



UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS

UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS

**INSTITUTO DE INVESTIGACIÓN EN CIENCIAS BÁSICAS Y APLICADAS
CENTRO DE INVESTIGACIÓN EN INGENIERÍA Y CIENCIAS APLICADAS**

**APLICACIÓN DE LA DIRECTA E INVERSA DE LA RED
NEURONAL ARTIFICIAL PARA COLECTORES
SOLARES DE CANAL PARABÓLICO DE BAJO COSTO**

**TESIS PARA OBTENER EL GRADO DE DOCTORADO EN INGENIERÍA Y
CIENCIAS APLICADAS**

PRESENTA:

EDUI DASSAEV REYES TÉLLEZ

DIRECTORES:

DR. JOSÉ ALFREDO HERNÁNDEZ PÉREZ

DRA. ARIANNA PARRALES BAHENA

SINODALES:

DR. ARMANDO HUICOCHEA RODRÍGUEZ

DR. FERNANDO ZENAIDO SIERRA ESPINOSA

DR. ALBERTO ARMANDO ÁLVAREZ GALLEGOS

DR. OBED CORTÉS ABURTO

DR. SALVADOR ANTONIO ARROYO DÍAZ

CUERNAVACA, MORELOS

Mayo 2020



25/05/2020

IMG-20200525-WA0003.jpg

INSTITUTO DE INVESTIGACIÓN EN CIENCIAS BÁSICAS Y APLICADAS
 Maestría de Posgrado en Ingeniería y Ciencias Aplicadas

Cuernavaca, Morelos, a 13 de mayo de 2020.

DR. ROSENBERG JAVIER ROMERO DOMÍNGUEZ
COORDINADOR DEL POSGRADO EN
INGENIERÍA Y CIENCIAS APLICADAS
PRESENTE

Atendiendo a la solicitud para emitir DICTAMEN sobre la revisión de la TESIS titulada "APLICACIÓN DE LA DIRECTA E INVERSA DE LA RED NEURONAL ARTIFICIAL PARA COLECTORES SOLARES DE CANAL PARABÓLICO DE BAJO COSTO" que presenta el alumno EDUI DASSAEV REYES TÉLLEZ, para obtener el título de DOCTORADO EN INGENIERÍA Y CIENCIAS APLICADAS.

Nos permitimos informarle que nuestro voto es:

NOMBRE	DICTAMEN	FIRMA
DR. FERNANDO ZENAIDO SIERRA ESPINOSA	Aprobatorio	
DR. ARMANDO HUICOCHEA RODRÍGUEZ	Aprobatorio	
DR. ALBERTO ARMANDO ÁLVAREZ GALLEGOS	Aprobatorio	
DR. OBED CORTÉS ABURTO (UPPue)		
DR. SALVADOR ANTONIO ARROYO DÍAZ (UPPue)		
DRA. ARIANNA PARRALES SAHENA		
DR. JOSÉ ALFREDO HERNÁNDEZ PÉREZ	Aprobatorio	

PLAZO PARA LA REVISIÓN 20 DÍAS HÁBILES (A PARTIR DE LA FECHA DE RECEPCIÓN DEL DOCUMENTO)

NOTA. POR CUESTIÓN DE REGLAMENTACIÓN LE SOLICITAMOS NO EXCEDER EL PLAZO SEÑALADO, DE LO CONTRARIO LE AGRADECEREMOS SU ATENCIÓN Y NUESTRA INVITACIÓN SERÁ CANCELADA.

Av. Universidad 1001 Col. Chantla, Cuernavaca Morelos, México. 62209
 Tel. (777) 328 70 00 ext. 6209 / correo: correo@uaem.mx

Una universidad de excelencia

UAEM

<https://mail.google.com/mail/u/0/#inbox/FMfcgxHNWDqSSfhvknHINmIKSLZxwQb?projector=1&messagePartId=0.1>

1/1



Cuernavaca, Morelos, a 13 de mayo de 2020.

DR. ROSENBERG JAVIER ROMERO DOMÍNGUEZ
COORDINADOR DEL POSGRADO EN
INGENIERÍA Y CIENCIAS APLICADAS
P R E S E N T E

Atendiendo a la solicitud para emitir DICTAMEN sobre la revisión de la TESIS titulada "APLICACIÓN DE LA DIRECTA E INVERSA DE LA RED NEURONAL ARTIFICIAL PARA COLECTORES SOLARES DE CANAL PARABÓLICO DE BAJO COSTO" que presenta el alumno **EDUI DASSAEV REYES TÉLLEZ**, para obtener el título de **DOCTORADO EN INGENIERÍA Y CIENCIAS APLICADAS**.

Nos permitimos informarle que nuestro voto es:

NOMBRE	DICTAMEN	FIRMA
DR. FERNANDO ZENAIDO SIERRA ESPINOSA		
DR. ARMANDO HUICOCHEA RODRÍGUEZ		
DR. ALBERTO ARMANDO ÁLVAREZ GALLEGOS		
DR. OBED CORTÉS ABURTO (UPPue)		
DR. SALVADOR ANTONIO ARROYO DÍAZ (UPPue)		
DRA. ARIANNA PARRALES BAHENA		
DR. JOSÉ ALFREDO HERNÁNDEZ PÉREZ		

PLAZO PARA LA REVISIÓN 20 DÍAS HÁBILES (A PARTIR DE LA FECHA DE RECEPCIÓN DEL DOCUMENTO)

NOTA. POR CUESTION DE REGLAMENTACIÓN LE SOLICITAMOS NO EXCEDER EL PLAZO SEÑALADO, DE LO CONTRARIO LE AGRADECEMOS SU ATENCIÓN Y NUESTRA INVITACIÓN SERÁ CANCELADA.

Se expide el presente documento firmado electrónicamente de conformidad con el ACUERDO GENERAL PARA LA CONTINUIDAD DEL FUNCIONAMIENTO DE LA UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS DURANTE LA EMERGENCIA SANITARIA PROVOCADA POR EL VIRUS SARS-COV2 (COVID-19) emitido el 27 de abril del 2020.

El presente documento cuenta con la firma electrónica UAEM del funcionario universitario competente, amparada por un certificado vigente a la fecha de su elaboración y es válido de conformidad con los LINEAMIENTOS EN MATERIA DE FIRMA ELECTRÓNICA PARA LA UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS emitidos el 13 de noviembre del 2019 mediante circular No. 32.

Sello electrónico

JOSE ALFREDO HERNANDEZ PEREZ | Fecha:2020-05-29 16:05:30 | Firmante
JwnnB.JkzVIM1HXqkqfGIGBvyUIGG3JXKz4ZBAG2BJPYLuf5KELSu1n5WH2GdXRvEQkZ2gJXNjIvEryN8Jv+TMJgTkckGJM+AW8BE7GivYqItvgbZEGzy4XGUj3mg3q9
1f4nt1viSXevci5ET7N94G3FbpJwLAIa6+pSyb1Xw7LuXUQC+PEWfGRDhVZYKRTTcXN17LbEty1xOe63+I02lrPLcncV07qcpVTFoty26zprhFB5mUoJScVWOqMzeur4Br
q8xzPdaqlwfm16mqo4yV0xfoolY6vl+MME576yul6sdCZJBBMzQ+j7+G4+JaQdJZhl85w9gtLltw==

ARMANDO HUICOCHEA RODRIGUEZ | Fecha:2020-05-29 16:26:04 | Firmante
e77GBtOBtuX7KA52F2ycLj2cfjWjhw32RC07ungxbL9Xy8V0Y2rY4gfnikOlpb0Bm919rhuxLAtL.SbZL9DY+98c/nA5kPpqky6GUw8IzqH1HKeoA35n9IFoz+zbsbVOR9IGQ0
KRJ5DhJq2D5587J7gMJ+WJG3kLKYzFIPyaabyALK2f3g0Gnkww6w/FzC3xCRXHyfKJZUMDwP/mNFEaFcDdMYT.JaulJbKx0/D4woKikAAeaFT83EWwJNQiPeRazuFWIP
mH9+KbsKVJOpK5umWQt+nLT+7umIgsSx8nGe2BLUVZZad0lcaDd4pggMAMoyvnA8Wov1008Q==

ALBERTO ARMANDO ALVAREZ GALLEGOS | Fecha:2020-05-29 19:31:01 | Firmante
tvpIME2k1k3166FBSEAF+pxzK1uxwC11FRyR.JWUP1NvITDh0Pz4mlw8z2dE1z2E9BITsTSXSAcIwCPu0MrpUxpXBZzelhGoXVIs8+RAvHVeVnsBRyo45y6jsX0ld2UpBM
+LHbkbNOioO4xnYLeEcR4RIGEAvfwhUsluBSZTAInmsqOY9sLoEuzKhPliCspIhIR4e+SbNihBNB9IBpB4zPK/OtSsOM8chbGvllZLpxJZ3xIFkiwF067QvK35cylq0cwjXaR
nm5piyhIdIRCXt15frUWQmHOXSgpHIXrzJIQc2u/NT0pgq3mXW8zcx5cdrkfwfJZoudT4EM7Q==

FERNANDO ZENAIDO SIERRA ESPINOSA | Fecha:2020-05-30 14:55:14 | Firmante
CusyRaswWYP8vSIOu4x+cGcA89PtqebvwxNw5TVKgn4+Jh0iF4UJ0df3etu/FihWCZuL7V+eCYf3pHg+F9HwdsM09/c196LBcjh5EEZzsrwnSZPwQ590Fi+MrkgSqXS
zRQ26X9CaigXUfPNUce28wXyxrPQKbsPOBBBuL25yNyTsOaluZMX6pggt5dxdx122hGGLhGxo3ndpJlU0WWhTQmQIzeWX0Kthx0S8UDbIWLqW566IBISjdaOihReB8qP0
93Kx3u23blcABichc222VGCpIIRxj0E0PMvqshNmceEopzTl0lFV2MKS/xC2d82Ffz0z1ncuk2g==

Puede verificar la autenticidad del documento en la siguiente dirección electrónica o
escaneando el código QR ingresando la siguiente clave:



KGH9NX

<https://etirma.uaem.mx/noRepudio/7MyDTCM15b1PodQ6EkoNZ97ymDdtpMJQ>



UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS



UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS



INSTITUTO DE INVESTIGACIÓN EN CIENCIAS BÁSICAS Y APLICADAS
Defensoría de Posgrado en Ingeniería y Ciencias Aplicadas

Cuernavaca, Morelos, a 13 de mayo de 2020.

DR. ROSENBERG JAVIER ROMERO DOMÍNGUEZ
COORDINADOR DEL POSGRADO EN
INGENIERÍA Y CIENCIAS APLICADAS
PRESENTE

Atendiendo a la solicitud para emitir DICTAMEN sobre la revisión de la TESIS titulada "APLICACIÓN DE LA DIRECTA E INVERSA DE LA RED NEURONAL ARTIFICIAL PARA COLECTORES SOLARES DE CANAL PARABÓLICO DE BAJO COSTO" que presenta el alumno **EDUI DASSAEV REYES TÉLLEZ**, para obtener el título de **DOCTORADO EN INGENIERÍA Y CIENCIAS APLICADAS**.

Nos permitimos informarle que nuestro voto es:

NOMBRE	DICTAMEN	FIRMA
DR. FERNANDO ZENAIDO SIERRA ESPINOSA		
DR. ARMANDO HUICOHEA RODRÍGUEZ		
DR. ALBERTO ARMANDO ÁLVAREZ GALLEGOS		
DR. OBED CORTÉS ABURTO (UPPue)	Aprobado	
DR. SALVADOR ANTONIO ARROYO DÍAZ (UPPue)	Aprobado	
DRA. ARIANNA PARRALES BAHENA	Aprobado	
DR. JOSÉ ALFREDO HERNÁNDEZ PÉREZ		

PLAZO PARA LA REVISIÓN 20 DÍAS HÁBILES (A PARTIR DE LA FECHA DE RECEPCIÓN DEL DOCUMENTO)

NOTA. POR CUESTION DE REGLAMENTACIÓN LE SOLICITAMOS NO EXCEDER EL PLAZO SEÑALADO, DE LO CONTRARIO LE AGRADECEMOS SU ATENCIÓN Y NUESTRA INVITACIÓN SERÁ CANCELADA.

Av. Universidad 1001 Col. Chamilpa, Cuernavaca Morelos, México, 62209
Tel. (777) 329 70 00, ext. 6208 / raquel.sotelo@uaem.mx

UA
EM



Instituto de
Investigación en
Ciencias
Básicas y
Aplicadas

ANÁLISIS DE LA TRANSFERENCIA DE ENERGÍA EN COLECTORES SOLARES
DE CANAL PARABÓLICO APLICANDO REDES NEURONALES ARTIFICIALES

i Resumen

La presente tesis de investigación se divide en tres etapas, todas enfocadas hacia un mismo objetivo, eficientar un sistema colector solar de canal parabólico para incrementar el nivel térmico del agua para uso residencial. Para lograrlo, se emplearon diversas estrategias, tanto a nivel experimental como a nivel de optimización apoyado con herramientas de computación como modelado empírico. La primera etapa (presentada en el capítulo 4), se enfoca a describir cómo se puede modelar, mediante redes neuronales artificiales, un sistema renovable, y como aumentar la precisión del modelo a través de un algoritmo de entrenamiento (Regularización Bayesiana). En la segunda etapa (capítulo 5), se optimiza el sistema con ayuda de la directa e inversa de la red neuronal artificial y algoritmos genéticos, lo cual puede incrementar la temperatura del agua con un mínimo de requerimientos. Por último, se analizan las funciones de transferencia/activación como una herramienta más que permite aumentar la precisión del modelo empírico obtenido por las redes neuronales artificiales. Cabe resaltar que, para todos los modelos desarrollados, se tuvo que confeccionar una base de datos robusta y que fueron sometidos a varias pruebas estadísticas con el fin de verificar la fiabilidad y adecuación de los datos simulados con sus contrapartes experimentales lo que conlleva a reducir la incertidumbre de los modelos dados por buenos.

ii Abstract

This research thesis is divided into three stages, all focused towards same objective, to make a parabolic trough solar collector system efficient to increase the thermal level of water for residential use. To achieve this, various strategies were used, both at the experimental level and at the optimization level supported by computer tools such as empirical modeling. The first stage (presented in Chapter 4), focuses on describing how a renewable system can be modeled using artificial neural networks, and how to increase the precision of the model through a training algorithm (Bayesian Regularization). In the second stage (Chapter 5), the system is optimized with the help of the direct and inverse of the artificial neural network and genetic algorithms, which can increase the water temperature with a minimum of requirements. Finally, the transfer / activation functions are analyzed as one more tool that allows increasing the precision of the empirical model obtained by artificial neural networks. It should be noted that, for all the models developed, a robust database had to be prepared and they were subjected to various statistical tests in order to verify the reliability and adequacy of the simulated data with its experimental counterparts, which leads to reducing the uncertainty of the models considered good.

iii Agradecimientos

Elaborar esta tesis fue más difícil de lo que parecía en un principio. Por eso quiero agradecer de manera muy especial a mis dos asesores. Al Dr. José Alfredo Hernández que fue más que un asesor, fue lo que los japoneses denominan 'sensei', ya que no se me ocurre un término más adecuado, y que me ayudo con sus palabras y ánimos a concluir esta etapa. Y si el Dr. Alfredo es un sensei, entonces la Dra. Arianna Parrales es una 'gurú' porque siempre tenía la palabra adecuada en el momento justo, tanto para la realización del trabajo como en la vida cotidiana.

Al Dr. Armando Huicochea al que considero como un amigo además de profesor, quien siempre me impulsa a arriesgarme más en todos los ámbitos con el fin de ser mejor cada día.

A mis profesores dentro del CIICAp, que con pusieron a disposición sus conocimientos para enseñarme. Concretamente, aún recuerdo las primeras veces que entre al salón a tomar clases con los Dres. Álvarez y Sierra.

A Edgar Cardoso y Giselle Ramírez que, a través de sus trabajos realizados, se pudo realizar esta tesis. Asimismo, a Roberto Conde, ahora ya doctor, que sin más me ayudo y enseñó cuando se lo necesite, y al cual estimo mucho.

A mis amigos que se han ido del CIICAp y a los que todavía quedan, que sin ellos no habríamos podido disfrutar del fútbol y su amistad. A mis colegas del laboratorio Carolin, Ulises, Elusaí y Wassila porque siempre era un gusto y privilegio estar a su lado compartiendo estos momentos tan agradables.

A mi esposa e hija, porque el tiempo que pasó realizando esta tesis es tiempo que no estoy con ellas, por su comprensión, paciencia y ánimo. A mis padres por apoyarme de cualquier forma y sin ataduras.



UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS



Al CIICAp por permitirme estar, estudiar y trabajar en sus instalaciones para la realización esta tesis. Igualmente, a CONACyT por el apoyo económico que me permitió tener estabilidad.

A todos ustedes, en verdad muchas gracias, porque sin ustedes no habría sido posible.



Instituto de
Investigación en
Ciencias
Básicas y
Aplicadas

ANÁLISIS DE LA TRANSFERENCIA DE ENERGÍA EN COLECTORES SOLARES
DE CANAL PARABÓLICO APLICANDO REDES NEURONALES ARTIFICIALES

iv

iv Dedicatorias

A mis padres Emma Y Rafael que me siguen apoyando gracias a ellos que me dieron la oportunidad de estudiar y su crianza que me han hecho la persona que soy hasta ahora, que pienso seguir honrando.

A Gloria, mi esposa, que es mi contraparte en todo pero a la que admiró por la mujer que es y que no sabe en lo que se convertirá. Por compartir su vida conmigo y aceptarme y apoyarme tal cual soy,

A mi hija Hadassa, que es la persona más especial en vida, aunque yo no lo pudiera ni imaginar. Que, siendo muy pequeña, comprende que en ocasiones no puedo estar con ella.

A mis hermanos Diego y Gersaín, y sus respectivas familias que me hacen ver que la vida es más que solo libros y videojuegos.

A mis amigos quienes fueron ese gran escape y son un gran apoyo emocional. Que esas conversaciones que nunca llevan a nada no se acaben y que pasamos horas riéndonos del pasado, pero aceptando nuestro presente.

Aunque no soy muy religioso, también agradezco a Dios por todo lo que me ha dado, porque una persona sin fe cae en la soberbia y pierde el juicio.

A mis profesores que ponen su tiempo a disposición de nosotros los estudiantes y a veces no sabemos valorar. Por no desistir de enseñar.

A mis asesores que me ayudaron bastante y no tengo como agradecerles lo suficiente.

A los doctores que conforman mi comité sinodal quienes estudiaron mi tesis y la aprobaron.

A todos los que me apoyaron para escribir y concluir esta tesis. Para ellos es esta dedicatoria de tesis, pues es de ustedes esta tesis más que mía. Por su apoyo incondicional y por su confianza depositada en mí.



v Índice

<i>i</i>	<i>Resumen</i>	<i>i</i>
<i>ii</i>	<i>Abstract</i>	<i>ii</i>
<i>iii</i>	<i>Agradecimientos</i>	<i>iii</i>
<i>iv</i>	<i>Dedicatorias</i>	<i>v</i>
<i>v</i>	<i>Índice</i>	<i>vi</i>
<i>vi</i>	<i>Lista de Tablas</i>	<i>ix</i>
<i>vii</i>	<i>Lista de Figuras</i>	<i>xi</i>
<i>viii</i>	<i>Nomenclatura</i>	<i>xiii</i>
1	INTRODUCCIÓN	1
1.1	Antecedentes	1
1.2	Justificación.....	9
1.3	Objetivo General	11
1.3.1	Objetivos específicos:	11
2	DESARROLLO DE UN MODELO DE RED NEURONAL	12
2.1	Reseña sobre Redes Neuronales Artificiales.....	12
2.2	Principio de funcionamiento de las neuronas biológicas y artificiales.....	15
2.3	Elementos de una Red Neuronal Artificial.....	17
2.3.1	Capas	17
2.3.2	Conexiones de peso	18
2.3.3	Funciones de transferencia o de activación.....	19
2.3.4	Pre-procesamiento de datos	21
2.3.5	Algoritmo de aprendizaje	22
2.3.6	Sobre-aprendizaje	22
2.3.7	División de los datos	23
2.3.8	Método (algoritmo) de aprendizaje: Levenberg-Marquardt (LM)	25
2.3.9	Análisis estadístico.....	27
2.3.10	Pruebas de significancia	28
2.3.11	Prueba de linealidad	29
2.3.12	Análisis de sensibilidad de variables	32
2.4	La inversa de las Redes Neuronales Artificiales	33
2.4.1	Algoritmos Genéticos	36
2.4.2	Población, función fitness/objetivo y operadores genéticos	39





3	<i>Red neuronal artificial aplicada al rendimiento del sistema de energía renovable.</i>	45
3.1	Descripción del equipo experimental	45
3.2	Modelo de la red neuronal	47
3.2.1	Condiciones de operación	48
3.3	Resultados	53
3.4	Conclusiones parciales	59
4	<i>Condición de funcionamiento óptima para un colector tipo cilindro parabólico con componentes de bajo costo utilizando una red neuronal inversa y resuelto por algoritmo genético.</i>	60
4.1	Descripción del equipo experimental.	61
4.2	Condiciones de operación	62
4.3	Desarrollo del modelo neuronal	63
4.4	Resultados del modelo de red neuronal	66
4.4.1	Análisis de Variables	71
4.5	Desarrollo del modelo de red neuronal inverso	71
4.5.1	Estimación del valor adecuado mediante algoritmos genéticos.	74
4.6	Comparación del PTC de bajo costo	77
4.7	Conclusiones parciales	79
5	<i>Análisis de funciones de transferencia y normalizaciones en un modelo ANN que predice el transporte de energía en un colector solar de canal parabólico con tubo absorbedor ranurado</i>	81
5.1	Descripción del equipo experimental	81
5.2	Condiciones de operación	82
5.3	Desarrollo del modelo de red neuronal	85
5.3.1	Análisis estadístico	88
5.3.2	Pruebas de significancia	89
5.3.3	Prueba de linealidad	89
5.3.4	Análisis de sensibilidad de variables	89
5.4	Modelo de red neuronal	89
5.5	Resultados	91
5.6	Conclusiones parciales	110





6	<i>Conclusiones finales y recomendaciones</i>	112
6.1	Conclusiones finales	112
6.2	Recomendaciones	114
	Referencias	115



vi Lista de Tablas

Tabla 1.1. Consumo diario de energía en diferentes etapas del desarrollo humano (en MJ) modificado de [González Velasco 2009].	1
Tabla 1.2. Tabla de aplicaciones de baja y media temperatura [Pranesh <i>et al.</i> 2019].	6
Tabla 1.3. Trabajos con redes neuronales en sistemas térmicos.	8
Tabla 2.1. Tipos de funciones de activación/transferencia.	20
Tabla 2.2. Expresiones genéticas y su equivalencia en Algoritmos Genéticos [Ponce Cruz 2010]	37
Tabla 3.1. Intervalos de condiciones de operación experimentales.	49
Tabla 3.2. Comparación del análisis estadístico para determinar el número de neuronas en la capa oculta.	52
Tabla 3.3. Comparación entre parámetros estadísticos de R^2 , $RMSE$, y $MAPE$ utilizando el entrenamiento de algoritmos LM y BR.	54
Tabla 3.4. Prueba de linealidad de pendiente e intercepto.	56
Tabla 3.5. Parámetros de pesos y bias obtenidos con el entrenamiento del algoritmo BR.	57
Tabla 4.1. Intervalos de las condiciones experimentales utilizadas para obtener los valores de T_{Salida} .	62
Tabla 4.2. Pruebas con diferentes arquitecturas ANN.	67
Tabla 4.3. Aplicación de las pruebas de significancia de las pruebas de Fisher y t de Student.	68
Tabla 4.4. Parámetros de pesos y bias obtenidos para el modelo ANN desarrollado.	69
Tabla 4.5. Parámetros usados para la corrida de Algoritmos Genéticos (AG).	74
Tabla 4.6. Pruebas experimentales aplicadas al modelo ANNi-AG.	75
Tabla 4.7. Comparación entre el PTC de bajo costo con respecto otros tipos de colectores.	77
Tabla 5.1. Características geométricas del sistema experimental PTC.	82
Tabla 5.2. Condiciones de operación del sistema experimental PTC.	84
Tabla 5.3. Calcular el día del año.	85
Tabla 5.4. Ecuaciones de las funciones de transferencia TANSIG y LOGSIG.	88
Tabla 5.5. Análisis estadístico del número de neuronas en la capa oculta para el modelo base ANN.	92
Tabla 5.6. Pruebas de significancia del modelo base.	93



Tabla 5.7. Análisis estadístico de los modelos ANN con los intervalos de normalización propuestos para TANSIG y LOGSIG.	96
Tabla 5.8. Pruebas de significancia para la comparación de los mismos intervalos de TANSIG y LOGSIG.	97
Tabla 5.9. Valores de la prueba de linealidad para la comparación de los mismos intervalos de TANSIG y LOGSIG.	97
Tabla 5.10. Análisis estadístico para la comparación a los mismos intervalos de TANSIG y LOGSIG con salida normalizada.	101
Tabla 5.11. Pruebas de significancia para la comparación a los mismos intervalos de TANSIG y LOGSIG con salida normalizada.	102
Tabla 5.12. Valores de la prueba de linealidad (prueba pendiente-intercepto) para los modelos con salida normalizada.	105
Tabla 5.13. Matriz de pesos y bias para el mejor modelo ANN.	107



vii Lista de Figuras

Figura 1.1. Demanda de suministro de energía mundial por energético en 2017 [IEA 2019].	2
Figura 1.2. Clasificación de los sistemas de conversión de energía solar.	5
Figura 2.1. Partes de una neurona biológica.	16
Figura 2.2. Partes de una neurona artificial.	17
Figura 2.3. Red neuronal artificial con multicapas.	18
Figura 2.4. Comportamiento de los datos usando las funciones LOGSIG y TANSIG.	21
Figura 2.5. Análisis del fenómeno de "sobre-aprendizaje" de la red [Flores López y Fernández 2008].	23
Figura 2.6. Esquema de una Red Neuronal Artificial Inversa	34
Figura 2.7. Representación esquemática de los términos usados en Algoritmos Genéticos.	38
Figura 2.8. Representación de la selección por ruleta y de la selección por muestreo universal estocástico.	41
Figura 2.9. Cruzamiento en un punto de dos padres y su respectiva descendencia.	43
Figura 2.10. Cruzamiento en dos puntos de dos padres y su respectiva descendencia.	43
Figura 2.11. Cruzamiento uniforme con la creación de una máscara aleatoria para generar la descendencia de dos padres.	44
Figura 3.1. Vistas frontal y lateral con las dimensiones del sistema PTC.	46
Figura 3.2. Diagrama esquemático del sistema experimental PTC con componentes de bajo costo.	47
Figura 3.3. Metodología para el entrenamiento del modelo ANN para el sistema PTC.	50
Figura 3.4. Esquema del modelo ANN del sistema PTC.	51
Figura 3.5. Comparación estadística y regresión lineal entre los datos de temperatura de salida simulados (ANN) y experimentales.	54
Figura 3.6. Simulación de la temperatura de salida del colector con los modelos propuestos para diferentes días.	55
Figura 3.7. Importancia relativa de las variables de entrada del modelo ANN del sistema PTC.	58
Figura 4.1. Diagrama esquemático del sistema experimental PTC con componentes de bajo costo.	62
Figura 4.2. Diagrama de flujo para el modelo ANN del PTC.	64



Figura 4.3. Arquitectura aplicada a la red para los valores de T_{salida} y procedimiento usado para el aprendizaje de la red neuronal. 66

Figura 4.4. Comparación estadística entre los datos de T_{salida} experimentales y los simulados. 68

Figura 4.5. Comparación de la temperatura de salida del agua caliente del PTC entre los datos experimentales y los simulados por la ANN, con respecto al tiempo de un día a) para Mayo, b) para Junio. 70

Figura 4.6. Análisis de sensibilidad de variables para determinar la variable con mayor influencia en este modelo. 71

Figura 4.7. Arquitectura de la ANNi para determinar la $T_{suministro}$ óptima y encontrar el valor máximo de T_{salida} 73

Figura 4.8. Condiciones óptimas de la $T_{suministro}$ en el PTC con respecto a la T_{salida} simulada mediante ANNi, para el mes de a) Mayo, b) Junio y c) Enero. 76

Figura 5.1. Características geométricas del tubo raurado. 82

Figura 5.2. Diagrama esquemático del sistema experimental PTC 83

Figura 5.3. Diferentes configuraciones del sistema experimental PTC a) Con tapa, b) con tapa y distancia focal optimizada, C) sin tapa, y d) sin tapa y distancia focal optimizada. 84

Figura 5.4. Arquitectura del modelo neuronal del PTC con nueve entradas. 92

Figura 5.5. Prueba de linealidad del modelo base ANN. 93

Figura 5.6. Comparación entre los datos experimentales y simulados del modelo base ANN..... 94

Figura 5.7. Diagrama de flujo para el proceso de aprendizaje de los modelos ANN. 95

Figura 5.8. Intervalos de normalización para las funciones de transferencia TANSIG y LOGSIG..... 96

Figura 5.9. Comparación entre los datos experimentales y simulados de los modelos TANSIG..... 99

Figura 5.10. Comparación entre los datos experimentales y simulados de los modelos LOGSIG..... 100

Figura 5.11. Comparación entre los datos experimentales y simulados de los modelos TANSIG con salida normalizada. 103

Figura 5.12. Comparación entre los datos experimentales y simulados de los modelos LOGSIG con salida normalizada. 104

Figura 5.13. Comparación entre los datos experimentales y simulados del mejor modelo ANN..... 106

Figura 5.14. Importancia relativa de las variables de entrada del mejor modelo ANN del sistema PTC. 109

viii Nomenclatura

a Vector de datos simulados obtenidos de la ANN,

a Intercepto [$a = \bar{x}_{sim} - b\bar{x}_{exp}$]

b Pendiente $\left[b = \frac{\sum_{i=1}^n \{(x_{exp(i)} - \bar{x}_{exp})(x_{sim(i)} - \bar{x}_{sim})\}}{\sum_{i=1}^n (x_{exp(i)} - \bar{x}_{exp})^2} \right]$

$b_{(1,J)}$ Bia capa oculta

$b_{(2,K)}$ Bia capa de salida

E_w Suma de los pesos al cuadrado [$E_w = \frac{1}{n} \sum_{j=1}^n W^2$]

e Vector de errores de la red ($\mathbf{e} = \mathbf{t} - \mathbf{a}$)

F Función de rendimiento, Prueba de Fisher [$F = \frac{S_{exp}^2}{S_{sim}^2}$]

g Gradiente descendiente

H_0 Hipótesis nula

H_1 Hipótesis alterna

I Matriz identidad

i Día del calendario

I_p Importancia relativa de la variable de entrada

$$I_p = \frac{\sum_{j=1}^J \left(\left(\frac{|W_{(P,j)}|}{\sum_{r=1}^R |W_{(r,j)}|} \right) \times |W_{(j,K)}| \right)}{\sum_{r=1}^R \left\{ \sum_{j=1}^J \left(\left(\frac{|W_{(r,j)}|}{\sum_{r=1}^R |W_{(r,j)}|} \right) \times |W_{(j,K)}| \right) \right\}}$$

J Matriz Jacobiana

n Número de muestras

m Número de coeficientes (pesos y bias)

$MAPE$ Error Porcentual Medio Absoluto (Mean Absolute Percent Error)

$$[MAPE = \frac{\sum_{i=1}^n \left| \frac{x_{exp(i)} - x_{sim(i)}}{x_{exp(i)}} \right|}{n} \times 100 (\%)]$$

MSE Suma de los Errores al Cuadrado (Mean Square Error)

$$[MSE = \frac{1}{n} \sum_{j=1}^n (x_{exp} - x_{sim})^2]$$

n_j Término de entrada ANN

$$[n_j = Wi_{(j,1)} \cdot P_{(1)} + Wi_{(j,2)} \cdot P_{(2)} + \dots + Wi_{(j,R)} \cdot P_{(R)} + b_{(1,j)}]$$

$Output$ Valor de salida de la ecuación ANN

P Variables de entrada

$RMSE$ Raíz del Error Cuadrático Medio (Root Mean Square Error)

$$[RMSE = \sqrt{\frac{\sum_{i=1}^n (x_{sim(i)} - x_{exp(i)})^2}{n}}]$$

R^2 Coeficiente de determinación $[R^2 = 1 - \frac{\sum_{i=1}^n (x_{exp(i)} - x_{sim(i)})^2}{\sum_{i=1}^n (x_{exp(i)} - \bar{x}_{exp})^2}]$

S Desviación estándar combinada

$$[S = \sqrt{\frac{\sum_{i=1}^{n_{exp}} (x_{exp(i)} - \bar{x}_{exp(i)})^2 + \sum_{i=1}^{n_{sim}} (x_{sim(i)} - \bar{x}_{sim(i)})^2}{(n_{exp} + n_{sim} - 2)}}]$$

S^2 Varianza

S_b Coeficiente de regresión lineal de la pendiente

$$[S_b = \sqrt{\frac{\sum_{i=1}^n (x_{sim(i)} - \hat{x}_{sim(i)})^2}{(n-2) \sum_{i=1}^n (x_{exp(i)} - \bar{x}_{exp(i)})^2}}]$$

S_a Coeficiente de regresión lineal del intercepto

$$S_a = \sqrt{\frac{\sum_{i=1}^n (x_{sim(i)} - \hat{x}_{sim(i)})^2 \sum_{i=1}^n x_{exp(i)}^2}{n(n-2) \sum_{i=1}^n (x_{exp(i)} - \bar{x}_{exp})^2}}$$

T Temperatura

t Vector target (datos experimentales)

$$t \text{ Prueba t de Student } \left[t = \frac{|\bar{x}_{exp} - \bar{x}_{sim}|}{s \left(\sqrt{\left(\frac{1}{n_{exp}} + \frac{1}{n_{sim}} \right)} \right)} \right]$$

$t_{(n-2)}$ Valor crítico t de Student de tablas con $n - 2$ grados de libertad

W Peso correspondiente a la posición de las neuronas

W_i Pesos entre las capas de entrada-oculta

W_o Pesos entre la capa oculta y la capa de salida

x Valor de datos, variable a optimizar

\bar{x} Media de los datos de muestra

x_k Vector actual de pesos y bias

x_{k+1} Vector subsecuente de pesos y bias

$\hat{x}_{sim(i)}$ Valor de x_{sim} calculado por regresión lineal para el mismo valor en

$$x_{sim(i)} \left[\hat{x}_{sim} = a + b \bar{x}_{exp(i)} \right]$$

y Valor límite del intervalo de normalización

Letras griegas

- α Nivel de significancia, ángulo de anchura de ranura, parámetros del algoritmo BR, tasa de aprendizaje
- β Ángulo de inclinación de la ranura, parámetros del algoritmo BR
- δ Función de transferencia LOGSIG
- λ Función de transferencia PURELIN
- μ Intervalo de confianza, tamaño de paso algoritmo LM
- τ Función de transferencia TANSIG

Subíndices

- 98ts* Tabla t Student a nivel de significancia a 98%
- 99ts* Tabla t Student a nivel de significancia a 98%
- a* Intercepto
- almacén* Tanque de almacenamiento de agua caliente
- amb* Ambiente
- b* Pendiente
- entrada* Entrada del fluido
- ext* Extrapolación
- exp* Datos experimentales
- i* Valor i-esimo
- int* Interpolación
- J* Neuronas en la capa oculta,



- K número de neuronas en la capa de salida
- k valor de cada iteración
- max Máximo
- min Mínimo
- $norm$ Normalización
- P Variables de entrada
- R Neuronas en la capa de entrada
- $R \in x$ Número de neurona correspondiente a la variable a optimizar
- $reflector$ Superficie de reflector
- sim Datos simulados
- $suministro$ Tanque de alimentación
- $vidrio$ Superficie de vidrio

Superíndices

- T Matriz Transpuesta
- -1 Matriz Inversa
- cv Intervalo de confianza

Acronimos

- ADN Ácido Desoxirribonucleico
- AG Algoritmos Genéticos



UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS



ANN Redes Neuronales Artificiales (Artificial Neuronal Networks)
ANNi inversa de la Red Neuronal Artificial (Artificial Neuronal Network inverse)
BR Regularización Bayesiana (Bayesian Regularization)
CIICAp Centro de Investigación en Ingeniería y Ciencias Aplicadas
CONACyT Consejo Nacional de Ciencia y Tecnología
CPC Colector Parabólico Compuesto
CST Colectores Solares Térmicos
GA Algoritmos Genéticos (Genetic Algorithms)
LFR Colector de Reflector de Fresnel Lineal (Linear Fresnel Reflector)
LM Levenberg-Marquardt
LOGSIG Función de transferencia Logarítmica sigmoidea
PTC Colector Solar de Canal Parabólico (Parabolic Trough Collector)
PURELIN Función de transferencia lineal
TANSIG Función de transferencia Tangencial hiperbólica o sigmoidea
UAEM Universidad Autónoma del Estado de Morelos

1 Introducción

1.1 Antecedentes

Desde la antigüedad, la energía ha sido el insumo fundamental para el desarrollo de las actividades humanas. Con el crecimiento de la población, el consumo de energía ha alcanzado una magnitud excesiva que necesita satisfacerse. Hoy en día, se necesitan aproximadamente unos 937.2 MJ/día en promedio por cada persona [González Velasco 2009]. La Tabla 1.1 muestra el incremento en el consumo de energía desde el hombre primitivo hasta el hombre actual.

Tabla 1.1. Consumo diario de energía en diferentes etapas del desarrollo humano (en MJ) modificado de [González Velasco 2009].

Era del Hombre	Alimentos	Industria Agricultura	Comercial Residencial	Transporte	Total
Recolector	10.8784	0	0	0	10.8784
Cazador	12.552	0	8.368	0	20.92
Agricultor primitivo	16.736	16.736	16.736	4.184	54.392
Agricultor avanzado	14.644	29.288	50.208	4.184	100.416
Industrial	14.644	100.416	133.888	58.576	309.616
Actual (tecnológico)	14.644	380.744	276.144	263.592	937.216

En la actualidad, el avance tecnológico y el crecimiento económico de las naciones dependen principalmente de la transformación y el uso adecuados de las fuentes de energía proporcionadas por el medio ambiente: renovables y no renovables [Kumar y Hasanuzzaman 2019].

En los últimos años, la mayoría de la energía consumida en el mundo es no renovable; es decir, que una vez consumida, no hay forma de recuperarse.

Ejemplos de fuentes no renovables son la energía nuclear y la quema de combustibles fósiles como el petróleo, el carbón, el gas natural, entre otras. Durante 2017, la producción total de energía primaria fue de 587.617 EJ (o 14,035 Millones de toneladas equivalentes de petróleo, 1 toe=41,868 MJ), de los cuales aproximadamente el 81% provino de combustibles fósiles. La demanda de suministro fue de 584.979 EJ [IEA 2019], como se observa en la Figura 1.1.

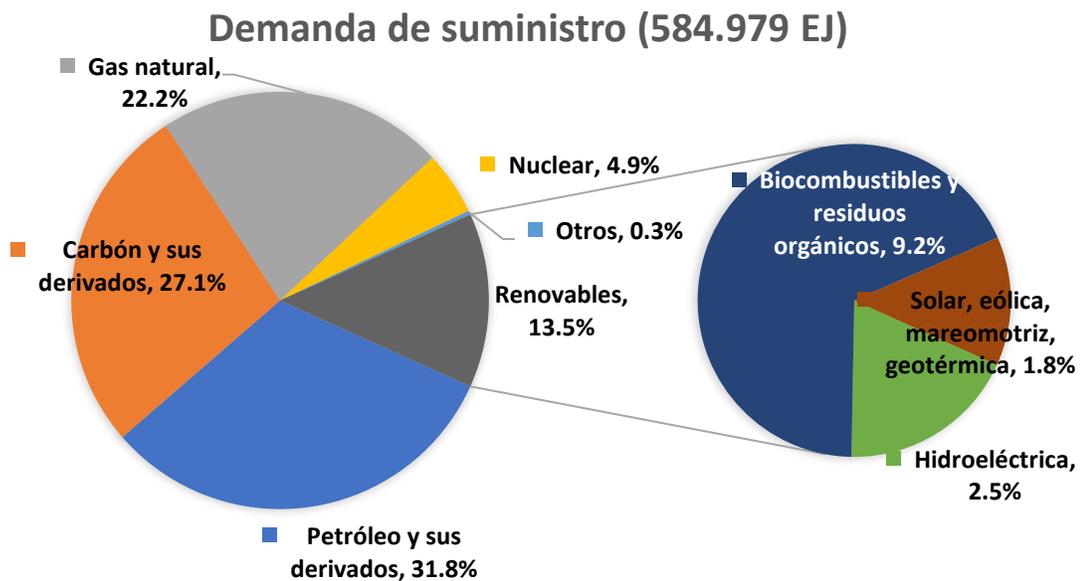


Figura 1.1. Demanda de suministro de energía mundial por energético en 2017 [IEA 2019].

Con el aumento en el uso de combustibles fósiles, causado por el desarrollo acelerado de la industria, se han generado consecuencias negativas en el medio ambiente. Como ejemplo, la liberación de dióxido de carbono a la atmósfera que induce altas temperaturas de calor en algunas regiones, contaminación del agua, sequías, lluvias ácidas, inundaciones, entre otros [Adam y Nsiah 2019].

