



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE ESTUDIOS SUPERIORES CUAUTITLÁN

**Cálculo no-iterativo del factor de fricción en tuberías para fluidos
Herschel-Bulkley, utilizando el modelo obtenido por redes
neuronales artificiales.**

TESIS

QUE PARA OBTENER EL TÍTULO DE:

INGENIERA EN ALIMENTOS

PRESENTA:

MARÍA GUADALUPE ESQUIVEL BAZÁN

ASESOR: I.A. FRANCISCO JAVIER LÓPEZ MARTÍNEZ.

CUAUTITLÁN IZCALLI, ESTADO DE MÉXICO, 2015.



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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



**FACULTAD DE ESTUDIOS SUPERIORES CUAUTITLÁN
UNIDAD DE ADMINISTRACIÓN ESCOLAR
DEPARTAMENTO DE EXÁMENES PROFESIONALES**

UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MÉXICO

U. N. A. M.
ASUNTO: VOTO APROBATORIO

**M. en C. JORGE ALFREDO CUÉLLAR ORDAZ
DIRECTOR DE LA FES CUAUTITLAN
PRESENTE**

**ATN: M. EN A. ISMAEL HERNÁNDEZ MAURICIO
Jefe del Departamento de Exámenes Profesionales
de la FES Cuautitlán.**



Con base en el Reglamento General de Exámenes, y la Dirección de la Facultad, nos permitimos comunicar a usted que revisamos el: Trabajo de Tesis

Cálculo no-iterativo del factor de fricción en tuberías para fluidos Herschel-Bulkley, utilizando el modelo obtenido por redes neuronales artificiales

Que presenta la pasante: María Guadalupe Esquivel Bazán
Con número de cuenta: 095266143 para obtener el Título de la carrera: Ingeniería en Alimentos

Considerando que dicho trabajo reúne los requisitos necesarios para ser discutido en el EXAMEN PROFESIONAL correspondiente, otorgamos nuestro VOTO APROBATORIO.

ATENTAMENTE
"POR MI RAZA HABLARÁ EL ESPÍRITU"
Cuautitlán Izcalli, Méx. a 27 de Febrero de 2015.

PROFESORES QUE INTEGRAN EL JURADO

	NOMBRE	FIRMA
PRESIDENTE	I.A. Francisco Javier López Martínez	
VOCAL	Dra. Abigail Martínez Estrada	
SECRETARIO	Dra. Virginia Agustina Delgado Reyes	
1er. SUPLENTE	I.A. Verónica Romero Arreola	
2do. SUPLENTE	I.A. Fátima Abigail Galván Ballesteros	

NOTA: los sinodales suplentes están obligados a presentarse el día y hora del Examen Profesional (art. 127).

AGRADECIMIENTOS

A Dios, por el regalo de la vida y por todos los dones que me ha otorgado para vivirla.

A mis Padres, que con su invaluable apoyo e infinito amor, han hecho posible que alcance cada uno de mis sueños.

A mi *alma mater*, la *Universidad Nacional Autónoma de México*, por brindarme la grandiosa oportunidad de formarme y desarrollarme como mejor ser humano, dentro de la Máxima Casa de Estudios, y con ello otorgarme, el privilegio y la responsabilidad de ser mejor mexicana.

A mi asesor, I.A. Francisco Javier López Martínez, por quien siento gran respeto, gracias por su apoyo en la realización de este proyecto, por siempre brindarme valiosas enseñanzas, por su dedicación como Profesor, su enorme paciencia y tiempo.

A mis sinodales, por su tiempo, sus valiosas aportaciones y observaciones para mejorar esta Tesis.

A todos los profesores, que durante toda mi vida académica, me inspiraron, instruyeron y guiaron con sus enseñanzas.

DEDICATORIAS

A mis Padres: Rosalino y Teresa; quienes son mi inspiración, mi guía y mi fuerza, los amo profundamente y me siento muy orgullosa de ser su hija.

A mis hermanos: Elsa, Anaid y Homero; quienes son mis guías, mis amigos, mis compañeros y mis más duros críticos, así como mis más leales soportes en los días más oscuros, los respeto y amo profundamente.

A mis amados sobrinos: Daniela, Paula, Hugo, Valeria, Diego, Fernanda y Camila; quienes son mi fuente de inspiración y motivación, para ser una mejor persona cada día y dejar en mi camino lo mejor para ustedes.

A toda mi querida familia: quienes son todos aquellos que siempre estamos juntos, en los días alegres, tristes ó amargos, gracias por su comprensión, cariño y apoyo.

In Memóriam a mis abuelos, por sus enseñanzas, ejemplo y alegrías, que siempre me acompañarán y que jamás olvidaré.

A mis queridas amigas: Sandy L., Maribel T., Berenice M., Ingrid R., Carmen P., Paola S. y Karen F., quienes siempre han estado ahí para apoyarme (desde antes del inicio de este proyecto, ó a regresar cuando se abandonó y hasta concluir el proceso de titulación), gracias por compartirme su conocimiento, alentarme e inspirarme, es un placer y una gran fortuna el haber coincidido en este camino y es un gran honor ser su amiga.

A todos mis compañeros y amigos de la carrera de Ingeniería en Alimentos (en especial a la generación 22), por las enseñanzas compartidas, las experiencias vividas y las alegrías brindadas. Ha sido un privilegio compartir ese tiempo con ustedes.

Finalmente a todas las personas que durante mi vida personal, académica y laboral han depositado en mí su confianza, apoyado y guiado con sus consejos, hasta lograr mis metas; lo que me exige corresponderles, dando mi mejor esfuerzo en el ejercicio responsable de mi profesión y en colaborar afanosamente en el desarrollo justo de nuestra sociedad, que es el cuerpo vivo de nuestra gran nación, México.

“Por mi raza hablará el espíritu”

Hay una fuerza motriz más poderosa que el vapor,
la electricidad y la energía atómica: la voluntad.

ALBERT EINSTEIN

No estudio para saber más, sino para ignorar menos.

SOR JUANA INÉS DE LA CRUZ

Los hombres geniales empiezan grandes obras,
los hombres trabajadores las terminan.

LEONARDO DA VINCI

La mejor vida no es la más larga, sino
la más rica en buenas acciones.

MARIE CURIE

El éxito no es el final, el fracaso no es la ruina,
lo que cuenta es el valor para continuar.

WINSTON CHURCHILL

ÍNDICE

	Páginas
ÍNDICE DE FIGURAS.....	IV
ÍNDICE DE TABLAS.....	VI
NOMENCLATURA.....	VII
RESUMEN.....	IX
INTRODUCCIÓN.....	X
OBJETIVOS.....	XII
CAPÍTULO 1. Antecedentes	
1.1 Generalidades reológicas de los fluidos alimenticios.....	1
1.1.1 Esfuerzo y deformación.....	3
1.1.2 Fluidos.....	6
1.1.3 Fluidos newtonianos.....	7
1.1.4 Fluidos no-newtonianos: características y modelos.....	9
1.14.1 Comportamiento de fluidos dependientes del tiempo.....	11
1.1.4.2 Comportamiento de fluidos independientes del tiempo.....	13
1.1.5 Regímenes de flujo: laminar y turbulento.....	18
1.2 Balance mecánico de energía para un fluido en una tubería en estado estacionario.....	22
1.2.1 El factor de fricción en el diseño de sistemas de bombeo en los procesos alimenticios.....	28
1.2.2 Ecuaciones para el cálculo de factor de fricción en régimen laminar y turbulento para fluidos Herschel-Bulkley.....	32

1.2.2.1	Método iterativo de cálculo para fluidos Herschel-Bulkley en régimen laminar.....	35
1.3	Las redes neuronales artificiales: definición, fundamentos y generalidades.....	39
1.3.1	Definición.....	39
1.3.2	Antecedentes históricos y fundamentos.....	40
1.3.2.1	Antecedentes históricos.....	40
1.3.2.2	Fundamentos.....	43
1.3.3	Estructura del sistema neuronal artificial.....	45
1.3.3.1	Micro-estructura, unidades de proceso: la neurona artificial.....	46
1.3.3.1.1	El modelo genérico de la neurona artificial.....	47
1.3.3.2	Meso-estructura, la red neuronal artificial.....	56
1.3.4	Aprendizaje o entrenamiento de la RNA.....	61
1.3.4.1	Mecanismos de aprendizaje: supervisado y no supervisado.....	61
1.3.4.2	Algoritmos de aprendizaje.....	63
1.3.4.3	Selección del Algoritmo de aprendizaje.....	67
1.3.5	Realización de redes neuronales artificiales.....	67
1.3.6	Generalidades de las RNA y aplicaciones en la industria alimentaria.....	69
1.3.6.1	Clasificación de las RNA.....	69
1.3.6.2	Ventajas y desventajas de usar las RNA.....	69
1.3.6.3	Aplicaciones de las RNA en la industria alimentaria.....	71
CAPÍTULO 2. Diseño del modelo matemático para el cálculo no-iterativo del factor de fricción en tuberías para fluidos Herschel-Bulkley en régimen laminar y turbulento.		
2.1	Desarrollo del modelo de la RNA.....	74
2.1.1	Metodología.....	74

2.1.2	Descripción del Cuadro metodológico.....	76
CAPÍTULO 3. Resultados		
3.1	Análisis de resultados.....	87
3.2	Validación del modelo obtenido de la RNA.....	98
3.3	Determinación de las ecuaciones explícitas para el cálculo del factor de fricción de fluidos Herschel-Bulkley en régimen laminar y turbulento.....	102
3.3.1	Ejemplo de cálculo del factor de fricción para un fluido Herschel-Bulkley, con las ecuaciones obtenidas a través de la RNA.....	105
CONCLUSIONES.....		108
RECOMENDACIONES.....		110
REFERENCIAS.....		111
GLOSARIO.....		115
ANEXO 1.	Instrucciones de uso del simulador Neural Network Toolbox del entorno Matlab® R2008 ^a	118
ANEXO 2.	El Método de Bisección.....	125
ANEXO 3.	Algoritmo de aprendizaje Levenberg-Marquardt.....	127

ÍNDICE DE FIGURAS

	Páginas
Figura 1.1 Componentes cartesianas de las tensiones en cada una de las caras de un cubo.....	3
Figura 1.2 Cizalladura de un cubo.....	5
Figura 1.3 Gráfica de esfuerzo de corte σ , en función de la velocidad de corte $\dot{\gamma}$ de un fluido newtoniano.....	8
Figura 1.4 Clasificación de los fluidos de comportamiento viscoso, en base a su comportamiento reológico.....	10
Figura 1.5 Reogramas de fluidos que exhiben comportamiento (a)Tixotrópico y (b)Antitixotrópico.....	12
Figura 1.6 Curvas para fluidos típicamente independientes del tiempo.....	13
Figura 1.7 Gráfica de esfuerzo de corte σ en función de la velocidad de corte $\dot{\gamma}$ para modelos que describen un comportamiento (a) Espesante (dilatante) y (c) Fluidificante (pseudoplástico).....	16
Figura 1.8 Flujo laminar en un ducto circular.....	18
Figura 1.9 Velocidad de distribución de un fluido viscoso entre dos placas paralelas.....	19
Figura 1.10 Perfil de velocidades definido de un fluido en una tubería.....	20
Figura 1.11 Una red neuronal artificial como un grafo dirigido y sus propiedades.....	40
Figura 1.12 Diagrama de la estructura de una RNA con tres capas (entrada-oculta-salida).....	46
Figura 1.13 Modelo genérico de una neurona artificial.....	47
Figura 1.14 Funciones de activación (a) Lineal y (b) Escalón (en este caso se considera que el umbral es cero, en el caso de que no lo fuera el escalón quedaría desplazado).....	52
Figura 1.15 Función de activación lineal a tramos (mixta).....	53
Figura 1.16 Funciones de activación (a) Sigmoidales y (b) Gausiana.....	55

Figura 1.17	Modelo de la función procesadora de una única neurona de tipo estándar, donde la regla de propagación es la suma ponderada y la función de salida es de tipo lineal).....	56
Figura 1.18	Estructura de una red multicapa con conexión unidireccional o propagación hacia delante.....	57
Figura 1.19	Estructura de una red monocapa (en este caso una red de Hopfield).....	58
Figura 1.20	Red neuronal con propagación hacia atrás (recurrente).....	60
Figura 1.21	Aprendizaje de corrección de errores, diagrama de una RNA enfatizando una sola neurona en la capa de salida.....	66
Figura 2.1	Diagrama del cuadro metodológico.....	75
Figura 2.2	Diagrama general de la RNA a nivel meso-estructura (RNA multicapa, con flujo de información de atrás hacia adelante y unidireccional), a utilizar.....	80
Figura 2.3	Modelo genérico del procesamiento de información de una neurona i de la capa oculta, donde hay presencia de un bias (una tendencia, que se registra como una entrada extra), la regla de propagación es la suma ponderada; la función de transferencia es la tangente hiperbólica y la función de salida es la lineal.....	81
Figura 3.1	Valores del logaritmo del factor de fricción calculado por ecuaciones “Log (f_{fc})” vs valores del logaritmo del factor de fricción obtenido por la RNA “Log (f_{fp})” del modelo de 20 neuronas.....	98
Figura 3.2	Valores del logaritmo del factor de fricción obtenido por ecuaciones “Log (f_{fc})” vs valores del logaritmo del factor de fricción obtenido por la RNA “Log (f_{fp})”, validación para el modelo de RNA de 20 neuronas.....	100
Figura 3.3	Valores del logaritmo del factor de fricción obtenido por ecuaciones “Log (f_{fc})” vs valores del logaritmo del factor de fricción obtenido por la RNA “Log (f_{fp})”, validación para el modelo RNA de 4 neuronas.....	100
Figura 3.4	Diagrama esquemático de la estructura de Red neuronal artificial de 4 neuronas en la capa oculta.....	101

ÍNDICE DE TABLAS

	Páginas
Tabla 1.1	Factor de corrección para la energía cinética para flujo laminar en tuberías..... 25
Tabla 2.1	Parámetros estadísticos a utilizar para la evaluación de la RNA..... 84
Tabla 3.1	Resultados prueba 1..... 94
Tabla 3.2	Resultados prueba 2..... 95
Tabla 3.3	Resultados prueba 3..... 96
Tabla 3.4	Resultados de validación de los mejores modelos..... 99
Tabla 3.5	Algunos resultados del factor de fricción (f_f) para fluidos Herschel-Bulkley..... 107

NOMENCLATURA

Símbolo	Significado
A	área [m ²]
A_g	constante de la función Gaussiana
a_j	estado de activación de la neurona j
D	diámetro de tubería [m]
d_k	señal de respuesta deseada de la neurona k
E	estados de activación
E_f	perdida de energía por unidad de masa debida a la fricción [J/Kg]
e_k	error de señal de la neurona k
F	fuerza [newton]
F_{sj}	función de salida de la neurona j
f_f	factor de fricción de Fanning [adimensional]
f_{fp}	factor de fricción predicho por la red neuronal artificial
f_{act}	función de activación
G	módulo de rigidez, tensión ó cizalladura [Pa]
g	aceleración debida a la gravedad [m/s ²]
H	función unitaria Heaviside
H_a	altura de un cubo [m]
He	numero de Hedstrom [adimensional]
K	índice de consistencia [Pa.s ⁿ]
K_f	coeficiente de resistencia por fricción [adimensional]
L	longitud de tubería [m]
n	índice de comportamiento al flujo [adimensional]
n_t	tiempo discreto
net_i	entrada neta de la neurona i
P	presión [Pa]
Q	velocidad de flujo volumétrico en una tubería [m ³ /s]
r	radio de la tubería [m]
Re	número de Reynolds [adimensional]

Re_c	número crítico de Reynolds, [adimensional] $\frac{Dv\rho}{\eta}$
Re_G	número de Reynolds generalizado [adimensional]
Re_{He}	número de Reynolds como una función implícita del número de Hedstrom (Hanks, 1978).
v	velocidad volumétrica, ó velocidad promedio de flujo [m/s]
t	tiempo en un instante dado [s]
W	trabajo ó energía necesaria para realizar un trabajo por unidad de masa [J/Kg]
\mathbf{W}	Matriz de pesos sinápticos de entrada a la neurona artificial
w_{ij}	peso de la sinapsis de la neurona i a la j
\mathbf{X}	vector de elementos de entrada a una neurona
x_j	señal de entrada a la neurona j
y_j	señal de salida de la neurona j
z	Altura de acuerdo a un punto de referencia [m]
z_n	operador de retraso (unidad de almacenamiento)
Letras griegas	
α	Factor de corrección a la energía cinética [adimensional]
β	ángulo de deformación
$\dot{\gamma}$	velocidad de corte [s ⁻¹]
ε	Índice de desempeño
η	coeficiente de viscosidad newtoniana [Pa.s]
η_{ap}	coeficiente de viscosidad aparente [Pa.s]
η_A	constante de la rapidez de aprendizaje ó parámetro de rapidez de aprendizaje
θ_i	umbral de activación de la neurona i
ζ_0	relación de esfuerzo inicial y esfuerzo en la pared [adimensional]
ρ	densidad [Kg/m ³]
ΔP	diferencial de presión ó caídas de presión debidas a la fricción [Pa]
$\sigma_{n i}$	regla de propagación de la neurona i
σ	esfuerzo de corte [Pa]
σ_0	esfuerzo de corte inicial [Pa]
σ_w	esfuerzo de cizalla en la pared interna de una tubería, valor de σ en r_w
ψ	función de flujo laminar

RESUMEN

En las industrias de procesamiento de alimentos, el transporte de fluidos en tuberías es una actividad esencial, de ahí la importancia de diseñar de sistemas hidráulicos eficientes, los cuales dependen en parte de la determinación del factor de fricción. Por otro lado muchos alimentos fluidos como la cajeta, el chocolate fluido, jarabes, mermeladas, etcétera; presentan un comportamiento reológico no-newtoniano representado por el modelo Herschel-Bulkley. El cálculo del factor de fricción para un fluido Herschel-Bulkley implica un proceso iterativo que en muchas ocasiones se considera complicado, porque no conocen herramientas computacionales que les facilite ese cálculo. En este trabajo se utilizó una herramienta computacional denominada Red Neuronal Artificial, con la cual se obtiene un modelo matemático ordinario lineal, en donde solo se requieren tres valores para calcular el factor de fricción sin la necesidad de realizar un proceso iterativo.

Este tipo de sistemas de procesamiento fueron inspirados en la estructura de los cerebros biológicos y tienen como característica primordial la capacidad de generalizar modelos a través de ejemplos provistos.

Utilizando las redes neuronales artificiales se obtuvo un modelo que es capaz de calcular el factor de fricción (f_f) tanto en régimen laminar como turbulento, teniendo como entrada tres parámetros: el número de Reynolds (Re), el índice de comportamiento al flujo (n), el número de Hedstrom (He). El modelo obtenido permite calcular el f_f dentro de los siguientes intervalos $Re=100$ a 1×10^7 , $n=0.1$ a 1 y $He \geq 10$ a $He < 1 \times 10^6$.

INTRODUCCIÓN

En la industria de alimentos es muy común que ingredientes y productos se encuentren en presentación fluida, ya sea en alguna etapa o en todo el proceso. Esto hace indispensable la implementación de sistemas hidráulicos adecuados para el buen manejo de fluidos. Gran parte de los fluidos que se utilizan en la industria de alimentos se comportan como fluidos newtonianos, y otros fluidos alimenticios muestran un comportamiento no-newtoniano, como lo indican Steffe & Morgan (1986). Los sistemas de bombeo son parte esencial de muchas operaciones de proceso para alimentos, por ende es necesario que los Ingenieros en Alimentos tengan el conocimiento de los principios del diseño de tuberías de distribución para fluidos no-newtonianos. Los alimentos fluidos del tipo no-newtoniano, son de diversas clases y propiedades, como por ejemplo: jarabes, cajeta, mermeladas, salsas, pulpas de frutas, chocolate fluido, etc., el modelo Herschel-Bulkley, representa el comportamiento de algunos de estos fluidos alimenticios y es utilizado al diseñar redes de tuberías para el transporte de estos.

Durante el diseño de tuberías de distribución y selección de bombas, es un requisito necesario calcular la caída de presión en tuberías debida a la fricción, esta última es en términos generales la resistencia de un fluido a ser transportado en una tubería. Para que un fluido se mueva es necesario aplicar energía mecánica al sistema, por lo tanto si las pérdidas de energía mecánica debidas a la fricción (generalmente representadas por caídas de presión) son grandes, representarían un deficiente o nulo transporte del fluido (Singh & Heldman, 2009). Sin embargo un adecuado diseño del sistema hidráulico, incluida la selección de las bombas garantiza el transporte del fluido y minimiza el consumo de energía, además de ser importante para el buen control del proceso, ya que todo esto repercute directamente en la calidad del alimento y en el costo final de éste.

El análisis hidráulico de redes de tuberías como lo menciona Sablani *et al.* (2003), requiere de la implementación de procedimientos de consumo de tiempo largos y tediosos, por lo que es necesario el uso extensivo de computadoras. Esto último con el fin de hacerlos más rápidos y más prácticos, y por otro lado por la necesidad de un rápido y eficiente procesamiento de información.

Además con el desarrollo tecnológico constante como lo indican Hilera & Martínez (1995), actualmente la tecnología de la información se encuentra en una fase donde se busca reproducir artificialmente muchas funciones neuronales, con la principal finalidad del desarrollo de nuevos tipos de ordenadores, con capacidad de realizar procesos con cierta inteligencia. Esto ha dado origen a las redes neuronales artificiales (RNA) que son modelos computacionales inspirados en una simplificación del funcionamiento estructural del cerebro. Sablani & Shayya (2003), indican que estos modelos son capaces aprender desde ejemplos a través de iteraciones sin requerir a priori del conocimiento de la relaciones de los parámetros de procesamiento y que son capaces de adaptarse a cambios en el entorno y también de tratar datos inciertos o con ruido y con relaciones no lineales, entre estos.

Desde hace ya más de una década las RNA como lo señala Islamoglu (2003), han sido usadas para resolver una amplia variedad de problemas complejos ya sean científicos, de ingeniería y negocios. Y su uso en la modelación química y de procesos alimenticios también ha sido amplio, algunas de las aplicaciones realizadas han sido las siguientes:

La predicción de los tiempos de congelamiento para alimentos (Mittal & Zhang, 2000), así como en la obtención de un modelo para predecir la humedad en quesos (Jiménez-Márquez *et al.*, 2003). De igual forma en la obtención de un método de predicción de caídas de presión en tuberías de fluidos alimenticios no-newtonianos (Adhikari & Jindal, 2000), y para realizar cálculos no-iterativos del factor de fricción en tuberías (Sablani & Shayya, 1998), asimismo en el cálculo explícito del factor de fricción en tuberías de fluidos plásticos de Bingham (Sablani *et al.*, 2003) y fluidos de la potencia (Sablani & Shayya, 2003). Además como lo señalan Martín & Sanz (1997), se han aplicado en control y aseguramiento de la calidad en las industrias, por ejemplo en la detección de niveles de contaminantes a partir de datos espectroscópicos y en la evaluación del grado de pureza del jugo de naranja.

En este trabajo se presenta un procedimiento no-iterativo basado en las RNA para estimar el factor de fricción f_f en tuberías para fluidos Herschel-Bulkley en régimen laminar y turbulento, y lograr reducir el tiempo de cálculo por iteración, esto considerando la metodología propuesta por Sablani & Shayya (2003). Además mostrar el panorama y potencialidades de la aplicación de las RNA en la Ingeniería en Alimentos y su viabilidad para implementarse como una herramienta alternativa.

OBJETIVOS

OBJETIVO GENERAL:

Desarrollar un modelo matemático mediante el uso de las redes neuronales artificiales que permita calcular de manera no-iterativa el factor de fricción en tuberías para fluidos Herschel-Bulkley en régimen laminar y turbulento, para disminuir el tiempo dedicado a los cálculos iterativos y generar una nueva herramienta de cálculo.

OBJETIVO PARTICULAR 1:

Seleccionar la estructura de la red neuronal artificial que permita calcular el factor de fricción en el menor tiempo posible en comparación con el empleado por el método iterativo.

OBJETIVO PARTICULAR 2:

Evaluar el desempeño del modelo de la red neuronal artificial para el cálculo del factor de fricción comparando sus resultados con el método iterativo para validar el funcionamiento de este modelo y obtener las ecuaciones lineales asociadas a este.

CAPÍTULO 1. Antecedentes

En los procesos industriales de alimentos, el manejo de fluidos es básico y por lo tanto es indispensable la implementación de sistemas hidráulicos eficientes y adecuados para resolver los problemas de flujo de cada tipo de los fluidos que puedan intervenir ó utilizar durante todo el proceso. Para comprender, diseñar e implantar sistemas de transporte de fluidos adecuados a las necesidades del proceso, es fundamental el conocimiento de dos ciencias: la mecánica de fluidos y la reología. Estas dos ciencias están estrechamente ligadas e incluso ambas pueden utilizarse para interpretar y estudiar el cálculo de sistemas idénticos, sin embargo tienen objetivos diferentes.

El objetivo de la mecánica de fluidos es el de predecir los movimientos de la materia que obedecen a una cierto número de fuerzas características (presión, peso, fricción, etc.), conociendo las ecuaciones del movimiento y la ecuación del estado reológico del fluido. Por otro lado, el objetivo de la reología es determinar la ecuación del estado reológico del fluido al efectuar medidas relativas de fuerzas, desplazamientos y en utilizar sus resultados para la ayuda de las ecuaciones del movimiento.

1.1 Generalidades reológicas de los fluidos alimenticios.

La reología es considerada como la ciencia de la deformación y el flujo de materia, es decir es el estudio de la manera en la cual los materiales responden al aplicarles un esfuerzo¹ ó deformación² (Steffe, 1996). Otra definición de reología, es la que la define como la ciencia que estudia el flujo y las deformaciones de sólidos y fluidos, bajo la influencia de fuerzas mecánicas.

¹ En el texto original en inglés el término es “stress”.

² En el texto original en inglés el término es “strain”.

Todos los materiales tienen *propiedades reológicas*, estas son relevantes en muchos campos de estudio, como la geología, la mecánica de suelos, el procesado de polímeros, dispersión de pinturas, pigmentos, bioingeniería, estructura de materiales, cosmética, etc. En los alimentos entender la reología es crítico en los esfuerzos de optimización de desarrollo de productos, en la metodología de proceso y en la calidad del producto final.

Las propiedades reológicas están basadas sobre las respuestas de los materiales al flujo y deformación cuando son sometidos a un esfuerzo. El estudio de la reología de fluidos alimenticios requiere conocimiento de tensores y de principios básicos del flujo de fluidos tales como la *ecuación de continuidad* (conservación de la masa) y de la *ecuación de movimiento* (conservación de momentum) como lo indica Méndez *et al.*, 2001.

Además en el diseño de todo proceso es necesario conocer, entre otras cosas las características físicas de las corrientes que lo forman. Una de estas características es el comportamiento reológico del fluido que se procesa. Y una propiedad de gran importancia de los fluidos es la viscosidad, esta se utiliza en el cálculo de parámetros en los fenómenos de transporte de cantidad de movimiento y energía, así como es utilizada como parámetro para el control de calidad de ciertos productos. Debido a ello en las ecuaciones del modelo matemático planteado para las diversas operaciones que forman un determinado proceso, intervienen las constantes reológicas de los fluidos que se procesan y que se deben determinar, generalmente mediante experimentación en cada caso particular.

Por lo tanto es de suma importancia la caracterización reológica de los diferentes fluidos y la deducción de ecuaciones que permitan calcular directamente las constantes reológicas en función del alimento considerado y de las variables de operación. Específicamente los parámetros reológicos de los alimentos son necesarios en la industria para (Steffe, 1996):

- a) Realizar cálculos necesarios en la ingeniería de procesos que implican una amplia variedad de equipos en la operación, como el diseño de tuberías, bombas, mezcladores, intercambiadores de calor, homogenizadores, equipos de filtración, tolvas, etc.
- b) Conocer la influencia de un ingrediente (tipo funcional) sobre el desarrollo del producto.
- c) Control de calidad del producto.

- d) Evaluación de la textura del alimento mediante la relación con datos sensoriales.
- e) Pruebas de vida de anaquel, etc.

En la industria se efectúan con frecuencia medidas reológicas sobre las materias primas y los productos en elaboración que son de gran utilidad para el control de los procesos.

1.1.1 Esfuerzo y deformación.

Al aplicar una fuerza a un cuerpo este se deformará, sin embargo la respuesta que presente será distinta según sea la naturaleza de cada material, existen dos ejemplos típicos de deformación que son el alargamiento y la compresión (Muller, 1973).

El *esfuerzo*³ (σ) aplicado a un material se define como la fuerza (F) por unidad de área (A), es decir F/A (newton/m² ó Pa). Cuando se aplica un esfuerzo (σ) a un material (vector de primer orden), este puede ser de tres tipos: tensión, corte ó cizalla y volumétrico dando como resultado un movimiento con cierta intensidad, sentido y dirección que se puede manifestar en tres direcciones x , y ó z que dependen de la superficie en la que fue aplicada dicha fuerza (Steffe, 1996). Si se supone un cuerpo como el representado en la Figura 1.1, se observa que sobre cada cara de dicho cuerpo pueden aplicarse tres esfuerzos, uno para cada dirección.

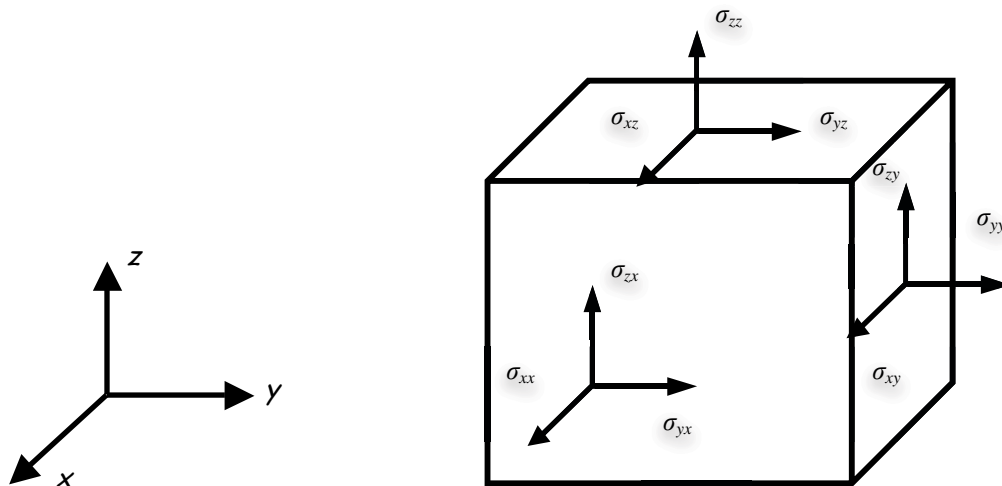


Figura 1.1 Componentes cartesianas de las tensiones en cada una de las caras de un cubo.

³ En algunas fuentes este término se puede encontrar representado con la letra griega (τ).

Por lo tanto, existirán nueve componentes del esfuerzo, cada una de las cuales se designará con dos subíndices, expresado el primero de ellos la sección sobre la que se aplica y el segundo la dirección. Así, σ_{ij} es un esfuerzo aplicado sobre la sección (ó cara) i en la dirección j . Las nueve componentes del esfuerzo forman lo que se denomina el *tensor esfuerzo*:

$$\sigma_{ij} = \begin{bmatrix} \sigma_{xx} & \sigma_{yx} & \sigma_{zx} \\ \sigma_{xy} & \sigma_{yy} & \sigma_{zy} \\ \sigma_{xz} & \sigma_{yz} & \sigma_{zz} \end{bmatrix} \quad (1.1)$$

El tensor esfuerzo es simétrico, lo que supone que $\sigma_{ij} = \sigma_{ji}$. Por ende, de las nueve componentes del tensor esfuerzo seis son independientes. Los esfuerzos pueden ser tanto normales, como tangenciales. Los esfuerzos tangenciales son aquellos en que $i \neq j$, mientras que en los normales $i = j$.

En la mayoría de los casos el movimiento se da en una sola dirección; cuando esta se da en la misma dirección que el esfuerzo se le conoce como *esfuerzo normal* σ_{xx} , σ_{yy} , σ_{zz} (vector de segundo grado ó tensor). Dentro de los normales cabe distinguir los de tracción y compresión. Los esfuerzos normales de tracción son positivos ($\sigma_{ii} > 0$), mientras que los de compresión son negativos ($\sigma_{ii} < 0$). Las deformaciones producidas por los esfuerzos normales son longitudinales y pueden ser alargamientos o acortamientos, dependiendo si el esfuerzo es de tracción o de compresión.

En el caso de los esfuerzos tangenciales, estos producen deformaciones angulares, observemos el cubo de la Figura 1.2 que está siendo deformado por la tensión aplicada tangencialmente (esfuerzo de corte) σ , la deformación viene dada por la tangente del ángulo β , L/H_a . Al módulo, en este tipo de deformación, se le llama módulo de tensión ó cizalladura, G . La ecuación representativa es:

$$\sigma = G \tan (\beta) \quad (1.2)$$

Donde σ es el esfuerzo de corte y β el ángulo de la deformación.

Cuando una fuerza es aplicada a un material sólido y la deformación relativa es directamente proporcional al esfuerzo aplicado, entonces se dice que el material obedece a la ley de Hooke (la

ecuación típica de los sólidos elásticos⁴).

La relación para el esfuerzo de corte y la deformación relativa γ puede establecerse como (Steffe, 1996):

$$\sigma = G \gamma \quad (1.3)$$

Donde G es el modulo de tensión, γ representa la constante de proporcionalidad entre el esfuerzo de corte y la deformación relativa, por ende también puede ser llamada *modulo de elasticidad* para los materiales que obedecen la Ley de Hooke, los cuales no fluyen y son linealmente elásticos.

Al esfuerzo de corte se le puede definir como un esfuerzo que provoca la deformación del material, y esta depende de la magnitud del esfuerzo, y en consecuencia se presentan modificaciones reológicas del material (Muller, 1973).

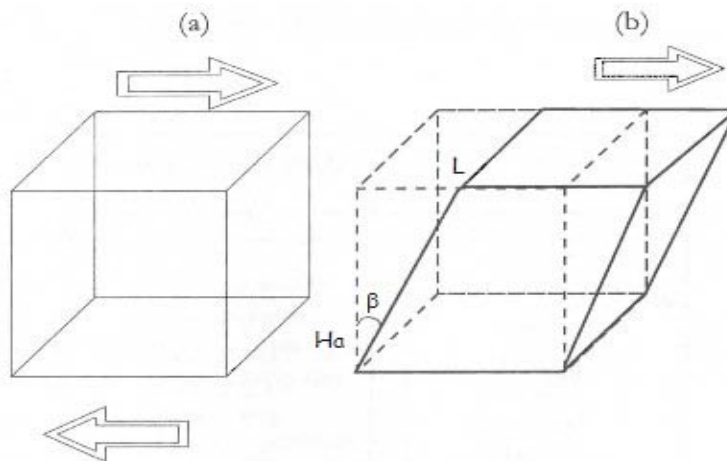


Figura 1.2 Cizalladura de un cubo (Méndez *et al*, 2001).

En el caso de los líquidos, la deformación se conoce como “flujo”. De manera similar a los sólidos, los líquidos obedecen a una ecuación de estado. La ecuación reológica de estado de los sólidos, relaciona la deformación con el esfuerzo de corte, la de los líquidos establece la

⁴ Es un cuerpo ideal, cuando se le aplica una fuerza, este se deforma, pero cuando deja de actuar dicha fuerza, el sólido recupera su forma inicial.

relación entre el esfuerzo de corte y la *velocidad de corte*⁵, la ecuación es (Méndez *et al.*, 2001):

$$\sigma = \dot{\gamma} \eta \quad (1.4)$$

Donde, σ es esfuerzo de corte, $\dot{\gamma}$ es velocidad de corte y η es el coeficiente de viscosidad. La velocidad de corte ($\dot{\gamma}$), se define cómo la magnitud de la deformación dividida por el tiempo que esta trata en producirse, sus unidades son (1/s).

Podemos notar que en la ecuación de estado de los líquidos (1.4), la constante es el “coeficiente de viscosidad” (η), equiparable al “modulo de elasticidad” de los sólidos.

1.1.2 Fluidos.

Se puede definir a un fluido como una sustancia que presenta la propiedad de que una porción de la misma puede desplazarse respecto a otra, es decir puede fluir de manera espontánea dentro de los recipientes que los contienen, venciendo fuerzas de atracción entre moléculas que originan una resistencia interna a este desplazamiento relativo (Resnick 1994). Por su parte Geankoplis, (1998), define a un fluido como una sustancia que no resiste de manera permanente, la deformación causada por una fuerza y por tanto cambia de forma o simplemente fluye.

En contraste los sólidos, solo se deforman hasta cierto punto, puesto que los cuerpos sólidos no son absolutamente rígidos porque las fuerzas externas pueden deformarlos y en realidad todos los materiales, incluso el acero más duro son elásticos en mayor ó menor grado, como lo que señalan Wilson & Buffa (2003).

En el caso de los alimentos estos son reológicamente complejos, porque aunque algunos son fácilmente definidos como líquidos como el agua, la leche, etc. Otros como las pastas, la mayonesa, el chicle, la masa de pan; son más complicado definirlos como fluidos ya que se podría argumentar que son sólidos deformables. En realidad, la clasificación de fluido o sólido deformable no es muy clara en muchos casos; frecuentemente sólo puede hablarse del grado en que un fluido se acerca a uno u otro tipo de comportamiento. En ese sentido, el comportamiento de los fluidos se puede acotar entre dos extremos, siendo el comportamiento viscoso newtoniano

⁵ También se puede encontrar como velocidad de deformación, intensidad de cizalladura ó gradiente de velocidad.

un extremo y el comportamiento sólido de Hooke⁶ el otro extremo. Ambas son sustancias ideales y por lo tanto inexistentes en la vida real.

1.1.3 Fluidos newtonianos

a) La viscosidad newtoniana.

Se denomina viscosidad a la fricción ó rozamiento interno del líquido (es decir, su resistencia interna al flujo). Difiere sin embargo de la fricción ó rozamientos de los sólidos puesto que al tratarse de un verdadero líquido la más mínima fuerza aplicada de un modo continuo, durante un largo periodo de tiempo, produce un flujo continuo (Méndez *et al*, 2001).

El coeficiente de viscosidad (η) es, una constante de proporcionalidad, un término cuantitativo; y se define como el coeficiente del esfuerzo de corte σ , dividida por la velocidad de corte es decir (Steffe, 1996):

$$\eta = \frac{\sigma}{\dot{\gamma}} \quad (1.5)$$

Si η es una constante independiente del valor de $\dot{\gamma}$, esta ecuación es la de un *líquido newtoniano*, éste se define como aquel para el que la representación del esfuerzo de corte en función de la velocidad de corte es una línea recta.

Usando las unidades de fuerza (N), área (m²), longitud (m) y velocidad (m/s) nos da *viscosidad* como Pa.s el cual es 1 poiseuille (Pl) que equivale a 10 poises, 1Pa.s son 1000 mPa.s, 1 mPa.s es igual a 1cP, éstas son algunas de las unidades en las cuales se puede encontrar.

El fluido newtoniano⁷ carece de propiedades elásticas, es incomprensible, isótropo⁸ y como el sólido de Hooke, ideal (Méndez *et al.*, 2001). Además cabe señalar que su comportamiento

⁶ Es el sólido ideal y teórico en el cual la deformación es proporcional a la fuerza (o tracción) que la produce. Carece de viscosidad y estructura (no tiene átomos), es isótropo (exhibe idénticas propiedades en todas las direcciones). Además su comportamiento reológico no se ve afectado con respecto al tiempo. Y obedece la Ley de Hooke.

⁷ El término *newtoniano* se usa en memoria de Sir Isaac Newton, quien fue el primero en definir el flujo viscoso en estos términos: la resistencia ocasionada por la “no deslizabilidad” de un líquido es proporcional, siempre que los demás factores se mantengan constantes, a la velocidad a que se separan las distintas partes del fluido.

⁸ Isótropo: dicese de los cuerpos cuyas propiedades físicas son idénticas en todas las direcciones.

reológico es independiente del tiempo.

