

UNIVERSIDAD CATÓLICA DE LA SANTÍSIMA CONCEPCIÓN

FACULTAD DE INGENIERÍA
DEPARTAMENTO DE INGENIERÍA ELÉCTRICA



UCSC

Análisis de Datos operacionales de Grupos de Celdas de Electro obtención (EW) de cobre & Desarrollo de modelos de Predicción de Performance

Sebastián Ignacio Figueroa Santibáñez

Informe de Habilitación Profesional para optar al título de:
Ingeniero Civil Eléctrico

Profesor Patrocinante:
Dr. Samuel Vergara Rojas.

Comisión Evaluadora:
Dr. Muhammad Nisar.
Dr. Eduardo Espinosa.

Concepción, abril 2024

UNIVERSIDAD CATÓLICA DE LA SANTISIMA CONCEPCION
Facultad de Ingeniería
Departamento de Ingeniería Eléctrica

Profesor Patrocinante:
Dr. Samuel Vergara Rojas.

Análisis de Datos operacionales de Grupos de Celdas de Electroobtención (EW) de cobre & Desarrollo de modelos de Performance.

Sebastian Ignacio Figueroa Santibáñez

Informe de Habilitación Profesional
para optar al Título de

Ingeniero Civil Eléctrico

Abril, 2024

Resumen

La electroobtención (EW) de cobre es un proceso crítico en la industria minera que involucra la extracción de cobre a partir de soluciones de sulfato de cobre utilizando celdas electrolíticas. En este estudio, se lleva a cabo un análisis exhaustivo de los datos operacionales de grupos de celdas de electro obtención de cobre con el objetivo de comprender mejor el rendimiento del proceso e identificar áreas de mejora.

El análisis de datos operacionales se realiza utilizando técnicas avanzadas de análisis de datos, incluido el procesamiento de datos en tiempo real, la minería de datos y el análisis predictivo. Se examinan una variedad de variables operativas, como la corriente eléctrica, la temperatura, el flujo de electrolito y la concentración de cobre en la solución, para identificar patrones, tendencias y relaciones significativas.

Además del análisis descriptivo, se desarrollan modelos de rendimiento utilizando técnicas de modelado predictivo, como la regresión y el aprendizaje automático. Estos modelos permiten predecir el rendimiento del proceso de electro obtención bajo diferentes condiciones operativas y optimizar el funcionamiento de las celdas.

Los resultados de este estudio proporcionan información valiosa para la optimización de la operación de los grupos de celdas de electro obtención de cobre, Antes del análisis de datos la eficiencia de energía es de aproximado 87%, luego del análisis se puede estimar una eficiencia del 92%, lo cual dejaría una mejora del 5%. junto con lo anterior las mejorar económicas y de productividad de la planta se verían beneficiadas, pero ello escapa del estudio del presente proyecto.

A los alumnos del pasado, presente y futuro de la carrera de Ingeniería Civil Eléctrica de la U.C.S.C.

Agradecimientos

Durante este desafiante y gratificante trayecto hacia la habilitación profesional como Ingeniero Civil Eléctrico, he tenido el privilegio de conocer a una gran cantidad de personas que han dejado una marca indeleble en mi recorrido. A todas ellas, quiero expresar mi más sincero agradecimiento.

En primer lugar, a mi familia por su amor, y comprensión a lo largo de mi carrera. A mi padre que siempre me dijo que siguiera adelante ante cualquier adversidad o problema que tuviera y a mi madre por estar apoyando en esos días que uno se siente sin ganas o energías de seguir en el duro camino de estudiar ingeniería y por último a Francisca mi polola, quien me apoyo mucho en la última etapa de mi proyecto animándome día a día.

Agradezco al personal académico y administrativo de la carrera por su apoyo y dedicación, que son los principales elementos que necesitamos los alumnos. Mención para el profesor Aníbal quien me inspiro bastante en esta última etapa de mi carrera y me hizo mejorar como alumno y persona.

Y por último al Profesor Samuel Vergara quien me apoyo en mi proyecto final, justo cuando sentía que todo estaba perdido, su experiencia y dedicación me ayudaron a culminar mi proyecto de manera satisfactoria.

Agradecer al equipo de Manufactura Avanzada, Martin Castillo y Enrique Nusdel y Felipe Nova su amistad ha sido un verdadero regalo a lo largo de mi trayecto académico. Agradezco profundamente su apoyo y orientación en cada etapa de este camino, así como cada gesto de amistad, cada conversación y cada momento compartido.

Tabla de Contenidos

LISTA DE TABLAS	X
ABREVIACIONES	XI
CAPÍTULO 1. INTRODUCCIÓN	1
1.1. INTRODUCCIÓN GENERAL	1
1.2. TRABAJOS PREVIOS	2
1.3. DISCUSIÓN	5
1.4. OBJETIVOS	7
1.5. OBJETIVO GENERAL.....	7
<i>Objetivos Específicos.....</i>	<i>7</i>
1.6. ALCANCES Y LIMITACIONES	7
1.7. TEMARIO.....	8
1.8. METODOLOGÍA	8
CAPÍTULO 2. ELECTROOBTENCIÓN DE CÁTODOS DE COBRE.....	9
2.1. INTRODUCCIÓN	9
2.2. PROCESO DE ELECTROOBTENCION DE CÁTODOS DE COBRE.....	9
2.3. ELECTRODOS	10
2.4. ELECTROLITO	10
2.5. NAVE DE ELECTROOBTENCION	11
2.6. RECTIFICADOR DE ALTAS CORRIENTES.....	11
2.7. CELDA DE ELECTROOBTENCION	12
2.8. SISTEMA DISTRIBUCIÓN DE CORRIENTE.....	12
2.9. INDICADORES CLAVES DE DESEMPEÑO (KPI)	12
<i>Eficiencia de Corriente.....</i>	<i>13</i>
2.10. CONSUMO ESPECIFICO DE ENERGÍA.....	14
<i>Voltaje de Ciclo</i>	<i>15</i>
2.11. REACCIONES QUÍMICAS	16
CAPÍTULO 3. ANÁLISIS ESTADÍSTICO	17
3.1. RECOPIACIÓN Y PREPARACIÓN DE DATOS	17
3.2. VARIABLES RECOPIADAS	18
3.3. ANÁLISIS DESCRIPTIVO	19
<i>Variables de Estudio.....</i>	<i>19</i>

3.4.	CORRELACIÓN Y COVARIANZA ENTRE VARIOS.....	22
3.5.	COMPARACIÓN DE DATOS ESTADÍSTICOS	25
CAPÍTULO 4. MODELOS DE PREDICCIÓN.....		26
4.1.	REGRESIÓN LINEAL	26
	<i>Regresión Lineal Simple</i>	<i>27</i>
	<i>Regresión Lineal Múltiple</i>	<i>27</i>
4.2.	REDES ANFIS.....	28
	<i>Variables de entrada.....</i>	<i>29</i>
	<i>Selección de Datos.....</i>	<i>29</i>
	<i>Datos de Entrenamiento y Prueba</i>	<i>29</i>
	<i>Parámetros de un modelo</i>	<i>30</i>
	<i>Arquitectura.....</i>	<i>32</i>
CAPÍTULO 5. RESULTADOS		33
5.1.	INTRODUCCIÓN	33
5.2.	SELECCIÓN DE DATOS	33
	<i>Base de datos para modelos</i>	<i>34</i>
5.3.	MODELOS FINALES REGRESIÓN LINEAL MÚLTIPLE.....	35
	<i>Modelo Regresión lineal Conjunto de datos 1</i>	<i>35</i>
5.4.	MODELO RED ANFIS	35
	<i>Modelo Red ANFIS conjunto de datos 1</i>	<i>36</i>
	<i>Modelo Red ANFIS conjunto de datos 2.....</i>	<i>37</i>
5.5.	RESULTADOS MODELO REGRESIÓN LINEAL MÚLTIPLE.....	38
	<i>Error de Predicción:.....</i>	<i>38</i>
5.6.	RESULTADOS MODELO RED ANFIS	39
	<i>Predicciones Modelo Redes ANFIS:.....</i>	<i>39</i>
	<i>Errores Obtenidos del Modelo.....</i>	<i>59</i>
CAPÍTULO 6. CONCLUSIÓN		63
6.1.	SUMARIO	63
6.2.	CONCLUSIONES	63
6.3.	TRABAJOS FUTUROS	64
CAPÍTULO 7. REFERENCIA.....		65
BIBLIOGRAFÍA.....		65

Lista de Figuras

Fig. 2-3: Cosecha de cátodos de Cobre	10
Fig. 2-1: Nave de Electroobtencion Anglo American.	11
Fig. 2-2: Topologia Rectificador de altas Corrientes.....	11
Fig. 3-1: Diagrama de analisis descriptivo.....	17
Fig. 3-2: Grafico Dispersión de Corriente v/s Consumo Específico de Energía	23
Fig. 3-3: Grafico Dispersión de Corriente v/s Eficiencia de Corriente	23
Fig. 3-4: Grafico Dispersión de Corriente v/s Voltaje Celda Ciclo.	24
Fig. 4-1: Arquitectura ANFIS para inferencia Takagi-Sugeno de Primer Orden [7].....	32
Fig. 5-1: Arquitectura Red ANFIS para 3 variables y 3 funciones de membresia.....	36
Fig. 5-2: Arquitectura Red ANFIS para 4 variables y 3 funciones de membresia.....	37
Fig. 5-3:Grafico prediccion v/s valores reales (6 funciones de membresia) PA4-Pre1	39
Fig. 5-4: Grafico prediccion v/s valores reales (6 funciones de membresia) PA4-Pre2	39
Fig. 5-5: Grafico prediccion v/s valores reales (6 funciones de membresia) PA4-Pre3	40
Fig. 5-6: Grafico prediccion v/s valores reales (6 funciones de membresia) PA4-Pre4	40
Fig. 5-7: Grafico prediccion v/s valores reales (5 funciones de membresia) PA4-Pre1	41
Fig. 5-8: Grafico prediccion v/s valores reales (5 funciones de membresia) PA4-Pre2	41
Fig. 5-9: Grafico prediccion v/s valores reales (5 funciones de membresia) PA4-Pre3	42
Fig. 5-10: Grafico prediccion v/s valores reales (5 funciones de membresia) PA4-Pre4	42
Fig. 5-11: Grafico prediccion v/s valores reales (4 funciones de membresia) PA4-Pre1	43
Fig. 5-12: Grafico prediccion v/s valores reales (4 funciones de membresia) PA4-Pre2	43
Fig. 5-13: Grafico prediccion v/s valores reales (4 funciones de membresia) PA4-Pre3	44
Fig. 5-14: Grafico prediccion v/s valores reales (4 funciones de membresia) PA4-Pre4	44
Fig. 5-15: Grafico prediccion v/s valores reales (3 funciones de membresia) PA4-Pre1	45
Fig. 5-16: Grafico prediccion v/s valores reales (3 funciones de membresia) PA4-Pre2	45
Fig. 5-17: Grafico prediccion v/s valores reales (3 funciones de membresia) PA4-Pre3	46
Fig. 5-18: Grafico prediccion v/s valores reales (3 funciones de membresia) PA4-Pre4	46
Fig. 5-19: Grafico prediccion v/s valores reales (2 funciones de membresia) PA4-Pre1	47
Fig. 5-20: Grafico prediccion v/s valores reales (2 funciones de membresia) PA4-Pre2	47
Fig. 5-21: Grafico prediccion v/s valores reales (2 funciones de membresia) PA4-Pre3	48
Fig. 5-22: Grafico prediccion v/s valores reales (2 funciones de membresia) PA4-Pre4	48
Fig. 5-23: Grafico prediccion v/s valores reales (6 funciones de membresia) PA3-Pre1	49
Fig. 5-24: Grafico prediccion v/s valores reales (6 funciones de membresia) PA3-Pre2	49

Fig. 5-25: Grafico prediccion v/s valores reales (6 funciones de membresia) PA3-Pre3	50
Fig. 5-26: Grafico prediccion v/s valores reales (6 funciones de membresia) PA3-Pre4	50
Fig. 5-27: Grafico prediccion v/s valores reales (5 funciones de membresia) PA3-Pre1	51
Fig. 5-28: Grafico prediccion v/s valores reales (5 funciones de membresia) PA3-Pre2	51
Fig. 5-29: Grafico prediccion v/s valores reales (5 funciones de membresia) PA3-Pre3	52
Fig. 5-30: Grafico prediccion v/s valores reales (5 funciones de membresia) PA3-Pre4	52
Fig. 5-31: Grafico prediccion v/s valores reales (4 funciones de membresia) PA3-Pre1	53
Fig. 5-32: Grafico prediccion v/s valores reales (4 funciones de membresia) PA3-Pre2	53
Fig. 5-33: Grafico prediccion v/s valores reales (4 funciones de membresia) PA3-Pre3	54
Fig. 5-34: Grafico prediccion v/s valores reales (4 funciones de membresia) PA3-Pre4	54
Fig. 5-35: Grafico prediccion v/s valores reales (3 funciones de membresia) PA3-Pre1	55
Fig. 5-36: Grafico prediccion v/s valores reales (3 funciones de membresia) PA3-Pre2	55
Fig. 5-37: Grafico prediccion v/s valores reales (3 funciones de membresia) PA3-Pre3	56
Fig. 5-38: Grafico prediccion v/s valores reales (3 funciones de membresia) PA3-Pre4	56
Fig. 5-39: Grafico prediccion v/s valores reales (2 funciones de membresia) PA3-Pre1	57
Fig. 5-40: Grafico prediccion v/s valores reales (2 funciones de membresia) PA3-Pre2	57
Fig. 5-41: Grafico prediccion v/s valores reales (2 funciones de membresia) PA3-Pre3	58
Fig. 5-42: Grafico prediccion v/s valores reales (2 funciones de membresia) PA3-Pre4	58

Lista de Tablas

Tabla 1: Estadística de la Densidad de Corriente.20
 Tabla 2: Estadística de la Eficiencia de Corriente21
 Tabla 3: Estadística del Consumo específico.21
 Tabla 4: Resumen estadístico Tecnologías.....22
 Tabla 5: Error de predicción Conjunto de datos 138
 Tabla 6: Error de predicción Conjunto de datos 238
 Tabla 7: Comparación error de Entrenamiento y Prueba Red ANFIS para Conjunto 1.....59
 Tabla 8: Comparación error de Entrenamiento y Prueba Red ANFIS para Conjunto 2.....60
 Tabla 9: Comparación error promedio de Entrenamiento61
 Tabla 10: Comparación error promedio de Prueba.....61

Abreviaciones

CC	: Corto Circuito Metalúrgico.
EW	: Electrowining.
ICBB	: Barra Intercelda.
Vcell	: Voltaje de Celda.
Isd	: Dispersión de Corriente.
ER	: Reacción de descomposición.
EC	: Sobrepotencial catódico.
EA	: Sobrepotencial Anódico.
E ohm	: Resistencia óhmica del electrolito.
EK	: Resistencia del hardware.
FM	: Funciones de Membresía.
Ep	: Épocas de Entrenamiento.
RMSE	: Raiz de error cuadrado Medio

Capítulo 1. Introducción

1.1. Introducción General

La Electroobtención de cátodos de cobre es un proceso de mucha importancia en la Industria minera, ya que permite obtener cobre con un 99,99% de pureza, el cual es recuperado de las soluciones de cobre obtenidas en el proceso previo de lixiviación.[1]

Las celdas de electroobtención de cobre sufren constantemente problemas de corto circuito y circuito abierto, pero estos mismos problemas son difíciles de detectar debido a que no existen sensores que puedan monitorear el flujo de la corriente en las celdas, ya que en la dimensión y la solución que se encuentra sumergidos los cátodos, no es posible sumergir sensores.[1]

El objetivo de este proyecto es crear modelos que permitan prever cómo variará el rendimiento del proceso en función de las condiciones operativas específicas. El proceso de desarrollo de los modelos de predicción incluirá la preparación y limpieza de los datos, la definición de las variables de entrada y salida, el diseño y entrenamiento de las redes ANFIS, y la evaluación de la precisión de los modelos generados. Al implementar estos modelos en operaciones reales de electroobtención, se espera lograr una mayor eficiencia en el proceso, reduciendo costos y optimizando los recursos. [5]

La industria minera y metalúrgica enfrenta constantemente el desafío de optimizar sus procesos para mejorar la eficiencia y la calidad del producto final. En el caso específico de la electroobtención de cobre, uno de los métodos más utilizados para la extracción de este metal, se busca continuamente innovar para reducir costos, mejorar la pureza del cobre y minimizar el impacto ambiental.

