de Países Exportadores de Petróleo (OPEP), y cuyo principal efecto es que, tras alcanzar el pico de máxima producción, la extracción de crudo disminuye inexorablemente hasta su final, encareciendo y dificultando su extracción en la etapa final para producir crudo aun que este sea cada vez de peor calidad (Laherrère, 2001 y Garrido, 2010).

Por otra parte las concentraciones de CO₂ en la atmósfera a lo largo de los últimos 100 años han aumentado una tercer parte, principalmente por el uso de combustibles fósiles y la tala de bosques. Más de la mitad del efecto invernadero procede de CO₂ y tres cuartas partes de este CO₂ vienen de la combustión de los combustibles fósiles. Si continuamos como hasta ahora, la concentración atmosférica de CO₂ hacia mediados del siglo XXI será más de dos veces de lo que era antes de la revolución industrial. Según el panel de expertos para el cambio climático, la temperatura de la tierra aumentará entre 1.5 °C y 4.5 °C desde ahora hasta 2100, causando grandes cambios climatológicos y su impacto socioeconómico sobre la humanidad (Pachauri y Reisinger, 2007).

Viendo este panorama, es evidente la necesidad en apoyarse en los fuentes de las energía limpias menos contaminantes y que sustituyen a los combustibles fósiles como es el caso de las Energías Renovables (ER). Es por lo antes mencionado que el desarrollo y la

investigación para el mejoramiento y el aumento de la eficiencia de estas fuentes Renovables se vuelve una necesidad prioritaria y una gran responsabilidad y desafío para las personas que se dedican a la ciencia y al avance tecnológico.

Actualmente el uso de energías renovables está experimentando un fuerte impulso en la generación de energía eléctrica. Dentro de este contexto, se encuentra la energía geotérmica, la cual constituye una de las fuentes alternas que ha mostrado una mayor madurez tecnológica (Demirbas, 2006; Lund, 2007). La capacidad geotermoeléctrica instalada a nivel mundial alcanzó los 10,715 MWe en 39 países (~0.4% de la generación mundial) y se pronostica que para el 2015, la capacidad instalada alcanzará los 18,500 MWe, esto es ~75,000 GWh con un factor de capacidad de ~77-85% (Bertani, 2010; Lund *et al.*, 2010 y Santoyo, 2011). Otros Pronósticos más optimistas indican que la capacidad geotermoeléctrica podría alcanzar los 140 GWh en el 2050 (1.104 TWh/a con un factor de capacidad del 90%), implicando un crecimiento de ~14.5 veces la capacidad instalada actual (Fridleifsson *et al.*, 2008).

En México la actual producción de energía eléctrica a partir de los recursos geotérmicos es de 964.5 Mwe, la cual está generada en cuatro campos geotérmicos: Cerro Prieto (Baja California, 720 Mw), Los Azufres (Michoacán, 188 Mw), Los Humeros (Puebla, 40 Mw) y Las Tres Vírgenes (Baja California, 10 Mw). Esta producción representa ~3% de la capacidad total instalada en México (177.80 TWh; Bertani 2005; Gutiérrez-Negrín y Quijano-León 2005, CFE 2009), y coloca al país en el cuarto lugar mundial, detrás de Estados Unidos, Filipinas e Indonesia.

La utilización de los recursos geotérmicos ha sido posible debido a la disponibilidad de yacimientos, en la mayoría de los casos a partir de **sistemas hidrotermales convectivos** (Dickson y Fanelli, 2005), en los cuales las condiciones geológicas facilitan el transporte de calor a través de un fluido (fase líquida o vapor), desde zonas profundas de alta temperatura hacia zonas menos profundas (Gutiérrez-Negrín y Quijano-León, 2005).

Los Sistemas de Roca Seca Caliente (En inglés: Enhanced Geothermal Systems) o **Sistemas Geotérmicos Mejorados** (SGM) son evaluados actualmente como una nueva generación de sistemas con alto potencial térmico (Santoyo y Barragán-Reyes, 2010;

Santoyo y Torres-Alvarado, 2010; Santoyo, 2011), caracterizados principalmente por la disponibilidad de una fuente de calor (roca caliente) y la ausencia de fluidos debido a las características geológicas de estos sistemas, generalmente rocas cristalinas, no permeables (Tester *et al.*, 2007; Haring *et al.*, 2008). La explotación de un SGM requiere la creación artificial de un yacimiento (reservorio), lo cual se crea experimentalmente mediante el bombeo de agua de alta presión a través de pozos hasta un cuerpo de rocas calientes y compactas, causando así su fracturamiento hidráulico. El agua circula en estas fracturas artificiales, extrayendo el calor de la roca circundante que actúa como un yacimiento natural. Este yacimiento es posteriormente alcanzado por un segundo pozo que se emplea para extraer el agua calentada. Por lo tanto, el sistema consiste en: (i) el pozo empleado para fracturamiento hidráulico a través del cual se inyecta el agua fría; (ii) el reservorio artificial; y (iii) el pozo utilizado para extraer el agua caliente. Todo el sistema, conjuntamente con la planta de utilización en la superficie del terreno, conforma un circuito cerrado (Dickson y Fanelli, 2005; Santoyo, 2011; ver Figura 1).

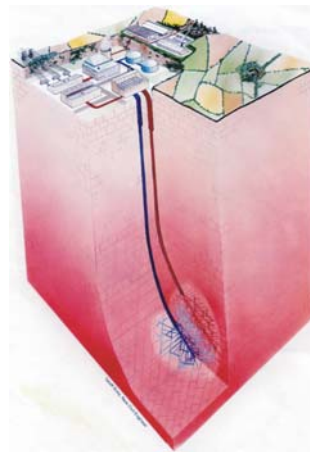


Figura 1 Esquema de un sistema de roca seca caliente a escala comercial (Dickson y Fanelli, 2005).

El modo de evaluar estos sistemas es mediante la modelación geotérmica cuyo objetivo principal es dar a conocer la temperatura interna y la distribución de densidad de flujo de calor de una zona de estudio, a través de un modelo conceptual del yacimiento. La temperatura y la transferencia de calor están involucradas en todos los procesos geológicos y geofísicos de la Tierra (Combs, 1980; Pollack y Henry, 1985; Chapman y Furlong, 1992; Pollock *et al.*, 1993; Hochstein y Bromley, 2005; Majorowicz y Grasby, 2010).

Nuestra comprensión del régimen térmico interno de la Tierra se basa principalmente en tres factores: en primer lugar con mediciones directas, en segundo lugar, a través de observaciones geológicas, y en tercer lugar, mediante modelos geofísicos (Jokinen, 2000). Las mediciones directas incluyen los registros de temperaturas en los pozos perforados, los flujos de calor durante la producción y las propiedades de transporte y de calor de rocas. Las mediciones de las propiedades térmicas y de transporte de rocas se realizan en laboratorios especializados bajo condiciones de temperaturas y presión elevadas.

El conocimiento de las condiciones térmicas en la litosfera se basa en modelos teóricos de transferencia de calor limitados por datos geológicos y geofísicos. Una de los problemas importantes se centra en las incertidumbres asociados con la temperatura calculada y los resultados de densidad de flujo de calor, así como el efecto de las incertidumbres de las propiedades térmicas de las rocas y las condiciones de frontera.

Para cumplir con el objetivo principal planteado de esta tesis, que es crear herramientas computacionales para obtener un modelo conceptual confiable de un sistema geotérmico, es necesario cumplir con una serie de objetivos específicos que pueden ser útiles tanto en las fases de exploración, como de explotación de estos sistemas. Estos objetivos son de carácter termodinámico y consisten en: (a) las estimaciones de las temperaturas estabilizadas de formación; (b) la determinación de los perfiles de presión en pozos geotérmicos productivos; y (c) la evaluación de los flujos de calor conductivo en sistemas geotérmicos.

En cuanto a las Temperaturas Estabilizadas de Formación (TEF) se han desarrollado una serie de trabajos tales como los de Hasan *et al.*, 1994; García *et al.*, 1998, Santoyo *et al.*, 2000; Kutasov *et al.*, 2003; Andaverde *et al.*, 2005 y Espinoza-Ojeda *et al.*, 2011. En la presente tesis se obtuvo una ecuación confiable mediante redes neuronales artificiales (RNA) para estimación de dicha propiedad. Por otra parte, se pretendió mediante la misma herramienta (RNA), y utilizando datos reales de pozos, determinar perfiles de presión de pozos geotérmicos productores, a partir de los cuales se puede estimar el flujo de fluidos en pozos productores. Por último, se desarrolló una metodología numérico-estadística para calcular el flujo de calor en un sistema geotérmico, en donde el mecanismo de transferencia de calor conductivo es el predominante. Tanto en el cálculo de las TEF, como en el flujo de calor, se evaluaron las incertidumbres asociados con estas variables.

En la figura 2 se representa esquemáticamente el objetivo principal de esta tesis, así como sus tres objetivos específicos. Cabe resaltar que las justificaciones de dicho proyecto doctoral se centran en un modelo que puede servir para la estimación de reservas energéticas geotérmicas, el desarrollo de programas de explotación de sistemas geotérmicos y su posible aplicación en el uso de bombas de calor geotérmicas para el acondicionamiento climático.

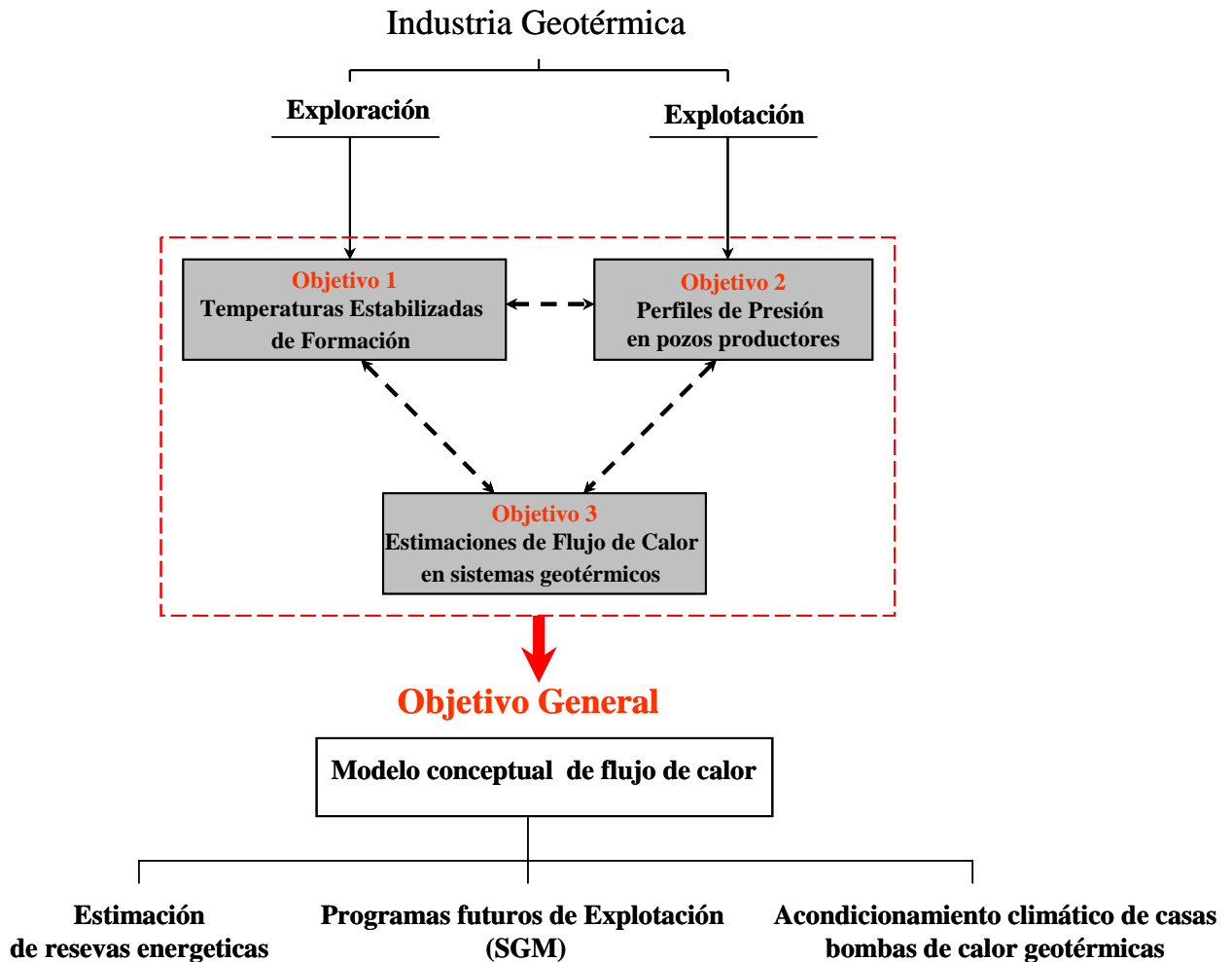


Figura 2 Objetivos y justificaciones del proyecto doctoral.

A continuación se describe la estructura de esta tesis. En la introducción, se expone la posición mundial actual de la energía geotérmica, los sistemas geotérmicos mejorados considerados como una nueva línea de investigación y el futuro de esta energía, la importancia de la estimación de flujos de calor en estos sistemas para la evaluación de los recursos y futuros planes de exploración, así como el objetivo general y los objetivos particulares de este proyecto. El primer capítulo presenta conceptos generales de las

herramientas computacionales utilizadas en el desarrollo de esta investigación (Redes Neuronales Artificiales y el método de Monte Carlo). El segundo capítulo describe el desarrollo de una nueva ecuación basada en la red neuronal para la estimación de las temperaturas estabilizadas de formación y los resultados de su validación. El tercer capítulo presenta un modelo neuronal para la predicción de los perfiles de presión dentro de los pozos geotérmicos productores, así como un análisis estadístico comparativo con los resultados obtenidos de un simulador de pozos y los datos de medición registrados en el campo. El cuarto capítulo presenta una nueva metodología estadístico-numérica basada en simulaciones Monte Carlo para la estimación de flujos de calor conductivos sus incertidumbres, así como la aplicación de esta nueva metodología en el campo geotérmico Los Humeros (México). Finalmente, se presentan los apéndices, y copias de los trabajos publicados producto de esta investigación.

Capítulo 1

Redes Neuronales Artificiales y Simulación Monte Carlo

1.1 Introducción

La evolución tecnológica, cuyos orígenes se remontan a los comienzos de la historia humana, ha tenido, a partir de la década de los cuarenta, un impulso extraordinario gracias al diseño y la progresiva popularización de las computadoras. Estos equipos han sido definidos como máquinas capaces de realizar y controlar a gran velocidad cálculos y procesos complicados que requieren una rápida toma de decisiones. Su función consiste por tanto, en tratar la información que se le suministra y proveer los resultados requeridos. Sin embargo este enunciado tan simple está definiendo un hito tradicional.

La computadora es, hoy por hoy, incapaz de hacer algo para lo que no ha sido programada. Sin embargo, el progreso que representa en la capacidad para realizar un número muy grande de cálculos a una alta velocidad (milésimas de segundo), permitió el desarrollo de modelos y algoritmos computacionales robustos donde se requiere de muchas operaciones (p.ej., algoritmos de inteligencia artificial: las redes neuronales artificiales, la lógica difusa, los algoritmos genéticos, entre otros, y los métodos estadísticos avanzados como es el caso de la simulación de Monte Carlo).

En el presente capítulo se presenta una definición de las redes neuronales artificiales (RNA), sus diferentes arquitecturas y la forma en que se entrenan. Al mismo tiempo se explica que es el método estadístico de Monte Carlo (MC) y los detalles de su simulación. Estas dos herramientas computacionales (RNA y MC) se utilizaron en el desarrollo de este proyecto doctoral con aplicaciones para la solución de los geo-problemas planteados en la introducción (Capítulos 2,3 y 4).

1.2 Redes Neuronales Artificiales (RNA)

Las primeras explicaciones teóricas sobre el cerebro y el pensamiento fueron dadas por Platón (427-347 a.C.) y Aristóteles (348-422 a.C.). Las mismas ideas las mantuvieron Descartes (1569-1650) y los filósofos empiristas del siglo XVIII. La clase de las llamadas máquinas cibernéticas, a la cual la computación neuronal pertenece, tiene más historia de la que se cree cuando Herón (100 a.C) construyó un autómatas hidráulico.

En 1936, Alan Turing fue el primero en estudiar el cerebro como una forma de ver el mundo de la computación. Sin embargo, los primeros teóricos que concibieron los fundamentos de la computación neuronal fueron Warren McCulloch, un neurofisiólogo, y Walter Pitts, un matemático, quienes en 1943, lanzaron una teoría sobre de la forma en la que trabaja las neuronas ‘Un Cálculo Lógico de la Inminente Idea de la Actividad Nerviosa’ (McCulloch y Pitts, 1987). Ellos modelaron una red neuronal simple mediante circuitos eléctricos. En 1949, Donald Hebb (Hebb, 1989) escribió el libro: *La organización del comportamiento*, en el que se establece una conexión entre psicología y fisiología. Donald Hebb fue el primero en explicar los procesos del aprendizaje el elemento básico de la inteligencia humana, desde un punto de vista psicológico, desarrollando una regla de como el aprendizaje ocurría. Aún hoy, este es el fundamento de la mayoría de las funciones de aprendizaje que pueden hallarse en una red neuronal. Su idea fue que el aprendizaje ocurría cuando ciertos cambios en una neurona eran activados. También intentó encontrar semejanzas entre el aprendizaje y la actividad nerviosa. Los trabajos de Hebb formaron las bases de la teoría de RNA. En 1957, Frank Rosenblatt planteó el desarrollo del Perceptron, el cual es concebido como la red neuronal más antigua, la cual se utiliza hoy en día como el reconocedor de patrones. Este modelo era capaz de generalizar, es decir, después de haber aprendido una serie de patrones podía reconocer otros similares, aunque no se le hubiesen presentado anteriormente, aunque tenía una serie de limitaciones. En 1959, escribió el libro *Principios de Neurodinámica* (Rosenblatt, 1989), en el que confirmó que, bajo ciertas condiciones, el aprendizaje del Perceptron convergía hacia un estado finito conocido como el teorema de Convergencia del Perceptron. En 1960, Bernard Widrow y Marcial Hoff (Widrow y Hoff, 1989) desarrollaron el modelo ADALINE (ADAPtative LINear Elements), que se aplica por primera vez a un problema real (filtros adaptativos para eliminar ecos en las líneas telefónicas) utilizado comercialmente durante varias décadas. En 1974, Paul Werbos

(Werbos, 1974) desarrolló la idea básica del algoritmo numérico de aprendizaje de propagación hacia atrás (conocido como Backpropagation) y cuyo significado quedó definitivamente aclarado en 1986, cuando John David Rumelhart y Geoffrey Hinton replantearon este algoritmo (Rumelhart *et al.*, 1986). A partir de 1986, el panorama fue alentador con respecto a las investigaciones, el desarrollo y la aplicación de las redes neuronales. En la actualidad, son numerosos los trabajos que se realizan y publican cada año, las aplicaciones nuevas que surgen (sobre todo en el área de simulación y control) y las empresas que lanzan al mercado productos nuevos, tanto hardware como software.

1.2.1 Definición de las RNA

Una RNA es una estructura compuesta de un número de unidades interconectadas (neuronas artificiales). Cada unidad posee una característica de entrada/salida e implementa una computación local o función. La salida de cualquier unidad está determinada por su característica de entrada/salida, su interconexión con otras unidades, y (posiblemente) de sus entradas externas. La RNA desarrolla usualmente una funcionalidad general a través de una o más formas de entrenamiento (Schalkoff, 1997).

El cerebro humano contiene más de 100 billones de elementos de procesos llamados neuronas, que se comunican a través de conexiones llamadas sinapsis. Cada neurona esta compuesta por tres partes fundamentales: el cuerpo, dendritas y axón. El cuerpo en su capa externa tiene la capacidad única de generar impulsos nerviosos. Las dendritas que son como las ramas que salen del cuerpo, poseen algunas conexiones sinápticas en donde se reciben señales que generalmente vienen de otros axones. El axón se encarga de activar o inhibir otras neuronas las cuales a su vez son activadas por cientos o miles de otras neuronas. El funcionamiento de una RNA está basado en este diseño.

Básicamente, Un modelo de una neurona simple consiste en aplicar un conjunto de entradas, cada una representando la salida de otra neurona, o una entrada del medio externo, realizar una suma ponderada con estos valores y “filtrar” este valor con una función como se puede observar en la Figura 1.1.

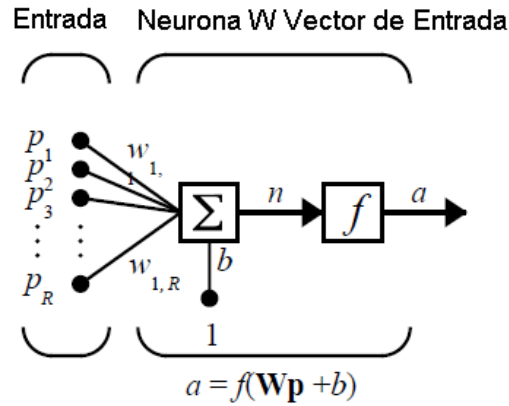


Figura 1.1 Presentación de una simple neurona (Demuth y Beale, 1998).

Cada neurona artificial recibe un vector In de entrada que corresponde a todas aquellas señales $p_1, p_2, p_3 \dots p_R$ que llegan a la sinapsis. Cada una de estas señales se multiplica por un peso asociado $w_1, w_2, w_3 \dots w_R$. Al conjunto de pesos se le denomina vector W . Cada peso representa la “intensidad” o fuerza de conexión de una sinapsis en una neurona biológica. Los resultados de estas multiplicaciones se suman y simulan vagamente el cuerpo de una neurona biológica (ver Figura 1.1).

1.2.2 Funciones de activación

Una vez que la entrada neta ha sido calculada, se transforma en el valor de activación y una vez hecho esto se puede aplicar la función de salida que es la encargada de transformar el valor de la entrada neta, en el valor de salida del nodo. La función de activación o transferencia f puede ser lineal o no lineal. Existen diferentes funciones de activación típicas que determinan distintos tipos de neuronas:

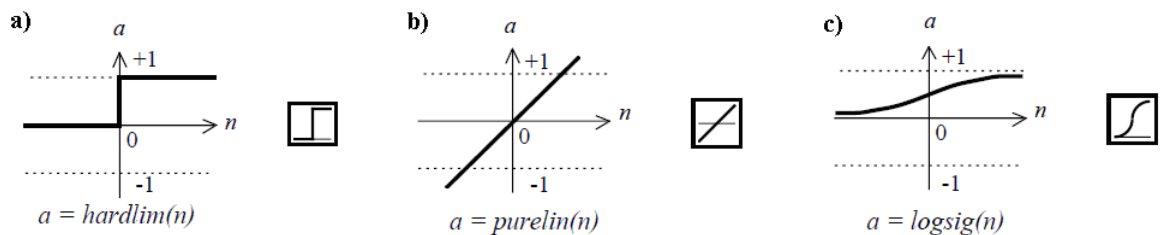


Figura 1.2 Diferentes tipos de funciones de activación.

-Neurona de función escalón: La función escalón se asocia a neuronas binarias en las cuales cuando la suma de las entradas es mayor o igual que el umbral de la neurona, la activación es 1, si es menor, la activación es 0 (ó -1), ver Figura 1.2 (a). Las redes formadas por este tipo de neuronas son fáciles de implementar en hardware, pero sus capacidades están limitadas.

-Neurona de función lineal: La función lineal corresponde a la función $a=n$. En las neuronas con función mixta si la suma de las señales de entrada es menor que un límite inferior, la activación se define como 0 (ó -1). Si dicha suma es mayor o igual que el límite superior, entonces la activación es 1. Si la suma de entrada está comprendida entre ambos límites, la activación se define como una función lineal de suma de las señales de entrada, (Figura 1.2 (b)).

-Neurona de función continua (también conocida como sigmoideal): Con este tipo de función para la mayoría de los valores del estímulo de entrada, el valor dado por la función es cercano a uno de los valores asintóticos. Esto hace posible que en la mayoría de los casos, el valor de salida esté comprendido en la zona alta o baja del sigmoide. De hecho cuando la pendiente es elevada, esta función tiende a la función escalón. La importancia de esta función es que su derivada es siempre positiva y cercana a cero para los valores grandes positivos o negativos; además toma su valor máximo cuando n es cero. Esto hace que se puedan utilizar las reglas de aprendizaje definidas para la función escalón, con la ventaja respecto a esta función, que la derivada está definida para todo el intervalo. La función escalón no podía definir la derivada en el punto de transición y esto no ayuda a los métodos de aprendizaje en los cuales se usan derivadas (Figura 1.2 (c)). Esta función es comúnmente usada en la red de retro propagación 'Backpropagation' (Demuth y Beale, 1998).

1.2.3 Clasificación de las RNA

Las neuronas se relacionan entre sí formando redes que pueden llegar a ser tan complejas como el neocognitron, o tan simples como el Perceptron (Gómez, 1999).

La red neuronal monocapa, es el modelo más simple según se observa en la Figura 1.3, ya que se tiene una capa de neuronas que proyectan las entradas a una capa de neuronas de salida donde se realizan diferentes cálculos. La capa de entrada, por no realizar ningún

cálculo, no se cuenta.

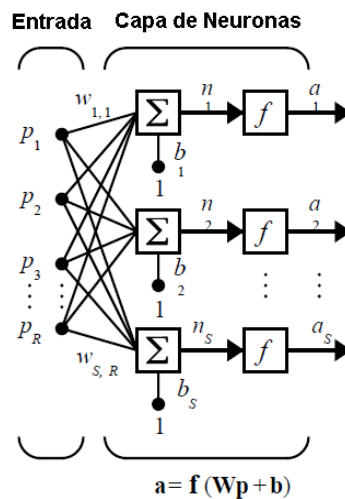


Figura 1.3 Modelo RNA monocapa (Demuth y Beale, 1998).

La RNA de varias capas (multicapas) se pueden visualizar como lo muestra la Figura 1.4. Si existen varios niveles (capas), las redes neuronales multicapas son muy poderosas, por ejemplo, una red de dos capas, donde la primera capa (capa oculta) es sigmoide y la segunda capa (capa de salida) es lineal, esta RNA puede ser entrenada para aproximar cualquier función (con un número finito de discontinuidades). Este tipo de RNA de dos capas se utiliza extensivamente en los casos de Backpropagation (Demuth y Beale, 1998).

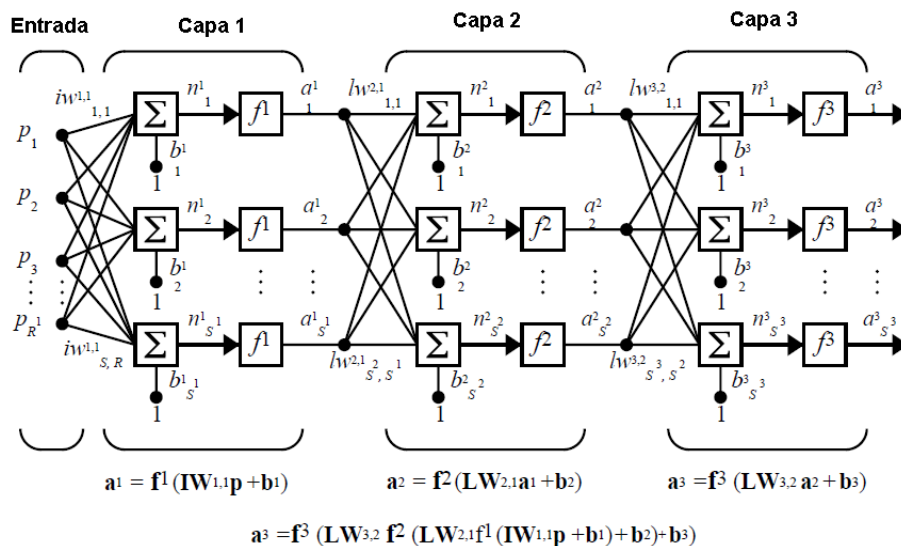


Figura 1.4 Modelo RNA multicapas (Demuth y Beale, 1998).

1.2.4 Métodos de aprendizaje

En una RNA es necesario definir un procedimiento por el cual las conexiones de las neuronas varíen para proporcionar la salida deseada (algoritmo de aprendizaje). Los métodos de aprendizaje se pueden dividir en las siguientes categorías (Rojas, 1995 y Antonio *et al.*, 2009; ver Figura 1.5).

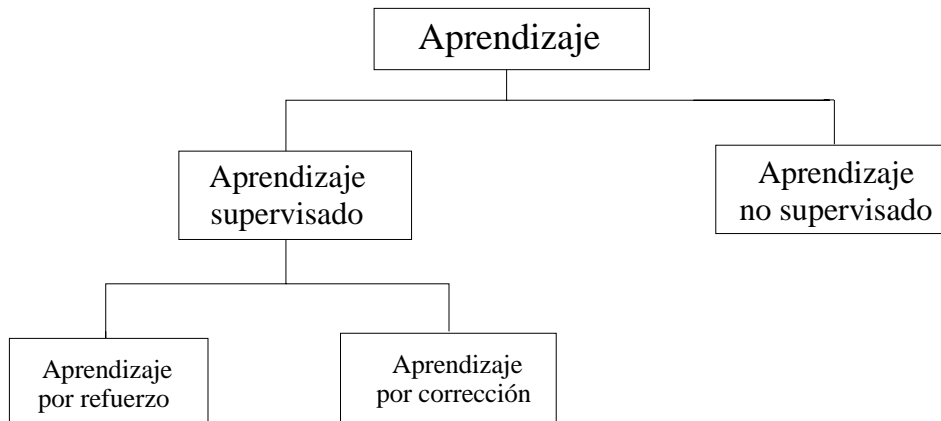


Figura 1.5 Modelos de aprendizaje de la RNA (Antonio *et al.*, 2009).

La primera gran división en los métodos de aprendizaje es entre algoritmos supervisados y no supervisados. En los algoritmos no supervisados no se conoce la señal que debe dar la red (salida deseada). La red en este caso se organiza ella misma agrupando, según sus características, las diferentes señales de entrada.

El aprendizaje supervisado presenta a la red las salidas que debe proporcionar ante las salidas que se presentan. Se observa la salida de la red y se determina la diferencia entre esta y la salida deseada. Posteriormente, los pesos de la red son modificados de acuerdo con el error cometido. Este aprendizaje admite dos variantes: aprendizaje por refuerzo o aprendizaje por corrección. En el aprendizaje por refuerzo solo conocemos si la salida de la red se corresponde o no con la salida deseada, es decir, nuestra información es de tipo booleana (verdadero o falso). En el aprendizaje por corrección conocemos la magnitud del error y esta determina la magnitud en el cambio de los pesos (Antonio *et al.*, 2009).

1.2.5 Modelo de Retro-propagación ‘Backpropagation’

Este es un algoritmo de aprendizaje que es utilizado para entrenar redes de varias capas (Rumelhart *et al.*, 1986). Este algoritmo posee una base matemática bastante sólida y es considerado como una generalización de la regla delta. Esta técnica minimiza el error promedio al cuadrado entre la salida real y la esperada, aplicando el concepto de gradiente descendiente.

El objetivo del algoritmo de Retro-propagación es que los pesos de las capas ocultas generen una representación interna adecuada del problema a resolverse. Estas características y su porcentaje de éxito lo han convertido en uno de los algoritmos de aprendizaje más populares.

1.2.6 Ventajas y desventajas de la RNA

Ventajas

- Una red neuronal puede realizar tareas que con un programa lineal no se puede.
- Cuando un elemento de la red neuronal falla, la simulación puede continuar sin ningún problema debido a la naturaleza paralela en que esta formada la RNA.
- Una red neuronal aprende, así no debería ser reprogramada.
- Se puede implementar sin ningún problema en cualquier aplicación.

Desventajas

- La red neuronal necesita de entrenamiento para operar.
- La arquitectura de una red neuronal es diferente de la arquitectura de microprocesadores por lo tanto, debe ser emulada.
- Se requiere de mucho tiempo de procesamiento para las grandes arquitecturas.
- La falta de reglas definitorias que ayuden a construir una red para un problema dado.
- En el modelado con la red y en algunos casos o aplicaciones, la RNA se presenta como una caja negra ya que no se sabe cual son los variables que forman parte del modelo o cuales se utilizaron finalmente para modelar comparándose con un modelo físico que es totalmente lo contrario.

1.3 Simulación de Monte Carlo

El método de Monte Carlo (MC) ha encontrado aplicaciones en diversas áreas de conocimiento (física estadística, biología, química, etc.) debido a que muchos fenómenos naturales presentan cierto comportamiento aleatorio y a que su naturaleza sugiere una posible aproximación estadística. Se emplean en situación donde los métodos analíticos son poco eficientes o involucran cálculo complicado. Los métodos Monte Carlo, son métodos numéricos, útiles para resolver problemas por medio de simulación de variables aleatorias, deben su nombre a la ruleta Monte Carlo, que es una excelente generadora de números aleatorios. Creado por Jhon Von Neumann, Stanislaw Ulam y Nicolas Metropolis como un posible acercamiento estadístico para resolver el problema de la difusión de neutrones en un material fisionable (Neumann *et al.*, 1947). En general y a través de múltiples ensayos, estos métodos nos permiten aproximar un resultado gracias a la teoría de la probabilidad, en particular a la ley de los grandes números y al teorema del límite central, pero es muy importante tener en cuenta que se trata de una aproximación y por ende, suele usarse en problemas donde no se exige mucha precisión. La característica fundamental de estos métodos, es que se realizan a través de generación de números pseudo-aleatorios. Actualmente, casi todas las computadoras poseen un generador de números pseudo-aleatorios bastante bueno, por lo que no es necesario desarrollarlos.

1.3.1 Propagación de incertidumbres con Monte Carlo

El método de Monte Carlo es uno de los métodos que se utilizan para el análisis de propagación de la incertidumbre, donde el objetivo es determinar cómo la variación aleatoria de los variables, la falta del conocimiento del fenómeno estudiado, o el error afectan a la sensibilidad, el rendimiento o la fiabilidad del sistema que se está modelando (Wittwer, 2004).

La simulación de Monte Carlo se clasifica como un método de muestreo, debido a que las variables de entrada se generan aleatoriamente a partir de muestras de una población real, para formar distribuciones de probabilidad y simular el proceso. Es decir, tratamos de escoger una distribución de las variables que más se acerque al comportamiento de los datos que ya tenemos, o que mejor represente el estado actual del conocimiento sobre el

fenómeno estudiado. Los datos generados a partir de la simulación (resultados) se pueden representar como distribuciones de probabilidad (histogramas) o se convierten en barras de error, predicciones de fiabilidad, zonas de tolerancia, o intervalos de confianza (Figura 1.6).

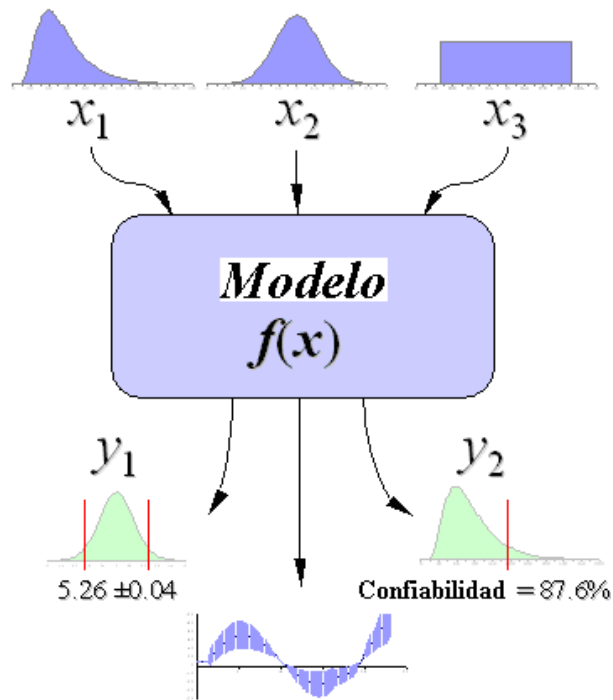


Figura 1.6 Propagación de incertidumbres mediante Monte Carlo de un modelo estocástico (Wittwer, 2004).

Los pasos de simulación de MC correspondiente a la propagación de la incertidumbre mostrada en la Figura 1.6 se pueden implementar fácilmente en un programa para la modelación, y son básicamente cinco:

Paso 1: Crear un modelo matemático con variables, $y = f(x_1, x_2, \dots, x_q)$.

Paso 2: Generar un conjunto de datos aleatorios, $x_{i1}, x_{i2}, \dots, x_{iq}$.

Paso 3: Evaluar el modelo y guardar los resultados como vector de resultados, y_i .

Paso 4: Repetir los pasos 2 y 3 para $i = 1$ hasta n .

Paso 5: Analizar los resultados mediante histogramas, intervalos de confianza, media representativa y desviación estándar, etc.

1.3.2 Números aleatorios

Los números aleatorios deben tener igual probabilidad de salir elegidos y no debe existir correlación serial. Se generan por tablas, o por dispositivos especiales: ruleta (Rand, 1955). En la práctica la generación de los números pseudo-aleatorios que sustituyen a los números aleatorios generalmente se realiza con códigos de computación fiables, como las dadas por Abramowitz y Stegun (1972) y Press *et al.* (1990) donde se debe asegurar la existencia de secuencias largas y densas. Existen varios métodos para generar los números pseudo-aleatorios entre los cuales de encuentran:

- *El método de centros cuadrados:* p.ej., $44^2 = 1936 \Rightarrow 93$
- *Los métodos congruenciales:* $X_n = (aX_{n-1} + c) \bmod m$

donde: X_0 es la semilla, (a) el multiplicador, (c) la constante aditiva y (m) el módulo

- *El método de la transformación inversa:* $x = F^{-1}(x)$ siendo $F(x) = \text{Prob}(X \leq x)$

1.3.3 Función de distribución (Gaussiana)

La transformación de los números aleatorio con distribución uniforme a una distribución Gaussiana se consigue utilizando el método de Box-Müller que emplea funciones trigonométricas (Press *et al.*, 1990).

La función de la distribución Gaussiana es:

$$F(x) = \int_{-\infty}^x dx' p(x) = \frac{1}{2} + \text{erf}\left(\frac{x - \bar{x}}{\sigma}\right) \quad (1.1)$$

donde: $\text{erf}(x)$ es la función de error (Abramowitz y Stegun, 1972), σ es la varianza del parámetro x y \bar{x} es el promedio de los parámetros (valor esperado).

La función de densidad de probabilidad de la distribución normal se expresa como lo siguiente:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\bar{x}}{\sigma}\right)^2\right) \quad (1.2)$$

1.3.4 Software utilizado

En este proyecto de investigación doctoral (Capítulo 4), la generación de las distribuciones de probabilidad de los variables se basa en un generador de números aleatorios llamado (randtool) que proporciona la plataforma Matlab (Martin, 2006), la ventaja de este software es que el utilizador puede seleccionar el tipo de distribución que se requiere para los variables, especificar el número de las muestras de las distribuciones y definir los límites de confianza que se requiere para su modelo.

Capítulo 2

Desarrollo de un nuevo método para estimación de las temperaturas estabilizadas de formación en sistemas geotérmicos usando redes neuronales artificiales

2.1 Introducción

La explotación de los recursos geotérmicos para la producción de electricidad requiere de perforaciones de pozos profundos en las regiones térmicas adecuadas de los campos geotérmicos (Saito *et al.*, 1998; Davis y Michaelides, 2009). La perforación de pozos geotérmicos es un proceso complejo debido a una constante anomalía térmica causada por la utilización de lodos de perforación. Esta anomalía afecta la estimación de la temperatura estabilizada de formación en campos geotérmicos (Fomin *et al.*, 2003, 2004). Una vez que la perforación de un pozo se ha completado, la recuperación térmica se evalúa mediante el análisis de las temperaturas de fondo del pozo (TFP) y de los tiempos de recuperación térmica registrados durante el proceso de perforación (Santoyo *et al.*, 2000). Las TFP son medidas en diferentes tiempos durante las operaciones de perforación. (Espinosa-Paredes y Espinosa-Martínez, 2009), este proceso de medición es costoso debido a los sofisticados equipos de medición utilizados y sobre todo porque se requiere el paro total del proceso de perforación en cada medición (Wisian *et al.*, 1998; Fomin *et al.*, 2005).

La determinación de las temperaturas estabilizadas de formación (TEF) a partir de las mediciones de la temperatura de fondo del pozo (TFP) es una tarea crucial para la evaluación de los sistemas geotérmicos (Espinosa *et al.*, 2001; Kutasov y Eppelbaum, 2003). La estimación de las TEF a partir de registros de TFP en tiempos cortos de perforación nos permite conocer la temperatura virgen de formación meses antes que pueda ser medida con precisión. Este enfoque es muy importante en términos de los programas de exploración, desarrollo y explotación de los sistemas geotérmicos. El conocimiento de las TEF es fundamental para: (1) la determinación de los gradientes geotérmicos para la asignación de exploración; (2) la determinación de los flujos de calor

de las zonas geotérmicas; (3) la interpretación de los registros del pozo; (4) el diseño óptimo de los programas de perforación y terminación; (5) la localización de las zonas permeables o de pérdida por circulación; (6) la evaluación de las propiedades termo-físicas de formación; y (7) la estimación de las reservas de calor en sistemas geotérmicos.

La estimación de las TEF en pozos geotérmicos se realiza mediante el uso de métodos analíticos simplificados, basándose en modelos complejos de transferencia de calor en función a las condiciones térmicas de fondo del pozo. Las mediciones de las temperaturas de fondo del pozo (TFP) tienden a reflejar las anomalías térmicas causadas por la circulación de fluidos de perforación en el pozo.

Los métodos analíticos más utilizados para la estimación de la temperatura estabilizada de formación (TEF) son: (i) el método de Manetti (MM: Manetti, 1973), quien propone que la fuente de calor es cilíndrica y el flujo de calor convectivo; (ii) el método de línea fuente conocido como el método de Horner (HM: Dowdle y Cobb, 1975); (iii) el método de Leblanc (LM: Leblanc *et al.*, 1981), donde la fuente de calor es cilíndrica y el flujo de calor conductivo; (iv) el método de Brennan (BM: Brennan, 1984), quien propone una fuente de calor radial con flujo de calor conductivo; (v) el método de Hasan y Kabir (1994), donde la fuente de calor es cilíndrica y el flujo de calor es conductivo-convectivo (CHSM); (vi) el método de Ascencio *et al.* (1994) quienes proponer que la fuente de calor sea en forma de esfera con un flujo de calor radial (SRM); (vii) el método de Horner generalizado o el método de Kutasov-Eppelbaum (KEM: Kutasov y Eppelbaum, 2005).

La mayoría de estos métodos analíticos requieren por lo menos tres o más mediciones de TFP en función del tiempo y de la profundidad del pozo. La TEF se determina mediante una solución matemática-analítica de cada método, utilizando las mediciones de TFP y el tiempo de recuperación térmica (Verma *et al.*, 2006a,b).

A pesar de los últimos avances logrados en esta área de investigación se han reportado grandes discrepancias entre los resultados de TEF utilizando los diferentes métodos analíticos encontrados en literatura (Andaverde *et al.*, 2005). Por lo tanto el desarrollo de nuevos métodos confiables para estimar la temperatura estabilizada de formación (TEF) sigue siendo un gran desafío. La industria geotérmica aun está interesada en la estimación de la TEF en los sistemas geotérmicos donde el fenómeno de convección domina. La

transferencia de calor por convección dentro de los sistemas geotérmicos se debe básicamente a la presencia de zonas permeables (donde se observa grandes pérdidas de lodo durante el proceso de perforación) en los yacimientos. La determinación de la TEF en estas condiciones requiere el uso de modelos analíticos complejos o simuladores de pozos. Sin embargo, estas herramientas requieren de mayor información como: las propiedades termo-físicas, de transporte y de cementación durante la perforación, así como también el conocimiento exacto de los tiempos de circulación del lodo generalmente es desconocido o difícil de determinar bajo condiciones de perforación (Hermanrud *et al.*, 1990).

Debido a este escenario complejo, la industria geotérmica requiere de herramientas nuevas y prácticas que utilicen los datos disponibles (series de TFP y tiempo de recuperación térmica) para estimar la TEF con exactitud y precisión aceptables. En este sentido, una nueva aplicación de redes neuronales artificiales (RNA) fue desarrollada en este proyecto doctoral como nuevo método confiable para el cálculo de la temperatura estabilizada de formación en pozos geotérmicos.

La RNA ha sido aplicada como una poderosa herramienta computacional en la solución de diversos problemas complejos en las diferentes áreas de ciencia, sobre todo para la predicción (Zhang *et al.*, 1998). El uso de redes neuronales artificiales ha sido perceptible en ciencias de la tierra (p. ej., Goutorbe *et al.*, 2006; Hsieh *et al.*, 2009; Leite y Filho, 2009a,b; Morton, 2009). En los últimos años, la aplicación de esta nueva herramienta en problemas de ingeniería geotérmica y petrolera ha sido objeto de estudio (ej., Farshad *et al.*, 2000; Bayram, 2001; Can, 2002; Yilmaz *et al.*, 2002; Spichak y Goidina, 2005, Spichak, 2006; Díaz-González *et al.*, 2008; Serpen *et al.*, 2009).

La metodología y validación de un nuevo método basado en las redes neuronales artificiales para la estimación de la TEF en pozos geotérmicos se describen con mas detalle a continuación.

2.2 Metodología

El nuevo método requiere como datos de entrenamiento: las mediciones de temperatura de fondo (TFP), los tiempos de recuperación térmica, los gradientes de temperatura transitoria

y las estimaciones de TEF normalizadas inferidas por siete métodos analíticos. El modelo de la RNA fue entrenada con una base de datos de pozos geotérmicos que contiene estimaciones de TEF estadísticamente normalizadas. Esta base de datos está formada con registros de recuperación térmica obtenidos durante la perforación de diferentes pozos geotérmicos en el mundo. La recuperación térmica contempla: mediciones de temperaturas de fondo (TFP), registros de tiempo medidos, gradientes de temperatura transitoria y estimaciones de TEF calculadas por siete métodos analíticos comúnmente usados en la industria geotérmica.

La normalización de las estimaciones de TEF se llevó a cabo utilizando las pruebas estadísticas para la detección y rechazo de valores discordantes (Verma, 2005). El valor promedio de la TEF normalizada, así como su desviación estándar se calcularon con la metodología estadística mejorada y desarrollada por Verma *et al.* (2009). La RNA junto con este enfoque estadístico permite obtener las estimaciones más confiables de las TEF. El nuevo modelo neuronal predictivo se evaluó a través de cuatro series (pruebas) de TFP con sus registros de tiempo, dos series sintéticas medidas en laboratorio y dos series reales medidas en pozos donde las temperaturas verdaderas de formación (TVF) son conocidas. Estas pruebas de validación se consideraron en la evaluación de precisión del método.

2.2.1 Creación de base de datos para el desarrollo del método

La base de datos de recuperación geotérmica fue realizada con valores de TFP con sus respectivos tiempos registrados (Δt) durante la perforación de diez pozos geotérmicos. Esta serie de datos se obtuvieron de informes de perforación en diferentes campos geotérmicos del mundo (ver Tabla 2.1). Los diez pozos fueron los siguientes: (1) Los Humeros, México (MXCO1., Andaverde *et al.*, 2005); (2) Los Humeros, México (MXCO2., Andaverde *et al.*, 2005); (3) Oklahoma, EUA (SGIL., Schoeppel y Gilarranz, 1966); (4) Mississippi wellbore, EUA (USAM., Kutasov, 1999); (5) Leyte, Filipinas (PHIL., Andaverde *et al.*, 2005); (6) Kyushu, Japon (JAPN., Hyodo y Takasugi, 1995); (7) Chipilapa, El Salvador (CH-A1., González-Partida *et al.*, 1997); (8) Chipilapa, El Salvador (CH-A2., Iglesias *et al.*, 1995); (9) Larderello, Italia (ITAL., Da-Xin, 1986); y (10) Norton Sound field, Alaska (COST., Cao *et al.*, 1988a).

Tabla 2.1 Base de datos utilizada para el entrenamiento y validación de la RNA. TFP: temperatura de fondo de pozo; t: tiempo de recuperación térmica; tc: tiempo de circulación; m: número de mediciones.