En consecuencia, se ha buscado la cooperación internacional para reducir el cambio climático a través de 17 objetivos de desarrollo sustentable establecidos por la Organización de las Naciones Unidas, ONU [IRENA 2019]. A través del protocolo Kioto, del cual México es parte, la comunidad

internacional intenta reducir los gases de efecto invernadero (dióxido de carbono, metano, óxido nitroso, hidrofluorocarbonos, perfluorocarbonos y hexafluoruro de azufre) y va encaminado a promover acciones más amigables con el ambiente, el cual se complementa con el Plan de Acción de Bali, aprobado en 2007 [Carpenter 2008]. Este Plan de Acción tiene cuatro elementos claves para limitar el calentamiento global a menos de 2 °C por encima de la temperatura preindustrial, que son: **mitigación** de los gases invernadero, **adaptación** al cambio climático (ajuste en los sistemas naturales o humanos a los estímulos climáticos reales o esperados, o a sus efectos, que modera el daño o aprovecha las oportunidades beneficiosas.), las **finanzas** para hacer frente a esta mitigación y adaptación del cambio climático; y la **tecnología** que se necesita para tener un futuro (y presente) con baja emisiones y tener procesos cada vez más ecológicos [UNDP 2008]. El 12 diciembre del 2015, el Acuerdo de París fue adoptado por 195 países como una versión mejorada del Plan de Acción de Bali.

Por lo tanto, para cumplir con el Acuerdo de París, las naciones están desarrollando tecnologías enfocadas en obtener fuentes de energía renovables y reemplazar los combustibles fósiles. Con esto, se espera que las energías renovables representen dos tercios del suministro total de energía global y que el 60% de las emisiones de dióxido de carbono a la atmósfera se reduzcan para el año 2050 [IRENA 2019]

Wahab y su grupo de trabajo coinciden en que la energía producida por los recursos de energía renovable es la más sustentable con una emisión mínima de gases de efecto invernadero [Wahab *et al.* 2019]. Las energías renovables, también conocidas como energías limpias, son aquellas que proporciona la naturaleza, no se agotan por su uso y se renuevan continuamente o en períodos bajos, las cuales son: hidráulica, solar, eólica, mareomotriz o marina (marea, olas y energía del océano) y biomasa [Sharif *et al.* 2019].



En 2018, la proporción del uso de energías renovables en todo el mundo ha ascendido a 2356.34 GW, equivalente a un sexto del suministro total de energía. La energía hidroeléctrica representa la mayoría del total con una participación de más del 50% (1295.317 GW) seguido de la energía eólica con el 22.59% (563.659 GW). Como tercer lugar, la energía solar (fotovoltaica y térmica) es del 19.48% (486.085 GW) y, finalmente, otras energías que en conjunto representa casi el 6% del total, que son: bioenergía (117.828 GW), energía geotérmica (13.277 GW) y energía marina (529 MW) [IRENA 2019].

En particular, la energía solar es la fuente de energía renovable más limpia con un gran potencial y una posible solución sustentable comparada con cualquier otro consumo de energía [Moriarty y Honnery 2019]. Además, basados en las estimaciones de los científicos de la NASA, el Sol tiene menos de la mitad de su vida útil, lo que significa que durará unos 6,500 millones de años más hasta que se reduzca a una enana blanca [NASA 2019]. La energía solar se obtiene de las interacciones en el núcleo del Sol, donde el hidrógeno se convierte en helio y es transmitida a la Tierra por fotones conocidos como radiación solar [Foster *et al.* 2010].

El Sol produce una enorme cantidad de energía todos los días en cualquier lugar del mundo, pero la radiación solar que atraviesa la atmósfera varía dependiendo de factores tales como la posición geográfica, la latitud, la estación del año, el día, la hora y las condiciones atmosféricas como la nubosidad [Moriarty y Honnery 2019]. La radiación solar proporciona dos energías primarias: luz y calor que impulsan todos los procesos ambientales y que actúan directa o indirectamente en la superficie de la Tierra.

Como se esperaba, la radiación solar tiene un gran potencial para el consumo humano debido a que el flujo anual de energía solar es casi 100 veces mayor que el consumo anual de energía global. De acuerdo con Rypkema, la recolección de solo el 1% de la radiación solar incidente podría servir para satisfacer las necesidades de energía de cada persona en la Tierra. En cuanto



a las formas de aprovechar la energía solar, hay dos maneras de recolectar: ya sea energía solar térmica y energía eléctrica o ambas [Rypkema 2018].

Para aprovechar la energía solar se utilizan dos sistemas de conversión principales, las celdas fotovoltaicas y los colectores solares. En la Figura 1.2 se da la clasificación de los principales sistemas de conversión de energía solar.

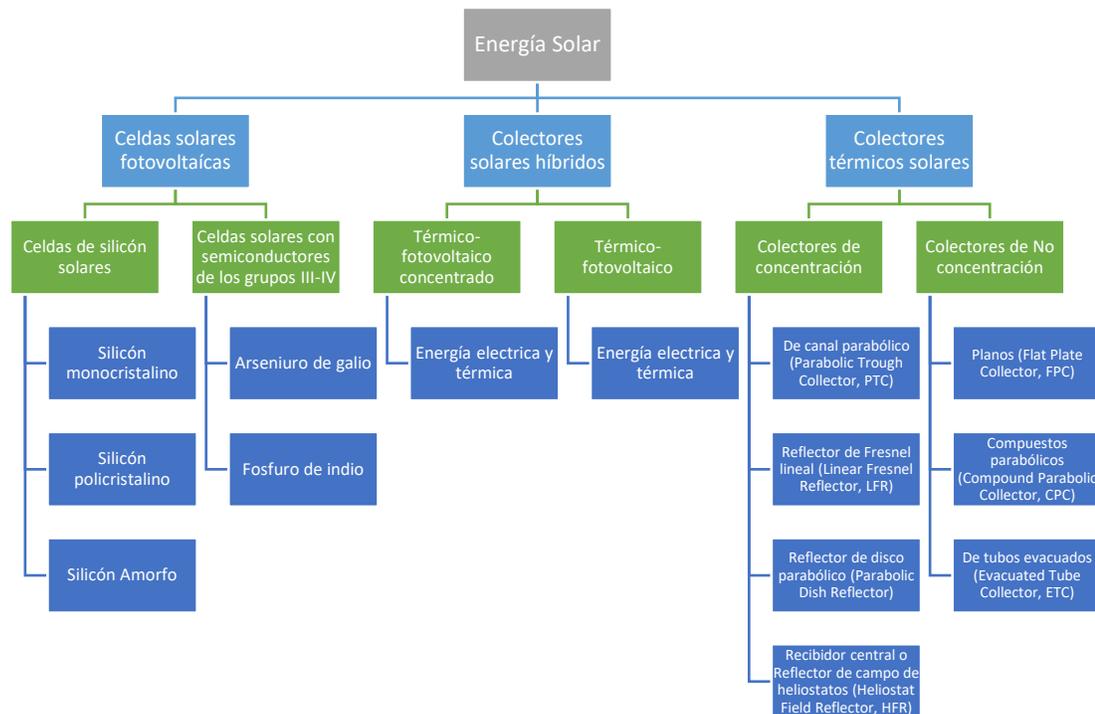


Figura 1.2. Clasificación de los sistemas de conversión de energía solar.

Específicamente, la energía solar térmica puede ser capturada por diversos tipos de colectores solares térmicos (CST). Los CST son un tipo particular de intercambiadores de calor que se utilizan para absorber la energía solar entrante, transformarla en calor y transferirla a un fluido de trabajo. El fluido de trabajo transporta la energía solar recolectada ya sea para ser utilizada o para depositarla en un tanque de almacenamiento aislado térmicamente [Shafieian *et al.* 2019]. Actualmente, estos sistemas pueden generar calor para fines de agua caliente sanitaria, calentamiento de zonas, cocción (hornos solares),

tratamientos de agua (desalinización) y algunos procesos industriales, entre otros [Evangelisti *et al.* 2019].

En concreto, el colector solar de canal parabólico (Parabolic Trough Collector, PTC) ha destacado por su sistema tecnológico de bajo costo en comparación con otras tecnologías de concentración solar [Jaramillo *et al.* 2016]. Además, los PTC tienen una amplia gama de disponibilidad térmica que permite cumplir con los requisitos de energía térmica de algunos procesos industriales o la generación de energía eléctrica.

La Tabla 1.2 muestra diferentes campos de aplicaciones de energía solar térmica (sectores industriales, aplicaciones de calefacción doméstica y sistemas de sorción) basados en diferentes niveles de temperatura de media a baja/alta temperatura.

Tabla 1.2. Tabla de aplicaciones de baja y media temperatura [Pranesh *et al.* 2019].

	Sector	Proceso	Intervalo de temperatura (°C)
A. Sectores industriales			
1	Lechería	Presurización	60–80
		Esterilización	100–120
		Secado	120–180
		Concentrados	60–80
		Agua de alimentación para calderas	60–90
2	Alimentos enlatados	Esterilización	110–120
		Pasteurización	60–80
		Cocinar	60–90
		Blanqueamiento	60–90
3	Textil	Blanqueamiento, teñido	60–90
		Secado, desengrasado	100–130
		Teñido	70–90
		Fijación	160–180
		Presionando	80–100
4	Papel	Cocinado, secado	60–80
		Agua de alimentación para calderas	60–90
		Blanqueamiento	130–150
5	Químicos	Jabones	200–260
		Caucho sintético	150–200
		Procesamiento de calor	120–180
		Precalentamiento de agua	60–90
6	Carnes	Lavado, esterilización	60–90
		Cocción	90–100
7	Bebidas	Lavado, esterilización	60–80

	Presurización	60–70
8	Harinas y producto	Esterilización 60–80
9	Subproductos de la madera	Vigas de difusión térmica 80–100
	Secado	60–100
	Pre calentamiento de agua	60–90
	Preparación de pulpa	120-170
10	Ladrillos y bloques	Curado 60–140
11	Plásticos	Preparación 120–140
	Destilación	140–150
	Separación	200–220
	Extensión	140-160
	Secado	180–200
	Mezcla	120–140
B.	Aplicaciones domésticas de agua caliente	Temperatura (°C)/ Consumo (l)/ Calor útil (Wh)
	Lavavajillas (por día, por persona)	50/12–15/550–700
	Lavado de manos (por día, por persona)	37/3–5/95–160
	Bañera	40/250*/5,200
	Ducha (5 min, por persona) *	37/150*/940–1400
	Lavado de cabello (por persona)	37/10–15/310–470
	Desalinización	60 a 90°C/-/-
C	Aplicaciones de enfriamiento/calentamiento por sorción	Temperatura (°C) (enfriamiento/calentamiento) y (COP_{enfriamiento}/COP_{calentamiento})
	Medio efecto H ₂ O-LiBr	67/108 y 0.35/1.33
	Simple efecto H ₂ O-LiBr	92/- y 0.72/-
	Doble efecto H ₂ O-LiBr	142/- y 1.29/-
	Simple efecto NH ₃ - H ₂ O	86/140 y 0.6/1,52
	Adsorción H ₂ O- gel de silicio	90/120 y 0.41/1.26

*[Clickgasoil 2017]

Pero hasta el día de hoy es difícil modelar analíticamente los sistemas de energías renovables, incluidos los colectores solares, porque dependen de muchos factores externos. Estos sistemas dependen de las condiciones climáticas las cuales suelen ser difíciles de predecir, por tal motivo se requiere de una herramienta computacional especializada para poder modelar estos sistemas. La herramienta propuesta que se propone para este trabajo es la de Redes neuronales artificiales, que ha sido ampliamente usada con éxito en este campo. Además, nos permite optimizar parámetros mediante la Red neuronal artificial inversa.

Los datos de los sistemas de energías renovables son inherentemente con ruido, por eso son buenos candidatos los problemas de este tipo para ser



manejados con redes neuronales [Kalogirou 2001]. Algunos de los trabajos más relevantes de redes neuronales artificiales aplicados a colectores solares se resumen en la **¡Error! No se encuentra el origen de la referencia.** Tabla 1.3.

Tabla 1.3. Trabajos con redes neuronales en sistemas térmicos.

Autor	Arquitectura de la red	Aplicación	Parámetros predichos	Número de neuronas
Kalogirou(1999)	RMFF	SCAg	Energía extraída útil y el incremento de la temperatura en el tanque de almacenamiento.	8-18-18-18-2
Kalogirou(2000)	RMFF	SCAg	Energía de salida. Energía de salida y agua caliente por mes.	13-5-5-5-1 14-7-7-7-2
Farkas (2003)	RMFF	SCAg	Temperatura de salida.	3-7-7-1
Kalogirou(2005)	RMFF	SCAg	Ecuación de Rendimiento (sin viento). Ecuación de Rendimiento con viento). Incidencia del ángulo modificador. Constante de tiempo. Temperatura de estancamiento. Capacidad calorífica.	25-7-7-7-3 25-7-7-7-3 12-6-6-6-2 12-4-4-4-1 12-4-4-4-1 15-6-6-6-1
Sözen <i>et al.</i> (2008)	RMFF	SCAg	Determina la eficiencia de un colector de placa plana.	7-20-20
Caner (2011)	RMFF	SCAir	Rendimiento del colector.	8-20-1
Benli (2013)	RMFF	SCAir	Rendimiento del colector.	3-3-1
Kalogirou(2013)	EMFF	SCAg	Energía diaria en condiciones típicas. Nivel de energía en el tanque de almacenamiento al final de un día de operación.	3-5-5-52 4-7-7-7-2
Hüseyin (2013)	RMFF	SCAg	Tipos de superficie.	6-3-1
L.V. Kamble (2014)	RMFF	THILP	Número de Nusselt y Coeficiente de transferencia de calor.	3-5-2
May Tzuc <i>et al.</i> (2017)	RMFF-GA	PTC	Eficiencia térmica y flujo óptimo.	7-3-1
Díez <i>et al.</i> (2019)	RMFF	FPC	Temperatura de salida	4-5-1, 4-3-1
Heng <i>et al.</i> (2019)	RMFF-SP	PTC	Temperatura en la salida para un solo pulso	3-68-68-1

*Sistema de Colectores de Agua (SCAg), Sistema de Colectores de Aire (SCAir), Red Multicapa "FeedForward" (RMFF), Grupo de Método de Datos Manejables (GMDH), Algoritmos Genéticos (GA), Tubos Horizontales Inmersos en Lechos de Partículas (THILP), Colector de Canal parabólico (PTC), Principio de superposición (SP)

Conjuntamente, se propone aumentar la eficiencia del colector con el uso de un tubo internamente ranurado, como tubo absorbedor, con el fin de aumentar la transferencia de calor hacia el fluido de trabajo. Los tubos ranurados o con aletas internas, son los que tienen mejor rendimiento termo-hidráulico (mayor transferencia de calor con menor pérdida de presión) comparado con las técnicas de insertos trenzados, tubos corrugados y tubos con hoyuelos, donde se demuestra que estos últimos tienen mayor caída de presión (o penalizaciones hidráulicas) mientras que para los insertos trenzados tienen su

mejor desempeño en flujos laminares y transitorios (donde son mejores usarlos con fluidos de alta viscosidad) [Wen-Tao *et al.* 2015]. Cuando el fluido de trabajo es gas, los insertos trenzados la eficiencia es menor debido a que se tienen mayores caídas de presión [Wen-Tao *et al.* 2017]. Solo se tienen diferencias significativas entre la forma geométrica circular y trapezoidal con Re superiores a 30,000 [Bilen *et al.* 2009]. Por consiguiente, las ranuras son mejor opción porque son más fáciles de diseñar y menos costosas.

Por último, se analizan las funciones de transferencia/activación y la repercusión que tienen los intervalos de normalización en cada una de estas funciones. Con motivo de obtener un modelo de red neuronal con mayor precisión, puesto que los resultados de una red se utilizan para optimizar la red neuronal artificial inversa.

1.2 Justificación

Estos estudios pretenden modelar el proceso de un sistema de energía renovable, en específico, el calentamiento de agua en un colector solar de bajo costo, a través de la medición de diferentes parámetros para predecir la temperatura de salida. Optimizar dichos parámetros de entrada para conseguir la mayor temperatura disponible con un mínimo de requerimientos. Además de, aumentar la temperatura de salida o calor absorbido mediante técnicas de mejoramiento de transferencia de calor. Esto con el fin de hacer más eficiente el equipo, con lo cual, se pretende disminuir el uso de los combustibles fósiles y/o reducir el consumo de energía de la red eléctrica (para ser más sustentables).

A su vez, el tener un colector de bajo costo en casa nos ayuda a reducir aún más los costos en la instalación de un colector solar típico, debido a que los materiales utilizados son más baratos.

Esta investigación aplicará un modelo matemático empírico basado en la herramienta computacional llamada Redes Neuronales Artificiales (Artificial Neuronal Networks, ANN), la cual utiliza una base de datos, que se adquiere a partir de varias mediciones experimentales, y se usará para poder estimar la temperatura del agua a la salida del colector.

Posteriormente, se realizará un Red Neuronal Artificial Inversa (Artificial Neuronal Network inverse, ANNi), mediante algoritmos genéticos para tener una temperatura de salida del agua óptima.

Además, se añadirá una técnica de mejoramiento de transferencia de calor, tubos ranurados en concreto, para aumentar la energía que se transfiere al fluido de trabajo (agua) y así, aumentar su temperatura.

Finalmente, realizar un análisis de funciones de transferencia y normalizaciones en un modelo ANN para mejorar la precisión de la red.

El uso de la ANN es porque el fenómeno es muy complejo y no lineal, debido a que hay parámetros que varían con el tiempo por diversos factores que han demostrado ser muy poderosas en el modelado de colectores solares y sistemas térmicos. Pueden ser más veloces y simples comparados con métodos algorítmicos convencionales, asimismo la red tiene la capacidad para aprender de ejemplos. Combinándolo con electrónica de control, se puede automatizar el sistema para que usuario disponga del agua a una temperatura deseada restringida a un intervalo de operación.

1.3 Objetivo General

Analizar la transferencia de energía en colectores solares de canal parabólico considerando tubo absorbedor liso y ranurado mediante el uso de redes neuronales artificiales.

1.3.1 Objetivos específicos:

- 1.- Desarrollar redes neuronales artificiales, así como su aplicación al rendimiento de un sistema de energía renovable (colectores solares de canal parabólico).
- 2.- Condiciones de funcionamiento óptimos para un colector tipo cilindro parabólico con componentes de bajo costo utilizando una red neuronal inversa y resuelto por algoritmo genético.
- 3.- Analizar funciones de transferencia y normalizaciones en un modelo ANN que predice el transporte de energía en un colector solar de canal parabólico con tubo absorbedor ranurado.

2 Desarrollo de un modelo de red neuronal

2.1 Reseña sobre Redes Neuronales Artificiales

El sistema nervioso, el sistema más complejo del cuerpo humano, comprende muchos miles de millones de células interconectadas llamadas neuronas organizadas en vías anatómicas que están involucradas en funciones sensoriales, motoras y cognitivas. Es el patrón de conexiones entre estas neuronas a través de un complejo proceso químico lo que determina las funciones cerebrales y, por ende, la capacidad intelectual [Fischer *et al.* 2012]. La organización anatómica en una secuencia de neuronas y su capacidad para responder a estímulos externos ha inspirado a los investigadores a crear modelos que reproduzcan este comportamiento biológico.

Los antecedentes relacionados con la estructura y la fisiología del sistema de aprendizaje humano han servido de base para proponer sistemas expertos en conocimientos incorporados en sí mismo, conocidos como sistemas conexionistas. De esta manera, la investigación en el campo neurobiológico ha progresado en estrecha colaboración con la comunidad matemática e informática. El principio central del conexionismo es que los fenómenos pueden describirse mediante bloques de unidades simples y frecuentemente iguales que se interconectan generando lo que hoy conocemos como modelos de **redes neuronales artificiales ANN (Artificial Neural Networks)** [Sugavaneswaran 2017].

Históricamente, el uso de modelos de redes neuronales comenzó con Warren McCulloch y Walter Pitts en 1943, quienes propusieron un modelo de elemento informático para representar la activación de una neurona como función umbral. Como su homólogo biológico, la neurona realizaba una suma ponderada de las entradas seguida de una operación lógica de umbral. Sin embargo, el principal inconveniente de este modelo de cálculo era que la

neurona no podía aprender de ejemplos, ya que establecía valores fijos para las conexiones [McCulloch y Pitts 1943].

Años más tarde, en 1949, Donald Hebb demostró que el concepto de percepción está representado en el cerebro por un conjunto de neuronas activas simultáneamente [Glanzman 2005]. Fue el primero en explicar los procesos del aprendizaje, un elemento esencial de la inteligencia humana, desde el punto de vista psicológico. Este descubrimiento permitió el desarrollo de un procedimiento que mostraría cómo las neuronas fortalecían las conexiones a través de un esquema de aprendizaje iterativo, el cual ajustaba los valores de conexión. Este esquema se convertiría en una regla fundamental de aprendizaje en la literatura sobre redes neuronales [Yegnanarayana 2006].

En 1957 se publicó el primer modelo de red neuronal conocido como perceptron, el cual fue desarrollado por Frank Rosenblatt e inspirado en los estudios de Donald Hebb. Perceptrones fue el nombre genérico asignado a una familia de modelos de redes neuronales artificiales teóricas y experimentales que Rosenblatt propuso en el período 1957-1962 [Kanal 2001]. El perceptron es un sistema clasificador capaz de aprender algo a partir del reconocimiento de patrones geométricos y abstractos. Como el primer algoritmo simple de aprendizaje, el perceptron marco el inicio del uso de esquemas neuronales para manejar con éxito problemas de clasificación linealmente separables [Gudivada *et al.* 2016]. Actualmente, los nombres de Perceptrón y Perceptrón Multicapa se utilizan para referirse a estructuras de ANN específicas basadas en los perceptrones de Rosenblatt.

Hoy en día, las redes neuronales artificiales se enmarcan dentro del campo de la inteligencia artificial (IA). La IA se define como la capacidad de imitar la inteligencia humana en los ámbitos de la comprensión, percepción, resolución de problemas y toma de decisiones con la finalidad de reproducirlos mediante neuronas artificiales diseñadas y entrenadas en relación a un objetivo;

teniendo como característica fundamental la habilidad de autoaprendizaje [Shabbir y Anwer 2015]. Con los avances en la inteligencia artificial y la informática se hace posible que las ANN aprendan a resolver problemas complejos en un tiempo razonable.

Las ANN se han definido como una herramienta de modelado matemático de datos estadísticos no lineal que puede dar solución a problemas complejos en distintos campos de conocimiento [Penm *et al.* 2013]. Las ANN son capaces de aprender relaciones fuertemente complejas no lineales y multivariantes a partir del conjunto de ejemplos para entrenamiento asociados con fenómenos que son difíciles de resolver con métodos convencionales [Jensen y Abonyi 2005].

Están compuestas de un gran número de simples elementos de procesamiento o nodos altamente interconectados en niveles operando en paralelo [Elmolla *et al.* 2010] inspirados en la forma, conexión y procesamiento de información de las **neuronas biológicas**. Estos elementos procesan y envían información entre sí modificando sus respuestas dinámicas frente a variables externas mediante la experiencia acumulada. La arquitectura paralela permite que procesen eficientemente volúmenes de datos de alta dimensión y a menudo incompletos. Son aptas para resolver problemas que no poseen un algoritmo claramente definido para transformar una entrada en una salida; por lo que, mediante la combinación de parámetros de ajuste en un proceso de aprendizaje automático, las ANN aprenden, reconocen y aplican relaciones entre variables.

Las ANN son una excelente herramienta para el tratamiento de la información destacándose por 3 principios importantes: el aprendizaje adaptativo, la generalización y la abstracción [Du y Swamy 2014].

- a) **Aprendizaje adaptativo:** La experiencia de la red se obtiene a partir del aprendizaje adaptativo. Las ANN aprenden y se forman a sí mismas

adaptando su comportamiento en función del entorno para con ello producir un valor de salida consistente.

- b) Generalización:** Las redes ofrecen respuesta a datos que no ha aprendido con anterioridad pero que pueden inferir con base a su aprendizaje durante el entrenamiento. Las redes pueden reconocer patrones de información con ruido, distorsión o incompletos.
- c) Abstracción:** Las ANN son capaces de abstraer las características principales de un conjunto de datos que aparentemente no presentan aspectos comunes o relativos.

2.2 Principio de funcionamiento de las neuronas biológicas y artificiales

Las neuronas son las células principales del tejido nervioso que se interconectan para formar redes de comunicación que transmiten señales neuroeléctricas (impulsos nerviosos) por sectores específicos del sistema nervioso. Esta interacción entre redes de neuronas da como resultado funciones complejas del sistema nervioso que serían imposibles de realizar para una sola neurona en particular [Koenig 2015]. Para poder comunicarse, las neuronas biológicas utilizan cuatro regiones estructurales, como se ve en la Figura 2.1:

- **Dendritas:** Son prolongaciones que se extienden desde el soma con múltiples ramificaciones que conforman una gran área superficial, aumentando la capacidad de conexiones sinápticas con otras células nerviosas que generalmente vienen de otros axones.
- **Soma o cuerpo celular:** además de albergar al núcleo, en su capa externa tiene la capacidad única de generar impulsos nerviosos. Integra las diversas señales eléctricas que recibe de las dendritas, y si tienen la magnitud suficiente, producirá una señal eléctrica rápida llamada potencial de acción. En el **núcleo** se tiene el ADN y toda la información

celular además de dirigir el funcionamiento celular. En las neuronas maduras, los cromosomas ya no se duplican y sólo funcionan en la expresión genética.

- **Axón:** Nace del soma, conduciendo el estímulo o potencial de acción desde el soma hacia la terminal sináptica, donde hace contacto con otra célula nerviosa, muscular o glandular. Es la responsable de activar, inhibir e interactuar con otras neuronas que, a su vez, se activan en conjunto por miles. Un axón podría extenderse desde la columna vertebral hasta un dedo del pie. Forman los nervios que van desde el cerebro y medula espinal hacia todo el cuerpo.
- **Sinapsis:** Son regiones terminales que viene desde el axón. Desde aquí se liberan los neurotransmisores en respuesta al potencial de acción que envían la información para estimular a la célula contigua que puede ser nerviosa, muscular o glandular [Audesirk et al. 2008].

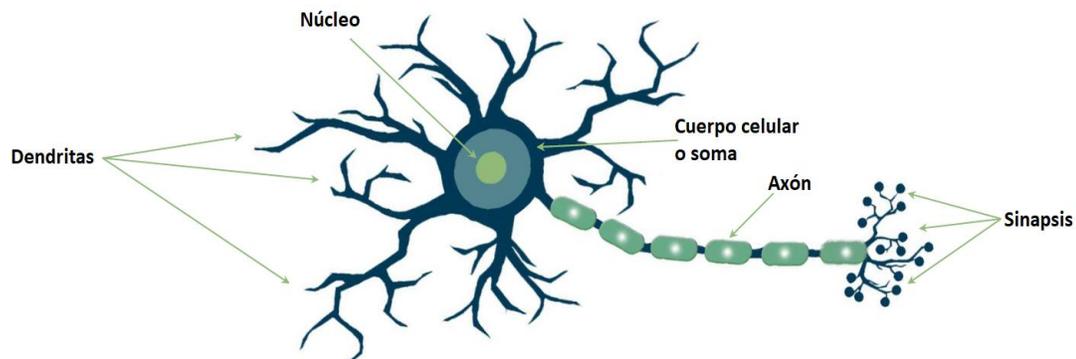


Figura 2.1. Partes de una neurona biológica.

Debido a la inspiración en la célula fundamental del sistema nervioso, los elementos de procesamiento reciben el nombre de neuronas artificiales. Están organizadas e interconectadas de forma similar a como se presentan en el cerebro biológico, donde cada neurona recibe y combina señales provenientes de neuronas previas y se comunica hacia otras neuronas, que son sometidas a diversas operaciones con la finalidad de alcanzar cierto umbral y producir

valores de respuesta de salida específica [Glinski *et al.* 2011]. La analogía de una neurona artificial con respecto a una biológica se muestra en la Figura 2.2.

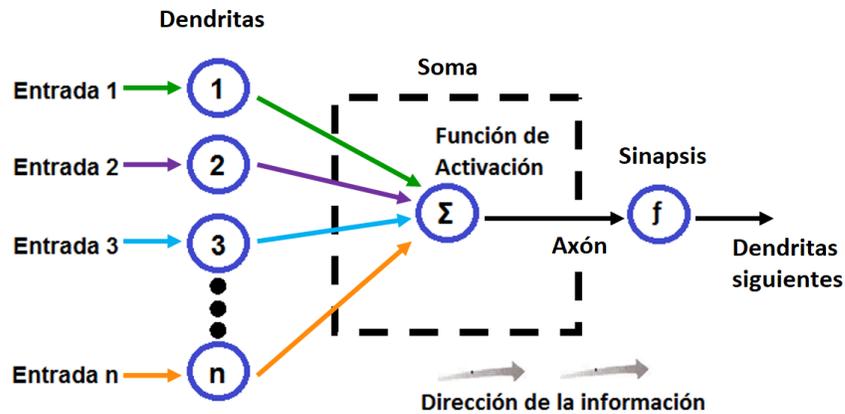


Figura 2.2. Partes de una neurona artificial.

2.3 Elementos de una Red Neuronal Artificial

2.3.1 Capas

La ANN no es un elemento aislado, sino que está compuesto por varias neuronas artificiales en interrelación. Generalmente en una red neuronal artificial, las neuronas están organizadas en tres grupos principales llamados niveles o capas [Ghorbani *et al.* 2013]: la capa de entrada, la capa oculta y la capa de salida. Dos capas establecen conexiones con el exterior:

- **Capa de entrada:** Es donde se presentan los datos de entrada al modelo de red. Estos datos son un valor finito de la representación directa referente a las mediciones de un proceso, sistema o fenómeno delimitadas por el intervalo de operación.
- **Capa de salida:** En esta sección, se presenta la respuesta de la red con respecto a la información de entrada.
- **Capa(s) oculta(s):** Pueden ser una o varias y se encuentran entre la capa de entrada y la capa de salida [Parrales *et al.* 2018]. Constituyen en general la mayor parte de la red neuronal. Una

representación general de una red neuronal artificial se muestra en la Figura 2.3, donde cada nodo circular representa una neurona artificial y cada flecha representa una conexión desde la salida de cada neurona a la entrada de otras ubicadas en las distintas capas.

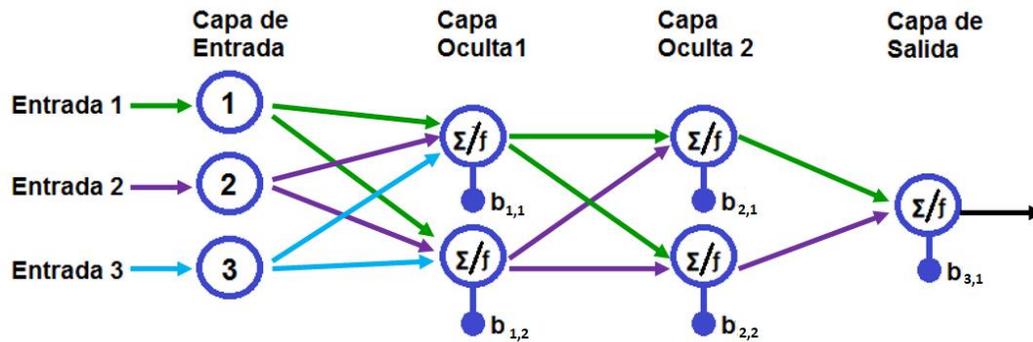


Figura 2.3. Red neuronal artificial con multicapas.

2.3.2 Conexiones de peso

El interés de las ANN reside principalmente en las conexiones interneuronales y en la transmisión de información que se lleva a cabo entre estas. Como se conoce biológicamente, cuando la neurona es estimulada a través de sus dendritas se presenta un cambio en el potencial eléctrico. A veces no es suficiente un potencial para llevar a la neurona al umbral, pero puede sumarse junto con otros para desencadenar un potencial de acción y lograr transformarse de manera transitoria en una señal. Entonces, la neurona receptora suma o integra todas las señales inhibitorias y excitatorias que recibe para decidir si activar o no una señal hacia el axón llamada potencial de acción. Esta señal se transmite hacia una neurona receptora permitiendo la transmisión de paquetes de información unidireccional, proceso conocido como sinapsis [Nierhaus 2009]. Este comportamiento dinámico de señalización representa un factor clave en el procesamiento eficiente de la información.

En el caso de la neurona artificial, el análogo al potencial eléctrico es representado por un valor numérico llamado **peso** sináptico, el cual puede ser

positivo si la neurona aumenta el estado de activación de la neurona adyacente o negativo si lo inhibe. Por consiguiente, el peso se define como un parámetro ajustable que establece la influencia que tiene una neurona sobre otra representando la intensidad de la señal. Dentro de la red, cada neurona artificial realiza el cálculo matemático de la suma ponderada de los valores de entrada junto a su factor correspondiente de los pesos sinápticos para lograr producir un valor de salida simulando el potencial de acción. De esta forma, la conexión entre neuronas y capas tiene asociado un valor de peso y un valor de compensación conocido como **bia**, que permite que la ANN aprenda [Hagan *et al.* 2014]. Entonces, el valor de entrada para una neurona en la capa oculta se asigna mediante la siguiente expresión:

$$n_j = W_{i(J,1)} \cdot P_{(1)} + W_{i(J,2)} \cdot P_{(2)} + \dots + W_{i(J,R)} \cdot P_{(R)} + b_{(1,J)} \quad (2.1)$$

Donde W_i son los coeficientes de los pesos de conexión entre la capa de entrada y la capa oculta, P son las variables de entrada, J es el número de neuronas en la capa oculta, R es el número de neuronas de entrada y $b_{(1,J)}$ es la bia correspondiente a la neurona en la capa oculta.

2.3.3 Funciones de transferencia o de activación

En la sinapsis artificial, la neurona calcula su grado de conectividad a partir de la función de activación o función de transferencia, la cual representa la tasa potencial a la cual una neurona se activa o no. Al igual que el axón transmite cierta información hacia otras neuronas biológicas, la función de transferencia emite una señal de salida en un determinado intervalo de amplitud; que dependerá de la función utilizada que es el resultado modificado de la suma de entradas. La señal emitida será la entrada a la siguiente neurona de la capa subsecuente de la red neuronal.

Existen diferentes tipos de funciones de transferencia o activación, como funciones inversas sigmoidales, lineales, sigmoidales, sigmoidales, gaussianas y gaussianas, mostradas en la Tabla 2.1. Los más destacados son

para la capa oculta, la función tangencial-sigmoidal (TANSIG, τ) y la logarítmica-sigmoidal (LOGSIG, δ). Normalmente, se utiliza una función de transferencia lineal (PURELIN, λ) para la capa de salida [Qin 1997].

Tabla 2.1. Tipos de funciones de activación/transferencia.

Nombre de la función	Descripción matemática	Icono	Acrónimo
Escalón	$a = 0$ si $n_j < 0$ $a = 1$ si $n_j \geq 0$		hardlim
Escalón simétrico	$a = -1$ si $n_j < 0$ $a = 1$ si $n_j \geq 0$		hardlims
Lineal	$a = n_j$		purelin
Lineal saturada	$a = 0$ si $n_j < 0$ $a = n_j$ si $0 \leq n_j \leq 1$ $a = 1$ si $n_j > 1$		satlin
Lineal simétrica saturada	$a = -1$ si $n_j < -1$ $a = n_j$ si $-1 \leq n_j \leq 1$ $a = 1$ si $n_j > 1$		satlins
Lineal positiva	$a = 0$ si $n_j < 0$ $a = n_j$ si $n_j \geq 0$		poslin
Logarítmica sigmoidea	$a = \frac{1}{1 + e^{-n_j}}$		logsig
Tangencial hiperbólica o sigmoidea	$a = \frac{e^{n_j} - e^{-n_j}}{e^{n_j} + e^{-n_j}} = \frac{2}{1 + e^{-2n_j}} - 1$		tansig
Competitiva	$a = 1$ si n_j es máximo $a = 0$ en otro caso		compet
Base Radial o Gaussiana	$a = e^{-n_j^2}$		radbas
Complemento gaussiano	$a = 1 - e^{-n_j^2}$		-
Base triangular	$a = 1 - n_j $, si $-1 \leq n_j \leq 1$ $a = 0$ en otro caso		tribas

2.3.4 Pre-procesamiento de datos

Como se ha mencionado antes, la red neuronal artificial puede ser entrenada con datos que contienen algún tipo de ruido; sin embargo, el pre-procesamiento de datos podría ser necesario. Una buena práctica ha sido eliminar datos desviados (outliers) a través de técnicas estadísticas [Verma 2005]. Además, si la aplicación o sistema lo permiten, se puede reducir el número de decimales de los valores experimentales adquiridos.

Es esencial usar la normalización (o estandarización) de los datos para ajustar a la misma escala; lo que permite comparar sus proporciones en lugar de sus magnitudes. Asimismo, la normalización de la base de datos previene la saturación matemática, ya que las funciones de transferencia sigmoideas tienden a saturarse cuando tienen valores superiores a +3 o menores a -3, como se puede ver en la Figura 2.4. Entonces, al tener un número alto en los valores de las variables de entrada, la cantidad asignada a los pesos debe ser mínima y, por lo tanto, esto haría que el entrenamiento sea muy lento [Han *et al.* 2012]. El intervalo de normalización de las variables de entrada debe ser relacionada con la función de activación que se decide a utilizar. Usualmente, se realiza en intervalos de -1 a 1, para lo cual existen diferentes ecuaciones en la literatura y se discutirá más a fondo en el Capítulo 5 de esta tesis.

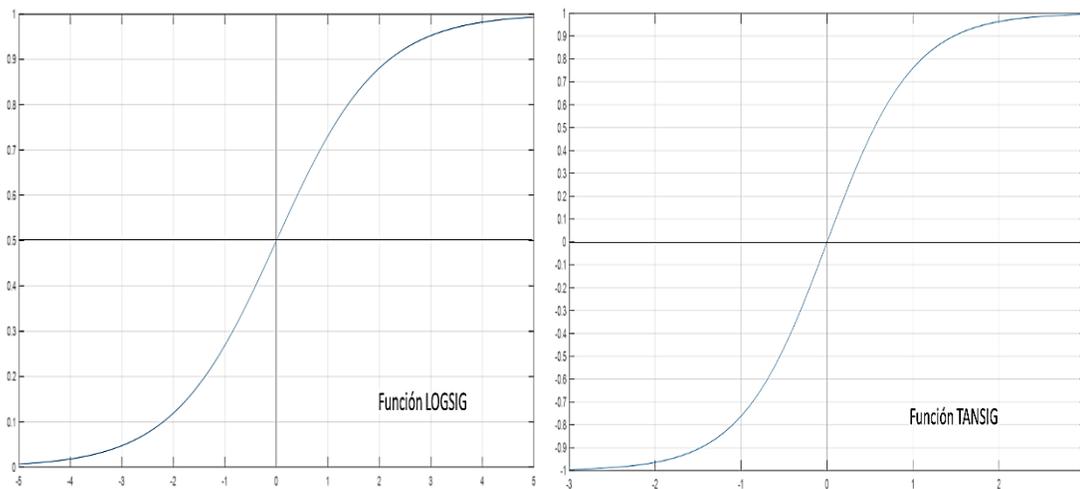


Figura 2.4. Comportamiento de los datos usando las funciones LOGSIG y TANSIG.

2.3.5 Algoritmo de aprendizaje

Después del pre-procesamiento de los datos, el aprendizaje de una red neuronal se lleva a cabo modificando y, de ser necesario, se pueden crear nuevas conexiones interneuronales. Por lo tanto, la red aprende examinando patrones de datos y generando la predicción mediante el ajuste de los valores asignados a pesos y las bias [Jafar *et al.* 2010]. Este proceso de aprendizaje se logra con el algoritmo de retropropagación (backpropagation).

El algoritmo de retropropagación consiste en introducir las variables de entrada a la red, obteniendo una salida final en la última capa y comparar este resultado con el resultado esperado (aprendizaje supervisado) [Curcio y Iorio 2013]. Luego, los valores de los pesos y las bias en las capas (salida, oculto y entrada) se modifican hacia atrás para minimizar el error entre la salida deseada y la salida generada (con el método del gradiente descendente) [Jothilakshmi y Gudivada 2016]. Este proceso se repite consecutivamente hasta que la red alcanza uno o más criterios de detención [Chakraborty *et al.* 2019], tales como: reducir el error entre el valor deseado y el real alcanzar un número definido de iteraciones u obtener una mayor precisión.

2.3.6 Sobre-aprendizaje

La proporción de cada contribución dependerá de los resultados del ajuste de pesos y bias obtenidos por la red para alcanzar un valor de salida con alta precisión. Es importante mencionar que, durante el entrenamiento, se debe evitar el sobre-aprendizaje [Bakr y Negm 2012]. Debido a que la red no será capaz de generalizar problemas con nuevos datos e indicará resultados erróneos. Una práctica común para prevenir el sobre-aprendizaje ha sido mantener un bajo número de neuronas en la capa oculta, en relación con el número de muestras de la base de datos, y dependiendo de los errores obtenidos durante el proceso de aprendizaje. Además, se debe considerar que

cada neurona en la capa oculta aumenta significativamente el número de coeficientes en el modelo de salida.

También se puede medir el “error de generalización” mediante un conjunto de pruebas estadísticas durante las fases de prueba y validación, y no solo durante el entrenamiento, como se muestra en la Figura 2.5. El sobre-aprendizaje puede ser evitado usando procesos de validación cruzada, es decir, entrenando, probando y validando la red a la vez, a fin de detectar un punto óptimo de aprendizaje. Estos procesos de validación son ampliamente aplicados en la fase de desarrollo en redes de Perceptrón Multicapa [Flores López y Fernández 2008].

