



Universidad Nacional Mayor de San Marcos

Universidad del Perú. Decana de América
Facultad de Química e Ingeniería Química
Escuela Profesional de Ingeniería Química

**Diseño de un sensor virtual para la predicción de la
eficiencia de conversión de SO₂ a SO₃ en un reactor
catalítico mult-ilechos mediante redes neuronales
artificiales**

TESIS

Para optar el Título Profesional de Ingeniero Químico

AUTOR

Sergio Hernán ESPINOZA QUIROZ

ASESOR

Jorge Luis CÁRDENAS RUIZ

Lima, Perú

2022



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

Referencia bibliográfica

Espinoza, S. (2022). *Diseño de un sensor virtual para la predicción de la eficiencia de conversión de SO₂ a SO₃ en un reactor catalítico mult-ilechos mediante redes neuronales artificiales*. [Tesis de pregrado, Universidad Nacional Mayor de San Marcos, Facultad de Química e Ingeniería Química, Escuela Profesional de Ingeniería Química]. Repositorio institucional Cybertesis UNMSM.

ANEXO 04

Información complementaria

DATOS DEL AUTOR	
Nombres y apellidos	Sergio Hernán Espinoza Quiroz
Tipo de documento de identidad	DNI
Número de documento de identidad	10459072
URL de ORCID	No aplica
DATOS DEL ASESOR	
Nombres y apellidos	Jorge Luis Cárdenas Ruiz
Tipo de documento de identidad	DNI
Número de documento de identidad	06627695
URL de ORCID	0000-0001-9313-905X
DATOS DEL JURADO	
Presidente del jurado	
Nombres y apellidos	Gregorio Roberto Álvarez Contreras
Tipo de documento de identidad	DNI
Número de documento de identidad	06009529
Miembro del jurado 1	
Nombres y apellidos	Jorge Eduardo Loayza Perez
Tipo de documento de identidad	DNI
Número de documento de identidad	08714366
DATOS DE INVESTIGACIÓN	
Línea de investigación	C.0.5.8. Inteligencia artificial
Grupo de investigación	No aplica
Agencia de financiación	No aplica
Ubicación geográfica de la investigación	Universidad Nacional Mayor de San Marcos Latitud: -12.060366° Longitud: -77.083700°
Año o rango de años en que se realizó la investigación	2021-2022
URL de disciplinas OCDE	https://purl.org/pe-repo/ocde/ford#2.04.00



UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS
Universidad del Perú, Decana de América
FACULTAD DE QUÍMICA E INGENIERÍA QUÍMICA

" Año del Fortalecimiento de la Soberanía Nacional "

ESCUELA PROFESIONAL DE INGENIERIA QUÍMICA

A C T A DE SUSTENTACIÓN DE TESIS VIRTUAL

Los suscritos Miembros del Jurado Evaluador de Sustentación de Tesis Virtual nombrados por la Dirección de la Escuela Profesional de Ingeniería Química, bajo la Presidencia del **M.Sc. Gregorio Roberto Alvarez Contreras**, el **Mg. Jorge Eduardo Loayza Pérez** (Miembro) y el **Ing. Jorge Luis Cárdenas Ruiz** (Asesor), habiendo presentado para el efecto la **TESIS**, titulada "**DISEÑO DE UN SENSOR VIRTUAL PARA LA PREDICCIÓN DE LA EFICIENCIA DE CONVERSIÓN DE SO₂ A SO₃ EN UN REACTOR CATALÍTICO MULT-ILECHOS MEDIANTE REDES NEURONALES ARTIFICIALES**", después de sustentada de manera virtual mediante el enlace ([https:// meet.google.com/sqr-cmjpg-vsd](https://meet.google.com/sqr-cmjpg-vsd)) y habiendo aprobada la Tesis elaborada por el Bachiller en Ingeniería Química: **SERGIO HERNÁN ESPINOZA QUIROZ**; para optar el **TITULO PROFESIONAL DE INGENIERO QUÍMICO**, el **Jurado Evaluador** acuerda calificarlo con la **NOTA** de:

DIECIOCHO

18

.....
(LETRAS)

.....
(NÚMEROS)

Lima, 08 de agosto del 2022

M.Sc. Gregorio Roberto Álvarez Contreras
Presidente de Jurado

Mg. Jorge Eduardo Loayza Pérez
Docente Revisor

Ing. Jorge Luis Cárdenas Ruiz
Asesor



Firmado digitalmente por PALOMINO
INFANTE Ruben Alfredo FAU
20148092282 soft
Motivo: Soy el autor del documento
Fecha: 12.08.2022 16:55:36 -05:00

Dr. R. Alfredo Palomino
Director de la EPIQ

AGRADECIMIENTO

Ante todo, quiero agradecer a Dios por haberme acompañado siempre en mi vida profesional, pues sin su ayuda no lo hubiera logrado. A mis queridos padres Sergio e Inés por su amor y apoyo incondicional que siempre me han brindado. A mi esposa Monica y mis hijos Alejandro, Romina y Renata por su paciencia y tiempo que me cedieron para concluir este trabajo.

Gracias a ellos fue posible la realización de la tesis.

Sergio Hernán Espinoza Quiroz.

Lima, agosto de 2022

ÍNDICE GENERAL

ÍNDICE GENERAL	i
LISTA DE FIGURAS	iii
LISTA DE TABLAS	vii
RESUMEN	viii
ABSTRACT	ix
CAPITULO I: INTRODUCCIÓN	1
1.1 Situación problemática	1
1.2 Formulación del problema	3
1.2.1 Problema general	3
1.2.2 Problemas específicos	3
1.3 Objetivos	4
1.3.1 Objetivo general	4
1.3.2 Objetivos específicos	4
1.4 Hipótesis	5
1.5 Justificación e importancia	5
1.6 Limitaciones del estudio	6
CAPITULO II: MARCO TEORICO	7
2.1 Antecedentes de la investigación	7
2.2 El proceso de conversión catalítica en la Planta de Ácido Sulfúrico	8
2.2.1 Área de limpieza de gases	8
2.2.2 Área de contacto	9
2.2.3 La eficiencia de conversión:	11
2.2.4 Influencia de las variables:	14
2.3 Los sensores virtuales y la simulación de procesos	16
2.4 Las redes neuronales artificiales (RNAs)	20
2.3.1 El perceptrón	23
2.3.2 El perceptrón multicapa	24
2.3.3 Número de capas ocultas	25
2.3.4 Número de neuronas ocultas	26
2.3.5 El algoritmo de retropropagación (Backpropagation Algorithm)	26
2.3.6 Algoritmos de segundo orden.	28
2.3.7 Inicialización de la red	29
2.3.8 Generalización de la red	30
CAPÍTULO III: MARCO METODOLÓGICO	33

3.1	Introducción.....	33
3.2	Preprocesamiento de datos.....	34
3.2.1	La limpieza de datos	37
3.2.2	La integración de datos	37
3.2.3	Reducción de datos	37
3.2.4	Escalado de datos:.....	38
3.3	Selección de las variables de entrada.....	38
3.4	Selección de la estructura neuronal.....	40
3.5	Prueba final de validación	41
CAPITULO IV: RESULTADOS Y DISCUSIÓN		43
4.1	Resultados del sensor virtual basado en RNA	43
4.1.1	Colección y filtrado de datos	43
4.1.2	Selección de las variables de entrada	50
4.1.3	Selección de la estructura neuronal.....	59
4.1.4	Prueba final de la predicción de concentración de salida de SO₂.....	66
4.1.5	Cálculo de la eficiencia de conversión.....	67
4.2	Resultados del sensor virtual derivado de leyes	69
CAPÍTULO V. CONCLUSIONES.....		73
RECOMENDACIONES		74
REFERENCIAS BIBLIOGRÁFICAS.....		75
ANEXOS.....		79
	ANEXO 1: Ecuación de estimación de la eficiencia de conversión derivado de leyes.....	79
	ANEXO 2: Redes de retropropagación en Matlab	84
	ANEXO 3: Programa en VBA de Excel.....	88
	ANEXO 4: Ejemplo de cálculo del tiempo de retardo.....	93
	ANEXO 5: Ejemplo de cálculo en Excel de los indicadores de desempeño de los modelos.....	94
	ANEXO 6: Indicadores de desempeño de los modelos de predicción de la concentración de SO₂ de salida, variando el número de variables de entrada y el número de neuronas en la capa oculta.....	95

LISTA DE FIGURAS

Figura 1	Ilustración del concepto de modelos de caja blanca, caja gris y caja negra.	2
Figura 2	Etapas de la limpieza de gases	8
Figura 3	Área de contacto	9
Figura 4	Etapas de cálculo de la eficiencia de conversión	11
Figura 5	Metodología de cálculo de la eficiencia de conversión mediante el modelo derivado en leyes.	12
Figura 6	Variación de las temperaturas en los lechos del convertidor durante un típico arranque de una planta de ácido sulfúrico.	16
Figura 7	Esquema simple de un sensor virtual.	17
Figura 8	Ejemplo un simulador de procesos avanzado en tiempo real que integra la aplicación de los sensores virtuales (soft sensor) y también calcula parámetros de rendimiento/eficiencia para el proceso de endulzamiento de nafta.	19
Figura 9	La neurona biológica.	20
Figura 10	Neurona biológica y neurona artificial	21
Figura 11	La neurona artificial.	22
Figura 12	El Perceptrón simple.	23
Figura 13	El Perceptrón multicapa.	24
Figura 14	Funciones de activación.	25
Figura 15	Mecanismo de entrenamiento supervisado.	27
Figura 16	Tipos de generalización: (a) Con buenas propiedades; (b) Con escasas propiedades	30
Figura 17	División de conjuntos de datos de entrenamiento, validación y prueba final.	31
Figura 18	Ilustración de la regla de detención temprana basada en la validación durante el proceso de aprendizaje.	32
Figura 19	Metodología del sensor virtual para el cálculo de la eficiencia de	33

	conversión.	
Figura 20	Complemento Excellink de Matlab.	34
Figura 21	Etapas de diseño del modelo neuronal para la predicción de la Concentración de SO ₂ de salida.	34
Figura 22	Ejemplos de dos señales con diferentes promedios y desviaciones estándar.	35
Figura 23	Tareas principales en el preprocesamiento de los datos	36
Figura 24	Reducción de las variables de entrada al modelo	38
Figura 25	Esquema que muestra el periodo de tiempo de los datos colectados, en condiciones de operación normal (azul) y condiciones transitorias (verde).	44
Figura 26	Resumen descriptivo de las etapas de diseño del modelo neuronal para la predicción de la concentración de SO ₂ , y el software empleado en cada etapa.	45
Figura 27	Histograma de frecuencias - % de SO ₂ de entrada, total de datos colectados.	46
Figura 28	Histograma de frecuencias - % de SO ₂ de entrada, filtrando los datos en el punto cero por parada o recirculación de la planta.	46
Figura 29	Distribución de los datos y análisis de datos atípicos que están fuera Del rango de tres veces la desviación estándar de la muestra.	47
Figura 30	Histograma de frecuencias – Concentración de Oxígeno del gas de entrada, %.	49
Figura 31	Histograma de frecuencias – Flujo volumétrico del gas de entrada, m ³ /s.	49
Figura 32	Ilustración de cálculo en Excel del coeficiente de correlación entre la concentración de entrada y la concentración de salida a diferentes tiempos de retardo de la concentración de salida. Para este esquema solo se muestra el retardo de hasta 2 minutos. Para más tiempo de retardo solo se sigue la misma metodología ilustrado en el esquema.	51
Figura 33	Coeficiente de correlación muestral de Pearson a diferentes tiempos de retardo entre la concentración de SO ₂ de entrada y la concentración de SO ₂ de salida, a condición de operación normal, condición de operación transitoria y ambas condiciones juntas.	52
Figura 34	Gráfica de dispersión entre la concentración de SO ₂ de entrada y la concentración de SO ₂ de salida a condiciones de operación normal y transitorias.	54

Figura 35	Coeficiente de correlación muestral de Pearson a diferentes tiempos de retardo entre la concentración de O ₂ de entrada y la concentración de SO ₂ de salida.	55
Figura 36	Coeficiente de correlación muestral de Pearson a diferentes tiempos de retardo entre la concentración de SO ₂ de entrada y la temperatura de salida del primer lecho.	56
Figura 37	Coeficiente de correlación muestral de Pearson a diferentes tiempos de retardo entre la concentración de SO ₂ de entrada y la temperatura de salida del segundo lecho.	56
Figura 38	Coeficiente de correlación muestral de Pearson a diferentes tiempos de retardo entre la concentración de SO ₂ de entrada y la temperatura de salida del tercer lecho.	57
Figura 39	Coeficiente de correlación muestral de Pearson a diferentes tiempos de retardo entre la concentración de SO ₂ de entrada y la temperatura de salida del cuarto lecho.	58
Figura 40	Coeficiente de correlación muestral de Pearson a diferentes tiempos de retardo entre el flujo de entrada y la concentración de SO ₂ de salida.	58
Figura 41	Coeficiente de correlación muestral de Pearson a diferentes tiempos de retardo entre la presión en el primer lecho y la concentración de SO ₂ de salida.	59
Figura 42	Estructura óptima del modelo neuronal para la predicción de la concentración de SO ₂ de salida.	64
Figura 43	Estructura y parámetros de entrenamiento la red neuronal en Matlab	65
Figura 44	Entrenamiento de la estructura 9-9-1 con el método de detención temprana (early stopping) en Matlab.	66
Figura 45	Prueba final del modelo de predicción de la concentración de SO ₂ de salida del reactor catalítico multi-lechos.	67
Figura 46	Sensor virtual obtenido para la predicción de la eficiencia de conversión.	67
Figura 47	Resultados de predicción de la eficiencia de conversión del reactor catalítico multi-lechos.	69
Figura 48	Coeficiente de correlación de Pearson entre la concentración de SO ₂ de entrada y la temperatura de salida del primer lecho a diferentes tiempos de retardo.	70

Figura 49	Coeficiente de correlación de Pearson entre la concentración de SO ₂ de entrada y la temperatura de salida del primer lecho a diferentes tiempos de retardo (suavizado).	70
Figura 50	Coeficiente de correlación de Pearson entre la concentración de O ₂ de entrada y la temperatura de salida del primer lecho a diferentes tiempos de retardo.	71
Figura 51	Coeficiente de correlación de Pearson entre la concentración de O ₂ de entrada y la temperatura de salida del primer lecho a diferentes tiempos de retardo (suavizado).	71

LISTA DE TABLAS

Tabla 1	Algunas aplicaciones de RNA en la predicción de procesos químicos.	8
Tabla 2	Algunas diferencias y semejanzas entre los sensores virtuales y la simulación.	17
Tabla 3	Algunas equivalencias entre la estadística y las redes neuronales	22
Tabla 4	Significados del coeficiente de correlación de Pearson.	40
Tabla 5	Valores promedio de las variables de entrada y salida, obtenidos a partir de 13 409 datos de cada variable.	48
Tabla 6	Determinación de la razón señal/ruido en las variables.	50
Tabla 7	Coefficientes de correlación de la concentración de SO ₂ de entrada con la concentración de SO ₂ de salida a diferentes tiempos de retardo.	53
Tabla 8	Indicadores de desempeño de los modelos de predicción de la concentración de SO ₂ de salida, considerando la combinación del número de variables en la capa de entrada y el número óptimo de neuronas en la capa oculta.	63
Tabla 9	Resultados de desempeño de predicción de la concentración de SO ₂ de salida, mediante la red neuronal.	66
Tabla 10	Resultados del desempeño de predicción de la eficiencia de conversión mediante el sensor virtual basado en RNA.	68
Tabla 11	Resultados del desempeño de predicción de la eficiencia de conversión mediante el sensor virtual derivado de leyes.	72

RESUMEN

El presente estudio examina la posibilidad de diseñar un sensor virtual basado en Redes Neuronales Artificiales (RNA) para predecir la eficiencia de conversión de un reactor catalítico de una planta de ácido sulfúrico. Para el diseño del sensor virtual se utilizaron datos de planta, que fueron preprocesados con el fin de seleccionar los datos útiles, y mediante análisis de correlación se determinan preliminarmente las variables de entrada relevantes para el modelo de predicción, siendo: La concentración de dióxido de azufre (SO_2) en el ingreso, las temperaturas de salida de los lechos, la concentración de oxígeno (O_2) en la entrada y el flujo volumétrico del gas en la entrada.

El diseño del sensor virtual basado en RNA requirió previamente el desarrollo de un modelo neuronal para la predicción de la concentración de salida de SO_2 en el último lecho del convertidor. Con el fin de obtener el mejor modelo de predicción se realizan entrenamientos y validaciones de diversas estructuras de redes neuronales cambiando su configuración en el número de variables de la capa de entrada y el número de neuronas de la capa oculta. La configuración final óptima del modelo para la predicción de la concentración de SO_2 de salida tiene nueve variables en la capa de entrada y nueve neuronas en la capa oculta. Los indicadores de desempeño del modelo de predicción de la concentración de SO_2 son: El error porcentual medio absoluto de 12 %, el índice refinado de concordancia de 0,832 y el coeficiente de correlación de 0,945. Luego a partir de los resultados de la concentración de salida de SO_2 se calcularon las eficiencias de conversión de SO_2 , obteniéndose el sensor virtual, cuyos indicadores de desempeño de predicción de la eficiencia de conversión son: El error porcentual medio absoluto de 0,2 %, índice refinado de concordancia de 0,902 y el coeficiente de correlación de 1,00.

Además, el sensor virtual basado en RNA es comparado con otro sensor virtual derivado de leyes en estado estacionario, este último se determinó calculando en cada lecho del convertidor con los datos de la planta de ácido. Los resultados indican que el sensor virtual basado en RNA tiene un mejor desempeño en comparación con el sensor virtual derivado de leyes en estado estacionario, que tuvo un bajo desempeño.

Palabras clave: Sensores virtuales, redes neuronales, convertidor catalítico.

ABSTRACT

The present study examines the possibility of designing a virtual sensor based on Artificial Neural Networks (ANN) to predict the conversion efficiency of a catalytic reactor of a sulfuric acid plant. For the development of the virtual sensor, plant data were used, which were preprocessed in order to select the useful data, and through correlation analysis, the relevant input variables for the prediction model are preliminarily determined, being: The concentration of sulfur dioxide (SO₂) at the inlet, the outlet temperatures of the beds, the concentration of oxygen (O₂) at the inlet and the volumetric flow of the gas at the inlet.

The design of the virtual sensor based on ANN previously required the development of a neural model for the prediction of the outlet concentration of SO₂ in the last bed of the converter. In order to obtain the best prediction model, various neural network structures are trained and validated by changing their configuration in the number of variables in the input layer and the number of neurons in the hidden layer. The final optimal configuration of the model for the prediction of the outlet SO₂ concentration has nine variables in the input layer and nine neurons in the hidden layer. The SO₂ concentration prediction performance indicators are: The mean absolute percentage error of 12 %, the refined index of agreement of 0,832, and the correlation coefficient of 0,945. Then, from the results of the SO₂ outlet concentration, the conversion efficiencies were calculated whose performance indicators of the virtual sensor indicate a mean absolute percentage error of 0,2 %, a refined index of agreement of 0,902 and a correlation coefficient of 1,00.

In addition, the virtual sensor based on ANN is compared with another virtual sensor derived from laws in steady state, the latter was determined by calculating in each bed of the converter with plant data in real time. The results indicate that the virtual sensor based on ANN has a better performance compared to the virtual sensor derived from laws in steady state, which had a lower performance.

Keywords: Virtual sensors, neural networks, catalytic converter.

CAPITULO I: INTRODUCCIÓN

1.1 Situación problemática

En los últimos años, se viene dando un gran salto tecnológico en el desarrollo e integración de las tecnologías digitales, la conectividad y la inteligencia artificial. Este nuevo cambio tecnológico en la industria ha sido denominado la cuarta revolución industrial o industria 4.0, y trae consigo nuevas tecnologías como el internet de las cosas (IoT), los grandes datos (big data), los dispositivos móviles, sensores inalámbricos, sensores inteligentes, etc. Actualmente, estas tecnologías se están integrando en la industria química, y permiten obtener mayores beneficios económicos logrando la implementación de sistemas inteligentes e innovadores, que optimizan el proceso de fabricación dando una mayor flexibilidad y eficiencia.

Este cambio tecnológico ha traído consigo nuevas técnicas para el modelado y simulación de procesos químicos. Actualmente, se puede acceder fácilmente a una mayor cantidad de datos, utilizar recursos computacionales de mayor velocidad, lo cual ha permitido el desarrollo de simuladores de procesos avanzados. Estos simuladores son diseñados a partir de una combinación de modelos derivados de leyes, modelos dinámicos y modelos basados en datos, de este último, una de las técnicas más conocidas son las redes neuronales artificiales (RNA), utilizado para estimar parámetros que por su complejidad son difíciles de modelar. Asimismo, los simuladores avanzados además de simular el comportamiento dinámico de los procesos, son también utilizados para la simulación con datos de planta en tiempo real, obteniendo un error aceptable e inclusive son usados para el entrenamiento de los operadores de la planta (Chowdhary A. R., 2014)

Los sensores virtuales son modelos diseñados con el fin de predecir una variable de un proceso químico. Por más de cuarenta años los sensores virtuales han respaldado las operaciones en la toma de decisiones y control de procesos. En sus inicios, los sensores virtuales han sido diseñados a partir de modelos de caja blanca, es decir, modelos basados en leyes, modelos dinámicos y modelos estadísticos. Los sensores virtuales diseñados a partir de modelos derivados de leyes y modelos dinámicos combinan principios de dinámica de fluidos, termodinámica, cinética química, física, etc.; tienen gran ventaja debido a que las leyes que lo gobiernan tienen dominios de validez muy amplios, sin embargo, estos modelos se obtienen a menudo realizando simplificaciones de la realidad y presentan un mayor grado de dificultad ante procesos más complejos. Tradicionalmente

los sensores virtuales han sido diseñados a partir de modelos estadísticos de caja negra utilizando técnicas clásicas como la regresión polinomial, regresión exponencial, etc. siendo utilizados especialmente cuando los procesos son poco comprendidos, en el sentido de que las leyes fundamentales responsables de su comportamiento dinámico se conocen poco o son completamente desconocidas. Estos modelos utilizan datos de entrada-salida y permiten predecir una variable. La figura 1 ilustra el concepto de los modelos de caja blanca, caja gris y caja negra.

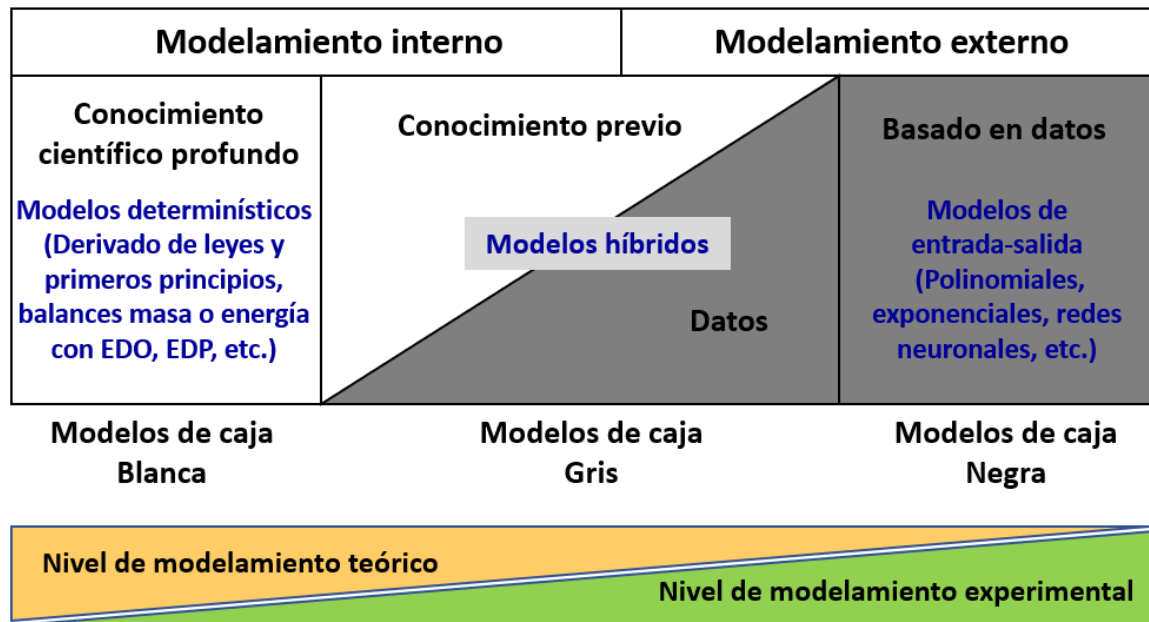


Figura 1 Ilustración del concepto de modelos de caja blanca, caja gris y caja negra. Fuente: Adaptado de Madsen H., Bacher P., Juhl R., Kloppenborg J. (2015).

El desafío actual de los diseñadores de sensores virtuales es obtener un modelo con mayor grado de robustez y eficiencia (desde el punto de vista de rapidez en el diseño y ejecución del modelo) y, para lograrlo, posiblemente se requieran mejoras algorítmicas adicionales y la explotación de nuevos paradigmas informáticos. Es por ello, que desde los años 90's se han diseñado modelos de caja negra basados en Redes neuronales artificiales (RNAs), siendo aplicados en diversos campos de la ingeniería, entre ellas, en la predicción de variables de procesos químicos, logrando un mejor desempeño en la predicción en comparación con los modelos estadísticos tradicionales (polinomiales, exponenciales, etc.). Uno de los tipos de redes neuronales más utilizados históricamente en la predicción de variables son las redes de tipo multicapa perceptrón (MLP), este tipo de red es conocido por ser un aproximador universal. Los MLPs son llamados también redes neuronales

superficiales cuando tienen una sola capa oculta, son diseñados a partir de datos de entrada y salida, durante su diseño se lleva a cabo una gran cantidad de esquemas de entrenamiento/validación, con procedimientos de prueba y error con la finalidad de encontrar la estructura neuronal óptima con menor error de predicción. A pesar de ello, el modelamiento mediante RNA es más rápida en comparación con los modelos derivados de leyes y los modelos dinámicos. Entre las ventajas de las RNA están el caracterizar sistemas no lineales donde intervienen un gran número de variables, requieren poco tiempo de desarrollo y son capaces de operar en tiempo real.

Actualmente, no hay trabajos científicos publicados en nuestro país sobre la aplicación de sensores virtuales basados en RNA en el proceso de conversión catalítica, sin embargo, existen otras aplicaciones a plantas de proceso, por ejemplo, en la predicción de la concentración de salida de una columna de destilación reactiva (Kataria G., Singh K., Kumar R., 2016), predicción del rendimiento de destilados en el proceso de visbreaking (Foddi O., Grosso M., Tronci S., Baratti R, 2007), predicción de la concentración de H_2S en una planta de endulzamiento de gases (ZareNezhad B., Aminian A., 2012), entre otros.

El modelamiento del reactor catalítico es complejo, ya que intervienen muchos parámetros que son desconocidos, ocurren un número de reacciones potenciales y diferentes estados de oxidación del vanadio durante el ciclo catalítico (Lapina O., Zhinimaev B., Boghosian S., Eriksen K y Fehrmann. 1999). Por las razones mencionadas, por las características del proceso, el número de variables que intervienen y su comportamiento no lineal, hacen que las redes neuronales sea la herramienta más adecuada para diseñar un sensor virtual para predecir la eficiencia de conversión de SO_2 a SO_3 en un reactor catalítico de múltiples lechos, mediante su aplicación con datos de planta en tiempo real.

1.2 Formulación del problema

1.2.1 Problema general

¿Será posible predecir la eficiencia de conversión de una planta de ácido sulfúrico utilizando un sensor virtual basado en RNA?

1.2.2 Problemas específicos

1. ¿Cuál sería la relación entre las variables del proceso de conversión catalítica?

2. ¿Cuál sería el desempeño de predicción del sensor virtual basado en RNA para predecir la eficiencia de conversión?
3. ¿Cuál sería la estructura óptima de la red neuronal con mejor desempeño en la predicción de la concentración de salida de SO₂?
4. ¿Cuál sería el desempeño de predicción de la eficiencia de conversión aplicando un sensor virtual derivado de leyes en estado estacionario?
5. ¿Cuál de los sensores virtuales, basado en RNA o derivado de leyes tendrá un mejor desempeño en la predicción de la eficiencia de conversión?

1.3 Objetivos

1.3.1 Objetivo general

Diseñar un sensor virtual basado en RNAs para predecir la eficiencia de conversión de un reactor catalítico de una planta de ácido sulfúrico.

1.3.2 Objetivos específicos

Para alcanzar el Objetivo General antes enunciado, progresiva y concatenadamente debemos lograr los siguientes propósitos específicos:

- Determinar la relación existente entre las variables del proceso de conversión catalítica utilizando datos de planta.
- Obtener la estructura neuronal óptima (con menor error) para la predicción de la concentración de SO₂ a la salida del reactor catalítico mediante la selección de diversas estructuras neuronales y variables de entrada. El resultado del modelo de predicción de SO₂ permitirá calcular la eficiencia de conversión de SO₂ a SO₃.
- Determinar el desempeño de predicción de la eficiencia de conversión utilizando un sensor virtual derivado de leyes en estado estacionario.
- Comparar el desempeño de predicción del sensor virtual basado en RNA y el sensor virtual derivado de leyes en estado estacionario.

1.4 Hipótesis

Predecir la eficiencia de conversión de un reactor catalítico en una planta de ácido sulfúrico a partir del diseño de un sensor virtual basado en RNA.

1.5 Justificación e importancia

La finalidad del presente trabajo de investigación es diseñar un sensor virtual para predecir la eficiencia de conversión del reactor catalítico. Primero, el modelo neuronal diseñado predice la concentración de salida de SO₂ en situaciones cuando el analizador de SO₂ de salida se encuentra fuera de operación o en mantenimiento, de tal manera que el sensor virtual podría utilizarse para estimar en tiempo real la concentración de SO₂ de salida del reactor y con ello determinar la eficiencia de conversión, siendo información útil para la toma de decisiones del operador de planta y el supervisor ambiental. Una disminución significativa de la eficiencia de conversión indica una alta concentración de SO₂ de salida respecto al rango normal de operación; y podría deberse a las siguientes causas:

- Debido a una pobre conversión de SO₂ a SO₃.
- Desvío del gas SO₂ en una o más camas del catalizador a causa de una filtración por los tubos del intercambiadores de calor.

El operador de la sala de control procede a regular las temperaturas de los lechos, disminuyendo la concentración de SO₂ en los gases de entrada a fin de disminuir su aporte de calor. En caso no se logre reducir la concentración de SO₂ de salida, se realiza la toma de muestras del gas a la entrada y salida de los intercambiadores de calor y los lechos de catalizador, a fin de determinar si existe filtración en algunos de los tubos.

Respecto a su importancia económica, la aplicación del sensor virtual en reemplazo del analizador de SO₂ existente, podría generar ahorros e inclusive puede extenderse su aplicación de modo permanente, para ello, es necesario el mantenimiento periódico del sensor virtual mediante el reentrenamiento del modelo. En la comparación de sus costos entre el sensor virtual y el analizador, el sensor virtual podría ofrecer ahorros significativos de aproximadamente 50 % menos del costo de capital y 10 a 20 % de los costos de operación y mantenimiento del analizador de SO₂, con la misma precisión y calidad en los datos. (Antson O., Pellikka T. 2014).

Otros aportes que se obtienen mediante el presente trabajo de investigación son:

- Utilizar la metodología empleada en el diseño de sensores virtuales para su aplicación en otros procesos industriales donde la instrumentación física sea costosa y difícil de instalar.
- En el contexto de industria 4.0 permitirá contribuir con el acervo científico para los estudiantes de ingeniería química mediante el aporte del conocimiento de modelos basados en datos y la técnica de las RNAs.

1.6 Limitaciones del estudio

En la presente investigación se diseña un modelo con datos de operación, siendo en su mayoría datos a condiciones de operación normal y en menor proporción con datos a condiciones de operación transitorias debido a que estas condiciones son muy inusuales.

Los datos históricos empleados para el diseño contienen información representativa por un período de tiempo limitado, es por ello, que cada cierto tiempo, se debe alimentar nuevos datos de entrenamiento en el sensor virtual con el fin de mantener el desempeño de predicción (Souza F., Araújo R., Mendes J. 2016).

CAPITULO II: MARCO TEORICO

2.1 Antecedentes de la investigación

Existen pocas investigaciones en el desarrollo de sensores virtuales para el proceso de conversión de un reactor catalítico. La razón, posiblemente sea debido a su grado de complejidad existente para lograr que se ajuste bien a una planta real.

2.1.1 A nivel nacional:

No se han realizado estudios en este tema a nivel nacional.

2.1.2 A nivel internacional:

Estudio realizado por los autores: Jianjun He, Junfeng Zhang y Helen Shang de la Escuela de Ingeniería de Bharti, Laurentian University, Canadá y publicado en The Canadian Journal of Chemical Engineering, Volume 95, June 2017. Canada.

Han desarrollado un sensor virtual derivado de leyes para predecir la razón (ratio) de conversión y la concentración de SO₂ a la salida del convertidor catalítico de una planta de ácido sulfúrico. El sensor virtual se diseñó aplicando un modelo sobre la base del equilibrio de masa y el balance de energía en estado estacionario; y puede ser utilizado con datos de planta en tiempo real. A partir del sensor virtual propuesto, se puede estimar la razón de conversión y la concentración de SO₂ en cada lecho del convertidor catalítico.

El análisis dinámico indica que la temperatura de salida del convertidor varía lentamente con la fracción molar de alimentación de SO₂. Las dos variables están bastante correlacionadas y el coeficiente de correlación alcanza el máximo a un tiempo de retardo de nueve minutos. La aplicación de un filtro exponencial a la fracción molar de alimentación de SO₂ permite que la señal filtrada se sincronice con la temperatura de salida del primer lecho. Se desarrolló un sensor virtual para estimar la razón de conversión como función de las variables medidas, que incluyen a la temperatura de salida del primer lecho y la fracción molar de alimentación. Al reemplazar las variables de alimentación con las variables filtradas, el modelo genera una estimación útil de las variables (la razón de conversión y las concentraciones de salida).

Esta metodología de diseño del sensor virtual derivado de leyes servirá de referencia en el presente estudio para diseñar un sensor virtual que permita predecir la eficiencia global de conversión, ya que se aplicará la metodología a los cuatro lechos del convertidor catalítico.

Finalmente, no se ha encontrado estudios sobre el desarrollo sensores virtuales basados en redes neuronales para la predicción de la eficiencia del convertidor catalítico en plantas de ácido sulfúrico. Sin embargo, se pueden citar algunos estudios de predicción de otros procesos químicos basados en RNA, mostrados en la tabla 1.

Tabla 1. *Algunas aplicaciones de RNA en la predicción de procesos químicos*

Autor(s)	Año	País	Aplicación
Kataria G., Singh K., Kumar R.	2016	India	Modelo de sensor virtual para una columna de destilación reactiva (An based soft sensor model for reactive distillation column).
ZareNezhad B., Aminian A.	2012	Irán	Un modelo de red neuronal artificial para predecir el rendimiento de eliminación de H ₂ S de los solventes de piperazina en plantas de endulzamiento de gas. (An artificial neural network model for predicting the H ₂ S removal performance of piperazine solvents in gas sweetening plants).
Boger Z.	2009	Israel	Modelado de redes neuronales artificiales para reducir la contaminación industrial del aire (Artificial Neural Networks Modeling to Reduce Industrial Air Pollution).
Foddi O., Grosso M., Tronci S., Baratti R.	2007	Italia	Predicción de rendimientos de destilados en un proceso de Visbreaking mediante redes neuronales (Prediction of Distillate Yields in a Visbreaking Process by Neural Networks).
Al-Duwaish H., Ghouti L., Halawani T., Mohandes M	2002	Bélgica	Uso de analizadores de procesos mediante redes neuronales artificiales: un caso de estudio (Use of Artificial Neural Networks Process Analyzers: A Case Study).