Datos de pozos para el entrenamiento de la RNA						Datos sintéticos para la validación						
Serie	t (h)	TFP (°C)	Serie	t (h)	TFP (°C)	Serie	t (h)	TFP (°C)	Serie	t (h)	TFP (°C)	
MXCO1 tc = 2.5 h m = 6	6	186.3	PHIL tc = 15 h m = 14	2.58	93	ITAL tc = 5 h m = 40	5.7	84.5	SHBE tc = 5 h m = 8	2.5	56.6	
	12	206.9		3.58	88		6.0	93.1		5	61.3	
	18	219.0		4.58	99		6.5	95.0		7.5	64.3	
	24	231.9		5.58	108		7.0	96.2		10	66.6	
	30	239.8		6.58	112		7.5	97.5		15	69.6	
	36	247.7		7.58	117		8.0	98.4		20	71.7	
MXCO2 tc = 2.5 h m = 7	6	178.6		8.58	120		8.5	99.3		30	74.1	
	12	198.9		9.58	126		9.0	100.3		40	75.5	
	18	211.4		10.58	133		9.5	101.1				
	24	225.8		11.58	133		10.0	102.0				
	30	235.1		12.58	134		10.5	102.7				
	36	240.4		13.58	137		11.0	103.4	CLAH tc = 5 h m = 15	2	91.7	
42	247.1		14.58	141		11.5	104.2	4		98.1		
			15.58	146		12.0	104.7	6		102.4		
						12.5	105.4	8		105.5		
						13.0	106.1	10		107.9		
						13.5	106.7	12		109.8		
SGIL tc = 3 hr m = 12	0	62.78	JAPN tc = 5 hr m = 10	5.5	92.0		14.0	107.6		14	111.3	
	1	81.94		6.5	98.5		14.5	108.3		16	112.5	
	2	86.56		7.5	103.0		15.0	108.9		18	113.6	
	3	88.84		8.5	107.4		15.5	109.6		20	114.4	
	4	90.76		9.5	110.4		16.0	110.2		22	115.2	
	5	92.29		12.5	119.6		16.5	110.8		24	115.8	
	6	93.44		15.5	126.6		17.0	111.3		30	117.1	
	7	94.21		18.5	132.8		17.5	111.9		40	118.4	
	8	94.59		24.5	142.4		18.0	112.4		50	119.1	
	9	94.98		72.5	170.9		18.5	112.9				
	10	95.36					19.0	113.3		CH-A4 tc = 5 h m = 5	6	85
	11	95.74		CH-A1 tc = 5 hr m = 6	6	90		19.5	113.8			12
12	96.13	26.57	116			20.0	114.2		26.57	122		
USAM tc = 5 h m = 14	1	138.75		47.1	130		20.5	114.7		47.1	139	
	2	139.55		95.27	145		21.0	115.1		95.27	152	
	5	140.89	CH-A2 tc = 5 hr m = 6	6.0	231		21.5	115.4				
	10	142.03		12.0	231		22.0	115.7		KJ-21 tc = 2.5 h m = 6	2	50
	15	142.72		26.6	235		22.5	116.1			30	125
	20	143.21		47.1	237		23.0	116.5			52	158
	25	143.60		95.3	239		24.0	117.1			96	184
	30	143.92		190.5	240		25.0	117.7			242	218
	40	144.42				26.0	117.9		556		225	
	50	144.81	COST tc = 4 hr m = 6	8.5	45.00		27.0	118.7				
	70	145.40		10.0	49.44							
	100	146.04		16.5	55.56							
150	146.76	34.5		56.11								
200	147.27	38.0		57.22								
		75.5		56.11								

Por otra parte, cuatro series de datos fueron recopilados para las pruebas de precisión y validación del modelo: (i) datos seleccionados de dos trabajos experimentales reportados en literatura (SHBE., Shen y Beck, 1986; CLAH., Cao *et al.*, 1988b); y (ii) datos reales de dos pozos geotérmicos donde se conoce la TEF (CH-A4., Iglesias *et al.*, 1995; KJ-21.,

Steingrímsson y Guðmundsson, 2006). Estos datos fueron utilizados para la evaluación de la predicción del modelo porque las TEF son conocidas (SHBE., TEF=80.0°C; CLAH., TEF=120.0°C; CH-A4., TEF=169°C y KJ-2., TEF=240°C).

2.2.2 Estimación inicial de las Temperaturas Estabilizadas de Formación normalizadas

Los siete métodos analíticos utilizados para el cálculo inicial de las TEF fueron los siguientes : (i) el método de Horner (HM), (ii) el método Brennan (BM), (iii) el método Kutasov-Eppelbaum (KEM), (iv) el método de Leblanc (LM), (v) el método de Manetti (MM), (vi) el método esférico-radial (SRM) y (vii) el método conductivo-convectivo de la fuente cilíndrica (CSM). Las soluciones analíticas obtenidos de los siete métodos anteriores se resumen en la Tabla 2.2.

Las TEF se estimaron utilizando la regresión lineal ordinaria (RLO), la regresión lineal ponderada (RLP) y la regresión cuadrática (RC); ver Andaverde *et al.*, 2005. La teoría de la regresión lineal ordinaria (OLR) establece la existencia de errores homocedásticos en el eje de la ordenada, en este caso para la variable TFP, mientras que los valores en el eje de las abscisas están libres de error (las funciones de tiempo en este caso; ver Tabla 2.2). Por otra parte, la regresión lineal ponderada (WLR) considera la existencia de errores heteroscedásticos en ambos ejes de la curva (s_x y s_y) en la estimación del intercepto (a) y la pendiente (b); Verma (2005).

De acuerdo con el procedimiento del análisis numérico utilizado para cada método, TEF se infiere de los valores de intercepto (a) de las regresiones OLR, WLR y QR (ej., Santoyo *et al.*, 2000; Andaverde *et al.*, 2005).

Tabla 2.2 Soluciones numéricas simplificadas de diferentes métodos analíticos (HM, BM, KEM, LM, MM, SRM y CSM) utilizados para la determinación de las TEF en pozos geotérmicos.

El método analítico	La Solución Numérica
BM	$TFP(t) = TEF_{BM} - b_{BM} \left(\frac{1}{\Delta t + pt_c} \right)$
CSM	$TFP = TEF_{CSM} + (b_{CSM}) \cdot [F(t_{cD} + \Delta t_D) - F(t_D)]$
HM	$TFP(t) = TEF_{HM} + (b_{HM}) \cdot \ln\{(t_c + \Delta t) / \Delta t\}$
KEM	$TFP(t) = TEF_{KEM} + b_{KEM} \ln(X)$
LM	$TFP(t) = TEF_{LM} - b_{LM} \left[1 - \exp\left(-\frac{r_w^2}{4\alpha\Delta t}\right) \right]$
MM	$TFP(t) \approx TEF_{MM} + b_{MM} \ln\left(\frac{\Delta t}{\Delta t - t_c}\right)$
SRM	$TFP(t) = TEF_{SRM} + (b_{SRM}) \cdot \left(\frac{1}{\sqrt{\Delta t}}\right)$

TFP: Temperatura del fondo del pozo; Δt : tiempo de registro; TEF son inferidas a partir del valor del intercepto para cada método analítico: TEF_{BM} , TEF_{CSM} , TEF_{HM} , TEF_{KEM} , TEF_{LM} , TEF_{MM} , y TEF_{SRM} . Para mas detalles de las variables primarias y secundarias ver las fuentes originales de los métodos (indicados en la sección 2.2.2).

2.2.3 Procedimiento general para la estimación de las TEF con RNA

El procedimiento utilizado para la estimación de TEF utilizando RNA se muestra en la Figura 2.1. Cinco secciones fundamentales se incluyeron en el cálculo de la temperatura estabilizada de formación: (i) lectura de los datos de entrada (10 series, con 120 datos de TFP y sus tiempos de registro); (ii) selección de siete métodos analíticos (cálculo de funciones de tiempo);(iii) estimación de TEF utilizando los siete métodos analíticos con tres modelos de regresión RLO, RLP y RC; (iv) la aplicación de pruebas estadísticas de discordancia para calcular la media de las TEF y sus desviaciones estándares; y (v) el cálculo de los gradientes de temperatura usando los datos de la recuperación térmica de los pozos.

Las TEF estimadas de los diez pozos (MXCO1, MXCO2, SGIL, USAM, PHIL, JAPN, CH-A1, CH-A2, ITAL, y COST) se calcularon de forma individual a partir de sus TF y sus tiempos de registro. Las funciones de tiempo de cada método analítico (HM, BM, KEM, LM, MM, SRM, y CSM) fueron calculadas por medio de sus ecuaciones correspondientes (Tabla 2.2). Se obtuvieron tres estimaciones de TEF como resultado de los tres modelos de regresión antes mencionados (RLO, RLP y RC). Finalmente, se obtuvieron 21 valores calculador de TEF.

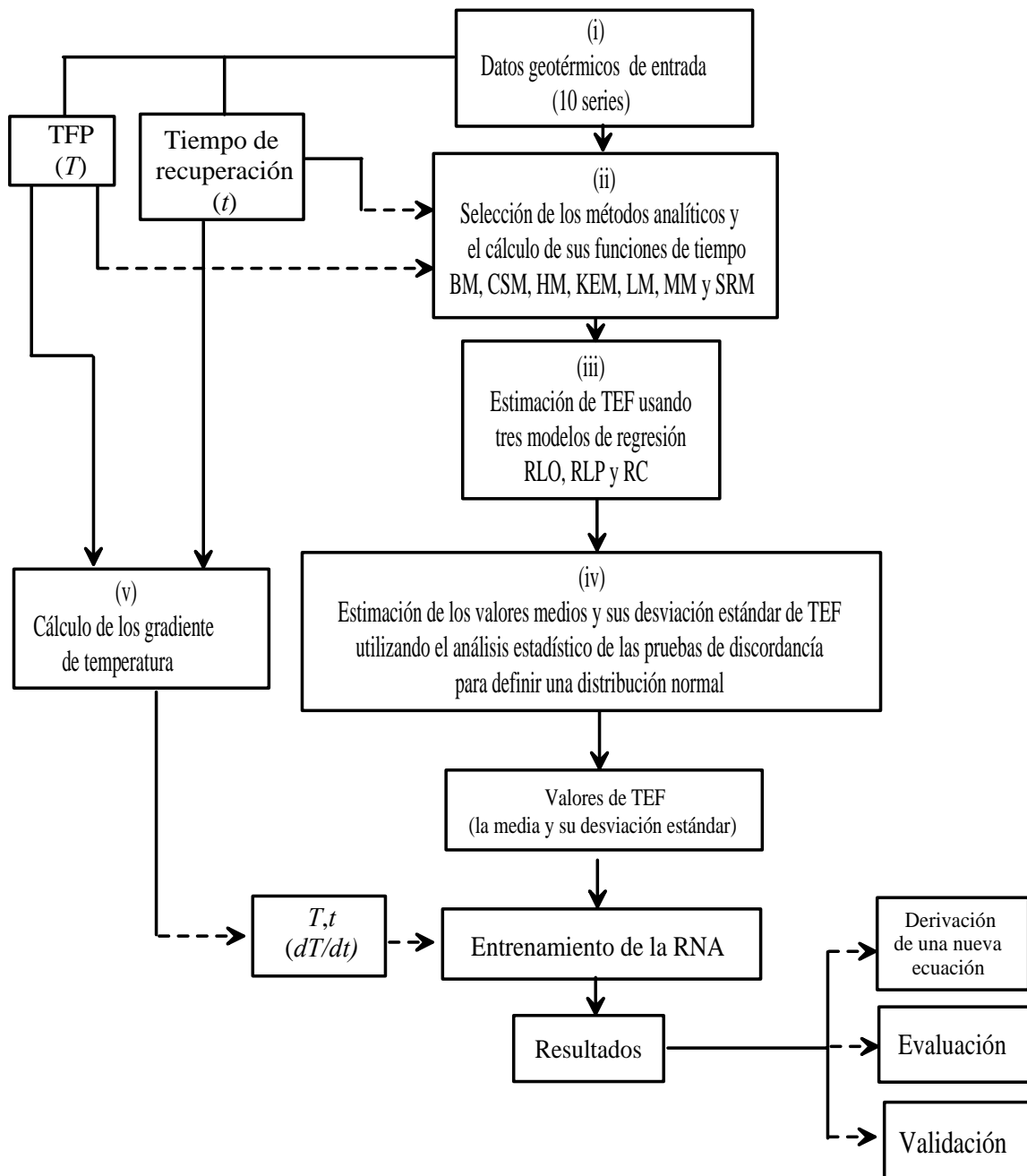


Figura 2.1 Esquema general de la metodología desarrollada.

En el objetivo de obtener estimaciones confiables, los valores de las 21 TEF antes calculadas fueron normalizados estadísticamente utilizando el software DODESYS (Verma *et al.*, 2009). Nueve pruebas de discordancia (N1, N2, N4, N7, N8, N9, N10, N14 y N15 propuestos por varios autores: ver Verma, 2005) fueron aplicadas a las 21 TEF para identificar y rechazar los valores extremos discordantes así como para calcular la media de TEF (\bar{x}_{TEF}) y su desviación estándar asociada (s_{TEF}).

El método de normalización estadística permite detectar y rechazar con eficacia los errores resultantes de los métodos analíticos para seleccionar las mejores estimaciones de TEF. La Tabla 2.3 resume los resultados obtenidos después de utilizar el procedimiento de normalización estadística.

Tabla 2.3 Valores de TEF estimadas con los siete métodos analíticos utilizados. Los datos fueron normalizados con el software DODESYS utilizando tres diferentes modelos de regresión (RLO, RLP, y RC).

Serie de Datos	TEF	TEF
	(Métodos Analíticos)	(Predicción de la RNA)
	$\bar{x}_{TEF} \pm s_{TEF}$	$\bar{x}_{RNA} \pm s_{RNA}$
MXCO1	263 ± 15	256 ± 20
MXCO2	262 ± 18	253 ± 15
USAM	146 ± 1	139 ± 5
ITAL	132 ± 10	133 ± 8
PHIL	197 ± 28	185 ± 25
JAPN	176 ± 18	180 ± 11
CH-A1	147 ± 5	151 ± 7
CH-A2	241 ± 18	243 ± 15
COST	57 ± 3	56 ± 2
SGIL	100 ± 3	102 ± 9

Los valores medios de las TEF y sus desviaciones estándares de cada serie, después de haber sido tratados con DODESYS, se utilizaron para el entrenamiento de la RNA. Tres valores de las TEF fueron empleados para reflejar la incertidumbre de esta variable (el

valor medio: \bar{x}_{TEF} ; la media más la desviación estándar: $\bar{x}_{TEF} + s_{TEF}$; y el valor de la media menos su desviación estándar: $\bar{x}_{TEF} - s_{TEF}$): ver Tabla 2.3.

Los gradientes transitorios de las temperaturas fueron calculados utilizando la siguiente ecuación:

$$\frac{dT}{dt} = \frac{(TFP_{i+1} - TFP_i)}{t_{i+1} - t_i} \quad (2.1)$$

donde TFP es la temperatura del fondo del pozo registrada durante un tiempo (t), y el subíndice (i) proporciona los valores transitorios de los variables.

El gradiente transitorio de la temperatura (dT/dt) es un parámetro importante en el entrenamiento de la RNA. Cuando el gradiente tiende a cero en un tiempo infinito ($t \rightarrow \infty$) significa que la temperatura de formación geotérmica se aproximó a un equilibrio térmico y por lo tanto:

$$\lim_{t \rightarrow \infty} \frac{dT}{dt} = 0 \quad \therefore \quad (TFP \approx TEF) \quad (2.2)$$

2.2.4 Arquitectura propuesta de la RNA

En el presente trabajo, se formó una RNA de tres capas: (i) capa de entrada con tres variables los cuales son: Las temperaturas del fondo de pozo, los tiempos de registro (t) y el gradiente transitorio (dT/dt); (ii) capa oculta contiene neuronas de conexión; y (iii) capa de salida formada por una variable de salida (en este caso son las estimaciones de las TEF). Ver Figura 2.2.

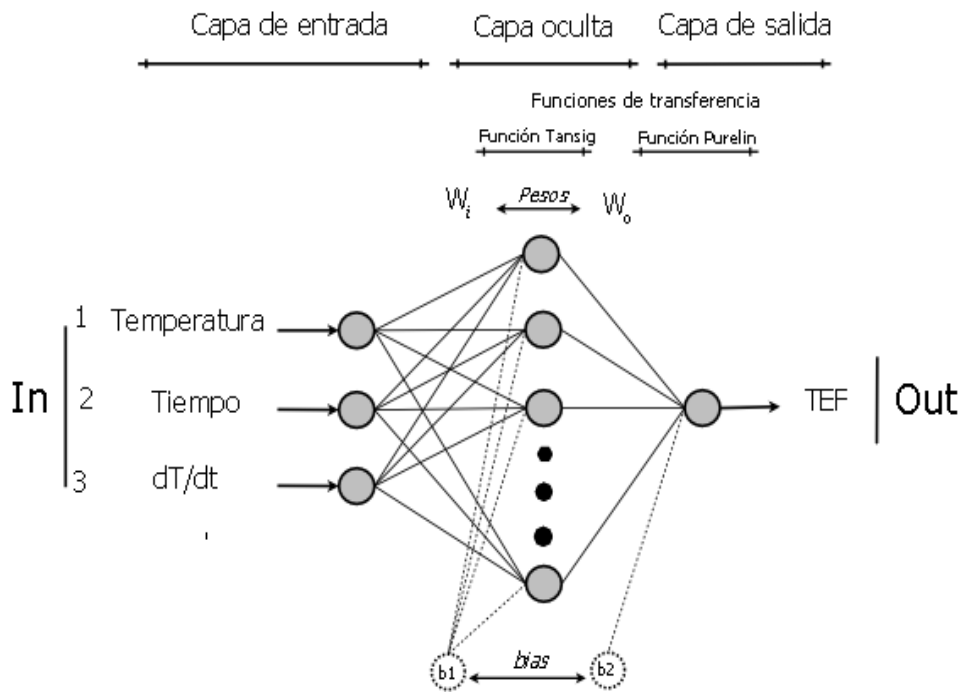


Figura 2.2 Arquitectura general de la RNA.

Normalmente el número óptimo de neuronas en la capa oculta (S) es difícil de precisar y por lo general depende de la complejidad del problema. Este número se determina durante el proceso del entrenamiento y la evaluación de la RNA. Cada neurona en la capa oculta tiene un (bias), $b_1(s)$, que se añade a la suma del producto (los pesos de cada neurona multiplicados por sus variables de entrada debidamente normalizadas).

Los coeficientes asociados por cada capa se agrupan en una matriz de pesos (W) y un vector de los biases (b) correspondientes a las neuronas que forman esta capa, así se distinguen las matrices W_i , W_o y los vector $b_1(s)$, $b_2(l)$ para la capa oculta y la capa de salida respectivamente. Utilizando la notación de matrices antes mencionada, la salida de la RNA (Out) puede ser expresada en general por la siguiente ecuación:

$$Out = g(W_o \cdot f(W_i \cdot In + b_1) + b_2) \quad (2.3)$$

donde f y g son las funciones de transferencia de cada capa, usando las siguientes funciones: tangente hiperbólica en la capa oculta ($f = \text{tansig}$) y lineal en la capa de salida ($g = \text{purelin}$), la misma Ecuación (2.3) puede ser escrita de esta forma:

$$Out_{(l)} = \sum_{s=1}^S \left[W_o(l,s) \cdot \left(\frac{2}{1 + \exp\left(-2 \cdot \left(\sum_{k=1}^K (W_i(s,k) \cdot In(k)) + b_l(s) \right)\right)} - 1 \right) \right] + b_{2(l)} \quad (2.4)$$

donde K es el número de variables de entrada, S es el número de neuronas en la capa oculta y l es el número de las variables de salida en la arquitectura de la RNA.

2.3 Modelo de la RNA

Un algoritmo de aprendizaje (también conocido como proceso de entrenamiento de la red) se define como un proceso que consiste en ajustar los coeficientes (pesos y biases) de una arquitectura para minimizar la función de error entre las salidas estimadas (calculadas por la red) y las salidas deseadas (ya conocidas). El gradiente de la función de error puede ser calculado por el modelo clásico de retro-propagación (Rumelhart *et al.*, 1986). Recientemente varios algoritmos de aprendizaje han utilizado este gradiente para la optimización de la máxima pendiente, pero los resultados han demostrado que los métodos de optimización de segundo orden son mucho más eficientes (Demuth y Beale, 1998).

En el modelo propuesto se utilizó el algoritmo de Levenberg-Marquardt (LM) para la optimización de la función de error. Este algoritmo está descrito en el libro de “*Matlab Neural Network Toolbox*”, (Demuth y Beale, 1998) y es básicamente una aproximación del método de Newton (para más detalles, ver el párrafo 3 del **Anexo A**).

La estimación del error se realizó con el estadístico *RMSE* (acrónimo del inglés *Root-Mean-Square-Error*; Ecuación 2.5) utilizando los datos experimentales, $Exp(i)$ y simulados por la RNA, $Sim(i)$. Este cálculo sirvió como criterio de optimización para la adecuación del modelo (Ver Figura 2.3).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Exp(i) - Sim(i))^2} \quad (2.5)$$

Durante el entrenamiento de la RNA 60% de los datos experimentales fueron utilizados para el aprendizaje y 40% de los datos, fueron utilizados para la evaluación. El objetivo de esta división fue obtener una muestra estadísticamente representativa del problema.

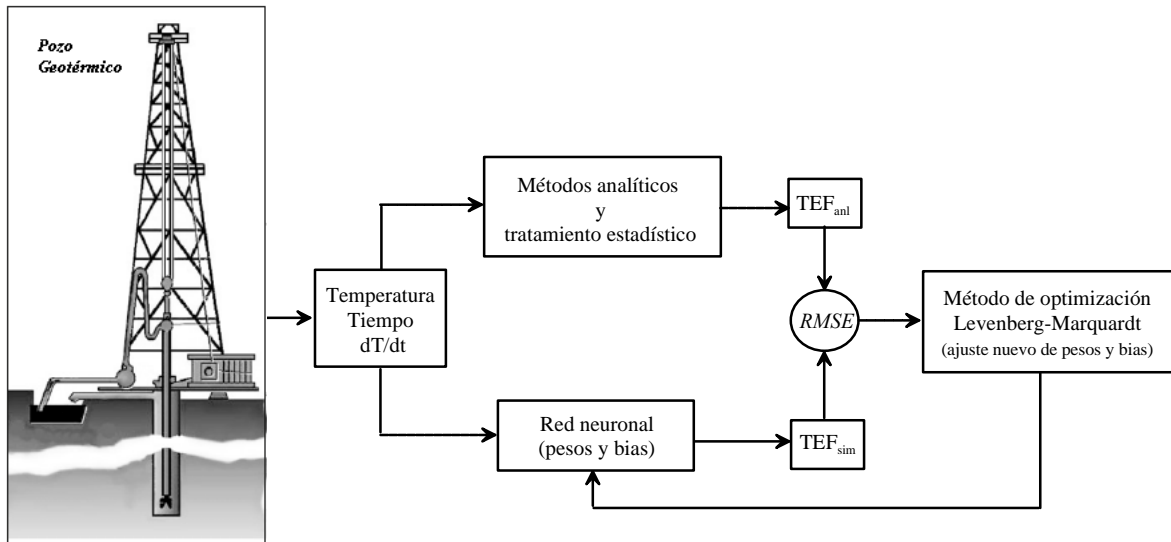


Figura 2.3 Algoritmo utilizado para el proceso de la optimización y aprendizaje de RNA.

2.3.1 Entrenamiento computacional del modelo

En el proceso de aprendizaje, el algoritmo del modelo neuronal (RNA) fue evaluado examinando los valores de la $RMSE$ en cada iteración, y aumentando el número de las neuronas en la capa oculta. Los resultados obtenidos demostraron que el error típico de aprendizaje disminuye cuando el número de neuronas en la capa oculta aumenta. Sin embargo, uno de los problemas que pueden presentarse durante el entrenamiento de redes neuronales se conoce como “sobre aprendizaje” (ej., Hernández-Pérez *et al.*, 2004).

La comparación de los valores calculados de $RMSE$ durante el aprendizaje y la evaluación de la RNA, resultó ser un buen criterio para optimizar el número de iteraciones y evitar el problema de sobre aprendizaje. Acorde a los resultados de la $RSME$, el número óptimo de neuronas en la capa oculta se fijó en cinco neuronas. Para este proceso, los valores de $RMSE$ fueron muy pequeños y casi invariables cuando se aumentó a seis el número de las neuronas en la capa oculta.

El nuevo modelo contiene 20 pesos (15 elementos para la matriz W_i y 5 para la matriz de salida W_o), y 6 bias en total ($b_1=5$ y $b_2=1$; ver Tabla 1.4). Estos coeficientes se utilizaron para calcular los valores de las TEF (Ecuación 2.6), y validar el modelo con los datos no incluidos en el entrenamiento.

$$TEF = \sum_{j=1}^S \left[W_o(1, j) \cdot \left(\frac{2}{1 + \exp\left(-2 \cdot \left(\sum_{k=1}^K (W_i(j, k) \cdot In(k)) + b_1(j)\right)\right)} - 1 \right) \right] + b_2 \quad (2.6)$$

Los valores predictivos de las TEF calculados por la RNA junto con sus incertidumbres están reportados en la Tabla 2.3.

Se encontró un buen acercamiento entre las estimaciones de TEF deducidas con los métodos analíticos y los datos simulados TEF_{RNA} obtenidos de la RNA (Figura 2.4). Se pudo observar también que los datos simulados (TEF_{RNA}) presentan la mejor relación lineal con respecto a las estimaciones deducidas con los métodos analíticos ($y = 8,5631+0.94376x$; $R^2 = 0,95$).

Tabla 2.4 Parámetros óptimos de la RNA.

Coeficientes de la RNA		$S=5$ y $K=3$				
W_i		3.9757	1.4220	-2.9996		
		5.0716	13.9674	-2.3583		
		-6.9071	-2.7251	-34.0318		
		-6.3881	2.3519	-6.0095		
		-5.2657	3.6209	-4.5029		
W_o	-1.4449	0.1627	-0.3257	0.4320	-0.5981	
b_1	-5.6937	-10.096	1.6998	-0.5554	-2.6091	
b_2			-1.4001			

K es el número de variables de entrada, S es el número de neuronas en la capa oculta.

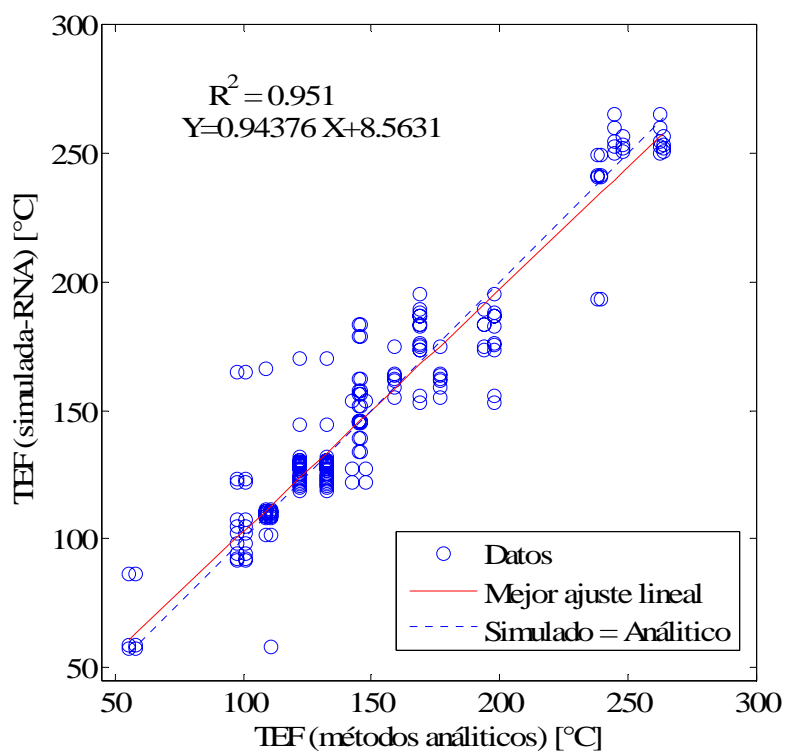


Figura 2.4 Comparación estadística entre las TEF estimadas por los métodos analíticos y las simuladas por RNA.

2.3.2 Validación numérica del modelo

El nuevo modelo (Ecuación 2.6) fue validado realizando cuatro pruebas (series de datos): dos de ellas sintéticas (SHBE: TVF = 80 °C, y CLAH: TVF = 120 °C), y dos series de mediciones reales en pozo geotérmico conociendo las temperaturas verdaderas de formación (CH-A4: TVF = 169 °C, y KJ-21: TVF = 240 °C; ver Tabla 1.1). Los resultados de precisión de estas cuatro pruebas mostraron un buen ajuste entre la TEF medida (también llamado TVF) y la calculada por el modelo (Figura 2.5).

Para las series sintéticas (Prueba 1 y 2) la RNA predijo valores de TEF entre 74.5 °C y 83.4 °C, con un valor medio (TEF_{RNA}) de 79.1 °C para la serie de datos SHBE, y un intervalo desde 117.8 °C hasta 125.6 °C, con un valor medio de 122.1 °C para la serie CLAH. Mientras que para las dos series de datos reales (Prueba 3 y 4): CH-A4 y KJ-21, el modelo predijo valores de TEF entre 167.6 °C y 173.5 °C, con un valor promedio de 171.2

° C, y valores de TEF desde 239.1 °C a 247.1 °C, con un valor medio de 243.3 °C respectivamente.

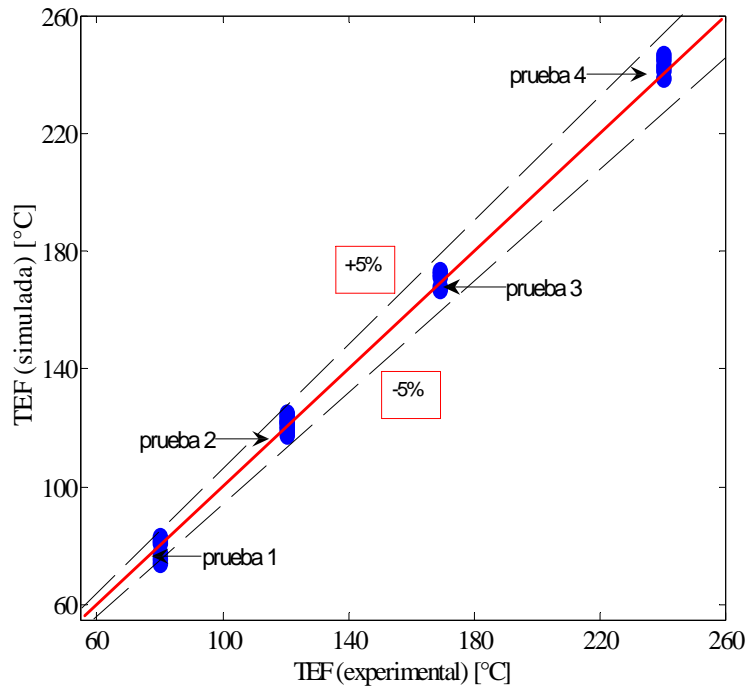


Figura 2.5 Evaluación de la precisión de la RNA utilizando una comparación estadística entre datos de TEF simulados por la red y las TEF verdaderas.

Con estos resultados queda demostrado que el nuevo modelo de la RNA (presentado por la Ecuación 1.6) es confiable y proporciona estimaciones aceptables de la TEF con errores típicos de $\pm 5\%$ los cuales son aceptables en las aplicaciones de herramientas de inteligencia artificial en problemas de ingeniería (ej., Leite y Filho, 2009a).

Finalmente, los pequeños residuos calculados en el proceso de validación y el coeficiente de regresión obtenido ($R^2 = 0.95$) permitieron confirmar la aproximación de los datos de TEF medidos y estimados por la RNA (ver Figura 2.5). Más aún el modelo desarrollado en esta tesis es capaz de predecir valores de TEF durante la perforación de pozos geotérmicos siempre y cuando los datos experimentales se encuentren entre el rango de 45 °C y 263 °C.

Capítulo 3

Determinación de gradientes de presión en pozos geotérmicos usando redes neuronales artificiales y un simulador numérico de pozos: Estudio comparativo

3.1 Introducción

La evaluación de la producción de pozos geotérmicos es una tarea de ingeniería crucial para la explotación óptima de sus recursos (p. ej., Khasani *et al.*, 2002; García-Valladares *et al.*, 2006; Al Jahanandish *et al.*, 2011). Esta tarea de ingeniería es necesaria para varias aplicaciones geotérmicas, tales como: (i) el diseño óptimo de la geometría del pozo y los equipos de superficie (Gunn *et al.*, 1992.); (ii) la identificación de las zonas permeables en el reservorio geotérmico (Grant *et al.*, 1983); (iii) la determinación del punto de flasheo (García-Valladares *et al.*, 2006); (iv) la predicción de las curvas de productividad para el análisis del problema de disminución de la producción (Iglesias y Moya, 1998); (v) la identificación temprana de las problemas de producción al interior de los pozos productivos (Grant *et al.*, 1983); (vi) la detección temprana de problemas relacionadas con el efecto de ‘Skin’ (Aragón *et al.*, 2008); (vii) la identificación de zonas secundarias de alimentación en pozos de producción (Bjornsson y Bodvarsson, 1987); (viii) la explotación sostenible de los recursos geotérmicos para la protección de la inversión de proyectos comerciales (Axelsson, 2010); entre otros.

El flujo de fluidos y el calor dentro de los pozos geotérmicos juegan un papel importante para una mejor optimización de sus recursos (p. ej., Hadgu y Freeston, 1990; Hadgu *et al.*, 1995; Garg *et al.*, 2004; Hasan y Kabir, 2010). Los datos de pozos de producción se obtiene generalmente a partir de mediciones estables que se llevan a cabo en el campo (García-Valladares *et al.*, 2006). En este contexto, el conocimiento preciso de la temperatura y la presión es de vital importancia para determinar las tareas de ingeniería de yacimientos anteriormente descritas (Gunn *et al.*, 1992). Los gradientes de temperatura y

de presión se mide generalmente mediante técnicas de registro complejas y costosas (Jung *et al.*, 2001).

La modelación numérica es una metodología bien conocida que se utiliza para predecir la mayoría de estos parámetros termodinámicos. Simuladores numéricos de pozos son ampliamente reconocidos como las herramientas más eficaces, prácticas, y económicas para el análisis de los procesos de fluidos y de flujo de calor, en función de los gradientes de presión y de temperatura.

Varios simuladores de pozo, basados en modelos de fluidos monofásico y bifásico, se han desarrollado para predecir estos gradientes (p. ej., Barelli *et al.*, 1982; Ortíz Ramírez, 1983; Hadgu y Freeston, 1990; Gunn *et al.*, 1992; Hadgu *et al.*, 1995; García *et al.*, 1999; García-Valladares, 2007; Hasan y Kabir, 2007, 2010, entre otros). Estas herramientas de simulación numérica tiene la ventaja de minimizar el número de mediciones de campo en pozos geotérmicos (Gunn *et al.*, 1992). Sin embargo, es importante reconocer que el uso de simuladores numéricos todavía presenta algunas limitaciones debido a las diferencias que comúnmente se encuentran entre los datos simulados y medidos (Takacs, 2001; Lu *et al.*, 2006; Álvarez del Castillo, *et al.*, 2010). Estas diferencias son causadas por varios efectos, tales como:

- (1) el uso de ecuaciones empíricas no adecuadas para la descripción de las condiciones que en realidad se encuentran en el fluido y los flujos de calor dentro de los pozos geotérmicos (Dalkilic *et al.*, 2009),
- (2) el uso de hipótesis que no describen los fenómenos físicos reales que ocurren dentro del sistema geotérmico (Gunn *et al.*, 1992),
- (3) el número limitado de mediciones de campo y sus errores experimentales asociados (Jung *et al.*, 2001),
- (4) las propiedades termodinámicas, de transporte de fluidos geotérmicos y de los materiales de pozo (p. ej., la formación, cimentación de la tubería, y fluidos de perforación) (Tian y Finger, 2000), y
- (5) el efecto de las temperaturas estabilizadas de formación (o gradientes geotérmicos) sobre los perfiles de presión en los pozos geotérmicos productores, especialmente para el análisis de los datos de producción del pozo bajo condiciones de flujo no adiabáticas (Grant *et al.*, 1983); entre otros.

Por otro lado, las técnicas de cálculo basados en herramientas de inteligencia artificial (p. ej., RNA, la lógica difusa, entre otros) han comenzado a aplicarse en diferentes problemas de ingeniería relacionados con estudios de explotación de petróleo y de los sistemas geotérmicos (p. ej., Farshad et al., 2000; Zellou y Ouenes, 2003; Spichak y Goidina, 2005; Jeirani y Mohebbi, 2006; Bassam *et al.*, 2010; Álvarez del Castillo *et al.*, 2011; Jahanandish *et al.*, 2011).

Farshad *et al.*, (2000) diseñaron un novedoso enfoque de redes neuronales artificiales (RNA) para la predicción de los perfiles de temperatura del flujo dentro de pozos petroleros. Zellou y Ouenes (2003) desarrollaron una nueva metodología que combina la RNA y la lógica difusa para la caracterización de las fracturas dentro de los reservorios. Jeirani y Mohebbi (2006) utilizaron un modelo entrenado de la RNA para estimaciones de la presión inicial del yacimiento, la permeabilidad de la formación, y el factor de daño. Bassam *et al.*, (2010) desarrollaron una nueva ecuación basada en RNA para la predicción de temperaturas estabilizadas de formación (TEF) en pozos geotérmicos. Álvarez del Castillo *et al.*, (2011) desarrollaron una nueva correlación empírica para calcular la fracción volumétrica de vapor mediante el método de RNA para el análisis de los flujos bifásicos en pozos geotérmicos. Jahanandish *et al.*, (2011) reportó un modelo de RNA para la predicción de la presión de fondo y las caídas de presión en los pozos multifásicos.

En cuanto a esta revisión, se encontró que las redes neuronales se han aplicado principalmente para la evaluación de algunas características de pozos petroleros (Farshad *et al.*, 2000; Jeirani y Mohebbi, 2006; y Jahanandish *et al.*, 2011). Sin embargo, hoy en día, la aplicación de esta técnica en los estudios geotérmicos ha sido limitada, por lo que representa un área de interés trascendente en el que se podrían realizar contribuciones científicas novedosas.

En este capítulo, el objetivo es presentar un nuevo enfoque de las técnicas de RNA para una determinación confiable de los gradientes de presión en pozos geotérmicos, con geometrías inclinadas y verticales. La ventaja de este nuevo estudio es demostrar la habilidad de los modelos RNA utilizando un número limitado de variables de entrada, que son comúnmente medidas en el campo. Con este propósito, en este estudio se presenta una metodología computacional en términos de diferentes arquitecturas desarrolladas de RNA, de algoritmos numéricos, de procedimientos de evaluación y de entrenamiento, y de

análisis estadísticos realizados para la validación de los resultados previstos. Adicionalmente, se llevó a cabo un análisis de sensibilidad de las mejores arquitecturas de RNA con el objeto de identificar las variables de entrada más importantes. Los gradientes de presión obtenidos de los modelos RNA fueron comparados estadísticamente con datos reales medidos en el campo y con los resultados numéricos calculados con el simulador numérico de pozos GEOWELLS. A continuación se describe a detalle los resultados numéricos de esta investigación.

3.2 Marco teórico

3.2.1 Gradientes de presión

Hoy en día, se han desarrollado numerosos trabajos de investigación y simuladores numéricos de pozos para la predicción de los gradientes de presión en pozos geotérmicos (p. ej., Freeston y Hadgu, 1988; Hasan y Kabir, 2010). La mayoría de estos simuladores determinan los gradientes de presión, a través de la solución de las ecuaciones de balance de cantidad de movimiento, tomando en cuenta un conjunto de condiciones iniciales y de frontera en el contorno del pozo (García-Valladares *et al.*, 2006). De acuerdo a Currie (1974), la ecuación general de cantidad de movimiento está dada por la siguiente ecuación:

$$\rho \left[\frac{\partial \vec{v}}{\partial t} + (\vec{v} \cdot \nabla) \vec{v} \right] = -\nabla P + \nabla \cdot \tilde{\tau} + \rho \vec{f} \quad (3.1)$$

Una simplificación detallada de esta ecuación se presenta en el **Apéndice A**. Generalmente el gradiente global de la presión en los pozos geotérmicos se estima bajo las condiciones unidimensionales de acuerdo con la siguiente ecuación:

$$-\left(\frac{dP}{dz} \right)_T = \left[\rho_m f \frac{V^2}{2D} \right]_f + \left[\rho_m V \frac{dV}{dz} \right]_a + [\rho_m g \sin \theta]_g \quad (3.2)$$

Los tres términos del lado derecha en la ecuación (3.2) representan las caídas de presión debido a las fuerzas de fricción, aceleración y gravitacional (Wallis, 1969). Las variables

de esta ecuación, incluyendo sus unidades de medición y el factor de fricción de Darcy (f) se describen adicionalmente en los **Apéndices A y B**.

Para obtener una predicción confiable de los gradientes de presión, la solución de la Ecuación (3.2) requiere de un conocimiento correcto de algunos parámetros cruciales del flujo bifásico (p. ej., fracción volumétrica de vapor, y las propiedades termodinámicas y de transporte del fluido). Comúnmente, cuando se tiene flujo bifásico, las dos fases (líquida y vapor), fluyen con diferentes velocidades, en este caso, la fase de vapor (que tiene una densidad menor), fluye más rápido que la fase líquida.

En este contexto el proceso de flujo de fluido se define como la retención de líquido por una fracción, mientras que la diferencia en la velocidad de las fases se conoce como deslizamiento (Hasan y Kabir, 2007). El conocimiento de estos parámetros de flujo bifásico es muy difícil de obtener en las condiciones geotérmicas, lo que ha llevado a buscar otros métodos prácticos u analíticos para predecir de forma confiable los gradientes de presión dentro del pozo (p. ej., Garg *et al.*, 2004; Álvarez del Castillo, *et al.*, 2011).

3.2.2 Simulador de pozos geotérmicos (GEOWELLS)

García-Valladares *et al.* (2006) desarrollaron el simulador geotérmico GEOWELLS con las siguientes características numéricas: (a) modela flujos monofásicos y bifásicos; (b) la simulación de los pozos inicia de la parte inferior hacia la superficie o cabezal del pozo con una malla de profundidad ajustable. Sin embargo, una simulación inversa podría ser fácilmente realizada usando un método iterativo con condiciones de frontera adecuadas; (c) considera cambios múltiples en la geometría del pozo (carcasa, diámetros, y la rugosidad), incluyendo pozos desviados; (d) estima las pérdidas de calor entre el pozo y la formación; y (e) despliega en pantalla los datos simulados de salida, como son: presión, temperatura, entalpía, calidad del vapor, flujo másico, velocidades de las fases, flujo de calor, entre otros; así como los gradientes de fricción, la aceleración y la gravitación de presión) a cada profundidad del pozo.

La obtención de estos datos de salida de producción, se llevan a cabo considerando las siguientes suposiciones físicas y termodinámicas:

- (i) modelación unidimensional de los fluidos para determinar los gradientes de presión y temperatura [es decir, $T(z, t)$, $p(z, t)$, etc.];

- (ii) el flujo bifásico puede ser analizado mediante modelos de flujo homogéneo o cualquier modelo de flujo derivado;
- (iii) las propiedades termodinámicas de los fluidos geotérmicos son iguales a las propiedades del agua;
- (iv) se desprecian tanto el intercambio de calor por radiación como la conducción de calor axial en el interior del pozo; y
- (v) la conducción de calor a través de los materiales del pozo (tubería, cemento y la formación de roca) se calculan mediante la resolución axial tridimensional de una distribución simétrica de la temperatura.

Las ecuaciones gobernantes (balances de masa, momento y energía) que se utilizan para describir el fluido y los flujos de calor dentro de un volumen de control (CV) de un pozo geotérmico, están descritos en detalle por García-Valladares *et al.*, (2006). En el presente estudio, los gradientes de presión del flujo en la zona bifásica (Ecuación 3.2) se calcularon utilizando el modelo clásico de flujo homogéneo, que ha sido ampliamente empleado en trabajos anteriores (p. ej., Hasan y Kabir, 2007). El uso de modelo de flujo homogéneo se justifica sobre todo en condiciones de flujo burbuja, en donde las velocidades de las fases son similares, y se reduce la complejidad del problema (Hasan y Kabir, 2007; Timlin, 2008).

La densidad de la mezcla bifásica (ρ_m) se estimó mediante la siguiente ecuación:

$$\rho_m = \rho_l(1 - \alpha_H) + \rho_s \alpha_H \quad (3.3)$$

donde, ρ_l y ρ_s son las densidades del líquido y de vapor saturado, respectivamente, y α_H es la fracción volumétrica de vapor, en el caso del modelo homogéneo se determina en función de la calidad del vapor (x) y de las densidades de las fases de acuerdo a la siguiente ecuación:

$$\alpha_H = \frac{1}{1 + \left(\frac{1-x}{x}\right) \left(\frac{\rho_s}{\rho_l}\right)} \quad (3.4)$$

3.2.3 Redes neuronales artificiales

En este estudio se diseñó una estructura RNA con 3 capas básicamente (capa de entrada, capa oculta y capa de salida) para la estimación de los gradientes de presión dentro de los pozos. En la Figura (3.1) se presenta el algoritmo de entrenamiento de la RNA y los criterios estadísticos de su evaluación. El proceso de aprendizaje se llevó a cabo mediante el ajuste de los pesos de conexión entre las neuronas. Se utilizó el algoritmo de Levenberg-Marquardt que es una aproximación del método de Newton, para la minimización de la función de error entre los valores calculados por la RNA y los datos medidos.

3.3 Metodología Computacional

Se desarrolló una metodología computacional para la determinación de los gradientes de presión dentro de los pozos geotérmicos (ver Figura 3.2). La metodología consistió en las tareas siguientes:

- (i) Creación de una base de datos a partir de registros de pozos geotérmicos productores de diferentes campos geotérmicos del mundo;
- (ii) Desarrollo y evaluación de modelos de RNA (RNA_1 y RNA_2) para la predicción de los gradientes de presión dentro de los pozos geotérmicos;
- (iii) Determinación de los gradientes de presión obtenidos con el simulador GEOWELLS, y los modelos desarrollados en este estudio (RNA_1 y RNA_2), y
- (iv) Un análisis estadístico comparativo entre los gradientes de presión medidos y las estimaciones (deducidas por GEOWELLS y los modelos de la RNA).

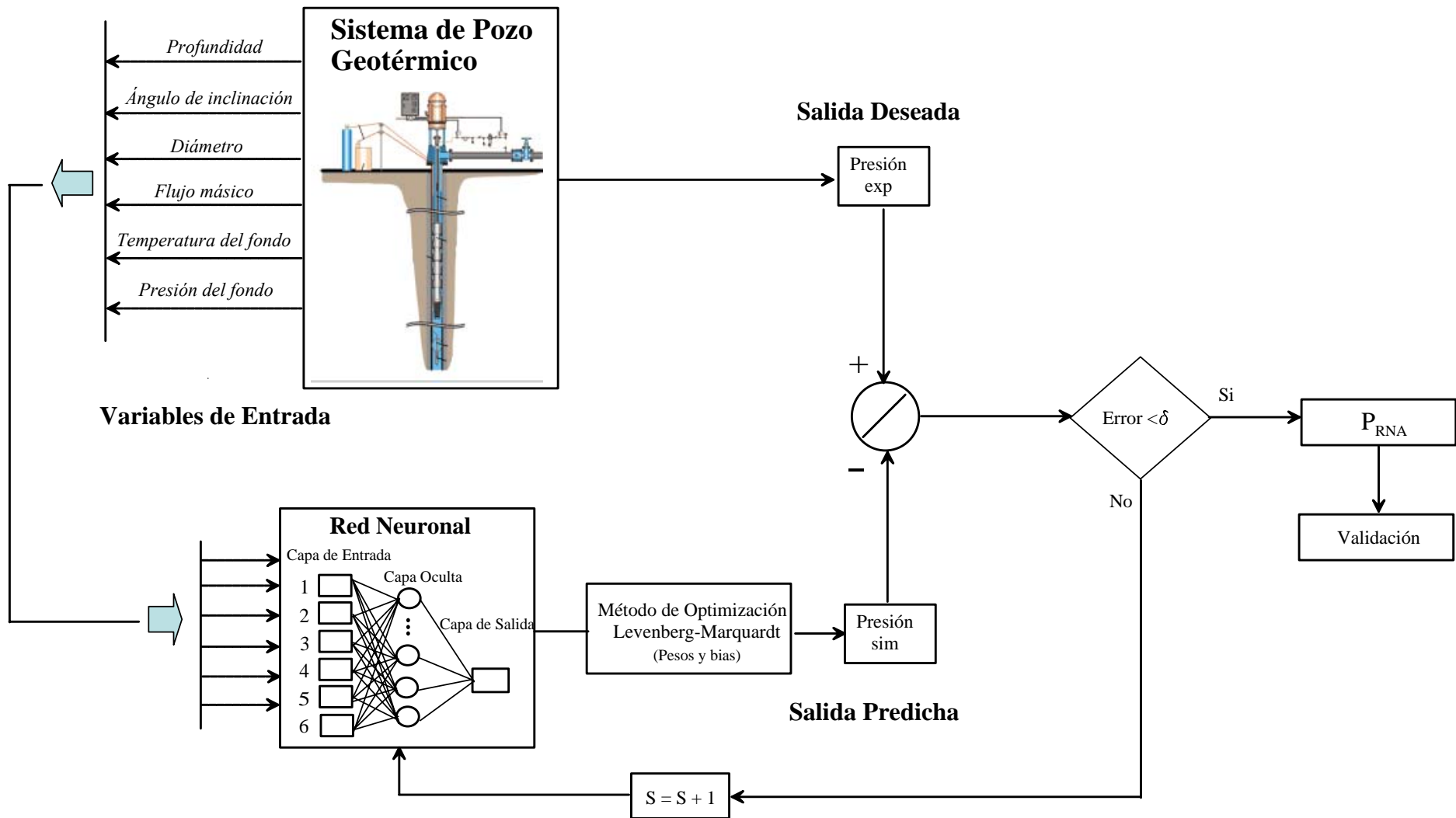


Figura 3.1 Algoritmo de entrenamiento de las RNA.