Este proyecto busca contribuir al avance de la tecnología de electroobtención de cobre al combinar análisis de datos operacionales y técnicas de inteligencia artificial. Los resultados obtenidos ayudarán a obtener mejoras en los KPI de la industria de electroobtención de cobre.

1.2. Trabajos Previos

- ♣ M. Tucker, C. Dorfling , M. Tadie. (2021). *Investigating an approach to parameter fitting for the development of a semi-empirical electrowinning model*. Stellenbosch University.

Este documento Explora diferentes enfoques para ajustar los parámetros del modelo de electrowinning de Tucker, el cual explica que ajustes realizados por el autor te sirven en caso de estar obteniendo un modelo para una mejor obtención de modelo, principalmente ajusta curvas de sobrepotencial catódico y anódico con el fin de tener un mejor rendimiento en el Voltaje de Celda de una piscina de EW de cobre.

- ♣ Mark E. Sclesinger Matthew J.King Kathryn C.Sole William G. Davenport. (2011). *Extractive Metallurgy of Copper 5ta Edition*. ELSERVIER.

Este informe detalla el proceso de fabricación de cobre. Se analizan minuciosamente los principales enfoques de extracción, que abarcan la pirometalurgia, hidrometalurgia y electrometalurgia, con investigaciones específicas para cada método. Respecto a la electroobtención, se expone en detalle los requisitos eléctricos esenciales para ejecutar el proceso electroquímico entre el fluido y los electrodos, junto con las especificaciones de materiales y dimensiones necesarias para la construcción de la celda electroquímica, anodos y cátodos.

- ♣ Cuba, C. B. (2016). *Electroobtencion de Cobre Empleando el Metodo Electrodialitico Reactivo*. Huancayo - Perú.

El documento presenta información acerca de la termodinámica y cinética electroquímica en el proceso de electroobtencion de cobre crucial para la deposición de cobre en el cátodo, realizando análisis para las reacciones catódica, anódica e indeseadas en el proceso. Además, de detallar la influencia de la densidad de corriente en la eficiencia, consumo de energía y producción.

- ♣ Eduardo Wiechmann F.(2010) Evaluación Industrial Tecnología Optibar (IF)

En este documento, se examina detalladamente el desempeño y la viabilidad de la tecnología Optibar en aplicaciones industriales específicas. Se presentan resultados de pruebas realizadas en entornos industriales reales, con el objetivo de demostrar la eficacia y utilidad de esta tecnología en la industria. Además, se proporcionan recomendaciones prácticas para la implementación efectiva de la tecnología Optibar en diferentes contextos industriales."

- ♣ Ricardo Rojas Boitano "Curso de Capacitacion Operacon de Plantas de Electroobtencion" Universidad de Santiago de Chile, 2006.

En este informe se explican los fundamentos básicos del proceso de electrobtención de cátodos de cobre, realizando análisis de la celda de electrolisis, reacciones químicas producidas en los cátodos y ánodos como su diseño en material y dimensiones. Además, presenta como afecta la densidad de corriente a los diferentes indicadores claves de desempeño, entregando análisis sobre cada KPI y como este indicador se ve afectado frente al fenómeno de neblina acida.

- ♣ Jang, J. S. (1993) "ANFIS Adaptive Network based Fuzzy Inference Systems". IEEE Transactions on Systems, Man & Cybernetics Vol. 23, pp 665-685

En este estudio, Jang explora cómo estas redes pueden adaptarse de manera efectiva a diferentes entornos y datos, mejorando así su capacidad para realizar inferencias precisas en situaciones complejas. El artículo probablemente proporciona detalles técnicos sobre la arquitectura y el funcionamiento de estos sistemas, así como ejemplos de su aplicación en diversos campos, como el control de sistemas, la inteligencia artificial y la toma de decisiones.

- ♣ Mora Flórez, J. J., Pérez Hernández, L. P., & Pérez Londoño, S. M. (2006). Utilización de redes ANFIS y señales de corriente para localización de la zona de falla en sistemas de distribución de energía eléctrica. Ingeniería e Investigación, 26(3), 93-99.

El documento detalla como se implementa esta Tecnología incluyendo la recopilación de datos y procesamiento de datos de corriente la configuración y entrenamiento de las redes ANFIS, y la interpretación de los resultados para determinar la ubicación de las fallas. Los datos Utilizados fueron reales.

- ♣ Correa-Henao, G. J., & Montoya-Suárez, L. M. (2013). APLICACIÓN DEL MODELO ANFIS PARA PREDICCIÓN DE SERIES DE TIEMPO. Lámpsakos, (9),12-25.

El artículo presenta la descripción y posterior aplicación de una metodología de Redes Neuro-Difusas aplicadas al problema de predicción de series de tiempo en el mercado de capitales de corto plazo. Estas aplicaciones proporcionan criterios de referencia para inversión especulativa en la bolsa de valores de Colombia, en la medida que complementan la realización de análisis técnicos y fundamentales.

- ♣ Bastos Guerrero, D, Illera, M y Sepúlveda, S. (2020). Sistema adaptativo de inferencia neuro-difusa (ANFIS) para la estimación de la radiación solar global. Investigación e Innovación en Ingenierías.

A partir de registros históricos del Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM) se obtuvieron las variables: temperatura ambiente, brillo solar, velocidad del viento, humedad relativa, hora de medición y radiación solar global entre los años 2005 y 2015. Usando la herramienta Fuzzy Logic Toolbox de Matlab y los datos del IDEAM se crearon redes ANFIS con diferentes configuraciones.

- ♣ Montero Granados. R (2016): Modelos de regresión lineal múltiple. Documentos de Trabajo en Economía Aplicada. Universidad de Granada. España.

La regresión lineal múltiple se enfoca en ajustar modelos lineales o que pueden ser linealizados entre una variable dependiente y múltiples variables independientes. En estos modelos, es crucial evaluar la heterocedasticidad, la multicolinealidad y la especificación. Este curso tiene como objetivo introducirnos en el ámbito de la modelización, incluyendo la creación de variables dummy, la configuración de un individuo de referencia, la utilización de factores de ponderación, y el manejo de variables de interacción e interrelación, entre otros aspectos.

1.3. Discusión

El trabajo de M. Tucker, C. Dorfling, M. Tadie (2021) explora diversos métodos para ajustar los parámetros del modelo de electrowinning, centrándose en la optimización de las curvas de sobrepotencial catódico y anódico. Estos ajustes son fundamentales para mejorar el rendimiento del voltaje de la celda en una piscina de electroobtención de cobre. Este enfoque se complementa con el análisis de Cuba, C. B. (2016), que examina la termodinámica y cinética electroquímica en la electroobtención de cobre, con un énfasis en cómo la densidad de corriente afecta la eficiencia, el consumo de energía y la producción. Además, Ricardo Rojas Boitano (2006) ofrece una visión integral de los fundamentos del proceso de electroobtención de cátodos de cobre, incluyendo el diseño de las celdas de electrólisis y el impacto de la densidad de corriente en los indicadores clave de desempeño, abordando también el fenómeno de la neblina ácida.

Eduardo Wiechmann F. (2010) evalúa la viabilidad y el desempeño de la tecnología Optibar en aplicaciones industriales, presentando resultados de pruebas en entornos reales. Las recomendaciones prácticas para la implementación de esta tecnología pueden ser aplicadas para mejorar la eficiencia en procesos industriales, incluyendo la electroobtención de cobre. La integración de estas recomendaciones con las metodologías de optimización descritas por Tucker, Cuba y Rojas Boitano puede resultar en un mejor rendimiento global de los procesos industriales.

El trabajo de Jang, J. S. (1993) sobre sistemas adaptativos de inferencia neuro-difusa (ANFIS) proporciona una base teórica para la adaptación de redes a diversos entornos y datos. Esta teoría se aplica en diversos contextos industriales y de investigación. Por ejemplo, Mora Flórez, J. J., Pérez Hernández, L. P., & Pérez Londoño, S. M. (2006) implementan redes ANFIS para la localización de fallas en sistemas de distribución eléctrica, mientras que Correa-Henao, G. J., & Montoya-Suárez, L. M. (2013) utilizan estas redes para la predicción de series de tiempo en el mercado de capitales. Asimismo, Bastos Guerrero, D., Illera, M., & Sepúlveda, S. (2020) aplican redes ANFIS para estimar la radiación solar global, demostrando la versatilidad y precisión de estas redes en aplicaciones de energías renovables y estudios meteorológicos.

Montero Granados, R. (2016) proporciona un marco teórico esencial sobre la regresión lineal múltiple, abordando la evaluación de la heterocedasticidad, la multicolinealidad y la especificación. Estos principios son fundamentales para entender y aplicar técnicas de modelización en diversos contextos industriales y de investigación. La comprensión y aplicación de la regresión lineal múltiple facilita el desarrollo de modelos más complejos, que pueden ser utilizados para optimizar procesos industriales como los descritos en los estudios de Tucker, Cuba y Rojas Boitano.

La interrelación de estos documentos se manifiesta en la capacidad de utilizar técnicas avanzadas de modelización y optimización para mejorar procesos industriales. La optimización de la electroobtención de cobre y la implementación de tecnologías como Optibar pueden beneficiarse de las metodologías de redes ANFIS y regresión lineal múltiple. La integración de datos y resultados de pruebas con modelos matemáticos avanzados proporciona un enfoque integral para la mejora continua y la toma de decisiones basada en datos, optimizando la eficiencia y el rendimiento en diversas aplicaciones industriales.

1.4. Objetivos

1.5. Objetivo General

Desarrollar modelos predictivos de rendimiento para grupos de celdas de electroobtención de cobre utilizando redes ANFIS, a partir del análisis de datos operacionales, con el propósito de optimizar la eficiencia y el control del proceso de obtención de cobre.

Objetivos Específicos

- Mapeo del proceso para identificar variables, parámetros e indicadores clave de desempeño.
- Construcción de base de datos de mediciones de celdas EW industriales.
- Integración, limpieza, etiquetado de datos.
- Análisis multivariable, clusterización de datos, análisis de correlaciones.
- Construcción de modelo basado en datos de línea base de operación de celdas.
- Construcción de modelo de predicción de rendimiento de celda usando redes ANFIS.
- Comparación de rendimiento de modelos.

1.6. Alcances y Limitaciones

- Proyecto considera trabajo con simulaciones en ambiente MATLAB.
- Proyecto no considera pruebas experimentales en laboratorio o terreno.
- Proyecto utilizará como input distintos sets de mediciones de variables y parámetros eléctricos,
- químicos, físicos y de productividad realizadas en terreno en Planta de EW de cobre – AMSA Minera Zaldívar (Antofagasta).

1.7. Temario

En el capítulo 2, se presenta el marco teórico de comentando los principales conceptos detrás del proceso de electroobtención de cátodos de cobre. El cual menciona equipos utilizados, variables presentes en el proceso y junto con ello KPI del proceso.

En el capítulo 3, se presenta un análisis estadístico el cual estudia los KPI del proceso de electroobtención de cobre; este análisis estadístico estudia dos tecnologías y su comportamiento Walker y Optibar, analizando comportamiento de la densidad de corriente, eficiencia de corriente y consumo específico. Para luego correlacionar variables y determinar cuál es la que más predomina en el Proceso de EW de cobre.

En el capítulo 4, se hace una breve introducción de los dos métodos predictivos que se van a utilizar para la predicción de la variable obtenida en el Capítulo 3, este capítulo más que nada es teórico, solo explicación de sus configuraciones y/o modelo matemático.

El capítulo 5, corresponde a los Resultados en este capítulo se muestran las configuraciones de los modelos de predicción aplicados con el fin de que el lector pueda entender y comprender el proceso que se está aplicando, además se agregan los errores y resultados obtenidos durante la predicción de estos, con el fin de poder analizar de mejor manera los resultados.

1.8. Metodología

Primero se realizó una revisión sobre el comportamiento estadístico de las variables predominantes en la industria de EW de cobre; Eficiencia de Corriente, Consumo Específico, Dispersión de corriente y Voltaje de Celda. Una vez realizado ese paso se determinará la variable que más afecta al sistema. Como segunda parte se crearon modelos predictivos mediante regresión Lineal y red anfis, con el que se predijo el comportamiento de la dispersión de corriente.

Para finalmente realizar una comparación de ambos modelos y ver que modelo presenta menor error al momento de predecir datos.

Capítulo 2. Electroobtención de cátodos de Cobre

2.1. Introducción

La técnica de electroobtención de cátodos de cobre, que se utiliza extensamente en la industria minera, permite la obtención de cobre de alta pureza a partir de soluciones de sulfato de cobre. Este proceso es el tema central de este capítulo.

Este capítulo se sumerge en el estudio detallado de este proceso electroquímico, que tiene un papel determinante en la producción de cobre de alta calidad. A lo largo del capítulo, se examinarán los elementos esenciales de la electroobtención, desde los fundamentos de la electrodeposición hasta los factores clave que determinan la eficiencia y la calidad del producto final.

2.2. Proceso de Electroobtención de cátodos de cobre

El proceso de electroobtención de cobre implica la inmersión de un ánodo y un cátodo en una solución de sulfato de cobre, donde al aplicar una corriente eléctrica, los iones de cobre se depositan en el cátodo como cobre metálico, mientras que en el ánodo ocurre la oxidación del cobre metálico, liberando iones de cobre en la solución.

El cobre electrolítico obtenido en el cátodo se recolecta y purifica, mientras que la solución de sulfato de cobre restante se regenera para su reutilización, formando así un ciclo continuo que permite la extracción eficiente y sostenible de cobre de soluciones acuosas.

2.3. Electroodos

Los electrodos que se utilizan en las celdas de electroobtención están compuestos por ánodos y cátodos, cada uno desempeñando un papel crucial en el proceso de extracción de metales. Los cátodos, en particular, son esenciales ya que son las superficies sobre las cuales se depositan los metales deseados, como el cobre, durante el proceso de electroobtención.



Fig. 2-1: Cosecha de cátodos de Cobre

2.4. Electrolito

El electrolito es una solución crucial en el proceso de electroobtención. Esta solución líquida, compuesta principalmente de sulfato de cobre ($CuSO_4$) disuelto en agua (H_2O) contiene iones de cobre (Cu^{2+}) en estado disuelto. Además del sulfato de cobre, el electrolito puede incluir otros compuestos químicos con el objetivo de optimizar la eficiencia y calidad del proceso. Por ejemplo, se puede añadir ácido sulfúrico para mantener el pH en un nivel óptimo.