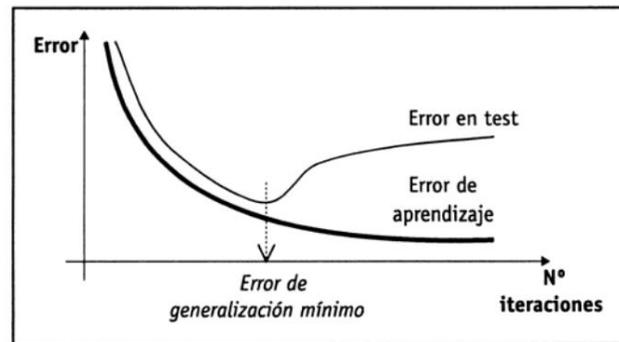


Figura 2.5. Análisis del fenómeno de "sobre-aprendizaje" de la red [Flores López y Fernández 2008].

2.3.7 División de los datos

Dividir los datos experimentales sirve para prevenir el sobre-entrenamiento (también llamado sobreajuste o sobre-aprendizaje). Se debe decidir qué porcentaje de la base de datos se deben tomar para validación, en vez de decidir cuantos datos para entrenamiento se deben dejar. A este método de dejar conjuntos de datos para entrenamiento, validación y prueba se le conoce como validación cruzada (cross-validation). El conjunto de datos para validación sirve para *parar* el entrenamiento y evitar el sobre-aprendizaje, esto sucede cuando el error en los datos de validación empieza a subir sin importar que el error en los datos de entrenamiento siga bajando, con lo cual la red puede generalizar mejor con nuevos datos [Beale *et al.* 2017].

El conjunto de datos para prueba sirve para comprobar que la división de los datos estuvo bien realizada. Cuando esta división es deficiente, el error mínimo que se tiene para los datos de validación y para los datos de prueba se alcanzan con una gran diferencia de iteraciones, es decir, el error mínimo para los datos de prueba se lograra mucho después de que el error mínimo para validación sea alcanzado, lo que indica que la red no generaliza bien (aunque no haya sobre-entrenamiento). Estos datos no se deben utilizar por ningún motivo para entrenar la red, ya que estos datos nos darán una idea de cómo la red se comportará en el futuro con datos nuevos [Demuth *et al.* 2014].

Entonces cual debería ser la división óptima de la base de datos. Desde que los conjuntos de datos para validación y para prueba comparten características similares (comprobar que la red generaliza de manera adecuada), Amari y comañia sugieren las siguientes divisiones [Amari *et al.* 1997]:

- Cuando se tienen pocos datos experimentales, alrededor de 200 o menos, y el número de coeficientes (pesos y bias) son altos ($m = 108$), dividir dichos datos en 20% para validación incrementa el error de generalización, comparado con dividirla al 33% y 42%.
- Cuando hay alrededor de 286 datos experimentales, los errores de generalización son similares para una división del 20 y 33%. En el intervalo de entre 250 datos y 550 datos, dividir los datos al 20% tiene un error ligeramente mayor comparado con dividirlos al 33 o 42%.
- A partir de 550 datos hasta 20000 datos, elegir una división del 20% mejora ligeramente el error de generalización de la red neuronal [Amari *et al.* 1997].

De aquí que se decide dividir el 25% de los datos para validación (y por ende un 25% extra para prueba) como una manera más generalizada de dividir los datos, obviando un poco el número de datos experimentales, siempre que sean suficientes, y que el número de coeficientes no sobrepase a los datos.

2.3.8 Método (algoritmo) de aprendizaje: Levenberg-Marquardt (LM)

Actualmente, las investigaciones en ANN se centran en las mejoras relacionadas con el aprendizaje del modelo. Para garantizar que el ANN aprende de manera eficiente, ha sido necesario el desarrollo de algoritmos de entrenamiento que proporcionen a la red la capacidad de discriminar y generalizar los datos [Turkson *et al* 2008]. Algunos de los algoritmos de entrenamiento u optimización que se pueden considerar son Gradiente conjugado, cuasi-Newton y Levenberg-Marquardt [Sözen *et al.* 2007]. La arquitectura ANN más utilizada es el perceptrón multicapa entrenado con el algoritmo de retropropagación en una red de avance (FeedForward) a través del método de entrenamiento Levenberg-Marquardt (LM) [Jozanikohan *et al.* 2015].

Para entender mejor cómo funciona el método LM, se explicará desde el método de gradiente descendiente. El método de gradiente descendiente aplica el gradiente más grande en cada paso lo que puede resultar en una convergencia no muy rápida. El método de gradiente conjugado usa la dirección negativa o positiva para poder converger más rápido que el gradiente descendiente normal. Además, la mayoría de estos métodos necesita un valor en su tasa de aprendizaje (tamaño de paso) para poder actualizar el valor del peso de la red, este tamaño de paso es ajustado en cada iteración de acuerdo a una búsqueda realizada en la dirección del gradiente conjugado para minimizar la función de rendimiento [Demuth y Beale 2005].

Los métodos newtonianos son métodos alternativos a los de los gradientes conjugados que frecuentemente convergen más rápidos que los de gradiente conjugados. Utiliza una matriz Hessiana (matriz cuadrada de segundas derivadas parciales) como función de rendimiento de los pesos y bias. Pero esta matriz Hessiana es compleja de calcular y cuesta muchos recursos computacionales para redes FeedForward. Los métodos Cuasi-newtonianos

utilizan una función (creada para dicho método) en vez de calcular la matriz Hessiana pero que se aproxima a esta [Demuth y Beale 2005].

El algoritmo LM es parecido a los métodos quasi-Newton, y está diseñado para aprovechar la velocidad de las segundas derivadas parciales sin necesidad de calcular la matriz Hessiana. Este diseño se aprovecha en las redes Backpropagation, debido a que, la función de rendimiento en dichas redes está basada en los errores medios cuadráticos (mean square errors), como se muestra:

$$x_{k+1} = x_k - \alpha_k g_k \quad (2.2)$$

Donde x_k es el vector actual de pesos y bias, g_k es el actual gradiente descendiente, α_k es la tasa de aprendizaje y el subíndice k es el valor de cada iteración [Demuth y Beale 2005].

Los pasos (del gradiente) para los métodos cuasi-newtonianos son:

$$x_{k+1} = x_k - A_k^{-1} g_k \quad (2.3)$$

Donde A_k es la matriz Hessiana.

En el método LM, puesto que se tiene una forma cuadrática en la función de rendimiento, la matriz de Hessiana puede aproximarse de la siguiente forma:

$$A = H = J^T J \quad (2.4)$$

Y el gradiente se calcula como

$$g = J^T e \quad (2.5)$$

Donde J es la matriz Jacobiana que incluye las primeras derivadas de los errores de la red con respecto a los pesos y bias, y e es un vector de errores de la red ($e = t - a$, aquí t es el target o valor del dato experimental y a es el valor simulado obtenido de la red). La matriz Jacobiana es más fácil de calcular y requiere menos costo computacional que la matriz Hessiana [Hagan y

Menhaj 1994]. De tal modo que, la función de rendimiento usando el método LM queda:

$$x_{k+1} = x_k - [J^T J + \mu I]^{-1} J^T e \quad (2.6)$$

El objetivo del método LM es dirigirse lo más rápido posible hacia el método Newton, esto sucede cuando el *escalar* μ es cero (o muy próximo a cero). En caso contrario, cuando μ es grande, se transforma en un método de gradiente descendiente con un tamaño de paso que depende de μ , que por lo general es pequeño. El valor de μ *disminuye* con cada paso exitoso, es decir cuando el error disminuye, o *umentará* solo si el error es mayor que en la iteración anterior. El método de Newton es más rápido y preciso cerca de un error mínimo, de aquí la importancia de que μ sea muy próxima a cero [Hagan *et al.* 2014]. I es la matriz identidad. Cuando se tiene una base de datos muy extensa, se requerirá almacenar un gran número matrices Jacobianas, lo que conlleva a un gran uso de memoria (computacional), lo que supone una desventaja [Mateo 2012].

2.3.9 Análisis estadístico

Para validar la precisión y la adaptabilidad del modelo propuesto, se realizan pruebas estadísticas para analizar los datos experimentales y los datos simulados obtenidos. Los parámetros de prueba que se usan comúnmente son la Raíz del Error Cuadrático Medio (Root Mean Square Error, *RMSE*), el Error Porcentual Medio Absoluto (Mean Absolute Percent Error, *MAPE*) y el coeficiente de determinación (R^2) [Verma 2005]. El modelo óptimo es el que tiene el R^2 más alto, y el *MAPE* y *RMSE* más bajos.

$$R^2 = 1 - \frac{\sum_{i=1}^n (x_{exp(i)} - x_{sim(i)})^2}{\sum_{i=1}^n (x_{exp(i)} - \bar{x}_{exp})^2} \quad (2.7)$$

$$MAPE = \frac{\sum_{i=1}^n \left| \frac{x_{exp(i)} - x_{sim(i)}}{x_{exp(i)}} \right|}{n} \times 100 (\%) \quad (2.8)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_{sim(i)} - x_{exp(i)})^2}{n}} \quad (2.9)$$

2.3.10 Pruebas de significancia

Para corroborar el ajuste del modelo, las pruebas de significación se aplican a la relación de bases de datos experimentales y simuladas. La prueba F o la prueba de Fisher y la prueba t de Student son las pruebas de significación más utilizadas para analizar la hipótesis nula H_0 o la hipótesis alternativa H_1 . El H_0 establece: *si las dos bases de datos tienen la misma varianza o varianza similar (prueba F), o si la media de dos bases de datos es igual o proviene de la misma población (prueba t)*. Del mismo modo, el H_1 determina que la varianza o la media de las bases de datos es diferente. Para calcular la prueba F y la prueba t, se utilizan las siguientes ecuaciones [Verma 2005]:

$$S = \sqrt{\frac{\sum_{i=1}^{n_{exp}} (x_{exp(i)} - \bar{x}_{exp(i)})^2 + \sum_{i=1}^{n_{sim}} (x_{sim(i)} - \bar{x}_{sim(i)})^2}{(n_{exp} + n_{sim} - 2)}} \quad (2.10)$$

$$F = \frac{S_{exp}^2}{S_{sim}^2} \quad (2.11)$$

$$t = \frac{|\bar{x}_{exp} - \bar{x}_{sim}|}{S \left(\sqrt{\left(\frac{1}{n_{exp}} + \frac{1}{n_{sim}} \right)} \right)} \quad (2.12)$$

Donde S es la desviación estándar combinada, S^2 es la varianza, n el número de muestras y \bar{x} es la media de la muestra. Para ambos casos, el valor crítico calculado debe ser menor al valor crítico que se da en tablas para aceptar la hipótesis nula (H_0), de lo contrario se debe aceptar la hipótesis alterna (H_1). Las tablas tienen su nivel de significancia (o confianza). Se debe seleccionar, de preferencia, un nivel de significancia de 0.01 (99% de confianza), lo indica que tenemos un error aceptable del 1%. Usualmente se seleccionan niveles de confianza $\geq 95\%$. Para la prueba F, la F calculada siempre debe ser ≥ 1 ,

en caso de que sea menor, se debe de calcular con los términos invertidos. El numerador siempre debe ser más grande o igual al denominador. Cuando se pasa la prueba F, se procede a realizar la prueba t de Student. De lo contrario ya no se aplica y el modelo debe ser rechazado [Verma 2005].

Además de estas pruebas, el **p-valor** nos dice cuál es el nivel de significancia más pequeño para aceptar la hipótesis alternativa (H_1), esta se calculó mediante un software estadístico (para más información sobre como calcular el p-valor manualmente <https://www.youtube.com/watch?v=pUPGoIV2AE0>). En 2016, la Asociación de Estadística Americana determinó que, el p-valor por sí solo no es una medida apropiada de la evidencia de un modelo hipótesis, por lo que se requieren análisis más apropiados y factibles. Por esta razón se requieren muchas pruebas estadísticas. Si el p-valor es menor que el nivel de significancia (α), entonces se acepta la hipótesis alterna (H_1) y el modelo debe ser rechazado [Wasserstein y Lazar 2016]

La prueba de Wilcoxon-Mann-Whitney (WMW) es la prueba no paramétrica equivalente a la prueba t Student. Sirve para comparar las medias de dos muestras independientes, que provienen de la misma población, con lo que se comprueba si dos medias muestrales son iguales o no. También se le conoce con los nombres de la prueba U de Mann-Whitney o prueba de suma de rangos Wilcoxon. Esta prueba se usa cuando los datos son ordinales o cuando no se cumplen los supuestos de la prueba t [Bergmann *et al.* 2000]. En este trabajo no se utilizó puesto que se calculó la prueba-t de Student.

2.3.11 Prueba de linealidad

Para validar la relación entre los datos simulados con el modelo ANN y los datos experimentales, se realizan pruebas de linealidad, para reducir la incertidumbre y ver si los datos son “adecuados para el propósito” (*fitness-for-purpose*) en lo que se utilizan [Verma 2005]. Se sabe de antemano que el análisis de regresión lineal se utiliza para obtener una relación entre dos variables (una dependiente y la otra independiente) y así estimar los

coeficientes de la pendiente y el intercepto, que determina la ecuación de la línea. Estadísticamente, la pendiente debe ser igual a 1 y la intersección igual a 0.

La prueba de linealidad pendiente-intersección, también llamada desviación estándar de la regresión lineal o error estándar, establece los límites de confianza de la pendiente y el intercepto. Estos límites o intervalos deberían ser idealmente uno y cero para la pendiente y la intersección, respectivamente [Rodríguez *et al.* 2013]. Mediante las siguientes ecuaciones, se estiman los errores en los coeficientes de regresión lineal:

$$S_b = \sqrt{\frac{\sum_{i=1}^n (x_{sim(i)} - \hat{x}_{sim(i)})^2}{(n-2) \sum_{i=1}^n (x_{exp(i)} - \bar{x}_{exp(i)})^2}} \quad (2.13)$$

y

$$S_a = \sqrt{\frac{\sum_{i=1}^n (x_{sim(i)} - \hat{x}_{sim(i)})^2 \sum_{i=1}^n x_{exp(i)}^2}{n(n-2) \sum_{i=1}^n (x_{exp(i)} - \bar{x}_{exp(i)})^2}} \quad (2.14)$$

Donde $x_{sim(i)}$ es el valor simulado y $\hat{x}_{sim(i)}$ el valor de x_{sim} calculado por regresión lineal para el mismo valor en $x_{sim(i)}$ ($\hat{x}_{sim} = a + b\bar{x}_{exp(i)}$) [Rocabrundo-Valdés *et al.* 2019]. La expresión $x_{sim(i)} - \hat{x}_{sim(i)}$ se llama residual, que proviene de los datos originales con respecto a la línea de regresión.

En este punto, es importante recordar que la pendiente b viene dada por:

$$b = \frac{\sum_{i=1}^n \{(x_{exp(i)} - \bar{x}_{exp(i)})(x_{sim(i)} - \bar{x}_{sim(i)})\}}{\sum_{i=1}^n (x_{exp(i)} - \bar{x}_{exp(i)})^2} \quad (2.15)$$

Y el intercepto a para:

$$a = \bar{x}_{sim} - b\bar{x}_{exp} \quad (2.16)$$

Donde \bar{x}_{exp} y \bar{x}_{sim} son la media de los datos experimentales y simulados, respectivamente.

Por lo tanto, los intervalos de confianza o límites de la pendiente y el intercepto se definirán como:

$$\text{para pendiente } \{b - (t_{(n-2)}s_b)\} < \mu_b < \{b + (t_{(n-2)}s_b)\} \quad (2.17)$$

$$\text{para intercepto } \{a - (t_{(n-2)}s_a)\} < \mu_a < \{a + (t_{(n-2)}s_a)\} \quad (2.18)$$

Donde $t_{(n-2)}$ es el valor crítico t de Student que proviene de tablas estadísticas, en un cierto nivel de confianza o significación de 95% o 99% (recomendado) con $n - 2$ grados de libertad. Para n datos bivariados, los grados de libertad se reducirán en dos debido a que se requieren al menos dos datos o puntos para dibujar una línea recta de regresión [Verma 2005]. Sin embargo, es necesario un ajuste cuando los datos exceden las mil muestras; este ajuste es principalmente para niveles de 98% y 99% (para ambos lados). Las expresiones son las siguientes [Verma y Cruz-Huicochea 2013]:

Menos o igual que 1000 datos y un nivel de confianza del 98%:

$$\begin{aligned} (t_{98ts}^{cv})_{int} = & (4.94002103 \pm 0.00012294) - (4.5022554 \pm 0.0018112) \\ & * (\ln(\ln(n))) + (2.785764 \pm 0.008079) * (\ln(\ln(n)))^2 \\ & - (0.485943 \pm 0.016415) * (\ln(\ln(n)))^3 \\ & - (0.3051890 \pm 0.017282) * (\ln(\ln(n)))^4 \\ & + (0.216257 \pm 0.009455) * (\ln(\ln(n)))^5 \\ & - (0.0506048 \pm 0.0022587) * (\ln(\ln(n)))^6 \\ & + (0.00132479 \pm 0.000062583) * (\ln(\ln(n)))^8 \end{aligned} \quad (2.19)$$

Para más de 1000 datos y un nivel de confianza del 98%:

$$\begin{aligned}
 (t_{98ts}^{cv})_{ext} = & (4.94237021 \pm 0.00014897) - (4.5412456 \pm 0.0007948) \\
 & * (\ln(\ln(n))) + (2.9690047 \pm 0.0013459) * (\ln(\ln(n)))^2 \\
 & - (0.8658555 \pm 0.0008820) * (\ln(\ln(n)))^3 + (0.09497881 \\
 & \pm 0.00019659) * (\ln(\ln(n)))^4
 \end{aligned}
 \tag{2.20}$$

Para menos o igual a 1000 datos y un nivel de confianza del 99%:

$$\begin{aligned}
 (t_{99ts}^{cv})_{int} = & (6.48557253 \pm 0.00017748) - (7.3462223 \pm 0.0026146) \\
 & * (\ln(\ln(n))) + (5.418770 \pm 0.011663) * (\ln(\ln(n)))^2 - 1.721741 \\
 & \pm 0.023697) * (\ln(\ln(n)))^3 - (0.0930165 \pm 0.024949) \\
 & * (\ln(\ln(n)))^4 + (0.274850 \pm 0.013649) * (\ln(\ln(n)))^5 \\
 & - (0.0779639 \pm 0.0032607) * (\ln(\ln(n)))^6 + (0.00221624 \\
 & \pm 0.00009035) * (\ln(\ln(n)))^8
 \end{aligned}
 \tag{2.21}$$

Y para más de 1000 datos y un nivel de confianza del 99%:

$$\begin{aligned}
 (t_{99ts}^{cv})_{ext} = & (2.94230900 \pm 0.00003147) - (1.1213450 \pm 0.00018085) \\
 & * (\ln(\ln(\ln(n)))) + (0.78201429 \pm 0.00031542) \\
 & * (\ln(\ln(\ln(n))))^2 + (0.3639632 \pm 0.0009774) * (\ln(\ln(\ln(n))))^3 \\
 & - (0.20510287 \pm 0.00048796) * (\ln(\ln(\ln(n))))^4 \\
 & - (0.1945693 \pm 0.012343) * (\ln(\ln(\ln(n))))^5 \\
 & - (0.04167330 \pm 0.00042639) * (\ln(\ln(\ln(n))))^6
 \end{aligned}
 \tag{2.22}$$

2.3.12 Análisis de sensibilidad de variables

El análisis de sensibilidad depende exclusivamente del peso que se le da a cada conexión durante el entrenamiento, es decir, los valores de pueden variar de modelo a modelo con similares R^2 , $RMSE$, y $MAPE$ e igual arquitectura. Esto se debe a que durante el entrenamiento los valores se ajustan para poder representar la variable de salida experimental. Se usó la ecuación de Garson

[Parrales *et al.* 2019] para determinar la importancia de las variables de entrada dentro de la red. Esta ecuación se basa en la partición de las conexiones de peso.

$$I_P = \frac{\sum_{j=1}^J \left(\left(\frac{|W_{(P,j)}|}{\sum_{r=1}^R |W_{(r,j)}|} \right) \times |W_{(j,K)}| \right)}{\sum_{r=1}^R \left\{ \sum_{j=1}^J \left(\left(\frac{|W_{(r,j)}|}{\sum_{r=1}^R |W_{(r,j)}|} \right) \times |W_{(j,K)}| \right) \right\}} \quad (2.23)$$

Donde I_P es la importancia relativa de la variable de entrada P sobre la variable de salida, J es el número de neuronas en la capa oculta, R es el número de neuronas de entrada, K es el número de neuronas en la capa de salida y W es el peso correspondiente a la posición de las neuronas.

2.4 La inversa de las Redes Neuronales Artificiales

Aunque esta terminología se haya usada desde hace casi 30 años, cuando se marcaban por primera vez las diferencias entre una arquitectura de red Backpropagation y una recurrente [Narendra y Parthasarathy 1990] [Shukla y Paul 1996], fue usada para identificar y controlar variables discretas dentro de la misma estructura que se generan en las ANN, y es conocido como “control inverso de redes neuronales” (Neural Network Inverse Control) [Qibing *et al.* 1998] [Du y Wu 2006].

Desde 2009, la inversa de las Redes Neuronales Artificiales (Artificial Neuronal Network inverse, ANNi), se ha establecido como una metodología de optimización de variables partiendo de los pesos y bias calculados de una ANN ya existente y validada [Hernández 2009] [Hernández *et al.* 2009] [Córtes *et al.* 2009]. Su objetivo principal es mejorar la variable de salida/target mediante una función objetivo de la variable de entrada elegida para llevar a cabo la

optimización. Esto se da gracias a que la salida y los parámetros de entrada son bien conocidos [Hernández 2009].

Si el modelo ANN tiene una sola neurona en la capa oculta, es posible despejar la variable de entrada para encontrar valores óptimos para una salida conocida o deseada. Pero resulta imposible realizar el despeje cuando hay más de una neurona oculta, como en la mayoría de los casos, por lo que resulta necesario implementar un método/algorithm numérico de optimización que busque y encuentre valores óptimos de entrada para resolver el proceso [Hernández 2009].

La Figura 2.6 muestra cómo se aplica la metodología ANNi, estableciendo una función objetivo que pueda ser minimizada lo más cercana a cero para un conjunto de condiciones de entrada conocidas y una salida también conocida que sea igual o muy próximo a un valor deseado.

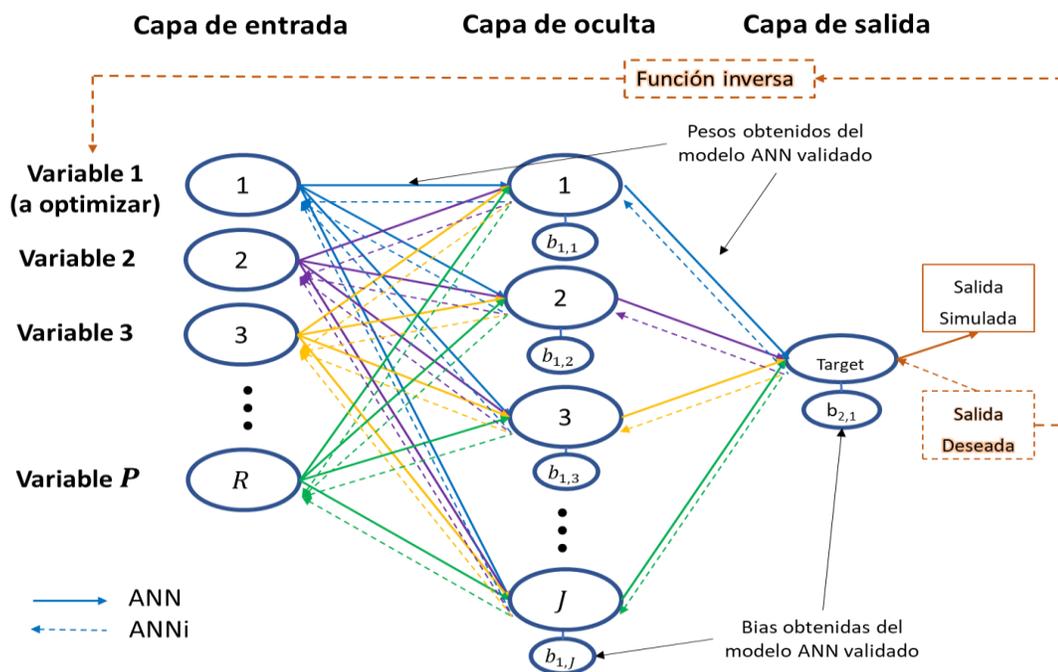


Figura 2.6. Esquema de una Red Neuronal Artificial Inversa

Como se mencionó antes, primero es necesario establecer la ecuación general del modelo ANN validado. Dependiendo de las funciones utilizadas en las capas oculta y de salida, la ecuación general tendrá diferentes formas, en este caso se presentan solo dos:

- Con una función de transferencia **tangencial-sigmoidal** en la capa oculta y una lineal a la salida,

$$Output = \sum_{j=1}^J \left\{ W_{o(K,j)} \left[\frac{2}{1+e^{-2 \sum_{r=1}^R (W_{i(j,r)} \cdot P_{(r)} + b_{(1,j)})}} - 1 \right] \right\} + b_{(2,K)} \quad (2.24.a)$$

$$Output = b_{(2,K)} + \sum_{j=1}^J \left[\frac{2 \cdot W_{o(K,j)}}{1+e^{-2 \sum_{r=1}^R (W_{i(j,r)} \cdot P_{(r)} + b_{(1,j)})}} \right] - \sum_{j=1}^J W_{o(K,j)} \quad (2.24.b)$$

- Con una función de transferencia **logarítmica-sigmoidal** en la capa oculta y una lineal a la salida.

$$Output = \sum_{j=1}^J \left\{ W_{o(K,j)} \left[\frac{1}{1+e^{-\sum_{r=1}^R (W_{i(j,r)} \cdot P_{(r)} + b_{(1,j)})}} \right] \right\} + b_{(2,K)} \quad (2.25.a)$$

$$Output = b_{(2,K)} + \sum_{j=1}^J \left[\frac{W_{o(K,j)}}{1+e^{-\sum_{r=1}^R (W_{i(j,r)} \cdot P_{(r)} + b_{(1,j)})}} \right] \quad (2.25.b)$$

Donde W_o son los pesos entre la capa oculta y la capa de salida, W_i son los pesos entre las capas de entrada-oculta. b_1 son las bias en la capa oculta y b_2 es la bia en la capa de salida. Los subíndices R , J y K refieren al número de neuronas en las capas de entrada, oculta y de salida, respectivamente.

Como el modelo ANNi depende de los pesos y bias del modelo ANN, es importante que este tenga alta precisión, de modo tal que, la ecuación 2.26 muestra la variable de entrada despejada como una función a optimizar, para el caso de una función de transferencia tangencial-sigmoidal y lineal en la capa oculta y de salida, respectivamente.

$$f(x) = b_{(2,K)} + \sum_{j=1}^J \left[\frac{2 \cdot W_{o(K,j)}}{1 + e^{-2[(W_{i(j,R \in x)} \cdot x) + \sum_{r \neq R \in x}^R (W_{i(j,r)} \cdot P_{(r)} + b_{(1,j)})]}} \right] - \sum_{j=1}^J W_{o(K,j)} \quad (2.26)$$

– Output

Una vez planteada la ecuación ANNi en función del valor de entrada a optimizar (x) y un valor de salida (*Output*) que se pretende conseguir, se procede a resolver dicha ecuación, donde la función objetivo ($f(x)$) debe minimizarse lo más próximo a cero o un valor que sea satisfactorio [Cruz-Jacobo 2018].

Como el método ANNi fue desarrollado para optimizar las condiciones de operación en tiempo real, se debe resolver la ecuación ANNi lo más rápido posible. Para esto es necesario un método de optimización que logre resolver la ecuación rápidamente para mejorar el resultado del proceso en línea. Cuando se desarrolló la estrategia ANNi, se propuso utilizar el método Simplex (nelder-Mead) para poder resolver la función objetivo y por ende, la ecuación [Hernández 2009] [Hernández *et al.* 2009] [Córtes *et al.* 2009] [Hamzaoui *et al.* 2011]. Sin embargo, las investigaciones hechas por Hamzaoui y sus colaboradores en 2012, demostraron que es mejor utilizar algoritmos genéticos debido a que puede dar soluciones globales, tiene gran velocidad de convergencia, además de alta precisión y poder estimar más de un parámetro a la vez [Hamzaoui *et al.* 2012].

2.4.1 Algoritmos Genéticos

Los algoritmos genéticos (Genetic Algorithms, GA) son métodos de optimización que se basan en el proceso de selección natural (supervivencia del más apto) que conduce a la evolución biológica [Bodenhofer 2003]. Tiene diferencias significativas respecto a otros métodos de búsqueda y optimización, que se resumen principalmente en las siguientes 4 [Fung-Bao 2008] [Chipperfield *et al.*]:

- Los GA buscan una población de puntos en cada iteración, de forma paralela, para encontrar el más se aproxime a una solución óptima; en vez de solo un punto.

- No requieren de información adicional, solo basta con la función objetivo (función fitness) y sus correspondientes parámetros para determinar la dirección de la búsqueda.
- Usan reglas de transición probabilísticas durante todo el proceso, que pueden incluir la *selección*, el *cruce* y la *mutación*; en vez de reglas deterministas.
- Como se debe codificar la información a código binario, no se necesita especificar nada más a fin de encontrar una solución final.

Al igual que las ANN, los GA tienen su base biológica, que a continuación se describe su analogía:

Tabla 2.2. Expresiones genéticas y su equivalencia en Algoritmos Genéticos
[Ponce Cruz 2010]

Evolución natural	Algoritmo genético
Genotipo	Carácter (o cadena de caracteres) codificada de toda la población
Fenotipo	Dato sin codificar de la población real
Cromosoma	Una sola cadena de carácter codificada
Gen	Posición dentro de la cadena
Alelo	Valor del gen
Función fitness (aptitud)	Valor de la función objetivo

El genotipo es la información genética contenida en el ADN de un organismo o individuo, mientras que el fenotipo es la representación externa del genotipo, es decir, todas las características que se pueden observar a simple vista. El cromosoma es la molécula única del ADN, el gen especifica los caracteres hereditarios de un individuo y se encuentran en un orden específico dentro de un cromosoma; mientras que un alelo es el valor que toma el gen (0 o 1 para código binario). La función fitness sirve para encontrar los cromosomas óptimos, estos se combinan y producen una nueva generación que sea mejor que la anterior. [Ponce Cruz 2010].

Como se mencionó anteriormente, la codificación del GA se hace comúnmente en código binario [Holland 1975], lo que conlleva a definir un algoritmo con el cual se pueda realizar una búsqueda de parámetros a fin de resolver un problema. Los pasos de dicho algoritmo son [Morales *et al.* 2015]:

- Se inicializa una población de cromosomas (individuos) aleatoriamente.
- Evaluar cada cromosoma con la función fitness, para categorizarlos de mejor a peor.
- Se seleccionan los mejores cromosomas para reproducirse (o se elige al mejor mediante *elitismo*).
- Durante la reproducción de un par de cromosomas, se utiliza el cruce para intercambiar genes y crear una nueva generación de cromosomas.
- El selecciona un cromosoma para mutación para introducirlo en la nueva generación, puede emplearse conjuntamente o no con el cruce [MathWorks 2013].

Los pasos del 2 al 5 se repiten hasta que la función fitness converja a un valor predefinido, o cuando se alcanza el número de generaciones (iteraciones), o cuando no hay mejoras en la población después de un cierto número de iteraciones. En la Figura 2.7 se resumen como se representa un AG.

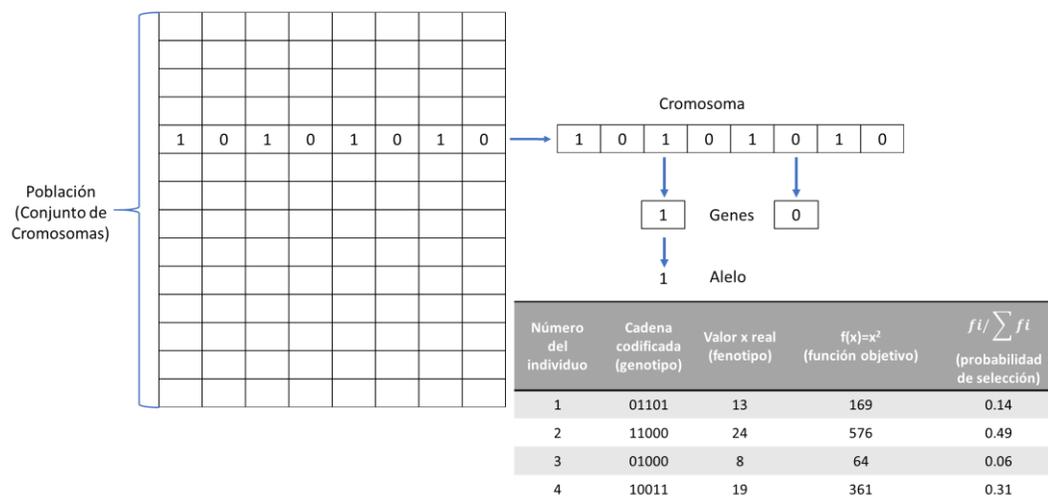


Figura 2.7. Representación esquemática de los términos usados en Algoritmos Genéticos.

2.4.2 Población, función fitness/objetivo y operadores genéticos

La **población** inicial puede suponer un problema de que tan grande o pequeña debe de ser, una población muy grande hará muy lento el computo del GA, mientras que una pequeña corre el riesgo de no cubrir el espacio de búsqueda adecuadamente. Ante esto se presentan dos escenarios, el primero es encontrar el tamaño de población ideal mediante prueba y error; el segundo plantea que el tamaño de población debe ser del mismo tamaño o hasta el doble de los bits empleados en el código binario, es decir, si el tamaño de la cadena es igual l , el tamaño de la población comprenderá desde l hasta $2l$ [Alander 1992]. Lo mejor es tener una población inicial que sea generada aleatoriamente, porque si se utiliza técnica heurística o de optimización local se corre el riesgo de tener óptimos locales [Tutorials Point 2016].

Técnicamente, la **función fitness** transforma el valor de la función objetivo, que mide cómo se desempeñan los individuos en el dominio del problema, en una medida *fitness* relativa [Chipperfield *et al*]. La función fitness y la función objetivo es la misma en la mayoría de los casos (algunos programadores y matemáticos prefieren llamarla función costo) [Haupt y Haupt 2004]. El objetivo principal de dicha función es ser una medida cuantitativa de que tan buena es una solución de las muchas que hay en la población. Debe ser lo suficientemente rápida de calcular porque se repite muchas veces a fin de encontrar los individuos más aptos y así tener mayor probabilidad de reproducción. Una función objetivo difícil de calcular puede ser lenta lo que conlleva gran costo computacional. Para casos realmente complejos es preferible tener múltiples funciones objetivos o una evaluación aproximada de la función objetivo [Tutorials Point 2016]. En la mayoría de los casos, la función costo es una función matemática (aunque también puede ser un experimento o un juego), y dependerá del problema si se requiere maximizar o minimizar la función [Haupt y Haupt 2004].

Los operadores genéticos son tres: selección, cruce (o cruzamiento) y mutación. Juntos se involucran en la reproducción de la nueva generación a partir de dos padres seleccionados basados en la función fitness.

Selección. Es el proceso por el cual se seleccionan los padres y recombinan sus genes para crear descendientes de la próxima generación. A fin de tener diversidad en todo el espacio de búsqueda, los algoritmos de selección presentan variaciones respecto al número de veces que se tomarán los mejores y peores individuos (cromosomas) para su reproducción. Si solo los mejores individuos predominan en la población, al cabo de pocas generaciones el modelo convergerá, lo que se conoce como **convergencia prematura**, lo cual es indeseable debido a que se tiene óptimos locales. Al admitir que individuos menos aptos sean elegidos para cruzarse, el espacio de búsqueda de soluciones se amplía, permitiendo encontrar un óptimo global (o uno más cercano) [Gestal *et al.* 2010]. Algunos de los métodos de selección más usados son:

- **Selección por ruleta.** Para este tipo de selección, se considera una rueda circular que se divide en partes, de acuerdo al número de individuos (cromosomas) en la población. Cada individuo tiene una parte proporcional de acuerdo a su valor *fitness*, de modo tal que, los mejores individuos tendrán un segmento más grande en la ruleta (ver Figura 2.8). Habitualmente, los segmentos más grandes se ubican al principio de la rueda (parte superior). Para seleccionar un individuo, se genera un número aleatorio entre los intervalos de la población (generalmente entre 0 y 1), lo que simula el “giro de la ruleta”, y devolverá al individuo que este ubicado en ese lugar, se repite el mismo proceso para elegir al segundo padre. Este método a pesar de ser sencillo, se vuelve ineficiente a medida que el tamaño de la población crece. Un método similar es el *muestreo universal estocástico*, solo que se utilizan múltiples marcadores igualmente espaciados para

seleccionar a todos los padres en un único “giro” [Tutorials Point 2016] [Gestal *et al.* 2010].

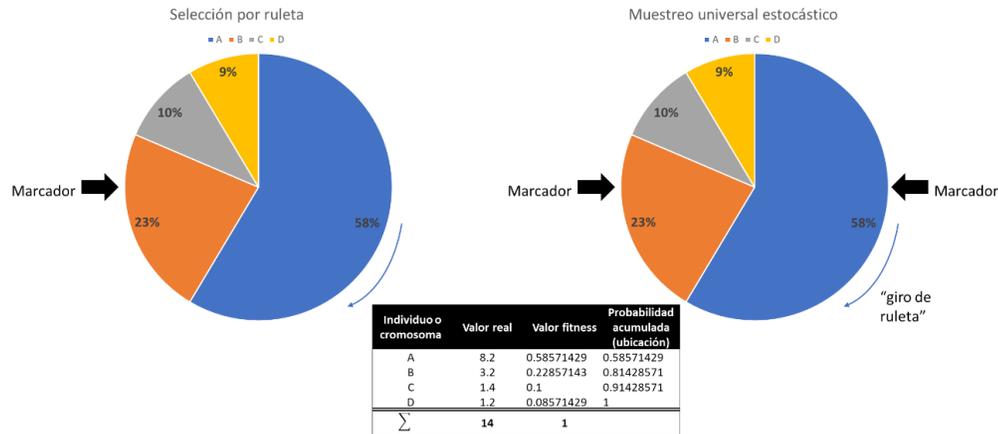


Figura 2.8. Representación de la selección por ruleta y de la selección por muestreo universal estocástico.

- Selección por torneo.** Este método consiste en seleccionar un pequeño subconjunto de cromosomas (generalmente dos o tres) para elegir al más apto o escoger a alguno de ellos, basado en probabilidad, para ser padre. Si se decide hacerlo de forma probabilística, se debe generar un número al azar (de 0 a 1), y si es mayor a un parámetro fijado (de 0.5 a 1) para todo el proceso, se escoge al individuo más apto, en caso contrario se elige al peor. Este proceso se repite hasta tener todos los padres que generarán la descendencia de la próxima generación. Si el subconjunto tiene muchos cromosomas, los peores individuos casi no tienen oportunidad de reproducción, y viceversa [Gestal *et al.* 2010]. Este tipo de selección funciona para grandes poblaciones porque el ordenar los individuos de mayor a menor consume mucho tiempo de cómputo para poblaciones muy grandes [Haupt y Haupt 2004]. Puede funcionar incluso con valores negativos [Tutorials Point 2016]. Una variante de la selección por torneo es el *elitismo global*, donde todos los individuos de la población participan el

torneo, en la cual, si la población es pequeña, los peores individuos tienen más oportunidad de reproducción [Gestal *et al.* 2010].

Función escalada. La función escalada de fitness sirve para convertir los valores calculados con la función fitness, a un valor dentro de un intervalo que es más sencillo de manejar, de acuerdo a su valor fitness. El intervalo de los valores sin escalar afecta al rendimiento del algoritmo genético, intervalos muy amplios conlleva a una convergencia prematura, mientras que intervalos con variaciones muy pequeñas hace que el muy lento el proceso [Chipperfield *et al.*]. Dentro del software Matlab® existen diferentes funciones escaladas, el más usado es Rango (Rank) debido a que ofrece más diversidad para la búsqueda en el espacio de soluciones. La Escalada por Rango, ordena de mayor a menor a los individuos dentro de la población proporcional a su valor fitness [MathWorks 2019].

Cruce. Una vez seleccionados los padres, estos combinaran sus genes para crear a nuevos individuos en la nueva generación. Los métodos de ajuste más utilizados son:

- **Cruce elitista.** Se escoge al mejor o los mejores individuos de una población sin sufrir ninguna modificación. Este tipo de cruce tiene bajo porcentaje de reproducción para garantizar la diversidad [MathWorks 2013] [Gestal *et al.* 2010].
- **Cruce en 1 punto:** se corta a los padres en un punto aleatorio para dividirlos en 2 partes sus cromosomas. Un hijo heredará la primera mitad del primer padre y la segunda mitad del segundo padre, y el segundo heredará las demás partes, como se muestra en la Figura 2.9 [Gestal *et al.* 2010] [Ponce Cruz 2010].