Sin embargo muchos líquidos ofrecen, comportamiento newtoniano dentro de un amplio intervalo de esfuerzos de corte y se llegan a considerar como líquidos newtonianos. En estos casos la única propiedad necesaria para caracterizar el sistema es la viscosidad. El flujo de un líquido newtoniano puede ilustrarse gráficamente representado esfuerzo de corte σ (Pa) ó el coeficiente de viscosidad η (Pa.s), en función de la velocidad de corte $\dot{\gamma}$ (s^{-1}), ver Figura 1.3, para decidir si un líquido es ó no newtoniano basta con observar si al duplicarse σ se duplica también $\dot{\gamma}$. Si esto no sucede, el líquido no es newtoniano.

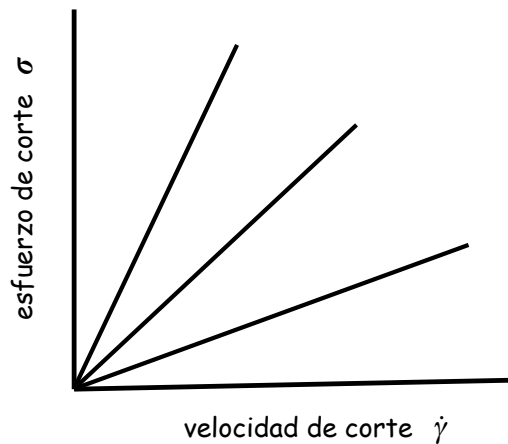


Figura 1.3 Gráfica de esfuerzo de corte σ , en función de la velocidad de corte $\dot{\gamma}$ de un fluido newtoniano.

Relación entre viscosidad y temperatura.

La viscosidad de los líquidos desciende notoriamente al aumentar la temperatura y cuando más alto es el coeficiente de viscosidad más se verá afectado por la temperatura. El coeficiente de viscosidad depende de la temperatura debido a que cuando la temperatura aumenta, el grado de movimiento molecular se incrementa reduciendo la distancia entre las fuerzas atractivas de las moléculas, por lo tanto la viscosidad del fluido disminuye. Han sido numerosos los intentos de relacionar cuantitativamente el coeficiente de viscosidad con la temperatura: los primeros trabajos al respecto se basaban en el hecho de que al aumentar la temperatura también aumenta el volumen de líquido, de donde se deducía que disminuía el número de moléculas por unidad de volumen (Méndez *et al.*, 2001; Rao, 1999).

Sin embargo la viscosidad de un líquido no solamente es función del volumen molecular; también juegan un papel importante las interacciones intermoleculares en el seno del líquido, cuya intensidad varía con la composición química de este.

El descenso de la viscosidad del agua, por ejemplo es del orden 1-3.5% por grado centígrado, dependiendo de la temperatura a la que se efectúe la medida (Méndez *et al.*, 2001). En el caso de los fluidos no-newtonianos, cualquier parámetro con las dimensiones de la viscosidad ó el mismo significado físico, presentara una dependencia similar a la de la viscosidad en fluidos que siguen un comportamiento newtoniano (Rao, 1999).

b) Estructura de los fluidos newtonianos.

Los líquidos simples y las disoluciones verdaderas suelen ser newtonianos, aunque son pocas las suspensiones ó disoluciones macromoleculares que presentan este comportamiento. En términos generales, las disoluciones ofrecen un flujo newtoniano si la cadena macromolecular tiene menos de 1000 átomos. Sin embargo también influye la concentración; a concentraciones bajas, pueden ofrecer comportamiento newtonianos incluso las disoluciones macromoleculares (Méndez *et al.*, 2001).

1.1.4 Fluidos no-newtonianos: características y modelos.

Un fluido no-newtoniano se define como aquel que exhibe flujo uniforme, pero para el cual no es constante la relación entre esfuerzo de corte y velocidad de corte (Méndez *et al.*, 2001). En otras palabras la viscosidad no es constante, pues deja de existir dicha proporcionalidad, ya que depende del esfuerzo de corte y las propiedades del flujo están influenciadas por la velocidad de corte (Rao, 1999), siempre en condiciones que ofrezcan un flujo continuo, incluso cuando la fuerza aplicada sea mínima.

Para estos casos en lugar del coeficiente de viscosidad η , se utiliza un “coeficiente de viscosidad aparente” η_{ap} , que se define como el esfuerzo de corte dividido por la velocidad de corte.

$$\eta_{ap} = f(\dot{\gamma}) = \frac{\sigma}{\dot{\gamma}} \quad (1.6)$$

Es claro que la viscosidad aparente no es una constante, depende del esfuerzo de corte.

Gran variedad de fluidos empleados en la industria alimentaria, son sustancias con un comportamiento no-newtoniano, y suelen tratarse de dispersiones de materiales poliméricos de alto peso molecular y suspensiones de sólidos o emulsiones de líquidos en el seno de un líquido, en los cuales el comportamiento no-newtoniano se incrementa con el aumento de concentración de sólidos, especialmente si los sólidos tienden a hincharse, solvatare o asociarse con la fase líquida. Las partículas dispersas pueden interactuar entre sí, al igual que con el medio, en cuyo seno se encuentran. La viscosidad aparente se define a partir del tipo de fluido de cual se trate y cambia de manera específica para cada uno de los fluidos.

Las propiedades de los fluidos no-newtonianos pueden ser clasificadas básicamente como independientes y dependientes del tiempo (Singh & Heldman, 2009). En la Figura 1.4, se muestra una clasificación de acuerdo el comportamiento reológico de un material, donde se consideran los parámetros reológicos obtenidos a partir de reómetros y la utilización de reogramas.

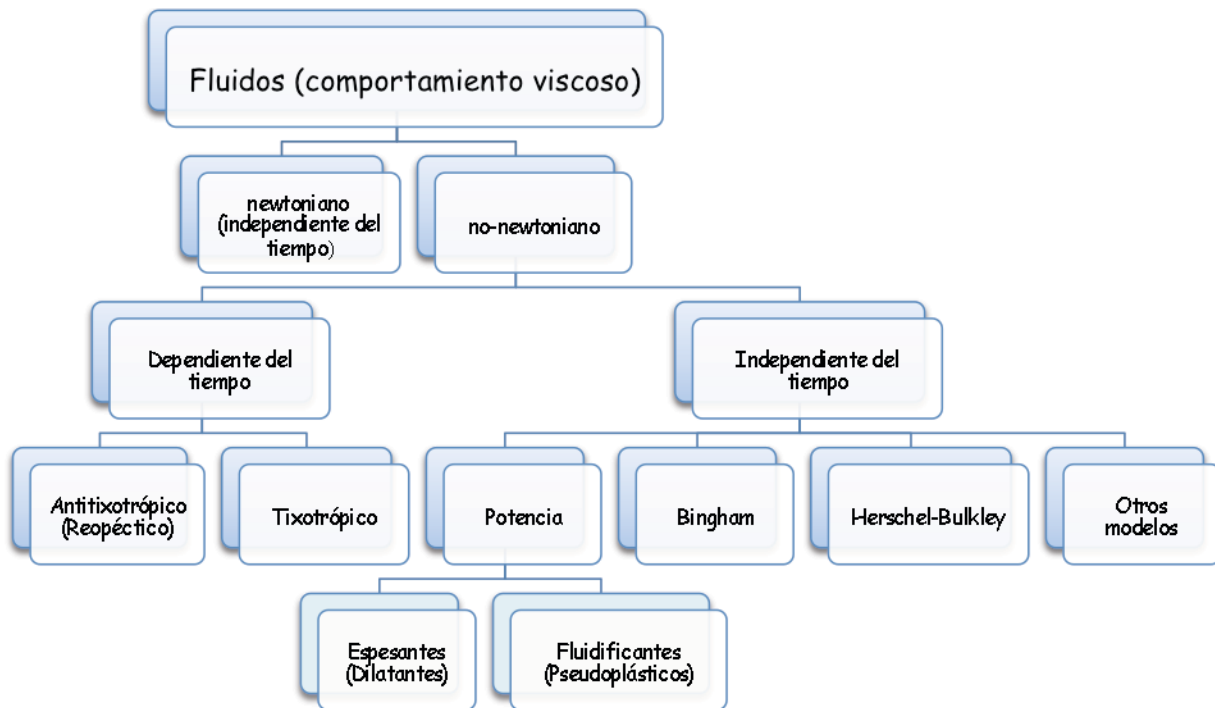


Figura 1.4 Clasificación de los fluidos de comportamiento viscoso, en base a su comportamiento reológico (Steffe, 1996).

Los fluidos no-newtonianos se dividen en base en su comportamiento al esfuerzo de corte y velocidad de corte, en dos categorías principales (Steffe, 1996; Singh & Heldman, 2009):

- a) Fluidos en los que el esfuerzo de corte es dependiente del tiempo ó duración de la acción cortante (dependientes del tiempo).
- b) Fluidos en los que el esfuerzo de corte es independiente del tiempo ó duración de la acción cortante (independiente del tiempo).

1.1.4.1 Comportamiento de fluidos dependientes del tiempo.

Los fluidos que siguen un comportamiento dependiente del tiempo son más complejos de analizar, debido a que su viscosidad aparente varía con respecto al tiempo, al gradiente de velocidad, de la historia de flujo y a la temperatura, a este tipo de fluidos se les clasifica como tixotrópicos y antitixotrópicos.

a) Tixotrópicos.

La tixotropia puede describirse como un decremento continuo de la viscosidad aparente con respecto al tiempo y la subsecuente recuperación de la viscosidad cuando el flujo se suspende (Méndez *et al.*, 2001). La representación gráfica que le corresponde es una curva similar a la pseudoplasticidad, en cuanto que η_{ap} desciende a medida que aumenta la velocidad de corte $\dot{\gamma}$. Difiere sin embargo, de ella en que el descenso de η_{ap} no está relacionado exclusivamente con la velocidad de corte sino también con el tiempo que dura la deformación.

En ambos casos cuando la velocidad de corte desciende o durante un periodo de reposo, la viscosidad aparente η_{ap} retorna a su valor inicial, es decir el producto comienza a espesarse lentamente hasta recobrar su comportamiento original. Este fenómeno es semejante a la dependencia al esfuerzo de corte y es resultado de la reorganización estructural que se pone en evidencia con una disminución ó aumento de la resistencia al flujo (Muller, 1973; Rosenthal, 2001). Otra característica es que cuando la velocidad de corte $\dot{\gamma}$ es constante, este esfuerzo de corte tiende a un valor límite que depende de la velocidad de corte.

La representación gráfica de σ en función de $\dot{\gamma}$ se obtiene aumentando $\dot{\gamma}$ a intervalos definidos, hasta un máximo y descendéndola luego del mismo modo hasta que alcance un valor de cero. Es

preciso mantener constante el intervalo de tiempo entre dos lecturas, al igual que las propias lecturas. Si las representaciones de σ en función de $\dot{\gamma}$ son dependientes del tiempo, terminan formando un bucle conocido como *histéresis* (Figura 1.5). El termino histéresis se aplica al fenómeno del comportamiento elástico aplazado; de ahí que la definición general de histéresis es: un proceso de deformación en que las fases de carga y descarga no coinciden, por lo que forman el bucle (Méndez *et al.*, 2001).

Se dice que estos fluidos tienen memoria reológica, es decir que como consecuencia directa con su dependencia con el tiempo es que la viscosidad aparente η_{ap} , se ve afectada por todas manipulaciones que implica el flujo del material previas a un ensayo reológico como por ejemplo la manera de verter el material en llenado del cuerpo de medida de un reómetro ó el muestreo, por ello se deben realizar siempre de un modo exactamente igual (Rosenthal, 2001).

Entre los alimentos que pueden ejemplificar este tipo de comportamiento reológico se encuentran, el yogurt, las mayonesas, leche condensada azucarada y algunos almidones.

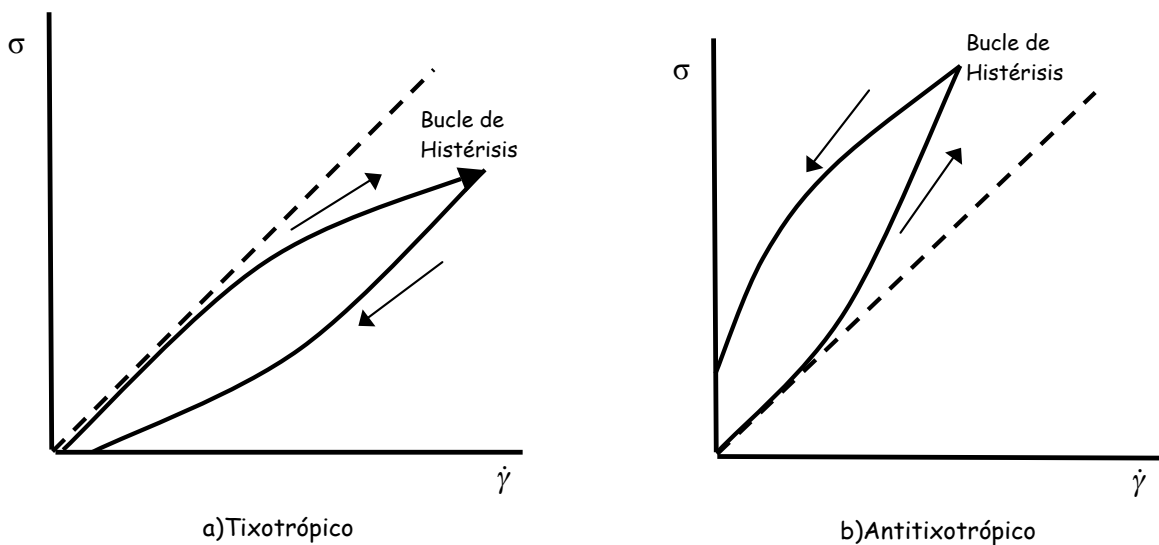


Figura 1.5 Reogramas de fluidos que exhiben comportamiento (a)Tixotrópico y (b)Antitixotrópico (Ibarz & Barbosa-Cánovas, 2005).

b) Antitixotrópicos

La antitixotropía es un espesamiento dependiente del tiempo, en el que el diagrama $\sigma - \dot{\gamma}$ es curvilíneo y en el que estos dos parámetros no están unívocamente relacionados (Figura 1.5 b). Existen dos valores de $\dot{\gamma}$ para cada valor de σ , y a la inversa, si se hacen rutinariamente pruebas cíclicas. La antitixotropía (también conocida como reopexia) es el fenómeno inverso de la tixotropía, por lo que es igualmente necesario un riguroso control del tiempo para las determinaciones (Méndez *et al.*, 2001; Singh & Heldman, 2009).

1.1.4.2 Comportamiento de fluidos independientes del tiempo.

Como su nombre lo dice, los fluidos independientes del tiempo, tienen una viscosidad que no varía con el tiempo, a cualquier esfuerzo cortante dado (Mott, 2006). Es decir que la viscosidad aparente η_{ap} depende solamente de la velocidad y del esfuerzo de corte aplicado y no de la duración de la deformación. Además los fluidos independientes del tiempo responden inmediatamente con flujo, tan pronto como una pequeña cantidad de esfuerzo cortante es aplicado, a diferencia los fluidos newtonianos, la relación entre esfuerzo corte y velocidad de corte es no-lineal ver Figura 1.6 (Singh & Heldman, 2009).

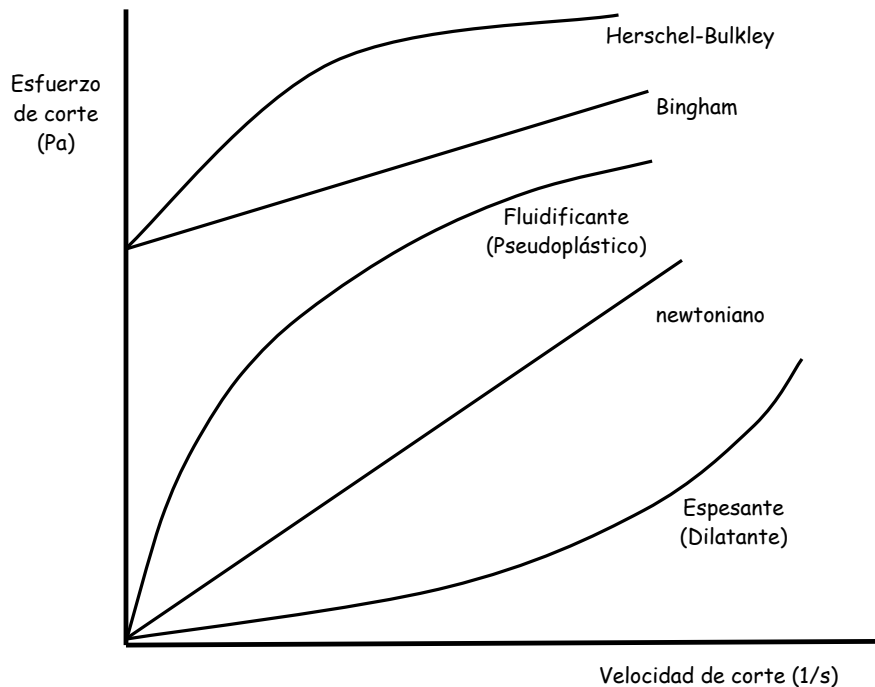


Figura 1.6 Curvas para fluidos típicamente independientes del tiempo (Singh & Heldman, 2009).

Dentro de los fluidos independientes del tiempo se encuentran:

a) Fluidos Ley de la Potencia.

Son los fluidos que presentan un comportamiento espesante a la cizalla y fluidificante a la cizalla, en los sistemas espesantes a la cizalla (dilatantes) y fluidificantes a la cizalla (pseudoplásticos) la relación en función del esfuerzo de corte σ y de la velocidad de corte $\dot{\gamma}$, no está expresada por una línea recta, sin embargo esta línea es única. A cada valor de esfuerzo de corte, le corresponde un solo valor de velocidad de corte. Por esa razón la gráfica se puede describir matemáticamente de la siguiente forma (Rao, 1999):

$$\sigma = K\dot{\gamma}^n \quad (1.7)$$

Donde K y n son constantes, K ha sido denominada como *índice de consistencia* en unidades (Pa.sⁿ) y n como *índice de comportamiento al flujo* (adimensional). Este último constituye una medida del grado de desviación del comportamiento newtoniano y ayuda a describir convenientemente los comportamientos que dependen del valor del índice de comportamiento al flujo n , es decir:

Cuando $n = 1$, el producto es newtoniano y $K =$ coeficiente de viscosidad (η).

Si $n > 1$, se produce un espesamiento, el producto es espesante a la cizalla (dilatante).

Si $n < 1$, se produce adelgazamiento y el material es fluidificante a la cizalla (pseudoplástico).

La viscosidad aparente para los fluidos de la Ley de la potencia se define como (Steffe, 1996):

$$\eta = f(\dot{\gamma}) = \frac{K(\dot{\gamma})^n}{\dot{\gamma}} = K(\dot{\gamma})^{n-1} \quad (1.8)$$

Es importante mencionar que el modelo de la Ley de la Potencia, sólo es válido para un intervalo medido. La ecuación no permite la extrapolación gráfica K y n , no tienen bases físicas, son una pura descripción matemática de una gráfica experimental. Sin embargo, tienen valor práctico: este modelo debe usarse si el esfuerzo de corte σ y la velocidad de corte $\dot{\gamma}$ no están unívocamente relacionados, porque en tal caso a cada valor de esfuerzo de corte σ le corresponden varios valores de velocidad de corte $\dot{\gamma}$ (Muller, 1973).

a.1) Pseudoplasticidad (Fluidificantes a la cizalla).

Los fluidos pseudoplásticos⁹ son menos espesos cuando se someten a altas velocidades de corte que cuando se cizallan lentamente. La viscosidad aparente η_{ap} depende en ellos de la velocidad de corte, pero no del tiempo durante el que están sometidos al esfuerzo de corte. La gráfica σ en función de $\dot{\gamma}$ no es una línea recta (Figura 1.7), la velocidad de corte aumenta en proporciones más altas que el esfuerzo de corte, de manera que la viscosidad aparente desciende a medida que aumenta la velocidad de corte (intensidad de cortadura). A cada valor de σ le corresponde otro de $\dot{\gamma}$ y a la inversa. Por ende el fenómeno se describe como de “estado estacionario”.

Algunos ejemplos comunes de alimentos fluidificantes son la leche condensada, los purés de frutas, mayonesa, mostaza, etc. Existen diferentes razones para explicar el comportamiento fluidificante, por ejemplo un fluido que aparentemente es homogéneo a simple vista, podría de hecho contener microscópicamente partículas sumergidas en él. Cuando estos fluidos son sometidos a un esfuerzo cortante, las partículas que están distribuidas en el líquido al azar se orientan en la dirección del flujo. De forma similar partículas agrupadas pueden deformarse en el sentido de flujo y alargarse o romper en grupos más pequeños. Este comportamiento es habitualmente reversible, así que cuando la acción de cizalladura se detiene, después de un tiempo determinado, las partículas regresan a su forma original (Singh & Heldman, 2009).

a.2) Dilatancia (espesante a la cizalla).

Es un fenómeno de espesamiento independiente del tiempo que se da a altas velocidades de corte (conocida también como cizalla); se trata del fenómeno opuesto a pseudoplasticidad. La gráfica no es lineal, pero los valores de σ y $\dot{\gamma}$ están unívocamente relacionados: η_{ap} aumenta al incrementar $\dot{\gamma}$ (Figura 1.7).

⁹ El calificativo de pseudoplástico deriva, del hecho de que en algunas muestras en los esfuerzos de corte (tensiones) bajos corresponden velocidades de corte muy pequeñas, de manera que las gráficas se parecen mucho a las de un Plástico de Bingham (Méndez, *et al.*, 2001).

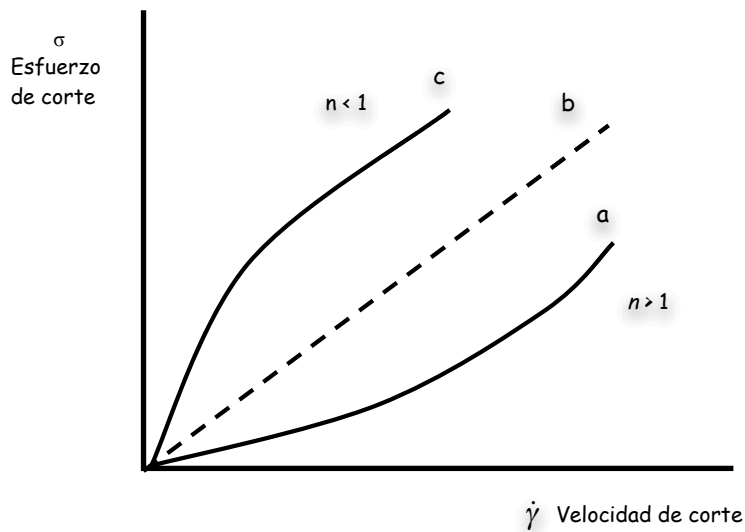


Figura 1.7 Gráfica de esfuerzo de corte σ en función de la velocidad de corte $\dot{\gamma}$ para modelos que describen un comportamiento (a) Espesante (dilatante) y (c) Fluidificante (pseudoplástico).

Ejemplo de fluidos espesantes incluyen suspensiones al 60% de almidón de maíz en agua, la mayoría de estos fluidos son suspensiones (partículas de sólidos en un líquido que actúa como plastificante). A bajas velocidades de corte, el líquido es suficiente para conservar las partículas sólidas bien lubricadas, y la suspensión fluye casi como un fluido newtoniano, pero cuando la velocidad de corte aumenta, las partículas de sólidos comienzan a separarse formando capas gruesas mientras incrementan el volumen total (Singh & Heldman, 2009).

b) Fluidos Bingham.

Estos fluidos necesitan de un *esfuerzo inicial* σ_0 para poder fluir, una vez iniciado el flujo la viscosidad aparente del fluido no cambia es constante, es decir se comporta como un fluido newtoniano. Si se graficará el esfuerzo de corte en función de la velocidad de corte, obtendríamos una pendiente lineal de la gráfica. El modelo para un fluido plástico de Bingham que exhibe un esfuerzo inicial es (Rao, 1999):

$$\sigma = \eta_{plB} \dot{\gamma} + \sigma_0 \quad (1.9)$$

Donde η_{plB} es la viscosidad plástica de Bingham.

c) Fluidos Herschel-Bulkley.

Estos fluidos tienen en común con los fluidos Bingham en necesitar de un esfuerzo de corte inicial σ_0 , sin embargo difieren en el comportamiento que seguirán después de que inicia el flujo, la viscosidad aparente en los fluidos Herschel-Bulkley no será constante, más bien se parecerá al comportamiento de un fluido tipo Ley de la Potencia (de tipo fluidificante). Además se puede decir que la diferencia principal entre un fluido que sigue un comportamiento Bingham a uno Herschel- Bulkley radica en el parámetro n (índice de comportamiento al flujo), y por supuesto en el comportamiento de cada fluido respecto a su viscosidad aparente. Otra característica importante a considerar sobre todo en cálculos de ingeniería es que por de bajo del esfuerzo de corte inicial el fluido se comporta como un sólido (Steffe, 1996).

La presencia del esfuerzo de corte inicial σ_0 , es un fenómeno conocido como “fuerza de cedencia”, definido como el esfuerzo crítico por debajo del cual no existe el flujo. El modelo Herschel-Bulkley se expresa como (Rao, 1999):

$$\sigma = \sigma_0 + k(\dot{\gamma})^n \quad (1.10)$$

Respecto a la viscosidad aparente de un fluido Herschel-Bulkley, esta se puede describir como una función de la velocidad de corte (Steffe, 1996):

$$\eta_{ap\ HB} = f(\dot{\gamma}) = \frac{K(\dot{\gamma})^n + \sigma_0}{\dot{\gamma}} = K(\dot{\gamma})^{n-1} + \frac{\sigma_0}{\dot{\gamma}} \quad (1.11)$$

En fluidos Herschel-Bulkley, la viscosidad aparente decrecerá con altas velocidades de corte cuando $0 < n < 1$, pero se tendrá el fenómeno opuesto cuando $n > 1$.

Estos fluidos que requieren esfuerzo de corte inicial (fuerza de cedencia) para fluir, pueden verse como si tuvieran una red inter-partículas o intermolecular que resiste bajos niveles esfuerzos cortantes cuando esta en reposo. A bajos niveles de esfuerzo de corte inicial, el material actúa como un sólido, es solo cuando se aplica esfuerzo que excede las fuerza que sostienen junta a la red que el material comienza a fluir (Singh & Heldman, 2009).

Los comportamientos antes descritos son los principales comportamientos de fluidos viscoso, y como ya se mencionó el comportamiento de los materiales se puede acotar entre dos extremos, el

viscoso newtoniano en un extremo y el sólido de Hooke (elástico) en el otro extremo, por ende los materiales reales exhiben tanto comportamientos viscosos como elásticos, aunque un tipo de comportamiento es frecuentemente el dominante (Steffe, 1996).

1.1.5 Regímenes de flujo: laminar y turbulento.

Para comprender las características de flujo de un fluido viscoso (líquido), imaginemos un tubo en el cual fluye líquido lentamente, el líquido del tubo constituye, un sistema continuo, pero al fluir se le puede imaginar que es formado por una serie de delgadas láminas concéntricas (Figura 1.8). Si el flujo es lento, la lámina fluida central se moverá más rápido y la más externa permanecerá en reposo, las intermedias ofrecerán velocidades de flujo dentro el intervalo delimitado por ambos valores y producirá un rozamiento entre ellos.

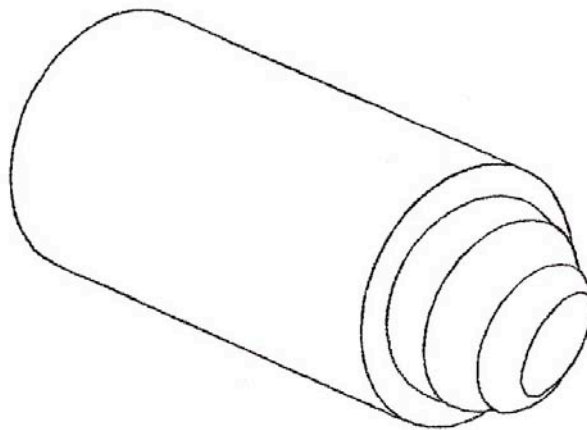


Figura 1.8 Flujo laminar en un ducto circular.

A este tipo de flujo se le llama en régimen laminar, que se define como el deslizamiento de entidades independientes de numerosas láminas de fluido infinitamente delgadas.

Otro ejemplo del desplazamiento de un fluido en régimen laminar es el que se observa en la Figura 1.9, donde un fluido se desliza entre dos placas paralelas una de ellas en movimiento y otra inmóvil, esto provoca que el fluido se desplace en laminas a diferentes velocidades.

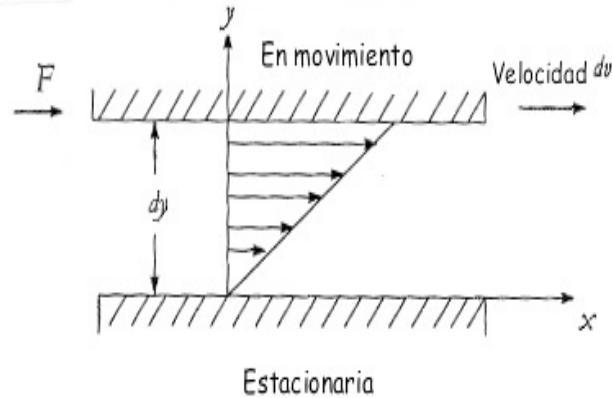


Figura 1.9 Velocidad de distribución de un fluido viscoso entre dos placas paralelas.

Además del régimen laminar, existe el régimen de flujo turbulento, entre ellos hay una zona denominada de zona de transición. El flujo laminar sólo se da a velocidades inferiores antes de que aparezca la turbulencia, es decir cuando el movimiento del fluido no sea irregular ó distorsionado, esto es mientras exista un perfil de velocidades definido.

Como ya se mencionó el flujo laminar de un fluido es en forma de líneas rectas paralelas, y en un conducto circular correspondería a anillos concéntricos de fluido de radios diversos que se trasladan siguiendo una trayectoria recta y uniforme (Figura 1.8). En ese caso la velocidad del flujo varía con el radio, es decir su valor máximo se tiene el centro y su valor mínimo (el cual suponemos es cero), en las paredes (Figura 1.10). Y por lo tanto la variación de la velocidad con la posición a lo largo del tubo no es lineal (Resnick, 1994).

El régimen de flujo (laminar ó turbulento) en una tubería depende de tres parámetros físicos que describen condiciones flujo. El primer parámetro es una escala de longitud del campo de flujo, es decir el diámetro de una tubería. Si dicha escala de longitud es lo bastante grande, una perturbación de flujo podría aumentar y el flujo se convertiría en turbulento. El segundo parámetro es una escala de velocidad tal como el promedio espacial de la velocidad, si la velocidad es lo bastante grande, el flujo podría ser turbulento. El tercer parámetro es la viscosidad cinemática¹⁰, si esta es lo bastante pequeña, el flujo podría ser turbulento. Los tres parámetros

¹⁰ Es el coeficiente de viscosidad medido directamente en los viscosímetros capilares de tubo en U, en los que el esfuerzo de corte lo proporciona el propio líquido de alimentación, y este esfuerzo depende tanto de la altura de columna y la densidad del líquido (Pineda & Romero, 2003).

pueden combinarse en uno solo y puede servirnos para predecir los regímenes de flujo (Porter, 1998).

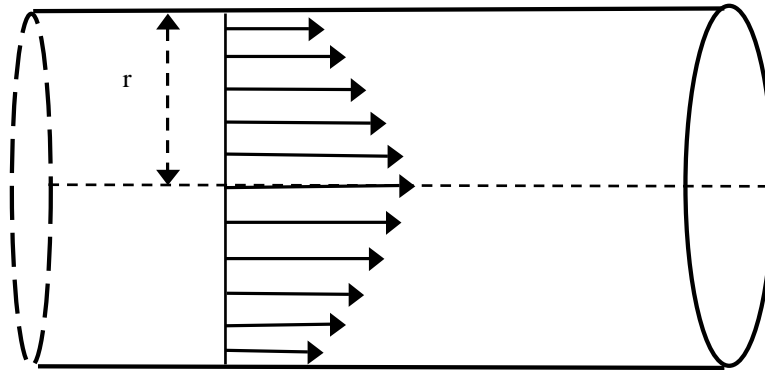


Figura 1.10 Perfil de velocidades definido de un fluido en una tubería.

El primer investigador que estudio científicamente a la turbulencia fue Osborne Reynolds, y demostró que el régimen de flujo en tuberías, tanto laminar como turbulento depende del diámetro de tubería, de la densidad, la viscosidad del fluido y de la velocidad del flujo.

Estas variables se combinan en la expresión del número de Reynolds, y este puede considerarse como la relación de las fuerzas de la masa del fluido con respecto a los esfuerzos de deformación ocasionados por la viscosidad, es decir:

$$Re = \frac{\text{fuerzas de inercia}}{\text{fuerzas viscosas}} \quad (1.12)$$

La magnitud del número de Reynolds nos permite saber previamente si el flujo será laminar o turbulento, cabe señalar que el número de Reynolds es un número adimensional y se puede calcular para los fluidos newtonianos con la ecuación:

$$Re = \frac{Dv\rho}{\eta} \quad (1.13)$$

Donde:

Re Número de Reynolds (adimensional)

D Diámetro de tubería (m)

ρ	Densidad del fluido (Kg/m^3)
η	Viscosidad del fluido (Pa.s)
v	Velocidad promedio del fluido (m/s)

De acuerdo a la ecuación anterior sabemos que cuando los flujos tienen un número Reynolds grande, puede ser debido a una alta velocidad o a una baja viscosidad o a ambas por lo que el régimen es turbulento. Aquellos fluidos que poseen una alta viscosidad y/o se mueven a baja velocidades tendrán un número de Reynolds pequeño, por lo que el flujo tenderá a ser laminar (Mott, 1996).

Si el fluido se encuentra en régimen laminar, Re alcanza un valor máximo del orden de 2100 (Steffe, 1996; Geankoplis, 1998), aunque algunos otros autores proponen distintos intervalos de valores para determinar tanto el régimen laminar y el régimen turbulento. La mayor aceptación es que hasta un valor de Re menor a 2100 para un fluido newtoniano en una tubería circular el régimen es laminar y a partir de Re mayor a 4000, el flujo se tornará turbulento excepto para unos casos especiales. Entre estos dos valores, existe la región de transición y el flujo puede ser viscoso o turbulento, dependiendo de los detalles del sistema, que no se pueden predecir (Geankoplis, 1998).

También hay que considerar que las características pseudoplásticas tienden a estabilizar el flujo por lo que la turbulencia solo se produce a valores de Re más altos. Si el coeficiente de viscosidad es extremadamente alto, el flujo puede no ser nunca turbulento, por resultar imposible alcanzar experimentalmente velocidades de flujo suficientemente altas. Para líquidos viscoelásticos, suelen determinarse valores de Re muchos mas bajos. Las partículas alargadas, frecuentemente presentes en estas suspensiones, actúan como núcleos de turbulencia. Generalmente, Re decrece a medida que aumenta el tamaño y las concentraciones de las macromoléculas (Méndez *et al.*, 2001)

Es importante señalar que la ecuación para obtener el número de Reynolds toma una forma diferente cuando se trata de conductos con secciones transversales no circulares, canales abiertos ó cuando el flujo de fluidos que tiene un comportamiento diferente al newtoniano, en el caso de el número de Reynolds para un fluido plástico de Bingham, la viscosidad plástica se sustituye por la viscosidad del fluido y se utiliza la ecuación (1.13), sin embargo cabe resaltar que al tratarse fluidos no-newtonianos no se puede calcular Re , porque a diferencia de la viscosidad newtoniana,

el coeficiente de viscosidad es variable (Muller, 1973). Pero se realiza otro cálculo equiparable como en el caso de fluidos de la Ley de la Potencia y Herschel-Bulkley que se calcula el Reynolds generalizado:

$$Re_G = \frac{D^n v^{(2-n)} \rho}{8^{(n-1)} K} \left(\frac{4n}{1+3n} \right)^n \quad (1.14)$$

Donde K , es el índice de consistencia y sus unidades son (Pa.sⁿ).

1.2 Balance mecánico de energía para un fluido en una tubería en estado estacionario.

Para conocer la cantidad de energía necesaria para mover un fluido, es necesario hacer un balance total de energía en el sistema, pues un fluido solo comenzara a moverse cuando la resultante de las fuerzas que actúan sobre él, sea diferente de cero. En el caso del flujo de fluidos el balance más útil es el *balance general de energía mecánica*. La ecuación del balance general de energía es una derivación del principio de conservación de *momentum*, comúnmente llamado la ecuación de ingeniería de Bernoulli (Geankoplis, 1998; Steffe & Singh, 1997; Singh & Heldman, 2009). Bajo condiciones estacionarias de flujo, la fuerza resultante que actúa sobre un líquido debe ser igual a la tasa neta de cambio en cantidad de movimiento (*momentum*), estos conceptos son básicos para derivar la ecuación de Bernoulli.

$$P_1 + \frac{1}{2} \rho v_1^2 + \rho g z_1 = P_2 + \frac{1}{2} \rho v_2^2 + \rho g z_2 \quad (1.15)$$

La ecuación de Bernoulli recibe su nombre del matemático Daniel Bernoulli, y la aplicación de esta ecuación permite un estudio profundo de los problemas de flujo de fluidos, sin embargo si no se cumplen las suposiciones usadas para la obtención de la ecuación, probablemente se obtendrán resultados erróneos, estas suposiciones son (Singh & Heldman, 2009):

- Las posiciones 1 y 2 se encuentran en la misma línea de corriente
- El fluido tiene densidad constante, y por tanto es incomprensible
- El fluido tiene viscosidad cero

- El flujo es estacionario
- No existe trabajo realizado sobre o por el fluido
- No tiene transmisión de calor entre el fluido y sus alrededores

Al desarrollar la ecuación del balance general de la energía mecánica se han hecho distintas suposiciones; entre ellas no hay que olvidar, una sola fase del fluido, un material con propiedades uniformes y una presión equivalente uniforme.

Sin embargo, al momento de tratarse del transporte de un fluido real, no se debe ignorar la viscosidad. Para superar las fuerzas viscosas, comúnmente llamadas **fricción** del fluido, es necesario realizar una determinada cantidad de trabajo. Debido a este trabajo de fricción, existe una transformación de energía en calor, con el consiguiente aumento de temperatura. Sin embargo, el aumento de temperatura es normalmente muy pequeño y de escaso valor práctico, y el trabajo de fricción se considera a menudo una pérdida de energía útil y se puede expresar como E_f , además en aquellos problemas de bombeo de fluidos se puede utilizar a W para expresar el trabajo realizado por la bomba (energía mecánica) (Geankoplis, 1998; Singh & Heldman, 2009).

Todas las formulas prácticas para el flujo de fluidos se derivan del teorema de Bernoulli, pero con modificaciones para tener en cuenta las pérdidas por fricción, en particular el balance de energía mecánica en estado estacionario para un fluido incompresible en una tubería puede ser escrito como (Singh & Heldman, 2009):

$$\frac{P_2}{\rho_2} + \frac{1}{2}v_2^2 + gz_2 + E_f = \frac{P_1}{\rho_1} + \frac{1}{2}v_1^2 + gz_1 + W \quad (1.16)$$

Reorganizando los términos de la ecuación (1.17) para obtener una expresión para los requerimientos de energía de la bomba (W) por unidad de masa, y si estamos tratando con un fluido incompresible, $\rho_2 = \rho_1 = \rho$,

$$W = \frac{P_2 - P_1}{\rho} + \frac{1}{2}(v_2^2 - v_1^2) + g(z_2 - z_1) + E_f \quad (1.17)$$

Los subíndices indican los dos puntos del balance es decir el punto de entrada y el punto de salida del sistema, y cada término de la ecuación representa un diferente tipo de energía involucrada (Geankoplis, 1998; Singh & Heldman, 2009):

$\Delta P / \rho$ Es la disipación de energía relativa al cambio de presión, entre los puntos 1 y 2, y representan la energía de presión, donde P es la Presión (Pa) y ρ es la densidad del fluido (Kg/m^3).