2.5. Nave de Electroobtencion

Son complejos industriales en los que se efectúa el procedimiento electroquímico para la deposición de cobre en los cátodos. Dichos complejos están formados por un conjunto de celdas electrolíticas interrelacionadas donde se desarrolla el proceso de electroobtención. Estas celdas están compuestas por ánodos y cátodos que se sumergen en una solución de sulfato de cobre.



Fig. 2-2: Nave de Electroobtencion Anglo American.

2.6. Rectificador de Altas corrientes

Los Rectificadores de Altas corrientes son dispositivos que convierten la corriente Alterna (AC) en corriente continua (DC). En el proceso de electroobtención de cobre, estos Rectificadores de 12 Pulsos en base a tiristores, son esenciales ya que proporcionan la corriente continua necesaria, permiten un control preciso del proceso, mejoran la eficiencia energética, garantizan la fiabilidad y durabilidad del sistema en su conjunto.

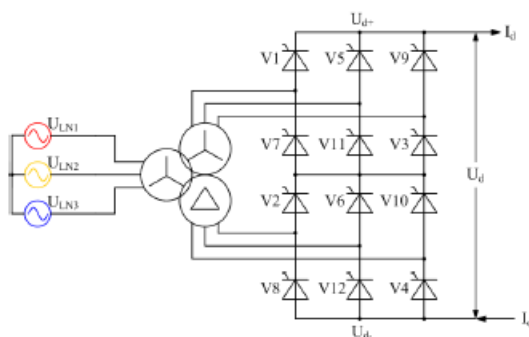


Fig. 2-3: Topologia Rectificador de altas Corrientes

2.7. Celda de Electroobtencion

Las celdas de electroobtención son estructuras geométricas diseñadas específicamente para el proceso de obtención de metales mediante la aplicación de corriente eléctrica. En estas celdas, los ánodos y cátodos (65) se sumergen alternadamente en una solución electrolítica. Esta disposición facilita la transferencia de iones metálicos desde el ánodo (donde se disuelven) hacia el cátodo (donde se depositan), permitiendo así la recuperación del metal deseado con alta pureza.

2.8. Sistema Distribución de Corriente

El sistema de distribución de corriente, también conocido como barras interceldas, facilita el flujo de corriente de una celda a la siguiente, asegurando una distribución uniforme de corriente en el cátodo. Esta uniformidad es fundamental para obtener depósitos de cobre de alta calidad y homogéneos.

En la industria actual, existen dos tecnologías principales: Walker y Optibar. Estas tecnologías difieren en términos de eficiencia, consumo específico de energía y voltaje por ciclo de celda.

La densidad de corriente controla 3 importantes variables:

- a. Eficiencia de corriente
- b. Consumo Especifico de Energia
- c. Voltaje Celda Ciclo

♣ Estas Variables son indicadores para medir el desempeño de celdas de electroobtención.

2.9. Indicadores Claves de desempeño (KPI)

Los KPI son esenciales para evaluar el rendimiento de cualquier proceso, incluyendo el proceso de Electroobtención de Cobre. Los Indicadores como la eficiencia de corriente, el consumo específico de energía y el voltaje de celda ciclo son particularmente importantes en este contexto. Estos indicadores permiten una evaluación precisa y basada en datos del rendimiento del proceso.

Al monitorear estos KPI, se pueden identificar rápidamente las áreas de mejora y tomar medidas correctivas de manera oportuna, lo que a su vez puede llevar a una mayor eficiencia operativa y a una mejor calidad del producto final.

Eficiencia de Corriente

La eficiencia de corriente (η) representa la eficiencia de la reacción de electrolisis en las celdas de Electroobtención (ecuación 8). Normalmente corresponde a un valor entre 91% y 93%. Esto significa que una fracción de la corriente de las celdas no contribuye a depositar cobre.

La corriente de proceso no utilizada en depositar cobre se desperdicia como calor en cortocircuitos entre electrodos, como corrientes parásitas a tierra o en reacciones parásitas en las celdas (por ejemplo, depositar hierro en cátodos).

La densidad de corriente de los cátodos debe mantenerse lo más cerca de la referencia de proceso para optimizar la eficiencia de corriente. Esto se consigue disminuyendo la dispersión de densidad de corriente entre los cátodos.

$$\eta = \frac{\text{ciclo de Produccion de Cobre}}{\text{Producción Teorica de Cobre}} = \frac{x}{K * I_{dc} * t_{ciclo}} (\%) \leq 100\% \quad (1)$$

Con:

$$K: \text{constante de Eficiencia} \left[\frac{\text{Ton}}{\text{kAh}} \right]$$

$$I_{dc} = \text{Corriente aplicada en la celda} [A]$$

$$t_{ciclo} = \text{tiempo de la corriente de ciclo} [s]$$

- Se Considera x como el cobre depositado efectivamente en los cátodos en el ciclo producción de duración t_{Ciclo} , para una corriente continua promedio I_{dc} aplicada a las celdas durante el ciclo.

2.10. Consumo Especifico de Energía

El índice clave de desempeño estándar utilizado para evaluar el resultado global del proceso de Electroobtención de cobre en términos de producción y consumo de energía es el consumo específico de energía. Este índice está definido como la energía eléctrica requerida para producir 1 tonelada de cátodos de cobre.

Producción y energía se complementan con evaluación de calidad física y química. El consumo específico de energía (w) de celdas de electroobtención de cobre puede ser expresada en términos del voltaje promedio de celda (v_{cell}) y la eficiencia de corriente del proceso (η) de la siguiente forma ($k = 1,18576 * 10^{-3} \left[\frac{Ton}{kAh} \right]$):

- Energía específica [KWh/Ton]

$$W = \frac{\text{consumo de energia de Ciclo}}{\text{Ciclo de Produccion de Cobre}} = \frac{V_{celda}}{K * \eta} \quad (2)$$

$$W = \text{consumo especifico de Energia} \left[\frac{KWh}{Ton} \right]$$

$$V_{cell} = \text{Voltaje promedio celda [V]}$$

$$\eta = \text{Eficiencia de corriente [\%]}$$

Voltaje de Ciclo

El voltaje de Celda indica la cantidad de energía que necesita el sistema para poder depositar la masa de cobre en los cátodos en el ciclo de electroobtención. La idea es que este valor sea el menor posible ya que implica directamente en la eficiencia del proceso.

La idea es que el sistema siempre consuma la menor energía posible, debido a que el mismo podría generar menor costo de Producción en la generación de cátodos de cobre.

El Voltaje de Celda depende de los potenciales de reacción EW, las sobretensiones de hardware y otros factores los cuales se mencionarán a continuación:

- Reacción de descomposición (ER)
- Sobrepotencial Catódico (EC)
- Sobrepotencial Anódico (EA)
- Resistencia óhmica del electrolito (E_{ohm})
- Resistencia del hardware (E_K)

Por ende, el potencial total se expresa como:

$$E_{total} = ER + E_a + E_c + E_{ohm} + E_k \quad (3)$$

El valor de la E_{total} siempre tiende a oscilar entre los valores de 2 a 2,5 $\left[\frac{V}{celda}\right]$, como se había mencionado anteriormente mientras mas bajo el valor de la Energía total, mejor son los KPI en General debido a que el consumo sufre una disminución y junto con ello la eficiencia energética.

2.11. Reacciones Químicas

En el Proceso de Electroobtención del cobre se induce una corriente continua promedio de $326 \left[\frac{A}{m^2} \right]$, lo cual al entrar en contacto con la solución electrolítica genera reacciones químicas en el cátodo y ánodo sumergidos en el electrolito.

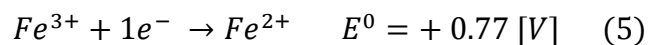
A continuación, se muestran las reacciones generadas en los Electroodos:

- El cobre rico del electrolito se galvaniza en las superficies de los cátodos.

Reducción de Cobre:

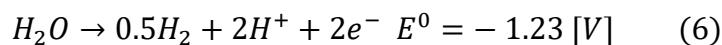


Reacción parasitaria:

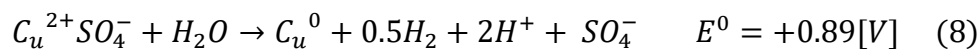
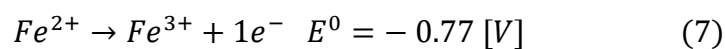


- Generación de Oxígeno y ácido sulfurico (O_2 y H_2SO_4)

Descomposición del Agua:



Reacción Parasitaria:



Capítulo 3. Análisis Estadístico

En este capítulo se analizarán datos operacionales del proceso de EW de cobre los cuales servirán como estudio previo, para la determinación de una variable predominante en el consumo energético de las Celdas de EW de cobre.

Desde la monitorización de variables clave hasta la identificación de tendencias y la optimización de parámetros, el análisis estadístico proporciona una herramienta invaluable para comprender y mejorar continuamente este proceso industrial.

A través de técnicas avanzadas de análisis de datos, este capítulo abordará cómo el análisis estadístico puede ayudar a identificar patrones, predecir comportamientos y tomar decisiones informadas para aumentar la eficiencia y la rentabilidad en la electroobtención del cobre.

3.1. Recopilación y preparación de datos

En el Proyecto se utilizaron como input distintos sets de mediciones de Variables y Parámetros Eléctricos, Químicos, Físicos y de Productividad realizadas en Terreno en Planta de EW de Cobre-AMSA Minera Zaldívar (Antofagasta). separan en tecnologías Walker y optibar, ambas con 65 mediciones correspondiente a la cantidad de cátodos de las celdas electroquímicas.

A continuación, se muestra el proceso de recopilación de datos que se utilizó el siguiente Diagrama el cual indica el proceso de elaboración propia; por el cual paso la base de datos entregada por Minera Zaldívar

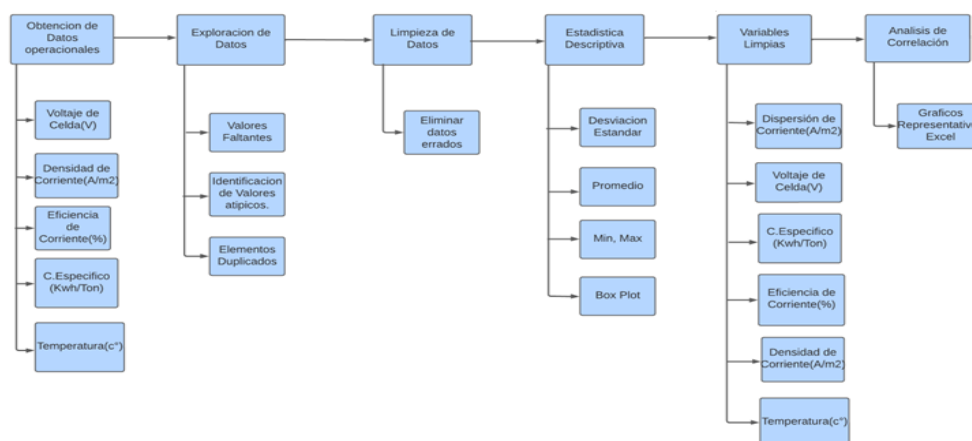


Fig. 3-1: Diagrama de analisis descriptivo

Como primer punto se tiene la obtención de datos Operacionales, en el cual se seleccionan las variables de interés para nuestro modelo, en este caso se tiene el Voltaje de Celda, Densidad de Corriente, Eficiencia de corriente, consumo específico y Temperatura. Para luego explorar los datos en este caso se hizo una verificación de datos así identificar valores faltantes, Valores atípicos; los cuales no permiten un buen análisis de datos y junto con ello Identificar los Valores Faltantes en ciertos días o Variables. Una vez identificados se procede a hacer la limpieza de datos solo dejando lo más importante para el siguiente paso.

Estadística descriptiva, en este paso se realizaron análisis de desviación estándar, máximos, mínimos y promedios. Para así analizar el comportamiento de variables en ciertos tiempos de operación, junto con ello se aplicaron Box plot para filtrar ciertos datos que a simple vista no son los ideales para la operación de estos.

3.2. Variables recopiladas

En el proceso de recopilación de datos se utilizó Excel ya que es el software que es más amigable para la filtración de datos que no sean aporte para el análisis de variables dependientes del sistema, en el cual se seleccionaron como se mencionó en el punto 3.1, dos grupos diferenciados por su Tecnología, los cuales tienen las mismas variables pero se diferencian en sus valores y junto con ello habían días en los cuales no habían registros; por lo tanto había que filtrar ciertos días para que los datos pudieran ser consistentes.

Variables:

- Voltaje de Celda (V)
- Dispersión de Corriente(A/m^2)
- Consumo Especifico(Kwh/Ton)
- Eficiencia de Corriente(%)
- Densidad de Corriente(A/m^2)

Las Variables seleccionadas son necesarias para el desarrollo de los modelos esto debido a que son las únicas que nos permiten avanzar en el estudio debido a su relación matemática con los modelos.

3.3. Análisis Descriptivo

El análisis descriptivo de datos operacionales es una parte fundamental del análisis de datos en una organización o empresa. Este tipo de análisis se enfoca en describir y resumir los datos relacionados con las operaciones diarias o procesos de la empresa. Su objetivo es proporcionar una visión clara y comprensible de los datos para ayudar en la toma de decisiones y la identificación de patrones o tendencias relevantes.

En el punto 3.2 se mencionan las variables a analizar, en este punto analizaremos las variables de mayor interés como es la Densidad de corriente, Consumo Especifico, Eficiencia de Corriente, Dispersión de corriente y Voltaje de Celda en el ciclo.

A continuación, se mostrarán Tablas la cuales entregarán el comportamiento de ciertas variables:

Variables de Estudio

Densidad de Corriente:

La densidad de corriente es una medida que se utiliza en Proceso de electro obtención para expresar la cantidad de corriente eléctrica aplicada por unidad de área en el cátodo. Se mide en amperios por metro cuadrado (A/m^2). La densidad de corriente es un parámetro crítico, ya que influye directamente en la velocidad de deposición del cobre en el cátodo. Un valor alto de densidad de corriente acelerará la obtención de cobre, pero también puede llevar a problemas como la formación de dendritas o un depósito de cobre no uniforme. Por otro lado, una densidad de corriente baja puede ralentizar el proceso de obtención de Cobre.

El comportamiento de las corrientes es bastante similar, en cambio sí nos fijamos en su dispersión en Walker tiene una diferencia muy grande a comparación de Optibar, de la cual podemos concluir muchas Teorías, pero solo quedarían en ello, ya que no podemos aseverar nada por ende se pasará a hacer un análisis de las siguientes variables, para tener un mayor contexto de las variables.

Tabla 1: Estadística de la Densidad de Corriente.

Tecnología de Barra Intercelda	Densidad de Corriente Mínima[A/ m²]	Densidad de Corrientes Promedio[A/ m²]	Densidad de Corrientes Máxima[A/ m²]	Dispersión de Corrientes (%)
Walker	299,1	326,7	354,6	14,3
Optibar	299,4	329,9	361,8	9,7

Eficiencia de Corriente:

La eficiencia de corriente es una medida que indica qué proporción de la corriente eléctrica aplicada se utiliza efectivamente para la obtención de cobre. Se expresa como un porcentaje y se calcula dividiendo la cantidad de cobre depositado en el cátodo por la cantidad teórica de cobre que debería haberse obtenido con la corriente eléctrica aplicada.

Esta se ve afectada por diversos factores, como la calidad de la solución electrolítica, la concentración de iones de cobre y la eficiencia de los cátodos. Una alta eficiencia de corriente es deseable para reducir el consumo de energía y obtener un producto de alta calidad.