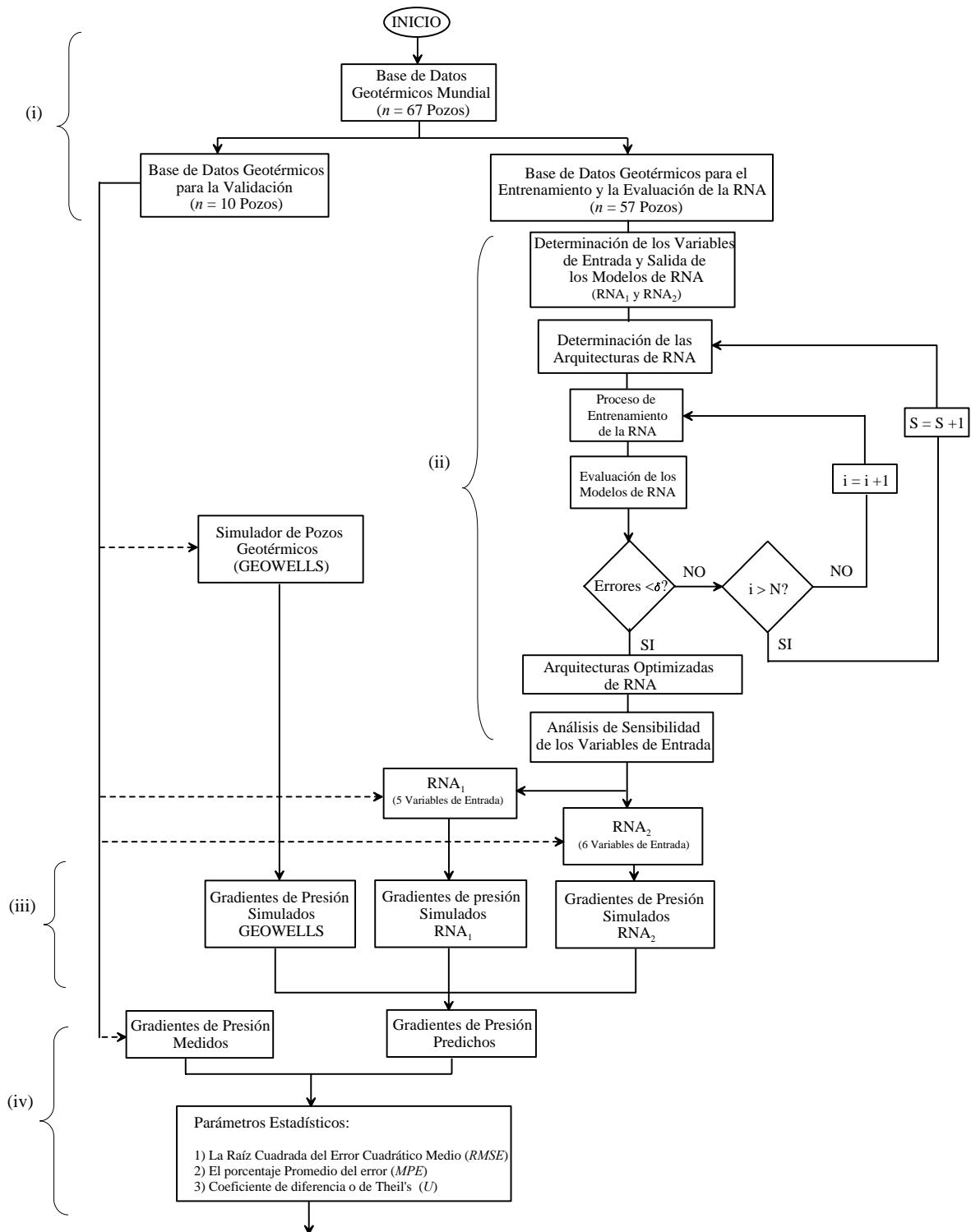


Figura 3.2 Esquema general de la metodología desarrollada.

(i) *Creación de una base de datos con registros de producción de pozos geotérmicos:* Se compiló a partir de una gran variedad de fuentes bibliográficas internacionales (p. ej., revistas científicas, actas de congresos de la energía geotérmica y reportes técnicos), una

base de datos general (BDG) que contiene 67 pozos geotérmicos de diferentes partes del mundo. La (BDG) incluye registros de producción de pozo geotérmicos de diferentes regiones, como de África (Djibouti: Sánchez-Upton, 2000); China (Freeston y Hadgu, 1988); Islandia (Bjornsson, 1987); Italia (Freeston y Hadgu, 1988; Barelli *et al.*, 1982); Japón (Garg y Combs, 2002; Garg *et al.*, 2004); México (Bjornsson, 1987; Aragón *et al.*, 1999; Sánchez-Upton, 2000; Álvarez del Castillo *et al.*, 2010); Nueva Zelanda (Freeston y Hadgu, 1988); y Filipinas (Ambastha y Gudmundsson, 1986). Cada serie de datos de un solo pozo contiene un promedio de 19 registros de presiones a diferentes profundidades. En general la BDG esta formada por una matriz de 1300 filas (registros) y 7 columnas (variables geométricas y termodinámicas de los pozos). A partir de esta estructura, una base de datos de 57 pozos geotérmicos aleatoriamente escogidos de la BDG, se utilizó para el desarrollo de los modelos de la RNA (proceso del entrenamiento), lo datos de los 10 pozos restantes se emplearon para la validación de los mismos modelos. Una distribución estadística de los variables de la BDG utilizados como parámetros de entrada y salida en el desarrollo de los modelos de la RNA se resumen en la Tabla 3.1, mínimos y máximos, la tendencia central (media y mediana) y las dispersiones (desviación estándar ‘s*’) están reportados.

Tabla 3.1 distribución estadística de la BDG utilizada en el proceso de entrenamiento y la evaluación de los modelos de la RNA.

Parámetros	n	mín.	máx.	media	mediana	s*	unidad
<i>Parámetros de entrada:</i>							
Profundidad (z)	1128	0	2600	637.77	600	46.40	[m]
Ángulo de inclinación (θ)	111	0	32.10	5.16	0.00	7.28	[°]
Diámetro (D)	110	0.10	0.38	0.24	0.22	0.07	[m]
Flujo másico (W)	57	1.10	202.00	53.17	34.00	47.32	[kg/s]
Temperatura del fondo (T_f)	57	104.4	364.1	232.1	224.5	39.7	[°C]
Presión del fondo (P_f)	57	6.7	196.3	47.0	29.7	39.5	[bar]
<i>Parámetros de salida:</i>							
Presión (P_{med})	1128	2.0	95.0	30.0	18.5	28.8	[bar]

*Desviación estándar

La geometría del pozo (profundidad, ángulo de inclinación, y el diámetro), y los datos de producción (presión y temperatura del fondo, y flujo másico) se utilizaron como variables de entrada principales. Los registros de presión de cada pozo fueron utilizados como la única variable de salida (variable deseada). Los datos de entrada antes mencionados se utilizaron también para realizar simulaciones numéricas con GEOWELLS. Es muy importante mencionar que todos los variables de entrada y salida involucradas en este estudio tienen una fuerte dependencia a los fenómenos que describen el fluido dentro de los pozos geotérmicos (Ecuación 3.2).

(ii) *Desarrollo y evaluación de los modelos de RNA.* Se desarrollaron dos diferentes modelos de RNA (RNA_1 y RNA_2) para la predicción de gradientes de presión en pozos geotérmicos. El modelo RNA_1 está caracterizado con cinco variables en la capa de entrada, una capa oculta, y una sola variable en la capa de salida, mientras que el modelo RNA_2 considera una capa de entrada con seis variables y una arquitectura en la capa oculta y de salida similar al modelo RNA_1 , usando las funciones ‘tansig’ en las neuronas de la capa oculta y lineal ‘purelin’ en la neurona de la capa de salida (ver Figura 3.3). La evaluación de la red multi-capas con diferentes combinaciones de funciones de activación era una estrategia adecuada de cálculo para encontrar relaciones lineales y no lineales entre las variables de entrada y salida (Spichak y Popova, 2000; Cortés *et al.*, 2009).

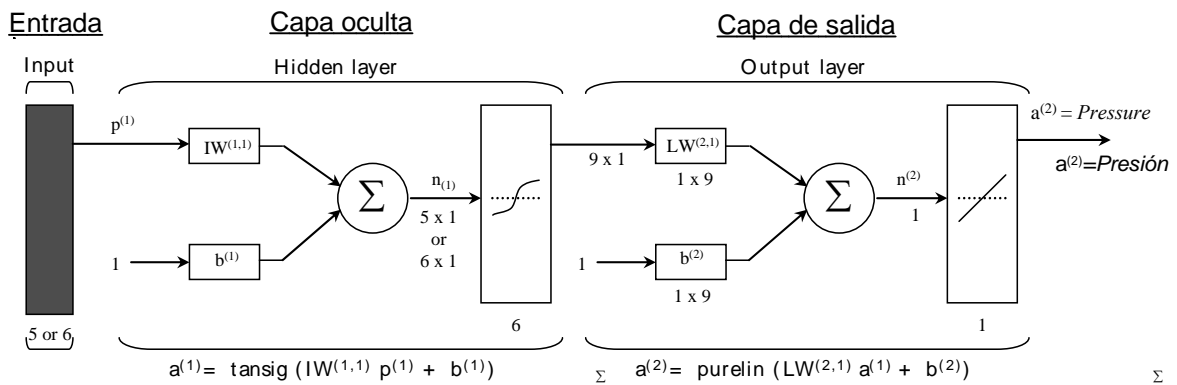


Figura 3.3 Arquitectura de modelos de las RNA.

Las variables de entrada de los modelos neuronales fueron normalizados dentro del rango [0,1] (debido al intervalo en que opera la función ‘tansig’ utilizada en la capa oculta) Demuth y Beale (2003).

(iii) *Determinación de los gradientes de presión con el simulador GEOWELLS y los modelos de RNA.* A partir de los 10 pozos de validación seleccionados de la BDG (ver Figura 2.3), se calcularon los perfiles de presión por el simulador GEOWELLS y los dos modelos desarrollados de la RNA con fines de ser estadísticamente comparados.

(iv) *El análisis estadístico comparativo entre la estimación de gradientes de presión a través de los modelos (RNA₁, RNA₂ y GEOWELLS) y los datos de campo.* Los resultados numéricos obtenidos con el simulador GEOWELLS y los modelos de la RNA fueron estadísticamente comparados con datos de campo medidos. Este análisis se realizó mediante la aplicación de las pruebas estadísticas siguientes: el estadístico *RMSE* (acrónimo del inglés *Root-Mean-Square-Error*), el promedio de los porcentajes de los errores (*MPE*), y el coeficiente de diferencia Theil’s U.

El parámetro *RMSE* es una medida de uso frecuente para la evaluación de las diferencias entre los valores pronosticados por el modelo y los que realmente observados (Verma, 2005). El Parámetro *MEP* que es el promedio de los porcentajes de los errores (%), el conocimiento de este parámetro estadístico ayuda a evaluar si las predicciones subestiman o sobreestiman a los datos reales. El coeficiente de *Theil’sU* se utiliza para medir la eficiencia de un modelo de predicción cuando se compara con otros modelos, por lo tanto, si el coeficiente tiene valores inferiores a 1, esto significa que los errores obtenidos por la predicción del modelo (en evaluación) son inferiores a los errores obtenidos por otros modelos (ver Tabla 3.2). Los detalles sobre el uso de estos parámetros estadísticos se describen a detalle en Álvarez del Castillo *et al.* (2012).

Tabla 3.2 Diferentes pruebas estadísticas utilizadas para la evaluación de los modelos desarrollados.

Parámetros estadísticos	Ecuación
La raíz cuadrada del error cuadrático medio (RMSE)	$RMSE = \sqrt{\frac{\sum_{i=1}^n (P_{sim(i)} - P_{med(i)})^2}{n}}$
Promedio de los porcentajes de los errores (MPE)	$MPE = \frac{1}{n} \sum_{i=1}^n \left(\frac{P_{sim(i)} - P_{med(i)}}{P_{sim(i)}} \right) 100$
Coficiente de diferencia (U)	$U = \frac{\sqrt{\sum_{i=1}^n (P_{sim(i)}^{RNA_2} - P_{med(i)})^2}}{\sqrt{\sum_{i=1}^n (P_{sim(i)}^c - P_{med(i)})^2}}$

Donde P^c es el valor de presión estimado por GEOWELLS o RNA₁.

3.4 Resultados y discusión

3.4.1 Evaluación de los modelos desarrollados

La realización del proceso de entrenamiento y la validación de ambos modelos RNA₁ y RNA₂ (ver Figura 3.1), se llevo a cabo considerando el 70% de los datos de producción para entrenamiento y el 30% de los datos para validación, esto con el objetivo de obtener una distribución representativa aceptable de los datos y evitar sesgo (Rumelhart, 1986). La base de datos del entrenamiento se utilizó para calcular los pesos y bias óptimos de los modelos, mientras que la base de datos de validación sirvió para evaluar la eficiencia de estos modelos. Evidentemente la capacidad de predicción de los modelos de RNA es optimizada aumentando el número de iteraciones para así reducir los errores al mínimo durante el proceso de aprendizaje y evaluación.

En el proceso de aprendizaje de los modelos, la selección de las funciones de activación $a^{(1)}$ y $a^{(2)}$ para la estructura de las capa oculta y de salida de los modelos de RNA (ver Figura 3.2), se lleva a cabo mediante una evaluación preliminar de varias combinaciones de funciones de transferencia (es decir, lineal-lineal, lineal-tansig, tansig-lineal, y tansig-

tansig). Como resultado de esta evaluación, se ha encontrado que la combinación de las funciones tansig-lineal era la más conveniente como estructura para la RNA. Además en ambos modelos (RNA₁ y RNA₂) se encontró que los mejores resultados del entrenamiento fueron obtenidos con estructuras de 9 neuronas en la capa oculta, (ver Figura 3.3). El objetivo de utilizar el mismo número de las neuronas en la capa oculta (en ambas arquitecturas RNA₁ y RNA₂) es por evaluar el efecto de la variable de entrada adicional (en este caso, la presión de fondo considerada por la RNA₂) en la predicción de los gradientes de presión. La ecuación general que presenta la salida de los modelos obtenidos RNA₁ y RNA₂ en función de las entradas es la siguiente:

$$a^{(2)} = \sum_m \left[LW^{ho}_{(n,m)} \left(\frac{2}{1 + \exp \left[-2 \left(\sum_k IW^{ih}_{(k,m)} \cdot In_{(k)} + b_{1(m)} \right) \right]} - 1 \right) \right] + b_{2(n)} \quad (3.5)$$

En donde el número total de las neuronas en la capa de entrada para ambos modelos RNA₁ y RNA₂ fueron $k=5$ y $k=6$, respectivamente; mientras que el número de neuronas en la capa oculta y la capa de salida fueron $m=9$ y $n=1$, respectivamente. Los términos $IW^{ih}_{(k,m)}$, $LW^{ho}_{(n,m)}$ corresponden a los pesos; y los términos $b_{1(m)}$ y $b_{2(n)}$ representan las bias de cada neurona (ver Tabla 3.3).

De acuerdo a la metodología descrita, en la Figura 3.1, los coeficientes de la Ecuación (3.5) se calcularon después de $\sim 10,000$ iteraciones, lo que permitió a la variable de salida simulada (P_{sim}) aproximarse a los valores deseados (en este caso los valores de registros de presión medidos durante la producción) con errores mínimos de $RMSE$ (definida en la Tabla 3.2). Cuando los valores del estadístico $RMSE$ son menores a los criterios de convergencia (δ), la variable de salida simulada (P_{sim}) es definida como la variable de salida optimizada por la RNA (P_{RNA}).

La capacidad de predicción de los modelos RNA₁ y RNA₂ fue evaluada individualmente por un análisis de regresión lineal ($y=a+bx$) entre los datos simulados de presión (P_{RNA}) y medidos (P_{med}) (utilizando el coeficiente de determinación lineal: R^2 ; intercepto: a; la pendiente: b, y sus incertidumbres correspondientes). Los resultados obtenidos de los modelos RNA₁ y RNA₂ se muestran gráficamente en las Figuras 3.4 y 3.5,

respectivamente. En el modelo RNA₁, la mejor ecuación de regresión lineal fue dada por los parámetros estadísticos siguientes: $a = 0.72 \pm 0.20$, $b = 0.972 \pm 0.004$ y una $R^2 = 0,9708$; mientras que para RNA₂, fueron $a = 0.39 \pm 0.13$ y $b = 0.988 \pm 0.003$ con una $R^2 = 0,9874$. De acuerdo a estos resultados, el modelo RNA₂ presentó una mejor predicción de los gradientes de presión que el modelo RNA₁. Claramente esto se observa cuando las variables P_{med} y P_{RNA} son comparadas a través de la regresión lineal, donde lo óptimo es cuando se obtiene $y=x$. Si los datos entre P_{med} y P_{RNA} son similares entonces éstos tienden a concentrarse en los límites de la línea ideal y los parámetros de regresión se aproximan a $b = 1$ y $a = 0$.

Tabla 3.3 Parámetros óptimos resultado del entrenamiento de los modelos RNA₁ y RNA₂.

RNA₁ ($k = 5, m = 9$ y $n = 1$) ‘ k ’, ‘ m ’ y ‘ n ’ son los números de neuronas en las capas de entrada, oculta y de salida.									
$IW^{ih}_{(k,m)}$	$IW^{ih}_{(k,1)}$	$IW^{ih}_{(k,2)}$	$IW^{ih}_{(k,3)}$	$IW^{ih}_{(k,4)}$	$IW^{ih}_{(k,5)}$	$IW^{ih}_{(k,6)}$	$IW^{ih}_{(k,7)}$	$IW^{ih}_{(k,8)}$	$IW^{ih}_{(k,9)}$
	0.3204	2.4256	4.3627	0.4890	-2.2497	9.6849	0.7743	-0.432	-4.7039
	1.3233	-14.285	-18.744	-1.201	-3.946	7.9212	1.5533	-3.320	121.9431
	3.3179	4.2232	2.8880	-2.4452	47.7616	70.9132	-15.055	25.447	58.4813
	-2.826	9.9721	12.4938	1.9567	-60.235	-22.673	7.2476	-60.95	-117.711
	9.0871	1.3565	-1.2250	-4.4872	-1.3651	60.0984	6.9697	-22.67	-62.7870
$LW^{ho}_{(n,m)}$	$LW^{ho}_{(n,1)}$	$LW^{ho}_{(n,2)}$	$LW^{ho}_{(n,3)}$	$LW^{ho}_{(n,4)}$	$LW^{ho}_{(n,5)}$	$LW^{ho}_{(n,6)}$	$LW^{ho}_{(n,7)}$	$LW^{ho}_{(n,8)}$	$LW^{ho}_{(n,9)}$
	63.1218	-43.241	61.0326	147.7915	-22.059	27.0020	44.4230	19.1116	20.2681
$\mathbf{b}_{1(m)}$	$\mathbf{b}_{1(1)}$	$\mathbf{b}_{1(2)}$	$\mathbf{b}_{1(3)}$	$\mathbf{b}_{1(4)}$	$\mathbf{b}_{1(5)}$	$\mathbf{b}_{1(6)}$	$\mathbf{b}_{1(7)}$	$\mathbf{b}_{1(8)}$	$\mathbf{b}_{1(9)}$
	-8.4671	-7.3651	-5.3948	5.5466	-13.665	-59.516	-1.6910	11.1369	44.5506
$\mathbf{b}_{2(n)}$	$\mathbf{b}_{2(1)}$								
	-52.051								
RNA₂ ($k = 6, m = 9$ y $n = 1$)									
$IW^{ih}_{(k,m)}$	$IW^{ih}_{(k,1)}$	$IW^{ih}_{(k,2)}$	$IW^{ih}_{(k,3)}$	$IW^{ih}_{(k,4)}$	$IW^{ih}_{(k,5)}$	$IW^{ih}_{(k,6)}$	$IW^{ih}_{(k,7)}$	$IW^{ih}_{(k,8)}$	$IW^{ih}_{(k,9)}$
	1.1540	-2.9835	-0.6190	1.9215	-2.1278	-1.0858	-2.0430	2.2851	-2.8870
	9.7532	7.4337	53.3913	6.5991	-3.2155	2.0629	0.1668	0.0954	0.9138
	-21.930	4.5341	4.2321	18.7966	-11.930	1.3945	-0.7688	0.7352	2.4754
	-13.456	14.4807	10.5773	-51.4960	34.4555	-13.940	2.8517	-1.3015	-1.7862
	-15.632	-7.3232	-6.4325	16.2066	-17.170	48.9586	-7.4408	1.5838	2.9125
	14.3470	-28.909	-19.609	1.7344	8.9383	7.4871	10.8223	2.3558	-2.5235
$LW^{ho}_{(n,m)}$	$LW^{ho}_{(n,1)}$	$LW^{ho}_{(n,2)}$	$LW^{ho}_{(n,3)}$	$LW^{ho}_{(n,4)}$	$LW^{ho}_{(n,5)}$	$LW^{ho}_{(n,6)}$	$LW^{ho}_{(n,7)}$	$LW^{ho}_{(n,8)}$	$LW^{ho}_{(n,9)}$
	-43.534	-57.668	92.7373	-16.9546	-22.205	44.7903	77.4224	27.6844	-91.4983
$\mathbf{b}_{1(m)}$	$\mathbf{b}_{1(1)}$	$\mathbf{b}_{1(2)}$	$\mathbf{b}_{1(3)}$	$\mathbf{b}_{1(4)}$	$\mathbf{b}_{1(5)}$	$\mathbf{b}_{1(6)}$	$\mathbf{b}_{1(7)}$	$\mathbf{b}_{1(8)}$	$\mathbf{b}_{1(9)}$
	10.4633	13.4605	9.3992	-17.0724	13.6713	-40.546	4.9108	-2.3395	0.6071
$\mathbf{b}_{2(n)}$	$\mathbf{b}_{2(1)}$								
	19.2366								

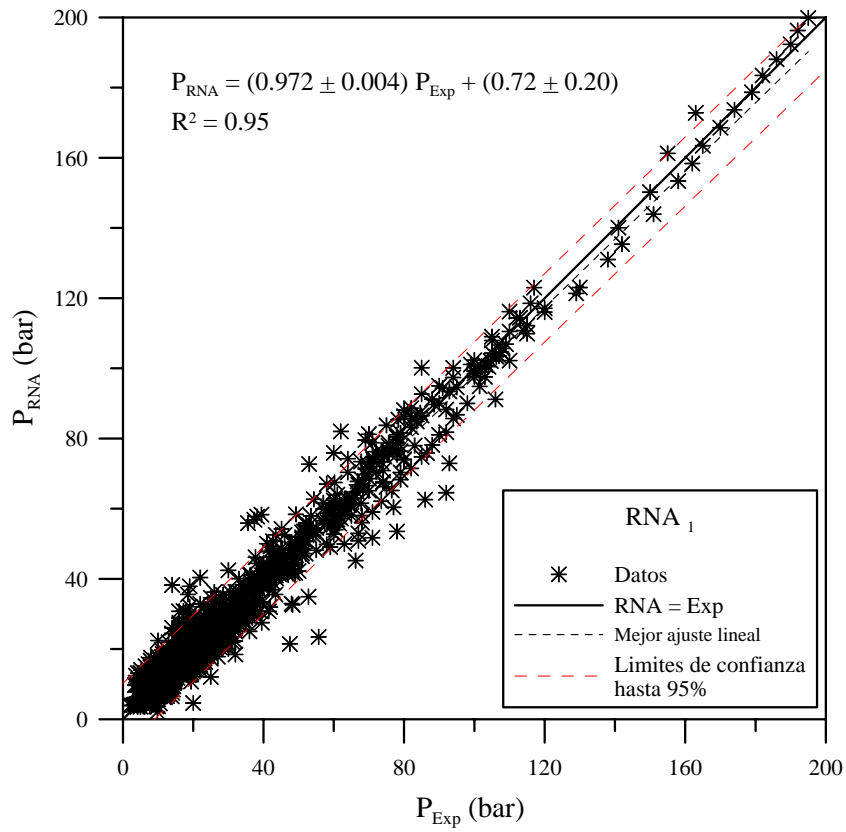


Figura 3.4 Comparación entre valores de presión medidos y los simulados por la RNA₁.

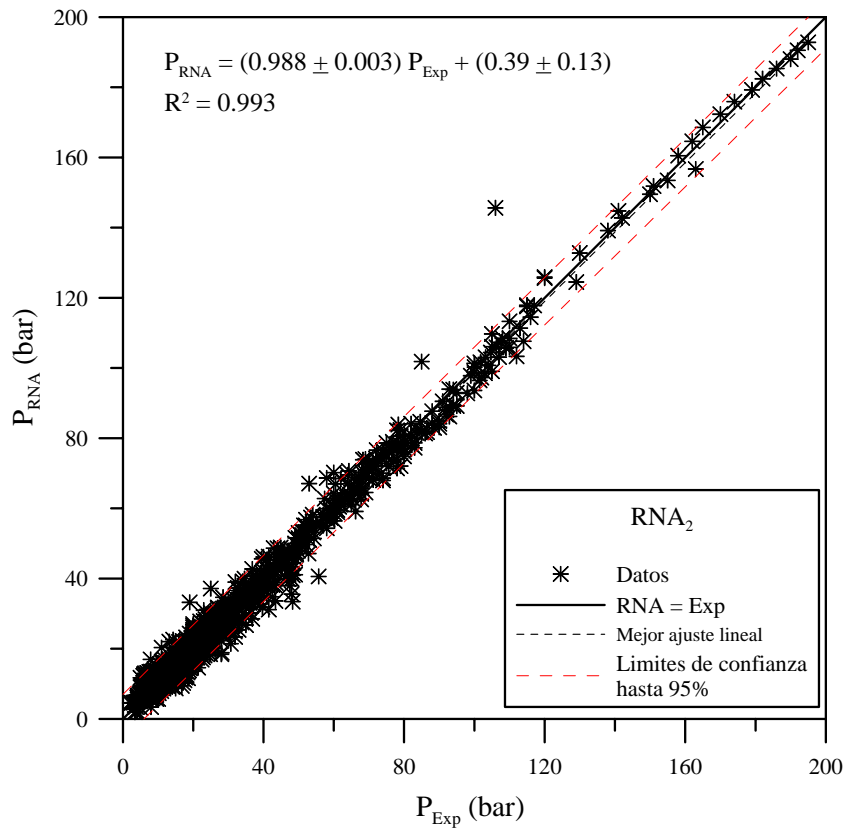


Figura 3.5 Comparación entre valores de presión medidos y los simulados por la RNA₂.

3.4.2 Análisis de sensibilidad

La evaluación de la importancia relativa de las variables de entrada en los dos modelos desarrollados de RNA (en particular, el efecto de la variable de presión de fondo en el modelo de RNA₂), se utilizaron las matrices de los pesos y la ecuación de Garson para el análisis de sensibilidad (Elmolla *et al.*, 2010). Garson (1991) sugiere la ecuación siguiente:

$$I_j = \frac{\sum_{m=1}^{N_h} \left(\left(|IW_{(j,m)}^{ih}| / \sum_{k=1}^{N_i} |IW_{(k,m)}^{ih}| \right) \times |LW_{(n,m)}^{ho}| \right)}{\sum_{k=1}^{N_i} \left\{ \sum_{m=1}^{N_h} \left(\left(|IW_{(k,m)}^{ih}| / \sum_{k=1}^{N_i} |IW_{(k,m)}^{ih}| \right) \times |LW_{(n,m)}^{ho}| \right) \right\}} \quad (3.6)$$

donde, I_j es la importancia relativa de la variable de entrada j^{th} ; N_i y N_h son el número de las neuronas en las capas de entrada y la oculta, respectivamente, y IW y LW son los pesos de conexión entre las capas de entrada-oculta y oculto-salida, respectivamente (ver Tabla 3.3). Los subíndices 'i', 'h' y 'o' se refieren a las capas de entrada, oculta y de salida, respectivamente, mientras que 'k' subíndices 'm' y 'n' se refieren a las neuronas de las capas de entrada, oculta y de salida, respectivamente.

Los resultados del análisis de sensibilidad obtenido para ambos modelos de RNA se muestran en la Tabla 3.4. Dos grupos diferentes de los variables de entrada están claramente definidos: las variables geométricas (la profundidad del pozo, el ángulo de inclinación, y el diámetro del pozo) y las variables termodinámicas (flujo másico, la temperatura y la presión del fondo del pozo). Desde un punto de vista general, la mayoría de los variables de entrada mostraron un efecto importante en la predicción de los gradientes de presión (porcentajes desde 18.12% hasta 28.40% para la RNA₁, y de 10.51% hasta 25.24% para la RNA₂), excepto el variable de profundidad del pozo (que proporcionó porcentajes de sensibilidad más bajos: 4.77% y 8.26% para ambos modelos).

Por otro lado, cuando los resultados de sensibilidad son analizados individualmente por cada modelo de RNA, y se puede observar que la temperatura de fondo (28.40%) y la presión de fondo (25.24%) son las variables de entrada más importantes para la RNA₁ y RNA₂, respectivamente. Cuando los resultados del análisis de sensibilidad de los modelos RNA₁ y RNA₂ se comparan entre sí, las variables termodinámicas siempre tienen porcentajes más altos que las variables geométricas, y en particular la presión de fondo

resulta ser la más importante para la predicción de los gradientes de presión de pozos productores. Probablemente, este parámetro explica por qué el modelo RNA₂ predice mejor los datos medidos de presión ($R^2 = 0.9874$): Figura 3.5.

Tabla 3.4 Análisis de sensibilidad de los variables de entrada en los modelos de la RNA.

Parametros	Importancia relativa (%)	
	RNA₁	RNA₂
<i>Variables geométricas</i>		
Profundidad (z)	4.77	8.26
Ángulo de inclinación (θ)	18.12	15.73
Diámetro del pozo (D)	23.59	10.51
<i>Variables termodinámicas</i>		
Flujo másico (W)	25.13	17.60
Temperatura del fondo (T_b)	28.40	22.66
Presión del fondo (P_b)	-	25.24

3.4.3 Validación de los gradientes de presión calculados por las RNA y GEOWELLS

La evaluación de la capacidad de predicción de ambos modelos de RNA aquí desarrollados, fue calificada mediante la reproducción de los gradientes de presión de producción de los 10 pozos de validación (es decir, Okoy-7 y Ngawha-11: Ambastha y Gudmundsson, 1986; M-39 y KW-2: Bjornsson, 1987; A-12, C-2 y KE1-11: Garg *et al.*, 2004; Az-42: Sánchez-Upton, 2000; N-22T: Garg y Combs, 2002; y W-1: Barelli *et al.*, 1982). Esta base de datos de validación no fue incluida en el proceso de entrenamiento de los modelos de la RNA. En el análisis comparativo se utilizó el simulador GEOWELLS para determinar los mismos gradientes de presión.

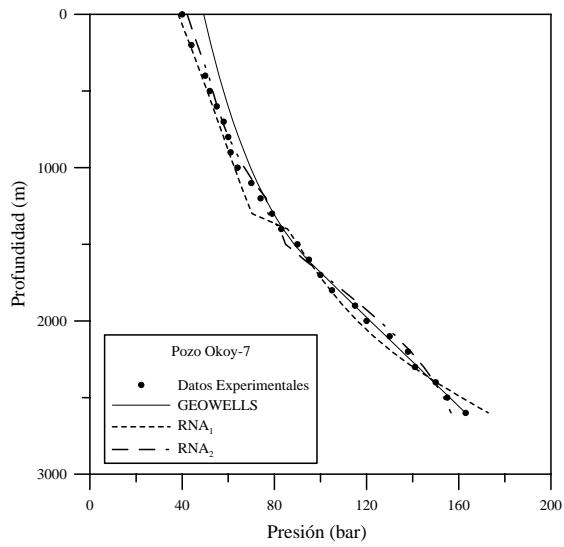


Figura 3.6 Perfiles de presión predichos mediante GEOWELLS y RNA vs. medidos para el pozo Okoy-7.

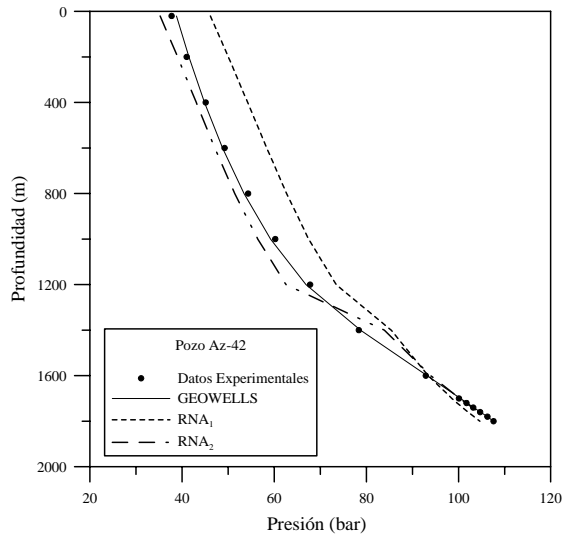


Figura 3.7 Perfiles de presión predichos mediante GEOWELLS y RNA vs. medidos para el pozo Az-42.

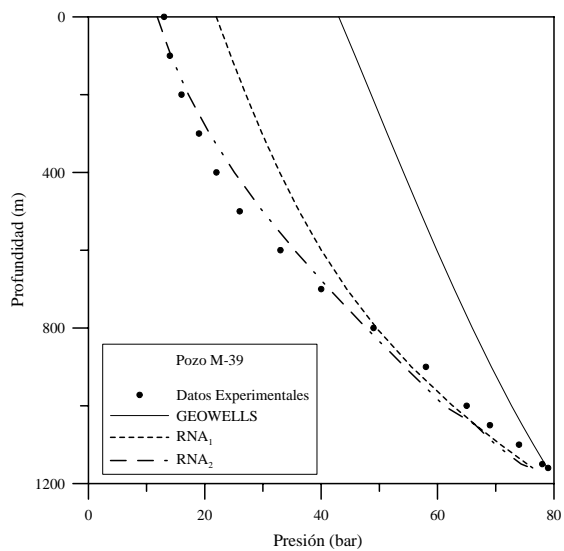


Figura 3.8 Perfiles de presión predichos mediante GEOWELLS y RNA vs. medidos para el pozo M-39.

Los resultados numéricos de los gradientes calculados por los modelos RNA_1 , RNA_2 y GEOWELLS juntos con datos de medición de campo de tres pozos geotérmicos (Okoy-7, Az-42, y M-39) son gráficamente mostrados en las Figuras 3.6, 3.7 y 3.8. Generalmente se puede observar claramente en estas figuras que los gradientes de presión calculados por el modelos RNA_2 (P_{RNA}) se ajustan mejor a los datos medidos en el campo P_{med} (con errores menores a 2.3%); mientras que los resultados obtenidos por el modelo RNA_1 fueron menos precisos (varían desde 2.5% para el pozo Okoy-7 hasta 15.1% para el pozo M-39). Por otro lado, el simulador GEOWELLS con el modelo de flujo homogéneo, calcula los gradientes de presión con mejor precisión, menos del 4%) para ambos pozos: Okoy-7 (Figura 3.6) y Az-42 (Figura 3.7); mientras que para el pozo M-39 (Figura 3.8), GEOWELLS con el modelo de flujo homogéneo presenta resultados con menor precisión (errores de predicción hasta del 35%), esto probablemente se atribuye a un problema de geometría del pozo (posiblemente causada por un cambio en el diámetro del pozo debido a un problema de escala mencionado por Goyal *et al.*, 1980 y Bjornsson, 1987).

3.4.4 Análisis estadístico comparativo entre datos de predicción y mediciones de campo

Un análisis estadístico global de los gradientes de presión de los 10 pozos geotérmicos utilizados para la validación de los tres modelos (RNA_1 , RNA_2 y GEOWELLS) se llevó a cabo mediante los parámetros descritos en la Tabla 3.2. Los resultados de esta comparación estadística se presentan en la Tabla 3.5.

Después de analizar el parámetro estadístico (MPE), se encontró que los promedios de los porcentajes de los errores de los gradientes de presión simulados (en valor absoluto) llegaron hasta valores de 41.8%, 14.3% y 80.7% para los modelos RNA_1 , RNA_2 y GEOWELLS, respectivamente. Esta comparación estadística fue confirmada también mediante el análisis de los valores de parámetro $RMSE$, donde se obtuvieron rangos de 1.3 a 11.4 bares; de 0.5 a 3.8 bares; y de 0.2 a 22.8 bares para predicciones de RNA_1 , RNA_2 , y GEOWELLS, respectivamente. Según el análisis estadístico completo de los residuos (es decir, MPE y $RMSE$), se confirmó que los gradientes de presión fueron sistemáticamente mejor predichos por el modelo RNA_2 . Estos resultados de predicción fueron validados estadísticamente mediante el tercer parámetro *Theil's U* (Tabla 3.2).

Tabla 3.5 Comparación estadística de los gradientes de presión calculados por (RNA₁, RNA₂ y GEOWELLS). Homogéneo

Pozo geotérmico (País) / Parámetros estadísticos de comparación	Herramientas Computacionales		
	RNA ₂	RNA ₁	GEOWELLS Modelo de flujo homogéneo
KW-2 (Islandia)			
<i>RMSE</i>	3.68	11.42	3.08
<i>MPE</i>	2.60	-8.26	-0.34
<i>THEIL'S U (r)</i>	-	0.32	1.19
Okoy-7 (Filipinas)			
<i>RMSE</i>	3.68	4.22	4.01
<i>MPE</i>	2.60	-2.54	4.14
<i>Theil's U (r)</i>	-	0.80	0.85
A-12 (Japón)			
<i>RMSE</i>	1.57	1.73	0.58
<i>MPE</i>	-10.96	-5.32	3.51
<i>Theil's U (r)</i>	-	0.90	2.72
Az-42 (México)			
<i>RMSE</i>	2.59	6.11	0.56
<i>MPE</i>	-2.34	6.35	-0.06
<i>Theil's U (r)</i>	-	0.42	4.66
C-2 (Japón)			
<i>RMSE</i>	0.52	1.32	0.21
<i>MPE</i>	-3.55	-10.79	1.04
<i>Theil's U (r)</i>	-	0.40	2.56
KE1-11 (Japón)			
<i>RMSE</i>	2.03	3.82	8.47
<i>MPE</i>	-2.83	9.77	-80.70
<i>Theil's U (r)</i>	-	0.53	0.24
M-39 (México)			
<i>RMSE</i>	3.80	7.13	22.77
<i>MPE</i>	-0.50	15.15	35.20
<i>Theil's U (r)</i>	-	0.53	0.17
N-22T (Japón)			
<i>RMSE</i>	3.48	2.02	9.32
<i>MPE</i>	14.27	-2.63	30.74
<i>Theil's U (r)</i>	-	1.72	0.37
Ngawha-11 (Nueva Zelanda)			
<i>RMSE</i>	2.96	7.41	3.75
<i>MPE</i>	3.53	-41.79	-2.89
<i>Theil's U (r)</i>	-	0.40	0.79
W-1 (Italia)			
<i>RMSE</i>	3.07	4.88	11.10
<i>MPE</i>	-2.06	-0.39	-48.46
<i>Theil's U (r)</i>	-	0.63	0.28

La interpretación final del coeficiente *Theil's U* muestra que los valores obtenidos con la RNA₁ y el simulador GEOWELLS son en su mayoría menores que 1, lo cual significa que el modelo RNA₂ proporciona datos de predicción de presión estadísticamente más cercanos a los datos medidos en campo.

Por último, Se debe mencionar que el simulador GEOWELLS fue evaluado en este estudio comparativo utilizando correlaciones de modelo de flujo homogéneo y que este programa proporciona en algunos pozos y utilizando las correlaciones de modelos de flujo separado perfiles de presión de producción con mejores resultados que la RNA.

Capítulo 4

Estimaciones de Flujo de calor y sus incertidumbres mediante un modelo de transferencia de calor simplificado y el método de Monte Carlo

4.1 Introducción

En las estimaciones de las reservas energéticas de una región geotérmica, ya sea las tradicionalmente explotadas a partir de fluidos o de roca seca caliente (esta última concebida actualmente como una nueva línea de investigación en la industria geotérmica), es necesario determinar el flujo de calor conductivo a diferentes profundidades. Existe en la literatura muchos trabajos donde se propone el cálculo de este flujo conductivo mediante la ecuación de Fourier considerando un sistema de una sola capa o de múltiples capas y a partir de valores experimentales de temperaturas de formación (TEF) en diferentes profundidades y de conductividad térmica de rocas (Haenel *et al.*, 1988; Pollock *et al.*, 1993; Springer Michael, 1999 y Morin *et al.*, 2010). La determinación del flujo de calor en la literatura internacional no es comúnmente reportada con sus errores, y esto es debido a que los parámetros que intervienen en su cuantificación no son estimados de manera estadísticamente correcta, es decir no se reportan con sus incertidumbres.

El método de Monte Carlo (MC) ha alcanzado últimamente un creciente interés en la geofísica, y no sólo por el mejoramiento constante de la eficiencia de las computadoras. Una de las primeras aplicaciones geotérmicas de este método fue dada por Cermak (1971) quien estudió el historial de las temperaturas de la Tierra a partir de los sitios de perforación. Otro historial de temperaturas en la capa de hielo de Groenlandia fue realizado con el mismo método por Dahl-Jensen *et al.* (1998). Royer y Danis (1988) aplicaron la variación aleatoria para el flujo de calor del manto y calcularon los intervalos de confianza del campo térmico. Lamontagne y Ranalli (1996) estudiaron los efectos de la incertidumbre de los parámetros térmicos en las propiedades reológicas de una zona sísmicamente activa. Jokinen (2000) presentó un trabajo completo sobre las estimaciones de la conductividad térmica de rocas a partir de registros de flujo de calor, en donde

propuso las simulaciones de MC para la solución del problema inverso. Sari *et al.* (2010) también aplicaron el método de MC para la estimación de un rango confiable para el parámetro Hoek–Brown de rocas.

El objetivo central de este estudio fue el desarrollo de una metodología que considera el problema de las incertidumbres de los variables antes mencionado para una estimación confiable de los flujos de calor conductivos utilizando la propagación de errores a través de la teoría de simulaciones de Monte Carlo. Esta metodología fue validada con datos reportados de flujo de calor en diferentes zonas geotérmicas del mar de Marmara en Turquía, metodología que fue exitosamente aplicada en la parte del colapso central de la caldera del campo geotérmico mexicano de Los Humeros, Puebla. La selección de este campo se sustentó debido a que se cuenta con información suficiente de registros de temperaturas estabilizadas de formación, de la estratigrafía de los pozos perforados y de las propiedades termofísicas de las rocas, específicamente la conductividad térmica.

4.2 El campo geotérmico Los Humeros

El Campo Geotérmico de Los Humeros (CGLH) está situado en el estado de Puebla cerca de los límites con Veracruz, en la parte Este del Cinturón Volcánico Mexicano (CVM, Mexican Volcanic Belt; Aguilar-y-Vargas, 1985). Este campo produce electricidad a partir de fluidos contenidos en rocas volcánicas y es reconocido como el tercer campo geotérmico en importancia en la República Mexicana, después de Cerro Prieto en Baja California y Los Azufres en Michoacán (Ferriz, 1985; Cedillo-Rodríguez, 2000). El campo se encuentra a una elevación promedio de 2800 metros sobre el nivel del mar y se encuentra localizado entre las coordenadas geográficas 19° 35' y 19° 45' latitud Norte, y 97° 23' y 97° 35' de longitud Oeste (Figura 4.1). Es pertinente señalar que existen una serie de calderas del CVM que han sido identificadas y en las cuales se pueden realizar estudios de exploración con fines geotérmicos. Algunas de estas calderas se señalan en la Figura 4.1.

El CGLH ha sido un lugar de intensa actividad volcánica, de gran interés y ampliamente estudiado por diversas disciplinas como: geología (Pérez-Reynoso 1979; Ferriz 1985), geoquímica (Ferriz, 1985; Ferriz y Mahood, 1987; Barragan *et al.*, 1991), geocronología

(Ferriz y Mahood, 1984), geofísica (Flores *et al.*, 1978; González-Morán y Suro-Pérez, 1982; Campos-Enríquez y Arredondo-Fragoso, 1992; Campos-Enríquez *et al.*, 2005), alteración hidrotermal (González-Partida *et al.*, 2001; Martínez-Serrano, 2002; Bienkowski *et al.*, 2005), propiedades termofísicas (Contreras *et al.*, 1990), yacimiento y pozos geotérmicos (Medina-Martínez, 2000; Tello-Hinojosa, 1992; Cedillo-Rodríguez, 1999, 2000), y modelado térmico (Prol y González-Morán, 1982; Campos-Enríquez y Duran, 1986; Castillo-Román *et al.*, 1991; Verma y Andaverde, 2007).

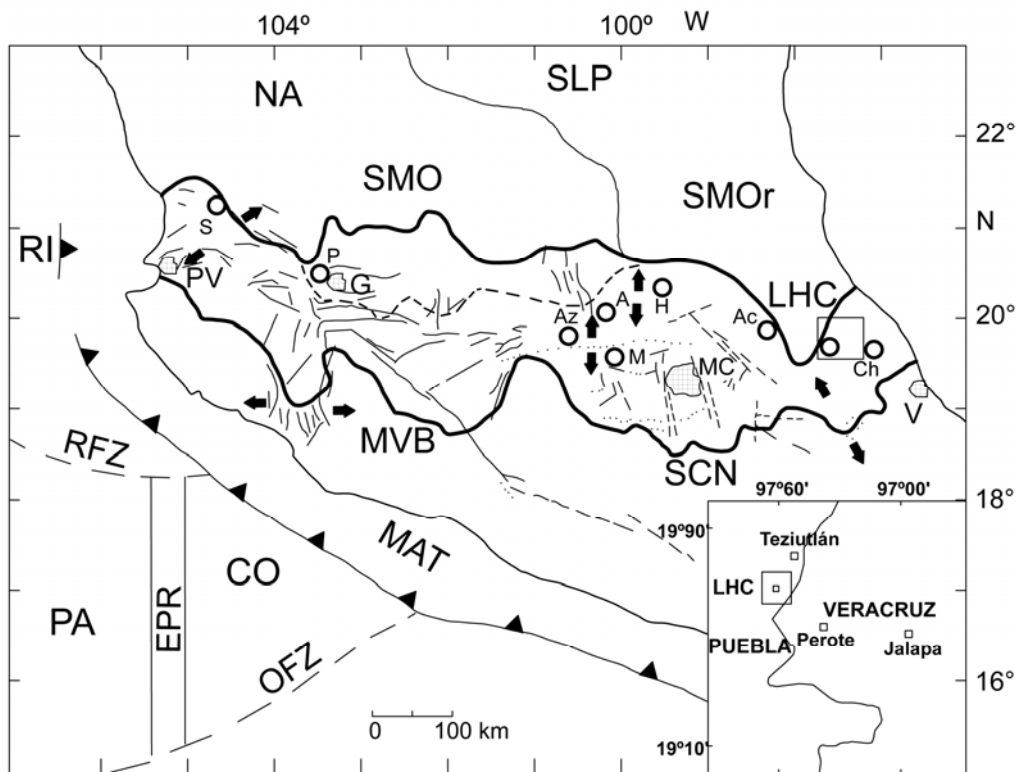


Figura 4.1 Localización y escenario tectónico del Cinturón Volcánico Mexicano (CVM), Verma (1985). Abreviatura: MAT= Trinchera de la Cordillera de América; EPR = Elevación del Pacífico Este; OFZ = Zona de Fractura Orozco; RFZ = Zona de Fractura Rivera; CO = Plato de Cocos; RI = Plato de la Rivera; PA = Plato del Pacífico; NA = Plato de América del Norte; SMO = Provincia Sierra Madre Occidental; SMOr = Provincia Sierra Madre Oriental; SCN = Campo monogenético de la Sierra de Chichinautzin; SLP = Área de San Luís Potosí. Calderas del CVM: S = Santa María del Oro; P = La Primavera; Az = Los Azufres; A = Amealco; M = Mazahua; H = Huichapan; Ac = Acoculco; Hu = Los Humeros; Ch = Chiconquiaco. Ciudades: MC = Ciudad de México; V = Veracruz; PV = Puerto Vallarta; G = Guadalajara.

4.2.1 Evolución volcánica

Durante los últimos 460,000 años, tres grandes erupciones plinianas y dos periodos de emplazamiento de derrames de lava periódicamente muestrearon la cámara magmática de Los Humeros (Ferriz, 1985). La primera actividad volcánica está representada por la erupción de lavas andesíticas y ferrobásalticas de la Formación Teziutlán (3.6 y 1.6 Ma). Poco después (0.46 Ma) tuvo lugar una erupción que llevó al emplazamiento de la Ignimbrita Xáltipan, dando origen al colapso de la Caldera Los Humeros, cuya superficie ha sido estimada en 315 km². La Ignimbrita Xáltipan que representa aproximadamente 115 km³ de magma es predominantemente riolítica. Posterior al colapso, se formaron varios domos silíceos de 0.3 Ma. Una tercera etapa de erupción tuvo lugar hace 0.1 Ma, 12 km³ de magma formaron la Ignimbrita Zaragoza, lo que provocó el colapso de la Caldera de Los Potreros, la cual tiene un diámetro promedio de 10 km y está anidada dentro de la Caldera de Los Humeros. Las emisiones de andesitas, andesitas basálticas y lavas, pómez y cenizas (0.08 – 0.04 Ma) dieron origen al colapso de la Caldera Xalapazco, de 1.7 km de diámetro, en el extremo sur de la Caldera de Los Potreros. La última actividad volcánica está representada por la eyección de aproximadamente 0.25 km³ de basaltos de olivino, durante los últimos 0.02 Ma, en los pisos de las Calderas de Los Potreros y El Xalapazco. En relación con los 0.46 Ma de vida activa del sistema de estas erupciones representan eventos esencialmente instantáneos, de tal manera que es razonable suponer que las variaciones en composición de cada una de estas unidades representa el zoneamiento composicional que prevalecía en la cámara magmática cuando la erupción tuvo lugar.

4.2.2 Termodinámica del yacimiento del CGLH

Los estudios de geoquímica de gases y de ingeniería de yacimientos indican la existencia de dos sistemas geotérmicos en el CGLH (Barragan *et al.*, 1991; Cedillo-Rodríguez, 1999, 2000), uno superior no agresivo con pH neutro (7-9) y fluido de tipo bicarbonatado-sódico alojado en andesita de augita; el otro inferior de alta temperatura y pH ácido (5-6), alojado en andesita de hornblenda con fluidos clorurados-sódicos. Las características e información del segundo yacimiento, conllevan a que la disposición del fluido presenta algunas problemas ya que al extraerlos, provocan la corrosión y desgaste de la tubería del pozo geotérmico (lo que implica gastos de mantenimiento), quedando explotable sólo el yacimiento superior. El CGLH ha sido ampliamente estudiado, haciendo referencia para

finés de este trabajo de investigación básicamente al área de la termodinámica del yacimiento, en el cual trabajos desarrollados han sido enfocados al estudio e interpretación de la distribución de presión y temperaturas no perturbadas de los fluidos alojados en los yacimientos del CGLH.