Fuente: Elaboración propia.

2.2 El proceso de conversión catalítica en la Planta de Ácido Sulfúrico

2.2.1 Área de limpieza de gases

Los gases procedentes de los procesos metalúrgicos contienen polvo y deben eliminarse antes de pasar a la etapa de oxidación catalítica del SO₂. Si no se elimina

el polvo, podría taponear las capas de catalizador y bloquear el flujo de gas, asimismo el arsénico presente en el polvo podría envenenar la actividad del catalizador. Las etapas de la limpieza de gases se muestran en la figura 2:

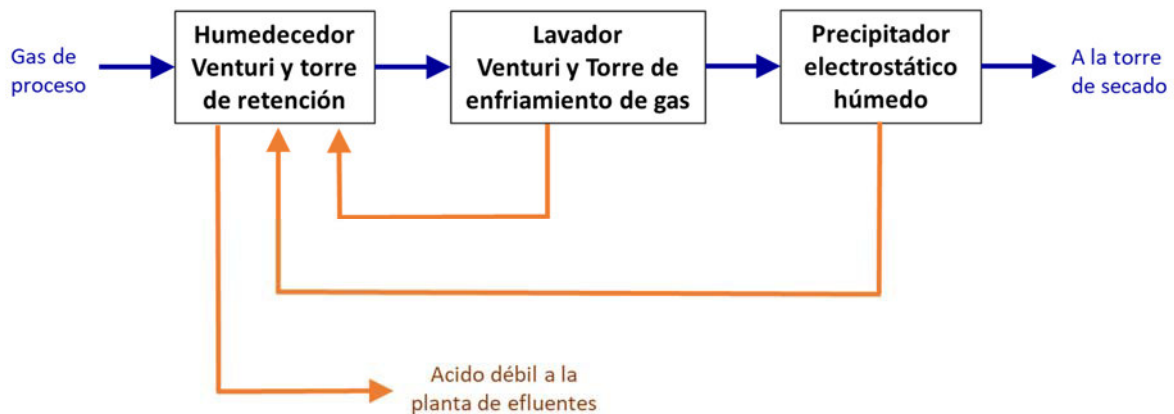


Figura 2 Etapas de la limpieza de gases. Fuente: Elaboración propia.

De acuerdo a su diseño, después de las etapas de limpieza, el gas resultante contiene una concentración de polvo de aproximadamente un miligramo por metro cúbico a condiciones normales (mg/m^3), quedando listo para las operaciones de secado, enfriamiento y oxidación catalítica del SO_2 .

2.2.2 Área de contacto

En el área de contacto los gases pasan por una torre de secado con el objetivo de eliminar la humedad del gas. Luego el gas es transportado hasta el convertidor catalítico mediante sopladores y es enfriado antes de su ingreso a los lechos del convertidor, mostrado en la figura 3.

Los gases ingresan al primero, segundo y tercer lecho del convertidor, pasando por etapas de enfriamiento antes del ingreso a cada lecho. Previo a su ingreso al cuarto lecho, el gas ingresa a una torre de absorción intermedia con el objetivo de liberar el SO_3 presente en el gas y, de esta manera, favorecer el equilibrio de reacción hacia una mayor conversión de SO_2 y que se obtiene en el cuarto lecho del convertidor. Al remover el SO_3 de la corriente de gas, el equilibrio cambiará para producir más SO_3 a fin de compensar la pérdida de concentración del SO_3 .

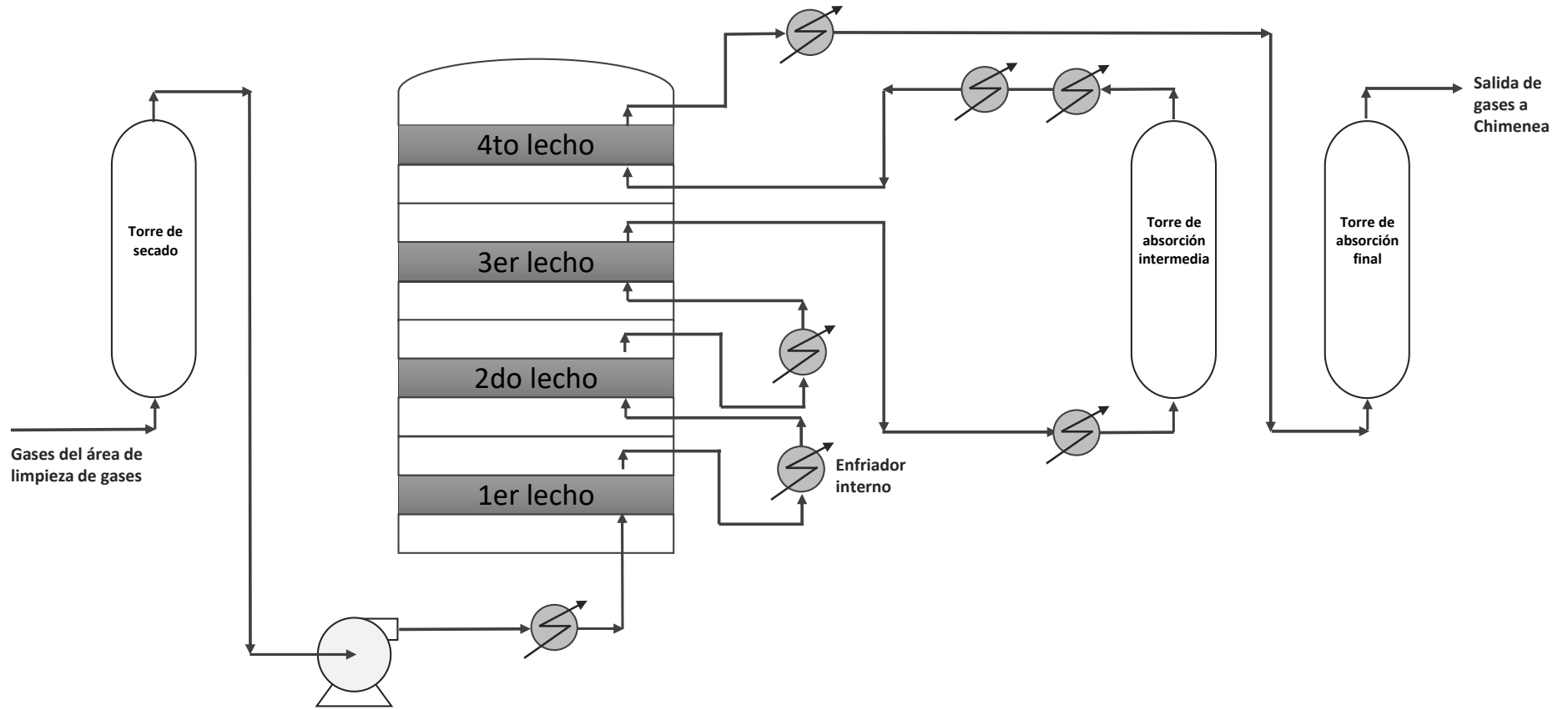
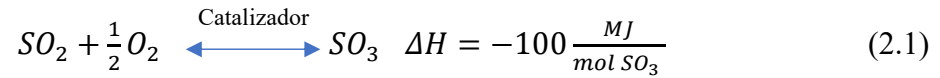


Figura 3. Área de contacto. Fuente: Elaboración propia.

La reacción de oxidación en el convertidor catalítico esta dado por la ecuación 2.1:



2.2.3 La eficiencia de conversión:

La determinación de la eficiencia de conversión en el reactor catalítico se realiza mediante la instalación de analizadores de SO₂ en línea ubicados a la entrada y salida del reactor catalítico. El analizador a la salida también permite medir el desempeño ambiental del proceso.

Estudios realizados por He J., Zhang J., Shang H. (2017) desarrollaron ecuaciones para estimar la eficiencia de conversión basado en el balance de materia y energía en el estado estacionario. La resolución de las ecuaciones por He et al. se muestran en el anexo 1.

Antes de utilizar las ecuaciones, He et al. realizan un análisis de correlación con el objetivo de determinar de los diferentes tiempos de retardo entre las variables, luego, para lograr la sincronización de las variables utiliza la técnica de suavización exponencial.

La ecuación para la estimación de la eficiencia de conversión basado en el balance de materia en estado estacionario es:

$$\Phi = \frac{X_{SO_2}^{in} - X_{SO_2}}{X_{SO_2}^{in} - \frac{1}{2}X_{SO_2}X_{SO_2}^{in}} \quad (2.2)$$

Donde:

$X_{SO_2}^{in}$ = Concentración de SO₂ a la entrada del lecho del convertidor.

X_{SO_2} = Concentración de SO₂ a la salida del lecho del convertidor.

θ = Eficiencia de conversión.

La ecuación 2.2 se utiliza para calcular la eficiencia de conversión, pero antes, se debe conocer con la concentración de salida del lecho del convertidor X_{SO_2} .

Otra ecuación basada en el balance de energía en estado estacionario, permite estimar la eficiencia de conversión, mostrado en la ecuación 2.3:

$$\Phi = \frac{(X_{SO_2}^{in} k_{p,SO_2} + X_{O_2}^{in} k_{p,O_2} + X_{SO_3}^{in} k_{p,SO_3} + X_{re}^{in} k_{p,N_2})(T - T^{in})}{\left[(k_{p,SO_2} + \frac{1}{2}k_{p,O_2} - k_{p,SO_3})T + (b_{p,SO_2} + \frac{1}{2}b_{p,O_2} - b_{p,SO_3}) \right] X_{SO_2}^{in}} \quad (2.3)$$

Donde:

$k_{p,SO_2}, k_{p,O_2}, k_{p,SO_3}, k_{p,N_2}, b_{p,SO_2}, b_{p,O_2}, b_{p,SO_3}, b_{p,N_2}$ son los coeficientes e interceptos empíricos para el cálculo de la entalpia del gas (King M., Davenport W., Moats M., 2013)

La concentración de SO_2 a la salida del lecho del convertidor se puede expresar en términos de eficiencia de conversión y concentración de entrada:

$$X_{SO_2} = \frac{(1-\Phi)X_{SO_2}^{in}}{1-\frac{1}{2}\Phi X_{SO_2}^{in}} \quad (2.4)$$

Para el cálculo de la concentración de salida en cada lecho, primero se calcula la eficiencia de conversión mediante la ecuación 2.3, y segundo se utiliza la ecuación 2.4 para calcular la concentración de salida en cada lecho. Las etapas de cálculo se muestran en la figura 4:



Figura 4 Etapas de cálculo de la eficiencia de conversión. Fuente: Elaboración propia.

Esta metodología de cálculo se aplica de manera similar a las demás componentes de los gases (O_2 , SO_3 y N_2) y también se calcula secuencialmente en cada lecho del convertidor desde el primero hasta el tercer lecho de manera similar, y en el cuarto lecho se calcula considerando la recuperación del SO_3 antes del ingreso al cuarto lecho, y de este modo se obtiene la eficiencia de conversión en el último lecho del

convertidor, y por ende la eficiencia global de conversión en el convertidor catalítico, como se muestra en la figura 5.

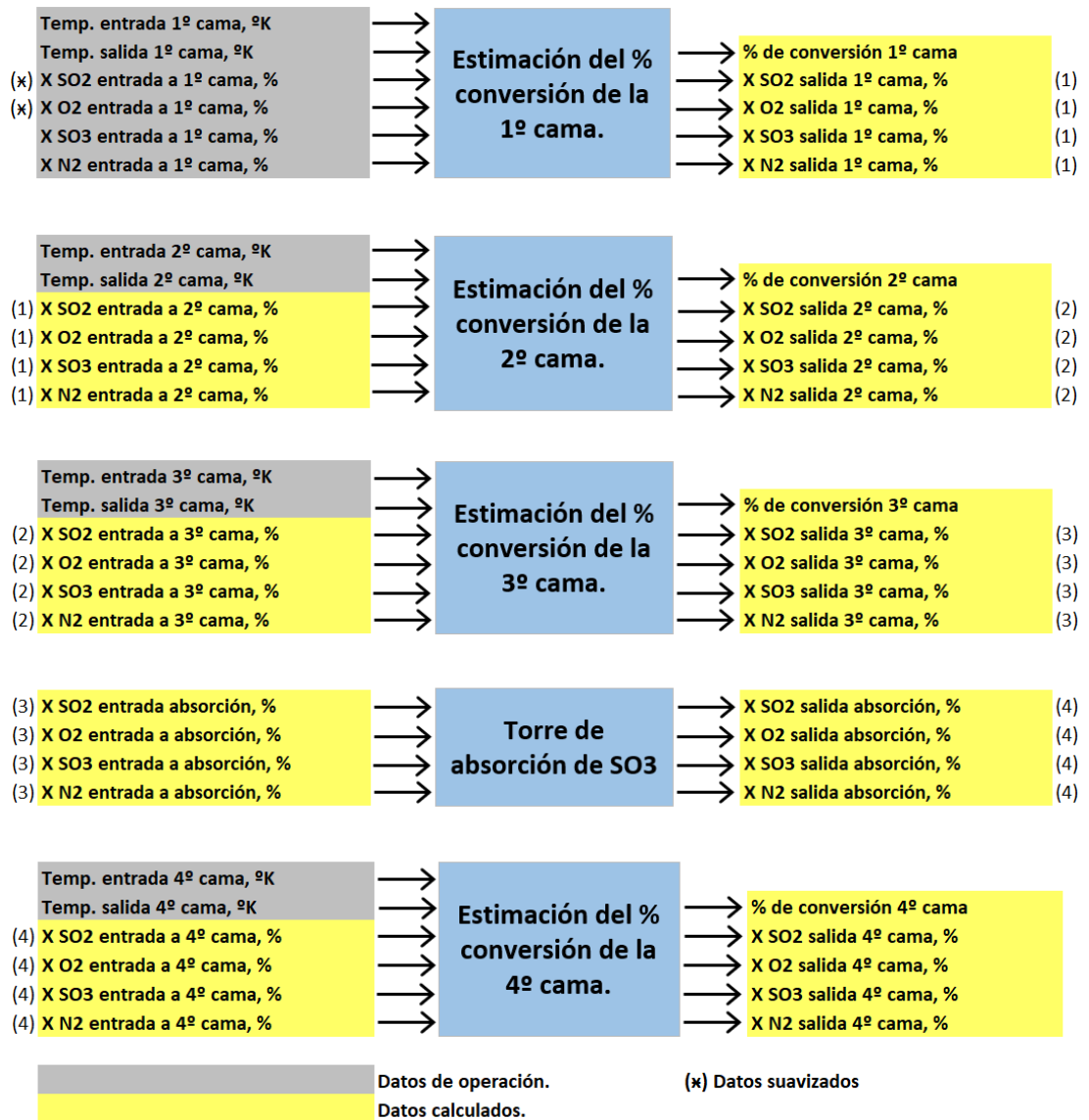


Figura 5 Metodología de cálculo de la eficiencia de conversión mediante el modelo derivado en leyes. Fuente: Elaboración propia.

2.2.4 Influencia de las variables:

2.2.4.1 El catalizador

El proceso de conversión requiere de un catalizador de pentóxido de vanadio, $V_2O_5-K_2SO_4$. El catalizador promueve que la reacción de oxidación del SO_2 se realice más rápido. La oxidación del SO_2 se lleva a cabo en la película fundida del catalizador, donde el SO_2 y el O_2 se absorben y el SO_3 se produce y desorbe rápidamente. Una película fundida comienza a formarse cuando el óxido de vanadio y los piro sulfatos de potasio, sodio, cesio se combinan para formar una capa fundida sobre el sustrato de sílice entre 660 K y 690 K. El SO_2 pasará a través del lecho sin reaccionar. El catalizador se vuelve completamente activo entre 680 K y 700 K. El gas de alimentación que ingresa al convertidor debe ser de 680 K o más; de lo contrario, el gas enfriará y solidificará la película fundida en el catalizador, desactivándolo. A temperaturas superiores a 900 K, el catalizador fundido y el sustrato sólido forman irreversiblemente un líquido viscoso inactivo (Caron S., 2019).

La actividad del catalizador es una medida de su capacidad para oxidar rápidamente el SO_2 a SO_3 . En la práctica, una ligera pérdida en la actividad del catalizador debido su envejecimiento puede compensarse operando con una temperatura de gas de alimentación más alta al ingreso del lecho.

2.2.4.2 La temperatura

Es difícil lograr un buen control de la temperatura del lecho del catalizador cuando se tratan gases de alimentación intermitentes y muy variables.

El aumento de la temperatura de los gases en la entrada del lecho aumenta la velocidad de reacción y por ende la eficiencia de conversión. Sin embargo, es importante evitar el aumento excesivo de la temperatura, puesto que podría ocasionar un sobrecalentamiento del catalizador, y para ello se realiza un control de la temperatura dentro rango especificado de diseño mediante intercambiadores de calor que enfrían los gases entre los lechos.

Deben evitarse las bajas temperaturas del gas de entrada porque pueden enfriar, solidificar y desactivar el catalizador de un lecho.

Deben evitarse las altas temperaturas del gas de entrada porque pueden elevar la temperatura de salida del gas por encima de la temperatura de desactivación del catalizador (King M., Davenport W., Moats M., 2013).

2.2.4.3 La presión

La influencia de la presión tiene poco efecto sobre las curvas de equilibrio. La temperatura de intercepción en la curva de equilibrio y el % de SO₂ oxidado aumentan ligeramente al aumentar la presión (King M., Davenport W., Moats M., 2013)

2.2.4.4 Composición del gas

La composición del gas de alimentación influye en la cantidad de SO₂ oxidado en el equilibrio. El incremento de la concentración de SO₂ y O₂ en los gases de alimentación desplaza el equilibrio de reacción hacia una mayor formación de SO₃.

De modo que, una disminución de la concentración de SO₃ mediante un proceso de absorción intermedia, permite desplazar el equilibrio de la reacción hacia una mayor formación de SO₃.

2.2.4.5 Condiciones de operación transitorias

Las condiciones de operación transitorias ocurren durante el arranque, la parada y el cambio de alimentación en la planta de ácido sulfúrico, ocasionando alteraciones en el proceso de manera desfavorable disminuyendo la eficiencia de conversión. El comportamiento dinámico del catalizador es de gran importancia en estas situaciones, ya que el catalizador requiere ciertas temperaturas mínimas para oxidar el SO₂ de manera eficiente a SO₃. (Sørensen, Christensen, 2016).

Durante el arranque los lechos de catalizador deben precalentarse por encima de su temperatura de ignición con aire caliente antes de la introducción del gas de alimentación. La temperatura de ignición representa la temperatura de entrada del gas más baja para una situación operativa adiabática especificada que mantendrá una aproximación bastante cercana al equilibrio en la salida. A esta temperatura, la mezcla de sal en el catalizador funde y disuelve parte del vanadio como

compuestos activos de vanadio (V) y permite la disolución y difusión de SO_2 , O_2 y SO_3 (Christensen, Cavalca, Beato, Helveg. 2014)

Mann R. (1986) demuestra la variación de las concentraciones de SO_2 durante las primeras horas luego del arranque de una planta de ácido sulfúrico y analiza el comportamiento de las temperaturas de salida de los lechos del convertidor, como se muestra en la figura 6. Se observa que el primer lecho llega rápidamente a su condición de diseño en aproximadamente treinta minutos. La temperatura de salida del segundo lecho desciende y luego aumenta debido al funcionamiento del sobrecalentador (superheater) entre el primer y segundo lechos.

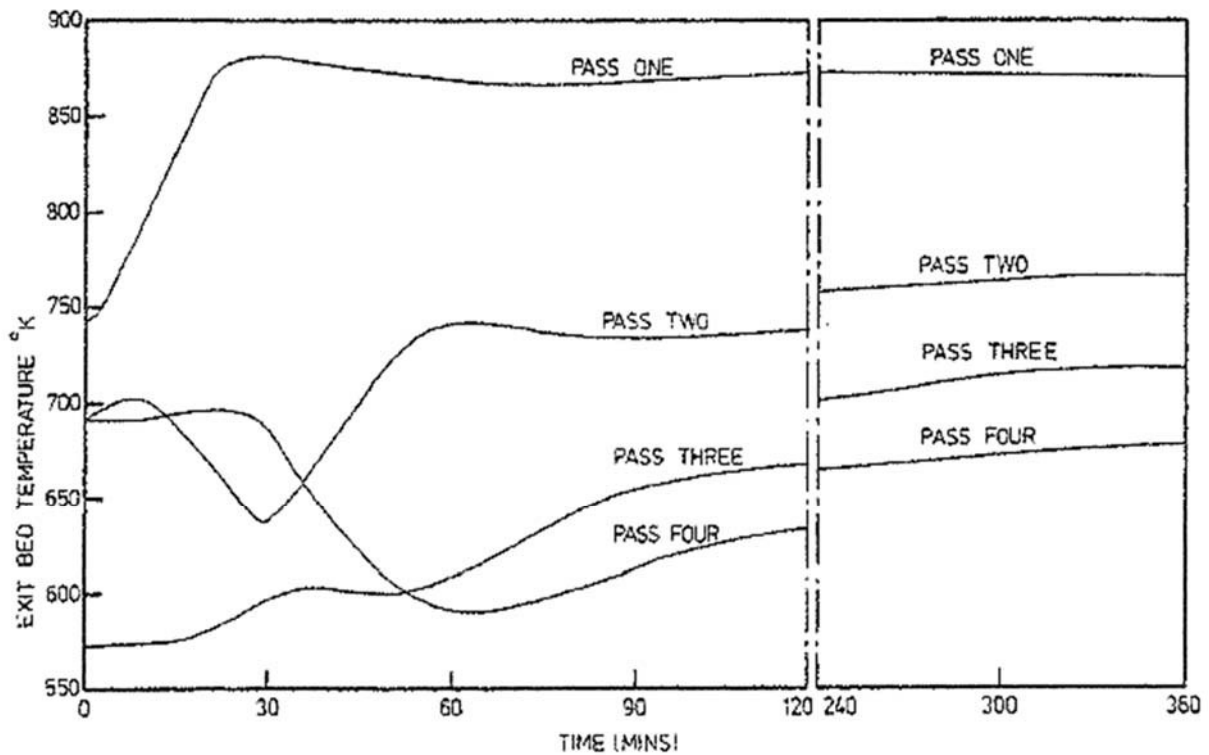


Figura 6 Variación de las temperaturas en los lechos del convertidor durante un típico arranque de una planta de ácido sulfúrico. Fuente: Mann R. (1986)

2.3 Los sensores virtuales y la simulación de procesos

Los sensores virtuales son modelos utilizados para predecir una variable de proceso, especialmente aquellas variables que no pueden ser medidas por la instrumentación física debido a su alto costo, variables que no pueden ser medido de forma continua debido al método de medición (puntual) o porque requiere un posterior análisis de la muestra en laboratorio. La figura 7 muestra un esquema general del sensor virtual.

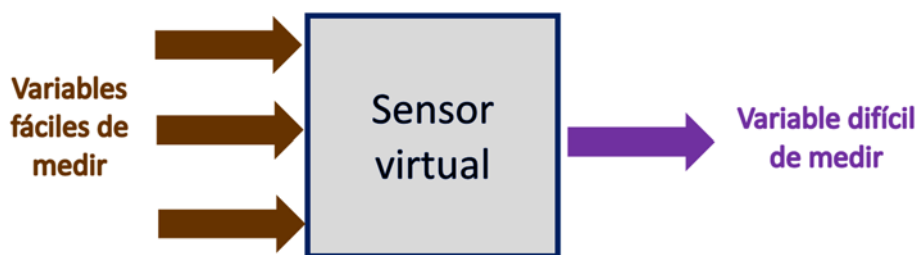


Figura 7 Esquema simple de un sensor virtual. Fuente: Elaboración propia.

Asimismo, en el campo de la simulación de procesos los sensores virtuales aportan en la predicción de variables y puede ser parte de un programa completo de simulación de procesos. La tabla N° 2 muestra las semejanzas y diferencias entre los sensores virtuales y la simulación de procesos.

Tabla 2. Algunas diferencias y semejanzas entre los sensores virtuales y la simulación.

Diferencias	Semejanzas
<ul style="list-style-type: none"> – Un sensor virtual modela una variable específica de un proceso, mientras que la simulación modela todo un sistema o proceso químico donde involucran más variables, utilizando submodelos y especificaciones fisicoquímicas de los componentes. – Un sensor virtual obedece a la necesidad de predecir el comportamiento de una variable, mientras la simulación obedece la necesidad de predecir el comportamiento de todas las variables, indicadores claves de eficiencia/rendimiento de equipos y sirve para el entrenamiento del operador. 	<ul style="list-style-type: none"> – El sensor virtual y los simuladores de proceso utilizan las mismas técnicas de modelado, ya sea mediante técnicas analíticas derivado de leyes y primeros principios, o mediante técnicas basados en datos como los modelos estadísticos y modelos de aprendizaje máquina (machine learning).

Fuente: Elaboración propia.

Durante décadas la simulación en estado estacionario ha sido utilizado en el diseño de plantas químicas, asimismo la simulación dinámica ha permitido capturar la dinámica de los procesos para conocer el comportamiento de las variables; sin embargo, actualmente se viene desarrollando simuladores de proceso avanzado de alta fidelidad que integran modelos de sensores virtuales para la simulación completa de la planta.

Los simuladores de proceso avanzados operan con información de la planta en tiempo real, utilizando modelos de sensores virtuales para predecir las variables que faltan o son difíciles de medir en el proceso. El simulador ayuda también a determinar los indicadores clave de rendimiento y eficiencia en los equipos, siendo una información que no se pueden medir directamente en la planta de procesos (Chowdhary A. R., 2014).

Un ejemplo de aplicación de simulación avanzada de alta fidelidad aplicado al proceso de endulzamiento de nafta fue desarrollado en el simulador DYNsimTM de SimSci-Esscor (Chowdhary A. R., 2014), mostrado en la figura 8.

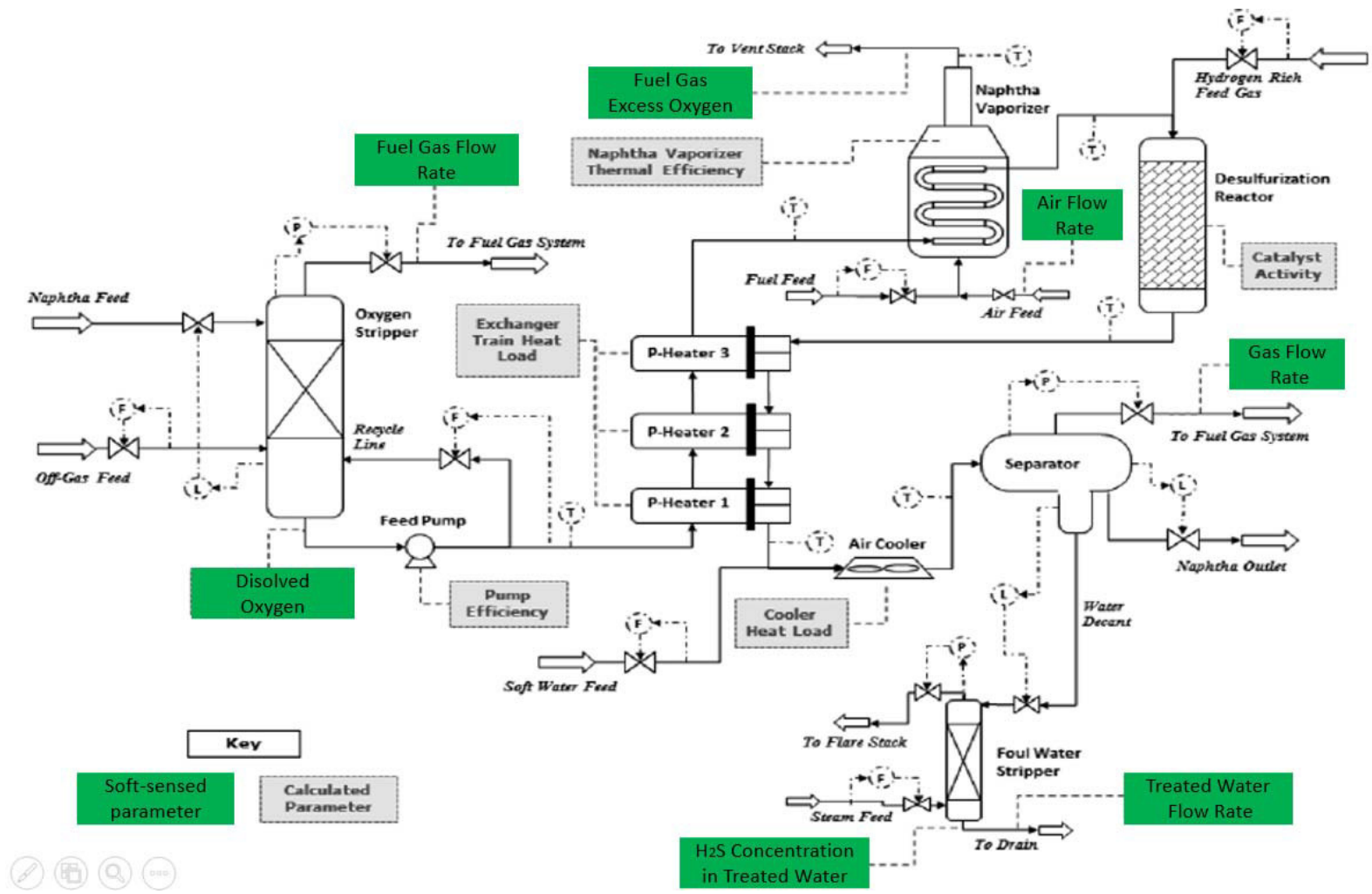


Figura 8 Ejemplo un simulador de procesos avanzado en tiempo real que integra la aplicación de los sensores virtuales (soft sensor) y también calcula parámetros de rendimiento/eficiencia para el proceso de endulzamiento de nafta. Fuente: Chowdhary A. R., 2014.

2.4 Las redes neuronales artificiales (RNAs).

Las RNAs son modelos inspirados en el funcionamiento de una neurona biológica. La neurona es la unidad fundamental del sistema nervioso mostrado en la figura 9. Cada neurona tiene un núcleo y una ramificación de salida llamado axón, donde mediante él se propagan los impulsos electroquímicos. También, tiene un gran número de ramificaciones de entrada, llamadas dendritas, por donde se propaga la señal al interior de la neurona (Isasi, Galván. 2004). La sinapsis es el punto de contacto entre dos neuronas.

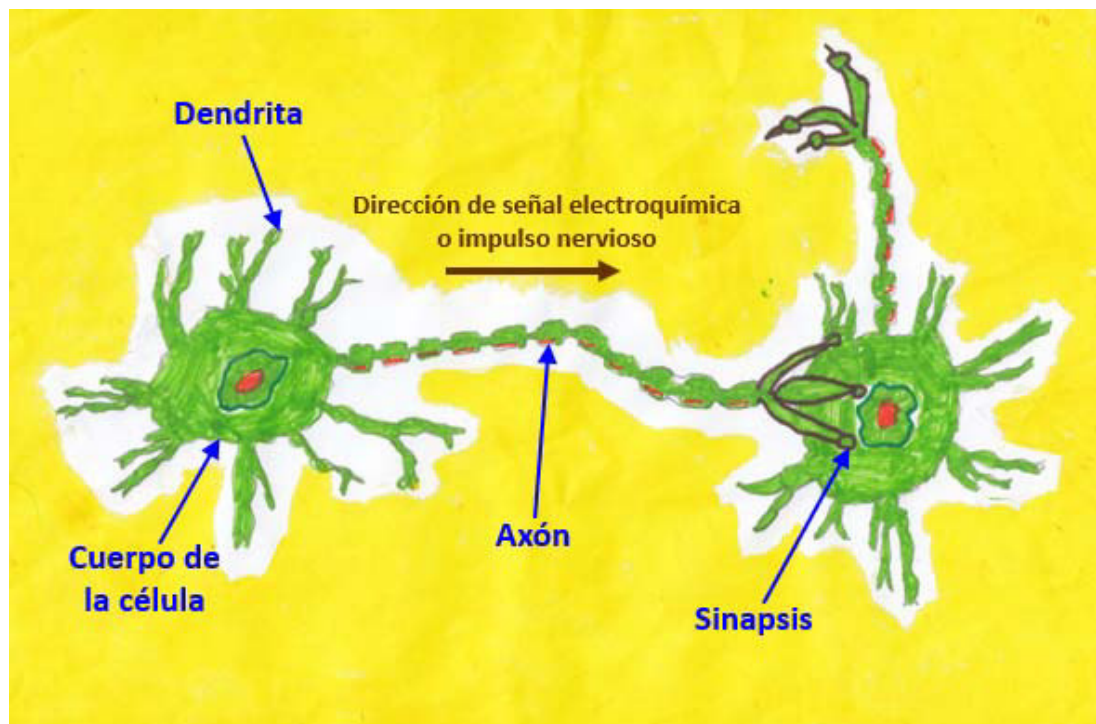


Figura 9 Neurona biológica. Fuente: Elaboración propia.

neurona artificial es un dispositivo simple de cálculo que a partir de un vector de entrada procedente del exterior o de otras neuronas, proporciona una única respuesta o salida.

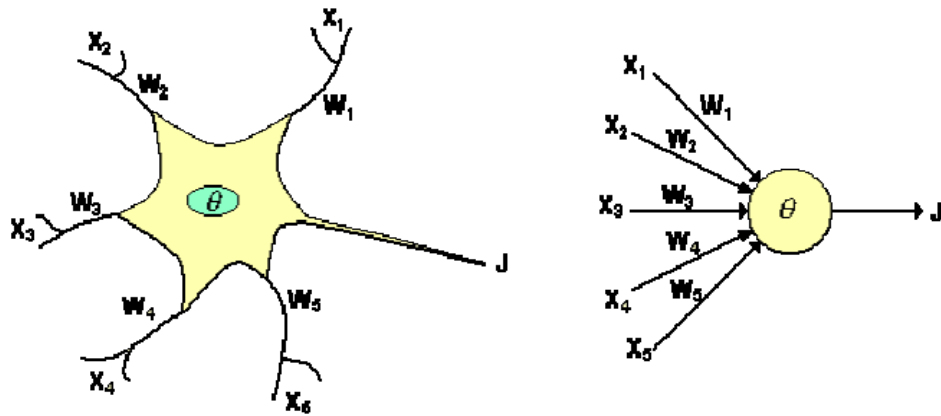


Figura 10 Neurona biológica y neurona artificial. Fuente: Reyes, 2010.

La comparación entre la neurona biológica y la neurona artificial se muestra en la figura 10. Las analogías que existen entre ambas son (Reyes, 2010):

- Las entradas x_i representan las señales provenientes de otras neuronas y que son capturadas por las dendritas.
- Los pesos w_i son la intensidad de la sinapsis que conectan dos neuronas; tanto x_i como w_i son valores reales.
- θ es la función umbral que la neurona debe sobrepasar para activarse, este proceso ocurre biológicamente en el cuerpo de la célula.

El procesamiento de la neurona se realiza primero con la sumatoria de los valores de las entradas por su peso aleatorio, luego compara este resultado con el valor del umbral, de tal manera que si lo iguala o supera envía una activación o salida. La notación de una neurona artificial está dada por los siguientes componentes mostrados en la figura 11. (Reyes, 2010):

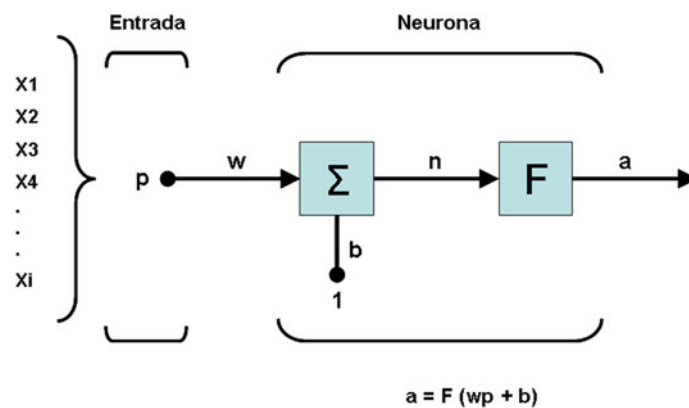


Figura 11 Neurona artificial. Fuente: Reyes, 2010.