En los resultados obtenidos del estudio se puede observar que la tecnología Optibar tiene la eficiencia de corriente levemente mayor lo cual implica un mejor producto final, con menos consumo de Energía. Esto se puede a debe directamente a la tecnología aplicada en el sistema

Tabla 2: Estadística de la Eficiencia de Corriente

Tecnología de Barra Intercelda	Eficiencia de Corrientes Mínima [%]	Eficiencia de Corrientes Promedio [%]	Eficiencia de Corrientes máxima [%]
Walker	91,0%	92,3%	93,3%
Optibar	91,4%	92,6%	93,6%

Consumo Especifico:

Es un indicador importante de la eficiencia energética del proceso de electro obtención. Un consumo específico bajo significa que el proceso está utilizando la energía de manera más eficiente para obtener el cobre, lo que se traduce en menores costos operativos y una menor huella ambiental. La reducción del consumo específico es un objetivo clave en la optimización de obtención de cátodos de cobre.

En la tabla se evidencia la diferencia de consumos, aunque se vea levemente mayor en Walker, hay que recordar que solo estamos midiendo el consumo por hora; ahora si lo vemos semanal, mensual o anual significaría una diferencia abismal, por ende, es otro factor en el cual hay que enfocarse.

Tabla 3: Estadística del Consumo específico.

Tecnología de Barra Intercelda	Consumo Especifico Mínimo[kWh/Ton]	Consumo Especifico Promedio[kWh/Ton]	Consumo Especifico Maximo[kWh/Ton]
Walker	1929,5	1957,9	2003,8
Optibar	1862,4	1913,1	1957,5

A continuación, se mostrará una tabla resumen con los promedios obtenidos en las 5 variables seleccionadas los cuales nos dan como sistema con mejores números a Optibar, viendo principales diferencias voltaje, dispersión de corriente y consumo específico.

Tabla 4: Resumen estadístico Tecnologías.

Variables	Walker	Optibar
Voltaje celda	2,14(V)	2,099(V)
Eficiencia de Corriente	92,3%	92,6%
Dispersión de Corriente	14,11(A/ m ²)	9,64 (A/ m ²)
Densidad de corriente Promedio	326,67(A/ m ²)	329,1(A/ m ²)
Consumo Especifico	1957,88 (kWh/Ton)	1917,68(kWh/Ton)

3.4. Correlación y Covarianza entre Varios

El análisis de correlación de variables es una herramienta estadística utilizada para evaluar la relación entre dos o más variables en un conjunto de datos. Su objetivo principal es determinar si existe una asociación o dependencia entre estas variables y, en caso afirmativo, la fuerza y dirección de esa relación.

Este nos permite aplicar nuestra estadística descriptiva vista en el punto anterior y visualizar los datos de manera gráfica, analizar las tendencias entre variables, analizar con qué frecuencia se reiteran ciertos eventos, segmentación de datos, entre otras.

A continuación, se muestran gráficos en los cuales se hacen correlaciones y como variable fija se deja la dispersión de corriente, ya que es el factor que genera mayor intercambio de energía y como variables cambiantes tenemos el consumo específico, voltaje de celda y eficiencia de Corriente.

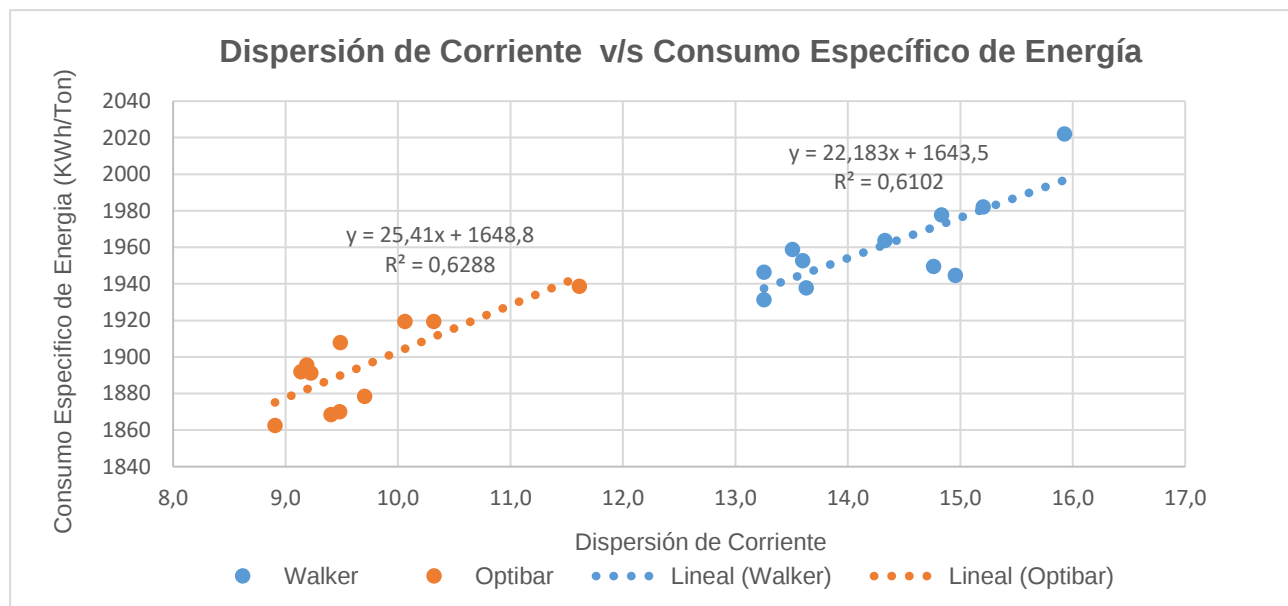


Fig. 3-2: Grafico Dispersión de Corriente v/s Consumo Específico de Energía

En el grafico se presenta la correlación existente entre dispersión de corriente y Consumo Especifico, el cual nos indica que para Walker el sistema presenta mayor dispersión de corriente; por ende, un mayor consumo especifico de Energía, ahora en cambio para Optibar podemos decir que presenta menor dispersión y su consumo especifico es menor.

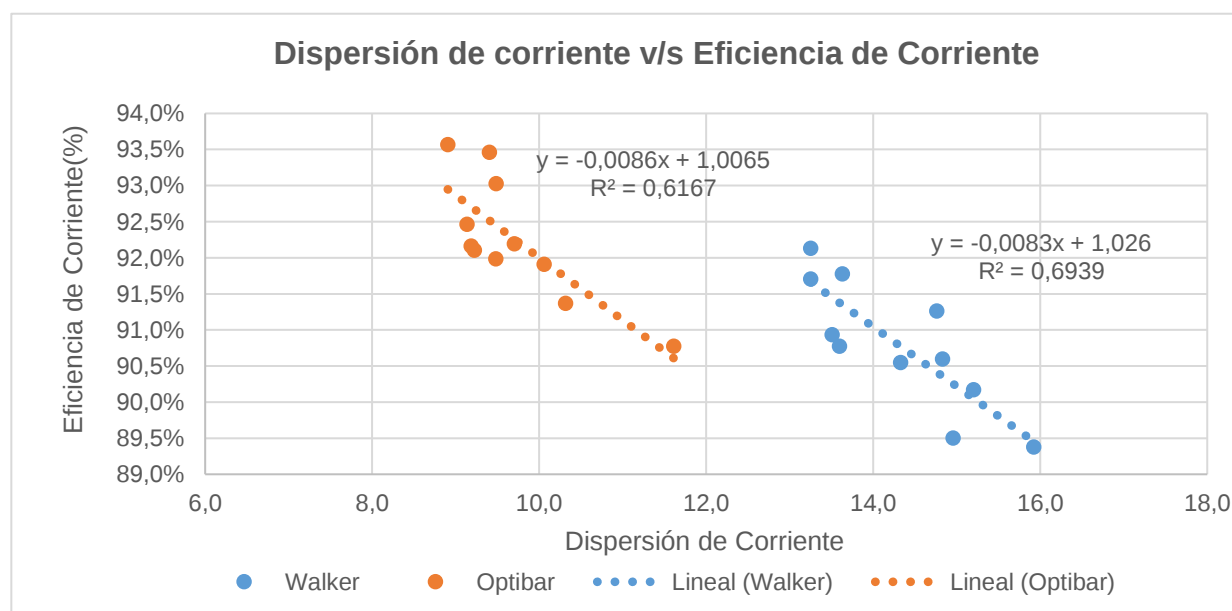


Fig. 3-3: Grafico Dispersión de Corriente v/s Eficiencia de Corriente

En la correlación de dispersión de corriente con eficiencia de corriente se puede ver una similitud de eficiencia de corriente, pero esto es engañoso debido a que Optibar presenta una menor variación de eficiencia por ende el sistema será más estable y esto implica que el mismo no pueda tener un correcto funcionamiento. Se puede ver que Walker tiene 2 fechas en que su eficiencia de corriente es baja lo que implica fallos en el sistema, esto se puede deber a corto circuitos en el sistema.

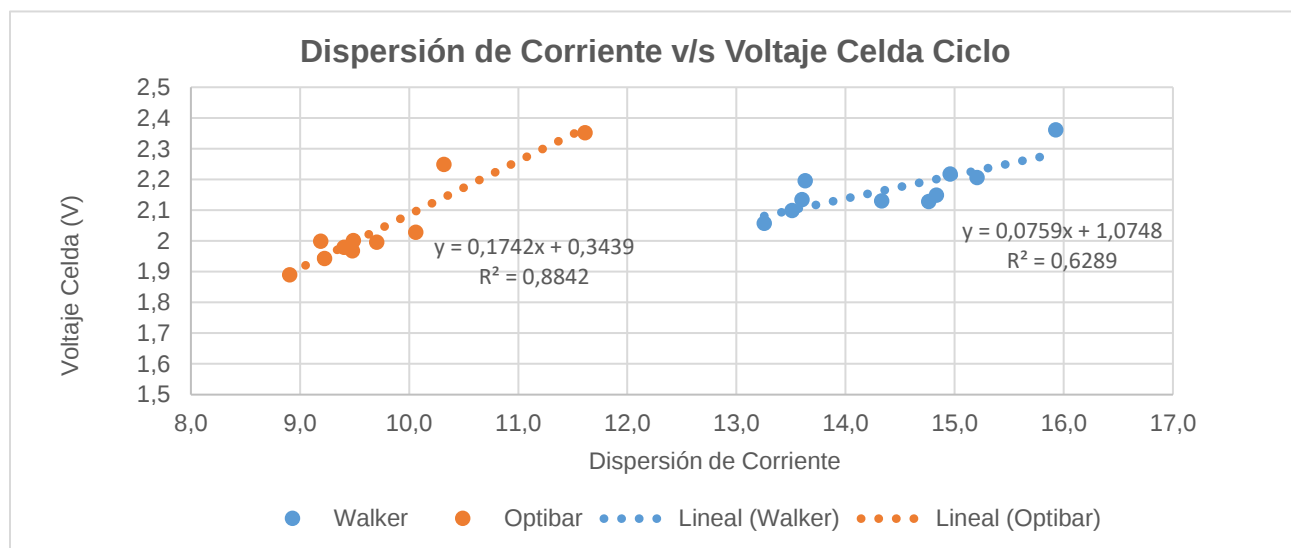


Fig. 3-4: Grafico Dispersión de Corriente v/s Voltaje Celda Ciclo.

Se puede observar un incremento del valor del voltaje de celda a medida que aumenta la dispersión de corriente, pero cabe recalcar que los valores de la Tecnología son mucho mayores en tema de la dispersión de corriente lo cual indica una mayor inestabilidad en el sistema y esto podría causar muchos problemas en el proceso de EW de cobre o Mayores consumos energéticos.

3.5. Comparación de Datos Estadísticos

En el capítulo 3 pudimos evidenciar que en el proceso de EW de cobre existen 3 variables que determinan la calidad y Funcionamiento del proceso:

- Consumo Especifico
- Eficiencia Energética
- Voltaje de Celda

Pero estas 3 variables por sí mismo no es posible controlarlas de manera autónoma, por lo tanto, en los puntos 3.3 y 3.4 se analizaron las variables y se correlacionaron con variables comunes dentro del sistema. Y como resultado se obtuvo que la Variable no Principal que más afecta a las principales es la Dispersión de Corriente.

Capítulo 4. Modelos de Predicción

En este capítulo se mencionará los métodos predictivos que se van a aplicar, específicamente de la Regresión Lineal y Redes ANFIS explicando en que consiste cada uno y la estructura matemática que lo compone.

En el ámbito de la ciencia de datos y el aprendizaje automático, los modelos de predicción son herramientas fundamentales utilizadas para predecir el valor de una variable objetivo basándose en datos históricos y patrones identificados en esos datos. Estos modelos son ampliamente empleados en diversas industrias y disciplinas para hacer pronósticos y tomar decisiones informadas.

4.1. Regresión Lineal

La regresión lineal es un método estadístico utilizado para modelar la relación entre una variable dependiente (o respuesta) y una o más variables independientes (o predictores). Su objetivo principal es entender y predecir cómo cambia la variable dependiente cuando cambian las variables independientes.

La regresión lineal busca encontrar la mejor lineal recta que represente la relación entre las variables involucradas. Esta Línea se determina minimizando la suma de los errores cuadráticos entre los valores observados y los valores predichos por el modelo.

Tipos de Regresiones

Regresión Lineal Simple

Estructura de un Modelo de Regresión Lineal Simple:

$$Y = \beta_0 + \beta_1 X + \varepsilon \quad (9)$$

Donde;

$$\beta_0 = \text{interseccion eje y (cte)}$$

$$\beta_1 = \text{pendiente recta (cte)}$$

$$\varepsilon = \text{error}$$

En la Expresión () se puede observar que los factores o causas que puedan influir en la respuesta Y. Se pueden dividir en dos grupos:

- Variable Explicativa X
- Error (ε)

Regresión Lineal Múltiple

La regresión lineal múltiple permite estudiar como varias Variables están relacionadas con una variable de interés y como estas variables combinadas afectan a la variable dependiente.

A continuación, se muestra un modelo Base:

- Estructura de un Modelo de Regresión Lineal múltiple

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots \dots + \beta_n X_n + \varepsilon \quad (10)$$

Donde;

- Y es la Variable dependiente que queremos predecir o modelar.
- $X_1, X_2, \dots, X_n$ son las Variables Independientes que se Utilizan para predecir Y .
- $\beta_0, \beta_1, \dots, \beta_n$ son los Coeficientes de regresión, que representan el efecto de cada Variable independiente en la variable dependiente.
- ε es el termino de error, que representa la variabilidad no explicada por el modelo.

4.2. Redes ANFIS

Las Redes ANFIS son un tipo de red neuronal artificial que combina las técnicas de la lógica difusa y las redes neuronales para crear un sistema de inferencia difusa basado en el aprendizaje. Se utiliza este tipo de red neuronal ya que, teóricamente es el que más se acomoda al control y predicción de datos.

Definiciones Fundamentales:

- **Lógica Difusa:** es un enfoque matemático que permite modelar y resolver problemas que implican incertidumbre, imprecisión y ambigüedad en los datos. A diferencia de la lógica clásica, que se basa en valores booleanos (verdadero o falso), la lógica difusa maneja grados de verdad que pueden ser cualquier valor en un intervalo de $[0, 1]$, lo que refleja la idea de que las afirmaciones pueden ser parcialmente verdaderas o falsas.
- **Redes Neuronales:** son modelos computacionales inspirados en el funcionamiento del cerebro humano que se utilizan para el aprendizaje automático y el procesamiento de datos. Estas redes están compuestas por unidades básicas llamadas neuronas artificiales o nodos, que están organizadas en capas interconectadas.
- **Validación Cruzada:** es una técnica utilizada en el aprendizaje automático para evaluar el rendimiento de un modelo predictivo. Su objetivo principal es estimar cómo se desempeñará

un modelo en datos no vistos, es decir, datos que no se utilizaron durante el entrenamiento del modelo.