El primer yacimiento (zona productora superior) se encuentra localizado a una profundidad entre 1,800 y 1,100 m.s.n.m. (metros sobre el nivel del mar) y es un yacimiento de líquido dominante (Jiménez y Herrera, 2000). La segunda zona productora se encuentra localizada entre los 900 y 200 m.s.n.m. y es un yacimiento de baja saturación del líquido, encontrándose una saturación del agua menor al 10 % en donde el vapor predomina y controla la presión (Jiménez y Herrera, 2000). Los datos de presión reportados en la literatura indican que esta variable termodinámica varía de 90 hasta 176 bares. En cuanto a datos de temperatura, estudios realizados han reportado que en la sección de la zona productora superior se han registrado temperaturas en los pozos entre los 290 y 300 °C.

Los registros de pozos localizados en la zona inferior muestran temperaturas mayores entre 300 y 400 °C (Hernández, 1995; Jiménez y Herrera, 2000; Gutiérrez, 2009). La información antes mencionada está resumida en la Tabla 4.1 La información de datos de presión y temperatura de los fluidos geotérmicos ha permitido reafirmar la existencia de los dos yacimientos, siendo actualmente la zona más rentable para la explotación y generación de energía eléctrica la que corresponde al yacimiento superior. Esta zona productora se localiza en el colapso Central (dentro de la caldera de Los Potreros) y en el corredor de Maztaloya. Cabe mencionar que a la fecha, la producción total de los pozos perforados en el CGLH es de 88, 150, 730 toneladas de fluido al año, de los cuales el 82.1% es vapor y el 17.9% es líquido (Gutiérrez, 2009). A la fecha se han perforado 40 pozos en el campo con profundidades que oscilan entre los 1,450 a 3,250 m (Herrera y Jiménez, 1999; Jiménez y Herrera, 2000; López, 2006; Tabla 4.1). Los pozos con mayor producción de gasto de vapor (t/h) son el H9, H7, H12, H20, H34, H31, H32, H15, H6, H19 y H37 (en orden de producción de 60 a 30 t/h). El pozo H1 (primer pozo perforado) es el único en el campo que produce a la fecha un considerable gasto de agua (66.20 t/h) y 18.4 t/h de vapor.

Tabla 4.1 Características termodinámicas de las zonas productoras del CGLH (tomado de Jiménez y Herrera, 2000).

Yacimiento	Litología	Profundidad (m.s.n.m.)	Presión (Bares)	Temperatura (°C)	Permeabilidad (mD)
Superior	Andesita de auguita	1800 - 1100	90 - 176	290 - 330	1 - 26
Inferior	Andesita de Horblenda	900 - 200	90 - 176	300 - 350	1 - 10

4.2.3 Geología del subsuelo del CGLH

Con el fin de conocer la distribución y límites en el subsuelo de los diferentes tipos de rocas (espesores y litología), se han realizado trabajos encaminados a dar respuesta a estas tareas geológicas (Cedillo-Rodríguez, 1999, 2000). Las características geológicas del subsuelo del CGLH han sido de gran utilidad como herramienta de apoyo a las áreas de la ingeniería de yacimientos, la perforación de pozos, la geoquímica, la geofísica y a la hidrogeología, mismas que han aportado valiosa información la cuestión de modelos conceptuales del campo así como para desarrollar y aplicar simuladores de yacimientos geotérmicos.

Con base en determinaciones litológicas y muestras de núcleos de pozos geotérmicos, la geología del subsuelo del CGLH se ha definido inicialmente en cuatro secciones litológicas principales (Viggiano y Camacho, 1998; Jiménez y Herrera, 2000; Cedillo -Rodríguez 1999, 2000), lo que ha permitido hacer referencia de la existencia de nueve unidades litológicas distribuidas (Tabla 4.2):

- I. Esta sección está constituida por una unidad de pómez, basaltos y andesitas, el espesor de esta varía entre 200 y 300 m. Tienen la característica hidrológica de formar acuíferos superficiales fríos o calientes además detener una alta permeabilidad.
- II. Esta sección está evidenciada por tres unidades litológicas (que subyacen a la primera), una de tobas líticas con espesores que varían entre los 250 y 450 m de espesor con poca permeabilidad. La segunda unidad de ignimbritas con espesores de 200 y 300 m,

con una permeabilidad casi nula. Una tercera unidad de intercalaciones de andesitas e ignimbritas con espesores entre 200 y 250 m, con baja permeabilidad.

Tabla 4.2 Geología del subsuelo del CGLH.

Sección litológica	Unidad litológica	Litología	Permeabilidad
I	1	Pómez, Basaltos y Andesitas	Alta
II	2	Tobas Líticas	Media
	3	Ignimbritas Líticas y	Baja
	4	Tobas Andesitas e Ignimbritas	Baja
II	5	Andesita de Auguita	Media
	6	Toba Vítreo Humeros	Baja
	7	Andesita de Horblenda	Media
	8	Basaltos	Media
IV	9	Calizas e Intrusivos	Baja

III. Esta sección está constituida por cuatro unidades litológicas. La primera de andesita de augita, donde se localiza el yacimiento superior del CGLH, esta unidad se caracteriza por tener alta permeabilidad, con espesores entre los 500 y 700 m. La segunda unidad es la toba vítrea Humeros, tiene la función de ser una frontera impermeable entre la andesita de augita y la andesita de horblenda (tercera unidad), con espesores entre los 150 y 250 m. La tercera unidad es la andesita de horblenda, donde se localiza el yacimiento inferior del CGLH, tiene una permeabilidad moderada y espesores entre los 200 y 350 m. La cuarta unidad de basaltos, con espesores entre 200 y 300 m, y una permeabilidad baja.

IV. Esta sección está constituida por una unidad litológica de calizas e intrusivos con espesores superiores a 200 m, y una permeabilidad baja.

La información anterior, describe las unidades litológicas perforadas por los pozos geotérmicos, quedando como incógnita el material existente a una mayor profundidad, llevando a suponer la existencia de rocas intrusivas.

4.3 Metodología usada para la determinación del flujo de calor

Se considera que la determinación de flujo de calor conductivo una tarea clave en la estimación de reservas energéticas en los campos geotérmicos y que es para el cálculo de este flujo se requiere tener un conocimiento amplio de las variables involucrados en esta determinación: Asaber, las Temperaturas Estabilizadas de Formación (TEF) y la conductividad térmica de rocas (variables explicadas ampliamente en las siguientes secciones). Existe muchos métodos para la estimación de las temperaturas estabilizadas de formación (Manetti *et al.*, 1973; Dowdle y Cobb, 1975; Ascensio *et al.*, 1994; Andaverde *et al.*, 2005 y Bassam *et al.*, 2010) estos métodos tratan de dar un aproximación de las TEF con errores. La fuente de estos errores puede deberse o estar relacionada con: (1) las hipótesis consideradas en las ecuaciones diferenciales que gobiernan la transferencia de calor, (2) las incertidumbres en los instrumentos que registran los valores de las variables de campo y, (3) las regresiones estadísticas aplicadas para resolver las ecuaciones analíticas o simplificadas de difusión de calor.

Por otra parte, las determinaciones experimentales de la conductividad térmica de las rocas en el laboratorio no reflejan de manera exacta lo que sucede en la roca al encontrarse en el subsuelo, ya que esta variable cambia debido a las condiciones de presión y temperatura y a los fluidos que en ellas están contenidos. Esto provoca, por lo tanto, que cada valor determinado experimentalmente en el laboratorio debe de ser expresado con las respectivas incertidumbres que provocan la inexactitud propia de los instrumentos de medición, las heterogeneidades propias de las rocas, así como los ya mencionados cambios en las condiciones de las rocas al modificar su condición *in situ* a el laboratorio.

De lo antes expuesto se puede concluir que es necesario desarrollar una metodología que involucre las incertidumbres en los parámetros de entrada de las ecuaciones de flujo de calor conductivo para transportarse en los resultados, a través de las estimaciones de la media y sus las incertidumbres.

4.3.1 Temperaturas Estabilizadas de Formación

Se define la TEF como la temperatura virgen de un sistema geotérmico que existió antes de la perturbación térmica inducida por la perforación de un pozo (al circular un fluido de baja temperatura con el objeto de lubricar y enfriar la barrena, así como para extraer recortes de la formación). En la industria geotérmica se considera que la TEF es fundamental para: (i) La estimación de los gradientes geotérmicos, (ii) La determinación del flujo de calor; (iii) La interpretación de registros de pozos; (iv) El diseño óptimo de programas de perforación y terminación de pozos; (v) La estimación del cambio de temperatura en regiones “permafrost”; y (vi) La estimación de las reservas de calor en los sistemas geotérmicos. Se ha reportado durante décadas en la literatura diferentes métodos de cálculo de esta propiedad (Manetti *et al.*, 1973 y Ascencio *et al.*, 1994) el más utilizado es el método de Horner (MH) reportado por Dowdle y Cobb (1975). Es una de las herramientas más comúnmente usadas para estimar TEF, la cual consiste en graficar la temperatura de fondo a pozo cerrado (sin circulación) contra el logaritmo del tiempo adimensional de Horner. Este método se fundamenta en el concepto de una línea fuente/sumidero de calor infinitamente largo sin sección transversal para representar los procesos de circulación y paro. El modelo matemático simplificado del método de Horner esta dado por la siguiente ecuación:

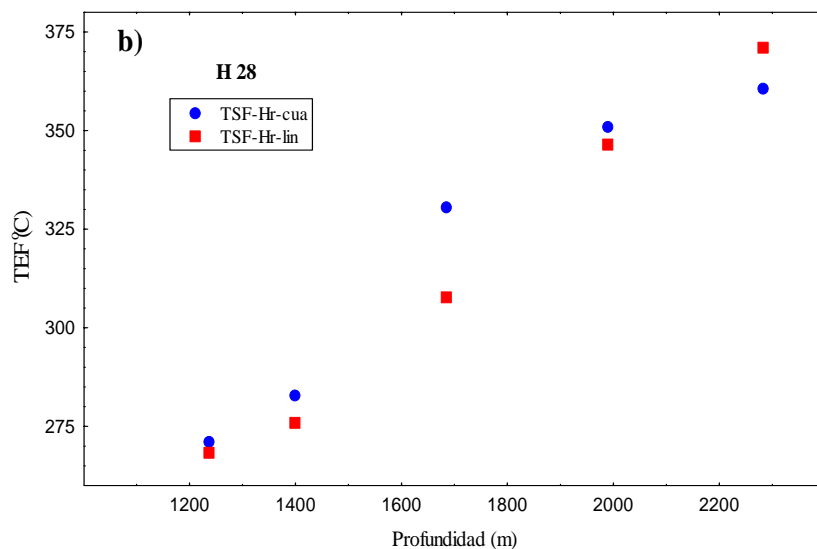
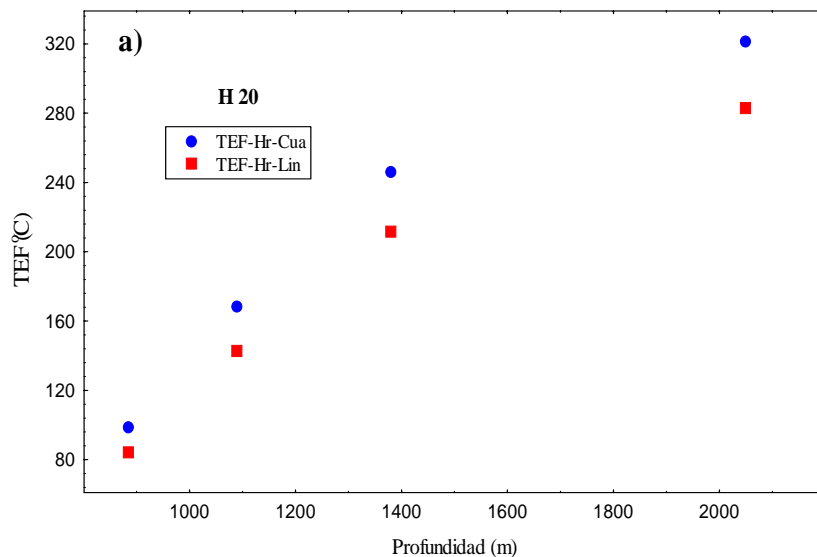
$$TFP(t) = TEF_{HM} + (b_{HM}) \cdot \ln\{(t_c + \Delta t) / \Delta t\} \quad (4.1)$$

Donde t_c es el tiempo de circulación, Δt es el tiempo de recuperación, TFP es la temperatura del pozo en reposo, $[(t_c + \Delta t) / \Delta t]$ es el tiempo adimensional de Horner y TEF_{HM} es la temperatura inicial o estabilizada de la formación. Así, una gráfica de TFP contra $[(t_c + \Delta t) / \Delta t]$ describe una línea recta con pendiente b_{HM} y ordenada al origen TEF_{HM} .

Aunque el método de Horner ha sido usado ampliamente en la industria geotérmica, se considera que sistemáticamente subestima las temperaturas de las formaciones y se ve limitado cuando se tienen gradientes de temperaturas muy pequeños y tiempos de circulación muy cortos (Dowdle y Cobb, 1975). Debido a este problema Andaverde *et al.* (2005) reportaron una modificación al método de Horner basándose en un ajuste

cuadrático de la solución del modelo conductivo de la línea fuente, lo que evita la subestimación de TEF que se observa en este método.

En el presente trabajo se utilizó el método de Horner (lineal y cuadrático) para calcular las TEF en los pozos que se ubican en el colapso central del campo Geotérmico de Los Humeros en México y eso a partir de series de TFP y tiempos registrados durante las perforaciones (Gutiérrez, 2009). Los pozos estudiados en este caso son H11, H16, H20, H28, H29, H31 y H32, todos están ubicados en la zona del colapso central del CGLH. En las Figuras 4.2 (a, b y c) se presentan ejemplos de valores de TEF en diferentes profundidades y es notoriamente evidente la subestimación de las TEF obtenidas con el método de Horner lineal en los pozos H20, H28 y H31.



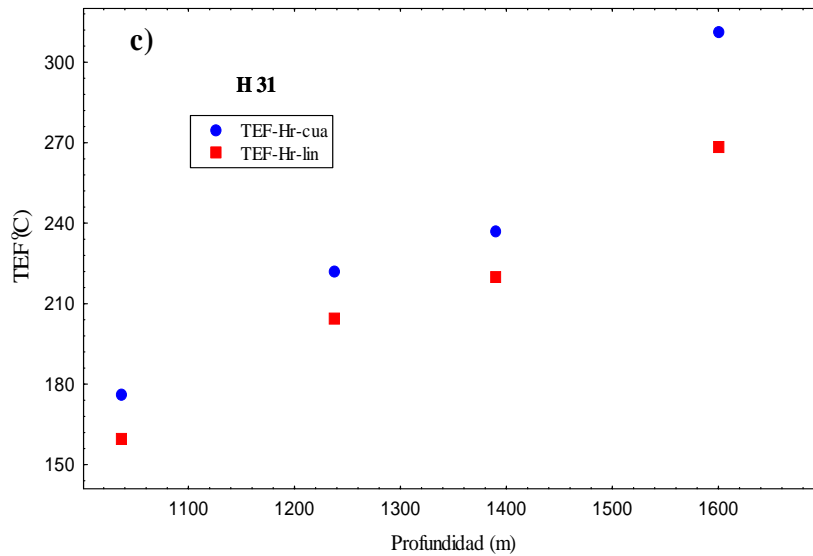


Figura 4.2 (a, b y c) Gradientes de TEF calculadas por el método de Horner (Lineal y Cuadrático) para los pozos H20, H28 y H31.

4.3.2 Conductividad térmica de rocas y su error

La conductividad térmica de roca es una de las propiedades térmicas importantes para calcular el flujo de calor, se determina experimentalmente en los laboratorios y existe diferentes métodos en la literatura que describen como se calcula esta propiedad, Carslaw y Jaeger (1959); Huenges *et al.* (1990); Pribnow y Sass (1995).

En el caso del CGLH se efectuaron determinaciones de la conductividad térmica en 15 especímenes, a condiciones de roca seca, como para roca saturada (Contreras *et al.*, 1990). Cada espécimen ensayado se obtuvo de un diferente núcleo. Las mediciones de la conductividad térmica de roca se efectuó mediante un procedimiento experimental propuesto por Carslaw y Jaeger (1959). Este procedimiento se basa en la solución de la ecuación de difusión de calor para un medio infinito dentro del cual existe una fuente lineal que emite calor uniformemente y axialmente con una rapidez de generación constante. Para representar el medio infinito se emplearon muestras cilíndricas de roca de aproximadamente 10 cm de diámetro por 12 cm de longitud.

Los resultados experimentales de la conductividad térmica de rocas secas y saturadas están reportados en la Tabla 4.3. El análisis de los 15 núcleos comprende seis tipos de roca.

Se analizaron diferentes muestras de la misma roca y se sacó una media y su desviación estándar como valor confiable de esta propiedad (Tabla 4.3).

Tabla 4.3 Conductividad térmica de roca seca y saturada (Contreras *et al.*, 1990).

Nombre de roca	Número de muestras	Conductividad térmica (W/m°C)	
		Roca seca X ± Sd	Roca saturada X ± Sd
Toba litica	3	0.973 ± 0.030	1.54 ± 0.05
Ignimbritas liticas y vitreas	3	1.161 ± 0.035	1.96 ± 0.06
Andesitas e ignimbritas	3	1.022 ± 0.030	1.96 ± 0.06
Andesitas cuaternarias (de augita)	7	1.28 ± 0.23	1.89 ± 0.23
Tobas vitreas	2	1.61 ± 0.09	2.36 ± 0.09
Andesitas terciarias (de hornblenda)	3	1.32 ± 0.38	2.08 ± 0.43

X: media , sd: desviación estándar

4.3.3 Estimación del Flujo de calor mediante un modelo de transferencia de calor simplificado y el método Monte Carlo

La conducción de calor puede ser presentada por la primera ley de Fourier (Haenel *et al.*, 1988), Que se ha establecido de forma experimental:

$$Q_{cond} = -k\nabla T = -k.gradT \quad (4.2)$$

donde: Q representa el flujo de calor conductivo, k es la conductividad térmica, y T es la temperatura.

Esta ecuación es utilizada en los sistemas de pozos geotérmicos para el cálculo de flujo de calor conductivo en un medio multicapas y por lo tanto podemos expresarla a través de la siguiente ecuación:

$$Q_{[i,i+1]cond} = \frac{TEF_{i+1} - TEF_i}{\sum_n \frac{L_j}{K_j}} \quad (4.3)$$

donde: $Q_{[i,i+1]_{cond}}$ es el flujo de calor entre dos puntos (i) y ($i+1$) (mW/m^2); TEF es la temperatura estabilizada de formación ($^{\circ}C$); K es la conductividad térmica de rocas ($W/m^{\circ}C$) y L es el espesor de la capa de roca (m).

Los valores de flujo de calor a determinarse con la Ecuación 4.3 tienen asociados errores debido a las incertidumbres individuales de las variables termofísicas que intervienen en la Ecuación 4.2. Normalmente se puede propagar los errores de esta ecuación de transferencia de calor usando la teoría de propagación a través de las ecuaciones de Bevington (Verma, 2005), cuando la estructura de la ecuación de flujo de calor conductivo es simple (Ecuación 4.3). A medida que esta ecuación se extiende a más de 3 capas el cálculo de FC se torna complejo y costoso. El alto costo se deriva de tener que medir las propiedades térmicas de las rocas en al menos una muestra por cada tipo de capa diferente presente en el sistema. Las características complejas se refieren a las operaciones de cálculo, las cuales pueden ser apoyadas con sistemas de cómputo y el método de Monte Carlo.

En la Figura 4.3 se muestra la metodología desarrollada y que consiste en los siguientes pasos:

- (i) Ubicación de perfil de TEF dentro del pozo y la determinación geológica del tipo de roca que se ubican dentro de dos TEF medidas,
- (ii) Determinar la ecuación de flujo de calor conductivo en un medio multicapas,
- (iii) Generar aleatoriamente para cada variable de la Ecuación 4.3 unas distribuciones normales de 100000 elementos con media 0 y desviación estándar 1,
- (iv) Normalizar estas distribuciones según el valor de la media y su desviación estándar de cada variable (Ecuación 4.4):

$$X_N = X_T + (S_T) \cdot (N_{aleatorio}) \quad (4.4)$$

donde, X_N es la variable normalizada, X_T es la media y S_T es la desviación estándar del variable.

- (v) Propagar el error de estas variables dentro de la Ecuación 4.3 y sacar una distribución de flujo de calor,

- (vi) Calcular la media representativa de flujo de calor y su desviación estándar. Estas etapas se repiten (N-1) veces para estimar el perfil de flujo de calor conductivo dentro de un pozo (N es el número de registros de TEF que tenemos en cada pozo) y se calcula el flujo para todos los pozos seleccionados y,

- (vii) Representar los valores de FC con su desviación estándar en una distribución espacial para su análisis.

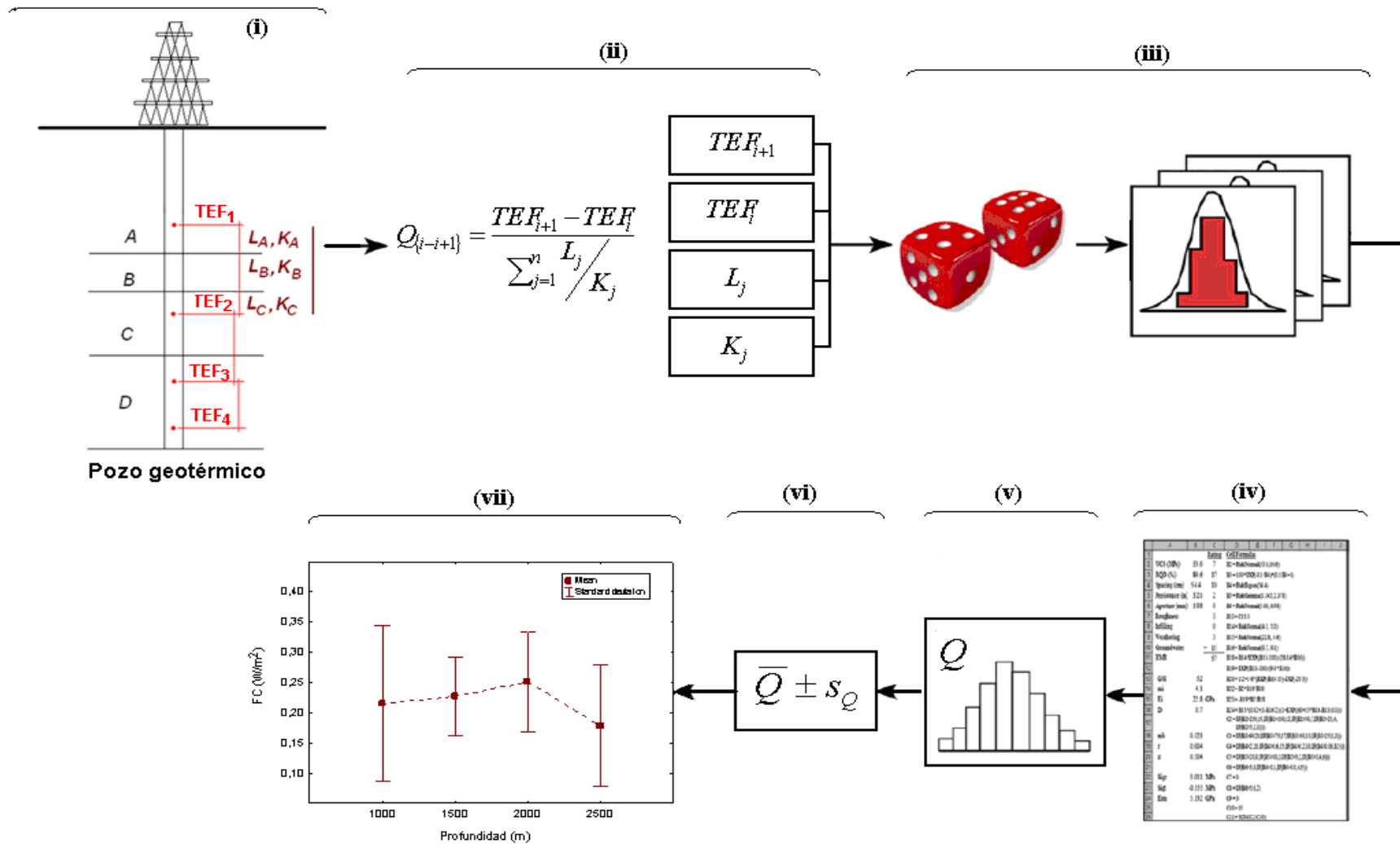


Figura 4.3 Metodología desarrollada para la estimación de flujo de calor conductivo en pozos.

4.4 Análisis de los resultados y discusiones

4.4.1 Validación de la metodología

La metodología antes de ser aplicada al CGLH fue numéricamente usando los siguientes casos de modelación: (a) aplicando las ecuaciones de flujo de calor en modelos de pocas capas (hasta 4 capas) y a través de una comparación estadística con los resultados obtenidos al propagar los errores con las ecuaciones de Bevington (Verma, 2005) , y (b) comparación estadística con valores valores y errores de flujo de calor reportados en la literatura internacional para pozos petroleros en las regiones de los mares de Marmara y Egeo en Turquía (Pfister *et al.*, 1998). En esta etapa se utilizaron errores de gradiente de temperatura entre ± 0.2 y ± 1.7 °C/km y de valores la conductividad térmica de la roca entre ± 0.2 y ± 0.3 W/m. °C los cuales fueron sugeridos por el mismo autor en su trabajo

En la Tabla 4.4, se presentan los resultados de la comparación de los casos modelados de validación en donde es evidente que no existe diferencia significativa entre los valores reportados y los calculados por ambos métodos (Bevington y Monte Carlo).

Para comparación visual de los resultados obtenidos se presenta la Figura 4.4. Se observa que los valores de media de flujo de calor tanto los reportados por (Pfister *et al.*, 1998) como los calculados mediante el MC son similares, situación que es demostrada a través del análisis estadístico de la regresión lineal ($Y = -1.196 + 1.017 X$) con una pendiente cerca de 1. En cuanto a los errores (desviación estándar) es muy claro que existe simetría, es decir los errores de los valores reportados por (Pfister *et al.*, 1998) y los calculados en el presente trabajo no presentan diferencias significativas.

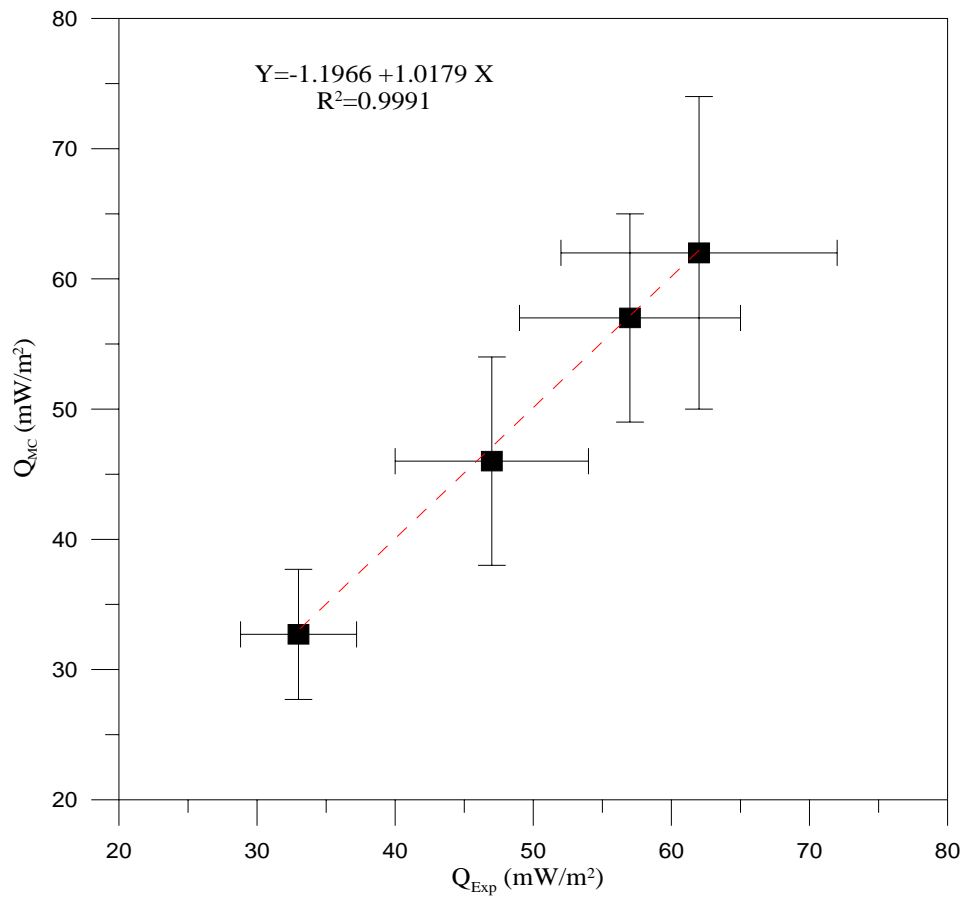


Figura 4.4 Comparación de flujos de calor reportados en la literatura por Pfister et al. (1998) vs. los calculados por el método de Monte Carlo.

Tabla 4.4 Validación de la metodología valores de flujo de calor conductivo medidos en el campo vs. Flujo de calor calculado por el método de MC.

Nombre del pozo	Ubicación Longitud / latitud	Γ	K	Estratigrafía	$Q_{Reportado}$	$Q_{Bevington}$	Q_{MC}
		$X \pm sd$ (°C/km)	$X \pm sd$ (W/m °C)		$X \pm sd$ (mW/m ²)	$X \pm sd$ (mW/m ²)	$X \pm sd$ (mW/m ²)
Karapurcek 1	26.9806° E / 41.2252° N	20.5 ± 0.6	2.3 ± 0.3	Oilgoceno	47 ± 8	47.1 ± 6.2	46 ± 7
Corlu 1	27.8759° E / 41.1005° N	27.1 ± 1.3	2.3 ± 0.3	Oilgoceno	62 ± 12	62 ± 9	62 ± 10
Maltepe1	26.7108° E / 41.0214° N	20.1 ± 1.7	3.1 ± 0.2	Eoceno	62 ± 10	62 ± 7	62 ± 8
Kadamis 1	27.2328° E / 41.1197° N	14.2 ± 0.2	2.3 ± 0.3	Oilgoceno	33 ± 5	32.6 ± 4.3	32.7 ± 4.2
Vakiflar 1	27.6633° E / 41.2650° N	24.6 ± 0.4	2.3 ± 0.3	Oilgoceno	57 ± 8	57 ± 7	57 ± 8

X: media, sd: desviación estándar.

4.4.2 Aplicación de la metodología al CGLH

La metodología descrita en la sección 4.3 se aplicó en el CGLH y a continuación se describen los resultados obtenidos. La aplicación se presenta en tres etapas: **I) Selección de los pozos conductivos, II) Propagación de errores en el FC, y III) Distribución en dirección vertical de FC.**

En el diagrama de flujo de la Figura 4.5 se presenta esquemáticamente el procedimiento usado en las tres etapas antes mencionadas. Con respecto a la selección de pozos con flujo de calor conductivo:

La Etapa I. Inicia con la creación de una base de datos de series de TFP y tiempo de recuperación térmica que incluye los pozos localizados en el área de colapso central del CGLH. Posteriormente se calculan las temperaturas estabilizadas de formación (TEF) con el método de Horner utilizando regresiones lineales y cuadráticas para evaluar el impacto de los métodos de cálculo de TEF sobre las estimaciones de FC. Con los datos de las TEF calculadas se graficaron los perfiles de temperaturas de cada pozo y a partir de ellas se seleccionaron siete pozos donde se presume que el flujo de calor es conductivo dado que su perfil es lineal.

En la Etapa II se centra la atención en el cálculo del flujo de calor y su incertidumbre. Para lograr lo anterior se requieren de: a) las TEF y sus errores de los pozos seleccionados; b) la ubicación de los puntos dentro de los pozos donde fueron medidas las TFP; c) los espesores de las capas o formaciones geológicas (perfiles estratigráficos de los pozos); y d) las propiedades termofísicas de las formaciones geológicas (conductividades térmicas presentadas en la Tabla 4.2). Con los datos anteriores y aplicando la metodología de propagación de errores mediante Montecarlo descrita en la sección 4.3 se obtuvieron los valores de flujo de calor y sus incertidumbres.

Como tercera y última etapa (Etapa III) se procedió a analizar los resultados de FC en cuanto a su distribución estadística y en dirección vertical. Para el análisis estadístico se aplicaron a los datos pruebas de datos desviados y la construcción de diagramas de distribución (histogramas). En cuanto a la distribución espacial solo se considera la ubicación de los datos en cuanto a su profundidad.

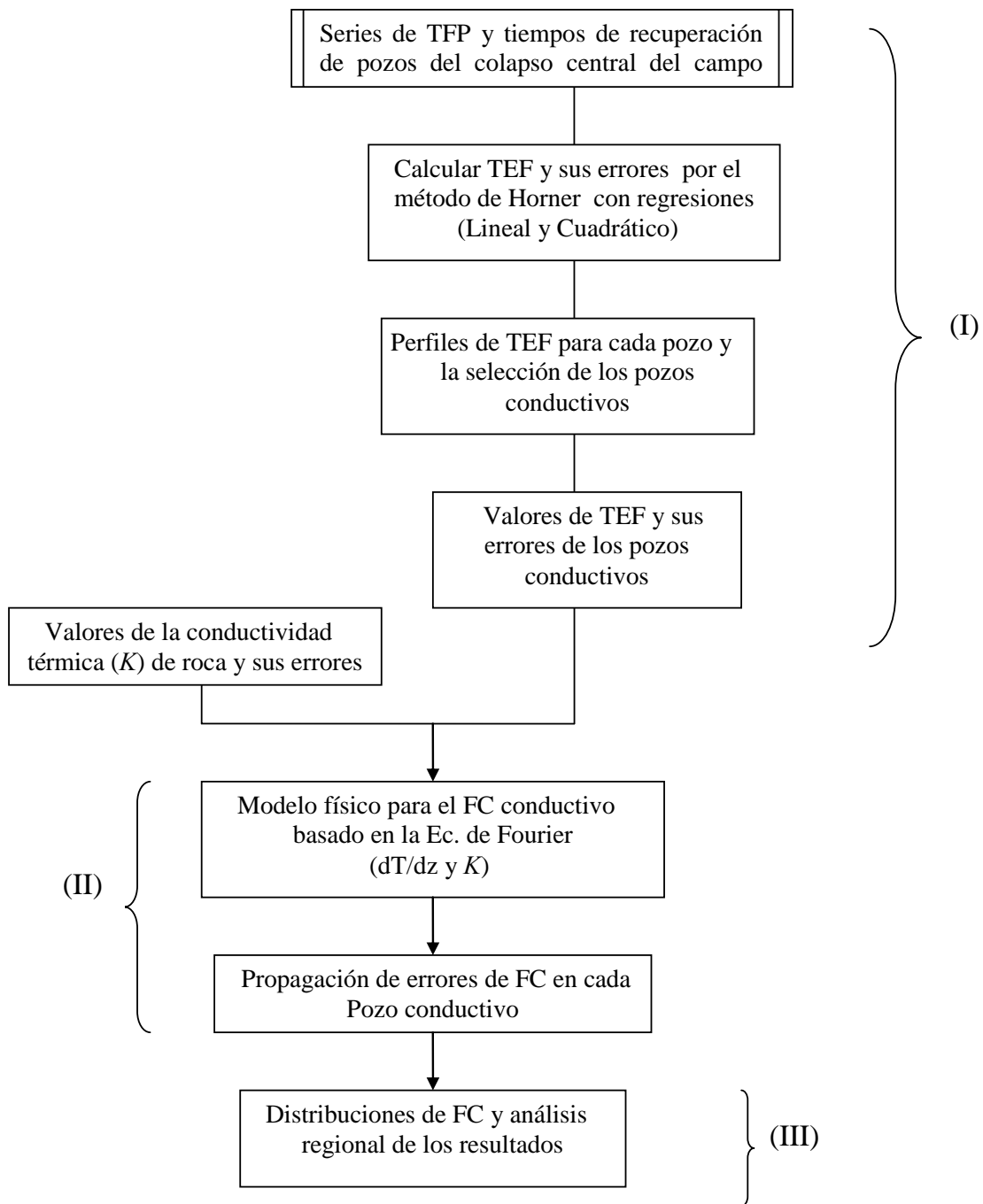


Figura 4.5 Propagación de errores del FC con el método de MC en el campo geotérmico de Los Humeros.

A continuación se presentan en el mismo orden de las etapas antes mencionadas los principales resultados obtenidos de la aplicación de esta metodología de propagación de errores y análisis de flujos de calor.

A) Selección de los pozos conductivos (Etapa I): La base de datos de de TFP y tiempo de recuperación térmica del CGLH esta formada de 39 series. Cada serie contiene al menos tres registros de TFP y en casos excepcionales llega hasta siete, teniendo en promedio 247 series. El total de pozos con registros en el colapso central es de 12 pozos, de los cuales se seleccionaron siete (H11, H16, H20, H28, H29, H31, H32).

Con el método de Horner (descrito en la sección 4.3) se calcularon las temperaturas estabilizadas de formación en los pozos en donde se ha presumido que el mecanismo de transporte de calor es conductivo y en la Tabla 4.5 se presentan los resultados obtenidos. El método de Horner se realizó aplicando regresiones tanto lineales (LIN) como cuadráticas (CUA). Estos resultados se discutirán en las siguientes secciones (*B* y *C*).

Como se ha mostrado en los ejemplos de la Figura 4.2-(a, b y c), el perfil de temperatura de los pozos seleccionados tiene un comportamiento lineal, lo que hace suponer que el flujo de calor es conductivo.

B) Propagación de errores en el FC: A partir de las TEF y las propiedades termofísicas de las rocas se calcularon los flujos de calor y sus incertidumbres para los pozos seleccionados (Tabla 4.5) tanto para el modelo lineal como cuadrático del método MH. Cabe aclarar que los valores de flujo de calor calculados son entre dos puntos contiguos de medición de TFP en cada pozo.

Sin embargo, se realizaron cálculos de flujos de calor entre puntos tanto contiguos como no contiguos. A modo de ejemplo se presenta en la Tabla 4.6 los resultados obtenidos para el pozo H20. En este caso se tiene cuatro secciones de medición (A, B, C, D) de

Tabla 4.5 Valores de TEF y flujo de calor conductivo en diferentes profundidades de los pozos estudiados del colapso central.

Pozo	Prof. (m.s.n.m.)	*TEF ± Sd (°C)		Q _{cond.} ± Sd (mW/m ²)		Pozo	Prof. (m.s.n.m.)	*TEF ± Sd (°C)		Q _{cond.} ± Sd (mW/m ²)	
		CUA	LIN	CUA	LIN			CUA	LIN	CUA	LIN
28	1399	282.8 ± 4.1	244 ± 8			11	549	138 ± 5	111 ± 8		
	1685	330.7 ± 3.9	292 ± 8	260 ± 70	260 ± 90		1871	308 ± 13	284 ± 6	190 ± 120	200 ± 120
	1990	351 ± 7	349.6 ± 3.8	110 ± 70	310 ± 80	16	656	153.4 ± 1.6	122 ± 5		
	2283	360.8 ± 0.1	387 ± 6	59 ± 7	220 ± 80		1410	281 ± 7	245 ± 7	270 ± 60	260 ± 60
29	790	195.8 ± 2.4	175 ± 5			1735	312.5 ± 1.3	289 ± 5	160 ± 50	220 ± 60	
	990	255 ± 12	218 ± 9	400 ± 200	320 ± 110	20	885	99 ± 7	83.2 ± 3.7		
	1187	270 ± 5	248 ± 5	120 ± 20	240 ± 60		1088	168.5 ± 1.1	142 ± 6	490 ± 190	420 ± 160
31	1037	176.4 ± 1.3	159.7 ± 3.6			1380	246.3 ± 3.4	212 ± 8	410 ± 100	370 ± 100	
	1238	222.2 ± 3.0	204 ± 5	330 ± 70	320 ± 70	2048	321 ± 7	283 ± 7	180 ± 50	170 ± 50	
	1390	237 ± 9	220 ± 5	390 ± 90	280 ± 80	32	997	205 ± 7	174 ± 7	340 ± 130	360 ± 130
	1600	311.1 ± 3.2	269 ± 7	600 ± 150	210 ± 60		1401	302 ± 6	275 ± 5		

*Para el cálculo de las TEF se utilizó el método de Horner (1951). Aproximación Lineal (LIN), aproximación cuadrática (CUA)

Tabla 4.6 Análisis de FC conductivo para el pozo H20.

Prof. (m.s.n.m.)	TEF ± Sd	Q _{cond.} ± Sd	Q _{tot.} ± Sd	TEF ± Sd	Q _{cond.} ± Sd	Q _{tot.} ± Sd
	(°C)	(mW/m ²)	(mW/m ²)	(°C)	(mW/m ²)	(mW/m ²)
	LIN	LIN		CUA	CUA	
A 885	83.2 ± 3.7			99 ± 7		
B 1088	142 ± 6	*Q _{AB} = 420 ± 160 *Q _{BC} = 370 ± 100	402±121	168.5 ± 1.1	*Q _{AB} = 490 ± 190 *Q _{BC} = 410 ± 100	442 ± 134
C 1380	212 ± 8	*Q _{AC} = 426 ± 130 Q _{BD} = 392 ± 111		246.3 ± 3.4	*Q _{AC} = 437 ± 120 Q _{BD} = 395 ± 130	
D 2048	283 ± 7	Q _{AD} = 399 ± 112 Q _{CD} = 170 ± 50		321 ± 7	Q _{AD} = 405 ± 173 Q _{CD} = 180 ± 50	

Aproximación Lineal (LIN), aproximación cuadrática (CUA)

los cuales se obtienen seis combinaciones (Q_{AB} , Q_{AC} , Q_{AD} , Q_{BC} , Q_{BD} , Q_{CD}) con igual número de FC. Aplicando pruebas de datos desviados (Verma, 2005) se observó que los valores de FC pertenecen a una misma población (370 – 426 mW/m^2 para lineal y 395 – 490 mW/m^2 para cuadrático), excepto por el valor Q_{CD} (170 y 180 mW/m^2 lineal y cuadrático respectivamente), tanto en el caso de flujos calculados a partir TEF obtenidas con regresiones lineales como cuadráticas. El hecho de que el valor de flujo de calor (Q_{CD}) no pertenece a la población se puede deber explicar a que los puntos C y D se encuentran probablemente en los límites del yacimiento, y que al menos uno de ellos se encuentra en un régimen convectivo. Como resultado de que el punto D se encuentra en el yacimiento, solo se toman en cuenta para el cálculo de flujo de calor los puntos A, B y C es decir Q_{AB} , Q_{AC} y Q_{BC} , obteniéndose en este caso $402 \pm 121 \text{ mW/m}^2$, $442 \pm 134 \text{ mW/m}^2$ como promedio total y desviación estándar del flujo de calor para este pozos según los modelos de regresión lineal y cuadrático, respectivamente (Tabla 4.6).

C) Distribución en dirección vertical del FC: Los valores de flujo presentados en la Tabla 4.5 se llevaron a un proceso de análisis estadístico iniciando con sus histogramas (Figura 4.6). En ellos se observa que para los flujos calculados con las TEF obtenidas con regresión cuadrática su distribución no es normal, más bien tiende a ser uniforme. Por otra parte, en el caso de los flujos calculados con las TEF obtenidas con regresión lineal su distribución es evidentemente normal o Gausiana. Lo anterior es significativo, dado que los flujos calculados con temperaturas inferidas por modelo de regresión lineal indican que su dispersión es reducida, lo que significa que el régimen de flujo de calor conductivo es homogéneo, por otra parte las aproximaciones obtenidas de los cálculos de FC usando temperaturas estabilizadas inferidas por modelo de regresión cuadrática indican la existencia de diferentes regiones de flujos en relación con la profundidad, es decir el flujo de calor presenta estratificación.

Con la finalidad de comparar los datos de flujo de calor en cuanto a su distribución en dirección vertical se construyeron las distribuciones de la Figura 4.6. En ellas se observa que para los flujos calculados usando temperaturas estabilizadas inferidas por modelo de regresión cuadrática la dispersión de los valores medios estimados es alta, mientras que para los flujos calculados con TEF inferidos por el modelo de regresión lineal su dispersión es baja. Por otra parte se muestra que para el caso de flujos calculados usando temperaturas estabilizadas inferidas por modelo de regresión cuadrática tienden a ser menores al

régimen de transferencia de calor. En cambio para los flujos calculados usando temperaturas estabilizadas inferidas por modelo de regresión lineal su reducción con la profundidad no es pronunciada, es decir no es observable el cambio en los valores de FC. Siendo evidente que en la parte poco profunda donde la transferencia de calor es conductiva ($\sim 800 - 1600$) y en la parte profunda (>1600) domina la convección.

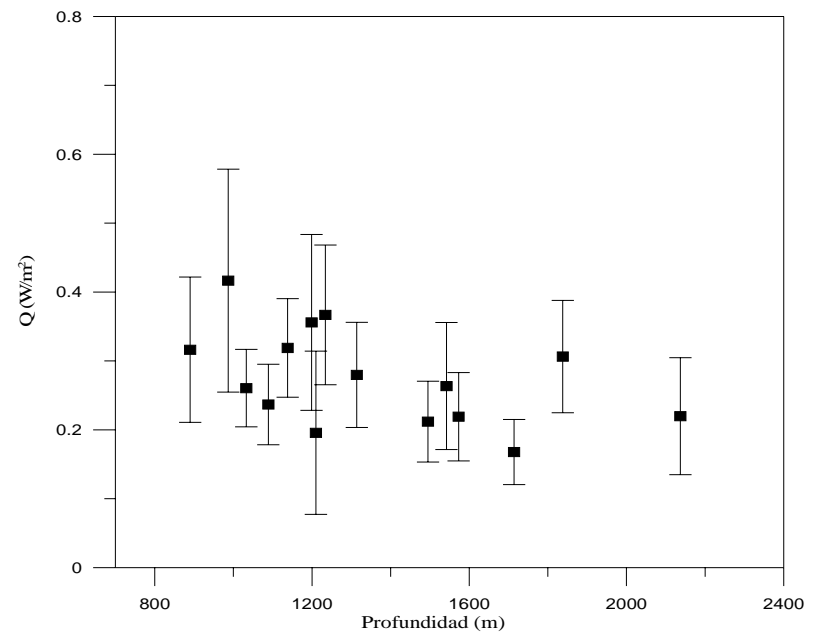
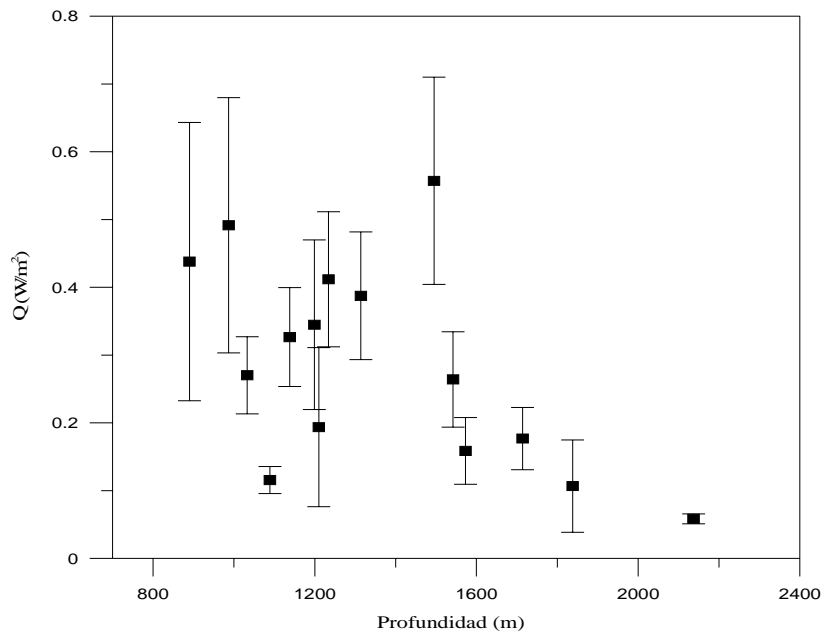
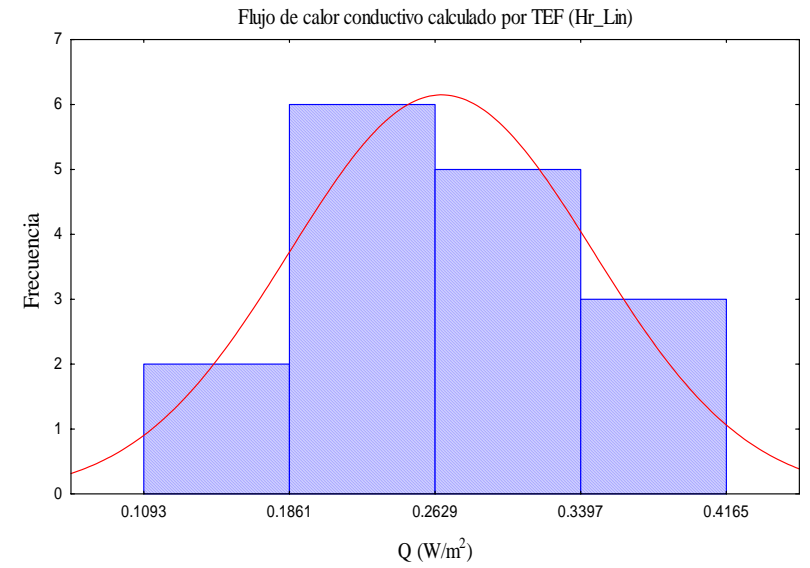
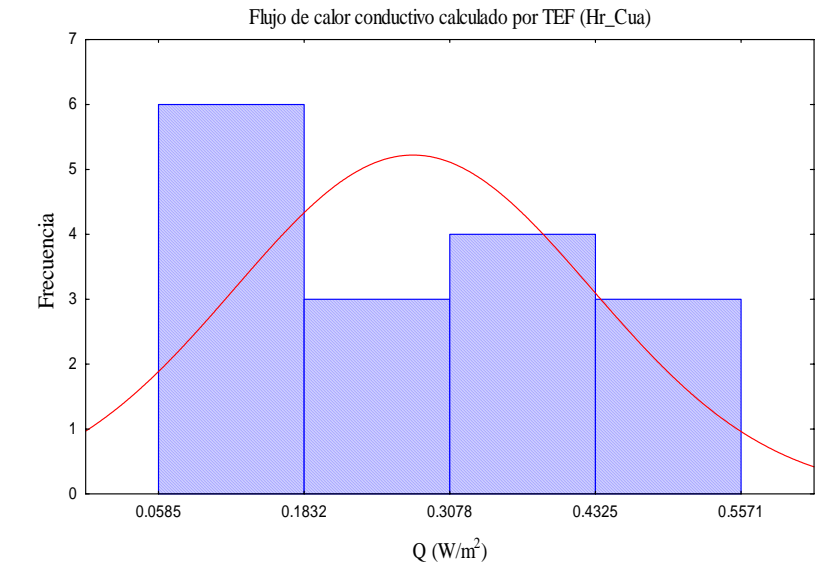


Figura 4.6 Distribuciones de FC calculado con MC usando las TEF estimadas con el método de Horner con regresiones (lineales y cuadráticas).