Donde:

x_i = Variables de entrada.

p = Vector de variables de entrada x_i .

w = Vector de pesos w_i .

n = Salida de la red, es decir la salida del sumador.

b = Ganancia que refuerza la salida del sumador.

F = Función de activación que transforma “ n ” en un valor de salida “ a ”.

a = Valor de salida obtenido por la función de activación.

Las redes neuronales artificiales y la literatura estadística tienen varios de los mismos conceptos, pero generalmente con diferente terminología. En algunos casos el mismo término se usan en ambas literaturas, pero tienen diferentes significados. La tabla 1 muestra las similitudes y equivalencias de términos utilizados en el campo de la estadística y de las RNA (Sarle W., 1994 y Montaña J., 2002):

Tabla 3. *Algunas equivalencias entre la estadística y las redes neuronales*

Terminología estadística	Terminología de RNA
Observación	Patrón
Muestra	Datos de entrenamiento
Muestra de validación	Datos de validación
Variables explicativas	Variables de entrada
Variables de respuesta	Variable de salida
Modelo	Arquitectura
Residual	Error
Error aleatorio	Ruido
Interpolación	Generalización
Coefficientes	Pesos de conexión
Constante	Peso umbral

Fuente: Sarle W., 1994 y Montaña J., 2002.

2.3.1 El perceptrón

La arquitectura del perceptrón tiene una sola neurona con múltiples entradas X_p , pesos ajustables W_p , un umbral θ y la salida representado por “y”, como se muestra en la figura 12.

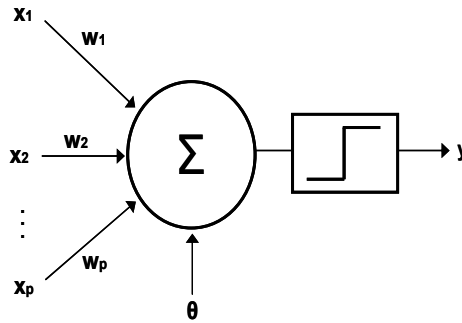


Figura 12 El perceptrón simple. Fuente: Elaboración propia.

El perceptrón tiene un parámetro denominado umbral indicado por θ . Este umbral es un factor de comparación para producir la salida. Si la suma ponderada de las entradas y los pesos superan este umbral, la salida de la red produce una señal aplicando la función de activación escalón (Isasi P. y Galván I. 2004):

$$y' = \sum_{i=1}^n w_i x_i \quad (2.5)$$

$$y = F(y', \theta) \quad (2.6)$$

$$F(s, \theta) = \begin{cases} 1 & \text{si } s > \theta \\ -1 & \text{en caso contrario} \end{cases} \quad (2.7)$$

Haciendo pasar el término θ al otro lado de la ecuación, se puede escribir la salida en una sola ecuación y el valor de F ya no depende de ningún parámetro y se introduce el valor de θ con un valor constante de -1 (Isasi P. y Galván I. 2004):

$$y = F(\sum_{i=1}^n w_i x_i + \theta) \quad (2.8)$$

$$F(s) = \begin{cases} 1 & \text{si } s > \theta \\ -1 & \text{en caso contrario} \end{cases} \quad (2.9)$$

El perceptrón obtiene una función de salida binaria de 1 y -1. Esta característica de dos dimensiones es una desventaja por la incapacidad para resolver problemas de clasificación que no sean linealmente separables (Haykin S., 1999).

2.3.2 El perceptrón multicapa

El perceptrón multicapa (MLP) permitió resolver la desventaja del perceptrón. El MLP una técnica inicialmente desarrollada por Paul Werbos (1974). Utiliza un algoritmo de entrenamiento de corrección del error conocido como retropropagación (backpropagation) que se basa en el cálculo del gradiente distribuido en los diferentes componentes de la red. Su arquitectura está compuesta de tres capas: La capa de entrada, la capa oculta y la capa de salida, mostrado en la figura 13.

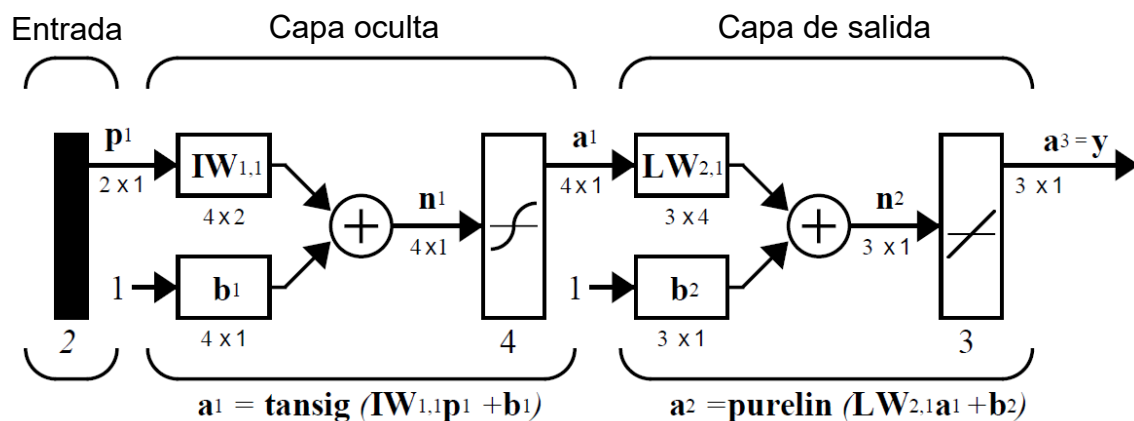


Figura 13 El Perceptrón Multicapa. Fuente: Matlab.

La capa de entrada se encarga de recibir las señales de exterior y se propagan en la capa oculta. La capa oculta realiza un procesamiento no lineal y la capa de salida permite obtener la respuesta del modelo (Isasi P. y Galván I. 2004)

En el perceptrón multicapa, las conexiones entre neuronas son siempre hacia delante y no hay conexiones laterales, por ello, la información siempre se transmite desde la capa de entrada hacia la capa de salida.

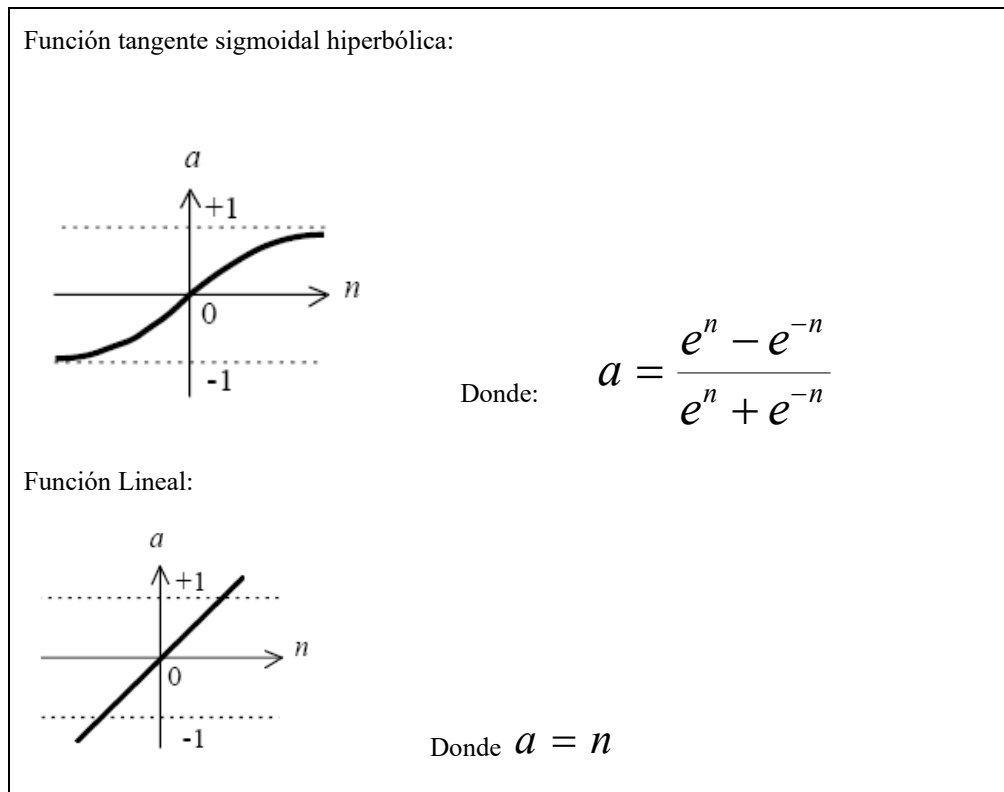


Figura 14 Funciones de activación. Fuente: Matlab.

Las funciones de activación definen la salida de la neurona en función del potencial de activación que viene a ser la salida del sumador. Existen diversos tipos de funciones de activación, sin embargo, para el presente trabajo de investigación se han empleado las funciones de activación tangente sigmoïdal hiperbólica y lineal, véase en la figura 14.

2.3.3 Número de capas ocultas

De acuerdo al número de capas ocultas, las redes neuronales de una sola capa oculta son llamados redes neuronales superficiales y las redes con más de una capa oculta son llamados redes neuronales profundas.

La cantidad de capas ocultas dependerá de la complejidad del problema. En muchas situaciones donde la desviación de la linealidad presente en los datos no es severa, con baja dimensionalidad y un número relativamente pequeño de características las RNA con una sola capa oculta producen resultados aceptables (Mhaskar H., Poggio T., 2016). Este tipo de redes posee la propiedad de ser un aproximador universal.

Podría emplearse redes neuronales con más de una capa oculta en caso de problemas más complejos, con alta dimensionalidad y gran cantidad de datos de

entrada. En los últimos años, estas redes han sido utilizados con mucho éxito en el reconocimiento de voz e imagen.

Estudios de modelos neuronales aplicados al tráfico de red malicioso comparan las redes con una y más capas ocultas. Los resultados mostraron que las redes neuronales de una capa oculta podrían identificar con más éxito que las redes neuronales profundas (Kim D., 2017).

2.3.4 Número de neuronas ocultas

Los factores que influyen en la selección del número de neuronas ocultas son (Arahal M., Berenguel M. y Rodríguez F., 2006):

- El número de patrones de entrenamiento.
- El número de neuronas de entrada y de salida
- La cantidad de ruido en el dato.
- La complejidad de la función a ser aprendida.
- El tipo de función de activación oculto.
- El algoritmo de entrenamiento.

Si hay muy pocas neuronas ocultas, podría ocurrir un subajuste (under-fitting) con un mayor error de entrenamiento y de generalización. De otra manera, si muchas neuronas ocultas son usadas el resultado sería un sobreajuste (over-fitting) con un menor error de entrenamiento, pero un mayor error de generalización. Por ello, no existe una respuesta definitiva respecto al número óptimo de neuronas ocultas, y en la mayor parte de las aplicaciones prácticas, el número de neuronas ocultas se determinan mediante prueba y error (Isasi P., Galván I., 2004), este número optimo se determina encontrando el error de validación más bajo variando el número de neuronas ocultas (Priddy K., Keller P., 2005).

2.3.5 El algoritmo de retropropagación (Backpropagation Algorithm)

El entrenamiento de una red multicapa se suele realizar utilizando un algoritmo de retropropagación de errores (Backpropagation Algorithm).

El algoritmo de retropropagación es un tipo de entrenamiento supervisado. Para realizar el entrenamiento primero se definen un conjunto de pares de patrones (X_i ,

Yi) que se usarán durante el entrenamiento y se define una función de error que permitirá comparar la diferencia entre la salida deseada y la obtenida durante el entrenamiento. Una vez obtenido dicho error se actualizan nuevamente los pesos hasta que el error alcance el mínimo requerido. Un esquema del mecanismo de entrenamiento se muestra en la figura 15.

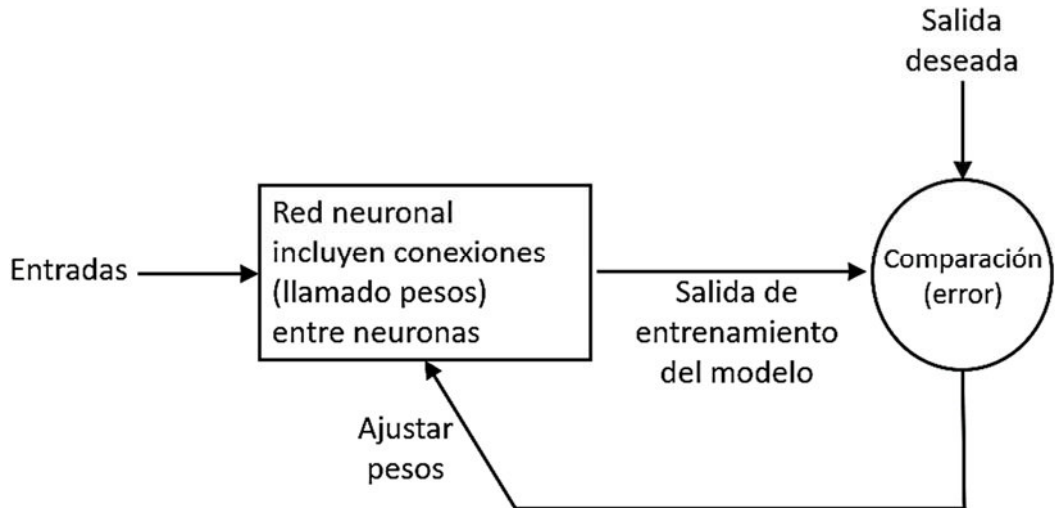


Figura 15 Mecanismo de entrenamiento supervisado. Fuente: Matlab.

El procedimiento que se emplea para minimizar el error es el *descenso en la dirección del gradiente*: una manera muy eficiente de implementarlo es a través de un procedimiento equivalente a computar la red hacia atrás. Esto da lugar al algoritmo de retropropagación:

$$E_T = \sum_{p=1}^P E_p = \frac{1}{2} \sum_{p=1}^P (d_p - O_p^S)^2$$

Donde:

E_T = Error total de salida.

E_p = Error de la salida p.

p = Número de neuronas de la última capa.

O_p^S = Salida obtenida en la neurona p de la capa S (la de salida).

d_p = Salida esperada en la neurona p.

El algoritmo de aprendizaje es:

$$\omega_{ij}^L(k+1) = \omega_{ij}^L(k) - \mu \frac{\partial E_T}{\partial \omega_{ij}^L(k)} \quad (2.10)$$

Donde:

$w_{ij}^L =$ Peso de la entrada i de la neurona j en la capa L .

$k =$ Iteración actual.

$\mu =$ Factor de aprendizaje.

La derivada del error es equivalente a:

$$\frac{\partial E_T}{\partial w_{ij}^L(k)} = \sum_{p=1}^P \frac{\partial E_p}{\partial w_{ij}^L(k)} \quad (2.11)$$

La salida es una función sigmoidea de la combinación lineal de las variables de entrada:

$$O_j^L = \frac{1}{1 + e^{-y_j^L}} \quad ; \quad y_j^L = \sum_{i=1}^{N_L} \omega_{ij}^L \cdot O_i^{L-1} \quad (2.12)$$

Donde:

$N_L =$ Número de neuronas de la capa L .

$O_j^L =$ Salida de la neurona j de la capa L .

$y_j^L =$ Estado de activación de la neurona j de la capa L .

Las redes multicapa alimentadas hacia delante (feed-forward) en ocasiones se denominan redes de retropropagación (backpropagation networks) porque el algoritmo de entrenamiento que se utiliza está basado en la retropropagación del error (Berzal F. 2018)

2.3.6 Algoritmos de segundo orden.

Existen diversos métodos alternativos al de retropropagación dirigidos a estimar los pesos de la red de una forma mucho más rápida y eficaz. Muchos de estos métodos alternativos son extensiones de la propia técnica de retropropagación.

El algoritmo de retropropagación utiliza la técnica de descenso del gradiente, mediante el cálculo de las derivadas de la superficie de error respecto de cada peso

$\frac{\partial E_T}{\partial \omega_{ij}^L(k)}$ obteniéndose la dirección de máxima pendiente de la superficie del

error y es utilizado para actualizar los pesos; pero no se puede asegurar que mediante esta técnica se logre camino más rápido hacia el mínimo (Martin del Brío B., Sanz A. 2002).

Por ello, el hacer uso de la derivada de segundo orden nos permite informar acerca de la curvatura de dicha superficie. Así, si la primera derivada representa la velocidad de decremento, la segunda derivada representa la deceleración del error.

Los algoritmos de segundo orden no solo hacen uso de la información de la pendiente, sino también la información del ritmo de cambio de la pendiente (Martin del Brío B., Sanz A. 2002). Entre algunos algoritmos de segundo orden tenemos de Levenberg-Marquardt, gradientes conjugados, gradientes conjugados escalados, Newton y cuasi-Newton.

2.3.7 Inicialización de la red.

Los pesos se encuentran conectados entre las neuronas y son donde las RNAs almacenan la información para describir el problema y cuantificar las interdependencias entre las entradas y salidas. La meta del proceso de entrenamiento es el calcular los valores de los pesos que corresponden al valor mínimo local de la función de error. En las redes neuronales del tipo alimentadas hacia delante (feedforward neuronal networks), esta función es usualmente la suma de los cuadrados de la diferencia entre los datos actuales y los calculados por las RNA, y siendo los pesos parámetros desconocidos. Los valores iniciales de los pesos determinan el punto de inicio del algoritmo de entrenamiento y directamente afecta la velocidad de entrenamiento y el error de entrenamiento. Si estos valores iniciales no son cercanos al mínimo global o están cercanos a un área con muchos mínimos locales de la función superficie, el algoritmo de entrenamiento podría ser atrapado en uno de los mínimos locales y por lo tanto el entrenamiento sería muy lento y produciría malos resultados. Para evitar tales problemas, la inicialización de pesos se efectúa al azar. Esto resulta de la necesidad de entrenar las mismas redes neuronales en múltiples tiempos, usando diferentes inicializaciones, para

asegurar que el rendimiento del modelo neuronal dependa fundamentalmente de la arquitectura y no de los valores iniciales de los pesos. Consecuentemente, el diseñador del modelo se ve envuelto en un proceso repetitivo que requiere más tiempo siendo importante la experiencia e intuición.

Hay diferentes métodos de inicialización. El método de Nguyen - Widrow propone realizar la inicialización de los pesos en el intervalo de $-0,5$ a $0,5$ y usando una distribución uniforme. La ventaja de este método comparado con el método de inicialización puramente al azar es que pocas neuronas son desechadas y el entrenamiento es más rápido porque el área de entrada tiene más neuronas asociadas. Estudios realizados por Pavelka A. y Procházka A. (2004) compararon diez métodos de inicialización de pesos en modelos de redes neuronales multicapa, obteniendo mejor resultado con el método de Nguyen-Widrow.

2.3.8 Generalización de la red

La capacidad de generalización se define como la capacidad de la red en responder correctamente ante patrones diferentes o nuevos; para esto la red debe extraer las características de las muestras durante el proceso de aprendizaje.

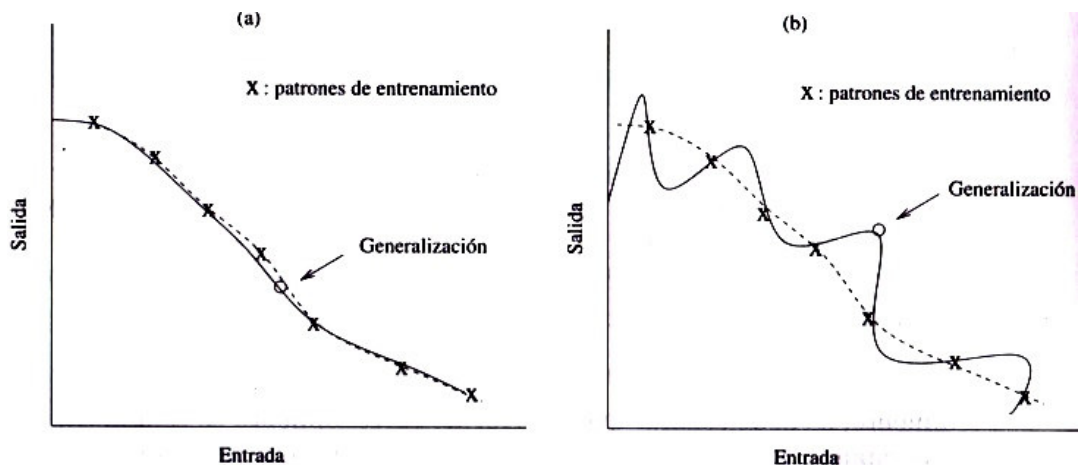


Figura 16 Tipos de generalización: (a) Con buenas propiedades; (b) Con escasas propiedades. Fuente: Isasi y Galván, 2004.

La figura 16 muestra dos tipos de generalización. Los puntos con cruces son los patrones de entrenamiento; la línea punteada son la función que se debe aproximar y la línea continua la salida del perceptrón luego del entrenamiento. La figura 16a muestra una buena generalización del modelo y la figura 16b muestra una mala

generalización producto de un entrenamiento inapropiado o sobre entrenamiento. (Isasi y Galván, 2004)

El sobreentrenamiento ocasiona un sobreaprendizaje (*overfitting*), anulando la capacidad de generalización de la red. En este caso, el error de entrenamiento se lleva a valores muy pequeños, pero en la etapa de prueba final utilizando nuevos datos el error que se obtiene es grande. La razón se debe a que en este caso la red memoriza los ejemplos de entrenamiento, pero no aprende a generalizar nuevas situaciones. Por ello, a fin de evitar este sobre entrenamiento se utiliza en paralelo un conjunto de datos de validación a fin de evaluar el error durante el entrenamiento. Un método utilizado para mejorar la generalización es la detención temprana (*early stopping*). En esta técnica se toma una muestra de datos y se dividen en tres conjuntos aleatorios: Conjunto de entrenamiento, conjunto de validación y conjunto de prueba final, como se muestra en la figura 17.

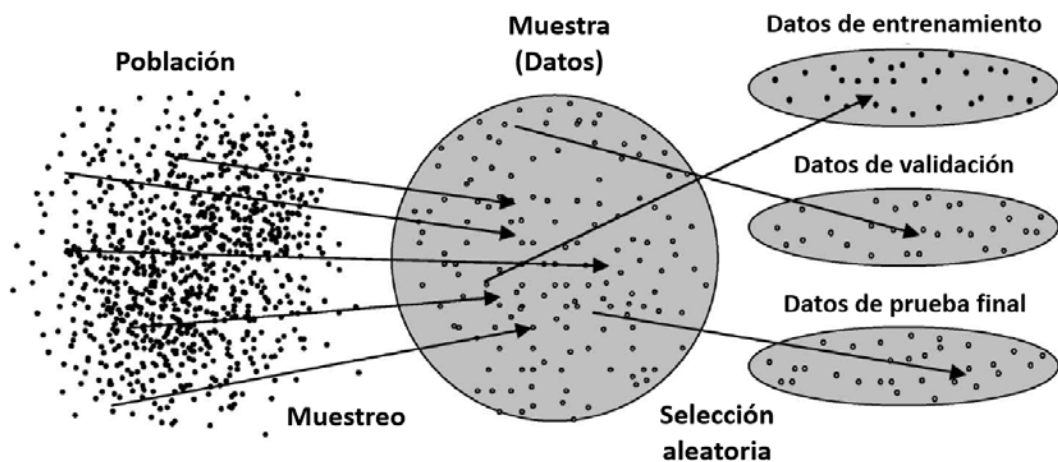


Figura 17 División de conjuntos de datos de entrenamiento, validación y prueba final. Fuente: Priddy K., Keller P., 2005.

Los porcentajes de división con respecto a la muestra, en el caso de los datos de entrenamiento varían entre 50% y 80%, mientras que los datos de validación y de prueba final generalmente tienen igual proporción, ejemplos: 15%-15%, 10%-10%, etc.

El primer conjunto de datos es el conjunto de entrenamiento, que se usa para computar el gradiente y actualizar los pesos y tendencias (bias) de la red. El segundo conjunto es el conjunto de validación. La figura 18 muestra como el error

del conjunto de validación es monitoreado durante el proceso de entrenamiento. El error de validación normalmente disminuye durante la fase inicial de entrenamiento, así como el error de entrenamiento. Sin embargo, cuando la red inicia el sobre ajuste de los datos, el error de validación normalmente empieza a aumentar. Cuando el error de validación incrementa para un número especificado de iteraciones, el entrenamiento se detiene y se retorna los pesos y tendencia en el mínimo error de validación, es decir en el punto de detención temprana.

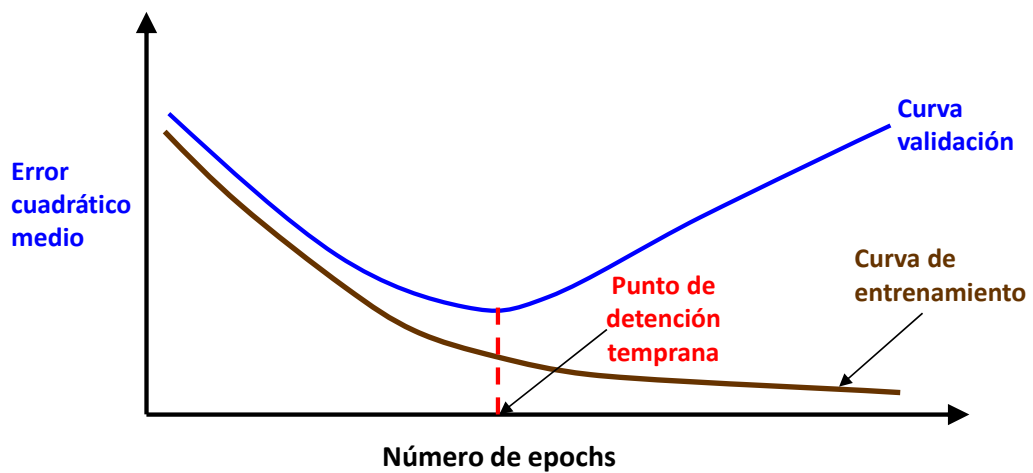


Figura 18 Ilustración de la regla de detención temprana basada en la validación durante el proceso de aprendizaje. Fuente: Elaboración propia.

A diferencia del conjunto de entrenamiento y validación que están sesgados, el conjunto de prueba final es un conjunto de prueba independiente que se utiliza para determinar el error de generalización. Usualmente los resultados de la prueba final se utilizan para reportar la eficacia del modelo.

El conjunto de prueba final nunca debe usarse para elegir la mejor configuración de la red neuronal, de modo que siga siendo una estimación no sesgada del error de generalización de la red (Priddy K., Keller P., 2005).

CAPÍTULO III: MARCO METODOLÓGICO

3.1 Introducción

En este capítulo se describe la metodología empleada para el diseño del sensor virtual basados en RNAs. La figura 19 muestra en resumen la metodología empleada en el diseño del sensor virtual para la predicción de la eficiencia de conversión, se puede simplificar en dos etapas: Primero aplicando la técnica de las RNAs se predice la concentración de SO₂ de salida, una vez determinado, se calcula la eficiencia de conversión mediante la ecuación 2.2.

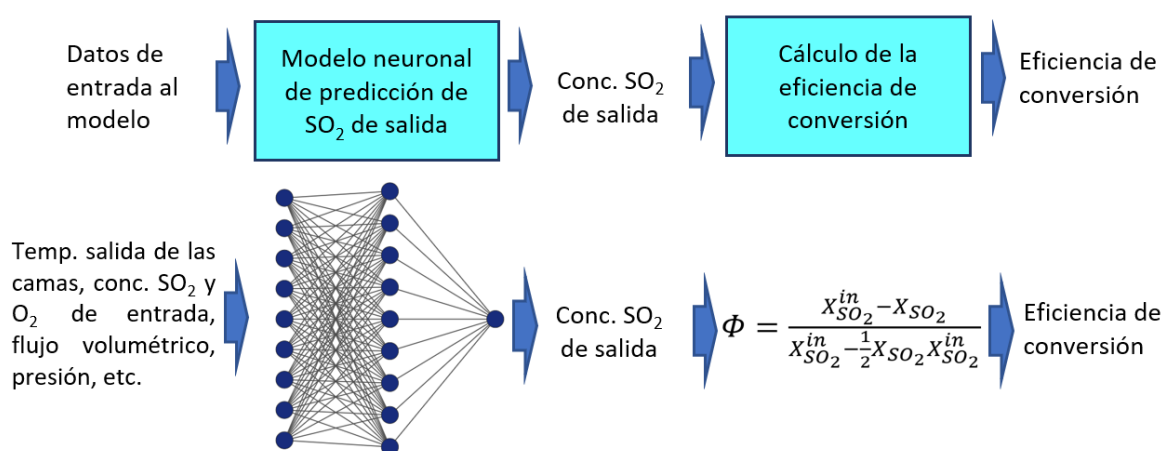


Figura 19 Metodología del sensor virtual para el cálculo de la eficiencia de conversión. Fuente: Elaboración propia.

El estudio se realizó en una computadora de escritorio con procesador AMD Ryzen 5 3400G, tarjeta gráfica Radeon Vega Graphics 3.7 GHz, memoria RAM de 8.00 Gb y disco duro de 1 terabyte. Se utilizó el sistema operativo Windows 10 Pro © 2019 Microsoft Corporation.

Los entrenamientos de las redes neuronales se realizaron mediante los softwares Matlab R2018b (Matlab 9.5) y Excel office 2019. Se habilitó en Excel el complemento *Excellink* que permite operar el Matlab desde una hoja de cálculo o macro de Excel, es decir, intercambiando y sincronizando datos entre Excel y MATLAB sin salir del entorno de Excel. Un esquema de este complemento se muestra en la figura 20.

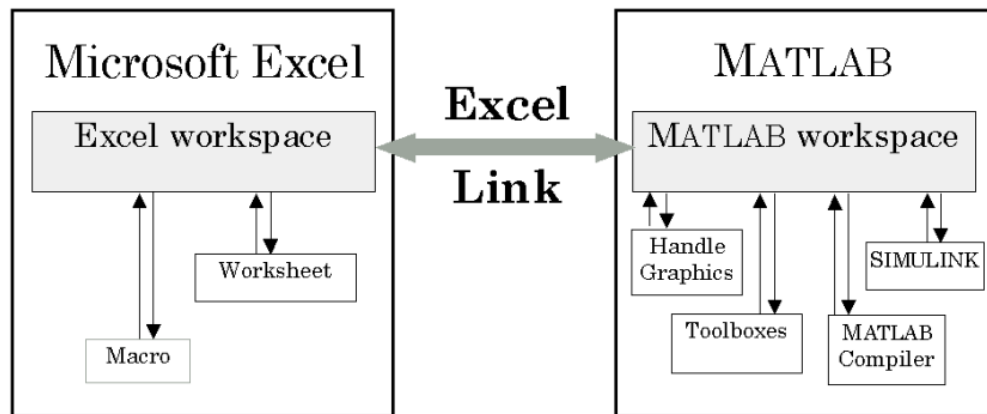


Figura 20 Complemento Excellink de Matlab. Fuente: Matlab.

Las etapas de diseño del modelo neuronal para la predicción de la concentración de SO_2 de salida es mostrado en la figura 21:

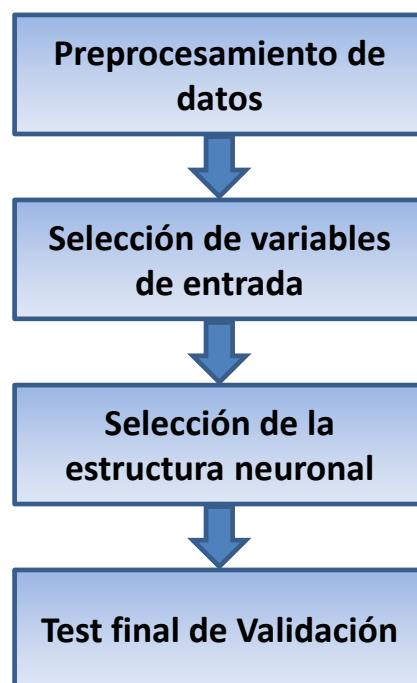


Figura 21 Etapas de diseño del modelo neuronal para la predicción de la concentración de SO_2 de salida. Fuente: Elaboración propia.

3.2 Preprocesamiento de datos

Se debe de tomar atención a los datos utilizados, ya que podrían llevar a elaborar un modelo con una pobre capacidad de estimación y con malas propiedades de generalización. Para asegurarse que el conjunto de datos brinde una representación precisa del proceso, se

debe examinar en todo su conjunto. Asimismo, se recomienda que el conjunto de datos cumpla con las siguientes pautas:

- Debe tener una buena representación del proceso. Para esto debe tener un tamaño razonable de mediciones, debe tener un suficiente rango que el proceso puede asumir y debe considerar un gran número de cambios del proceso (Roffel y Betlem, 2006, p. 292).
- La razón de la señal y el ruido debería ser más de cinco. Esto significa que los cambios de la variable de un proceso en particular deberían ser mínimo cinco veces el valor del ruido del proceso para aquel variable (Roffel y Betlem, 2006, p. 297).

Para una señal la magnitud del ruido es convenientemente definido como la desviación estándar de numerosas mediciones de la señal, y la señal esta dado por el promedio de las mediciones (Smith S. 1999), la ecuación es la señal/ruido es:

$$\frac{\text{Señal}}{\text{Ruido}} = \frac{\text{Promedio}}{\text{Desviación estandar}}$$

La figura 22 muestra dos ejemplos de datos con diferentes relaciones señal/ruido. En el caso (a) la relación señal/ruido es igual a $0.5/1 = 0.5$ y en el caso (b) la relación señal/ruido es igual a $3/0.2 = 15$.

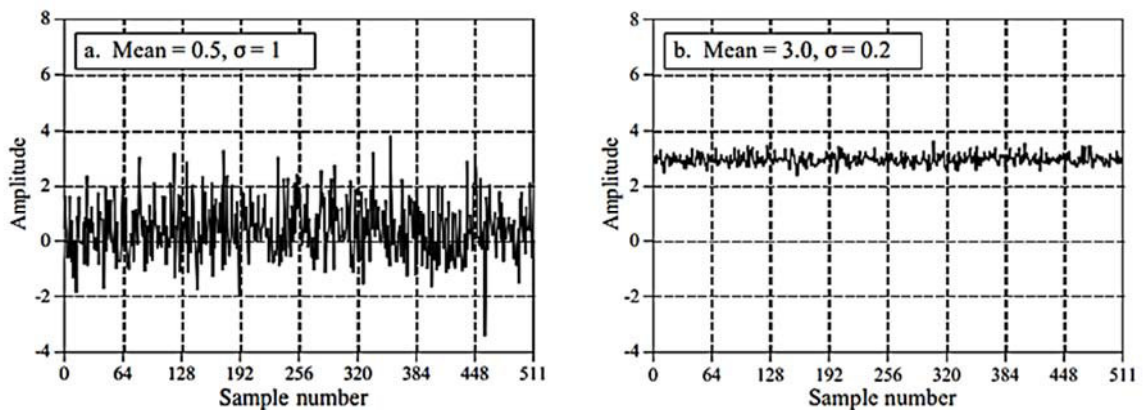


Figura 22 Ejemplos de dos señales con diferentes promedios y desviaciones estándar. Fuente: Smith S. 1999.

El modelo podría dar resultados erróneos cuando el ruido es demasiado grande en comparación con la señal de salida (Marlin T., 2015).

- Debe tener una buena distribución estándar o uniforme.

Estos requisitos son aconsejables, aunque no imprescindibles, ya que la RNA en algunos casos podría dar buenos resultados, pero invertirá un tiempo mayor de uno a dos órdenes de magnitud en conseguirlo.

La calidad del modelo de predicción dependerá de la calidad de los datos. En algunas ocasiones la preparación de los datos toma más del 80% del tiempo de preparación del modelo y es necesario además tener un buen conocimiento del caso de estudio a modelar. Las tareas principales a realizar en el preprocesamiento de los datos se muestran en la figura 23.

