

**HERRAMIENTA SOFTWARE PARA LA ESTIMACIÓN DE VELOCIDADES DE
CORROSIÓN POR CO₂ Y H₂S MEDIANTE EL USO DE TÉCNICAS BASADAS
EN INTELIGENCIA ARTIFICIAL**

CHRISTIAN ANDRES ARIZA QUITIAN

UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FISICOMECÁNICAS
ESCUELA DE INGENIERIA DE SISTEMAS E INFORMÁTICA
BUCARAMANGA

2012

**HERRAMIENTA SOFTWARE PARA LA ESTIMACIÓN DE VELOCIDADES DE
CORROSIÓN POR CO₂ Y H₂S MEDIANTE EL USO DE TÉCNICAS BASADAS
EN INTELIGENCIA ARTIFICIAL**

CHRISTIAN ANDRES ARIZA QUITIAN

Trabajo de grado para optar al título de
Ingeniería de Sistemas

Director:

MSc. Lola Xiomara Bautista Rozo

Codirector:

Ing. Laura Viviana Galvis Carreño

UNIVERSIDAD INDUSTRIAL DE SANTANDER

FACULTAD DE INGENIERÍAS FISICOMECÁNICAS

ESCUELA DE INGENIERIA DE SISTEMAS E INFORMÁTICA

BUCARAMANGA

2012

AGRADECIMIENTOS

A la energía indescriptible, Dios.

A mi familia, mi madre Etelvina, mi padre Gerardo, mi tía Aura Rosa, mis primas Jaqueline, Eliana y Gabriela y a mi novia Jessica por estar conmigo, apoyarme y darme las fuerzas para seguir adelante en cada etapa de mi vida.

Al Dr. Orlando Aya Ramírez (Dr. Rer. Nat.) por su gran apoyo, liderazgo, paciencia y confianza en la elaboración de la herramienta.

A todo el personal de la C.I.C., especialmente a los Ingenieros Alexi y Beatriz, por la oportunidad de desarrollar este trabajo, por su ayuda y los conocimientos compartidos.

A los ingenieros Iván, Edgar y Laura así como a la profesora Lola por su apoyo en la elaboración de este trabajo.

A todos mis amigos en especial Camilo y Daniel, por esos momentos inolvidables de nuestras vidas y dar larga vida a la fuente de vida que nos mueve, el rock y el metal.

CONTENIDO

INTRODUCCIÓN	15
1. GENERALIDADES	16
1.1 PLANTEAMIENTO DEL PROBLEMA.....	16
1.2 OBJETIVOS	18
1.2.1 Objetivo General	18
1.2.2 Objetivos Específicos	18
2. MARCO TEÓRICO	19
2.1 CORROSIÓN	19
2.1.1 Introducción a la corrosión	19
2.1.2 Proceso corrosivo.....	20
2.1.3 Corrosión por CO ₂ y H ₂ S.....	21
2.1.4 Formas de medir la corrosión	22
2.1.5 Monitoreo de corrosión.....	22
2.1.5.1 Pérdida de peso en cupones.....	23
2.1.5.2 Probeta.....	25
2.2 APROXIMACIÓN FUZZY AL PROCESAMIENTO DE DATOS: CURVAS Y SUPERFICIES FUZZY	27
2.2.1 introducción	27
2.2.2 Curvas Fuzzy de Primer Grado (o primer nivel)	28
2.2.3 Curvas Fuzzy de segundo grado.....	32
2.2.4 Superficies Fuzzy de primer grado (o primer nivel)	33
2.2.5 Superficies Fuzzy de segundo grado	34
2.2.6 Procedimiento de selección de las variables de entrada basado en metodología fuzzy	36
2.3 REDES NEURONALES ARTIFICIALES.....	36
2.3.1 Introducción.....	36

2.3.2 Relación redes neuronales artificiales y neuronas biológicas	38
2.3.2.1 Estructura de la Red Neuronal Artificial.....	38
2.3.2.2 Modo de Operación de la Red	39
2.3.3 Especificación de la estructura de la red neuronal	40
2.3.4 Entrenamiento de la Red.....	41
2.3.4.1 Algoritmo de retro-propagación (back-propagation)	41
2.3.4.2 Algoritmo de Levenberg Marquardt	42
3 DESARROLLO DE LA HERRAMIENTA COMPUTACIONAL	44
3.1 FASE I	45
3.1.1 Cupones en Cabezas de Pozo (Well Head).....	47
3.1.2 Cupones en Manifold.....	49
3.1.3 Cupones en la línea principal	52
3.1.4 Cupones Unidad Central de Procesamiento	55
3.1.5 Análisis de la información	55
3.2 FASE II (Selección de Variables).....	57
3.3 FASE III	59
3.3.1 Normalización de datos	59
3.3.2 Implementación de la red neuronal	60
3.3.3 Ratificación en la estructura neuronal	64
3.4 FASE IV	66
3.4.1 Diagramas UML	67
4 RESULTADOS	72
4.1 Base de Datos Cupones.....	72
4.2 Pre-procesamiento fuzzy de las variables de entrada	73
4.3 Red Neuronal	81
5 CONCLUSIONES	86

6 RECOMENDACIONES.....	87
7 BIBLIOGRAFÍA.....	88
8 ANEXOS.....	90

LISTA DE TABLAS

Tabla 1: Constante K para perdida de espesor.....	25
Tabla 2: Contante <i>k</i> para tasa de corrosión	25
Tabla 3: Variables de entrada y Variable de salida.....	45
Tabla 4: Caso de uso Pre-procesar y Generar Red.....	68
Tabla 5: Caso de uso Generar Red	69
Tabla 6: Caso de uso Determinar Velocidades.....	69
Tabla 7: Caso de uso Cargar Datos.....	70
Tabla 8: Caso de uso Guardar Resultados	70
Tabla 9: Caso de uso Cargar Red	71
Tabla 10: Información recopilada.....	72
Tabla 11: Resultados Curvas Fuzzy de primer Grado para las variables recopiladas.....	76
Tabla 12: Resultados de Curvas Fuzzy de Segundo Grado para las variables recopiladas.....	77
Tabla 13: Índice único de Efectividad Curvas Fuzzy en los datos recopilados	78
Tabla 14: Eliminación índice único de Efectividad C.F.	78
Tabla 15: Índice Efectividad Superficies Fuzzy de primer Grado en Datos Peor Caso.....	79
Tabla 16: Índice único efectividad Superficies Fuzzy de primer Grado en Datos Promedio.....	80
Tabla 17: Eliminación variables por índice único de efectividad Superficies Fuzzy	80
Tabla 18: Variables más significativas	81
Tabla 19: Resultados redes neuronales	85

LISTA DE FIGURAS

Figura 1: Formas de corrosión	19
Figura 2: Oxidación y Reducción	20
Figura 3: Celda de Corrosión	21
Figura 4: Corrosión Interna	21
Figura 5: Cupón sin usar y Cupón corroído	23
Figura 6: Instalación de Cupón	24
Figura 7: Probetas	26
Figura 8: Sensor e Instrumento.....	26
Figura 9: Puntos (Xi, Y) m casos.....	29
Figura 10: Función de membresía	30
Figura 11: Promedio local Pesado	31
Figura 12: Neurona Biológica.....	37
Figura 13: Estructura Red Neuronal Artificial	38
Figura 14: Funciones de activación, tansig, logsig y purelin	40
Figura 15: Fases del desarrollo de la herramienta.....	44
Figura 16: a) Cabeza de pozo, b) Manifold, c) Línea Principal, d) Unidad central de procesamiento.....	46
Figura 17: Archivos Generados	48
Figura 18: Ejemplo Manifold con 3 líneas de pozo	49
Figura 19: Línea Principal	52
Figura 20: Datos recopilados de cupones (peor caso).....	57
Figura 21: Diagrama de Flujo Selección de Variables	58
Figura 22: Diagrama de Flujo del entrenamiento de la red	63
Figura 23: Regresión lineal	66
Figura 24: Diagrama casos de uso de la herramienta	67
Figura 25: Archivos Excel con Datos Recopilados.....	73
Figura 26: CO ₂ y H ₂ S en Datos de Peor Caso.....	74

Figura 27: CO2 y H2S en Datos Promedio	75
Figura 28: Diferencia en la velocidad de corrosión entre datos simulados y experimentales	82
Figura 29: Velocidad de corrosión simulada Vs Experimental	82
Figura 30: Diferencia entre datos Simulados y Experimentales (8 variables de entrada)	83
Figura 31: Velocidad de corrosión simulada Vs Experimental	84
Figura 32: Rendimiento de las redes dependiendo de las variables de entrada...	85
Figura 33: Ventana Principal de la Herramienta	91
Figura 34: Ventana Principal Indicando Funcionalidad de la Opción	92
Figura 35: (a) Selección Información, (b) Información en documento Excel	92
Figura 36: Validar Información Cargada	93
Figura 37: (a) Resultado Metodología Fuzzy, (b) Almacenamiento Resultado Metodología Fuzzy	94
Figura 38: (a) Progreso Red, (b) Mensaje indicando el final del entrenamiento ...	94
Figura 39: Resultados Entrenamiento Red Neuronal.....	95
Figura 40: (a) Ingresar nombre para la carpeta que albergará la red entrenada, (b) Ubicación de la carpeta generada.....	96
Figura 41: Carga de archivo Excel con variables de entrada, (b) validación de la información cargada	97
Figura 42: Ubicación de la carpeta que contiene la Red Entrenada	97
Figura 43: Resultado del cálculo de Velocidad de Corrosión.....	98

LISTA DE ANEXOS

8 ANEXO A: MANUAL DE USUARIO	90
8.1 Introducción	90
8.2 Selección actividad	90
8.3 Opción Pre-Procesamiento y Generación de la Red	92
8.4 Opción Generar Red.....	96
9 ANEXO B: POSTER.....	99

RESUMEN

TÍTULO: HERRAMIENTA SOFTWARE PARA LA ESTIMACIÓN DE VELOCIDADES DE CORROSIÓN POR CO₂ Y H₂S MEDIANTE EL USO DE TÉCNICAS BASADAS EN INTELIGENCIA ARTIFICIALⁱ

AUTOR: CHRISTIAN ANDRES ARIZA QUITIANⁱⁱ

PALABRAS CLAVE: Redes Neuronales, Curvas y Superficies Fuzzy, Corrosión, Velocidad de Corrosión, CO₂, H₂S.

DESCRIPCIÓN: La corrosión es un fenómeno natural que genera preocupación en las grandes industrias, por ejemplo la petrolera, debido a la gran cantidad de problemas que causa y los gastos económicos que conlleva.

En la industria petrolera la estructura interna de los depósitos y tuberías que albergan crudo se ven afectados por la corrosión a causa de diferentes factores, entre otros las características del fluido, entre las cuales el contenido de CO₂ y H₂S son los más influyentes.

En la actualidad se conocen diversidad de herramientas de uso general que ayudan en el cálculo de velocidades de corrosión por CO₂ y H₂S, pero que presentan problemas al momento de ser aplicadas en los campos locales.

El objetivo de este proyecto es emplear la técnica de redes neuronales para estimar velocidades de corrosión en los campos locales, teniendo en cuenta los buenos resultados obtenidos mediante la aplicación de redes neuronales a la estimación y predicción de procesos en diferentes áreas tecnológicas. En el proyecto se propone la técnica de curvas y superficies fuzzy para pre-procesamiento de las variables de entrada en la red neuronal.

Finalmente se genera una librería de esta herramienta en lenguaje de programación C Sharp, con interfaz grafica de fácil manejo para el usuario, la cual será vinculada al software de la Corporación para la investigación de la Corrosión (CIC).

Esta librería incluye el pre procesamiento de las variables de entrada mediante técnica fuzzy y hace uso del código abierto: Fast Artificial Neural Network (FANN), para generación y entrenamiento de la estructura neuronal y determinación de las velocidades de corrosión.

ⁱ Proyecto de Grado

ⁱⁱFacultad de Ingeniería Físico mecánicas. Escuela de Ingeniería de Sistemas e Informática.
Director: Lola Xiomara Bautista Roza. Codirector: Laura Viviana Galvis Carreño.

SUMMARY

TITLE: NEURAL NETWORKS ARTIFICIAL INTELLIGENCE SOFTWARE TOOL FOR ESTIMATION OF CO₂ AND H₂S CORROSION SPEEDSⁱ

AUTHOR: CHRISTIAN ANDRES ARIZA QUITIANⁱⁱ

KEYWORDS: Neural Networks, Fuzzy Curves and Surfaces, Corrosion, Speed Corrosion, CO₂, H₂S.

DESCRIPTION: Corrosion processes generate great concern in major industries such as Oil industry, due to the great number of problems and economic costs involved.

Oil recipients and pipelines in petroleum industry are affected by corrosion caused by several factors, including the characteristic of fluid such as CO₂ and H₂S content.

Several general purpose tools are currently used for calculation of CO₂ and H₂S corrosion speeds. However most of these tools are not readily applied to local fields.

Based on the reported performance of neural networks for estimation and prediction of processes in different areas, this project aims using neural network techniques to estimate corrosion speeds in local fields. Fuzzy curves and surfaces procedures will be used for pre-processing of the input variables in the neural network.

As a final step the project will generate a library in C# language with user easy handling graphical interface, to be linked to the general corrosion software of the Corporación para la investigación de la Corrosión (CIC).

This library contains pre-processing of the input variables through fuzzy techniques and uses Fast Artificial Neural Network (FANN) open code (generated in C++ language and easy to use), for generation and training of the neural structure and estimation of corrosion rates.

ⁱWork Degree

ⁱⁱFaculty of Physicomechanical Engineering, Systems and Informatics School.
Director: Lola Xiomara Bautista Rozo. Codirector: Laura Viviana Galvis Carreño.

INTRODUCCIÓN

La corrosión es un fenómeno que genera graves problemas tecnológicos e importantes pérdidas económicas a numerosas industrias: automotriz, aeronáuticas, acueductos, petroleras, entre otras [1]. Por ésta razón la investigación sobre el origen del proceso y las posibilidades de controlarlo o atenuarlo, es un importante campo de actividad en las industrias implicadas.

La industria petrolera ha determinado que es posible correlacionar la velocidad de corrosión en sus estructuras con las condiciones operacionales y las características del crudo. Se ha comprobado que el CO₂ y el H₂S, gases que normalmente acompañan al crudo, aceleran el proceso de corrosión [2].

Ante la urgente necesidad de estudiar el proceso de corrosión, se han propuesto diversidad de procesos y tecnologías. Generalmente estos modelos están contruidos bajo condiciones específicas para casos particulares y se encuentran grandes dificultades para aplicarlos en casos más generales.

El modelado de procesos con inteligencia artificial, se ha consolidado como un procedimiento apropiado para la solución de problemas cada vez más complejos e importantes [7].

Modelos como las redes neuronales artificiales, y la lógica difusa (fuzzy logic) ampliamente probadas en diversidad de áreas de desarrollo tecnológico, se han implementado también en diferentes aspectos de la investigación en corrosión con resultados óptimos. Esto ha motivado a la Corporación para la Investigación de la Corrosión (CIC), para incluir entre sus procedimientos de modelado del proceso y predicción de velocidades de corrosión en pozos petroleros, el procedimiento basado en Redes Neuronales Artificiales (RNA).

1. GENERALIDADES

1.1 PLANTEAMIENTO DEL PROBLEMA

La corrosión es responsable de las fallas más representativas en el campo petrolero; tan solo en Estados Unidos el manejo de la corrosión demanda inversiones del orden de 1.4 billones de dólares [1].

Esto ha generado el desarrollo de diversos modelos experimentales y mecanicistas que buscan determinar la velocidad del proceso con el fin de determinar las variables de importancia en el proceso, así como desarrollar estrategias para mitigar el desgaste y las malformaciones producidas en los revestimientos industriales.

En los modelos de corrosión existentes se tienen en cuenta principalmente las condiciones operacionales y el tipo de crudo así como sus variables químicas y físicas.

La corporación para la investigación de la corrosión (CIC), en la actualidad utiliza algunos de los modelos de uso general en la industria petrolera mundial para la predicción de velocidades de corrosión por agentes como CO₂ (dióxido de carbono) y H₂S (ácido sulfhídrico). El CO₂ y H₂S se disuelven en el agua para formar ácidos débiles que son corrosivos. Adicionalmente diversos factores tales como condiciones operacionales, propiedades físico-químicas del agua y composición iónica del agua influyen en la determinación de la velocidad de corrosión [3].

Generalmente los modelos están contruidos bajo condiciones específicas para casos particulares y se encuentran grandes dificultades para aplicarlos en la industria petrolera local, encontrando, en muchos casos grandes diferencias entre

los valores de velocidad de corrosión simulada por los modelos y los valores experimentales medidos en campo por cuponesⁱ.

Dado el impacto negativo que causa la corrosión y la urgente necesidad de disminuirlo, la industria se ve en la necesidad de buscar técnicas alternativas para la construcción de modelos de velocidad de corrosión por CO₂ y H₂S.

Las redes neuronales artificiales han demostrado un buen desempeño en modelos de aproximación y estimación en diversos tipos de problemas, y se han reportado ya implementaciones exitosas en problemas de corrosión.

Adicionalmente se propone utilizar la técnica de curvas y superficies fuzzy, para determinar la sensibilidad del procedimiento RNA a las variables de entrada propuestas y seleccionar así las de mayor importancia en el fenómeno.

ⁱLos cupones son dispositivos que se insertan en instalaciones industriales como forma de medir los niveles de corrosión.

1.2 OBJETIVOS

1.2.1 Objetivo General

Desarrollar una herramienta software basada en la técnica de redes neuronales artificiales, para la estimación de velocidades de corrosión por CO₂ y H₂S utilizando datos provenientes de monitoreo con cupones.

1.2.2 Objetivos Específicos

- ✓ Determinar las variables asociadas al proceso de velocidad de corrosión por CO₂ y H₂S, utilizando datos provenientes de monitoreo realizado en cupones por la Corporación para la Investigación de la Corrosión (CIC).
- ✓ Recopilar la información existente en el conjunto de datos de trabajo de cupones.
- ✓ Aplicar la técnica de curvas y superficies Fuzzy para seleccionar las variables de entrada más importantes para la configuración de redes neuronales.
- ✓ Configurar y optimizar el diseño de redes neuronales artificiales para la determinación de velocidades de corrosión por CO₂ y H₂S utilizando los datos provenientes de monitoreo con cupones.
- ✓ Desarrollar la herramienta de simulación para el uso de los modelos implementados en la librería de cálculo utilizada en la CIC.
- ✓ Implementar los modelos basados en redes neuronales artificiales para el cálculo de velocidades de corrosión desarrollados, como un módulo en la librería de simulación de la CIC.

2. MARCO TEÓRICO

2.1 CORROSIÓN

2.1.1 Introducción a la corrosión

La corrosión es un proceso de degradación que ocurre en materiales procesados; se puede representar como un proceso electroquímico donde se tiene un material que cede electrones (ánodo) y otro que los recibe (cátodo).

La corrosión toma diversas formas en el material, las más comunes son:



Figura 1: Formas de corrosión

Fuente: Cortés María Teresa and Ortiz Pablo. *CORROSIÓN. Hipótesis/apuntes científicos Uniandinos. Dic. 2004.*

- Corrosión uniforme: se presenta de manera superficial en las estructuras, es la más general, la que mayores pérdidas produce y la más fácil de trabajar.
- Corrosión por picaduras: se presenta de manera localizada en la parte interna de las estructuras, crece rápidamente y es difícil de trabajar.
- Corrosión galvánica: se presenta en la unión de metales de diferente potencial eléctrico [9].

- Corrosión por Esfuerzos: esta falla se presenta a causa de corrosión existente asistida por un esfuerzo.
- Corrosión Intergranular: debido a que los metales están compuestos de granos, la corrosión se presenta en las fronteras de estos a causa de impurezas acumuladas [9].
- Corrosión por Hendiduras: ocurre en los espacios que quedan al unir dos metales así como en los espacios donde se presente estancamiento de alguna especie de líquido o solución.