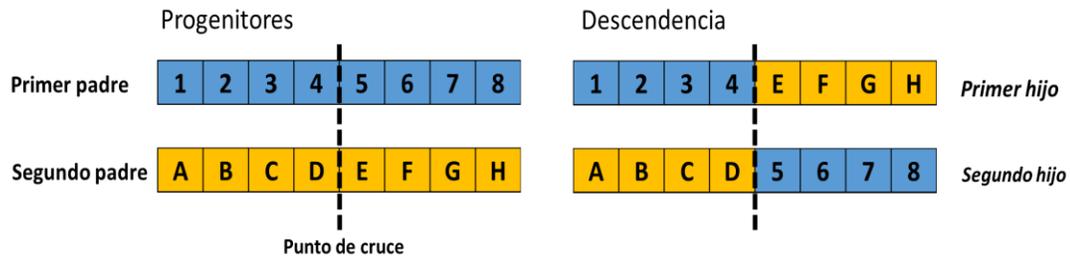


Figura 2.9. Cruzamiento en un punto de dos padres y su respectiva descendencia.

- **Cruce en 2 puntos.** Parecido al cruce en 1 punto, pero en vez de cortar en un solo punto se corta en dos para dividir a los padres en tres partes, como se muestra en la Figura 2.10. Para generar a los nuevos individuos, se elegirá la parte central de un padre y las partes laterales del otro padre. De manera general, se pueden añadir más puntos de cruce, pero muchos puntos de cruce pueden hacer que pierdan su valor fitness, lo que ralentiza el rendimiento del GA. Una ventaja es que se amplía el espacio de búsqueda [Gestal *et al.* 2010].

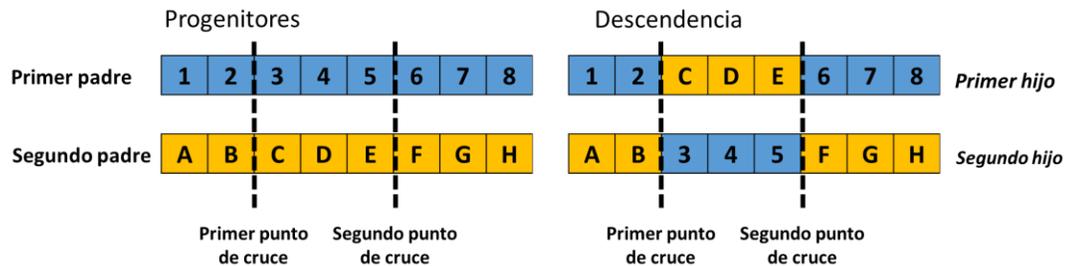


Figura 2.10. Cruzamiento en dos puntos de dos padres y su respectiva descendencia.

- **Cruce uniforme.** Se decide aleatoriamente cada gen de uno u otro padre para el nuevo individuo. Para realizar esto, se genera una máscara o plantilla de cruce, la cual actúa como un interruptor para decidir de que padre recibirá un gen el nuevo individuo como se muestra en la Figura 2.11. El segundo individuo heredará la parte contraria de los genes del primer individuo. La máscara puede ser fija o totalmente aleatoria a lo largo de todo el proceso. Si se decide por la forma aleatoria, se debe generar una máscara por cada cruce [Gestal *et al.*

2010] [Ponce Cruz 2010]. Matlab® llama a este tipo de cruce *función de cruce dispersa* (scattered) [MathWorks 2019].

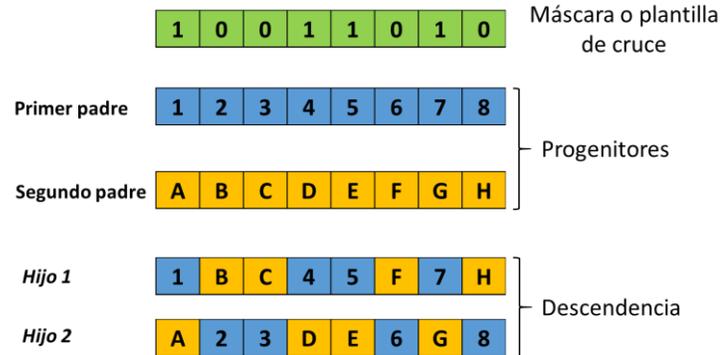


Figura 2.11. Cruzamiento uniforme con la creación de una máscara aleatoria para generar la descendencia de dos padres.

Mutación. Es provocar el cambio de valor en un gen (usualmente en un solo gen) de algún individuo. Se usa para garantizar la diversidad de la población en el algoritmo genético. Para el caso de cromosomas en código binario, consiste únicamente en negar un bit o alelo. Pueden seleccionarse directamente los individuos a mutarse, pero lo común es que se haga de forma aleatoria. Puede usarse conjuntamente con el operador de cruce, pero Matlab® usa por default solo al resto de la población que no se cruzó [MathWorks 2013]. La probabilidad de mutación debe ser baja para garantizar la diversidad, de otra forma buscara por otros espacios de solución lo que hace que el algoritmo gaste más recursos de los necesarios, lo que conllevará a que tarde mucho tiempo en converger nuestro algoritmo [Tutorials Point 2016] [Gestal *et al.* 2010] [Ponce Cruz 2010] [Chipperfield *et al.*].

Matlab® usa por default una población inicial de 20 individuos y un 5% de cruce elitista [MathWorks 2013].

3 Red neuronal artificial aplicada al rendimiento del sistema de energía renovable.

En este capítulo de esta tesis, se describirá el desarrollo de un modelo neuronal ANN para predecir la temperatura de salida de un colector solar de bajo costo. Este tipo de colector se elige específicamente como una opción viable para su uso en áreas residenciales, ya que pueden fabricarse con materiales económicos y factibles para su instalación en comunidades de bajos ingresos. Los materiales utilizados para su fabricación se pueden encontrar en cualquier región (tubo de cobre y concentrador de acero inoxidable) y tiene una larga vida útil debido a su resistencia. Adicionalmente, su coste de manufacturación puede ser más bajo que el de un CPC, debido a la doble parábola que maneja suele ser en ocasiones más caro de realizar. Investigaciones recientes han demostrado que los LFR pueden ser viables en cuanto a costo [Morin *et al.* 2012], pero los PTC siguen teniendo mejor eficiencia [Kincaid *et al.* 2018].

Este modelo ANN considera parámetros esenciales que influyen en el comportamiento térmico del colector solar de cilindro parabólico. La temperatura de entrada del fluido, mes, día, hora, temperatura ambiente, la temperatura de la superficie del vidrio y la temperatura de la superficie reflectante fueron las variables de entrada. Además, los valores calculados se compararon con datos experimentales para demostrar la efectividad de la capacidad de la red.

3.1 Descripción del equipo experimental

El sistema experimental estudiado fue un PTC estático el cual está compuesto principalmente de una placa reflectante empotrada en un concentrador parabólico que tiene dimensiones de 85 cm de largo por 57 cm de diámetro. La placa reflectante está hecha de acero inoxidable 316L calibre 16 que se

usa para concentrar los rayos del Sol a lo largo de una línea focal en la cual está colocada el tubo receptor.

El tubo receptor está hecho de cobre de $\frac{1}{2}$ pulgada de diámetro y 90 cm de largo, el cual, tiene un recubrimiento selectivo (pintado de negro) para tener una absorptividad más alta, como se muestra en la Figura 3.1. A través del tubo pasa agua como fluido de trabajo. Una segunda superficie, que es concéntrica al tubo receptor, lo cubre externamente y consiste de 2 tubos concéntricos de vidrio con diferentes diámetros. Para evitar pérdidas convectivas y tomar ventaja del efecto invernadero, el PTC tiene una cubierta de vidrio de 85 cm de largo por 57 cm de ancho son un marco de aluminio.

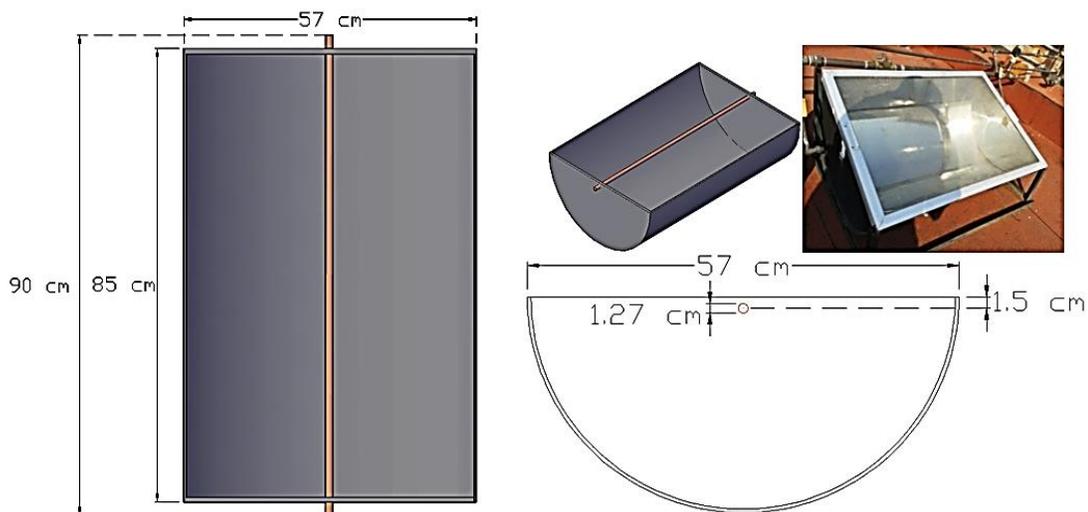


Figura 3.1. Vistas frontal y lateral con las dimensiones del sistema PTC.

El PTC fue instrumentado con termopares tipo T debidamente calibrados para la medición de temperatura en las regiones de interés: a la entrada del tubo receptor ($T_{entrada}$), en el medio ambiente (T_{amb}), en la superficie de la cubierta vidrio (T_{vidrio}), y en la superficie reflectante ($T_{reflector}$). Igualmente, se registró la temperatura a la salida del tubo receptor (T_{salida}), la cual fue la variable de salida o target en el modelo ANN. Una tarjeta de adquisición de datos Agilent

34970A y el software Agilent Benchlink Data Logger fueron usados para adquirir y registrar los datos experimentales.

El equipo experimental en operación se muestra en la Figura 3.2, donde se observa que el tanque de alimentación (fría) y de almacenamiento (caliente) son de 750 y 450 L, respectivamente. Están interconectados para tener una recirculación y así incrementar la temperatura del tanque de alimentación cuando el tanque de agua caliente está en desuso. Se instalaron válvulas y una moto-bomba a presión modelo AQC-15 con una potencia de 2/15 HP para controlar el flujo. Un flujómetro Blue-White F-400 con capacidad de 20 LPM fue usado para medir el flujo volumétrico a la salida del tubo.

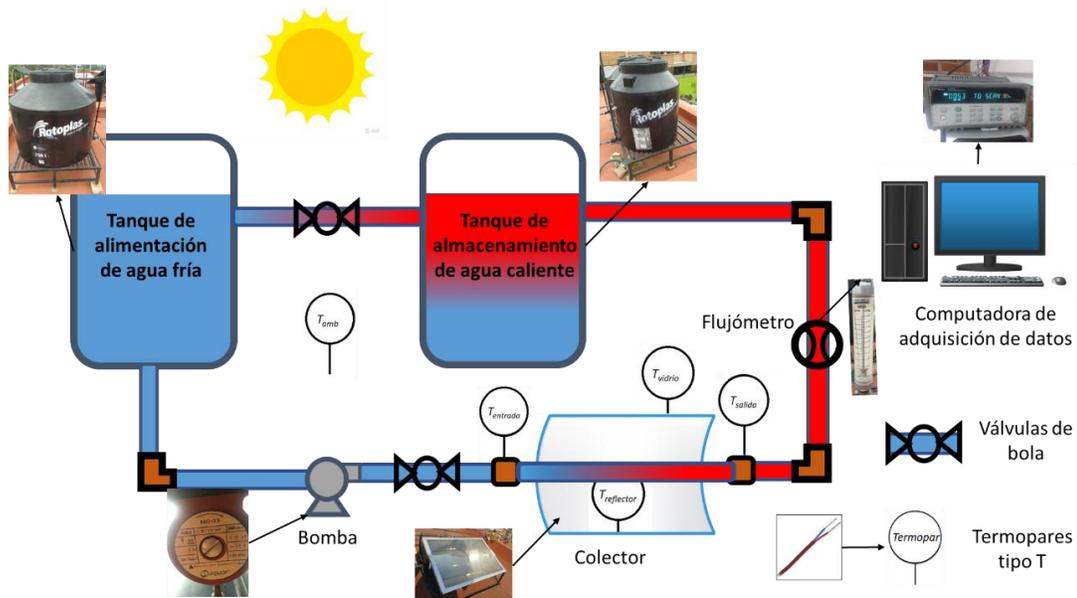


Figura 3.2. Diagrama esquemático del sistema experimental PTC con componentes de bajo costo.

3.2 Modelo de la red neuronal

Para desarrollar el modelo ANN se requieren tres tareas fundamentales. La primera tarea fue la formación de una base de datos de trabajo proporcionada por la variación de las variables de operación y adquirida a través de equipos

de instrumentación. Es esencial mencionar que las variables consideradas para el modelo fueron representativas del fenómeno para reducir el tamaño de la red y evitar la redundancia en los datos de capacitación. En este punto, se debe incluir la normalización del conjunto de datos. La segunda tarea fue obtener una configuración de red adecuada para alcanzar una predicción precisa. En este paso, se consideró la cantidad de neuronas ocultas, el algoritmo de entrenamiento y la función de activación como parámetros esenciales para el entrenamiento de la red neuronal. De modo que más tarde, en la tercera tarea, se estableció la correlación entre la temperatura de salida experimental y la predicha del colector solar.

3.2.1 Condiciones de operación

El PTC está localizado en las instalaciones del CIICAp-UAEM con coordenadas de latitud y longitud de 18.981655 -99.23418, respectivamente. El PTC tiene una inclinación de 19° y está orientado de Este-Oeste con la cara de la placa reflectante apuntado hacia el Sur para aprovechar mejor la incidencia de la radiación solar.

Para formar una base de datos de trabajo que alimentaron al modelo ANN, se obtuvieron 67,691 pruebas experimentales durante Mayo y Junio del 2014, de 8:00 a 19:00 h, con intervalos de 10 segundos entre mediciones. De entre todas las mediciones realizadas, se seleccionaron siete variables que son las que tienen un impacto más significativo para el modelo del colector solar. Las variables de entrada fueron: la temperatura de entrada del fluido ($T_{entrada}$), mes, día, hora, temperatura ambiente (T_{amb}), temperatura de la superficie del vidrio (T_{vidrio}) y temperatura de la superficie reflectante ($T_{reflector}$). La variable de interés para simular y cotejarla con su contraparte experimental fue la temperatura de salida (T_{salida}). La Tabla 3.1 muestra los intervalos de operación a los cuales el sistema PTC está sujeto.

Tabla 3.1. Intervalos de condiciones de operación experimentales.

Variables experimentales	Intervalo de operación
Día	1-31
Mes	1-12
Hora, <i>s</i>	20267-69085
$T_{entrada}, ^\circ C$	15.43-50.00
$T_{amb}, ^\circ C$	9.94-70.28
$T_{vidrio}, ^\circ C$	14.5-51.4
$T_{reflector}, ^\circ C$	16.00-77.92
$T_{salida}, ^\circ C$	15.59-50.66

Los datos de entrada se normalizaron en un intervalo de entre 0.1 a 0.9, debido a que la función utilizada en la capa oculta era la función de transferencia sigmoidea tangente hiperbólica TANSIG. Los datos normalizados de [0.1,0.9] han mostrado mejores resultados en comparación con otras normalizaciones [Hernández y Colorado 2013]. Se sabe que la normalización representa una influencia directa en el modelo y ofrece beneficios tales como resultados adecuados y una disminución considerable en el tiempo de cálculo [Millán-Ocampo *et al.* 2018]. La ecuación que se utilizó para la normalización de datos [Sözen *et. al* 2007] es la siguiente:

$$x_{i,norm} = 0.8 \left(\frac{x_i - x_{min}}{x_{max} - x_{min}} \right) + 0.1 \quad (3.1)$$

Después del proceso de normalización, el conjunto de datos se dividió aleatoriamente para entrenamiento (50%), prueba (25%) y validación (25%). En la Figura 3.3, se representa el proceso de aprendizaje del modelo ANN.

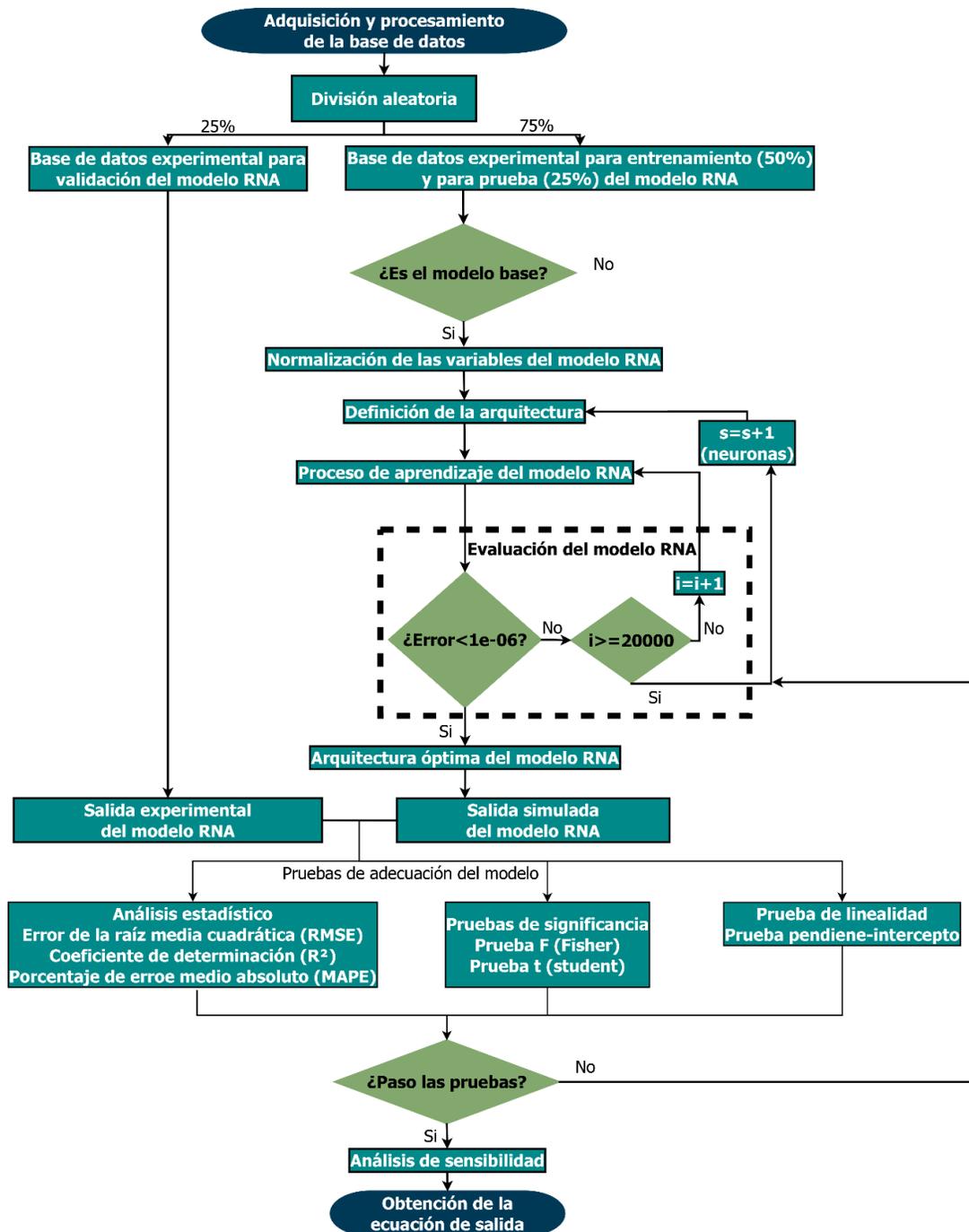


Figura 3.3. Metodología para el entrenamiento del modelo ANN para el sistema PTC.

Para obtener una configuración de red adecuada y lograr la predicción correcta de la temperatura de salida del colector solar lo más confiable posible, se deben desarrollar diferentes arquitecturas ANN. Para esto, se utilizó un modelo

de red de retroalimentación (FeedForward) con la implementación de la estructura de retropropagación (Backpropagation). La función de transferencia tangente hiperbólica sigmoidea y la función de transferencia lineal se utilizaron en la capa oculta y la capa de salida, respectivamente. Se utilizó el algoritmo de entrenamiento Levenberg-Marquardt (LM) que es más rápido por no tener que calcular la matriz de Hessiana [Díaz-Gómez *et al.* 2014]. Además, el algoritmo de LM tiene un error cuadrático medio (*mse*) más pequeño en comparación con otros algoritmos de retropropagación [Arabameri *et al.* 2015]. Como se mencionó anteriormente, este algoritmo se usa ampliamente para tener un tiempo de procesamiento computacional más corto. El desarrollo del modelo neural PTC se llevó a cabo utilizando el software Matlab y el kit de herramientas ANN Toolbox [Demuth y Beale 2005] [Beale *et al.* 2017]. La Figura 3.4 muestra un esquema del modelo ANN.

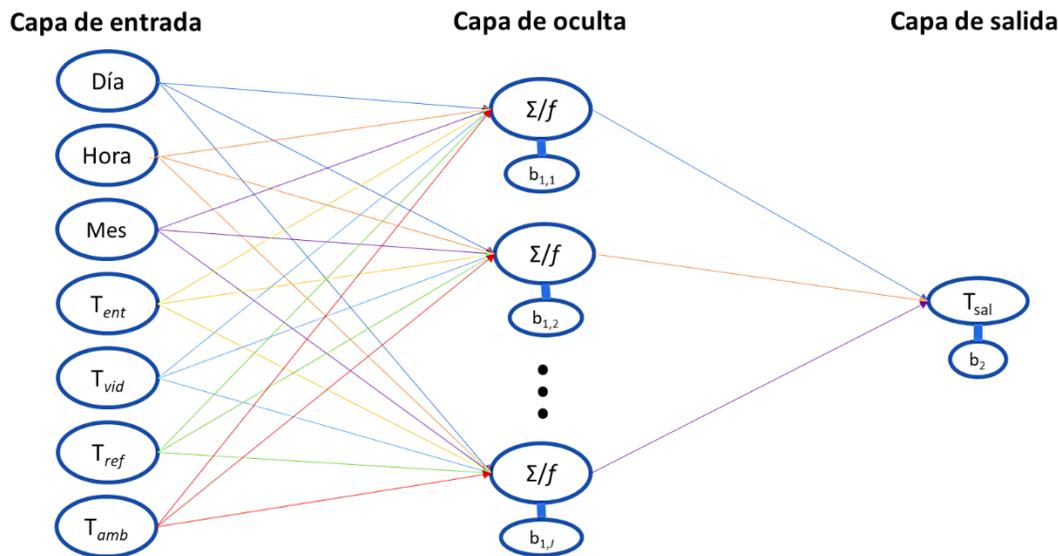


Figura 3.4. Esquema del modelo ANN del sistema PTC.

Para la última tarea, se evaluó el desempeño de la red utilizando los parámetros estadísticos como el coeficiente de determinación (R^2), el Error Porcentual Medio Absoluto (Mean Absolute Percent Error, *MAPE*) y la Raíz del Error Cuadrático Medio (Root Mean Square Error, *RMSE*). Estos parámetros

se describieron en la Sección 2.3.9 de esta tesis. La mejor arquitectura se presentó con 19 neuronas en la capa oculta, y la Tabla 1.1Tabla 3.2, muestra la comparación del análisis estadístico. Como puede verse, al aumentar el número de neuronas ocultas por encima de 20, los valores obtenidos de R^2 , $RMSE$ y $MAPE$ no muestran variaciones significativas. Por lo tanto, el aumento continuo del número de neuronas en la capa oculta reduciría muy poco el error.

Tabla 3.2. Comparación del análisis estadístico para determinar el número de neuronas en la capa oculta.

Arquitectura ANN	Número de neuronas	Coeficiente de Determinación	Raíz del Error Cuadrático Medio	Error Porcentual Medio Absoluto
		R^2	RMSE	MAPE
7-1-1	1	0.8833	1.9824	1.2122
7-2-1	2	0.9092	1.8182	1.1209
7-3-1	3	0.9407	2.0009	1.0951
7-4-1	4	0.9460	1.9167	1.0526
7-5-1	5	0.9577	1.7930	1.0065
7-6-1	6	0.9654	1.7793	0.9871
7-7-1	7	0.9664	1.8196	0.9832
7-8-1	8	0.9704	1.8501	0.9729
7-9-1	9	0.9725	1.8345	0.9888
7-10-1	10	0.9773	1.7538	0.9563
7-11-1	11	0.9972	1.7835	0.9592
7-15-1	15	0.9865	0.7700	0.0555
7-16-1	16	0.9865	0.7701	0.0603
7-17-1	17	0.9867	0.7649	0.0541
7-18-1	18	0.9883	0.7176	0.0542
7-19-1	19	0.9894	0.6829	0.0405
7-20-1	20	0.9886	0.7078	0.0462
7-25-1	25	0.9932	0.5452	0.0319
7-30-1	30	0.9915	0.6119	0.0353

Además, considerar un número de neuronas ocultas más alto involucraría una cantidad más significativa de coeficientes en la ecuación resultante. Por lo tanto, el mejor modelo se presentó con una arquitectura 7-19-1 con un coeficiente de determinación R^2 de 0.9894. Es importante destacar que el error

de generalización de $MAPE$ es mínimo, lo que indica que la red puede generalizar a partir de nuevos datos para obtener buenas predicciones.

Para mejorar la función de rendimiento, se utilizó el método de entrenamiento de Regularización Bayesiana (Bayesian Regularization, BR), que es una modificación del algoritmo Levenberg-Marquardt, el cual mejora la generalización de la ANN [Kamble *et al.* 2015]. Este método consiste en alterar la función de rendimiento, dado por:

$$F = \beta E_d + \alpha E_w \quad (3.2)$$

Donde α y β son parámetros que el algoritmo BR optimiza automáticamente, E_d es la suma de los errores al cuadrado de la red (Mean Square Error, $MSE = \frac{1}{n} \sum_{j=1}^n (x_{exp} - x_{sim})^2$), y E_w es el promedio de la suma de los pesos al cuadrado ($E_w = \frac{1}{n} \sum_{j=1}^n W^2$). Este algoritmo proporciona una medida de qué parámetros (pesos y bias) están siendo utilizados efectivamente por la red. Además de mejorar la generalización de la red, previene el sobre-entrenamiento, y puede ser usado en base de datos pequeñas porque no requiere separar la base de datos para validación. Sus principales desventajas son que tarda más en converger y que no es tan bueno en problemas de reconocimiento de patrones como el método LM [Demuth y Beale 2005].

3.3 Resultados

La comparación entre los parámetros estadísticos de R^2 , $RMSE$, y $MAPE$ usando ambos algoritmos de entrenamiento se describe en la Tabla 3.3. Con estos resultados, se puede ver que el algoritmo de Regularización Bayesiano puede minimizar aún más los errores logrados que con el Algoritmo Levenberg-Marquardt ($MAPE = 0.0372$). Aunque el error en $RMSE$ es ligeramente mayor ($RMSE = 0.6829$), el error de generalización $MAPE$ es

mínimo. Por lo tanto, indica que la red podría generalizar a partir de nuevos datos y obtener predicciones correctas.

Tabla 3.3. Comparación entre parámetros estadísticos de R^2 , $RMSE$, y $MAPE$ utilizando el entrenamiento de algoritmos LM y BR .

Arquitectura ANN	Número de neuronas	Coefficiente de Determinación	Raíz del Error Cuadrático Medio	Error Porcentual Medio Absoluto
		R^2	$RMSE$	$MAPE$
7-19-1	19	0.9894	0.6829	0.0405
7-19-1	19	0.9909	0.6285	0.0372

La Figura 3.5 muestra la comparación de la temperatura de salida experimental y la simulada por los modelos ANN con el método de entrenamiento de Regularización Bayesiana.

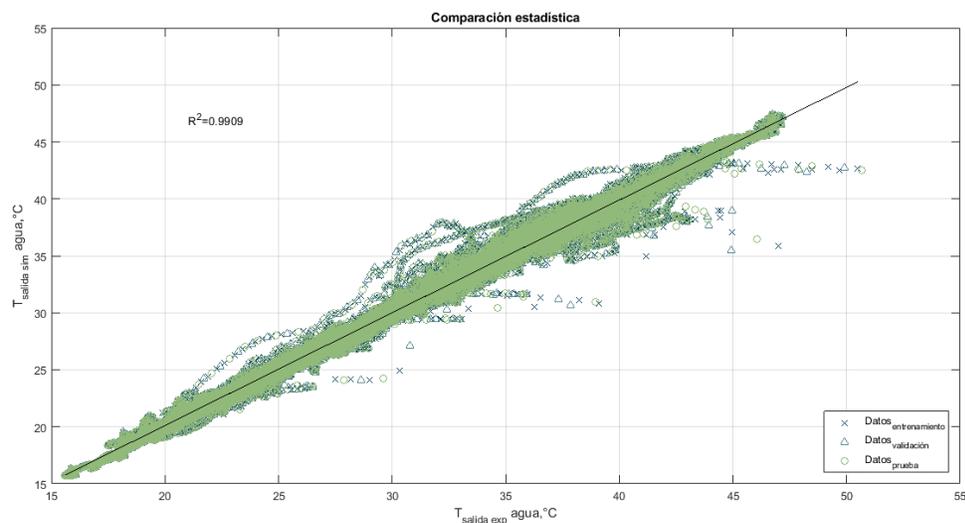


Figura 3.5. Comparación estadística y regresión lineal entre los datos de temperatura de salida simulados (ANN) y experimentales.

La Figura 3.6 muestra la simulación de la temperatura de salida del colector con los modelos propuestos para diferentes días. Se muestra que, en general, los modelos ANN siguen un patrón definido de calentamiento de agua en la salida del colector a pesar de las variaciones que puedan existir. Las variaciones están presentes porque, en los sistemas de energía renovable, la

naturaleza está expuesta a diferentes factores que cambian con respecto al tiempo y al espacio, causando ruido en las mediciones. Es decir, generan la existencia de datos lejos del promedio de las muestras.

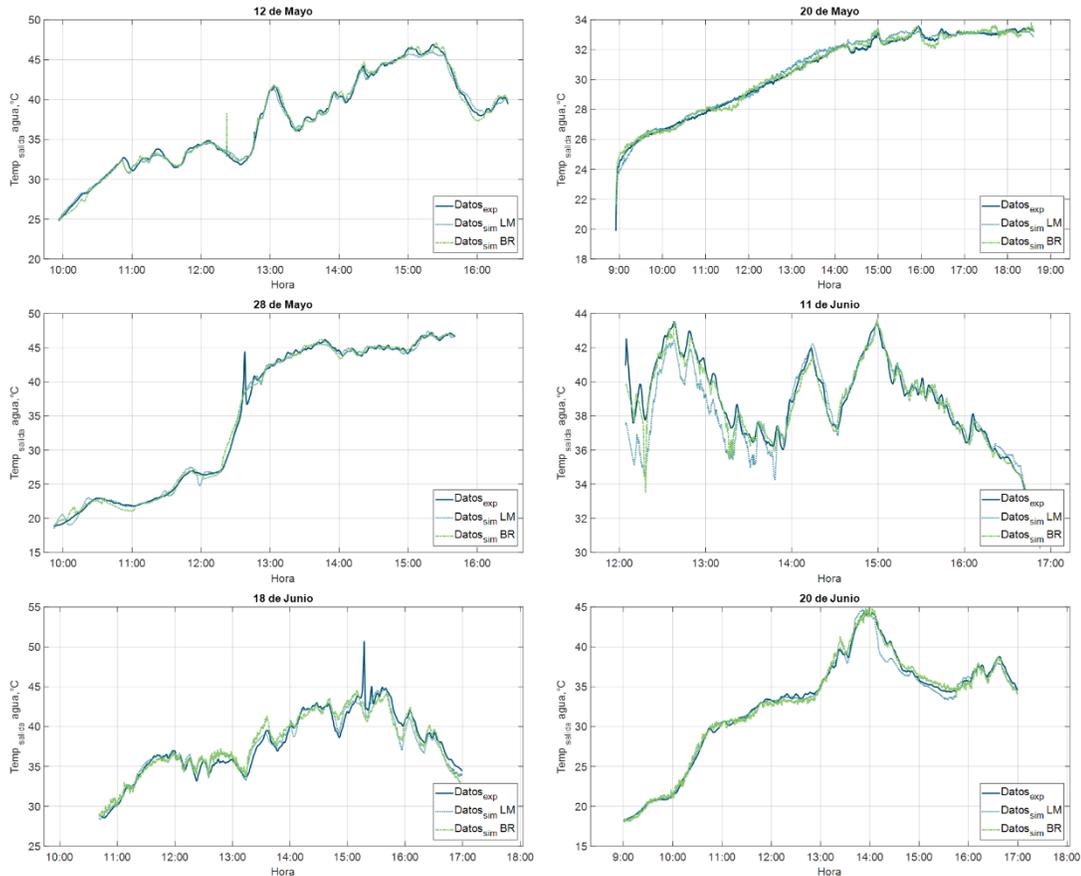


Figura 3.6. Simulación de la temperatura de salida del colector con los modelos propuestos para diferentes días.

Sin embargo, las predicciones de la red se pueden mejorar a través del algoritmo de Regularización Bayesiano. Se observa en detalle que este algoritmo sigue más el patrón principal de los datos experimentales y trata de discriminar los picos que se generan en comparación con su contraparte. Los datos simulados con el algoritmo de BR están, en su mayor parte, menos dispersos y cuando hay un patrón complicado a seguir, se las arregla para mantenerse dentro del promedio.

Para verificar la linealidad y establecer los límites de confianza para los parámetros de regresión, se realizó la prueba de linealidad. En la prueba de linealidad, el valor de la pendiente debe ser 1 (o cercano a 1), y la intersección debe ser 0 (o cercana a 0), como se estableció en la Sección 2.3.11. La Tabla 3.4 muestra la prueba de linealidad para ambos algoritmos de entrenamiento y, como se puede ver, la Regularización Bayesiana obtuvo mejores resultados en comparación con el algoritmo Levenberg-Marquardt, ya que sus parámetros están más cerca de 1 y 0, para pendientes e interceptos, respectivamente.

Tabla 3.4. Prueba de linealidad de pendiente e intercepto.

Método de entrenamiento	R^2	99% de confianza			
		Pendiente superior	Pendiente inferior	Intercepto superior	Intercepto inferior
LM	0.9894	0.9904	0.9884	0.3529	0.2902
BR	0.9909	0.9914	0.9901	0.3013	0.2435

Por lo tanto, el mejor modelo para predecir la temperatura del colector solar fue con la arquitectura 7-19-1 con el algoritmo de entrenamiento Regularización Bayesiana. El modelo ANN está representado en términos generales de acuerdo con las funciones de activación seleccionadas, que fueron [Rocabrano-Valdés *et al.* 2015]:

$$n_j = Wi_{(j,1)} \cdot P_{(1)} + Wi_{(j,2)} \cdot P_{(2)} + Wi_{(j,3)} \cdot P_{(3)} + \dots + Wi_{(j,R)} \cdot P_{(R)} + b_{(1,j)} \quad (3.3.a)$$

$$n_j = \sum_{r=1}^R Wi_{(j,r)} \cdot P_{(r)} + b_{(1,j)} \quad (3.3.b)$$

$$tansig(n_j) = \tau(n_j) = \frac{2}{1+e^{-2n_j}} - 1 \quad (3.4.a)$$

$$\tau(n_j) = \frac{2}{1+e^{-2(\sum_{r=1}^R Wi_{(j,r)} \cdot P_{(r)} + b_{(1,j)})}} - 1 \quad (3.4.b)$$

$$Output = purelin\left[\sum_{j=1}^J (Wo_{(K,j)} \cdot \tau + b_{(2,K)})\right] \quad (3.5.a)$$

$$Output = \sum_{j=1}^J \left\{ Wo_{(K,j)} \left[\frac{2}{1+e^{-2\left[\sum_{r=1}^R (Wi_{(j,r)} \cdot P_{(r)} + b_{(1,j)})\right]}} \right] \right\} + b_{(2,K)} \quad (3.5.b)$$

Por lo tanto, la temperatura de salida del fluido (T_{salida}) se obtuvo reemplazando los 172 coeficientes obtenidos (pesos y bias) en la siguiente ecuación:

$$T_{salida} = \sum_{j=1}^J \left[W_{o(K,j)} \left(\frac{2}{1 + \exp(-2(\sum_{r=1}^R (W_{i(j,r)} \cdot P_{(r)}) + b_{(1,j)}))} - 1 \right) \right] + b_{(2,K)} \quad (3.6)$$

Donde W_o son los pesos entre la capa oculta y la capa de salida, W_i son los pesos entre las capas de entrada-oculta. b_1 son las bias en la capa oculta y b_2 es la bia en la capa de salida. Los subíndices R , J y K refieren al número de neuronas en las capas de entrada, oculta y de salida, respectivamente. La formulación de la ecuación no es un proceso complicado ya que solo se requieren operaciones aritméticas simples. La Tabla 3.5 muestra los coeficientes resultantes del entrenamiento de Regularización Bayesiana.

Tabla 3.5. Parámetros de pesos y bias obtenidos con el entrenamiento del algoritmo BR.

Neuronas	Pesos								Bias	
				W_i				W_o		
	Día	Mes	Hora	T_{input} °C	T_{env} °C	T_{glass} °C	$T_{reflect}$ °C	T_{output} °C	b_1	b_2
1	-6.6343	6.2872	3.4883	-2.2966	0.0203	-1.5021	3.8636	-12.7682	-6.0673	
2	0.9612	-2.0988	-4.4438	15.1499	16.4538	-0.0580	-6.4706	40.2968	-2.7195	
3	42.6016	26.0394	78.3668	-48.4249	-14.4120	36.1544	11.4646	0.5659	-73.3716	
4	-5.8659	0.9408	3.9892	-10.6760	-17.9960	2.4015	2.7706	3.6064	9.0308	
5	-29.4409	-6.4120	-5.3869	5.7846	13.8036	-2.5486	-0.4248	-4.9614	18.4303	
6	9.2342	5.7294	0.7812	1.7514	1.8980	-3.6860	-2.1114	-8.1805	-9.4105	
7	8.6283	-6.2575	6.5926	-6.6689	12.5297	2.5046	2.5926	2.2803	-1.6101	
8	-6.9416	-12.3673	2.6605	-5.0401	-7.9831	-1.2154	3.6659	-11.1146	7.5469	
9	-2.1297	-1.5680	0.0114	2.9297	7.6631	-0.4124	3.2281	10.8348	-0.4840	
10	2.2005	0.1645	-0.7058	-0.1350	-0.7118	0.7017	-1.4186	-45.4173	0.2979	2.3190
11	-5.3189	4.0159	1.4266	-0.5928	4.6552	-0.4311	1.1626	-54.5733	-2.8135	
12	6.5120	-1.6923	-1.2098	-0.5772	-6.8965	0.9929	-0.3655	-40.2758	0.4191	
13	-209.2692	30.1575	-44.5119	33.9495	155.4319	70.3200	-123.0939	-2.6301	24.9688	
14	-10.8554	-26.1708	-64.2758	28.8494	-60.9620	-11.5636	-5.2275	-1.0686	79.0484	
15	9.9964	0.6544	-13.9439	24.0715	47.3801	39.6620	-37.4032	1.4398	-17.3910	
16	6.6736	11.0591	-1.4247	-0.0704	-6.0150	1.2079	-1.1522	47.0567	-1.1056	
17	-0.8569	1.8382	4.0467	-13.8967	-16.0811	0.0087	5.5881	46.2467	2.9249	
18	-191.9632	-70.4659	-0.8560	-0.0997	-12.3246	0.1633	0.0273	-6.8981	185.159	
19	10.2503	0.7817	4.6163	-9.3261	0.9948	0.5215	3.5090	-4.5800	-8.1016	

Con la ecuación de Garson fue posible discernir la influencia de las variables del Colector Solar de Canal Parabólico en el modelo ANN. Los porcentajes finales de las variables de entrada se muestran en la Figura 3.7.

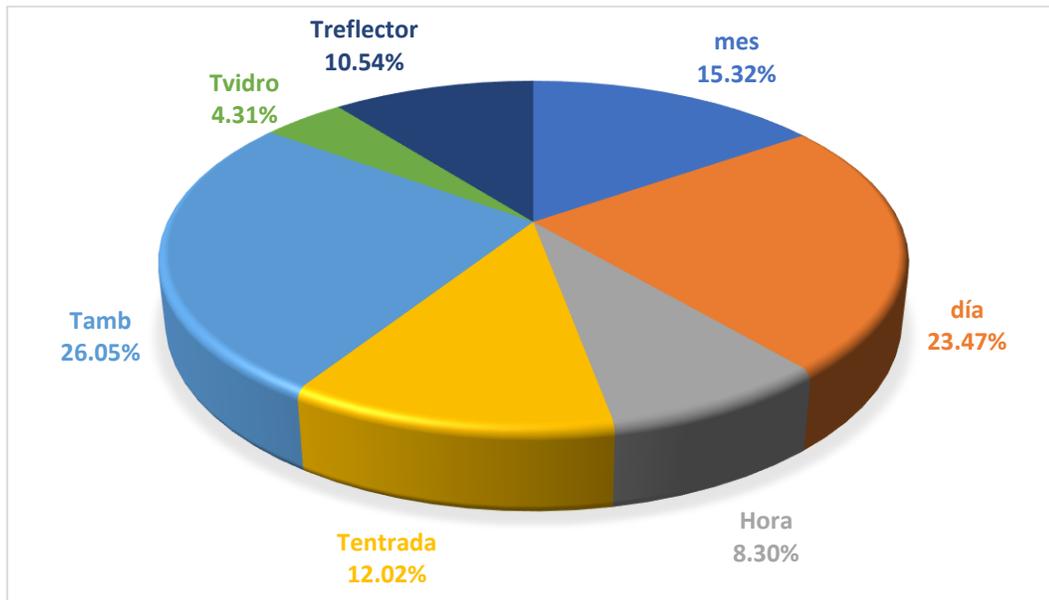


Figura 3.7. Importancia relativa de las variables de entrada del modelo ANN del sistema PTC

El resultado del análisis de sensibilidad sirvió para determinar que variable tiene mayor preponderancia sobre el modelo, la cual fue la temperatura ambiente con 26.05%. Dado que la temperatura ambiente obtenida durante todo el día está estrechamente relacionada con la radiación solar, siendo sus valores máximos alrededor de las 12:00 h y mínimas cercanas a las 19:00 h.