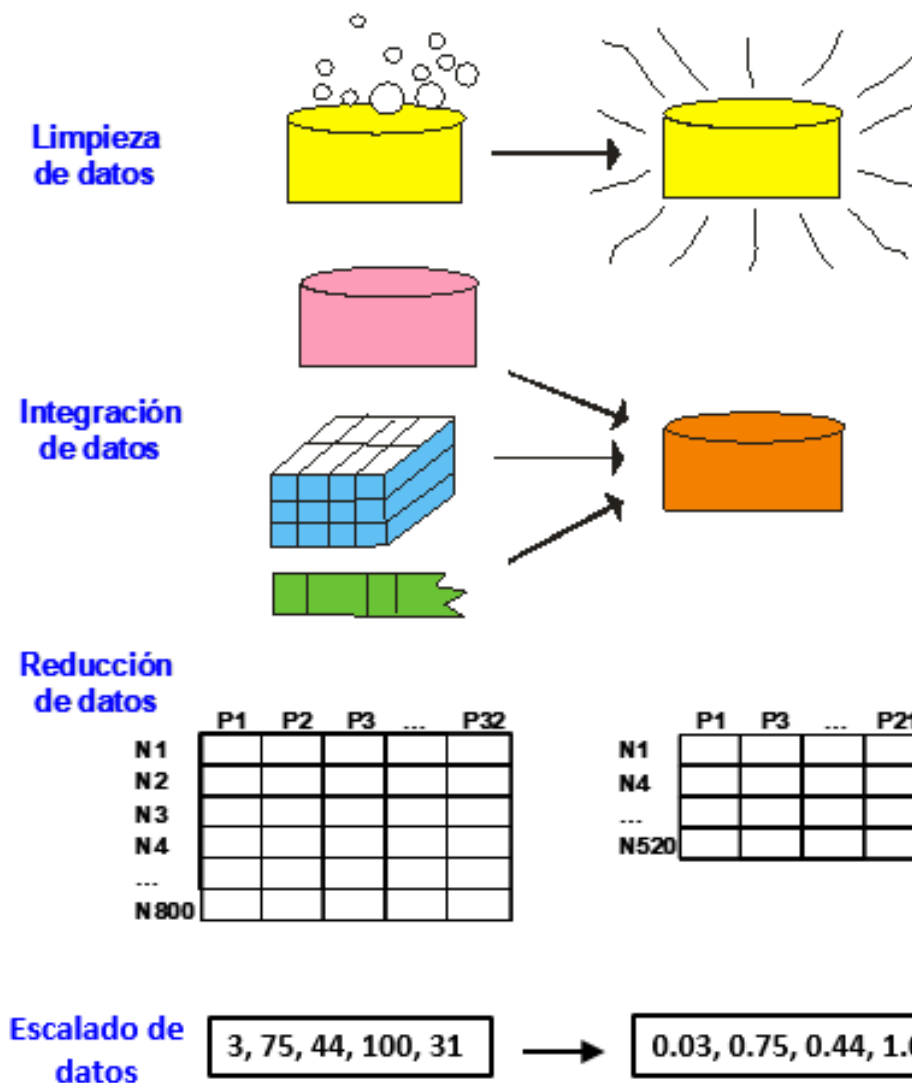


Figura 23 Tareas principales en el preprocesamiento de los datos. Fuente: Adaptado de Han J., Kamber M. (2006)

3.2.1 La limpieza de datos

Se debe realizar la limpieza de los datos debido a falsas medidas. Esta tarea podría involucrar el llenado de los datos faltantes, suavizado de los errores de los datos, la eliminación o corrección de datos inconsistentes (Han J., Kamber M. 2006).

Es importante no confundir con falsas medidas la aparición de valores inusuales o poco frecuentes. (Arahal M., Berenguel M. y Rodríguez F. 2006).

Las causas de estas falsas medidas podrían deberse al mal funcionamiento de los sensores, o proceden de registros de datos durante las actividades de mantenimiento y verificación de los sensores.

Una técnica para la detección de datos atípicos es la regla de tres sigmas (tres veces la desviación estándar). El experto con conocimiento del proceso toma la decisión final de considerar como dato válido (inusual) o descartarlo como falsa medida.

Los datos de eventos inusuales representan una fuente preciosa de información para el modelo (Fortuna et al., 2007).

3.2.2 La integración de datos

Cuando los datos proceden de diferentes bases de datos, estos requieren ser combinados dentro de un almacenamiento coherente en función del tiempo y lugar de muestreo.

Durante la integración de múltiples bases de datos podría ocurrir la redundancia entre las variables, esto sucede cuando la variable tiene diferentes nombres en diferentes bases de datos o cuando la variable es derivada de otra variable de la misma base de datos. La redundancia se puede detectar mediante un análisis de correlación entre las variables (Han J., Kamber M., 2006).

3.2.3 Reducción de datos

La reducción de datos se realiza con el fin de obtener una representación reducida en volumen, pero produciendo los mismos resultados o similares en el análisis (Clifton C., 2004). Durante la reducción de datos se puede detectar y eliminar aquellos datos que sean poco relevantes, repetitivos o redundantes.

3.2.4 Escalado de datos:

Para modificar el rango de valores de las entradas se aplica un escalamiento o también llamado normalización de datos. El objetivo es acotar sus valores entre $[0,+1]$ o entre $[-1,1]$, como límites para la función de activación del modelo neuronal, y evitar la saturación de esta al realizar las sumas ponderadas de las entradas. Para ello se busca el mínimo y el máximo del rango de las entradas y se transforma el intervalo en forma lineal.

3.3 Selección de las variables de entrada

Se debe elegir cuidadosamente para reducir el número de las variables de entrada a utilizar en el modelo, véase la figura 24.

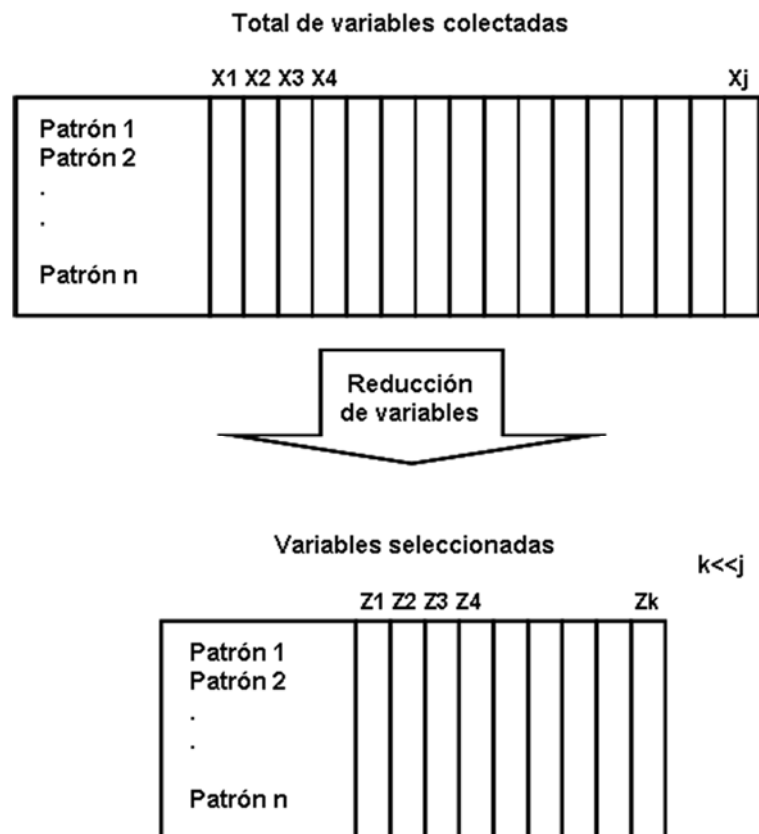


Figura 24 Reducción de las variables de entrada al modelo. Fuente: Colmenares, 2004.

Un excesivo número de variables podría traer las siguientes desventajas:

- Requiere mayor número de casos de entrenamiento.
- Requiere en su estructura un mayor número de nodos, número de pesos y mayor consumo de tiempo computacional.

- Requiere mayor esfuerzo humano excesivo conducente a múltiples pruebas por ensayo y error. Esto haría difícil la interpretación de los resultados.

Un mayor número de variables puede introducir ruido en el modelo que oculte las variables más significativas y aumentando el error. Por otra parte, un menor número de variables podría provocar una falta de información.

Los procesos químicos tienen un proceso dinámico y se comportan de manera similar a una serie de tiempos, ocasionando en muchos casos la ocurrencia de tiempos de retardo entre las variables, es decir, cuando ocurre un cambio de una variable, este no tiene un impacto instantáneo en otro punto operativo y se requiere de un lapso.

Esto significa que se debe hacer una estimación del tiempo que toma antes de que un cambio en la entrada alcance su mayor impacto en la salida. Esto podría determinarse mediante un análisis de correlación de Pearson entre las variables a diferentes tiempos de retardo.

Para estimar la relación entre las variables se puede aplicar el coeficiente de correlación de Pearson, mostrado en la ecuación (3.1).

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (3.1)$$

Donde: r = Coeficiente de correlación de Pearson.

X_i, Y_i = Valor individual de las variables.

\bar{X}, \bar{Y} = Valor promedio de las variables.

Para interpretar el resultado del coeficiente de correlación y su significado, se clasifica de acuerdo con la tabla 4.

Tabla 4. Significados del coeficiente de correlación de Pearson.

Valor	Significado
-1	Correlación negativa perfecta
-0,9 a -0,99	Correlación negativa muy alta
-0,7 a -0,89	Correlación negativa alta
-0,5 a -0,69	Correlación negativa moderada
-0,3 a -0,49	Correlación negativa baja
-0,01 a -0,29	Poca o ninguna correlación
0	Correlación nula
0,01 a 0,29	Poca o ninguna correlación
0,3 a 0,49	Correlación positiva baja
0,5 a 0,69	Correlación positiva moderada
0,7 a 0,89	Correlación positiva alta
0,9 a 0,99	Correlación positiva muy alta
1	Correlación positiva perfecta

Fuente: Adaptado de Hinkle D., Wiersma W., Jurs S. (2013). Applied statistics for the behavioral sciences.

3.4 Selección de la estructura neuronal

El objetivo principal es seleccionar la estructura neuronal que permita obtener la máxima precisión alcanzable en el modelo de predicción. Es importante considerar un número mínimo de parámetros que represente mejor el sistema y con un bajo tiempo computacional. Entre los parámetros están el número de capas ocultas, número de nodos en la capa oculta, número de entradas en el modelo, las funciones de activación de los nodos, etc.

Para elegir cuidadosamente la arquitectura más adecuada, se debe de tomar en cuenta la dimensionalidad del problema, la cantidad y la calidad de los datos disponibles, así como también la experiencia del usuario. Hoy en día, no existe un marco único para la selección de la mejor estructura neuronal y en la mayoría de los casos se inicia a partir de una estructura elegida y se acude a procedimientos de ensayo y error variando el número de capas, el número de variables de entrada y el número de nodos en la capa oculta.

La selección del número de neuronas en la capa oculta no es una tarea simple. Una red con muy pocas neuronas ocultas proporcionará una estimación lineal de la tendencia real, mientras que demasiadas neuronas ocultas harán que la red siga el ruido debido a la sobre parametrización. A pesar de que hay muchas reglas generales, lo recomendable es comenzar con un número pequeño de neuronas ocultas y aumentar gradualmente este

número hasta cumplir con los requisitos de precisión del modelo a fin de obtener la estructura óptima (Roffel, Betlem, 2006).

En el presente estudio el número de variables de entrada y el número de unidades en la capa oculta se determina mediante pruebas de ensayo y error, a fin de optimizar la estructura del modelo, encontrando aquella que permita predecir con el menor error.

3.5 Prueba final de validación

El objetivo de la validación es establecer la credibilidad de un modelo para un propósito especificado. La validación es importante para ver si el modelo puede predecir un caso real, utilizando datos diferentes a los usados en la calibración del modelo.

Actualmente no existe ningún protocolo para la selección de indicadores de evaluación y comparación de modelos. Entre los indicadores más utilizados la mayoría de los estudios optan por el uso del coeficiente de correlación, el error cuadrático medio, el error absoluto medio, el error absoluto medio normalizado, el índice refinado de concordancia y el sesgo. Para la evaluación del presente estudio se ha optado por utilizar el coeficiente de correlación, el error porcentual medio absoluto y el índice refinado de concordancia.

Utilizamos el coeficiente de correlación para evaluar la linealidad entre el valor real y el valor estimado, pero no la concordancia entre ellas. Para ello, utilizamos el error porcentual medio absoluto y el índice refinado de concordancia que mide el grado del error en la estimación del modelo, y no como una medida de correlación entre los valores reales y estimadas.

- **Error porcentual medio absoluto (MAPE):** Es el promedio del error porcentual absoluto entre el valor real y el valor estimado, expresado como un porcentaje de los valores reales. Se utiliza la siguiente ecuación:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|O_i - P_i|}{|O_i|} \quad (3.2)$$

Donde:

n = Número de datos.

O_i = Valor real.

P_i = Valor estimado.

EA = Error porcentual absoluto

- **Índice refinado de concordancia (IOA):** Se utiliza la siguiente ecuación propuesta por Willmott et al., (2012):

$$IOA = \left\{ \begin{array}{l} 1 - \frac{\sum_{i=1}^n |P_i - O_i|}{2 \sum_{i=1}^n |O_i - \bar{O}|}, \text{ Cuando } \sum_{i=1}^n |P_i - O_i| \leq 2 \sum_{i=1}^n |O_i - \bar{O}| \\ \frac{2 \sum_{i=1}^n |O_i - \bar{O}|}{\sum_{i=1}^n |P_i - O_i|} - 1, \text{ Cuando } \sum_{i=1}^n |P_i - O_i| > 2 \sum_{i=1}^n |O_i - \bar{O}| \end{array} \right\} \quad (3.3)$$

Donde:

n = Número de datos

O_i = Valor real.

\bar{O} = Promedio de los valores reales.

P_i = Valor estimado.

IOA = Índice refinado de concordancia.

El rango del IOA , es de -1 a 1. Un IOA de 1 indica una concordancia perfecta entre el modelo y la observación, y un IOA de -1 indica falta de acuerdo entre el modelo y la observación.

CAPITULO IV: RESULTADOS Y DISCUSIÓN

4.1 Resultados del sensor virtual basado en RNA

4.1.1 Colección y filtrado de datos

Los sensores de medición se encuentran ubicados en el área de limpieza de gases y el área de contacto de la planta de ácido sulfúrico, asimismo cuentan con un programa de mantenimiento y verificación. Los datos son colectados mediante el software de visualización y acceso a datos *ActiveFactory*. Los datos corresponden a un periodo de tiempo de doce días y son datos promedio de un minuto. Durante este periodo se cuenta con información ideal para el estudio ya que incluyen episodios de operación normal, parada y puesta en marcha a diferentes tiempos de parada, como se muestra en la figura 25.

El resumen de las etapas de diseño del modelo neuronal para la predicción de la concentración de SO₂, así como el software empleado en cada etapa son mostrados en la figura 26.

Los códigos fuentes de los toolbox de Matlab aplicados a redes neuronales y la interfase para trabajar con Excel se describen en el anexo 2.

El programa desarrollado en VBA de Excel para el modelo de predicción de la concentración de SO₂ de salida se muestra en el anexo 3.

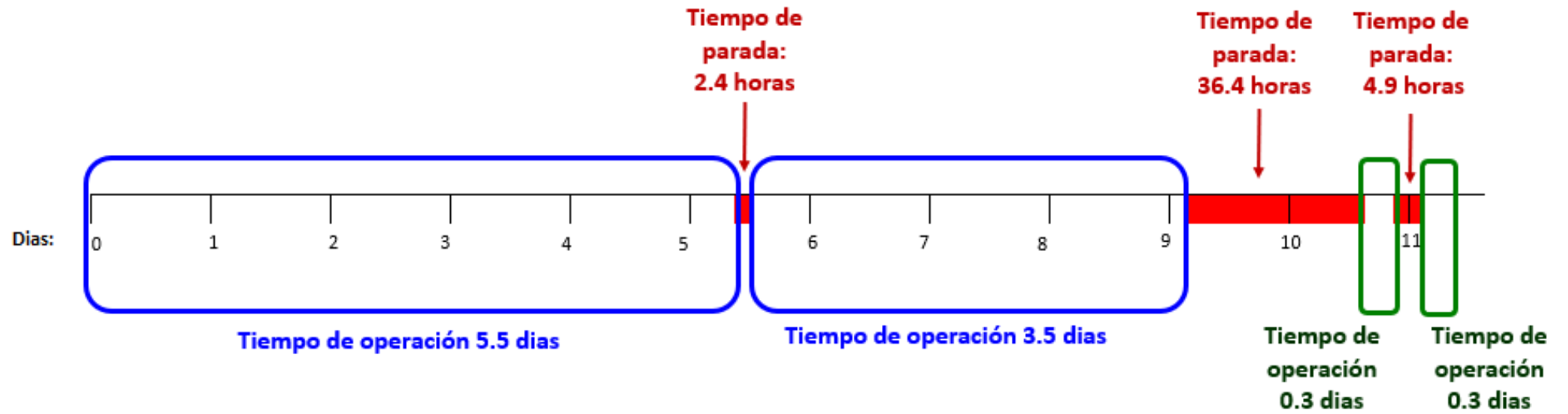


Figura 25 Esquema que muestra el periodo de tiempo de los datos colectados, en condiciones de operación normal (azul) y condiciones transitorias (verde). Fuente: Elaboración propia.



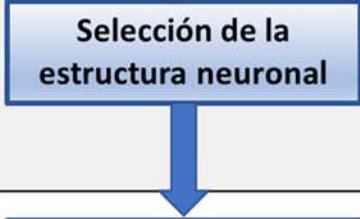

Descripción	Software	Etapas de diseño
<ul style="list-style-type: none"> • Limpieza de datos (falsas medidas) • Integración de datos (De diferentes bases de datos) • Reducción de datos (Poco relevantes repetitivos o redundantes) • Escalado de datos (En el rango de -1 a 1) 	Excel Excel Excel Matlab/Excel	
<ul style="list-style-type: none"> • Selección preliminar de variables de entrada (Aplicación del coeficiente de correlación de Pearson, tiempo de retardo) 	Excel	
<ul style="list-style-type: none"> • Se establecen los parámetros fijos de la red (Número de capas ocultas, algoritmo de entrenamiento, método de inicialización de pesos, funciones de activación) • Optimización de la red variando el número de variables de entrada y el número de neuronas en la capa oculta. 	Matlab/Excel Matlab/Excel	
<ul style="list-style-type: none"> • Prueba final de validación del modelo óptimo (Predicción de la concentración de SO₂ de salida) 	Matlab/Excel	

Figura 26 Resumen descriptivo de las etapas de diseño del modelo neuronal para la predicción de la concentración de SO₂, y el software empleado en cada etapa. Fuente: Elaboración propia.

En la figura 27 se muestra un histograma de frecuencia con 17 272 datos de la concentración de SO₂ a la entrada, observándose que los datos presentan una distribución bimodal debido a una frecuencia alta en el punto de concentración cero, esto debido a la concentración de SO₂ igual a cero durante los tiempos de parada y recirculación de gases en la planta de ácido. Estos datos en el punto cero son redundantes, no aportarían al modelo e incrementaría el tiempo computacional durante entrenamiento. Por tal motivo, los datos en el punto cero fueron eliminados y se redujeron a una cantidad de 13 409 datos. Luego se elabora un nuevo histograma de frecuencia mostrado en la figura 28.

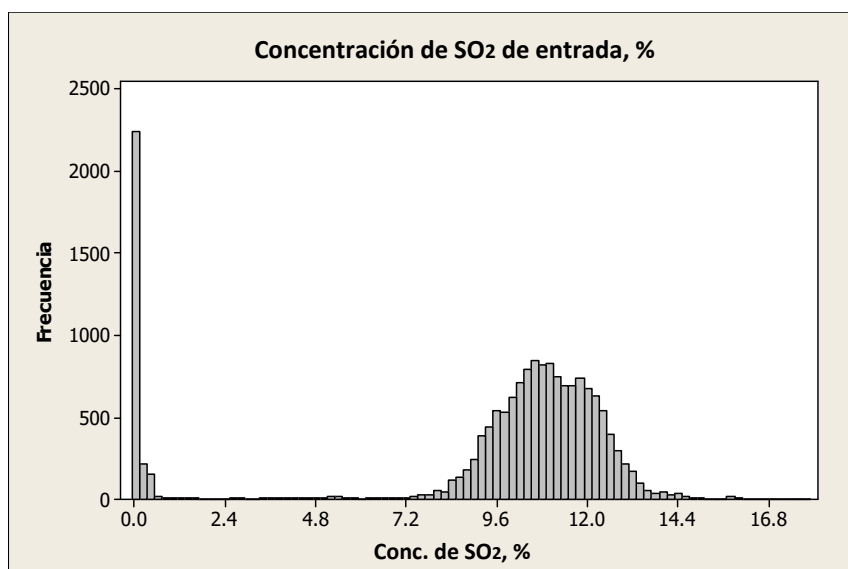


Figura 27 Histograma de frecuencias - % de SO₂ de entrada, total de datos colectados. Fuente: Elaboración propia.

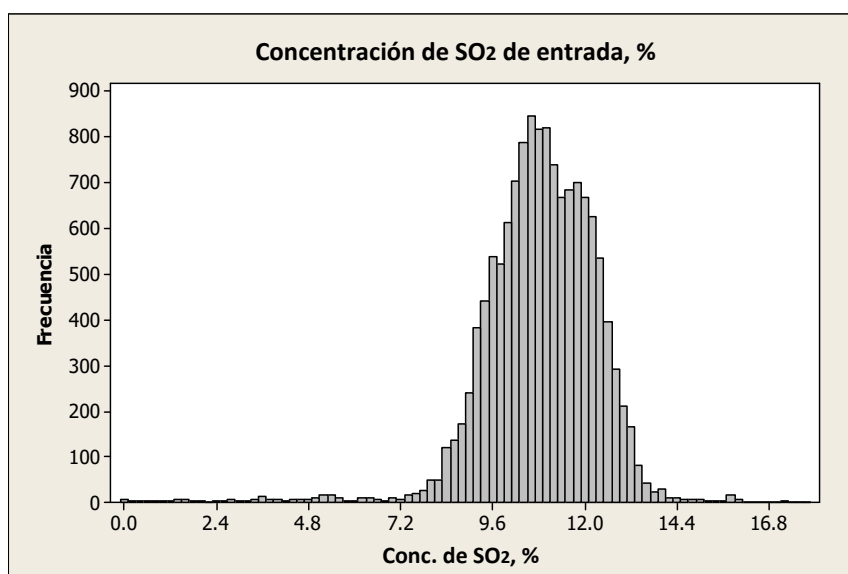


Figura 28 Histograma de frecuencias - % de SO₂ de entrada, filtrando los datos en el punto cero por parada o recirculación de la planta de ácido. Fuente: Elaboración propia.

Los datos graficados en la figura 28 se utilizarán como base para el diseño del sensor virtual, es decir, en adelante se filtrarán los datos en el punto cero retirando los datos durante la parada y la etapa de recirculación en la planta de ácido. Los datos seleccionados proporcionan información de la planta a condiciones de operación normal y en condiciones de operación transitoria.

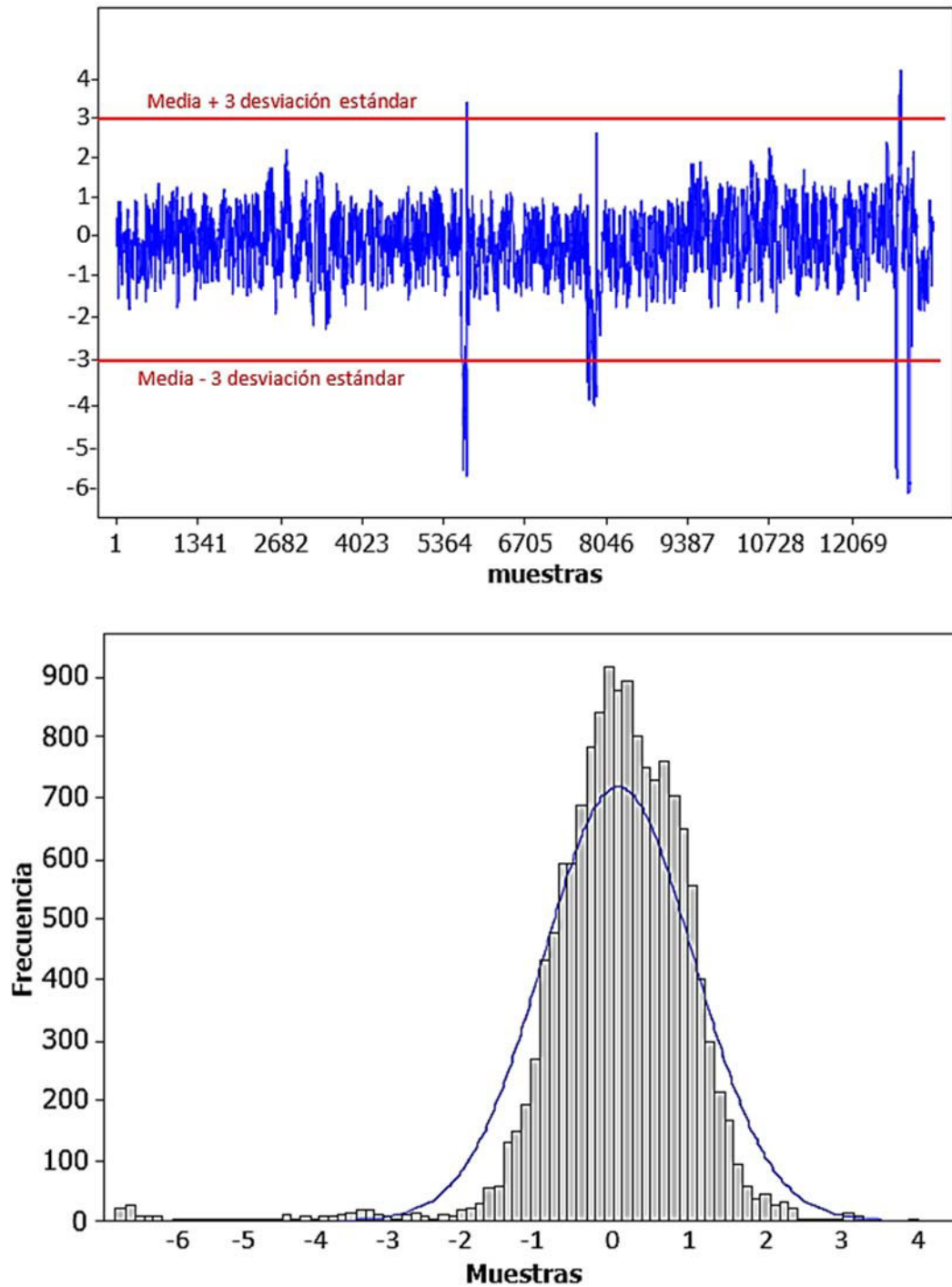


Figura 29 Distribución de los datos de concentración de SO₂ de entrada y análisis de datos atípicos fuera del rango de tres veces la desviación estándar de la muestra. Fuente: Elaboración propia.

La figura 29 muestra la aplicación del método de 3 sigma (tres veces la desviación estándar de la muestra) para detectar valores atípicos de la concentración de SO₂ de entrada. La figura muestra la existencia de datos fuera del rango de 3 sigma, siendo datos atípicos y válidos, ya que corresponden a condiciones de operación inusuales o poco frecuentes de la planta. Estas condiciones transitorias se deben a la mayor variación de la concentración de los gases de entrada propias del proceso metalúrgico o debido a la variación de la concentración durante las condiciones de arranque de planta.

Los valores promedio de operación de las variables de entrada y salida se muestran en la tabla 5.

Tabla 5. Valores promedio de las variables de entrada y salida, obtenidos a partir de 13 409 datos de cada variable.

Variable	Promedio
Temperatura de salida del primer lecho (°K)	897,6
Temperatura de salida del segundo lecho (°K)	802,7
Temperatura de salida del tercer lecho (°K)	706,2
Temperatura de salida del cuarto lecho (°K)	713,1
Concentración de SO ₂ de entrada (%)	10,8
Concentración de O ₂ de entrada (%)	15,5
Flujo de gas de entrada (m/s)	84,3
Concentración de SO₂ de salida (ppm)	316,0

Fuente: Elaboración propia.

En la figura 30 se muestra el histograma de frecuencias de la concentración de oxígeno en el gas de entrada, observándose una distribución bimodal.

En la figura 31 se muestra el histograma de frecuencias de flujo volumétrico de gas en la entrada, observándose una distribución con sesgo hacia la izquierda.

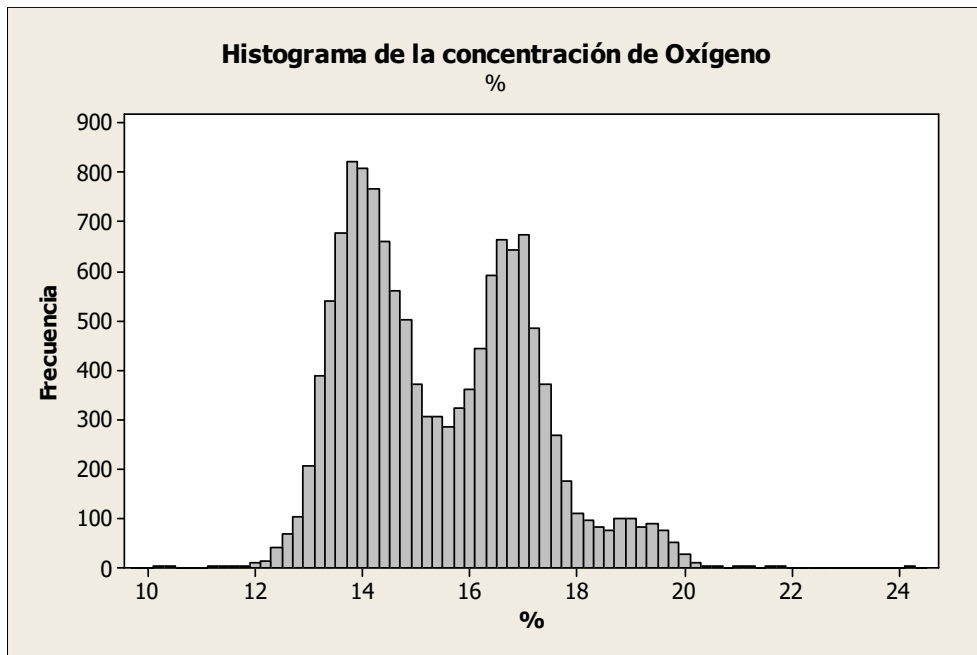


Figura 30 Histograma de frecuencias – Concentración de Oxígeno del gas de entrada, %. Fuente: Elaboración propia.

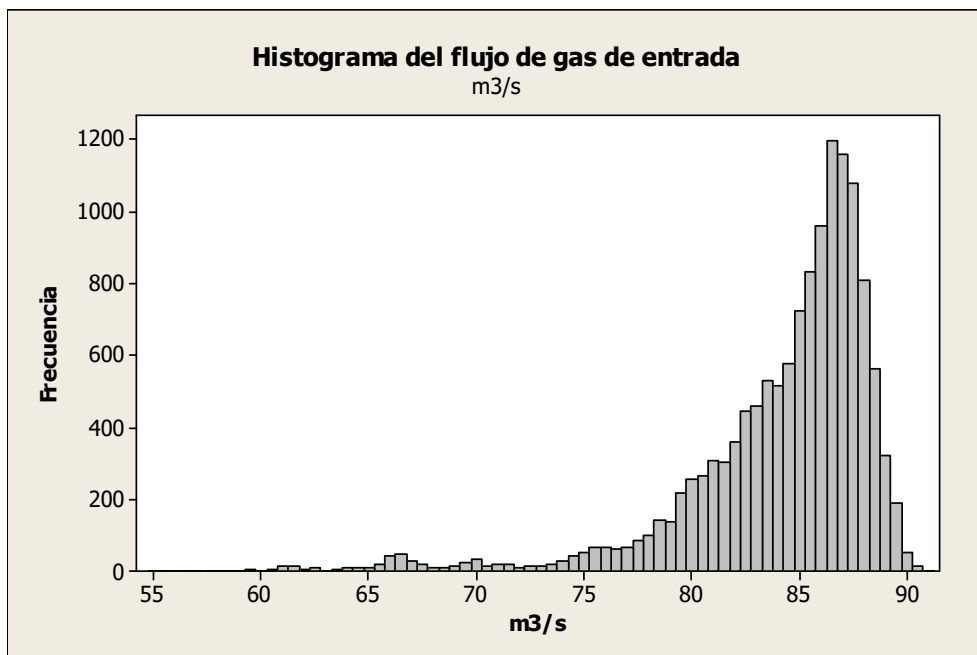


Figura 31 Histograma de frecuencias – Flujo volumétrico del gas de entrada, m³/s. Fuente: Elaboración propia

Asimismo, se verifica que la razón señal/ruido de los datos medidos sean mayor de 5. Los resultados se muestran en la tabla 6.

Tabla 6. Determinación de la razón señal/ruido en las variables.

Variable	Promedio ó señal	Desv. Estándar ó ruido	Razón Señal/Ruido
Temperatura de salida del primer lecho (°K)	897.6	17.7	51
Temperatura de salida del segundo lecho (°K)	802.7	16.8	48
Temperatura de salida del tercer lecho (°K)	706.2	14.0	50
Temperatura de salida del cuarto lecho (°K)	713.1	13.7	52
Concentración de SO ₂ de entrada (%)	10.8	1.6	7
Concentración de O ₂ de entrada (%)	15.5	2.0	8
Flujo de gas de entrada (m/s)	84.3	4.5	19
Concentración de SO₂ de salida (ppm)	316.0	207.6	2

Fuente: Elaboración propia.

La tabla 6 muestra que la señal o variable de salida “Concentración de SO₂ de salida” tiene una razón señal/ruido igual a 2 (menor de 5). Lo cual indica que el ruido es significativamente alto con respecto a la magnitud medida.

4.1.2 Selección de las variables de entrada

Se realiza una selección preliminar del número de variables de entrada utilizando el coeficiente de correlación de Pearson. Se utilizan pares de datos, en este caso entre las variables de entrada: Temperaturas, concentraciones de entrada de SO₂ y O₂ y el flujo; y en la variable de salida: La concentración de salida de SO₂.

El análisis de correlación de Pearson se determina entre las dos variables a diferentes tiempos de retardo con el fin de encontrar el lapso de tiempo para la sincronización de las variables, donde el cambio de la variable de entrada alcanza su mayor impacto en la salida.

Un esquema ilustrativo del cálculo se muestra en la figura 32. Los cálculos se realizan en una hoja Excel. Los datos son ordenados y se determinan los coeficientes de correlación para cada retardo de la concentración de salida. Un ejemplo de cálculo del tiempo de retardo se muestra en el anexo 4.

A continuación, se calculan y muestran gráficamente los resultados de los análisis de correlación entre las variables y con diferentes tiempos de retardo, mediante la aplicación del coeficiente de correlación de Pearson.