2.1.2 Proceso corrosivo

La corrosión es un proceso electroquímico, en el cual ocurre transferencia de electrones. La pérdida de electrones se denomina Oxidación y la ganancia de electrones se denomina Reducción. La Figura 2 muestra el proceso en el cual la Oxidación ocurre en el Ánodo y la Reducción en el Cátodo.

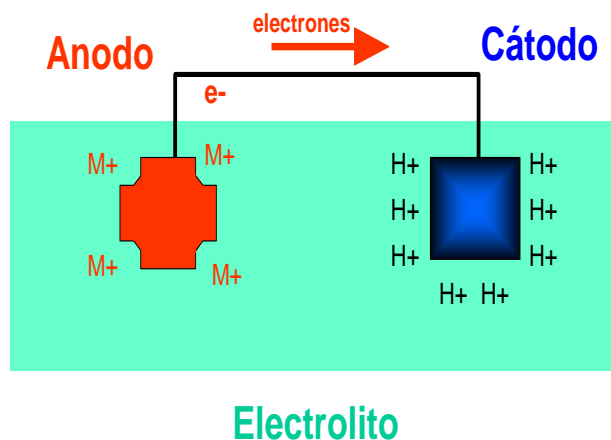


Figura 2: Oxidación y reducción

Fuente: Nace International, 2004. Curso de corrosión Básica, Manual del Estudiante.

El proceso se puede describir mediante el modelo de celda de corrosión, donde el área que cede electrones y el área que los recibe se encuentran en una solución

liquida (electrolito) la cual permite el flujo de átomos o moléculas cargadas llamadas iones.

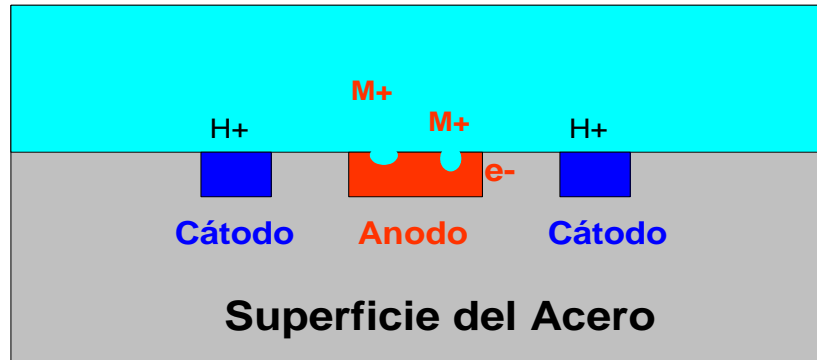


Figura 3: Celda de corrosión

Fuente: Nace International, 2004. Curso de corrosión Básica, Manual del Estudiante

2.1.3 Corrosión por CO_2 y H_2S



Figura 4: Corrosión interna

Fuente: Corporación para la Investigación de la Corrosión (CIC)

En la extracción de crudo de las profundidades de la tierra, además de crudo, se extrae agua y gas; lo cual da lugar a flujo multifásico. El gas presente en el sistema puede contener pequeñas cantidades de dióxido de carbono (CO₂) o sulfuro de hidrogeno (H₂S) que se disuelven en el agua formando especies reactivas como el ácido carbónico, ácido sulfhídrico, iones bicarbonato, iones carbonato iones hidrosulfuro, iones sulfuro e iones hidrógeno(H⁺), los cuales participan en la interacción electrónica con la estructura metálica del sistema de transporte.

Estos productos generan procesos electroquímicos en el sistema, los cuales dan paso a la corrosión, que en sus primeras etapas produce un gran deterioro de la estructura [3].

2.1.4 Formas de medir la corrosión

La medición de la corrosión es un proceso de cuantificación mediante el cual se puede determinar qué tan corrosivo es un determinado sistema y el ambiente que lo rodea, así como el desgaste que ocurre en la estructura. Este proceso de medición también permite optimizar las técnicas de mantenimiento y prevención de la corrosión así como determinar la vida útil de la estructura del sistema.

Existen numerosas formas para medir la corrosión, como, el uso de Radiografía, Ultrasonido, Termografía, marranos inteligentes, Análisis químico (Medición de pH, análisis microbiológico, gases disueltos) y operacional (tasa de flujo, presión, temperatura, pH), Monitoreo por cupones y monitoreo por probetas, entre otras.

2.1.5 Monitoreo de corrosión

Entre las técnica más comunes en el monitoreo de la corrosión se encuentran la medición por pérdida de peso en Cupones y resistencias eléctricas en probetas.

2.1.5.1 Pérdida de peso en cupones

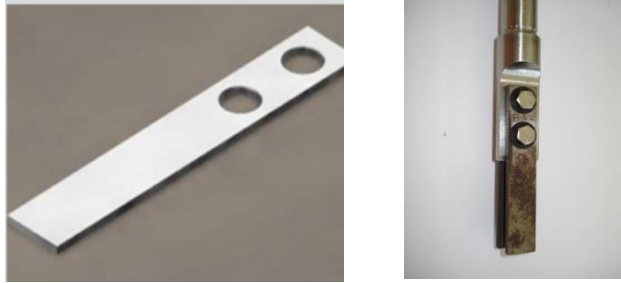


Figura 5: Cupón sin usar y cupón corroído

Fuente: Corporación para la Investigación de la Corrosión (CIC)

Es una de las técnicas más usadas en la industria por sus ventajas, como:

- La simplicidad de su uso.
- Posible inspección visual y análisis de materiales adheridos.
- Aplicable a cualquier ambiente (fluidos con partículas sólidas, líquidos y Gas).
- Fácil determinación de la rata de corrosión por medio de la pérdida de peso del cupón.
- Fácil fabricación.

El cupón es una pequeña muestra del mismo material que la estructura a monitorear, orientado de forma conveniente dentro de la estructura; por lo general se orienta de forma paralela al flujo para tener las mismas condiciones de las paredes de la estructura que lo contiene.



Figura 6: Instalación de Cupón

Fuente: Corporación para la Investigación de la Corrosión (CIC)

El Cupón es expuesto por un tiempo prolongado (por lo general mayor a 4 semanas) en un sitio estratégico dentro de la estructura en un montaje apropiado (porta-cupón) como muestra la figura 6. Allí estará bajo las condiciones operacionales y demás que intervienen en el proceso degenerativo del material. Al momento de retirar el cupón, y luego de ser limpiado de todas las impurezas adheridas, se puede determinar el peso perdido(PP), con el cual se halla la pérdida de espesor (PE) y el promedio de la rata de corrosión(RC)[14].

$$PE = \frac{PP * K}{DM * AM} \quad (2.1)$$

PP=Peso perdido (g)

DM=Densidad del material (g/cm³)

AM=Área del material expuesta (A)

T=Tiempo de exposición en horas

K = constante, definida por la tabla1.

$$RC = \frac{PP*k}{DM*AM*T} \quad (2.2)$$

k = constante, definida en la tabla 2.

Unidad(PE)	Unidad del Área (A)	K
Mils	In ²	61.02
Mils	Cm ²	393.7
Milímetros	Cm ²	10.0

Tabla 1: Constante **K** para perdida de espesor

Unidad(PE)	Unidad del Área (A)	k
mils/año(mpy)	In ²	5.34x10 ⁵
mils/año(mpy)	Cm ²	3.45 x10 ⁵
milímetros/año(mmy)	Cm ²	8.75 x10 ⁵

Tabla 2: Contante **k** para tasa de corrosión

Una desventaja del cupón es cuando llegue a generarse un alto valor en la velocidad de corrosión para algún momento durante el tiempo de exposición, no será posible determinar la magnitud del desgaste y el momento preciso en que ocurrió.

2.1.5.2 Probeta

Tiene las mismas ventajas del cupón además de poder expresar en tiempo real el desgaste de la estructura expuesta a condiciones propicias para el proceso de corrosión.



Figura 7: Probetas

Fuente: Corporación para la Investigación de la Corrosión (CIC)

La probeta contiene un elemento sensor de material similar al que se quiere analizar, y está conectada a un instrumento en donde se encuentra un elemento de referencia.

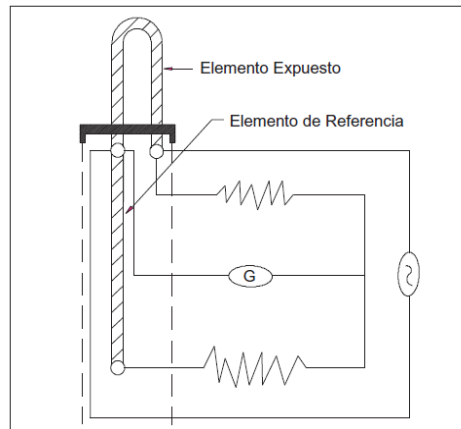


Figura 8: Sensor e Instrumento

Fuente: Metal Samples. Monitoreo por Resistencia Eléctrica. Disponible en:

<http://www.alspi.com/spanish.htm>

La probeta funciona midiendo la diferencia de resistencia eléctrica, entre el sensor expuesto y el elemento de referencia que contiene el cupón. El desgaste en la sección transversal del sensor por el proceso de corrosión, es proporcional al crecimiento de la resistencia, debido al principio,

$$R = r \frac{L}{A} [ohm] \quad (2.3)$$

Donde,

R= Resistencia eléctrica de un elemento

L= Longitud de del elemento

A= Área del corte transversal

r= Resistencia especifica del material

Entonces la diferencia en la resistencia eléctrica se atribuye a la pérdida del material expuesto [15].

2.2 APROXIMACIÓN FUZZY AL PROCESAMIENTO DE DATOS: CURVAS Y SUPERFICIES FUZZY

2.2.1 introducción

Cuando se tiene un conjunto de datos de entrada y salida para aplicar redes neuronales, es conveniente preseleccionar las variables de entrada según su grado de importancia; esto debido a la creciente complejidad de los procesos de aproximación en las redes por uso de variables superfluas, aparición de variables interrelacionadas y la posibilidad de la red para quedar atrapada en mínimos locales.

El manejo de las redes neuronales se puede simplificar considerablemente, y su rendimiento se puede optimizar, particularmente en el caso de sistemas no

lineales, con un manejo apropiado de los datos de entrada, que permita seleccionar los datos más significativos y excluir datos interdependientes.

Recientemente se ha desarrollado un procedimiento basado en la aplicación de curvas y superficies fuzzy para clasificar los datos de entrada en orden de importancia para la determinación del valor de un comportamiento buscado, así como para eliminar parámetros no independientes [11], [13].

2.2.2 Curvas Fuzzy de Primer Grado (o primer nivel)

El procesamiento de datos mediante Curvas Fuzzy de grado uno parte de una hipótesis: Las variables de entrada más importantes permiten la mejor aproximación de la variable de salida.

Para explicar el funcionamiento de las curvas Fuzzy, supongamos un problema [11] para el cual se tienen N variables de entrada X^i , $(X^1, X^2, \dots, X^i, \dots, X^N)$ relacionadas con una variable de salida Y, y m casos experimentales para entrenamiento,

$$\begin{aligned} & (X_1^1, X_1^2, X_1^3, \dots, X_1^i, \dots, X_1^N, Y_1), \\ & (X_2^1, X_2^2, X_2^3, \dots, X_2^i, \dots, X_2^N, Y_2), \dots, (X_m^1, X_m^2, X_m^3, \dots, X_m^i, \dots, X_m^N, Y_m) \end{aligned} \quad (2.4)$$

Para cada variable de entrada X^i se grafican los m puntos (X_k^i, Y_k) correspondientes a los m casos ejemplo sobre el espacio (X^i, Y) , como se muestra en la figura 9.

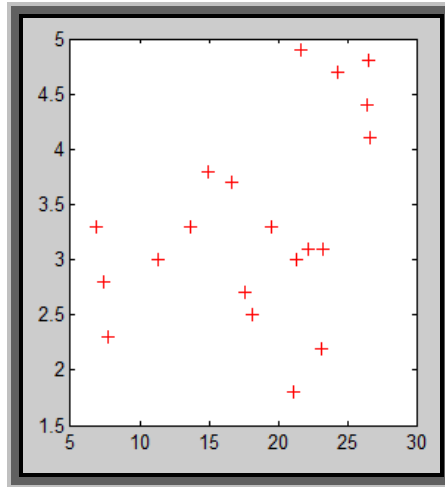


Figura 9: Puntos (X^i, Y) m casos.

Fuente: Autor del proyecto

Para cada punto experimental k (X_k^i, Y_k) se define la función de membresía u_k^i en el espacio (X^i, Y) así:

$$u_k^i(X^i) = \exp\left(-\left(\frac{X_k^i - X^i}{b}\right)^2\right) \quad (2.5)$$

$b = 0.2 * [Max(X^i) - min(X^i)]$, = 20% del intervalo de la variable X^i .

La función de membresía en el espacio (X^i, Y) , presenta forma de campana de Gauss con máximo en el punto experimental (X_k^i, Y_k) , como se ve en la figura 10 que representa el producto $Y_k * u_k^i(X^i)$ vs X^i para cada punto experimental k en el espacio (X^i, Y) .

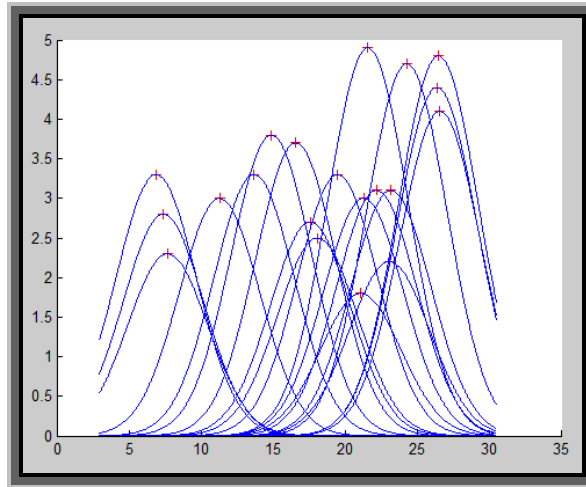


Figura 10: Función de membresía, X^i vs $u_k^i(X^i) * Y$

Fuente: Autor del proyecto

Se debe notar:

- Puntos del espacio (X^i, Y) con coordenadas X^i , iguales a los valores de entrada X_k^i tienen valor $u_k^i = 1$.
- Puntos con $X^i - X_k^i = b$ tienen valores de membresía $u_k^i = \frac{1}{e}$.
- Puntos con coordenadas X^i en las cercanías de valores de entrada X_k^i ; $(X^i - X_k^i) < b$ tienen valores de membresía $u_k^i > \frac{1}{e}$.
- Puntos con $(X^i - X_k^i) > b$ tienen valores de membresía $u_k^i(X^i) < 1/e$.

A continuación se defuzzifica cada variable de entrada X^i , para así generar la curva Fuzzy Y_c^i (figura 11) :

$$Y_c^i(X^i) = \frac{\sum_{k=1}^m Y_k \times u_k^i(X^i)}{\sum_{k=1}^m u_k^i(X^i)} \quad (2.6)$$

Cada punto de la curva Y_c^i puede ser interpretado como un promedio local (en un punto sobre la curva X^i vs Y) pesado de Y , a lo largo de los m casos, para cada variable X^i .

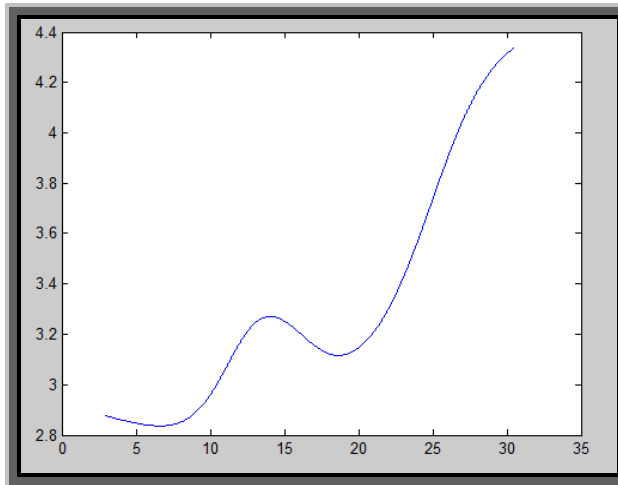


Figura 11: Promedio local Pesado, $Y_c^i(X^i)$ vs X^i

Fuente: Autor del proyecto

Las variables de entrada más importantes serán las que tengan los valores de promedios locales pesados más cercanos a Y_k . Entonces como criterio de selección se define un índice de efectividad, $P_{y_c^i}$ para las variables de entrada,

$$P_{y_c^i} = \frac{1}{mv_y} \sum_{k=1}^m (Y_c^i(X^i) - Y_k)^2 \quad (2.7)$$

Dónde v_y es la varianza de la variable de salida.

$$v_y = \sum_{k=1}^m (Y_k - Y_{prom})^2 \quad (2.8)$$

A menor valor de $P_{y_c^i}$, más importante es la variable X^i .

Es así como mediante las curvas fuzzy de grado uno, rápidamente se clasifican las variables de entrada en orden de importancia. Sin embargo hasta aquí no se han descartado las variables dependientes o relacionadas.

2.2.3 Curvas Fuzzy de segundo grado

La clasificación de variables con base en curvas Fuzzy de segundo nivel $v_c^i(X^i)$ parte de la hipótesis: Variables de entrada relacionadas causalmente con la variable de salida permiten la mejor aproximación de Y . Lo contrario ocurre con las variables de entrada relacionadas en forma aleatoria con Y .

Las curvas Fuzzy de segundo nivel representan un estimado local pesado (promedio local pesado) de la varianza del valor de salida Y a lo largo de los m casos para cada variable X^i :

$$v_c^i(x^i) = \frac{\sum_{k=1}^m (Y_c^i(X_k^i) - Y_k)^2 \mu_k^i(X^i)}{\sum_{k=1}^m \mu_k^i(X^i)} \quad (2.9)$$

Variables de entrada aleatoriamente relacionadas con la variable de salida Y presentarán valores de la varianza local $v_c^i(X^i)$ cercanas a la varianza global v_y de la variable de salida.

Variables de entrada X^i relacionadas de forma causal con Y presentaran valores de $v_c^i(X^i)$ y v_y diferentes en muchas situaciones. Por lo tanto el índice de efectividad de curvas fuzzy de segundo grado, se define como:

$$Pv_c^i = \frac{\sum_{k=1}^m (v_c^i(X_k^i) - v_y)^2}{m v_y^2} \quad (2.10)$$

El cual mide el comportamiento causal de la variable de entrada con respecto a la variable de salida. Aquí los valores más elevados de Pv_c^i corresponden a las variables más significativas, comportamiento contrario al observado para Pv_c^i .

Para seleccionar las variables de entrada más importantes y con comportamiento causal con respecto a la variable de salida, se combinan los índices de efectividad definidos con base en curvas fuzzy de primer grado, $P y_c^i$ y curvas fuzzy de segundo grado, $P v_c^i$, en un índice único de efectividad ($P c^i$):

$$P c^i = \frac{P y_c^i}{1 + P v_c^i} \quad (2.11)$$

Este índice presenta valores en el intervalo: $0 < P c^i < 1$ (0 es bueno y 1 es malo)

$P c^i \rightarrow 0 \rightarrow X^i$ importante y causal.

$P c^i \rightarrow 1 \rightarrow X^i$ irrelevante y aleatoria.

2.2.4 Superficies Fuzzy de primer grado (o primer nivel)

Las superficies fuzzy, están basadas en una segunda idea simple: variables de entrada independientes hacen un mejor trabajo de aproximación de la variable de salida que las entradas interdependientes. Las superficies fuzzy de primer grado se definen asíⁱ:

$$Y_s^{i,j}(X^i, X^j) = \frac{\sum_{k=1}^m Y_k \times \mu_k^i(X^i) \times \mu_k^j(X^j)}{\sum_{k=1}^m \mu_k^i(X^i) \times \mu_k^j(X^j)} \quad (2.12)$$

Cada punto de la superficie $Y_s^{i,j}(X^i, X^j)$ puede ser interpretado como un promedio local (en un punto sobre la superficie Y vs X^i, X^j) pesado de Y a lo largo de los m casos.