Además, se demostró que tanto el día como el mes son importantes, con 23.47% y 15.32%, respectivamente; esto reitera la dependencia del modelo con la estación del año y puede considerarse como variables influyentes. Las variables de temperatura de entrada del fluido (12.02%), la temperatura de la superficie reflectante (10.54%) y el tiempo (8.30%) muestran un pequeño impacto con una proporción similar. Finalmente, la variable menos importante es la temperatura de la superficie del vidrio con 4.31%.

3.4 Conclusiones parciales

Se desarrolló un modelo de red neuronal artificial para predecir la temperatura de salida del agua en un colector solar de bajo costo mediante mediciones experimentales de un proceso de calentamiento de agua. El modelo considera 7 variables de entrada, que fueron: la temperatura de entrada del fluido, mes, día, hora, temperatura ambiente, la temperatura de la superficie del vidrio y la temperatura de la superficie reflectante.

El modelo de red neuronal fue entrenado con el algoritmo Levenberg-Marquardt. La mejor arquitectura encontrada fue con 19 neuronas en la capa oculta. Según las pruebas estadísticas, se obtuvieron los siguientes resultados: $R^2 = 0.9894$, $RMSE = 0.6829$ y $MAPE = 0.0405\%$. Se mejoró la red obtenida con el algoritmo de entrenamiento Regularización Bayesiana, obteniendo los siguientes resultados estadísticos: $R^2 = 0.9909$, $RMSE = 0.6285$ y $MAPE = 0.0372\%$. Los resultados obtenidos concluyen que el modelo ANN obtenido fue capaz de simular los comportamientos no lineales de la temperatura de salida del colector para el período de Mayo a Junio.

Con el análisis de sensibilidad, se demostró que todas las variables influyen en el modelo. Sin embargo, la temperatura ambiente es la variable que más influyó además de las variables relacionadas con la temporada. La temperatura de entrada del agua y la temperatura de la superficie reflectante dependen de los materiales y el diseño del PTC, por lo que estas variables pueden considerarse como posibles oportunidades para aumentar la temperatura de salida.

Cuando se aplica un modelo ANN para predecir la temperatura de salida del colector, no se pretende eliminar el uso de modelos matemáticos para obtener el rendimiento de un colector solar, sino que se debe considerar como una opción que facilita la comprensión del comportamiento del calentamiento del fluido en colectores solares.

4 Condición de funcionamiento óptima para un colector tipo cilindro parabólico con componentes de bajo costo utilizando una red neuronal inversa y resuelto por algoritmo genético.

Los colectores de canal parabólico (PTC) son una opción posible para su uso en áreas residenciales, ya que pueden fabricarse con materiales económicos y son factibles para su instalación en comunidades de bajos ingresos. En este trabajo, este tipo de colector se elige específicamente porque los materiales utilizados se pueden encontrar en cualquier región (tubo de cobre y concentrador de acero), con el fin de tener una larga vida útil por su resistencia, en comparación con otros materiales frágiles.

En este trabajo, se estudia la descripción del equipo experimental PTC con componentes de bajo costo. Además, se simulan los datos experimentales obtenidos con un modelo de red neuronal artificial (ANN). Una vez simulados los datos experimentales utilizando ANN, es necesario conocer las condiciones óptimas en las variables de entrada para el PTC para obtener mejores resultados. La Red Neuronal Artificial inversa (Artificial Neuronal Networks inverse, ANNi) es un elemento esencial para calcular las condiciones óptimas de operación. Este modelo utiliza los mismos pesos y bias obtenidos durante el modelo ANN de entrenamiento y se ha utilizado en equipos experimentales para mejorar el rendimiento, tales como: Laidi y compañía utilizaron este tipo de metodología para determinar el valor óptimo de COP solar de un sistema de refrigeración por adsorción asistido por energía solar que funciona con carbón/metanol activado como pares de trabajo, donde esta metodología se aplicó para encontrar el parámetro de entrada óptimo para el COP solar requerido [Laidi y Hanini 2013]. Hattab junto a su equipo de trabajo

desarrollaron y probaron un modelo de Red Neuronal Artificial inversa para la predicción del tratamiento óptimo del suelo para reducir la toxicidad del cobre (Cu) evaluada por una concentración objetivo determinada de Cu en las hojas de frijol enano de los insumos seleccionados del suelo [Hattab y Motelica-Heino 2014]. Morales y sus colaboradores desarrollaron un modelo de Red Neuronal Artificial (Artificial Neuronal Network, ANN) para predecir el coeficiente de rendimiento (COP) de un transformador de calor de absorción con un nuevo diseño físico que consta de componentes compactos y aplicaron un modelo de Red Neuronal Artificial inversa para optimizar el valor de COP, que muestra un aumento en COP de 0.35–0.40, encontrando las condiciones óptimas en el generador de temperaturas de entrada y para el evaporador [Morales *et. al* 2015].

Se elige la temperatura de salida de agua (T_{salida}) del PTC como un valor requerido, ya que se considera una variable general para determinar el sistema de rendimiento, pero para conocer este valor, depende de las variables del sistema, tales como: la temperatura del tanque de alimentación ($T_{suministro}$), la temperatura del tanque de almacenamiento del agua caliente ($T_{almacén}$), la temperatura de entrada al colector solar ($T_{entrada}$), mes (m), día (d), tiempo total (t), temperatura ambiente (T_{amb}), la temperatura de la superficie al colector solar de vidrio (T_{vidrio}) y la temperatura de la superficie del reflector ($T_{reflector}$), que algunas de estas variables de entrada pueden optimizarse para aumentar el rendimiento del PTC.

4.1 Descripción del equipo experimental.

El sistema experimental utilizado fue el mismo que se explicó en la Sección 3.1, con la variante de que se anexaron las mediciones de las temperaturas del tanque de alimentación (agua fría) y del tanque de almacenamiento de agua caliente (termo-tanque), como se muestra en la Figura 4.1.

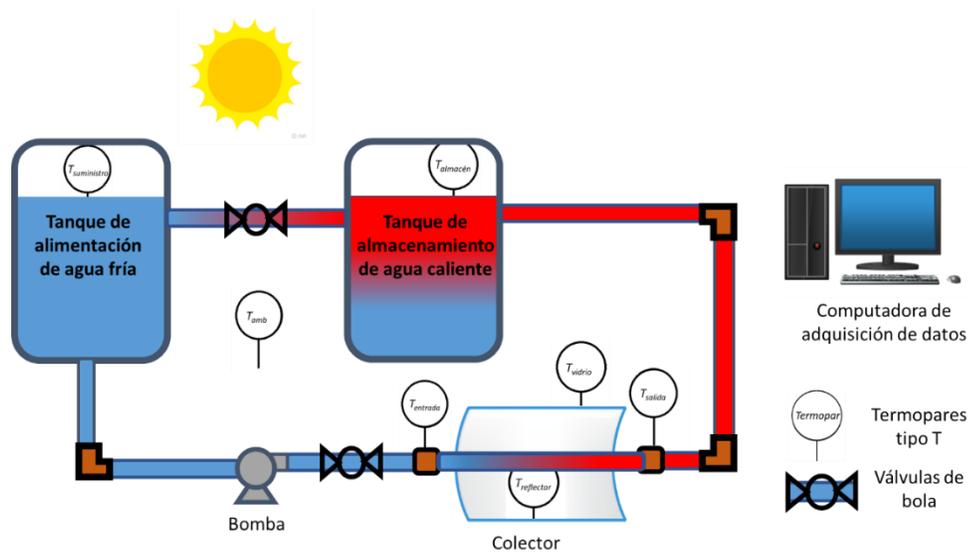


Figura 4.1. Diagrama esquemático del sistema experimental PTC con componentes de bajo costo.

4.2 Condiciones de operación

La ubicación del colector de canal parabólico se describe en la Sección 3.2.1. Se realizaron pruebas experimentales con un caudal de agua sanitaria de 1 LPM para uso común. Para esto, las válvulas de cierre fueron reguladas y de acuerdo con el medidor de flujo para mantener este flujo constante durante toda la duración de la prueba. La Tabla 4.1 muestra las condiciones de operación que se aplicaron en el PTC experimental.

Tabla 4.1. Intervalos de las condiciones experimentales utilizadas para obtener los valores de T_{Salida} .

VARIABLES EXPERIMENTALES	INTERVALO DE OPERACIÓN	UNIDADES
Entradas:		
Mes	1-12	Adimensional
Día	1-31	Adimensional
Hora	20267-69085	s
Temp. Entrada	9-50.01	°C
Temp. Ambiente	9.07-70.28	°C
Temp. Superficie Vidrio	9.53-51.4	°C
Temp. Reflector	9.05-77.92	°C
Temp. Suministro	13.14-30.35	°C
Temp. Almacenamiento	13.57-54.07	°C
Salida/Target:		
Temp. Salida	9.73-50.67	°C

Las horas, en las que se iniciaron y terminaron las pruebas, fueron variadas, tomando la primera hora del día a las 5:30 a.m., mientras que la última fue a las 19:10 h. El inicio y final más comunes de la prueba fue a las 9:00 a las 17:30 h. El intervalo entre mediciones fue de 10 s. Estas pruebas se realizaron en los meses de Mayo y Junio (primavera y verano) de 2014 y Enero de 2016 (invierno) [Reyes-Téllez 2015].

4.3 Desarrollo del modelo neuronal

El modelo de red neuronal artificial (ANN) se integró con R entradas, donde cada entrada (P) se asignó con un coeficiente apropiado llamado peso (Wi). La suma ponderada de las entradas y bias b genera el argumento (n_j) que se aplicará en una función de transferencia que generará una salida, esta ecuación es la misma ecuación que se dio en la Ec. 2.1 y Ec. 3.3.a.

$$n_j = Wi_{(j,1)} \cdot P_{(1)} + Wi_{(j,2)} \cdot P_{(2)} + Wi_{(j,3)} \cdot P_{(3)} + \dots + Wi_{(j,R)} \cdot P_{(R)} + b_{(1,j)} \quad (4.1)$$

Los pesos (Wi) y las bias ($b_{(1,j)}$) son coeficientes conectados con las capas ocultas de entrada que se agrupan en matrices. Para la capa de salida, la suma ponderada de las señales proporcionadas por la capa oculta y los coeficientes asociados se agruparon en matrices Wo y $b_{(2,K)}$. En la capa oculta, se aplicó la función de transferencia de tangente hiperbólica ($TANSIG$) y en la capa de salida se usó la función de transferencia lineal ($PURELIN$) para tener una mejor predicción con respecto a los datos experimentales. Todos los cálculos se realizaron con el software matemático Matlab con la caja de herramientas ANN.

Para el desarrollo del modelo ANN, se realizó a través de tres tareas generales [Bassam *et al.* 2014] (ver Figura 4.2).

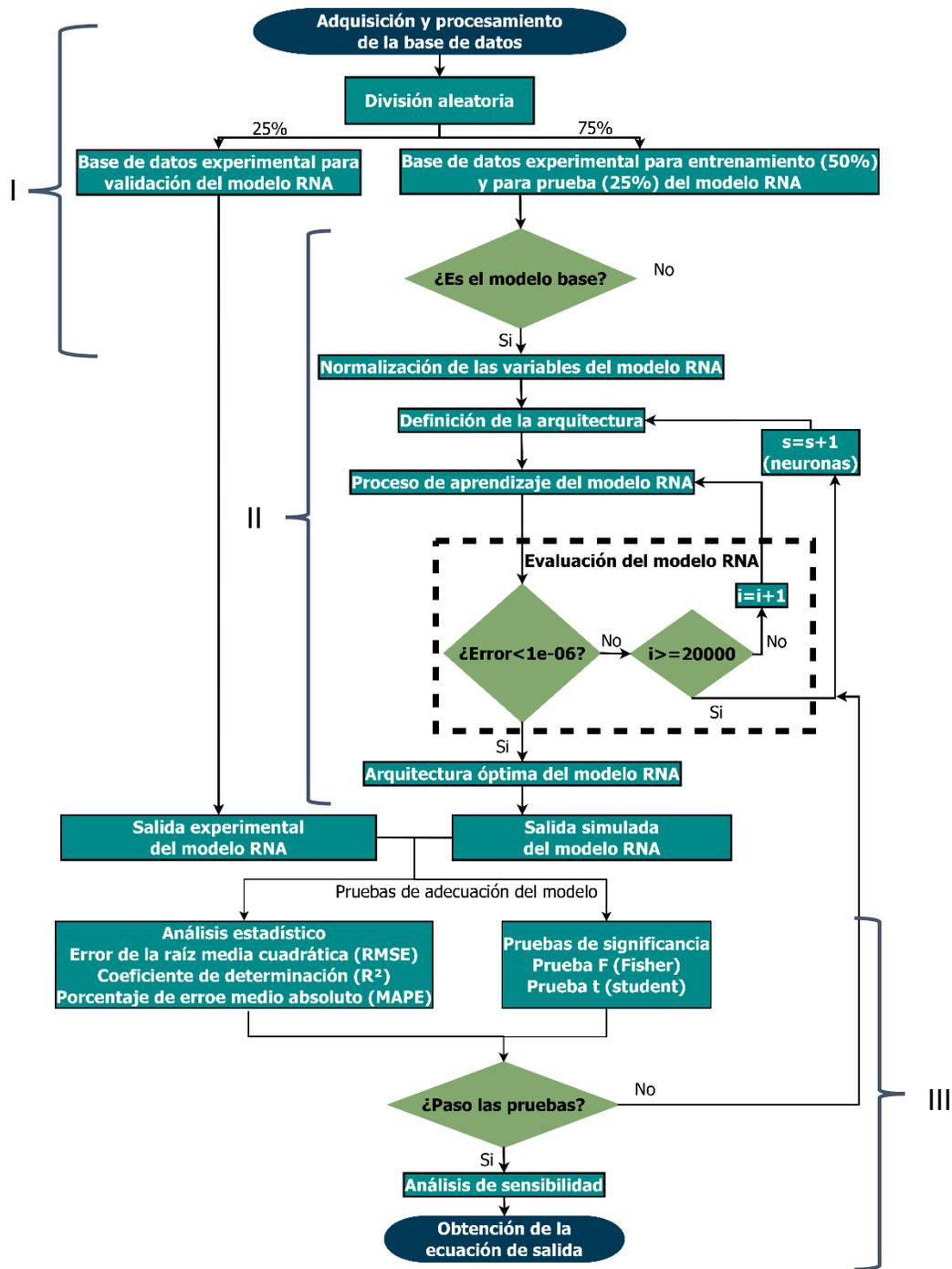


Figura 4.2. Diagrama de flujo para el modelo ANN del PTC.

- I. Divide aleatoriamente la base de datos experimental (69,047) en capacitación (50%), pruebas (25%) y validación (25%). Estos datos fueron suficientes para entrenar y probar el modelo ANN.
- II. Desarrollo y evaluación de modelos ANN, para la predicción confiable de la temperatura de salida del agua caliente de PTC.
- III. Análisis estadístico comparativo entre datos experimentales y T_{salida} predicho (inferido de la aplicación de las herramientas ANN).

Es importante mencionar que los parámetros de entrada se normalizaron ($x_{i,norm}$) en el intervalo [0, 0.9] de la siguiente manera [Rocabruno-Valdés *et al.* 2015], esta ecuación fue dada en la Sección 3.2.1 (Ec. 3.1):

$$x_{i,norm} = 0.8 \cdot \left(\frac{x_i - x_{min}}{x_{max} - x_{min}} \right) + 0.1 \quad (4.2)$$

La ANN fue normalmente entrenada para dirigir una entrada particular para predecir un objetivo de salida. Para minimizar las diferencias entre la salida *target* (dado por datos experimentales) y la salida simulada (generada por un proceso de ajustes de pesos), se consideró los algoritmos de optimización de retropropagación (Backpropagation) usando el algoritmo de entrenamiento Levenberg-Marquardt [Hagan y Menhaj 1994] [Hagan *et al.* 2014] como el más eficiente y ampliamente preferido (ver Figura 4.3) [Çelik *et al.* 2016]. La raíz del error medio cuadrático (*RMSE*) se aplicó como un criterio de optimización en la capacitación para el ajuste del modelo ANN.

Con el fin de corroborar la adaptabilidad del modelo ANN, se compararon los datos experimentales con los datos simulados aplicando los siguientes parámetros de prueba estadística: Raíz del Error Medio Cuadrático (*RMSE*), Error Porcentual Medio Absoluto (*MAPE*) y el Coeficiente de Determinación (R^2) [Verma 2005], que se mostraron en la Sección 2.3.9.

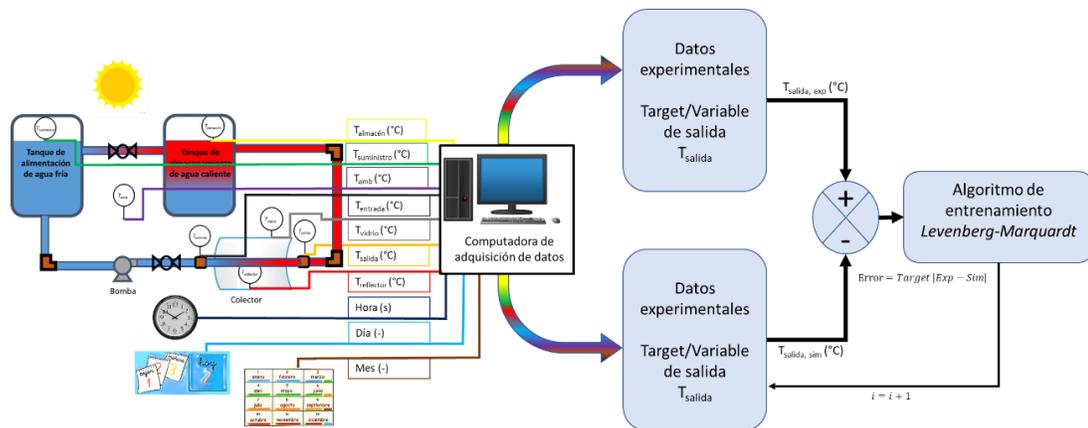


Figura 4.3. Arquitectura aplicada a la red para los valores de T_{salida} y procedimiento usado para el aprendizaje de la red neuronal.

Finalmente, se realizó la comparación de la base de datos simulada mediante ANN y la base de datos experimental aplicando las pruebas de las relaciones F de Fisher (misma o similar varianza) y t de Student (misma o similar media poblacional) con el fin de confirmar la adecuación del modelo [Verma 2005]. Las ecuaciones se mostraron en la Sección **2.3.10** de esta tesis.

4.4 Resultados del modelo de red neuronal

Se descubrió que un modelo ANN con nueve neuronas en la capa oculta es eficiente para predecir los valores de temperatura de salida (T_{salida}), para un colector de canal parabólico (PTC) con componentes de bajo costo. Se probaron diferentes números de neuronas en la capa oculta para determinar la mejor arquitectura del modelo ANN. La Tabla 4.2 proporciona una comparación entre la función de transferencia (TANSIG) y el error para T_{salida} predicho por un modelo ANN.

Tabla 4.2. Pruebas con diferentes arquitecturas ANN.

Arquitectura ANN	Número de neuronas	Épocas	Coeficiente de determinación	Raíz del Error Cuadrático Medio	Error Porcentual Medio Absoluto	Raíz del Error Cuadrático Medio	Error Porcentual Medio Absoluto
			(R^2)	($RMSE$)	($MAPE$)	($RMSE$)	($MAPE$)
TANSIG						Datos de entrenamiento	
9-1-1	1	1000	0.9158	1.9373	0.3969	1.3697	0.3975
9-2-1	2	1000	0.9461	1.5492	0.2131	1.0956	0.2140
9-3-1	3	1000	0.9587	1.3560	0.1729	0.9586	0.1728
9-4-1	4	1000	0.9693	1.1695	0.1350	0.8277	0.1353
9-5-1	5	1000	0.9712	1.1335	0.1296	0.8018	0.1285
9-6-1	6	1000	0.9751	1.0532	0.1041	0.7458	0.1057
9-7-1	7	1000	0.9786	0.9764	0.1000	0.6905	0.0995
9-8-1	8	1000	0.9808	0.9244	0.0897	0.6540	0.0909
9-9-1	9	1000	0.9854	0.8055	0.0586	0.5700	0.0603
9-10-1	10	1000	0.9851	0.8140	0.0651	0.5763	0.0673

Acorde a los datos obtenidos por el entrenamiento ANN, los mejores valores para $RMSE$, $MAPE$ y R^2 fueron 0.8055, 0.0586 y 0.9854, respectivamente. Figura 4.4 presenta una comparación entre los datos predichos y experimentales de los valores de T_{salida} , utilizando todos los datos (entrenamiento, prueba y validación). Los datos experimentales de $T_{salida.exp}$ y los simulados $T_{salida.sim}$ se compararon satisfactoriamente mediante un modelo de regresión lineal [$T_{salida.sim} = 0.98542 * T_{salida.exp} + 0.44038$].

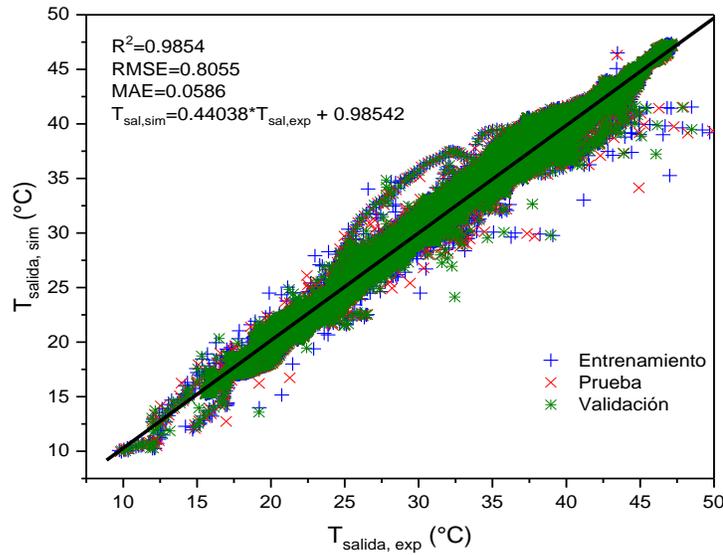


Figura 4.4. Comparación estadística entre los datos de T_{salida} experimentales y los simulados.

La Tabla 4.3 muestra la prueba de la relación t de Student y Fisher, donde se mostró la hipótesis nula H_0 para ambas pruebas. Los resultados en las dos muestras tienen la misma varianza y las medias de las muestras son iguales, lo que indica que la muestra proviene de la misma población.

Tabla 4.3. Aplicación de las pruebas de significancia de las pruebas de Fisher y t de Student.

Prueba F de Fisher	Prueba t de Student
F (Obtenida de la ecuación 2.11) 1.0146	t (Obtenida de la ecuación 2.12) 0.1161
F _c (Obtenida de tablas estadísticas 99%) 1.0178	t _c (Obtenida de tablas estadísticas 99%) 2.3264
Comparación 1.0146 < 1.0178	Comparación 0.1161 < 2.3264
La hipótesis nula H_0 es aceptada " $S_x^2 \approx S_y^2$ "	La hipótesis nula H_0 es aceptada " $\bar{x} \approx \bar{y}$ "

La Tabla 4.4 presenta los parámetros ajustables (W_i , W_o , $b_{(1,j)}$, $b_{(2,K)}$) del modelo propuesto. En consecuencia, estos coeficientes se utilizan en el modelo ANN para simular valores de T_{salida} .

Tabla 4.4. Parámetros de pesos y bias obtenidos para el modelo ANN desarrollado.

Número de entradas (r)	Pesos de la capa oculta $W_{i(r,j)}$								
	(j=1)	(j=2)	(j=3)	(j=4)	(j=5)	(j=6)	(j=7)	(j=8)	(j=9)
1	-7.47223	-45.43951	-23.35323	-28.32771	8.04169	6.66546	6.94243	-9.03039	132.6431 1
2	-4.33701	0.01231	6.69262	-3.32465	0.42654	0.80512	3.23766	1.83490	-0.08478
3	0.19145	0.66056	-0.98259	1.55410	-1.23535	-0.16123	0.45906	0.57498	0.64693
4	-49.39705	11.99704	-5.02621	-36.71076	-19.01847	2.39624	25.20531	27.59983	13.66471
5	13.39573	12.11535	7.49851	-15.02677	-9.06184	-2.68765	8.22856	2.31844	4.25276
6	-30.41051	-37.72298	-50.21982	16.61051	-7.70863	8.37961	-27.39636	-43.31220	266.32813
7	8.39629	-1.98483	-1.20287	11.43457	5.90166	-1.24893	-4.12611	-8.70573	-4.95515
8	4.88839	-5.00822	3.21137	8.65863	4.98912	-2.18178	-3.74167	-7.26927	-12.09112
9	-46.48324	6.28701	-25.30987	10.58262	-20.10860	-0.48387	-9.75767	18.70274	-14.01995
Pesos de la capa de salida (K = 1) $W_{o(j,K)}$									
1, j	1.96232	6.55676	21.17485	3.29298	-2.98869	-1.59993	3.95455	4.29601	-2.13722
Bia $b_{(1,j)}$	28.05543	-2.35558	-0.75726	37.70280	-18.33455	18.67116	-1.07319	0.33432	52.19837
Bia $b_{(2,K)}$	22.14903								

* j es el número de neuronas en la capa oculta, r es el número de neuronas en la capa de entrada, K es el número de neuronas en la capa de salida (K= 1).

Basándose en la mejor estructura de ANN (9 neuronas en la capa oculta) y una función de transferencia (*TANSIG*), el modelo propuesto puede representarse analíticamente mediante la siguiente ecuación (igual a la Ec. 2.26 de la Sección 3.3):

$$T_{salida} = \sum_{j=1}^J \left[W_{o(K,j)} \left(\frac{2}{1 + \exp(-2(\sum_{r=1}^R (W_{i(j,r)} P_{(r)}) + b_{(1,j)})} - 1 \right) \right] + b_{(2,K)} \quad (4.3)$$

Donde J es el número de neuronas en la capa oculta (J = 9), R es el número de neuronas en la capa de entrada (R = 9), y W y b son los pesos y bias, respectivamente (ver Tabla 4.4).

Este modelo, después de ser comparado y analizado estadísticamente, se puede utilizar con éxito para la simulación y el control en línea del equipo experimental, como se muestra en la Figura 4.5, que compara los valores de T_{salida} con respecto a la hora del día, donde Se observa que los valores más altos de T_{salida} alcanzaron en el período de 12 a 16 h, debido a la región y la luz solar. Con el modelo ANN, se observó un buen acuerdo entre los valores pronosticados y los puntos de datos experimentales. Esto confirmó la

adecuación del ANN para llevar a cabo la simulación desde la temperatura de salida del agua caliente para un colector de canal parabólico con componentes de bajo costo.

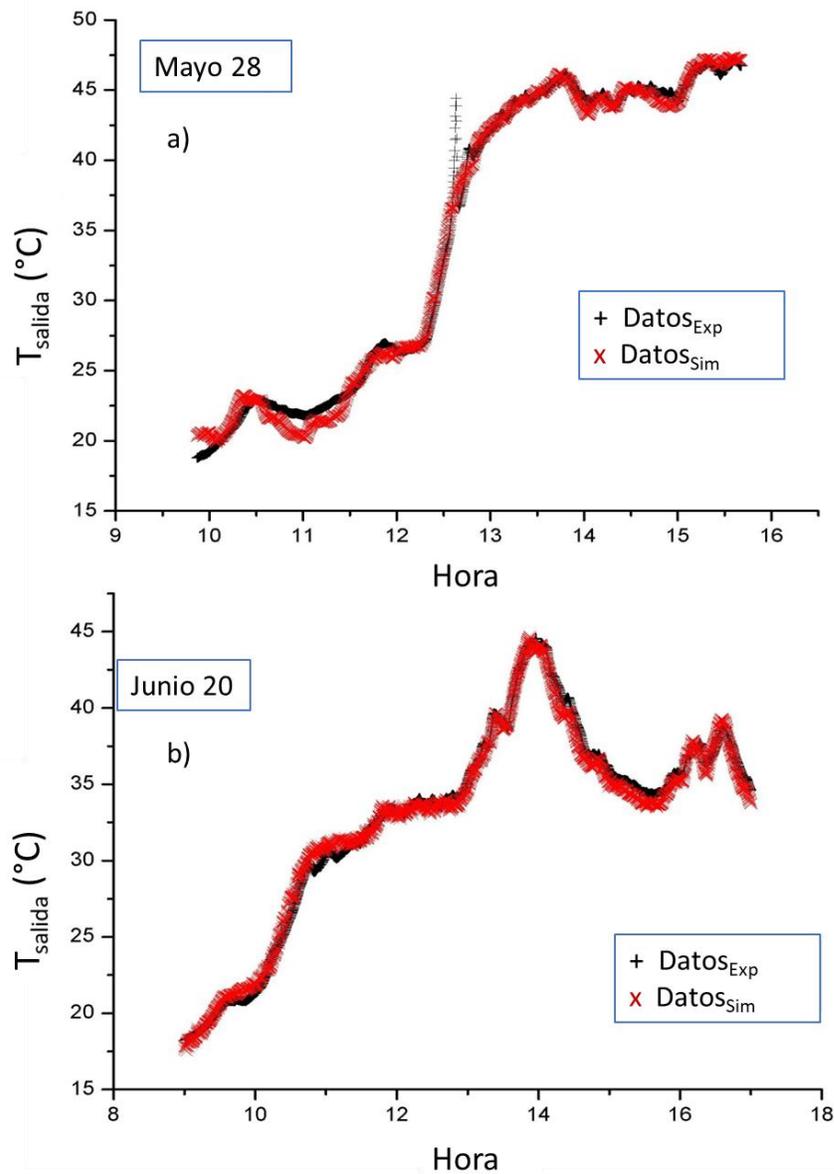


Figura 4.5. Comparación de la temperatura de salida del agua caliente del PTC entre los datos experimentales y los simulados por la ANN, con respecto al tiempo de un día a) para Mayo, b) para Junio.

4.4.1 Análisis de Variables

Para determinar cuál variable tiene mayor preponderancia, se utilizó la ecuación de Garson, descrita en la Sección 2.3.12. Esto también sirve para saber si la variable a optimizar mediante ANNi tenga una gran repercusión en el modelo. En la Figura 4.6 se muestra que la temperatura de entrada (19.69%) es la de mayor importancia, seguida de la hora (14.13%), temperatura ambiente (12.80%) y temperatura de almacenamiento (12.51%). Es importante señalar que la influencia de temperatura de la superficie de vidrio es la más baja de todas las variables elegidas para realizar el modelo ANN, pero no lo es tanto para ser descartada.

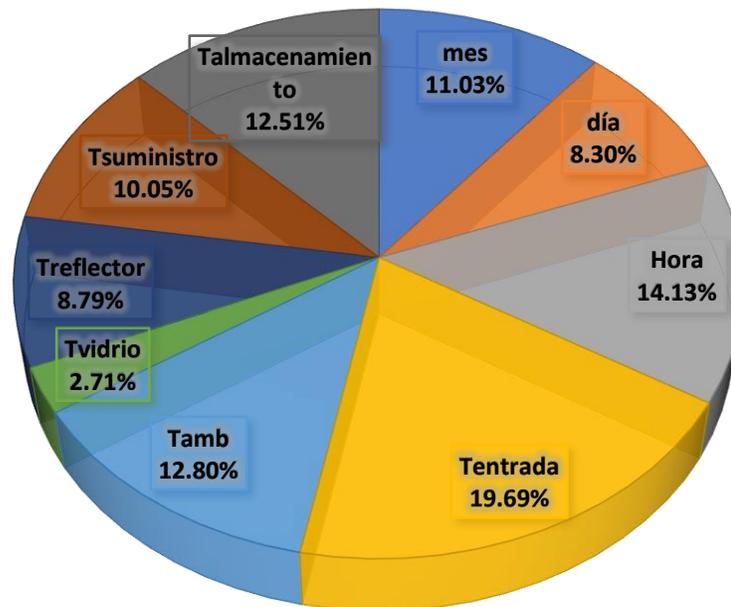


Figura 4.6. Análisis de sensibilidad de variables para determinar la variable con mayor influencia en este modelo.

4.5 Desarrollo del modelo de red neuronal inverso

Esta metodología se utilizó para calcular las condiciones óptimas de funcionamiento en las variables de entrada del equipo experimental para obtener mejores resultados [Hernández 2009]. A partir de la ec. 4.3, si la salida

T_{salida} es conocida, se puede encontrar una variable de entrada óptima (R), como se muestra en los siguientes pasos: Primero se desarrolla la ecuación del T_{salida} del modelo ANN, con nueve neuronas, con sus pesos y bias, respectivamente.

$$\begin{aligned}
 T_{salida} = 2 & \left[\frac{Wo_{(1,1)}}{1 + e^{(x_1)}} + \frac{Wo_{(1,2)}}{1 + e^{(x_2)}} + \frac{Wo_{(1,3)}}{1 + e^{(x_3)}} + \frac{Wo_{(1,4)}}{1 + e^{(x_4)}} + \frac{Wo_{(1,5)}}{1 + e^{(x_5)}} + \frac{Wo_{(1,6)}}{1 + e^{(x_6)}} \right. \\
 & \left. + \frac{Wo_{(1,7)}}{1 + e^{(x_7)}} + \frac{Wo_{(1,8)}}{1 + e^{(x_8)}} + \frac{Wo_{(1,9)}}{1 + e^{(x_9)}} \right] - (Wo_{(1,1)} + Wo_{(1,2)} + Wo_{(1,3)} \\
 & + Wo_{(1,4)} + Wo_{(1,5)} + Wo_{(1,6)} + Wo_{(1,7)} + Wo_{(1,8)} + Wo_{(1,9)}) \\
 & + b_{(2,1)}
 \end{aligned} \tag{4.4}$$

Donde en (x) puede elegir cualquiera de las 9 variables de entrada disponibles (R), para conocer su valor óptimo y observar los resultados obtenidos con respecto a T_{salida} .

$$\begin{aligned}
 x_1 = -2 \cdot & (Wi_{(1,1)} \cdot P_1 + Wi_{(1,2)} \cdot P_2 + Wi_{(1,3)} \cdot P_3 + Wi_{(1,4)} \cdot P_4 + Wi_{(1,5)} \cdot P_5 \\
 & + Wi_{(1,6)} \cdot P_6 + Wi_{(1,7)} \cdot P_7 + Wi_{(1,8)} \cdot P_8 + Wi_{(1,9)} \cdot P_9 + b_{(1,1)}) \\
 & \vdots \\
 x_9 = -2 \cdot & (Wi_{(9,1)} \cdot P_1 + Wi_{(9,2)} \cdot P_2 + Wi_{(9,3)} \cdot P_3 + Wi_{(9,4)} \cdot P_4 + Wi_{(9,5)} \cdot P_5 + \\
 & Wi_{(9,6)} \cdot P_6 + Wi_{(9,7)} \cdot P_7 + Wi_{(9,8)} \cdot P_8 + Wi_{(9,9)} \cdot P_9 + b_{(1,9)}) \tag{4.5}
 \end{aligned}$$

En este trabajo, fue importante identificar los parámetros de operación a través de ANNi, para establecer un valor óptimo. En este caso, se eligió la temperatura del tanque de alimentación, ya que esta variable puede ser regulada directamente por una válvula, que a su vez repercuten en la temperatura de entrada. De acuerdo con la optimización, se puede variar la temperatura del agua, dejando entrar agua más caliente o más fría en el equipo, según sea necesario. Las nuevas temperaturas establecidas

permitirán un funcionamiento eficiente del sistema y pueden aumentar el valor de T_{salida} .

Aplicando la metodología ANNi (ver Figura 4.7), la función objetivo se puede minimizar a cero para establecer la condición de entrada óptima que se obtiene en la ecuación 4.6, que en este trabajo se resolvió mediante un método de optimización a partir de algoritmos genéticos.

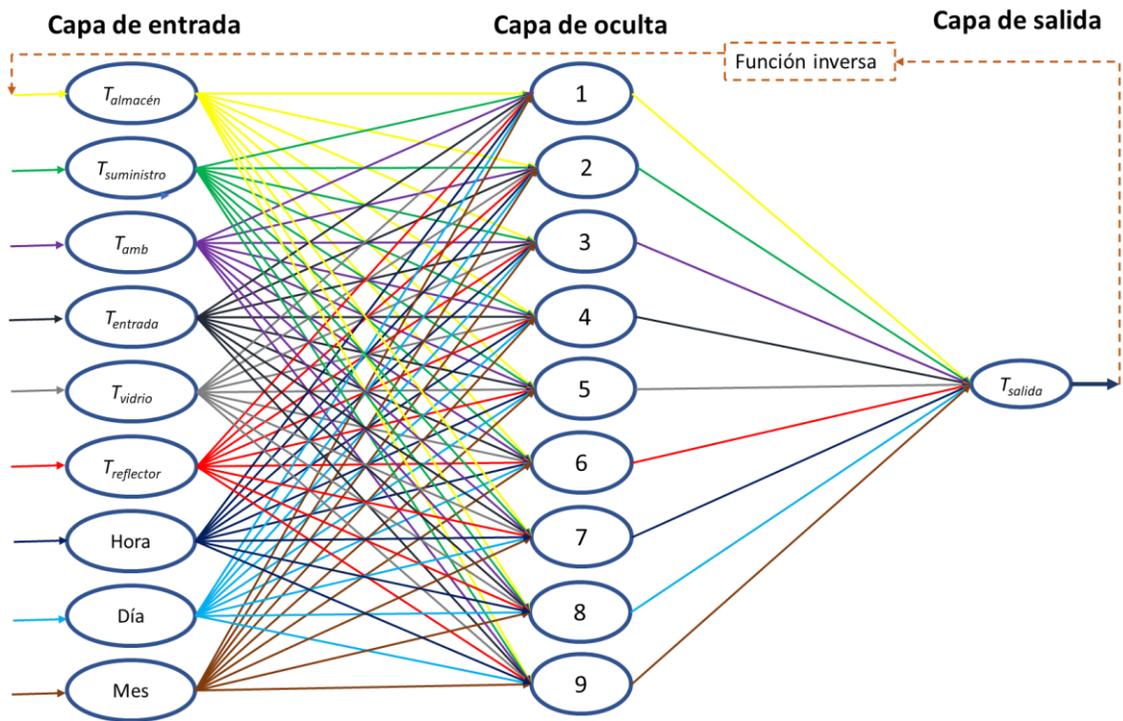


Figura 4.7. Arquitectura de la ANNi para determinar la $T_{suministro}$ óptima y encontrar el valor máximo de T_{salida} .

La ecuación de la función objetivo, queda de la siguiente forma:

$$f(T_{suministro}) = b_{(2,K)} + \sum_{j=1}^J \left[\frac{2 \cdot Wo_{(K,j)}}{1 + e^{\{-2[(Wi_{(j,R)} \cdot T_{suministro}) + \sum_{r=1}^R (Wi_{(j,r)} \cdot P_{(r)} + b_{(1,j)})]\}}} \right] - \sum_{j=1}^J Wo_{(K,j)} - T_{salida} \quad (4.6)$$

4.5.1 Estimación del valor adecuado mediante algoritmos genéticos.

Para resolver la ecuación 4.6, se utilizó una herramienta computacional llamada algoritmos genéticos (AG). AG fue un algoritmo de optimización estocástico que emplea una población de cromosomas y cada uno de ellos representa una posible solución. Considerando los operadores genéticos, cada mejora incremental sucesiva en un cromosoma se convierte en la base de la próxima generación. El proceso continúa hasta que se haya completado el número deseado de generaciones o se haya alcanzado el valor de aptitud predefinido [Liu 2008].

Este algoritmo se programó para resolver y encontrar un valor desconocido, su propósito era encontrar el valor máximo de la función objetivo minimizado a cero y, por lo tanto, a partir de un número aleatorio, el algoritmo traza una tendencia de comportamiento entre los valores de entrada optimizados con respecto al valor de T_{salida} . La eficiencia de un AG dependía en gran medida de sus parámetros de ajuste; las combinaciones para controlar los parámetros utilizados para ejecutar AG se muestran en la Tabla 4.5.

Tabla 4.5. Parámetros usados para la corrida de Algoritmos Genéticos (AG).

Algoritmo	Parámetros	Valor
AG	Creación de población	Aleatorio
	Número de generaciones	100
	Función escalada	Rango
	Función de selección	Rueda de ruleta
	Probabilidad de reproducción de cruce	0.2
	Función de mutación	Uniforme
	Tasa de mutación	0.2
	Función de cruce	Dispersa

El modelo ANNi se aplicó a tres pruebas experimentales utilizadas al azar para determinar las condiciones óptimas de $T_{almacén}$, con respecto a un valor deseado de T_{salida} , como se muestra en la Tabla 4.6. ANNi mostró una

tendencia a la cual, de acuerdo con un valor óptimo $T_{almacén}$, se logró un valor máximo de T_{salida} .