Variables de entrada

Cada sistema o modelo predictivo debe considerar variables con las cuales el sistema va a poder ser entrenado y validado, por ende, es de suma importancia que la cantidad de datos que contenga cada variable sea la mayor posible ya que esto permitirá tener menor errores en los valores que debamos predecir.

Selección de Datos

Una vez que se tienen definidas las Variables de entrada es importante distribuir los datos, para que el modelo seleccionado pueda entrar en aprendizaje.

Datos de Entrenamiento y Prueba

Los datos de entrenamiento se utilizan para ajustar los parámetros del modelo ANFIS para que se ajuste mejor a los datos disponibles, mientras que los datos de prueba se utilizan para evaluar qué tan bien el modelo generaliza a nuevos datos que no fueron utilizados durante el entrenamiento.

Estos pueden ser seleccionados de la siguiente Manera:

- División de Datos
- Aleatoriedad
- Validación Cruzada

Parámetros de un modelo

Es importante mencionar que para la creación de un modelo de cualquier red ANFIS es necesario ajustar ciertos parámetros para que el sistema pueda estar ajustado y dar una mejor salida; respecto a la entrada entregada por el creador del modelo.

En el General de los modelos se pueden encontrar 5 tipos de parámetros [5]:

1. Parámetros de las Funciones de Membresía

- Parámetros de forma: Estos son los parámetros que definen la forma y posición de las funciones de membresía utilizadas en las variables difusas de entrada.

2. Parámetros de los nodos tipo Takagi-Sugeno

- Coeficientes lineales: En un nodo tipo Takagi-Sugeno, que es comúnmente utilizado en las redes ANFIS, los parámetros incluyen los coeficientes de la regresión lineal asociada con cada regla difusa. Estos coeficientes determinan cómo se combinan las salidas de las funciones de membresía difusas para producir una salida final.

3. Parámetros de aprendizaje

- Coeficientes de adaptación: Estos son los parámetros que se ajustan durante el entrenamiento del modelo para minimizar el error entre las predicciones del modelo y los datos de entrenamiento. Los métodos de optimización, como el descenso de gradiente, se utilizan para ajustar estos parámetros.

4. Parámetro de las Funciones de pertenencia

- Centros y anchuras: En las funciones de pertenencia utilizadas en las variables difusas de entrada, los parámetros como los centros y anchuras determinan la posición y extensión de las funciones.
- Parámetros de reglas y combinación: Pesos de las reglas: En las redes ANFIS, cada regla difusa tiene asociado un peso que indica la importancia o contribución de esa regla a la salida final del sistema. Estos pesos son ajustados durante el entrenamiento.

5. Parámetros de entrenamiento:

- Número de épocas: Este parámetro define cuántas veces se iterará sobre el conjunto de datos durante el proceso de entrenamiento.
- Tasa de aprendizaje: Es el factor que determina la magnitud de los ajustes realizados en los parámetros durante cada paso de optimización.

Arquitectura

La Arquitectura de la Red es de suma importancia debido a que nos indica la cantidad de funciones

La arquitectura está compuesta por 5 capas [7]:

- 1) Capa 1: las entradas corresponden a x y y . La salida del nodo es el grado de pertenencia para el que la variable de entrada satisface el término asociado a este nodo.
- 2) Capa 2: cada nodo calcula el grado de activación de su regla asociada. Ambos nodos están representados con una T en la Figura 6, debido que pueden representar cualquier t-norma para modelar la operación lógica and. Los nodos de esta capa son conocidos como de reglas.
- 3) Capa 3: cada nodo en esta capa está representado por una N en la Figura 4-1, para indicar la normalización de los grados de activación. La salida del nodo es el grado de activación normalizado de la regla i , con respecto a la suma de los grados de activación.
- 4) Capa 4: la salida de los nodos corresponde al producto entre el grado de activación normalizado por la salida individual de cada regla.
- 5) Capa 5: el único nodo de esta capa calcula la salida total del sistema (agregación) como la suma de sus entradas individuales.

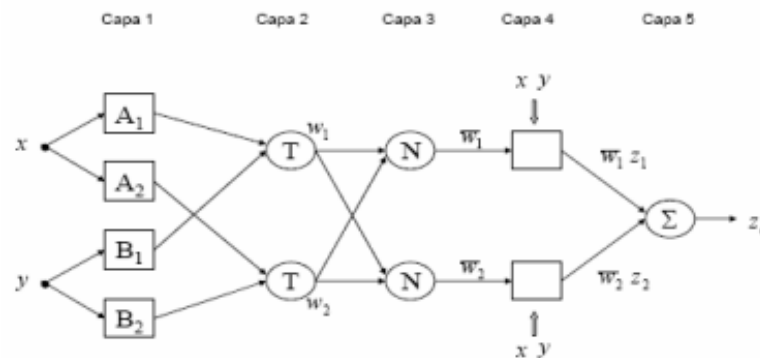


Fig. 4-1: Arquitectura ANFIS para inferencia Takagi-Sugeno de Primer Orden [7].

Capítulo 5. Resultados

5.1. Introducción

En este capítulo se menciona el origen de los datos utilizados en la creación de los modelos de predicción de regresión lineal y Redes Anfis, para luego presentar los parámetros utilizados en los modelos y como último paso entregar los errores de los modelos con el propósito de compararlos y ver en que condiciones y cual se comporta mejor.

5.2. Selección de datos

Se seleccionaron 6 bases de datos, las cuales toman la primera medición del día para la primera celda de EW de cobre, correspondientes a la medición de la densidad de corriente en los cátodos:

- Base 1: Contiene 7 días continuos de medición.
 - Base 2: Contiene 7 días continuos de medición.
 - Base 3: Contiene 15 días continuos de medición.
 - Base 4: Contiene 7 días continuos de medición.
 - Base 5: Contiene 7 días continuos de medición.
 - Base 6: Contiene 8 días continuos de medición.
- ♣ No se consideraron mediciones menores a 6 días debido a que como requisito mínimo para predecir 4 días es requisito tener un mínimo de 7 días de estudio, se consideraron datos operacionales en condiciones normales, es decir, no se consideraron corto circuitos ni circuitos abiertos.

Base de datos para modelos

Luego de seleccionar las densidades de corriente de nuestra base, con ayuda de Excel se calcula la dispersión de corriente diaria, que es calcular la desviación estándar de los 65 cátodos de cada celda:

Por ende, nuestra base se verá modificada, quedando compuesta por 4 bases:

Conjunto de datos n°1:

El primer conjunto de datos se compone de 4 bases las que están compuestas por 4 días previos de mediciones para poder predecir uno al futuro, las bases del conjunto 1 están compuestas con 5 columnas de las cuales 4 son de mediciones pasadas y la última representa el día futuro.

- PA4-PRE1
- PA4-PRE2
- PA4-PRE3
- PA4-PRE4

Conjunto de datos n°2:

El segundo conjunto de datos se compone de 4 bases las que están compuestas por 3 días previos de mediciones para poder predecir uno al futuro, las bases del conjunto 1 están compuestas con 4 columnas de las cuales 3 son de mediciones pasadas y la última representa el día futuro.

- PA3-PRE1
- PA3-PRE2
- PA3-PRE3
- PA3-PRE3

- ♣ El conjunto 1 cuenta con cuatro bases y cada una de ellas cuenta con 4 entradas y una salida.
- ♣ El conjunto 2 cuenta con cuatro bases y cada una de ellas cuenta con 3 entradas y una salida.

5.3. Modelos Finales Regresión Lineal múltiple

Modelo Regresión lineal Conjunto de datos 1

- Modelo matemático de Regresión Lineal múltiple para conjunto de datos 1:

$$Y = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \varepsilon \quad (11)$$

- Modelo matemático de Regresión Lineal múltiple para conjunto de datos 2:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon \quad (12)$$

5.4. Modelo Red ANFIS

Los modelos de Red ANFIS contienen una configuración muy amplia debido a que se puede controlar las variables de entrada y salida mediante Funciones de membresía número de épocas y opciones que buscar minimizar el error de la salida de este.

En el presente caso se crearon dos modelos Base, los que se basaron el conjunto de datos 1 y 2 respectivamente

Modelo Red ANFIS conjunto de datos 1

Para este modelo se utilizó:

- 3 Variables de entrada; correspondiente a las densidades de corriente del conjunto 1.
- 2,3,4,5 y 6 Funciones de Membresia, con el fin de ver la diferencia del error RMS al aumentar el numero de funciones.
- la base de datos fue dividida en dos para tener un conjunto de Entrenamiento (70% datos) y test (30% datos).
- 4000 Epocas.
- ErrorGoal 0,1; se utilize en el Código de MATLAB, con el fin de hacer que el Sistema se detenga cuando el error de entrenamiento llegue a 0,1. Esto con el fin de disminuir los tiempos de proceso del codigo y limitar el error de entrenamiento.

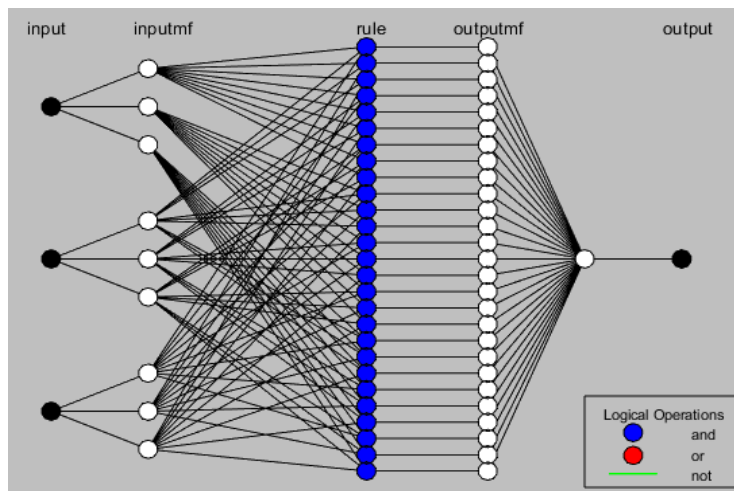


Fig. 5-1: Arquitectura Red ANFIS para 3 variables y 3 funciones de membresia.

Modelo Red ANFIS conjunto de datos 2

Para este modelo se utilizó:

- 4 Variables de entrada; correspondiente a las densidades de corriente del conjunto 2.
- 2,3,4,5 y 6 Funciones de Membresia, con el fin de ver la diferencia del error RMS al aumentar el numero de funciones.
- la base de datos fue dividida en dos para tener un conjunto de Entrenamiento (70% datos) y test (30% datos).
- 4000 Epocas.
- ErrorGoal 4; se utilizo en el Código de MATLAB, con el fin de hacer que el Sistema se detenga cuando el error de entrenamiento llegue a 4. Esto con el fin de disminuir los tiempos de proceso del codigo y limitar el error de entrenamiento.

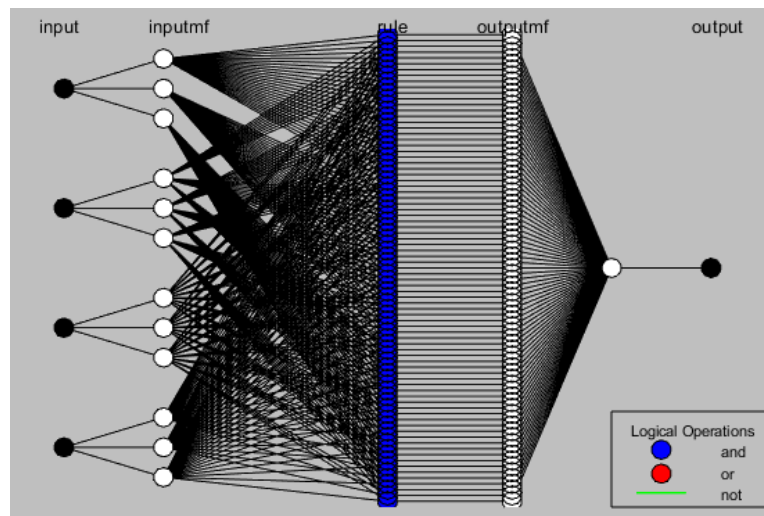


Fig. 5-2: Arquitectura Red ANFIS para 4 variables y 3 funciones de membresia.

5.5. Resultados Modelo Regresión Lineal múltiple

Error de Predicción:

Errores modelo conjunto de datos1:

Tabla 5: Error de predicción Conjunto de datos 1

Base de datos	RMSE de Entrenamiento	RMSE de Prueba
PA4-Pre1	7,399	10,227
PA4-Pre2	9,499	13,784
PA4-Pre3	10,54	4,5
PA4-Pre4	6,232	4,803

Errores modelo Conjunto de datos 2:

Tabla 6: Error de predicción Conjunto de datos 2

Base de datos	RMSE de Entrenamiento	RMSE de Prueba
PA3-Pre1	11,608	7,448
PA3-Pre2	10,66	11,669
PA3-Pre3	10,819	4,366
PA3-Pre4	5,683	9,134

Análisis:

Para tener una noción previa, el promedio de las dispersiones utilizadas en las predicciones es de 68 (A/m^2). Al analizar los resultados obtenidos en los Conjuntos 1 y 2, se puede inferir que los valores de RMSE no son satisfactorios, ya que alcanzan hasta 13 (A/m^2), lo cual es considerable, representando casi un cuarto del valor promedio. Además, se observa que los valores obtenidos en el Conjunto 1 son más consistentes que los obtenidos en el Conjunto 2.

5.6. Resultados Modelo Red ANFIS

Predicciones Modelo Redes ANFIS:

A continuación, se muestran gráficamente los resultados del modelo de red ANFIS, para tener una perspectiva visual del comportamiento de las predicciones del modelo de red ANFIS.