ⁱ**Fuente:** LIN Yinghua and CUNNINGHAM George. Nonlinear System Input Structure Identification: Two Stage Fuzzy Curves and Surfaces. IEEE. 1998.

Los valores promedio local pesado sobre puntos de superficies X^i, X^j correspondientes a pares de variables de entrada X^i, X^j independientes entre sí serán los más cercanos a valores experimentales Y_k .

Partiendo del orden de importancia establecido por las curvas Fuzzy de primer grado, se toman pares X^i, X^j . Cada vez X^i la variable más importante y X^j la siguiente en orden de importancia.

Se define un índice de efectividad ($P y_S^{i,j}$) para $Y_S^{i,j}(X^i, X^j)$ como:

$$P y_S^{i,j} = \frac{\sum_{k=1}^m (Y_S^{i,j}(X_k^i, X_k^j) - Y_k)^2}{m v_y} \quad (2.13)$$

Los valores más pequeños de $P y_S^{i,j}$ corresponden a la mayor independencia del par de variables X^i y X^j .

Las curvas y superficies fuzzy de primer grado permiten seleccionar las variables de entrada más importantes e independientes y descartar las poco relevantes y dependientes.

2.2.5 Superficies Fuzzy de segundo grado

Las curvas y superficies fuzzy de segundo grado permiten diferenciar entre causalidad y aleatoriedad de las variables de entrada con respecto a la variable de salida.

Las superficies fuzzy de Segundo grado se define, como:

$$v_S^{i,j}(X^i, X^j) = \frac{\sum_{k=1}^m (Y_S^{i,j}(X_k^i, X_k^j) - Y_k)^2 \mu_k^i(X^i) \mu_k^j(X^j)}{\sum_{k=1}^m \mu_k^i(X^i) \mu_k^j(X^j)} \quad (2.14)$$

$v_s^{i,j}(X^i, X^j)$, puede ser interpretada como un promedio ponderado local (en un punto en una superficie (X^i, X^j)) de la varianza a lo largo de los m casos para cada superficie (X^i, X^j) .

El índice de efectividad para las superficies Fuzzy de segundo nivel está definido como:

$$Pv_s^{i,j} = \frac{\sum_{k=1}^m (v_s^{i,j}(X_k^i, X_k^j) - v_y)^2}{mv_y^2} \quad (2.15)$$

Contrario al índice de efectividad basado en las superficies fuzzy de grado uno ($Py_s^{i,j}$), el valor más alto de $Pv_s^{i,j}$ es más importante.

Definiendo un índice único de efectividad para superficies fuzzy de primer y segundo grado:

$$P_S^{i,j} = \frac{Py_s^{i,j}}{1 + Pv_s^{i,j}} \quad (2.16)$$

Donde el rango es: $0 < P_S^{i,j} < 1$.

Valores cercanos a cero corresponden a un buen comportamiento (independencia y causalidad). Valores cercanos a uno corresponden a un mal comportamiento (dependencia y aleatoriedad).

$P_S^{i,j} \rightarrow 0 \rightarrow X^i, X^j$ independientes y causalmente relacionadas con la variable de salida (Y).

$P_S^{i,j} \rightarrow 1 \rightarrow X^i, X^j$ interdependientes y aleatoriamente relacionadas con Y.

Las Curvas y Superficies Fuzzy permiten seleccionar un conjunto de variables de entrada importante, independiente y causal sin utilizar complicadas técnicas para modelar sistemas no-lineales.

2.2.6 Procedimiento de selección de las variables de entrada basado en metodología fuzzy

Este procedimiento se usa para aislar las variables de entrada más importantes e independientes [13].

1. Generando y agregando una variable aleatoria X^R al grupo de variables de entrada y además teniendo la siguiente consideración $0.99 \leq \alpha \leq 1$.
2. Se generan las curvas fuzzy, se generan los índices de efectividad y con esos datos se ordenan las variables según el índice unico de efectividad Pc^i .
3. La variable de entrada con el menor Pc^i , es la variable más importante. Utilizando el índice de efectividad (Pc^R) de la variable aleatoria (X^R) se eliminan todas las variables para las cuales $Pc^i/Pc^R > \alpha$. Ya que aparentemente solo son variables de entrada relacionadas aleatoriamente con la variable de salida.
4. Usando la variable más importante del proceso anterior (X^i), se generan secuencialmente pares (X^i, X^j) con las variables restantes $(X^j), (j \neq i)$ para generar las superficies fuzzy. La variable con el índice unico de efectividad $Ps^{i,j}$ más pequeño es la siguiente variable más importante. Las variables para las cuales $Ps^{i,j}/Ps^{i,R} > \alpha$ son eliminadas porque se consideran variables aleatorias.
5. Se repite el paso 3 hasta que no existan más variables que Eliminar.

2.3 REDES NEURONALES ARTIFICIALES

2.3.1 Introducción

Para abordar el tema de Redes Neuronales Artificiales, es importante hacer una revisión de la estructura y funcionamiento de las neuronas biológicas, que son el modelo de las redes neuronales artificiales.

El descubrimiento de la estructura neuronal que conforma el sistema nervioso se debe al Científico español Santiago Ramón y Cajal [6], quien observó que el sistema nervioso estaba compuesto por una cadena de células, denominadas neuronas, separadas entre sí por un diminuto espacio, sinapsis (figura 12). En el sistema neuronal la información fluye desde las dendritas hacia el axón (unidireccional).

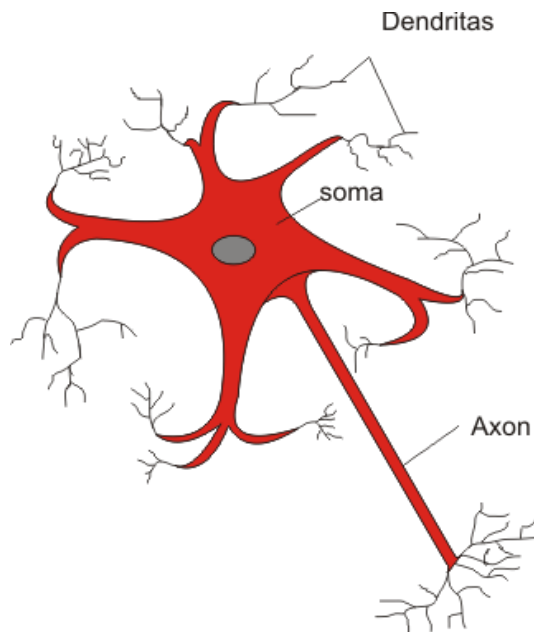


Figura 12: Neurona Biológica

Fuente: Autor del proyecto

Una neurona puede interactuar intercambiando información con muchas otras, esta información se transmite en forma de señales químicas, que dentro de la neurona generan señales eléctricas.

La neurona que envía la señal libera una serie de sustancias químicas (neurotransmisores) en el espacio que une las neuronas (Sinapsis), si la neurona que recibe la señal es sensible a estos componentes, permite el ingreso de cierto

número de iones, creando una corriente que da paso a pequeños potenciales que se combinan e integran en el soma. Si la suma de los potenciales supera un umbral de -45mV se genera un impulso de corriente que pasa por el axón y libera en la sinapsis los neurotransmisores contenidos en los extremos del axón [7].

2.3.2 Relación redes neuronales artificiales y neuronas biológicas

La red neuronal artificial, con una estructura que imita el comportamiento de la neurona biológica siguiendo el esquema de trabajo del cerebro, se propone para tratar de resolver determinados problemas de manera eficiente.

2.3.2.1 Estructura de la Red Neuronal Artificial

La Estructura propuesta de la Red Neuronal Artificial consta de capa de entrada, capas ocultas y capa de salida (figura 13).

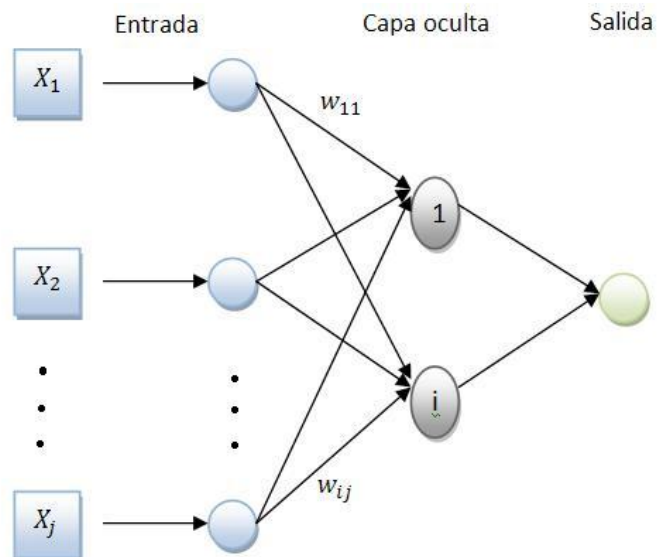


Figura 13: Estructura Red Neuronal Artificial

Fuente: Autor del proyecto

La capa de entrada consta de un número de neuronas igual al número de parámetros de entrada. Las capas ocultas, que pueden ser una o más, son los sitios de procesamiento de la información. La capa de salida es la salida de la red. A esta estructura se le denomina perceptrón multicapa, figura 13.

2.3.2.2 Modo de Operación de la Red

2.3.2.2.1 Entrada Al igual que la neurona biológica, la neurona artificial puede recibir información externa o de otras neuronas con las cuales se encuentre conectada.

La información de entrada se define como un vector $X = [X_1, X_2, \dots, X_j,]$ (la información puede ser continua o digital) [6].

2.3.2.2.2 Pesos Sinápticos En el sistema biológico, la conexión entre dos neuronas se denomina sinapsis; la neurona(j) que envía la información $X(j)$, se denomina pre-sináptica y la que la recibe(i) se denomina post-sináptica. La intensidad de la conexión sináptica entre la neurona i y la neurona j se define como peso sináptico W_{ij} [7].

2.3.2.2.3 Acumulado de Señales Cada neurona artificial i recibe las señales de entrada X_j procedentes del exterior o de otras neuronas, las modula con los correspondientes pesos sinápticos W_{ij} y genera una suma ponderada de las señales recibidas:

$$u_i = \sum_j W_{ij} X_j \quad (2.17)$$

2.3.2.2.4 Activación y salida Así como la neurona biológica se activa dependiendo de un umbral, la neurona artificial utiliza un proceso, el cual consiste en la aplicación de una función de activación a la suma ponderada para generar un dato de salida $f(u_i)$, que va a otra neurona o al exterior [6], [7].

$$f(u_i) = f(\sum_j W_{ij}X_j) \quad (2.18)$$

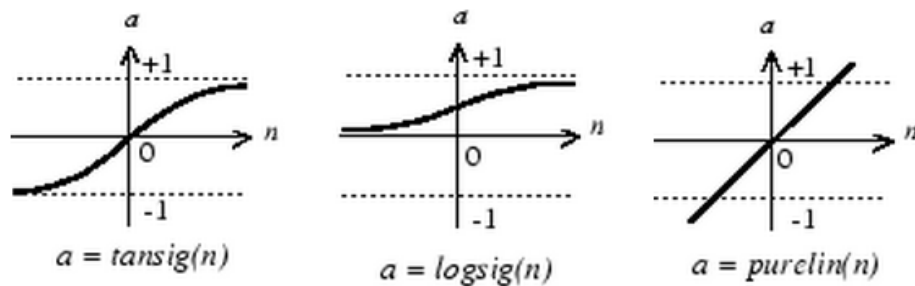


Figura 14: Funciones de activación, tansig, logsig y purelin

Fuente: M. Beale, M.Haga, H. Demuth. Neural Network Toolbox™ User's Guide, MathWorks Inc.

2.3.3 Especificación de la estructura de la red neuronal

Las dimensiones de la capa de entrada y la capa de salida de la red son fáciles de estimar. El número de neuronas de la capa de entrada es igual al número de parámetros independientes de entrada. La capa de salida consta de tantas neuronas como variables de salida, cada neurona de la capa de salida entrega el valor de una variable de salida. En nuestro caso solo hay un valor de salida; la velocidad de corrosión.

Para las capas ocultas, es difícil precisar el número de neuronas; se hacen pruebas con número variable de neuronas y el número apropiado se selecciona con base en el análisis de los resultados [7].

2.3.4 Entrenamiento de la Red

Cuando se le presenta a la red neuronal un grupo de datos de entrada y el correspondiente grupo de datos de salida, la red se entrena modificando gradualmente los pesos sinápticos en función de la diferencia entre los valores correctos de los datos de salida y los valores de salida simulados por la red. El proceso avanza hasta que la salida simulada tienda a la salida correcta. El proceso se detiene cuando la diferencia entre la salida deseada y la simulada por la red es inferior a un valor estipulado, o cuando el número de iteraciones supera un número también estipulado por el programador de la red. A este procedimiento se le denomina aprendizaje supervisado.

En el entrenamiento de la red neuronal artificial se establece una correlación matemática entre las variables de entrada de la red y las salidas deseadas; esta correlación está dada por los pesos sinápticos y las funciones de activación que unen las neuronas.

2.3.4.1 Algoritmo de retro-propagación (back-propagation) Algoritmos de entrenamiento como el de retro-propagación de errores (back-propagation), muy utilizado por sus buenos resultados y en el cual se basan muchos otros algoritmos, son utilizados para hallar los valores adecuados de los pesos sinápticos que unen a las neuronas, y así optimizar el desempeño de la red.

El algoritmo de retro-propagación, realiza conceptualmente el siguiente proceso: Seleccionados los datos con que se entrenará la red neuronal, y las respectivas salidas deseadas, se construye un vector de datos de entrada que se entrega a la capa de entrada. La información de entrada, procesada como ya se indicó por los pesos sinápticos y las funciones de activación, se propaga por la estructura de la red hasta llegar a la capa de salida de la red como valor de salida.

En cada iteración se calcula el error del proceso como la diferencia entre el valor de salida calculado por la red y el valor correcto (experimental); con base en éste error, se corrigen los pesos sinápticos:

$$W(t + 1) = W(t) + \Delta W(t) \quad (2.19)$$

Donde, $W(t + 1)$ es el peso corregido (iteración t+1), $W(t)$ el peso actual (iteración t) y $\Delta W(t)$ es la corrección calculada como función del error.

El proceso termina si se cumple una de las condiciones [6]:

- Error global E inferior a un valor estipulado

$$E = \frac{1}{2} \sum_{k=1}^m \sum_{q=1}^M (e_{kq})^2 \quad (2.20)$$

$$e_{kq} = (d_{kq} - Y_{kq}) \quad (2.21)$$

Donde, m es el número de casos (ejemplos) para entrenar la red, M el número de neuronas en la capa de salida, Y_{kq} valores simulados de datos de salida y d_{kq} valores experimentales (reales) de datos de salida.

- Número de iteraciones inferior a un número determinado por el programador de la red.

2.3.4.2 Algoritmo de Levenberg Marquardt

Es un algoritmo con una carga computacional mayor que el algoritmo de back-propagation, pero con la ventaja de converger más rápido. Este algoritmo es efectivo si se manejan redes con arquitecturas de pocas neuronas y pocas conexiones.

El algoritmo se basa en la combinación de dos métodos: el de gradiente descendente y el método de Newton. En este último la forma de disminuir el error global E es actualizando los pesos de la red representados por la matriz W .

$$W(t + 1) = W(t) - \frac{E'}{E''} \quad (2.22)$$

Donde,

E' , es el vector gradiente

E'' , es la matriz Hessiana

Para cada neurona en la capa de salida el vector gradiente es la derivada parcial del error (e) con respecto a cada uno de los pesos (W_{ij}) que conforman la red.

Y la Hessianaⁱ se calcula con las segundas derivadas parciales del error con respecto a los pesos [6].

ⁱMatrix cuadrada de las segundas derivadas parciales.

3 DESARROLLO DE LA HERRAMIENTA COMPUTACIONAL

El proyecto se desarrolla en una secuencia de cuatro fases. La finalización de una fase dará inicio a la siguiente, figura 15.

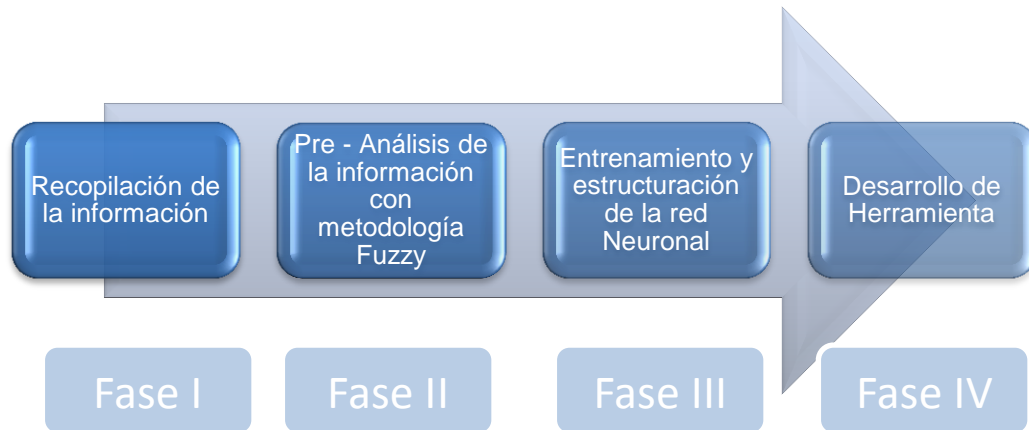


Figura 15: Fases del desarrollo de la herramienta

Fuente: Autor del proyecto

Inicialmente se deben seleccionar las variables asociadas al proceso corrosivo por CO₂ y H₂S en dispositivos de acero al carbón que transportan flujo multifásico. La experiencia del personal de la CIC determina que las variables asociadas son: presión, temperatura, flujo de crudo, flujo de gas, flujo de agua, pH y API, presentadas en la tabla 3.

De esta manera se inicia el proceso de recopilación de la información (Fase I), que es uno de los objetivos del proyecto. La calidad de la información recopilada en ésta fase es crucial para el éxito del proceso total.

Tipo de característica	Variable	Unidades
Condiciones operacionales	Presión	PSI
	Temperatura	°F
	Flujo de Crudo	BPD
	Flujo de Gas	mmscf/d
	Flujo de Agua	BPD
	API	
Fisicoquímicas	pH	
Gases Corrosivos	CO2	%v/v
	H2S	Ppm
Corrosión	Vel. Corrosión	Mpy

Tabla 3: Variables de entrada y Variable de salida

3.1 FASE I

Apoyado en un documento guía generado por la CIC con los datos de cupones, identificación, rango de vida del cupón y punto de monitoreo, se recopila información de las variables antes descritas.

Los puntos de monitoreo existentes en el documento guía por tener instalados cupones para su estudio son, Cabezas de pozo, manifold, líneas principales y Unidades centrales de procesamiento, figura 16.



(a)



(b)

Fuente: <http://www.ingenieriadepetroleo.com>

Fuente: <http://www.soldadurasasi.com>



(c)



(d)

Fuente: Corporación para la Investigación de la Corrosión

Fuente: Corporación para la Investigación de la Corrosión

Figura 16: a) Cabeza de pozo, b) Manifold, c) Línea Principal, d) Unidad central de procesamiento

En el desarrollo, se denominó “rango de vida del cupón” el intervalo de tiempo que dura expuesto el cupón a condiciones de trabajo, desde la fecha de puesta hasta la fecha de retiro; el rango de vida del cupón se encuentra en el documento guía,

el cual debe ser comparado con los registros almacenados en los documentos de la CIC.

Con esto en mente se procede a reunir la información por puntos. En caso de no existir información para cualquiera de las variables dentro del rango de vida del cupón se procede a tomar el registro con la fecha más cercana (La experiencia del personal de la CIC, argumenta que las condiciones no varían de manera significativa).