Tabla 4.6. Pruebas experimentales aplicadas al modelo ANNi-AG.

VARIABLES EXPERIMENTALES	1ra. prueba	2da. prueba	3ra. prueba
Entradas:			
Mes	Mayo	Junio	Enero
Día	28	20	22
Hora	55193	50400	42648
Temp. Entrada	43.66	40.76	34.06
Temp. Ambiente	63.43	26.04	39.97
Temp. Superficie Vidrio	48.17	46.70	32.06
Temp. Reflector	75.33	69.27	33.51
Temp. Suministro	Variable a optimizar		
Temp. Almacenamiento	24.35	24.32	19.62
Salida/Target:			
Temp. Salida	Buscar el valor máximo		

Una vez resuelta la Ecuación 4.6 para las tres pruebas experimentales de AG, mostramos en la Figura 4.8 el valor máximo de T_{salida} que fue en Mayo, alcanzando un valor para la salida del colector de 49°C, seguido de Junio con 44.5°C y Enero con 32.6°C, con entradas óptimas en el tanque de alimentación de 19°C, 20°C y 13°C, respectivamente. Estas condiciones predichas por el modelo ANNi eran lógicas, ya que en el mes de Mayo fue donde se produce la mayor parte de la temperatura ambiente en el área y, por lo tanto, es consecuencia directa de que hay más radiación solar. Por otro lado, en Enero la temperatura en invierno era más baja. En consecuencia, en las estaciones cálidas (primavera y verano) el valor de T_{salida} será alto, pero en las estaciones de invierno (otoño e invierno), el valor de T_{salida} será bajo.

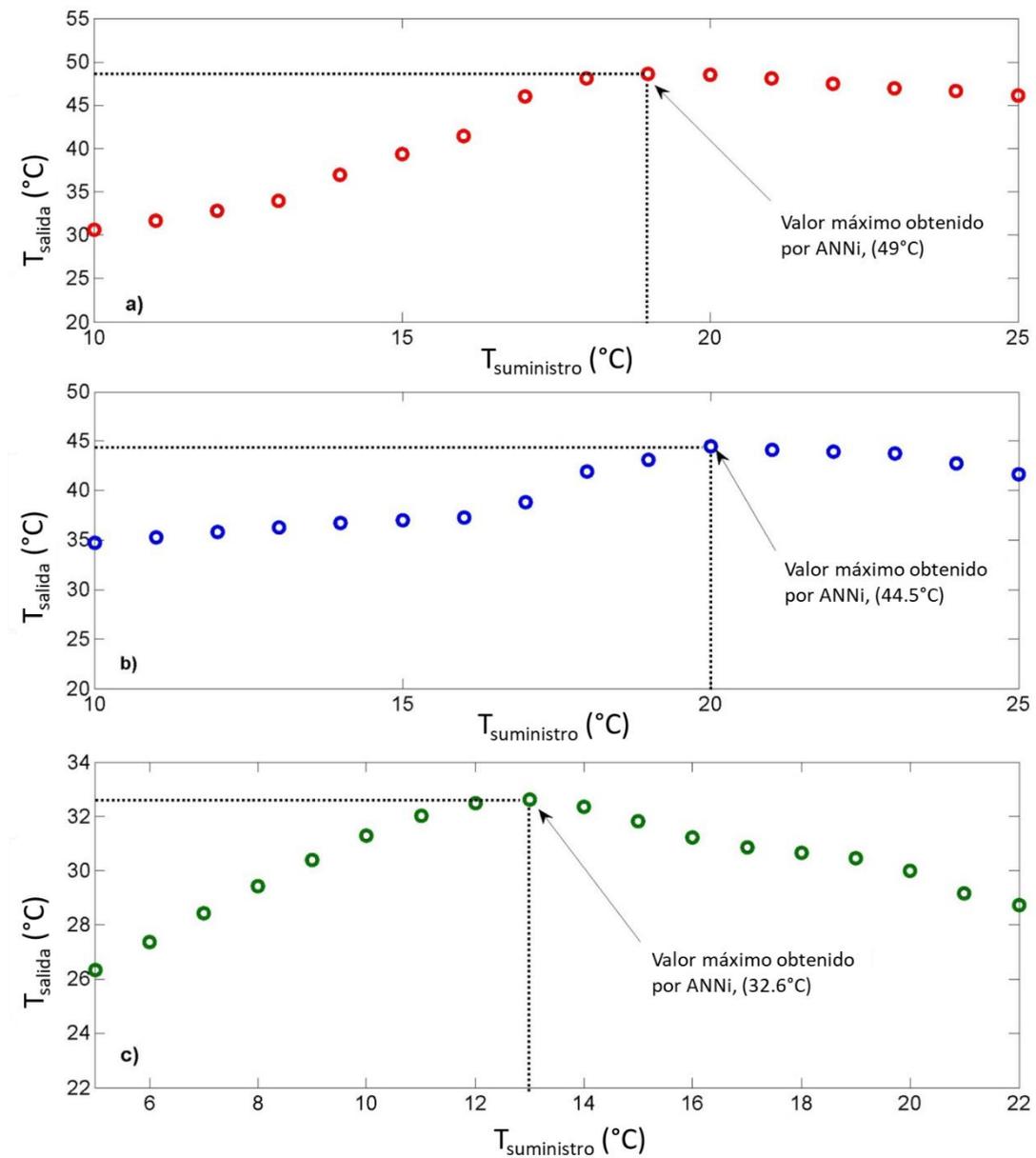


Figura 4.8. Condiciones óptimas de la $T_{suministro}$ en el PTC con respecto a la T_{salida} simulada mediante ANNi, para el mes de a) Mayo, b) Junio y c) Enero.

Ahora, el uso de bajas temperaturas en el tanque de alimentación (menos de 20°C) demuestra la capacidad de PTC para calentar el agua a partir de bajas temperaturas.

4.6 Comparación del PTC de bajo costo

Finalmente, en la Tabla 4.7 se ofrece una comparación entre las ventajas y desventajas del PTC estudiado y su metodología de simulación-optimización con respecto a otros tipos de colectores.

Tabla 4.7. Comparación entre el PTC de bajo costo con respecto otros tipos de colectores.

Referencia del equipo experimental	Tipo de colector	Ventaja	Desventaja
[Kaushika y Reddy 2000]	Disco parabólico	Alta eficiencia, temperatura más alta en un solo punto. Eficiencia de la conversión del vapor a 450°C es del 70-80%	El sistema de tracking usado resulta caro, el costo total del colector es de 38000 Rupias (568,48 USD) (fue construido en 1992), actualmente \$1USD≈71 INR
[Sobhansarbandi y Atikol 2015]	Colector de Placa Plana (FPC) y Colector Parabólico Compuesto (CPC)	La temperatura de salida del fluido del FPC está entre 25-75 °C, mientras que para el CPC está en un intervalo de entre 25-95°C en invierno. Las simulaciones sugieren que un CPC de 2m ² puede igualar la demanda de uso doméstico que un FPC de 8 m ² .	Ambos equipos están modelados mediante ecuaciones matemáticas para simular las condiciones climáticas de invierno. Ambos modelos no proporcionan un análisis económico de los materiales utilizados. Ambos sobrepasan la temperatura deseada (aprox. 24°C), lo cual requiere de un control.
[Hsieh 1981]	Colector Parabólico Compuesto (CPC)	Buena eficiencia (aprox. 56%). Temperaturas de salida de 50 hasta 140° con temperaturas de entrada de 20 hasta 120. La curva de eficiencia para el CPC tiene una pendiente mucho menor que la de un colector de placa plana, lo que hace que el CPC sea	El modelo no compara sus resultados obtenidos con respecto al equipo experimental, probando solo bajo las condiciones del modelo. No se contempla un análisis económico de los materiales utilizados, por lo tanto, el equipo podría ser costoso



		especialmente adecuado para aplicaciones de alta temperatura.	(enchaquetado del receptor con un recubrimiento selectivo).
[Bansal y Uhlemann 1984]	Colector Solar poroso para calentamiento de aire	<p>Eficiencias de hasta 60%,</p> <p>Diferencias de temperaturas de la entrada con respecto a la salida de hasta 23°C.</p> <p>Es muy económico y con larga vida, lo que lo hace rentable.</p> <p>Se usa un absorbedor textil (100% poliestireno).</p>	<p>La eficiencia del colector es notablemente reducida por el tamaño del poro del absorbedor textil.</p> <p>El aislamiento trasero debe ser el idóneo para que no haya pérdidas térmicas significativas (hay mejor eficiencia con un aislamiento de 6 cm que con uno de 3 cm)</p>
Trabajo actual	Colector Parabólico Compuesto (CPC)	<p>El tubo de cobre resulta ser un material muy económico y con una buena conductividad. Costo del total del colector es \$292 USD (2013).</p> <p>Las propiedades térmicas del acero inoxidable con pulimento espejo lograron una mejor eficiencia durante todo el día, ya que afecta directamente la mayor absorción de la luz directa y difusa.</p> <p>El modelo ANNi se utilizó para optimizar las variables de entrada de acuerdo con los datos experimentales del PTC, a fin de lograr controlar las variables de los equipos y, a su vez, mejorar la temperatura de salida.</p>	<p>La eficiencia del colector no es muy alta.</p> <p>El modelo ANN y LA ANNi están limitadas para predecir y optimizar otras variables solo del PTC estudiado, es necesario la adquisición de más datos provenientes de otros tipos de colector solar para tener una mejor metodología.</p>

4.7 Conclusiones parciales

Se presentó un colector de canal parabólico con componentes de bajo costo como una posible opción para usar en áreas residenciales, con una buena estabilidad durante la operación y una buena característica en los materiales que le permiten durar varios años, con un mínimo de mantenimiento.

El PTC se modeló para predecir el T_{salida} mediante un modelo ANN, con una estructura de 9 neuronas en la capa de entrada, 9 neuronas en la capa oculta y 1 neurona en la capa de salida. Este modelo fue entrenado con una amplia gama de datos experimentales, utilizando como variables de entrada: la temperatura del tanque de alimentación ($T_{suministro}$), la temperatura del tanque de almacenamiento del agua caliente ($T_{almacén}$), la temperatura de entrada del agua al colector solar ($T_{entrada}$), mes (m), día (d), tiempo total (s), temperatura ambiente (T_{amb}), temperatura de la superficie al colector solar de vidrio (T_{vidrio}) y la temperatura de la superficie del reflector ($T_{reflector}$). Los resultados del modelo ANN mostraron una buena relación con los datos experimentales que pasaron satisfactoriamente las pruebas de significancia (prueba Fisher y t de Student) y un valor más bajo de $RMSE$ de 0.8055 y 0.57 durante el entrenamiento.

En consecuencia, utilizando los pesos y las bias del modelo ANN, se desarrolló una estrategia para optimizar las variables de entrada para aumentar el valor de T_{salida} . Una vez planteado el modelo ANNi, la función objetivo resultante se resuelve mediante algoritmos genéticos, mostrando un aumento en T_{salida} de 49°C para Mayo, encontrando las condiciones óptimas en la temperatura del tanque de alimentación a 19°C, condición adecuada para regiones donde se alcanzó una temperatura ambiente alta.

Con este parámetro optimizado, permite que el PTC funcione de manera eficiente. Para ajustar la temperatura del tanque de alimentación ($T_{suministro}$),



era necesario regular la válvula que recircula el agua desde el tanque de almacenamiento (agua caliente) hacia el tanque de alimentación (agua fría).

Finalmente, esta temperatura máxima alcanzada por el PTC, utilizando la metodología de ANNi, puede satisfacer las necesidades básicas de las áreas residenciales, debido a que, con esta temperatura de salida de agua caliente, se podrían realizar actividades como: ducharse, precalentar alimentos, o para piscinas.

Las combinaciones de métodos computacionales descritos en este documento fueron una alternativa que nos permite controlar el proceso en línea (ANN), y también optimizar cualquiera de las variables de entrada (ANNi). La aplicación de AG fue útil para la búsqueda de un valor desconocido a partir de una población aleatoria y, a su vez, encontrar un valor máximo.



5 Análisis de funciones de transferencia y normalizaciones en un modelo ANN que predice el transporte de energía en un colector solar de canal parabólico con tubo absorbedor ranurado

El presente trabajo tuvo como objetivo desarrollar un modelo preciso basado en ANN para determinar la temperatura de salida del fluido en un colector solar de canal parabólico (PTC), que utilizaba un tubo ranurado como tubo absorbedor. El modelo de precisión se llevó a cabo considerando el análisis en diez modelos ANN a los cuales se les aplicaron diferentes funciones de transferencia, así como diferentes normalizaciones. Las principales contribuciones de este trabajo han sido: (1) el desarrollo de un modelo basado en ANN para determinar la temperatura de salida del fluido del PTC con un tubo absorbedor ranurado. (2) La evaluación de diferentes combinaciones de funciones de activación y normalizaciones para obtener un modelo óptimo con el mínimo error. Y (3) el análisis estadístico para determinar el modelo que tuvo el mejor rendimiento.

5.1 Descripción del equipo experimental

El PTC es el mismo presentado en la Sección 3.1 pero con un tubo receptor con ranuras internas para aumentar la transferencia de calor. El tubo receptor está hecho de un tubo de cobre ranurado de $\frac{1}{2}$ pulgada de largo y pintado de negro para tener una mayor capacidad de absorción. Las características del PTC y el absorbedor ranurado se muestran en la Tabla 5.1 y la Figura 5.1, respectivamente.

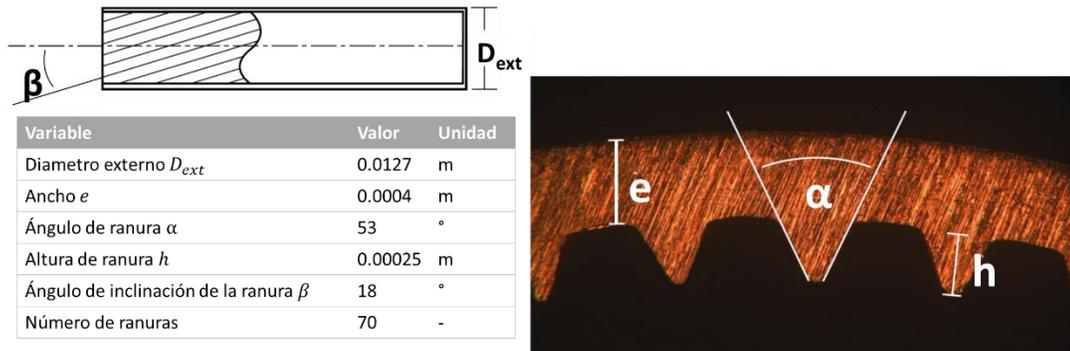


Figura 5.1. Características geométricas del tubo raurado.

Tabla 5.1. Características geométricas del sistema experimental PTC.

Variable	Valor	Unidad
Longitud del tubo receptor	0.85	m
Diámetro interno del receptor	0.0123	m
Diámetro externo del receptor	0.0127	m
Diámetro interno de la 1ra. cubierta, C1	0.034	m
Diámetro externo de la 1ra. cubierta, C1	0.036	m
Diámetro interno de la 2da. cubierta, C2	0.048	m
Diámetro externo de la 2da. cubierta, C2	0.05	m
Absorbancia del receptor	0.906	-
Emitancia del receptor	0.14	-
Transmitancia de la cubierta de vidrio	0.95	-
Superficie reflectiva	0.93	-
Factor de forma	0.92	-
Ángulo de inclinación	19	°
Ángulo de incidencia modificado	1	-
Eficiencia óptica	0.56	-
Distancia Focal	0.025	m
Emitancia de cubierta (vidrio ventana)	0.95	-
Factor de vista	1	-

5.2 Condiciones de operación

El PTC está ubicado en las instalaciones del CIICAp-UAEM Morelos, México, con coordenadas de latitud y longitud de 18.981655 -99.23418, respectivamente. El PTC tiene una inclinación de 19 ° y está orientado hacia el Este-Oeste con la cara de la placa reflectante apuntando hacia el sur para aprovechar mejor la incidencia de la radiación solar, como se había

mencionado en la Sección 3.2.1. La Figura 5.2 muestra un diagrama esquemático del sistema experimental PTC con componentes de bajo costo.

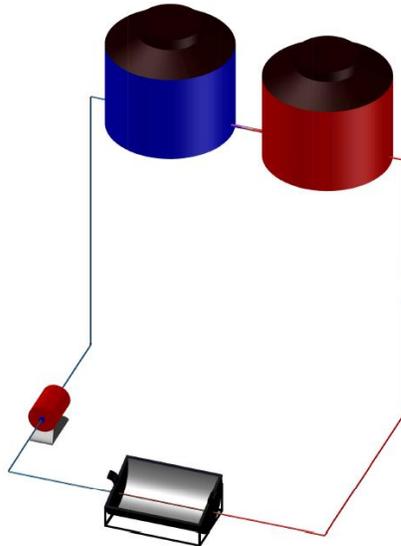


Figura 5.2. Diagrama esquemático del sistema experimental PTC

Como se puede ver en la Figura 5.3, se realizaron 4 configuraciones en la disposición de PTC: a) con cubierta de vidrio que protege al PTC del clima y el polvo, b) sin cubierta de vidrio para evitar pérdidas ópticas, c) con tapa de vidrio y tubo receptor ubicado en una segunda línea focal, optimizada mediante trazados de rayos láser [Ramírez-Ramos 2017], y d) sin tapa de vidrio y tubo receptor situado en la segunda línea focal.

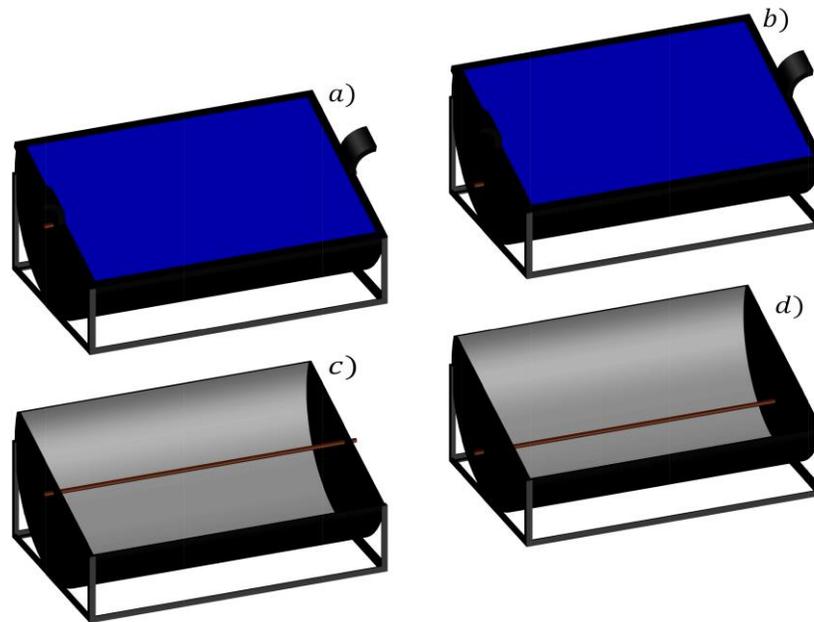


Figura 5.3. Diferentes configuraciones del sistema experimental PTC a) Con tapa, b) con tapa y distancia focal optimizada, C) sin tapa, y d) sin tapa y distancia focal optimizada.

Las pruebas experimentales se llevaron a cabo a diferentes flujos volumétricos: 0.7, 1, 2, 3, 4 y 5 LPM utilizando agua como fluido de trabajo. Se registraron un total de 1,155 muestras durante Octubre y Noviembre de 10:00 a 18:00 h, con intervalos de 5 minutos entre mediciones. La Tabla 5.2 muestra los intervalos operativos a los que se sometió el sistema PTC.

Tabla 5.2. Condiciones de operación del sistema experimental PTC.

Variabes experimentales	Intervalo de trabajo
Entradas	
Día del año	287-318
Flujo (LPM)	0.7-5
Configuración	1-4
Hora (min)	595-1085
Irradiancia (W/m^2)	2-1043
$T_{entrada}$ ($^{\circ}C$)	20.01-38.11
T_{amb} ($^{\circ}C$)	19.27-39.02
$T_{suministro}$ ($^{\circ}C$)	18.18-28.95
$T_{almacén}$ ($^{\circ}C$)	18.75-30.72
Salida	
T_{salida} ($^{\circ}C$)	20.02-41.99

Los datos de irradiancia fueron recolectados usando un piranómetro ubicado en una Estación Meteorológica [SMN 2016]. El cálculo del día del año se realizó con base en la Tabla 5.3, donde i es el calendario del día del mes, y la hora del día se estableció en minutos.

Tabla 5.3. Calcular el día del año.

Mes	Día del año
Enero	i
Febrero	$31 + i$
Marzo	$59 + i$
Abril	$90 + i$
Mayo	$120 + i$
Junio	$151 + i$
Julio	$181 + i$
Agosto	$212 + i$
Septiembre	$243 + i$
Octubre	$273 + i$
Noviembre	$304 + i$
Diciembre	$334 + i$

5.3 Desarrollo del modelo de red neuronal

Las redes neuronales artificiales ANN son algoritmos computacionales inspirados en el proceso de aprendizaje del cerebro humano. Al igual que una red neuronal biológica, los ANN están formados por unidades de procesamiento en paralelo llamadas neuronas, que generan y envían señales de información para producir una salida de respuesta específica [Parrales *et al.* 2018]. En diferentes campos del conocimiento, se ha definido como una herramienta de modelado matemático eficiente de problemas de regresión multivariante lineal y no lineal [Khayet y Cojocarú 2012].

A partir de una cierta cantidad de información, los ANN tienen la capacidad de discernir la relación entre las variables de un proceso y reconocer patrones de comportamiento cuando son difíciles de describir matemáticamente [Khataee

y Kasiri 2010]. Son útiles cuando los datos provienen de procesos no lineales, que inherentemente tienen ruido; específicamente, los datos provienen de sistemas de energía renovable, ya que es difícil identificar una función matemática precisa cuando algunos parámetros varían según el tiempo y los factores ambientales [Kalogirou 2001].

Desarrollar un modelo ANN requiere tres tareas fundamentales. La primera tarea es la creación de una base de datos de trabajo proporcionada por la medición de variables de operación a través de equipos de instrumentación. Es esencial mencionar que las variables consideradas para el modelo deben ser preferiblemente representativas del fenómeno para reducir el tamaño de la red y evitar la redundancia en los datos de capacitación. Además, se recomienda su análisis con procesos estadísticos [Abiodun *et al.* 2018].

La segunda tarea es obtener una configuración de red adecuada para alcanzar una predicción precisa. En este paso, el tamaño y la normalización del conjunto de datos, el número de neuronas ocultas, el algoritmo de entrenamiento y la función de activación son parámetros esenciales a considerar. Para que luego, en la tercera tarea, se establezca la correlación entre las variables obtenidas y las experimentales.

Tres capas caracterizan principalmente la configuración del modelo de red neuronal. La primera capa, conocida como la capa de entrada, contiene las variables que se considerarán como neuronas de entrada en el modelo. Se debe aplicar un proceso de normalización o estandarización a los datos de entrada para ajustarlos en la misma escala. Este proceso permite comparar sus proporciones en lugar de sus magnitudes y ofrece una disminución considerable en el tiempo de cálculo [Martínez-Martínez *et al.* 2017].

La normalización de la base de datos evita la saturación matemática ya que las funciones de transferencia sigmoideas tienden a saturarse cuando se presentan valores superiores a +3 o inferiores a -3. Al tener un alto número

de valores en las variables de entrada, la cantidad asignada a los pesos debe ser mínima, lo que haría que el entrenamiento sea muy lento [Millán-Ocampo *et al.* 2018]. El intervalo de normalización de las variables de entrada debe estar relacionado con la función de activación que se decide utilizar. Por lo general, se realiza en intervalos de -1 a 1 , para lo cual existen diferentes ecuaciones en la literatura.

La segunda capa es la capa oculta, donde hay una o más unidades de procesamiento llamadas neuronas ocultas. La predicción precisa del modelo ANN está determinada por el número de estas unidades asociadas con su función de activación. Por último, en la capa de salida, a través de una matriz de pesos y bias, se interconectarán todas las neuronas del modelo para generar las variables a predecir.

Para cada neurona oculta se requiere el término de entrada de red n_j que consiste en un grupo de enlaces de conexión llamados sinapsis; donde cada uno de estos tiene un valor de conexión ajustable, llamado peso sináptico W_i , que puede ser positivo si la neurona aumenta el estado de activación de la neurona adyacente o negativa si la inhibe. El producto de peso y su respectiva entrada P se suman para cada neurona agregando un valor de compensación b conocido como bias [Han *et al.* 2012], como se mostró en las ecuaciones **2.1** y **3.3.a**, de las Secciones **2.3.2** y **3.3**, respectivamente:

$$n_j = W_{i(J,1)} \cdot P_{(1)} + W_{i(J,2)} \cdot P_{(2)} + \dots + W_{i(J,R)} \cdot P_{(R)} + b_{(1,J)} \quad (5.1)$$

Donde W_i son los coeficientes de los pesos de conexión entre la capa de entrada y la capa oculta, P son las variables de entrada, J es el número de neuronas en la capa oculta, R es el número de neuronas de entrada y b es la bia correspondiente a cada neurona en la capa oculta.

Se aplica una función de activación o transferencia a n_j para determinar la magnitud de la respuesta de cada neurona oculta. Existen diferentes funciones de transferencia o activación como sigmoide de límite duro, lineal, sigmoide

logarítmico, sigmoide tangente hiperbólico, gaussiano, entre otras. Las funciones más destacadas para la capa oculta son la función tangencial-sigmoidal (*TANSIG*, τ) y la función logarítmica-sigmoidal (*LOGSIG*, δ). Sus ecuaciones se muestran en la Tabla 5.4. Normalmente, se utiliza una función de transferencia lineal (*PURELIN*, λ) para la capa de salida [Bassam *et al* 2014]. Un modelo ANN generalmente se representa de acuerdo con la función de activación seleccionada.

Tabla 5.4. Ecuaciones de las funciones de transferencia *TANSIG* y *LOGSIG*.

Función de activación	TANSIG	LOGSIG
Forma matemática	$\tau = \frac{e^{n_j} - e^{-n_j}}{e^{n_j} + e^{-n_j}} = \frac{2}{1 + e^{-2n_j}} - 1$	$\delta = \frac{1}{1 + e^{-n_j}}$
Ecuación general	$\tau(n_j) = \frac{2}{1 + e^{-2\sum_{j=1}^J (W_{i(j,r)} \cdot P_{(r)} + b_{(1,j)})}} - 1$	$\delta(n_j) = \frac{1}{1 + e^{-[\sum_{j=1}^J (W_{i(j,r)} \cdot P_{(r)} + b_{(1,j)})]}}$
Ecuación simplificada de salida	$Output = \lambda(W_{0(K,J)} \cdot \tau(n_j) + b_{(2,K)})$	$Output = \delta(W_{0(K,J)} \cdot \delta(n_j) + b_{(2,K)})$
Ecuación de salida general	$Output = \sum_{k=1}^K \left\{ W_{0(K,J)} \left[\frac{2}{1 + e^{-2\sum_{j=1}^J (W_{i(j,r)} \cdot P_{(r)} + b_{(1,j)})}} - 1 \right] \right\} + b_{(2,K)}$	$Output = \sum_{k=1}^K \left\{ W_{0(K,J)} \left[\frac{1}{1 + e^{-[\sum_{j=1}^J (W_{i(j,r)} \cdot P_{(r)} + b_{(1,j)})]}} \right] \right\} + b_{(2,K)}$

Para el proceso de aprendizaje, la arquitectura más común utilizada ha sido el perceptrón multicapa entrenado con el algoritmo de entrenamiento de propagación hacia atrás Levenberg-Marquardt en una red multicapa FeedForward [Rocabruno-Valdés *et al.* 2015], que se ha recomendado para el análisis del intercambiador de calor [Hagan *et al.* 2014] y el rendimiento de un colector solar [Mohanraj *et al.* 2015]

5.3.1 Análisis estadístico

Para validar la precisión y la adaptabilidad del modelo propuesto, se realizan pruebas estadísticas para analizar los datos experimentales y los datos simulados obtenidos. Los parámetros de prueba que se usan comúnmente están descritos en la Sección 2.3.9 de esta tesis.

5.3.2 Pruebas de significancia

Las pruebas de significación se aplican a la relación de bases de datos experimentales y simuladas para corroborar el ajuste del modelo. La prueba F o la prueba de Fisher y la prueba t de Student son las pruebas de significancia más utilizadas para aceptar o rechazar la hipótesis nula (H_0), y están descritas en la Sección 2.3.10.

5.3.3 Prueba de linealidad

La prueba de linealidad pendiente-intersección establece los límites de confianza de la pendiente y la intersección que deberían ser idealmente uno y cero para la pendiente y la intersección, respectivamente [Rocabruno-Valdés 2019]. Para realizar esta prueba consultar la Sección 2.3.11.

5.3.4 Análisis de sensibilidad de variables

La ecuación de Garson [Parrales *et al.* 2019] se usa para determinar la importancia de las variables de entrada dentro del modelo obtenido mediante ANN. Esta ecuación está descrita en la Sección 2.3.12.

5.4 Modelo de red neuronal

Para que un modelo de red neuronal aprenda a calcular la variable de salida con alta precisión es necesario establecer la topología o arquitectura con precisión. La normalización de los datos de entrada, el número de neuronas en la capa oculta y el grado de conectividad obtenido a través de la función de transferencia son parámetros esenciales para determinar.

Al igual que una neurona biológica, la neurona artificial calcula su grado de conectividad a partir de la función de activación. En términos simples, una función de transferencia o activación de una neurona representa la velocidad potencial a la que se activa o no una neurona. Su propósito principal es relacionar una respuesta de salida con las señales de entrada que varían en el intervalo de -1 a 1 , dependiendo de la función de transferencia

seleccionada. Por lo tanto, el intervalo de normalización de las variables de entrada debe estar relacionado con la función de activación que se decide utilizar.

Como se mencionó anteriormente, el proceso de normalización evita la saturación matemática cuando se utilizan funciones de transferencia. Cuando los valores de las variables de entrada son muy altos, los valores de peso obtenidos serán muy pequeños. Por lo tanto, el gradiente descendente en el proceso será corto, lo que lleva a un entrenamiento prolongado [Millán-Ocampo *et al.* 2018]. Este comportamiento es independiente de los que tienen las bias, ya que estas solo actúan como un valor de compensación o contrapeso para la red.

En la literatura, se han aplicado varias ecuaciones de normalizaciones para algunas funciones de transferencia. En un artículo publicado por Hernández *et al.* [Hernández y Colorado 2013], se evaluaron diferentes funciones de transferencia en intervalos variables de normalización en un transformador de calor de absorción con reciclaje de energía. Se determinó que la función *TANSIG* con un intervalo de normalización $[0,1]$ mostró mejores resultados en comparación con otros. Sin embargo, existe cierta incertidumbre al elegir la función de transferencia adecuada y el proceso de normalización, así como el número óptimo de neuronas en la capa oculta. Por lo tanto, en este trabajo, para discernir la mejor topología, fue necesario hacer un análisis de las funciones de transferencia a diferentes intervalos de normalización y encontrar el mejor modelo capaz de estimar la temperatura de salida de PTC.

Se desarrolló un primer modelo de red que se utilizó como modelo base a partir de 1,155 datos experimentales concordantes. Se seleccionaron nueve variables, que se consideraban significativas para nuestro sistema PTC. Las variables de entrada fueron: la temperatura de entrada del fluido, temperatura ambiente, irradiancia, hora, día, configuración, flujo volumétrico, temperatura de alimentación y temperatura de almacenamiento.

Este modelo fue entrenado con el 50% de los datos, validado con el 25% y probado con el 25% restante, basado en lo planteado en la Sección 2.3.7. Se usó una arquitectura de propagación hacia atrás de múltiples capas con nueve neuronas de entrada en un intervalo de normalización de -1 a 1 , el intervalo especificado para la función *TANSIG*. La forma general de la ecuación de normalización se representa de la siguiente manera:

$$x_{i,norm} = (y_{max} - y_{min}) \cdot \left(\frac{x_i - x_{min}}{x_{max} - x_{min}} \right) + y_{min} \quad (5.2)$$

Donde y_{max} e y_{min} son los valores límite del intervalo de normalización. Para este modelo, los valores de y_{max} e y_{min} fueron 1 y -1 , respectivamente. Una vez que se obtuvo el modelo neuronal, la comparación entre los datos simulados y los datos experimentales se realizó utilizando las pruebas estadísticas R^2 , *MAPE* y *RMSE* descritas anteriormente.

5.5 Resultados

La Tabla 5.5 muestra la comparación del análisis estadístico para determinar el número de neuronas en la capa oculta. Se observa que tanto el *RMSE* como el *MAPE* muestran una ligera disminución para las configuraciones de 5 y 6 neuronas. Por lo tanto, continuar aumentando el número de neuronas en la capa oculta reduciría en valores mínimos el error y aumentaría el número de coeficientes en la ecuación de salida, lo que genera un ajuste excesivo. La Figura 5.4 muestra la arquitectura empleada en el modelo neural de T_{salida} con seis neuronas ocultas.

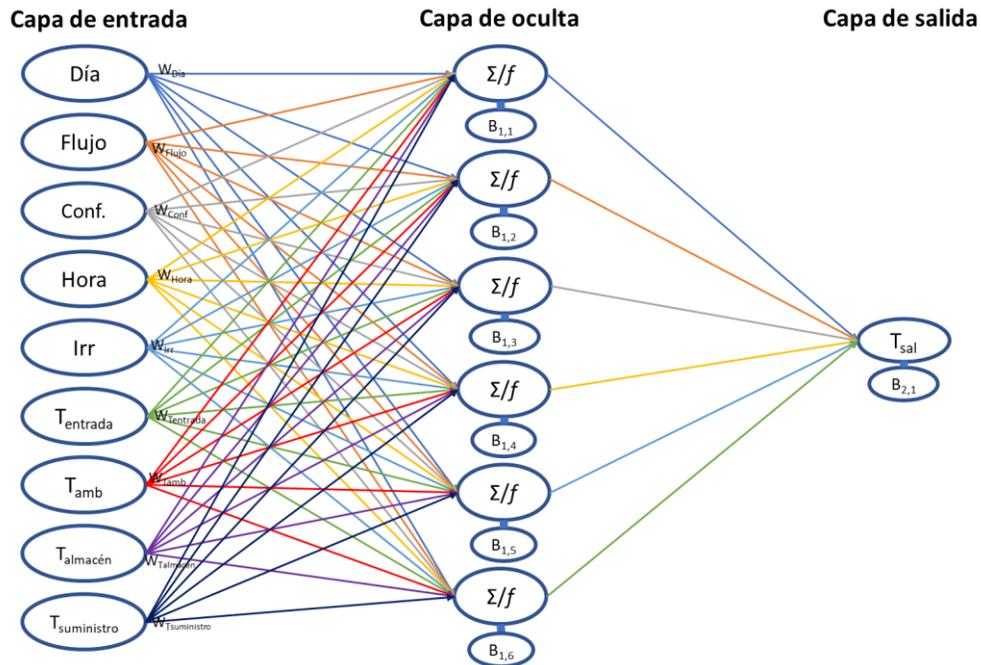


Figura 5.4. Arquitectura del modelo neuronal del PTC con nueve entradas.

Tabla 5.5. Análisis estadístico del número de neuronas en la capa oculta para el modelo base ANN.

Número de neuronas	R^2	RMSE	MAPE
1	0.970366963	0.411245901	0.022058509
2	0.977060605	0.361024753	0.028778328
3	0.984686936	0.294146581	0.040628118
4	0.988624436	0.253367799	0.019870286
5	0.995204523	0.164814324	0.009307217
6	0.996191094	0.147062448	0.007954328

Además, las pruebas de significación de la prueba F y la prueba t se realizaron en el modelo ANN. Se seleccionó un nivel de significancia de 0.01 (99% de confianza) para indicar un error aceptable del 1%, y se determinó el valor p para establecer el nivel mínimo de significancia para aceptar el modelo. La Tabla 5.6 muestra los valores obtenidos de las pruebas F de Fisher, t-student y el p-valor.

Tabla 5.6. Pruebas de significancia del modelo base.

	Prueba Fisher	Prueba t de Student
Valor calculado (Eq. 2.11 & Eq. 2.12)	1.00583512	0.44966639
Valor crítico (de tablas)	1.14688062	2.32958361
Si valor calculado < valor crítico	Si	Si
p-valor		0.92129576

La prueba de pendiente-intercepto se realizó para verificar que los datos simulados están dentro del intervalo definido por los límites del objetivo y no fuera de él. Los valores de la prueba pendiente-intercepto se presentan en la Figura 5.5, donde se ratifica el modelo base. Donde la intersección inferior debe ser menor que cero, y la intersección superior debe ser mayor que cero. Del mismo modo, la pendiente inferior y superior debe ser menor y mayor que 1, respectivamente. Al promediar los valores resultantes deben estar cerca de cero para la intersección y uno para la pendiente.

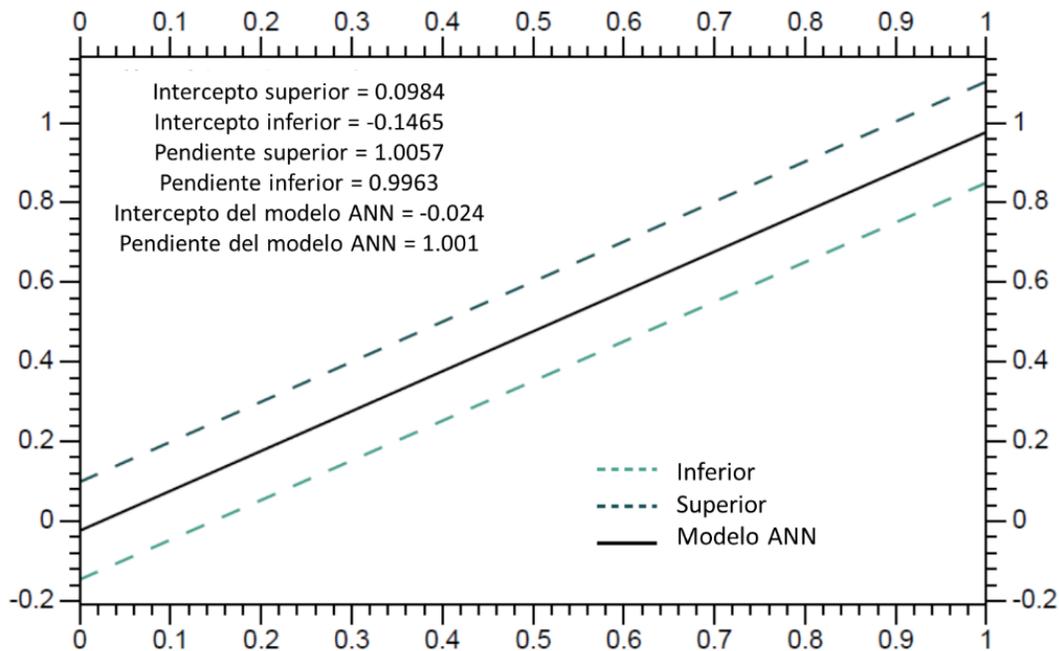


Figura 5.5. Prueba de linealidad del modelo base ANN.

La Figura 5.6 muestra la simulación de la temperatura de salida del colector utilizando el modelo base el 20 y 27 de Octubre, donde se observa que el modelo se ajusta satisfactoriamente a los datos experimentales a pesar del ruido en las mediciones.