$\frac{1}{2}\Delta v^2$ Representa la energía cinética, debido al movimiento de traslación o rotación de la masa, donde v es la velocidad de flujo (m/s) con respecto al límite del sistema en cierto punto.

$\Delta z g$ Representan la energía potencial, debido a la posición de la masa en un campo gravitacional, donde z es la altura relativa (m) desde un plano de referencia y g es la constante de aceleración debida a la gravedad (9.81 m/s^2).

E_f Representa las pérdidas de energía por fricción para un fluido que fluye por una tubería, y están compuestas por pérdidas mayores y pérdidas menores (J/Kg).

W Representa el trabajo por unidad de masa (J/Kg), realizado por la bomba sobre el fluido.

Cabe señalar que debido a los efectos viscosos la velocidad nunca es uniforme en la sección de la tubería, sino que varía, por lo tanto se debe usar un factor de corrección, α y modificar el término de la energía cinética de la siguiente manera:

$$\frac{v_2^2 - v_1^2}{2\alpha} \quad (1.18)$$

Donde α es un factor de corrección adimensional que depende del régimen de flujo y del tipo de fluido que esta siendo transportado.

En régimen turbulento, el valor de $\alpha = 0.945$; por lo tanto, siempre se toma el valor de la unidad para α , y esto es para cualquier tipo de fluido (Ibarz & Barbosa-Cánovas, 2005). Sin embargo

Steffe, (1996) y Rao, (1999) indican que el valor de α en flujo turbulento para cualquier fluido independiente del tiempo es $\alpha = 2$.

En el caso del régimen laminar el valor de α cambia con respecto al tipo de fluido, como se muestra a continuación (Osorio & Steffe, 1984; Steffe, 1996).

Tabla 1.1 Factor de corrección para la energía cinética para flujo laminar en tuberías.

Valores de α en flujo laminar
Fluidos newtonianos:
$\alpha = 1$
Fluidos de la Ley de la potencia:
$\alpha = \frac{2(2n + 1)(5n + 3)}{3(3n + 1)^2}$
Plásticos de Bingham:
$\alpha = \frac{2}{2 - \xi_0}$
Fluidos Herschel-Bulkley:
$\alpha = \frac{[2(1 + 3n + 2n^2 + 2n^2\xi_0 + 2n\xi_0 + 2n^2\xi_0^2)^3(2 + 3n)(3 + 5n)(3 + 4n)]}{\left[\left[(1 + 2n)^2(1 + 3n)^2 \right] \left[18 + n(105 + 66\xi_0) + n^2(243 + 306\xi_0 + 85\xi_0^2) + n^3(279 + 522\xi_0 + 350\xi_0^2) + n^4(159 + 390\xi_0 + 477\xi_0^2) + n^5(36 + 108\xi_0 + 216\xi_0^2) \right] \right]}$

Siendo $n =$ índice de comportamiento al flujo y $\xi_0 = \frac{\sigma_0}{\sigma_w}$ (el cual se explicará más adelante en la sección 1.2.2), ambas características son intrínsecas al fluido y sus condiciones de transporte en la tubería.

En el caso de fluidos Herschel-Bulkley, en general el valor numérico de α va de un intervalo de 0.74 a 2 (Steffe & Morgan, 1986; Steffe, 1996). Y el valor mínimo de $\alpha=0.74$ ocurre cuando $\xi_0=0$, y n se acerca a infinito.

Aunque exista una ligera divergencia entre, cual es el valor de α , la realidad es que las diferencias en la energía cinética son usualmente pequeñas y frecuentemente ignoradas al evaluar los requerimientos de energía al seleccionar una bomba (Steffe, 1996). Las velocidades de v_1 y v_2 , pueden ser muy cercanas, a menos que exista un cambio de diámetro durante el proceso y por lo tanto haya cambios de velocidad de flujo. En la práctica con valores constantes de n (en un intervalo de 0.2 a 0.6 que es el típico para materiales alimenticios), la variación de α es pequeña sobre el sistema de bombeo (Steffe & Morgan, 1986).

También hay que mencionar que para unidades inglesas, los términos de energía cinética y potencial de la ec. (1.18) se dividen entre g_c que es un factor de proporcionalidad, que tiene un valor de 32.174 ft lb_f/ lb_m y de 1.0 en el Sistema internacional de unidades (SI).

Pérdidas totales de energía debidas a la fricción.

Las pérdidas de energía debidas a la fricción¹¹ de un líquido que fluye por una tubería E_f , esta compuesta por pérdidas mayores y menores:

$$E_f = E_{f_{mayor}} + E_{f_{menor}} \quad (1.19)$$

Las pérdidas mayores se deben al flujo del líquido viscoso en las porciones rectas de la tubería, esta se explicará ampliamente en la siguiente sección (1.2.1). Por otro lado las pérdidas menores, se deben a varios componentes que se usan en las conducciones: expansiones y contracciones en la tubería, accesorios (tales como válvulas, conexiones, aparatos de medición, etc.), así como a cambios en la dirección y/o diversificación del flujo del fluido, debidos generalmente a codos, tes, yes, y equipos (Ibarz & Barbosa-Cánovas, 2005; Geankoplis, 1998). Los equipos de proceso que pueden ser instalados en el sistema de transporte de fluido, tienen normalmente asociada una determinada pérdida de carga por fricción, así que también las pérdidas de energía debidas a equipos especiales (como intercambiadores de calor) debe ser añadido a E_f (Steffe & Singh, 1997).

¹¹ A las pérdidas de energía debida a la fricción, también se le conoce como Hfs y pueden ser tanto de accesorios ($Hf_{S_{Acc}}$) y de tuberías ($Hf_{S_{Tub}}$) la suma de ambas sería lo que mismo que E_f . $E_f = Hfs = \frac{\Delta P}{\rho}$

En general, las pérdidas debidas a cada uno de estos elementos son pequeñas e inclusive podrían despreciarse, pero en tuberías que contengan un número considerable de estos accesorios, dichas pérdidas pueden llegar a ser importantes (Ibarz & Barbosa-Cánovas, 2005).

Las pérdidas de energía debidas a la fricción que se presentan tanto en tuberías rectas, accesorios y equipos, pueden expresarse de acuerdo a la ecuación (1.20), el primer término de la ecuación se refiere a las pérdidas de presión en tuberías y el segundo se refiere a las pérdidas de presión en accesorios:

$$E_f = \frac{2 f_f L v^2}{D} + \sum_{a=1}^b \frac{K_f v^2}{2} \quad (1.20)$$

Donde:

- D Diámetro de tubería (m)
- E_f Pérdida de energía por unidad de masa debida a la fricción (J/Kg)
- f_f Factor de fricción de Fanning (adimensional)
- K_f Coeficiente de resistencia por fricción en accesorios (adimensional)
- L Longitud de tubería (m)
- v Velocidad volumétrica en una tubería (m/s)

En el primer término cabe mencionar la importancia del f_f del cual se profundizara más en la siguiente sección, en el segundo término de la ecuación (1.20), b es el número total de válvulas ó accesorios (contracciones, expansiones, etc.), hay que señalar que el coeficiente de resistencia (K_f), se refiere a las pérdidas adicionales en términos de una cabeza de velocidad equivalente debida al accesorio, el cual es único para cada accesorio en particular. Para determinar los valores de K_f , es necesario datos experimentales, la mayoría de los valores están publicados en Crane (1996) u otros manuales de ingeniería. Y cabe señalar que el valor de K_f aunque debería ser constante para accesorios que son geoméricamente similares, este decrece conforme se incrementa el diámetro, parte de esta variación se debe a la falta de un escalamiento geométrico y también es debido al análisis que ignora una posible dependencia con el número de Reynolds y el tipo de fluido que se transporta (Steffe, 1996), por lo cual es importante considerar siempre estos aspectos al consultar estas tablas y en caso de no encontrar con datos correspondiente al K_f en las condiciones requeridas, Steffe (1996) recomienda una guía heurística de cómo estimar el

valor de K_f . Por último también hay que considerar que los valores de v , K_f y f_f cambiaran siempre que en el sistema de bombeo haya cambios del diámetro de tuberías (Steffe & Morgan, 1986).

1.2.1. El factor de fricción en el diseño de sistemas de bombeo en los procesos alimenticios.

En la industria de alimentos es indispensable contar con líneas para el transporte de fluidos, adecuadas para el manejo de los alimentos. Transportar un fluido desde un punto del sistema de transporte de fluidos a otro generalmente implica utilizar una bomba que nos permita crear un diferencial de presiones (ΔP) entre ambos puntos, por ello es primordial conocer las pérdidas de energía debidas a la fricción del fluido con las paredes de la tubería, equipos ó accesorios por los que pasa ya que tales pérdidas de energía mecánica tienen como resultado una disminución ó caída de presión entre dos puntos del sistema de flujo, lo que representaría un deficiente o nulo transporte del fluido.

Tener un cálculo preciso de como cambia la presión en un punto a otro implica conocer el factor de fricción del fluido dentro de la tubería, que no es otra cosa que la cantidad de energía mecánica perdida debido a la fricción, y esta última es la resistencia de un fluido a ser transportado por una tubería. En el cálculo del factor de fricción es evidente la relación entre las propiedades reológicas del fluido, las características del sistema de tuberías y condiciones a las que se transporta, pues todas y cada una de ellas ejerce una influencia única al sistema de transporte.

La influencia de las fuerzas de fricción en el sistema, se expresan en la forma de un **factor de fricción**, pues es necesario cuantificar esas pérdidas de energía por fricción en el sistema, mediante la medición de la velocidad y las diferencias de presión de un fluido en movimiento en una distancia establecida (Singh & Heldman, 2009; Steffe & Singh, 1997).

Comenzando con el desarrollo matemático, el factor de fricción de Fanning (f_f) es definido, desde las consideraciones de un análisis dimensional, como el radio de la tensión de cizalladura

de la pared en una tubería a la energía cinética por unidad de volumen según lo señalan Steffe & Singh (1997).

$$f_f = \frac{2\sigma_w}{\rho v^2} \quad (1.21)$$

Si se considera la definición de esfuerzo de corte en la pared, derivado de la deducción de la ecuación de Rabinowitsch-Mooney (Steffe, 1996):

$$\sigma_w = \frac{(\Delta P)r}{2L} \quad (1.22)$$

f_f puede ser considerado en términos de caída de presión por la sustitución de la definición de esfuerzo de corte en la pared:

$$f_f = \frac{(\Delta P)r}{\rho L v^2} = \frac{(\Delta P)D}{2\rho L v^2} \quad (1.23)$$

Esta es la ecuación que define el factor de fricción para cualquiera modelo reológico y régimen de flujo, donde:

- f_f factor de fricción de Fanning (adimensional)
- ΔP caídas de presión, $\Delta P = P_1 - P_2$ (Pa)
- D diámetro (m)
- L longitud (m)
- ρ densidad de fluido (Kg/m^3)
- v velocidad promedio (m/s)

de acuerdo a la ecuación anterior la energía de presión pérdida por unidad de masa (necesaria en el balance de energía mecánica) puede ser expresada en términos de f_f

$$\frac{(\Delta P)}{\rho} = \frac{f_f 2L v^2}{D} \quad (1.24)$$

Cabe señalar que también se puede calcular la energía de presión pérdida con el factor de fricción de Darcy y que hay una diferencia entre el factor de fricción de Fanning y el de Darcy y aunque

esta es mínima hay que tomarla en cuenta para evitar errores de cálculo, y esto es debido a la procedencia de la ecuación del factor de fricción de Darcy, ya que aunque resulta del mismo tratamiento que se hizo para la ecuación de Fanning, la única diferencia radica en que el factor de Darcy equivale a cuatro veces el factor de Fanning, y el factor de cuatro tiene su base en un concepto denominado radio hidráulico¹². Tampoco hay que olvidar que las caídas de Presión se podrían ajustarse por esta diferencia (Steffe, 1996; Singh & Heldman, 2009).

Las ecuaciones de factor de fricción de Fanning que se presentan a continuación son solo para tuberías sin rugosidades, es decir la tubería necesaria para sistemas de bombeo sanitario que es el indicado para transportar alimentos.

En flujo laminar, los valores de f_f pueden ser determinados desde ecuaciones que describen la relación entre caída de presión y velocidad de flujo para un fluido en particular. Considerar, por ejemplo a un fluido newtoniano ($\sigma = \eta \dot{\gamma}$) y usando la ecuación Hagen-Poiseuille para fluido newtoniano (la cual indica que el radio tiene una gran influencia sobre el comportamiento del sistema desde que este es elevado a la cuarta potencia) de acuerdo Steffe (1996):

$$Q = \frac{\pi \Delta P r^4}{8 L \eta} \quad (1.25)$$

Entonces, el promedio de la velocidad volumétrica para este material en un tubo en flujo laminar puede ser calculado:

$$v = \frac{Q}{\pi r^2} = \frac{1}{\pi r^2} \left(\frac{\pi (\Delta P) r^4}{8 L \eta} \right) = \frac{(\Delta P) D^2}{32 L \eta} \quad (1.26)$$

La simplificación nos da una expresión para la caída de presión por unidad de longitud:

$$\frac{\Delta P}{L} = \frac{32 v \eta}{D^2} \quad (1.27)$$

¹² Es la relación entre la superficie transversal del flujo y el perímetro mojado de la superficie, en los conductos circulares y trabajando en presión, en función de diámetro se define $R_h = \frac{A}{P_e} = \frac{\frac{\pi D^2}{4}}{\pi D} = \frac{D}{4}$, donde A= área (m²), P_e =perímetro (m), D=diámetro (m).

Sustituyendo esta expresión dentro de la ecuación (1.23) da lo siguiente:

$$f_f = \left[\frac{\Delta P}{L} \right] \left(\frac{D}{2\rho(v)^2} \right) = \left[\frac{32(v)\eta}{D^2} \right] \left(\frac{D}{2\rho(v)^2} \right) = \frac{16}{Re} \quad (1.28)$$

El cual es apropiado para predecir factores de fricción para fluidos newtonianos en régimen de flujo laminar, $Re < 2100$.

$$f_f = \frac{16}{Re} \quad (1.29)$$

Usando el mismo procedimiento, para fluidos en régimen laminar, pueden calcularse el factor de fricción para fluidos de la Ley de la Potencia con la ecuación (1.29) y para plástico de Bingham¹³ desde la siguiente ecuación:

$$\frac{1}{Re_B} = \frac{f_f}{16} - \frac{He}{6(Re_B)^2} + \frac{(He)^4}{3f_f^3(Re_B)^8} \quad (1.30)$$

El factor de fricción para flujo laminar de fluidos newtonianos, Ley de la Potencia, Bingham y Herschel-Bulkley pueden ser obtenidos desde simples relaciones. Sin embargo existen numerosas correlaciones para estimar el factor de fricción bajo las condiciones de régimen turbulento (Garcia & Steffe, 1987). En flujo turbulento, el factor de fricción puede ser determinado por ecuaciones empíricas formuladas a partir de datos experimentales según lo cita Steffe (1996). En el caso de condiciones de flujo en transición y turbulento las derivaciones matemáticas llegan a ser altamente complejas (Singh & Heldman, 2009). Para fluidos tipo Herschel- Bulkley han existido varios desarrollos para el cálculo de factor de fricción en tuberías lisas (sanitarias), sobre todo para condiciones de flujo en transición o turbulento, sin embargo los de mayor aceptación son las ecuaciones desarrolladas por Hanks (1963) para régimen laminar y la ecuación desarrollada por Torrance (1963), en régimen turbulento como lo señalan García & Steffe (1987).

¹³ donde el $Re_B = \frac{Dv\rho}{\eta_{ap}}$

1.2.2 Ecuaciones para el cálculo de factor de fricción en régimen laminar y turbulento para fluidos Herschel-Bulkley.

a) Factor de fricción en régimen laminar para fluidos Herschel-Bulkley.

En el caso de fluidos Herschel-Bulkley en condiciones de flujo laminar el factor de fricción puede ser encontrado por un método que fue descrito por Hanks en 1978, como lo mencionan Garcia & Steffe (1987):

$$f_f = \frac{16}{\psi Re_G} \quad (1.31)$$

En la cual Re_G , es el Reynolds generalizado (que puede ser utilizado tanto para fluidos Herschel-Bulkley y Ley de la Potencia) es definido como lo muestra la ecuación (1.14) y ψ (psi), que es una función del flujo laminar que esta descrita como:

$$\psi = (1 + 3n)^n (1 - \xi_0)^{1+n} \left[\frac{(1 - \xi_0)^2}{(1 + 3n)} + \frac{2 \xi_0 (1 - \xi_0)}{(1 + 2n)} + \frac{\xi_0^2}{(1 + n)} \right]^n \quad (1.32)$$

Donde ξ_0 (ξ_0) es la relación adimensional del esfuerzo cortante inicial y es fuerza en la pared:

$$\xi_0 = \frac{\sigma_0}{\sigma_w} = \frac{\sigma_0}{(D\Delta P/4L)} = \frac{\sigma_0}{(\frac{1}{2} f_f \rho v^2)} \quad (1.33)$$

ξ_0 también puede ser escrita como una función implícita de Re_G y una generalización del número de Hedstrom (He) como:

$$Re_{He} = 2He \left(\frac{n}{1 + 3n} \right)^2 \left(\frac{\psi}{\xi_0} \right)^{\frac{2-n}{n}} \quad (1.34)$$

Donde:

$$He = \frac{D^2 \rho}{K} \left(\frac{\sigma_0}{K} \right)^{\frac{2-n}{n}} \quad (1.35)$$

Para fluidos de la ley de la potencia y newtonianos, el factor de fricción puede ser estimado directamente de la ecuación (1.31) a partir de valores de $\xi_0=0$ y $\psi=1$ donde $\sigma_0=0$. Para fluidos plástico de Bingham y Herschel-Bulkley, sin embargo, ξ_0 debe ser calculado a través de iteración de la ecuación (1.34). El intervalo de ξ_0 para fluido laminar es $1 > \xi_0 > \xi_{0c}$ donde ξ_{0c} es valor de transición laminar-turbulento de ξ_0 . Una vez que ξ_0 es conocido, el factor de fricción es calculado desde la ecuación (1.31), (1.14) y (1.32).

También a través de la manipulación algebraica de las ecuaciones (1.31), (1.14), (1.32), (1.33) y (1.35) es posible de calcular el f_f , como una función de Re_G , n y He ver ecuación (1.36).

$$f_f = \frac{16}{Re_G \left(1 - \frac{2}{f_f} \left(\frac{He^n (4n)^{2n}}{Re_G^2 (1+3n)^{2n} 8^{2(n-1)}} \right)^{\frac{1}{2-n}} \right)^{1+n} \left[\frac{1+n \left(3+2n+2 \left(\frac{2}{f_f} \left(\frac{He^n (4n)^{2n}}{Re_G^2 (1+3n)^{2n} 8^{2(n-1)}} \right)^{\frac{1}{2-n}} \right) \right)}{(1+2n)(1+n)} \left(1+n+n \left(\frac{2}{f_f} \left(\frac{He^n (4n)^{2n}}{Re_G^2 (1+3n)^{2n} 8^{2(n-1)}} \right)^{\frac{1}{2-n}} \right) \right) \right]^n} \quad (1.36)$$

b) Factor de fricción en régimen turbulento para fluidos del modelo Herschel-Bulkley.

Para el cálculo de factor de fricción en régimen turbulento consideramos las ecuaciones de Torrance (1963), las cuales desarrolló siguiendo el método de análisis de Clapp (1961), y logró una relación para fluidos Herschel-Bulkley (Garcia & Steffe, 1987).

Usando la constante de Clapp, Torrance obtuvo:

$$\frac{1}{\sqrt{f_f}} = 0.45 - \frac{2.75}{n} + \frac{1.97}{n} \ln(1 - \xi_0) + \frac{1.97}{n} \ln \left(Re_G \left(\frac{1+3n}{4n} \right)^n f_f^{1-(n/2)} \right) \quad (1.37)$$

Donde Re_G y ξ_0 son dados por la ecuación (1.14) y (1.33) respectivamente. Combinando la ecuación (1.14), (1.33) y (1.35), ξ_0 puede ser escrita como:

$$\xi_0 = \frac{16(2He)^{n/(2-n)} \left(\frac{n}{1+3n} \right)^{2n/(2-n)}}{Re_G^{2/(2-n)} f_f} \quad (1.38)$$

Entonces la ecuación (1.37) y (1.38) da f_f como una función de n , Re_G y He , tras la manipulación algebraica, ver ecuación (1.39). Además note que la ecuación (1.38) es solo una manipulación de números adimensionales y es valida para cualquier tipo de condiciones de flujo. Si f_f es sustituido dentro de la ecuación (1.38) usando la ecuación (1.31), (1.34) se obtiene la ecuación para flujo laminar.

$$f_f = \left(\frac{1}{0.45 - \frac{2.75}{n} + \frac{1.97}{n} \ln \left(1 - \frac{16(2He)^{\frac{n}{2-n}} \left(\frac{n}{1+3n} \right)^{\frac{2n}{2-n}}}{Re_G^{\frac{2}{2-n}} f_f} \right) + \frac{1.97}{n} \ln \left(Re_G \left(\frac{1+3n}{4n} \right)^n f_f^{1-(\frac{n}{2})} \right)} \right)^2 \quad (1.39)$$

1.2.2.1 Método iterativo de cálculo para fluidos Herschel- Bulkley en régimen laminar.

Suponiendo que se tiene que calcular el *factor de fricción* de un fluido con las siguientes características y condiciones:

$$n=0.6$$

$$K=4.8 \text{ Pa}\cdot\text{s}^n$$

$$\sigma_0=10 \text{ Pa}$$

$$\rho=1320.9 \text{ Kg/m}^3$$

$$v=1.84 \text{ m/s}$$

Diámetro nominal = 3 pulgadas, cedula 40

Diámetro interno= 0.0779 m

Al observar que el fluido tiene σ_0 de inmediato se determina que no es de tipo newtoniano y puesto que el valor de $n < 1$ se sabe que no se trata ni un fluido Bingham ni Casson, por lo tanto es un fluido tipo Herschel-Bulkley.

Entonces se inicia el cálculo utilizando la ecuación (1.14):

$$Re_G = \frac{(0.0779 \text{ m})^{0.6} (1.84 \text{ m/s})^{(2-0.6)} (1320.9 \text{ Kg/m}^3)}{8^{(0.6-1)} (4.8 \text{ Pa}\cdot\text{s}^{0.6})} \left(\frac{4(0.6)}{(1+3(0.6))} \right)^{0.6} = 292.66$$

Como $Re_G < 2,100$ por ende el fluido se encuentra en régimen laminar.

Utilizando la ecuación (1.35) se calcula el valor del número de Hedstrom:

$$He = \frac{(0.0779 \text{ m})^2 (1320.9 \text{ kg/m}^3)}{4.8 \text{ Pa}\cdot\text{s}^{0.6}} \left(\frac{10 \text{ Pa}}{4.8 \text{ Pa}\cdot\text{s}^{0.6}} \right)^{\frac{2-0.6}{0.6}} = 9.25708$$

Para calcular el valor de ζ_0 es necesario encontrarlo por iteración de la ecuación (1.34), usando las ecuaciones (1.14), (1.31), (1.32), (1.33) y (1.35); como se desconoce el valor de ζ_0 se supone este valor, que puede ir desde un intervalo de $0 < \zeta_0 < 1$, entonces en este caso se inicia con una suposición de 0.5 y se sustituye en la ecuación (1.32) para estimar el valor de ψ

$$\psi = (1 + 3(0.6))^{0.6} (1 - 0.5)^{1+0.6} \left[\frac{(1 - 0.5)^2}{(1 + 3(0.6))} + \frac{2(0.5)(1 - 0.5)}{(1 + 2(0.6))} + \frac{(0.5)^2}{(1 + 0.6)} \right]^{0.6} = 0.402529$$

Como el objetivo final es el cálculo de f_f y al final todos los valores deben de coincidir, en la secuencia de cálculo, entonces sustituyendo en la ecuación (1.31), se obtiene

$$f_f = \frac{16}{(0.402529)(292.660)} = 0.1358$$

y a su vez se sustituye el resultado en la ecuación (1.33)

$$\xi_0 = \frac{2(10Pa)}{(0.1358)(1320.9kg/m^3)(1.84m/s)^2} = 0.03293$$

Este nuevo valor de ξ_0 se utiliza para evaluar nuevamente el valor de ψ

$$\psi = (1 + 3(0.6))^{0.6} (1 - 0.03293)^{1+0.6} \left[\frac{(1 - 0.03293)^2}{(1 + 3(0.6))} + \frac{2(0.03293)(1 - 0.03293)}{(1 + 2(0.6))} + \frac{(0.03293)^2}{(1 + 0.6)} \right]^{0.6}$$

$$\psi = 0.95813$$

Con estos nuevos valores de ξ_0 y ψ se calcula el valor de Re_{He} con la ecuación (1.34)

$$Re_{He} = 2(9.25708) \left(\frac{0.6}{1 + 3(0.6)} \right)^2 \left(\frac{0.95813}{0.03293} \right)^{\frac{2-0.6}{0.6}} = 2211.11$$

Si la suposición es correcta entonces los valores tanto de Re_G y Re_{He} deben de coincidir, como en este caso no coinciden se vuelve a calcular el f_f con la ecuación (1.31) utilizando el nuevo valor de ψ

Paso 1:

$$f_f = \frac{16}{(0.95813)(292.660)} = 0.05706$$

Paso 2:

$$\xi_0 = \frac{2(10Pa)}{(0.05706)(1320.9kg/m^3)(1.84m/s)^2} = 0.07837$$

Paso 3:

$$\psi = (1 + 3(0.6))^{0.6} (1 - 0.07837)^{1+0.6} \left[\frac{(1 - 0.07837)^2}{(1 + 3(0.6))} + \frac{2(0.07837)(1 - 0.07837)}{(1 + 2(0.6))} + \frac{(0.07837)^2}{(1 + 0.6)} \right]^{0.6} = 0.9005$$

Paso 4:

$$Re_{He} = 2(9.25708) \left(\frac{0.6}{1 + 3(0.6)} \right)^2 \left(\frac{0.9005}{0.0783} \right)^{\frac{2-0.6}{0.6}} = 251.751$$

Como puede notarse esta vez el valor de Re_{He} es más cercano al valor de Re_G para obtener un valor mucho más cercano es necesario una iteración más, es decir repetir el procedimiento desde el paso 1 al paso 4 tomando el nuevo valor de ψ como inicio, y así se obtiene:

Paso 1:

$$f_f = \frac{16}{(0.9005)(292.660)} = 0.06071$$

Paso 2:

$$\xi_0 = \frac{2(10Pa)}{(0.06071)(1320.9kg/m^3)(1.84m/s)^2} = 0.07366$$

Paso 3:

$$\psi = (1 + 3(0.6))^{0.6} (1 - 0.07366)^{1+0.6} \left[\frac{(1 - 0.07366)^2}{(1 + 3(0.6))} + \frac{2(0.07366)(1 - 0.07366)}{(1 + 2(0.6))} + \frac{(0.07366)^2}{(1 + 0.6)} \right]^{0.6} = 0.90651$$

Paso 4:

$$Re_{He} = 2(9.25708) \left(\frac{0.6}{1 + 3(0.6)} \right)^2 \left(\frac{0.90651}{0.07366} \right)^{\frac{2-0.6}{0.6}} = 297.2709$$

Aunque este valor, ya es muy cercano a $Re_G = 292.66$, es necesaria otra iteración más para depurar el valor y/o descartar errores.

Paso 1:

$$f_f = \frac{16}{(0.90651)(292.660)} = 0.0603$$

Paso 2:

$$\xi_0 = \frac{2(10Pa)}{(0.0603)(1320.9kg/m^3)(1.84m/s)^2} = 0.07415$$

Paso 3:

$$\psi = (1 + 3(0.6))^{0.6} (1 - 0.07415)^{1+0.6} \left[\frac{(1 - 0.07415)^2}{(1 + 3(0.6))} + \frac{2(0.07415)(1 - 0.07415)}{(1 + 2(0.6))} + \frac{(0.07415)^2}{(1 + 0.6)} \right]^{0.6} = 0.90589$$

Paso 4:

$$Re_{He} = 2(9.25708) \left(\frac{0.6}{1 + 3(0.6)} \right)^2 \left(\frac{0.90589}{0.07415} \right)^{\frac{2-0.6}{0.6}} = 292.199$$

Los nuevos valores de $\xi_0 = 0.7415$ y $\psi = \mathbf{0.90589}$, dan como resultado un valor de $Re_{He} = 292.199$ que es un valor muy cercano a $Re_G = \mathbf{292.66}$, existe una diferencia de apenas 0.15% entre Re_G y Re_{He} , por lo tanto es un valor aceptable pues entra en el intervalo de tolerancia generalmente aceptado, pues estas diferencias pueden deberse sobre todo al número de decimales que se maneja en cada valor de la secuencia de cálculo.

Sustituyendo el último valor aceptable de ψ en la ecuación (1.31), se obtiene finalmente el valor de f_f :

$$f_f = \frac{16}{(0.90589)(292.660)} = 0.06035$$

Como puede notarse el proceso de iteración puede ser engorroso, repetitivo, y con un considerable consumo de tiempo, y por supuesto esto puede variar dependiendo que tan cercana sea la primera suposición del proceso iterativo.

1.3 Las redes neuronales artificiales: definición, fundamentos y generalidades.

En esta sección, se explicará que son y como funcionan las Redes neuronales artificiales, desde sus fundamentos hasta las aplicaciones realizadas en la industria alimentaria.

1.3.1 Definición.

Las **redes neuronales artificiales** son un sistema procesador basado en la estructura del sistema neuronal biológico; son básicamente modelos computacionales, los cuales simulan habilidades muy simples de nuestro cerebro (Mittal & Zhang, 1991; Afaghi *et al.*, 2001).

Fueron inspiradas en la arquitectura de los sistemas nerviosos biológicos y en la analogía matemática; la unidad básica de procesamiento es similar a las neuronas de los sistemas biológicos y es conocida como **nodo ó neurona artificial**, una red neuronal artificial es un conjunto de neuronas artificiales interconectadas masivamente por sinapsis artificiales (simbolizadas por una matriz de números) y con organización jerárquica que funcionan en paralelo para la solución de problemas específicos (Luera Peña & Minim, 2001); a través de modelos matemáticos, tienen la capacidad de simular propiedades como: aprendizaje, organización, reconocimiento, procesamiento y reproducción de información.

La definición formal de un red neuronal artificial, esta basada en el concepto matemático de **grafo**¹⁴, una red neuronal artificial es un *grafo dirigido*¹⁵ (Figura 1.11), con las siguientes propiedades (Müller & Reinhardt 1990):

1. A cada nodo i se asocia una variable de estado a_i .
2. A cada conexión (i, j) de los nodos i y j se asocia un peso $w_{ij} \in \mathfrak{R}$.
3. A cada nodo i se asocia un umbral θ_i .

¹⁴ Objeto consistente en un conjunto de nodos (vértices), más un conjunto de conexiones establecidas entre ellos. Típicamente, un grafo se representa gráficamente como un conjunto de puntos (vértices o nodos) unidos por líneas (aristas), otra forma de representación es mediante una matriz de conexiones, un algoritmo, etc. Los grafos permiten estudiar las interrelaciones entre unidades que interactúan unas con otras (Martín & Sanz, 1997; <http://es.wikipedia.org/wiki/Grafo>, 2015).

¹⁵ Se le llama dirigido, cuando las conexiones tienen asignado un sentido.

4. Para cada nodo i se define una función $f_{act\ i}(a_j, w_{ij}, \theta_i)$, que depende de los pesos de sus conexiones, del umbral y de los estados de los nodos j a él conectados. Esta función proporciona el nuevo estado del nodo.

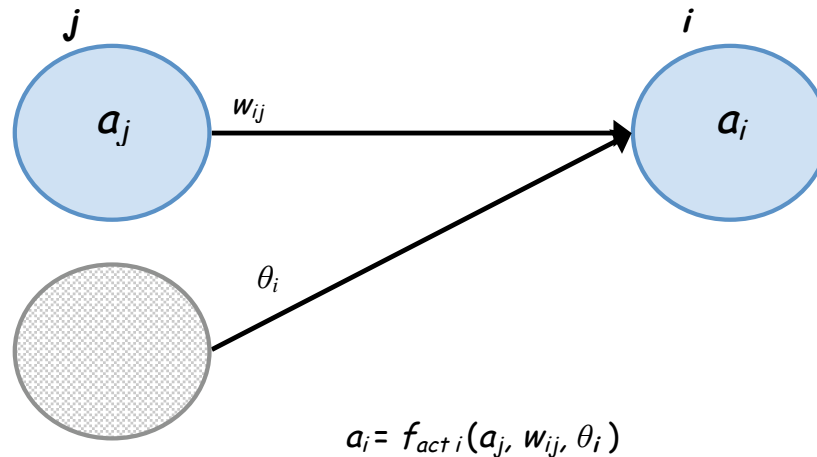


Figura 1.11 Una red neuronal artificial como un grafo dirigido y sus propiedades.

1.3.2 Antecedentes históricos y fundamentos.

Las redes neuronales artificiales, representan una tecnología que tiene su origen en muchas disciplinas: neurociencias, matemáticas, estadística, física, ciencia computacional e ingeniería. Para comprender más claramente su origen es necesario un breve repaso sobre los acontecimientos más sobresalientes en la historia de esta tecnología.

1.3.2.1 Antecedentes históricos.

A lo largo de la historia de la humanidad el conseguir diseñar y construir máquinas capaces de realizar procesos con cierta inteligencia ha sido uno de los principales objetivos y preocupaciones de los científicos, la mayoría de estos intentos no han sido con el afán de que compitan con el ser humano si no que realicen tareas de cierto rango intelectual que sean de ayuda al ser humano. Y así cada uno de los múltiples intentos a lo largo de la historia humana ha dado un pequeño paso para alcanzar este fin.

La era moderna de las redes neuronales comenzó con el trabajo pionero de Warren McCulloch (psiquiatra y neuroanatomista) y Walter Pitts (matemático) en 1943, quienes lanzaron una teoría acerca de la forma de como trabajan las neuronas y como estas podían aprender, el modelo formal fue asumido y posteriormente seguido de la ley “todo o nada”, en esta exponen la creación de un modelo abstracto de neurona, en el cual la probabilidad de que una neurona se active dependía de la señal de entrada a esta y de la sinapsis de conexión.

En 1949 Donald Hebb publica su libro *La Organización del Comportamiento*, donde describe un método por el cual las neuronas podrían aprender, cuando la entrada desde la sinapsis a la neurona y la salida de esta fuesen ambas activadas, entonces la fuerza sináptica de la conexión era reforzada.

En 1958 Frank Rosenblatt (Psicólogo) en su intento de ilustrar algunas de las propiedades fundamentales de los sistemas inteligentes, creía (a diferencia de McCulloch-Pitts) que la herramienta de análisis más apropiada era la teoría de probabilidades, y desarrolló una teoría de separabilidad estadística que utilizó para caracterizar las propiedades más visibles de redes neuronales de interconexión aleatoria. Gracias a esto realizó el *Perceptron*, que es el modelo más antiguo de una red neuronal, y que aun se usa hoy en día en varias formas para la aplicación de reconocimiento de patrones. Este modelo fue capaz de generalizar es decir, después de haber aprendido una serie de patrones era capaz de reconocer otros similares, aunque no se les hubiera presentado anteriormente.

En 1960 Bernard Widrow y Marcial Hoff, de Stanford, desarrollaron el *Algoritmo de aprendizaje de mínimos cuadrados “LMS”* (Least Mean-Square) y usaron este para crear *Adaline* (Adaptive Linear Element), utilizaron las matemáticas de procesamiento de señal adaptativa para producir la primer red neuronal comercial. La diferencia de entre el Perceptron y la Adaline radica en procedimiento de entrenamiento.

Una de las primeras redes neuronales con capas con múltiples elementos adaptativos fue la *Madaline* (multiple-adaline), estructura propuesta por Widrow y sus estudiantes en 1962.

Durante el periodo clásico de las perceptron en los 1960's estas redes neuronales fueron vistas como si podrían realizar cualquier cosa. Pero tras la publicación del libro de Marvin Minsky y

Seymour Papert, del Instituto Tecnológico de Massachusetts (MIT) en 1969, donde se demostró matemáticamente que hay límites fundamentales¹⁶ sobre cuales las perceptrons de una capa pueden computar y por lo tanto consideraban que la extensión a perceptrons multicapa era completamente estéril. Como las limitaciones eran importantes, y sobre todo había una incapacidad para resolver problemas más complejos, gran parte de las investigaciones dieron un giro hacia la inteligencia artificial, que prometía más en aquel entonces.

Estas críticas junto con otras razones, como la falta de financiamiento y las limitaciones tecnológicas (por ejemplo que no había computadoras personales ó estaciones de trabajo para la experimentación); frenaron hasta 1980's el crecimiento que estaban experimentando las investigaciones sobre redes neuronales artificiales. Esa fue la razón principal que en 1980's las mayores contribuciones a la teoría y diseño de redes neuronales fueron hechas desde varios frentes que hicieron resurgir el interés por las redes neuronales.

En 1980 Stephen Grossberg, retomó su trabajo del aprendizaje competitivo¹⁷, y estableció un nuevo principio de auto-organización conocido como *teoría de resonancia adaptativa "ART"* (Adaptative resonance theory), por medio de la cual la arquitectura de la red puede simular habilidades del cerebro como la memoria a corto y a largo plazo (un ejemplo de esto es la amplificación y prolongación de la actividad neuronal).

En 1982 John Hopfield presentó su trabajo sobre redes neuronales en la Academia Nacional de Ciencia, en el describe con claridad y rigor matemático una red, que es una variación del asociador lineal (un modelo lineal que consistía en unos elementos integradores lineales que sumaban sus entradas), pero además mostró como pueden trabajar las redes y que pueden hacer. Hopfield uso la idea de función de energía de activación para formular un nuevo camino para entender el desempeño de la computación de las redes recurrentes con conexiones sinápticas simétricas. Además él estableció el isomorfismo tal entre una red recurrente y un *modelo Ising*¹⁸ utilizado en la estadística física. Esta analogía mostró el camino para un diluvio de teorías físicas

¹⁶ Esto se explica con el Teorema de aproximación universal (ver más adelante en el texto).

¹⁷ Es un algoritmo de aprendizaje, que como su nombre lo indica las neuronas de salida de una red neuronal compiten entre ellas para llegar a activarse, en este aprendizaje solo una sola neurona es activada en un solo tiempo.

¹⁸ Modelo físico propuesto para estudiar el comportamiento de partículas con espín que pueden adoptar únicamente dos estados hacia arriba ó hacia abajo (ver glosario).

(y físicos) se adentraron a la modelación neuronal, y por lo tanto transformaron el campo de las redes neuronales.

Además en 1982 se celebró la U.S.-Japan *Joint Conference on Cooperative Neural Networks*, a este tipo de reuniones siguieron otras como la reunión anual de *Neural Networks for Computing* en 1985, esta fue la primera de muchas otras posteriormente. Ese año también se formó la *International Neural Networks Society (INNS)*. Desde entonces el interés por esta área de la investigación y el desarrollo se ha ido incrementando de forma notable, como lo demuestran el gran número de congresos y reuniones así como las revistas y otras publicaciones científicas sobre el tema.

Las redes neuronales artificiales tienen ya un largo camino desde sus tempranos días de McCulloch y Pitts. Además se lograron establecer como un sujeto interdisciplinario con profundas raíces en las neurociencias, psicología, matemáticas y las ciencias físicas e ingeniería. Sin duda continuarán creciendo en teoría, diseño y aplicaciones.

1.3.2.2 Fundamentos.

La idea principal que dio origen a la estructura de las Redes Neuronales Artificiales (RNA)¹⁹ es que al imitar la estructura de las redes neuronales biológicas (el hardware del sistema nervioso), se desea construir sistemas de procesamiento de información que puedan presentar un cierto comportamiento *inteligente*, y los conceptos clave de los sistemas nerviosos biológicos que se pretende emular son: *adaptabilidad* al entorno, *memoria distribuida* y *paralelismo de cálculo* (Martín & Sanz, 1997).