3.1.1 Cupones en Cabezas de Pozo (Well Head)

- **Variables de condiciones operacionales**

En caso de existir un solo registro dentro de la documentación de la CIC, se almacena este como dato.

Si existe más de un registro, se almacenan dos tipos de datos,

1. El registro de las variables con el valor de flujo de agua más alto (la experiencia en campo ha identificado que la corrosión es proporcional a la cantidad de agua en el sistema).
2. El promedio de los registros para cada variable. En caso de existir diversos valores en un rango de tiempo se puede suponer que el desgaste por corrosión es un proceso continuo que ocurrió desde la puesta del cupón hasta la fecha de medición (promedio).

- **Variables fisicoquímicas (PH)**

Se promedian los valores de pH en los registros existentes dentro del rango de vida del cupón.

- **Variables de Gases Corrosivos**

Se estudia el registro de las variables existentes dentro del rango de vida del cupón. En caso de encontrar más de un registro, se toman dos tipos de datos:

1. Las variables con el valor de CO2 más alto.
2. El promedio de los registros existentes dentro del intervalo de vida del cupón.

De esta manera se generan dos archivos (figura 17) con información de cupones con los siguientes nombres:

1. “Peor caso”, que utiliza los registros para condiciones operacionales correspondientes al mayor flujo de agua, para gases corrosivos el valor más alto de CO2 y para fisicoquímicas el promedio hallado.
2. “Promedio”, almacena la misma información que el anterior solo que utiliza los datos promediados para condiciones operacionales y gases corrosivos, para los casos con más de un registro.

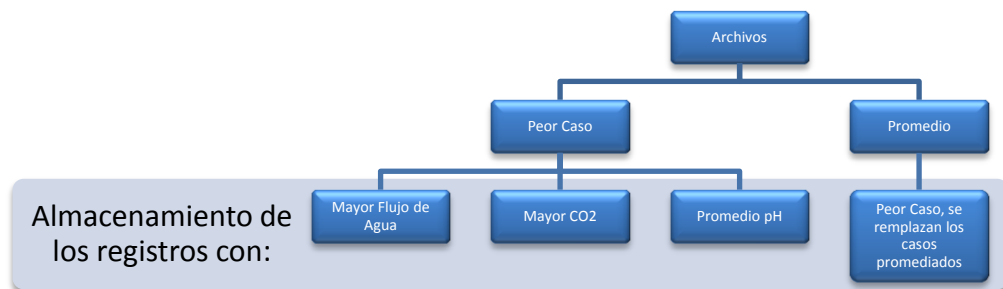


Figura 17: Archivos Generados

Fuente: Autor del proyecto

3.1.2 Cupones en Manifold

La Figura15 es útil para el proceso que se lleva a cabo, se muestran 3 pozos (pueden entrar uno o más pozos) que llegan a un manifold,

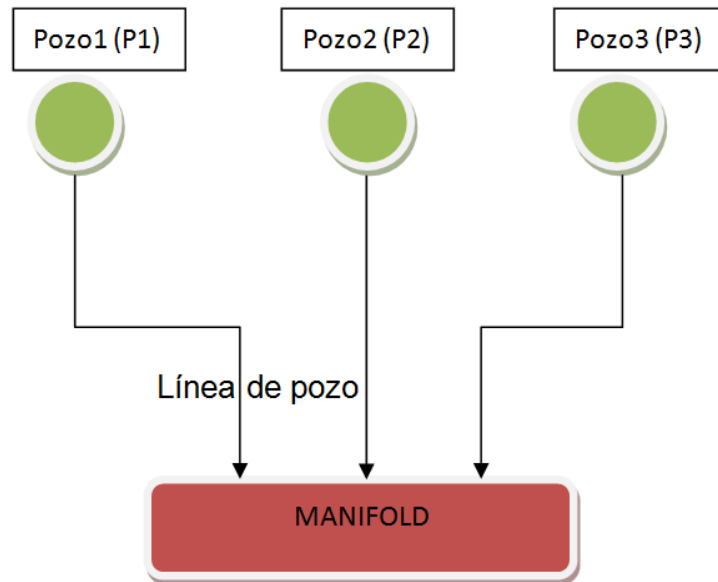


Figura 18: Ejemplo Manifold con 3 líneas de pozo

Fuente: Autor del proyecto

La información en cada manifold se halla por cálculos en los cupones, ya que las características del flujo en la cabeza de pozo varían muy poco al entrar al manifold (es lo que se conoce experimentalmente, además de ser la única información que se tiene); La forma de hallar la información en los cupones es descrita en la sección 3.1.1.

- **Variables condiciones operacionales**

Si entra solo una línea de pozo, se toma la información de la cabeza de pozo correspondiente a la línea. Pero si entra más de una línea de pozo, se toman de la siguiente manera:

- Flujo de crudo total en el manifold, es la suma de los flujos de crudo de cada línea de pozo que entra al manifold.
- Flujo total de Gas, es la suma de los flujos de gas de cada línea de pozo.
- Flujo total de Agua, es la suma de los flujos de agua de cada línea de pozo.
- El API, se determina promediando los datos existentes de cada línea de pozo.

Temperatura: Con la suma de los flujos de agua, gas y crudo para cada línea de pozo (con base en la figura 15), se desarrollan las siguientes operaciones:

$$flujo_{p1} = flujogas_{p1} + flujocrudo_{p1} + flujoagua_{p1} \quad (3.1)$$

$$flujo_{p2} = flujogas_{p2} + flujocrudo_{p2} + flujoagua_{p2} \quad (3.1.1)$$

$$flujo_{p3} = flujogas_{p3} + flujocrudo_{p3} + flujoagua_{p3} \quad (3.1.2)$$

Se calcula el total de los flujos,

$$flujo_{total} = flujo_{p1} + flujo_{p2} + flujo_{p3} \quad (3.2)$$

Que da pasó a un factor X_{pozo} (aporte fraccionario de cada pozo al flujo total), para cada línea de pozo,

$$X_{p3} = (flujo_{p3}/flujo_{total}) \quad (3.3)$$

Así la temperatura se calcula sumando los productos de la temperatura de pozo (T_{pozo}) y el correspondiente factor (X_{pozo}), para cada línea que entra al manifold,

$$T = X_{p1} * T_{p1} + X_{p2} * T_{p2} + X_{p3} * T_{p3} \quad (3.4)$$

Presión La presión total en el manifold se calcula mediante un software especializado, el cual no posee la CIC. Debido a esto se registra la presión más baja de las líneas de pozo, como la presión total (en teoría, se nota una diferencia en la mayoría de los casos no muy relevante entre el flujo más bajo y el flujo total).

- **Variables Fisicoquímicas (pH)**

Para el pH es necesario hallar el flujo total de agua ($flujo_{aguatotal}$),

$$flujo_{aguatotal} = flujo_{aguap1} + flujo_{aguap2} + flujo_{aguap3} \quad (3.5)$$

Que posteriormente se utiliza para calcular un factor (X_{wp}) en cada línea de pozo:

$$X_{wpi} = (flujo_{aguapi} / flujo_{aguatotal}) \quad (3.6)$$

i= representa cada línea de pozo que entra al manifold.

Es importante, que para el valor de PH total se considere la siguiente ecuación,

$$pH = - \text{Log}(cH^+) \rightarrow cH^+ = - \text{Log}pH \quad (3.7)$$

La cual calcula la concentración de iones hidronio (cH^+) disueltos en cada línea de pozo, así se calcula la concentración de hidronio total, y a continuación el pH total en el manifold,

$$cH_{manifold}^+ = X_{wp1} * cH_{p1}^+ + X_{wp2} * cH_{p2}^+ + X_{wp3} * cH_{p3}^+ \quad (3.8)$$

$$pH = -\log cH_{manifold}^+ \quad (3.9)$$

- **Variables de Gases Corrosivos**

Los valores de concentración de gases corrosivos se calculan en forma similar a la concentración de hidronio, teniendo en cuenta los flujos de gas.

$$X_{gp1} = (flujo_{gasp1}/flujo_{gastotal}) \quad (3.10)$$

Se halla para cada línea de pozo y se multiplica por el CO₂ y H₂S correspondiente,

$$CO_2 = X_{gp1} * CO_{2p1} + X_{gp2} * CO_{2p2} + X_{gp3} * CO_{2p3} \quad (3.11)$$

$$H_2S = X_{gp1} * H_{2Sp1} + X_{gp2} * H_{2Sp2} + X_{gp3} * H_{2Sp3} \quad (3.12)$$

3.1.3 Cupones en la línea principal

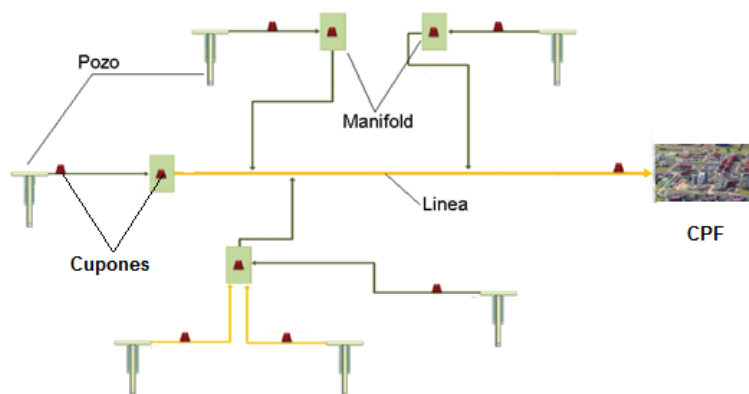


Figura 19: Línea Principal

Fuente: Corporación para la investigación de la corrosión

Se recopila la información de cada manifold que entra a la línea principal. El proceso en el caso de la línea principal es el mismo que ha sido desarrollado en la sección 3.1.2:

- Variables de condiciones operacionales.
- Flujo total de crudo en la línea, se calcula sumando los flujos de crudo de cada uno de los manifold que llegan a la línea.
- Flujo total de Gas, se calcula sumando los flujos de gas de cada manifold.
- Flujo total de agua, se suma los flujos de agua de cada manifold.
- El API se calcula promediando los datos de cada manifold.

Temperatura Con la suma de los flujos de agua, gas y crudo para cada manifold que llega a la línea y suponiendo la existencia de dos manifold que entran a una línea, el proceso es el siguiente:

$$flujo_{m1} = flujogas_{manifold1} + flujocrudo_{manifold1} + flujoagua_{manifold1} \quad (3.13)$$

$$flujo_{m2} = flujogas_{manifold2} + flujocrudo_{manifold2} + flujoagua_{manifold2} \quad (3.13.1)$$

Con el cálculo de la suma total de los flujos,

$$flujo_{mtotal} = flujo_{m1} + flujo_{m2} \quad (3.14)$$

Se determina un factor $X_{manifold}$, en cada uno de los manifold que llegan a la línea,

$$X_{mi} = (flujo_{mi}/flujo_{mtotal}) \quad (3.15)$$

$i = 1,2,\dots$ total de manifold

Con esto la temperatura total en la línea principal es la suma de los productos, temperatura manifold ($T_{manifold}$) y el correspondiente factor ($X_{manifold}$),

$$T = X_{m1} * T_{m1} + X_{m2} * T_{m2} \quad (3.16)$$

Presión Para determinar la presión en la línea, se tiene el mismo inconveniente presentado anteriormente en los manifold, por tanto se hace el mismo proceso.

- **Variables Fisicoquímicas (pH)**

En el caso del pH, se calcula el flujo total de agua en la línea

$$flujo_{agualinea} = flujo_{aguam1} + flujo_{aguam2} \quad (3.17)$$

Y para cada manifold de la línea, se calcula un factor:

$$X_{wm1} = (flujo_{aguam1}/flujo_{agualinea}) \quad (3.18)$$

$$X_{wm2} = (flujo_{aguam2}/flujo_{agualinea}) \quad (3.18.1)$$

Como se mencionó en la sección 3.1.2, el valor del pH está asentado en la concentración de iones Hidronio ($cH_{manifold}^+$). De tal manera que el pH en la línea es,

$$cH_{linea}^+ = X_{wm1} * cH_{m1}^+ + X_{wm2} * cH_{m2}^+ \quad (3.19)$$

$$pH = -\log cH_{linea}^+ \quad (3.20)$$

- **Variables de Gases Corrosivos**

Con el flujo total de gas en la línea,

$$flujo_{gaslinea} = flujo_{gasm1} + flujo_{gasm2} \quad (3.21)$$

Se determina el factor,

$$X_{gm1} = (flujo_{gasm1}/flujo_{gaslinea}) \quad (3.22.1)$$

$$X_{gm2} = (flujo_{gasm2}/flujo_{gaslinea}) \quad (3.22.2)$$

Y se calcula el total de cada variable de gas disuelto en el fluido, con la suma del producto para cada manifold,

$$CO2 = X_{gm1} * CO2_{m1} + X_{gm2} * CO2_{m2} \quad (3.23)$$

$$H2S = X_{gm1} * H2S_{m1} + X_{gm2} * H2S_{m2} + X_{gm3} * H2S_{m3} \quad (3.24)$$

3.1.4 Cupones Unidad Central de Procesamiento

La Unidad Central de Procesamiento comprende varios componentes en cada uno de los cuales debe existir un cupón. Esta información no estaba disponible en la CIC, y por tanto se descartó este conjunto de cupones.

3.1.5 Análisis de la información

Los datos obtenidos de CO₂, están dados como porcentaje de gas disuelto en el sistema, para el cálculo se debe hallar el CO₂ disuelto en el agua del sistema. Para esto se usó la ley de Henryⁱ, la cual enuncia que “a una temperatura

ⁱDisponible en: http://en.wikipedia.org/wiki/Henry's_law

constante, la cantidad de gas disuelta en un líquido es directamente proporcional a la presión parcial que ejerce ese gas sobre el líquido”.

$$K_h = \frac{PCO_2}{[CO_2]} \quad (3.25)$$

Donde:

K_h , es la constante de Henry.

PCO_2 , es la presión parcial de CO₂ en la fase de gas.

$[CO_2]$, es la concentración de CO₂ en la fase líquida.

- o La presión parcial de CO₂, se calcula así:

$$PCO_2 = \%CO_2 * Presión \quad [Atm] \quad (3.26)$$

La constante de Henry (K_h) Para el caso del CO₂, esta definida en función de la temperatura estándar (T =298 K):

$$K_h = 24,41 * e^{-2400 * (\frac{1}{T} - \frac{1}{298})} \quad [L * Atm/mol] \quad (3.27)$$

- o Concentración de CO₂

Con lo anterior ya desarrollado, finalmente se obtiene la concentración del CO₂ en agua,

$$[CO_2] = \frac{PCO_2}{K_h} \quad [mol/L] \quad (3.28)$$

En unidades [ppm= mg/L):

$$1 \text{ mol de CO}_2 = (16 \times 2) + (12)$$

$$1 \text{ mol de CO}_2 = 44 \text{ [g]} = 44000 \text{ [mg]}$$

Multiplicar $[\text{CO}_2] \text{ (mol/L)} \times 44000 \text{ [mg/mol]}$ da como resultado la concentración de CO_2 en partes por millón [ppm] en el agua.

3.2 FASE II (Selección de Variables)

En esta fase se seleccionan las variables de entrada más importantes, independientes y causales utilizando la metodología fuzzy (sección 2.2). Dicha metodología se aplica a la información de cupones recopilada en la fase anterior (fase I), figura 20.

	A	B	C	D	E	F	G	H	I	J
1	Presión (PSI)	Temperatura (°F)	Flujo Crudo (BPD)	Flujo Gas (mmscf/d)	Flujo Agua (BPD)	CO2 (ppm)	H2S (ppm)	PH	API	Rata Corrosión (mpy)
53	1267	172	478	43,641	109	2105,7047	5	5,9	42,6	1,13
54	1246	171	458	43,285	113	2664,30386	5	5,9	42,4	1,44
55	935	191	579	62,124	5878	1156,9365	3,2	6,97	30,4	8,92
56	1473	198	380	48,208	4621	2037,94226	3,4	6,76	30,3	4,39
57	1344	198	396	44,07	4904	1859,46667	4,8	6,79	30,3	10,34
58	777	155	797	6,46	714	1418,38739	5	6,87	39,6	0,50
59	662	145	391	27,662	104	1140,22289	4,8	6,41	35,4	0,40
60	715	152	457	58,635	229	1486,08228	4,2	6,89	36,8	0,48
61	796	171	243	18,777	1327	1458,92172	3	6,74	38,1	2,81
62	788	165	285	22,73	965	1542,47751	3	6,41	38,7	2,73
63	736	165	266	23,045	950	1344,64367	4	7,22	39,7	4,33
64	846	149	751	16,155	332	1489,61782	4	6,8	43,4	0,37
65	867	155	416	14,378	618	1899,21524	4	6,83	42,8	1,12
66	919	150	392	10,769	525	2132,56819	4	6,5	41,3	0,20

Figura 20: Datos recopilados de cupones (peor caso)

Fuente: Autor del proyecto

El uso de la metodología fuzzy se presenta de manera gráfica en un diagrama de flujo (figura 21) de acuerdo con el cual primero se generan las curvas fuzzy y el respectivo índice único de efectividad utilizado para determinar las variables más

importantes, descartando las variables de menor importancia y que aparentemente están aleatoriamente relacionadas con la variable de salida.

Con base en la clasificación dada por las curvas fuzzy se generan superficies fuzzy y el respectivo índice único de efectividad con el cual se descartan las variables que no hacen un buen trabajo de aproximación con respecto a la variable de salida. Así se obtiene la clasificación definitiva para usar con las redes neuronales.

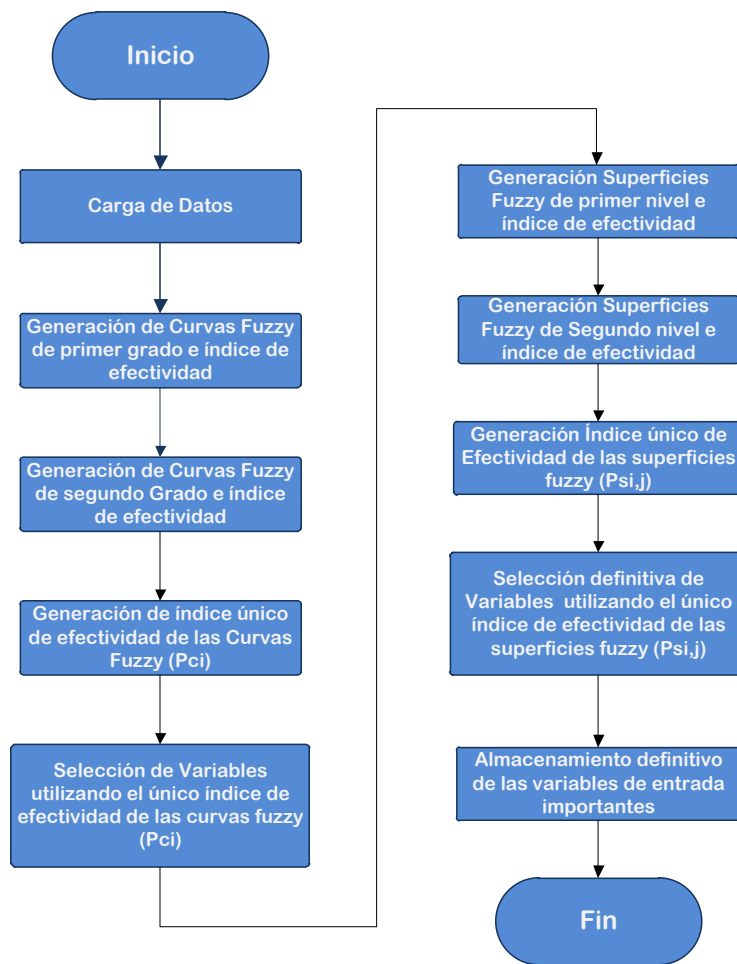


Figura 21: Diagrama de Flujo Selección de Variables

Fuente: Autor del proyecto

- **Cargar Datos:** Carga el archivo excel que comprende las variables de entrada y salida con los casos recopilados, figura 20.
- **Generar índice Único de Efectividad Curvas Fuzzy(PCi):** se combinan los índices de efectividad de curvas fuzzy de primer nivel Pv_c^i y de curvas fuzzy de segundo nivel Pv_c^i en un índice único de efectividad (Pc^i).
- **Generar Índice Único de Efectividad de las Superficies ($Ps^{i,j}$):** Con la variable más importante (X^i) generada en el índice único de efectividad de curvas fuzzy (Pc^i) se generan secuencialmente pares (X^i, X^j) con las variables restantes (X^j), ($j \neq i$) para generar las superficies fuzzy. La variable con el índice único de efectividad $Ps^{i,j}$ más pequeño es la siguiente variable más importante.
- **Almacenamiento definitivo de las variables de entrada importantes:** se guarda un archivo de texto con las variables más importantes. Esta información se usará en el entrenamiento de la red neuronal.