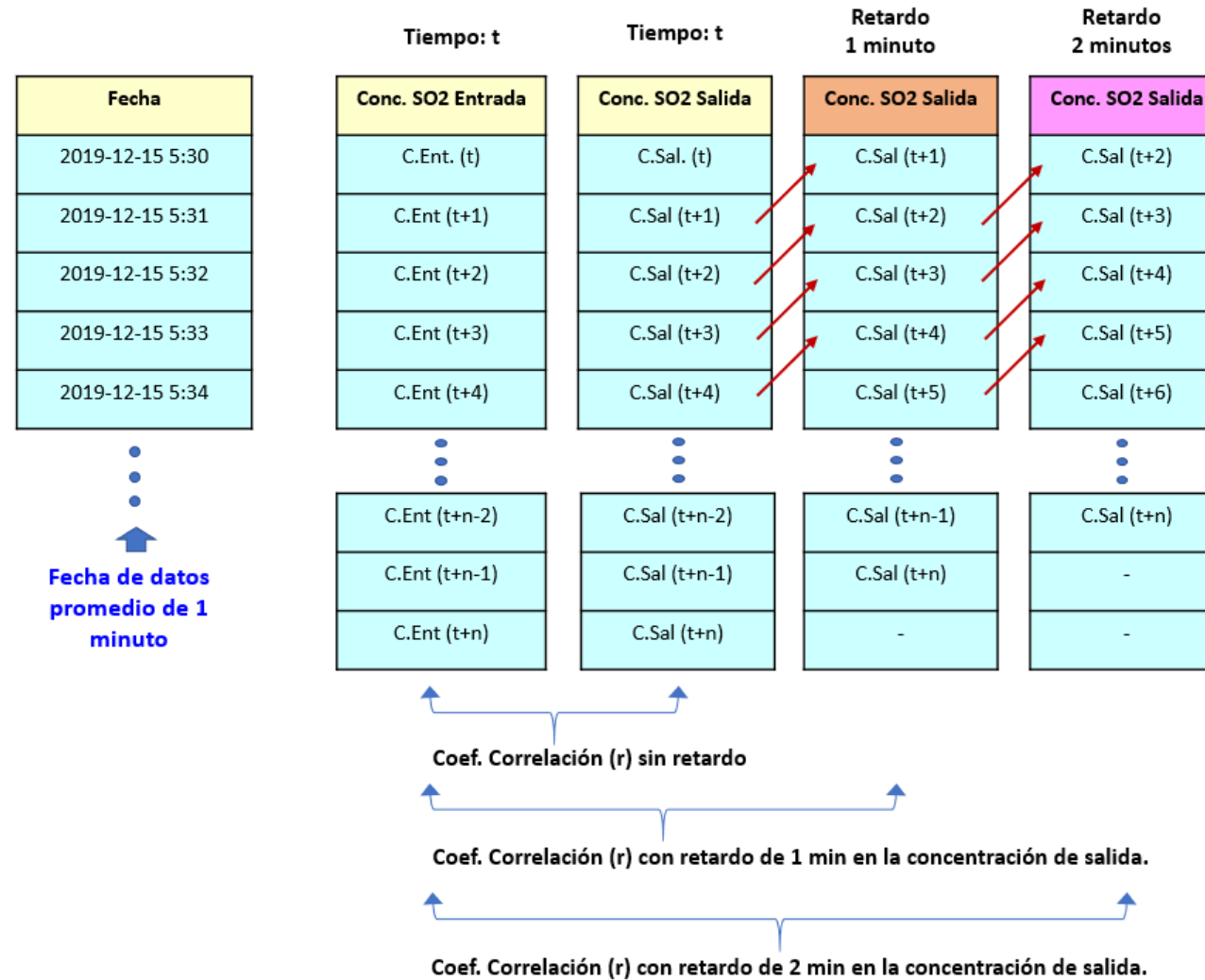


Figura 32 Ilustración de cálculo en Excel del coeficiente de correlación entre la concentración de entrada y la concentración de salida a diferentes tiempos de retardo de la concentración de salida. Para este esquema solo se muestra el retardo de hasta 2 minutos. Para más tiempo de retardo solo se sigue la misma metodología ilustrado en el esquema. Fuente: Elaboración propia.

Correlación entre la concentración de entrada de SO₂ y la concentración de salida de SO₂:

Se realiza un análisis de correlación de los datos considerando tres escenarios:

- Datos a condiciones de operación normal.
- Datos a condiciones de operación transitorias.
- Datos totales. (ambas condiciones juntas).

Para cada escenario, se determinan los coeficientes de correlación entre la concentración de entrada y la concentración de salida. Para ello se determina la mejor correlación variando el tiempo de retardo de la concentración de salida.

De acuerdo a la figura 33 se muestra una mayor correlación a condiciones de operación normal.

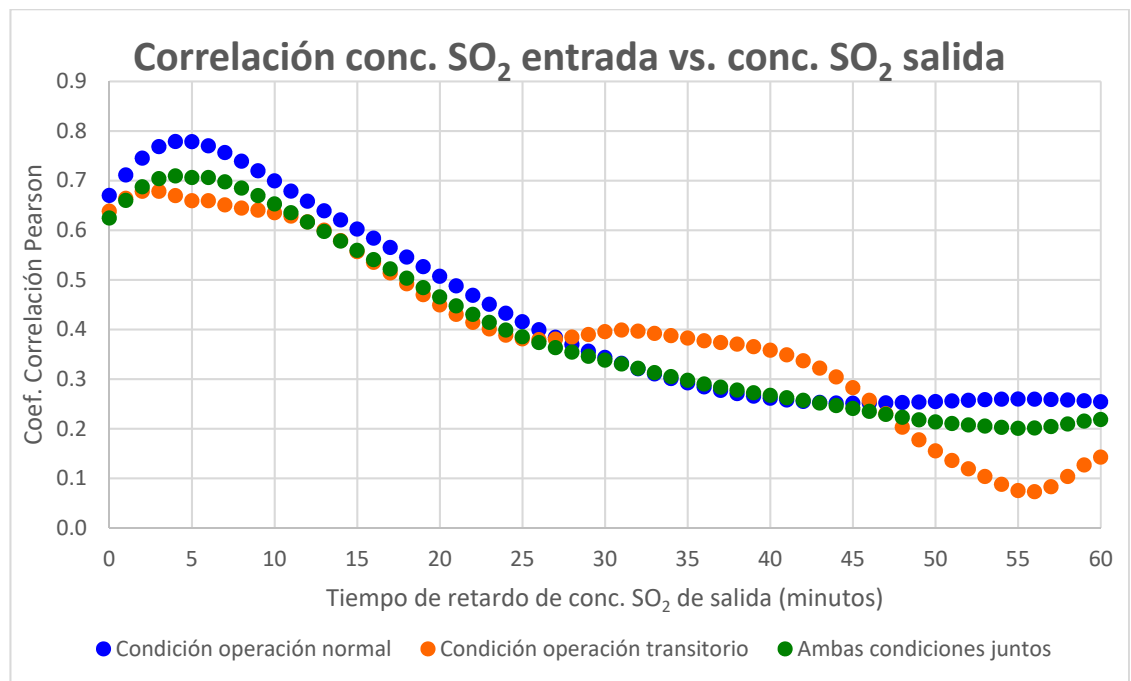


Figura 33 Coeficiente de correlación muestral de Pearson a diferentes tiempos de retardo entre la concentración de SO₂ de entrada y la concentración de SO₂ de salida, a condición de operación normal, condición de operación transitoria y ambas condiciones juntas. Fuente: Elaboración propia.

La figura 33 muestra la variación de las correlaciones a diferentes tiempos de retardo. Los resultados son mostrados en la tabla 7, donde se muestra una correlación positiva alta a condiciones de operación normal.

Tabla 7. *Coefficientes de correlación de la concentración de SO₂ de entrada con la concentración de SO₂ de salida a diferentes tiempos de retardo.*

Tipo de datos	Tiempo de retardo	Coefficiente de correlación
Condición de operación normal	4 minutos	0,779
Condición de operación transitoria	2 minutos	0,678
Datos totales (ambas condiciones)	4 minutos	0,709

Fuente: Elaboración propia.

El análisis de la dispersión de los datos a condición de operación normal y condición de operación transitoria se muestra en la figura 34. Se observa que en condiciones de operación normal la concentración de salida se incrementa notoriamente a medida que el valor de la concentración de entrada aumenta. Por otro lado, en condiciones de operación transitoria hay una mayor dispersión de los datos y el valor de la concentración de salida se incrementa más a medida que la concentración de entrada aumenta.

De acuerdo con la figura 34, se observa que ambas condiciones normal y transitorio siguen la misma tendencia. Por ello, con el fin de utilizar la mayor cantidad de datos y enriquecer más el modelo, se considerará ambas condiciones como un solo conjunto de datos a utilizarse para el desarrollo del sensor virtual.

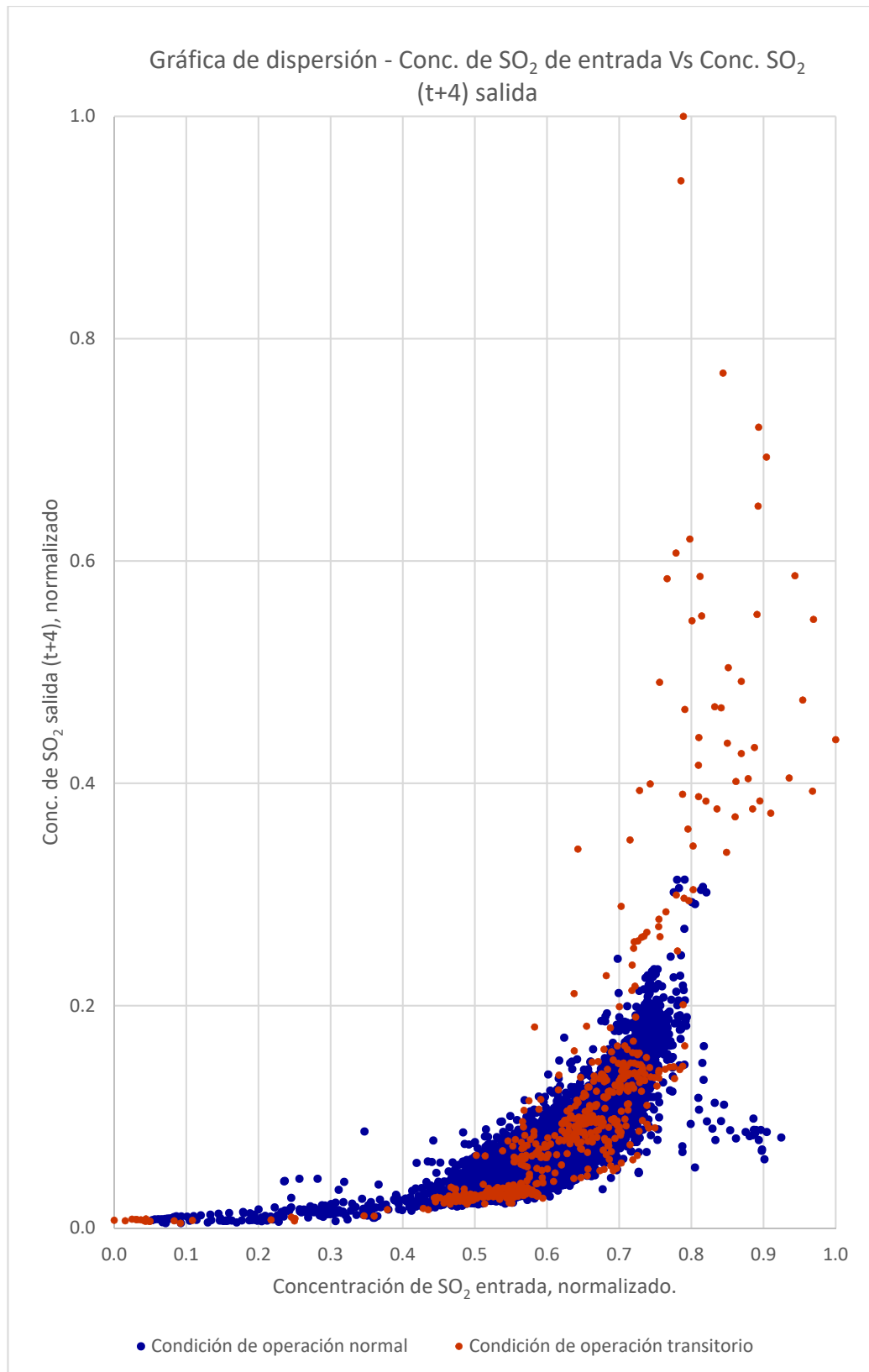


Figura 34 Gráfica de dispersión entre la concentración de SO₂ de entrada y la concentración de SO₂ de salida a condiciones de operación normal y transitorias. Fuente: Elaboración propia.

Correlación entre la concentración de entrada de O_2 y la concentración de salida de SO_2 :

Se realiza un análisis de correlación entre los datos de concentración de O_2 y la concentración de salida de SO_2 . De acuerdo con la figura 35 se observa una mayor correlación a un tiempo de retardo de cinco minutos, obteniendo un coeficiente de correlación de Pearson de -0,572, siendo una correlación negativa moderada.

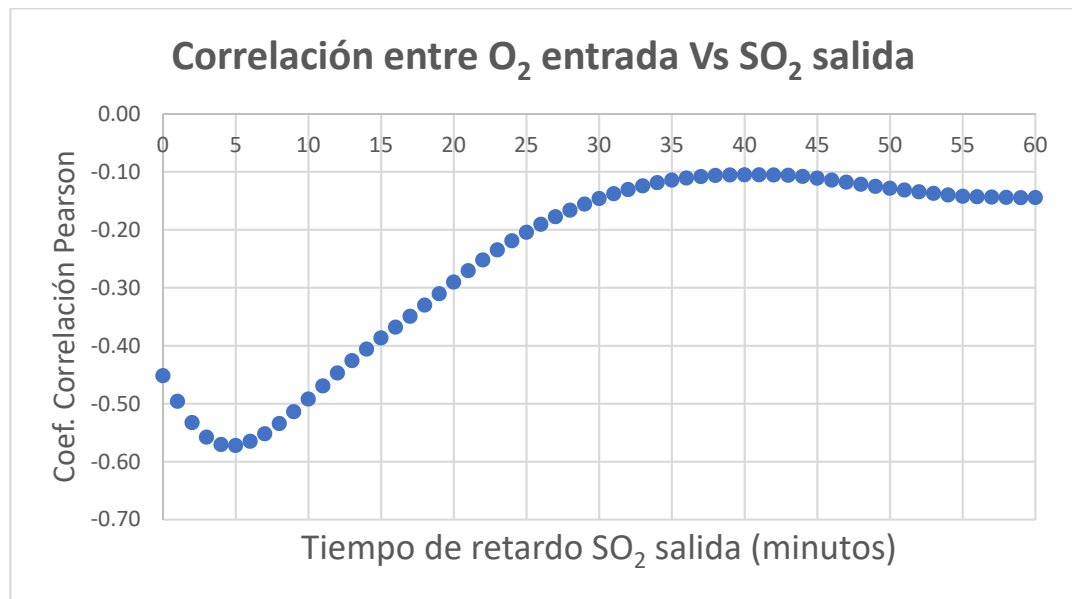


Figura 35 Coeficiente de correlación muestral de Pearson a diferentes tiempos de retardo entre la concentración de O_2 de entrada y la concentración de SO_2 de salida. Fuente: Elaboración propia.

Correlación entre la concentración de entrada de SO_2 y la temperatura de salida del primer lecho:

De acuerdo con la figura 36 se observa una mayor correlación a un tiempo de retardo de catorce minutos, obteniendo un coeficiente de correlación de Pearson de 0,754, siendo una correlación positiva alta.

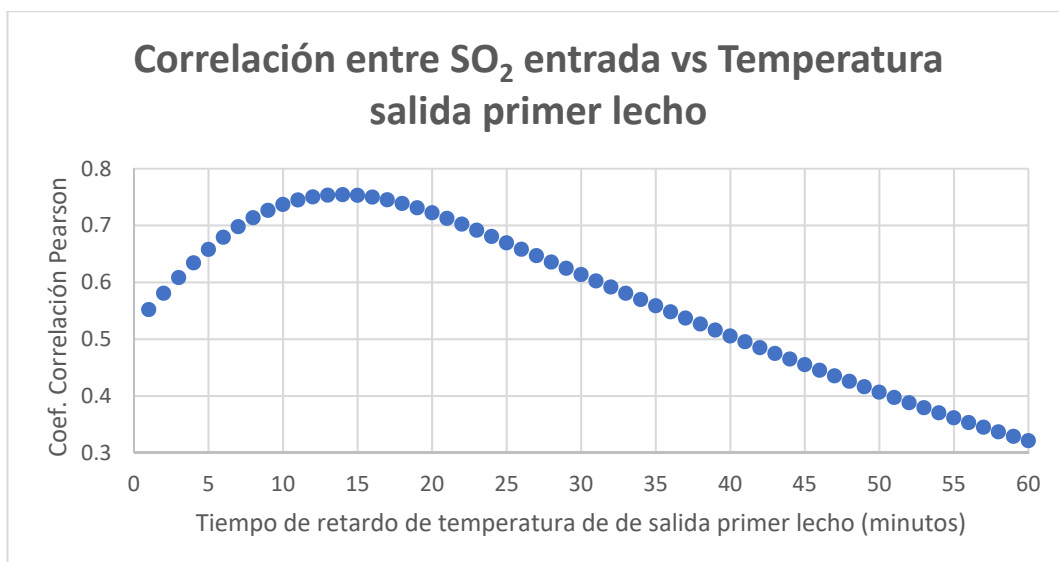


Figura 36 Coeficiente de correlación muestral de Pearson a diferentes tiempos de retardo entre la concentración de SO₂ de entrada y la temperatura de salida del primer lecho. Fuente: Elaboración propia.

Correlación entre la concentración de entrada de SO₂ y la temperatura de salida del segundo lecho:

De acuerdo con la figura 37 se observa una mayor correlación a un tiempo de retardo de veinte minutos, obteniendo un coeficiente de correlación de Pearson de 0,730, siendo una correlación positiva alta.

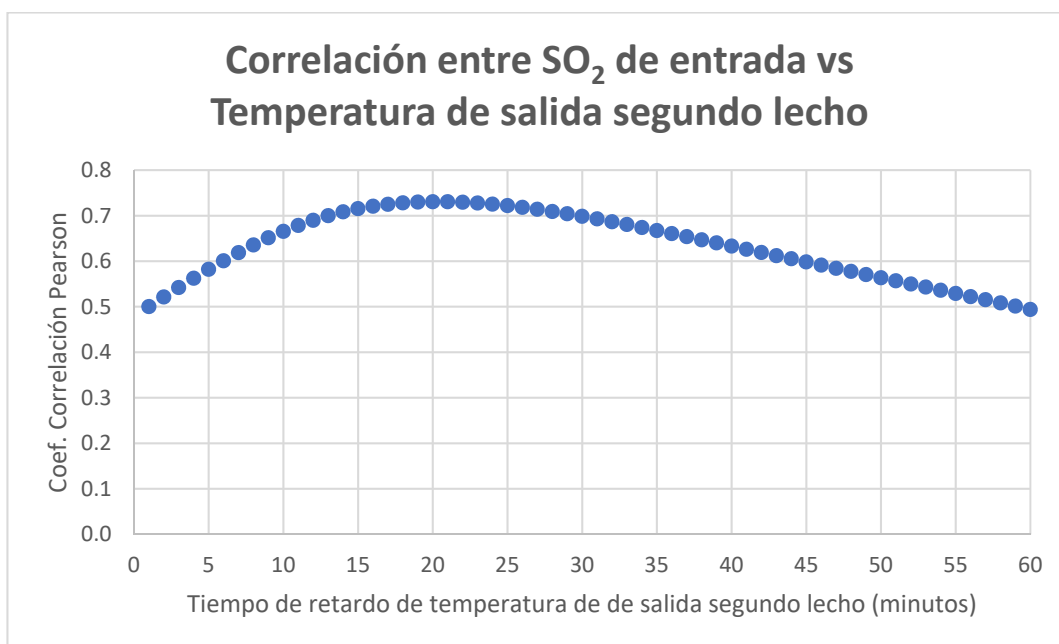


Figura 37 Coeficiente de correlación muestral de Pearson a diferentes tiempos de retardo entre la concentración de SO₂ de entrada y la temperatura de salida del segundo lecho. Fuente: Elaboración propia.

Correlación entre la concentración de entrada de SO₂ y la temperatura de salida del tercer lecho:

De acuerdo con la figura 38 se observa una mayor correlación a un tiempo de retardo de treinta y ocho minutos, obteniendo un coeficiente de correlación de Pearson de 0,611, siendo una correlación positiva moderada.

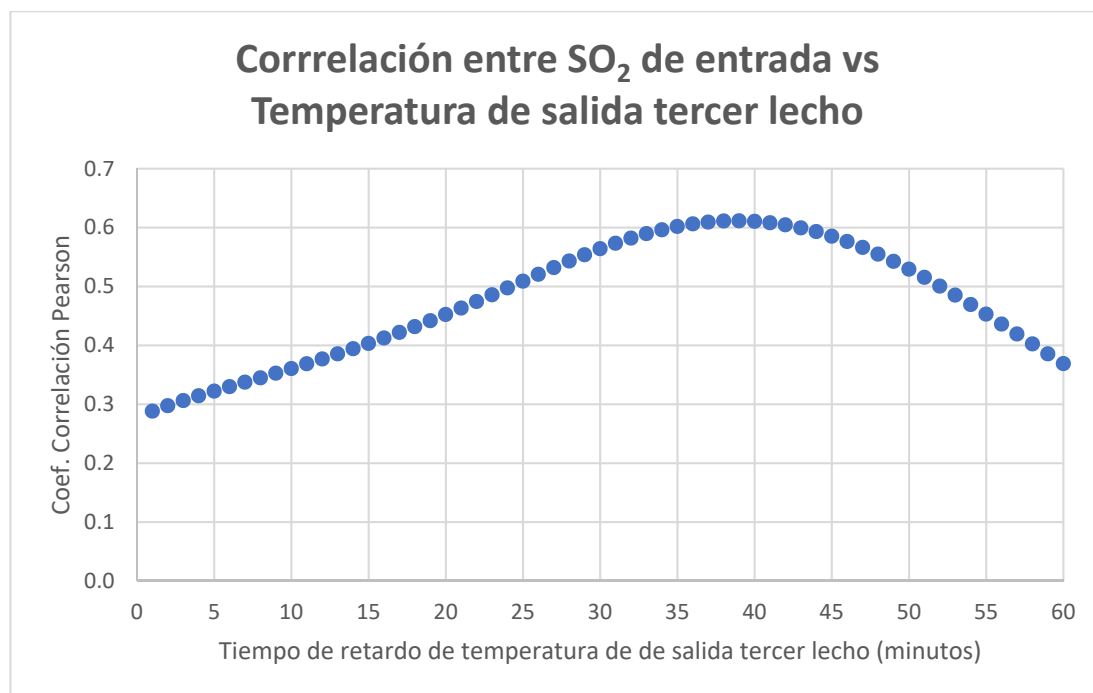


Figura 38 Coeficiente de correlación muestral de Pearson a diferentes tiempos de retardo entre la concentración de SO₂ de entrada y la temperatura de salida del tercer lecho. Fuente: Elaboración propia.

Correlación entre la concentración de SO₂ en la entrada y la temperatura de salida del cuarto lecho:

De acuerdo con la figura 39 se observa una mayor correlación a un tiempo de retardo de treinta y tres minutos, obteniendo un coeficiente de correlación de Pearson de 0,578, siendo una correlación positiva moderada.

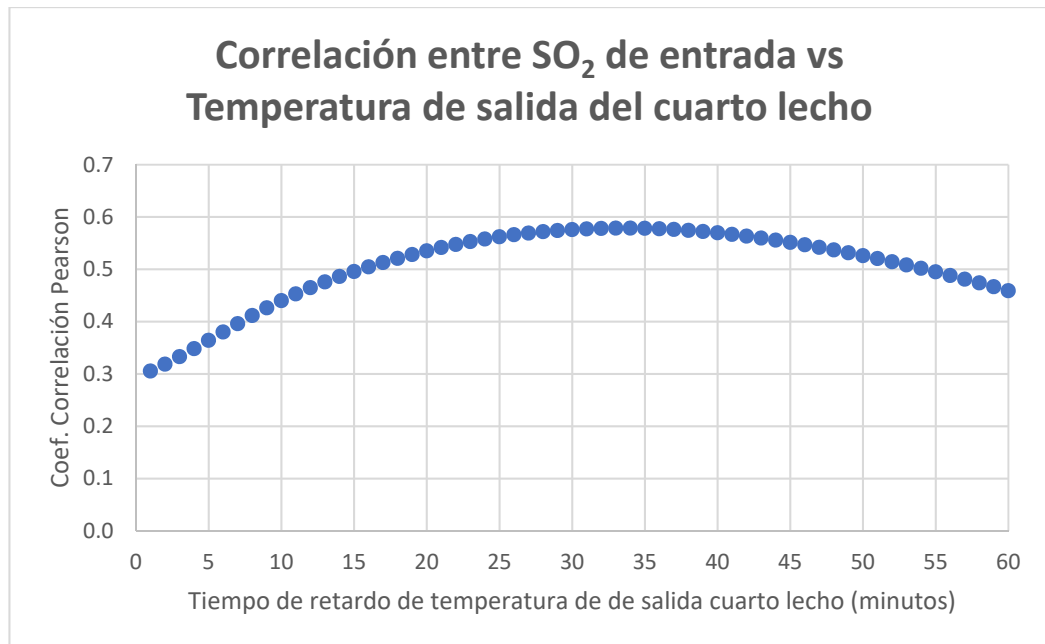


Figura 39 Coeficiente de correlación muestral de Pearson a diferentes tiempos de retardo entre la concentración de SO₂ de entrada y la temperatura de salida del cuarto lecho. Fuente: Elaboración propia.

De acuerdo a la figura 40 el coeficiente de correlación entre la concentración de salida de SO₂ y el flujo de entrada, tiene una correlación positiva baja.

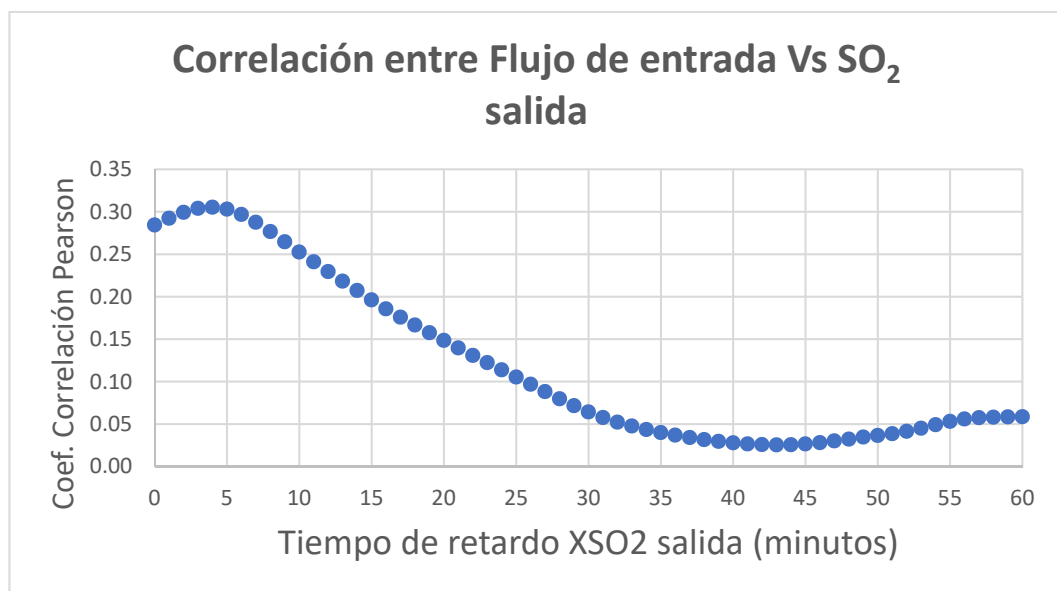


Figura 40 Coeficiente de correlación muestral de Pearson a diferentes tiempos de retardo entre el flujo de gas de entrada y la concentración de SO₂ de salida. Fuente: Elaboración propia.

De acuerdo a la figura 41 el coeficiente de correlación entre la concentración de salida de SO₂ y la presión, tiene poca o ninguna correlación lineal.

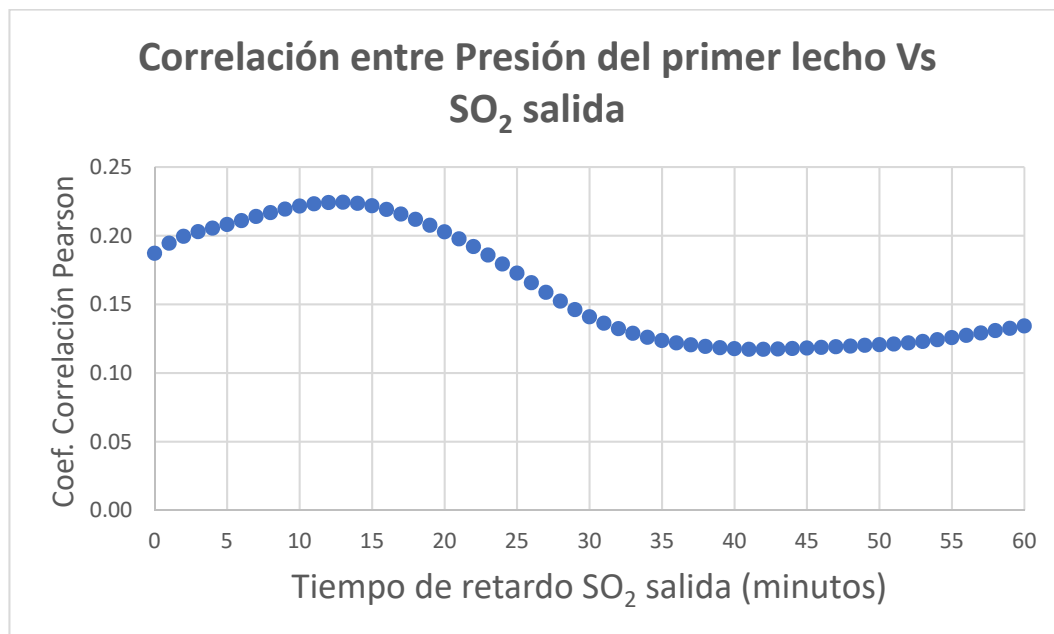


Figura 41 Coeficiente de correlación muestral de Pearson a diferentes tiempos de retardo entre la presión en el primer lecho y la concentración de SO₂ de salida. Fuente: Elaboración propia.

De acuerdo con el análisis de correlación de Pearson se establece preliminarmente las variables relevantes a ser considerados en el diseño de sensor virtual: La concentración de entrada de SO₂, la concentración de entrada de O₂, las temperaturas de salida de los lechos del convertidor, y finalmente también se considerará en el flujo de gas de entrada.

4.1.3 Selección de la estructura neuronal

Mediante pruebas de ensayo y error se selecciona la mejor estructura de la red neuronal, siendo el óptimo aquella estructura que realice la predicción con el menor error. La capacidad de predicción de cada estructura neuronal se determina mediante el cálculo de los indicadores de desempeño. Los diseños de la red neuronal y los entrenamientos de cada estructura se realizaron mediante el software Matlab R2018b (Matlab 9.5). Se establecen los parámetros fijos y variables para el desarrollo de los modelos, entre los parámetros fijos establecidos en el presente estudio son el número de capas ocultas, el algoritmo de entrenamiento y la función de activación. Asimismo, los parámetros variables establecidos en el presente estudio son el número de variables o neuronas en la capa de entrada y el número

de neuronas en la capa oculta. Los indicadores de desempeño empleados en la evaluación de cada estructura neuronal son el error porcentual medio absoluto, el índice refinado de concordancia y el coeficiente de correlación.

Partición de ejemplos de datos

Con el fin de aplicar la técnica de detención temprana (early stopping) se realiza la partición de los ejemplos de datos, siendo un total de 13 409 ejemplos. La proporción en porcentaje y la cantidad de ejemplos particionados para el entrenamiento, validación y prueba final son:

- Entrenamiento (80%): 10 727 ejemplos.
- Validación (10%): 1 341 ejemplos.
- Prueba final (10%): 1 341 ejemplos.

Número de capas ocultas

Existen diversas aplicaciones para la predicción en procesos químicos usando redes de tipo multicapa perceptrón donde se utiliza solamente una capa oculta y obteniéndose buenos resultados de predicción. Para el presente estudio se elige una capa oculta para minimizar el tiempo de procesamiento y tomando como referencia lo recomendado por Mhaskar H., Poggio T., (2016) y la aplicación de una capa oculta por realizado por Kim D., (2017).

Algoritmo de entrenamiento

Para el entrenamiento del modelo se elige el algoritmo de entrenamiento de Levenberg – Marquardt; ya que es considerado como uno de los algoritmos de entrenamiento más eficientes, probado en varios problemas de aproximación de funciones, ha sido comparado con otros algoritmos como el gradiente conjugado y un algoritmo de tasa de aprendizaje variable. (Hagan M. T., Menhaj M., 1994).

Funciones de activación

En la capa oculta se elige la función de activación del tipo tangente sigmoideal hiperbólica, siendo una función continua no lineal y con un rango comprendido entre -1 y 1. En la capa de salida se elige la función de activación del tipo lineal, con una función de una línea recta de pendiente unitaria. Para la elección de las funciones de activación se utiliza como referencia otros modelos con una capa

oculta y con resultados satisfactorios, como el modelo de Barrera R., Mora J. y Perez S. (2009) para la estimación del ángulo del rotor de un generador sincrónico, o el trabajo de Tarafdar A., Chandra N., Singh A. y Sirohi R. (2017), para el modelamiento de la actividad del agua.

Método de inicialización de los pesos

Se elige el uso del método de Nguyen-Widrow tomando como referencia el trabajo realizado por Pavelka A. y Procházka A. (2004), que compararon diez métodos de inicialización de pesos en modelos de redes neuronales multicapa, obteniéndose el mejor resultado con el método de Nguyen-Widrow.

Número preliminar de variables en la capa de entrada:

Como punto de partida se establece la estructura neuronal con doce variables o neuronas en la capa de entrada:

1. Temperatura de salida del primer lecho (t)
2. Temperatura de salida del primer lecho (t-1)
3. Temperatura de salida del segundo lecho (t)
4. Temperatura de salida del segundo lecho (t-1)
5. Temperatura de salida del tercer lecho (t)
6. Temperatura de salida del cuarto lecho (t)
7. Concentración de SO₂ de entrada (t-2)
8. Concentración de SO₂ de entrada (t-3)
9. Concentración de SO₂ de entrada (t-4)
10. Concentración de SO₂ de entrada (t-5)
11. Concentración de O₂ de entrada (t-5)
12. Flujo de gas de entrada (t-4)

No se considera al catalizador como una variable de entrada debido a que el reactor opera con un solo tipo de catalizador durante un periodo largo de tiempo (2 años aprox.). En ese sentido, si consideramos el tiempo de envejecimiento del catalizador podría durar meses, lo cual indica que el catalizador no es una variable de corto plazo, considerando que los datos utilizados en el modelo corresponden a

un periodo de tiempo de 12 días. Es por ello que se recomienda el mantenimiento periódico del sensor virtual, es decir, se debe reentrenar el modelo, ya sea con datos nuevos o incrementando más datos al modelo existente.

Variación del número de variables en la capa de entrada y el número de neuronas en la capa oculta:

Mediante la aplicación de ensayos de prueba y error, se realizan corridas de entrenamiento y validación de diversas estructuras neuronales, variando el número de variables en la capa de entrada y el número de neuronas en la capa oculta.

El objetivo es encontrar el número óptimo de variables de entrada y el número de unidades en la capa oculta mediante pruebas de ensayo y error, esto con el fin de optimizar o encontrar la mejor estructura del modelo, que nos permita predecir con el menor error. Los detalles se muestran en el anexo 6. En cada corrida de los modelos se aplicó el método de detención temprana. A partir de los resultados obtenidos en cada corrida se determinan los indicadores de desempeño de los modelos y se elige el mejor modelo neuronal, siendo el que tenga un mejor indicador de desempeño. Un ejemplo del cálculo en Excel de los indicadores de desempeño se muestra en el anexo 5.

Los indicadores de desempeño obtenidos son mostrados en la tabla 8, entre ellos se aprecia dos escenarios con un mejor indicador de desempeño: El modelo con estructura neuronal 9-9-1 y el otro con la estructura neuronal 12-16-1. Finalmente, entre ambos modelos se elige el modelo con estructura 9-9-1 que tiene un menor error porcentual medio absoluto de 12,9 % y tiene una estructura simple con una capa oculta, lo cual permite al modelo un tiempo de procesamiento más rápido.