Conclusiones

Las principales conclusiones del presente trabajo se presentan a continuación en cuatro bloques referentes a: a) temperaturas estabilizadas de formación, b) perfil de presión en pozos geotérmicos productivos, c) estimación de flujo de calor conductivo y, d) la contribución del presente trabajo en la industria geotérmica.

Con respecto a

A) Temperaturas estabilizadas de formación

A1.- Se desarrolló exitosamente por primera vez en la literatura mundial, una nueva ecuación para la estimación de las TEF en pozos geotérmicos utilizando RNA su aplicación requiere como datos iniciales las TFP y sus tiempos de recuperación térmica. Esta característica la convierte en una herramienta ventajosa sobre los actuales métodos analíticos que dependen de otras variables complejas y de difícil acceso, (p. ej., el tiempo de circulación del lodo de perforación). La RNA se entrenó con una base de datos experimental y se evaluó con series (pruebas) de validación. La validez de las TEF calculadas por el modelo fue confirmada a través de una comparación estadística de los valores medidos y simulados (errores de $\pm 5\%$).

A2.- La nueva ecuación se puede utilizar como herramienta práctica para una determinación confiable de TEF en pozos geotérmicos, especialmente cuando los registros de TF están dentro del intervalo de entrenamiento (de 45 °C a 263 °C). El intervalo de la aplicación de la RNA puede ser ampliado si se utiliza una base de datos más extensa. La precisión de la predicción también se podría mejorar mediante el uso de patrones de entrenamiento con un gran número de mediciones de TFP, así como sus incertidumbres.

A3.- La idea de aplicar la RNA para la estimación de TEF no pretende eliminar el uso o el desarrollo de los métodos complejos de transferencia de calor para calcular esta variable, sino para proporcionar una herramienta adicional práctica, rápida y económica para la industria geotérmica. Además el modelo desarrollado puede tener una implementación

como sensor inteligente para la determinación en tiempo real de TEF en los pozos geotérmicos durante la perforación.

b) Perfil de presión en pozos geotérmicos productivos

B1.- Se presentó una nueva aplicación de las técnicas de redes neuronales artificiales para la industria geotérmica. Se logró desarrollar con éxito dos arquitecturas optimizadas de RNA (RNA₁ y RNA₂) para una determinación confiable de los gradientes de presión dentro de los pozos geotérmicos. La estructura computacional de los modelos RNA₁ y RNA₂ se caracterizó por utilizar un número limitado de variables de producción del pozo (como, la geometría del pozo y los datos de producción: registros de presión, temperatura y flujo másico). El número de las neuronas en la capa oculta como los datos experimentales utilizados en el proceso de entrenamiento han sido optimizados para obtener modelos de RNA con predicción confiables con el objetivo de reducir el tiempo y el costo de las mediciones que se realizan en campo.

B2.- Los modelos desarrollados fueron capaces de predecir los gradientes de presión con una precisión aceptable (en especial los proporcionados por RNA₂ con errores alrededor de 2,3%). Los gradientes de presión predichos por ambos modelos de la RNA y el simulador GEOWELLS fueron estadísticamente comparados con datos experimentales de campo. A partir de este análisis, se encontró que las predicciones del modelo RNA₂ proporcionan resultados más aceptables en comparación con los resultados obtenidos con el modelo RNA₁ y el simulador GEOWELLS utilizando modelo de flujo homogéneo. El resultado del análisis de sensibilidad de las variables de entrada de ambos modelos de la RNA, muestran que el uso de la presión del fondo como variable de entrada adicional impactó en los resultados de predicción de estos modelos. Esta variable de entrada (que por lo general se registra durante las mediciones de campo), se considera como un parámetro clave en el proceso de entrenamiento y evaluación de las RNA.

B3.- Las técnicas de RNA y los simuladores numéricos de pozos también pueden ser utilizados como herramientas adecuadas para el estudio complejo de las predicciones de los gradientes de presión dentro de los pozos bajo diferentes condiciones del fluido y flujo de calor. Por lo tanto es importante recalcar la eficiencia de los modelos de RNA mejoran

en la predicción cuando se entrenan con una base de datos de mediciones mas grande y representativa. Bajo estas condiciones, los modelos de RNA constituyen, sin duda alguna, una herramienta práctica que puede tener una implementación física como sensores de medición de bajo costo para la estimación de los gradientes de presión dentro de los pozos geotérmicos productores.

c) Estimación de flujo de calor conductivo

C1.- Se han calculado temperaturas estabilizadas de formación las cuales se reportan. Estas temperaturas corresponden a pozos localizados en el colapso central del campo geotérmico de Los Humeros en donde se ha supuesto que el mecanismo de transferencia de calor dominante en la formación es el flujo de calor conductivo.

C2.- Una nueva metodología numérica estadística basada en la propagación de errores mediante el método Monte Carlo se ha desarrollado exitosamente y utilizado para calcular el flujo de calor en el CGLH, tanto con las TEF obtenidas por el método de Horner mediante regresión lineal como cuadrática.

C3.- Algo fundamental que resaltar es que la propagación se realizó tomando en cuenta la distribución de las principales capas geológicas en los pozos, es decir que se utilizó un modelo multicapas en donde tiene lugar la transferencia de calor.

C4.- Es evidente que el modelo de regresión dentro del método de Horner impacta en los resultados de flujo de calor calculados. Para los flujos de calor calculados con temperaturas estabilizadas de formación inferidas del modelo de regresión cuadrática se puede distinguir zonas o estratos según el régimen de transferencia de calor gobernante. De lo anterior se puede deducir que la regresión cuadrática propuesta por Andaverde *et al.* (2005) es fiable para evaluar recursos geotérmicos.

d) Contribución del presente trabajo en la industria geotérmica.

D1.- Las herramientas desarrolladas en el presente proyecto doctoral pueden ser utilizadas en la industria geotérmica para la estimación y optimización de los recursos geotérmicos,

además de contribuir al ahorro en los programas de explotación actuales y los que se pudieran desarrollar en sistemas geotérmicos mejorados.

D2.- Durante el desarrollo de este trabajo se ha observado que en la actual industria geotérmica no se tiene cuidado en la calidad de mediciones en cuanto a sus incertidumbres, por lo que se sugiere que en las mediciones físicas de campo se incluya el registro de ellas y se reporten a la comunidad científica y tecnológica, esto con el fin de tener una evaluación mas fiable de estos sistemas geotérmicos.

Apéndice A

Desde el punto de vista teórico, los gradientes de presión y temperatura se obtienen generalmente a partir de un balance de la energía mecánica del sistema involucrando la ecuación de continuidad y el balance del momentum apropiado. Según Currie (1993) la ecuación general de momentum es:

$$\rho \left[\frac{\partial \vec{v}}{\partial t} + (\vec{v} \cdot \nabla) \vec{v} \right] = -\nabla P + \nabla \cdot \vec{\tau} + \rho \vec{f} \quad (\text{A.1})$$

donde

$$\frac{\partial \vec{v}}{\partial t} + (\vec{v} \cdot \nabla) \vec{v} = \frac{\partial \vec{v}}{\partial t} + v_r \frac{\partial \vec{v}}{\partial r} + \frac{v_\theta}{r} \frac{\partial \vec{v}}{\partial \theta} + v_z \frac{\partial \vec{v}}{\partial z} \quad (\text{A.2})$$

$$\nabla \cdot \vec{\tau} = \frac{1}{c_1 c_2 c_3} \left[\frac{\partial}{\partial x_1} (c_2 c_3 \tau_{x_1 x_1}) + \frac{\partial}{\partial x_2} (c_1 c_3 \tau_{x_2 x_2}) + \frac{\partial}{\partial x_3} (c_1 c_2 \tau_{x_3 x_3}) \right] \quad (\text{A.3})$$

y

$$\begin{aligned} c_1 &= 1 & c_2 &= r & c_3 &= 1 \\ x_1 &= r & x_2 &= \theta & x_3 &= z \end{aligned}$$

Para las Ecuaciones (A.1-A.3), ρ es la densidad del fluido; v la velocidad; r, θ y z representan las coordenadas cilíndricas, C_1, C_2 y C_3 son constantes; t es el tiempo, \vec{f} representa las fuerzas por fricción; P es la presión y $\vec{\tau}$ es el esfuerzo cortante.

Si se representa el flujo del fluido geotérmico en la tubería en coordenadas cilíndricas, y tomando en cuenta que la componente de la dirección del eje vertical de la tubería es mucho mayor que en las otras direcciones, la ecuación del momentum (A.1) para la componente z en estado permanente, se representa de la siguiente manera:

$$\rho \left[\frac{\partial v_z}{\partial t} + v_r \frac{\partial v_z}{\partial r} + \frac{v_\theta}{r} \frac{\partial v_z}{\partial \theta} + v_z \frac{\partial v_z}{\partial z} \right] = -\frac{\partial P}{\partial z} + \frac{1}{r} \left[\frac{\partial}{\partial r} r \tau_{rz} + \frac{\partial}{\partial \theta} \tau_{\theta z} + \frac{\partial}{\partial z} r \tau_{zz} \right] + \rho g \sin \theta \quad (\text{A.4})$$

Una solución simplificada de esta ecuación aceptada para el caso de un pozo geotérmico se tiene bajo consideración de las siguientes suposiciones:

- a. Flujo estable y unidimensional. Asumiendo que la porción de flujo que viaja por la tubería es corta y que el flujo alcanza las condiciones estables relativamente rápido. Así mismo, se asume que la dirección predominante del flujo es vertical.
- b. Las fases se encuentran en equilibrio termodinámico, es decir las fases de vapor y de líquido se consideran que se encuentran a la misma temperatura.
- c. Las propiedades del fluido permanecen constantes dentro de un intervalo.
- d. En el caso del último término ($\rho g \text{sen}\theta$) se utiliza $\text{sen}\theta$ cuando se refiere al ángulo de inclinación con respecto a la horizontal y $\text{cos}\theta$ cuando el ángulo de inclinación es con respecto a la vertical.

Lo cual implica lo siguiente:

$$\frac{\partial v_z}{\partial t} = 0, \quad v_r = v_\theta = 0, \quad y \quad \tau_{\theta z} = \tau_{zz} = 0$$

Considerando las condiciones anteriores, la ecuación A.4 puede expresarse como:

$$\rho v_z \frac{\partial v_z}{\partial z} = -\frac{\partial P}{\partial z} + \frac{1}{r} \frac{\partial}{\partial r} r \tau_{rz} + \rho g \text{sen}\theta \quad (\text{A.5})$$

Donde, g es la aceleración debida a la gravedad, y θ representa el ángulo de inclinación del pozo. La ecuación del momentum, comúnmente usada para estimar los perfiles de presión, se puede reescribir como una ecuación explícita simplificada en la siguiente forma (Wallis, 1969):

$$\left[\frac{dP}{dz} \right]_T = \left[\frac{dP}{dz} \right]_f + \left[\frac{dP}{dz} \right]_a + \left[\frac{dP}{dz} \right]_g \quad (\text{A.6})$$

El término del lado izquierdo representa el gradiente de presión total (Ecuación A.6.1) de un pozo o tubería y los tres términos del lado derecho representan las componentes de

fricción (Ecuación A.6.2), aceleración (Ecuación A.6.3) y gravedad (Ecuación A.6.4) del gradiente de presión, respectivamente, es decir,

$$\left[\frac{\partial P}{\partial z} \right]_T = \frac{\partial P}{\partial z} \quad \text{Gradiente de presión total} \quad (\text{A.6.1})$$

$$\left[\frac{\partial P}{\partial z} \right]_f = -\frac{1}{r} \frac{\partial}{\partial r} r \tau_{rz} \quad \text{Gradiente de presión por fricción} \quad (\text{A.6.2})$$

$$\left[\frac{\partial P}{\partial z} \right]_a = -\rho v_z \frac{\partial v_z}{\partial z} \quad \text{Gradiente de presión por aceleración} \quad (\text{A.6.3})$$

$$\left[\frac{\partial P}{\partial z} \right]_g = -\rho g \text{sen} \theta \quad \text{Gradiente de presión gravitacional} \quad (\text{A.6.4})$$

Tomando en cuenta las condiciones de frontera: $\tau_{rz} = \tau_m$ y haciendo algunas simplificaciones, se tiene:

$$\left[\frac{dP}{dz} \right]_f = -\frac{1}{r} \frac{\partial}{\partial r} r \tau_{rz} = -\frac{P_{er}}{A} \tau_m \quad , \quad \left[\frac{dP}{dz} \right]_a = -\frac{A \rho v_z}{A} \frac{\partial v_z}{\partial z} = -\frac{W}{A} \frac{dv_z}{dz} \quad \text{y}$$

$$\left[\frac{dP}{dz} \right]_g = -\rho_m g \text{sen} \theta$$

donde P_{er} es el perímetro de la tubería, A es la sección transversal ; τ_m es el esfuerzo mixto, y W es el flujo másico. Sustituyendo estos términos en la ecuación A.6, se tiene finalmente:

$$\left[\frac{dP}{dz} \right]_T = -\frac{P_{er}}{A} \tau_m - \frac{W}{A} \frac{dv}{dz} - \rho_m g \text{sen} \theta \quad (\text{A.7})$$

Según Hasan y Kabir (2010), la Ecuación A.7 puede ser expresada en términos del factor de fricción, las densidades y velocidades de fluido:

$$-\left(\frac{dP}{dz} \right)_T = \rho_m f \frac{v^2}{2D} + \rho_m v \frac{dv}{dz} + \rho_m g \text{sen} \theta \quad (\text{A.8})$$

donde D representa el diámetro del pozo. Esta ecuación unidimensional generalmente ha sido usada por numerosos investigadores para describir las características del flujo bifásico

en pozos geotérmicos (Wallis, 1969; Ambastha y Gudmundsson, 1986; Tian y Finger, 2000; Hasan y Kabir, 2010; entre otros).

Apéndice B

Las formulas matemáticas usadas para la descripción del fluido dentro del pozo geotérmico están descritas por Garcia-Valladares *et al.* (2006). A partir de la ecuación 2.2 el variable f corresponde al factor de fricción de Darcy. Para un régimen de flujo monofásico, f se calcula con la siguiente ecuación:

$$f = \frac{64}{Re} \quad (\text{B.1})$$

donde, Re es el número de Reynolds:

$$Re = \frac{\rho v D}{\mu} \quad (\text{B.2})$$

donde μ es la viscosidad del fluido. La ecuación (B.1) corresponde a régimen de flujo laminar ($Re \leq 2300$).

Para un régimen de flujo transitorio ($2300 < Re \leq 4000$), el factor de fricción se expresa con lo siguiente:

$$f = 10^{[260.67 - 228.62 Re + 66.307 Re^2 - 6.3944 Re^3]} \quad (\text{B.3})$$

Mientras para un flujo turbulento ($Re > 4000$), el factor de fricción esta dado por la correlación de Colebrook (Colebrook, 1939). Esta ecuación es comúnmente utilizada para resolver el factor de fricción de Darcy–Weisbach f , y se presenta como lo siguiente:

$$\frac{1}{\sqrt{f}} = 2 \log \left(\frac{D}{e} \right) + 2 \log \left(\frac{9.34}{\left(\frac{e}{D} \right) Re \sqrt{f}} \right) + 1.14 \quad (\text{B.4})$$

donde e es la rugosidad absoluta de la tubería y D es su diámetro.

Para el flujo bifásico, el factor de fricción (f) es calculado con las correlaciones del flujo monofásico (Ecuaciones B.1-B.3) multiplicados por un factor de corrección (ϕ), y eso según Beattie (1973):

$$\phi = \left[1 + x \left(\frac{\rho_l}{\rho_s} - 1 \right) \right]^{0.8} \left\{ 1 + x \left[\frac{3.5 \mu_s + 2 \mu_l}{(\mu_s + \mu_l) \rho_s} - 1 \right] \right\}^{0.2} \quad (\text{B.5})$$

donde μ_l , μ_s , ρ_l y ρ_s son las viscosidades y las densidades de líquido y el vapor saturado, respectivamente.

Referencias

- Abramowitz, M., Stegun, I. A., 1972. Handbook of mathematical functions. Dover Publications, Inc., New York.
- Aguilar-y-Vargas, V.H., Verma, S.P., 1987. Composición química (elementos mayores) de los magmas en el Cinturón Volcánico Mexicano. *Geofísica Internacional* 26(2), 195-272.
- Álvarez del Castillo, A., Santoyo, E., García-Valladares, O., 2011. A new empirical void fraction correlation inferred from artificial neural networks for modeling two-phase flow in geothermal wells. *Comput. Geosci.* 41, 25-39.
- Ambastha, A.K., Gudmundsson, J.S., 1986. Pressure profiles in two-phase geothermal wells: comparison of field data and model calculations. *Proceedings 11th Workshop on Geothermal Reservoir Engineering*. Stanford, California, pp. 183-188.
- Andaverde, J., Verma, S.P., Santoyo, E., 2005. Uncertainty estimates of static formation temperature in borehole and evaluation of regression models. *Geophysical Journal International* 160, 1112-1122.
- Andaverde, J., 2007. Simulación numérica en 3-D de temperaturas y elementos químicos mayores en el campo geotérmico de Los Humeros, Puebla. Tesis Doctoral, Instituto de Geofísica, UNAM, 162 p.
- Aragón, A., Moya, S.L., García-Gutiérrez, A., Arellano, V. 2008. A new method of damage determination in geothermal wells from geothermal inflow with application to Los Humeros, Mexico. *Geof. Int.* 47, 371-382.
- Ascencio, F., García, A., Rivera, J., Arellano, V., 1994. Estimation of undisturbed formation temperatures under spherical radial heat flow conditions. *Geothermics* 23, 317-326.
- Axelsson, G., 2010. Sustainable geothermal utilization - case histories; definitions; research issues and modelling. *Geothermics* 39, 283-291.
- Barelli, A.R., Corsi Del Pizzo, G., Scali, C., 1982. A two-phase flow model for geothermal wells in the presence of non-condensable gas. *Geothermics* 11, 175-191.
- Barragán, R.M., Nieva, D., Santoyo, E., González, E., Verma, M., López-Mendiola, J.M., 1991. Geoquímica de fluidos del campo geotérmico de Los Humeros, Puebla (México). *Geotermia, Revista Mexicana de Geoenergía* 7: 23-48.
- Bassam, A., Santoyo, E., Andaverde, J., Hernández, J.A., Espinoza-Ojeda, O.M., 2010. Estimation of static formation temperatures in geothermal wells by using an artificial neural network approach. *Comput. Geosci.* 36, 1191-1199.
- Bayram, A.F., 2001. Application of an artificial neural network model to a Na/K geothermometer. *Journal of Volcanology and Geothermal Research* 112, 75-81.
- Beattie, D.R.H., 1973. A note of the calculation of two-phase pressure losses. *Nucl. Eng. Des.* 25, 395-402.
- Bertani, R., 2005. World geothermal power generation in the period 2001-2005: *Geothermics* 34, 651-690.
- Bertani, R., 2010. Geothermal Power Generation in the World 2005–2010 Update Report. *Proceedings of the World Geothermal Congress 2010, Bali, Indonesia*, 41 p.
- Bienkowski, R., Torres-Alvarado, I.S., Hinderer, M., 2005. Geochemical modeling of acid fluids in Los Humeros geothermal field, México, in *Proceedings World Geothermal Congress 2005, Antalya, Turkey*.
- Bjornsson, G.A., 1987. Multi-feedzone geothermal wellbore simulator. M. Sci. Thesis, Lawrence Berkeley Laboratory. University of California, 102 p.

- Bjornsson, G.A., Bodvarsson, G.S., 1987. A Multi-feedzone wellbore simulator. *TrGRC* 11, 503-507.
- Brennand, A.W., 1984. A new method for the analysis of static formation temperature test. In: *Proceedings 6th New Zealand Geothermal Workshop*, New Zealand, pp. 45-47.
- Campos-Enríquez, J.O., Durán, M.F., 1986. Determinación preliminar del campo de temperaturas en Los Humeros, Pue. *Geotermia Revista Mexicana de Geoenergía* 2, 141-152.
- Campos-Enríquez, J.O., Arredondo-Fragoso, J.J., 1992. Gravity study of Los Humeros caldera complex, Mexico: Structure and associated geothermal system. *Journal of Volcanology and Geothermal Research* 49, 69-90.
- Campos-Enríquez, J.O., Domínguez-Méndez, F., Lozada-Zumaeta, M., Morales-Rodríguez, H.F., Andaverde-Arredondo, J., 2005. Application of the Gauss theorem to the study of silicic calderas: The calderas of La Primavera, Los Azufres and Los Humeros (Mexico). *Journal of Volcanology and Geothermal Research* 147, 39-67.
- Can, I., 2002. A new improved Na/K geothermometer by artificial neural networks. *Geothermics* 31, 751-760.
- Cao, S., Hermanrud, C., Lerche, I., 1988a. Estimation of formation temperature from bottom-hole temperature measurements: COST #1 well, Norton Sound, Alaska. *Geophysics* 53, 1619-1621.
- Cao, S., Lerche, I., Hermanrud, C., 1988b. Formation temperature estimation by inversion of borehole measurements. *Geophysics* 53, 979-988.
- Carlsaw, H. S., Jaeger, J. C., 1959. *Conduction of Heat in Solids*. Oxford Science Publications, Clarendon Press, Oxford, 2nd ed.
- Castillo-Román, J., Verma, S.P., Andaverde, J., 1991. Modelación de temperaturas bajo la caldera de Los Humeros, Puebla, México, en términos de profundidad de la cámara magmática. *Geofísica Internacional* 30 (3), 149-172.
- Cedillo-Rodríguez, F., 1999. Modelo hidrogeológico de los yacimientos geotérmicos de Los Humeros, Pue., México. *Geotermia, Revista Mexicana de Geoenergía* 15 (3), 159-170.
- Cedillo-Rodríguez, F., 2000. Geología del subsuelo del campo geotérmico de Los Humeros, Puebla, México. *Geotermia, Revista Mexicana de Geoenergía* 16 (1-3), 7-16.
- Cermák, V., 1971. Underground temperature and inferred climatic temperature of the past millennium. *Paleogeography, Paleoclimatol, Paleoecol* 10, 1-19.
- CFE, 2009. Comisión Federal de Electricidad, Generación de electricidad, SENER (en línea) México: actualización 12 de junio de 2009.
- Chapman, D. S., Furlong, K. P., 1992. Thermal state of the continental lower crust. In: Fountain, D. M., Arculus, R., Kay, R. W. (Eds.), *Continental Lower Crust*. *Developments in Geotectonics* 23, Elsevier, Amsterdam, pp. 179-199.
- Colebrook, C.F., 1939. Turbulent flow in pipes with particular reference to the transition region between the smooth and rough pipe laws. *J. Inst. Civil Eng.* 11, 133-156.
- Combs, J., 1980. Heat flow in the coso geothermal area, inyo county, California. *J.Geophys Res.* 85 (B5), 2411-2424.
- Contreras, E., Dominguez, B., Rivera, M.O., 1990. Mediciones petrofísicas en núcleos de perforación del campo geotérmico Los Humeros. *Geotermia, Revista Mexicana de Geoenergía* 6, 9-42.
- Cortés, O., Urquiza, G., Hernández, J.A., 2009. Optimization of operating conditions for compressor performance by means of neural network inverse. *Appl. Energ.* 86, 2487-2493.
- Currie, I.G., 1974. *Fundamental Mechanics of Fluids*, ed. Mc. Graw Hill, Inc., USA.
- Dahl-Jensen, D., Mosegaard, K., Gundestrup, N., Clow, G. D., Johnsen, S. J., Hansen, A. W., Balling, N., 1998. Past temperatures directly from the Greenland ice sheet. *Science* 282, 268-271.

- Dalkilic, A.S., Laohalertdecha, S., Wongwises, S., 2009. Effect of void fraction models on the two-phase friction factor of R134a during condensation in vertical downward flow in a smooth tube. *Int. Commun. Heat Mass Transf.* 35, 921-927.
- Davis, A.P., Michaelides, E.E., 2009. Geothermal power production from abandoned oil wells. *Energy* 34, 866–872.
- Da-Xin, L., 1986. Non-linear fitting method of finding equilibrium temperature from BHT data. *Geothermics* 15, 657-664.
- Demirbas, A., 2006. Global renewable energy resources. *Energy Sources* 28, 779-792.
- Demuth, H., Beale, M., 1998. *Neural Network Toolbox for use with Matlab, User's Guide Version 3*, MathWorks, Inc, USA.
- Demuth, H., Beale, M., 2003. *Neural Network Toolbox for use with Matlab, User's Guide Version 4.1*, MathWorks, Inc, USA.
- Díaz-González, L., Santoyo, E., Reyes, J., 2008. In Spanish: Tres nuevos geotermómetros mejorados de Na/K usando herramientas computacionales y geoquimiométricas: aplicación a la predicción de temperaturas de sistemas geotérmicos (“Three new improved geothermometers of Na/K using computational and geochemometric tools: Application to the prediction of temperatures in geothermal systems”). *Revista Mexicana de Ciencias Geológicas* 25, 465-482.
- Dickson, M.H., Fanelli, M., 2005. *Geothermal Energy: Utilization and Technology*. Earthscan Publications Ltd., 226 p.
- Dowdle, W.L., Cobb, W.M., 1975. Static formation temperature from well logs – an empirical method. *Journal of Petroleum Technology* 11, 1326 – 1330.
- Elmolla, E. S. and Chaudhuri, M. and Eltoukhy, M. M., 2010. The use of artificial neural network (ANN) for modeling of COD removal from antibiotic aqueous solution by the Fenton process. *J. Hazard. Mater.* 179, 127-134.
- Espinosa, G., Garcia, A., Santoyo, E., Hernandez, I., 2001, TEMLOPI/V.2: a computer program for estimation of fully transient temperatures in geothermal wells during circulation and shut-in. *Comput. Geosci.* 27, 327-344.
- Espinosa-Paredes, G., Espinosa-Martinez, E.G., 2009. A feedback-based inverse heat transfer method to estimate unperturbed temperatures in wellbores. *Energy Conversion and Management* 50, 140-148.
- Espinoza-Ojeda, O.M., Santoyo, E., Andaverde, J., 2011. A new look at the statistical assessment of approximate and rigorous methods for the estimation of stabilized formation temperatures in geothermal and petroleum wells. *J. Geophys. Eng.* 8, 233.
- Farshad, F.F., Garber, J.D., Lorde, J.L., 2000. Predicting temperature profiles in producing oil wells using artificial neural networks. *Engineering Computations* 17, 735-754.
- Ferriz, H., 1985. Zoneamiento composicional y mineralógico en los productos eruptivos del entro volcánico de Los Humeros, Puebla, México. *Geofísica* 24, 97-157.
- Ferriz, H., Mahood, G.A., 1987. Strong compositional zonation in a silicic magmatic system: Los Humeros, Mexican Neovolcanic Belt. *Journal of Petrology* 28, 171-209.
- Flores, L.C., Álvarez, R., Singh, S.K., Urrutia, J., 1978. Aeromagnetic survey of Los Humeros caldera, Mexico. *Geofísica Internacional* 17, 415-428.
- Fomin, S., Chugunov, V., Hashida, T., 2003. Analytical modelling of the formation temperature stabilization during the borehole shut-in period. *Geophysical Journal International* 155, 469–478.
- Fomin, S., Hashida, T., Chugunov, V., Kuznetsov, A.V., 2005. A borehole temperature during drilling in a fractured rock formation. *International Journal of Heat and Mass Transfer* 48, 385-394.

- Freeston D.H., Hadgu T., 1988. Comparison of results from some wellbore simulators using a data bank. Proceedings 10th New Zealand Geothermal Workshop, pp. 299-305.
- Fridleifsson, I.B., 2006. Geothermal energy among the world's energy sources. Proceedings of the Workshop for Decision Makers on Geothermal Projects in Central America, organized by UNU-GTP and LaGeo in San Salvador, El Salvador, 10 p.
- García, A., Hernandez, I., Espinoza, G., Santoyo, E., 1998. TEMPLOPI: A thermal Simulator for estimation of drilling mud and formation temperatures during drilling of geothermal wells, *Comput. Geosci.* 42, 465-477.
- García, A., Ascencio, F., Espinosa, G., Santoyo, E., Gutiérrez, H., Arellano, V., 1999. Numerical modeling of high-temperature deep wells in the Cerro Prieto geothermal field, Mexico. *Geof. Int.* 38, 251-260.
- García-Valladares, O., Sánchez-Upton, P., Santoyo, E., 2006. Numerical modeling of flow processes inside geothermal wells: An approach for predicting production characteristics with uncertainties. *Energ. Convers. Manage.* 47, 1621-1643.
- García-Valladares, O., 2007. Numerical simulation and experimental validation of coiled adiabatic capillary tubes. *Appl. Therm. Eng.* 27, 1062-1071.
- Garg, S.K., Combs, J.A., 2002. Study of production/injection data from slim holes and large-diameter wells at the Okuaizu geothermal field, Tohoku, Japan. LLC Report, INEEL/EXT-02-01429, Idaho National Engineering and Environmental Laboratory, Bechtel BWXT, Idaho, USA, 257 p.
- Garg, S.K., Pritchett, J.W., Alexander, J.H., 2004. A new liquid hold-up correlation for geothermal wells. *Geothermics* 33, 795-817.
- Garson, G.D., 1991. Interpreting neural-network connection weights. *Artif. Int. Expert* 6, 7, 47-51.
- Garrido, D., 2010. La transición energética y el cambio climático. Creative Commons. Universidad de Zaragoza – EUITIZ. España.
- Gómez, G. P., 1999. Notas del curso Fundamentos de Redes Neuronales. Universidad de las Américas, Puebla, México.
- González-Morán, T., Suro-Pérez, V., 1982. Obtención de parámetros físicos de las estructuras principales localizadas en la caldera de los Humeros, Puebla. *Geofísica Internacional* 21, 57-68.
- González-Partida, E., García-Gutiérrez, A., Torres-Rodríguez, V., 1997. Thermal and petrologic study of the CH-A well from the Chipilapa – Ahuichapan geothermal area, El Salvador. *Geothermics* 26, 701-713.
- González-Partida, E., Tello-Hinojosa, E., Verma, M.P., 2001. Interacción agua geotérmica-manantiales en el campo geotérmico de Los Humeros, Puebla, México. *Ingeniería Hidráulica en México* 16, 185-194.
- Goutorbe, B., Lucazeau, F., Bonneville, A., 2006. Using neural networks to predict thermal conductivity from geophysical well logs. *Geophysical Journal International* 166, 115-125.
- Grant, M. A. Bixley, P. F., Donaldson, L. G., 1983. Internal flows in geothermal wells: their identification and effect on the wellbore temperature and pressure profiles. *Soc. Petrol. Eng. J.* 23, 168-176.
- Gunn, M.I.C., Freeston, H.D., Hadgu, T., 1992. Principles for wellbore simulator validation and calibration using matching analysis—I. Analytical techniques. *Geothermics* 21, 341-361.
- Gutiérrez, A.G., 2009. Estado térmico inicial del campo geotérmico de Los Humeros, Puebla, México. *Geotermia, Revista Mexicana de Geoenergía* 22 (1), 59-70.
- Gutiérrez-Negrín, L.C.A., Quijano-León, J.L., 2005. Update of Geothermics in Mexico: *Geotermia* 17, 21-30.

- Gutiérrez-Negrín, L.C.A., Maya-Gonzales, R., Quijano-León, J.L., 2010. Current Status of Geothermics in Mexico. Proceeding of World Geothermal Congress, Bali, Indonesia, 11 p.
- Hadgu, T., Freeston, D., 1990. A multi-purpose wellbore simulator. *TrGRC* 14, 1279-1286.
- Hadgu, T., Zimmerman, W.R., Bodvarsson, S.G., 1995. Coupled reservoir-wellbore simulation of geothermal reservoir behavior. *Geothermics* 24, 145-166.
- Haenel, R., Rybach, L., Stegena, L., (eds) 1988. *Handbook of Terrestrial Heat-Flow Density Determination*. Kluwer Academic Publishers, Dordrecht.
- Hagan, M.T., Menhaj, M., 1994. Training feed forward networks with the Marquardt algorithm. *IEEE Transactions on Neural Networks* 5, 989-993.
- Häring, M.O., Schanz, U., Ladner, F., Dyer, B.C., 2008. Characterisation of the Basel 1 Enhanced Geothermal System. *Geothermics* 37, 469-495.
- Hasan, A.R., Kabir, C.S., 1994. Static reservoir temperature determination from transient data after mud circulation. *Society of Petroleum Engineering, Drilling & Completion* 9, 17-24.
- Hasan, A.R., Kabir, C.S., 2007. A simple model for annular two-phase flow in wellbores. *Soc. Petrol. Eng.* 22, 168-175.
- Hasan, A.R., Kabir, C.S., 2010. Modeling two-phase fluid and heat flows in geothermal wells. *J. Petrol. Sci. Eng.* 71, 77-86.
- Haykin S., 1999. *Neural networks: A Comprehensive Foundation*, second ed. Prentice Hall, New Jersey, USA.
- Hebb, D. O., 1989. The first stage of perceptron: Growth and Assenbly. Aparece recopilado en *neurocomputing : foundation of research*, pp 45-56, MIT Press.
- Hermanrud, C., Cao, S., Lerche, I., 1990. Estimates of virgin rock temperature derived from BHT measurements: bias and error. *Geophysics* 55, 924-931.
- Hernández, A.L., 1995, Estudio regional volcánico y estructural del campo geotérmico de Los Humeros, Puebla, México: *Geotermia, Revista Mexicana de Geoenergía* 11 (2), 17-36.
- Hernández-Pérez, J.A., García-Alvarado, M.A., Trystram, G., Heyd, B., 2004. Neural networks for the heat and mass transfer prediction during drying of cassava and mango. *Food Science and Emerging Technologies* 5, 57-64.
- Herrera, A.P., Jiménez, M.M., 1999, Estudio y caracterización termodinámico del sector centro del campo geotérmico Los Humeros, Pue., México: *Geotermia, Revista Mexicana de Geoenergía*, 15 (3), 195-211.
- Hochstein. P. M., Bromley. J. C., 2005. Measurement of heat flux from steaming ground. *Geothermics* 34, 133-160.
- Hsieh, B.Z., Wang, C.W., Lin, Z.S., 2009. Estimation of formation strength index of aquifer from neural networks. *Comput. Geosci.* 35, 1933-1939.
- Huenges. E., Burkhardt. H., Erbas. K., 1990. Thermal conductivity profile of the KTB pilot borehole. *Sci. Drill.* 1, 224- 230.
- Hyodo, M., Takasugi, S., 1995. Evaluation of the curve-fitting method and the Horner plot for estimation of the true formation temperature using temperature recovery logging data. In: *Proceedings 20th Workshop on Geothermal Reservoir Engineering*, Stanford University, Stanford, USA, pp. 94-100.
- Iglesias, E.R., Campos-Romero, A., Torres, R.J., 1995. A reservoir engineering assessment of the Chipilapa, El Salvador, geothermal field. In: *Proceedings World Geothermal Congress*, Florence, Italy, pp. 1531-1536.
- Iglesias, E.R., Moya, S.L., 1998. Applicability of geothermal inflow performance reference curves to CO₂-bearing reservoirs, *Geothermics* 27, 305-315.
- Jahanandish, I., Salimifard, B., Jalalifar, H., 2011. Predicting bottomhole pressure in vertical multiphase flowing wells using artificial neural networks. *J. Petrol. Sci. Eng.* 75, 336-342.

- Jeirani, Z., Mohebbi, A., 2006. Estimating the initial pressure, permeability and skin factor of oil reservoirs using artificial neural networks. *J. Petrol. Sci. Eng.* 50, 11-20.
- Jiménez, M.M., Herrera, A.P., 2000. Estudio y caracterización termodinámica del sector noroeste del campo geotérmico Los Humeros, Puebla, México: *Geotermia, Revista Mexicana de Geoenergía* 16 (1), 71-82.
- Jokinen, J., 2000. Uncertainty analysis and inversion of geothermal conductive models using random simulation methods. Geological Survey of Finland, P.O. Box 96, FIN-02151 Espoo, Finland *Acta Univ. Oul. A* 343, Oulu, Finland.
- Jung, D.B., Wai, K.W., Howard, W.T., 2001. Geothermal flow metering errors. *TrGRC* 25, 23-25.
- Khasani, R.I., Tanaka, T., Fukuda, M., 2002. An analysis of pressure drops in wellbore under low flow rate conditions shown on a deliverability curve. *Proceedings of the 24th New Zealand Geothermal Workshop, New Zealand*, pp. 115-120.
- Kutasov, I.M., 1999. *Applied Geothermics for Petroleum Engineers*, Elsevier, Amsterdam 360 pp.
- Kutasov, I.M., Eppelbaum, L.V., 2003. Prediction of formation temperatures in permafrost regions from temperature logs in deep wells-field cases. *Permafrost and Periglacial Process* 14, 247-258.
- Kutasov, I.M., Eppelbaum, L.V., 2005. Determination of formation temperature from bottom-hole temperature logs — a generalized Horner method. *Journal of Geophysical Engineering* 2, 90-96.
- Laherrère, J., 2001. Estimates of Oil Reserves, paper presented at the EMF/IEA/IEW meeting in Laxenburg, Austria.
<http://www.iiasa.ac.at/Research/ECS/IEW2001/pdf/Papers/Laherrere-long.pdf>
- Lamontagne, M., Ranalli, G., 1996. Thermal and rheological constraints on the earthquake depth distribution in the Charlevoix, Canada, intraplate seismic zone. *Tectonophysics* 257, 55– 69.
- Leblanc, Y., Pascoe, L.J., Jones, F.W., 1981. The temperature stabilization of a borehole. *Geophysics* 46, 1301-1303.
- Leite, E.P., Filho, C.R. de S., 2009a. Probabilistic neural networks applied to mineral potential mapping for platinum group elements in the Serra Leste region, Carajás mineral province, Brazil. *Comput. Geosci.* 35, 675-687.
- Leite, E.P., Filho, C.R. de S., 2009b. TEXTNN – a MATLAB program for textural classification using neural networks. *Comput. Geosci.* 35, 2084-2094.
- Limin, F., 1995. *Neural Networks in Computer Intelligence*, McGraw-Hill, New York 460 pp.
- López-Romero, O., 2006. Actualización del modelo geoquímico de Los Humeros, Puebla, México: *Geotermia, Revista Mexicana de Geoenergía* 19(1), 23–32.
- Lu, X., Watson, A., Gorin, A.V., Deans, J., 2006. Experimental investigation and numerical modelling of transient two-phase flow in a geysiring geothermal well. *Geothermics* 3, 409-427.
- Lund, J.W., 2007. World geothermal generation in 2007: *Proceedings European Geothermal Congress* 11 p.
- Lund, J.W., Freeston, D.H., Boyd, T.L., 2010. Direct Utilization of Geothermal Energy 2010 Worldwide Review. *Proceedings of the World Geothermal Congress 2010, Bali, Indonesia*, 23 p.
- Majorowicz, J., Grasby, E. S., 2010. Heat flow, depth–temperature variations and stored thermal energy for enhanced geothermal systems in Canada. *J. Geophys. Eng.* 7, 232–241.
- Manetti, G., 1973. Attainment of temperature equilibrium in holes during drilling. *Geothermics* 2, 94-100.
- Martin, H.T., Hagan, T.M., Mohammad, N.B., 1994. Training feedforward networks with the Marquardt Algorithm. *IEEE Trans. Neural Netw.* 6, 989-993.

- Martin, H.T., 2006. *MATLAB Recipes for Earth Sciences*. Springer-Verlag Berlin Heidelberg, Netherlands.
- Martínez-Serrano, R.G., 2002. Chemical variations in hydrothermal minerals of the Los Humeros geothermal system, Mexico. *Geothermics* 31, 579-612.
- McCulloch, W.S., Pitts, W., 1987. A logical calculus of the immanent in nervous System. Recopilado en *Neurocomputing: Foundation of Research*, pp18-27, MIT Press.
- Medina-Martínez, M., 2000. Intervalos productores de fluidos en los pozos H-35 y H-39 del campo geotérmico Los Humeros, Puebla, México. *Geotermia, Revista Mexicana de Geoenergía* 16, 51-60.
- Metropolis, N., Ulam, S., 1949. The Monte Carlo Method. *J. Amer. Stat. Assoc.* 44, 335-341.
- Morin, R. H., Williams, T., Henrys, S. A., Magens, D., Niessen, F., Hansaraj, D., 2010. Heat Flow and Hydrologic Characteristics at the AND-1B borehole, ANDRILL McMurdo Ice Shelf Project, Antarctica. *Geosphere* 6, 370–378.
- Morton, J.C., 2009. Boosting a fast neural network for supervised land cover classification. *Comput. Geosci.* 35, 1280–1295.
- Neumann. V. J., Stanislaw. U., Metropolis. N., 1947. Statistical method en neutron diffusion. LAMS-551, Los Alamos National Laboratory.
- Ortíz-Ramírez J., 1983. Two-phase flow in geothermal wells: development and uses of a computer code. M.Sc. Thesis, University of Stanford, CA, USA, 89 p.
- Pachauri, R.K., Reisinger, A., 2007. Cambio climático 2007: Informe de síntesis. Contribución de los Grupos de trabajo I, II y III al Cuarto Informe de evaluación del Grupo Intergubernamental de Expertos sobre el Cambio. IPCC, Ginebra, Suiza.
- Pérez-Reynoso, J., 1979. Geografía y petrografía de la caldera de Los Humeros. *Geomimet* 91, 97-106.
- Pfister. M., Rybach. L., Simsek. S., 1998. Geothermal reconnaissance of the Marmara sea region (NW Turkey): surface heat flow density in an area of active continental extension. *Tectonophysics* 291, 77-89.
- Pollack, H. N., Henry, S. G., 1985. Heat flow in the presence of topography: numerical analysis of data ensembles. *Geophysics* 50 (8), pp. 1335-1341.
- Pollack, H.N., Hurter, S., Johnson, R. J., 1993. Heat flow from the Earth's interior — analysis of the global data set. *Rev.Geophys.* 31, 267–280.
- Press, W. H., Flannery, B. P., Teukolsky, S. A., Vetterling, W. T., 1990. *Numerical recipes in C, The art of science computing*. Cambridge University Press, Cambridge.
- Pribnow, D. F. C., Sass. J. H., 1995. Determination of thermal conductivity for deep boreholes. *J. Geophys. Res.* 100, 9981-9994.
- Prol, R.M., González-Morán, T., 1982. Modelo preliminar del régimen térmico conductivo en la caldera de Los Humeros, Puebla. *Geofísica Internacional* 21, 295-307.
- Rojas, R., 1995. *Neural Networks: A Systematic Introduction*. Springer Verlag .
- Rosenblatt, F., 1989. The Perceptron A Probabilistic Model for Information Storage and Organization in the Brain. *Neurocomputing: fondation research*, pp 92'113, MIT Press.
- Royer, J. J., Danis, M., 1988. Steady state geothermal model of the crust and the problem of the boundary conditions: application to a rift system, the southern Rhinegraben. *Tectonophysics* 156, 239–255.
- Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning internal representations by error propagation. *Parallel Distributed Processing* 1, 318-362.
- Rumethart, D.E., Hinton, .G.E., Williams, R.J., 1986. Learning representations by backpropagation errors. *Nature*, 323.