- Conjunto de datos n°1
- 6 funciones de membresía y 4000 épocas

Base 1:

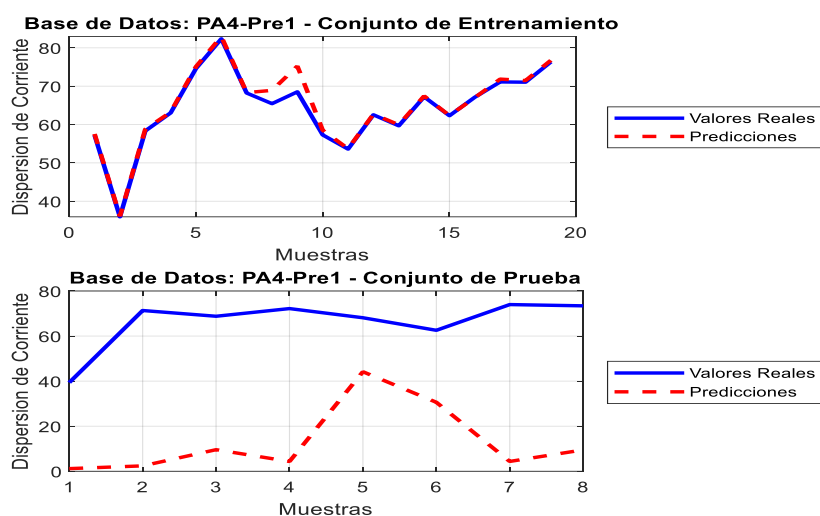


Fig. 5-3: Grafico predicción v/s valores reales (6 funciones de membresía) PA4-Pre1

Base 2:

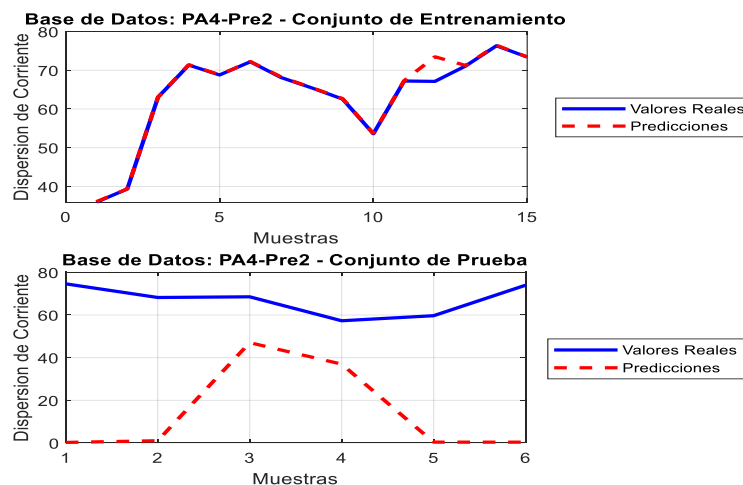


Fig. 5-4: Grafico predicción v/s valores reales (6 funciones de membresía) PA4-Pre2

Base 3:

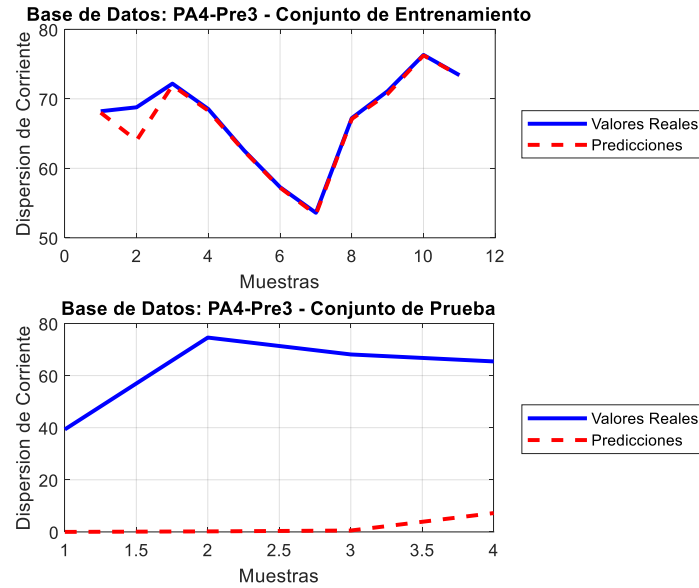


Fig. 5-5: Grafico predicción v/s valores reales (6 funciones de membresia) PA4-Pre3

Base 4:

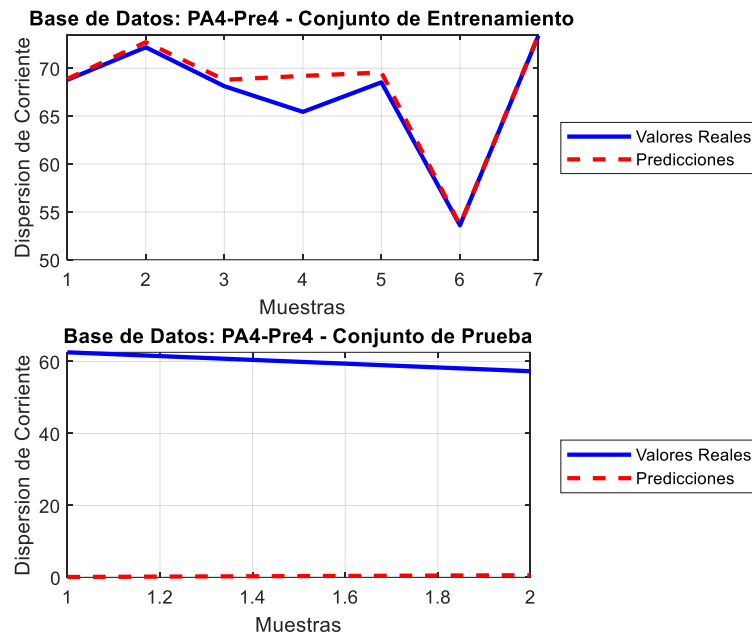


Fig. 5-6: Grafico predicción v/s valores reales (6 funciones de membresia) PA4-Pre4

- 5 funciones de Membresia y 4000 épocas

Base 1:

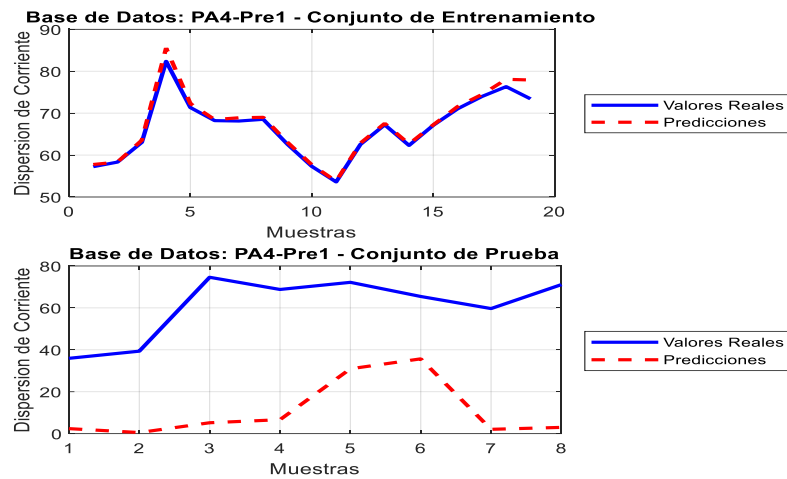


Fig. 5-7: Grafico predicción v/s valores reales (5 funciones de membresia) PA4-Pre1

Base 2:

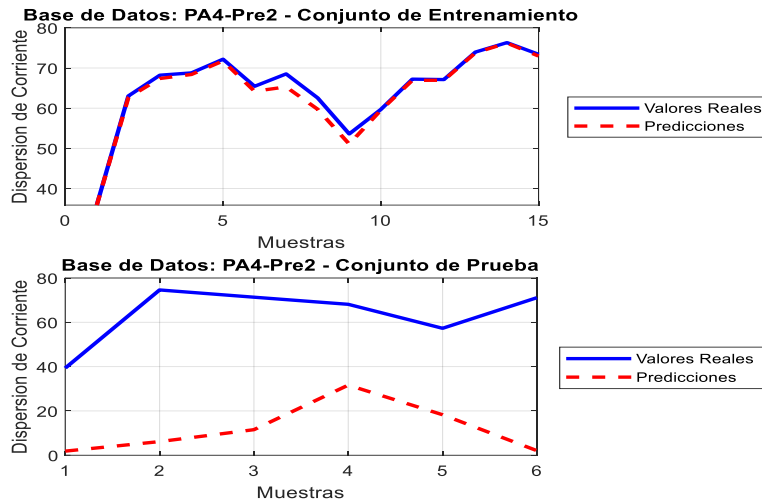


Fig. 5-8: Grafico predicción v/s valores reales (5 funciones de membresia) PA4-Pre2

Base 3:

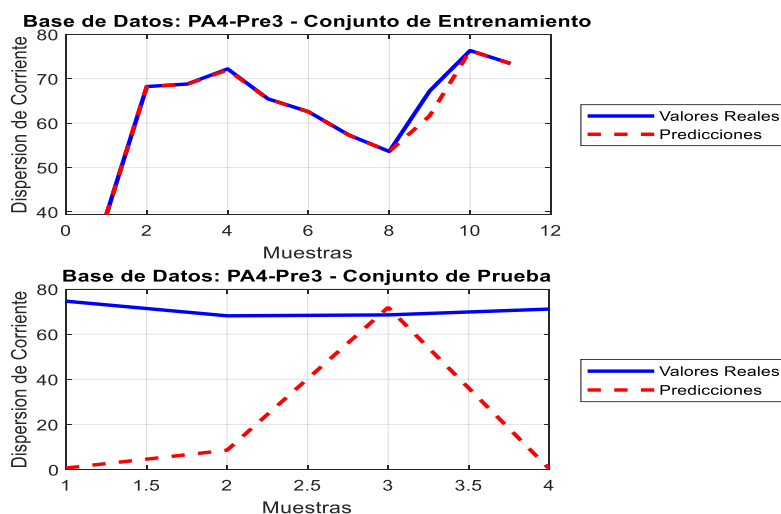


Fig. 5-9: Grafico prediccion v/s valores reales (5 funciones de membresia) PA4-Pre3

Base 4:

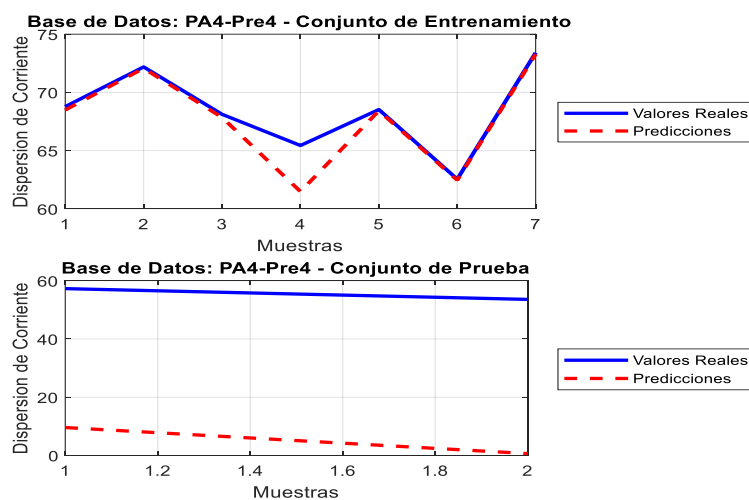


Fig. 5-10: Grafico prediccion v/s valores reales (5 funciones de membresia) PA4-Pre4

Comentarios:

Se observa una mejora en las predicciones tanto para el conjunto de prueba como para el de entrenamiento debido a la disminuci3n de funciones de membresía. Se aprecia un mejor resultado o ajuste en el conjunto de entrenamiento en la Base 1, en comparaci3n con la Base 4. Esto se debe a la mayor cantidad de datos que contiene la Base 1 en relaci3n con la Base 4.

- 4 funciones de membresia y 4000 epocas

Base 1:

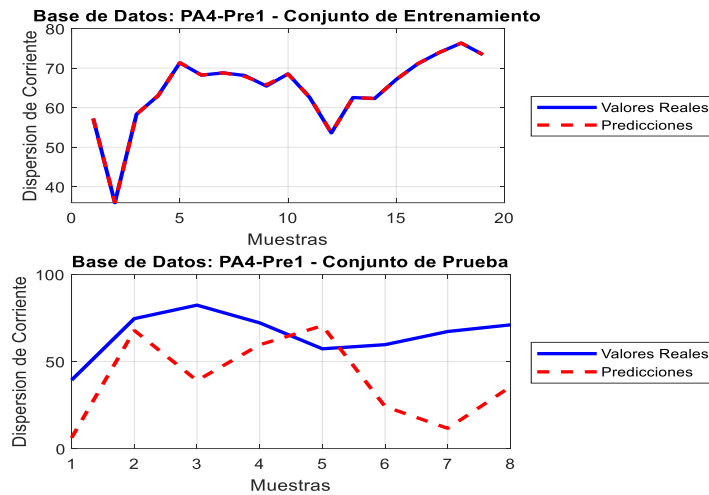


Fig. 5-11: Grafico prediccion v/s valores reales (4 funciones de membresia) PA4-Pre1

Base 2:

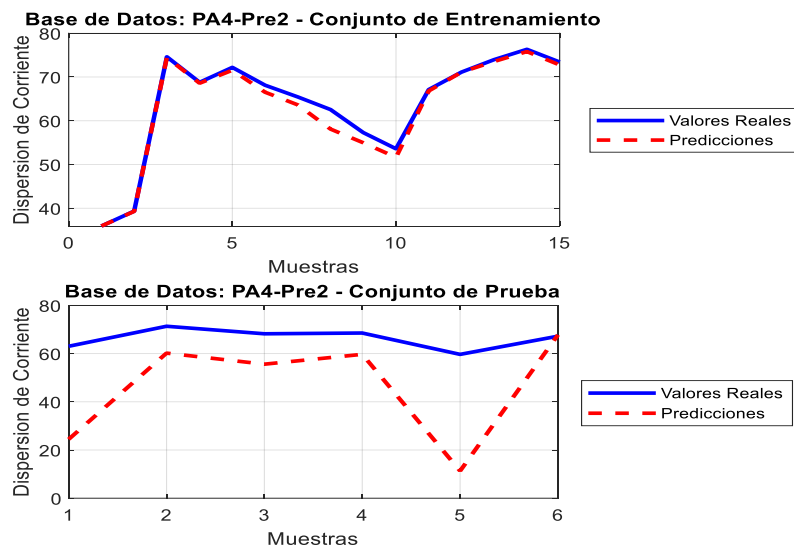


Fig. 5-12: Grafico prediccion v/s valores reales (4 funciones de membresia) PA4-Pre2

Base 3:

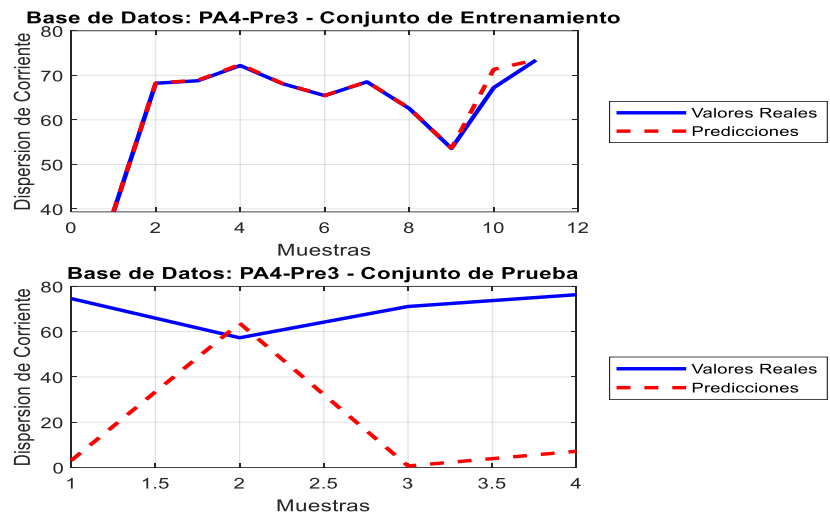


Fig. 5-13: Grafico prediccion v/s valores reales (4 funciones de membresia) PA4-Pre3

Base 4:

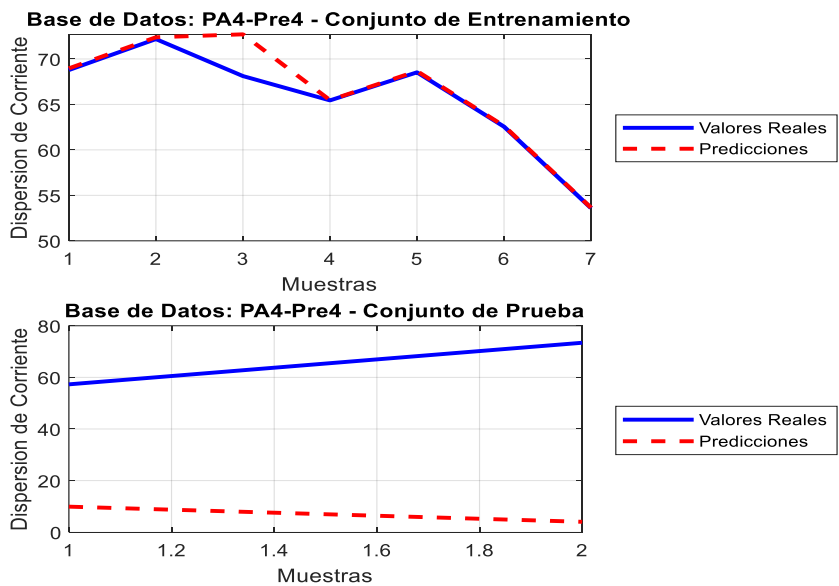


Fig. 5-14: Grafico prediccion v/s valores reales (4 funciones de membresia) PA4-Pre4