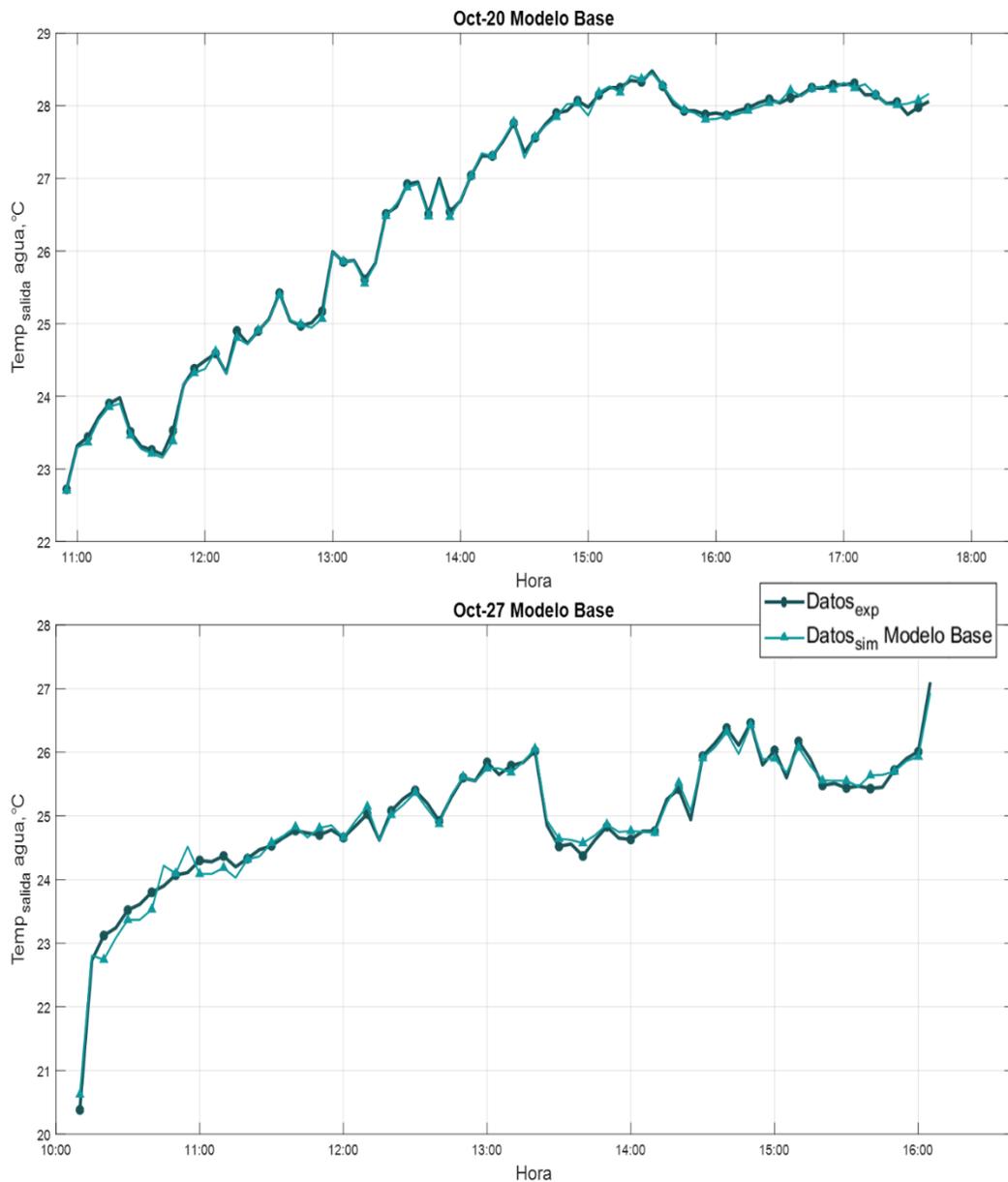


Figura 5.6. Comparación entre los datos experimentales y simulados del modelo base ANN.

Para determinar la influencia del proceso de normalización en las funciones de transferencia, se identificaron cinco áreas de interés para el uso de las funciones *TANSIG* y *LOGSIG*, que son las más empleadas. El proceso de entrenamiento para la comparación se muestra en la Figura 5.7, y los intervalos de normalización analizados se muestran en la Figura 5.8.

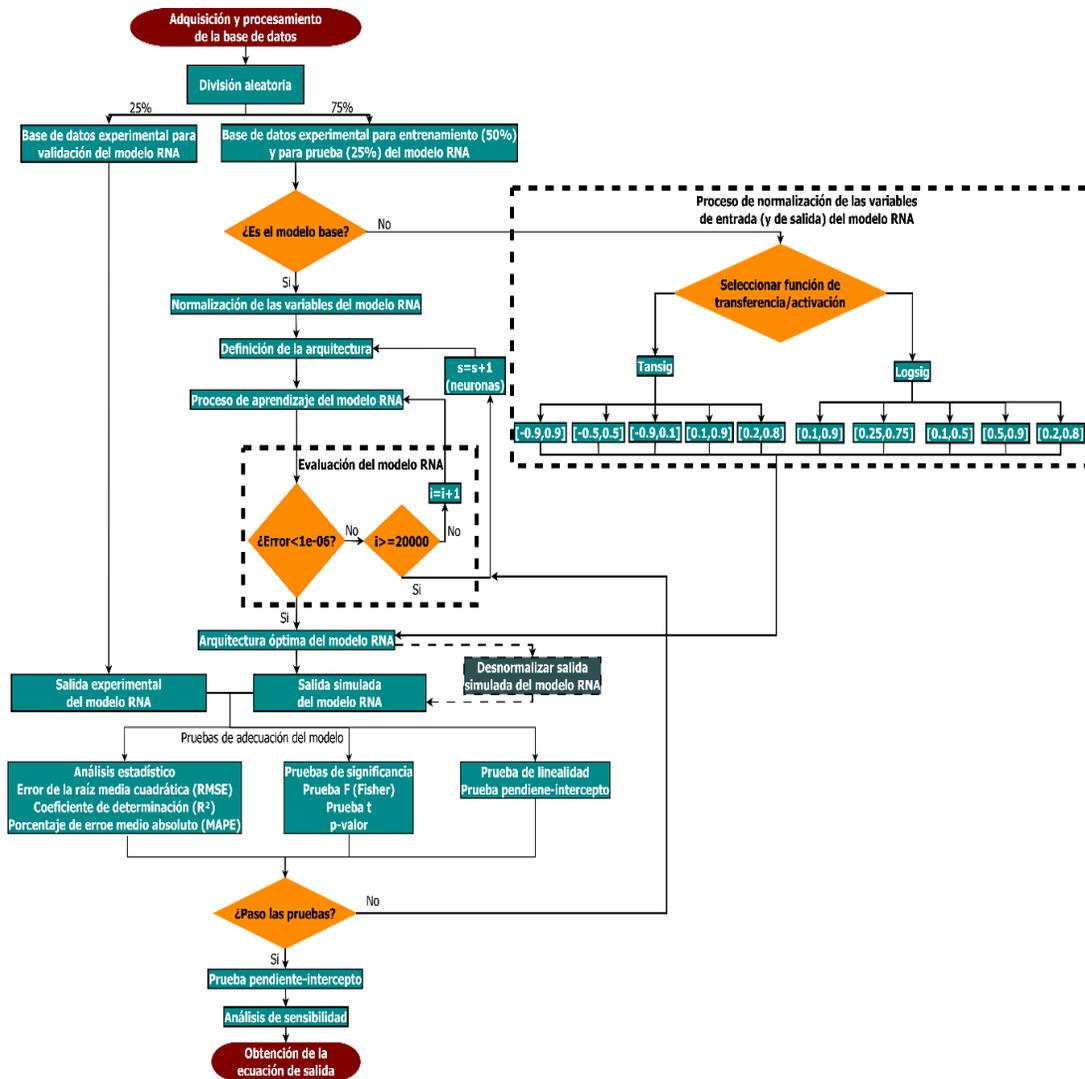


Figura 5.7. Diagrama de flujo para el proceso de aprendizaje de los modelos ANN.

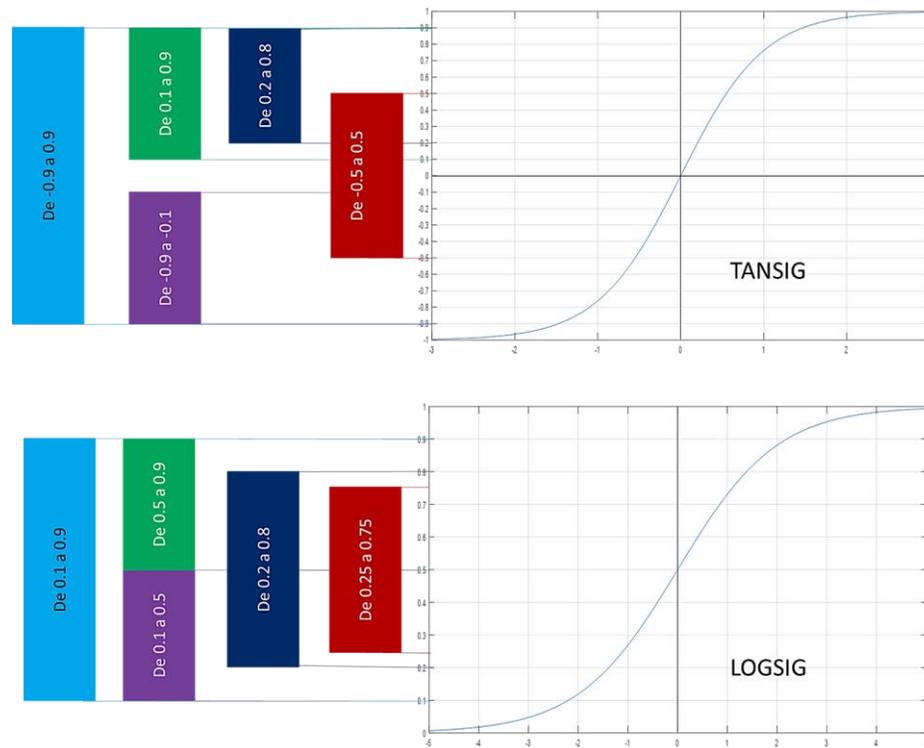


Figura 5.8. Intervalos de normalización para las funciones de transferencia TANSIG y LOGSIG.

En consecuencia, se desarrollaron diez modelos para el análisis de los intervalos de normalización propuestos utilizando la misma arquitectura que el modelo base. Del mismo modo, los parámetros estadísticos se calcularon para validarlos, y los resultados se muestran en la Tabla 5.7 para cada intervalo de normalización. Como puede verse, se obtuvieron valores similares de R^2 , $RMSE$ y $MAPE$ en comparación con el modelo base, e incluso algunos alcanzaron un mejor valor, como fue el caso con la normalización en el intervalo $[0.1, 0.9]$ para TANSIG o con $[0.1, 0.5]$ para LOGSIG.

Tabla 5.7. Análisis estadístico de los modelos ANN con los intervalos de normalización propuestos para TANSIG y LOGSIG.

Análisis estadístico	TANSIG					LOGSIG				
	[-0.9,0.9]	[-0.9,-0.1]	[0.1,0.9]	[0.2,0.8]	[-0.5,0.5]	[0.1,0.9]	[0.1,0.5]	[0.5,0.9]	[0.2,0.8]	[0.25,0.75]
R^2	0.996221316	0.995597257	0.997405465	0.9964935	0.99624876	0.99624533	0.99671914	0.99618892	0.9962321	0.99610716
RMSE	0.146385282	0.157763969	0.121237247	0.14086828	0.145702771	0.14636318	0.1361285	0.14724789	0.14737937	0.14836108
MAPE	0.04043158	0.01724422	5.9341E-05	0.00145046	0.02207994	0.008941	0.00252443	0.0162055	0.02743816	0.00199189

Las pruebas t de Fisher y Student se aplicaron y aprobaron con éxito para la mayoría de los modelos. Los resultados de las pruebas de significación se muestran en la Tabla 5.8. Con respecto a la prueba de pendiente-intercepción, el modelo *TANSIG* con normalización [-0.5, 0.5] y el modelo *LOGSIG* con [0.5, 0.9] no pasó la prueba con ningún nivel de confianza; a pesar de haber elegido 99%, 98% y 95% de confianza. Los resultados de la prueba de pendiente-intersección para todos los modelos se muestran en la Tabla 5.9.

Tabla 5.8. Pruebas de significancia para la comparación de los mismos intervalos de *TANSIG* y *LOGSIG*.

Función	Intervalos	F calculada (Ec. 2.11)	t de Student calculada (Ec. 2.12)	Si	p-valor ($\alpha = 0.01$)
				valor calculado < valor crítico [prueba F, prueba t]	
TANSIG	[-0.9, 0.9]	1.00084628	2.26123776	[si, si]	0.02393009
	[-0.9,-0.1]	1.00043807	1.05231189	[si, si]	0.29287677
	[0.1,0.9]	1.00383619	0.01545336	[si, si]	0.98767317
	[0.2,0.8]	1.00289051	0.05134121	[si, si]	0.95906251
	[-0.5, 0.5]	1.00711859	1.59383554	[si, si]	0.11124671
LOGSIG	[0.1,0.9]	1.00896994	0.65140059	[si, si]	0.51491762
	[0.1,0.5]	1.00028414	0.06101505	[si, si]	0.95135779
	[0.5,0.9]	1.01500877	0.54525118	[si, si]	0.58568614
	[0.2,0.8]	1.01312795	1.76914282	[si, si]	0.07713413
	[0.25,0.75]	1.00173861	0.05659824	[si, si]	0.95487504
Valor crítico (de tablas)		1.14688062	2.32958361	Nivel de significancia 99% ($\alpha = 0.01$)	p-value > α

Tabla 5.9. Valores de la prueba de linealidad para la comparación de los mismos intervalos de *TANSIG* y *LOGSIG*.

Intervalo	TANSIG					LOGSIG				
	[-0.9,0.9]	[-0.9,-0.1]	[0.1,0.9]	[0.2,0.8]	[-0.5,0.5]	[0.1,0.9]	[0.1,0.5]	[0.5,0.9]	[0.2,0.8]	[0.25,0.75]
Intercepto superior	0.191250	0.189170	0.08505	0.12565772	0.25423	0.057478	0.15247	0.36510	0.00948925	0.15183
<i>Intercepto</i>	0.069691	0.057926	-0.015916	0.00833601	0.13348	-0.064292	0.039133	0.24387	-0.11274612	0.028288
Intercepto inferior	-0.051872	-0.073319	-0.11688	-0.1089857	0.01274	-0.18606	-0.074204	0.12264	-0.2349815	-0.095257
Pendiente superior	1.00240	1.00260	1.00450	1.00419279	0.99923	1.0073	1.0029	0.99534	1.0093389	1.00370
<i>Pendiente</i>	0.99769	0.997580	1.00060	0.99968671	0.99459	1.0026	0.9985	0.99069	1.00464409	0.99892
Pendiente inferior	0.99302	0.992540	0.99674	0.99518063	0.98995	0.99791	0.99415	0.98603	0.99994929	0.99417



Como se muestra en la Figura 5.9 y en la Figura 5.10, la comparación entre todos los datos simulados por los modelos propuestos con los datos experimentales revela que, en este caso, la prueba de intercepción de pendiente no tiene importancia para obtener un modelo predictivo preciso de nuestro sistema. Sin embargo, como se puede observar, el modelo que obtuvo el menor error, aprobó todas las pruebas y mostró el mejor ajuste fue el modelo *TANSIG* con normalización [0.1,0.9].

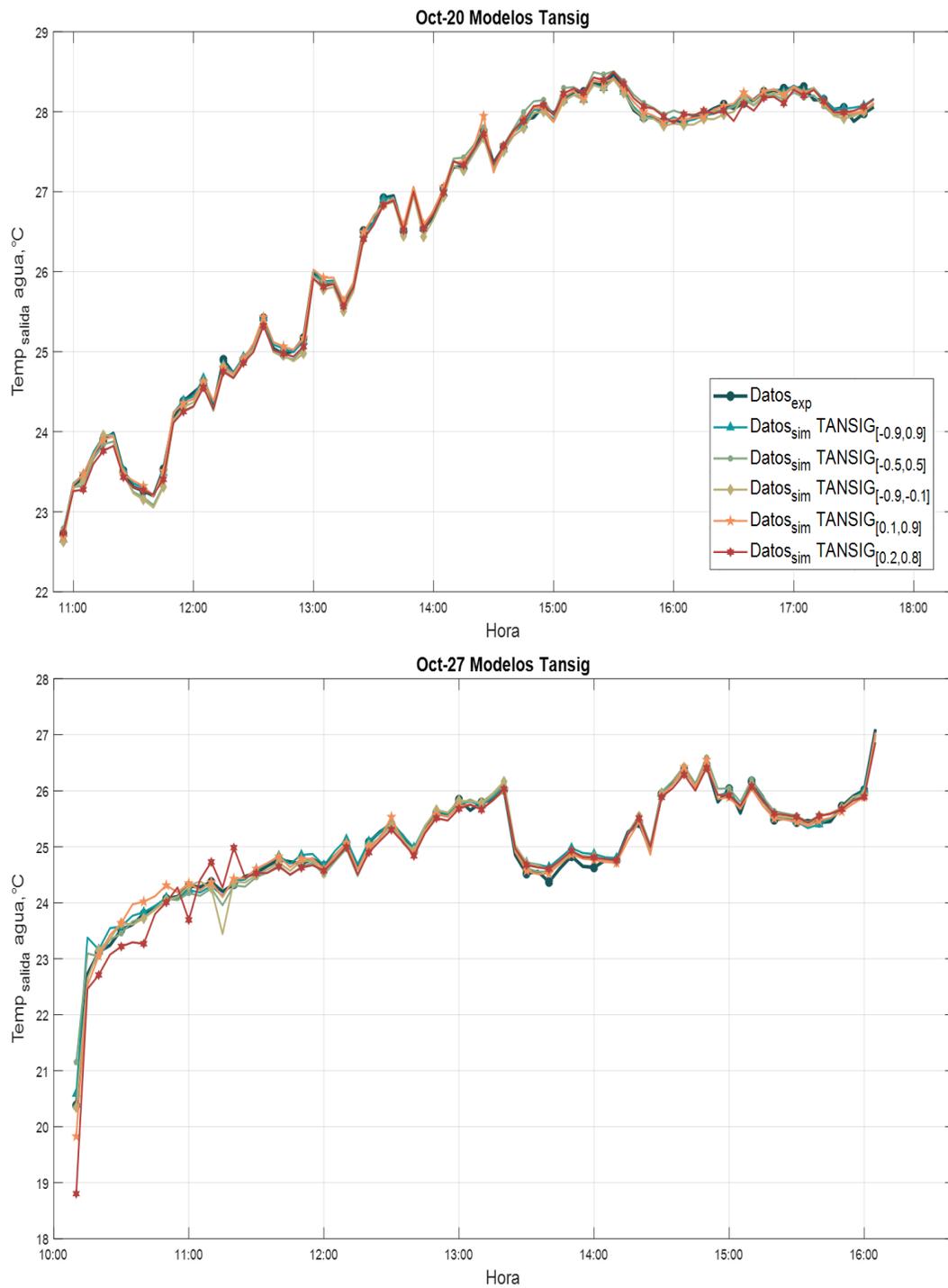


Figura 5.9. Comparación entre los datos experimentales y simulados de los modelos TANSIG.

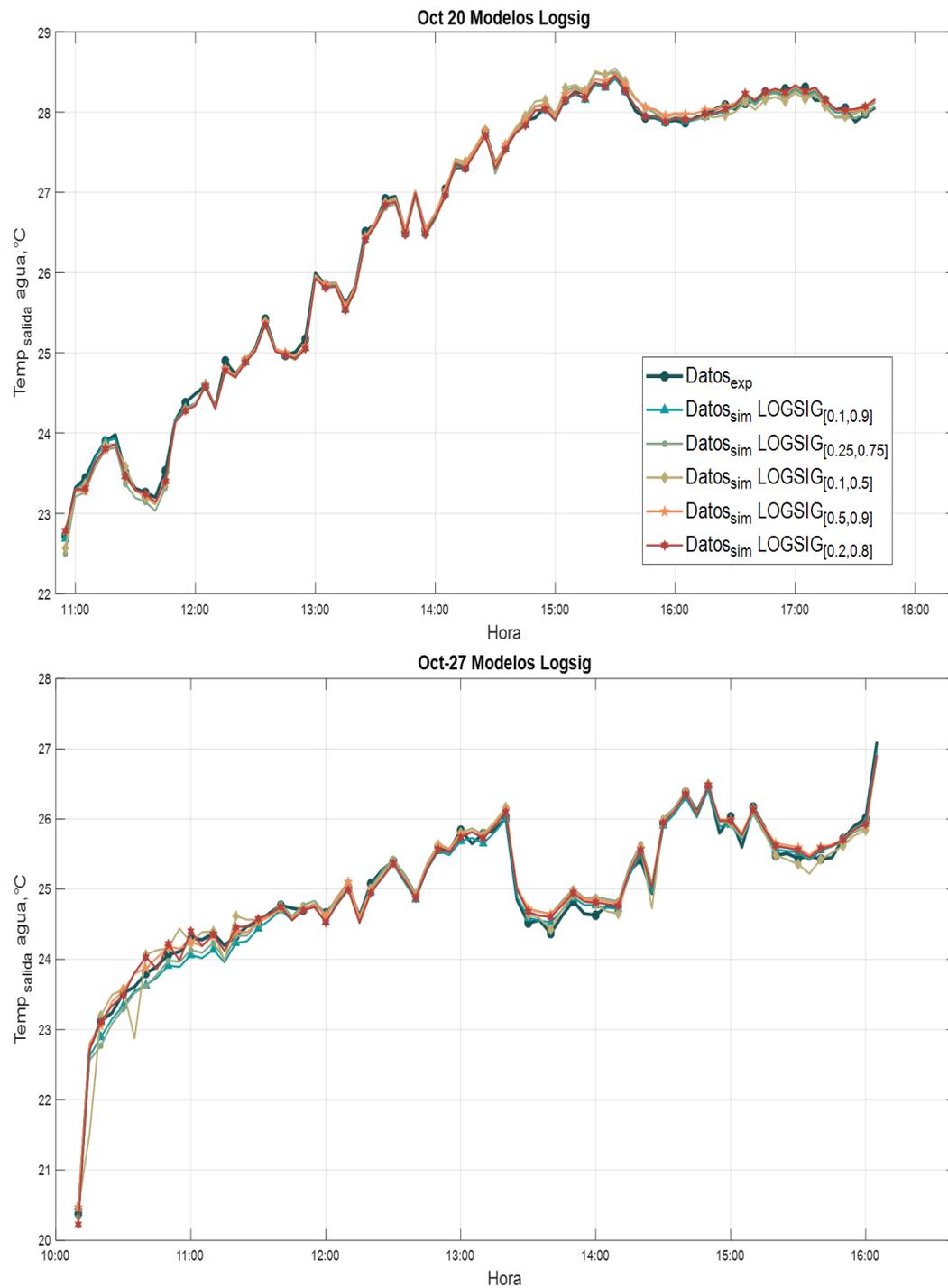


Figura 5.10. Comparación entre los datos experimentales y simulados de los modelos LOGSIG.

El proceso de normalización también ha sugerido su uso en las variables de salida. Por lo tanto, la normalización de la variable de salida se implementó en los modelos propuestos anteriormente. Este procedimiento no se usa con frecuencia debido a que la función de transferencia lineal *PURELIN* es la más manejada en la capa de salida. La función de transferencia *PURELIN* no tiene restricciones; es decir, puede tomar valores de $-\infty$ a $+\infty$. Sin embargo, la normalización de las variables de salida podría mejorar las predicciones del modelo, pero es necesario para desnormalizar los datos de salida para su uso. La validación de estos modelos se realizó mediante pruebas estadísticas, y la Tabla 5.10 muestra los resultados obtenidos.

Tabla 5.10. Análisis estadístico para la comparación a los mismos intervalos de TANSIG y LOGSIG con salida normalizada.

Análisis estadístico	TANSIG con salida normalizada					LOGSIG con salida normalizada				
	[-0.9,0.9]	[-0.9,-0.1]	[0.1,0.9]	[0.2,0.8]	[-0.5,0.5]	[0.1,0.9]	[0.1,0.5]	[0.5,0.9]	[0.2,0.8]	[0.25,0.75]
R²	0.996040	0.994909	0.996359	0.996011	0.996737	0.996021	0.996188	0.987329	0.997721	0.996639
RMSE	0.012302	0.006173	0.005227	0.004100	0.006179	0.005458	0.002689	0.004879	0.003111	0.003137
MAPE	0.476835	0.005938	0.038955	0.023999	0.321890	0.024467	0.054940	0.017070	0.019419	0.042913

Evidentemente, los modelos con salida normalizada obtuvieron resultados similares a los del modelo base. Tenga en cuenta que cuando el modelado con ANN utiliza resultados normalizados, el *RMSE* es muy pequeño. El modelo con normalización en el intervalo de [-0.9, 0.9] para *TANSIG* se descartó porque no aprobó las pruebas de significación, como se muestra en la Tabla 5.11. La Figura 5.11 y la Figura 5.12 muestran una comparación entre los datos simulados por modelos con salida normalizada y datos experimentales; Como se puede ver, el modelo rechazado presentó más dispersión en comparación con los otros modelos.



Tabla 5.11. Pruebas de significancia para la comparación a los mismos intervalos de TANSIG y LOGSIG con salida normalizada.

Función	Intervalos	F calculada (Eq. 2.11)	t de Student calculada (Eq. 2.12)	Si	
				valor calculado < valor crítico [prueba F, prueba t]	p-valor ($\alpha = 0.01$)
TANSIG con salida normalizada	[-0.9, 0.9]	1.01015006	2.73958746	[si, no]	0.00624644
	[-0.9,-0.1]	1.00806719	0.33356398	[si, si]	0.73876920
	[0.1,0.9]	1.00994680	0.08944406	[si, si]	0.92874453
	[0.2,0.8]	1.00000431	0.55675852	[si, si]	0.57780038
	[-0.5, 0.5]	1.00095651	0.7103265	[si, si]	0.47764521
LOGSIG con salida normalizada	[0.1,0.9]	1.00715098	0.39782846	[si, si]	0.69083021
	[0.1,0.5]	1.01597356	2.29983889	[si, si]	0.02163516
	[0.5,0.9]	1.00132610	0.84987004	[si, si]	0.39557363
	[0.2,0.8]	1.00649484	0.78383054	[si, si]	0.43330040
	[0.25,0.75]	1.00274888	1.48134423	[si, si]	0.13878785
Valor crítico (de tablas)		1.14688062	2.32958361	Nivel de significancia 99% ($\alpha = 0.01$)	p-valor > α

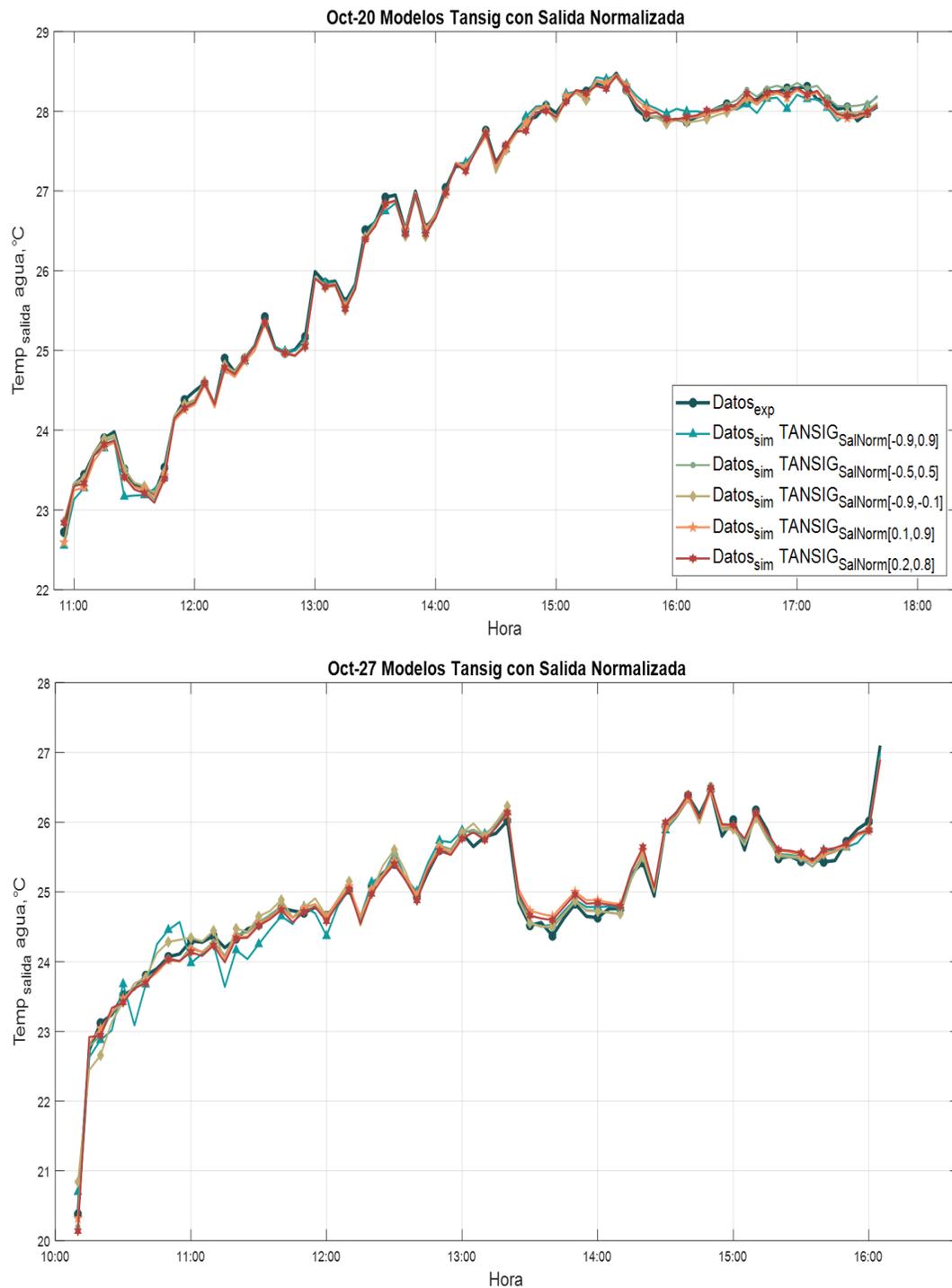


Figura 5.11. Comparación entre los datos experimentales y simulados de los modelos TANSIG con salida normalizada.

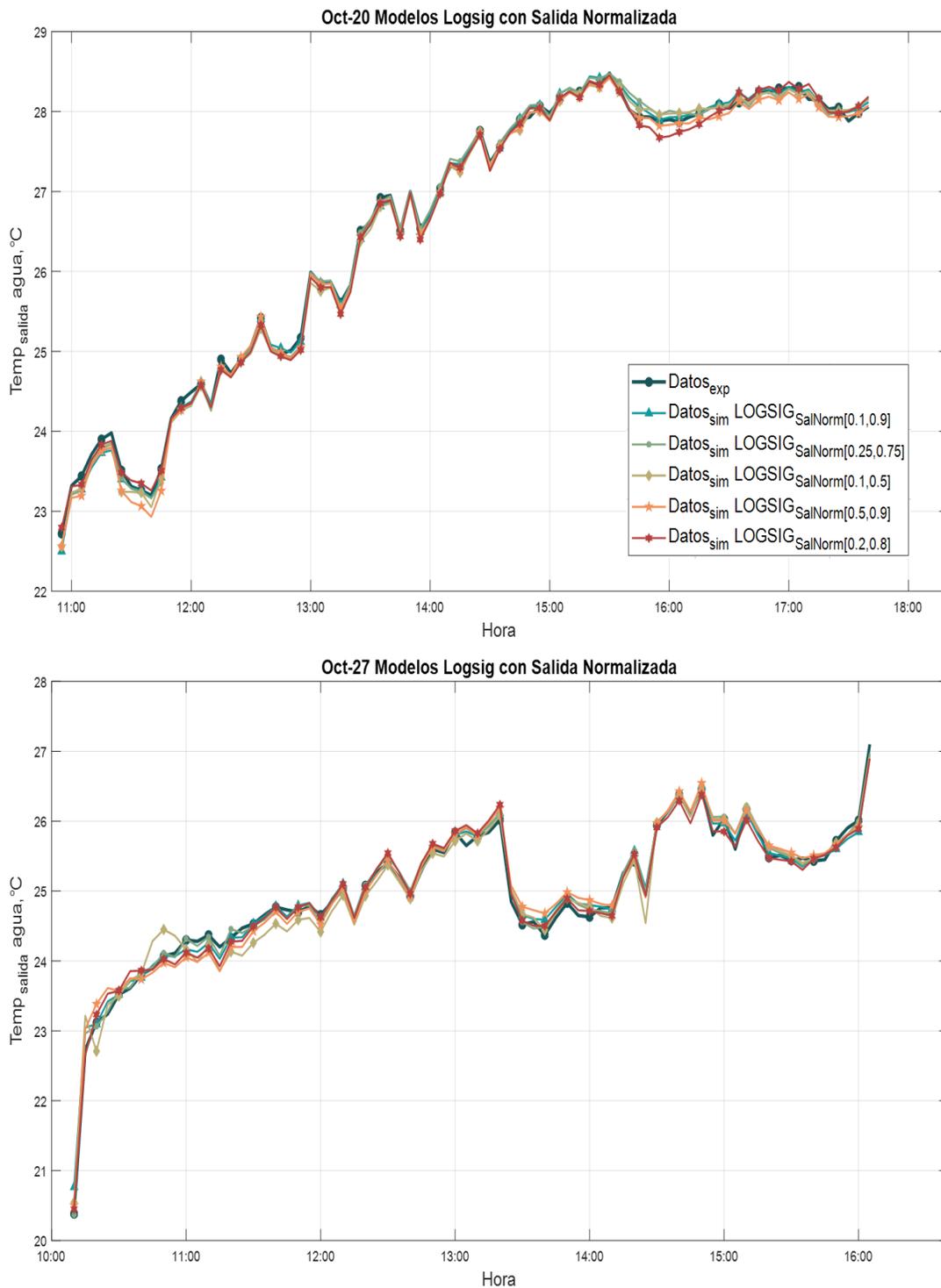


Figura 5.12. Comparación entre los datos experimentales y simulados de los modelos LOGSIG con salida normalizada.

Al contrario de lo que sucede con los modelos con salidas no normalizadas, la Tabla 5.12 muestra que hubo muy pocos modelos que aprobaron la prueba de intercepción de pendiente. Solo los modelos *TANSIG* de [-0.5, 0.5], *LOGSIG* de [0.5,0.9], *LOGSIG* de [0.2,0.8] y *LOGSIG* [0.25, 0.75] aprobaron la prueba. Es esencial resaltar que cuando el modelo ANN usa la función de transferencia *TANSIG*, evitando la normalización de salida, las normalizaciones que consideran intervalos negativos tuvieron un peor desempeño.

Tabla 5.12. Valores de la prueba de linealidad (prueba pendiente-intercepto) para los modelos con salida normalizada.

Intervalo	TANSIG con salida normalizada					LOGSIG con salida normalizada				
	[-0.9,0.9]	[-0.9,-0.1]	[0.1,0.9]	[0.2,0.8]	[-0.5,0.5]	[0.1,0.9]	[0.1,0.5]	[0.5,0.9]	[0.2,0.8]	[0.25,0.75]
Intercepto superior	-0.00172	-0.00070	0.00363	0.00255	0.00075	0.00324	0.00284	0.00855	0.00065	0.00302
Intercepto inferior	-0.00390	-0.00442	0.00214	0.00079	-0.00036	0.00168	0.00185	0.00334	-0.00069	0.00131
Pendiente superior	0.99775	0.99885	0.99781	1.00280	1.00220	0.99924	0.99487	1.00290	1.00574	1.00130
Pendiente inferior	0.99299	0.99345	0.99325	0.99800	0.99789	0.99446	0.99021	0.99430	1.00210	0.99695
Pendiente inferior	0.98823	0.98805	0.98869	0.99320	0.99355	0.98968	0.98556	0.98574	0.99846	0.99255

Como se puede observar, todos los modelos se ajustan bien, pero se ha considerado que el mejor modelo fue el modelo *TANSIG* con normalización en un intervalo de [0.1,0.9], que fue el que obtuvo el menor error por las pruebas estadísticas, aprobó las pruebas de significancia, y presentando menor dispersión. Tiene el R^2 más alto, el *MAPE* más pequeño y el *RMSE* mínimo sin considerar los modelos con salida estandarizada ($R^2 = 0.997405465$, $RMSE = 0.121237247$ y $MAPE = 5.9341 \times 10^{-5}$). De hecho, este intervalo de normalización ha sido ampliamente informado por algunos autores [Reyes-Téllez *et al.* 2017] [Martínez-Martínez *et al.* 2017] [Hernández y Colorado 2013] [Morales *et al.* 2015] [Bassam *et al.* 2014]. La Figura 5.13 muestra la verificación de precisión del mejor modelo ANN.

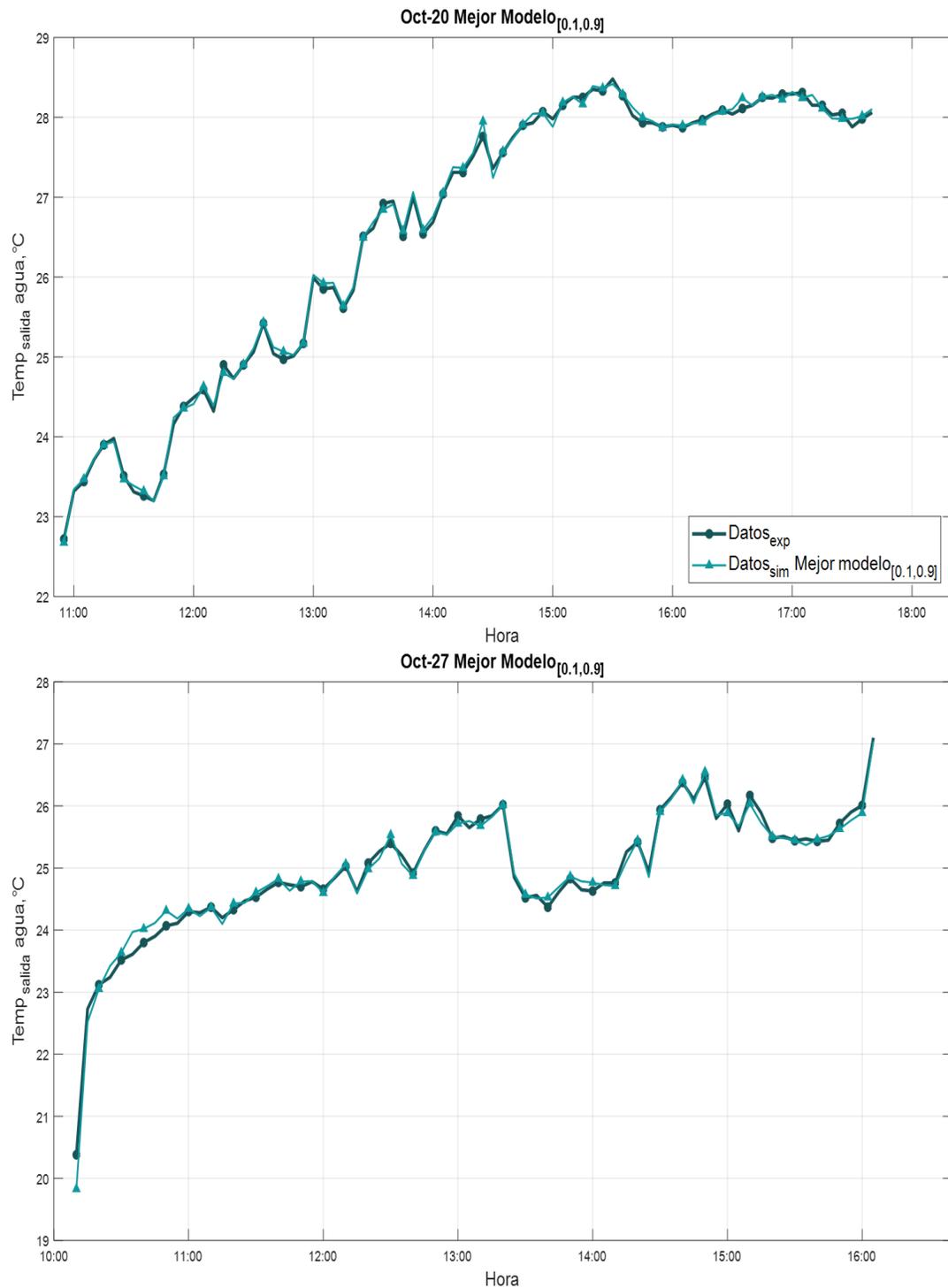


Figura 5.13. Comparación entre los datos experimentales y simulados del mejor modelo ANN.

Una vez que el modelo fue validado, se obtuvo la ecuación para determinar la temperatura de salida del colector. Como tiene una función de transferencia *TANSIG* en la capa oculta y una función de transferencia *PURELIN* en la capa de salida, la ecuación se describe a continuación:

$$T_{salida} = \sum_{k=1}^K \left\{ W_{o(J)} \left[\frac{2}{1+e^{-2 \sum_{j=1}^J (W_{i(J,R)} \cdot P_{(R)} + b_{(1,J)})}} - 1 \right] \right\} + b_{(2,K)} \quad (5.3)$$

Expresado en términos simples:

$$T_{salida} = 2 \left[\frac{W_{o(1,1)}}{1+e^{(x_1)}} + \frac{W_{o(1,2)}}{1+e^{(x_2)}} + \frac{W_{o(1,3)}}{1+e^{(x_3)}} + \frac{W_{o(1,4)}}{1+e^{(x_4)}} + \frac{W_{o(1,5)}}{1+e^{(x_5)}} + \frac{W_{o(1,6)}}{1+e^{(x_6)}} \right] - (W_{o(1,1)} + W_{o(1,2)} + W_{o(1,3)} + W_{o(1,4)} + W_{o(1,5)} + W_{o(1,6)}) + b_{(2,1)} \quad (5.4.a)$$

Donde:

$$x_1 = -2 \cdot (W_{i(1,1)} \cdot P_1 + W_{i(1,2)} \cdot P_2 + W_{i(1,3)} \cdot P_3 + W_{i(1,4)} \cdot P_4 + W_{i(1,5)} \cdot P_5 + W_{i(1,6)} \cdot P_6 + W_{i(1,7)} \cdot P_7 + W_{i(1,8)} \cdot P_8 + W_{i(1,9)} \cdot P_9 +) + b_{(1,1)}$$

⋮

$$x_6 = -2 \cdot (W_{i(6,1)} \cdot P_1 + W_{i(6,2)} \cdot P_2 + W_{i(6,3)} \cdot P_3 + W_{i(6,4)} \cdot P_4 + W_{i(6,5)} \cdot P_5 + W_{i(6,6)} \cdot P_6 + W_{i(6,7)} \cdot P_7 + W_{i(6,8)} \cdot P_8 + W_{i(6,9)} \cdot P_9 +) + b_{(1,6)} \quad (5.5.a)$$

Los valores en las ecuaciones 5.4.a y 5.5.a se reemplazan por pesos y bias que se dan en la Tabla 5.13.