3.3 FASE III

3.3.1 Normalización de datos

Por conveniencia se recomienda normalizar las variables, especialmente cuando los valores son altos, lo que sucede con la información recopilada. Cada una de las variables (X^i) se normaliza en un rango $[-1, 1]$, según la siguiente función,

$$X_k^i \text{ normalizada} = \frac{2*[X_k^i - \min(X^i)]}{\max(X^i) - \min(X^i)} - 1 \quad (3.29)$$

Donde, X_k^i es la variable a normalizar: $i =$ tipo de variable de entrada. $k = 1, 2, 3, \dots, 207$ casos experimentales, información recopilada en cupones.

3.3.2 Implementación de la red neuronal

El proceso comprende cuatro etapas generales: agrupamiento y pre-procesamiento de los datos disponibles, establecimiento de la estructura de la red neuronal; generación y optimización de la red neuronal.

- **I Etapa:** Agrupamiento de los datos disponibles

En el agrupamiento propone el 92% del total de datos existentes (207 casos) para entrenar la red (III Etapa) y el 8% restante para determinar velocidades de corrosión con la red ya estructurada y entrenada (IV Etapa); de esta manera se observa su efectividad.

- **II Etapa:** Establecimiento de capa oculta y numero de neuronas

Inicialmente se propone el uso de perceptrón multicapa de una capa oculta con función de transferencia sigmoidea y capa de salida con función de transferencia linealⁱ. Esta selección se basa en la experiencia con redes neuronales reportada en la literatura, la cual recomienda no utilizar más de una capa oculta, para reducir la carga computacional de la red [6]. Otros aportes como el teorema de Cibenko, demuestran que el perceptrón de una sola capa oculta y función de transferencia sigmoidea en las neuronas de la capa pueden aproximar funciones continuas [16].

Se han propuesto diversidad de técnicas para establecer el número de neuronas en la capa oculta de la red. Una de las formas más señaladas en la literatura consiste en iniciar con un número grande de neuronas, e ir disminuyéndolo; según el rendimiento obtenido durante el entrenamiento con cada número de neuronas en la capa oculta, se selecciona la mejor estructura neuronal.

Para determinar el número de neuronas en la capa oculta se utilizó inicialmente el toolbox para redes neuronales de MATLAB®, y finalmente la librería de código

ⁱ $f(x) = x$

abierto FANNⁱ. Los dos procedimientos arrojaron buenos resultados con números de neuronas menores a 40.

- **III Etapa:** Entrenamiento de la Red Neuronal

Para entrenar la red neuronal se hace uso de la librería de código abierto Fast Artificial Neural Network (**FANN**) escrita en lenguaje de programación C++; se procede a entrenar la red neuronal:

- La capa de entrada consta de nueve (9) neuronas que corresponden a las variables del problema, seleccionadas por metodología fuzzy: Presión, Temperatura, Flujo de Crudo, Flujo de Gas, Flujo de Agua, pH, CO₂, H₂S y API.
- La capa de salida está definida por una neurona que corresponde a la variable, velocidad de corrosión que es el objetivo del problema.
- La función de transferencia para la capa oculta es sigmoidea, *tansig*; la capa de salida tiene una función *lineal*.
- Los pesos sinápticos se inician de manera aleatoria en un rango de [-2, 2].
- Parámetros de parada para el entrenamiento: 60000 épocas y rendimiento (error cuadrático medio) de 0.001.
- El algoritmo de entrenamiento utilizado es iRPROP [19], el cual hace una ligera modificación en el algoritmo adaptativo RPROP (resilient propagation) mejorando la velocidad de aprendizaje. Seleccionado por ser un algoritmo de adaptación local que reduce de manera significativa el tiempo de entrenamiento y mejora el rendimiento de la red.

ⁱ Tomado de: <http://leenissen.dk/fann/wp/>

Rprop es propuesto por Riedmiller y Braun, es un algoritmo que modifica el valor de cada peso sináptico según el comportamiento de la secuencia signos de la derivada parcial en cada dimensión del espacio-peso [18].

- **IV Etapa:** Evaluación de la red

Durante el entrenamiento se optimiza la red por reducción del error cuadrático medio de los valores de las salidas simuladas por la red con respecto a los valores experimentales (algoritmo iRPROP).

Con el 8% de los datos experimentales que no fueron incluidos en la etapa anterior, se determinan velocidades de corrosión con la red ya estructurada y entrenada; de esta manera se determina el rendimiento de la red.

De forma grafica se muestra el proceso descrito (I hasta la IV Etapa) en un diagrama de flujo, figura 22. Posterior a la figura se explica con más detalle cada proceso en el diagrama de flujo.

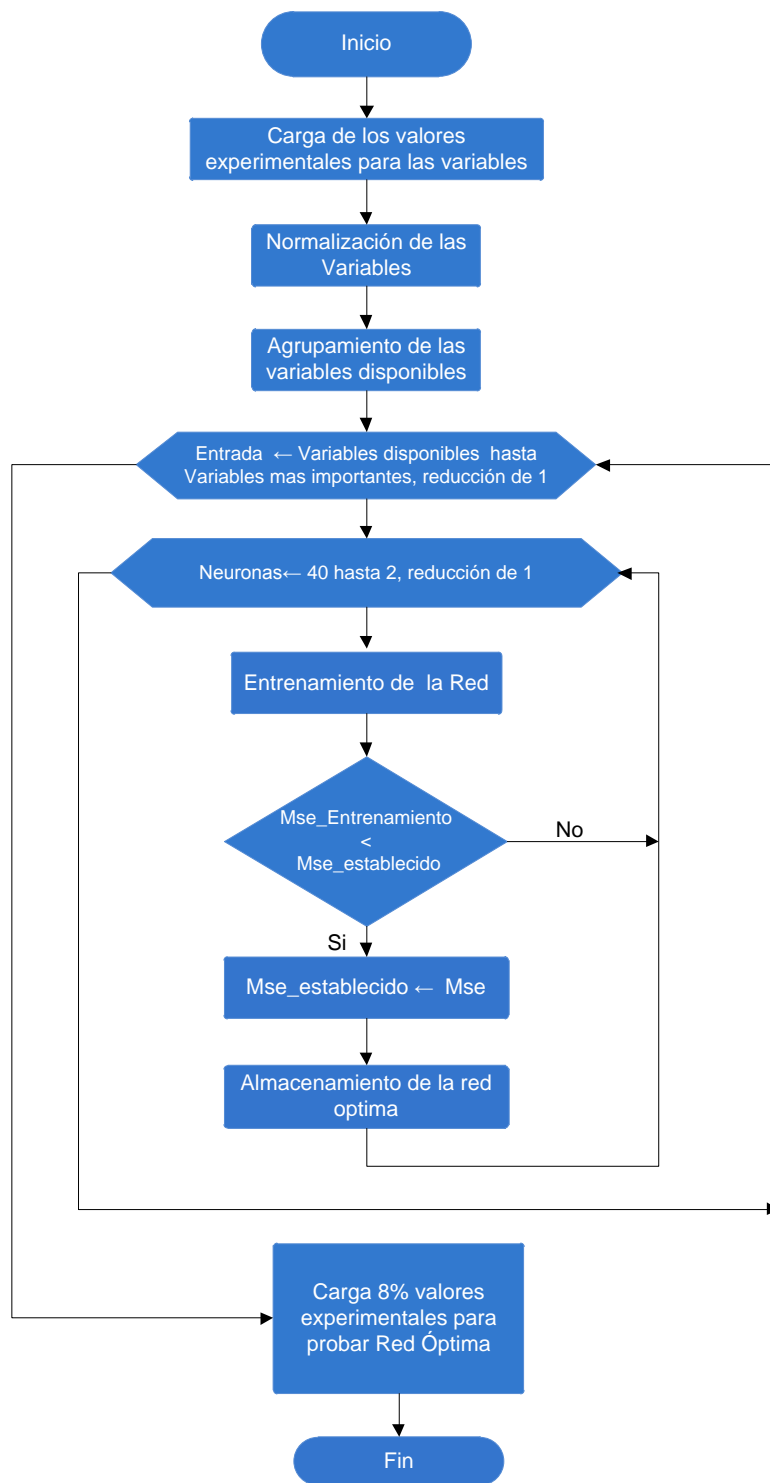


Figura 22: Diagrama de Flujo del entrenamiento de la red

Fuente: Autor del proyecto

- **Cargar de los datos experimentales:** Se carga el total de variables de entrada recopiladas y la variable de salida con el total de casos recopilados.
- **Entrada ← Variables disponibles hasta Variables más importantes, reducción de 1:** ciclo que busca entrenar la red utilizando inicialmente el total de las variables propuestas para el cálculo y luego gradualmente las variables más importantes seleccionadas por la metodología fuzzy.
- **Agrupamiento de los casos:** Se utiliza el 92% de los casos recopilados para entrenar la red. El 8% restante se usa para evaluar la efectividad de la red para determinar velocidades de corrosión.
- **Neuronas ← 40 hasta 2, reducción de 1:** Ciclo iterativo que entrena la red neuronal, inicialmente con un número elevado de neuronas en la capa oculta y que va reduciendo gradualmente éste número en 1 hasta llegar a 2 neuronas en la capa oculta.
- **Mse_establecido ← Mse:** Fijación de valor mínimo para error cuadrático medio (mse). Si el error cuadrático medio durante el entrenamiento de la red es menor que el mínimo propuesto, se cumple la condición de un buen rendimiento, y se sigue al paso siguiente, almacenamiento.
- **Almacenamiento de la red óptima:** Busca en la red con el mejor rendimiento; almacena la mejor estructura de la red y los pesos sinápticos del entrenamiento.
- **Carga 8% valores experimentales para probar Red Óptima:** determina velocidades de corrosión con el 8% de casos experimentales no incluidos en el entrenamiento de la red.

3.3.3 Ratificación en la estructura neuronal

Con el propósito de generar algunos prototipos de redes neuronales para ver su rendimiento empleando la información recopilada de cupones, se hace uso del

toolbox de MATLAB® como prueba de conceptos. De esta manera se logra idear la estructura y el funcionamiento de la red neuronal.

Un primer prototipo de red, con dos capas ocultas y 9 neuronas de entrada, presento: 32 neuronas en la primera capa oculta y 13 neuronas en la segunda. Las capas ocultas con función de transferencia sigmoidea ($tansig^i$) y la neurona de la capa de salida función de transferencia lineal.

Los resultados de la red de dos capas ocultas fueron buenos. Pero otro prototipo de red presento una capa oculta de 33 neuronas con mejores resultados. Esto se determinó con base en el error cuadrático medio (mse) [17] durante el entrenamiento de la red, el cual para la estructura con dos capas ocultas fue de 0,5572 mientras que para la estructura neuronal de una capa oculta fue de 0,3369.

$$mse = \frac{1}{N} \sum_{i=1}^N (t_i - y_i)^2 \quad (3.30)$$

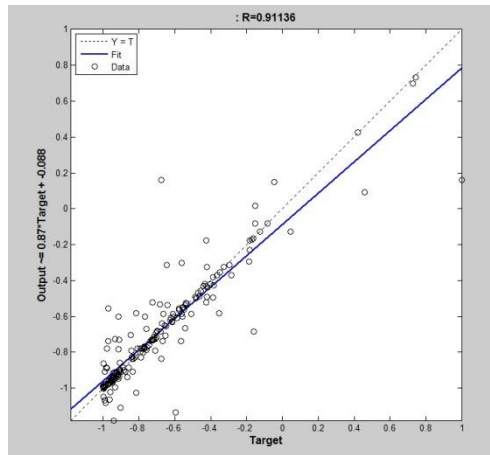
N , total de datos usados durante el entrenamiento

t_i , datos experimentales

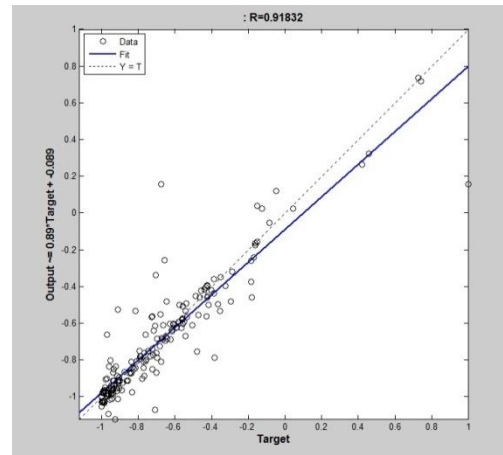
y_i , datos simulados por la red neuronal

La figura 23 muestra los resultados del cálculo de regresión lineal para ambas estructuras.

ⁱ $f(x) = \frac{1}{1+e^{-2x}} - 1$



(a) 2 capas ocultas



(b) 1 capa oculta

Figura 23: Regresión lineal

Fuente: Autor del proyecto

Sin embargo los resultados del error cuadrático medio (mse), muestran un mejor comportamiento de la red neuronal de una sola capa.

Adicionalmente el tiempo de trabajo para la red de dos capas ocultas fue de 9,55 horas mientras que para la estructura de una capa oculta fue 39,42 minutos.

Con los resultados anteriores, la idea de usar una red neuronal de una sola capa oculta deja de ser una idea y pasa a ser un hecho.

3.4 FASE IV

Se estructura una interfaz gráfica de fácil manejo en lenguaje C#, que enlaza el pre-procesamiento de la información (metodología fuzzy), además de la estructuración y entrenamiento de la red neuronal bajo lenguaje de programación C++. Este software generado se integra en el software de la corporación (CIC).

3.4.1 Diagramas UML

Se realiza un análisis previo de lo que se quiere hacer, estableciendo la funcionalidad propia del sistema por medio de los llamados casos de uso.

Los casos de uso representan las interacciones entre el usuario y el sistema, describiendo la función que ofrece el sistema como una serie de eventos visibles para el usuario.

Los diagramas de casos de uso esquematizan la funcionalidad del sistema desde el punto de vista del usuario. Cada caso de uso se representa como una elipse (figura 24) y define una forma de usar el sistema.

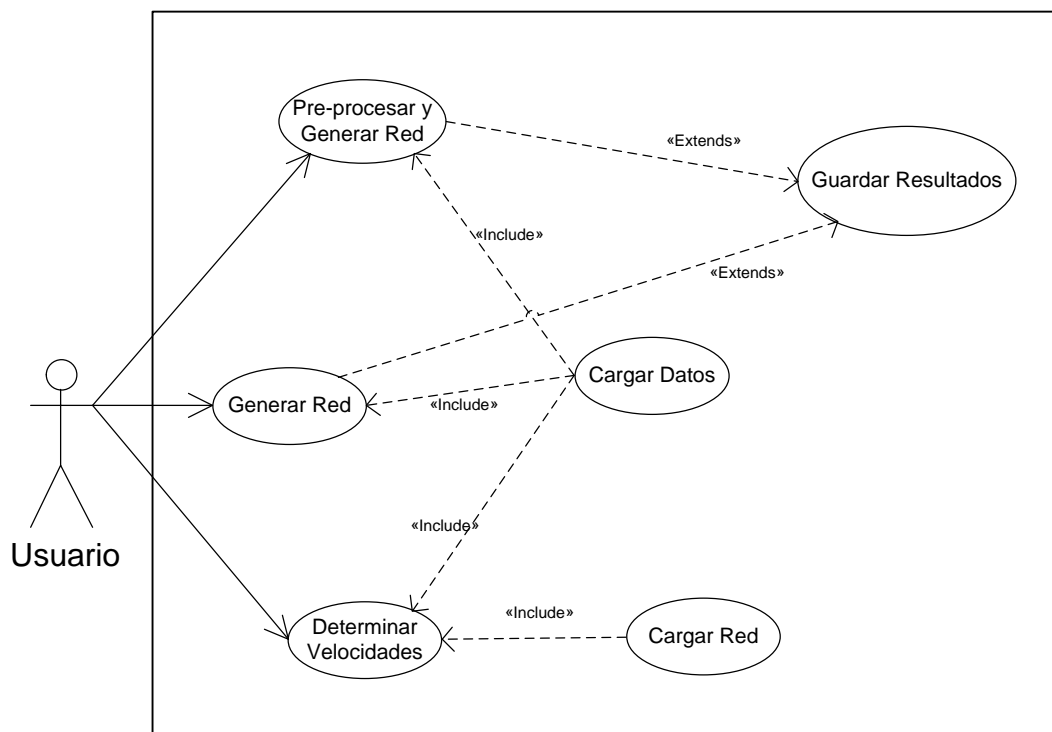


Figura 24: Diagrama casos de uso de la herramienta

Fuente: Autor del proyecto

Caso de Uso:	Pre-procesar y Generar Red
Tipo:	Básico
Actor:	Usuario
Propósito:	Realizar el pre-análisis de la información y generar la red con base en dicho análisis.
Resumen:	Este caso de uso busca analizar las variables de entrada y determinar cuáles son variables más importantes e independientes, para continuar con el entrenamiento de la red alterando las neuronas en su capa oculta y así seleccionar la red con el mejor rendimiento.
Precondición(es)	Ninguna
Flujo Principal:	<p>El usuario presiona el botón “Pre-procesamiento y Generación de la Red” en la pantalla principal (P-1) Figura 33.</p> <p>Se ejecuta el caso de uso “Cargar Datos”.</p> <p>Si se cargan los datos se realiza el análisis de los datos y posterior a esto, se ejecuta el sub-flujo, Entrenar con Análisis (S-1).</p>
Sub-flujos:	<p>S-1 Entrenar con Análisis</p> <p>Presenta al usuario un formulario con la información detallada del análisis de curvas y superficies Fuzzy, además presenta las opciones de “Continuar”, “Guardar” y “Cancelar”.</p> <p>Si selecciona “Continuar”, procede a entrenar la red con el total de variables de entrada variando el número de neuronas en la capa oculta. Luego entrena con las variables de entrada más importantes según el análisis fuzzy realizado y almacena la red con el mejor rendimiento obtenido en su entrenamiento.</p> <p>Finaliza con el sub-flujo Resultados Red (S-2).</p> <p>La opción “Guardar”, permite almacenar en el equipo el resultado del pre-procesamiento de la información.</p> <p>Seleccionar “Cancelar” enviara al usuario a la página principal (P-1).</p> <p>S-2 Resultados Red</p> <p>Presenta al usuario un formulario con información del entrenamiento y las opciones “Guardar” y “Cerrar”.</p> <p>Si la actividad seleccionada es “Guardar” se ejecuta el caso de uso Almacenar.</p> <p>Si se selecciona “Cerrar” enviara al usuario a la página principal (P-1).</p>

Tabla 4: Caso de uso Pre-procesar y Generar Red

Caso de Uso:	Generar Red
Tipo:	Básico
Actor:	Usuario
Propósito:	Generar Red neuronal sin Pre-Análisis de la información.
Resumen:	Este caso de uso se entrena la red con las variables cargadas en la información, variando el número de neuronas en la capa oculta.
Precondición(es)	Ninguna
Flujo Principal:	El usuario selecciona “Pre-procesamiento y Generación de la Red” en la pantalla principal (P-1) Figura 33. Se ejecuta el caso de uso “Cargar Datos”. Si se cargan los datos se entrena la red y posterior a esto se da paso al sub-flujo Resultados Red (S-2).
Sub-flujos:	Ninguno