Las RNA se consideran como **sistemas adaptativos**, porque en lugar de ser programadas para usarse en una aplicación particular; los modelos de las RNA generan sus propias reglas a través de un proceso de aprendizaje a partir de ejemplos y difieren de los programas convencionales en su habilidad para aprender acerca del sistema que está modelando, esto aun sin poseer un conocimiento previo de la naturaleza de la relación entre las variables de proceso, pero con capacidad de reaprender de acuerdo a nuevos datos (Mittal & Zhan, 2000; Afaghi *et al.*, 2001).

¹⁹ En inglés ANN (Artificial Neural Networks), además existen otras denominaciones, como sistemas neuronales artificiales o ANS (Artificial Neural Systems), neurocomputación (Neural computing) o sistemas conexionistas.

Por ello pueden adaptarse a cambios; y como las personas, aprender de la experiencia. También son capaces de tratar con relaciones no-lineales, que les brinda más ventajas comparativas con respecto a los modelos empíricos.

Otra característica es que las RNA son los primeros métodos computacionales con la capacidad inherente a la tolerancia a fallos, comparados con los sistemas computacionales tradicionales. Cuentan con dos aspectos distintos a la tolerancia a fallos: primero, las RNA pueden aprender a reconocer patrones con ruido²⁰, distorsionados, incompletos ó con inconsistencia en los datos, esta es una tolerancia a fallos con respecto a los datos y segundo si en las RNA se produce un fallo en un pequeño número de neuronas, aunque el comportamiento del sistema se vea influenciado, puede seguir realizando su función (con cierta degradación), la RNA no sufrirá una caída repentina; debido principalmente a que en los sistemas neuronales la información se encuentra distribuida por las sinapsis de la red, donde las fuerzas de conexión interneuronal (pesos sinápticos), son usadas para almacenar el conocimiento adquirido, a esto se le conoce como **memoria distribuida** (Haykin, 1999; Hiler & Martínez, 1995).

También se les considera un **sistema de procesado en paralelo** ya que normalmente se usa un gran número de neuronas de procesado en su estructura (similar a lo que serian las neuronas en un cerebro biológico) para un alto nivel de interconectividad y que funcionan al mismo tiempo, lo que las hace potencialmente rápidas para el computo de ciertas tareas.

Pero a pesar que el diseño de las RNA imita la estructura del sistema neuronal biológico para reproducir algunas de sus capacidades, las RNA dirigidas a una aplicación están en poco ligadas a las redes neuronales biológicas, por lo consiguiente se han de definir otras funcionalidades y estructuras de conexión, distintas a las de la perspectiva biológica.

²⁰ Cuando se habla de *ruido* por lo general se hace referencia que existe una sensible diferencia entre la señal original que produce la célula y la señal que se percibe a través de determinados sensores (ver glosario).

1.3.3 Estructura del sistema neuronal artificial.

En los sistemas neuronales biológicos, las neuronas son los elementos básicos, estas se agrupan en conjuntos compuestos por millones de ellas organizadas en capas, constituyendo así un sistema con funcionalidad propia. De manera similar en una RNA el elemento esencial de partida será la neurona artificial, que se organizará en capas; varias capas constituirán una RNA y por último una red neuronal (o un conjunto de ellas), junto con las interfaces de entrada y de salida, más los módulos convencionales adicionales necesarios, constituirán el sistema global de proceso. Formalmente un sistema neuronal artificial esta compuesto por los siguientes elementos:

- Un conjunto de neuronas artificiales.
- Un patrón de conectividad ó arquitectura.
- Una dinámica de activaciones.
- Una regla de aprendizaje.
- El entorno donde opera.

Es necesario para su mejor comprensión, describir a las RNA en términos de sus niveles estructurales, y examinar su estructura a través de su funcionamiento en cada nivel (Maren,1990; Hilera & Martinez, 1995):

1. Micro-estructura.-Se consideran en concreto las operaciones de una típica neurona artificial, así como cuales son las características de cada neurona en la red.
2. Meso-estructura.- Corresponde a la organización física y arreglo de las neuronas en la red, es decir el número de capas que tiene; así como lo relacionado a la función de la red (refiriéndose al conjunto de neuronas como tal), explicando cuales son los patrones de conexión y como es el flujo de información.
3. Macro-estructura.- Se refiere principalmente a la manera de cómo diferentes redes pueden conjuntamente realizar diferentes tareas ó dar dirección a problemas de gran escala.

A continuación se explicará el funcionamiento de los dos primeros niveles estructurales de las RNA, ya que es fundamental para su comprensión; respecto al funcionamiento de la macro-estructura, no se profundizará más, pero si se desea consulte Maren, (1990).

1.3.3.1 Micro-estructura, unidades de proceso: la neurona artificial.

Se denomina **neurona artificial**²¹ a un dispositivo simple de cálculo que, a partir de un vector de entrada procedente del exterior ó de otras neuronas, proporciona una única respuesta ó salida (Martín & Sanz, 1997). Existen tres tipos de neuronas (Figura 1.12):

- a) **Neurona de entrada**, son aquellas neuronas sin sinapsis (conexiones) entrantes, lo que significa que estas reciben estímulos externos del entorno y toman la información de entrada que pueden ser señales provenientes de sensores o de otros sectores del sistema (estas son en realidad la entrada a la red).
- b) **Neurona de salida**, son aquellas neuronas sin conexiones salientes, una vez que ha finalizado el periodo de procesado la información llega a las o la neurona de salida la cual envía la señal fuera del sistema, estas señales pueden controlar directamente otros sistemas.
- c) **Neurona oculta**, son los elementos internos por los cuales la información de entrada se transmite, para su procesado. Es en las sinapsis y neuronas correspondientes a este nivel donde se genera cualquier tipo de representación interna de información, estas neuronas no tienen contacto con el exterior.

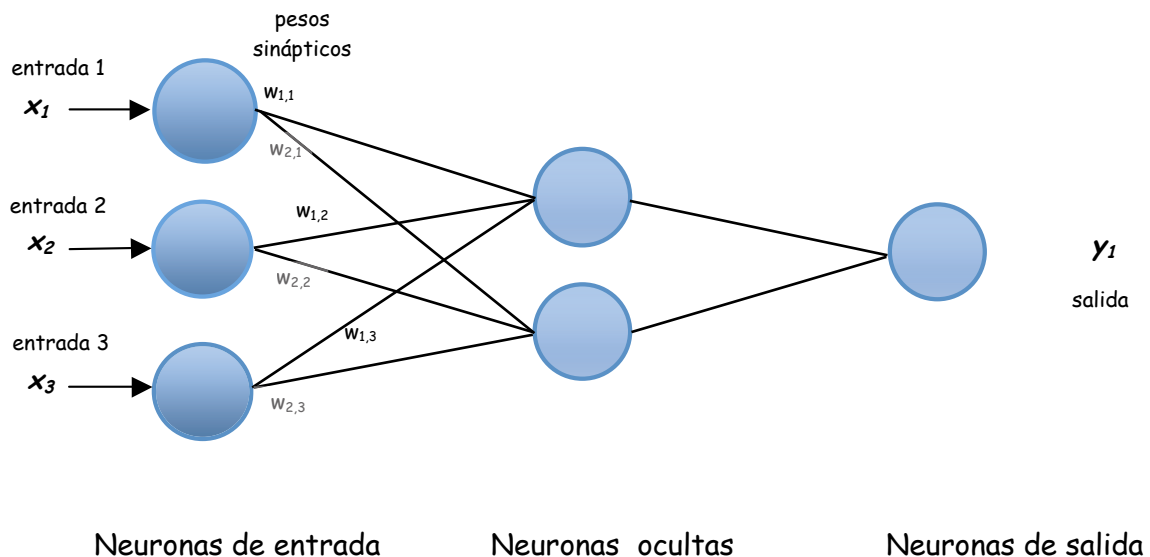


Figura 1.12 Diagrama de la estructura de una RNA con tres capas (entrada-oculta-salida).

²¹ También llamado procesador elemental ó nodo.

1.3.3.1.1 El modelo genérico de la neurona artificial.

Cada neurona j -ésima esta caracterizada en cualquier instante por un valor numérico denominado *estado de activación* $a_j(t)$; asociado a cada unidad, existe una *función de salida*, F_{sj} , que transforma el estado actual de activación en una *señal de salida*, y_j . Dicha señal es enviada a través de los canales de comunicación a otras unidades de la red; en estos canales la señal se modifica de acuerdo con la conexión (*el peso*, w_{ij}) asociada a cada uno de ellos según determinada regla. Las señales moduladas que han llegado a la unidad i -ésima se combinan entre ellas, generando así la *entrada total ó entrada neta*²², net_i .

$$net_i = \sum_j y_j w_{ij} \quad (1.40)$$

Ahí una *función de activación* f_{act} , determina el nuevo estado de activación $a_i(t)$ de la *neurona* i -ésima, teniendo en cuenta la entrada total calculada y el anterior estado de activación $a_i(t-1)$. A partir de este momento será necesario que el proceso vuelva a comenzar en la unidad i -ésima (Figura 1.13).

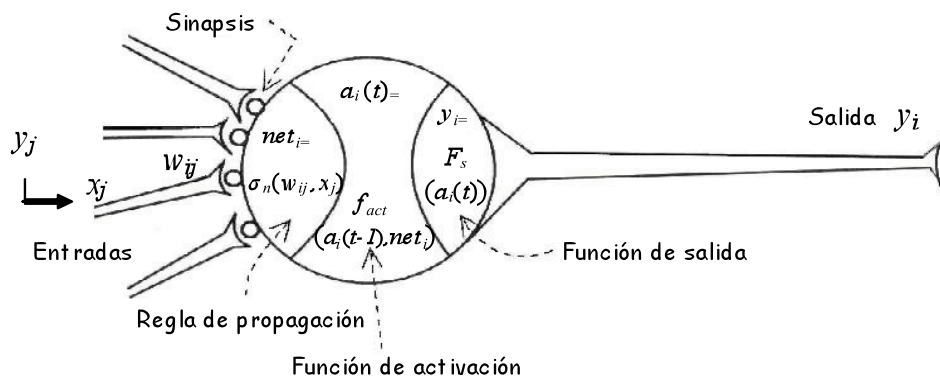


Figura 1.13 Modelo genérico de una neurona artificial (Martín & Sanz, 1997).

Los elementos que constituyen la neurona de etiqueta i son (Martín & Sanz, 1997):

- Conjunto de entradas, $x_j(t)$.

²² Aunque esta es la manera más usual de llamarla, su nombre correcto es el de potencial postsináptico, que se designa en este caso como, h_i , y representa el nivel de excitación de la neurona.

- Pesos sinápticos de la neurona i , w_{ij} que representan la intensidad de interacción entre cada neurona presináptica j y la neurona postsináptica i .
- Regla de propagación $\sigma_n(w_{ij}, x_j(t))$, que proporciona el valor de la entrada neta $net_i(t) = \sigma_n(w_{ij}, x_j(t))$ de la neurona i en función de sus pesos y entradas.
- Función de activación $f_{act\ i}(a_i(t-1), net_i(t))$, que proporciona el estado de activación actual $a_i(t) = f_{act\ i}(a_i(t-1), net_i(t))$ de la neurona i en función de su estado anterior $a_i(t-1)$ y de su entrada neta actual.
- Función de salida $F_{s\ i}(a_i(t))$, que proporciona la salida actual $y_i(t) = F_{s\ i}(a_i(t))$ de la neurona i en función de su estado de activación.

De este modo, la operación de la neurona i puede expresarse como:

$$y_i(t) = F_{s\ i} \left(f_{act\ i} \left[a_i(t-1), \sigma_{n\ i} \left(w_{ij}, x_j(t) \right) \right] \right) \quad (1.41)$$

Entradas y salidas

Las variables de entrada y salida pueden ser binarias (digitales) ó continuas (analógicas), dependiendo del modelo y aplicación. La neurona tiene muchas entradas pero solo tiene una salida (ver Figura 1.12), que a su vez se puede aplicar a muchas otras neuronas de la red. La entrada que recibe la neurona i -ésima procedente de la j -ésima neurona, se indica como x_j (obsérvese que este valor es también la salida del j -ésimo nodo²³; del mismo modo que la salida generada por el i -ésimo nodo se denota como x_i).

Regla de propagación

Una neurona recibe un conjunto de señales que le dan información del estado de activación de todas las neuronas a las cuales esta conectada, cada conexión (Figura 1.12) entre la neurona i y la neurona j esta ponderada por un peso w_{ij} . La regla de propagación²⁴ muestra el procedimiento a seguir para combinar los valores de entrada (x_j) a una neurona con los pesos de las conexiones

²³ El cual también es común que sea denotado como y_j

²⁴ También se le puede denominar, Función de combinación.

que llegan a esa unidad. Esta regla permite obtener, a partir de las entradas y los pesos, la entrada neta (net_i):

$$net_i(t) = \sigma_{n_i}(w_{ij}, x_j(t)) \quad (1.42)$$

generalmente, la función²⁵ es de tipo *lineal*, y se basa en *la suma ponderada* de las entradas con los pesos sinápticos, considerando que el efecto de cada señal es aditivo, de tal forma que la entrada neta que recibe una neurona (net_i) es la suma del producto de cada señal individual por el valor de la sinapsis que conecta ambas neuronas:

$$net_i(t) = \sum_j w_{ij} x_j \quad (1.43)$$

El **peso sináptico** w_{ij} , es el valor asociado a las conexiones que unen a las neuronas que forman una RNA, este es el que permite que la red adquiera conocimiento y define la intensidad de interacción entre la neurona presináptica j y la postsináptica i . Si el peso es positivo tenderá a excitar a la neurona postsináptica, si el peso es negativo tenderá a inhibirla. Por ende existen conexiones excitadoras e inhibitoras.

Formalmente también puede interpretarse net_i , como el *producto escalar* de los vectores de entradas y pesos (Hilera & Martínez):

$$net_i(t) = \sum_j w_{ij} x_j = \mathbf{w} \cdot \mathbf{x} \quad (1.44)$$

Por ejemplo, el net en un tiempo determinado se calcula con el producto punto (producto escalar) de entradas y pesos, expresados como un vector \mathbf{X} a los elementos de entrada (x), y como una matriz \mathbf{W} a los pesos de las conexiones entre neuronas.

$$net_i = \begin{bmatrix} w_{1,1} & w_{1,2} \dots & w_{1,R} \\ w_{2,1} & w_{2,2} \dots & w_{2,R} \\ \dots & \dots & \dots \\ w_{S,1} & w_{S,2} \dots & w_{S,R} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \dots \\ x_n \end{bmatrix} \quad (1.45)$$

²⁵ Puesto que también existen varias reglas de propagación ó funciones de combinación, del tipo no lineal (ver Barro *et al.*,1995).

En este caso, el subíndice de la fila en el elemento de la matriz \mathbf{W} indica el peso de la neurona destino, y el subíndice de la columna (segundo subíndice) indica cual es el origen del peso de entrada (input) ver Figura 1.12, así que el subíndice en $w_{1,2}$ dice la fuerza de señal desde el segundo elemento de entrada a la primera neurona.

Función de activación ó función de transferencia.

Todas las neuronas que componen la red se hallan en cierto estado. Básicamente podemos decir que hay dos posibles estados, *reposo* y *excitado*, a los que se les domina globalmente *estados de activación*, y a los cuales se les asignan un valor. El conjunto de los valores²⁶ de activación de las neuronas de una RNA, indica lo que esta representa en un momento dado. La regla que combina las entradas con el estado actual de la neurona para producir un nuevo estado de activación, es la *función de activación*, la cual proporciona el estado de activación actual $a_i(t)$ a partir del potencial postsináptico $net_i(t)$ y del propio estado de activación anterior $a_i(t-1)$.

$$a_i(t) = f_{act_i}(a_i(t-1), net_i(t)) \quad (1.46)$$

En algunos modelos de RNA, cuando la f_{act} (función de activación) es la *función identidad* se considera que el estado actual de la neurona no depende de su estado anterior, sino únicamente del actual puesto que el estado de activación de una neurona en t coincidirá con el net de la misma en $t-1$.

$$a_i(t) = f_{act_i}(net_i(t)) \quad (1.47)$$

En este caso, el parámetro que se le pasa a la función de salida, F_s , de la neurona será directamente el net (potencial postsináptico). El estado de activación anterior no se tiene en cuenta, entonces la función de activación también es llamada *función de transferencia*.²⁷ Tanto la función de activación, f_{act_i} , como la función de salida F_{si} , pueden tomar distintas formas, como por ejemplo una función identidad, $f(x)=x$, sin embargo generalmente son funciones umbrales, de tal forma que un elemento de proceso afecta a otro solo si el nivel de actividad del primero supera un cierto valor.

²⁶ Estos valores de activación pueden ser continuos, tomando valores reales cualesquiera ó valores dentro de un intervalo cerrado; también pueden ser discretos, generalmente binarios.

²⁷ Algunos autores denominan a la función de activación ó transferencia indistintamente, además suelen considerarla como función única, que establece tanto el estado de activación así como el valor de salida; cuando la función de salida obviamente es de tipo lineal.

Además es importante señalar que algunas arquitecturas neuronales especifican una función umbral la cual es adicionada a la suma de señales de entrada que llegan a la neurona. Este tipo de umbral es parecido a una *tendencia* (en inglés **bias**) este valor se denota por θ_i y representa el *umbral de activación*²⁸ de la neurona i (Figura 1.17). Este actúa parecido a una neurona más en la capa anterior con una salida constante el cual es conectado al nodo de interés, este usualmente es un valor ajustable; este nodo imaginario que provee la tendencia no tiene relación a otros nodos en la red. Además da un significado de adicionar un valor constante²⁹, a la suma de entradas, en forma de una entrada externa adicional, cuyo fin es el de poder aumentar o disminuir el umbral de excitación de la neurona dependiendo de que si es un valor positivo ó negativo, respectivamente y esto puede ser usado para adaptar el promedio de entradas dentro de un rango útil. Este θ_i (bias), se resta de la entrada neta, por lo que argumento de la función de activación quedaría (si la regla de propagación es la suma ponderada) de la manera siguiente:

$$f_{acti}(\sum_{j=1}^n w_{ij}x_j - \theta_i) \quad (1.48)$$

La función de activación establece límites en el intervalo de la salida de la neurona, existen diversas funciones de activación, las más utilizadas son cinco funciones de transferencia, que determinan distintos tipo de neuronas: función identidad, escalón, lineal a tramos (mixta), sigmoideal y gaussiana. Con la forma $y=f(x)$ de las funciones de activación en las RNA (donde x es la entrada neta y el estado de activación es y).

a) Función identidad o lineal.

Corresponde a la expresión $f(x)=x$, el conjunto de estados E puede contener cualquier número real ($y_i \in \mathfrak{R}$); el estado de activación coincide con la entrada total que a llegado a la neurona (Figura 1.14 a).

$$y_i(t) = net_i - \theta_i = f_{acti}(\sum_j w_{ij}x_j - \theta_i) \quad (1.49)$$

²⁸ Formalmente es una función umbral de pre-transferencia, también es denominado como valor umbral interno, tendencia (bias), polarización ó peso de sesgo.

²⁹ Es habitual el considerar la polarización de θ_i , como el peso sináptico de una entrada adicional cuyo valor de entrada siempre es -1, es decir $w_{i0} \equiv \theta_i$ y $x_0 \equiv -1$. Aunque algunos autores consideren que el valor de entrada de la neurona ficticia es la unidad, con un peso de valor $-\theta_i$.

b) Función de activación escalón³⁰

La forma más fácil de definir la activación de una neurona es considerar que es binaria. La función de transferencia escalón se asocia a neuronas binarias en las cuales cuando la suma de las entradas es mayor ó igual que el umbral de la neurona, la activación es 1; si es menor la activación es 0 (ó -1). Ejemplos de estas funciones son la función signo, $sign(x)$ y la función escalón estándar, $\theta(x)$ que son definidas como (Hilera & Martínez, 1995):

Si el conjunto de los estados de activación es $E = [0,1]$, la función es escalón estándar, también conocida como función unitaria de Heaviside $H(x)$ ver Figura 1.14b.1, donde:

$$y_i(t) = \begin{cases} 1 & \text{si } (net_i \geq \theta_i) \\ y_i(t-1) & \text{si } net_i = \theta_i \\ 0 & \text{si } (net_i < \theta_i) \end{cases} \quad (1.50)$$

Y será función signo (Figura 1.14b.2), si el conjunto es $E = [-1,1]$, tendremos que:

$$y_i(t) = \begin{cases} +1 & \text{si } (net_i \geq \theta_i) \\ y_i(t-1) & \text{si } net_i = \theta_i \\ -1 & \text{si } (net_i < \theta_i) \end{cases} \quad (1.51)$$

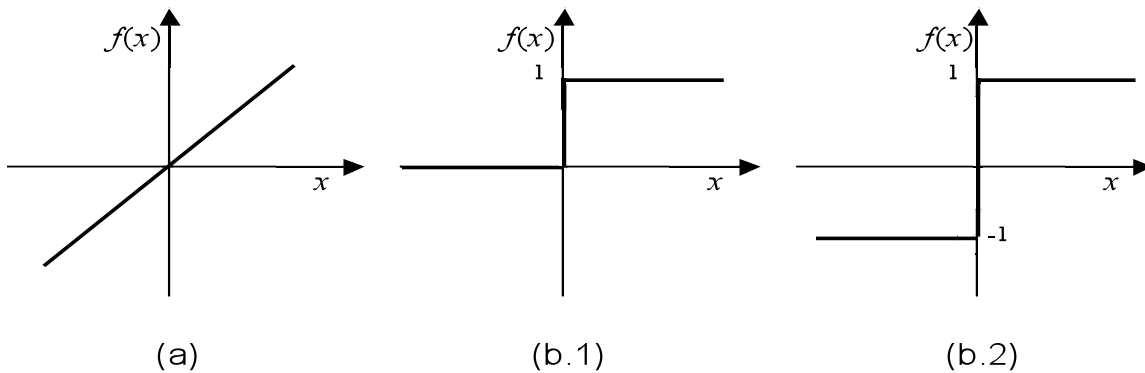


Figura 1.14 Funciones de activación (a) Lineal y (b) Escalón (en este caso se considera que el umbral es cero, en el caso de que no lo fuera el escalón quedaría desplazado).

³⁰ También es conocida como función umbral, función de lógico umbral ó función de paso.

c) Función de activación mixta o lineal a tramos.

En las neuronas con función mixta, si la suma de las señales de entrada es menor que un límite inferior, la activación se define como 0 (ó -1). Si dicha suma es mayor ó igual que el límite superior, entonces la activación es 1. Si la suma de entrada esta comprendida entre ambos limites, superior e inferior, entonces la activación se define como una función lineal de la suma de las señales de entrada. En esta función el estado de activación de la unidad está obligado a permanecer dentro de un intervalo de valores reales prefijados (Hilera & Martínez, 1995).

Con intervalo de [0,1] ver Figura 1.15a

$$y_i(t) = \begin{cases} 1 & \text{si } net \geq c \\ net/(2c) + 1/2 & \text{si } -c < net < c \\ 0 & \text{si } net \leq -c \end{cases} \quad (1.52)$$

Donde se toma el límite superior de la suma de todas las entradas de activación que afectan a la neurona durante el ciclo de operación (x) como c y el límite inferior como $-c$ (en el valor de ambos limites se suma inherentemente el valor de umbral interno θ_i).

Con intervalo de [-1,1], ver Figura 1.15b

$$y_i(t) = \begin{cases} -1 & \text{si } net < -c \\ 1 & \text{si } net > c \\ a \cdot net & \text{si } -c < net < c \end{cases} \quad (1.53)$$

Donde el parámetro a determina la pendiente de la recta definida en el tramo $-c < net < c$.

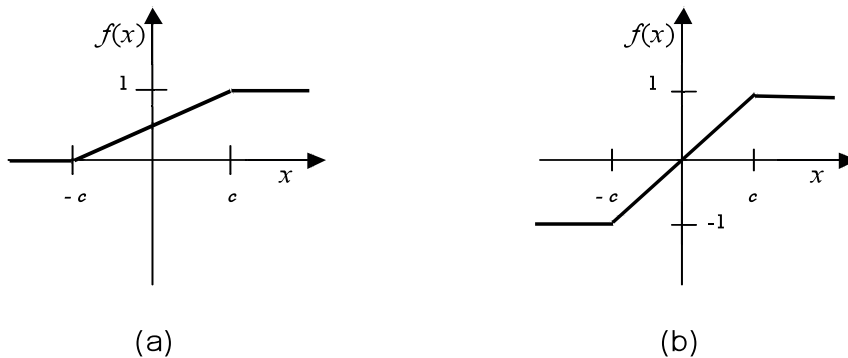


Figura 1.15 Función de activación lineal a tramos (mixta).

d) Función sigmoideal.

Es una función continua, por lo tanto el espacio de los estados de activación es un intervalo del eje real, da una salida gradual no lineal para las entradas, y a su vez representa un equilibrio adecuado entre el comportamiento lineal y el no lineal. Es la función más utilizada debido a que el intervalo de salida es continuo y aunque acotado es derivable, su derivada es siempre positiva y cercana a cero para los valores grandes positivos o negativos de x ; además toma su valor máximo cuando x es 0. Con la función sigmoideal para la mayoría de los valores del estímulo de entrada (variable independiente) el valor dado por la función es cercano a uno de los valores asintóticos. Esto hace que en la mayoría de los casos, el valor de salida este comprendido en la zona alta ó baja del sigmoide. De hecho cuando la pendiente es elevada, esta función tiende a la función escalón. Esto facilita que se puedan utilizar las reglas de aprendizaje definidas para la función escalón, con la ventaja de que la derivada está definida en todo el intervalo. Las funciones sigmoideales más populares son (Hilera & Martínez, 1995; Martín & Sanz, 1997):

- Función logística (Figura 1.16a.1) con un intervalo de [0,1]

$$y_i(t) = \frac{1}{(1 + e^{-(b \cdot net_i)})} \quad (1.54)$$

Donde b determina la pendiente de la función en el origen. Este valor define su forma, de modo que sí b tiende hacia infinito la función logística tiende a la función escalón.

- Función tangente hiperbólica (Figura 1.16a.2) con un intervalo de [-1,1]

$$y_i(t) = \tanh(b \cdot net_i) = \frac{e^{b \cdot net_i} - e^{-b \cdot net_i}}{e^{b \cdot net_i} + e^{-b \cdot net_i}} \quad (1.55)$$

Cuando b tiende hacia infinito, la función tangente hiperbólica tiende a la función signo.

e) Función de base radial (Función gaussiana).

Las funciones de base radial, son típicamente una función Gaussiana, esta es útil cuando se crea una RNA para funciones continuas de mapeo. Puesto que los centros y anchura de estas funciones pueden ser adaptados, lo cual la hace más adaptativa que las funciones sigmoideales (las cuales tiene una tendencia adicionada a ellas). Esta función suele utilizarse junto con otras reglas de propagación diferentes a la regla lineal de la suma ponderada (Hilera & Martínez, 1995).

Función gaussiana, donde el intervalo es de $[0,1]$:

$$y_i(t) = A_g e^{-B \cdot x^2} \quad (1.56)$$

Donde A_g y B son constantes, A_g determina el valor máximo del centro de la campana, que en este caso es 1, y B determina la pendiente de la función (Figura 1.16b).

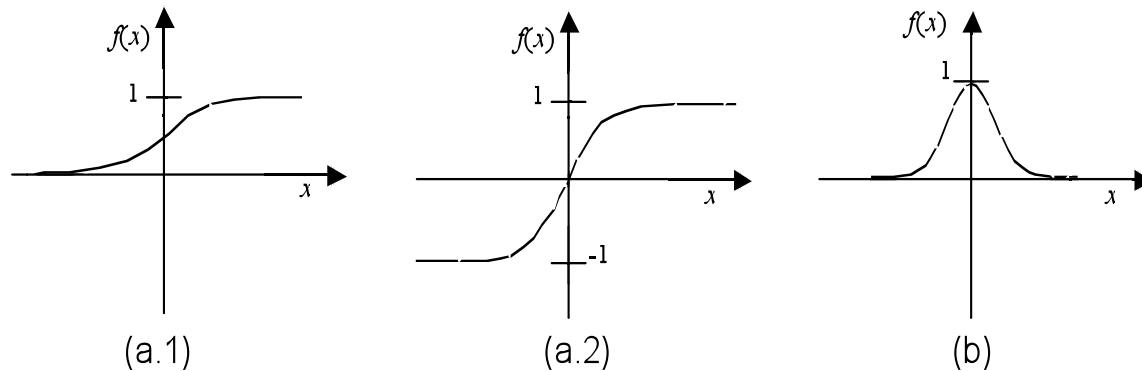


Figura 1.16 Funciones de activación (a) Sigmoidales y (b) Gaussiana.

Función de salida.

Cada unidad transmite señales a aquellas que están conectadas con su salida. Asociada con cada unidad U_i hay una función de salida $F_{si}(a_i(t))$, que transforma el estado actual de activación $a_i(t)$ en una señal de salida $y_i(t)$, es decir; esta función proporciona la salida global de la neurona $y_i(t)$ en función de su estado de activación actual $a_i(t)$. Muy frecuentemente la función de salida es simplemente la identidad $f(x)=x$ (Figura 1.17), de modo que el estado de activación de la neurona se considera como la propia salida (Hilera & Martinez, 1995):

$$y_i(t) = F_{si}(a_i(t)) = a_i(t) \quad (1.57)$$

Esto ocurre en los modelos más comunes, como el *Multilayer perceptron* “MLP” (Perceptron multicapa) ó la *Adaline*; la función identidad ó lineal equivale a no aplicar función de salida. Sin embargo la función de salida puede ser también de tipo escalón, lo que supone que la neurona no se dispara hasta que la activación supere un cierto umbral, la función escalón únicamente se

utiliza cuando las salidas de la red son binarias (dos posibles valores), las funciones mixta y sigmoideal son mas apropiadas cuando queremos como salida información analógica.

Σ

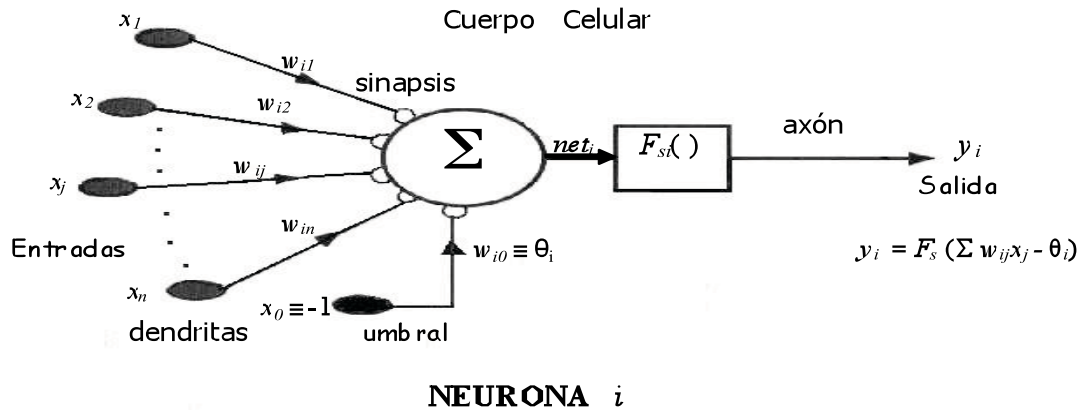


Figura 1.17 Modelo de la función procesadora de una única neurona de tipo estándar, donde la regla de propagación es la suma ponderada y la función de salida es de tipo lineal (Martín & Sanz, 1997).

1.3.3.2 Meso-estructura, la red neuronal artificial.

Una RNA consiste de un conjunto de varios elementos de proceso (neuronas) conectados de alguna forma, y generalmente agrupadas en *unidades estructurales denominadas capas* (Hilera & Martínez, 1995).

Dentro de la red neuronal se distinguen tres tipos de capas (Figura 1.18), como lo señalan Hilera & Martínez (1995):

1. **Capa de entrada.**- Es la capa que recibe directamente la información proveniente de las fuentes externas de la red, es decir del entorno (por ejemplo proporcionados por sensores); y es a través de esta capa que los datos son presentados a la red.
2. **Capa oculta o intermedia.**- Son internas a la red y no tienen contacto directo con el entorno exterior y el número de niveles ocultos puede estar entre cero y un número elevado. Las neuronas de las capas ocultas pueden estar interconectadas de distintas maneras, lo que determina junto con su número, las distintas topologías de la red. Además esta proporciona a la red neuronal grados de libertad adicionales, gracias a los cuales puede encontrar representaciones internas correspondientes a determinados rasgos del entorno, proporcionando una mayor riqueza computacional.

3. **Capa de salida.**- Es aquella capa cuyas neuronas transfieren información de la red hacia el exterior, es a través de esta que la red da respuesta a una entrada dada (sus neuronas pueden estar conectadas a efectores).

A su vez las neuronas de una capa pueden agruparse, formando grupos neuronales³¹. Finalmente el conjunto de una ó más capas constituye la red neuronal.

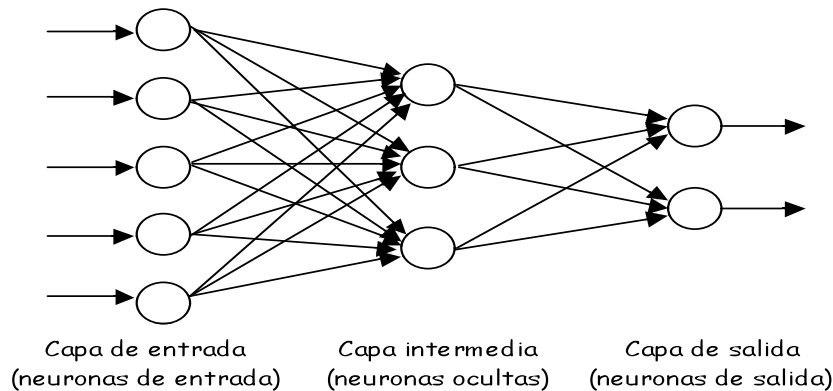


Figura 1.18 Estructura de una red multicapa con conexión unidireccional o propagación hacia delante (Martín & Sanz, 1997).

Además de la existencia de diferentes capas de neuronas para formar una RNA existen otros aspectos que considerar para comprender su funcionamiento a nivel meso-estructura, uno de ellos es determinar su arquitectura y puesto que las neuronas están interconectadas, el otro aspecto importante es conocer las rutas a través de las cuales se propagan las señales para procesar información, y de como se realiza este procesamiento (es decir como es el cambio de estado de las neuronas en dicho camino), pues de ello dependerá su aprendizaje, los modos de operación y recuerdo.

Arquitectura de las redes neuronales artificiales.

La arquitectura de las RNA consiste en la organización y disposición de las neuronas en la red que forman capas ó agrupaciones de neuronas entre la entrada y salida de la red. En este

³¹ En inglés, este término se corresponde con slabs (bloques) o clusters (grupos).

sentido, los parámetros fundamentales de la red son: el número de capas, el número de neuronas por capa, el grado de conectividad y el tipo de conexiones entre neuronas (Hilera & Martínez, 1995; Martín & Sanz, 1997).

Existen diferentes tipos de arquitecturas neuronales, la primera característica es en su relación a su estructura en capas, ó número de capas en una red, por lo cual se distinguen las redes con una capa ó redes monocapa (*single-layer network*) y las redes cuyas neuronas se organizan en varias capas, mejor conocidas como redes multicapa (*layered networks*) ver Figura 1.18.

Cabe mencionar que es el tipo de conectividad lo que permite crear diferentes estructuras y es por ello que se debe diferenciar entre tipos de conexiones intracapa e intercapa. Las *conexiones intracapa*, también denominadas laterales, tienen lugar entre las neuronas pertenecientes a una misma capa, mientras que las *conexiones intercapa* se producen entre las neuronas de diferentes capas.

Las **conexiones laterales ó intracapa**, se utilizan en el caso de redes de una sola capa, este tipo de conexión se usa generalmente en capas competitivas, donde cada nodo se conecta así mismo mediante un peso positivo (excitante) y a los demás nodos de la capa, con pesos negativos (inhibitorios); estas conexiones se diseñan como conexiones excitadoras para permitir la cooperación entre neuronas, ó como inhibidoras, estableciéndose una competición entre las neuronas correspondientes (Haykin, 1999; Hilera & Martínez, 1995), ver Figura 1.19.

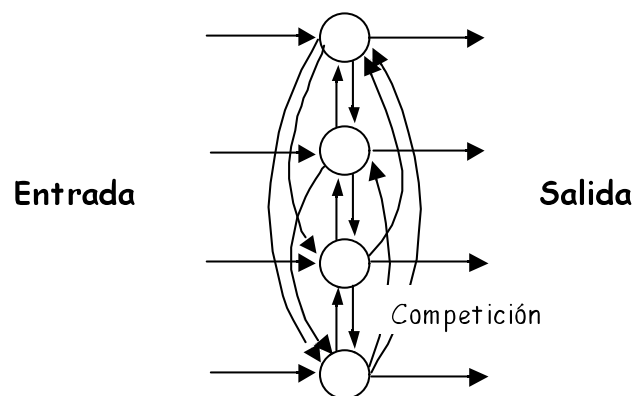


Figura 1.19 Estructura de una red monocapa (en este caso una red de Hopfield).

Las **conexiones intercapa** obviamente se dan en las redes multicapa. Hay que mencionar también que existe una conexión muy particular donde la señal de salida de una neurona puede ser una entrada de sí mismo lo que se denomina como es una *conexión autorrecurrente* ó *conexiones realimentadas* consigo mismas (Figura 1.20).

En las redes multicapa se dice que una red es totalmente conectada si todas las salidas desde un nivel llegan a todos y cada uno de los nodos del nivel siguiente. Nótese que en este caso habría muchas más conexiones que nodos.

Asimismo, con respecto al flujo de datos de la red neuronal podemos referirnos de redes unidireccionales ó redes con conexiones hacia delante (*feedforward*), que son aquellas donde la información circula en un sentido, y redes recurrentes ó redes con conexiones a hacia atrás (*feedback*), donde la información puede circular entre las capas en cualquier sentido incluyendo salida-entrada.

En las **redes unidireccionales o redes no recurrentes** (*feedforward*), la información fluye siempre en un determinado sentido, desde las neuronas de entrada hacia las de salida sin realimentaciones, ese es el sentido de avance de información (Figura 1.18).

Cabe señalar que la estructuración en capas incrementa notablemente el poder representativo de las RNA (o capacidad de la red para modelar una función específica), pero depende de la función de activación de sus neuronas; esta afirmación se basa en el **Teorema de Aproximación Universal** que establece que “*las redes multicapa no añaden capacidad computable a menos que la función de activación de las capas sea no lineal*” (Corchado *et al.*, 2000).

Una red se denomina **recurrente** (*feedback*) cuando en el flujo de datos entre las neuronas que consideramos de entrada y las que consideramos de salida existen realimentaciones (Figura 1.20); esta cierta cantidad de retroalimentación es usada para dar a la red un cierto carácter temporal, es decir, hacer que sea sensible al tiempo; si existe retroalimentación la información reverberará alrededor de la red a través de las capas, ó dentro de una de ellas hasta que alcance algún criterio de convergencia, momento en el que la información es pasada a la salida. Y es esta posibilidad de realimentación la que hace que la red sea mucho más potente computacionalmente hablando, ya que se le permite tener “memoria” aún en el caso de que se asuman elementos de proceso sin

memoria local y no se considere una integración temporal. Estas conexiones hacia atrás o recurrentes, son incorporadas a las redes para tratar modelos dinámicos y temporales, es decir modelos con memoria.

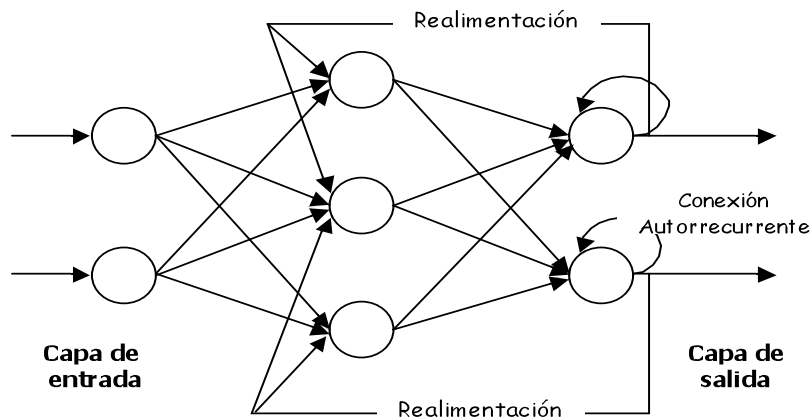


Figura 1.20 Red neuronal con propagación hacia atrás (recurrente).