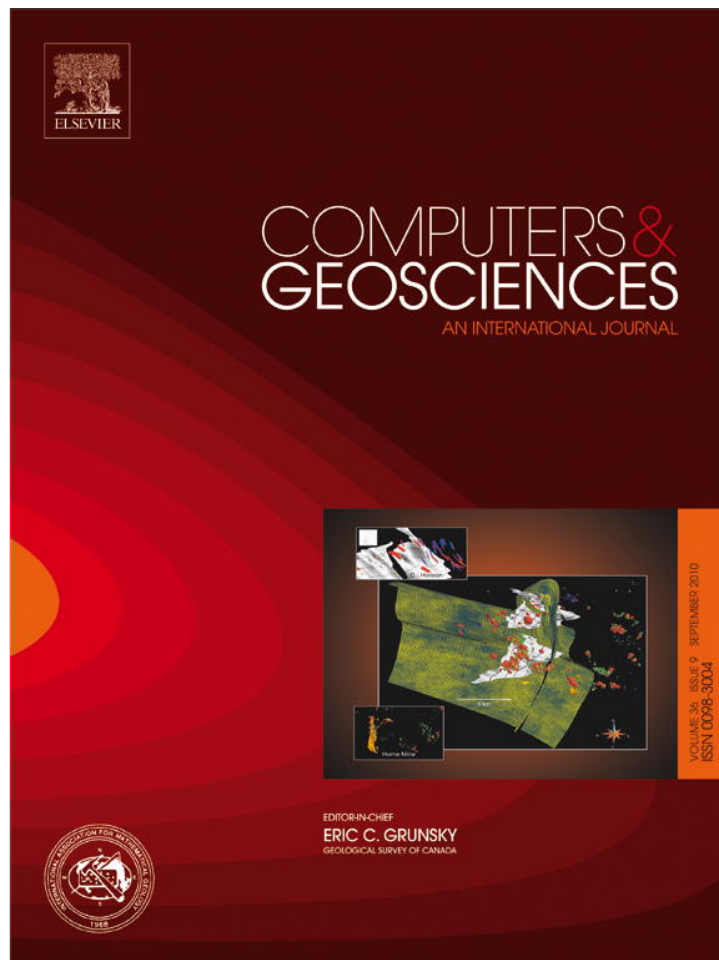
- Saito, S., Sakuma, S., Uchida, T., 1998. Drilling procedures, techniques and test results for a 3.7 km deep, 500°C exploration well, Kakkonda, Japan. *Geothermics* 27, 573-590.
- Sánchez-Upton, P., 2000. The wellbore simulator SIMU2000. In: *Proceedings of the World Geothermal Congress (PaperR0057)*, Kyushu-Tohoku, Japan, pp. 2851-2856.
- Santoyo, E., García, A., Espinosa, G., Hernández, I., Santoyo, S., 2000. STATIC_TEMP: A useful computer code for calculating static formation temperatures in geothermal wells. *Comput. Geosci.* 26, 201-217.
- Santoyo E., Barragán-Reyes, R.M., 2010. Energía geotérmica. En: *Volumen Especial sobre Energías Alternativas (abril-junio 2010) Ciencia, Revista de la Academia Mexicana de Ciencias* 61, 40-51.
- Santoyo E., Torres-Alvarado, I.S., 2010. Escenario futuro de explotación de la energía geotérmica: hacia un desarrollo sustentable. *Revista Digital Universitaria [en línea]*, Vol. 11, No. 10. Disponible en Internet: <http://www.revista.unam.mx/vol.11/num10/art95/index.html>, ISSN: 1607- 6079.
- Santoyo, E., 2011. Geotermia: Energía Renovable Inagotable del Futuro. XXI Congreso Nacional de Geoquímica, Monclova, México, 2 p.
- Sari. M., Karpuz. C., Ayday. C., 2010. Estimating rock mass properties using Monte Carlo simulation: Ankara andesites. *Computers & Geosciences* 36, 959–969.
- Schalkoff, R. J., 1997. *Artificial neural Networks*. McGraw , N.Y. USA.
- Schoepfel, R.J., Gilarranz, S., 1966. Use of well log temperatures to evaluate regional geothermal gradients. *Journal of Petroleum Technology* 18, 667-673.
- Serpen, G., Palabiyik, Y., Serpen, U., 2009. An artificial neural network model for Na/K geothermometer. In: *Proceedings 34th Workshop on Geothermal Reservoir Engineering*, Stanford University, Stanford, USA, pp. 1-12.
- Serrano, J. A., Soria, J. E., Martín, J. D., 2009. *Cursos Redes Neuronales artificiales*. Escuela Técnica Superior de ingeniería. Universidad de Valencia, España.
- Shen, P.Y., Beck, A.E., 1986. Stabilization of bottomhole temperature with finite circulation time and fluid flow. *Geophysical Journal of the Royal Astronomical Society* 86, 63-90.
- Spichak, V.V., Popova, I., 2000. Artificial neural network inversion of magnetotelluric data in terms of three-dimensional earth macroparameters. *Geophys. J. Int.* 142, 15-26.
- Spichak, V.V., Goidina, A.G., 2005. Temperature prediction in geothermal zones from borehole measurements using neural networks. *Phys. Solid Earth* 41, 844–852.
- Spichak, V.V., 2006. Estimating temperature distributions in geothermal areas using a neuronet approach. *Geothermics* 35, 181-197.
- Springer, M., Forster, A., 1998. Heat-flow density across the Central Andean subduction zone. *Tectonophysics* 291, 123–139.
- Springer, M., 1999. Interpretation of heat-flow density in the Central Andes. *Tectonophysics* 306, 377–395.
- Steingrímsson, B., Gudmundsson, A., 2006. Geothermal borehole investigations during and after drilling. In: *Proceedings at Workshop on Decision Makers on Geothermal Projects in Central America*, San Salvador, El Salvador, pp. 1-10.
- Takacs, G., 2001. Considerations on the selection of an optimum vertical multiphase pressure drop prediction model for oil wells. *Soc. Petrol. Eng. (Conference Paper)* DOI: 10.2118/68361-MS
- Tello-Hinojosa, E., 1992. Composición química de la fase líquida a descarga total y a condiciones de reservorio de pozos geotérmicos de Los Humeros, Puebla, México. *Geofísica Internacional* 31, 383-390.
- Tester, J., Blackwell, D., Petty, S., Richards, M., Moore, M., Anderson, B., Livesay, B., Augustine, Ch., DiPippo, R., Nichols, K., Veatch, R., Drake, E., Toksoz, N., Baria, R.,

- Batchelor, A.S., Garnish, J., 2007. The future of geothermal energy: an assessment of the energy supply potential of engineered geothermal systems (EGS) for the United States. Proceedings of the Thirty- Second Workshop on Geothermal Reservoir Engineering, Stanford University, Stanford, California, USA, SGP-TR-183, 1 p.
- Tian, F.S., Finger, T.J., 2000. Advanced geothermal wellbore hydraulics model. *J. Energ. Resour-ASME*. 122, 142-146.
- Timko, D. J., Fertl, W. H., 1972. How down-hole temperatures, pressures affect drilling, part 5, Predicting hydrocarbon environments with wireline data. *World Oil* 174, 73-88.
- Verma, S.P., 1985. On the magma chamber characteristics as inferred from surface geology and geochemistry: examples from Mexican geothermal areas: *Physics of the Earth and Planetary Interiors*, 41, 207-214.
- Verma, S.P., 2005. In Spanish: Estadística Básica para el Manejo de Datos Experimentales con la Aplicación en la Geoquímica - Geoquimiometría (Basic Statistics for Experimental Data Reduction with Application in Geochemistry – Geochemometrics). Universidad Nacional Autónoma de México, México, D.F., 186 pp.
- Verma, S.P., Andaverde, J., Santoyo, E., 2006a. Application of the error propagation theory in estimates of static formation temperatures in geothermal and petroleum boreholes. *Energy Conversion and Management* 47, 3659-3671.
- Verma, S.P., Andaverde, J., Santoyo, E., 2006b. Statistical evaluation of methods for the calculation of static formation temperatures in geothermal and oil wells using an extension of the error propagation theory. *Journal of Geochemical Exploration* 89, 398-404.
- Verma, S.P., Andaverde, J., 2007. Coupling of Thermal and Chemical Simulations in a 3-D Integrated Magma Chamber-Reservoir Model. *A New Energy Research Frontier*, Edited Book *Geothermal Energy Research Frontiers*, 42 p.
- Verma, S.P., Díaz-González, L., González-Ramírez, R., 2009. Relative efficiency of single-outlier discordancy tests for processing geochemical data on reference materials. *Geostandards and Geoanalytical Research* 33, 29-49.
- Viggiano, J.C.G., Camacho, J.R., 1998. Mineralogía hidrotermal en el campo geotérmico de Los Humeros, Pueb. I: sus usos como indicadores de temperatura y del régimen hidrológico. *Geotermia, Revista Mexicana de Geoenergía* 4 (1), 15-28.
- Wallis, G.B., 1969. *One-dimensional Two-Phase Flow*. ed. McGraw-Hill, USA.
- Werbos, P., 1974. *Beyond Regression: New Tools for Predicting and Analysis in the Behavioral Sciences*. Harvard University, Ph.D. Thesis, Reprinted by Willey & Sons, 1995.
- Widrow, B., Hoff, M.E., 1989. *Adaptive Switching Circuits*. Neurocomputing: foundation research, pp 126-134, MIT Press.
- Wisian, K.W., Blackwell, D.D., Bellani, S., Henfling, J.A., Normann, R.A., Lysne, P.C., Forster, A., Schrotter, J., 1998. Field comparasion of conventional and new technology temperature logging systems. *Geothermics* 27, 131-141.
- Wittwer, J.W., 2004. Monte Carlo Simulation Basics. From Vertex42.com, <http://www.vertex42.com/ExcelArticles/mc/MonteCarloSimulation.html>
- Yilmaz, S., Demircioglu, C., Akin, S., 2002. Application of artificial neural networks to optimum bit selection. *Computers & Geosciences* 28, 261–269.
- Zellou, A.M, Ouenes, A., 2003. Chapter 26 Integrated fractured reservoir characterization using neural networks and fuzzy logic: Three case studies. *Dev. Petrol. Sci.* 51, 583-602.
- Zhang, G., Patuwo, B.E., Hu, M.Y., 1998. Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting* 14, 35–62.

Anexo A

Bassam, A., Santoyo, E., Andaverde, J., Hernández, J.A., Espinoza-Ojeda, O.M., 2010. Estimation of static formation temperatures in geothermal wells by using an artificial neural network approach. *Comput. Geosci.* 36, 1191-1199.

Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at ScienceDirect

Computers & Geosciences

journal homepage: www.elsevier.com/locate/cageo

Estimation of static formation temperatures in geothermal wells by using an artificial neural network approach

A. Bassam^a, E. Santoyo^{a,*}, J. Andaverde^b, J.A. Hernández^b, O.M. Espinoza-Ojeda^a

^a Centro de Investigación en Energía, Universidad Nacional Autónoma de México, Privada Xochicalco s/n, Centro, 62580 Temixco, Morelos, México

^b Centro de Investigación en Ingeniería y Ciencias Aplicadas, Universidad Autónoma del Estado de Morelos, Av. Universidad 1001, Col. Chamilpa, 62209 Cuernavaca, Morelos, México

ARTICLE INFO

Article history:

Received 13 August 2009

Received in revised form

6 January 2010

Accepted 7 January 2010

Keywords:

Horner method

Geothermal energy

Levenberg–Marquardt algorithm

Borehole drilling

Bottom-hole temperature

Artificial intelligence

Shut-in time

ABSTRACT

An artificial neural network (ANN) approach was used to develop a new predictive model for the calculation of static formation temperature (SFT) in geothermal wells. A three-layer ANN architecture was successfully trained using a geothermal borehole database, which contains “statistically normalised” SFT estimates. These estimates were inferred from seven analytical methods commonly used in geothermal industry. Bottom-hole temperature (BHT) measurements and shut-in times were used as main input variables for the ANN training. Transient temperature gradients were used as secondary variables. The Levenberg–Marquardt (LM) learning algorithm, the hyperbolic tangent sigmoid transfer function and the linear transfer function were used for the ANN optimisation. The best training data set was obtained with an ANN architecture composed by five neurons in the hidden layer, which made possible to predict the SFT with a satisfactory efficiency ($R^2 > 0.95$). A suitable accuracy of the ANN model was achieved with a percentage error less than $\pm 5\%$. The SFTs predicted by the ANN model were statistically analyzed and compared with “true” SFTs measured in synthetic experiments and actual BHT logs collected in geothermal boreholes during long shut-in times. These data sets were processed both to validate the new ANN model and to avoid bias. The SFT estimates inferred from the ANN validation process were in good agreement ($R^2 > 0.95$) with the “true” SFT data reported for synthetic and field experiments. The results suggest that the new ANN model could be used as a practical tool for the reliable prediction of SFT in geothermal wells using BHT and shut-in time as input data only.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

The exploitation of geothermal resources for producing electricity requires the drilling of deep boreholes in the most suitable thermal regions of geothermal fields (Saito et al., 1998; Davis and Michaelides, 2009). The borehole drilling is a complex process in which a constant thermal anomaly (added to a circulating drilling mud) affects the original rock-formation temperature surrounding the borehole (Fomin et al., 2003, 2005). Once the borehole drilling is completed, the resulting thermal recovery is evaluated by analyzing build-up bottom-hole temperature (BHT) and shut-in time measurements (Santoyo et al., 2000). BHT data are measured at different shut-in times during the borehole drilling operations (Espinoza-Paredes and Espinoza-Martinez, 2009). BHTs are usually costly due to the use of sophisticated logging equipment, but mostly, because the borehole drilling progress needs to be stopped (Wisian et al., 1998; Fomin et al., 2005).

The determination of the static formation temperatures (SFTs) from BHT measurements constitutes a crucial task for the evaluation of geothermal systems (Espinoza et al., 2001; Kutasov and Eppelbaum, 2003). The estimation of SFT from BHT data at early drilling times offers the chance to know the virgin formation temperature several months before it can be measured with accuracy. This approach is valuable in terms of planning, exploration, developing, and exploitation programmes. The knowledge of the SFT is required for the: (1) determination of geothermal gradients for exploration mapping; (2) determination of heat flows of promising geothermal zones; (3) interpretation of borehole logging; (4) optimal design of borehole drilling and completion programs; (5) location of permeable or lost circulation zones; (6) evaluation of *in-situ* thermophysical formation properties; and (7) for the estimation of the heat reserves in geothermal systems.

The estimation of SFT in geothermal boreholes is usually performed by using simplified analytical methods based on complex heat transfer models concentrated at the bottom-hole conditions, where BHTs are actually measured. The BHT measurements tend to reflect the thermal anomalies caused by the drilling mud circulation to the rock-formation.

* Corresponding author.

E-mail address: esg@cie.unam.mx (E. Santoyo).

The most common analytical methods include: (i) the radial source with a conductive heat flow or Brennand method (BM: Brennand, 1984); (ii) the cylindrical heat source with a conductive-convective heat flow method (CSM) proposed by Hasan and Kabir (1994); (iii) the line-source or the well-known Horner-plot method (HM: Dowdle and Cobb, 1975); (iv) the generalized Horner or the Kutasov–Eppelbaum method (KEM: Kutasov and Eppelbaum, 2005); (v) the cylindrical source with a conductive heat flow or Leblanc method (LM: Leblanc et al., 1981); (vi) the cylindrical source with a conductive heat flow or Manetti method (MM: Manetti, 1973); (vii) the spherical and radial heat flow method (SRM) proposed by Ascencio et al. (1994); among others.

Most of these analytical methods require at least three or more BHT measurements carried out at the same borehole depth but at different shut-in times. The SFT is determined through the mathematical solution of each analytical method, using BHT and shut-in time measurements as input data, and the linear or nonlinear regressions between BHT data and the time functions of each method (Verma et al., 2006a, b).

Although considerable progress has been achieved in this area over recent years, large discrepancies among the SFT results provided by different methods have been reported (Andaverde et al., 2005). Therefore, the development of new reliable methods to estimate SFT is still a challenging area. The geothermal industry is interested in estimating SFT in systems where heat convection dominates. Heat convection is attributed to the presence of permeable zones in reservoirs and drilling mud losses into the formation. The determination of SFT under such conditions requires the use of complex analytical methods or wellbore simulators. These tools need thermophysical and transport properties of drilling, cementing, and formation materials, which are rarely available in the literature. In addition, the accurate knowledge of the mud circulation time, which is generally unknown or difficult to determine under drilling conditions, is also needed by most analytical methods (Hermanrud et al., 1990).

Considering this complex scenario, the geothermal industry requires new practical tools that use the commonly available data (i.e., BHT and shut-in time) to estimate the SFT with an acceptable accuracy. In this regard, a new application of artificial neural networks (ANN) is here proposed for developing a new reliable method for the calculation of SFTs in geothermal wells.

ANN have been suggested as powerful computational tools in many sciences for modelling and solving complex real-world problems, and mainly for forecasting applications (Zhang et al., 1998). The use of ANN is increasingly in Earth Sciences (e.g., Goutorbe et al., 2006; Hsieh et al., 2009; Leite and Filho, 2009a, b; Morton, 2009). In recent years, the application of ANN to geothermal and petroleum engineering problems has also been the subject of study (e.g., Farshad et al., 2000; Bayram, 2001; Can, 2002; Yilmaz et al., 2002; Spichak and Goidina, 2005; Spichak, 2006; Díaz-González et al., 2008; Serpen et al., 2009).

In this work, a three-layer ANN approach was successfully trained for obtaining a new predictive model to estimate SFT in geothermal wells. The new ANN model requires BHT and shut-in time measurements (as main input data), normalised SFT estimates (inferred from seven analytical methods), and transient temperature gradients.

2. Methodology

The ANN model was trained with a geothermal borehole database containing SFT estimates (statistically normalised). The geothermal database also contains build-up thermal recovery logs recorded during drilling operations carried out in some geothermal boreholes of the world. Build-up thermal recovery logs

consist of BHT and shut-in time measurements, and transient temperature gradients, together with SFT estimates inferred from the regression analysis of curvilinear relationships between BHT and shut-in time data using simplified analytical methods. SFT estimates were calculated with seven analytical methods commonly used in geothermal industry.

The normalisation of SFT estimates was carried out using statistical tests for discordant outlier detection and rejection (Verma, 2005). Mean SFT estimates (normalised) and standard deviations were computed using an improved outlier detection/rejection statistical method (Verma et al., 2009). This statistical approach coupled to the ANN enables the resulting errors to be effectively detected and rejected before selecting reliable SFT estimates. The new predictive ANN model was evaluated by using four BHT data sets (two synthetic and two actual borehole data sets) measured in experimental and long logging works where the SFT or the true formation temperature (TFT) was already known. These validation tests were considered as an accuracy evaluation of the ANN model.

2.1. Geothermal borehole databases for learning and validation tests

A geothermal borehole database containing BHT and shut-in time (Δt) measurements logged from drilling operations of 10 boreholes was created (Table 1). These thermal recovery data sets were collected from the reports of boreholes drilled in various world geothermal fields: (1) Los Humeros, Mexico [MXCO1, number of BHT measurements (m_{BHT})=6; shut-in times (Δt) up to 36 h after cessation of drilling mud circulation (Andaverde et al., 2005)]; (2) Los Humeros, Mexico [MXCO2, $m_{\text{BHT}}=7$; $\Delta t \leq 42$ h (Andaverde et al., 2005)]; (3) Oklahoma, USA [SGIL, $m_{\text{BHT}}=13$; $\Delta t \leq 12$ h (Schoeppel and Gilarranz, 1966)]; (4) Mississippi wellbore, USA [USAM, $m_{\text{BHT}}=14$; $\Delta t \leq 200$ h (Kutasov, 1999)]; (5) Leyte, Philippines [PHIL, $m_{\text{BHT}}=14$; $\Delta t \leq 15.58$ h (Andaverde et al., 2005)]; (6) Kyushu, Japan [JAPAN, $m_{\text{BHT}}=10$; $\Delta t \leq 72.5$ h (Hyodo and Takasugi, 1995)]; (7) Chipilapa, El Salvador [CH-A1, $m_{\text{BHT}}=5$; $\Delta t \leq 95.27$ h (González-Partida et al., 1997)]; (8) Chipilapa, El Salvador [CH-A2, $m_{\text{BHT}}=6$; $\Delta t \leq 190.5$ h (Iglesias et al., 1995)]; (9) Larderello, Italy [ITAL, $m_{\text{BHT}}=40$; $\Delta t \leq 27$ h (Da-Xin, 1986)]; and (10) Norton Sound field, Alaska [COST, $m_{\text{BHT}}=6$; $\Delta t \leq 75.5$ h (Cao et al., 1988a)].

Four thermal recovery data sets were additionally compiled for the validation or accuracy tests:

(i) two synthetic data sets collected from experimental works reported in the literature (SHBE, $m_{\text{BHT}}=8$: Shen and Beck, 1986; and CLAH, $m_{\text{BHT}}=15$: Cao et al., 1988b); and (ii) two data sets logged in some geothermal boreholes which were collected during long logging works (CH-A4 borehole, $m_{\text{BHT}}=5$: Iglesias et al., 1995; and KJ-21 borehole, $m_{\text{BHT}}=6$: Steingrimsson and Gudmundsson, 2006). These data sets were considered for the evaluation of accuracy because the TFTs were already known (SHBE=80.0 °C; CLAH=120.0 °C; CH-A4=169 °C; and KJ-21=240 °C, respectively).

2.2. Initial estimation of SFTs using analytical methods

Seven analytical methods based on complex heat transfer models were used for the initial calculation of SFTs. The seven analytical methods were: (i) the Horner-plot method, HM; (ii) the Brennand method, BM; (iii) the Kutasov–Eppelbaum method, KEM; (iv) the Leblanc method (LM); (v) the Manetti method (MM); (vi) the conductive spherical-radial method (SRM); and (vii) the conductive-convective cylindrical source method (CSM). The simplified solutions of these analytical methods are summarised in Table 2. SFTs were estimated using conventional

Table 1
Thermal recovery data sets used for ANN training and validation.

Data set	Borehole data for training the new ANN model						Synthetic data for the validation						
	t (h)	BHT (°C)	Data set	t (h)	BHT (°C)	Data set	t (h)	BHT (°C)	Data set	t (h)	BHT (°C)		
MXCO1 $t_c=2.5$ h $m=6$	6	186.3	PHIL $t_c=15$ h $m=14$	2.58	93	ITAL $t_c=5$ h $m=40$	5.7	84.5	SHBE $t_c=5$ h $m=8$	2.5	56.6		
	12	206.9		3.58	88		6.0	93.1		5	61.3		
	18	219.0		4.58	99		6.5	95.0		7.5	64.3		
	24	231.9		5.58	108		7.0	96.2		10	66.6		
	30	239.8		6.58	112		7.5	97.5		15	69.6		
	36	247.7		7.58	117		8.0	98.4		20	71.7		
MXCO2 $t_c=2.5$ h $m=7$	6	178.6		8.58	120		8.5	99.3		30	74.1		
	12	198.9		9.58	126		9.0	100.3		40	75.5		
	18	211.4		10.58	133		9.5	101.1					
	24	225.8		11.58	133		10.0	102.0					
	30	235.1		12.58	134		10.5	102.7					
	36	240.4		13.58	137		11.0	103.4		CLAH $t_c=5$ h $m=15$	2	91.7	
42	247.1	14.58	141	11.5	104.2	4	98.1						
SGIL $t_c=3$ h $m=13$	0	62.78	JAPN $t_c=5$ h $m=10$	15.58	146		12.0	104.7			6	102.4	
	1	81.94		5.5	92.0		12.5	105.4			8	105.5	
	2	86.56		6.5	98.5		13.0	106.1			10	107.9	
	3	88.84		7.5	103.0		13.5	106.7			12	109.8	
	4	90.76		8.5	107.4		14.0	107.6		14	111.3		
	5	92.29		9.5	110.4		14.5	108.3		16	112.5		
	6	93.44		12.5	119.6		15.0	108.9		18	113.6		
	7	94.21		15.5	126.6		15.5	109.6		20	114.4		
	8	94.59		18.5	132.8		16.0	110.2		22	115.2		
	9	94.98		24.5	142.4		16.5	110.8		24	115.8		
	10	95.36		72.5	170.9		17.0	111.3		30	117.1		
	11	95.74		CH-A1 $t_c=5$ h $m=5$	6		90	17.5		111.9	40	118.4	
12	96.13	12	102		18.0	112.4	50	119.1					
USAM $t_c=5$ h $m=14$	1	138.75			26.57	116		18.5	112.9	CH-A4 $t_c=5$ h $m=5$	6	85	
	2	139.55			47.1	130		19.0	113.3		12	94	
	5	140.89			95.27	145		19.5	113.8		26.57	122	
	10	142.03			6.0	231		20.0	114.2		47.1	139	
	15	142.72		$t_c=5$ h $m=6$	12.0	231		20.5	114.7		95.27	152	
	20	143.21		26.6	235	21.0		115.1					
	25	143.60		47.1	237	21.5		115.4	KJ-21 $t_c=2.5$ h $m=6$		2	50	
	30	143.92		95.3	239	22.0		115.7			30	125	
	40	144.42		190.5	240	22.5		116.1			52	158	
	50	144.81		COST $t_c=4$ h $m=6$	8.5	45.00		23.0			116.5	96	184
	70	145.40			10.0	49.44		24.0			117.1	242	218
	100	146.04			16.5	55.56		25.0			117.7	556	225
150	146.76	34.5	56.11		26.0	117.9							
200	147.27	38.0	57.22		27.0	118.7							
		75.5	56.11										

t_c —drilling mud circulation time; additional data reported for each data set: SGIL ($R=0.10028$, $k=2.25419$ W/m °C, $c=878.64$ J/kg °C, and $\rho=2306.6$ kg/m³); COST ($k=4.97$ W/m °C); SHBE ($r=0.108$ m, TFT=80 °C); CLAH ($r=0.108$ m, TFT=120 °C); CH-A4 (TFT=169 °C); KJ-21(TFT=240 °C). m —number of data.

Table 2
Listing of major numerical solutions of analytical methods (HM, BM, KEM, LM, MM, SRM, and CSM) used for determining SFT in geothermal boreholes.

Analytical method	Numerical solution
BM	$BHT(t) = T_{BM} - b_{BM} \left(\frac{1}{\Delta t + p t_c} \right)$
CSM	$BHT = T_{CSM} + (b_{CSM}) [F(t_{cD} + \Delta t D) - F(t_D)]$
HM	$BHT(t) = T_{HM} + (b_{HM}) \ln \{ (t_c + \Delta t) / \Delta t \}$
KEM	$BHT(t) = T_{KEM} + b_{KEM} \ln(X)$
LM	$BHT(t) = T_{LM} - b_{LM} \left[1 - \exp \left(- \frac{r_w^2}{4\alpha \Delta t} \right) \right]$
MM	$BHT(t) \approx T_{MM} + b_{MM} \ln \left(\frac{\Delta t}{\Delta t - t_c} \right)$
SRM	$BHT(t) = T_{SRM} + (b_{SRM}) \left(\frac{1}{\sqrt{\Delta t}} \right)$

BHT: Bottom-hole temperature; Δt : shut-in time; SFTs were inferred from the intercept value of each analytical method: T_{BM} , T_{CSM} , T_{HM} , T_{KEM} , T_{LM} , T_{MM} , and T_{SRM} . For more details of primary and secondary variables, the readers are referred to the original sources of the methods (indicated in the introduction section).

linear (ordinary and weighted) and quadratic (QR) regressions between the BHT and the time function data of each method (see Andaverde et al., 2005). Ordinary linear regressions (OLR) neglect the errors in the x (time functions) and y (BHT) variables for the calculation of the regression parameters: intercept (a) and slope (b), whereas the weighted linear regressions (WLR) consider the errors of x and y (s_x and s_y) for the estimation of the regression parameters (Verma, 2005).

According to the numerical procedure used by each analytical method, SFTs were inferred from the intercept values (a) of the OLR, WLR and QR (e.g., Santoyo et al., 2000; Andaverde et al., 2005).

2.3. General scheme for the SFT estimation

A schematic flow diagram showing the training procedure used by the ANN model is shown in Fig. 1. Five fundamental

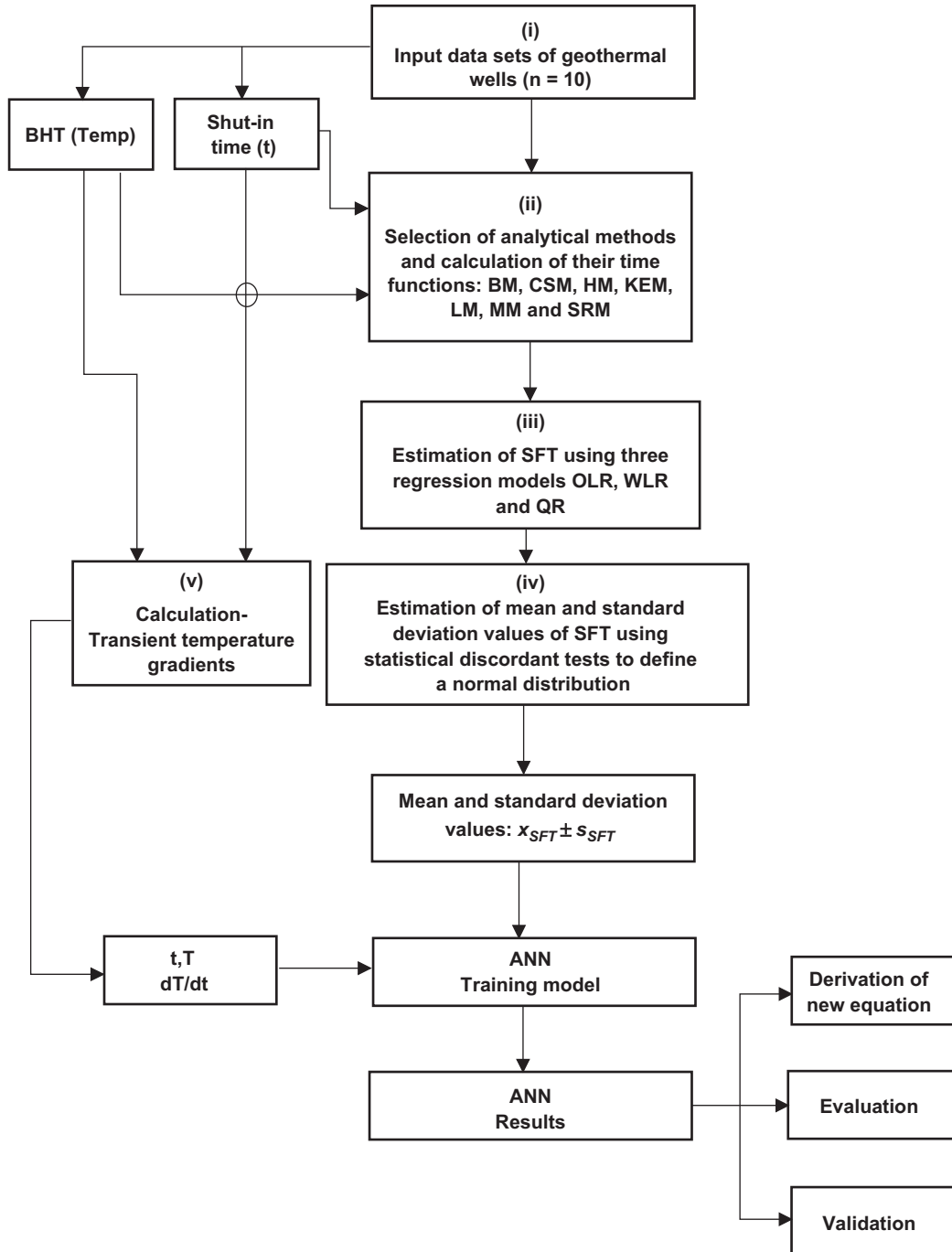


Fig. 1. Schematic flow diagram of ANN model used for determining SFT in geothermal boreholes.

sections were included in the training process: (i) the input data reading ($j=10$, containing 120 pairs of BHT and shut-in time data); (ii) the selection of analytical methods and calculation of their time functions per data set; (iii) the determination of the SFT by seven analytical methods using three different regression models: OLR, WLR, and QR per data set; (iv) the application of statistical discordant tests to calculate the mean values of the SFT estimates and their standard deviations per data set; and (v) the calculation of temperature gradients using the transient thermal recovery data of the boreholes.

SFT estimates were individually computed by using the 10 data sets (MXCO1, MXCO2, SGIL, USAM, PHIL, JAPN, CH-A1, CH-A2, ITAL, and COST), which contain BHT and shut-in time data. The time functions of each analytical method (HM, BM, KEM, LM, MM, SRM,

and CSM) were calculated by means of their corresponding equations (Table 2). Three SFT estimates were individually inferred from the regression analysis (OLR, WLR, and QR) using each analytical method and referred to each data set. A total number of 21 SFT estimates per data set were finally computed.

For training and correlating the borehole thermal recovery data sets with reliable SFT values, mean values of the 21 SFT estimates were computed per data set through a statistical normalisation using the software DODESYS (Verma et al., 2009). Nine discordant tests (N1, N2, N4, N7, N8, N9, N10, N14, and N15 proposed by several authors: see Verma, 2005) were applied to the 21 SFT estimates of each data set; firstly, to identify and to reject discordant outliers, and secondly, to calculate the mean SFT (\bar{x}_{SFT}) and its associated standard deviation (s_{SFT}).

This method of statistical normalisation enabled the large resulting errors derived from the analytical methods to be efficiently detected and rejected before selecting the best SFT estimates. Table 3 summarises the results obtained after using the statistical normalisation procedure.

The mean SFT estimates and discordant outliers detected after applying the statistical tests are plotted in Fig. 2. Thirty discordant outliers were detected and rejected. These outliers corresponded to the SFT estimates inferred from the SRM which systematically overestimate the SFT (Andaverde et al., 2005). The mean values of the SFT estimates (\bar{x}_{SFT}) and their standard deviations (s_{SFT}) per data set were used for the ANN training. Three SFT values were finally employed for transporting the uncertainties (the mean value: \bar{x}_{SFT} ; the mean value plus the standard deviation: $\bar{x}_{SFT} + s_{SFT}$; and the mean value minus the standard deviation: $\bar{x}_{SFT} - s_{SFT}$) (see Table 3).

Transient temperature gradients were computed using the following equation:

$$\frac{dT}{dt} = \frac{(BHT_{i+1} - BHT_i)}{t_{i+1} - t_i} \quad (1)$$

Table 3
Mean SFT estimates inferred from 21 SFT predictions of seven analytical methods (using three different regression models OLR, WLR, and QR), and after applying a statistical normalisation process.

Data set	SFT (analytical methods) $\bar{x}_{SFT} \pm s_{SFT}$	SFT (predicted—ANN) $\bar{x}_{ANN} \pm s_{ANN}$
MXCO1	263 ± 15	256 ± 20
MXCO2	262 ± 18	253 ± 15
USAM	146 ± 1	139 ± 5
ITAL	132 ± 10	133 ± 8
PHIL	197 ± 28	185 ± 25
JAPN	176 ± 18	180 ± 11
CH-A1	147 ± 5	151 ± 7
CH-A2	241 ± 18	243 ± 15
COST	57 ± 3	56 ± 2
SGIL	100 ± 3	102 ± 9

where BHT denotes the bottom-hole temperatures, t the shut-in times, and the subscripts i provides the transient values of these variables.

The transient temperature gradient (dT/dt) was an important parameter for training the ANN. When (dT/dt) tend to zero at infinite time ($t \rightarrow \infty$), it means that the geothermal formation temperature approaches to thermal equilibrium, and therefore:

$$\lim_{t \rightarrow \infty} \frac{dT}{dt} = 0 \quad \therefore (BHT \approx SFT) \quad (2)$$

The SFT_{ANN} was defined as the unique output parameter of the ANN model.

2.4. ANN architecture

The neurons are grouped into distinct layers and interconnected according to a given computational architecture. The network function is largely determined by the connections between neurons. Each connection among two neurons has a weighting coefficient. The standard network structure for an approximation function is the multiple-layer perception (or the well-known feed forward network). The feed forward network often has one or more hidden layers of sigmoidal neurons followed by an output layer of linear neurons. Multiple-layers of neurons with nonlinear transfer functions allow the network to learn nonlinear and linear relationships between input and output vectors. The linear output layer enables the network to generate values outside the -1 to $+1$ range (e.g., Limin, 1995). For the network, the appropriate notation is used in two-layer networks (e.g. Demuth and Beale, 1998). A simplified sketch of the network structure is schematically represented in Fig. 3. The number of neurons in the input and output layers is, respectively, given by the number of input and output variables of the process under investigation. In this work, the input layer can consist of a set of different variables, In_k (i.e., BHT, shut-in time, and dT/dt), while the output layer can contain the output or target variables of the model, Out_i (i.e., SFT), see Fig. 3.

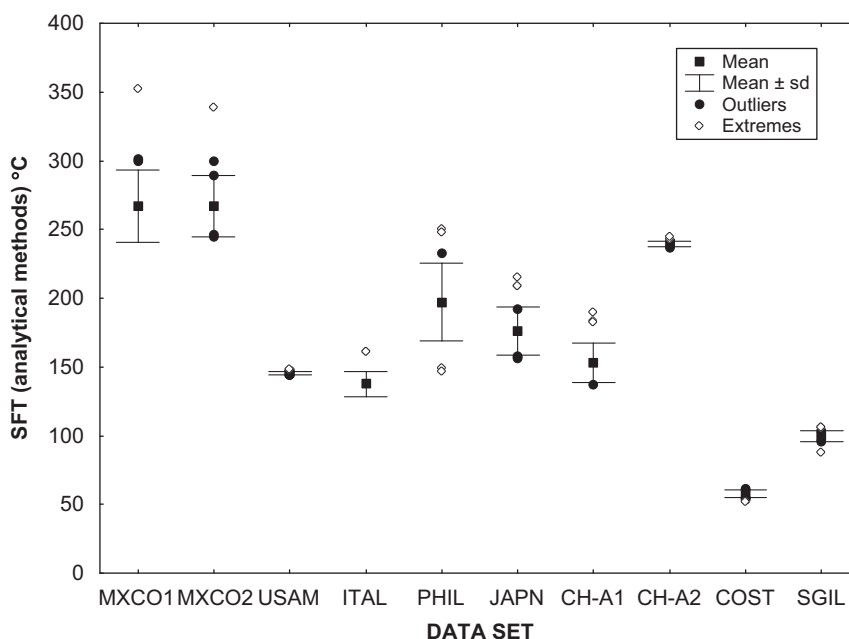


Fig. 2. Mean SFT estimates obtained from normal statistical samples after applying nine statistical discordant tests. Mean SFT estimates per data set were: MXCO1: 263 ± 15 °C; MXCO2: 262 ± 18 °C; USAM: 146 ± 1 °C; ITAL: 132 ± 10 °C; PHIL: 197 ± 28 °C; JAPN: 176 ± 18 °C; CH-A1: 147 ± 5 °C; CH-A2: 241 ± 18 °C; COST: 57 ± 3 °C; and SGIL: 100 ± 3 °C.

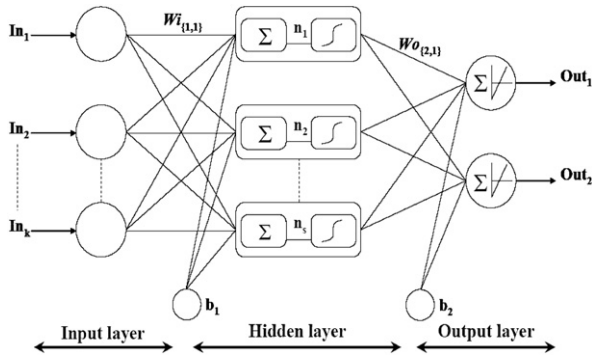


Fig. 3. General ANN architecture (k =input number; In=Input; Out=Output; Continuous lines=weights and bias).

The optimal number of neurons in the hidden layer (S) is difficult to specify, and it generally depends on the problem complexity. This number (S) is iteratively determined. Each neuron in the hidden layer has a bias, $b_1(s)$, which is added to the sum of the product (weights multiplied by its corresponding input variables). In general, the optimal neuron (n_s), is given by the following equation:

$$n_s = W_{i(s,1)}In_1 + W_{i(s,2)}In_2 + \dots + W_{i(s,k)}In_k + b_1(s) \quad (3)$$

The coefficients associated with the hidden layer are grouped into a matrix W_i (given by weights) and the vector $b_1(s)$ of biases. The output layer computes the weighted sum of the signals provided by the hidden layer, and the associated coefficients are grouped into a matrix W_o and the vector $b_2(l)$. Using the matrix notation, the network output can be expressed by the following general equation:

$$Out = g(W_o f(W_i In + b_1) + b_2) \quad (4)$$

Hidden layer neurons may use any differentiable transfer function to generate their output. In this work, a hyperbolic tangent sigmoidal transfer function ($f=tansig$) and a linear transfer function ($g=purelin$) were used. As a result, Eq. (4) can be re-written using these transfer functions to provide the following general equation:

$$Out_{(l)} = \sum_{s=1}^S \left[W_o(l,s) \left(\frac{2}{1 + \exp\left(-2\left(\sum_{k=1}^K (W_i(s,k)In(k) + b_1(s))\right)\right)} - 1 \right) \right] + b_{2(l)} \quad (5)$$

where K , S , and l are the number of input variables, the number of neurons in the hidden layer, and the number of the output variables in the ANN architecture, respectively.

3. Neural network learning and testing

A learning (or training) algorithm is defined as a process that consists of adjusting the coefficients (weights and biases) of a network, to minimize an error function (usually a quadratic one) between the network outputs (for a given set of inputs) and the correct outputs (already known). If smooth non-linearities are used, the error function gradient can be computed by the classical back propagation model (Rumelhart et al., 1986).

Previous learning algorithms have used this gradient directly in a steepest descent optimisation, but recent results have shown that second order methods are far more efficient. In this work, the Levenberg–Marquardt (LM) optimisation algorithm (described in the Matlab Neural Network Toolbox) was used (Demuth and Beale, 1998). This algorithm is basically an approximation of

Newton's method (Eq. (6)), which is suggested to approach a second order training speed without having to compute the Hessian matrix (Hagan and Menhaj, 1994)

$$x_{k+1} = x_k - [J^T J + \mu I]^{-1} J^T e \quad (6)$$

where J is the Jacobian matrix that contains first derivatives of the network errors with respect to the weights and biases; e is the vector of network errors; $J^T J$ is an approximation of the Hessian matrix; and $J^T e$ is the gradient. When the scalar μ is zero, Eq. (6) becomes the classical Newton's method using the approximate Hessian matrix. On the other hand, when μ is large, Eq. (6) becomes a more efficient descent gradient with a small step size. The Jacobian matrix can be computed through a standard back-propagation technique (see Hagan and Menhaj, 1994).

Despite the fact that computations involved in iterations were more complex than the steepest descent case, the convergence was faster (typically by a factor of 100). The root mean square error (RMSE) was calculated using the experimental and simulated ANN data (Eq. (7)). This calculation was used as an optimisation criterion for the model adequacy (Fig. 4).

The experimental database (see Table 1) were split into learning (60% of experimental data set; $n=180$) and testing (40% of experimental data set; $n=120$) databases to obtain a statistically representative sample of the problem

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Exp(i) - Sim(i))^2} \quad (7)$$

3.1. ANN training

The new neural network model trained for this work (Fig. 5) finally involved five neurons ($n_s=5$) in the hidden layer with 20 weights (15 elements of a matrix W_i and 5 elements of the matrix W_o), and 6 biases ($b_1=5$ and $b_2=1$) to calculate a SFT value

$$SFT = \sum_{j=1}^S \left[W_o(1,j) \left(\frac{2}{1 + \exp\left(-2\left(\sum_{k=1}^K (W_i(j,k)In(k) + b_1(j))\right)\right)} - 1 \right) \right] + b_2 \quad (8)$$

where S and K were already defined and $l=1$ (see Table 4).

During the learning process, the algorithm of the ANN model was evaluated by examining the RMSE value per iteration, together with a different number of neurons in the hidden layer (ranging from one to five neurons). The obtained results (data not shown) proved that the typical learning error decreased when the number of neurons in the hidden layer increased. This observation was evident because of the number of adjusted parameters increased. Nevertheless, one of the problems that could occur during feed-forward neural network training is the well-known "over-fitting" (e.g. Hernández-Pérez et al., 2004).

The comparison of the RMSE calculated for the learning and testing databases was a good criterion to optimize the number of iterations, and to avoid the over-fitting problem. In this neural network, the RMSE showed that for six neurons in the hidden layer, the learning database value was too small with respect to the testing database.

According to RMSE results, the optimal number of neurons in the hidden layer was fixed as five neurons (Fig. 5). The predicted SFT values computed with the ANN model together with their uncertainties are also included in Table 3. A good agreement between the SFT estimates (inferred from analytical methods: x -axis) and the simulated data, SFT_{ANN} (obtained from the ANN model: y -axis) was found (Fig. 6). It can be observed that the

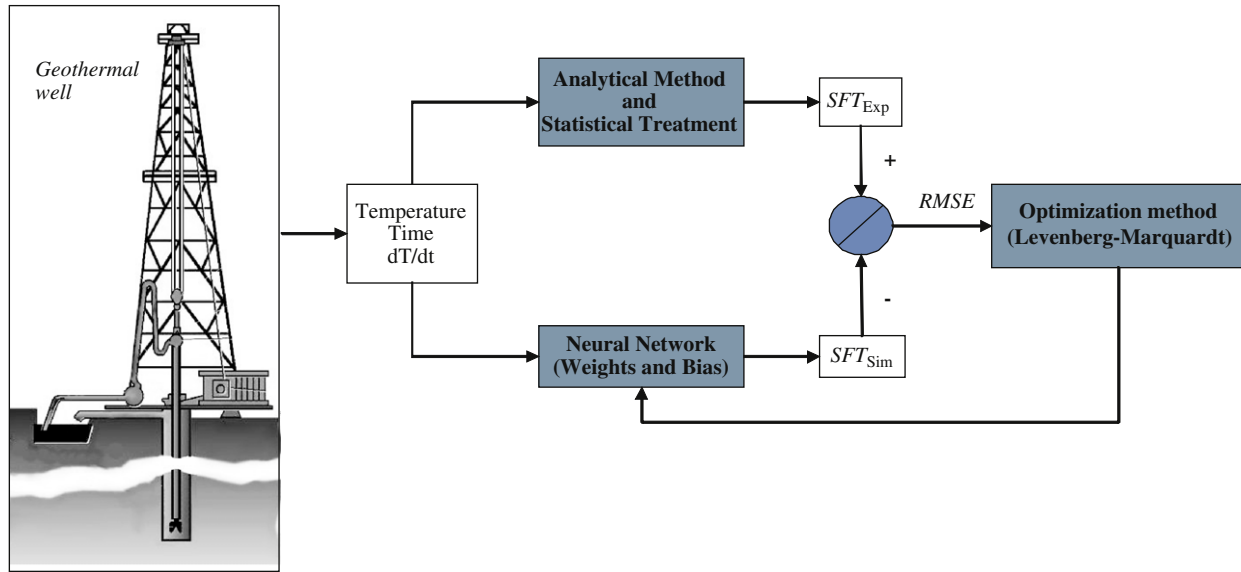


Fig. 4. Numerical algorithm of ANN used for learning and optimizing processes.

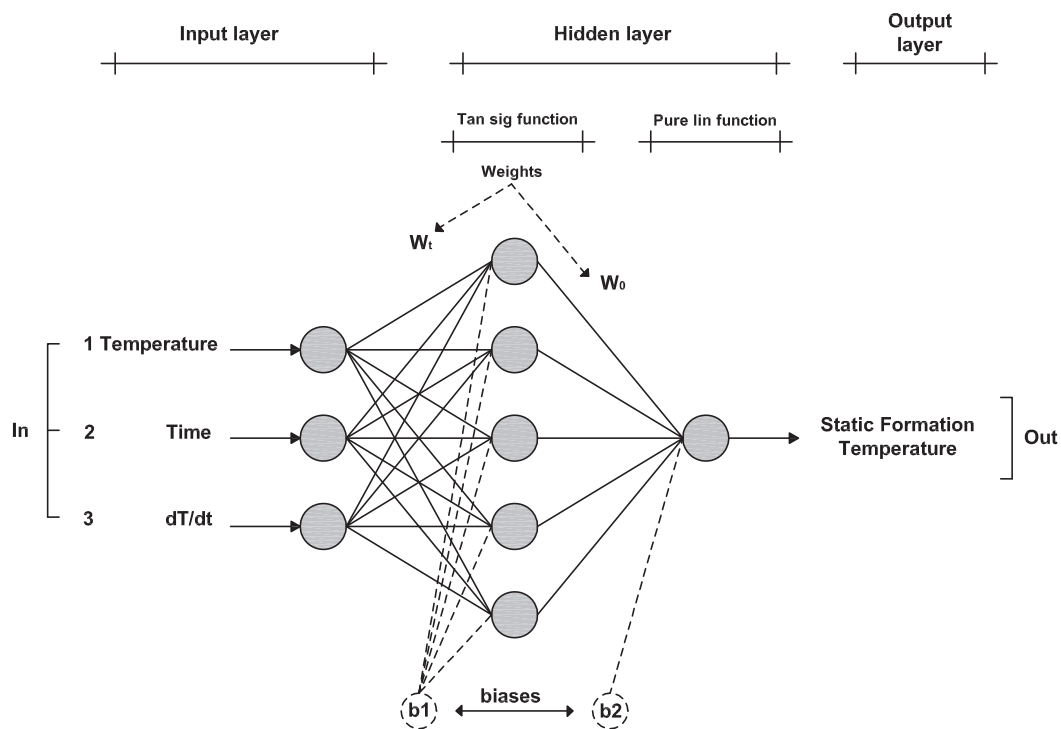


Fig. 5. ANN architecture used for estimating SFT in geothermal boreholes.

simulated SFT_{ANN} data exhibit a good linear relationship with respect to those estimates inferred from analytical methods ($y=0.94376x+8.5631$; $R^2=0.95$). This linear relationship was also confirmed by analyzing the small residuals computed between the measured and predicted SFT values.

The optimum weighting coefficients (W_i , W_o , b_1 , and b_2) obtained for the ANN model using five neurons in the hidden layer are reported in Table 4. These coefficients were used to validate the ANN model with testing data.

3.2. ANN testing

The new ANN model (Eq. (8)) was validated by using the testing data reported for synthetic experiments (SHBE: TFT=80 °C; and CLAH: true TFT=120 °C); and for actual geothermal borehole measurements (CH-A4: TFT=169 °C; and KJ-21: true TFT=240 °C) (see Table 1). The results from these four accuracy tests showed a good agreement (less than a $\pm 5\%$) between the measured SFT (also called TFT) and predicted SFT data (Fig. 7).

Table 4
Optimised weighting and bias parameters of new ANN model.

ANN parameters	With $l=1, S=5$ and $K=3$				
W_i		3.9757	1.4220	-2.9996	
		5.0716	13.9674	-2.3583	
		-6.9071	-2.7251	-34.0318	
		-6.3881	2.3519	-6.0095	
		-5.2657	3.6209	-4.5029	
W_o	-1.4449	0.1627	-0.3257	0.4320	-0.5981
b_1			-5.6937		
			-10.0967		
			1.6998		
			-0.5554		
			-2.6091		
b_2		-1.4001			

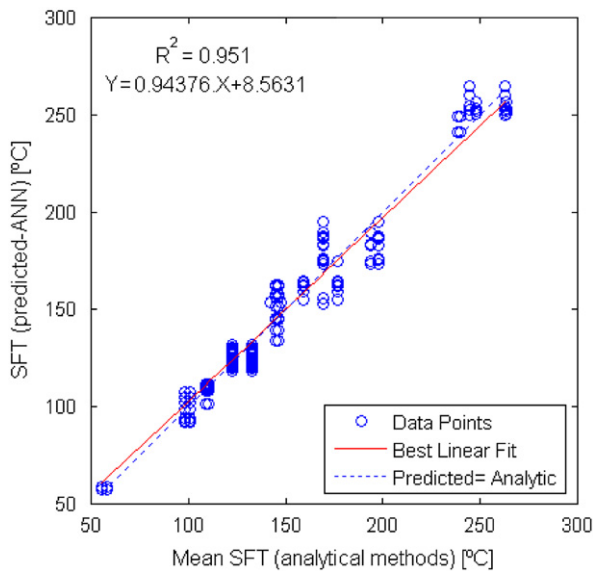


Fig. 6. Statistical comparison between SFT estimates (inferred from analytical methods) and ANN simulated SFT values.

For the synthetic data sets, the ANN model predicted SFT values ranging from 74.5 to 83.4 °C, with a mean SFT_{ANN} value of 79.1 °C for the SHBE data set; and SFT values in the interval 117.8–125.6 °C, with a mean SFT_{ANN} value of 122.1 °C for the CLAH data set. Whereas for the borehole data sets: CH-A4 and KJ-21, the ANN model predicted SFT values in the interval 167.6–173.5 °C, with a mean SFT_{ANN} value of 171.2 °C; and SFT values ranging from 239.1 to 247.1 °C, with a mean SFT_{ANN} value of 243.3 °C, respectively.

From these results, it can be reliably demonstrated that the new ANN model (given by the Eq. (8)) provides acceptable estimates of the SFT with typical errors commonly reported in artificial intelligence engineering applications (e.g. Leite and Filho, 2009a).

The small residuals calculated with the testing database, as well as the good regression coefficient obtained ($R^2=0.95$), enabled to confirm the satisfactory agreement between the measured and predicted SFT data (see Fig. 7). The ANN model developed was therefore able to predict the SFT throughout the experimental data domain used for the training process (45–263 °C).

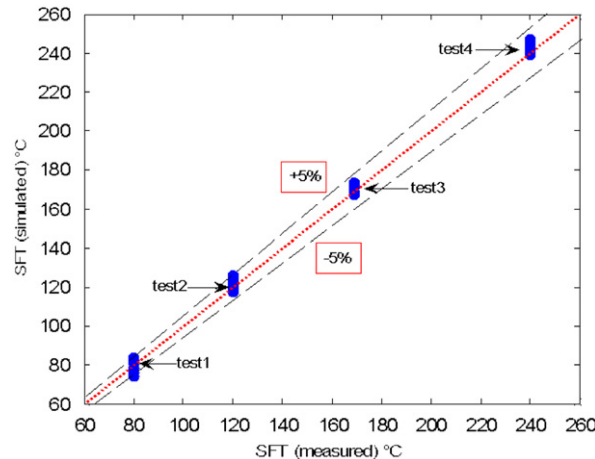


Fig. 7. Evaluation of ANN accuracy by using a statistical comparison between ANN simulated SFT values and true SFT data or TFT (reported in synthetic and actual borehole data sets).