Tabla 5: Caso de uso Generar Red

Caso de Uso:	Determinar Velocidades
Tipo:	Básico
Actor:	Usuario
Propósito:	Uso de la red ya entrenada para determinar velocidades de corrosión.
Resumen:	Este caso de uso carga la información de la red ya entrenada, para que el usuario determine velocidades de corrosión utilizando información recopilada del mismo tipo experimental.
Precondición(es)	Ninguna
Flujo Principal:	En la pantalla principal (P-1) Figura 33, el usuario selecciona “Pre-procesamiento y Generación de la Red”. Se ejecuta el caso de uso “Cargar Datos”, luego se ejecuta el caso de uso “Cargar Red”. Si se ejecutan con éxito los casos anteriores se cargan los datos, se entrena la red y posterior se da paso al sub-flujo Resultados Red (S-2). Si no se ejecutan con éxito y no se carga la información el usuario vuelve al formulario principal (P-1)
Sub-flujos:	Ninguno

Tabla 6: Caso de uso Determinar Velocidades

Caso de Uso:	Cargar Datos
Tipo:	Inclusión
Actor:	Usuario
Propósito:	Cargar información almacenada en archivos de Excel
Resumen:	Esta caso de uso muestra un formulario en el cual el usuario debe seleccionar el archivo en el que se encuentra la información que desea usar en la herramienta
Precondición(es)	Se requiere haber ejecutado alguno de los siguientes casos de uso “Pre-Análisis y Entrenamiento” , “Entrenar Red” o “Usar Red”
Flujo Principal:	Se presenta al usuario un formulario para buscar el archivo y las opciones “Aceptar” y “Cancelar”. Si selecciona el archivo y aceptar se cargara la información correspondiente para determinado caso de uso dela precondición. Seleccionar “Cancelar” devolverá al usuario al formulario principal (P-1).
Sub-flujos:	Ninguno

Tabla 7: Caso de uso Cargar Datos

Caso de Uso:	Guardar Resultados
Tipo:	Extensión
Actor:	Usuario
Propósito:	Almacenar información de la estructura y el análisis de resultados de la red neuronal entrenada.
Resumen:	Este caso de uso se almacena en una carpeta un archivo de texto con la información de los resultados del entrenamiento neuronal. También un script con la información de la estructura neuronal (capas y pesos).
Precondición(es)	Se requiere haber ejecutado alguno de los siguientes casos de uso “Pre-Análisis y Entrenamiento” o “Entrenar Red”.
Flujo Principal:	Al ejecutarse este caso de uso se presenta un formulario para dar nombre el archivo además de las opciones “Aceptar” y “Cancelar”. “Aceptar” almacenara los archivos. “Cancelar” enviara el usuario al formulario principal (S-1)
Sub-flujos:	Ninguno

Tabla 8: Caso de uso Guardar Resultados

Caso de Uso:	Cargar Red
Tipo:	Extensión
Actor:	Usuario
Propósito:	Cargar el script almacenado de la estructura neuronal.
Resumen:	Carga el la script con la información de la red neuronal ya entrenada.
Precondición(es)	Se requiere haber ejecutado caso de uso "Usar Red".
Flujo Principal:	<p>Se presenta al usuario un formulario para buscar el archivo y las opciones "Aceptar" y "Cancelar".</p> <p>Si selecciona el archivo y aceptar se cargara la información correspondiente para determinar velocidades de corrosión.</p> <p>Seleccionar "Cancelar" devolverá al usuario al formulario principal (P-1).</p>
Sub-flujos:	Ninguno

Tabla 9: Caso de uso Cargar Red

4 RESULTADOS

En este capítulo se muestran los resultados obtenidos en la recopilación de información para cupones y la aplicación de la metodología fuzzy a dicha información. También se muestran los resultados de redes neuronales generadas y su respectivo desempeño obtenido.

4.1 Base de Datos Cupones

La información recopilada de cupones comprende:

	WELL HEAD	MANIFOLD	LINEA	Total
Cupones Con Información	107	82	18	207
Cupones descartados	7	9	15	31
Total Cupones	114	91	33	238

Tabla 10: Información recopilada

Los cupones con información (207 en total), se albergaron en un documento Excel, que cuenta con dos hojas de datos (figura 25). En una se encuentran los datos recopilados que registraron sólo una fecha dentro del rango de vida del cupón (Peor Caso), y en la otra los mismos datos, remplazando los casos en que existía más de un registro (Promedio),

CUPONES - Microsoft Excel

	A	B	C	D	E	F	G	H	I	J
1	Presión (PSI)	Temperatura (°F)	Flujo Crudo (BPD)	Flujo Gas (mmscf/d)	Flujo Agua (BPD)	CO2 (ppm)	H2S (ppm)	PH	API	Rata Corrosión (mpy)
53	1267	172	478	43,641	109	2105,7047	5	5,9	42,6	1,13
54	1246	171	458	43,285	113	2664,30386	5	5,9	42,4	1,44
55	935	191	579	62,124	5878	1156,9365	3,2	6,97	30,4	8,92
56	1473	198	380	48,208	4621	2037,94226	3,4	6,76	30,3	4,39
57	1344	198	396	44,07	4904	1859,46667	4,8	6,79	30,3	10,34
58	777	155	797	6,46	714	1418,38739	5	6,87	39,6	0,50
59	662	145	391	27,662	104	1140,22289	4,8	6,41	35,4	0,40
60	715	152	457	58,635	229	1486,08228	4,2	6,89	36,8	0,48
61	796	171	243	18,777	1327	1458,92172	3	6,74	38,1	2,81
62	788	165	285	22,73	965	1542,47751	3	6,41	38,7	2,73
63	736	165	266	23,045	950	1344,64367	4	7,22	39,7	4,33
64	846	149	751	16,155	332	1489,61782	4	6,8	43,4	0,37
65	867	155	416	14,378	618	1899,21524	4	6,83	42,8	1,12
66	919	150	392	10,769	525	2132,56819	4	6,5	41,3	0,20

(a) Archivo con Datos de Peor Caso

CUPONES_Promedio - Microsoft Excel

	A	B	C	D	E	F	G	H	I	J
1	Presión (psig)	Temperatura (F)	Flujo Crudo	Flujo Gas	Flujo Agua	CO2(ppm)	H2S (ppm)	PH	API	Rata Corrosión (mpy)
53	1267	172	478	43,641	109	2105,70	5	5,9	42,6	1,13
54	1246	171	458	43,285	113	2378,84	5	5,9	42,4	1,44
55	1359	188,33	483	52,126	4665	1728,22	3,2	6,97	31,2	8,92
56	1473	198	380	48,208	4621	1868,11	3,3	6,76	30,3	4,39
57	1344	198	396	44,07	4904	1859,47	4,8	6,8	30,3	10,34
58	822	150,4285	701	5,45	665,1428	1581,67	5	6,87	39,1	0,50
59	692	146,5	368	27,092	104	1171,01	4,8	6,41	35,4	0,40
60	734	153,5	443,5	58,097	229	1499,44	4,2	6,89	36,8	0,48
61	1131,33	164,66	235,33	16,994	878,166	2037,65	3	6,74	40,35	2,81
62	788	165	285	22,73	965	1542,48	3	6,61	38,7	2,73
63	736	165	266	23,045	950	1344,64	4	7,22	39,7	4,33
64	1037	142,1428	467	12,571	288	1979,72	4	6,8	42,9	0,37
65	948,5	154	408	13,91	515	1891,51	3,75	6,83	43	1,12
66	1126,75	143,75	268	8,2	269,75	2670,71	5	6,5	43,3	0,20

(b) Archivos con Datos Promedio

Figura 25: Archivos Excel con Datos Recopilados

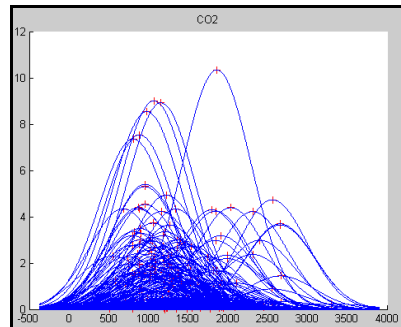
Fuente: Autor del proyecto

4.2 Pre-procesamiento fuzzy de las variables de entrada

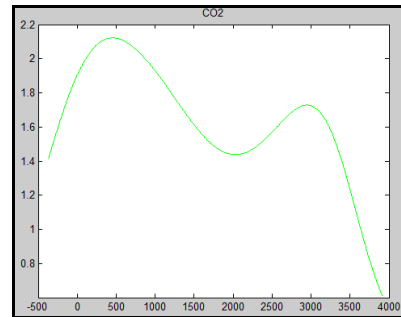
Se desarrolló pre-procesamiento de datos, que comprende selección de variables de entrada más importantes, independientes entre si y relacionadas causalmente con la variable de salida en los dos tipos de datos recopilados, Peor Caso y

Promedio.

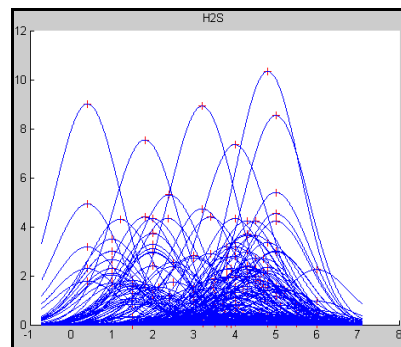
Enseguida se muestran como ejemplo de éste pre-procesamiento las gráficas de interés del proceso fuzzy correspondientes a las dos variables de entrada CO2 y H2S para los tipos de datos recopilados en la Fase I (figura 26 y 27).



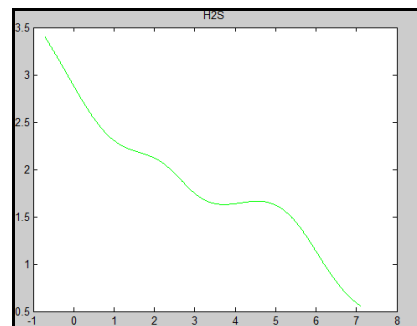
(a) Función de membresía CO2



(b) Curvas Fuzzy CO2

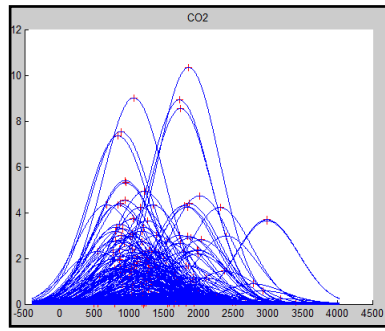


(c) Función de membresía H2S

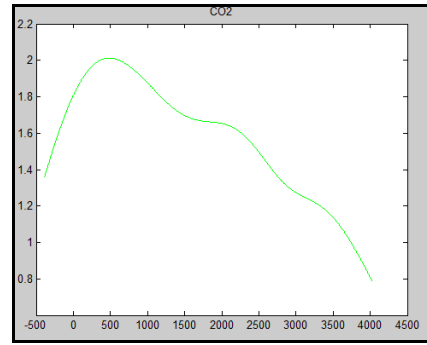


(d) Curvas Fuzzy H2S

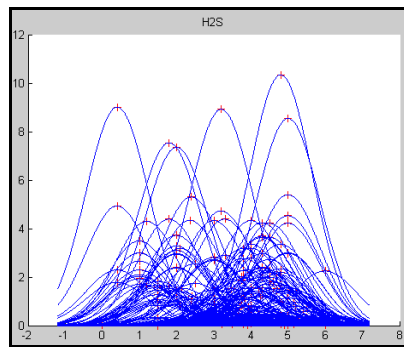
Figura 26: CO2 y H2S en Datos de Peor Caso



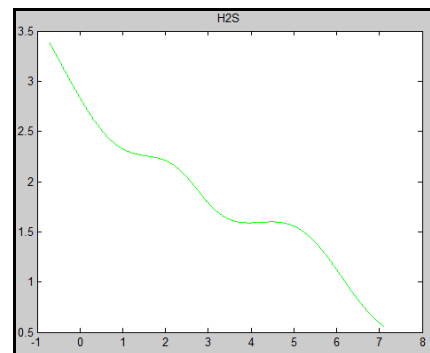
(a) Función de membresía CO2



(b) Curvas Fuzzy CO2



(c) Función de membresía H2S



(d) Curvas Fuzzy H2S

Figura 27: CO2 y H2S en Datos Promedio

Fuente: Autor del proyecto

Con respecto a la fase II del desarrollo (sección 3.2) se obtiene los siguientes resultados.

- **Clasificación de las variables de entrada con base en las Curvas Fuzzy de Primer Grado**

Se clasifican las posibles entradas en orden de importancia, las variables dependientes o relacionadas aún NO se descartan.

La tabla 11 muestra el análisis de curvas fuzzy de primer grado realizado a los dos tipos de información recopilada.

Datos Peor Caso	Datos Promedio
1. Temperatura	1. Temperatura
2. API	2. API
3. Flujo Agua	3. Flujo Agua
4. H2S	4. H2S
5. CO2	5. Presión
6. Presión	6. Flujo Gas
7. Flujo Gas	7. PH
8. PH	8. CO2
9. Flujo Crudo	9. Flujo Crudo

Tabla 11: Resultados Curvas Fuzzy de primer Grado para las variables recopiladas

Se puede notar que la variable más importante es la temperatura para los dos tipos de datos recopilados. Además las variables conservan el mismo orden hasta la variable H2S.

- **Clasificación de las variables de entrada con base en las Curvas Fuzzy de Segundo Grado**

Se presenta un estimado local de la varianza a lo largo de los 207 casos para cada variable (Tabla 12).

Datos Peor Caso	Datos Promedio
1. Flujo Crudo	1. Flujo Crudo
2. PH	2. PH
3. CO2	3. CO2
4. Presión	4. Flujo Gas
5. H2S	5. H2S
6. Flujo Gas	6. Presión
7. Flujo Agua	7. Flujo Agua
8. Temperatura	8. Temperatura
9. API	9. API

Tabla 12: Resultados de Curvas Fuzzy de Segundo Grado para las variables recopiladas

Las variables para ambos casos presentan un orden similar hasta la variable de entrada CO₂. Y el API es la variable con comportamiento no causal más pronunciado con respecto a la variable de salida.

- **Clasificación de las variables de entrada con base al índice único de efectividad de Curvas Fuzzy**

Con el resultado de las curvas fuzzy de primer y segundo grado se presenta un índice único de efectividad. En orden se presentan las variables más importantes y de comportamiento causal con respecto a la variable de salida (tabla 13).

Datos Peor Caso	Datos Promedio
1. Flujo Crudo	1. Flujo Crudo
2. H2S	2. H2S
3. PH	3. PH
4. CO2	4. CO2
5. Presión	5. Flujo Gas
6. Flujo Gas	6. Presión
7. Temperatura	7. Flujo Agua
8. Flujo Agua	8. Temperatura
9. API	9. API

Tabla 13: Índice único de Efectividad Curvas Fuzzy en los datos recopilados

- **Proceso de selección de las variables de entrada basado en el índice único de efectividad para superficies fuzzy**

VARIABLES DESCARTADAS LUEGO DE APLICAR EL ÍNDICE ÚNICO DE EFECTIVIDAD A LAS CURVAS FUZZY, POR SER VARIABLES POCO IMPORTANTES Y DE COMPORTAMIENTO NO CAUSAL CON RESPECTO A LA VARIABLE DE SALIDA (TABLA 14).

Datos Peor Caso	Datos Promedio
1. Flujo Gas	1. Flujo Gas
2. Temperatura	2. Presión
3. Flujo Agua	3. Flujo de Agua
4. API	4. Temperatura
	5. API

Tabla 14: Eliminación índice único de Efectividad C.F.

- **Clasificación de las variables de entrada con base en el índice único de efectividad para Superficies Fuzzy**

En orden las variables más independientes y de comportamiento causal para los datos de peor caso se presentan en la tabla 15 y para los datos de caso promedio se en la tabla 16.

Índice único de Efectividad Curvas	Índice único de Efectividad Superficies
1. Flujo Crudo 2. H2S 3. PH 4. CO2 5. Presión 6. Flujo Gas 7. Temperatura 8. Flujo Agua 9. API	1. Flujo Crudo 2. PH 3. H2S 4. CO2 5. Presión

Tabla 15: Índice Efectividad Superficies Fuzzy de primer Grado en Datos Peor Caso

Índice único de Efectividad Curvas	Índice único de Efectividad Superficies
1. Flujo Crudo 2. H2S 3. CO2 4. PH 5. Flujo Gas 6. Temperatura 7. Presión 8. API 9. Flujo Agua	1. Flujo Crudo 2. PH 3. CO2

Tabla 16: Índice único efectividad Superficies Fuzzy de primer Grado en Datos Promedio

- **Proceso de selección de las variables de entrada basado en el índice único de efectividad para superficies fuzzy**

Las variables descartadas luego de aplicar el índice de efectividad de las superficies fuzzy se muestran en la tabla 17.

Datos Peor Caso	Datos Promedio
1. Presión	1. H2S

Tabla 17: Eliminación variables por índice único de efectividad Superficies Fuzzy

En la tabla 18 se presenta el resultado final de la metodología fuzzy, mostrando las variables más importantes para los dos tipos de datos recopilados.

Datos Peor Caso	Datos Promedio
1. Flujo Crudo	1. Flujo Crudo
2. PH	2. PH
3. CO2	3. CO2
4. H2S	

Tabla 18: Variables más significativas

Los resultados del pre-análisis para ambos tipos de datos son similares.

Para el desarrollo de la red neuronal se utilizan las variables de entrada más importantes de acuerdo a los datos de peor caso ya que estos resultados incluyen las variables de entrada CO2 y H2S.

4.3 Red Neuronal

En esta sección se muestran resultados obtenidos con las redes neuronales.

Una primera búsqueda de la mejor estructura neuronal, con 9 neuronas en la capa de entrada (Presión, Temperatura, Flujo de Crudo, Flujo de Gas, Flujo de Agua, CO2, H2S, pH y API) y una neurona en la capa de salida (velocidad de corrosión) arroja los siguientes resultados.

37 neuronas en la capa oculta con un error cuadrático medio en el entrenamiento de 0,004713.

La diferencia entre los valores experimentales y los simulados para el total de casos (207) se muestran en la figura 28; la figura 29 muestra la evaluación de la red (sección 3.3.2 Etapa IV) con datos que no fueron presentados durante la etapa de entrenamiento.