También existen redes que disponen de conexiones tanto hacia delante como hacia atrás y son conocidas como redes **feedforward/feedback**. En este tipo de redes la información fluye tanto hacia delante, como hacia atrás durante el funcionamiento de la red., generalmente suelen ser bicapa, existiendo por tanto dos conjuntos de pesos: los correspondientes a las conexiones hacia delante (feedforward) de la primera capa (capa de entrada) hacia la segunda (capa de salida) y de los de las conexiones hacia atrás (feedback) de la segunda a la primera.

Algunas redes de este tipo tienen un funcionamiento basado en lo que se conoce como resonancia, de tal forma que las informaciones en la primera y segunda capas interactúan entre si hasta que alcanzan un estado estable: este funcionamiento permite un mejor acceso a las informaciones almacenadas en la red. También en este grupo de redes existen algunas que tienen conexiones laterales.

Dinámica de actualización del estado de las neuronas.

A partir de un estado inicial el estado de la red varia en el tiempo, aunque se puede mas bien considerar una evolución del estado de la red continua en el tiempo, dado un vector de entradas,

la red evolucionará mediante la aplicación de reglas de transición, las cuales estarán determinadas principalmente por el flujo de control de las neuronas, y por la regla de activación de estos. El flujo de control hace referencia al orden en que será actualizado el nivel de actividad interna de las neuronas, es decir sus estados. Existen dos dinámicas fundamentales: *síncrona* y *asíncrona*³².

En los modelos con *dinámica síncrona*, la información llega de forma continua, pero los cambios en los estados se realizan simultáneamente como si existiera un reloj interno. Habitualmente el proceso se lleva a cabo en capas, lo que significa que todas las neuronas pertenecientes a una misma capa se actualizan a la vez, comenzando desde la capa de entrada y continuando hasta la de salida. Esta es la dinámica más habitual. En los modelos de *dinámica asíncrona*, las neuronas evalúan su estado continuamente según les va llegando información, cada neurona tiene una probabilidad de modificar su estado en el instante t , y normalmente sólo una neurona cambiará su activación en dicho instante de tiempo, aunque lo hacen de forma independiente de las demás.

Las redes no recurrentes utilizan habitualmente una actualización síncrona de las neuronas de cada capa. Por otro lado existen redes recurrentes para cada tipo de funcionamiento asíncrono y síncrono.

1.3.4 Aprendizaje o entrenamiento de la RNA.

1.3.4.1 Mecanismos de aprendizaje: supervisado y no supervisado.

Una vez elegida la arquitectura de la red para un problema en particular, la red está lista para el proceso de aprendizaje. Y dado que la noción de “*aprendizaje*” es muy genérica, se adopta la siguiente definición para expresar lo que se entiende por aprendizaje en el contexto de las RNA: “*el aprendizaje es el proceso mediante el cual los parámetros de una RNA se adaptan como consecuencia de un proceso de estimulación llevado a cabo por el entorno en el que la red opera*” (Corchado *et al.*, 2000). Este aprendizaje se consigue a través de reglas que adaptan o cambian los pesos de las conexiones de la red en respuesta a los ejemplos de entrada dados y opcionalmente a la salida deseada para esas entradas (Fernández, 1993). Como durante el proceso

³² Existe la posibilidad de aplicar ambas dinámicas sobre una misma red neuronal, como sucede en el modelo de Hopfield discreto (Martín & Sanz, 1997).

de aprendizaje, los pesos de las conexiones de la red sufren modificaciones, se puede afirmar que este proceso ha terminado (la red ha aprendido) cuando los valores de los pesos permanecen estables.

Los criterios (tipos de regla de aprendizaje) que se siguen para cambiar el valor asignado a las conexiones cuando se pretende que la red aprenda una información nueva son básicamente dos, conocidos como: *aprendizaje supervisado* y *aprendizaje no supervisado*. Los métodos de aprendizaje se utilizan de acuerdo al tipo de información disponible.

En el **aprendizaje supervisado**, la regla de aprendizaje es provista con un grupo de ejemplos (el grupo de entrenamiento, ejemplos de entrenamiento) del comportamiento apropiado de la red (Demuth *et al.*, 1992).

$$\{p_1, t_1\}, \{p_2, t_2\}, \dots, \{p_Q, t_Q\} \quad (1.58)$$

donde p_Q es una entrada a la red, y t_Q es la salida objetivo³³ (la salida correcta correspondiente). Como las entradas son aplicadas a la red, las salidas de la red son comparadas con los objetivos (*target*). La regla de aprendizaje entonces es usada para ajustar a los pesos y tendencias (bias) de la red para mover la salida de la red lo más cercana a los objetivos. En este caso los pesos se obtienen usualmente minimizando alguna función de error que mide la diferencia entre los valores entrada y de salida deseados y los calculados por la red neuronal; la regla de aprendizaje del perceptron cae dentro de esta categoría.

En el **aprendizaje no supervisado**, los pesos y bias son modificados en respuesta a solo a las entradas de la red. No hay targets (objetivos) disponibles, la red tiene que descubrir por si misma patrones o categorías, esta se organiza así misma internamente de tal forma que cada elemento de proceso oculto responde muy activamente o muy débilmente a un grupo de estímulos de entrada, y gran parte del desempeño de esos algoritmos es agrupando operaciones, ellos categorizan los patrones de entrada dentro un número finito de clases. Este tipo de aprendizaje se encuadra dentro de las técnicas autoorganizativas o técnicas automáticas para descubrir la estructura de los datos. Este es especialmente útil en aplicaciones tales como cuantificación de

³³ En inglés target

vectores. Algunos métodos populares de aprendizaje no supervisado son: el aprendizaje Hebbiano, el aprendizaje Competitivo.

Una tercera clase de aprendizaje esta entre los dos anteriores, es el *aprendizaje reforzado*, en el cual un entrenador externo le pasa la información a la red, informando únicamente si la respuesta a una entrada es buena o mala (Martín & Sanz, 1997).

1.3.4.2 Algoritmos de aprendizaje.

La *regla de aprendizaje*, consiste en algoritmos matemáticos que modifican el valor de los pesos de las conexiones y las tendencias (bias) en función de las entradas disponibles para optimizar la respuesta de la red. El esquema de aprendizaje de una red es lo que determina el tipo de problemas que será capaz de resolver, además las RNA son sistemas de aprendizaje basados en ejemplos y por lo tanto la capacidad de una red para resolver un problema esta fundamentada en el tipo de ejemplos de que dispone en el proceso de aprendizaje (Hilera & Martínez, 1995; Isasi & Galvan, 2004; Martin & Sanz, 1997).

El conjunto de ejemplos a usar en el proceso de aprendizaje debe poseer las siguientes características (Isasi & Galvan, 2004; Martin & Sanz, 1997):

- Ser significativo. Debe haber un número suficiente de ejemplos. Si el conjunto de aprendizaje es reducido, la red no será capaz de adaptar sus pesos de forma eficaz. O de lo contrario si son demasiados ejemplos en el entrenamiento existirá el problema de sobre ajuste, este uno de los principales problemas del desarrollo de las RNA, sin embargo es conocido que a menor número de parámetros de entrada, menor será el número de ejemplos necesarios para entrenarla apropiadamente³⁴.
- Ser representativo. Los componentes del conjunto de aprendizaje deberán ser diversos. Si un conjunto de aprendizaje tiene muchos más ejemplos de un tipo que del resto, la red se especializara en dicho subconjunto de datos y no será de aplicación general. Es

³⁴ Debido a la ausencia de resultados teóricos que indiquen la arquitectura óptima para un problema dado, la intuición y la experiencia resultan fundamentales en el desarrollo del sistema neuronal que resuelva con eficacia para una aplicación práctica determinada (Martin & Sanz, 1997), en este caso problemas similares son muy útiles como referencia.

importante que todas las regiones significativas del espacio de estados estén suficientemente representadas en el conjunto de aprendizaje.

Una vez introducidos todos los ejemplos se comprueba si se ha cumplido cierto criterio de convergencia; de no ser así se repite el proceso y todos los ejemplos del conjunto vuelven a ser introducidos. El criterio de convergencia depende del tipo de red utilizado o del tipo de problema a resolver. El entrenamiento usual en las RNA consistirá en modificar los pesos según un determinado algoritmo de aprendizaje hasta obtener la respuesta conveniente.

La finalización del periodo de aprendizaje se puede determinar:

- a. Mediante un número fijo de ciclos. Se decide a priori cuantas veces será introducido todo el conjunto, y una vez superado dicho número se detiene el proceso y se da por aceptada la red resultante.
- b. Cuando el error descienda por debajo de una cantidad preestablecida. En este caso habrá que definir en primer lugar una función de error, bien a nivel patrón individual, bien a nivel de la totalidad del conjunto de entrenamiento: se decide a priori un valor aceptable para dicho error, y solo se detiene el proceso de aprendizaje cuando la red produzca un valor por debajo del prefijado. Para este criterio puede suceder que la red jamás consiga bajar por debajo del nivel prefijado, en cuyo caso se debe disponer de un criterio adicional de parada, por ejemplo un número de ciclos que de utilizarse por la red significará que esta no ha convergido.

Cual sea el tipo de aprendizaje usado es *la regla de aprendizaje*, una característica esencial de la red, pues esta indica cómo adaptar los pesos de las conexiones para aprender un ejemplo dado. Existen una gran cantidad de algoritmos de aprendizaje (Aprendizaje Hebbiano, aprendizaje competitivo, aprendizaje de memoria basada, etc.) que se han desarrollado a lo largo del tiempo y con la madurez en esta área, a continuación se explicará como funcionan los algoritmos de aprendizaje, tomando como ejemplo al Aprendizaje de corrección de errores.

Aprendizaje de corrección de errores.

Considerando un caso simple de una neurona k constituyendo sólo el nodo computacional en la capa de salida de una red neuronal con conexiones hacia adelante (*feedforward*) como se ilustra

en la Figura 1.18, la neurona k es manejada por un vector de señal $x(n_t)$ producido por una o más capas de neuronas ocultas, las cuales son así mismas manejadas (estimuladas) por un vector de entrada aplicado al origen de nodos (por ejemplo una capa de entrada) de la red neuronal. El argumento n_t denota tiempo discreto, o más precisamente el paso del tiempo de un proceso iterativo involucrado en ajustar el peso sináptico de la neurona k . La señal de salida $y_k(n_t)$, representa sólo la salida de la red neuronal, es comparada con una respuesta deseada o un objetivo de salida, denotado por $d_k(n_t)$. Consecuentemente un error de señal, denotado por $e_k(n_t)$, es producido. Por definición (Haykin, 1999), se tiene que:

$$e_k(n_t) = d_k(n_t) - y_k(n_t) \quad (1.59)$$

La señal de error $e_k(n_t)$ actúa como un mecanismo de control, el propósito del cual es para aplicar una secuencia de ajustes correctivos a el peso sináptico de la neurona k . Los ajustes correctivos son diseñados para hacer que la señal de salida $y_k(n_t)$ llegue lo más cercano a la respuesta deseada $d_k(n_t)$ en un manejo paso por paso. Este objetivo es llevado acabo para minimizar un costo de función o índice de desempeño $\varepsilon(n_t)$, que es definido en términos de la señal errónea $e_k(n_t)$ como:

$$\varepsilon(n_t) = \frac{1}{2} e_k^2(n_t) \quad (1.60)$$

Eso es $\varepsilon(n_t)$ que es el valor instantáneo del error de energía. El ajuste paso por paso a los pesos sinápticos de la neurona k , continúan hasta que el sistema alcance un estado estacionario (es decir cuando los pesos sinápticos son esencialmente estabilizados). En este punto el proceso de aprendizaje ha terminado.

El proceso de aprendizaje arriba descrito es obviamente referido como aprendizaje de corrección de errores. En la particular minimización del costo de función $\varepsilon(n_t)$, supongamos una regla de aprendizaje comúnmente referida como la *Regla Delta* ó *Regla Widrow-Hoff*.

Esto es $w_{kj}(n_t)$ denota el valor del peso sináptico w_{kj} de la neurona k excitada por el elemento $x_j(n_t)$ de la señal del vector $\mathbf{X}(n_t)$ en un paso de tiempo n_t . De acuerdo con la regla Delta, el ajuste $\Delta w_k(n_t)$ aplicado al peso sináptico w_{kj} en el paso del tiempo n_t es definido por:

$$\Delta w_{kj}(n_t) = \eta_A e_k(n_t) x_j(n_t) \quad (1.61)$$

donde η_A es una constante positiva que determina la rapidez de aprendizaje, como el procedimiento desde un paso en el proceso de aprendizaje a otro. Este es por lo tanto natural que nos refiere a η_A como el parámetro de rapidez de aprendizaje, en otras palabras, la regla delta podría ser así: “*el ajuste hecho a un peso sináptico de una neurona es proporcional a el producto de la señal de error y la señal de entrada de la sinapsis en cuestión*”.

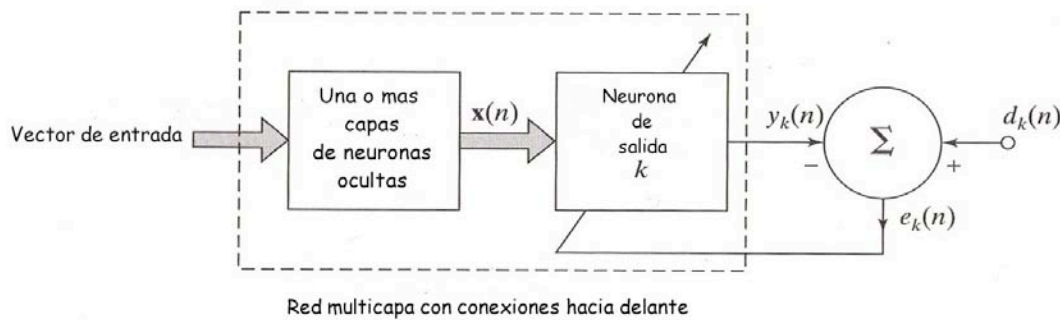


Figura 1.21 Aprendizaje de corrección de errores, diagrama de una RNA enfatizando una sola neurona en la capa de salida (Haykin,1999).

Teniendo en mente que la regla delta como se anuncia arriba, presume que el error de señal es directamente medible, en otras palabras la neurona k es visible a el mundo externo, como lo ilustra la Figura 1.21. Esto es simplemente decir que el ajuste sináptico hecho por la regla delta es localizado alrededor de la neurona.

Una vez computado el ajuste sináptico $\Delta w_{kj}(n_t)$, el valor de actualización de peso sináptico w_{kj} es determinado por

$$w_{kj}(n_t + 1) = w_{kj}(n_t) + \Delta w_{kj}(n_t) \quad (1.62)$$

En efecto, $w_{kj}(n_t)$ y $w_{kj}(n_t+1)$ podrían ser vistos como los valores anterior y nuevo del peso sináptico w_{kj} , respectivamente. En términos computacionales podríamos escribir:

$$w_{kj}(n_t) = z_n^{-1}[w_{kj}(n_t + 1)] \quad (1.63)$$

Donde z_n^{-1} es la unidad operador de retraso, que representa un elemento de almacenamiento.

1.3.4.3 Selección del Algoritmo de aprendizaje.

Como ya se mencionó anteriormente existen gran cantidad de algoritmos y cada uno tiene diferentes propiedades, que son más adecuadas para unas tareas que para otras, sin olvidar que todos los paradigmas computacionales, tienen sus limitaciones al igual que cualquier otro método numérico. La selección apropiada de un algoritmo de aprendizaje dependerá de la naturaleza del problema (es importante dominar el conocimiento suficiente acerca del problema práctico) y del tipo de datos involucrados. Aunque muchos algoritmos de aprendizaje podrían producir resultados, los mejores resultados dependerán de la adecuada selección del algoritmo) algunos factores a considerar son:

- 1) Propiedades de convergencia
- 2) Estabilidad
- 3) Exactitud en la clasificación de nuevos casos
- 4) Habilidad para interpretar resultados
- 5) Compatibilidad del fundamento del algoritmo con la aplicación de la RNA

1.3.5 Realización de redes neuronales artificiales.

Hay esencialmente tres maneras de realizar una RNA, cada una tiene características particulares, que pueden ser convenientes ó no, al usarlas (Martín & Sanz, 1997).

a) Simulación (Software) de redes neuronales artificiales.

La realización más simple e inmediata consiste en simular la red sobre un ordenador convencional mediante un software específico; los ordenadores actuales resultan suficientes para resolver numerosos problemas, al permitir simular redes de tamaño medio a una velocidad razonable. Es un procedimiento rápido, poco costoso e insustituible por el momento para realizar entrenamiento y evaluación de las redes, la mayor desventaja radica en el hecho de que intentar simular redes con alto grado de paralelismo sobre máquinas que ejecutan secuencialmente operaciones, valores intrínsecos no pueden obtenerse de esta forma, además considerando los recursos computacionales para entrenar una RNA de tamaño medio o alto, y la necesidad de respuesta en tiempo real en diversas aplicaciones, se debe superar en ocasiones la simple

simulación en ordenador y recurrir a estructuras hardware especialmente diseñadas para la emulación o realización de redes neuronales.

b) Emulación³⁵ (Hardware) de redes neuronales artificiales.

Con el término hardware se refiere a la utilización de una estructura de procesamiento con cierta capacidad de cálculo paralelo, especialmente diseñada para el trabajo con las RNA. Es una solución intermedia entre la simulación y la verdadera realización hardware. Se trata de utilizar una tarjeta coprocesadora dependiente de un cierto *host*,³⁶ construida tomando como base microprocesadores de altas prestaciones tales como chip neuronales, mediante un software de simulación adecuado, se pueden aprovechar las características de cálculo limitadamente paralelo de la tarjeta utilizada para acelerar la simulación de una RNA. Es una buena solución, tanto para el desarrollo y evaluación de un sistema neuronal como para aplicaciones complejas. Aunque no siempre resulta excesivamente económica y tampoco se aprovechan totalmente las características de cálculo paralelo de la red, pues subsiste el hecho de que el comportamiento real de la red sigue siendo simulado por una estructura ajena a la estructura intrínseca de una RNA.

c) Realización Hardware de redes neuronales artificiales.

Consiste en construir físicamente una RNA, mediante su implementación por uno o varios circuitos integrados específicos; un sistema cuya arquitectura refleje en cierta medida la estructura de la RNA, son los llamados *chips neuronales* (dentro de mismo chip se incluyen todas las interconexiones, lo que le proporciona velocidades superiores) y los *neurocomputadores*. Las neuronas y conexiones se emulan con dispositivos específicos, de forma que la estructura del circuito integrado refleja la arquitectura de la red, debido a ello el sistema hardware debe soportar un gran número de conexiones, en ocasiones globales y un flujo elevado de comunicaciones de datos. Estas circunstancias limitan considerablemente el tamaño de la RNA que puede ser implementada. Sin embargo de esta forma se consiguen realizaciones que funcionan a una alta velocidad, permitiendo en muchas ocasiones el proceso en tiempo real, pero a costa de una pérdida notable de flexibilidad. Este tipo de realización está especialmente indicado para aplicaciones especiales donde se precisa un alto rendimiento, ya que el diseño se realiza para cierto modelo de red neuronal y para una aplicación concreta.

³⁵ Término comúnmente utilizado en informática, ver glosario para más detalles.

³⁶ En español significa anfitrión, pero aquí se utiliza "host" como un término informático, ver glosario.

1.3.6 Generalidades de las RNA y aplicaciones en la industria alimentaria.

1.3.6.1 Clasificación de las RNA

Las RNA pueden ser clasificadas en diferentes maneras dependiendo de su estructura, modelo de neurona utilizada, algoritmo de aprendizaje y su función u objetivo. Una clasificación de acuerdo a su funcionalidad puede ser la siguiente, la cual divide a las redes en las categorías siguientes (Cohen & Hudson, 2000):

- **De clasificación:** los modelos de clasificación asignan datos detallados de entrada a dos o más categorías. Estos modelos podrían usar aprendizaje supervisado en el cual las categorías son conocidas o aprendizaje no supervisado en el cual las categorías no son conocidas.
- **De asociación:** los dos tipos de modelos de asociación son el de auto-asociación (aprendizaje no supervisado), el cual se enfoca sobre la recuperación de un objeto basado sobre parte del objeto en sí mismo; y la hetero-asociación (aprendizaje supervisado), la cual se enfoca sobre la recuperación de un objeto de un grupo usando un objeto en un grupo diferente.
- **De optimización (Generalización³⁷):** el objetivo de estos sistemas es el encontrar la mejor solución por un costo de función mínimo u otra medida.
- **De organización por si mismo (Auto-organización):** este aprovechamiento organiza la información usando facilidades del aprendizaje adaptivo, este es similar a ramificar o agrupar algoritmos, basado en técnicas de aprendizaje no supervisado.

1.3.6.2 Ventajas y desventajas de usar las RNA

Se debe analizar y considerar cuando es conveniente utilizarlas y cuando no son la mejor opción, para resolver el problema presentado (Martín & Sanz, 1997; Hilera & Martínez, 1995).

Ventajas de utilizar las redes neuronales:

- Son apropiadas, cuando no se dispone de un conjunto de reglas sistemáticas que describan completamente el problema, si no hay modelo específico que lo resuelva, pero sí se

³⁷ En inglés Fitting a function.

dispone, de muchos ejemplos o casos históricos (esta es una condición indispensable para la aplicación de las RNA).

- Capacidad de aprender a realizar tareas basadas en un entrenamiento o una experiencia inicial (Aprendizaje adaptativo).
- Autoorganización, puede crear su propia organización o representación de la información que recibe (a través del aprendizaje).
- Tolerancia a fallos, cuando los datos procedentes del problema son imprecisos, incoherentes o con ruido y también ante la destrucción parcial de la red y degradación de su estructura, algunas capacidades se pueden retener, aún sufriendo mucho daño.
- Operación en tiempo real. Las RNA pueden ser diseñados en paralelo y se fabrican máquinas con hardware especiales con esta capacidad.
- Fácil inserción dentro de la tecnología existente. Ya sea con un software ó consiguiendo chips especializados para RNA que mejoran su capacidad en ciertas tareas. Ello facilitará la integración modular en los sistemas existentes.
- Cuando se requiera una alternativa mucho más rápida y sencilla de desarrollar que otras técnicas convencionales.

Inconvenientes de utilizar las redes neuronales

Es obvio, que las RNA no son la panacea que resuelve todos los problemas, sino que están dirigidas a un determinado tipo de tareas, además las RNA tienen todavía limitaciones por resolver (Martín & Sanz, 1997; Hilera & Martínez, 1995):

- Si existe un algoritmo que resuelva con total precisión el problema, como es el caso de los problemas puramente algorítmicos o numéricos (multiplicación de números, inversión de matrices, etc.)
- Si **no** se dispone de un número adecuado de casos (ejemplos) para entrenar la red.
- Constituyen un método de resolución de problemas demasiado creativo, es decir, dadas las especificaciones de un problema, se desconoce la arquitectura de las RNA que la va a solucionar de forma más eficiente. Hay que analizar el método de prueba y error.

- Una vez entrenada una RNA, es difícil interpretar su funcionamiento, aún peor, no es fácil asegurar con que acierto responderá ante casos nunca vistos (aunque en la práctica es muy poco común que un RNA bien entrenada y verificada falle al emplearla en el mundo real).
- Es difícil averiguar porque la RNA no es capaz de ajustar los datos que se le proporcionan, algunas veces, puede resultar difícil encontrar porqué no funcionan correctamente.
- Los modelos neuronales precisan elevados requisitos de cómputo, sobre todo en el modo aprendizaje: no obstante, este inconveniente se contrarresta con su relativa facilidad de implementación en dispositivos hardware específicos.

1.3.6.3 Aplicaciones de las RNA en la industria alimentaria.

Las características de las RNA permiten que encuentran aplicaciones en campos tan diversos como la modelación, patrón de reconocimientos, identificación, clasificación, procesamiento de señales y control, etc. En particular las aplicaciones de las RNA a procesos químicos y bioquímicos, hasta ahora han sido diversos; actualmente las RNA son frecuentemente usadas para problemas relacionados con procesos alimenticios, y a continuación se describen algunos ejemplos agrupados por la principal aplicación:

- Clasificación de patrones.

Se han utilizado en la clasificación de objetos, por forma, color y textura en la inspección de productos industriales y agrícolas (Hilera & Martínez, 1995); y en la inspección de calidad de productos, todo esto basado en imágenes para la clasificación de manzanas y cítricos, reciclado de botellas, etc. También se han aplicado para el control de calidad y aseguramiento de la calidad en las industrias, por citar un ejemplo, en la detección de niveles de contaminantes a partir de datos espectroscópicos en plantas químicas y la evaluación del grado de pureza del jugo de naranja (Martín del Brío & Sanz, 1997). Y para predecir y clasificar los atributos de palatabilidad de la carne de res, usando características de espectros ultrasónicos como parámetros descriptivos (Mittal & Zhang, 2000).

- Control y optimización de proceso.

En esta área se desarrollo una modelo de RNA, para predecir condiciones óptimas de procesamiento térmico, y predecir óptimas temperaturas de esterilización, las cuales resultan en la menor degradación de la calidad (Afaghi *et al.*, 2001). Otra aplicación es modelos de RNA para proceso de cocción de pan, para predecir los factores de calidad de volumen, oscurecimiento y temperatura. El desarrollo de modelos de RNA fueron parte de la simulación de un controlador difuso del horno usado para el proceso de cocción. La aplicación de este controlador resulto en un decrecimiento del costo de calentamiento del horno sin ninguna pérdida de la calidad del pan (Afaghi *et al.*, 2001). Asimismo se utilizó una RNA para determinar las variables significantes en la cosecha y los efectos de proceso sobre la calidad del surimi, los cuales fueron después base para determinar factores que afectan la calidad final del producto en una operación multi-proceso (Afaghi, *et al.*, 2001).

- Generalización (Predicción).

En esta aplicación las RNA han tenido diversidad de usos como: el desarrollo de un modelo para la predicción de crecimiento microbiano considerando efectos de temperatura, pH, y actividad de agua (Mittal & Zhang, 2000). Asimismo han tenido uso para la predicción de propiedades reológicas de una masa galletera, usando dos técnicas de análisis de espectros, la transformación rápida de Fourier (FFT) y la densidad de potencia de espectro (PSD) (Mittal & Zhang, 2000). Así como la predicción de la producción de aceite de semillas de cacahuate en una prensa hidráulica, considerando variables de proceso como contenido de humedad, presión aplicada, tiempo de presión, y tiempo y temperatura de calentamiento, (Olajide *et al.*, 2007).

En esta área las RNA también se han utilizado como herramienta de cálculo para la predicción de los tiempos de congelamiento para productos alimenticios a partir de datos generados del modelo de Pham (Mittal & Zhang, 2000). También el predecir dos parámetros para procesamiento de alimentos a alta presión: la temperatura máxima o mínima alcanzada en la muestra después de presurización y el tiempo necesario para el re-equilibrio térmico en el sistema de alta presión, con el fin de optimizar procesos térmicos (Torrecilla *et al.*, 2004). Así como el desarrollo de un procedimiento explícito para calcular el factor de fricción bajo condiciones de flujo tanto laminar como turbulento de fluido plástico de Bingham, en conductos cerrados y redes de bombeo. El

procedimiento establece un relación explícita entre el número de Reynolds Re , número de Hedstrom He , y el factor de fricción f_f en ambas condiciones de flujo, además elimina la necesidad soluciones iterativas y de realizar cálculos complejos aunado a la reducción del consumo de tiempo (Sablani & Shayya, 2003).

Como se puede observar los usos han sido diversos, sin embargo de acuerdo a lo que se ha explicado anteriormente hay una metodología básica al desarrollar el modelo para una aplicación en particular. Y es precisamente esta última aplicación la que se considera como base para la metodología propuesta al desarrollar el modelo para calcular el factor de fricción de los fluidos Herschel-Bulkley de manera no-iterativa basado en RNA.

CAPÍTULO 2. Diseño del modelo matemático para el cálculo no-iterativo del factor de fricción en tuberías para fluidos Herschel-Bulkley en régimen laminar y turbulento

2.1 Desarrollo del modelo de la RNA.

2.1.1 Metodología.

La metodología se muestra en la Figura 2.1.

Objetivo: Desarrollar un modelo matemático mediante el uso de las RNA que permita calcular de manera no-iterativa el factor de fricción en tuberías para fluidos Herschel-Bulkley en régimen laminar y turbulento.

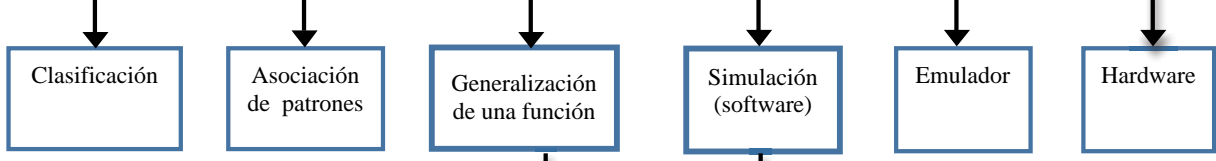
1.A. Generación de datos para entrenamiento de la RNA
 Se utilizan las ecuaciones de Hanks (régimen laminar) y Torrance (régimen turbulento), para calcular f_f para fluidos Herschel-Bulkley, en función de los parámetros $\{Re_G, n, He\}$; vía método de bisección.

- En régimen Laminar, del intervalo:
 $Re_G = [100 \text{ a } 2100]$
 $n = [0.1 \text{ a } 1]$
 $He = [10 \text{ a } 1 \times 10^7]$
- En régimen Turbulento, del intervalo:
 $Re_G = [3000 \text{ a } 1 \times 10^7]$
 $n = [0.1 \text{ a } 1]$
 $He = [10 \text{ a } 1 \times 10^7]$

1.B. Selección general de la RNA de acuerdo al problema

1.B.1. Propósito de la RNA

1.B.2. Modo de realizar la RNA



2.B. Definición de las características de la RNA

2.B.1. Arquitectura de la RNA

2.B.2. Tipo de aprendizaje de la RNA

Meso-estructura de la RNA
 Características:

- Número de capas en la RNA: red multicapa, de tres capas (entrada-oculta-salida).
- Conexión entre neuronas: intercapa y unidireccional.
- Flujo de información: en sentido, de atrás hacia adelante.

Micro-estructura de la RNA
 Características:

- Capa de entrada
 - Regla de propagación: no aplica.
 - Presencia de *bias* (umbral ó tendencia).
 - Función de transferencia: no aplica.
 - Función de salida: no aplica.
- Capa oculta (intermedia)
 - Regla de propagación: suma ponderada.
 - Presencia de *bias* (umbral ó tendencia).
 - Función de transferencia: tangente hiperbólica.
 - Función de salida: lineal.
- Capa de salida
 - Regla de propagación: suma ponderada.
 - Sin presencia de *bias* (umbral ó tendencia).
 - Función de transferencia: lineal.
 - Función de salida: lineal.

Supervisado

No supervisado

2.B.2.1. Algoritmo de aprendizaje: Aprendizaje de Corrección de errores

Regla Delta (regla Widrow-Hoff): Optimizada se deriva en el algoritmo Levenberg-Marquardt

2.A. Tratamiento previo de datos
 Los datos obtenidos se organizan en parejas: {datos de entrada, datos de salida}, donde a cada miembro de la pareja de datos se le considera como un sub-conjunto. Y a cada dato de entrada le corresponde una salida.
 Ya organizados en parejas, se le considera un conjunto (grupo) de ejemplos, estos serán datos de entrada y datos de la salida objetivo.
 El grupo de datos originados por las ecuaciones de Hanks y Torrance, es la base de datos para el entrenamiento, pero a estos se les aplicará un diferente tratamiento previo de datos, y posteriormente serán presentados a la RNA como ejemplos para que aprenda de ellos. Las tres variaciones entre los tratamientos previos de datos son:

- Grupo A.- Datos de entrada = $\{Re_G, n, He\}$; datos de salida = $\{f_f\}$.
- Grupo B.- Datos de entrada = $\{\text{Log}(Re_G), \text{Log}(n), \text{Log}(He)\}$; datos de salida = $\{f_f\}$.
- Grupo C.- Datos de entrada = $\{\text{Log}(Re_G), \text{Log}(n), \text{Log}(He)\}$; datos de salida = $\{\text{Log}(f_f)\}$.

3.B. Selección del modelo particular de la RNA
 Del modelo general de la RNA, se hace una variación del número de neuronas en la capa oculta, para elegir la última característica del modelo de la RNA de acuerdo a los resultados que se obtengan.
 Número de neuronas de la capa oculta, inicialmente a probar: 4, 8, 16, 20, 30 y 40.

4. Entrenamiento de la RNA

5. Evaluación de la RNA
 Comparación del factor de fricción obtenido por el método iterativo (f_f) vs el factor de fricción obtenidos con cada modelo de la RNA y sus condiciones de entrenamiento (f_{fp}), utilizando:

- Desviación estándar relativa (STDR),
- Desviación estándar absoluta (STDA),
- Error medio relativo (MRE)
- Error medio absoluto (MAE)
- Coefficiente de determinación (R^2).

6. Selección de mejores modelos de la RNA y condiciones

7. Validación de la RNA
 Con un grupo de datos nuevos, pero dentro del mismo intervalo de valores (condiciones) en los que se entrenó al modelo de RNA, se compara estadísticamente el factor de fricción (f_{fp}) obtenido por el modelo de RNA elegido vs el valor de factor de fricción calculado por las ecuaciones de Hanks y Torrance (f_{fc}).

- Desviación estándar relativa (STDR),
- Desviación estándar absoluta (STDA),
- Error medio relativo (MRE)
- Error medio absoluto (MAE)
- Coefficiente de determinación (R^2).

8. Análisis de resultados

9. Obtención de ecuaciones asociadas al modelo elegido (considerando su estructura, sus características de funcionamiento y las condiciones de entrenamiento).

Conclusiones

2.1.2 Descripción del Cuadro metodológico.

A continuación se describe la metodología seguida en este trabajo, con la finalidad de resolver el **Objetivo general:**

Desarrollar un modelo matemático mediante el uso de las redes neuronales artificiales que permita calcular de manera no-iterativa el factor de fricción en tuberías para fluidos Herschel-Bulkley en régimen laminar y turbulento, para disminuir el tiempo dedicado a los cálculos iterativos y generar una nueva herramienta de cálculo.

Para conseguirlo se realizan una serie de pasos o actividades requeridas para tal propósito. Por una parte se necesita seleccionar la RNA de manera general, de acuerdo a las características del problema que se desea resolver. Y por otro lado, esta la necesidad de generar una base de datos, que muestre de manera global las relaciones entre las variables del problema, (en este caso, el factor de fricción que se encuentra en función de las variables: número de Reynolds generalizado " Re_G ", el índice de comportamiento al flujo " n " y número de Hedstrom " He "), esta base de datos, servirá de ejemplo para la que la RNA aprenda acerca de las relaciones entre las variables que intervienen en el cálculo del factor de fricción de un fluido Herschel-Bulkley en una tubería, tanto en régimen de flujo, laminar como turbulento. Ambos procesos tanto el de generar un base de datos y el proceso de seleccionar un modelo general de RNA con las características necesarias para resolver el problema planteado, se llevarán acabo por separado y de manera paralela al inicio y más adelante, estos dos elementos se relacionaran para resolver el problema.

Primero se describirá el proceso de cómo generar la base de datos del fenómeno a generalizar.

➤ 1.A. Generación de datos.

Tener con una base de datos del fenómeno a generalizar, es requerida para el entrenamiento de la RNA (son los ejemplos de los cuales aprenderá la RNA). Y los aspectos primordiales a considerar al generar esta base de datos, es que sea representativa es decir que estén representadas todas las regiones significativas del espacio de estados y como se explicó en el tema (1.3.4.2) del presente trabajo, también debe contar con suficientes ejemplos para mostrar de manera general las variaciones de las relaciones entre variables involucradas en el fenómeno. En

este caso en particular, se trata de generalizar la relación de dependencia del factor de fricción en función de las variables $\{Re_G, n, He\}$, para un fluido Herschel- Bulkley en régimen laminar de acuerdo a la ecuación de Hanks (ecuación 1.36) y en régimen turbulento de acuerdo a la ecuación de Torrance (ecuación 1.39). La creación de la base de datos se inicia definiendo cuales serán los parámetros de entrada y por consiguiente de salida de la red e intervalos que abarcaran, además de definir la cantidad de ejemplos necesarios para el entrenamiento, cabe señalar que estos datos serán generados utilizando el método de bisección para la resolución de las ecuaciones (1.36) y (1.39) y así obtener los parámetros objetivo (salida). Estos datos generados se agruparán en dos sub-conjuntos de datos, el primer sub-conjunto={datos de entrada} representados por los valores de las variables $\{Re_G, n, He\}$ y el segundo sub-conjunto={datos de salida} representados por el valor de factor de fricción $\{f_f\}$, que juntos formaran una pareja donde un arreglo de datos de entrada corresponden a un único dato de salida.

Se proponen los siguientes intervalos para cada uno de parámetros de entrada $\{Re_G, n, He\}$:

- En régimen Laminar

Re_G con valores dentro del intervalo [100 a 2100]

n con valores dentro del intervalo [0.1 a 1]

He con valores dentro del intervalo [10 a 1×10^7]

- En régimen Turbulento

Re_G con valores dentro del intervalo [3000 a 1×10^7]

n con valores dentro del intervalo [0.1 a 1]

He con valores dentro del intervalo [10 a 1×10^7]

La variable dependiente en ambos casos es el factor de fricción (f_f), que en ambos regímenes de flujo queda en función de las variables $\{Re_G, n, He\}$. Y por lo tanto se considera el parámetro de salida, en adelante a este parámetro se le identificara como el factor de fricción calculado por ecuaciones (f_{fc}).

➤ 2.A. Tratamiento previo de datos.

Una vez obtenida la base de datos, estos se ordenarán en un conjunto (un grupo de datos) de la forma:

$$Datos = \{(p_1, t_1), (p_2, t_2), \dots, (p_Q, t_Q)\}$$

donde p_Q es una entrada, y t_Q es la salida objetivo. Y estos a su vez estarán agrupados en dos sub-conjuntos: Datos de entrada, $p = \{Re_G, n, He\}$ y Datos de salida, $t = \{f_{fc}\}$.

Estos valores, como ya se mencionó formaran una pareja {datos de entrada, datos de salida} y se presentarán a la RNA en forma matricial, donde un arreglo de datos de entrada corresponden a un único dato de salida. Antes de que estos datos se presenten a la RNA para que aprenda de ellos (proceso de entrenamiento), se hará un tratamiento previo de los datos, tanto para el sub-conjunto de datos de entrada como para el sub-conjunto de datos de salida, con tres diferentes combinaciones:

- Pareja de datos (Grupo A).- Valores normales para las matrices de entrada y salida.
Datos de entrada, $p_A = \{Re_G, n, He\}$ y datos de salida, $t_A = \{f_{fc}\}$.
- Pareja de datos (Grupo B).- Valores con transformación logarítmica (calcular el logaritmo base 10 a los valores) para el sub-conjunto de entrada y sin transformación logarítmica para el sub-conjunto de salida.
Datos de entrada, $p_B = \{\text{Log}(Re_G), \text{Log}(n), \text{Log}(He)\}$ y datos de salida $t_B = \{f_{fc}\}$.
- Pareja de datos (Grupo C).- Con transformación logarítmica para ambos sub-conjuntos.
Datos de entrada, $p_C = \{\text{Log}_{10}(Re_G), \text{Log}_{10}(n), \text{Log}_{10}(He)\}$ y datos de salida, $t_C = \{\text{Log}_{10}(f_{fc})\}$.

Ahora se describirá como se llevará acabo el proceso de seleccionar el modelo general de RNA más apropiado para resolver el problema en cuestión, este proceso es paralelo a la generación de la base de datos y su tratamiento previo.

Como parte del objetivo general, se generan los objetivos particulares necesarios para resolver el objetivo general, en este caso:

Objetivo particular 1: Seleccionar la estructura de la red neuronal artificial que permita calcular el factor de fricción en el menor tiempo posible en comparación con el empleado por el método iterativo.