- 3 funciones de membresia y 4000 epocas

Base1:

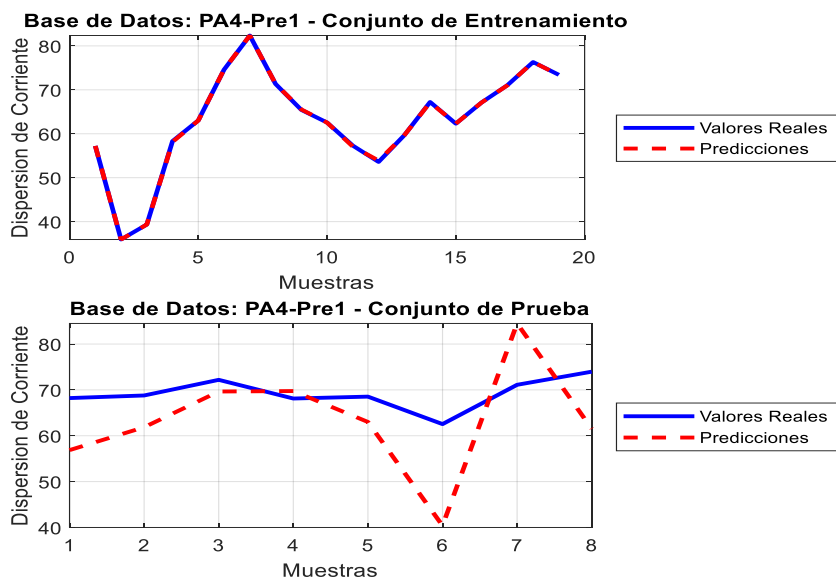


Fig. 5-15: Grafico prediccion v/s valores reales (3 funciones de membresia) PA4-Pre1

Base2:

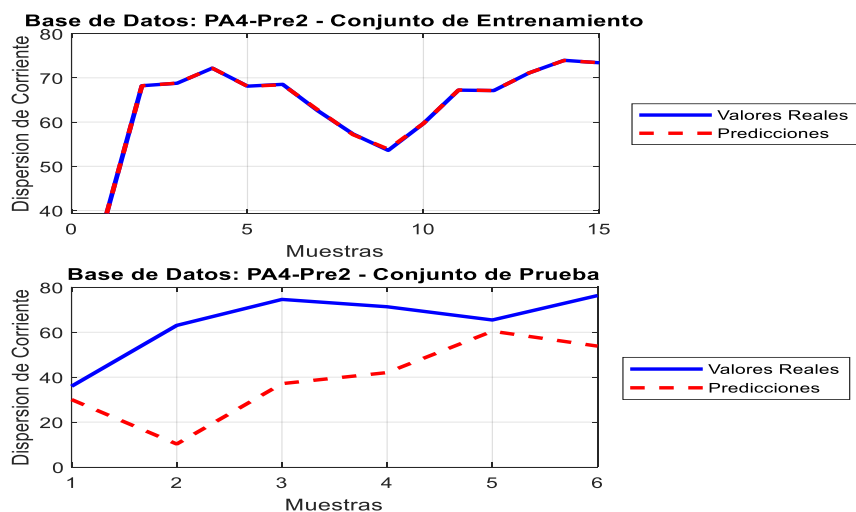


Fig. 5-16: Grafico prediccion v/s valores reales (3 funciones de membresia) PA4-Pre2

Base3:

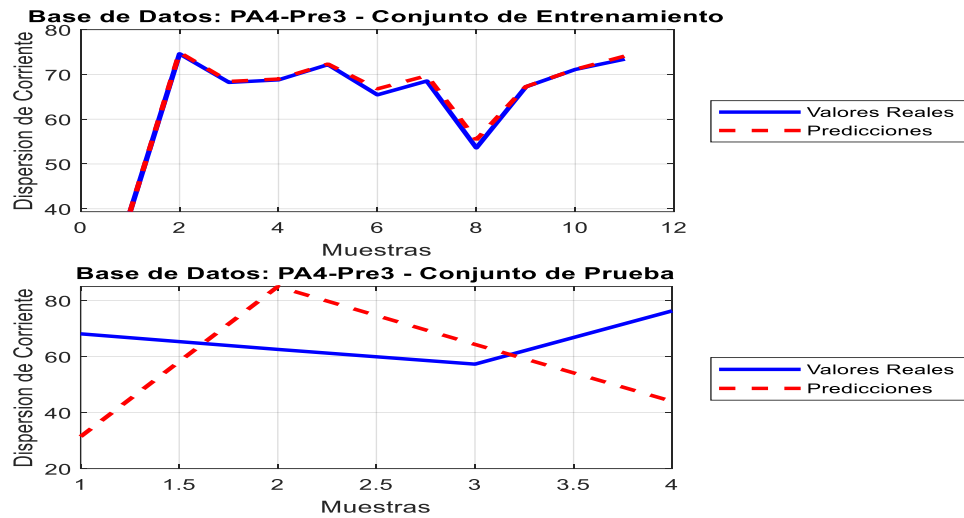


Fig. 5-17: Grafico prediccion v/s valores reales (3 funciones de membresia) PA4-Pre3

Base4:

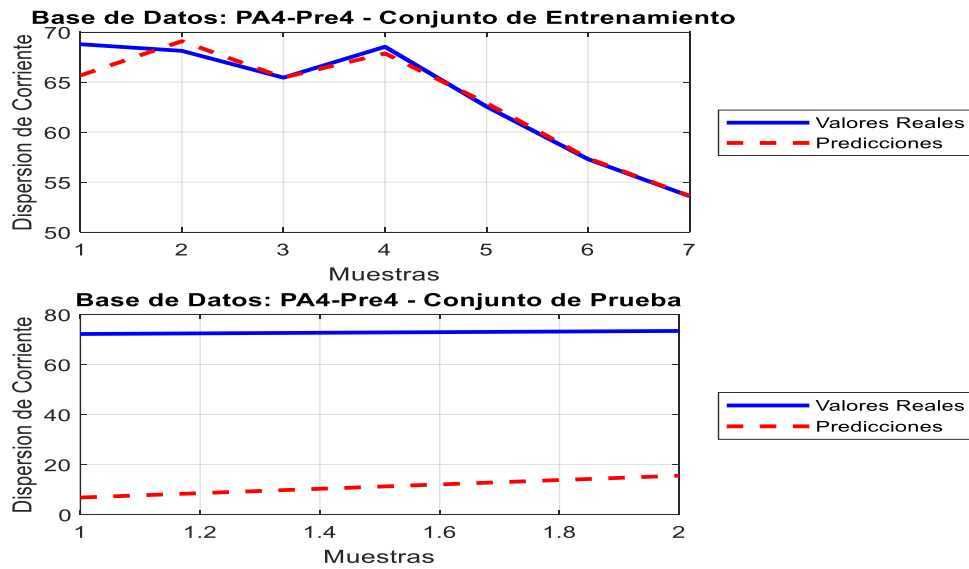


Fig. 5-18: Grafico prediccion v/s valores reales (3 funciones de membresia) PA4-Pre4

- 2 funciones de membresia y 4000 epocas

Base 1:

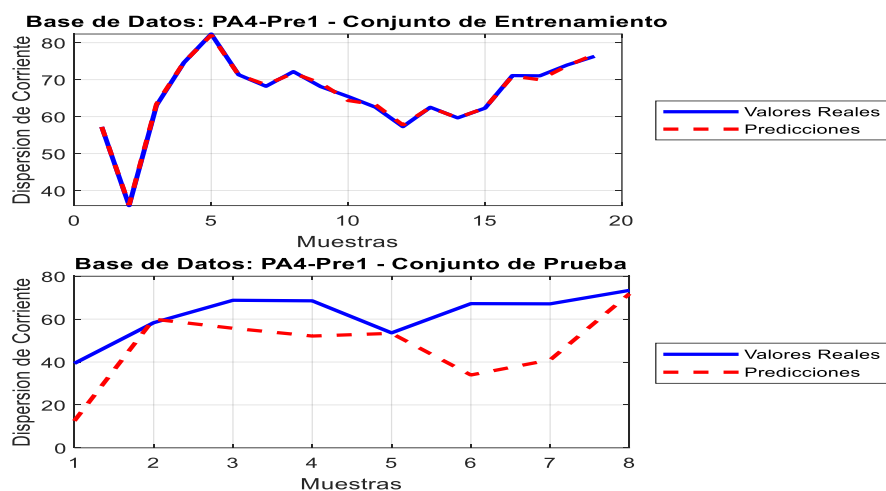


Fig. 5-19: Grafico prediccion v/s valores reales (2 funciones de membresia) PA4-Pre1

Base 2:

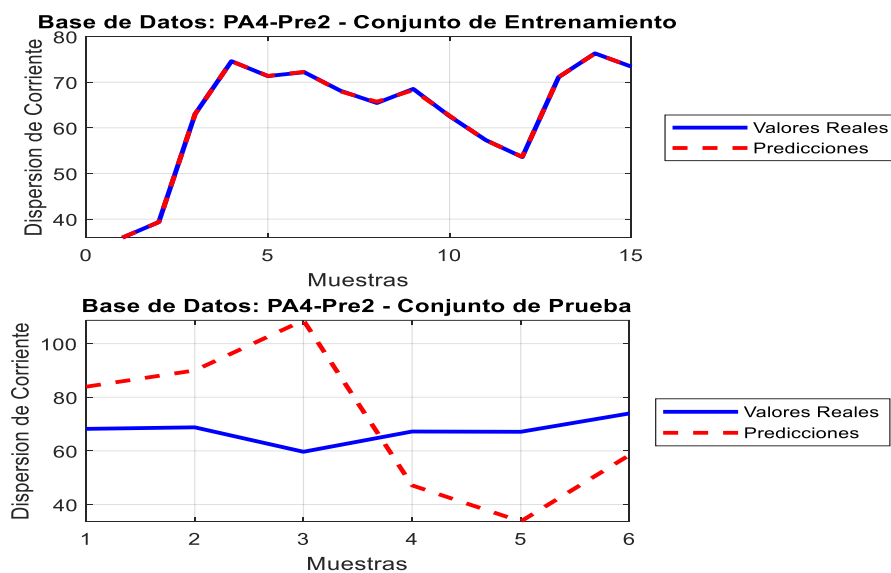


Fig. 5-20: Grafico prediccion v/s valores reales (2 funciones de membresia) PA4-Pre2

Base 3:

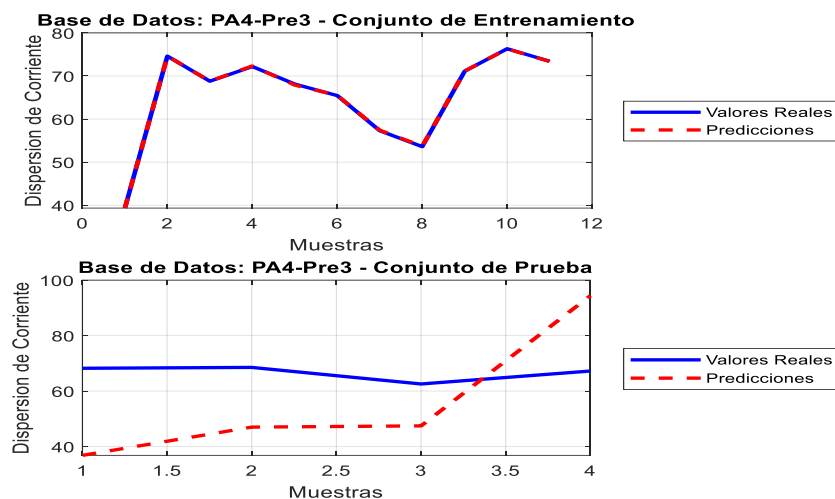


Fig. 5-21: Grafico prediccion v/s valores reales (2 funciones de membresia) PA4-Pre3

Base 4:

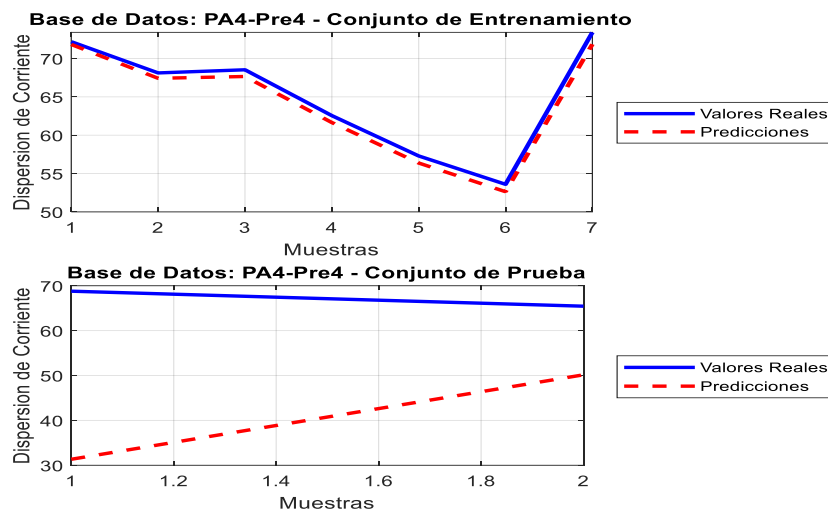


Fig. 5-22: Grafico prediccion v/s valores reales (2 funciones de membresia) PA4-Pre4

Comentarios:

Para el primer conjunto de datos se observó que las predicciones a los valores de entrenamiento son muy cercanas, rozando la perfección y esto se mantiene con todas las funciones de membresía aplicadas, junto con ello hay que recalcar que los valores se ajustan más a la base 1 de conjunto n°1, eso se debe a que la que cuenta con mayor cantidad de datos.

Por otra parte, en todas las bases que se utilizó 2 funciones los datos tienen tendencia a ajustarse más a los valores reales.