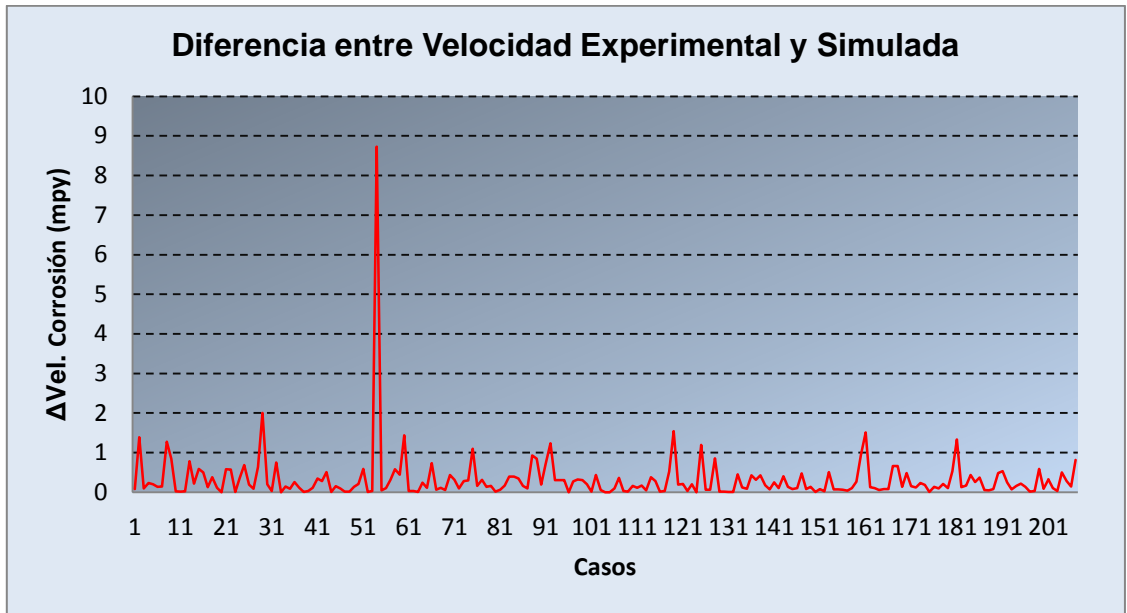


Figura 28: Diferencia en la velocidad de corrosión entre datos simulados y experimentales

Fuente: Autor del proyecto

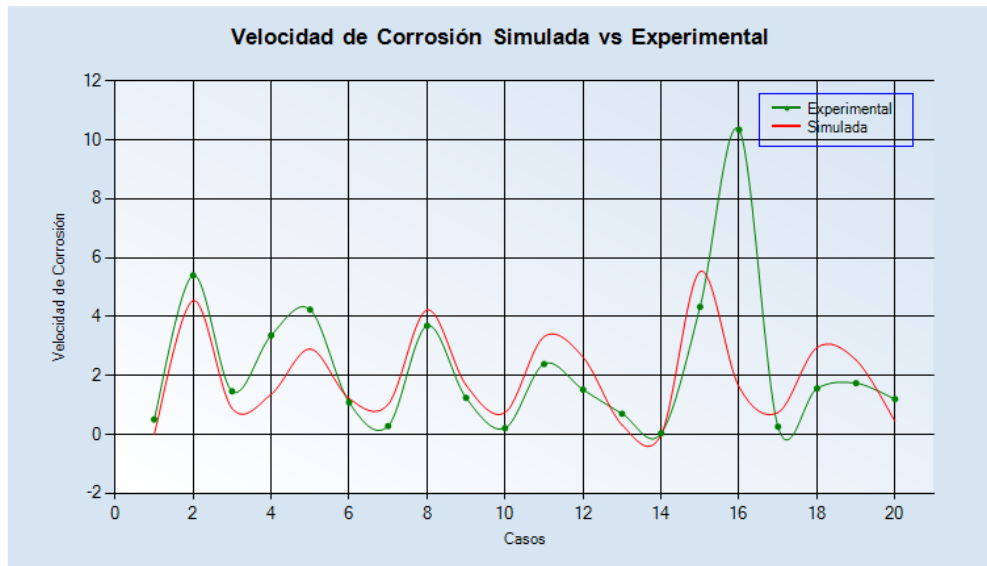


Figura 29: Velocidad de corrosión simulada Vs Experimental

Fuente: Autor del proyecto

Modificando la capa de entrada de la red según el análisis obtenido en la metodología fuzzy, se logra una red óptima con 8 neuronas en la capa de entrada correspondientes a las variables Presión, Temperatura, Flujo de Crudo, Flujo de Gas, Flujo de Agua, CO₂, H₂S y pH; 30 neuronas en la capa oculta y un error cuadrático medio en el entrenamiento de 0,0028 .

La figura 30 muestra los resultados de la diferencia entre los valores simulados por la red y los experimentales para los 207 casos de cupones. La figura 31 muestra los resultados de evaluar la red con datos que no fueron presentados durante la etapa de entrenamiento.

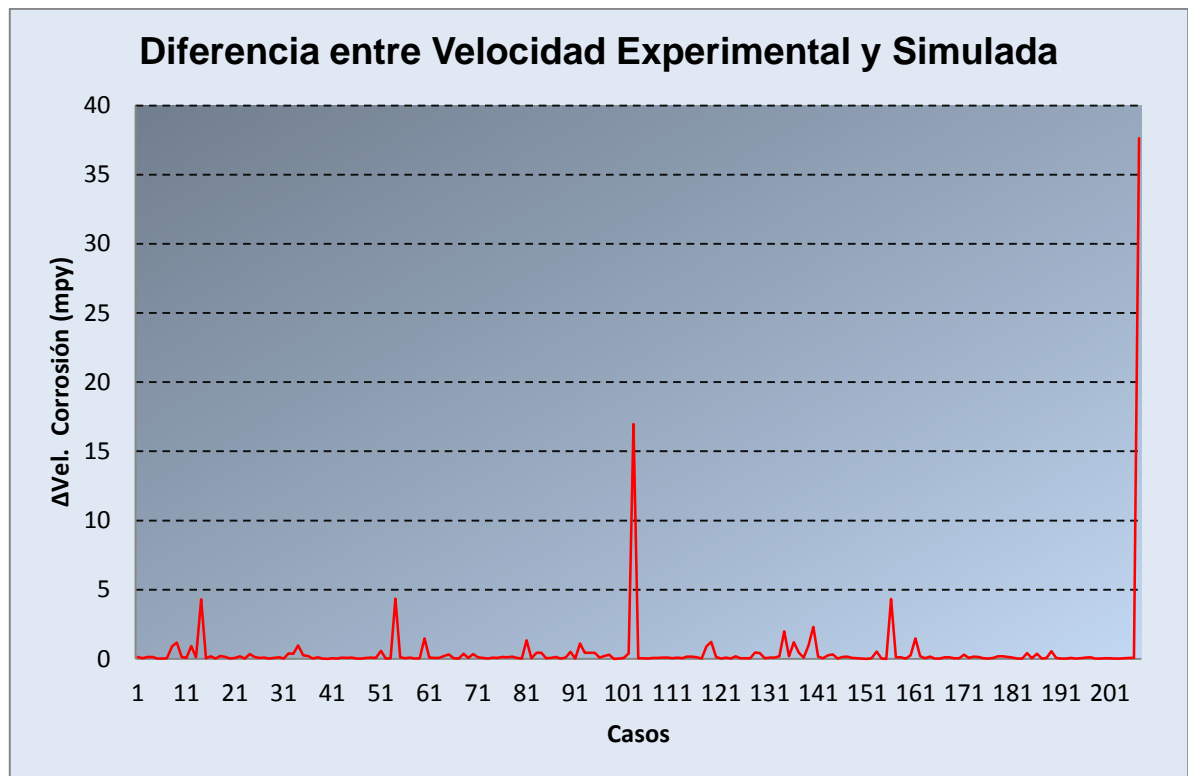


Figura 30: Diferencia entre datos Simulados y Experimentales (8 variables de entrada)

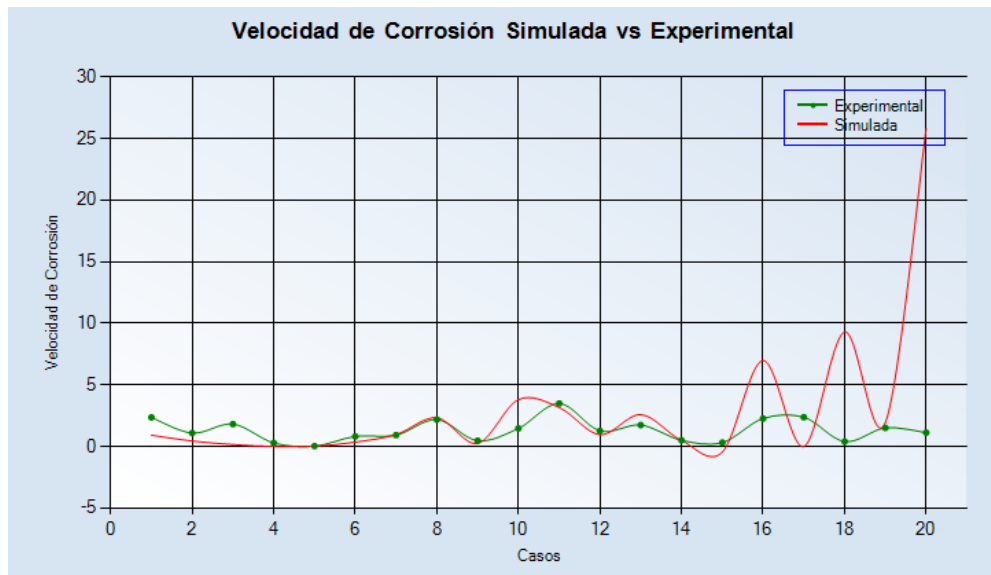


Figura 31: Velocidad de corrosión simulada Vs Experimental

Fuente: Autor del proyecto

Una mejor comparación de redes neuronales generadas (prototipos) se muestra en la figura 32 y la tabla 19; se anotan los resultados del mejor rendimiento obtenido durante el entrenamiento (MSE), además del número de neuronas en la capa oculta (NCO) y las variables en su capa de entrada según los resultados de la metodología fuzzy (variables en orden de importancia):

- 9 Variables de entrenamiento; no se descartan variables.
- 8 Variables de entrenamiento; se descarta API.
- 7 Variables de entrenamiento; se descartan: API y Flujo Agua.
- 6 Variables de entrenamiento; se descartan: API, Flujo Agua y Temperatura.
- 5 Variables de entrenamiento; se descartan: API, Flujo Agua, Temperatura y Flujo Gas.
- 4 Variables de entrenamiento; se descartan: API, Flujo Agua, Temperatura, Flujo Gas y Presión.

Neuronas capa entrada	9		8		7		6		5		4	
	NCO	MSE	NCO	MSE	NCO	MSE	NCO	MSE	NCO	MSE	NCO	MSE
1	23	0,0102	22	0,0033	24	0,0105	37	0,0109	37	0,0043	29	0,0134
2	24	0,0091	39	0,0034	27	0,0104	38	0,0105	36	0,0041	36	0,0124
3	25	0,0103	37	0,0031	37	0,0103	30	0,0106	40	0,0043	35	0,0137
4	26	0,0102	33	0,0033	36	0,0103	39	0,0105	33	0,0038	29	0,0120
5	25	0,01053	30	0,0028	30	0,0101	27	0,0058	35	0,0037	32	0,0131

Tabla 19: Resultados redes neuronales

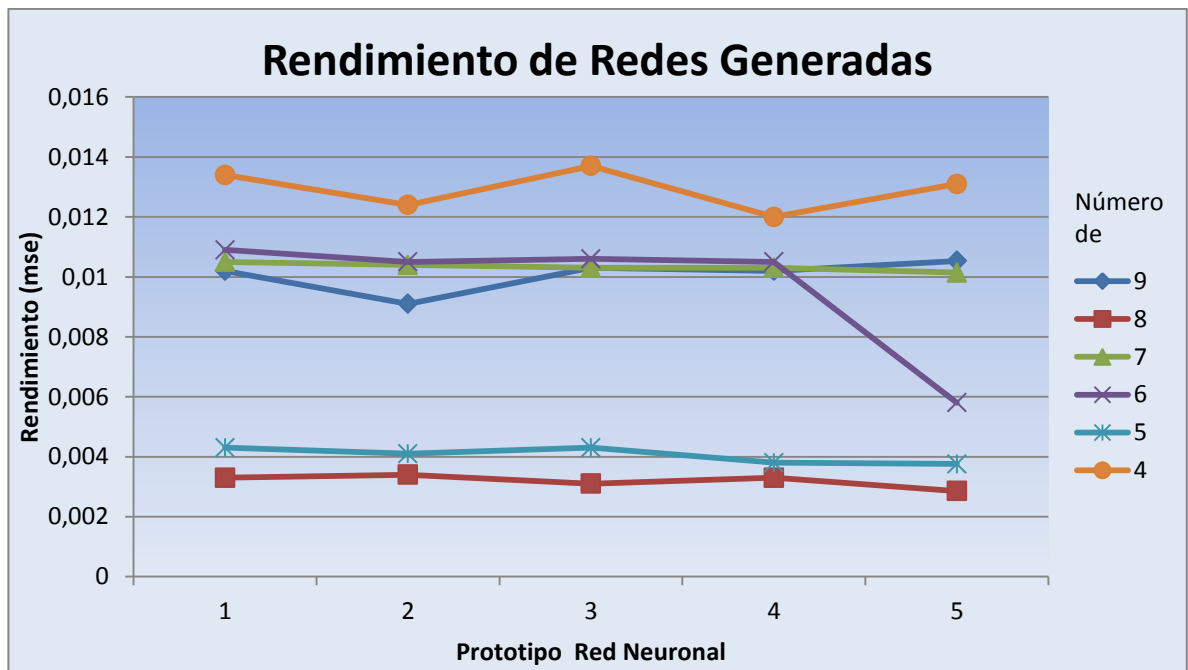


Figura 32: Rendimiento de las redes dependiendo de las variables de entrada

Fuente: Autor del Proyecto

5 CONCLUSIONES

- ✓ El uso de curvas y superficies fuzzy determinó las variables de entrada más importantes, independientes y no causales con respecto a la velocidad de corrosión. En este grupo de variables se encontraron el CO₂ y H₂S que son los agentes corrosivos fundamentales en este tipo de sistemas.

- ✓ El mejor rendimiento en las redes neuronales se encontró en la estructura con 8 variables de entrada (Presión, Temperatura, flujo de crudo, flujo de Gas, Flujo de Agua, CO₂, H₂S y pH). Resultado que valida la clasificación de variables establecida por la metodología fuzzy.

- ✓ La herramienta con relación a otros métodos utilizados para determinar velocidades de corrosión (Waard & Lotz 93, Waard & Lotz 95 y Norsok), proporcionó mejores resultados (Anexo B).

- ✓ La recopilación de la información presentó faltas debido a la no existencia en algunos casos de registros necesarios en puntos de monitoreo experimentales.

6 RECOMENDACIONES

- ✓ Mejorar el proceso de recopilación de datos para determinación de velocidades de corrosión.

- ✓ Continuar el desarrollo del proyecto introduciendo métodos modernos y apropiados para la búsqueda de la estructura de la red neuronal tales como los Algoritmos Genéticos y metodología fuzzy.

- ✓ Desarrollar la aplicación de estas metodologías para otros problemas de interés en la actividad de la CIC.

7 BIBLIOGRAFÍA

- [1] Corrosion Cost and Preventative Strategies in the United States, September 2001, Report FHWA-RD-01-156.
- [2] M.B.Kermany and D. Harrop, BP International, SPE 29784, 1995
- [3] Fuentes Carmelo, Peña D. (2007). "Predicción de la Corrosión por CO₂ + H₂S en Tuberías de Acero al Carbono", *Scientia et Technica* Año XIII, No 36, p. 881.
- [4] Wei You, Yaxiu Liu. IEEE Trans. "Predicting the Corrosion Rates of Steels in Sea Water Using Artificial Neural Network,"(2008): p. 101.
- [5]G. Kumar, R.G. Buchheit (2008). "Use of artificial Neural Network models to predict coated component life from short-term electrochemical impedance spectroscopy measurements," *NACE*, Vol. 64, No. 3, P. 241-254.
- [6] Caicedo, Eduardo y Jesús López. 2009. *Una aproximación practica a las Redes Neuronales Artificiales*. Cali: Editorial Universidad del Valle.
- [7]Molina, Alfredo y Bonifacio del Brio.2007. *Redes Neuronales y sistemas Borrosos*. México: Alfaomega Grupo Editor S.A.
- [8] Srdjan, Nesic et al.2001.Probabilistic modelling of CO₂ Corrosion laboratory data using neural networks. *CorrosionScience*. 43:1373
- [9] Cortés, María y Pablo Ortiz. 2004. Corrosión. *Hipótesis Apuntes Científicos Uniandinos*. No 4:14.
- [10] CIC Acerca de la CIC: Historia (2011). Disponible en: <http://corrosion.uis.edu.co>

- [11] Y. Lin, G. Cunningham and S. Coggeshall. 1996. Input Variable Identification – Fuzzy Curves and fuzzy Interfaces. *Fuzzy Sets and Systems*.82:65.
- [12] Y. Lin and G. Cunningham. 1994. A Fuzzy Approach to Input Variable Identification. *IEEE*. Page 2031.
- [13] Y. Lin, G. Cunningham, S. Coggeshall and R. Jones. September 1998. Nonlinear System Input Structure Identification: Two Stage Fuzzy Curves and Surfaces. *IEEE*. Vol. 28, No 5, Page 638.
- [14] “Cupones de Corrosión y Análisis de Perdida de Peso”. Disponible en <http://www.alspi.com/spanish.htm>
- [15] “Monitoreo por Resistencia Eléctrica”. Disponible en <http://www.alspi.com/spanish.htm>
- [16] G. Cybenko. 1989. "Approximation by Superpositions of Sigmoidal Function", *Mathematics of Control, Signals and Systems*, p. 203-314.
- [17] M. Beale, M.Haga, H. Demuth. *Neural Network Toolbox™ User's Guide*, MathWorks Inc.
- [18] M. Riedmiller, H. Braun. 1993. “A Direct Adaptive Method For Faster Backpropagation Learning: Ther RPROP Algorithm”. *IEEE*. Page 586.
- [19] C. Igel, M. Hüsken. 2000. “Improving the Rprop Learning Algorithm”. In *Proceedings of the Second ICSC International Symposium on Neural Computation (NC' 2000)*. Page 115.
- [20] W. Schiffmann, MJoots, R. Werner. 1993. “Comparison of Optimized Backpropagation Algorithms”. In Verleysen, editor, *Proceedings of the European Symposium on Artificial Neural Networks, ESANN*. Page 97.

8 ANEXO A: MANUAL DE USUARIO

8.1 Introducción

Neural Corrosión Tool es una herramienta software que utiliza la técnica de redes neuronales artificiales para estimar velocidades de corrosión en sistemas específicos que son atacados por este fenómeno, y que cuentan con determinados factores ligados al proceso que denominaremos variables de entrada.

Para determinar velocidades de corrosión primero se debe dar a conocer información de las variables de entrada y la velocidad de corrosión obtenida en cada caso, de esta manera la herramienta aprenderá a reconocer las variables, generando una estructura matemática que permitirá estimar velocidades futuras del sistema enseñado.

8.2 Selección actividad

Al ejecutar la herramienta se presenta la ventana principal (figura 33) con tres opciones para seleccionar. Al posar el indicador del mouse sobre alguna de estas se presenta un mensaje indicando su funcionalidad (figura 34).

- Pre-Procesamiento y Generación de la Red: inicia realizando un pre-procesamiento de la información cargada para generar la red, determinando que variables de entrada son más importantes independientes entre si y no aleatorias para estimar velocidades de corrosión.

Posterior a este análisis se entrena una red neuronal con las variables de entrada cargadas en la información, alternado el numero de neuronas en la capa oculta y descartando las variables poco importantes. Buscando la estructura neuronal con el mejor rendimiento.



Figura 33: Ventana Principal de la Herramienta

- Generación de la Red: no realiza pre-procesamiento de la información. Entrena una red neuronal con las variables de entrada cargadas en la información, variando el número de neuronas en la capa oculta, buscando la estructura neuronal con el mejor rendimiento.
- Determinación de la Red: Luego de tener una red neuronal entrenada y almacenada para un sistema específico, se usa esta opción para estimar velocidades de corrosión para variables con información del mismo sistema. NOTA: la información cargada en archivos de tipo Excel, debe contar con las variables de entrada que entreno la red. No se debe incluir la variable de salida.

8.3 Opción Pre-Procesamiento y Generación de la Red

Al ejecutar esta opción del formulario principal se muestra una ventana, figura 35.a para seleccionar la información designadas para la estructuración y entrenamiento de la red neuronal. Esta información se presenta en un documento Excel y debe contener las variables de entrada en la parte superior y los casos respectivos, sin espacio entre ellos como muestra la figura 35.b.