Tabla 8. Indicadores de desempeño de los modelos de predicción de la concentración de SO₂ de salida, considerando la combinación del número de variables en la capa de entrada y el número óptimo de neuronas en la capa oculta.

Nº de variables de entrada	Combinación de variables de entrada	Número óptimo de unidades ocultas	Error porcentual medio absoluto (MAPE)%	Índice refinado de concordancia (IOA)	Coef. de correlación "r"
1	Temp_sal_Lecho1(t)	1	26,7%	0,620	0,685
1	Conc_SO2entrada(t-4)	3	31,7%	0,708	0,870
2	Conc_SO2entrada(t-4), Temp_sal_Lecho1(t)	17	22,5%	0,809	0,876
3	Conc_SO2entrada(t-4), Temp_sal_Lecho1(t), Temp_sal_Lecho2(t)	15	25,4%	0,801	0,871
4	Conc_SO2entrada(t-4), Temp_sal_Lecho1(t), Temp_sal_Lecho2(t), Temp_sal_Lecho3(t)	1	12,6%	0,856	0,912
5	Conc_SO2entrada(t-4), Temp_sal_Lecho1(t), Temp_sal_Lecho2(t), Temp_sal_Lecho3(t), Temp_sal_Lecho4(t)	3	12,6%	0,856	0,923
6	Conc_SO2entrada(t-4), Temp_sal_Lecho1(t), Temp_sal_Lecho2(t), Temp_sal_Lecho3(t), Temp_sal_Lecho4(t), Conc_O2entrada(t-5)	5	22,3%	0,856	0,915
9	Conc_SO2entrada(t-2), Conc_SO2entrada(t-3), Conc_SO2entrada(t-4), Conc_SO2entrada(t-5), Temp_sal_Lecho1(t), Temp_sal_Lecho2(t), Temp_sal_Lecho3(t), Temp_sal_Lecho4(t), Conc_O2entrada(t-5).	9	12,9%	0,877	0,950
11	Conc_SO2entrada(t-2), Conc_SO2entrada(t-3), Conc_SO2entrada(t-4), Conc_SO2entrada(t-5), Temp_sal_Lecho1(t), Temp_sal_Lecho1(t-1), Temp_sal_Lecho2(t), Temp_sal_Lecho2(t-1), Temp_sal_Lecho3(t), Temp_sal_Lecho4(t), Conc_O2entrada(t-5).	9	20,4%	0,873	0,941
12	Conc_SO2entrada(t-2), Conc_SO2entrada(t-3), Conc_SO2entrada(t-4), Conc_SO2entrada(t-5), Temp_sal_Lecho1(t), Temp_sal_Lecho1(t-1), Temp_sal_Lecho2(t), Temp_sal_Lecho2(t-1), Temp_sal_Lecho3(t), Temp_sal_Lecho4(t), Conc_O2entrada(t-5), Flujo_entrada(t-4)	16	15,6%	0,880	0,941

Fuente: Elaboración propia

Número óptimo de neuronas en la capa de entrada y en la capa oculta:

Luego de realizar entrenamientos con diferentes estructuras y variando el número de neuronas ocultas, se determina la mejor estructura neuronal con mejor indicador de desempeño, siendo del tipo 9-9-1 mostrado en la figura 42. La estructura requiere nueve variables en la capa de entrada y nueve neuronas en la capa oculta. Las variables de la capa de entrada son:

1. Temperatura de salida lecho 1 (t)
2. Temperatura de salida lecho 2 (t)
3. Temperatura de salida lecho 3 (t)
4. Temperatura de salida lecho 4 (t)
5. Concentración de SO₂ de entrada (t-2)
6. Concentración de SO₂ de entrada (t-3)
7. Concentración de SO₂ de entrada (t-4)
8. Concentración de SO₂ de entrada (t-5)
9. Concentración de O₂ de entrada (t-5)

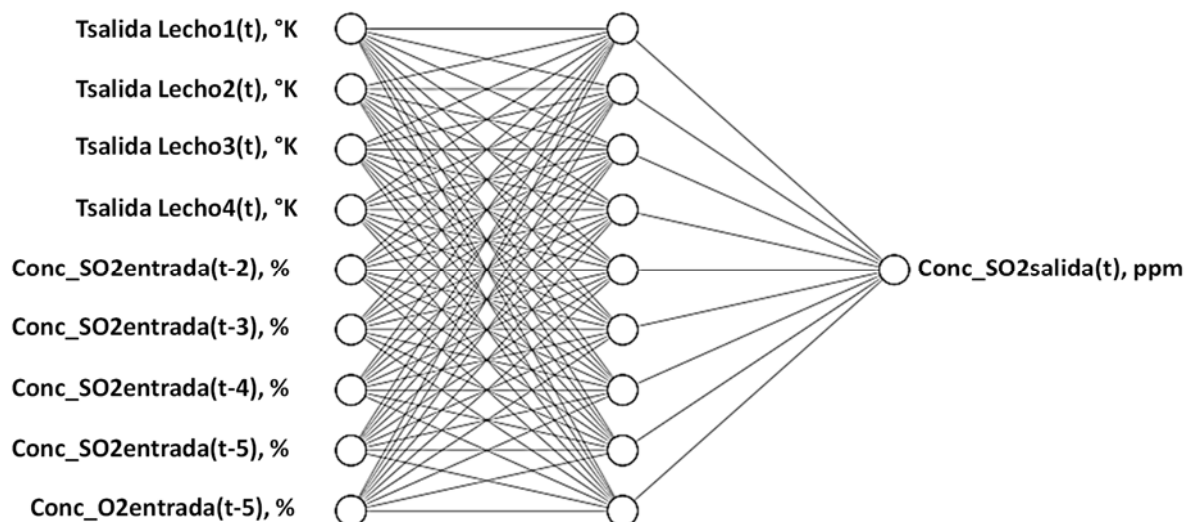


Figura 42 Estructura óptima del modelo neuronal para la predicción de la concentración de SO₂ de salida.

Fuente: Elaboración propia.

Las figuras 43 y 44 muestran las pantallas de visualización de Matlab durante la configuración de la red neuronal 9-9-1 y el performance de la red durante el entrenamiento mediante el método de detención temprana. De acuerdo a la figura 38 el entrenamiento se detuvo en la iteración (o época) 10, retornando los pesos obtenidos en la iteración 4 donde se obtuvo un mínimo error de validación, ya que luego se inicia el incremento del error de validación.

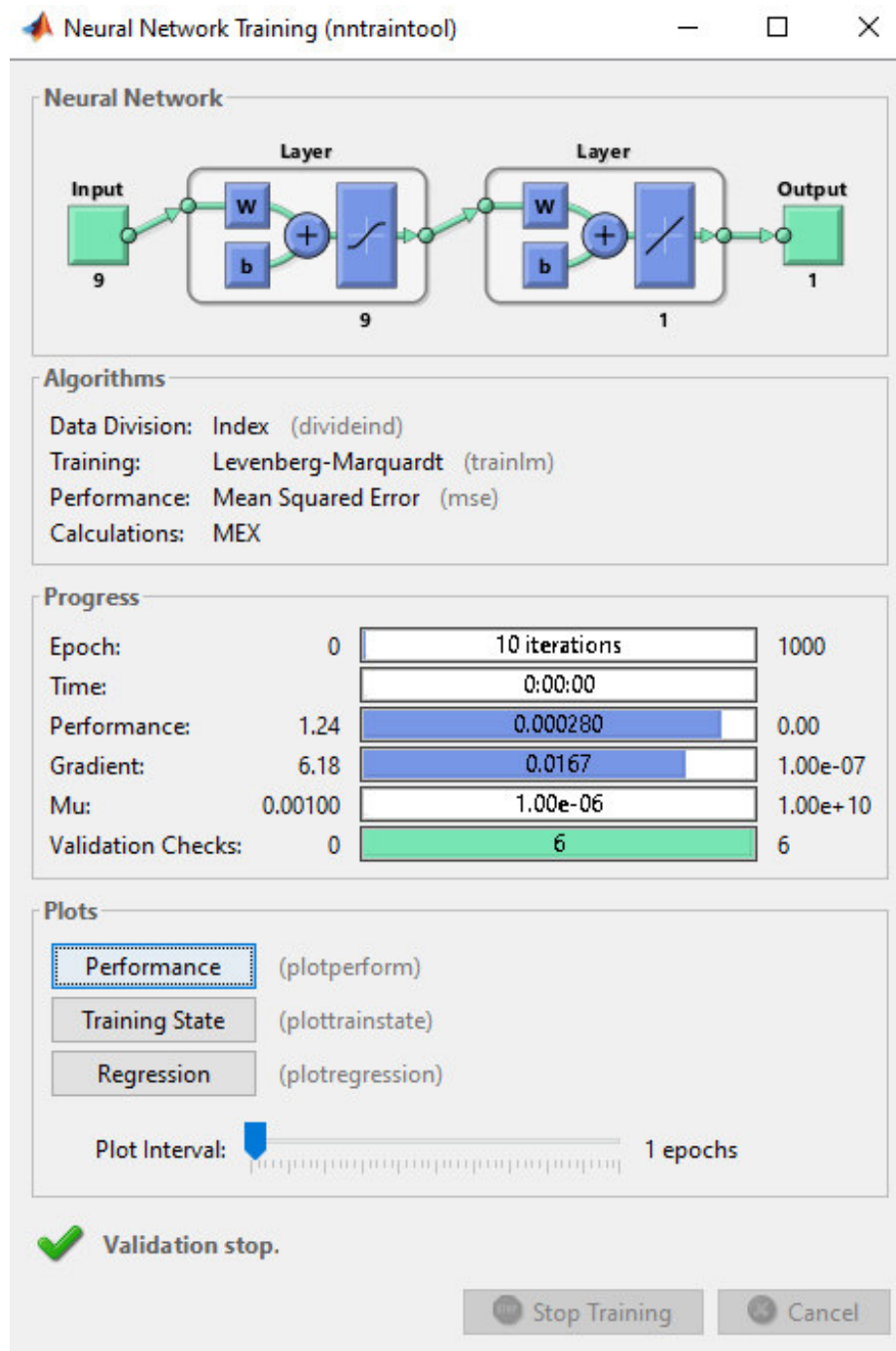


Figura 43 Estructura y parámetros de entrenamiento de la red neuronal en Matlab. Fuente: Elaboración propia.

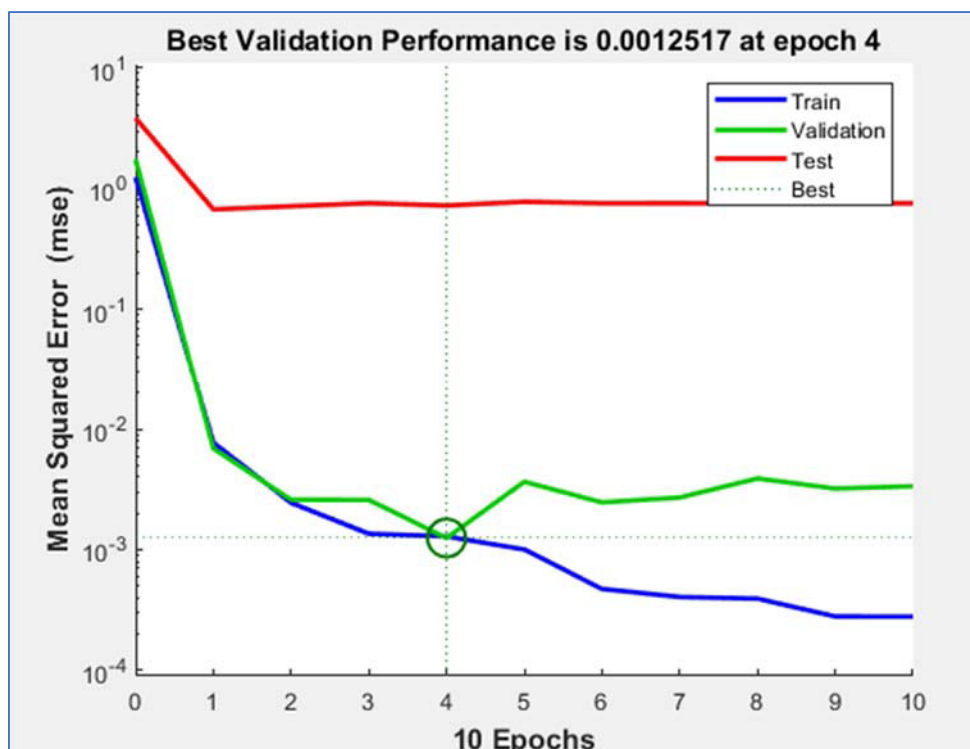


Figura 44 Entrenamiento de la estructura 9-9-1 con el método de detención temprana (early stopping) en Matlab. Fuente: Elaboración propia.

4.1.4 Prueba final de la predicción de concentración de salida de SO_2

Luego de determinar la estructura de la red óptima, se realiza una prueba final con nuevos datos independientes y que no fueron utilizados durante los procesos de entrenamiento y validación. Para la prueba final de la predicción de concentración de salida de SO_2 se utilizaron 1 341 ejemplos, obteniéndose los indicadores de desempeño mostrados en la tabla 9 y los resultados de predicción con los valores reales se muestran en la figura 45, donde se muestra que el modelo tiene mayor error de predicción a concentraciones altas.

Tabla 9. Resultados de desempeño de predicción de la concentración de SO_2 de salida, mediante la red neuronal

Estructura neuronal	Error porcentual medio absoluto (MAPE)%	Índice refinado de Concordancia (IOA)	Coefficiente de correlación (r)
9 – 9 – 1	12%	0,832	0,945

Fuente: Elaboración propia

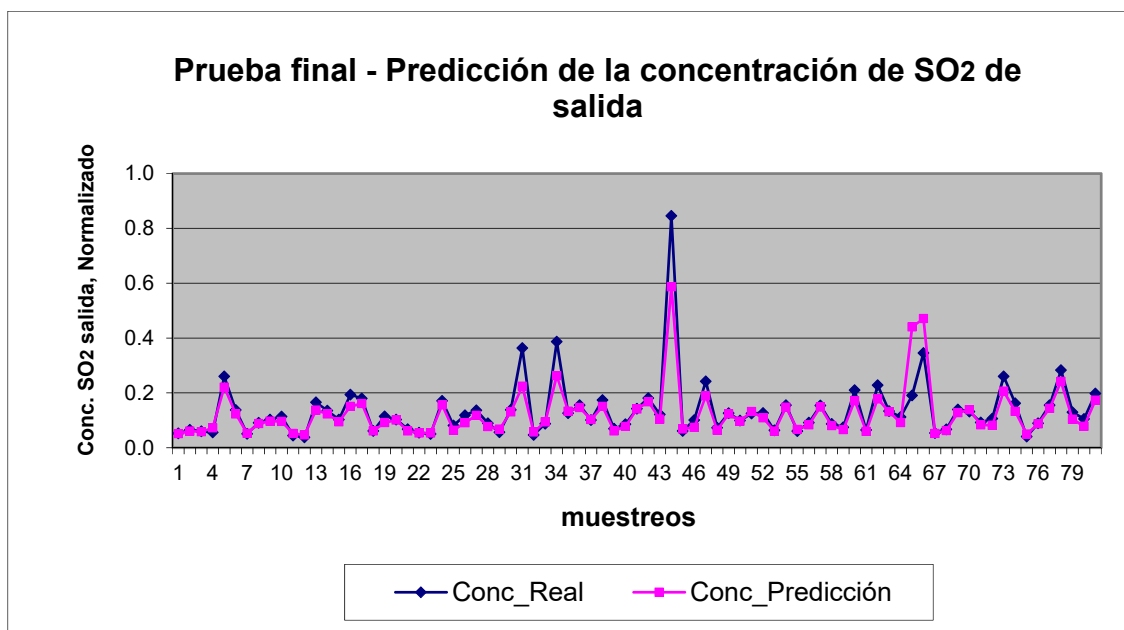


Figura 45 Prueba final del modelo de predicción de la concentración de SO₂ de salida del reactor catalítico multi-lechos. Fuente: Elaboración propia.

4.1.5 Cálculo de la eficiencia de conversión

El diagrama mostrado en la figura 46 muestra las etapas utilizadas para el cálculo de la eficiencia de conversión. En la primera etapa, aplicando RNA se realiza la predicción de la concentración de salida de SO₂. En la segunda etapa, ya con el resultado de la concentración de salida de SO₂ se calcula la eficiencia de conversión.



Figura 46 Sensor virtual obtenido para la predicción de la eficiencia de conversión. Fuente: Elaboración propia.

Ejemplo:

El gas ingresa al reactor catalítico con un porcentaje molar promedio de 10,8 % de SO₂. Luego de pasar por los cuatro lechos del reactor catalítico, el gas sale a una concentración promedio de 316 ppm de SO₂ equivalente a un porcentaje molar de 0,0316 % de SO₂. (Equivalencia: 1% es igual a 10 000 ppm).

Aplicando la fórmula de eficiencia de conversión:

$$\Phi = \frac{X_{SO_2}^{in} - X_{SO_2}}{X_{SO_2}^{in} - \frac{1}{2} X_{SO_2} X_{SO_2}^{in}} = \frac{0,108 - 0,000316}{0,108 - \frac{1}{2} 0,000316 * 0,108}$$

$$\Phi = \frac{0,107684}{0,10798294} = 0,99723$$

La eficiencia de conversión de SO₂ calculado es 99,723 %.

Los cálculos de la eficiencia de conversión son realizados para los 1 341 ejemplos o patrones de datos obtenidos de la muestra de prueba final (10 % de los datos). Finalmente, los resultados de eficiencia obtenidas del modelo son comparados contra los resultados de eficiencia real del reactor catalítico, y se determinan los indicadores de desempeño de la predicción de la eficiencia de conversión, los indicadores de desempeño del sensor virtual basados en RNA para la predicción de la eficiencia de conversión se muestran en la tabla 10 y se aprecia un coeficiente de correlación redondeado a dos decimales igual a 1,00. Los resultados de predicción de la eficiencia de conversión se muestran en la figura 47, donde se muestra un mayor error a menores eficiencias de conversión, esto debido a que el error de predicción de la concentración de salida (mostrado en la figura 45) ha sido trasladado en la predicción de la eficiencia de conversión.

Tabla 10. Resultados del desempeño de predicción de la eficiencia de conversión mediante el sensor virtual basado en RNA.

	Error porcentual medio absoluto (MAPE)%	Índice refinado de Concordancia (IOA)	Coefficiente de correlación (r)
Eficiencia de conversión con el sensor virtual basado en RNA.	0,20%	0,902	1,00

Fuente: Elaboración propia.

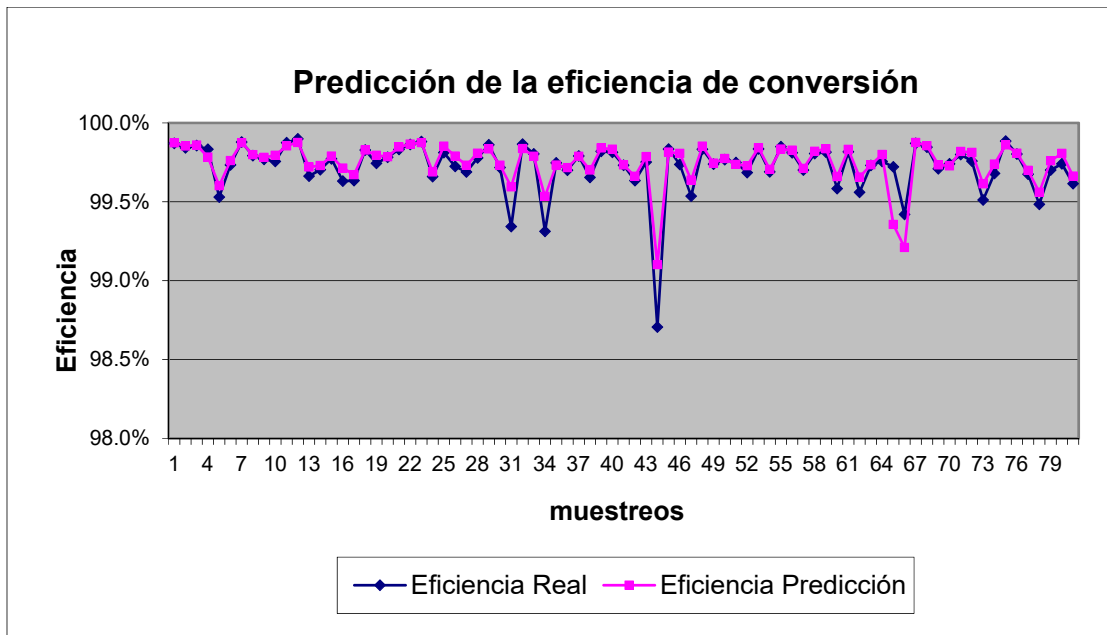


Figura 47 Resultados de predicción de la eficiencia de conversión del reactor catalítico multi-lechos. Fuente: Elaboración propia.

4.2 Resultados del sensor virtual derivado de leyes

- Sincronización de las variables aplicando suavización exponencial:

Las figuras 48, 49, 50 y 51 muestran la sincronización de los datos de concentración de SO₂ de entrada y O₂ de entrada con la temperatura de salida del primer lecho aplicando la técnica de suavización exponencial con un factor $\alpha = 0,97$.

$$X_{SO_2 \text{ entr. suavizado}}(t) = (1 - \alpha) X_{SO_2 \text{ entr.}}(t) + \alpha X_{SO_2 \text{ entr. suavizado}}(t-1)$$

$$X_{O_2 \text{ entr. suavizado}}(t) = (1 - \alpha) X_{O_2 \text{ entr.}}(t) + \alpha X_{O_2 \text{ entr. suavizado}}(t-1)$$

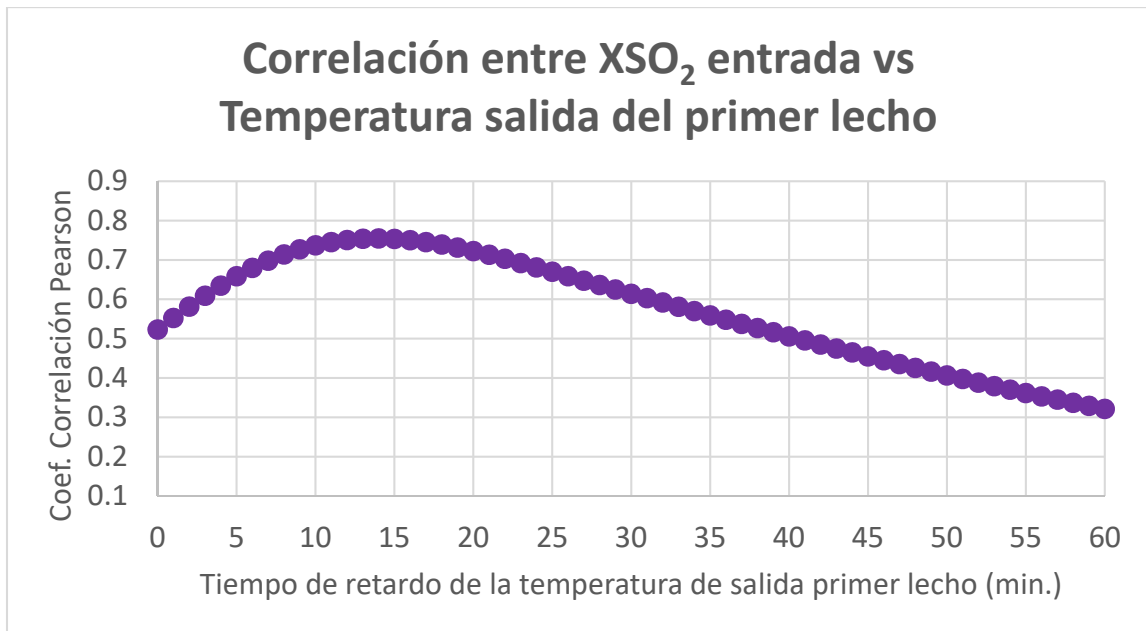


Figura 48 Coeficiente de correlación de Pearson entre la concentración de SO_2 de entrada y la temperatura de salida del primer lecho a diferentes tiempos de retardo. Fuente: Elaboración propia.

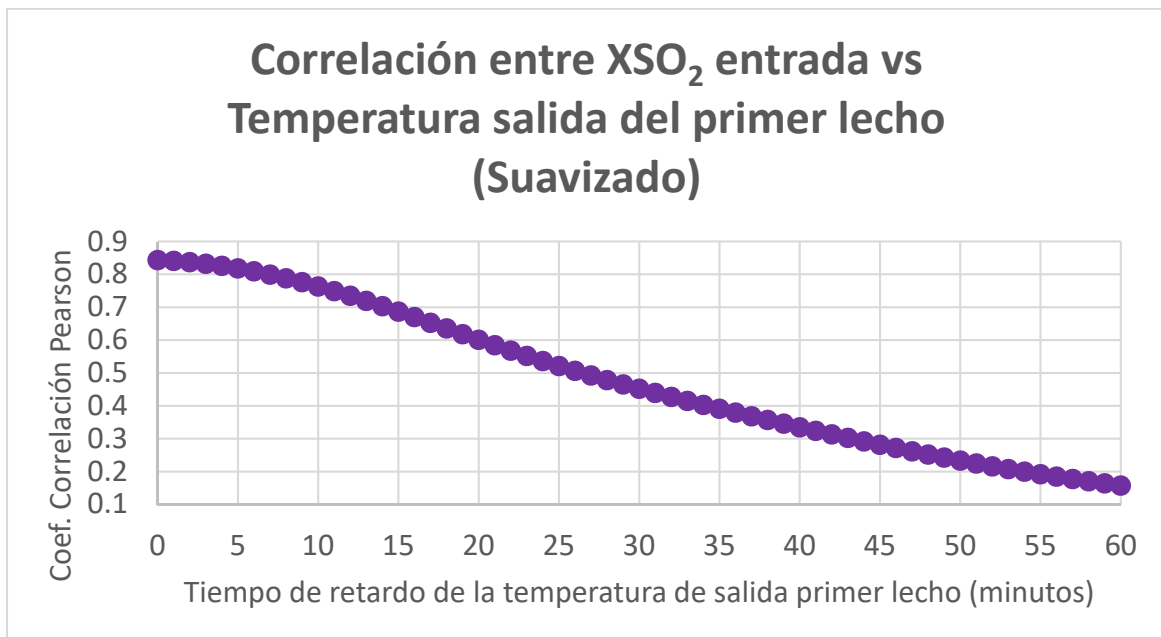


Figura 49 Coeficiente de correlación de Pearson entre la concentración de SO_2 de entrada y la temperatura de salida del primer lecho a diferentes tiempos de retardo (suavizado). Fuente: Elaboración propia.

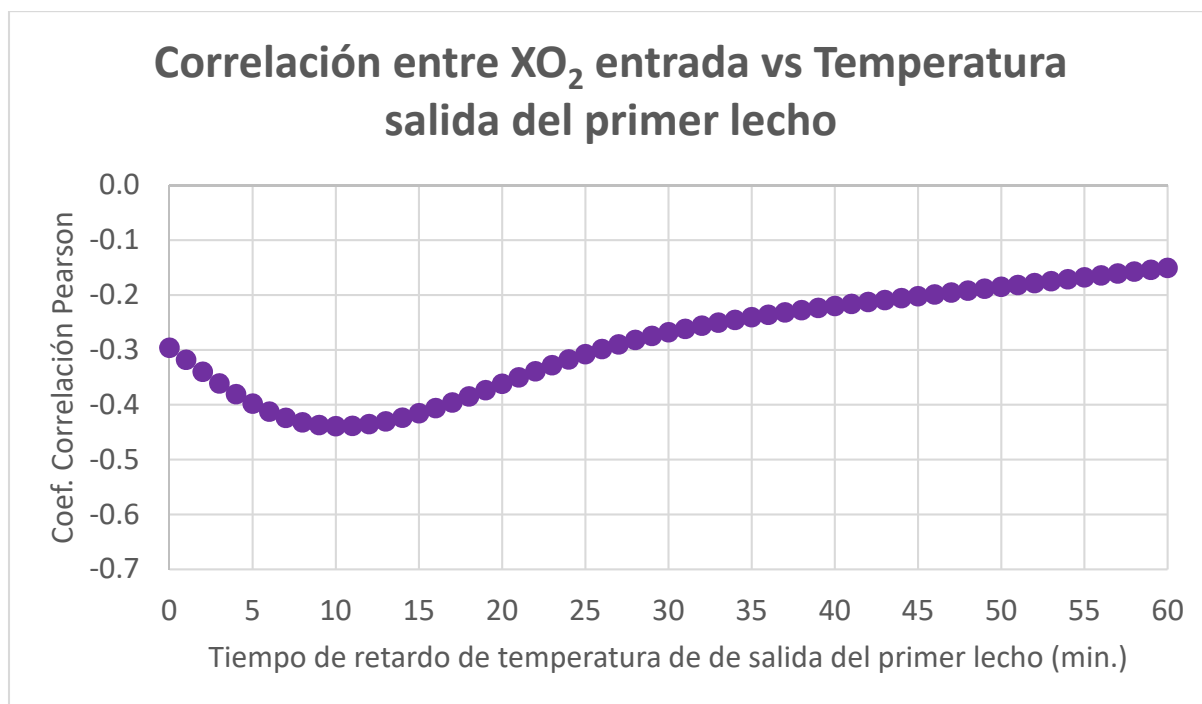


Figura 50 Coeficiente de correlación de Pearson entre la concentración de O_2 de entrada y la temperatura de salida del primer lecho a diferentes tiempos de retardo. Fuente: Elaboración propia.

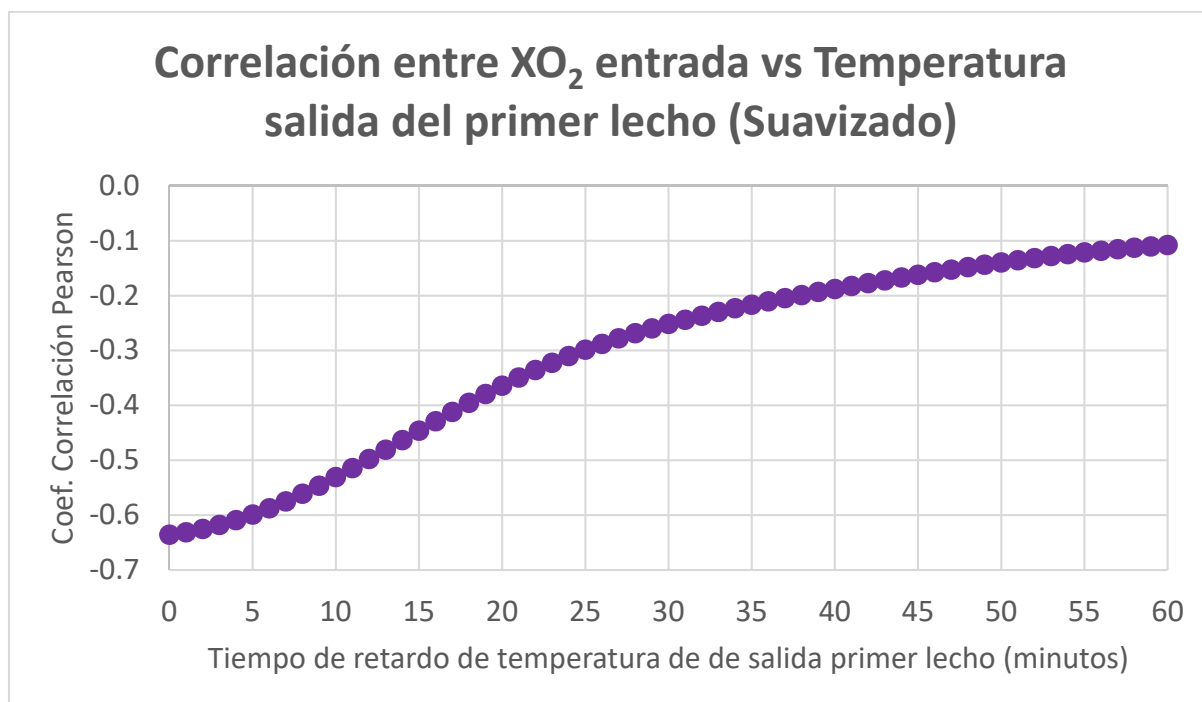


Figura 51 Coeficiente de correlación de Pearson entre la concentración de O_2 de entrada y la temperatura de salida del primer lecho a diferentes tiempos de retardo (suavizado). Fuente: Elaboración propia.

Ya con los datos sincronizados, se realiza el cálculo de la eficiencia de conversión en cada lecho del convertidor aplicando las ecuaciones 2.2, 2.3 y 2.4. En los resultados de

eficiencia se observa que el modelo en estado estacionario subestima o sobreestima la eficiencia de conversión en cada lecho, el desempeño global de predicción de la eficiencia de conversión se muestra en la tabla 11:

Tabla 11. Resultados de desempeño de predicción de la eficiencia de conversión mediante el sensor virtual derivado de leyes.

	Error porcentual medio absoluto (MAPE)%	Índice refinado de Concordancia (IOA)	Coefficiente de correlación (r)
Eficiencia de conversión con el sensor virtual derivado de leyes.	7,10%	-0,975	0,215

Fuente: Elaboración propia.

El modelo no muestra buen desempeño en la estimación de la eficiencia global de conversión. Esto se debe a que el modelo en estado estacionario posiblemente no se ajusta al comportamiento dinámico real del proceso de conversión y su complejidad debido a su cinética, lo cual requiere un mayor ajuste del modelo con información de los proveedores de catalizadores para describir con precisión el proceso de conversión. Asimismo, el modelo en estado estacionario no considera la comprensión de ciertos fenómenos de transferencia de calor y masa en sistemas porosos, siendo también limitado. (Werner S., 2020).

CAPÍTULO V. CONCLUSIONES

Se diseñó el sensor virtual basado en RNA para predecir la eficiencia de conversión con un error porcentual medio absoluto de 0,20%, un índice refinado de concordancia de 0,902 y un coeficiente de correlación redondeado a dos decimales igual a 1,00.

La concentración de SO₂ de entrada tiene una fuerte correlación con la concentración de SO₂ de salida ($r = 0,709$) a un tiempo de retardo de cuatro minutos, en comparación de las otras variables que tienen menor correlación con la concentración de SO₂ de salida.

La estructura óptima de la red neuronal para la predicción de SO₂ es 9-9-1. La estructura tiene nueve variables en la capa de entrada y nueve neuronas en la capa oculta. El modelo tiene un error porcentual medio absoluto de 12%, un índice refinado de Concordancia de 0.832 y un coeficiente de correlación de 0.945.

El sensor virtual derivado de leyes para la predicción de la eficiencia de conversión tiene un bajo desempeño con error porcentual medio absoluto de 7,10%. un índice refinado de concordancia de -0,975 y un coeficiente de correlación de 0,215.

El sensor virtual basado en RNA tiene un mejor desempeño de predicción de la eficiencia de conversión en comparación con el sensor virtual derivado de leyes en estado estacionario.

RECOMENDACIONES

Para mejorar la capacidad de estimación del sensor virtual basado en RNA se requiere aumentar la cantidad de datos de entrenamiento y validación a mayores concentraciones de SO₂ de salida.

Se recomienda realizar un mantenimiento al sensor virtual basado en RNA cada tres meses a fin de actualizar los datos de entrenamiento que garantice una representación actual del proceso, considerando cualquier cambio en la actividad del catalizador en el tiempo. El mantenimiento y ajuste del sensor virtual se podría realizar de 02 maneras: Actualizando el modelo existente con nuevos datos de entrada-salida o actualizando el modelo con datos antiguos y nuevos combinados (Brownlee J., 2021).