Other additional advantages of the ANN model included: (i) to use the ANN tool without any assumptions about the nature of underlying mechanisms; (ii) to have the ability to take into account non-linearities and iterations between input and output variables; (iii) to optimise the computing time because these models rely on simple mathematical operations (allowing faster calculations); and (iv) to obtain reliable and faster predictions of SFT in comparison with those inferred from current analytical methods.

4. Conclusions

A new ANN model was successfully developed to predict the SFT in geothermal boreholes. It only requires BHT and shut-in time measurements as main input data. This feature makes it an advantageous tool over current available analytical methods that strongly depend on other complex input variables, such as the drilling mud circulation time. The ANN model was effectively trained with an experimental database, and validated with an unbiased testing database. The validity of the SFT computed by the ANN model was confirmed through a statistical comparison between the measured and simulated values, showing a good agreement with errors less than $\pm 5\%$.

The new ANN model can be used as a practical tool for the reliable determination of SFT in geothermal boreholes, especially when the BHT data are in the training interval (45 to ~ 263 °C). The ANN application interval could be extended by using higher BHT measurements. The predicting accuracy of the ANN could also be improved by using training patterns with a large number of BHT measurements, and their uncertainties.

The idea of applying ANN for the estimation of SFT is to propose a new practical and faster tool for the geothermal industry. In fact, the ANN model here developed could be coupled with smart BHT sensors for on-line determination of SFT in geothermal boreholes.

Finally, the ANN model developed may enable the implementation of smart sensors for on-line determination (SFT) in geothermal boreholes.

Acknowledgements

The authors wish to thank to the Engineering Ph.D. Programme of UNAM and to CONACyT for the financial support provided.

Thanks are also due to Dr. A. Santoyo-Castelazo (University of Liverpool) for her valuable comments. We are grateful to the anonymous reviewer, Dr. Viacheslav V. Spichak, and Prof. Dr. E.C. Grunsky (Editor in Chief) for their helpful comments on an earlier version of this paper.

References

- Andaverde, J., Verma, S.P., Santoyo, E., 2005. Uncertainty estimates of static formation temperature in borehole and evaluation of regression models. *Geophysical Journal International* 160, 1112–1122.
- Ascencio, F., García, A., Rivera, J., Arellano, V., 1994. Estimation of undisturbed formation temperatures under spherical radial heat flow conditions. *Geothermics* 23, 317–326.
- Bayram, A.F., 2001. Application of an artificial neural network model to a Na/K geothermometer. *Journal of Volcanology and Geothermal Research* 112, 75–81.
- Brennand, A.W., 1984. A new method for the analysis of static formation temperature test. In: *Proceedings of the 6th New Zealand Geothermal Workshop*, Auckland, New Zealand, pp. 45–47.
- Can, I., 2002. A new improved Na/K geothermometer by artificial neural networks. *Geothermics* 31, 751–760.
- Cao, S., Hermanrud, C., Lerche, I., 1988a. Estimation of formation temperature from bottom-hole temperature measurements: COST #1 well, Norton Sound, Alaska. *Geophysics* 53, 1619–1621.
- Cao, S., Lerche, I., Hermanrud, C., 1988b. Formation temperature estimation by inversion of borehole measurements. *Geophysics* 53, 979–988.
- Davis, A.P., Michaelides, E.E., 2009. Geothermal power production from abandoned oil wells. *Energy* 34, 866–872.
- Da-Xin, L., 1986. Non-linear fitting method of finding equilibrium temperature from BHT data. *Geothermics* 15, 657–664.
- Demuth, H., Beale, M., 1998. *Neural Network Toolbox for Matlab-User's Guide Version 3*. The MathWorks Inc., Natick, MA 742 pp.
- Díaz-González, L., Santoyo, E., Reyes, J., 2008. Tres nuevos geotermómetros mejorados de Na/K usando herramientas computacionales y geoquímicas: aplicación a la predicción de temperaturas de sistemas geotérmicos ("Three new improved geothermometers of Na/K using computational and geochemometric tools: application to the prediction of temperatures in geothermal systems"). *Revista Mexicana de Ciencias Geológicas* 25, 465–482.
- Dowdle, W.L., Cobb, W.M., 1975. Static formation temperature from well logs—an empirical method. *Journal of Petroleum Technology* 11, 1326–1330.
- Espinosa, G., García, A., Santoyo, E., Hernandez, I., 2001. TEMLOPI/V.2: a computer program for estimation of fully transient temperatures in geothermal wells during circulation and shut-in. *Computers & Geosciences* 27, 327–344.
- Espinosa-Paredes, G., Espinosa-Martinez, E.G., 2009. A feedback-based inverse heat transfer method to estimate unperturbed temperatures in wellbores. *Energy Conversion and Management* 50, 140–148.
- Farshad, F.F., Garber, J.D., Lorde, J.L., 2000. Predicting temperature profiles in producing oil wells using artificial neural networks. *Engineering Computations* 17, 735–754.
- Fomin, S., Chugunov, V., Hashida, T., 2003. Analytical modelling of the formation temperature stabilization during the borehole shut-in period. *Geophysical Journal International* 155, 469–478.
- Fomin, S., Hashida, T., Chugunov, V., Kuznetsov, A.V., 2005. A borehole temperature during drilling in a fractured rock formation. *International Journal of Heat and Mass Transfer* 48, 385–394.
- González-Partida, E., García-Gutiérrez, A., Torres-Rodríguez, V., 1997. Thermal and petrologic study of the CH-A well from the Chipilapa-Ahuachapan geothermal area, El Salvador. *Geothermics* 26, 701–713.
- Goutorbe, B., Lucazeau, F., Bonneville, A., 2006. Using neural networks to predict thermal conductivity from geophysical well logs. *Geophysical Journal International* 166, 115–125.
- Hasan, A.R., Kabir, C.S., 1994. Static reservoir temperature determination from transient data after mud circulation. *Society of Petroleum Engineering, Drilling & Completion* 9, 17–24.
- Hagan, M.T., Menhaj, M., 1994. Training feed forward networks with the Marquardt algorithm. *IEEE Transactions on Neural Networks* 5, 989–993.
- Hermanrud, C., Cao, S., Lerche, I., 1990. Estimates of virgin rock temperature derived from BHT measurements: bias and error. *Geophysics* 55, 924–931.
- Hernández-Pérez, J.A., García-Alvarado, M.A., Trystram, G., Heyd, B., 2004. Neural networks for the heat and mass transfer prediction during drying of cassava and mango. *Food Science and Emerging Technologies* 5, 57–64.
- Hsieh, B.Z., Wang, C.W., Lin, Z.S., 2009. Estimation of formation strength index of aquifer from neural networks. *Computers & Geosciences* 35, 1933–1939.
- Hyodo, M., Takasugi, S., 1995. Evaluation of the curve-fitting method and the Horner plot for estimation of the true formation temperature using temperature recovery logging data. In: *Proceedings of the 20th Workshop on Geothermal Reservoir Engineering*, Stanford University, Stanford, USA, pp. 94–100.
- Iglesias, E.R., Campos-Romero, A., Torres, R.J., 1995. A reservoir engineering assessment of the Chipilapa, El Salvador, geothermal field. In: *Proceedings of the World Geothermal Congress*, Florence, Italy, pp. 1531–1536.
- Kutasov, I.M., 1999. *Applied Geothermics for Petroleum Engineers*. Elsevier, Amsterdam 360 pp.
- Kutasov, I.M., Eppelbaum, L.V., 2003. Prediction of formation temperatures in permafrost regions from temperature logs in deep wells-field cases. *Permafrost and Periglacial Process* 14, 247–258.
- Kutasov, I.M., Eppelbaum, L.V., 2005. Determination of formation temperature from bottom-hole temperature logs—a generalized Horner method. *Journal of Geophysical Engineering* 2, 90–96.
- Leblanc, Y., Pascoe, L.J., Jones, F.W., 1981. The temperature stabilization of a borehole. *Geophysics* 46, 1301–1303.
- Leite, E.P., Filho, C.R.de S., 2009a. Probabilistic neural networks applied to mineral potential mapping for platinum group elements in the Serra Leste region, Carajás mineral province, Brazil. *Computers & Geosciences* 35, 675–687.
- Leite, E.P., Filho, C.R.de S., 2009b. TEXTNN—a MATLAB program for textural classification using neural networks. *Computers & Geosciences* 35, 2084–2094.
- Limin, F., 1995. *Neural Networks in Computer Intelligence*. McGraw-Hill, New York 460 pp.
- Manetti, G., 1973. Attainment of temperature equilibrium in holes during drilling. *Geothermics* 2, 94–100.
- Morton, J.C., 2009. Boosting a fast neural network for supervised land cover classification. *Computers & Geosciences* 35, 1280–1295.
- Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning internal representations by error propagation. *Parallel Distributed Processing* 1, 318–362.
- Saito, S., Sakuma, S., Uchida, T., 1998. Drilling procedures, techniques and test results for a 3.7 km deep, 500 °C exploration well, Kakkonda, Japan. *Geothermics* 27, 573–590.
- Santoyo, E., García, A., Espinosa, G., Hernández, I., Santoyo, S., 2000. STATIC_TEMP: a useful computer code for calculating static formation temperatures in geothermal wells. *Computers & Geosciences* 26, 201–217.
- Schoeppel, R.J., Gilarranz, S., 1966. Use of well log temperatures to evaluate regional geothermal gradients. *Journal of Petroleum Technology* 18, 667–673.
- Serpen, G., Palabiyik, Y., Serpen, U., 2009. An artificial neural network model for Na/K geothermometer. In: *Proceedings of the 34th Workshop on Geothermal Reservoir Engineering*, Stanford University, Stanford, USA, pp. 1–12.
- Shen, P.Y., Beck, A.E., 1986. Stabilization of bottomhole temperature with finite circulation time and fluid flow. *Geophysical Journal of the Royal Astronomical Society* 86, 63–90.
- Spichak, V.V., 2006. Estimating temperature distributions in geothermal areas using a neuronet approach. *Geothermics* 35, 181–197.
- Spichak, V.V., Goidina, A.G., 2005. Temperature prediction in geothermal zones from borehole measurements using neural networks. *Physics of the Solid Earth* 4, 844–852.
- Steingrímsson, B., Gudmundsson, A., 2006. Geothermal borehole investigations during and after drilling. In: *Proceedings at Workshop on Decision Makers on Geothermal Projects in Central America*, El Salvador, pp. 1–10.
- Verma, S.P., 2005. *Estadística Básica para el Manejo de Datos Experimentales con la Aplicación en la Geoquímica—Geoquimiometría (Basic Statistics for Experimental Data Reduction with Application in Geochemistry—Geochemometrics)*. Universidad Nacional Autónoma de México, México, DF 186 pp.
- Verma, S.P., Andaverde, J., Santoyo, E., 2006a. Application of the error propagation theory in estimates of static formation temperatures in geothermal and petroleum boreholes. *Energy Conversion and Management* 47, 3659–3671.
- Verma, S.P., Andaverde, J., Santoyo, E., 2006b. Statistical evaluation of methods for the calculation of static formation temperatures in geothermal and oil wells using an extension of the error propagation theory. *Journal of Geochemical Exploration* 89, 398–404.
- Verma, S.P., Díaz-González, L., González-Ramírez, R., 2009. Relative efficiency of single-outlier discordancy tests for processing geochemical data on reference materials. *Geostandards and Geoanalytical Research* 33, 29–49.
- Wisian, K.W., Blackwell, D.D., Bellani, S., Henfling, J.A., Normann, R.A., Lysne, P.C., Forster, A., Schrotter, J., 1998. Field comparison of conventional and new technology temperature logging systems. *Geothermics* 27, 131–141.
- Yilmaz, S., Demircioglu, C., Akin, S., 2002. Application of artificial neural networks to optimum bit selection. *Computers & Geosciences* 28, 261–269.
- Zhang, G., Patuwo, B.E., Hu, M.Y., 1998. Forecasting with artificial neural networks: the state of the art. *International Journal of Forecasting* 14, 35–62.

Anexo B

Bassam, A., Alvarez del Castillo., Santoyo, E., Garcia-Valladares, O., 2011. Determination of flowing pressure gradients inside geothermal wells using artificial neural networks and wellbore simulation tools. *Journal of Petroleum Science and Engineering*, sumidito, en revisión.

Manuscript Number:

Title: Determination of flowing pressure gradients inside geothermal wells using artificial neural networks and wellbore simulation tools

Article Type: Research Paper

Keywords: geothermal energy, Levenberg-Marquardt algorithm, statistics, modeling pressure-drops, artificial intelligence, two-phase flow

Corresponding Author: Dr E. Santoyo, PhD

Corresponding Author's Institution: CIE-UNAM

First Author: A. Bassam, PhD Student

Order of Authors: A. Bassam, PhD Student; A. Álvarez del Castillo, PhD Student; E. Santoyo, PhD; E. Santoyo, PhD; O. García-Valladares, PhD

Abstract: A new predictive approach based on artificial neural networks (ANN) for the determination of flowing pressure gradients inside geothermal wells was successfully developed. Several ANN computational models based on the Levenberg-Marquardt optimization algorithm, and the hyperbolic tangent sigmoid and linear transfer functions were evaluated. Two ANN models (ANN1 and ANN2, characterized by using five and six input variables, respectively; and a common architecture of 9 neurons in the hidden layer) were found to be the most suitable tools for the reliable prediction of geothermal pressure gradients. These ANN models used a limited number of input variables which are commonly available in field measurements (i.e., wellbore geometry and production data: pressure, temperature and mass flow rate). Such ANN models were trained by using a production database which was compiled from several world geothermal fields. Additional wellbore simulation works were also carried out using the same production data and a numerical simulator (GEOWELLS). The pressure gradients predicted by using all the computing tools (ANNs and GEOWELLS) were statistically compared with measured field data. From this matching analysis, it was demonstrated that the ANN2 model provided the most acceptable results (with average prediction errors less than 2.3%) in comparison with those results inferred from ANN1 and GEOWELL tools. Details of the computational methodology developed in this study, as well as the numerical validation, and the comparative statistical analysis are fully described.



CENTRO DE INVESTIGACIÓN EN ENERGÍA
Universidad Nacional Autónoma de México



December 16th, 2011

Prof. Dr. Birol Dindoruk
Editor in Chief
Journal of Petroleum Science and Engineering

We would like to submit the manuscript “Determination of flowing pressure gradients inside geothermal wells using artificial neural networks and wellbore simulation tools” by A. Bassam, A. Álvarez del Castillo, E. Santoyo, O. García-Valladares”, for its possible publication as a paper in your prestigious journal *Journal of Petroleum Science and Engineering*. We are submitting on-line in the Web page of the journal, Word files containing text, tables and figures of the manuscript.

We sincerely hope you will find this manuscript in good shape and up to the high standards of the journal. Please acknowledge its receipt and inform us of its later progress.

Sincerely yours,

Dr. E. Santoyo
Leader of the Geoenergy Research Group/Energy Systems
CIE-UNAM

Corresponding author
Tel: +52-55-56229723; Fax: +52-55-56229791; e-mail: esg@cie.unam.mx

Determination of flowing pressure gradients inside geothermal wells using artificial neural networks and wellbore simulation tools

A. Bassam¹, A. Álvarez del Castillo¹, E. Santoyo^{2*}, O. García-Valladares²

¹Centro de Investigación en Energía (UNAM), Posgrado en Ingeniería (Energía-Geotermia), Privada Xochicalco s/n, Centro, Temixco, Mor., 62580, México; ²Centro de Investigación en Energía (UNAM), Sistemas Energéticos, Privada Xochicalco s/n, Centro, Temixco, Mor., 62580, México.

*Corresponding author: esg@cie.unam.mx

Office-Phone: +52 (55) 5622 9723

Fax-Phone: +52 (55) 5622 9721

Submitted to Journal of Petroleum Science and Engineering

December 16, 2011

Abstract

A new predictive approach based on artificial neural networks (ANN) for the determination of flowing pressure gradients inside geothermal wells was successfully developed. Several ANN computational models based on the Levenberg-Marquardt optimization algorithm, and the hyperbolic tangent sigmoid and linear transfer functions were evaluated. Two ANN models (ANN₁ and ANN₂, characterized by using five and six input variables, respectively; and a common architecture of 9 neurons in the hidden layer) were found to be the most suitable tools for the reliable prediction of geothermal pressure gradients. These ANN models used a limited number of input variables which are commonly available in field measurements (i.e., wellbore geometry and production data: pressure, temperature and mass flow rate). Such ANN models were trained by using a production database which was compiled from several world geothermal fields. Additional wellbore simulation works were also carried out using the same production data and a numerical simulator (GEOWELLS). The pressure gradients predicted by using all the computing tools (ANNs and GEOWELLS) were statistically compared with measured field data. From this matching analysis, it was demonstrated that the ANN₂ model provided the most acceptable results (with average prediction errors less than 2.3%) in comparison with those results inferred from ANN₁ and GEOWELL tools. Details of the computational methodology developed in this study, as well as the numerical validation, and the comparative statistical analysis are fully described.

Keywords: *geothermal energy, Levenberg-Marquardt algorithm, statistics, modeling pressure-drops, artificial intelligence, two-phase flow*

1. Introduction

The evaluation of the production performance of geothermal wells is a crucial engineering task required for the optimized exploitation of their resources (e.g., Khasani et al., 2002; García-Valladares et al., 2006; Jahanandish et al., 2011). This engineering task is normally required for several geothermal applications, such as: (i) the optimum design of the wellbore geometry and surface equipment (Gunn et al., 1992); (ii) the identification of permeable feed zones in the geothermal reservoir (Grant et al., 1983); (iii) the reliable determination of the flashing point (García-Valladares et al., 2006); (iv) the prediction of productivity curves for the analysis of a production declining problem (Iglesias and Moya, 1998); (v) the early identification of scaling problems inside the production wells (Grant et al., 1983); (vi) the timely detection of skin factor problems (Aragón et al., 2008); (vii) the identification of secondary feed zones in the production wells (Bjornsson and Bodvarsson, 1987); (viii) the sustainable exploitation of geothermal resources for protecting the investment of commercial projects (Axelsson, 2010); among others.

Fluid and heat flows inside geothermal wells play an important role for a better optimization of their resources (e.g., Hadgu and Freeston, 1990; Hadgu et al., 1995; Garg et al., 2004; Hasan and Kabir, 2010). Wellbore production data are commonly obtained from stable flowing measurements performed in the field (García-Valladares et al., 2006). Within this context, the accurate knowledge of temperature and pressure is crucially important to support the above described reservoir engineering tasks (Gunn et al., 1992). Temperature and pressure gradients are usually measured using complex and costly logging techniques (Jung et al., 2001).

Numerical modeling is a well-known methodology used for predicting most of these thermodynamic parameters. Wellbore numerical simulators are widely recognized as the most effective, practical, and cheapest tools for analyzing fluid and heat flow processes, in terms of the flowing pressure and temperature gradients.

Several wellbore simulators, based on single- and two-phase flow models, have been developed for predicting such gradients (e.g., Barelli et al., 1982; Ortíz-Ramírez, 1983; Hadgu and Freeston, 1990; Gunn et al., 1992; Hadgu et al., 1995; García et al., 1999; García-Valladares, 2007; Hasan and Kabir, 2007, 2010; among others). These numerical simulation tools have the advantage of minimizing the number of field measurements in geothermal wells (Gunn et al., 1992). However, it is important to recognize that the use of numerical simulators still present several limitations due to the differences commonly found between simulated or predicted and measured data (Takacs, 2001; Lu et al., 2006; Álvarez del Castillo et al., 2010). These differences are typically caused by several sources, such as:

- (1) the use of unsuitable empirical equations for describing fluid and heat flows that actually occur inside the geothermal wells (Dalkilic et al., 2009),
- (2) the use of unrealistic assumptions that do not describe the actual physical phenomena occurring inside the geothermal system (Gunn et al., 1992),
- (3) the limited number of field measurements and the associated experimental errors (Jung et al., 2001),
- (4) the thermodynamic and transport properties of geothermal fluids and wellbore materials (e.g., formation, pipe, cement, drilling fluids (Tian and Finger, 2000), and
- (5) the effect of static formation temperatures (or geothermal gradients) on the flowing pressure changes, especially for the analysis of production wellbore data under non-adiabatic flow conditions (Grant et al., 1983), among others.

On the other hand, powerful computational techniques of artificial intelligence (e.g., ANN, fuzzy logic, etc.) have started to be applied in some engineering studies related to the exploitation of petroleum and geothermal systems (e.g., Farshad et al., 2000; Zellou and Ouenes, 2003; Spichak and Goidina, 2005; Jeirani and Mohebbi, 2006; Bassam et al. 2010; Álvarez del Castillo et al., 2011; Jahanandish et al., 2011).

Farshad et al. (2000) designed a novel approach of ANNs for predicting temperature profiles in flowing petroleum wells. Zellou and Ouenes (2003) developed a new methodology for the characterization of integrated fracture reservoirs using neural networks and fuzzy logic techniques. Jeirani and Mohebbi (2006) used the training of ANN with pressure build-up data for the estimation of the initial reservoir pressure, the formation permeability, and the skin factor damage. Bassam et al. (2010) developed a new equation based on ANN for the prediction of static formation temperatures (SFT) in geothermal wells. Álvarez del Castillo et al. (2011) proposed a new empirical void fraction correlation using ANN tools for the analysis of two-phase flows of geothermal wells. Jahanandish et al. (2011) reported an ANN model for the prediction of the bottomhole flowing pressure and pressure drops in vertical multiphase flowing wells.

Regarding this literature review, it was found that the ANNs have been mostly applied for the evaluation of petroleum wells (Farshad et al., 2000; Jeirani and Mohebbi, 2006; and Jahanandish et al., 2011). Up to our knowledge, the application of ANN techniques in geothermal studies has been limited, and therefore, it still constitutes a fertile scientific field for novelty contributions.

The aim of this investigation is to report a new approach of ANN techniques for the reliable determination of flowing pressure gradients in geothermal wells with inclined and vertical geometries. The advantage of this new study was to demonstrate the applicability of the ANN models using a limited number of input variables, which are commonly measured in actual field works. With these purposes, a computational methodology is here reported in terms of the ANN architectures developed, the numerical algorithms, the learning-testing procedures, and the statistical analyses performed for validating the predicted results. A sensitivity analysis of the best ANN architectures was also performed to identify the most important weighted input variables. The pressure gradients inferred from the ANN models were statistically compared with both measured field data and numerical results predicted with a wellbore numerical simulator (GEOWELLS). These numerical results are outlined in detail.

2. Theoretical background

2.1 Flowing pressure gradients

Numerous research works and wellbore numerical simulators have been developed for the prediction of flowing pressure gradients in geothermal wells (e.g., Freeston and Hadgu 1988; Hasan and Kabir, 2010). To determine the flowing pressure gradients, most of these simulators solve momentum balance equations subjected to a set of initial and boundary wellbore conditions (García-Valladares et al., 2006). According to Currie (1974), the general momentum equation is given as:

$$\rho \left[\frac{\partial \vec{v}}{\partial t} + (\vec{v} \cdot \nabla) \vec{v} \right] = -\nabla P + \nabla \cdot \tilde{\tau} + \rho \vec{f} \quad (1)$$

As a simplified approach to this equation (see Appendix A), the total pressure gradient in producing geothermal wells has been usually estimated under one-dimensional conditions by using the following equation:

$$-\left(\frac{dP}{dz} \right)_T = \left[\rho_m f \frac{V^2}{2D} \right]_f + \left[\rho_m V \frac{dV}{dz} \right]_a + [\rho_m g \sin \theta]_g \quad (2)$$

Three right terms of Eq. (2) represents the flowing pressure gradients due to friction, acceleration, and gravitational forces (Wallis, 1969). The variables of this equation, including their measurement units, are described in the nomenclature section, whereas the Darcy friction factor (f) is additionally described in Appendix B.

For a reliable prediction of the flowing pressure gradients, the solution of Eq. (2) requires a correct knowledge of some crucial two-phase flow parameters (e.g., the liquid hold-up or void fraction, and the fluid thermodynamic-transport properties). In two-phase flow, liquid and gas phases could flow with different velocities. Gas phase (having lower density), flows faster than the liquid phase resulting a process where gas phase holds up the liquid phase.

The fluid flow process is therefore defined by a liquid hold-up fraction, whereas the difference in velocity is known as slip (Hasan and Kabir, 2007). The knowledge of these two-phase flow parameters is very difficult to obtain under geothermal conditions, which have pushed for searching other practical or analytical methods to predict reliably the wellbore pressure gradients (e.g., Garg et al., 2004; Álvarez del Castillo et al., 2011). A simplified method to estimate the two-phase flow parameters is also provided by the alternative use of the classical homogeneous model (Hasan and Kabir, 2007). Numerical schemes used for the solution of governing equations of wellbore simulators are additionally important for improving their prediction capability in terms of accuracy, stability, and computational speed.

2.2 GEOthermal WELLbore Simulator (GEOWELLS)

García-Valladares et al. (2006) developed the wellbore simulator GEOWELLS with the following numerical capabilities: (a) Single and two-phase flow can be modeled; (b) Wellbore simulations from the bottom to the wellhead section with an adaptable depth mesh (up to 0.1 m; without stability problems) can be analyzed. However, a reverse simulation could be easily performed using an iterative method with suitable boundary conditions; (c) Multiple changes of wellbore geometry (casing, diameters, and roughness), including deviated wells can be considered; (d) Transient heat losses between wellbore and the surrounding formation can also be estimated, and (e) Simulated data (pressure, temperature, enthalpy, steam quality, mass flow rate, flow regime, as well as friction, acceleration and gravitation pressure gradients) at each wellbore depth may also be quantified.

For obtaining such output production data, the following physical and thermodynamic considerations are assumed:

- (i) One-dimensional and transient fluid and heat flows are used to determine pressure and temperature gradients [i.e., $T(z,t)$, $p(z,t)$, etc.];
- (ii) Two-phase flow can be analyzed by using either homogeneous or drift flux models;
- (iii) Thermodynamic behavior of geothermal fluids is approached by using water properties;

- (iv) Both radiant heat exchange and axial heat conduction inside the fluid flow are neglected; and
- (v) The heat conduction through the wellbore materials (pipe, cement and rock formation) are calculated by solving a three-dimensional transient axial-symmetric temperature distribution.

The mathematical formulation (mass, momentum and energy balances) used for describing the fluid and heat flows inside a control volume (CV) of a typical geothermal well is fully described by Garcia-Valladares et al. (2006). In the present study, the pressure gradients of two-phase flow zone (Eq. 2) were calculated by using the classical homogenous model, which has been widely employed in previous simulation works (e.g., Hasan and Kabir, 2007). The use of homogenous flow model was especially justified under bubble flow pattern conditions, which reduced the complexity of the two-phase flow problem without introducing any significant inaccuracy (Hasan and Kabir, 2007; Timlin, 2008).

The density of the two-phase mixture (ρ_m) was estimated by using the conventional equation:

$$\rho_m = \rho_l(1 - \alpha_H) + \rho_s \alpha_H \quad (3)$$

where ρ_l and ρ_s are the liquid and steam densities, respectively; and α_H is the homogeneous void fraction which was determined as a function of the steam quality (x) as follows:

$$\alpha_H = \frac{1}{1 + \left(\frac{1-x}{x} \right) \left(\frac{\rho_s}{\rho_l} \right)} \quad (4)$$

2.3 Artificial neural networks (ANNs)

A neural network is a massively parallel distributed processor, which has a natural tendency for learning experimental knowledge, which can be later used in a wide variety of engineering applications (Haykin, 1999).

ANNs can be trained to solve non-linear and multivariate problems associated with complex physical phenomena that are difficult either to solve with conventional methods or to carry out under costly experimental conditions. ANNs are composed of simple elements operating in parallel (Demuth and Beale, 1998). These elements are inspired in biological nervous systems. As in nature, a network function is mostly determined by connections between neurons. An elementary neuron with R inputs is shown in Fig. 1. Each input is assigned with an appropriate weighting factor (W). The sum of the weighted inputs and the bias b produces the input for a transfer-function which will generate an output.

The learning process of an ANN is usually carried out by adjusting connection weights among neurons. ANNs are commonly trained for leading a particular input to predict a specific output (or output target): Fig. 2.

For minimizing the differences between the output target (given by measured field data) and the predicted or simulated output (i.e., generated by a weight adjusting process), the well-known Levenberg-Marquardt optimization algorithm (a modified version of the Newton's method) was used (Martin et al., 1994; Demuth and Beale, 2003).

3. Computational methodology

A numerical computational methodology was developed for the determination of flowing pressure gradients inside geothermal wells (Fig. 3). The methodology consists of the four general tasks:

- (i) Creation of a working database with production wellbore data collected from several world geothermal fields;
- (ii) Development and evaluation of ANN models (ANN_1 and ANN_2) for the reliable prediction of flowing pressure gradients of geothermal wells;
- (iii) Determination of flowing pressure gradients using both the wellbore simulator (GEOWELLS), and the two optimized ANN tools developed in this study (ANN_1 and ANN_2); and

- (iv) A comparative statistical analysis between measured field data and predicted pressure gradients (inferred from the application of the ANN and GEOWELLS tools).

(i) *Creation of a working database with production wellbore datasets.* A world geothermal database (WGDB) containing 67 production wellbore datasets was created. Such a WGDB was compiled from a wide variety of international literature sources (e.g., scientific journals, proceedings of geothermal congresses, and technical reports). WGDB includes production wellbore datasets collected from some world geothermal fields, such as Africa (Djibouti: Sánchez-Upton, 2000); China (Freeston and Hadgu, 1988); Iceland (Bjornsson, 1987); Italy (Freeston and Hadgu, 1988; Barelli et al, 1982); Japan (Garg and Combs, 2002; Garg et al., 2004); Mexico (Bjornsson, 1987; Aragón et al. 1999; Sánchez-Upton, 2000; Álvarez del Castillo et al., 2010); New Zealand (Freeston and Hadgu, 1988); and Philippines (Ambastha and Gudmundsson, 1986).

Each wellbore dataset contains an average number of 19 wellbore logs of pressure-depth. From a general point of view, WGDB contains a numerical matrix with a dimension of 1300 rows and 7 columns (not included in the paper but available on request from the authors). From this structure, 57 wellbore datasets were used for the development of the ANN models (here defined as the ANN input database), whereas the remaining 10 datasets were employed for numerical validation purposes. A statistical distribution of the input and output parameters used for the development of the ANN models is summarized in Table 1. Such a statistical distribution is given in terms of the working intervals (min and max), and central tendency (mean and median) and dispersion (standard deviation) parameters.

Wellbore geometry (depth, inclination angle, and diameter), and production data (bottomhole pressure and temperature data, and mass flow rate) were used as main input variables. Wellbore pressure logs were specifically utilized as the target or output variable. All these input data were also used for performing the wellbore numerical simulations with GEOWELLS. It is important to remark, that the input and output variables used in this study have a strong dependency on mechanistic models that actually describe the fluid flows inside geothermal wellbores (Eq. 2).

(ii) *Development and evaluation of ANN models.* Two different ANN models (ANN_1 and ANN_2) were developed for the reliable prediction of flowing pressure gradients in geothermal wells. ANN_1 model was a network characterized by an input layer with five variables, one hidden layer, and an output layer with a single variable (using activation functions that were evaluated during the learning/testing processes), whereas ANN_2 considered an input layer with six input variables and a similar architecture of ANN_1 for the hidden and output layers (Fig. 4). The evaluation of multiple layers of neurons with different combinations of activation functions was suggested as a suitable computational strategy to find out the linear and non-linear relationships between input and output variables (Cortés et al., 2009; Spichak and Popova, 2000).

The input variables of the ANN_1 and ANN_2 models were normalized to work in a closed computing domain $[-1,1]$ (i.e., with a mean equal zero and a standard deviation that depends on the input variable distribution), which constitutes the classical computational method proposed for these type of architectures (e.g., Demuth and Beale, 2003).

(iii) *Determination of flowing pressure gradients using GEOWELLS and ANN models.* From the original working database (see Fig. 3), 10 production wellbore datasets were randomly selected for application and validation purposes (here defined as the validation database). Flowing pressure gradients for these 10 geothermal wells were determined by using both the simulator GEOWELLS and the two ANN models developed.

(iv) *Comparative statistical analysis between measured field data and predicted pressure gradients.* The numerical results obtained from GEOWELLS simulator and the ANN models were statistically compared with measured field data. This analysis was performed by applying the following statistical test parameters: Root Mean Square Error (*RMSE*), the Mean Percentage Error (*MPE*), and Difference Coefficient Ratio: *Theil's U*.

RMSE parameter is a frequently-used measure of the differences between values predicted by a model and the values actually observed (Verma, 2005). *MPE* parameter is the computed average of errors (%) by which estimated predictions of a variable differ from their actual values. The knowledge of this statistical parameter aids to evaluate if the estimated predictions are either

underestimated or overestimated with respect to the actual or expected data. *Theil's U* parameter is used to measure the efficiency of a prediction model when it is compared with others. For *Theil's U* values lower than 1, it indicates, that the error obtained by the predicted model (under evaluation) is lower than those errors obtained by other models (see Table 2). Details on the use of these statistical parameters are fully described by Álvarez del Castillo et al. (2011).

4. Results and discussion

4.1 Evaluation of the two ANN models developed

To perform the learning/testing process of ANN₁ and ANN₂ models (Fig. 3), the input database was randomly divided into learning (with 70 % of the production datasets) and testing (with the remaining 30 %) for obtaining a good representation of the data distribution, and also to avoid bias (Rumelhart, 1986). The learning database was then used for calculating the optimal weighting coefficients and biases of each ANN model, whereas the testing database was employed to evaluate the efficiency of each ANN model. The prediction capability of the ANN models was evidently improved by minimizing the learning and testing errors, which tend to decrease as the number of computing iterations increases.

To select the best activation functions $a^{(1)}$ and $a^{(2)}$ for the structure of hidden-output layers of the ANN models (Fig. 4), a preliminary evaluation of several combinations of transfer functions was performed (i.e., *linear-linear*, *linear-tansig*, *tansig-linear*, and *tansig-tansig*). As a result of this assessment, the *tansig-linear* combination (also referred as hyperbolic tangent sigmoid-linear transfer functions) was found as the most suitable activation functions for the hidden-output layers. It was also found that both models ANN₁ and ANN₂ were efficiently trained using a hidden layer with 9 neurons (see Fig. 4). The objective of using the same number of neurons in the hidden layer (in both ANN₁ and ANN₂ architectures) was to evaluate the effect of an additional input variable (i.e., the bottomhole pressure which was only considered by the ANN₂ model) on the prediction of the pressure gradients. The optimized equation obtained for the ANN₁ and ANN₂ models was given by:

$$a^{(2)} = \sum_m \left[LW^{ho}_{(n,m)} \left(\frac{2}{1 + \exp \left[-2 \left(\sum_k IW^{ih}_{(k,m)} \cdot In_{(k)} + b_{1(m)} \right) \right]} - 1 \right) \right] + b_{2(n)} \quad (5)$$

where the number of neurons for the input layer of ANN₁ and ANN₂ models were $k=5$ and $k=6$, respectively; whereas the number of neurons for the hidden and output layers were $m=9$ and $n=1$, respectively. $IW^{ih}_{(k,m)}$, $LW^{ho}_{(n,m)}$ are the weighting coefficients; and $b_{1(m)}$ and $b_{2(n)}$ are the biases of the neurons (see Table 3).

According to the methodology described in Fig. 2, the coefficients of Eq. (5) were estimated after ~10,000 of computing iterations, which allowed the predicted output variable (P_{sim}) to be closely approached to the desired output (i.e., the measured pressure logs) with a minimum residual error ($RMSE$, defined in Table 2). When $RMSE$ value was less than a convergence error (δ), the predicted output variable (P_{sim}) was defined as the optimized output variable (P_{ANN}).

The prediction capability of the ANN₁ and ANN₂ models was individually evaluated by a linear regression analysis ($y=a+bx$) between the predicted (P_{ANN}) and measured (P_{meas}) pressure data (using the linear determination coefficient: R^2 ; the intercept: a ; the slope: b ; and their corresponding uncertainties. The results obtained for the ANN₁ and ANN₂ models are graphically shown in Figs. 5 and 6, respectively. For the ANN₁ model, the best linear regression equation was given by the following statistical parameters: $a=0.72 \pm 0.20$ and $b=0.972 \pm 0.004$ with a $R^2=0.9708$; whereas for the ANN₂, $a=0.39 \pm 0.13$ and $b=0.988 \pm 0.003$ with a $R^2=0.9874$. According to these results, the ANN₂ model provided a much better prediction than the ANN₁ model. This is clearly observed when the variables P_{meas} and P_{ANN} were statistically compared through the analysis of the ideal straight line ($y=x$). The more the two data sets agree (Fig. 6), the more the scatters tend to concentrate in the vicinity of the ideal line with regression parameters more approached to $a=0$ and $b=1$.

4.2 Sensitivity analysis

To evaluate the relative importance of the input variables on the two ANN models developed (particularly, the effect of the bottomhole pressure variable in the ANN₂ model case), the neural net weight matrix and the well-known Garson equation were used (Elmolla *et al.*, 2010). Garson (1991) suggests an equation based on the partitioning of connection weights given by:

$$I_j = \frac{\sum_{m=1}^{N_h} \left(\left(|IW_{(j,m)}^{ih}| / \sum_{k=1}^{N_i} |IW_{(k,m)}^{ih}| \right) \times |LW_{(n,m)}^{ho}| \right)}{\sum_{k=1}^{N_i} \left\{ \sum_{m=1}^{N_h} \left(\left(|IW_{(k,m)}^{ih}| / \sum_{k=1}^{N_i} |IW_{(k,m)}^{ih}| \right) \times |LW_{(n,m)}^{ho}| \right) \right\}} \quad (6)$$

where, I_j is the relative importance of the j^{th} input variable on the output variable (or ANN target); N_i and N_h are the number of input and hidden neurons, respectively; and IW and LW are the connection weights between the input-hidden and hidden-output layers, respectively (see Table 3). The superscripts ‘ i ’, ‘ h ’ and ‘ o ’ are referred to the input, hidden and output layers, respectively; whereas the subscripts ‘ k ’, ‘ m ’ and ‘ n ’ correspond to the input, hidden and output neurons, respectively.

The weighting factors among the neurons calculated during the learning/testing processes of the two ANN models are included in Table 3. The sensitivity analysis results obtained for the two ANN models are shown in Table 4. Two different groups of input variables (used by the ANN models) are clearly defined: geometric variables (wellbore depth, inclination angle, and wellbore diameter) and thermodynamic variables (mass flow rate, and the bottom-hole temperature and pressure).

From a general point of view, most of these input variables showed an important effect on the prediction of the flowing pressure gradients (ranging from 18.12% to 28.40% for the ANN₁; and from 10.51% to 25.24% for the ANN₂), except the wellbore depth variable (which provided the lowest sensitivity percentages: 4.77% and 8.26% for the ANN₁ and ANN₂, respectively).

If the sensitivity results are individually analyzed by each ANN models (considering their total number of variables), the bottomhole temperature (28.40%) and the bottomhole pressure (25.24%) are the most important input variables for the ANN₁ and ANN₂ models, respectively. When the sensitivity analysis results of the ANN₁ and ANN₂ models were together compared, the thermodynamic variables provided the highest values of relative percentages, being the bottomhole pressure the most important input variable for the prediction of the wellbore flowing pressure gradients. This parameter probably explains why the ANN₂ model provided the better prediction performance ($R^2=0.9874$) which was clearly observed during the learning/testing processes: Fig. 6.

4.3 Validation: determination of flowing pressure gradients using the ANN and GEOWELLS tools.

To evaluate the prediction capability of the two ANN models here developed, the validation database consisting of 10 wellbore datasets was used for the prediction of their production pressure gradients (i.e., Okoy-7 and Ngawha-11: Ambastha and Gudmundsson, 1986; M-39 and KW-2: Bjornsson, 1987; A-12, C-2 and KE1-11: Garg et al., 2004; Az-42: Sánchez-Upton, 2000; N-22T: Garg and Combs, 2002; and W-1 (Barelli et al., 1982). This validation database was not included for the learning/testing process of the two ANN models to avoid bias. For comparison purposes, GEOWELLS simulator was also applied for the determination of the flowing pressure gradients of the same geothermal wells. The numerical runs were carried out using both ANN and GEOWELLS tools. For simplicity and to save space in this paper, only the numerical results obtained for three geothermal wells (Okoy-7, Az-42, and M-39) are graphically shown in Figs. 7-9. All the validation results obtained for the remaining geothermal wells are available upon request from the authors (i.e., complete plots and datasets).

Flowing pressure gradients predicted by using ANN₁, ANN₂, and GEOWELLS, including their measured field data are together shown in Figs. 7-9. From these plots, it is clearly observed that the predicted pressure gradients with the ANN₂ model (P_{ANN}) generally provided the best matching with measured field data: P_{meas} (with average prediction errors less than 2.3%); whereas the results obtained from ANN₁ model were less accurate (ranging from 2.5% for the Okoy 7 well to 15.1% for the M-39 well).

On the other hand, the simulator GEOWELLS provided flowing pressure gradients with a satisfactory accuracy (less than 4%) for only two geothermal wells: Okoy 7 (Fig. 7) and Az-42 (Fig. 8); whereas for the well M-39, GEOWELLS predicted with a less accuracy (with prediction errors up to 35%), which was probably attributable to a wellbore geometry problem (possibly caused by a change in the wellbore diameter due to a scaling problem - previously cited by Goyal et al., 1980 and Bjornsson, 1987). According to the original production reports of these wells (Okoy 7 well: Ambastha and Gudmundsson, 1986; Az-42: Sánchez-Upton, 2000; M-39: Bjornsson, 1987), the bottomhole flow conditions were characterized by single-phase (Az-42 and Okoy 7) and two-phase (M-39) fluid feeds.

4.4 Comparative statistical analysis between predicted pressure gradients and measured field data.

A comprehensive matching analysis of geothermal pressure gradients for all the 10 geothermal wellbore datasets were also carried out. With this purpose, an analysis of residuals was performed using a statistical comparison between predicted and measured pressure gradients. *RMSE*, *MPE*, and *Theil's U* statistical parameters were estimated using all the simulation results provided by the ANN₁, ANN₂, and GEOWELLS tools. Table 5 summarizes all the statistics results obtained from these numerical tools.

After analyzing the statistical residual parameter (*MPE*), it was found that the average errors observed for the predicted pressure gradients (in absolute value) reached up to 41.8%, 14.3%, and 80.7% for the ANN₁, ANN₂ and GEOWELLS tools, respectively. This statistical behavior was also confirmed by analyzing the *RMSE* values, which range from 1.3 to 11.4 bars; 0.5 to 3.8 bars; and 0.2 to 22.8 bars for the ANN₁, ANN₂, and GEOWELLS predictions, respectively. According to the global statistical analysis of residuals (i.e., *MPE* and *RMSE*), it was confirmed that the pressure gradients were systematically better predicted by using the ANN₂ model. This prediction inference was statistically validated by analyzing the results of the *Theil's U* parameter (through the calculation of the difference coefficient ratio, *r*, Table 2).

The final interpretation of the r parameter show that the values obtained for the ANN₁ and GEOWELLS models are mostly lower than 1, which it statistically means that the ANN₂ model provides the closest prediction to the field measured pressure data.

5. Conclusions

A new application of the artificial neural network techniques was developed for the geothermal industry. An optimization of two ANN architectures (ANN₁ and ANN₂) for the reliable determination of flowing pressure gradients in geothermal wells was successfully achieved. The computational structure of the ANN₁ and ANN₂ models were characterized by using a limited number of wellbore production variables (i.e., wellbore geometry and production data: pressure, temperature and mass flow rate). The number of hidden layer neurons and experimental data points used for the training database were optimized to obtain reliable ANN prediction models that minimize costly field measurements and time.

The developed ANN models were able to predict flowing pressure gradients with an acceptable accuracy (especially those provided by ANN₂ with around 2.3%). The flowing pressure gradients predicted by ANN₁ and ANN₂ models, including the numerical results provided by the simulator GEOWELLS were statistically compared with field measurements. From this matching analysis, it was found that the ANN₂ predictions provided the most acceptable results in comparison with those results inferred from ANN₁ and GEOWELLS tools. Statistical results of the comparative and sensitivity analyses for both ANN₁ and ANN₂ models show that the use of the bottomhole pressure as input variable plays an important role for the prediction of the pressure gradients. This input variable (generally available from typical field measurements) is therefore considered as an unavoidable parameter for achieving an efficient learning/testing process in the ANN simulations.

ANN techniques and wellbore numerical simulations can be used as a suitable computing tool for the study of the complex prediction of geothermal pressure gradients under different fluid and heat flow conditions. However, it is also important to recognize that the prediction capability of the ANNs could be significantly improved by an appropriate training with a large number of field measurements. Under such conditions, ANN models could constitute a useful and practical tool for the implementation of smart sensors that enable the flowing pressure gradients in geothermal wells to be reliably determined at a lower cost.

Acknowledgements

The first and second authors wish to thank to the Engineering PhD Program of UNAM and CONACyT for the financial support provided.

Appendix A

From a theoretical approach, pressure gradients are generally obtained from a balance of mechanical energy involving the continuity and momentum equations. According to Currie (1974), a general momentum equation is given by:

$$\rho \left[\frac{\partial \vec{v}}{\partial t} + (\vec{v} \cdot \nabla) \vec{v} \right] = -\nabla P + \nabla \cdot \vec{\tau} + \rho \vec{f} \quad (\text{A.1})$$

where

$$\frac{\partial \vec{v}}{\partial t} + (\vec{v} \cdot \nabla) \vec{v} = \frac{\partial \vec{v}}{\partial t} + v_r \frac{\partial \vec{v}}{\partial r} + \frac{v_\theta}{r} \frac{\partial \vec{v}}{\partial \theta} + v_z \frac{\partial \vec{v}}{\partial z} \quad (\text{A.2})$$

$$\nabla \cdot \vec{\tau} = \frac{1}{c_1 c_2 c_3} \left[\frac{\partial}{\partial x_1} (c_2 c_3 \tau_{x_1 x_3}) + \frac{\partial}{\partial x_2} (c_1 c_3 \tau_{x_2 x_3}) + \frac{\partial}{\partial x_3} (c_1 c_2 \tau_{x_3 x_3}) \right] \quad (\text{A.3})$$

and

$$\begin{aligned} c_1 &= 1 & c_2 &= r & c_3 &= 1 \\ x_1 &= r & x_2 &= \theta & x_3 &= z \end{aligned}$$

For the Eqs. (A.1-A.3), ρ represents the fluid density; v is the velocity; r , θ and z are cylindrical coordinates; C_1 , C_2 and C_3 are constants; t is the time, f is the friction forces; P is the pressure; and $\vec{\tau}$ is the shear stress.

Using cylindrical coordinates and taking into account that the direction of the vertical axis of the pipe is greater than other directions, the momentum equation (A.1) for the component z under steady state conditions is given by:

$$\rho \left[\frac{\partial v_z}{\partial t} + v_r \frac{\partial v_z}{\partial r} + \frac{v_\theta}{r} \frac{\partial v_z}{\partial \theta} + v_z \frac{\partial v_z}{\partial z} \right] = -\frac{\partial P}{\partial z} + \frac{1}{r} \left[\frac{\partial}{\partial r} r \tau_{rz} + \frac{\partial}{\partial \theta} \tau_{\theta z} + \frac{\partial}{\partial z} r \tau_{zz} \right] + \rho g \sin \theta \quad (\text{A.4})$$

A simplified solution for this equation can be derived under the following assumptions:

- a) One-dimensional steady flow. It was assumed that the portion of flow that travels through the pipe is short and quickly reaches the stable conditions. It is also assumed that the vertical direction is the dominant flow direction.
- b) Fluid phases are in thermodynamic equilibrium, which means that vapor and liquid phases are considered at the same temperature.
- c) The inclination angle of the production pipe is given by the last term of Eq. (A.4), which it is referred to the horizontal axis.