➤ 1.B Selección general de la RNA de acuerdo al problema.

Una vez realizada la revisión bibliográfica, se determinará el propósito (función) que se desea que realice la RNA para resolver el objetivo planteado y la manera más adecuada para llevar a cabo la RNA.

1.B.1. Propósito de la RNA.

Las RNA comunes tienen tres propósitos (funciones) primordialmente: Clasificación de patrones, Asociación de patrones, Generalización (en este caso de una función matemática).

De estas funciones se elige la que sea más conveniente para resolver el problema, por lo tanto se elige \Rightarrow Generalización de una función.

1.B.2. Modo de realizar la RNA.

Paralelamente a la elección de la función que tendrá la RNA se debe elegir como se llevará a cabo la realización de la RNA, de acuerdo a las tres maneras generales de llevarse a cabo (Martin & Sanz, 1997): Simulación (software), Emulación (tarjeta co-procesadora) y Hardware (arreglo físico). De estas opciones se elige \Rightarrow Simulación (Software).

Cabe señalar que se trabajará con un simulador de RNA, el Neural Network Toolbox™ de la plataforma Matlab®.

➤ 2.B Características de la RNA.

Una vez definido el propósito que tendrá la RNA y la manera de cómo se llevará a cabo, se tiene que definir sus características generales: Arquitectura y Tipo de aprendizaje.

2.B.1. Arquitectura de la RNA (estructura).

La arquitectura consiste en la organización y disposición de neuronas en la RNA, por lo tanto define la mayoría de las características que tendrá la RNA, y se divide en dos niveles:

- Meso-estructura de la RNA, en la cual se define: el número de capas que tendrá la RNA, el tipo de conexiones entre neuronas y la dirección de flujo de información.

Definición de características de la RNA a nivel meso-estructura (Figura 2.2):

- Número de capas en la RNA \Rightarrow red multicapa, de tres capas (entrada-oculta-salida).
- Conexión entre neuronas \Rightarrow intercapa y unidireccional.
- Flujo de información \Rightarrow en sentido, de atrás hacia adelante.

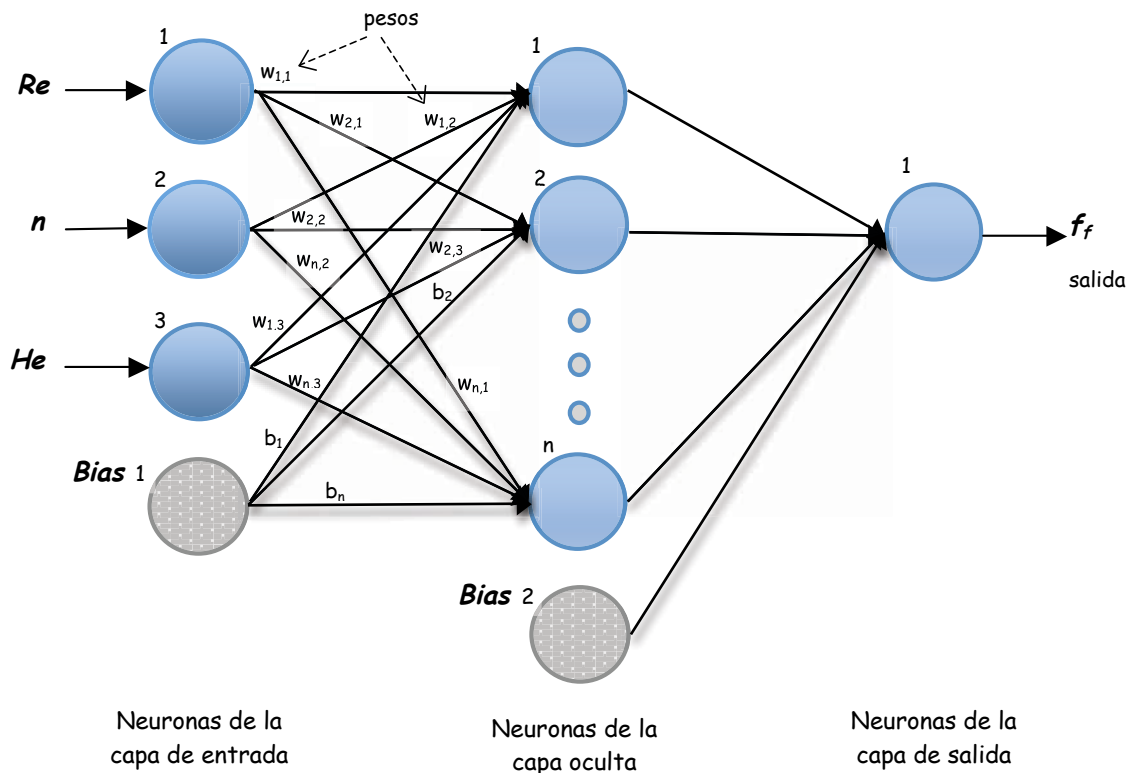


Figura 2.2 Diagrama general de la RNA a nivel meso-estructura (RNA multicapa, con flujo de información de atrás hacia adelante y unidireccional), a utilizar.

- Micro-estructura de la RNA, en la cual se define: la regla de propagación de las neuronas, uso de *bias* (valor de tendencias ó umbral), el tipo de función de transferencia y función de salida con la que trabajarán todas las neuronas de cada capa de la RNA (Figura 2.3).

Definición de características de la RNA a nivel micro-estructura:

1. Capa de entrada

- Regla de propagación \Rightarrow no aplica.
- Presencia de *bias* (umbral ó tendencia).
- Función de transferencia \Rightarrow no aplica.
- Función de salida \Rightarrow no aplica.

2. Capa oculta

- Regla de propagación \Rightarrow suma ponderada.
- Presencia de *bias* (umbral ó tendencia).
- Función de transferencia \Rightarrow tangente hiperbólica.
- Función de salida \Rightarrow lineal.

3. Capa de salida

- Regla de propagación \Rightarrow suma ponderada.
- Sin presencia de *bias* (umbral ó tendencia).
- Función de transferencia \Rightarrow lineal.
- Función de salida \Rightarrow lineal.

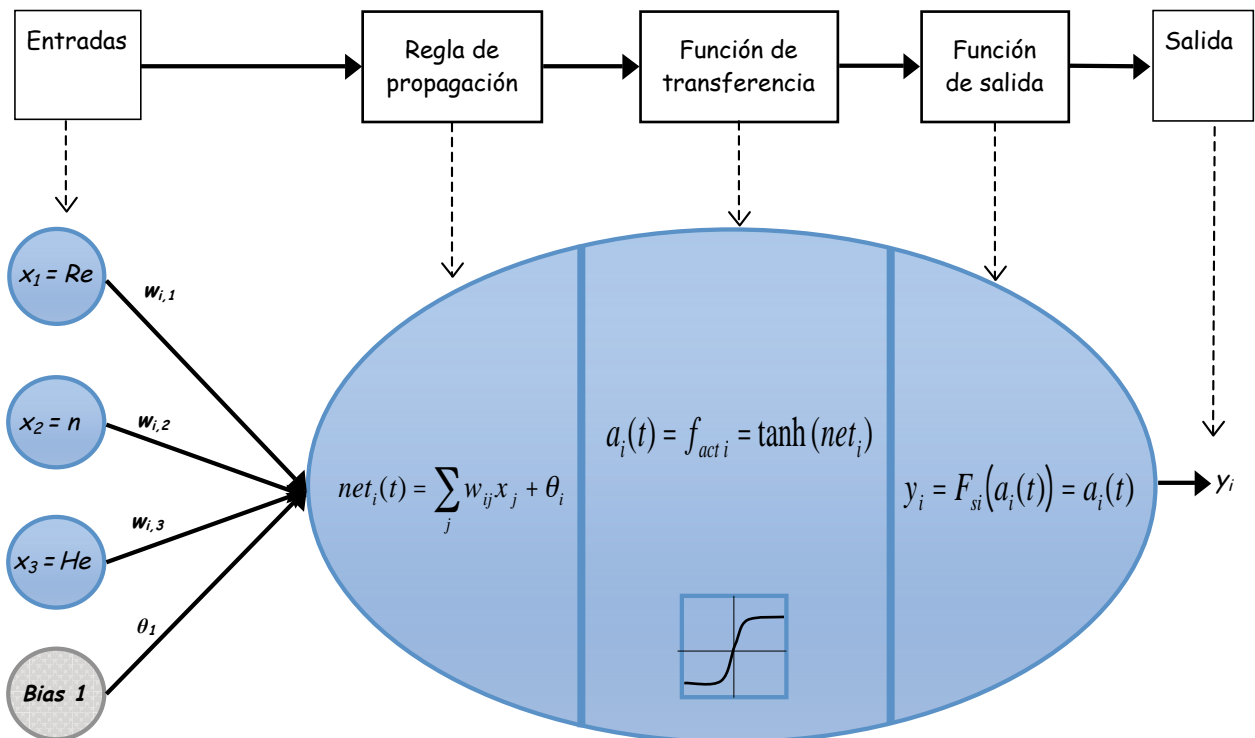


Figura 2.3 Modelo genérico del procesamiento de información de una neurona i de la capa oculta, donde hay presencia de un bias (una tendencia, que se registra como una entrada extra), la regla de propagación es la suma ponderada; la función de transferencia es la tangente hiperbólica y la función de salida es la lineal.

2.B.2. Tipo de Aprendizaje de la RNA.

El modo de Aprendizaje (entrenamiento) para las RNA, básicamente puede ser de dos tipos: Aprendizaje supervisado y no supervisado, se determina que se usará \Rightarrow El criterio de aprendizaje supervisado.

Donde se utiliza un grupo de ejemplos del comportamiento adecuado, que se desea de la RNA.

Los ejemplos se le proporcionarán a la RNA de la siguiente forma: $\{(p_1, t_1), (p_2, t_2), \dots, (p_Q, t_Q)\}$

donde p_Q es una entrada a la red, y t_Q es la salida objetivo. Las parejas de datos se agruparan en lotes de datos, y se presentaran (en Batch), es decir se le proveerá a la RNA de todos lo datos al mismo tiempo y por grupos.

2.B.2.1. Algoritmo de aprendizaje.

El algoritmo de aprendizaje, consiste en algoritmos matemáticos que modifican el valor de los pesos de las conexiones entre neuronas y sus respectivos *bias* (tendencias ó umbrales) en función de las entradas disponibles para optimizar la respuesta de la RNA. Existen gran variedad algoritmos de aprendizaje, sin embargo al elegir el modo de aprendizaje supervisado, se limita las opciones, y además el algoritmo debe elegirse de acuerdo al propósito que se desea de la RNA, en este caso la generalización de una función.

Entre los diferentes algoritmos de aprendizaje se elige \Rightarrow Aprendizaje de corrección de errores, en particular el de la Regla Delta (ó regla Widrow-Hoff), que al optimizarlo derivó en el algoritmo Levenberg-Marquardt.

➤ 3.B. Selección del modelo particular de la RNA.

Hasta este punto, se ha definido la mayoría de las características de la estructura general de la RNA. Sin embargo falta establecer la cantidad de neuronas (las unidades de procesamiento) en la capa oculta (intermedia) de la RNA. Por lo tanto del modelo general de la RNA, se hace una variación del número de neuronas en la capa oculta (intermedia), para elegir la última característica del modelo particular de la RNA de acuerdo a los resultados que se obtengan y si los resultados no son satisfactorios, modificarlas.

Se establece, inicialmente variar el número de neuronas de la capa oculta en: 4, 8, 16, 20, 30 y 40, respectivamente.

➤ 4. Entrenamiento de la RNA.

Una vez elegida la arquitectura de la RNA para un problema en particular, esta se encuentra lista para el proceso de aprendizaje. En este punto se llevará a cabo el entrenamiento de las propuestas de los modelos particulares de la RNA (modelos con características de funcionamiento y estructura iguales, pero únicamente con la variación entre sí del número de neuronas en la capa intermedia). Para el entrenamiento de estos modelos, se utilizarán los datos obtenidos después de un tratamiento previo, que quedaron agrupados en tres grupos: A, B y C, estos definirán las condiciones en las cuales la RNA aprenderá a generalizar la relación entre variables, de acuerdo al algoritmo de aprendizaje seleccionado previamente (Algoritmo Levenberg-Marquardt).

De las tres combinaciones de los datos acondicionados (A, B y C) y junto los seis modelos particulares de la RNA propuestos, las pruebas se organizan:

Prueba número 1:

- Grupo de datos (A).- Datos de entrada $p_A = \{Re_G, n, He\}$ y datos de salida $t_A = \{f_c\}$.
- Utilizando los modelos particulares de la RNA, que constan 4, 8, 16, 20, 30 y 40 neuronas en la capa intermedia, respectivamente.

Prueba número 2:

- Grupo de datos (B).- Datos de entrada de $p_B = \{\text{Log}(Re_G), \text{Log}(n), \text{Log}(He)\}$ y datos de salida $t_B = \{f_c\}$.
- Entrenamiento del modelo general de RNA con la variación del número de neuronas en la capa intermedia de: 4, 8, 16, 20, 30 y 40, respectivamente.

Prueba número 3:

- Grupo de datos (C).- Datos de entrada $p_C = \{\text{Log}(Re_G), \text{Log}(n), \text{Log}(He)\}$ y datos de salida $t_C = \{\text{Log}(f_c)\}$.

- Entrenamiento del modelo general de RNA con la variación del número de neuronas en la capa intermedia de 4, 8, 16, 20, 30 y 40, respectivamente.

➤ 5. Evaluación de la RNA.

Una vez entrenados los modelos particulares de la RNA, se realizará la evaluación estadística (MAE, STDA MRE STDR y R^2) comparando los resultados obtenidos por el modelo particular de la RNA (dichos resultados se identificarán como f_{fp}), contra el grupo de datos de f_{fc} , es decir los obtenidos por el método de bisección de la resolución de las ecuaciones (1.36) y (1.39) y que fueron utilizados para el entrenamiento de la RNA.

Tabla 2.1 Parámetros estadísticos a utilizar para la evaluación de la RNA.

SIGLA	SIGNIFICADO	ECUACIÓN	FINALIDAD
MAE	Error medio absoluto	$MAE = \frac{1}{N} \sum_{i=1}^N \Delta f_A$	Indica la imprecisión media en valores absolutos, que acompaña al resultado (el cálculo). Nos da idea de la sensibilidad del cálculo.
STDA	Desviación estándar muestral absoluta	$STDA = \sqrt{\frac{\sum_{i=1}^N (\Delta f_A - \overline{\Delta f_A})^2}{N - 1}}$	Indican el valor absoluto de cuanto se alejan (dispersan) las diferencias de los resultados de la muestra del valor promedio de error de la muestra.
MRE (%)	Error medio relativo	$MRE = \frac{1}{N} \sum_{i=1}^N \Delta f_R$	Sirve para comparar el valor verdadero y el aproximado, y conocer su porcentaje promedio de error.
STDR (%)	Desviación estándar muestral relativa	$STDR = \sqrt{\frac{\sum_{i=1}^N (\Delta f_R - \overline{\Delta f_R})^2}{N - 1}}$	Indican el porcentaje de cuanto se alejan (dispersan) los errores de la muestra del valor medio del error de la muestra.
R^2	Coefficiente de determinación	Regresión lineal	Determina la calidad del modelo para replicar los resultados, y la proporción de variación de los resultados que puede explicarse por el modelo (en este caso lineal).

Donde:

$$\Delta f_{Absoluto} = \Delta f_A = |f_{fp} - f_{fd}|$$

$$\Delta f_{\text{Relativo}} = \Delta f_R = \left| (f_{fp} - f_{fd}) / f_{fd} \right|$$

f_{fp} = factor de fricción predicho (aproximado).

f_{fd} = factor de fricción deseado (verdadero).

Objetivo particular 2:

Evaluar el desempeño del modelo de la red neuronal artificial para el cálculo del factor de fricción comparándolo sus resultados con el método iterativo para validar el funcionamiento de este modelo y obtener las ecuaciones lineales asociadas a este.

- 6. Selección de los mejores modelos de la RNA y sus condiciones.

En base a los resultados de la evaluación se podrá seleccionar cual de los modelos particulares de la RNA y junto con que condiciones de acondicionamiento de datos, se obtienen los mejores resultados del factor de fricción vía las RNA.

- 7. Validación del modelo de RNA.

Ya realizada la selección se continuará con la validación de ese modelo resultante con un grupo de datos nuevos, calculados con el mismo método y dentro de los mismos intervalos utilizados en la generación de datos (punto 1.A. de esta metodología), utilizados para el entrenamiento de la RNA. Este nuevo grupo de datos será en cantidad un 10% del total de los datos usados durante el entrenamiento.

La validación consistirá en comparar los valores del factor de fricción predicho (f_{fp}) que fueron obtenidos vía el modelo particular de RNA, contra los valores de factor de fricción calculado por el método de bisección (f_{fc}). Esta validación se basa en calcular parámetros estadísticos como: desviación estándar relativa, desviación estándar absoluta, error medio relativo, error medio absoluto y coeficiente de determinación (R^2).

- 8. Análisis de resultados.

- 9. Obtención de ecuaciones asociadas al modelo elegido y sus condiciones de entrenamiento.

Después de la validación de la RNA, se obtendrán las ecuaciones asociadas al mejor modelo que resuelva el objetivo general. Se consideraran las características generales y particulares del modelo seleccionado. Es decir en base a los ajustes ya obtenidos de los pesos de las conexiones de las neuronas (que son constantes) y el valor del *bias* (tendencia), y de cómo se procesa la información en la RNA, por capas y en cada una de las neuronas de cada capa, reproduciendo las operaciones matemáticas que realiza la regla de propagación, las funciones de transferencia y la función de salida usadas, todo esto aplicando, operaciones matemáticas básicas y algebra lineal (operaciones básicas entre matrices).

CAPÍTULO 3. Resultados

3.1 Análisis de resultados.

- 1.A. Generación de datos para el entrenamiento de la RNA.

Para la generación de datos que sirvieron de ejemplos para la Red neuronal artificial, se utilizaron las ecuaciones para calcular el factor de fricción en una tubería para un fluido tipo Herschel-Bulkley, presentadas por Hanks, $f_f = \frac{16}{\psi Re_G}$ para régimen laminar y Torrance,

$$\frac{1}{\sqrt{f_f}} = 0.45 - \frac{2.75}{n} + \frac{1.97}{n} \ln(1 - \xi_0) + \frac{1.97}{n} \ln \left(Re_G \left(\frac{1 + 3n}{4n} \right)^n f_f^{1-(n/2)} \right)$$

para régimen turbulento, que tradicionalmente se resuelven a través de iteraciones, sin embargo para contar con los suficientes datos representativos de los sistemas (pues de ello depende la capacidad de la RNA de generalizar el modelo para calcular el factor de fricción de fluidos Herschel-Bulkley) fue necesaria una gran cantidad de datos, los cuales era poco viables calcularlos por iteración, por lo tanto se optó por generarlos a través de la implementación de un procedimiento implícito como lo es el método de bisección, el cual es un método numérico para resolver ecuaciones algebraicas no-lineales y así poder resolver las ecuaciones para régimen laminar (1.36) y (1.39) para régimen turbulento. Estos datos fueron generados con ayuda de un programa básico para el entorno Matlab[®] del método de bisección (ver Anexo 2).

Los parámetros de entrada elegidos para la RNA fueron el número de Reynolds generalizado (Re_G), el índice de comportamiento al flujo (n) y el número de Hedstrom (He), pues en las ecuaciones (1.36) y (1.39) que corresponden a los regímenes de flujo laminar y turbulento respectivamente, el f_f queda en función de Re_G , n y He , que son las variables independientes que coinciden en ambos regímenes. Por supuesto el parámetro de salida de la RNA es f_f .

Los intervalos de datos considerados para cubrir ambos regímenes de flujo fueron:

- En régimen Laminar

Re_G con valores dentro del intervalo [100 a 2100]

n con valores dentro del intervalo [0.1 a 1]

He con valores dentro del intervalo [10 a 1×10^7]

- En régimen Turbulento

Re_G con valores dentro del intervalo [3000 a 1×10^7]

n con valores dentro del intervalo [0.1 a 1]

He con valores dentro del intervalo [10 a 1×10^7]

Estos intervalos se decidieron con la intención de cubrir el mayor espectro posible de f_f de acuerdo a las gráficas de relaciones de f_f en función de Re_G , n y He para fluidos Herschel-Bulkley basadas en las ecuaciones de Torrance y Hanks presentes en Garcia & Steffe (1986).

Las variables para la generación de datos queda de la siguiente manera:

- En régimen laminar, *variables independientes* $\{Re_G, n, He\}$; *variable dependiente* $\{f_f\}$.
- En régimen turbulento, *variables independientes* $\{Re_G, n, He\}$; *variable dependiente* $\{f_f\}$.

De las distintas combinaciones entre los parámetros se obtienen un total de **2660** datos dentro de los intervalos arriba establecidos, **1150 datos para régimen laminar** y **1510 para turbulento**.

➤ 2.A. Tratamiento previo de los datos

En esta etapa se siguió con lo propuesto en la metodología y para cada una de las tres pruebas de entrenamiento que se realizaron, se utilizó un tratamiento previo determinado.

- Pareja de datos (Grupo A).- Valores normales para las matrices de entrada y salida.
Datos de entrada, $p_A = \{Re_G, n, He\}$ y datos de salida, $t_A = \{f_{fc}\}$.
- Pareja de datos (Grupo B).- Valores con transformación logarítmica (calcular el logaritmo base 10 a los valores) para el sub-conjunto de entrada y sin transformación logarítmica para el sub-conjunto de salida.

Datos de entrada, $p_B = \{\text{Log}(Re_G), \text{Log}(n), \text{Log}(He)\}$ y datos de salida, $t_B = \{f_{fc}\}$.

- Pareja de datos (Grupo C).- Con transformación logarítmica para ambos sub-conjuntos.

Datos de entrada, $p_C = \{\text{Log}_{10}(Re_G), \text{Log}_{10}(n), \text{Log}_{10}(He)\}$ y datos de salida,

$t_C = \{\text{Log}_{10}(f_{fc})\}$.

- 1. B. Selección general de la RNA de acuerdo al problema.

De acuerdo al objetivo general del presente trabajo, se planteó la manera más apropiada para resolverlo, y con respecto a la Red neural artificial (RNA), primero se realizó la selección general de la RNA, en base a la definición de los siguientes puntos: Definir el propósito que tendría la RNA y la mejor manera de realizar la RNA.

1.B.1. Propósito de la RNA.

Después de haber realizado una revisión bibliográfica de las aplicaciones que hasta ahora se habían hecho de las RNA en la industria alimentaria y de los fundamentos del funcionamiento de las RNA, se decide que el propósito que se busca de la RNA, es la generalización de una función ó una optimización.

1.B.2. Modo de realizar la RNA.

En base a la revisión bibliográfica del tema y un estudio básico de la disponibilidad y costo de las RNA, se seleccionó la manera más adecuada para realizar la RNA, y se elige utilizar un simulador de RNA, debido a las diversas ventajas que ofrece en comparación a los otros modos de realización (emulador y hardware), ventajas de tipo económicas (es el de más bajo costo), tecnológicas (fue creado por expertos en el tema), accesibilidad (solo se necesita un ordenador de propósito general adecuado para su funcionamiento, que cumpla con los requerimientos del sistema y conocimiento muy básico en programación), disponibilidad (se puede descargar tras el pago correspondiente, desde la página <http://es.mathworks.com>), rapidez (solo hay que conseguirlo e instalarlo), etcétera. Por todas estas ventajas que ofrece se opta por un software comercial ampliamente utilizado como lo es el Neural Network Toolbox™ que forma parte del entorno MATLAB® en este caso versión R2008a.

➤ 2.B. Características de la RNA.

Las características generales de una RNA se definen en base a su arquitectura y tipo de aprendizaje.

2.B.1. Arquitectura de la RNA.

La arquitectura de la RNA, define la mayoría de las características que tendrá la RNA, pues define el flujo de la información y el procesamiento de esta. La arquitectura de una RNA se divide básicamente en dos niveles: meso-estructura y micro-estructura. Cabe indicar que se respetaron todas las características propuestas en la metodología.

- En la meso-estructura de la RNA, se opta por una red multicapa con propagación de la información unidireccional y hacia adelante (feed forward), las redes multicapa tiene gran capacidad para generalizar casi cualquier función, y se determina que solo tendrá en su estructura tres capas (capa de entrada-capa oculta-capa de salida), en la capa de entrada se tienen tres neuronas debido al número de parámetros de entradas pues cada neurona corresponde a un parámetro (Re_G , n y He) y en la capa de salida solo hay una correspondiente a f_{fp} que es el parámetro de salida de la RNA. El número de neurona en la capa intermedia (oculta) se define más adelante.
- En la micro-estructura de la RNA, al ser el propósito de la RNA el generalizar una función y al tratarse de una RNA multicapa, entonces como lo señala el Teorema de aproximación universal (ver sección 1.3.3.2) una RNA multicapa debe tener funciones de transferencia no lineales para agregar capacidad computable, por ende la función de transferencia en la capa intermedia se decide que será la tangente hiperbólica, la cual es una función no-lineal, y es descrita por la ecuación (1.55), además se elige que la regla de propagación sea la suma ponderada y como función de salida se elige la función lineal (ó función identidad). En la capa de entrada, no hay ninguna regla de propagación, ni función de transferencia, ni función de salida, pues la neuronas de la primer capa son los valores de las parámetros de entrada $\{Re_G, n, He\}$. En el caso de la capa de salida, la regla de propagación es la suma ponderada y la función de transferencia y la función de salida es la función lineal descrita por la ecuación (1.49). Además se decide que haya

presencia de *bias* (tendencia ó umbral) para las entradas a la capa intermedia y a la capa de salida.

2.B.2. Tipo de aprendizaje de la RNA.

Se designa realizar un entrenamiento de la RNA de tipo supervisado, debido a que se adecua perfectamente a los ejemplos disponibles para usarse en aprendizaje tipo supervisado, donde la organización de datos, que es descrita por la ecuación (1.58) presenta un acomodo de parejas de datos, de la manera {datos de entrada, datos de salida objetivo}, de los cuales se hace un grupo y se presenta a la RNA en un lote de datos (en Batch), es decir se le provee a la red de todos los datos al mismo tiempo y en un grupo, además esta manera es la más sencilla de entrenar la RNA.

2.B.2.1. Algoritmo de aprendizaje.

Se opta por el algoritmo de aprendizaje recomendado por los creadores Neural Network Toolbox™, en este caso el Levenberg-Marquardt, el cual es básicamente una regla de aprendizaje por corrección de errores, y es un modelo generalizado del algoritmo de aprendizaje de la regla Delta (también conocida como algoritmo Widrow-Hoff), que usa el gradiente descendente para reducir el error, pero a diferencia del algoritmo Widrow-Hoff es mucho más veloz para converger gracias a la optimización de técnicas numéricas (ver Anexo 3).

- 3.B. Selección del modelo particular de la RNA (estructura de la capa intermedia del modelo de la RNA).

Una vez obtenidos los datos (ejemplos) para el entrenamiento, y la estructura general de la RNA, que se utilizará, el siguiente paso es definir el modelo particular en lo concerniente al número de neuronas de la capa intermedia. Se proponen inicialmente seis modelos con variaciones del número de neuronas en la capa intermedia de: 4, 6, 8, 16, 20, 30 y 40 respectivamente. En esta etapa por las características inherentes a las RNA y al no contar con una guía clara, pues solo se conoce que entre más neuronas en la capa intermedia (oculta) mayor capacidad para generalizar, se prueban diferentes modelos en base a prueba y error.

➤ 4. Entrenamiento de la RNA.

Para crear la RNA se utilizó el software Neural Network Toolbox™, el cual es un simulador donde se fueron estableciendo paso a paso cada una de las características antes descritas, en particular para este propósito y con estas características de funcionamiento existe en el Neural Network Toolbox™ del entorno MATLAB®, el interfaz gráfico Neural Network Fitting Tool GUI (en el Anexo 1, se describe su operación).

Antes de iniciar el entrenamiento fue necesario hacer una división de los datos en porcentajes pues así lo requiere el software, ya que como parte del funcionamiento de Neural Network Toolbox, este divide a los datos aleatoriamente en tres partes una de ellas para su entrenamiento, otra para su validación interna (pues este parámetro le indica a la red que esta generalizando y detiene su entrenamiento antes de sobre ajustar ó memorizar los datos) y la última parte es para hacer una prueba de la generalización. La RNA inmediatamente después de ser entrenada, presenta los datos de los parámetros más importantes de su entrenamiento y esta información sirve para decidir si se acepta o no el entrenamiento dando la opción de re-entrenar ó modificar al algún parámetro como la proporción de las divisiones, el número de neuronas y aumentar el número de datos.

En esta etapa al no tener alguna referencia determinante al respecto (más allá del que se prefiere el menor error medio posible y el coeficiente de determinación de una regresión lineal, más cercano a uno), se opta por la prueba y error. Después de varias pruebas preliminares se determina que la mejor proporción de la división de datos era 80% para entrenamiento, 15% para la validación interna y 5% para la prueba para generalización (que es el porcentaje mínimo que acepta el software) además se decide en solo re-entrenar la RNA cuando fuera necesario, cuando los datos de su desempeño de entrenamiento, eran muy alejados de los resultados deseados y se probaría una y otra vez hasta mejorarlos, las demás opciones se mantienen estables.

De acuerdo a la metodología propuesta se realizó pruebas de cada grupo de datos, resultantes del tratamiento previo, organizados en: Grupos A, B y C; que se consideran las condiciones en las que se entrena la RNA y de las cuales tuvo que aprender, y por otro lado la estructura y características del funcionamiento de cada modelo en particular, modelos con la única diferencia entre sí del número de neuronas en la capa intermedia.

Prueba número 1

- *Variable independiente 1:* Estructura de la RNA, al hacer variación del número de neuronas en la capa intermedia en 4, 8, 16, 20, 30 y 40 neuronas.
- *Variable independiente 2:* Condiciones de entrenamiento, utilizando al Grupo A de datos, Datos de entrada= $\{ Re_G, n, He \}$ y datos de salida= $\{ f_{fc} \}$.
- *Variable dependiente:* factor de fricción predicho por la RNA (f_{fp}).

Prueba número 2

- *Variable independiente 1:* Estructura de la RNA, al hacer variación del número de neuronas en la capa intermedia en 4, 8, 16, 20, 30 y 40 neuronas.
- *Variable independiente 2:* Condiciones de entrenamiento, utilizando al Grupo B de datos, Datos de entrada= $\{ \text{Log}(Re_G), \text{Log}(n), \text{Log}(He) \}$; Datos de salida = $\{ f_{fc} \}$.
- *Variable dependiente:* factor de fricción predicho por la RNA (f_{fp}).

Prueba número 3

- *Variable independiente 1:* Estructura de la RNA, al hacer variación del número de neuronas en la capa intermedia en 2, 4, 8, 16 y 20 neuronas.
- *Variable independiente 2:* Condiciones de entrenamiento, utilizando al Grupo C, Datos de entrada= $\{ \text{Log}(Re_G), \text{Log}(n), \text{Log}(He) \}$, Datos de salida= $\{ \text{Log}(f_{fc}) \}$.
- *Variable dependiente:* factor de fricción predicho por la RNA (f_{fp}).

➤ 5. Evaluación de la RNA.

A continuación se muestran las variables de cada prueba y sus respectivos resultados de cuatro parámetros estadísticos y el coeficiente de determinación de una regresión lineal, entre los valores del factor de fricción calculados por el modelo de RNA (f_{fp}) y la salida deseada, es decir el valor del factor de fricción calculado por las ecuaciones de Hanks y Torrance (f_{fc}), estos dos valores fueron los utilizados para evaluar el desempeño de la RNA. Cabe señalar que se eligieron considerando las características del problema y tomando como referencia la metodología propuesta en Sablani & Shayya (2003).

Prueba número 1

En esta prueba se utilizan los datos obtenidos en la generación de datos sin ningún tipo de tratamiento previo y se agrupan en un conjunto de datos como el grupo A.

Como se puede observar en la Tabla 3.1, todos los resultados están por encima de valores deseados, pues los valores de MAE, STDA, MRE y STDR, son muy grandes y lo que se busca es lo contrario, que no haya grandes diferencias entre un método y otro. Por lo tanto estos resultados indican que ningún modelo se acerca a los resultados del método iterativo de cálculo. Cabe señalar que en los modelos con más neuronas en la capa intermedia, hay una pequeña mejoría en los resultados, sin embargo comparando el tamaño de los errores con la mejoría, esta es insignificante. Los errores tanto MAE y MRE son demasiado grandes para ser considerados, además están lejos de representar de manera favorable algún tipo de relación lineal entre si, por lo cual, se infiere que la relación entre parámetros es de tipo no-lineal pues los valores de R^2 son muy lejanos a 1, y este resultado no es adecuado para el propósito de este trabajo.

Tabla 3.1 Resultados prueba 1

	Variaciones de la estructura de la RNA en la capa intermedia					
	4 neuronas	8 neuronas	16 neuronas	20 neuronas	30 neuronas	40 neuronas
MAE	98.600	79.300	52.510	29.190	36.490	32.290
STDA	84.990	49.290	37.910	31.420	43.980	45.180
MRE (%)	302.200	246.900	123.190	98.710	127.460	136.740
STDR (%)	317.910	169.500	105.200	87.830	94.720	73.310
R^2	0.039	0.043	0.210	0.390	0.114	0.103

Prueba número 2

En esta prueba, se realiza un tratamiento previo de datos, al conjunto se nombra Grupo B. Y el tratamiento previo antes de presentarse a la RNA, consistió en aplicar al sub-conjunto de datos de entrada, el cálculo del Logaritmo en base 10, es decir datos de entrada= $\{\text{Log}_{10}(Re_G), \text{Log}_{10}(n), \text{Log}(He)\}$. Los resultados de esta prueba se muestran en la Tabla 3.2.

Con estos resultados, se puede afirmar que hay una mejoría notable en la linealización con los transformación logarítmica de los datos de entrada, no obstante los resultados aún están lejos de unos datos que puedan utilizarse para el objetivo del problema, pues a pesar de la mejoría de resultados, incluso los modelos con mayor número de neuronas en la capa intermedia, presentan evaluaciones poco favorables. Sin embargo nos afirman de acuerdo a lo que se conoce por las gráficas conocidas de factor de fricción en función del número de Reynolds, número de Hedstrom y n , que la relación entre ellos es de tipo logarítmica, y puesto que la modelación que hace la RNA es una “caja negra” en muchos aspectos, conocer o tener alguna idea de la relación ayuda a mejorar el desempeño de la RNA y por lo tanto la generalización.

Tabla 3.2 Resultados prueba 2

	Variaciones de la estructura de la RNA en la capa intermedia					
	4 neuronas	8 neuronas	16 neuronas	20 neuronas	30 neuronas	40 neuronas
MAE	0.1885	0.2167	0.0602	0.0224	0.0957	0.1861
STDA	0.1410	0.1911	0.0830	0.0228	0.1041	0.1654
MRE (%)	37.8441	44.4325	9.7786	5.8197	24.6831	43.2625
STDR (%)	56.5581	85.3656	19.0838	17.8483	82.6883	87.3865
R^2	0.5100	0.0790	0.7910	0.9570	0.4150	0.4130

Por último cabe mencionar que la mejor generalización de la relación de parámetros de entrada= $\{\text{Log}_{10}(Re_G), \text{Log}_{10}(n), \text{Log}(He)\}$ con f_f la obtuvo el modelo con 20 neuronas con un MAE de 0.022 y un MRE de 5.819 % y un R^2 de 0.957. Sí bien el resultado es mucho mejor que en la prueba 1, el tamaño del modelo es grande, y aún así no alcanza valores de los parámetros estadísticos muy adecuados, pues la STDR con 17.848 % se considera alto.

Prueba número 3

En esta prueba se utilizaron los datos del Grupo C, que contaban con un tratamiento previo, en el cual se calculó el Log_{10} a los sub-conjuntos de entrada y de salida (datos de entrada= $\{\text{Log}(Re_G), \text{Log}(n), \text{Log}(He)\}$, datos de salida= $\{\text{Log}(f_{fc})\}$).

En esta última prueba, teniendo el conocimiento a posteriori del desempeño de los modelos antes probados con respecto a las neuronas de la capa intermedia se opta por reducir los modelos a 2, 4, 8, 16 y 20 neuronas en la capa intermedia (oculta). Puesto que más de 20 neuronas en esa capa reflejan muy poca mejoría y además representan una desventaja por ser modelos muy grandes, y puesto que si bien es cierto que a mayor número de neuronas mayor capacidad computable, también significa que existiría mayor cantidad de constantes y en consecuencia sería un modelo poco viable para el objetivo de este trabajo.

Tabla 3.3 Resultados prueba 3

	Variaciones de la estructura de la RNA en la capa intermedia				
	2 neuronas	4 neuronas	8 neurona	16 neuronas	20 neuronas
MAE	0.1422	0.0740	0.0360	0.0218	0.0146
STDA	0.1312	0.0690	0.0423	0.0319	0.0290
MRE (%)	0.1829	0.1808	0.0646	0.0493	0.0231
STDR (%)	1.8385	3.4590	0.8097	1.1241	0.2801
R ²	0.9630	0.9900	0.9970	0.9980	0.9990

La mejoría del desempeño de todos los modelos de RNA, en esta prueba es notable con respecto a las anteriores pues en todos los modelos se tienen mejoría en términos generales con respecto al mejor de la prueba número 2, sobre todo en coeficiente de determinación y MRE pues en esta prueba todos los modelos están por debajo de 0.2%. Además la STDR de todos los modelos es más pequeña que el mejor de la prueba 2, pues están por debajo de 3.5%.

Los resultados de la prueba número 3, confirman que la relación entre parámetros más adecuada es la logarítmica y el mejor desempeño lo obtuvo el modelo de 20 neuronas con un MAE de 0.01463, un MRE de 0.02312%, un coeficiente de determinación de 0.999 y una STDR de 0.28%.

Como se puede observar en la Figura (3.1) el modelo de 20 neuronas es el mejor pues tiende a la linealización perfecta con una R² de 0.999 al graficar $\text{Log}(f_c)$ vs $\text{Log}(f_{fp})$, estando en el eje de

las abscisas el Log (f_{fc}) y en el de las ordenadas Log (f_{fp}). Cabe señalar que como se muestra en la gráfica hay un pequeño número de datos fuera de la linealización, estos datos corresponde a datos de entrada del He con valores mayores o iguales 1×10^6 , este error se presenta en todos los modelos de esta prueba, sin embargo no en todas las entradas con valor de He mayor o igual a 1×10^6 , pues solo ocurre en el 50% de las veces. Lo que significa que a partir de este valor del número de Hedstrom, la RNA es incapaz de lograr una generalización adecuada. No obstante para el resto de los intervalos para la que la RNA fue entrenada, esta puede realizar una generalización apropiada.

➤ 6. Selección de los mejores modelos de la RNA y condiciones.

De acuerdo a los resultados de las tres pruebas, en la prueba 3 se obtuvieron los mejores modelos y condiciones para la RNA capaz de calcular el factor de fricción de un fluido Herschel-Bulkley en ambos regímenes de flujo, de esta prueba el modelo 20 neuronas es el modelo con la mejor generalización, pero también es el modelo más grande de la prueba 3; pues hay que recordar que entre más grande es la RNA más aumenta su capacidad para la generalización de una función, sin embargo al ser el modelo más grande, inherentemente tendrá más neuronas y esto provoca que aumente el número de conexiones por neurona y pesos relacionados a estos, los que se convertirían en una gran cantidad de coeficientes (constantes), que aumentarían el número de cálculos, lo cual resultaría poco práctico pues tendría ecuaciones algebraicas muy largas, con lo cual no resolvería el problema de reducción de tiempo de cálculo del factor de fricción. Para cumplir con este fin los modelos de 4 neuronas y 8 neuronas de la prueba 3, serían más apropiados por su sencillez, además que considerando sus resultados en la evaluación estadística se encuentran dentro de intervalos aceptables, pues ambos tienen coeficientes de determinación mayores al 0.95 y el error medio relativo (MRE) es de 0.18% y 0.06% respectivamente, estos resultados de la evaluación los convierten en modelos viables, para que en base a ellos se desarrolle el modelo explícito del cálculo de factor de fricción.