Se recomienda documentar el aseguramiento de calidad de los sensores y analizadores de la planta de ácido, a fin de garantizar la confiabilidad de las mediciones de las variables de entrada utilizados en el sensor virtual. Es necesario la verificación periódica de los analizadores de gases (SO₂ de entrada, SO₂ salida y O₂ de entrada) utilizando gases patrones certificados y trazables. Asimismo, el mantenimiento y calibración de los sensores de temperatura que se encuentran en cada lecho del reactor catalítico (existen dos sensores de temperatura en cada lecho).

Se recomienda contar con más datos de concentración de salida de SO₂ a condiciones transitorias a fin de mejorar la predicción del modelo en condiciones transitorias y posiblemente se logre incrementar la razón señal/ruido mayor a 5.

REFERENCIAS BIBLIOGRÁFICAS

- Al-Duwaish H., Ghouti L., Halawani T., Mohandes M. (2002). *Use of artificial neural networks process analyzers: A case study*. King Fahd University of petroleum and minerals, Saudi Arabia.
- Antson O., Pellikka T. (2014). *Predictive Emission Monitoring Systems, PEMS and their acceptance and use in Europe*. Research Report VTT-R-04313-14.
- Arahal M., Berenguel M. y Rodríguez F. (2006). *Técnicas de predicción con aplicaciones en ingeniería*. España: Secretariado de publicaciones universidad de Sevilla.
- Barrera R., Mora J. y Perez S. (2009). *Estimador neuronal del ángulo del rotor de generadores sincrónicos a partir de mediciones de tensión y corriente en terminales*. Ingeniería y competitividad, Volumen 11, Número 1, página 9-20. Colombia.
- Berzal F. (2018) *Redes neuronales y Deep learning*. España: Editorial universidad de Granada.
- Boger Z. (2009). *Artificial neural networks modeling to reduce industrial air pollution*. Optimal – Industrial neural systems. Israel.
- Brownlee J. (March 5, 2021) *How to Update Neural Network Models With More Data*. Machine learning mastery. <https://machinelearningmastery.com/update-neural-network-models-with-more-data/>
- Caron S., (2019). *Temperature modeling and control of the sulfuric acid plant in an industrial smelter*. [Tesis de maestría, Laurentian university], Ontario, Canada.
- Chowdhary A. R. (2014). *Application of self-tuning high-fidelity dynamic simulation model for soft-sensing*. 8th International conference on foundations of computer-aided process design – FOCAPD. July 13-17, 2014, Cle Elum, Washington, USA.
- Christensen K., Cavalca F., Beato P., Helveg S. (2014). *From nano-scale studies of working sulfuric acid catalysts to improved industrial-scale sulfuric acid production*. Sulphur 2014 Conference proceedings, Paris 3-6th november.
- Clifton, C. (2004) *Introduction to data mining*. USA. Purdue University.

- Colmenares G. (2004). *Curso de redes neuronales, capitulo 5, pre procesamiento de datos*. University of the Andes. Venezuela. Recuperado de http://webdelprofesor.ula.ve/economia/gcolmen/programa/redes_neuronales/capitulo5_preprocesamiento_de_datos.pdf
- Foddi O., Grosso M., Tronci S., Baratti R. (2007). *Prediction of distillate yields in a visbreaking process by neural networks*. Università degli studi di Cagliari, Italy.
- Fortuna L., Graziani S., Rizzo A. y Xibilia M. (2007). *Soft sensors for monitoring and control of industrial processes*, Italia: Springer.
- Hagan M. T., Menhaj M. (1994), *Training feedforward networks with the Marquardt algorithm*, IEEE Transactions on Neural Networks, 5(6), 989–993.
- Han J., Kamber M. (2006). *Data mining, Concepts and Techniques*. Segunda edición. Elsevier.
- Haykin S. (1999). *Neural networks. A comprehensive foundation*. Canada: Prentice-Hall.
- He J., Zhang J., Shang H. (2017) *A soft sensor for the sulphur dioxide converter in an industrial smelter*. The Canadian journal of chemical engineering, volume 95, june 2017. Canada.
- He J. (2018). *Dynamic modelling of catalytic SO₂ converter in a sulfuric acid plant of an industrial smelter*. [Tesis de doctorado, Laurentian University], Ontario, Canada.
- Hinkle D., Wiersma W., Jurs S. (2013). *Applied statistics for the behavioral sciences*. Houghton Mifflin, 5a edición.
- Isasi P., Galván I. (2004). *Redes neuronales artificiales, un enfoque practico*, España: Pearson.
- Kataria G., Singh K., Kumar R. (2016) *ANN based soft sensor model for reactive distillation column*. Department of chemical engineering, Malaviya National Institute of Technology, Jaipur, India.
- Kim D. (2017). *Comparison of shallow and deep neural networks in network intrusion detection*. [Tesis de maestría, California State University]. USA.
- King M., Davenport W., Moats M. (2013) *Sulfuric Acid Manufacture. Analysis, Control, and Optimization*. USA: Elsevier.

- Lapina O., Zhinimaev B., Boghosian S., Eriksen K y Fehrmann M. (1999). *Progress on the mechanistic understanding of SO₂ oxidation catalysts*. *Catalysis Today* 51, 469-479.
- Madsen H., Bacher P., Juhl R., Kloppenborg J. (2015). *Grey-Box Modeling; An approach to combined physical and statistical model building*. Mathematical Sciences Collaboration in Energy Systems Integration. Technical University of Denmark.
- Mann R. (1986). Controlling SO₂ emissions during start-up of multi-bed sulphuric acid converters. *Transactions of the Institute of Measurement and Control*. Vol 8, N° 3, July-Sept 1986.
- Marlin T. (2015). *Process Control. Designing Processes and Control Systems for Dynamic Performance*. Copyright 2015 by Thomas Marlin.
- Martin del Brío B., Sanz A. (2002). *Redes neuronales y sistemas difusos*. España. Alfaomega. 2da edición.
- Mhaskar H., Poggio T. (2016). *Deep vs. shallow networks: an approximation theory perspective*. Thesis. California State University,
- Montaño J. (2002). *Redes neuronales aplicadas al análisis de datos*. Universitat de les Illes Balears, España.
- Priddy K., Keller P. (2005). *Artificial neural networks. An introduction*. USA: SPIE Publications.
- Pavelka A. y Procházka A. (2004). *Algorithms for initialization of neural network weights*. Institute of Chemical Technology, Prague, Czech republic.
- Reyes L. y Tapia L. (2010). *Diseño e implementación de un programa basado en redes neuronales para el control de máquinas trifásicas de inducción*. Ecuador.
- Roffel B., Betlem B. (2006). *Process dynamics and control, modelling for control and prediction*, Netherlands: Wiley.
- Sarle W. (1994) *Neural networks and statistical models*. Proceedings of the nineteenth annual SAS users group international conference.

- Smith S. (1999). *The Scientist and Engineer's Guide to Digital Signal Processing*, USA, California Technical Publishing.
- Sørensen A. y Christensen K. (2016). *Understanding dynamics and emissions during sulfuric acid converter startup*. 40th International phosphate fertilizer & sulfuric acid technology conference clearwater beach, Florida June 10th & 11th.
- Souza F., Araújo R., Mendes J. (2016). *Review of soft sensors methods for regression applications*.
- Tarafdar A., Chandra N., Singh A. y Sirohi R. (2018). *Artificial neural network modeling of water activity: A low energy approach to freeze drying*. Food and Bioprocess Technology volume 11, pages164–171. India.
- Werner S. (2020). *Chemical process models – from first principle to hybrid models*. Blog article. Navigance. Alemania.
- ZareNezhad B., Aminian A. (2012). *An artificial neural network model for predicting the H₂S removal performance of piperazine solvents in gas sweetening plants*. Ministry of science, research and technology, Semnan University, Iran.

ANEXOS

ANEXO 1: Ecuación de estimación de la eficiencia de conversión derivado de leyes.

Jianjun He, Junfeng Zhang y Helen Shang proponen un modelo matemático para estimar la eficiencia de conversión y lo desarrollaron en las ecuaciones del A3.1 al A3.21, siendo obtenidas a partir del balance de masa y de energía en estado estacionario, aplicado en el lecho del convertidor como se muestra en la figura A3.1. Para el desarrollo de las ecuaciones se asume que el convertidor es adiabático y que la diferencia de temperatura entre el gas y el lecho del catalizador es despreciable.

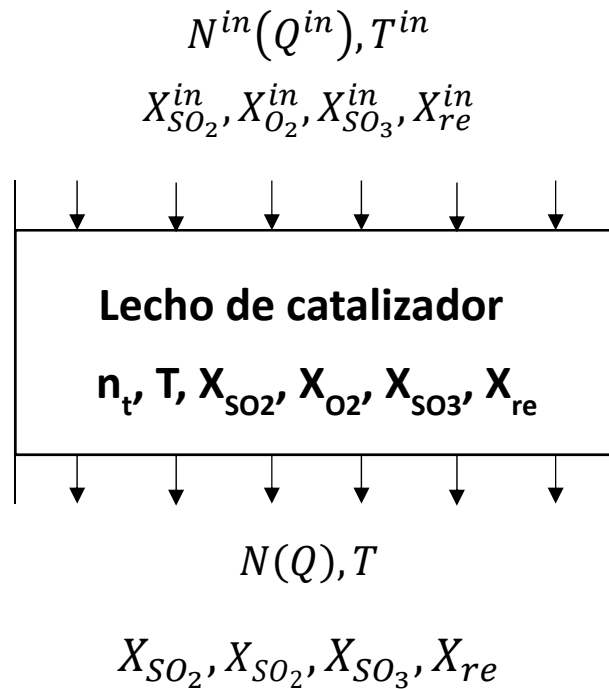


Figura A1.1 Esquema de variables en el lecho del convertidor

Donde:

N = Razón de flujo molar

Q = Razón de flujo volumétrico

X = Fracción molar (Concentración)

T = Temperatura

n_t = Cantidad molar de gas en el lecho del convertidor.

Aplicando un balance de masa particular al azufre y al oxígeno:

$$X_{SO_2}^{in} N^{in} + X_{SO_3}^{in} N^{in} = X_{SO_2} N + X_{SO_3} N \quad (A1.1)$$

$$2X_{SO_2}^{in} N^{in} + 2X_{O_2}^{in} N^{in} + 3X_{SO_3}^{in} N^{in} = 2X_{SO_2} N + 2X_{O_2} N + 3X_{SO_3} N \quad (A1.2)$$

Balance del gas inerte:

$$X_{re}^{in} N^{in} = X_{re} N \quad (A1.3)$$

Fracción molar del gas inerte:

$$X_{re}^{in} = 1 - X_{SO_2}^{in} - X_{O_2}^{in} - X_{SO_3}^{in} \quad (A1.4)$$

$$X_{re} = 1 - X_{SO_2} - X_{O_2} - X_{SO_3} \quad (A1.5)$$

Eficiencia de conversión de SO₂:

$$\Phi = \frac{X_{SO_2}^{in} N^{in} - X_{SO_2} N}{X_{SO_2}^{in} N^{in}} \quad (A1.6)$$

Resolviendo ecuaciones para expresar la eficiencia de conversión en términos de concentración de SO₂:

Despejando X_{SO_2} de la ecuación x:

$$X_{SO_2} = (1 - \Phi) X_{SO_2}^{in} \frac{N^{in}}{N} \quad (A1.7)$$

Reemplazando la ecuación (A3.7) en las ecuaciones (A3.1), (A3.2) y (A3.3):

$$X_{O_2} = \left(X_{O_2}^{in} - \frac{1}{2} \Phi X_{SO_2}^{in} \right) \frac{N^{in}}{N} \quad (A1.8)$$

$$X_{SO_3} = \left(X_{SO_3}^{in} + \Phi X_{SO_2}^{in} \right) \frac{N^{in}}{N} \quad (A1.9)$$

$$X_{re} = \left(1 - X_{SO_2}^{in} - X_{O_2}^{in} - X_{SO_3}^{in} \right) \frac{N^{in}}{N} \quad (A1.10)$$

De la ecuación estequiométrica por cada mol de SO₂ que se convierte, hay una reducción de medio mol en la cantidad total de moles.

$$N = \left(1 - \frac{1}{2} \Phi X_{SO_2}^{in} \right) N^{in} \quad (A1.11)$$

Expresando la relación de razón de flujo molar $\frac{N^{in}}{N}$:

$$\frac{N^{in}}{N} = \frac{1}{1 - \frac{1}{2} \Phi X_{SO_2}^{in}} \quad (A1.12)$$

Sustituyendo (A3.12) en las ecuaciones (A3.7) al (A3.9)

$$X_{SO_2} = \frac{(1-\Phi)X_{SO_2}^{in}}{1 - \frac{1}{2} \Phi X_{SO_2}^{in}} \quad (A1.13)$$

$$X_{O_2} = \frac{X_{O_2}^{in} - \frac{1}{2} \Phi X_{SO_2}^{in}}{1 - \frac{1}{2} \Phi X_{SO_2}^{in}} \quad (A1.14)$$

$$X_{SO_3} = \frac{X_{SO_3}^{in} + \Phi X_{SO_2}^{in}}{1 - \frac{1}{2} \Phi X_{SO_2}^{in}} \quad (A1.15)$$

$$X_{re} = \frac{1 - X_{SO_2}^{in} - X_{O_2}^{in} - X_{SO_3}^{in}}{1 - \frac{1}{2} \Phi X_{SO_2}^{in}} \quad (A1.16)$$

Sustituyendo (A3.11) en la ecuación (A3.6) se obtiene la ecuación de la eficiencia de conversión como función de la concentración de SO₂:

$$\Phi = \frac{X_{SO_2}^{in} - X_{SO_2}}{X_{SO_2}^{in} - \frac{1}{2}X_{SO_2}X_{SO_2}^{in}} \quad (A1.17)$$

Determinación de la eficiencia de conversión con ecuaciones de balance de energía:

Balance de energía en el lecho del convertidor:

$$\begin{aligned} & X_{SO_2}^{in} N^{in} H_{SO_2}^{in} + X_{O_2}^{in} N^{in} H_{O_2}^{in} + X_{SO_3}^{in} N^{in} H_{SO_3}^{in} + X_{re}^{in} N^{in} H_{re}^{in} \\ &= X_{SO_2} N H_{SO_2} + X_{O_2} N H_{O_2} + X_{SO_3} N H_{SO_3} + X_{re} N H_{re} \end{aligned} \quad (A1.18)$$

Donde:

H = entalpía/mol del gas.

Sustituyendo las ecuaciones del (A3.7) al (A3.10) en la ecuacion (A3.18) se obtiene:

$$\begin{aligned} & X_{SO_2}^{in} (H_{SO_2}^{in} - H_{SO_2}) + X_{O_2}^{in} (H_{O_2}^{in} - H_{O_2}) + X_{SO_3}^{in} (H_{SO_3}^{in} - H_{SO_3}) + X_{re}^{in} (H_{re}^{in} - H_{re}) \\ &= - \left(H_{SO_2} + \frac{1}{2} H_{O_2} - H_{SO_3} \right) \Phi X_{SO_2}^{in} \end{aligned} \quad (A1.19)$$

De acuerdo a Davenport (2006) las entalpias pueden aproximarse a una ecuacion lineal y son una funcion de la temperatura:

$$H = k_p T + b_p \quad (A1.20)$$

Donde :

$$k_{p,SO_2} = 0,05161 \text{ KJ}/(\text{mol. K}),$$

$$k_{p,O_2} = 0,03333 \text{ KJ}/(\text{mol. K}),$$

$$k_{p,SO_3} = 0,07144 \text{ KJ}/(\text{mol. K}),$$

$$k_{p,N_2} = 0,03110 \text{ KJ}/(\text{mol. K}),$$

$$b_{p,SO_2} = -314,3 \text{ KJ/mol}$$

$$b_{p,O_2} = -10,79 \text{ KJ/mol}$$

$$b_{p,SO_3} = -420,6 \text{ KJ/mol}$$

$$b_{p,N_2} = -9,797 \text{ KJ/mol}$$

Sustituyendo la ecuación (A3.20) en (A3.19) se obtiene la siguiente ecuación de la eficiencia de conversión como función de la temperatura y concentración:

$$\Phi = \frac{\left(X_{SO_2}^{in} k_{p,SO_2} + X_{O_2}^{in} k_{p,O_2} + X_{SO_3}^{in} k_{p,SO_3} + X_{re}^{in} k_{p,N_2}\right)(T - T^{in})}{\left[\left(k_{p,SO_2} + \frac{1}{2}k_{p,O_2} - k_{p,SO_3}\right)T + \left(b_{p,SO_2} + \frac{1}{2}b_{p,O_2} - b_{p,SO_3}\right)\right]X_{SO_2}^{in}} \quad (\text{A1.21})$$

ANEXO 2: Redes de retropropagación en Matlab

a) Introducción.

El toolbox de redes neuronales de Matlab, contempla las redes neuronales del tipo retropropagación y utiliza rutinas en código fuente donde el usuario puede adaptarlas a sus necesidades.

b) Funciones en redes neuronales superficiales de retropropagación.

- **Creación de una red alimentada hacia delante (feedforward)**

Newff: Función que crea una red neuronal tipo feed-forward usando el algoritmo de propagación hacia atrás como entrenamiento.

Sintaxis:

$$\text{net} = \text{newff}(\text{PR}, [\text{S1 S2} \dots \text{SN1}], \{\text{TF1 TF2} \dots \text{TFN1}\}, \text{BTF}, \text{PF}) \quad (\text{A2.1})$$

Descripción:

$$\text{newff}(\text{PR}, [\text{S1 S2} \dots \text{SN1}], \{\text{TF1 TF2} \dots \text{TFN1}\}, \text{BTF}, \text{PF}), \quad (\text{A2.2})$$

Donde:

PR: Matriz de $R \times 2$ de los valores mínimos y máximos de cada elemento vector de entrada.

Si : Tamaño de cada capa, número de capas.

TFi: Tipo de función de activación en cada capa, por defecto = 'tansig'.

BTF: Nombre del tipo de función de entrenamiento de propagación hacia atrás, por defecto = 'traingdx'.

PF: Función de Performance de la red, por defecto = 'mse'.

Las funciones de transferencia TFi pueden ser alguna función como tansig, logsig, or purelin.

La función de entrenamiento BTF puede ser alguna de las funciones de entrenamiento de propagación hacia atrás, como también están trainlm, trainbfg, trainrp, traingd, etc.

Precaución: trainlm es la función de entrenamiento por defecto porque es más rápido. Pero se requiere una gran cantidad de memoria para correr. Si al correr el entrenamiento

sale un error "out-of-memory" por falta de memoria, entonces trata de hacer lo siguiente:

- Se puede entrenar lentamente la función `trainlm`, tomando en cuenta los requerimientos de memoria reducida para esto se utiliza en los parámetros de aprendizaje seteando `net.trainParam.mem_reduc` igual a 2 o más.
- Usar `trainbfg`, el cual es más lento, pero de memoria más eficiente que `trainlm`.
- Usar `trainrp` el cual es más lento, pero de memoria más eficiente que `trainbfg`.

La función de rendimiento de la red puede ser la función `mse` o `msereg`.

- **Inicialización de pesos**

Antes de entrenar una red alimentada hacia delante (feedforward), los pesos y bias tienen que ser inicializados. Las bias y pesos iniciales son creados con el comando `init`. Esta función toma una red objeto como una entrada y retorna una red objeto con todos los pesos y bias inicializados. La red se inicializa con la siguiente función:

$$\text{net} = \text{init}(\text{net}) \quad (\text{A2.3})$$

La función de inicialización `init` es llamado por `newff`, por lo tanto, la red es automáticamente inicializado con los parámetros por default cuando esto es creado, y `init` no tiene que ser llamado separadamente.

Sin embargo, el método de Nguyen y Widrow realiza la inicialización de los pesos en el intervalo $[-0.5 \text{ a } 0.5]$ y con una distribución uniforme. La red se inicializa con la siguiente función:

$$\text{net} = \text{initnw}(\text{net}) \quad (\text{A2.4})$$

- **Entrenamiento de la red.**

Antes de entrenar la red se debe definir los siguientes parámetros de aprendizaje.

- `Net.trainParam.show`: Muestra los resultados en cada posición de las iteraciones.
- `Net.trainParam.lr`: Tasa de aprendizaje usado en algunos esquemas de gradiente. (`traingd`, `traingdm`, `traingda`, `traingdx`).
- `Net.trainParam.epochs`: Define el número máximo de iteraciones.
- `Net.trainParam.goal`: Tolerancia de error. Criterio de parada.

Train: Función que entrena una red neuronal de acuerdo a la función de entrenamiento en la red y a los parámetros de entrenamiento.

Sintaxis:

$$[\text{net}, \text{tr}] = \text{train}(\text{NET}, \text{P}, \text{T}) \quad (\text{A2.5})$$

Donde se ingresan:

NET: Red.

P: Entrada de la red.

T: Objetivo o salida de la red.

Y retorna:

net: Nueva red.

tr: Graba el entrenamiento (iteraciones y perfil).

- **Simulación de la red.**

Sim: Función que simula una red neuronal.

Sintaxis:

$$[\text{Y}, \text{Pf}, \text{Af}] = \text{sim}(\text{net}, \text{P}, \text{Pi}, \text{Ai}) \quad (\text{A2.6})$$

Donde se ingresan:

Net: Red que ha sido entrenado.

P: Entrada de la red.

Y retorna:

Y: Salida de la red.

c) Excellink.

Es un toolbox que tiene el Matlab y se utiliza para trabajar con el Excel a través de una interfase. La forma de operar el excellink en el Excel se reduce básicamente al uso de tres funciones: `mlputmatrix`, `mlevalstring`, y `mlgetmatrix`.

mlputmatrix: Esta función se encarga de mandar los datos al Matlab.

Ej: `=mlputmatrix("nombre de la variable", espacio que ocupan los datos)`.

mlgetmatrix: Esta función se encarga de traer los datos del Matlab a Excel.

Ej: `=mlgetmatrix("nombre de la variable", "espacio que ocupan los datos")`.

mlevalstring: Esta función se encarga de evaluar las variables.

Ej: `=mlevalstring("nombre de la variable")`.

ANEXO 3: Programa en VBA de Excel.

Function entrenamiento(aData, starget1, starget2, starget3, starget4, starget5, starget6, contador)

'Mandando los datos de excel al matlab

MLPutMatrix "data_all", aData

'Datos de Entrenamiento

MLEvalString "P=data_all (1:10727,1:9)"

MLEvalString "T=data_all (1:10727,13)"

'Datos de Validación

MLEvalString "a=data_all (10728:12068,1:9)"

MLEvalString "b=data_all (10728:12068,13)"

'Datos de Prueba final

MLEvalString "c=data_all (12069:13409,1:9)"

MLEvalString "d=data_all (12069:13409,13)"

'Normalización de los datos

MLEvalString "[pn, minp,maxp,tn, mint,maxt]=premmx(P',T)"

MLEvalString "[an, mina,maxa,bn, minb,maxb]=premmx(a',b)"

MLEvalString "[cn, minc,maxc,dn, mind,maxd]=premmx(c',d)"

'Creando la estructura de datos para el método de detención temprana (Early Stopping).

MLEvalString "v.P=an"

MLEvalString "v.T=bn"

MLEvalString "t.P=cn"

MLEvalString "t.V=dn"

'Contador de entrenamientos variando el número de unidades ocultas.

contador = Sheets("RNA").Range("M11").Value

'Arquitecturas de la red de 01 a 30 capas ocultas.

Select Case contador

Case Is = 0

```
MLEvalString "net=newff(minmax(pn), [1,1], {'tansig', 'purelin'}, 'trainlm')"
```

Case Is = 1

```
MLEvalString "net=newff(minmax(pn), [2,1], {'tansig', 'purelin'}, 'trainlm')"
```

Case Is = 2

```
MLEvalString "net=newff(minmax(pn), [3,1], {'tansig', 'purelin'}, 'trainlm')"
```

Case Is = 3

```
MLEvalString "net=newff(minmax(pn), [4,1], {'tansig', 'purelin'}, 'trainlm')"
```

Case Is = 4

```
MLEvalString "net=newff(minmax(pn), [5,1], {'tansig', 'purelin'}, 'trainlm')"
```

Case Is = 5

```
MLEvalString "net=newff(minmax(pn), [6,1], {'tansig', 'purelin'}, 'trainlm')"
```

Case Is = 6

```
MLEvalString "net=newff(minmax(pn), [7,1], {'tansig', 'purelin'}, 'trainlm')"
```

Case Is = 7

```
MLEvalString "net=newff(minmax(pn), [8,1], {'tansig', 'purelin'}, 'trainlm')"
```

Case Is = 8

```
MLEvalString "net=newff(minmax(pn), [9,1], {'tansig', 'purelin'}, 'trainlm')"
```

Case Is = 9

```
MLEvalString "net=newff(minmax(pn), [10,1], {'tansig', 'purelin'}, 'trainlm')"
```

Case Is = 10

```
MLEvalString "net=newff(minmax(pn), [11,1], {'tansig', 'purelin'}, 'trainlm')"
```

Case Is = 11

```
MLEvalString "net=newff(minmax(pn), [12,1], {'tansig', 'purelin'}, 'trainlm')"
```

Case Is = 12

```
MLEvalString "net=newff(minmax(pn), [13,1], {'tansig', 'purelin'}, 'trainlm')"
```


Case Is = 13

```
MLEvalString "net=newff(minmax(pn), [14,1], {'tansig', 'purelin'}, 'trainlm')"
```

Case Is = 14

```
MLEvalString "net=newff(minmax(pn), [15,1], {'tansig', 'purelin'}, 'trainlm')"
```

Case Is = 15

```
MLEvalString "net=newff(minmax(pn), [16,1], {'tansig', 'purelin'}, 'trainlm')"
```

Case Is = 16

```
MLEvalString "net=newff(minmax(pn), [17,1], {'tansig', 'purelin'}, 'trainlm')"
```

Case Is = 17

```
MLEvalString "net=newff(minmax(pn), [18,1], {'tansig', 'purelin'}, 'trainlm')"
```

Case Is = 18

```
MLEvalString "net=newff(minmax(pn), [19,1], {'tansig', 'purelin'}, 'trainlm')"
```

Case Is = 19

```
MLEvalString "net=newff(minmax(pn), [20,1], {'tansig', 'purelin'}, 'trainlm')"
```

Case Is = 20

```
MLEvalString "net=newff(minmax(pn), [21,1], {'tansig', 'purelin'}, 'trainlm')"
```

Case Is = 21

```
MLEvalString "net=newff(minmax(pn), [22,1], {'tansig', 'purelin'}, 'trainlm')"
```

Case Is = 22

```
MLEvalString "net=newff(minmax(pn), [23,1], {'tansig', 'purelin'}, 'trainlm')"
```

Case Is = 23

```
MLEvalString "net=newff(minmax(pn), [24,1], {'tansig', 'purelin'}, 'trainlm')"
```

Case Is = 24

```
MLEvalString "net=newff(minmax(pn), [25,1], {'tansig', 'purelin'}, 'trainlm')"
```

Case Is = 25

```
MLEvalString "net=newff(minmax(pn), [26,1], {'tansig', 'purelin'}, 'trainlm')"
```

Case Is = 26

MLEvalString "net=newff(minmax(pn), [27,1], {'tansig', 'purelin'}, 'trainlm')"

Case Is = 27

MLEvalString "net=newff(minmax(pn), [28,1], {'tansig', 'purelin'}, 'trainlm')"

Case Is = 28

MLEvalString "net=newff(minmax(pn), [29,1], {'tansig', 'purelin'}, 'trainlm')"

Case Is = 29

MLEvalString "net=newff(minmax(pn), [30,1], {'tansig', 'purelin'}, 'trainlm')"

Case Else

End Select

'Inicializacion de la red con el método de Nguyen and Widrow

MLEvalString "net=initnw(net)"

'Entrenamiento de la red con el método de detención temprana

MLEvalString " [net,tr]=train(net,pn,tn,[],[],v,t)"

'Simulando la red con los datos

MLEvalString "y=sim(net,pn)"

MLEvalString "val=sim(net,an)"

MLEvalString "test=sim(net,cn)"

'Convirtiendo los datos a su escala original

MLEvalString "t=postmnmx(y',mint,maxt)"

MLEvalString "val_t=postmnmx(val',minb,maxb)"

MLEvalString "test_t=postmnmx(test',mind,maxd)"

'Creando las tablas de resultados

MLEvalString "Real_Train=T"

MLEvalString "Pred_Train=t"

MLEvalString "Real_Verif=b"

```
MLEvalString "Pred_Verif=val_t"
```

```
MLEvalString "Real_Test_final=d"
```

```
MLEvalString "Pred_Test_final=test_t"
```

```
'Traendo al excel los resultados reales y de predicción.
```

```
Mlgetmatrix "Real_Train", starget1
```

```
Mlgetmatrix "Pred_Train", starget2
```

```
Mlgetmatrix "Real_Verif", starget3
```

```
Mlgetmatrix "Pred_Verif", starget4
```

```
Mlgetmatrix "Real_Test_final", starget5
```

```
Mlgetmatrix "Pred_Test_final", starget6
```

```
End Function
```

ANEXO 4: Ejemplo de cálculo del tiempo de retardo.

Ejemplo de cálculo en Excel del coeficiente de correlación de Pearson entre la concentración de entrada de SO₂ y la temperatura de salida de la cama 1, con un tiempo de retardo 0 minutos y 1 minuto.

Cálculo del coef. con tiempo de retardo de 0 minutos:

BS3		=COEF.DE.CORREL(\$D3:\$D13458;E3:E13458)					
	BR	BS	BT	BU	BV	BW	BX
1							
2	Concentración de SO ₂ entrada a 1ª cama	Temperatura Salida Cama Nº 1	Temperatura Salida Cama Nº 1_t+1	Temperatura Salida Cama Nº 1_t+2	Temperatura Salida Cama Nº 1_t+3	Temperatura Salida Cama Nº 1_t+4	Temperatura Salida Cama Nº 1_t+5
3	% SO ₂	0.523	0.552	0.581	0.608	0.634	0.658
4	Tiempo de retardo (min)	0	1	2	3	4	5

Datos:

	B	D	E	F	G	H	I	J
1								
2	Fecha	Concentración de SO ₂ entrada a 1ª cama, %	Temperatura Salida Cama Nº 1, °K	Temperatura Salida Cama Nº 1_t+1	Temperatura Salida Cama Nº 1_t+2	Temperatura Salida Cama Nº 1_t+3	Temperatura Salida Cama Nº 1_t+4	Temperatura Salida Cama Nº 1_t+5
3	2020-01-14 00:00	11.0	905.2	904.9	904.5	904.1	904.0	903.9
4	2020-01-14 00:01	10.9	904.9	904.5	904.1	904.0	903.9	903.9
5	2020-01-14 00:02	10.6	904.5	904.1	904.0	903.9	903.9	904.0
6	2020-01-14 00:03	10.7	904.1	904.0	903.9	903.9	904.0	904.2
7	2020-01-14 00:04	11.6	904.0	903.9	903.9	904.0	904.2	904.4
8	2020-01-14 00:05	11.8	903.9	903.9	904.0	904.2	904.4	904.5
9	2020-01-14 00:06	11.6	903.9	904.0	904.2	904.4	904.5	904.5
10	2020-01-14 00:07	11.2	904.0	904.2	904.4	904.5	904.5	904.6
11	2020-01-14 00:08	11.2	904.2	904.4	904.5	904.5	904.6	904.6
12	2020-01-14 00:09	11.0	904.4	904.5	904.5	904.6	904.6	904.5
13	2020-01-14 00:10	10.9	904.5	904.5	904.6	904.6	904.5	904.3
14	2020-01-14 00:11	10.8	904.5	904.6	904.6	904.5	904.3	903.8
15	2020-01-14 00:12	10.6	904.6	904.6	904.5	904.3	903.8	903.2
16	2020-01-14 00:13	10.7	904.6	904.5	904.3	903.8	903.2	902.7
17	2020-01-14 00:14	11.0	904.5	904.3	903.8	903.2	902.7	902.2
18	2020-01-14 00:15	10.3	904.3	903.8	903.2	902.7	902.2	901.7
19	2020-01-14 00:16	8.9	903.8	903.2	902.7	902.2	901.7	901.2
20	2020-01-14 00:17	8.4	903.2	902.7	902.2	901.7	901.2	900.8
21	2020-01-14 00:18	9.1	902.7	902.2	901.7	901.2	900.8	900.5
22	2020-01-14 00:19	9.8	902.2	901.7	901.2	900.8	900.5	900.5
23	2020-01-14 00:20	10.1	901.7	901.2	900.8	900.5	900.5	900.7

Cálculo del coef. con tiempo de retardo de 1 minuto:

BT3		=COEF.DE.CORREL(\$D3:\$D13458;F3:F13458)					
	BR	BS	BT	BU	BV	BW	BX
1							
2	Concentración de SO ₂ entrada a 1ª cama	Temperatura Salida Cama Nº 1	Temperatura Salida Cama Nº 1_t+1	Temperatura Salida Cama Nº 1_t+2	Temperatura Salida Cama Nº 1_t+3	Temperatura Salida Cama Nº 1_t+4	Temperatura Salida Cama Nº 1_t+5
3	% SO ₂	0.523	0.552	0.581	0.608	0.634	0.658
4	Tiempo de retardo (min)	0	1	2	3	4	5

ANEXO 5: Ejemplo de cálculo en Excel de los indicadores de desempeño de los modelos.

Se utilizan 12 pares de datos a modo de ejemplo.