These assumptions imply the following arguments:

$$\frac{\partial v_z}{\partial t} = 0, \quad v_r = v_\theta = 0, \quad \tau_{\theta z} = \tau_{zz} = 0$$

Considering the above conditions, Eq. (A.4) can be expressed as:

$$\rho v_z \frac{\partial v_z}{\partial z} = -\frac{\partial P}{\partial z} + \frac{1}{r} \frac{\partial}{\partial r} r \tau_{rz} + \rho g \sin \theta \quad (\text{A.5})$$

where g is the acceleration due to the gravity, and θ represents the wellbore inclination angle. According to Wallis (1969) the momentum equation, commonly used to estimate the profiles of pressure, could be rewritten as a simplified explicit equation:

$$\left[\frac{dP}{dz} \right]_T = \left[\frac{dP}{dz} \right]_f + \left[\frac{dP}{dz} \right]_a + \left[\frac{dP}{dz} \right]_g \quad (\text{A.6})$$

The term on the left side represents the total wellbore pressure gradient (Eq. A.6.1) and the three terms on the right side represent the components of friction, acceleration and gravity of pressure gradient, respectively

$$\left[\frac{\partial P}{\partial z} \right]_T = \frac{\partial P}{\partial z} \quad \text{Total pressure gradient} \quad (\text{A.6.1})$$

$$\left[\frac{\partial P}{\partial z} \right]_f = -\frac{1}{r} \frac{\partial}{\partial r} r \tau_{rz} \quad \text{Frictional pressure gradient} \quad (\text{A.6.2})$$

$$\left[\frac{\partial P}{\partial z} \right]_a = -\rho v_z \frac{\partial v_z}{\partial z} \quad \text{Acceleration pressure gradient} \quad (\text{A.6.3})$$

$$\left[\frac{\partial P}{\partial z} \right]_g = -\rho g \sin \theta \quad \text{Gravitational pressure gradient} \quad (\text{A.6.4})$$

Taking into account the boundary conditions: $\tau_{rz} = \tau_m$ and making some simplifications, we can therefore write the following:

$$\left[\frac{dP}{dz} \right]_f = -\frac{1}{r} \frac{\partial}{\partial r} r \tau_{rz} = -\frac{P_{er}}{A} \tau_m, \quad \left[\frac{dP}{dz} \right]_a = -\frac{A \rho v_z}{A} \frac{\partial v_z}{\partial z} = -\frac{W}{A} \frac{dv_z}{dz} \quad \text{and} \quad \left[\frac{dP}{dz} \right]_g = -\rho_m g \sin \theta$$

where P_{er} is the perimeter, A is the cross-section area; τ_m is the mixture shear stress, and W is the mass flow rate. Substituting these terms in the Eq. (A.6), the total pressure gradient is approach by:

$$\left[\frac{dP}{dz} \right]_T = -\frac{P_{er}}{A} \tau_m - \frac{W}{A} \frac{dv}{dz} - \rho_m g \sin \theta \quad (\text{A.7})$$

According to Hasan and Kabir (2010), Eq. A.7 can be also written in terms of the friction factors and fluid densities and velocities as follows:

$$-\left(\frac{dP}{dz} \right)_T = \rho_m f \frac{v^2}{2D} + \rho_m v \frac{dv}{dz} + \rho_m g \sin \theta \quad (\text{A.8})$$

where D represents the wellbore diameter. This one-dimensional equation generally has been used by many researchers in order to describe the main characteristics of the single- and two-phase flow in geothermal wells (Wallis, 1969; Ambastha and Gudmundsson, 1986; Tian and Finger, 2000; Hasan and Kabir, 2010; among others).

Appendix B

The mathematical formulation used for describing the fluid flow inside of a geothermal well is described by Garcia-Valladares et al. (2006). From Eq. (2) the variable f corresponds to the Darcy friction factor used in the simulator GEOWELLS. For single-phase flow the f is given by the following equation:

$$f = \frac{64}{Re} \quad (\text{B.1})$$

where Re is the Reynolds number:

$$Re = \frac{\rho v D}{\mu} \quad (\text{B.2})$$

where μ is the fluid viscosity. The Eq. (B.1) corresponds to a laminar flow ($Re < 2300$). For transition flow ($2300 < Re < 4000$), the friction factor is:

$$f = 10^{[26067 - 22862Re + 66.307Re^2 - 6.3944Re^3]} \quad (\text{B.3})$$

whereas for turbulent flow ($Re > 4000$), the friction factor is given by the Colebrook equation, which is an implicit equation that combines experimental results of studies of turbulent flow in smooth and rough pipes (Colebrook, 1939). The equation is used to iteratively solve the Darcy–Weisbach friction factor f . This equation is also known as the Colebrook–White:

$$\frac{1}{\sqrt{f}} = 2 \log \left(\frac{D}{e} \right) + 2 \log \left(\frac{9.34}{\left(\frac{e/D}{Re \sqrt{f}} \right)} \right) + 1.14 \quad (\text{B.4})$$

where e is pipe roughness and D is the diameter of the pipe.

For two-phase flow, the friction factor (f) is calculated by the single phase flow correlations (Equations B.1-B.3) and multiplied by a correction factor (ϕ), according to Beattie. (1973):

$$\phi = \left[1 + x \left(\frac{\rho_l}{\rho_s} - 1 \right) \right]^{0.8} \left\{ 1 + x \left[\frac{3.5 \mu_s + 2 \mu_l}{(\mu_s + \mu_l) \rho_s} - 1 \right] \right\}^{0.2} \quad (\text{B.5})$$

where μ_l and μ_s are the liquid and steam viscosity.

References

- Álvarez del Castillo, A., Santoyo, E., García-Valladares, O., 2011. A new empirical void fraction correlation inferred from artificial neural networks for modeling two-phase flow in geothermal wells. *Comput. Geosci.*, in press.
- Ambastha, A.K., Gudmundsson, J.S., 1986. Pressure profiles in two-phase geothermal wells: comparison of field data and model calculations. *Proceedings 11th Workshop on Geothermal Reservoir Engineering*. Stanford, California, pp. 183-188.
- Aragón, A., Moya, S.L., García-Gutiérrez, A., Arellano, V. 2008. A new method of damage determination in geothermal wells from geothermal inflow with application to Los Humeros, Mexico. *Geof. Int.* 47, 371-382.
- Axelsson, G., 2010. Sustainable geothermal utilization - case histories; definitions; research issues and modelling. *Geothermics* 39, 283-291.
- Barelli, A.R., Corsi Del Pizzo, G., Scali, C., 1982. A two-phase flow model for geothermal wells in the presence of non-condensable gas. *Geothermics* 11, 175-191.
- Bassam, A., Santoyo, E., Andaverde, J., Hernández, J.A., Espinoza-Ojeda, O.M., 2010. Estimation of static formation temperatures in geothermal wells by using an artificial neural network approach. *Comput. Geosci.* 36, 1191-1199.
- Beattie, D.R.H., 1973. A note of the calculation of two-phase pressure losses. *Nucl. Eng. Des.* 25, 395-402.
- Bjornsson, G.A., 1987. Multi-feedzone geothermal wellbore simulator. M. Sci. Thesis, Lawrence Berkeley Laboratory. University of California, 102 p.
- Bjornsson, G.A., Bodvarsson, G.S., 1987. A Multi-feedzone wellbore simulator. *TrGRC* 11, 503-507.
- Colebrook, C.F., 1939. Turbulent flow in pipes with particular reference to the transition region between the smooth and rough pipe laws. *J. Inst. Civil Eng.* 11, 133-156.
- Cortés, O., Urquiza, G., Hernández, J.A., 2009. Optimization of operating conditions for compressor performance by means of neural network inverse. *Appl. Energ.* 86, 2487-2493.
- Currie, I.G., 1974. *Fundamental Mechanics of Fluids*, ed. Mc. Graw Hill, Inc., USA.
- Dalkilic, A.S., Laohalertdech, S., Wongwises, S., 2009. Effect of void fraction models on the two-phase friction factor of R134a during condensation in vertical downward flow in a smooth tube. *Int. Commun. Heat Mass Transf.* 35, 921-927.
- Demuth, H., Beale, M., 1998. *Neural Network Toolbox for use with Matlab, User's Guide Version 3*, MathWorks, Inc, USA.
- Demuth, H., Beale, M., 2003. *Neural Network Toolbox for use with Matlab, User's Guide Version 4.1*, MathWorks, Inc, USA.
- Elmolla, E. S. and Chaudhuri, M. and Eltoukhy, M. M., 2010. The use of artificial neural network (ANN) for modeling of COD removal from antibiotic aqueous solution by the Fenton process. *J. Hazard. Mater.* 179, 127-134.

- Farshad, F.F., Garber, J.D., Lorde, J.N., 2000. Predicting temperature profiles in producing oil wells using artificial neural networks. *Eng. Comput.* 17, 735-754.
- Freeston D.H., Hadgu T., 1988. Comparison of results from some wellbore simulators using a data bank. *Proceedings 10th New Zealand Geothermal Workshop*, pp. 299-305.
- García, A., Ascencio, F., Espinosa, G., Santoyo, E., Gutiérrez, H., Arellano, V., 1999. Numerical modeling of high-temperature deep wells in the Cerro Prieto geothermal field, Mexico. *Geof. Int.* 38, 251-260.
- García-Valladares, O., Sánchez-Upton, P., Santoyo, E., 2006. Numerical modeling of flow processes inside geothermal wells: An approach for predicting production characteristics with uncertainties. *Energ. Convers. Manage.* 47, 1621-1643.
- García-Valladares, O., 2007. Numerical simulation and experimental validation of coiled adiabatic capillary tubes. *Appl. Therm. Eng.* 27, 1062–1071.
- Garg, S.K., Combs, J.A., 2002. Study of production/injection data from slim holes and large-diameter wells at the Okuaizu geothermal field, Tohoku, Japan. LLC Report, INEEL/EXT-02-01429, Idaho National Engineering and Environmental Laboratory, Bechtel BWXT, Idaho, USA, 257 p.
- Garg, S.K., Pritchett, J.W., Alexander, J.H., 2004. A new liquid hold-up correlation for geothermal wells. *Geothermics* 33, 795-817.
- Garson, G.D., 1991. Interpreting neural-network connection weights. *Artif. Int. Expert* 6. 7, 47-51.
- Grant, M. A. Bixley, P. F., Donaldson, L. G., 1983. Internal flows in geothermal wells: their identification and effect on the wellbore temperature and pressure profiles. *Soc. Petrol. Eng. J.* 23, 168-176.
- Gunn, M.I.C., Freeston, H.D., Hadgu, T., 1992. Principles for wellbore simulator validation and calibration using matching analysis—I. Analytical techniques. *Geothermics* 21, 341-361.
- Hadgu, T., Freeston, D., 1990. A multi-purpose wellbore simulator. *TrGRC* 14, 1279-1286.
- Hadgu, T., Zimmerman, W.R., Bodvarsson, S.G., 1995. Coupled reservoir-wellbore simulation of geothermal reservoir behavior. *Geothermics* 24, 145-166.
- Hasan, A.R., Kabir, C.S., 2007. A simple model for annular two-phase flow in wellbores. *Soc. Petrol. Eng.* 22, 168–175.
- Hasan, A.R., Kabir, C.S., 2010. Modeling two-phase fluid and heat flows in geothermal wells. *J. Petrol. Sci. Eng.* 71, 77–86.
- Haykin S., 1999. *Neural networks: A Comprehensive Foundation*, second ed. Prentice Hall, New Jersey, USA.
- Iglesias, E.R., Moya, S.L., 1998. Applicability of geothermal inflow performance reference curves to CO₂-bearing reservoirs, *Geothermics* 27, 305-315.
- Jahanandish, I., Salimifard, B., Jalalifar, H., 2011. Predicting bottomhole pressure in vertical multiphase flowing wells using artificial neural networks. *J. Petrol. Sci. Eng.* 75, 336–342.
- Jeirani, Z., Mohebbi, A., 2006. Estimating the initial pressure, permeability and skin factor of oil reservoirs using artificial neural networks. *J. Petrol. Sci. Eng.* 50, 11-20.

- Jung, D.B., Wai, K.W., Howard, W.T., 2001. Geothermal flow metering errors. *TrGRC* 25, 23-25.
- Khasani, R.I., Tanaka, T., Fukuda, M., 2002. An analysis of pressure drops in wellbore under low flow rate conditions shown on a deliverability curve. *Proceedings of the 24th New Zealand Geothermal Workshop, New Zealand*, pp. 115-120.
- Lu, X., Watson, A., Gorin, A.V., Deans, J., 2006. Experimental investigation and numerical modelling of transient two-phase flow in a geysiring geothermal well. *Geothermics* 3, 409-427.
- Martin, T., Hagan, T.M., Mohammad, N.B., 1994. Training feedforward networks with the Marquardt Algorithm. *IEEE Trans. Neural Netw.* 6, 989-993.
- Ortíz-Ramírez J., 1983. Two-phase flow in geothermal wells: development and uses of a computer code. M.Sc. Thesis, University of Stanford, CA, USA, 89 p.
- Rumelhart, D.E., Hinton, G.E., Williams R.J., 1986. Learning internal representations by error propagation. In: *Parallel distributed processing: explorations in the microstructure of cognition*. MIT Press Cambridge, MA, USA, 318-362.
- Sánchez-Upton, P., 2000. The wellbore simulator SIMU2000. In: *Proceedings of the World Geothermal Congress (PaperR0057)*, Kyushu-Tohoku, Japan, pp. 2851-2856.
- Spichak, V.V., Goidina, A.G., 2005. Temperature prediction in geothermal zones from borehole measurements using neural networks. *Phys. Solid Earth* 41, 844–852.
- Spichak, V.V., Popova, I., 2000. Artificial neural network inversion of magnetotelluric data in terms of three-dimensional earth macroparameters. *Geophys. J. Int.* 142, 15-26.
- Takacs, G., 2001. Considerations on the selection of an optimum vertical multiphase pressure drop prediction model for oil wells. *Soc. Petrol. Eng. (Conference Paper)* DOI: 10.2118/68361-MS
- Tian, F.S., Finger, T.J., 2000. Advanced geothermal wellbore hydraulics model. *J. Energ. Resour-ASME.* 122, 142-146.
- Verma, S.P., 2005. *Estadística básica para el manejo de datos experimentales: aplicaciones en Geoquímica (Geoquimiometría)*. 1st ed, UNAM, Mexico.
- Wallis, G.B., 1969. *One-dimensional Two-Phase Flow*. ed. McGraw-Hill, USA.
- Zellou, A.M, Ouenes, A., 2003. Chapter 26 Integrated fractured reservoir characterization using neural networks and fuzzy logic: Three case studies. *Dev. Petrol. Sci.* 51, 583-602.

LIST OF TABLES

- Table 1. Input and output data used for the ANN model training and the numerical simulations.
- Table 2. Optimized weighting and bias coefficients computed by the ANN models during the learning /testing process. The superscripts *ih* and *ho* corresponds to the *input-hidden* layers and *hidden-output* layers, respectively.
- Table 3. Results of the sensibility analysis and the ANN evaluation of the relative importance of input variables on the output prediction.
- Table 4. Definition of the statistical residual parameters used for the evaluation of the ANN models and the GEOWELLS simulator
- Table 5. Results of the comparative statistical analysis between the measured and the simulated data (ANN₁, ANN₂ and GEOWELLS).

FIGURE CAPTIONS

- Figure 1. An elementary neuron with R input.
- Figure 2. Numerical procedure used for the ANN learning process, and the iterative architecture used by the ANN model to predict flowing pressure gradients of geothermal wells (S : number of the neurons in the hidden layer, and δ : converge criteria).
- Figure 3. Computational methodology developed for the determination of the flowing pressure gradients inside the geothermal wells (N : number of iterations: 10000); S : Number of neuron in the hidden layer; and δ : converge criteria).
- Figure 4. Computing architecture of the ANN models developed.
- Figure 5. Evaluation of the ANN₁ model during the learning process using the linear regression analysis between measured and simulated pressure gradients.
- Figure 6. Evaluation of the ANN₂ model during the learning process using the linear regression analysis between measured and simulated pressure gradients.
- Figure 7. Flowing pressure gradients predicted for the geothermal well Okoy-7 by using ANN models and GEOWELLS simulator.
- Figure 8. Flowing pressure gradients predicted for the geothermal well Az-42 by using ANN models and GEOWELLS simulator.
- Figure 9. Flowing pressure gradients predicted for the geothermal well M-39 by using ANN models and GEOWELLS simulator.

Highlights

- New application of ANNs for the determination of geothermal pressure gradients
- A general computational methodology to obtain optimized ANNs is reported
- ANNs and wellbore simulations were together used to predict pressure gradients
- Computing results (from ANN and simulator) were compared with field measurements
- Pressure gradients predicted by ANN₂ model gave a better agreement with field data
- Analysis of residuals between predicted and field data were used as solid evaluation

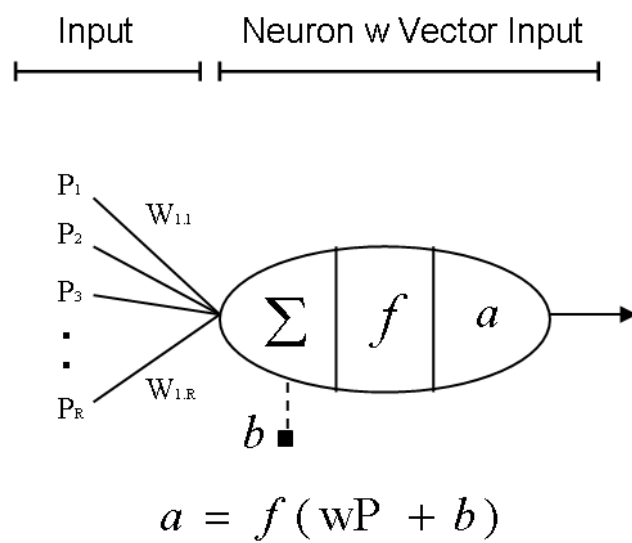


Fig. 1

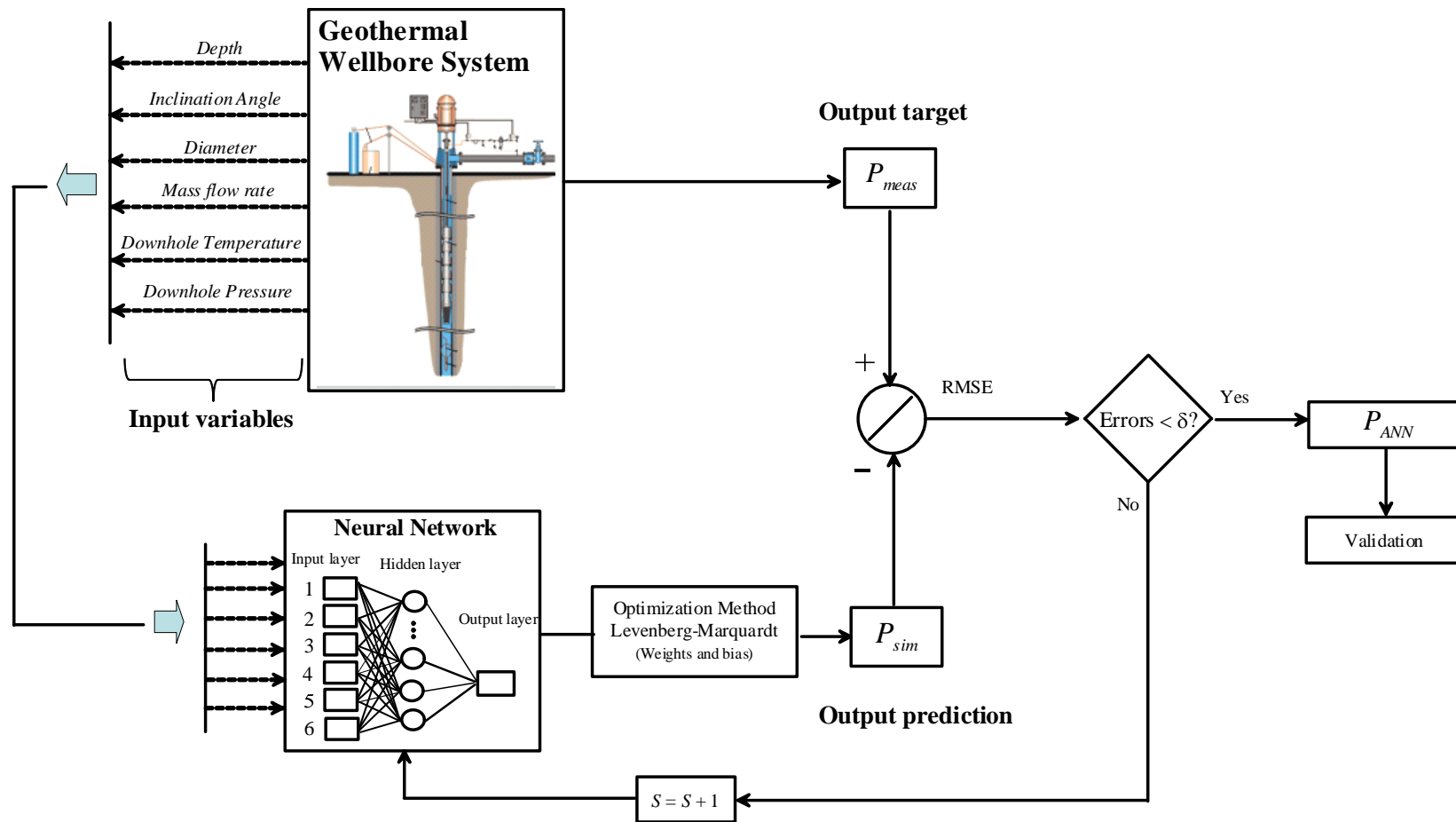


Fig. 2

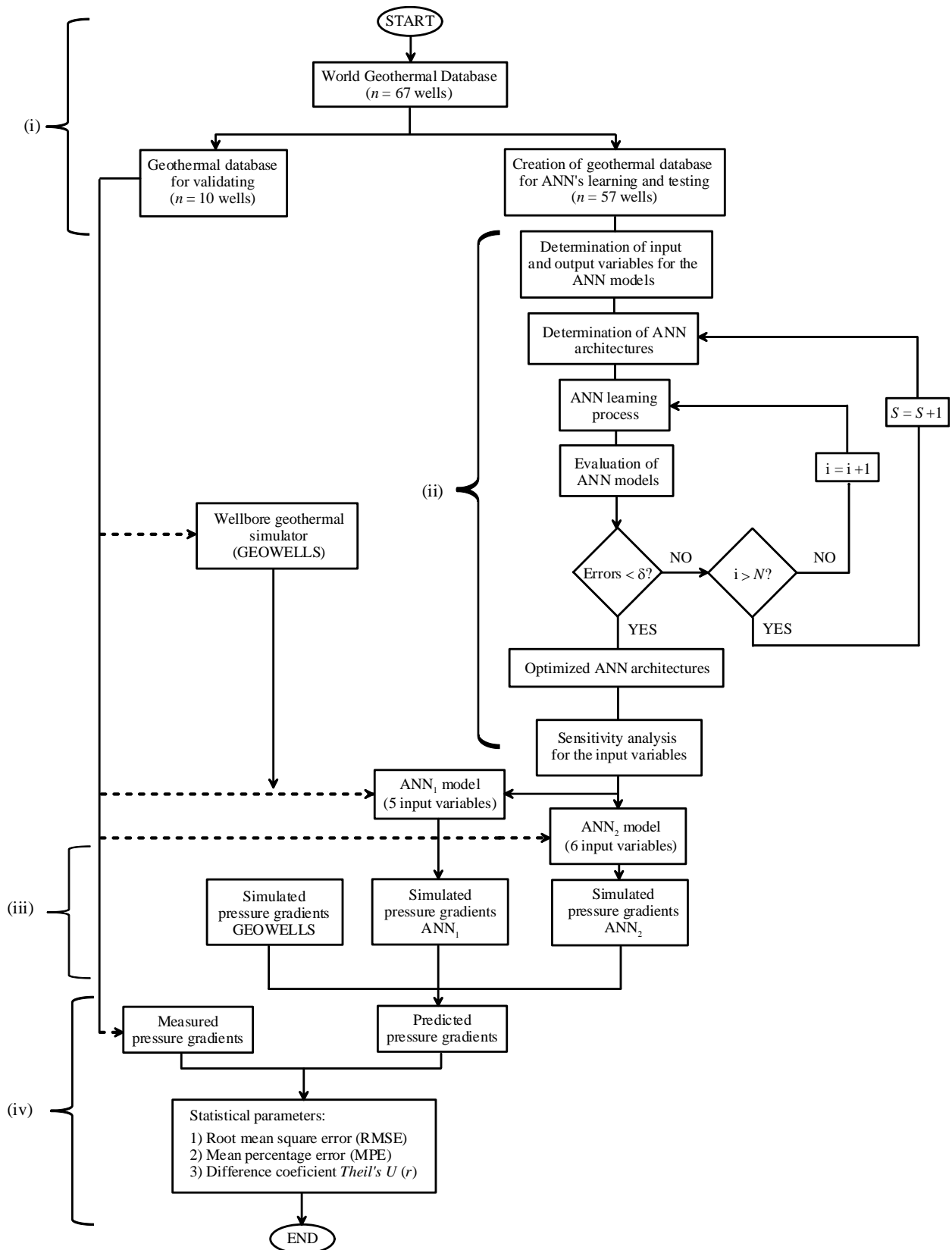


Fig. 3

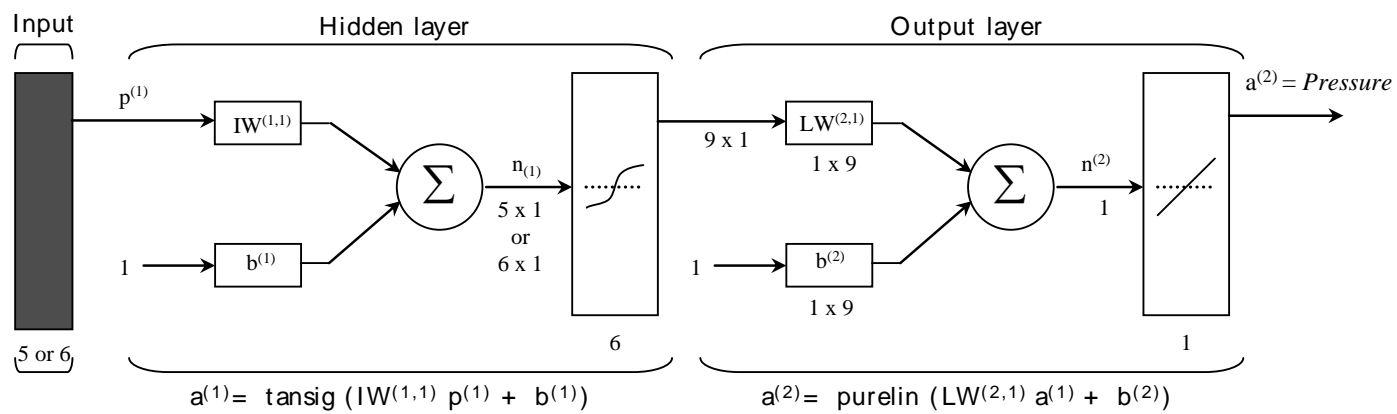


Fig. 4

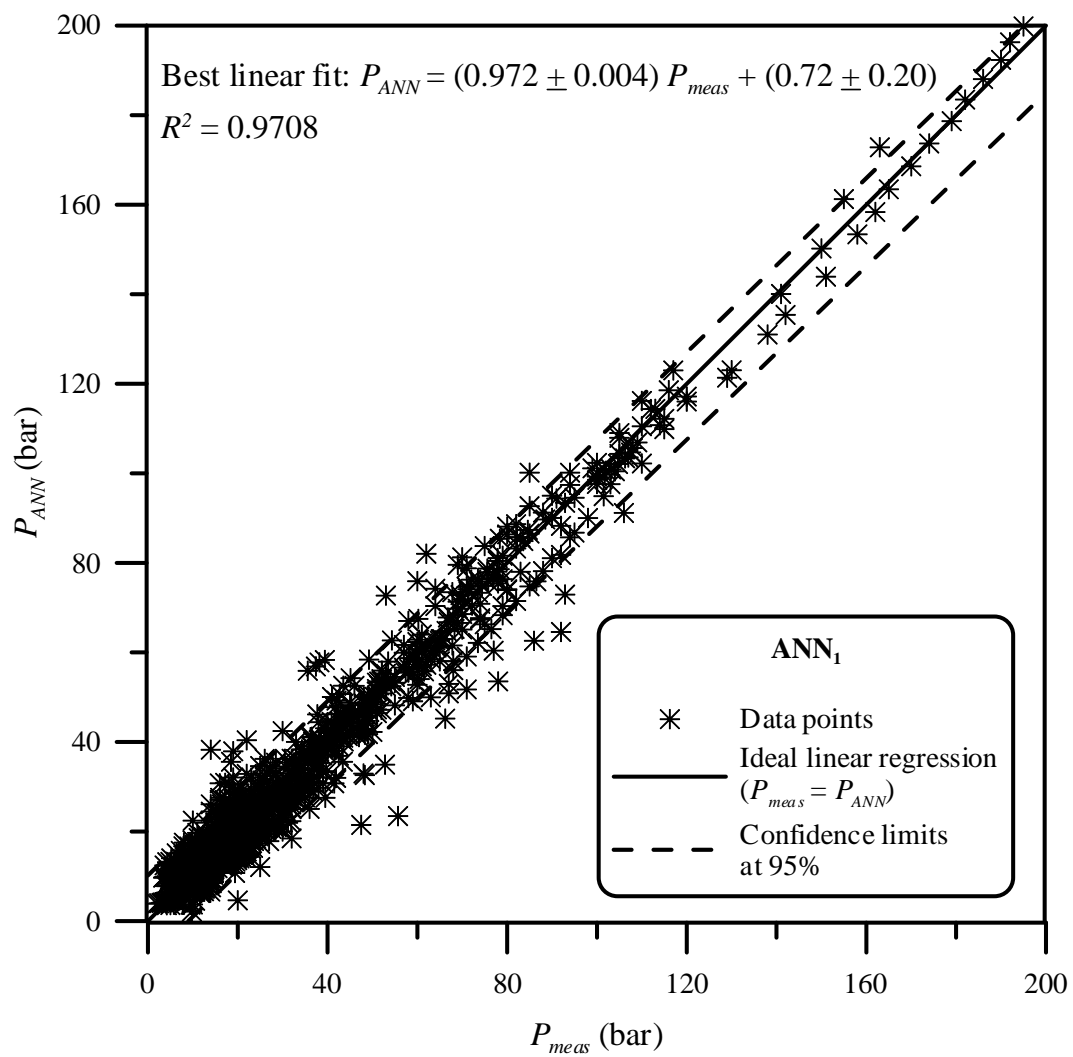


Fig. 5

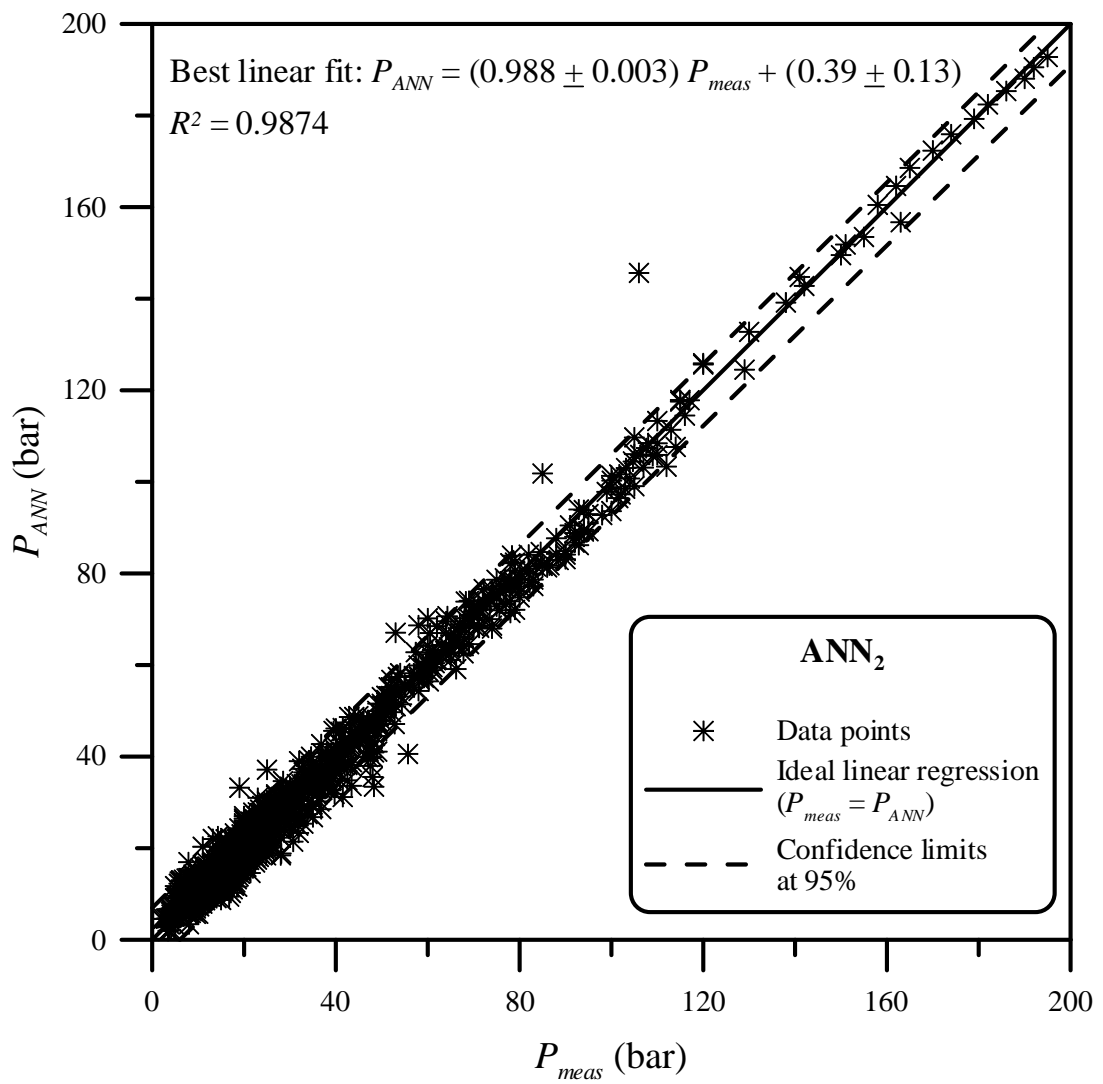


Fig. 6

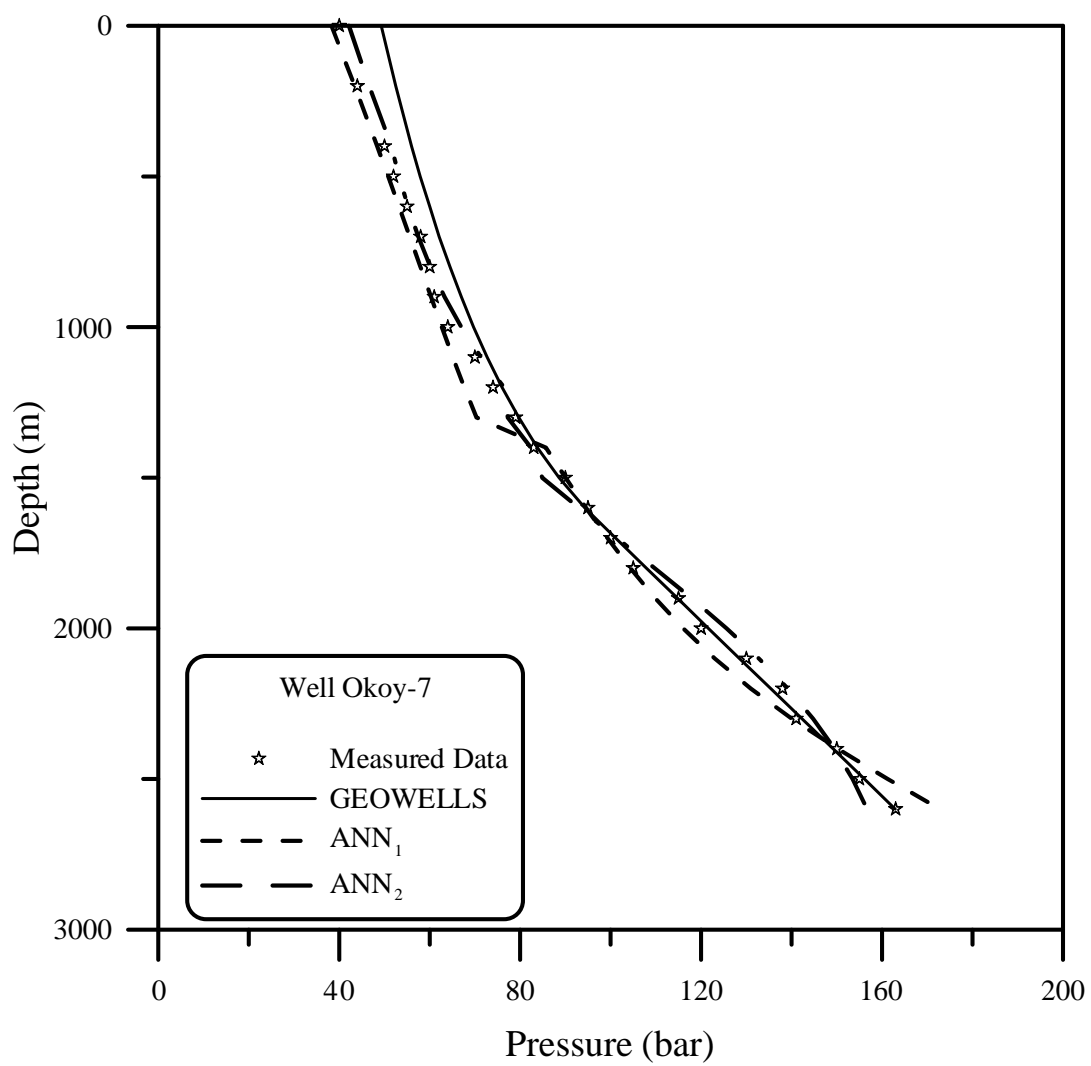


Fig. 7

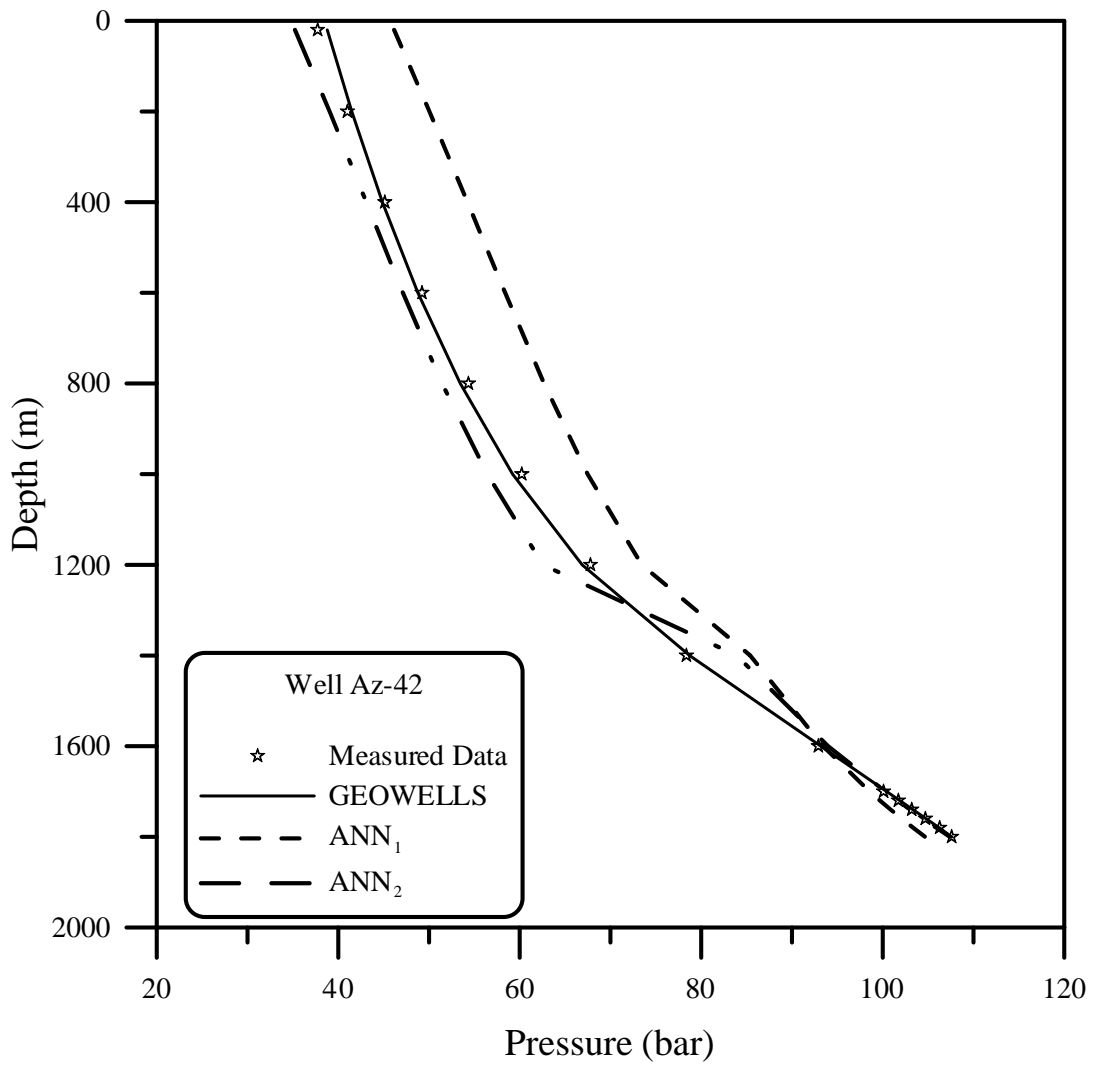


Fig. 8

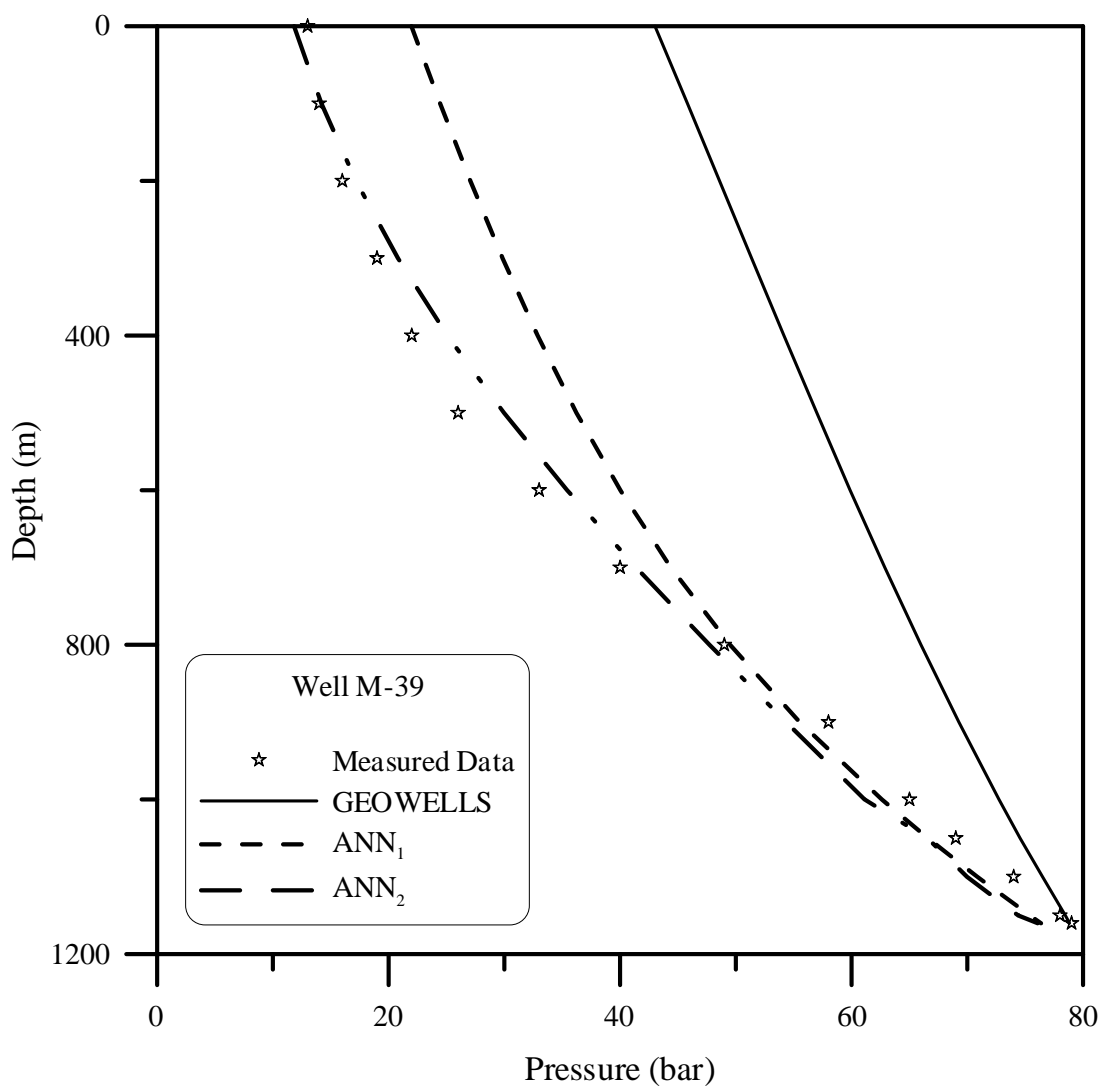


Fig. 9

Table 1

Parameters	n	min	max	mean	median	s*	units
<i>Input parameters:</i>							
Wellbore depth (z)	1128	0	2600	637.77	600	46.40	[m]
Inclination angle (θ)	111	0	32.10	5.16	0.00	7.28	[°]
Wellbore diameter (D)	110	0.38	0.10	0.24	0.22	0.07	[m]
Mass flow rate (W)	57	1.10	202.00	53.17	34.00	47.32	[kg/s]
Bottom-hole temperature (T_b)	57	104.4	364.1	232.1	224.5	39.7	[°C]
Bottom-hole pressure (P_b)	57	6.7	196.3	47.0	29.7	39.5	[bar]
<i>Output parameters:</i>							
Pressure (P_{meas})	1128	2.0	95.0	30.0	18.5	28.8	[bar]

*Standard deviation

Table 3

Parameters	Relative importance (%)	
	ANN₁	ANN₂
<i>Geometric variables</i>		
Wellbore depth (z)	4.77	8.26
Inclination angle (θ)	18.12	15.73
Wellbore diameter (D)	23.59	10.51
<i>Thermodynamic variables</i>		
Mass flow rate (W)	25.13	17.60
Bottom-hole temperature (T_b)	28.40	22.66
Bottom-hole pressure (P_b)	-	25.24

Table 4

Statistical residual parameters	Equation
Root of mean square error (RMSE)	$RMSE = \sqrt{\frac{\sum_{i=1}^n (P_{sim(i)} - P_{meas(i)})^2}{n}}$
Mean percentage error (MPE)	$MPE = \frac{1}{n} \sum_{i=1}^n \left(\frac{P_{sim(i)} - P_{meas(i)}}{P_{sim(i)}} \right) 100$
Difference coefficient (<i>Theil's U</i>) - r	$r = \frac{\sqrt{\sum_{i=1}^n (P_{sim(i)}^{ANN_2} - P_{meas(i)})^2}}{\sqrt{\sum_{i=1}^n (P_{sim(i)}^c - P_{meas(i)})^2}}$

Where P^c is the pressure value estimated either by GEOWELLS or ANN₁ model.

Table 5

Geothermal well (Country) / Statistical residual parameter	Computing tools		
	ANN ₂	ANN ₁	GEOWELLS
KW-2 (Iceland)			
<i>RMSE</i>	3.68	11.42	3.08
<i>MPE</i>	2.60	-8.26	-0.34
<i>THEIL'S U (r)</i>	-	0.32	1.19
Okoy-7 (Philippines)			
<i>RMSE</i>	3.68	4.22	4.01
<i>MPE</i>	2.60	-2.54	4.14
<i>Theil's U (r)</i>	-	0.80	0.85
A-12 (Japan)			
<i>RMSE</i>	1.57	1.73	0.58
<i>MPE</i>	-10.96	-5.32	3.51
<i>Theil's U (r)</i>	-	0.90	2.72
Az-42 (Mexico)			
<i>RMSE</i>	2.59	6.11	0.56
<i>MPE</i>	-2.34	6.35	-0.06
<i>Theil's U (r)</i>	-	0.42	4.66
C-2 (Japan)			
<i>RMSE</i>	0.52	1.32	0.21
<i>MPE</i>	-3.55	-10.79	1.04
<i>Theil's U (r)</i>	-	0.40	2.56
KE1-11 (Japan)			
<i>RMSE</i>	2.03	3.82	8.47
<i>MPE</i>	-2.83	9.77	-80.70
<i>Theil's U (r)</i>	-	0.53	0.24
M-39 (Mexico)			
<i>RMSE</i>	3.80	7.13	22.77
<i>MPE</i>	-0.50	15.15	35.20
<i>Theil's U (r)</i>	-	0.53	0.17
N-22T (Japan)			
<i>RMSE</i>	3.48	2.02	9.32
<i>MPE</i>	14.27	-2.63	30.74
<i>Theil's U (r)</i>	-	1.72	0.37
Ngawha-11 (New Zeland)			
<i>RMSE</i>	2.96	7.41	3.75
<i>MPE</i>	3.53	-41.79	-2.89
<i>Theil's U (r)</i>	-	0.40	0.79
W-1 (Italy)			
<i>RMSE</i>	3.07	4.88	11.10
<i>MPE</i>	-2.06	-0.39	-48.46
<i>Theil's U (r)</i>	-	0.63	0.28

Anexo C

Participaciones en congresos

Memorias en congresos

- ♣ A. Bassam, A. Álvarez del Castillo, O. García-Valladares, E. Santoyo. Determination of flowing pressure gradients in producing geothermal wells by using artificial neural networks. Water-Rock Interaction XIII, October 01, 2010 by CRC Press, Guanajuato, México.

http://www.crcpress.com/product/isbn/9780415604260;jsessionid=Pit2c6qgsVs62n5BsmWJrw**

- ♣ A. Bassam, J.A. Andaverde-Arredondo, E. Santoyo. Metodología numérico-estadística para la determinación del flujo de calor en sistemas geotérmicos usando simulación Monte Carlo para la propagación de errores. Congreso Nacional de Geoquímica por el Instituto Nacional de Geoquímica A.C. Monclova, Coahuila del 3-7 de Octubre, 2011. México.

<http://www.inageq.org.mx/inageq/actas.html>

Premios y Distinciones recibidas

Premio a la mejor ponencia de nivel Doctorado otorgada en el XXI Congreso Nacional de Geoquímica por el Instituto Nacional de Geoquímica A.C. celebrado en la Facultad de Metalurgia, Universidad Autónoma de Coahuila, Monclova, Coahuila del 3-7 de Octubre, 2011. México.