Tabla 5.13. Matriz de pesos y bias para el mejor modelo ANN.

Neurona oculta	Wi (pesos entre la capa de entrada y la capa oculta)									Wo		Bias	
	R=1	R=2	R=3	R=4	R=5	R=6	R=7	R=8	R=9	K=1	b1	b2	
J=1	-8.18103	2.32215	-19.31879	-2.23376	-1.30984	-0.44468	22.90262	-2.89844	-5.25285	2.46815	5.56966		
J=2	1.01091	-0.71149	-10.32036	1.08541	0.53503	0.03464	-4.42111	-0.24447	0.93712	-3.20533	1.84897		
J=3	17.25090	-15.34880	-0.04665	-32.67173	-41.37492	-25.45103	-1.45004	-3.12630	-7.31176	19.87490	27.53423	41.5084	
J=4	-0.24791	0.12391	0.06648	-0.00857	0.02901	-0.00472	-1.64105	0.29083	-0.02988	-19.91882	0.31405	57	
J=5	3.37239	-5.30957	1.55653	-2.07880	-3.86633	3.89572	-5.42197	-1.76573	12.47754	-0.18427	-1.76577		
J=6	-4.31978	1.76543	2.35392	1.96386	1.64968	0.87594	-20.49092	5.37128	2.67082	1.06170	1.91300		

Reemplazando:

$$T_{salida} = 2 \left[\frac{(2.468148)}{1+e^{(x_1)}} + \frac{(-3.205328)}{1+e^{(x_2)}} + \frac{(19.874896)}{1+e^{(x_3)}} + \frac{((-19.918823)}{1+e^{(x_4)}} + \frac{(-0.184265)}{1+e^{(x_5)}} + \frac{(1.061698)}{1+e^{(x_6)}} \right] - [(2.468148) + (-3.205328) + (19.874896) + (-19.918823) + (-0.184265) + (1.061698)] + 41.508457 \quad (5.4.b)$$

Y

$$x_1 = -2 \cdot [(-8.181033) \cdot P_1 + (2.322151) \cdot P_2 + (-19.318793) \cdot P_3 + (-2.233757) \cdot P_4 + (-1.30984) \cdot P_5 + (-0.444677) \cdot P_6 + (22.902622) \cdot P_7 + (-2.898441) \cdot P_8 + (-5.252849) \cdot P_9] + 5.569662 \quad (5.5.b)$$

$$x_2 = -2 \cdot [(1.010914) \cdot P_1 + (-0.711492) \cdot P_2 + (-10.320357) \cdot P_3 + (1.085409) \cdot P_4 + (0.535034) \cdot P_5 + (0.034638) \cdot P_6 + (-4.421108) \cdot P_7 + (-0.244466) \cdot P_8 + (0.937116) \cdot P_9] + 1.848974 \quad (5.5.c)$$

$$x_3 = -2 \cdot [(17.250895) \cdot P_1 + (-15.348804) \cdot P_2 + (-0.046645) \cdot P_3 + (-32.671733) \cdot P_4 + (-41.374916) \cdot P_5 + (-25.451028) \cdot P_6 + (-1.450043) \cdot P_7 + (-3.126300) \cdot P_8 + (-7.311760) \cdot P_9] + 27.534233 \quad (5.5.d)$$

$$x_4 = -2 \cdot [(-0.247912) \cdot P_1 + (0.123911) \cdot P_2 + (0.066479) \cdot P_3 + (-0.008568) \cdot P_4 + (0.029014) \cdot P_5 + (-0.004724) \cdot P_6 + (-1.641045) \cdot P_7 + (0.290829) \cdot P_8 + (-0.029878) \cdot P_9] + 0.314045 \quad (5.5.e)$$

$$\begin{aligned}
 x_5 = -2 \cdot [& (3.372385) \cdot P_1 + (-5.309574) \cdot P_2 + (1.556534) \cdot P_3 \\
 & + (-2.078800) \cdot P_4 + (-3.866334) \cdot P_5 + (3.895716) \cdot P_6 \\
 & + (-5.421965) \cdot P_7 + (-1.765726) \cdot P_8 + (12.477538) \cdot P_9] \\
 & + (-1.765766)
 \end{aligned}
 \tag{5.5.f}$$

$$\begin{aligned}
 x_6 = -2 \cdot [& (-4.319783) \cdot P_1 + (1.765434) \cdot P_2 + (2.353920) \cdot P_3 + (1.963862) \\
 & \cdot P_4 + (1.649682) \cdot P_5 + (0.875938) \cdot P_6 + (-20.490922) \cdot P_7 \\
 & + (5.371282) \cdot P_8 + (2.670823) \cdot P_9] + 1.913004
 \end{aligned}
 \tag{5.5.g}$$

La ecuación de Garson se utilizó para realizar el análisis de sensibilidad, que determina la importancia de las variables de entrada sobre el mejor modelo obtenido. El análisis de sensibilidad depende exclusivamente del peso dado a cada conexión durante el entrenamiento, es decir, los valores pueden variar de un modelo a otro con R^2 , $RMSE$ y $MAPE$ similares y la misma arquitectura. La razón es que, durante el entrenamiento, las conexiones se ajustan para poder representar la variable de salida experimental. Los resultados del análisis de sensibilidad para el mejor modelo obtenido se muestran en la Figura 5.14.

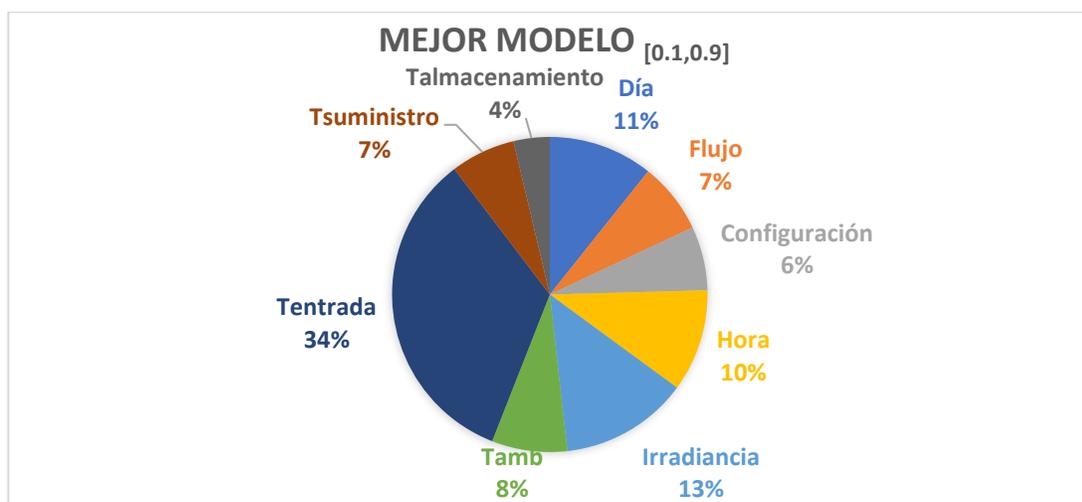


Figura 5.14. Importancia relativa de las variables de entrada del mejor modelo ANN del sistema PTC.

Del resultado del análisis de sensibilidad, se determinó que la variable con mayor preponderancia sobre el mejor modelo fue la temperatura de entrada, con un 34% como era de esperar. Se sabe que la temperatura de salida adquiere valores en relación con la temperatura de entrada del fluido. Las variables que siguen en orden de importancia, como la irradiancia y el día del año, reiteran cómo la estación del año representa una variable esencial para el sistema PTC. La posición de la Tierra con respecto al Sol influye significativamente en la cantidad de radiación absorbida por el PTC.

Las siguientes dos variables, que son la hora del día y la temperatura ambiente, representan claramente cómo el rendimiento del sistema PTC está estrechamente relacionado con el tiempo de exposición. Por otro lado, la configuración muestra un pequeño impacto sobre el modelo, lo que significa que las pérdidas por convección podrían ser comparables a las pérdidas ópticas.

Aunque la temperatura del tanque de alimentación y el flujo volumétrico tuvieron un porcentaje similar más bajo que otros, podrían ser variables a considerar para aumentar la temperatura de salida. Finalmente, la variable menos importante fue la temperatura de almacenamiento, que se puede depreciar porque se representa directamente por la temperatura de entrada debido a que hay un sistema en recirculación.

5.6 Conclusiones parciales

Se instaló un colector solar de canal parabólico de bajo costo con un tubo de absorción ranurado como una opción económica para uso residencial. Mediante mediciones experimentales de un proceso de calentamiento de agua, el PTC se modeló con redes neuronales artificiales a fin de predecir la temperatura de salida.

Para obtener un modelo con alta precisión, se realizó un análisis de entrenamiento usando las funciones de transferencia *TANSIG* y *LOGSIG* en la capa oculta, considerando diferentes intervalos de normalización en las variables de entrada y salida.

El mejor modelo fue el que obtuvo un comportamiento sobresaliente según las pruebas estadísticas, de significancia y de pendiente-intercepto. Este modelo utilizó *TANSIG* como una función de transferencia con un intervalo de normalización de [0.1,0.9]. Se propuso una arquitectura de propagación hacia atrás de múltiples capas (Multilayer Feedforward Backpropagation) con una estructura de 9 neuronas en la capa de entrada, seis neuronas en la capa oculta y una en la capa de salida. Las variables de entrada fueron: la temperatura de entrada del fluido, temperatura ambiente, irradiancia, hora, día, configuración, flujo volumétrico, temperatura de suministro y temperatura de almacenamiento. Los resultados estadísticos obtenidos fueron $R^2 = 0.9974$, $RMSE = 0.12123$ y $MAPE = 5.93 \times 10^{-5}$.

Cabe señalar que todos los modelos con salida normalizada tenían el *RMSE* más pequeño en comparación con los modelos de salida no normalizados. Además, la normalización adecuada de la base de datos acelera el entrenamiento y podría mejorar la precisión de los modelos ANN.

Con el análisis de sensibilidad, se demostró que la mayoría de las variables influyen en el modelo. Sin embargo, la temperatura de entrada es la variable que más influyó, además de las variables relacionadas con la estación del año. La temperatura en el tanque de alimentación y el flujo volumétrico son variables que pueden considerarse como posibles oportunidades para aumentar la temperatura de salida.

6 Conclusiones finales y recomendaciones

6.1 Conclusiones finales

Un colector de canal parabólico de bajo costo fue presentado como una opción para uso residencial. Posteriormente, se le añadió un tubo receptor ranurado para aumentar la ganancia de calor que se transfiere al fluido de trabajo. Se recopilaron los datos experimentales de distintas variables de operación para poder hacer el modelado mediante redes neuronales artificiales.

Las variables recurrentes en el PTC para hacer un modelado mediante redes neurales fueron: la temperatura de entrada del fluido, temperatura ambiente, temperatura de salida, hora, día y mes. Todas estas variables describen como se calienta la superficie terrestre a través de la época del año, así como en el transcurso del día, que son representación del movimiento de traslación y rotación de la Tierra. La irradiancia se puede medir de manera indirecta tomando las mediciones de la temperatura de las superficies, en este caso, de la superficie de la tapa de vidrio y de la superficie reflectante. En cuanto a las temperaturas del agua del tanque de alimentación y del tanque de almacenamiento serán necesarias medirlas para saber el incremento en la temperatura del agua (útiles cuando se usa ANNi)

Los modelos ANN fueron entrenados con una vasta base de datos, tomando el 50% de los datos para entrenamiento, 25% para prueba y el 25% restante para validación, que son porcentajes que tienden a evitar el sobre-aprendizaje e indica que generaliza bien. Los modelos aprobados tuvieron un R^2 mayor a 0.985, un $RMSE$ menor a 1 y un $MAPE$ menor a 1%, lo que indica una buena precisión de los modelos. Además, también pasaron satisfactoriamente las pruebas de significancia (Fisher y t de Student), lo que indica una buena concordancia y fiabilidad de los datos. Asimismo, aprobar la prueba de linealidad de pendiente-intercepto reduce la incertidumbre de los datos, lo que significa que son aptos para usarse. Los resultados de los modelos ANN

fueron capaz de simular el comportamiento no lineal referente al sistema experimental.

El análisis de sensibilidad revela que las variables seleccionadas influyen en el modelo matemático. Debido a que se usa la ecuación de Garson (ec. 2.23), dependiendo de los valores obtenidos de pesos y bias en el modelo ANN, la variable de mayor preponderancia será diferente, pero siempre es referente a una temperatura, ya sea la del ambiente o alguna que tome en cuenta que tan caliente este el entorno donde se encuentra el PTC. Asimismo, aunque la irradiancia es una variable muy importante, en ningún modelo se considera como la más importante puesto que físicamente toma un tiempo para poder incrementar el nivel térmico de un material o fluido.

Se utilizaron diferentes estrategias para optimizar el sistema PTC y la precisión del modelo ANN. Se utilizaron técnicas de mejoramiento de transferencia de calor (tubo ranurado) y diferentes configuraciones (con y sin tapa, y distancia focal modificada) para hacer más eficiente el sistema experimental. Asimismo, el uso de herramientas computacionales tales como Regularización Bayesiana, y un análisis de las funciones de transferencia/activación, así como de las normalizaciones sirvieron para aumentar la precisión en la ecuación ANN de la temperatura de salida simulada. Todo esto con el fin de optimizar la temperatura de salida y tener la temperatura más alta, esto se logra apoyándose con ANNi-AG, debido a que la optimización se realiza directamente con los parámetros obtenidos del modelo ANN y no con los datos experimentales. Por eso es tan importante ser los más precisos posibles y ajustarse bien a los datos experimentales porque muchos pequeños errores acumulados pueden desviarse de la solución óptima del modelo ANNi-AG.

Los métodos computacionales mostrados en esta tesis, fueron una alternativa para predecir la temperatura de salida o alguna otra variable de interés, que proporciona una comprensión del comportamiento del PTC; pero no se pretende descartar el uso de modelos matemáticos.

6.2 Recomendaciones

La metodología aquí utilizada puede ser fácilmente extrapolada a otros sistemas de intercambiadores de calor (y colectores solares), solo se necesitan seleccionar las variables que más repercuten en el sistema y, claro, tener mediciones experimentales para crear una base de datos.

Se debe tener un número bajo de neuronas en la capa oculta, en relación con el número de muestras, porque el modelo incrementara sus coeficientes en la ecuación de salida final del modelo ANN, lo que puede provocar sobreaprendizaje además de que sería un error tener más coeficientes que muestras experimentales.

La variable a optimizar, mediante ANNi-GA, siempre estará delimitada por los intervalos de operación, es decir, puede ser que haya un valor más optimo, pero si esta fuera del intervalo de operación será físicamente imposible ponerlo en práctica. Además de que se debe elegir una variable que sea fácil de manipular su valor en campo. Se podría realizar optimización multivariable en el PTC, pero al ser tan pocas las variables que se pueden manipular físicamente, es recomendable optimizar solo una.

Referente al análisis de sensibilidad de variables, existen múltiples ecuaciones y métodos en la literatura y es difícil saber cuál es la más fiable. Es necesario proponer una ecuación que no dependa exclusivamente de los pesos para, posteriormente, comprobarla experimentalmente. Esto con el fin de, saber la tendencia de cuáles son las variables más importantes (y cuales no), además de que no deben variar tanto de un modelo a otro.

Se sugiere hacer modelos con variables más globales que se suelen utilizar en dispositivos térmicos, tales como los números de Reynolds, Nusselt, Prandtl, Grashof, Rayleigh, entre otras. Esto con el fin de tener una sola ecuación general derivada de un modelo ANN con una buena precisión de predicción.

Referencias

Abiodun O. I., Jantan A., Omolara A. E., Dada K. V., Mohamed N. A., Arshad H. State-of-the-art in artificial neural network applications: A survey. *Heliyon* 4 (11) (2018) e00938.

Alander J.T. On optimal population size of genetic algorithms. *Proceedings CompEuro 1992, Computer Systems and Software Engineering, 6th Annual European Computer Conference* (1992) p. 65-70

Amari S., Murata N., Müller K. R., Finke M., Yang H. H. Asymptotic Statistical Theory of Overtraining and Cross-Validation. *IEEE Transactions On Neural Networks*, Vol. 8, No. 5, September (1997) p. 985-996.

Arabameri M., Javid A., Roudbari A. The Use of Artificial Neural Network (ANN) for Modeling of Ammonia Nitrogen Removal from Landfill Leachate by the Ultrasonic Process. *International Journal of Health Studies* 1 (3) (2015) 13-19.

Audesirk T., Audesirk G., Byers B.E. *Biología: La vida en la Tierra*. 8va. Edición. Editorial Pearson Educación. México (2008) P. 1024

Bakr M. H., Negm M. H. Chapter three - Modeling and design of high-frequency structures using artificial neural networks and space mapping. *Advances in Imaging and Electron Physics Volume 174* (2012) 223-260.

Banco mundial. Emisiones de CO2. Recuperado el 24 de noviembre del 2019. https://datos.bancomundial.org/indicador/EN.ATM.CO2E.KT?end=2014&most_recent_year_desc=true&start=1960&view=chart

Bansal N.K., Uhlemann R. Development and testing of low cost solar energy collectors for heating air. *Solar Energy* 33 (1984) 197–208.

Bassam A., Conde-Gutierrez R.A., Castillo J., Laredo G., Hernandez J.A. Direct neural network modeling for separation of linear and branched paraffins by adsorption process for gasoline octane number improvement. *Fuel* 124 (2014) 158–167

Bassam A., Conde-Gutierrez R.A., Castillo J., Laredo G., Hernandez J.A. Direct neural network modeling for separation of linear and branched paraffins by adsorption process for gaso-line octane number improvement. *Fuel* 124 (2014) 158–167.

Beale M. H., Hagan M. T., Demuth H. B. *Neural Network Toolbox- User's Guide*. The MathWorks, Inc. (2017).

Bergmann R., Ludbrook J., Spooren W. P. J. M. Different outcomes of the Wilcoxon-Mann-Whitney test from different statistics packages. *The American Statistician Association*. February, Vol. 54, No. 1 (2000) p. 72-77.

Bharadwaj P., Khondge A.D., Date A.W. Heat transfer and pressure drop in a spirally grooved tube with twisted tape insert. *International Journal of Heat and Mass Transfer* 52 (2009) 1938–1944.

Bilen K., Cetin M., Gul H., Balta T. The investigation of groove geometry effect on heat transfer for internally grooved tubes. *Applied Thermal Engineering* 29 (2009) 753–761.

Bodenhofer U., *Genetic Algorithms: theory and Applications Lecture Notes, Third Edition—Winter 2003/2004, Fuzzy Logic Laboratorium Linz-Hagenberg (2003).*

Caner M., Gedik E., Keçebas A. Investigation on thermal performance calculation of two type solar air collectors using artificial neural network, *Expert Syst. Appl.*, 38 (2011) 1668–1674.

Cardoso H. E. *Desarrollo y análisis experimental de concentradores solares con diferentes arreglos: Tesis de maestría. CIICAp. Cuernavaca, Morelos. México (2011).*

Carpenter C. *El Plan de Acción de Bali: Los temas principales en las negociaciones sobre el clima - Resumen para los encargados de la formulación de políticas. UNDP Environment & Energy Group. New York. USA (2008)*

Çelik O., Teke A., Yıldırım H.B. The optimized artificial neural network model with Levenberg- Marquardt algorithm for global solar radiation estimation in Eastern Mediterranean Region of Turkey. *J. Clean Prod.*, 116 (2016) 1–12.

Çengel Y. A., Ghajar A. J. *Transferencia de Calor y Masa: Fundamentos y aplicaciones. 4a Edición. McGrawHill. México D.F. (2011).*

Chakraborty K., Bhattacharyya S., Bag R., Hassanien A. A. Chapter book 7. *Sentiment Analysis on a Set of Movie Reviews Using Deep Learning Techniques. Book: Social Network Analytics - Computational Research Methods and Techniques. Academic Press - Elsevier (2019) 127-147.*

Chipperfield A., Fleming P., Pohlheim H., Fonseca C. *Genetic Algorithm TOOLBOX - For Use with MATLAB, User's Guide Version 1.2. Department of automatic control and systems Engineering, University of Sheffield*

Clickgasoil. ¿Cuánta agua gastamos en la ducha? Editado el 6 de abril del 2017 <https://www.clickgasoil.com/blog/cuanta-agua-gastamos-en-la-ducha>

Cortés O., Urquiza G., Hernández J.A. Optimization of operating conditions for compressor performance by means of neural network inverse. Applied Energy. Vo. 86, Issue 11 (2009) p. 2487-2493

Cruz-Jacobo U. Optimización multivariable de un transformador térmico con reciclado de calor usando Algoritmos Genéticos en la Red Neuronal Artificial Inversa: Tesis de maestría. CIICAp, Cuernavaca, Morelos, México (2018).

Curcio S., Iorio G. Chapter book 16 - Models of membrane reactors based on artificial neural networks and hybrid approaches. Book: Handbook of Membrane Reactors – Fundamental Materials Science, Design and Optimisation 1 (2013) 569-597.

Demuth H., Beale M. Neural Network Toolbox for use with Matlab, User's Guide Version 4: The Math Works, Inc., Massachusetts, USA (2005).

Díaz-Gómez J., Parrales A., Álvarez A., Silva-Martínez S., Colorado D., Hernández J.A. Prediction of global solar radiation by artificial neural network based on a meteorological environmental data. Desalination and Water Treatment (2014) 1-8.

Diez F.J., Navas-Gracia L.M., Martínez-Rodríguez A., Correa-Guimares A., L. Chico-Santamarta. Modelling of a flat-plate solar collector using artificial neural networks for different working fluid (water) flow rates. Solar Energy 188 (2019) 1320-1331.

Du K. L., Swamy M. N. S. Neural Networks and Statistical Learning. Springer (2014).

Du C., Wu A. Recurrent Neural Network-Based Inverse Model Learning Control of Manipulators. 6th World Congress on Intelligent Control and Automation, June 2006

Duffie J. A., Beckman W. A. Solar Engineering of Thermal Processes. 4a. Edición. John Wiley & Sons, Inc. New Jersey, USA (2013).

Elmolla E.S., Chaudhuri M., Eltoukhy M.M. The use of artificial neural network (ANN) for modeling of COD removal from antibiotic aqueous solution by the Fenton process. Journal of Hazardous Materials 179 (1-3) (2010) 127-134.

Evangelisti L., Vollaro R.L., Asdrubali F. Latest advances on solar thermal collectors: a comprehensive review. *Renewable and Sustainable Energy Reviews* 114 (2019) 109318.

Farkas I. y Géczy-Vig P. Neural network modelling of flat-plate solar collector. *Computers and electronics in agriculture* 40 (2003) 87-102.

Fischer S., Frey P., Drück H. A comparison between state-of-the-art and neural network modelling of solar collectors. *Solar Energy* 86 (2012) 3268-3277.

Flores López R., Fernández Fernández J. *Las redes neuronales artificiales: Fundamentos teóricos y aplicaciones prácticas*. Netbiblio. México (2008).

Foster R., Ghassemi M., Cota A. *Solar energy: renewable energy and the environment*. CRC Press Taylor & Francis Group (2010).

Fung-Bao L. A modified genetic algorithm for solving the inverse heat transfer problem of estimating plan heat source, *International Journal of Heat and Mass Transfer* 51 (2008) 3745–3752

Gestal M., Rivero D., Rabuñal J. R., Dorado J., Pazos A. *Introducción a los Algoritmos Genéticos y la Programación Genética*. Universidade da Coruña, Servizo de Publicacións. A Coruña (2010) p. 76

Ghorbani M.A., Khatibi R., Hosseini B., Bilgili M. Relative importance of parameters affecting wind speed prediction using artificial neural networks. *Theoretical and Applied Climatology* 114 (1-2) (2013) 107-114.

Glanzman D. L. Associative learning: Hebbian Flies. *Current Biology* 15 (11) (2005) R416-R419.

Glinski J., Horabik J., Lipiec J. *Basic concepts of neural networks*. Encyclopedia of agrophysics. Springer (2011).

González Velasco J. *Energías Renovables*. Editorial Reverté. Barcelona, España (2009). 656 p.

Gudivada V.N., Irfan M.T., Fathi E., Rao D.L. Chapter 5 - Cognitive analytics: going beyond big data analytics and machine learning. *Handbook of Statistics* 35 (2016) 169-205.

Hagan M. T., Demuth H. B., Beale M. H. y De Jesús O. *Neural Network Design*. 2da edición. PSW Publishing Co (2014) eBook.

Hagan M. T., Menhaj M. Training feedforward networks with the Marquardt algorithm. *IEEE Transactions on Neural Networks* 5 (1994) 989–993.

Hagan M.T., Menhaj M.B. Training feed forward networks with the Marquardt algorithm, *IEEE Trans. Neural Netw.*, 5(6) (1994) 989–993.

Hamzaoui Y. El., A. Bassam, Hernandez J. A., Aburto-Cortez O., Oubram O. Search for Optimum Operating Conditions for a Water Purification Process Integrated to a Heat Transformer with Energy Recycling using Artificial Neural Network Inverse Solved by Genetic and Particle Swarm Algorithms. *Chemical Product and Process Modeling: Vol. 7: Iss. 1, Article 3.* (2012).

Hamzaoui Y. El., Hernández J. A., Silva-Martínez S., Bassam A., Álvarez A., Lizama-Bahena C. Optimal performance of COD removal during aqueous treatment of alazine and gesaprim commercial herbicides by direct and inverse neural network. *Desalination* 277 (2011) 325–337

Han J., Kamber M., Pei J. *Data Mining: concepts and techniques.* Chapter book 3: Data preprocessing. Morgan Kaufmann-Elsevier. Third Edition (2012) 83-124.

Hattab N., Motelica-Heino M. Application of an inverse neural network model for the identification of optimal amendment to reduce copper toxicity in phytoremediated contaminated soils. *J. Geochem. Explor.*, 136 (2014) 14–23.

Haupt R. L., Haupt S. E. *Practical Genetic Algorithms.* John Wiley & Sons, Inc. Second edition. New Jersey, USA (2004)

Heng S.Y., Asako Y., Suwa T., Nagasaka K. Transient thermal prediction methodology for parabolic trough solar collector tube using artificial neural network. *Renewable Energy* 131 (2019) 168-179.

Hernández J.A. Optimum operating conditions for heat and mass transfer in foodstuffs drying by means of neural network inverse. *Food Control.* Vol. 20, Issue 4 (2009) p. 435-438.

Hernández J.A., Bassam A., Siqueiros J., Juárez-Romero D. Optimum operating conditions for a water purification process integrated to a heat transformer with energy recycling using neural network inverse. *Renewable Energy.* Volume 34, Issue 4 (2009) p. 1084-1091.

Hernández J.A., Colorado D. Uncertainty analysis of COP prediction in a water purification system integrated into a heat transformer using several artificial neural networks. *Desalination Water Treatment* 51 (2013) 1443-1456.

Holland J.H. *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor (1975). Republished by the MIT Press in 1992.

Hsieh C.K. Thermal analysis of CPC collectors. *Solar Energy* 27 (1981) 19–29.

Hüseyin B. Determination of thermal performance calculation of two different types solar air collectors with the use of artificial neural networks. *International Journal of Heat and Mass Transfer* 60 (2013) 1-7.

IEA. Data and statistics. Recuperado el 17 de noviembre del 2019. <https://www.iea.org/statistics/>

IEA. Renewables information: Overview. International Energy Agency (2019).

IEA. World Energy Balances: Overview. International Energy Agency (2019).

IRENA. *Climate Change and Renewable Energy: National policies and the role of communities, cities and regions* (Report to the G20 Climate Sustainability Working Group (CSWG)). International Renewable Energy Agency, Abu Dhabi (2019).

IRENA. *Estadísticas de energía Renovable 2019*. International Renewable Energy Agency, Abu Dhabi (2019).

IRENA. *Personas, Planeta y Prosperidad: Impulsando la ambición climática a través de las renovables*. International Renewable Energy Agency (2019)

IRENA. *Renewable capacity highlights*. International Renewable Energy Agency (2019).

Jafar R., Shahrour I., Juran I. Application of artificial neural networks (ANN) to model the failure of urban water mains. *Mathematical and Computer Modelling* 51 (9-10) (2010) 1170-1180.

Jaramillo O.A., Borunda M., Velazquez-Lucho K.M., Robles M. Parabolic trough solar collector for low enthalpy processes: an analysis of the efficiency enhancement by using twisted tape inserts. *Renewable Energy* 93 (2016) 125-141.

Jensen B. A., Abonyi J. Neural networks for process modeling. In *Instrument Engineers Handbook, Fourth Edition: Process Control and Optimization*. Vol. 2. CRC Press (2005) p. 253-264.

Jothilakshmi S., Gudivada V.N. Chapter Book 10 - Large scale data enabled evolution of spoken language research and applications. Book: Handbook of Statistics 35 (2016) 301-340.

Jozanikohan G., Norouzi G. H., Sahabi F., Memarian H., Moshiri B. The application of multilayer perceptron neural network in volume of clay estimation: case study of Shurijeh gas reservoir, Northeastern Iran. Journal of Natural Gas Science and Engineering 22 (2015) 119-131.

Kalogirou S. A. Artificial neural networks in renewable energy systems applications: a review. Renewable and Sustainable Energy Reviews 5 (2001) 373-401.

Kalogirou S. A. Long-term performance prediction of forced circulation solar domestic water heating systems using artificial neural networks. Applied Energy 66 (2000) 63-74.

Kalogirou S. A. Prediction of flat-plate collector performance parameters using artificial neural networks. Solar Energy 80 (2006) 248-259.

Kalogirou S. A., Panteliou S., Dentsoras A. Modeling of solar domestic water heating systems using artificial neural networks. Solar Energy 6 (1999) 335-342.

Kalogirou S.A., Mathioulakis E. y Belessiotis V. Artificial neural networks for the performance prediction of large solar systems. Renewable energy 63 (2013) 90-97.

Kamble L.V., Pangavhane D.R., Singh T.P. Neural network optimization by comparing the performances of the training functions -Prediction of heat transfer from horizontal tube immersed in gas–solid fluidized bed. International Journal of Heat and Mass Transfer 83 (2014) 337-344.

Kanal L.N. Perceptrons. International encyclopedia of the social & behavioral Sciences (2001) 11218-11221.

Kaushika N.D., Reddy K.S. Performance of a low cost solar paraboloidal dish steam generating system. Energ. Convers. Manage. 41 (2000) 713–726.

Khataee A.R., Kasiri M.B. Artificial neural networks modeling of contaminated water treatment processes by homogeneous and heterogeneous nanocatalysis. J. Mol. Catal. A-Chem. 331 (1-2) (2010) 86-100.

Khayet M., Cojocarú C. Artificial neural network modeling and optimization of desalination by air gap membrane distillation. *Sep. Purif. Technol.* 86 (2012) 171-182.

Kincaid N., Mungas G., Kramer N., Wagner M., Zhu G. An optical performance comparison of three concentrating solar power collector designs in linear Fresnel, parabolic trough, and central receiver. *Applied Energy* 231 (2018) 1109–1121.

Koenig C. Histología UC-Escuela de Medicina P. Universidad Católica de Chile. Recuperado el 20 de Mayo de 2015. <http://escuela.med.puc.cl/paginas/cursos/segundo/histologia/histologiaweb/paginas/ne35145.html>

Kumar L., Hasanuzzaman M., Rahim N.A. Global advancement of solar thermal energy technologies for industrial process heat and its future prospects: a review. *Energy Conversion and Management* 195 (2019) 885-908.

Laidi M., Hanini S. Optimal solar COP prediction of a solar-assisted adsorption refrigeration system working with activated carbon/methanol as working pairs using direct and inverse artificial neural network. *Inter. J. Refrig.*, 36 (2013) 247–257.

Liu F.B. A modified genetic algorithm for solving the inverse heat transfer problem of estimating plan heat source. *Inter. J. Heat. Mass Trans.* 51 (2008) 3745–3752.

Martínez-Martínez E., Escobedo-Trujillo B.A., Colorado D., Morales L.I., Huicochea A., Hernández J.A., Siqueiros J. Criteria for improving the traditional artificial neural network methodology applied to predict COP for a heat transformer. *Desalin. Water Treat.* 73 (2017) 90-100.

Mateo F.J. Redes neuronales y preprocesado de variables para modelos y sensores en bioingeniería. Tesis Doctoral. Departamento de Ingeniería Electrónica. Valencia, España (2012).

MathWorks. Genetic Algorithm Options. Recuperado el 20 de diciembre del 2019. <https://www.mathworks.com/help/gads/genetic-algorithm-options.html>

MathWorks. Global Optimization Toolbox User's Guide-R2013a, The MathWorks Inc, (2013).

May Tzuc O., Bassam A., Escalante Soberanis M. A., Venegas-Reyes E., Jaramillo O. A., Ricalde L. J., Ordoñez E. E., El Hamzaoui Y. Modeling and

optimization of a solar parabolic trough concentrator system using inverse artificial neural network. *Journal of Renewable and Sustainable Energy* 9, 013701 (2017).

McCulloch W.S., Pitts W. A logical calculus of the ideas immanent in nervous activity. *The bulletin of Mathematical Biophysics* 5 (4) (1943) 115-133.

Millán-Ocampo D.E., Parrales A., González-Rodríguez J.G., Silva-Martínez S., Porcayo-Calderón J., Hernández-Pérez J.A., Modelling of behavior for inhibition corrosion of bronze using artificial neural network (ANN), *Entropy* 20 (2018) 409-419.

Mohanraj M., Jayaraj S., Muraleedharan C. Applications of artificial neural networks for thermal analysis of heat exchangers-A review. *International Journal of Thermal Sciences*. 90 (2015) 150-172.

Morales L.I., Conde-Gutiérrez R.A., Hernández J.A., Huicochea A., Juárez-Romero D., Siqueiros J. Optimization of an absorption heat transformer with two-duplex components using inverse neural network and solved by genetic algorithm. *Applied Thermal Engineering* 85 (2015) 322-333

Moriarty P., Honnery D. Global renewable energy resources and use in 2050. *Academic Press Managing Global Warming* (2019) 221–235.

Morin G., Dersch J., Platzer W., Eck M., Häberle A. Comparison of Linear Fresnel and Parabolic Trough Collector power plants. *Solar Energy* 86 (2012) 1–12.

Narendra K. S., Parthasarathy K. Identification and Control of Dynamical Systems Using Neural Networks. *IEEE Transactions on Neural Networks*. Volume 1, Issue: 1 (1990) p. 4-27.

NASA Science Solar System Exploration. Our Sun - In depth. Recuperado el 20 de agosto del 2019. <https://solarsystem.nasa.gov/solar-system/sun/in-depth/>

Nierhaus G. Chapter 9 Artificial Neural Networks. *Algorithmic Composition-Paradigms of Automated Music Generation*. Springer Wien New York (2009) 205-224.

Parrales A., Colorado D., Díaz-Gómez J.A., Huicochea A., Álvarez A., Hernández J.A. New void fraction equations for two-phases flow in helical heat exchangers using artificial neural networks. *Applied Thermal Engineering* 130 (2018) 149-160.

Parrales A., Hernández-Pérez J. A., Flores O., Hernandez H., Gómez-Aguilar J. F., Escobar-Jiménez R., Huicochea A. Heat Transfer Coefficients Analysis in a Helical Double-Pipe Evaporator: Nusselt Number Correlations through Artificial Neural Networks. *Entropy* 21 (7) (2019) 689.

Penm J., Chaar B., Moles R., Penm J. Book chapter 37 - Predicting ASX health care stock index movements after the recent financial crisis using patterned neural networks. *Rethinking Valuation and Pricing Models* (2013) 599-610.

Ponce Cruz P. *Inteligencia artificial con aplicaciones a la ingeniería*. Primera Edición. Alfaomega Grupo Editor, S.A. de C.V. México (2010) Páginas: 376.

Pranesh V., Velraj R., Christopher S., Kumaresan V. A 50 year review of basic and applied research in compound parabolic concentrating solar thermal collector for domestic and industrial applications. *Solar Energy* 187 (2019) 293–340

Qibing J., Xiaodong Z., Shusheng G. Neural Network Inverse Control for Electric Heating Furnace Outlet Temperature. *IFAC Proceedings Volumes*. Volume 31, Issue 25 (1998) p. 233-237.

Qin S. J. Chapter 8 - Neural networks for Intelligent sensors and control — Practical issues and some solutions. *Neural Systems for Control* (1997) 213-234.

Ramírez-Ramos G. E. *Análisis de la transferencia de calor y su eficiencia en un colector solar de canal parabólico usando tubo ranurado: Tesis de maestría*. CIICAp, Cuernavaca, Morelos, México (2017)

Reyes-Téllez E.D. *Estimación de la temperatura de salida en un colector solar de bajo costo, Tesis de maestría*. CIICAp, Cuernavaca, Morelos, México (2015) p. 115.

Reyes-Téllez E.D., Conde-Gutiérrez R.A., Hernández J.A., Cardoso E., Silva-Martínez S., Sierra F. Z., Cortés-Aburto O. Optimal operating condition for a type W parabolic trough collector with low-cost components using inverse neural network and solved by genetic algorithm. *Desalin. Water Treat.* (2017) 1-10

Rocabruno-Valdés C.I., González-Rodríguez J.G., Díaz-Blanco Y., Juantorena A.U., Muñoz-Ledo J.A., El-Hamzaoui Y., Hernández J.A. Corrosion rate prediction for metals in biodiesel using artificial neural networks. *Renewable Energy* 140 (2019) 592-601.

Rocabruno-Valdés C.I., Ramírez-Verduzco L.F., Hernández J.A. Artificial neural network models to predict density, dynamic viscosity, and cetane number of biodiesel. *Fuel* 147 (2015) 9-17.

Rocas y minerales. Carbón. Recuperado el 24 de noviembre del 2019. <https://www.rocasym minerales.net/carbon/>

Rodríguez J.A., Hamzaoui Y. El., Hernández J.A., García J.C., Flores J.E., Tejada A.L. The use of artificial neural network (ANN) for modeling the useful life of the failure assessment in blades of steam turbines. *Engineering Failure Analysis* 35 (2013) 562–575.

Rypkema H.A. Chapter 2.1 - Environmental chemistry, renewable energy, and global policy. *Green Chemistry* (2018) 19-47.

SMN. Servicio Nacional Meteorológico. Recuperado en Octubre y Noviembre del 2016. <https://smn.cna.gob.mx/>

Shabbir J., Anwer T. Artificial intelligence and its role in near future. *Journal of Latex Class Files* 14 (8) (2015) 1-11.

Shafieian A., Khiadani M., Nosrati A. Strategies to improve the thermal performance of heat pipe solar collectors in solar systems: a review. *Energy Conversion and Management* 183 (2019) 307-331.

Sharif A., Raza S.A., Ozturk I., Afshan S. The dynamic relationship of renewable and nonrenewable energy consumption with carbon emission: a global study with the application of heterogeneous panel estimations. *Renewable Energy* 133 (2019) 685-691.

Shukla D., Paul F. W. Computationally Efficient Control of Nonlinear Systems Using Orthonormal Activation Function Based Neural Networks. *IFAC Proceedings Volumes*. Volume 29, Issue 1, (1996) p. 4995-5000.

Sobhansarbandi S., Atikol U. Performance of flat-plate and compound parabolic concentrating solar collectors in under-floor heating systems. *Journal Solar Energy* 137 (2015) 034501.

Sözen A., Özalp M., Arcaklioğlu E. Calculation for the thermodynamic properties of an alternative refrigerant (R508b) using artificial neural network. *Applied Thermal Engineering* 27 (2-3) (2007) 551-559.

Sözen A., Tayfun M., Sinan Ü. Determination of efficiency of flat-plate solar collectors using neural network approach. *Expert Systems with Applications* 35 (2008) 1533-1539.

Sugavaneswaran L. *Mathematical modeling of gene networks*. Elsevier Inc. (2017).

Turkson R. F., Yan F., Ali M. K. A., Hu J. Artificial neural network applications in the calibration of spark-ignition engines: an overview. *Engineering Science and Technology, an International Journal* 19 (3) (2016) 1346-1359.

Tutorials Point. *Genetic Algorithms, Tutorials Point- Simply easy learning* (2016), e-book

UNDP. *La Hoja de Ruta de Bali: Los temas claves en negociación*. United Nations Development Programme - Environment & Energy Group. New York. USA (2008)

Verma S.P. *Estadística básica para el manejo de datos experimentales: aplicación en la Geoquímica (Geoquimiometría)*. Universidad Nacional Autónoma de México. México (2005).

Verma S.P., Cruz-Huicochea R. Alternative approach for precise and accurate Student's t critical values and application in geosciences. *Journal of Iberian Geology* 39 (1) (2013) 31-56.

Wahab A., Hassan A., Qasim M.A., Ali H.M., Babar H., Sajid M.U.. Solar energy systems – potential of nanofluids. *Journal of Molecular Liquids* 289 (2019) 111049.

Wasserstein R. L., Lazar N. A. The ASA's Statement on p-Values: Context, Process, and Purpose. *The American Statistician* Vol. 70 No. 2 (2016) 129-133.

Wen-Tao J., Jacobi A. M., Ya-Ling H., Wen-Quan T. Summary and evaluation on single-phase heat transfer enhancement techniques of liquid laminar and turbulent pipe flow. *International Journal of Heat and Mass Transfer* 88 (2015) 735–754.

Wen-Tao J., Jacobi A. M., Ya-Ling H., Wen-Quan T. Summary and evaluation on the heat transfer enhancement techniques of gas laminar and turbulent pipe flow. *International Journal of Heat and Mass Transfer* 111 (2017) 467–483.

Yegnanarayana B. *Artificial neural networks*. Prentice-Hall of India Private Limited (2006).