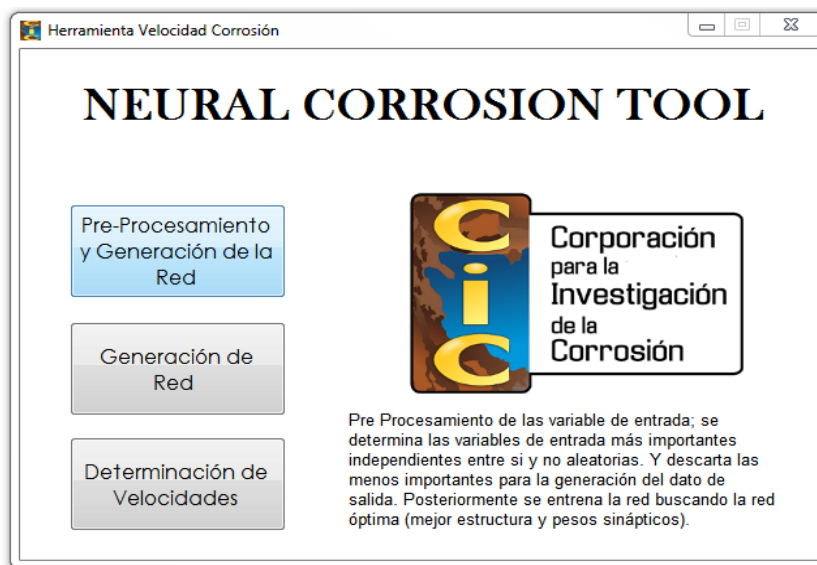
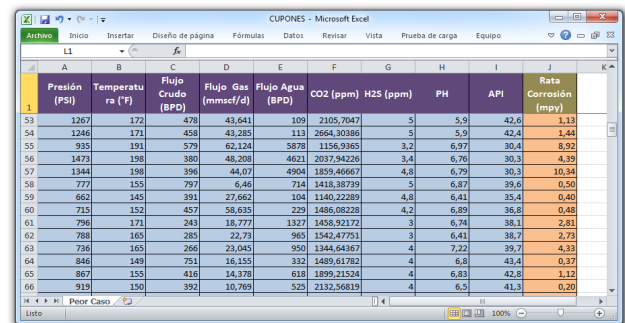


Figura 34: Ventana Principal Indicando Funcionalidad de la Opción



(a)



	A	B	C	D	E	F	G	H	I	J	K
	Presión (PSI)	Temperatura (°F)	Flujo Crudo (BPD)	Flujo Gas (mmscf/d)	Flujo Agua (BPD)	CO2 (ppm)	H2S (ppm)	PH	API	Rata Corrosión (mpy)	
53	1267	172	478	43,641	109	2105,7047	5	5,9	42,6	1,13	
54	1246	172	458	43,285	113	2664,30386	5	5,9	42,4	1,04	
55	935	191	579	62,124	5878	1156,9365	3,2	6,97	30,4	8,92	
56	1473	198	380	48,208	4621	2037,94226	3,4	6,76	30,3	4,39	
57	1344	198	396	44,07	4904	1859,46667	4,8	6,79	30,3	10,34	
58	777	155	797	6,46	714	1418,38739	5	6,87	39,6	0,50	
59	662	145	391	27,662	104	1140,22389	4,8	6,41	35,4	0,40	
60	715	152	457	58,635	229	1486,08228	4,2	6,89	36,8	0,48	
61	796	171	243	18,777	1327	1458,92172	3	6,74	38,1	2,81	
62	788	165	285	22,73	965	1542,47751	3	6,41	38,7	2,73	
63	736	165	266	23,045	950	1344,94867	4	7,22	39,7	4,33	
64	846	190	751	16,153	312	1489,61782	4	6,8	43,4	0,37	
65	867	155	416	14,378	618	1899,21524	4	6,83	42,8	1,12	
66	919	150	392	10,769	525	2132,56819	4	6,5	41,3	0,20	

(b)

Figura 35: (a) Selección Información, (b) Información en documento Excel

Para validar la información cargada, se muestra en pantalla un mensaje (figura 36) indicando la información detectada. Si la herramienta encuentra inconsistencias en la información muestra un mensaje de error y envía el usuario a la ventana principal.

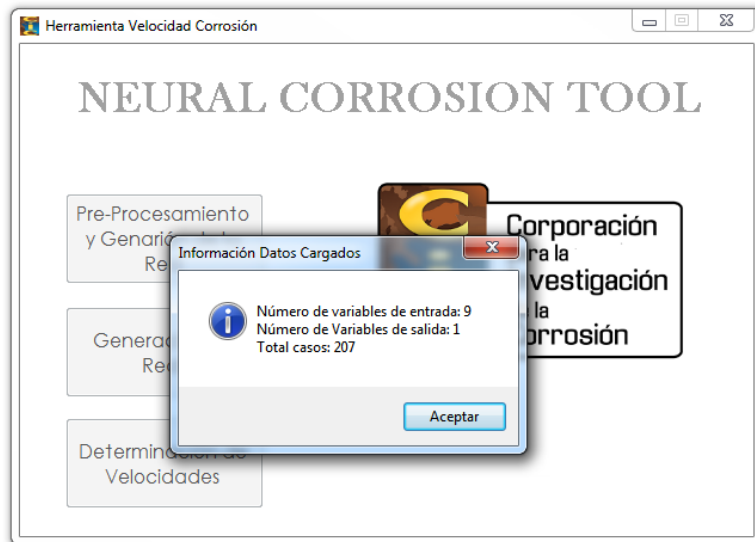
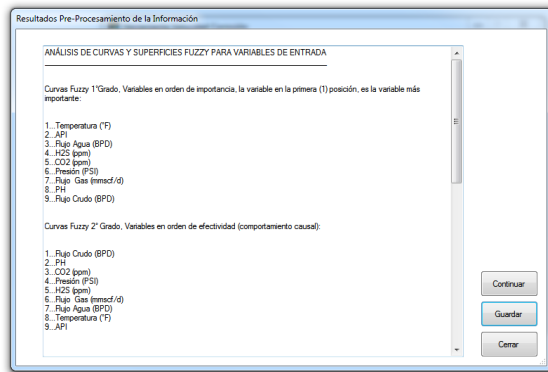


Figura 36: Validar Información Cargada

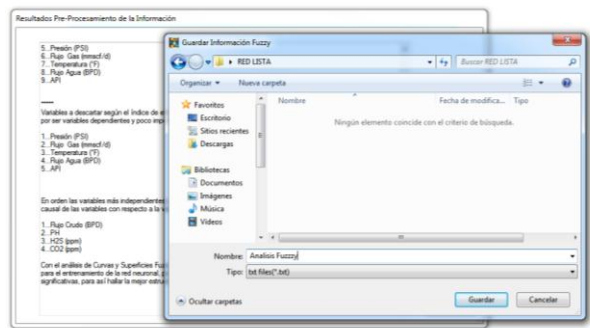
Si la información cargada es adecuada se procede al pre-procesamiento de la información aplicando metodología fuzzyⁱ. Los resultados de dicho pre-procesamiento se muestran en pantalla (figura 37.a), en una ventana que contiene tres opciones:

- Continuar: generar la red neuronal.
- Guardar: guardar la información resultante del pre-procesamiento en un ubicación especificada por el usuario, figura 37.b.
- Cerrar: volver a la ventana principal, figura 34.

i



(a)

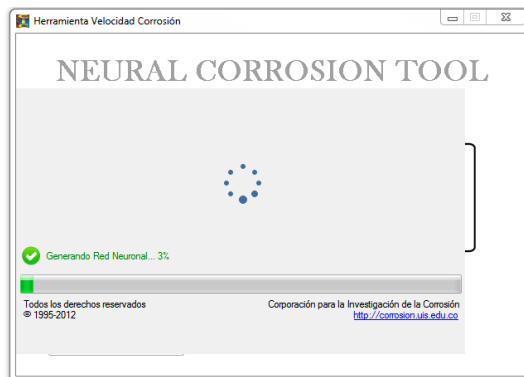


(b)

Figura 37: (a) Resultado Metodología Fuzzy, (b) Almacenamiento Resultado Metodología Fuzzy

La opción continuar da paso a la estructuración y generación de la red neuronal. Mientras se ejecuta este procedimiento que puede tardar varios minutos se presenta el avance del proceso en una ventana que contiene una barra de carga, figura 38.a.

Finalizando se muestra un mensaje indicando el éxito en la generación de la red neuronal (figura 38.b).



(a)



(b)

Figura 38: (a) Progreso Red, (b) Mensaje indicando el final del entrenamiento

El entrenamiento concluye mostrando una ventana (figura 39) que contiene:

- Una grafica comparativa entre la velocidad de corrosión simulada por la red neuronal y la experimental; las velocidad simulada se hace con datos no presentados a la red durante el entrenamiento (8% de la información cargada).
- Una tabla con la velocidad simulada y la experimental.
- Un mensaje con resultados de la red neuronal y el rendimiento obtenido.
- Botón cerrar para volver al formulario principal
- Botón guardar que almacena la red generada y archivos necesarios para su uso en una carpeta designada por el usuario (figura 40).

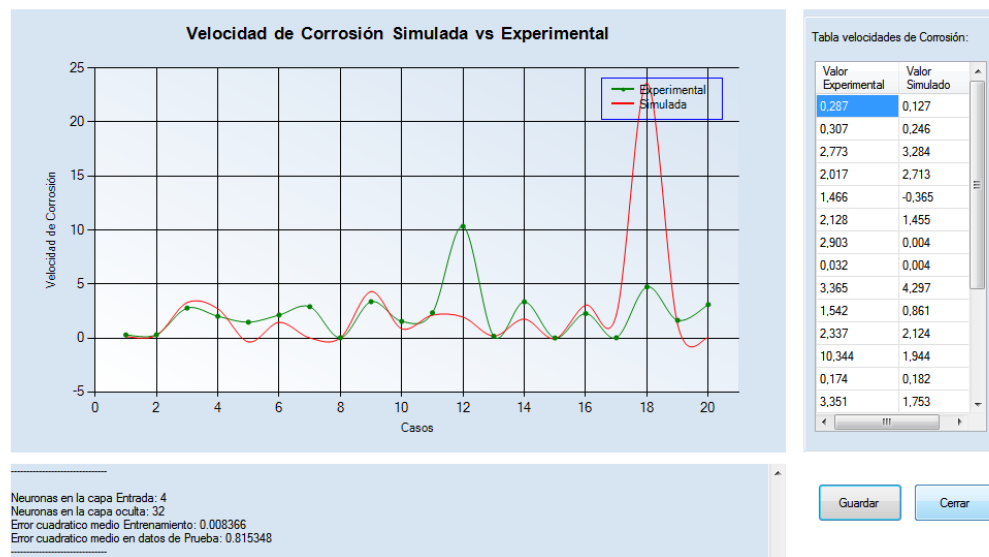
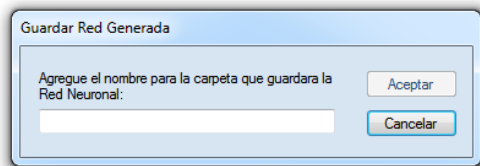
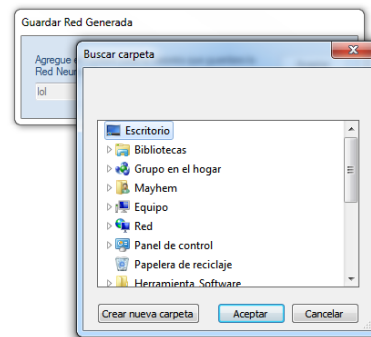


Figura 39: Resultados Entrenamiento Red Neuronal



(a)



(b)

Figura 40: (a) Ingresar nombre para la carpeta que albergará la red entrenada, (b) Ubicación de la carpeta generada

8.4 Opción Generar Red

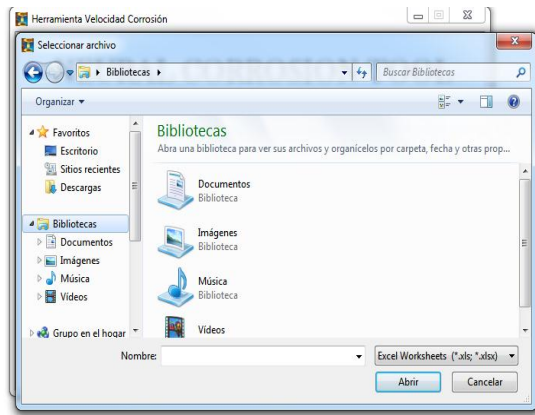
Entrena una red neuronal alterando el número de neuronas en la capa oculta. El procedimiento es el mismo que el descrito antes (Opción Pre-Procesamiento y Generación de la Red) excepto por el pre-procesamiento de la información que no se incluye.

8.5 Opción Calcular Velocidades

Se generan velocidades de corrosión a partir de una red ya creada.

Para este proceso es necesario cargar la información de las variables de entrada para las cuales se desea determinar velocidades de corrosión, figura 41.a; el documento debe contener únicamente las variables de entrada.

Para validar la información cargada se muestra un mensaje indicando los datos encontrados, figura 41.b. A demás se debe ubicar la carpeta con la información almacenada de la red generada (figura 42).



(a)



(b)

Figura 41: Carga de archivo Excel con variables de entrada, (b) validación de la información cargada

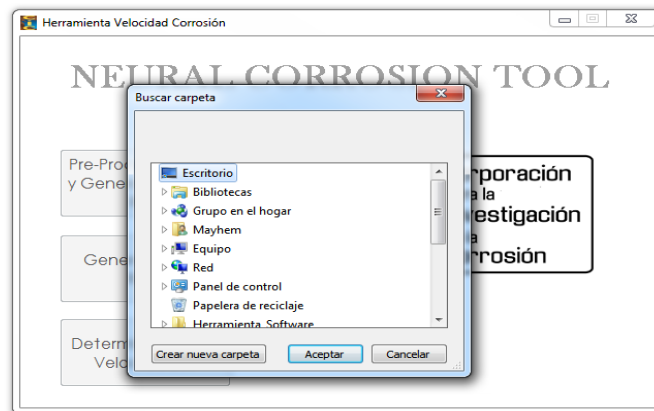


Figura 42: Ubicación de la carpeta que contiene la Red Entrenada

El proceso finaliza mostrando una ventana (figura 43) que contiene:

- La grafica de la velocidad de corrosión simulada por la red neuronal.
- Una tabla con la velocidad simulada.
- Mensaje con la estructura de la red neuronal implementada.
- Botón cerrar para volver al formulario principal, figura 33.

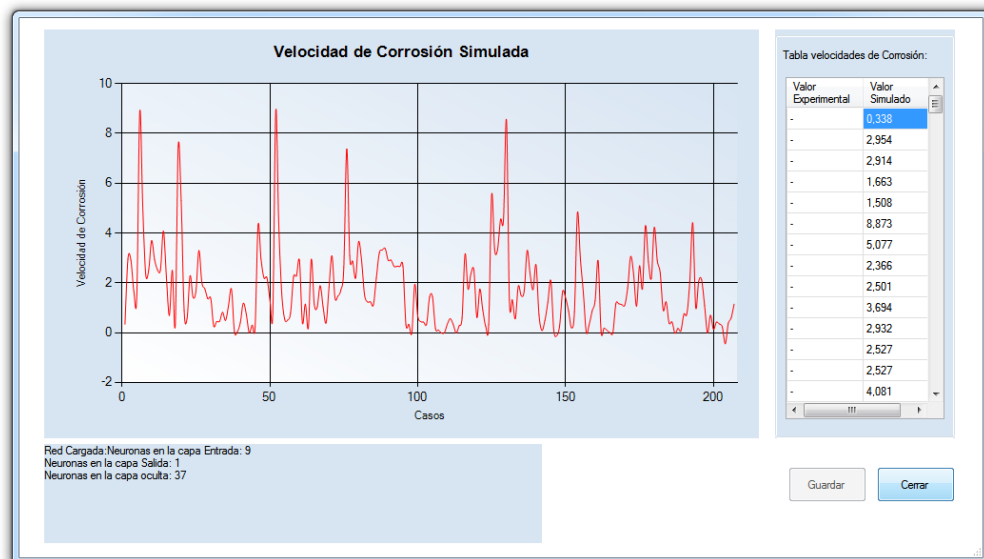


Figura 43: Resultado del cálculo de Velocidad de Corrosión

9 ANEXO B: POSTER

Modelo basado en Redes Neuronales Artificiales, para la determinación de velocidades de corrosión por CO₂ y H₂S en campo de producción petrolera.

C. Ariza, I. Rodríguez, E. Méndez, O. Aya

1. Resumen

Se propone el desarrollo e implementación de redes neuronales artificiales para el cálculo de velocidades de corrosión por CO₂ y H₂S a partir de los datos medidos en campos petroleros.

2. Introducción

Para estudiar el proceso de corrosión, se han propuesto diversidad de métodos y tecnologías, los cuales generalmente exigen la proposición de mecanismos fisicoquímicos para el proceso. La inteligencia artificial ha permitido desarrollar procedimientos que no exigen proposición de mecanismos. Entre éstos se destaca la técnica de redes neuronales artificiales (RNA), muy apropiada para solución de problemas complejos. Aquí presentamos los resultados preliminares del un procedimiento RNA para el cálculo de velocidades de corrosión y para la predicción de confiabilidad de los sistemas estudiados.

3. Metodología



Figura 1. Desarrollo e implementación del modelo basado en redes neuronales para el cálculo de velocidades de corrosión.

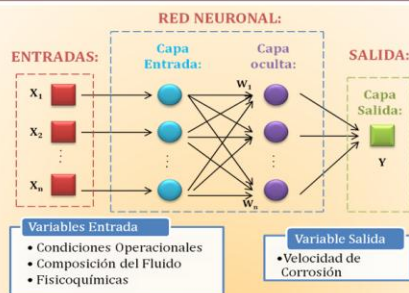


Figura 2. Esquema de la Red Neuronal

4. Resultados

Total Casos	Casos Entrenamiento	Casos Validación	Casos Prueba
175	123	26	26

Tabla 1. Especificación del número de casos para el entrenamiento y validación de la red neuronal artificial.

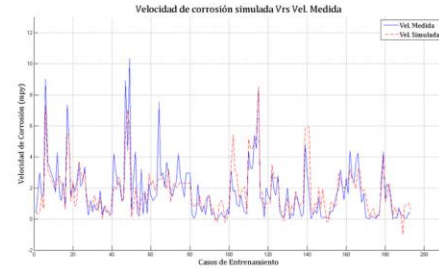


Figura 3. Resultados de la simulación de entrenamiento de la red neuronal artificial.

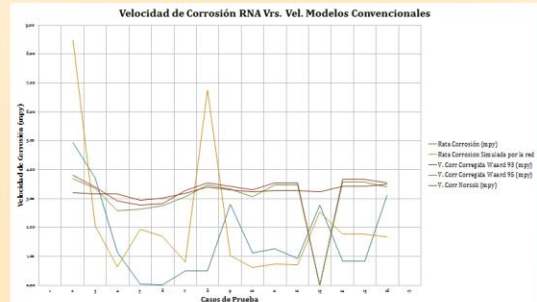


Figura 4. Resultados de la simulación de prueba con la red neuronal artificial y los modelos convencionales.

5. Conclusiones

Las redes neuronales artificiales operan con gran cantidad de datos experimentales.

Bajo éstas condiciones esperamos simulaciones más cercanas al comportamiento real de los sistemas estudiados y así mejorar el control sobre los procesos de corrosión y toma de decisiones.

6. Referencias

- Kermany, MB. and Harrop D. "The impact of Corrosion on Oil and Gas Industry" presented at Society of Petroleum Engineers, Middle East Oil Show, paper SPE 29784, Bahrain March 1995.
- You, W. and Liu, Y. "Predicting the Corrosion Rates of Steels in Sea Water Using Artificial Neural Network", IEEE Trans(2008) 101.
- Kumar, G. and Buchheit R.G. "Use of artificial Neural Network models to predict coated component life from short-term electrochemical impedance spectroscopy measurements" NACE(2008) Vol 64241

Christian Ariza
christianarizaq@gmail.com

Ing. Beatriz Pedraza
bpedraza@corrosion.uis.edu.co

Dr. Orlando Aya
oaya@uis.edu.co