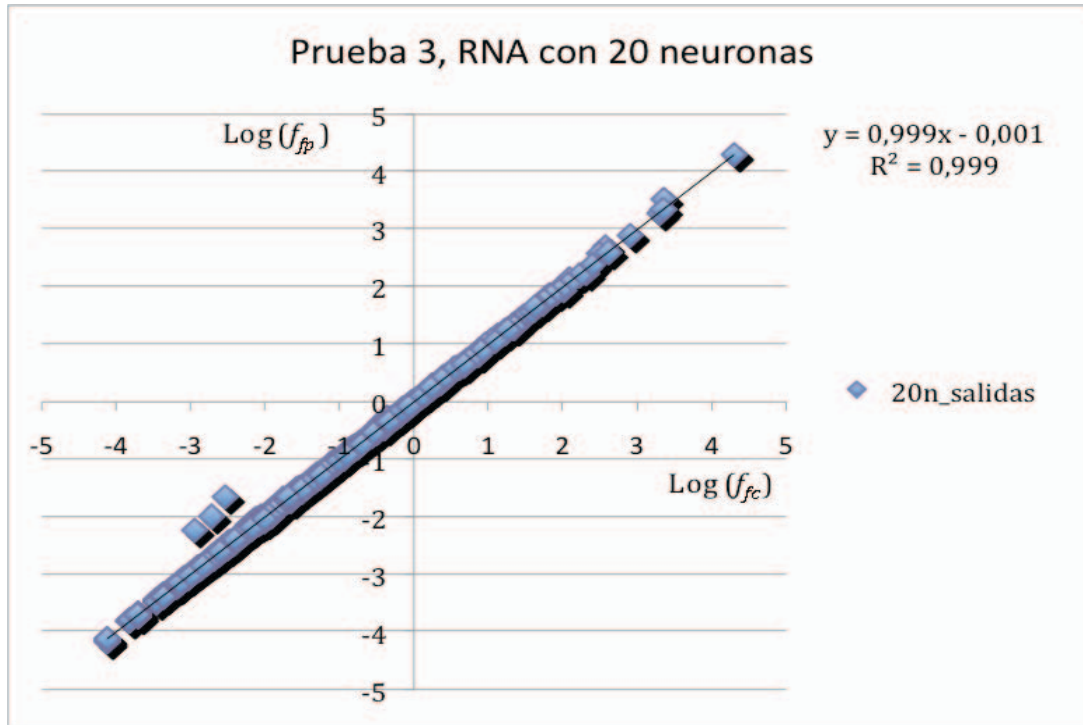


Figura 3.1 Valores del logaritmo del factor de fricción calculado por ecuaciones “Log (f_{fc})” vs valores del logaritmo del factor de fricción obtenido por la RNA “Log (f_{fp})” del modelo de 20 neuronas.

3.2 Validación del modelo de obtenido de la RNA.

➤ 7. Validación del modelo de RNA.

Para la validación fue necesario utilizar un grupo nuevo de datos, que no se hubieran ocupado anteriormente en el entrenamiento, estos datos fueron obtenidos de la misma manera que los del entrenamiento (por el método de la bisección resolviendo las ecuaciones (1.36) para régimen laminar y (1.39) para régimen turbulento) y por supuesto con los mismos parámetros y dentro de los mismos intervalos con los que la RNA fue entrenada. Esta vez se validaron los datos comparando, los datos obtenidos vía método de bisección (f_{fc}) vs datos calculados (predichos) por la RNA (f_{fp}).

Una vez elegidos los mejores modelos para cumplir con el objetivo general de este trabajo, se prosigue con la validación de los modelos, así como del mejor modelo de 20 neuronas, para tomarlo como una referencia y buscar siempre acercarse a estos parámetros.

Tabla 3.4 Validación de los mejores modelos

Resultados de validación			
	Variaciones de la estructura de la RNA en la capa intermedia		
	4 neuronas	8 neuronas	20 neuronas
MAE	0.0774	0.0514	0.0285
STDA	0.1338	0.1162	0.1132
MRE (%)	0.0600	0.0355	0.0164
STDR (%)	0.1198	0.0630	0.0464
R^2 de $\text{Log}(f_{fc})$ vs $\text{Log}(f_{fp})$	0.9540	0.9680	0.9730
R^2 de f_{fc} vs f_{fp}	0.9120	0.9900	0.9970

El error medio relativo (MRE) del modelo de 4 neuronas en la capa intermedia es de 0.06% y el de 20 neuronas es de 0.01%, recordando que este valor representa su porcentaje promedio de error, lo que refleja en ambos casos un porcentaje de error muy pequeño. También como se puede observar la mejor linealización en la validación fue del mejor modelo de 20 neuronas, sin embargo el modelo de 4 neuronas (el modelo más sencillo que se validó) se encuentra cercano a los valores del modelo de 20 neuronas en la capa oculta, con R^2 de 0.954. Sin olvidar que el error que existe de los puntos fuera de la linealización, son únicamente cuando el valor de $He \geq 1 \times 10^6$, valores que son poco probables que se presenten en los fluidos Herschel-Bulkley alimenticios, hasta ahora conocidos.

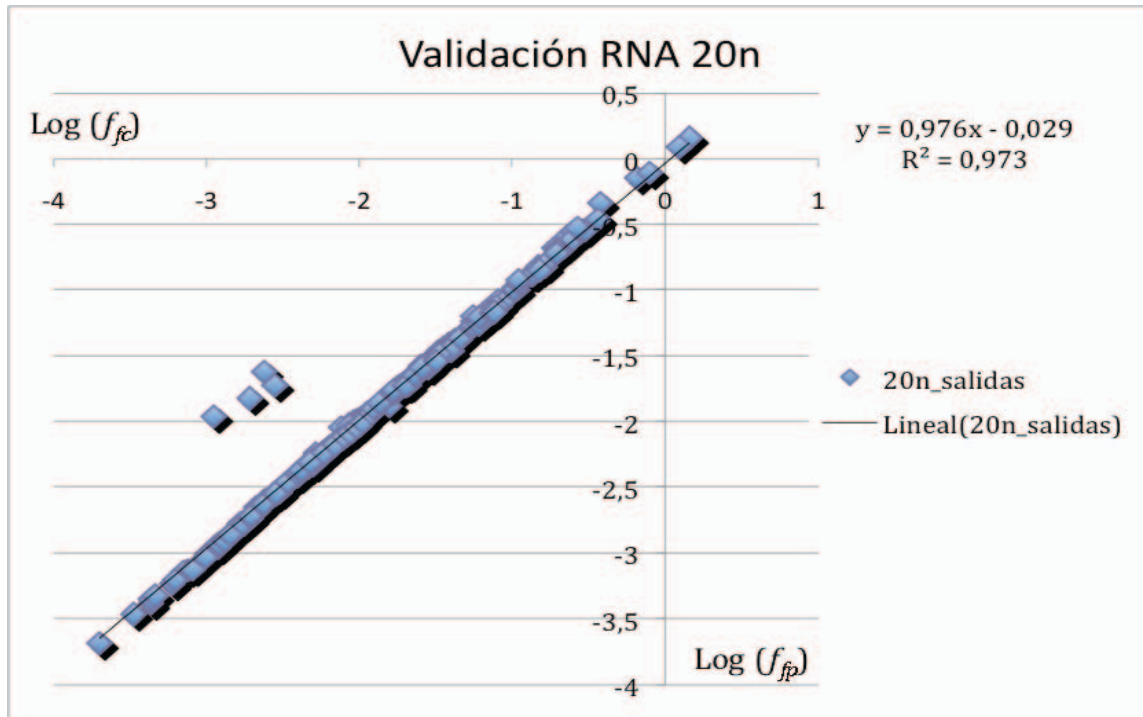


Figura 3.2 Valores del logaritmo del factor de fricción obtenido por ecuaciones “ $\text{Log}(f_{fc})$ ” vs valores del logaritmo del factor de fricción obtenido por la RNA “ $\text{Log}(f_{fp})$ ”, validación para el modelo RNA de 20 neuronas.

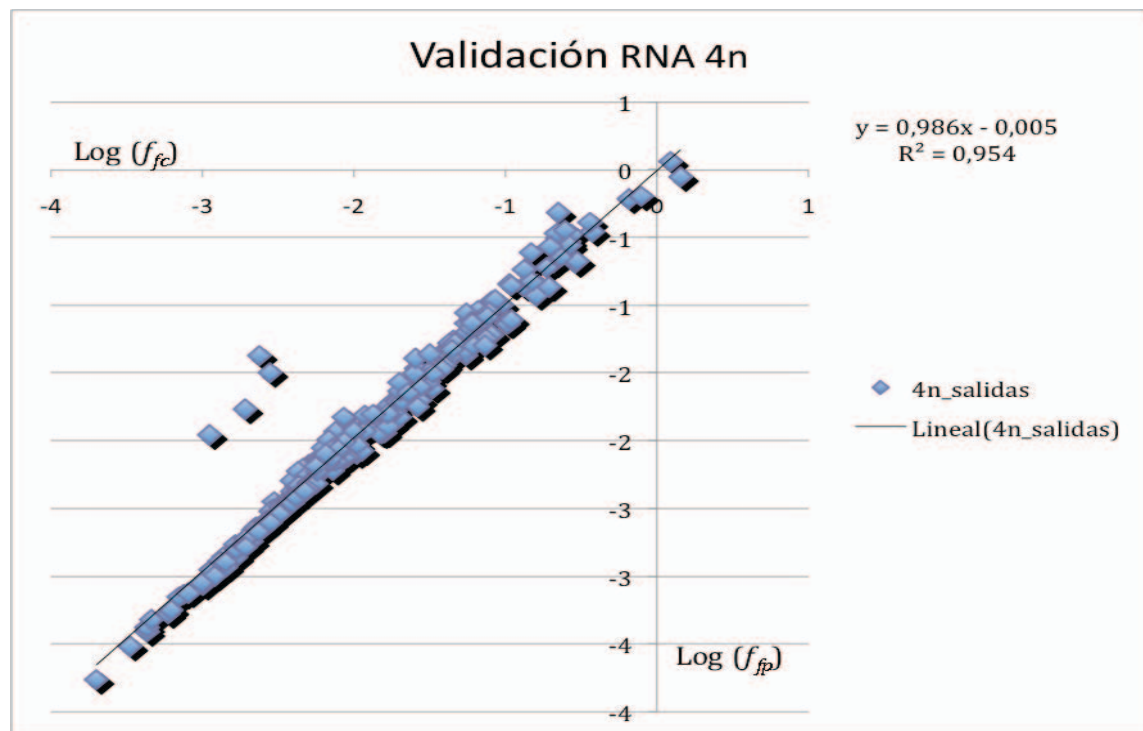


Figura 3.3 Valores del logaritmo del factor de fricción obtenido por ecuaciones “ $\text{Log}(f_{fc})$ ” vs valores del logaritmo del factor de fricción obtenido por la RNA “ $\text{Log}(f_{fp})$ ”, validación para el modelo RNA de 4 neuronas.

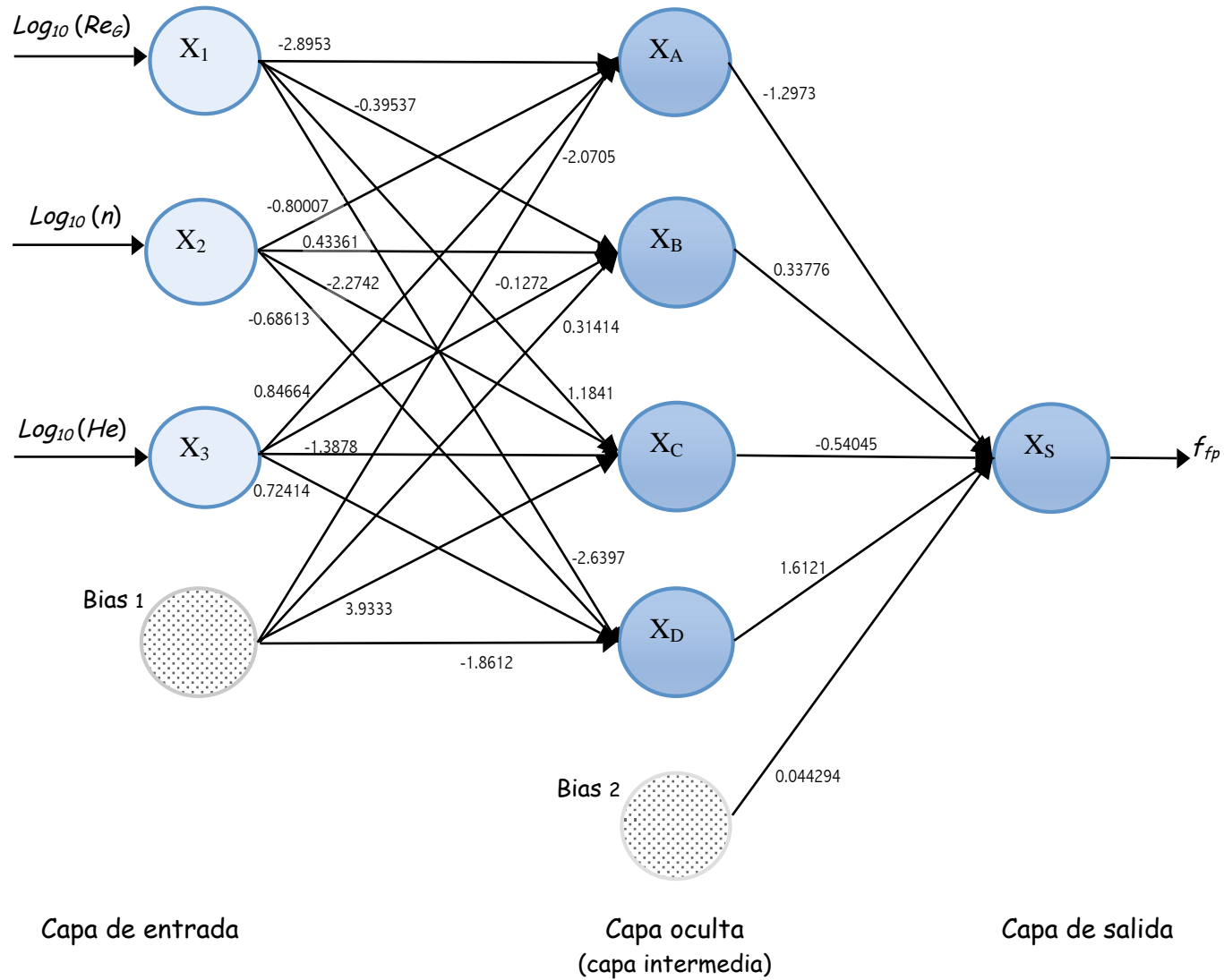


Figura 3.4 Diagrama esquemático de la estructura de Red neuronal artificial de 4 neuronas en la capa oculta.

3.3 Determinación de las ecuaciones explícitas para el cálculo del factor de fricción de fluidos Herschel-Bulkley en régimen laminar y turbulento.

- 9. Obtención de ecuaciones asociadas al modelo elegido y sus condiciones de entrenamiento.

Parte esencial de este trabajo es encontrar una alternativa al cálculo del factor de fricción por iteración y generar una herramienta de cálculo simple que solo requiera operaciones algebraicas lineales, así como reducir el tiempo que se consume al iterar por el método tradicional de prueba y error, utilizando las ecuaciones de Hanks y Torrance para el cálculo de f_f en régimen laminar y turbulento respectivamente. Considerando lo anterior se determina que el modelo más adecuado para cumplir con tal propósito es el modelo con 4 neuronas en la capa intermedia, debido a su mayor simplicidad y la valoración de la validación, con resultados aceptables de grado de error medio absoluto (MAE) de 0.0774, error medio relativo (MRE) de 0.06% y su R^2 de 0.954 que en comparación del mejor modelo con 20 neuronas en la capa intermedia tiene un R^2 de 0.973, lo que representa una diferencia muy pequeña entre ellos. La desviación estándar muestral absoluta (STDA) para este modelo es de 0.133 y la desviación estándar muestral relativa (STDR) 0.119%.

La obtención de las ecuaciones asociadas al modelo más apropiado (4 neuronas en la capa intermedia) para calcular el f_f de manera explícita y reducir del tiempo de cálculo por iteración, es la siguiente: Se toman las constantes desarrolladas por el modelo de RNA (que ya es capaz de generalizar adecuadamente la función, ver Figura 3.4) y se realizan las mismas operaciones algebraicas que realiza la RNA en su simulador, paso a paso se procesa la información desde las entradas, tratamiento previo de los datos, la entrada (*net*) a cada neurona, se aplica la función de transferencia (función tangente hiperbólica) y función de salida (función lineal) de cada neurona y posteriormente por capa, de acuerdo a las características del modelo.

La ecuación explícita para calcular de manera no-iterativa el factor de fricción en una tubería para fluidos Herschel-Bulkley tanto en régimen laminar como turbulento, esta en función de los siguientes parámetros:

El número de Reynolds generalizado, el índice de comportamiento al flujo y el número de Hedstrom, respectivamente representados con las siglas Re_G , n y He , y queda expresada de la siguiente forma:

$$f_{fp} = 10^{\left[4.211748 \left(\left(\tanh \left[\left([0.4(\text{Log}(Re_G) - 2)] - 1 \right) * (-2.8953) + \left([2(\text{Log}(n) + 1)] - 1 \right) * (-0.80007) + \left(\left[\frac{2}{7}(\text{Log}(He) - 1) \right] - 1 \right) * (0.84664) - 2.0705 \right] \right) * (-1.2973) + \right. \right. \\ \left. \left(\tanh \left[\left([0.4(\text{Log}(Re_G) - 2)] - 1 \right) * (-0.39537) + \left([2(\text{Log}(n) + 1)] - 1 \right) * (0.43361) + \left(\left[\frac{2}{7}(\text{Log}(He) - 1) \right] - 1 \right) * (-0.1272) + 0.31414 \right] \right) * (0.33776) + \right. \\ \left. \left(\tanh \left[\left([0.4(\text{Log}(Re_G) - 2)] - 1 \right) * (1.1841) + \left([2(\text{Log}(n) + 1)] - 1 \right) * (-2.2742) + \left(\left[\frac{2}{7}(\text{Log}(He) - 1) \right] - 1 \right) * (-1.3878) + 3.9333 \right] \right) * (-0.54045) + \right. \\ \left. \left(\tanh \left[\left([0.4(\text{Log}(Re_G) - 2)] - 1 \right) * (-2.6397) + \left([2(\text{Log}(n) + 1)] - 1 \right) * (-0.68613) + \left(\left[\frac{2}{7}(\text{Log}(He) - 1) \right] - 1 \right) * (0.72414) - 1.8612 \right] \right) * (1.6121) + 0.044294 + 1 \right] - 4.1215973 \right)$$

Donde, f_{fp} es el factor de fricción obtenido por el modelo elegido de las redes neuronales artificiales. Cabe señalar que esta ecuación al ser obtenida por RNA, solo puede ser utilizada dentro de las condiciones bajo las cuales fue entrenada la RNA y que es dentro de estas condiciones donde es completamente apta, los intervalos son: $Re_G = [100 \text{ a } 1 \times 10^7]$, $n = [0.1 \text{ a } 1]$ y $He \geq 10$ a $He < 10^6$.

Como puede notarse la ecuación es muy larga y poco práctica para ser utilizada así, debido a la gran cantidad de constantes y operaciones (que aunque simples son numerosas y continuas), pues esto eleva la posibilidad de que se cometan errores al resolverla, por lo tanto se optó por dividirla en pasos simples y continuos para una resolución menos compleja. A continuación se muestran los pasos:

Ecuaciones explícitas para el cálculo de f_f para fluidos tipo Herschel-Bulkley, aptas cuando $Re_G = 100$ a 1×10^7 , $n = 0.1$ a 1 y $He \geq 10$ a $He < 10^6$.

Donde:

Re_G = número de Reynolds generalizado

n = índice de comportamiento al flujo

He = número de Hedstrom

Paso 1. Entradas

$$Y_1 = \text{Log}_{10}(Re_G)$$

$$Y_2 = \text{Log}_{10}(n)$$

$$Y_3 = \text{Log}_{10}(He)$$

Paso 2. Escalamiento de parámetros de entrada.

$$X_1 = [0.4(Y_1 - 2)] - 1$$

$$X_2 = [2(Y_2 + 1)] - 1$$

$$X_3 = [2/7(Y_3 - 1)] - 1$$

Paso 3. Estimación de la salida de la capa oculta

$$X_A = \tanh[X_1(-2.8953) + X_2(-0.80007) + X_3(0.84664) - 2.0705]$$

$$X_B = \tanh[X_1(-0.39537) + X_2(0.43361) + X_3(-0.1272) + 0.31414]$$

$$X_C = \tanh[X_1(1.1841) + X_2(-2.2742) + X_3(-1.3878) + 3.9333]$$

$$X_D = \tanh[X_1(-2.6397) + X_2(-0.68613) + X_3(0.72414) - 1.8612]$$

Paso 4. Estimación de salida en la capa de salida

$$X_s = X_A(-1.2973) + X_B(0.33776) + X_C(-0.54045) + X_D(1.6121) + 0.044294$$

Paso 5. Reescalando la salida

$$f_{fp} = 10^{([4.2117482(Y_s + 1)] - 4.1215973)}$$

3.3.1 Ejemplo de cálculo del factor de fricción para un fluido Herschel-Bulkley, con las ecuaciones obtenidas a través de la RNA.

Para calcular el *factor de fricción*, utilizando las ecuaciones arriba presentadas, se retomará el ejemplo de la sección (1.2.2.1), donde se calculó el *factor de fricción* de un fluido con las siguientes características y condiciones:

$$n=0.6$$

$$K=4.8 \text{ Pa}\cdot\text{s}^n$$

$$\sigma_0=10 \text{ Pa}$$

$$\rho=1320.9 \text{ Kg/m}^3$$

$$v=1.84 \text{ m/s}$$

$$\text{Diámetro interno}= 0.0779 \text{ m}$$

De acuerdo al primer paso del cálculo de f_{fp} , para iniciar se requieren los valores de Re_G , n y He . En este caso como no se proporcionan Re_G y He , estos se calculan con las ecuaciones (1.14) para Re_G y (1.35) para He .

Utilizando las ecuaciones (1.14) y (1.35), se obtiene:

$$Re_G = \frac{(0.0779 \text{ m})^{0.6} (1.84 \text{ m/s})^{(2-0.6)} (1320.9 \text{ Kg/m}^3)}{8^{(0.6-1)} (4.8 \text{ Pa}\cdot\text{s}^{0.6})} \left(\frac{4(0.6)}{(1+3(0.6))} \right)^{0.6} = 292.66$$

$$He = \frac{(0.0779 \text{ m})^2 (1320.9 \text{ kg/m}^3)}{4.8 \text{ Pa}\cdot\text{s}^{0.6}} \left(\frac{10 \text{ Pa}}{4.8 \text{ Pa}\cdot\text{s}^{0.6}} \right)^{\frac{2-0.6}{0.6}} = 9.25708$$

Conociendo los valores de Re_G , n y He , se inicia el paso 1

$$Re_G = 292.66$$

$$n = 0.6$$

$$He = 9.25708$$

Paso 1. Entradas.

Se sustituyen los valores Re_G , n y He , en las ecuaciones del paso 1.

$$Y_1 = \text{Log}_{10}(292.66) = 2.46636$$

$$Y_2 = \text{Log}_{10}(0.6) = -0.221848$$

$$Y_3 = \text{Log}_{10}(9.25708) = 0.966474$$

Paso 2. Escalamiento de parámetros de entrada.

Los valores obtenidos del paso 1, Y_1 , Y_2 y Y_3 , se sustituyen en las ecuaciones del paso 2.

$$X_1 = [0.4(2.46636 - 2)] - 1 = -0.813454$$

$$X_2 = [2(-0.221848 + 1)] - 1 = 0.556302$$

$$X_3 = [2/7(0.966474 - 1)] - 1 = -1.00957$$

Paso 3. Estimación de la salida de la capa oculta.

Los valores de X_1 , X_2 y X_3 obtenidos del paso 2, se sustituyen en las ecuaciones del paso 3.

$$X_A = \tanh[(-0.813454)(-2.8953) + (0.556302)(-0.80007) + (-1.00957)(0.84664) - 2.0705] = -0.767877$$

$$X_B = \tanh[(-0.813454)(-0.39537) + (0.556302)(0.43361) + (-1.00957)(-0.1272) + 0.31414] = 0.763849$$

$$X_C = \tanh[(-0.813454)(1.1841) + (0.556302)(-2.2742) + (-1.00957)(-1.3878) + 3.9333] = 0.995997$$

$$X_D = \tanh[(-0.813454)(-2.6397) + (0.556302)(-0.68613) + (-1.00957)(0.72414) - 1.8612] = -0.678697$$

Paso 4. Estimación de salida en la capa de salida.

Los valores de X_A , X_B , X_C y X_D , obtenidos del paso 3 se sustituyen en las ecuaciones del paso 4.

$$X_s = (-0.767877)(-1.2973) + (0.763849)(0.33776) + (0.995997)(-0.54045) + (-0.678697)(1.6121) + 0.044294 = -0.333956$$

Paso 5. Reescalando la salida.

Finalmente el valor de X_s , obtenido del paso 4 se sustituye en la ecuación del paso 5, y así obtenemos el valor de f_{fp} .

$$f_{fp} = 10^{([4.2117482 (-0.333956+1)] - 4.1215973)} = 0.04826$$

$$f_{fp} = 0.04826$$

A continuación se muestran algunos ejemplos, de los resultados obtenidos del cálculo de factor de factor de fricción por iteración, en comparación con los resultados obtenidos con las ecuaciones de la RNA para la estimación del factor de fricción. Los datos de Re_G , n y He para fluidos Herschel-Bulkley utilizados, fueron considerados simplemente como referencias de condiciones del régimen laminar, y el último ejemplo es un caso hipotético de un fluido tipo Herschel-Bulkley en condiciones de régimen turbulento, tomando como referencia un punto de la representación gráfica de la relación de Torrance para fluidos Herschel-Bulkley encontrada en Garcia y Steffe (1986).

Tabla 3.5 Algunos resultados del factor de fricción (f_f) para fluidos Herschel-Bulkley.

DATOS			f_{fc}	f_{fp}
Re_G	N	He	(vía iteración)	(vía ecuaciones de RNA)
1260	0.4	100	0.0168	0.0150
120	0.2	100	0.2325	0.2280
810	0.9	1000	0.0251	0.0315
64.189	0.7	1.6003	0.2560	0.2770
7000	0.5	10000	0.0056	0.0053

CONCLUSIONES

Con el grupo de datos utilizado como ejemplos para entrenar a la red neuronal artificial (RNA), que tenían previamente una transformación logarítmica de todos los parámetros (tanto de entrada como de salida), se obtuvieron los mejores resultados, debido que representa mejor la relación entre parámetros de entrada con los de salida, es decir la relación $\{f_f\}$ en función de $\{Re_G, n, He\}$ en una tubería para un fluido tipo Herschel-Bulkley, es una relación de tipo logarítmica como lo pueden constatar no solo los resultados sino la gráficas que se han realizado con anterioridad por diversos autores al graficar f_f vs Re . Por lo tanto considerar el tipo de relación entre variables y proporcionar los ejemplos tomando en cuenta esta relación, mejora significativamente la generalización del problema, por parte de la RNA.

El modelo óptimo para calcular el factor de fricción de un fluido Herschel-Bulkley por una vía no-iterativa, es el modelo de 20 neuronas en la capa intermedia, de la tercera prueba originado por entradas y salida de datos logarítmica. De acuerdo a su evaluación tiene un error medio absoluto (MAE) de 0.0146, un error medio relativo (MRE) de 0.0231%, y un coeficiente de determinación (R^2) de 0.99, cabe mencionar que si este se aleja de 1 es a consecuencia de los valores que se salen de la linealización, debido a que pertenecen a valores de $He \geq 1 \times 10^6$. Sin embargo por su tamaño sería poco práctico utilizarlo para resolver el problema de consumo de tiempo en el cálculo por iteración, pues al ser de 20 neuronas en la capa intermedia, implica que tendría mucho más coeficientes y mucho más ecuaciones para representar en ecuaciones algebraicas la capa intermedia, lo que lo haría un procedimiento, largo y tedioso.

Esa es la razón por la que se decide optar por un modelo más simple pero con resultados cercanos al modelo óptimo, por lo que se elige al modelo con 4 neuronas en la capa intermedia de la prueba 3, este tiene valores de validación aceptables con un MAE de 0.077, un MRE de 0.06%, y una desviación estándar relativa (STDR) de 0.119%. Cabe señalar que existen diferencias entre el resultado del cálculo iterativo de factor de fricción

para un fluido Herschel–Bulkley utilizando las ecuaciones de Hanks (régimen laminar) y Torrance (régimen turbulento) y el cálculo no-iterativo con las ecuaciones explícitas obtenidas por una red neuronal artificial, esta disparidad de resultados son de esperarse pues son inherentes a la diferencias entre los procedimientos de cálculo. Sin embargo estas diferencias son mínimas y dependiendo de las necesidades de cálculo, podría optarse por un método u otro, sin embargo la ventaja de las ecuaciones explícitas derivadas del modelo de una RNA es que además de quitar el procedimiento iterativo y/o la necesidad de utilizar métodos numéricos para su resolución, las ecuaciones explícitas son de gran simplicidad y es por ello que el tiempo que se utiliza para su cálculo, suele ser mucho menor que el utilizado para calcular el factor de fricción de un fluido Herschel-Bulkley por iteración.

Por lo tanto el modelo desarrollado a través de las RNA cumple con las características necesarias para ser usado como una alternativa al proceso de iteración de prueba y error para el cálculo el factor de fricción de un fluido Herschel-Bulkley en una tubería tanto en régimen laminar como turbulento en un amplio intervalo de valores para de Re_G y He , y perfectamente apto dentro de los intervalos establecidos para su uso, que son $Re_G= 100$ a 1×10^7 , $n= 0.1$ a 1 y $He \geq 10$ a $He < 10^6$, para obtener resultados apropiados.

En adición, de manera indirecta se ha demostrado que el uso de las RNA puede tener diversidad de aplicaciones y son una herramienta que puede generar alternativas a métodos tradicionales, sin olvidar que en algunos problemas su uso puede ser apropiado y en otros no. No obstante, las RNA son una herramienta alternativa que los ingenieros en alimentos pueden implementar siempre y cuando las características del problema sean las adecuadas para la aplicación de las RNA, y sean una alternativa de solución viable.

RECOMENDACIONES

Para resolver el problema de los errores, cuando $He \geq 1 \times 10^6$, se podría corregir este error si así se desea en otro modelo con características similares pero con condiciones de entrenamiento corregidas, al aumentar el número de ejemplos suministrados en ese intervalo de valores, pues al parecer este error se genera precisamente por la falta de datos en ese intervalo, puesto que los intervalos que se utilizaron son sumamente amplios en régimen turbulento y no lograron generalizar el comportamiento de las variables a partir de ese intervalo.

REFERENCIAS

1. Adhikari, B. & Jindal, V. K. (2000), Artificial neural networks: a new tool for prediction of pressure drop of non-Newtonian fluid foods through tubes. *Journal of Food Engineering*, 46, 43-51.
2. Afaghi, M., Ramaswamy, H., & Prasher, S. (2001). Thermal process calculations using artificial neural network models. *Food Research International*, 34, 55-65.
3. Anderson, J. (2007). *Redes neuronales*. México, Alfaomega.
4. Barro, S., Fernández-Delgado, M., Regueiro, C., & Sánchez, E. (1995). Modelos básicos de redes neuronales artificiales. – en Barro, S., & Mira, J. (Eds.), *Computación Neuronal –Curso y Congreso de Universidad de Santiago Compostela*. España, Servicio de Publicaciones e Intercambio Científico, 181-213.
5. Burden R., & Faires, D. (2002). *Análisis numérico 7ª Ed.* México, Thomson Learning.
6. Butt, R. (2008). *Introduction to numerical analysis using MATLAB®*. Canada, Infinity Science Press LLC.
7. Castillo, E., Cobo, A., Gutiérrez, J. & Pruneda, R. (1999). *Introducción a las Redes Neuronales con Aplicaciones: Un nuevo paradigma neuronal*. España, Paraninfo.
8. Cohen, M. & Hudson, D. (2000). *Neural Networks and Artificial Intelligence for Biomedical Engineering*. USA, IEEE Press.
9. Corchado, J., Díaz, F., Borrajado, L., & Fernández, F. (2000). *Redes Neuronales Artificiales, un enfoque práctico*. España, Servicio de publicaciones Universidad de Vigo.
10. Crane, (1996). *Flujo de fluidos en válvulas, accesorios y tuberías*. 1ra Ed. México, Mc Graw Hill.
11. Demuth, H., Beale, M. & Hagan, M. *Neural Network Toolbox™ 6, User's Guide*. Matlab, Copyright (1992-2010) by The Math Works, Inc.
12. Fernández, E. (1993). Modelos Conexionistas: conceptos generales, origen y evolución. –en Olmeda, I., & Barba-Romero, S. (Eds.), *Redes Neuronales Artificiales, fundamentos y aplicaciones- Actas de un curso de la Universidad de Alcalá*, Junio de 1993. España, Servicio de Publicaciones de la Universidad de Alcalá, 3-21.
13. Freeman, J., & Skapura, D. (1991). *Neural Networks: Algorithms, Applications and Programming techniques*. USA: Addison-Wesley. (Versión en castellano: 1993. *Redes*

- Neuronales, Algoritmos, Aplicaciones y Técnicas de Programación. EUA, Addison-Wesley Iberoamericana/ Díaz de Santos).
14. García, E. & Steffe, J. (1987). Comparison of friction factor equations for non-newtonian fluids in pipe flow. *Journal of Food Process Engineering*, 9, 93-120.
 15. Geankoplis, C. J. (1998). Procesos de transporte y operaciones unitarias. 3a Ed. México, CECSA.
 16. Haykin, S. (1999). Neural Networks: A comprehensive Foundation. 2nd ed. USA, Prentice Hall.
 17. Hilera, G., & Martínez, H. (1995). Redes Neuronales Artificiales: Fundamentos, modelos y aplicaciones. España, Ra-Ma.
 18. Huang, S. & Xing, H. (2002). Extract intelligible and concise fuzzy rules from neural networks. *Fuzzy Sets and Systems*, 132, 233-243.
 19. Ibarz, A. & Barbosa-Cánovas, G. (2005). Operaciones unitarias en la ingeniería de alimentos. –Colección tecnología de alimentos. España, Ediciones Mundi Prensa.
 20. Isasi, P. & Galván, I. (2004). Redes de Neuronas Artificiales: Un Enfoque Práctico. España, Pearson Educación.
 21. Islamoglu, Y. (2003). A new approach for the prediction of the heat transfer rate of the wire-on tube type heat exchanger- use of an artificial neural network model. *Applied Thermal Engineering*, 23, 243-249.
 22. Jimenez-Marquez, S., Lacroix, C. & Thibault, J. (2003). Impact of modeling parameters on the prediction of cheese moisture using neural networks. *Computers and Chemical Engineering*, 27, 631-646.
 23. Johnson R. & Kuby P. (2012). Estadística elemental. 11ª Ed. México, Cengage Learning Inc.
 24. Leithold, L. 1998. El cálculo 7ed., Oxford University Press, México.
 25. Luera Peña, W. & Minim, L. (2001). Aplicación de las redes neuronales artificiales en la modelización del tratamiento térmico de alimentos. *Ciencia y Tecnología de los Alimentos*, 3(2), 81-88.
 26. Maren, A. (1990). Neural network structures: form follows function.- In Harston, C., Maren, A., & Pap, R. (Eds.), Handbook of Neuronal Computing Applications. USA, Academic Press. 45-57.
 27. Martín, D. B. & Sanz, M. A. (1997). Redes Neuronales y Sistemas Borrosos: Introducción, teoría y práctica. España, Ra-Ma.

28. Méndez M., Núñez, M. & Solorza, J. (2001). *Introducción a la Reología*. México, Instituto Politécnico Nacional.
29. Mittal, G. & Zhang, J. (2000). Prediction of freezing time for food products using a neural network. *Food Research International*, 33, 557-562.
30. Mott, R. (2006). *Mecánica de fluidos*. 6a Ed. México, Pearson Educación.
31. Muller, H.G. (1973). *Introducción a la reología de alimentos*. España, Acribia.
32. Müller, B. & Reinhardt, J. (1990). *Neural Networks, An Introduction*. Germany, Springer-Verlag.
33. Olajide, J.O., Igbeka, J.C., Afolabi T.J., & Emiola O.A. (2007). Prediction of oil yield from groundnut kernels in an hydraulic press using artificial neural network (ANN). *Journal of Food Engineering*, 81, 643-646.
34. Osorio, F. A. & Steffe, J. (1984). Kinetic energy calculations for non-newtonian fluids in circular tubes. *Journal of food science*, 49, 1295-1296, 1315.
35. Pineda, M. L. & Romero, A. V. (2003). *Manual de Mecánica de fluidos aplicada a los alimentos*. Tesis (Ingeniería en Alimentos), UNAM, Facultad de Estudios Superiores Cuautitlán.
36. Potter, M. & Wiggerti, D. (1998). *Mécanica de Fluidos*. 2da Ed., México, Prentice Hall.
37. Rao, M. A. (1999). *Rheology of fluid and semisolid Foods: Principles and applications*. Maryland, An Aspen Publication.
38. Resnick, R. H. (1994). *Física*. México, Editorial Continental.
39. Rosenthal, A. (2001). *Textura de los alimentos: medida y percepción*. España, Acribia.
40. Sablani, S., & Shayya, W. (1998). An artificial neural network for non-iterative calculation of the friction factor in pipeline flow. *Computers and electronics in agriculture*, 21, 219-228.
41. Sablani, S., & Shayya, W. (2003). Neural Networks based non- iterative calculation of the friction factor for power law fluids. *Journal of Food Engineering*, 57, 327-335.
42. Sablani, S., Shayya, W. & Kacimov, A. (2003). Explicit calculation of the friction factor in pipeline flow of Bingham plastic fluids: a neural networks approach. *Chemical Engineering Science* 58, 99-106.
43. Singh, P. & Heldman, D. (2009). *Introduction to Food Engineering* 4th ed. China, Academic Press.

44. Steffe, J. (1996). *Rheological Methods in Food Process Engineering*. 2nd ed. USA, Freeman Press.
45. Steffe, J. & Morgan, R. (1986). Pipeline Design and Pump Selection for Non-Newtonian Fluid Foods: Use of effective viscosity in calculations enables pumps to be selected using manufacturers' data for Newtonian fluids. *Food Technology*, 40, 78-85.
46. Steffe, J. & Singh, P. (1997). Pipeline Design Calculations for Newtonian and Non-Newtonian Fluids.- In Rotstein, E., Singh, P., & Valentas, K. (Eds.) *Handbook of Food Engineering Practice* (1997). USA, CRC Press. Chapter 1.
47. Torrecilla, J. S., Otero, L., & Sanz P.D. (2005). Artificial neural Networks: a promising tool to design and optimize high-pressure food processes. *Journal of Food Engineering*, 69, 299-306.
48. Williams, D. (1990). Problems of measurement and control in biotechnological processes.-In Winkler, M. (Ed.), *Chemical Engineering Problems in Biotechnology-Critical Reports on Applied Chemistry*, Vol. 29. London and USA, Elsevier Applied Science. 167-213.
49. Wilson, J. D. & Buffa, A. J. (2003). *Física*. 5a Ed. México, Pearson Educación.
50. Zuñiga, V. E. (1991). El estudio y aplicación de la caída de presión permisible como criterio de dimensionamiento preliminar para fluidos alimenticios newtonianos y no-newtonianos en régimen laminar. Tesis (Ingeniería en Alimentos), UNAM, Facultad de Estudios Superiores Cuautitlán.

Referencias de páginas electrónicas.