$\sum_{i=1}^n |P_i - O_i| = \text{SUMA}(F5:F16)$ $\sum_{i=1}^n |O_i - \bar{O}| = \text{SUMA}(G5:G16)$

Sumatorias para el IOA:
 274 848

Conc. Real	Conc. Pred.	$ P_i - O_i $	$ O_i - \bar{O} $	Error Porcentual Absoluto ((O-P)/O)
249	209	40.1	10.2	16%
260	261	1.7	20.9	1%
311	327	16.2	72.1	5%
370	354	15.7	131.2	4%
205	187	18.4	33.8	9%
337	282	54.2	97.8	16%
331	286	45.1	91.9	14%
145	158	13.0	93.5	9%
167	144	23.3	72.0	14%
181	151	30.7	57.4	17%
179	167	11.7	59.9	7%
131	128	3.5	107.5	3%

Coefficiente de correlación "r"	Índice refinado de Concordancia (IOA)	Error porcentual medio absoluto (MAPE)%
0.963	0.839	9%

Prom. Conc. Real: 239 $=\text{PROMEDIO}(C5:C16)$

$\frac{|O_i - P_i|}{|O_i|} = \text{ABS}((C5-D5)/C5)$

$|O_i - \bar{O}| = \text{ABS}(C5-C\$18)$

$|P_i - O_i| = \text{ABS}(D5-C5)$

$= \frac{1}{n} \sum_{i=1}^n \frac{|O_i - P_i|}{|O_i|} = \text{PROMEDIO}(H5:H16)$

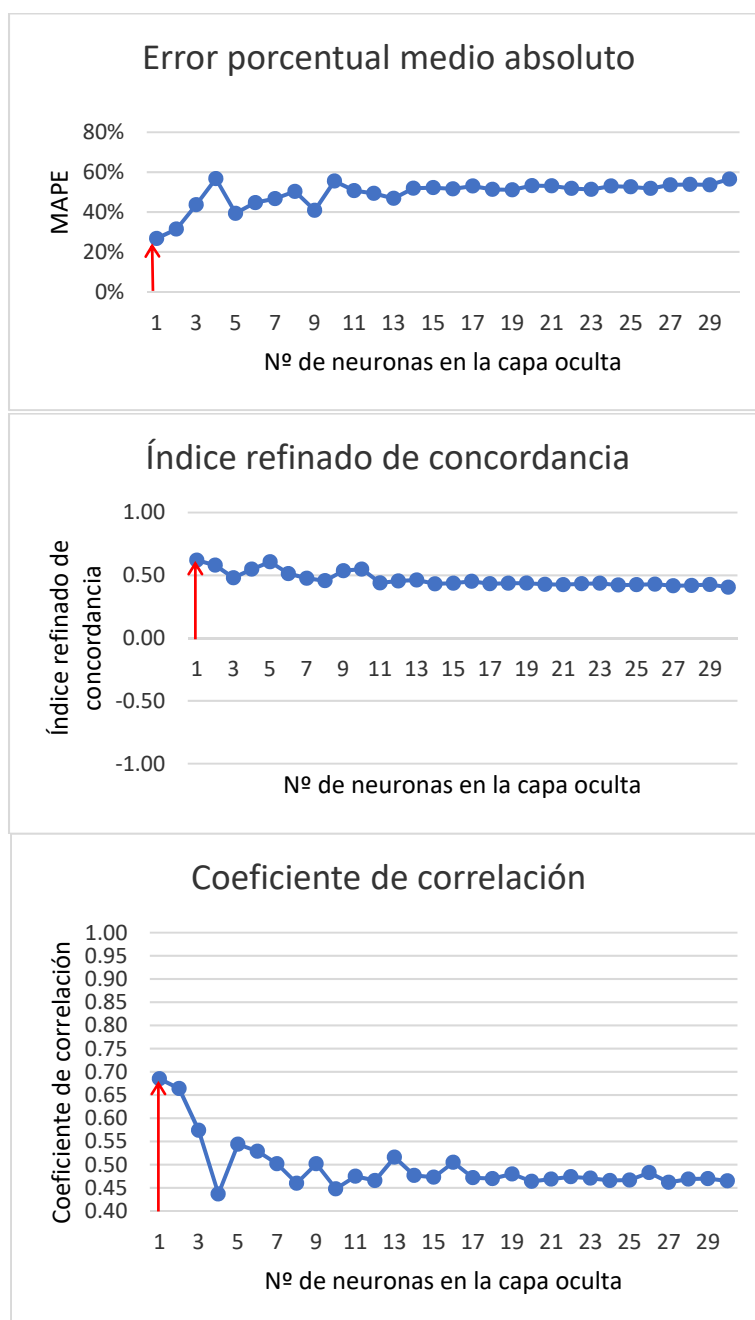
$= \begin{cases} 1 - \frac{\sum_{i=1}^n |P_i - O_i|}{2 \sum_{i=1}^n |O_i - \bar{O}|}, & \text{Cuando } \sum_{i=1}^n |P_i - O_i| \leq 2 \sum_{i=1}^n |O_i - \bar{O}| \\ \frac{2 \sum_{i=1}^n |O_i - \bar{O}|}{\sum_{i=1}^n |P_i - O_i|} - 1, & \text{Cuando } \sum_{i=1}^n |P_i - O_i| > 2 \sum_{i=1}^n |O_i - \bar{O}| \end{cases}$

$= \text{SI}(F3 > 2 * G3; ((2 * G3) / F3) - 1; 1 - (F3 / (2 * G3)))$

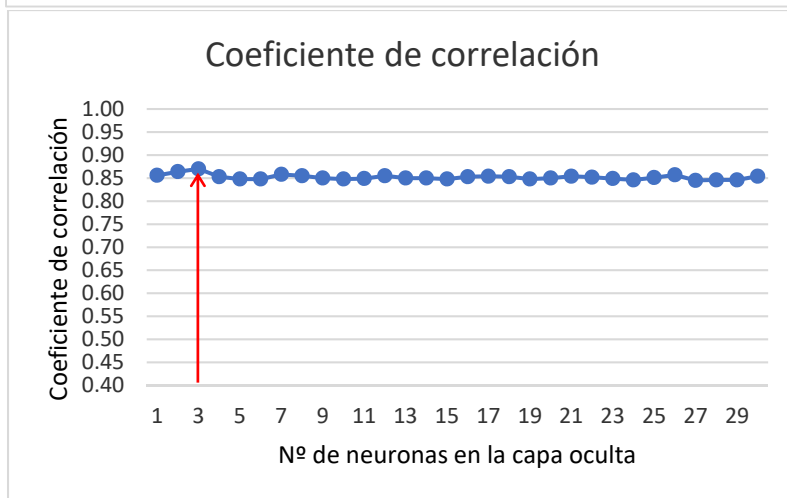
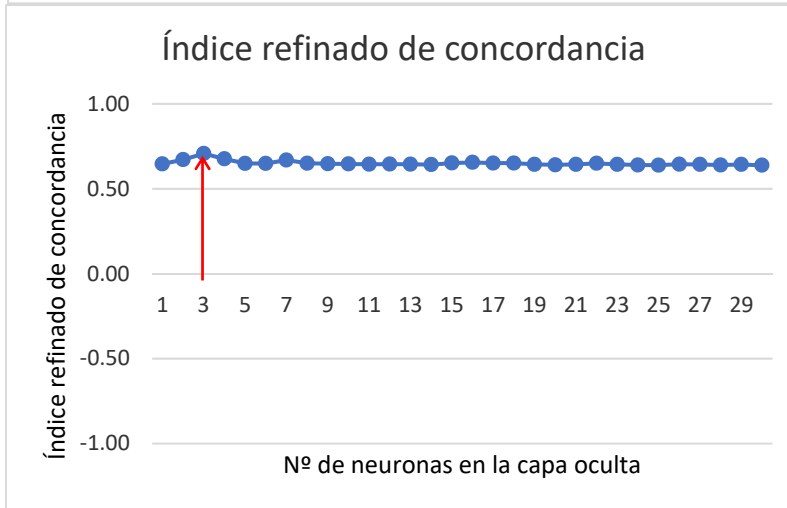
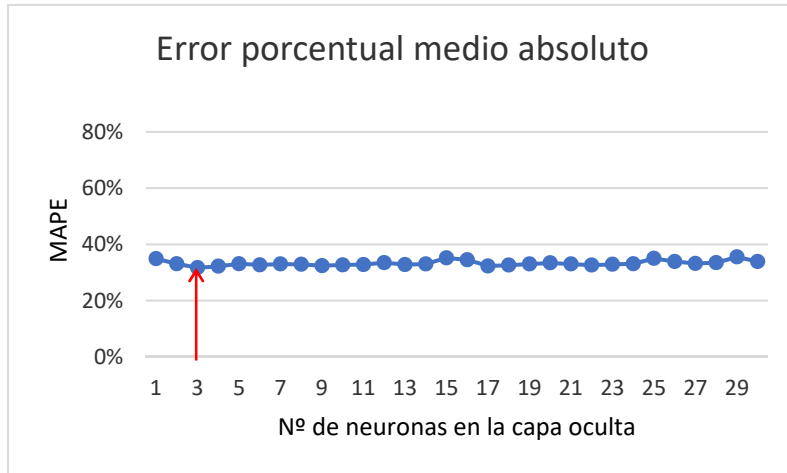
$= \text{COEF.DE.CORREL}(C5:C16; D5:D16)$

ANEXO 6: Indicadores de desempeño de los modelos de predicción de la concentración de SO₂ de salida, variando el número de variables de entrada y el número de neuronas en la capa oculta.

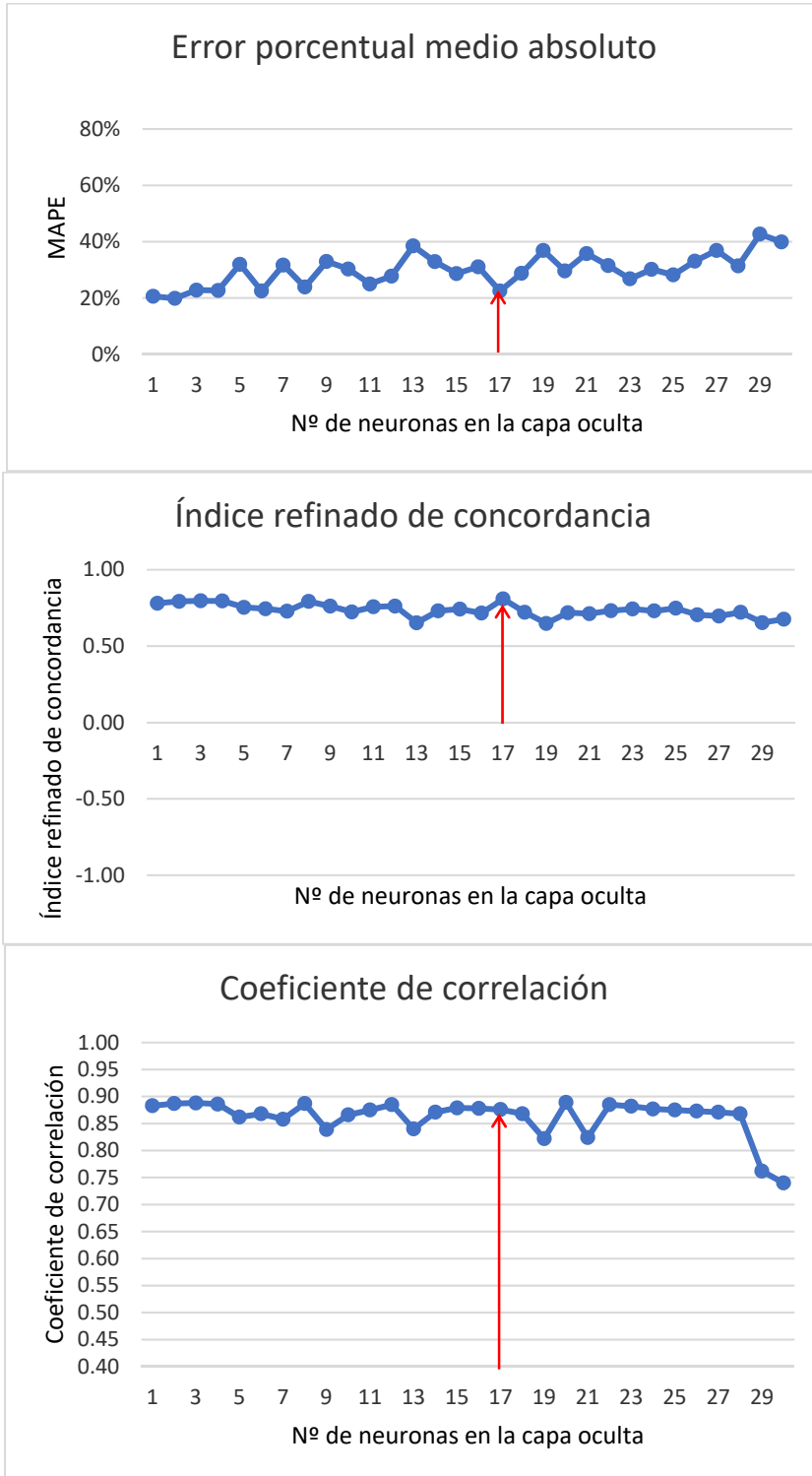
Nº de variables de entrada	Combinación de variables de entrada	Número óptimo de unidades ocultas	Error porcentual medio absoluto (MAPE)%	Índice refinado de concordancia (IOA)	Coef. de correlación "r"
1	Temp_sal_Lecho1(t)	1	26,7%	0,620	0,685



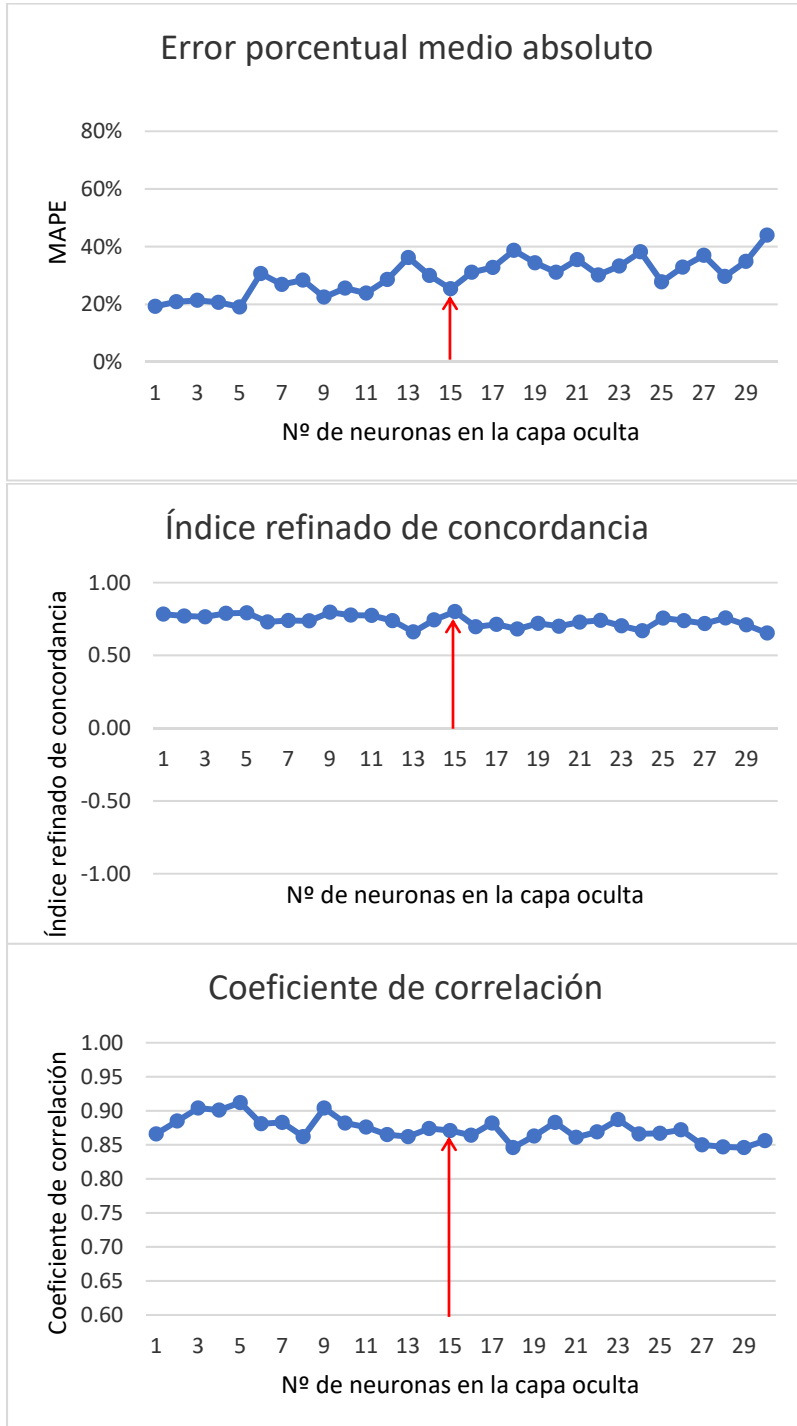
Nº de variables de entrada	Combinación de variables de entrada	Número óptimo de unidades ocultas	Error porcentual medio absoluto (MAPE)%	Índice refinado de concordancia (IOA)	Coef. de correlación "r"
1	Conc_SO2entrada(t-4)	3	31,7%	0,708	0,870



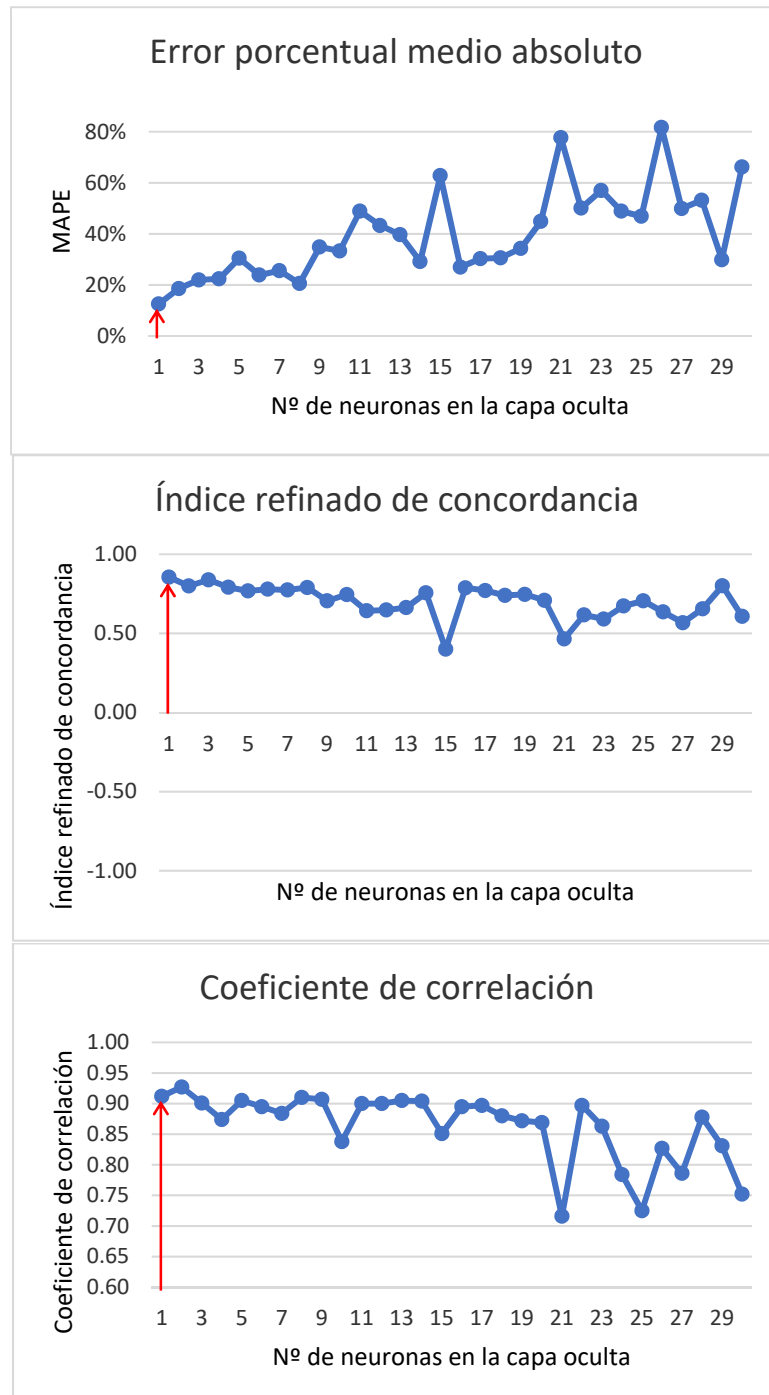
Nº de variables de entrada	Combinación de variables de entrada	Número óptimo de unidades ocultas	Error porcentual medio absoluto (MAPE)%	Índice refinado de concordancia (IOA)	Coef. de correlación "r"
2	Conc_SO2entrada(t-4), Temp_sal_Lecho1(t)	17	22,5%	0,809	0,876



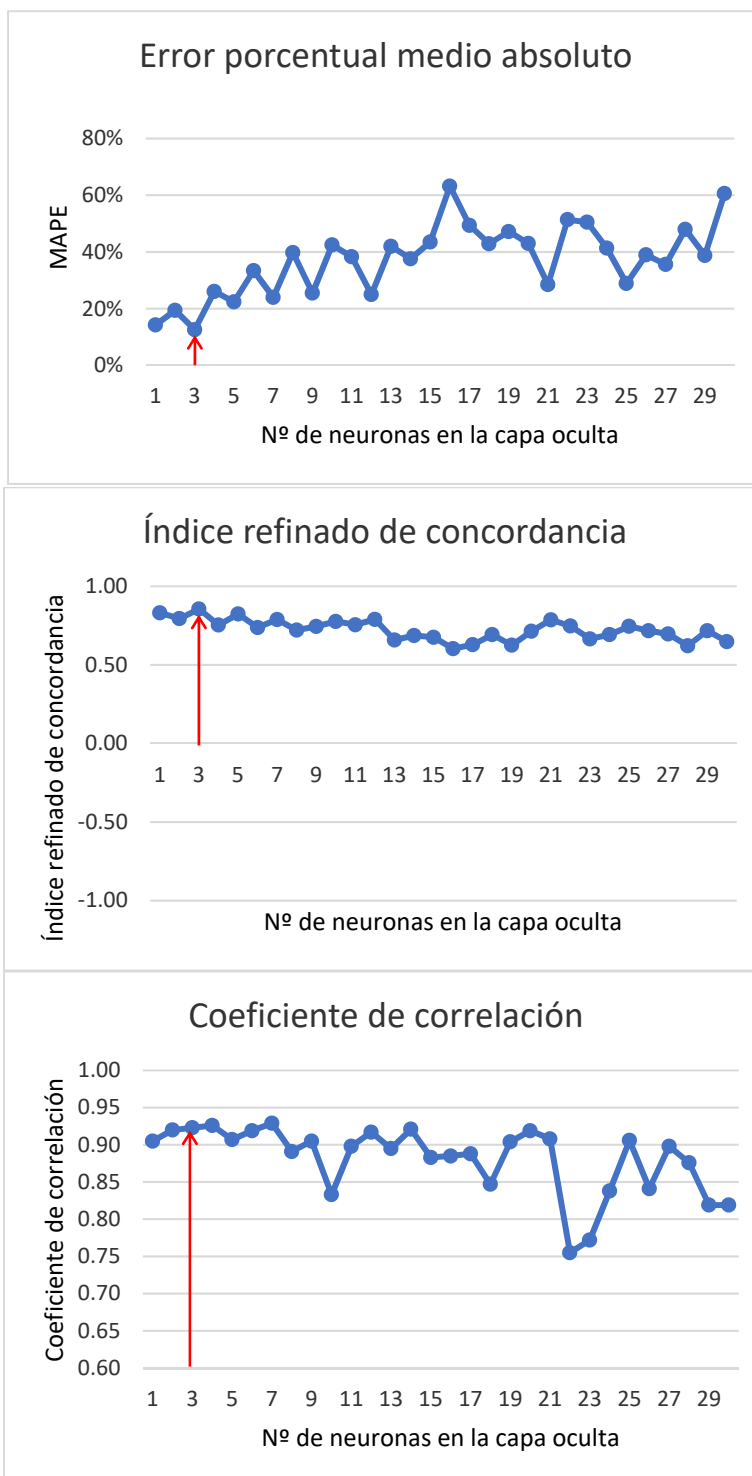
Nº de variables de entrada	Combinación de variables de entrada	Número óptimo de unidades ocultas	Error porcentual medio absoluto (MAPE)%	Índice refinado de concordancia (IOA)	Coef. de correlación "r"
3	Con_SO2entrada(t-4), Temp_sal_Lecho1(t), Temp_sal_Lecho2(t)	15	25,4%	0,801	0,871



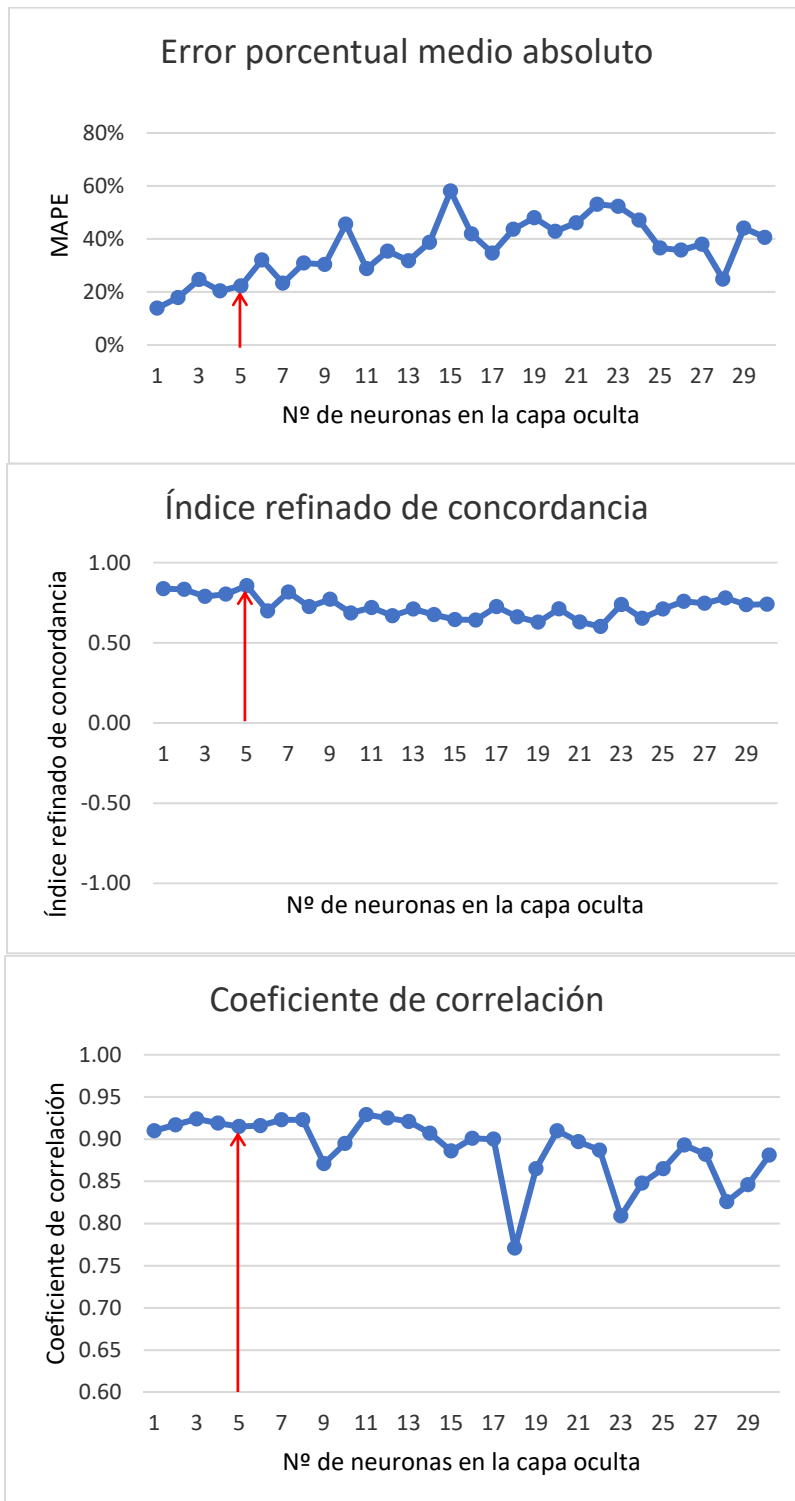
Nº de variables de entrada	Combinación de variables de entrada	Número óptimo de unidades ocultas	Error porcentual medio absoluto (MAPE)%	Índice refinado de concordancia (IOA)	Coef. de correlación "r"
4	Conc_SO2entrada(t-4), Temp_sal_Lecho1(t), Temp_sal_Lecho2(t), Temp_sal_Lecho3(t)	1	12,6%	0,856	0,912



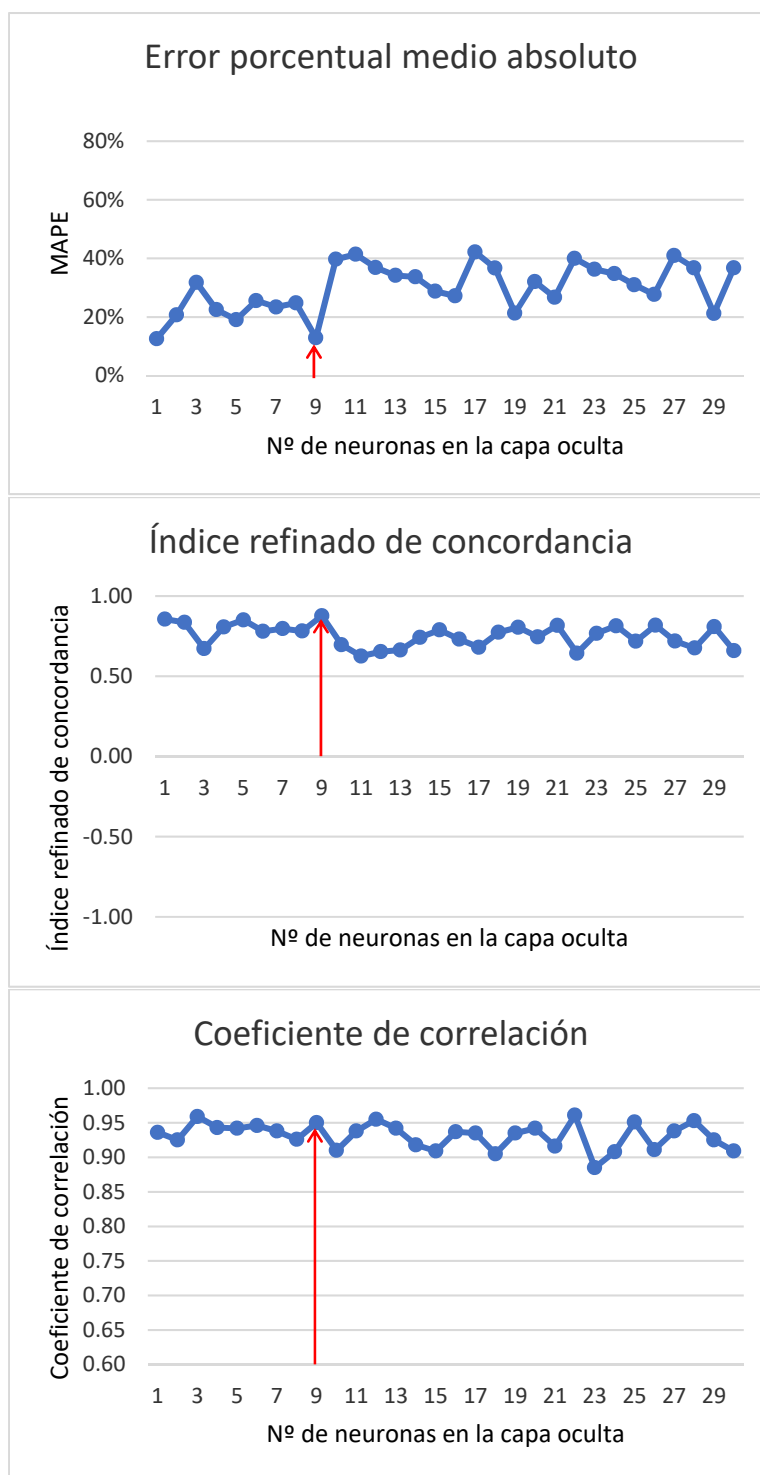
Nº de variables de entrada	Combinación de variables de entrada	Número óptimo de unidades ocultas	Error porcentual medio absoluto (MAPE)%	Índice refinado de concordancia (IOA)	Coef. de correlación "r"
5	Conc_SO2entrada(t-4), Temp_sal_Lecho1(t), Temp_sal_Lecho2(t), Temp_sal_Lecho3(t), Temp_sal_Lecho4(t)	3	12,6%	0,856	0,923



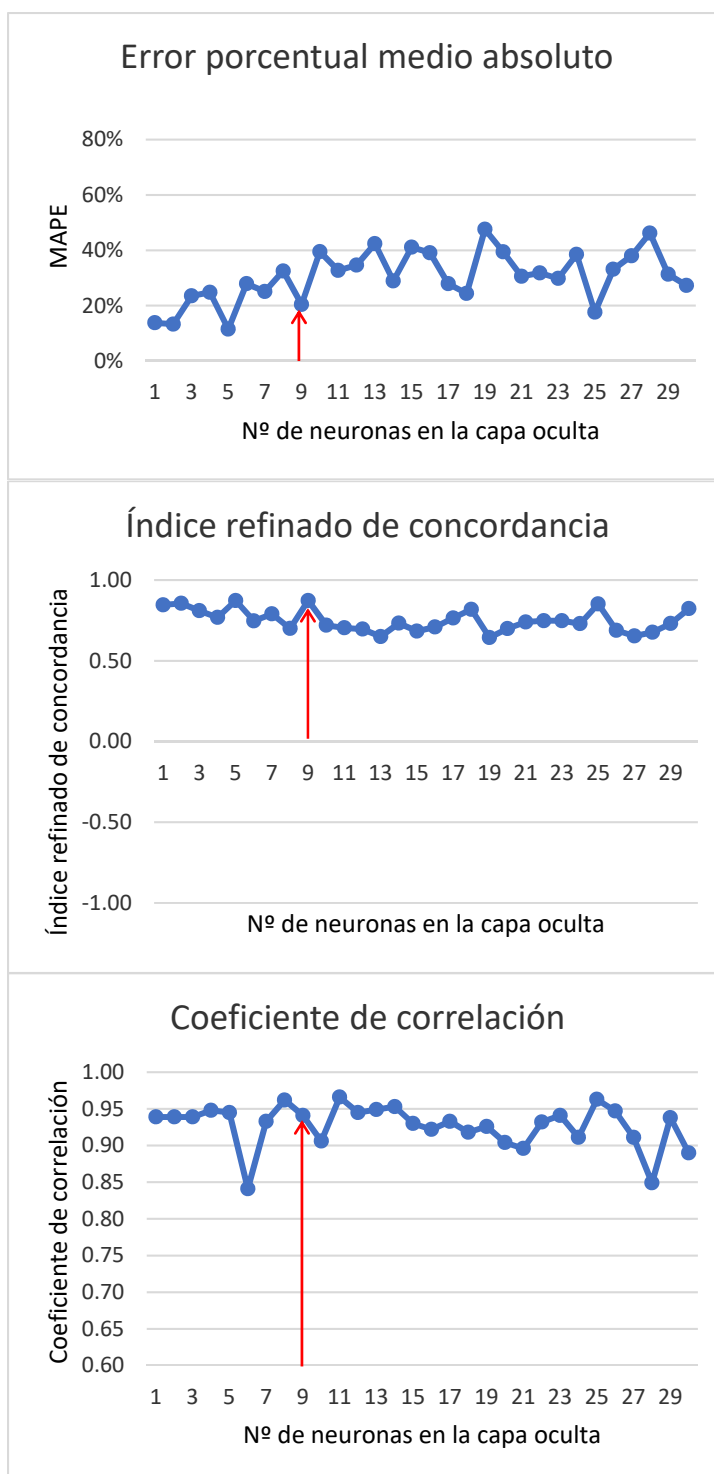
Nº de variables de entrada	Combinación de variables de entrada	Número óptimo de unidades ocultas	Error porcentual medio absoluto (MAPE)%	Índice refinado de concordancia (IOA)	Coef. de correlación "r"
6	Conc_SO2entrada(t-4), Temp_sal_Lecho1(t), Temp_sal_Lecho2(t), Temp_sal_Lecho3(t), Temp_sal_Lecho4(t), Conc_O2entrada(t-5)	5	22,3%	0,856	0,915



Nº de variables de entrada	Combinación de variables de entrada	Número óptimo de unidades ocultas	Error porcentual medio absoluto (MAPE)%	Índice refinado de concordancia (IOA)	Coef. de correlación "r"
9	Conc_SO2entrada(t-2), Conc_SO2entrada(t-3), Conc_SO2entrada(t-4), Conc_SO2entrada(t-5), Temp_sal_Lecho1(t), Temp_sal_Lecho2(t), Temp_sal_Lecho3(t), Temp_sal_Lecho4(t), Conc_O2entrada(t-5).	9	12,9%	0,877	0,950



Nº de variables de entrada	Combinación de variables de entrada	Número óptimo de unidades ocultas	Error porcentual medio absoluto (MAPE)%	Índice refinado de concordancia (IOA)	Coef. de correlación "r"
11	Conc_SO2entrada(t-2), Conc_SO2entrada(t-3), Conc_SO2entrada(t-4), Conc_SO2entrada(t-5), Temp_sal_Lecho1(t), Temp_sal_Lecho1(t-1), Temp_sal_Lecho2(t), Temp_sal_Lecho2(t-1), Temp_sal_Lecho3(t), Temp_sal_Lecho4(t), Conc_O2entrada(t-5).	9	20,4%	0,873	0,941



Nº de variables de entrada	Combinación de variables de entrada	Número óptimo de unidades ocultas	Error porcentual medio absoluto (MAPE)%	Índice refinado de concordancia (IOA)	Coef. de correlación "r"
12	Conc_SO2entrada(t-2), Conc_SO2entrada(t-3), Conc_SO2entrada(t-4), Conc_SO2entrada(t-5), Temp_sal_Lecho1(t), Temp_sal_Lecho1(t-1), Temp_sal_Lecho2(t), Temp_sal_Lecho2(t-1), Temp_sal_Lecho3(t), Temp_sal_Lecho4(t), Conc_O2entrada(t-5), Flujo_entrada(t-4)	16	15,6%	0,880	0,941