- Glosario de terminología reológica. Revista Iberoamericana de polímeros, Volumen 3(2), Abril 2002. <http://www.ehu.es/reviberpol/pdf/abr/reologia.pdf>, Consulta: Septiembre-2015.
- Grafo <http://es.wikipedia.org/wiki/Grafo>, Consulta: Mayo-2015.
- MATLAB and Simulink for technical computing - Mathworks España. <http://es.mathworks.com>, Consulta: Diciembre-2014.
- Redes neuronales artificiales, Introducción, características y ejemplos. www.pol.una.py/tesis/tesis1/index.htm, Consulta: Abril-2005.
- Teoría de Grafos en la vida real. Conceptos básicos. © Universitat Politècnica de València (UPV). Bloque del curso, Aplicaciones de la Teoría de Grafos a la vida real. Publicado el 24/11/2013. Autor: Jordan Lluch, C. https://www.youtube.com/watch?v=BJbKvJs_yF0, Consulta: Mayo-2015.

Glosario

Algoritmo.- Método y notación del cálculo, todo procedimiento de cálculo. Es un procedimiento que describe, sin ambigüedades, una serie finita de pasos a realizar en un orden específico. El objeto del algoritmo es poner en práctica un procedimiento para resolver un problema o aproximarse a una solución del problema.

Asíntota.- En matemática, se le llama **asíntota** de la gráfica de una función, a una recta a la que se aproxima continuamente la gráfica de tal función; es decir que la distancia entre las dos tiende a ser cero (0), a medida que se extienden indefinidamente.

Fuente: <http://es.wikipedia.org/wiki/As%C3%ADntota> Septiembre, 2014.

Bias.-Parámetro de una neurona artificial, que es sumado con los pesos de las entradas y pasa a través de la función de activación de la neurona para generar la salida de la neurona.

Cambio de estado en la red neuronal artificial.- El estado de todas y cada una de las neuronas se denomina estado de la red, y su variación en el tiempo se realiza de acuerdo con un determinado modo de operación ya sea síncrono o Asíncrono.

Emulación.-En informática, un **emulador** es un software que permite ejecutar programas o videojuegos en una plataforma (sea una arquitectura de hardware o un sistema operativo) diferente de aquella para la cual fueron escritos originalmente. A diferencia de un **simulador**, que solo trata de reproducir el comportamiento del programa, un emulador trata de modelar de forma precisa el dispositivo de manera que este funcione como si estuviese siendo usado en el aparato original.

Fuente: <http://es.wikipedia.org/wiki/Emulador> Mayo, 2015.

Host.-El término *host* ("**host**", en español) es usado en informática para referirse a las computadoras conectadas a una red, que proveen y utilizan servicios de ella. Los usuarios deben utilizar *anfitriones* para tener acceso a la red. En general, los *anfitriones* son computadores monousuario o multiusuario que ofrecen servicios de transferencia de archivos, conexión remota, servidores de base de datos, servidores web, etc. Los usuarios que hacen uso de los *anfitriones* pueden a su vez pedir los mismos servicios a otras máquinas conectadas a la red. Un host o anfitrión es un ordenador que funciona como el punto de inicio y final de las transferencias de datos.

Fuente: <http://es.wikipedia.org/wiki/Host> Mayo, 2015.

Isótropo.- En física, la **isotropía**, (cuya etimología está en las raíces griegas *ἴσος* [isos], equitativo o igual, y *τρόπος* [tropos], medio, espacio de lugar, dirección), es la característica de los cuerpos cuyas propiedades físicas no dependen de la dirección. Es decir, se refiere al hecho de que ciertas magnitudes vectoriales conmensurables, dan resultados idénticos con independencia de la dirección escogida para dicha medida.

Fuente: <http://es.wikipedia.org/wiki/Isotrop%C3%ADa> Julio, 2011.

Modelo de Ising.-El **modelo de Ising** es un modelo físico propuesto para estudiar el comportamiento de materiales ferromagnéticos. Se trata de un modelo paradigmático de la Mecánica Estadística, en parte porque fue uno de los primeros en aparecer, pero sobre todo porque es de los pocos modelos útiles (no sólo pedagógicamente) que tiene solución analítica exacta (esto es, sin cálculos aproximados). Esto lo hace muy útil para ensayar nuevos tipos de aproximaciones y luego comparar con el resultado real. Fue propuesto por Ernst Ising, quien intentaba demostrar que el sistema presentaba una transición de fase.

Fuente: http://es.wikipedia.org/wiki/Modelo_de_Ising, Mayo, 2010.

Matriz Hessiana.-La matriz Hessiana o hessiano de una función f de n variables, es la matriz cuadrada de $n \times n$, de las segundas derivadas parciales.

Fuente: http://es.wikipedia.org/wiki/Matriz_hessiana Mayo, 2015.

Matriz Jacobiana.-La **matriz jacobiana** es una matriz formada por las derivadas parciales de primer orden de una función. Una de las aplicaciones más interesantes de esta matriz es la posibilidad de aproximar linealmente a la función en un punto. En este sentido, el jacobiano representa la derivada de una función multi-variable.

Propiamente deberíamos hablar más que de matriz jacobiana, de diferencial jacobiana o aplicación lineal jacobiana ya que la forma de la matriz dependerá de la base o coordenadas elegidas. Es decir, dadas dos bases diferentes la aplicación lineal jacobiana tendrá componentes diferentes aún tratándose del mismo objeto matemático.

Fuente: <http://es.wikipedia.org/wiki/Jacobiano> Mayo, 2015.

Modelo Matemático.-En ciencias aplicadas, un **modelo matemático** es uno de los tipos de modelos científicos que emplea algún tipo de formulismo matemático para expresar relaciones, proposiciones sustantivas de hechos, variables, parámetros, entidades y relaciones entre variables y/o entidades u operaciones, para estudiar comportamientos de sistemas complejos ante situaciones difíciles de observar en la realidad.

Fuente: http://es.wikipedia.org/wiki/Modelo_matemático Marzo, 2015.

Expresar una situación del mundo real en términos de una relación funcional, se denomina modelo matemático de la situación (Leithold, 1998).

Levenberg-Marquardt.-Es una técnica, basada en un promedio ponderado del método de Newton y del descenso más rápido. La ponderación se inclina hacia el método del descenso rápido hasta que se descubre la convergencia; en ese momento el peso se inclina hacia el método de Newton que converge más rápidamente (Burden & Faires, 2002).

Origen de ruido en los instrumentos de medición y control.- En este caso el ruido puede ser internamente generado dentro de un circuito debido al azar, movimiento de electrones inducidos por temperatura y otras cargas llevadas a cabo en los resistores y semi-conductores (ruido térmico). Un segundo origen de ruido interno es llamado 'shot noise', y ocurre en transistores debido a fluctuaciones al azar en la velocidad a la cual se lleva a cabo la difusión a través de un empalme o conexión (Williams, 1990).

Ponderados.- Multiplicados, contrapeso, equilibrado (Freeman & Skapura, 1991).

Robusto.- Se dice de un programa sólido que trabaja correctamente bajo todas las condiciones normales y la mayoría de las anormales.

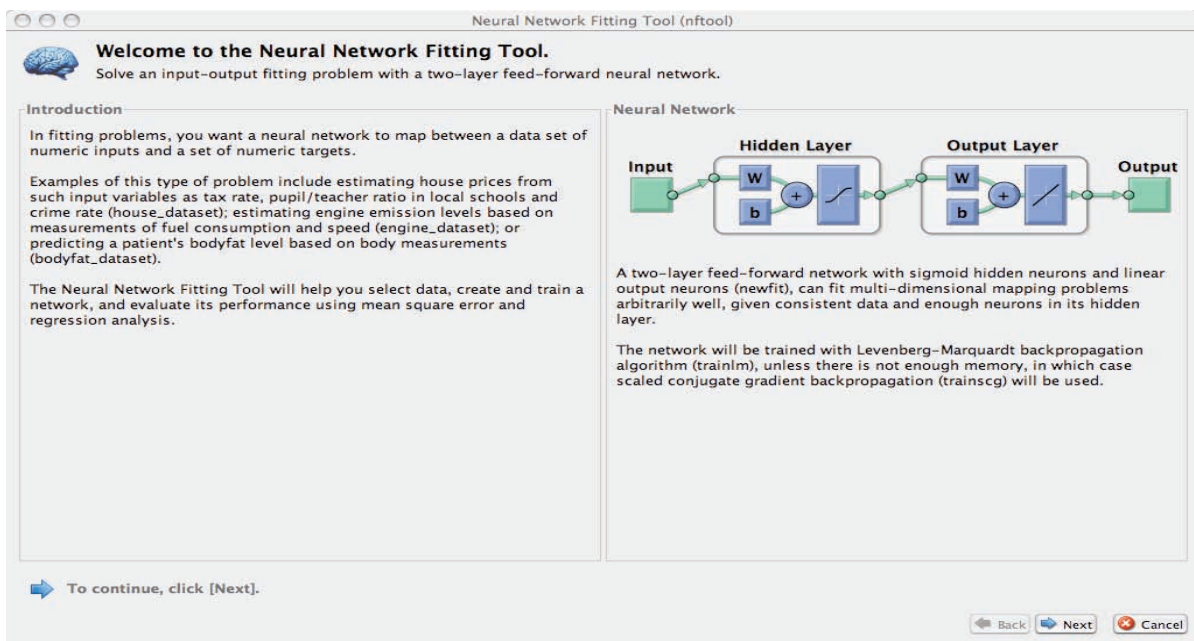
Ruido.- Si en un experimento o una operación que se repite bajo aproximadamente las mismas condiciones, los resultados observados no son completamente idénticos las fluctuaciones que ocurren de una repetición a otra se denominan: *ruido*; variaciones experimentales, error experimental o simplemente error. Es la variación producida por factores distorsionantes, tanto conocidos como desconocidos, normalmente solo una pequeña parte de el puede ser atribuido a error de medición y a menudo son variaciones inevitables y no esta asociado a culpa de ninguna clase.

En un conjunto de datos el ruido puede definirse como las variaciones aleatorias de una o más características de datos, en los problemas reales esto es muy habitual debido a la existencia de fuentes incontrolables generadoras de ruido, por ejemplo: fuentes internas como la tendencia de los circuitos y del conexionado a captar impulsos de líneas adyacentes, distorsiones introducidas por los sensores que miden los datos, ruido metido en un canal de comunicación a través del cual se transmite una información, fuentes externas como las condiciones meteorológicas, etc.

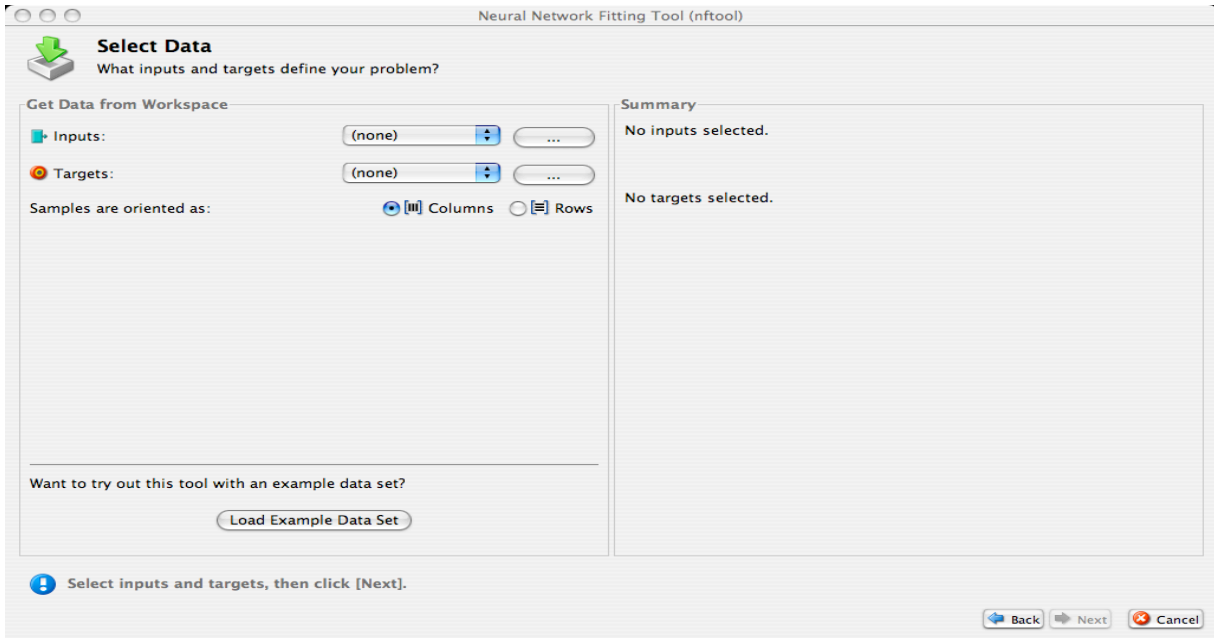
ANEXO 1. Instrucciones de uso del simulador Neural Network Toolbox de entorno Matlab® R2008^a

En Matlab existe un interfaz gráfico llamado Neural Network Fitting Tool GUI para resolver específicamente un problema de generalización de una función, este es el simulador diseñado por Matlab, que fue utilizado en este trabajo. Los pasos para utilizar el simulador son:

1. Desde la línea de comando, con el comando `nftool` se abre la siguiente pantalla.



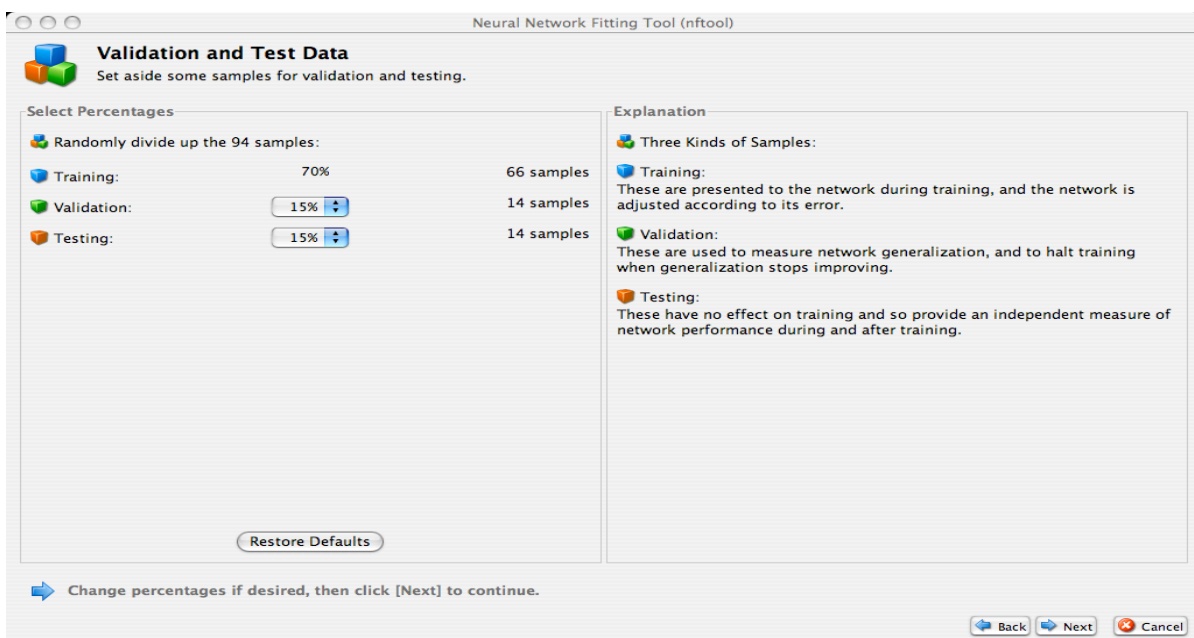
2. Se tecllea **Next** para proceder



3. Se da click para descargar el grupo de datos que servirán de ejemplo en la ventana de selección de datos. Y el grupo de datos elegidos se abrirá, después selecciónelos.

Nota: Es importante que para que el grupo de datos se descargue, estos grupos tanto entradas (Inputs) y salidas (Targets) deben estar abiertos en MATLAB® workspace.

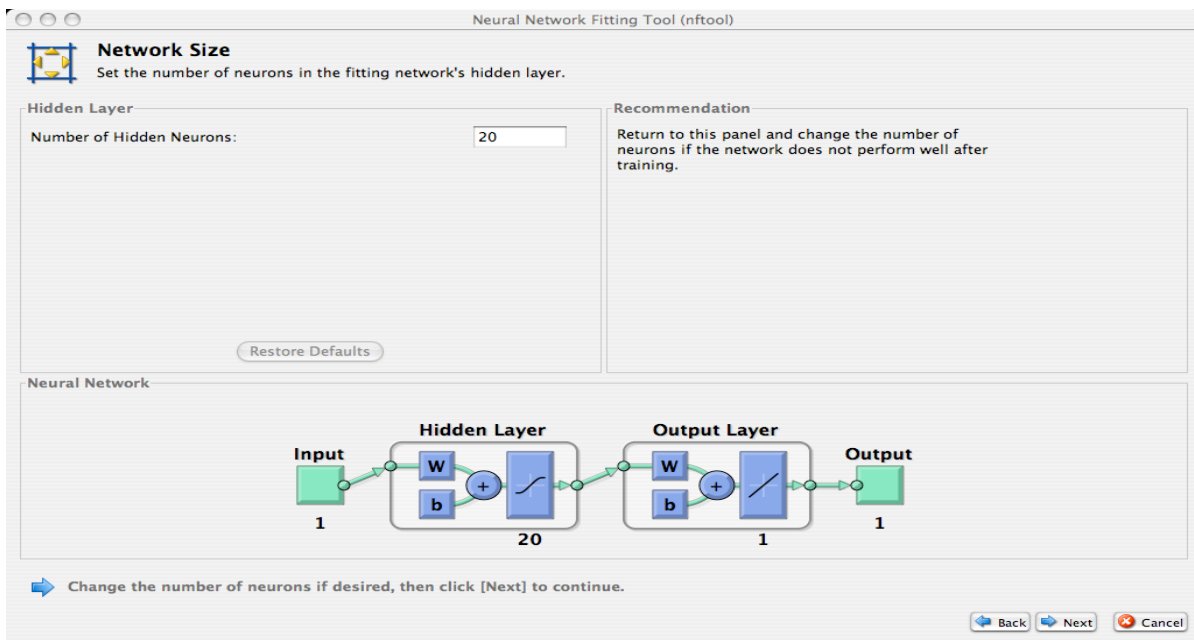
4. Una vez descargados teclee **Next** para mostrar la ventana de Validación y Prueba mostrada en la siguiente figura.



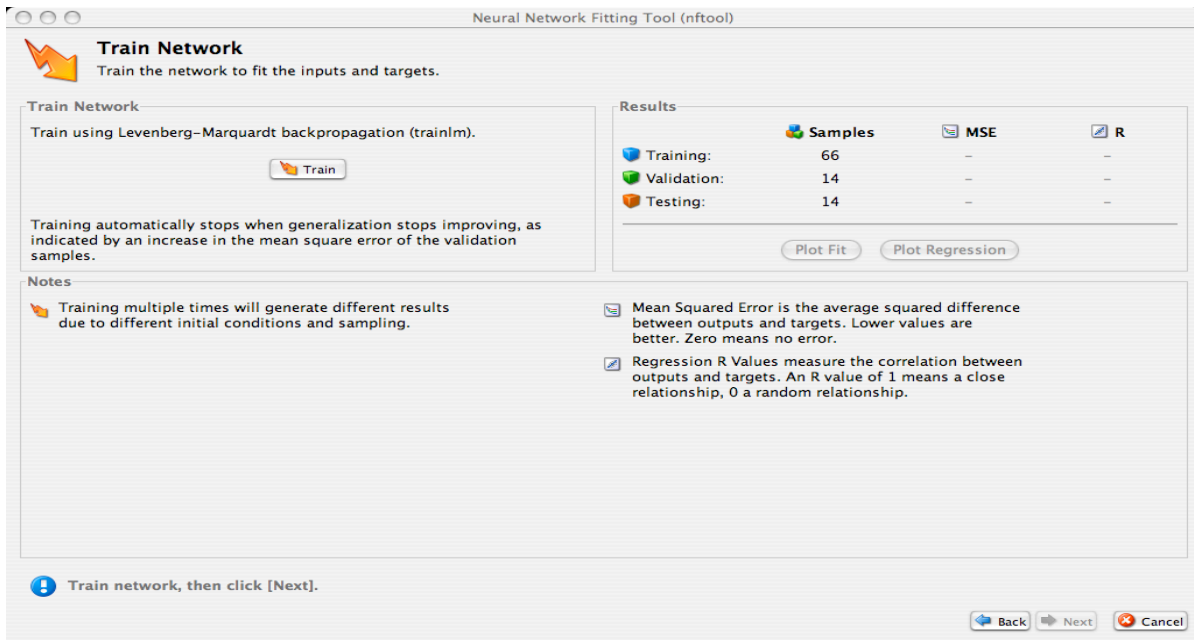
Los grupos de datos de validación y prueba son por default un 15% para cada uno del total de los datos originales, pero se pueden modificar de acuerdo a las necesidades.

5. De click en Next.

El número de neuronas ocultas es de 20 por default, pero se puede modificar este valor y probar otro número si así se desea. También se podría cambiar este número si la RNA no se desempeña tan bien como se espera.

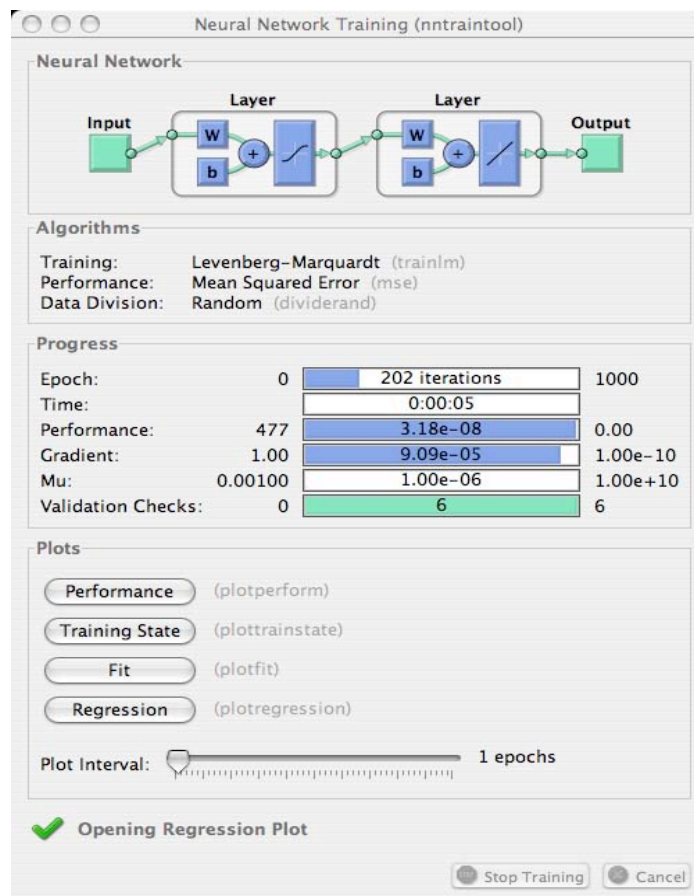


6. De click en Next



7. De click en **Train (entrenamiento)**

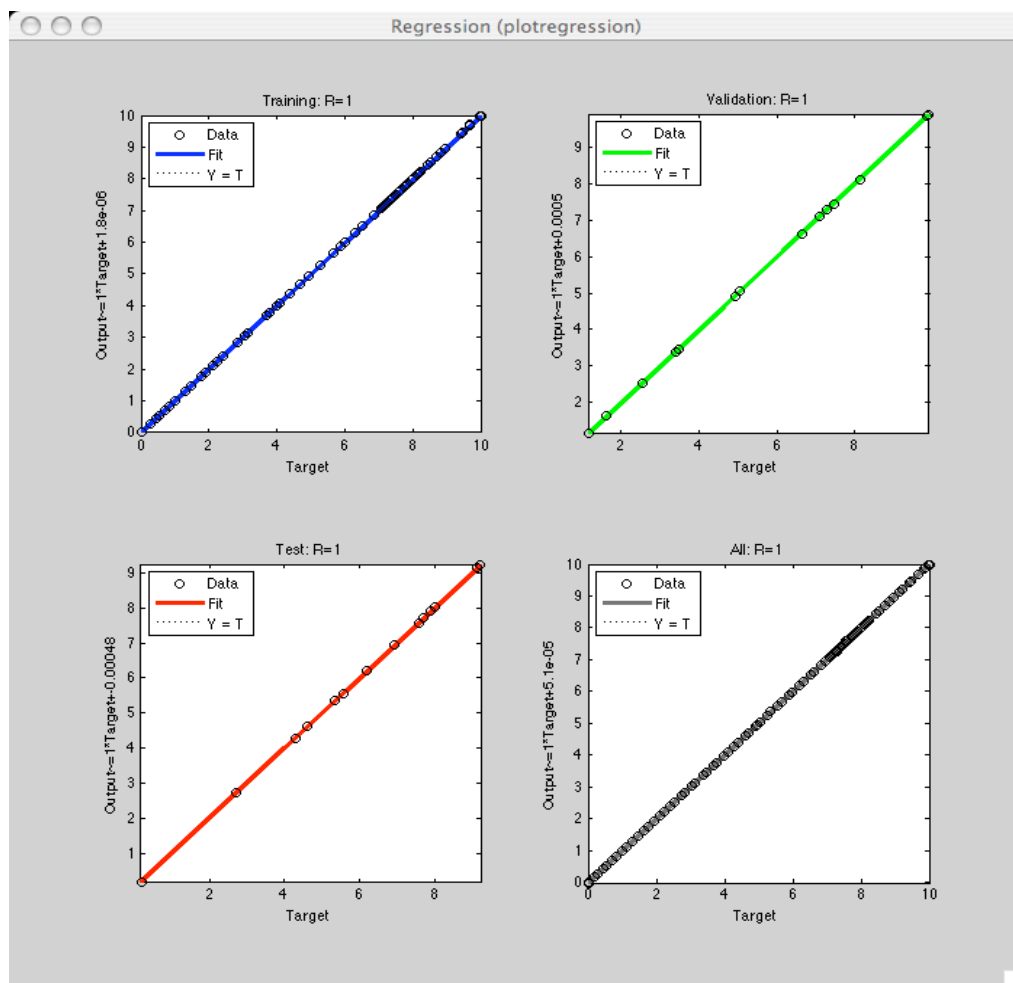
El tiempo de entrenamiento puede continuar hasta un máximo de 1000 iteraciones.



Esta ventana es muy útil porque señala en términos generales como se esta llevando acabo el desempeño del entrenamiento o como fue si la ventana se abre al final de este, también puede mostrar la gráfica del estado de entrenamiento, el desempeño del entrenamiento, la regresión, el tiempo que tarda, etc.

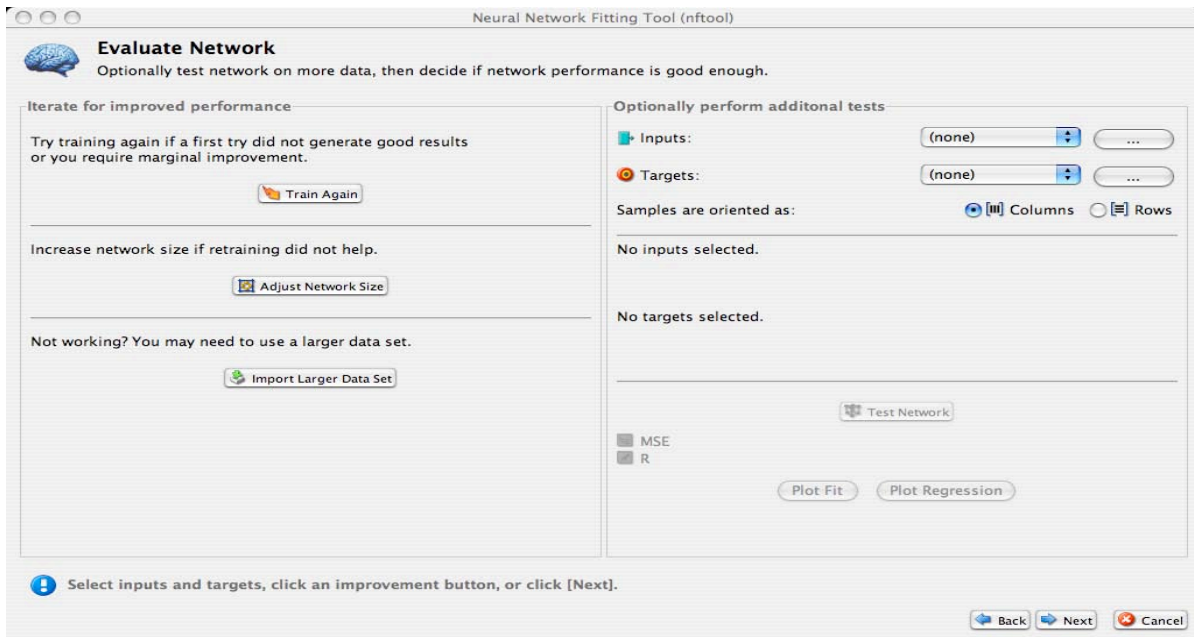
8. Debajo de **Plots**, de click en **Regression**.

En este ejemplo, se puede notar que la generalización ó ajuste es casi perfecto para los datos de entrenamiento, prueba y validación. Este solo es un ejemplo pero en general este parámetro es muy útil para elegir entre un entrenamiento y otro, pues muestra el valor de R, la gráfica y la ecuación de la gráfica.



Estas gráficas, son las gráficas de regresión para las salidas con respecto al entrenamiento, validación y prueba de datos.

9. De click en **Next** en el Neural Network Fitting Tool para evaluar la red.



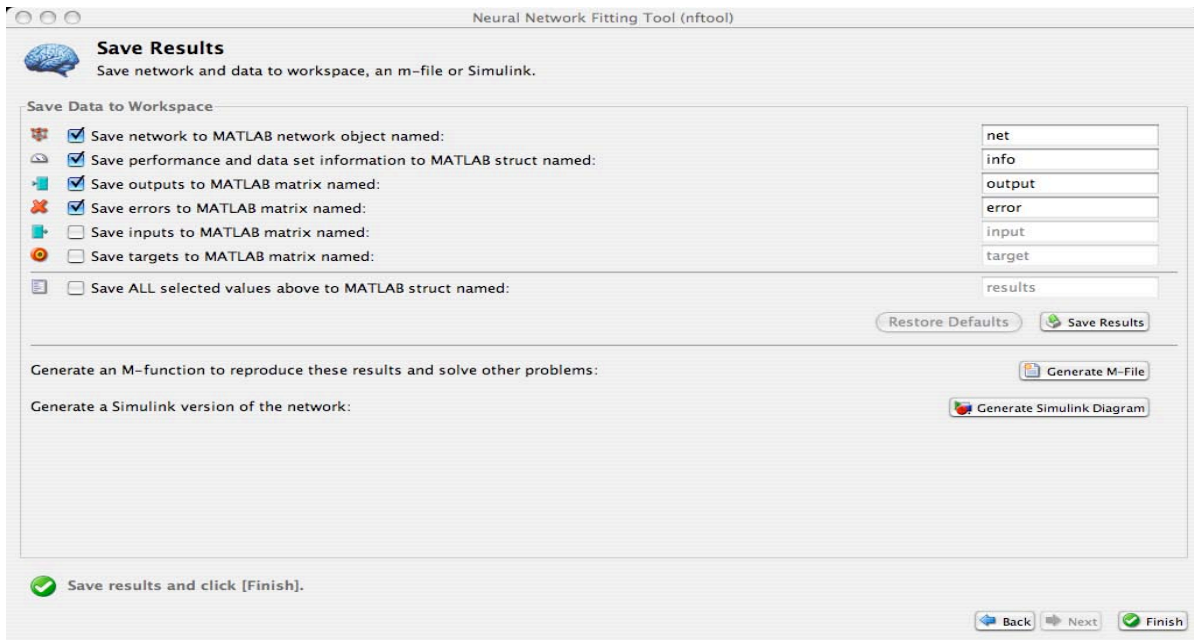
En este punto usted puede probar la red de nuevo contra nuevos datos.

Si usted esta insatisfecho con el desempeño de la red sobre los datos originales o nuevos datos usted puede seguir cualquiera de los siguientes pasos:

- Tratar de nuevo (entrenar de nuevo a la red)
- Aumentar el número de neuronas
- Conseguir un grupo de datos más grande

10. Si usted esta satisfecho con el desempeño de la red, de click en **Next**.

11. Use los botones en la apantalla para salvar sus resultados.



- Usted tiene la red salvada como net1 (por default) en el workspace de Matlab y puede guardarla desde ahí. Si desea puede realizar pruebas adicionales en este o ponerlo a trabajar sobre nuevos datos, usando la función sim. Un ejemplo de esto es:

```
a= sim(net,p)
```

Donde net es el nombre de la RNA y p son las entradas de la RNA.

- Usted puede también teclear **Generate M-File** para crear un M-file (un archivo) que puede ser usado para reproducir todos los pasos previos desde la línea de comando. Creando un M-file puede ser de ayuda si usted desea aprender como usar la funcionalidad de línea de comando del toolbox para acostumbrarse al proceso de entrenamiento.

12. Cuando usted haya salvado sus resultados, de click en **Finish**.

ANEXO 2. El Método de Bisección

Esta es una de las técnicas iterativa más simples para determinar raíces de ecuaciones no-lineales. Esta basado en le teorema de valor intermedio. La base de este método se ilustra con la ecuación:

$$y = f(x)$$

El objetivo es encontrar un valor de x para cual y es cero. Supongamos que f es una función continua definida en el intervalo $[a,b]$ con $f(a)$ y $f(b)$ de signos diferentes. De acuerdo con el teorema del valor intermedio, existe un número p en (a,b) tal que $f(p) = 0$. Si bien el procedimiento se aplica aunque exista mas de una raíz en el intervalo (a,b) , por razones de simplicidad suponemos que la raíz de este intervalo es única. El método requiere dividir varias veces a la mitad los sub-intervalos de $[a, b]$ y, en cada paso, localizar la mitad que contenga a p .

Entonces por la evaluación de la función en dos valores de x . Digamos a y b tales que

$$f(a) \cdot f(b) < 0$$

La implicación es que uno de los valores es negativo y el otro es positivo. Estas condiciones pueden ser fácilmente satisfechas por el intervalo cerrado de la función. Obviamente, la función es negativa en un punto final (endpoint), a , del intervalo y positivo en el otro punto final, b , y es continuo sobre $a \leq x \leq b$. Por lo tanto la raíz debe estar entre a y b (por el Teorema del valor intermedio) y una nueva aproximación de la raíz, α , puede ser calculado como

$$c = \frac{a+b}{2},$$

y en general

$$c_n = \frac{a_n + b_n}{2}, \quad n \geq 1.$$

Esta es la forma iterativa que se conoce como el método de bisección.

Si $f(c) \approx 0$, entonces $c \approx \alpha$ es la raíz deseada, y si no es así, entonces hay dos posibilidades. La primera, si $f(a) \cdot f(b) < 0$, entonces $f(x)$ tiene un cero entre el punto a y el punto c . El proceso

puede ser repetido sobre el nuevo intervalo $[a, b]$. Segundo si $f(a) \cdot f(c) > 0$, sigue que $f(b) \cdot f(c) < 0$ a partir de eso, es conocido que $f(b)$ y $f(c)$ tienen signos opuestos. Por lo consiguiente $f(x)$ tiene cero entre el punto c y el punto b y el proceso puede ser repetido con $[c, b]$. Vemos que después, en un paso del proceso, encontramos otro cero o un intervalo cerrado, el cual es precisamente la mitad de largo del original. El proceso continúa hasta la exactitud deseada como lograda. Esto nos da el método que se describe en el algoritmo.

Program 2.1 (Butt, 2008).

MATLAB m-file para el método de la bisección.

```
function sol = bisect (fn, a, b, tol)
    fa = feval(fn,a); fb = feval(fn,b);
    if fa * fb > 0; fprintf ('Endpoints have same sign') return
    end
    while abs (b - a) > tol
        c =(a + b)/2; fc =feval (fn,c);
        if fa *fc < 0; b = c; else a = c; end
    end; sol=(a + b)/2;
```

En MATLAB para el método de bisección, primero definimos la función m-file como `fn.m` para la ecuación como sigue:

```
function y = fn(x)
```

Entonces se usa el comando

```
>> s = bisect ('fn', a, b, tol)
```

ANEXO 3. Algoritmo de aprendizaje Levenberg-Marquardt

La generalización de la regla de aprendizaje Widrow-Hoff, para una red multicapa con funciones de transferencia no-lineal diferenciables, es el algoritmo Backpropagation. El vector de entrada (input) y el correspondiente vector objetivo (target), son usados para entrenar a una red hasta que esta pueda aproximar una función, y asociar los vectores de entradas con un específico vector de salida o clasificar el vector de entrada en una manera apropiada, como usted lo defina. Una RNA con biases (tendencias) una capa con función sigmooidal y con capa con una salida lineal son capaces de aproximar cualquier función con un finito número de discontinuidades.

El backpropagation estándar, es un algoritmo de gradiente descendiente, como es la regla de aprendizaje Widrow-Hoff. En el cual los pesos de la red son movidos a lo largo del gradiente negativo del desempeño de la función. El término backpropagation (de atrás hacia adelante) se refiere a la manera en la cual el gradiente es computado para redes multicapa no-lineales. Hay un número de variaciones sobre el algoritmo básico que están basados sobre otras técnicas de optimaciones estándar tales como el gradiente conjugado y el método de Newton. La más simple implementación de aprendizaje de backpropagation, actualiza los pesos y bias (tendencia) de la red en la dirección en la cual el desempeño de la función decrece más rápidamente, el gradiente negativo. Una iteración de este algoritmo puede ser escrita:

$$x_k + 1 = x_k - \alpha_k g_k$$

Donde x_k es un vector de pesos comunes y bias, g_k es el gradiente concurrente y α_k es la rapidez de aprendizaje.

Levenberg-Marquardt

El algoritmo Levenberg-Marquardt fue diseñado para aprovechar el entrenamiento de velocidad de segundo orden, sin tener computada la matriz Hessiana (matriz de segundas derivadas), cuando el desempeño de la función tiene la forma de una suma de cuadrados (como es típicamente en el entrenamiento de una red “feedforward” con conexiones hacia adelante), entonces la matriz Hessiana puede ser aproximada como:

$$H = J^T J$$

y el gradiente puede ser computado como:

$$g = J^T e$$

Donde J es la matriz Jacobiana que contiene la primera derivada de los errores de la red con respecto a los pesos y los bias (tendencias) y e es un vector de errores de la red. La matriz Jacobiana puede ser computada a través de una técnica backpropagation estándar (una retroalimentada), que es mucho menos compleja que computar la matriz Hessiana. El algoritmo Levenberg-Marquardt usa esta aproximación a la matriz Hessiana en la siguiente actualización tipo método de Newton:

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

Donde I es la matriz identidad y μ (parámetro mu) es el valor inicial para el escalar, cuando el escalar μ es cero, este se convierte en el método de Newton, usando la aproximación de la matriz Hessiana. Cuando μ (escalar) es grande, este llega a ser el gradiente descendente con un paso de tamaño pequeño. El método de Newton es más rápido y más exacto cerca de un error mínimo. Así el apuntar es ajustar hacia el método de Newton tan rápido como es posible. De esta manera, μ decrecerá después de cada paso exitoso (reducción en el desempeño de la función) e incrementa solo cuando un paso tentativo pueda incrementar el desempeño de la función. En este sentido, el desempeño de la función (performance) es siempre reducida en cada iteración del algoritmo.

El siguiente código reinicia la red previa y la re-entrena, usando el algoritmo Levenberg-Marquardt (trainlm). Los parámetros de entrenamiento para el trainlm son: epochs, show, goal, time, min_grad, max_fail, mu, mu_dec, mu_inc, mu_max, y mem_reduc. Los primeros seis parámetros determinan cuando se detiene el entrenamiento. El estatus del entrenamiento se muestra en la pantalla con show, para cada iteración del algoritmo, el entrenamiento se detiene si el número de iteraciones excede las epochs, si la función desempeño disminuye por debajo de goal, si la magnitud del gradiente es menos que min_grad, o si el tiempo de entrenamiento es más largo que los segundos de time. El parámetro max_fail, esta asociado con la técnica de paro adelantado. El parámetro mu, es el valor inicial de μ . Este valor es multiplicado por mu_dec siempre que la función desempeño es reducida por un paso. Si mu llega a ser más grande que mu_max, el algoritmo se detiene. El parámetro mem_reduc es usado para controlar de la cantidad de memoria usada por el algoritmo.