- Conjunto de datos n°2
- 6 funciones de membresia y 4000 epocas

Base 1:

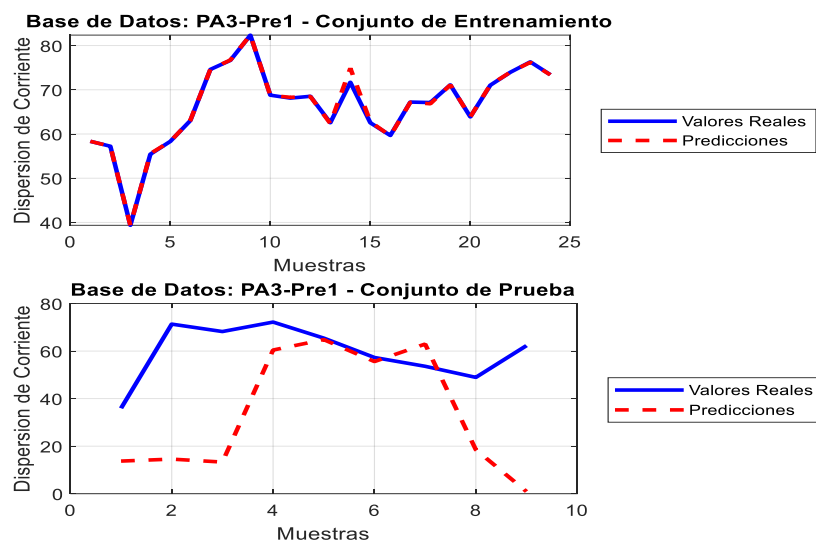


Fig. 5-23: Grafico prediccion v/s valores reales (6 funciones de membresia) PA3-Pre1

Base2:

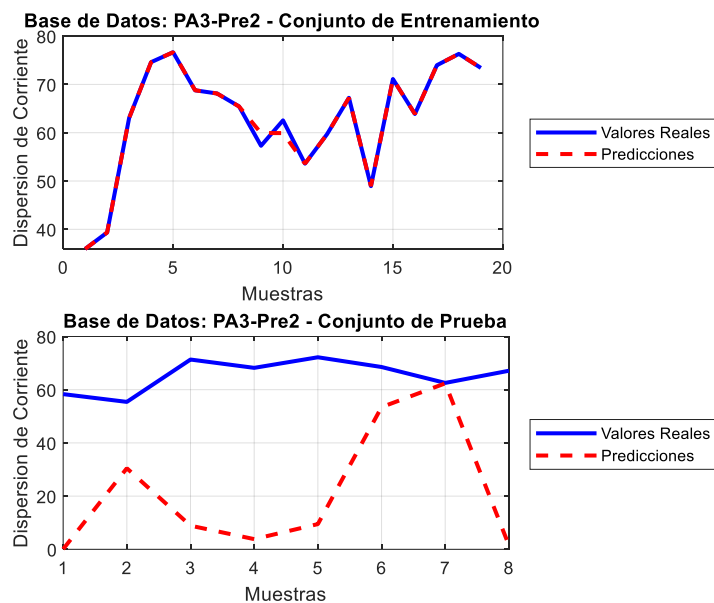


Fig. 5-24: Grafico prediccion v/s valores reales (6 funciones de membresia) PA3-Pre2

Base3:

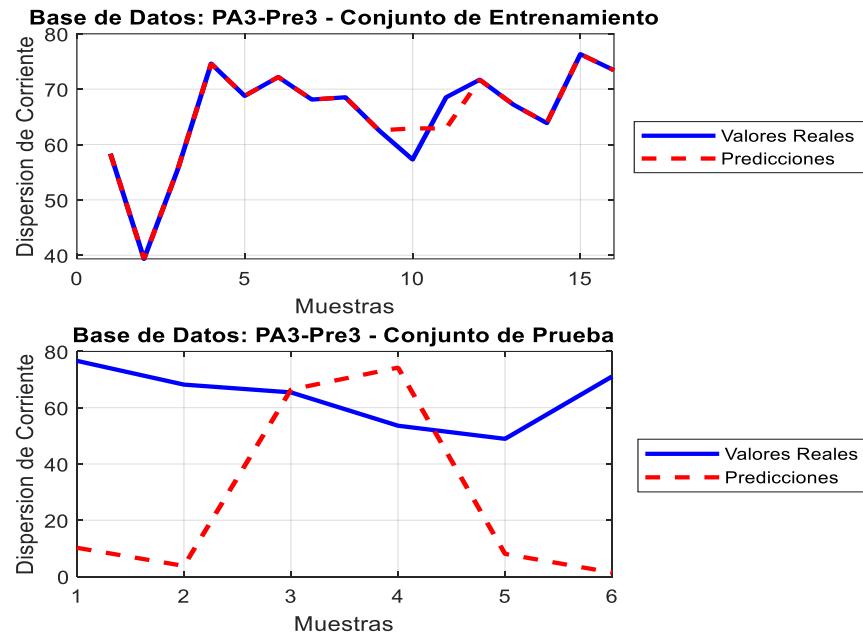


Fig. 5-25: Grafico prediccion v/s valores reales (6 funciones de membresia) PA3-Pre3

Base4:

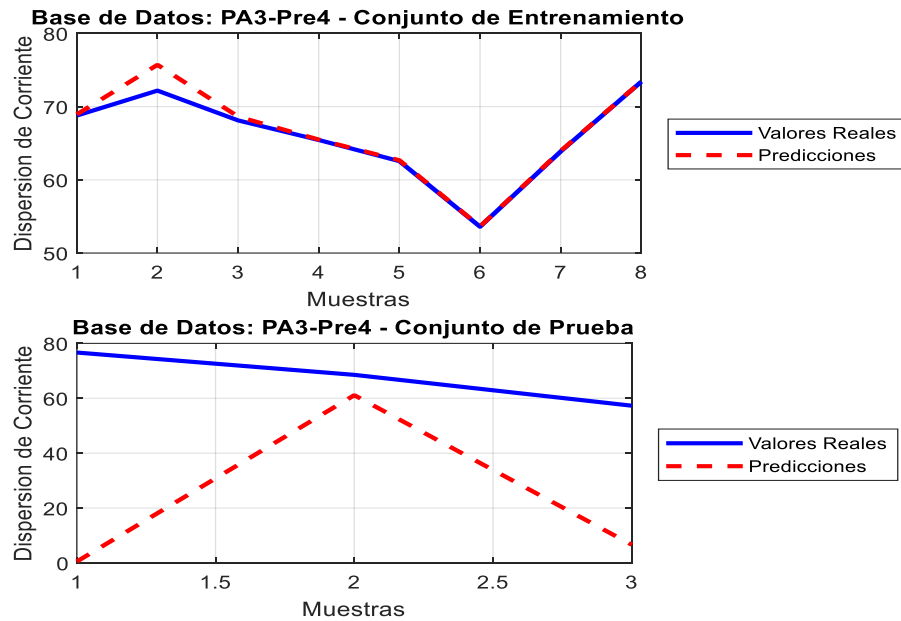


Fig. 5-26: Grafico prediccion v/s valores reales (6 funciones de membresia) PA3-Pre4

- 5 funciones de membresia y 4000 epocas

Base 1:

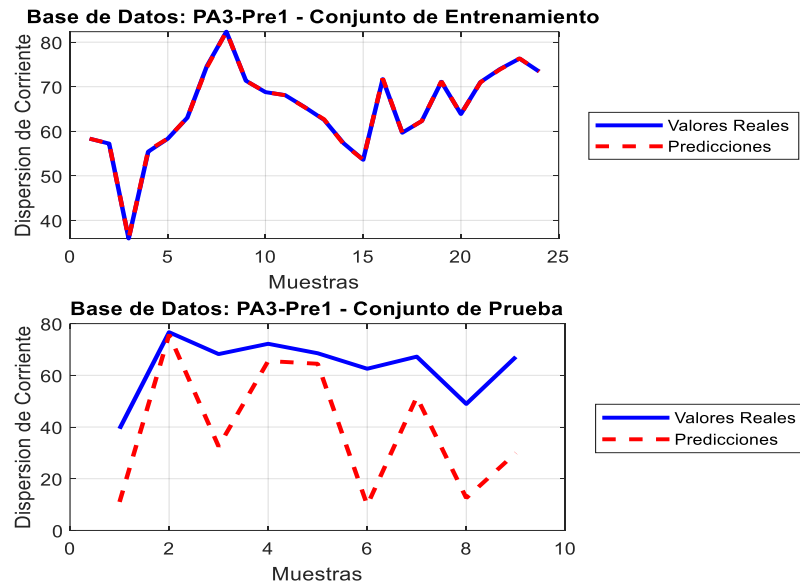


Fig. 5-27: Grafico prediccion v/s valores reales (5 funciones de membresia) PA3-Pre1

Base2:

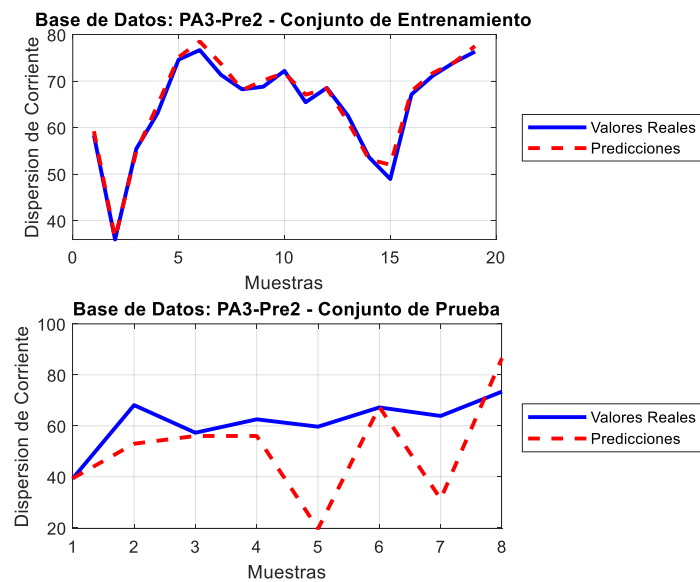


Fig. 5-28: Grafico prediccion v/s valores reales (5 funciones de membresia) PA3-Pre2

Base 3:

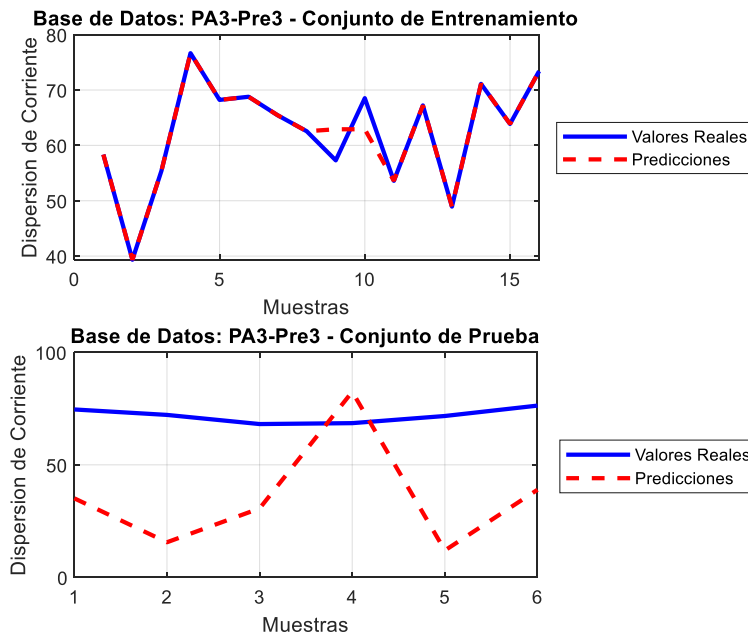


Fig. 5-29: Grafico prediccion v/s valores reales (5 funciones de membresia) PA3-Pre3

Base 4:

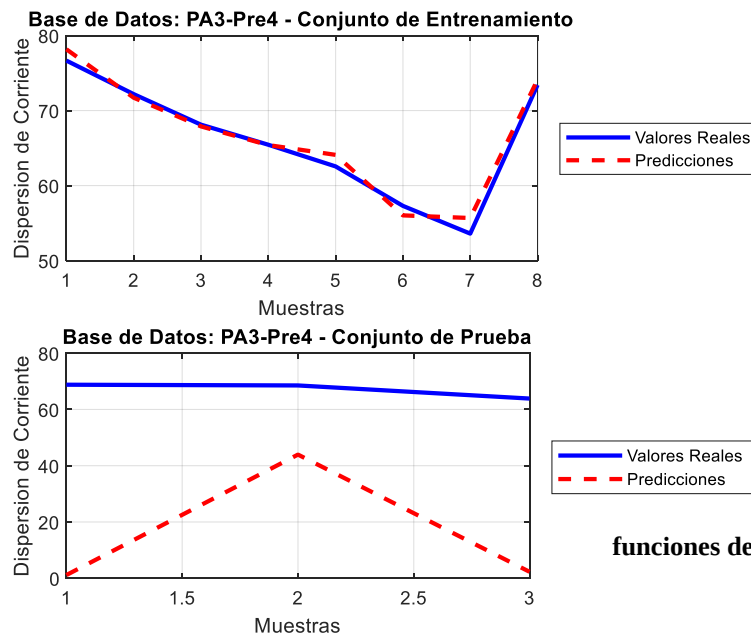


Fig. 5-30:
valores reales (5

Grafico prediccion v/s
funciones de membresia) PA3-Pre4

➤ 4 funciones de membresia y 4000 epocas

Base 1:

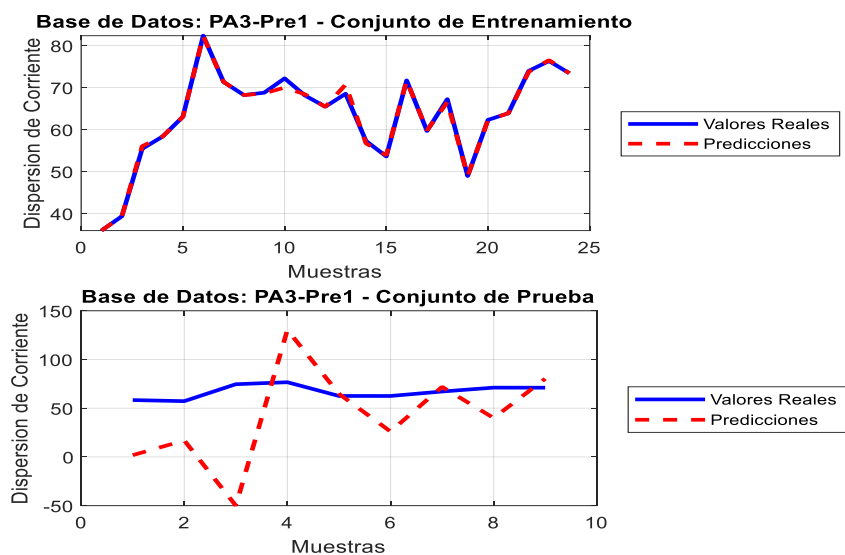


Fig. 5-31: Grafico prediccion v/s valores reales (4 funciones de membresia) PA3-Pre1

Base2:

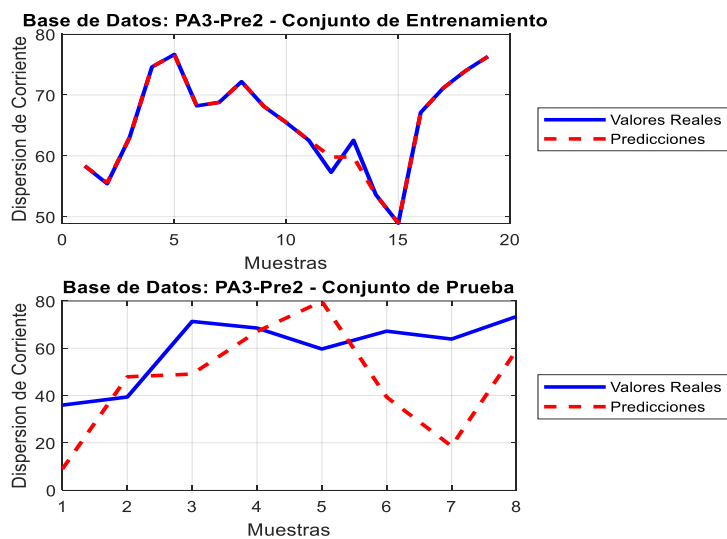


Fig. 5-32: Grafico prediccion v/s valores reales (4 funciones de membresia) PA3-Pre2

Base 3:

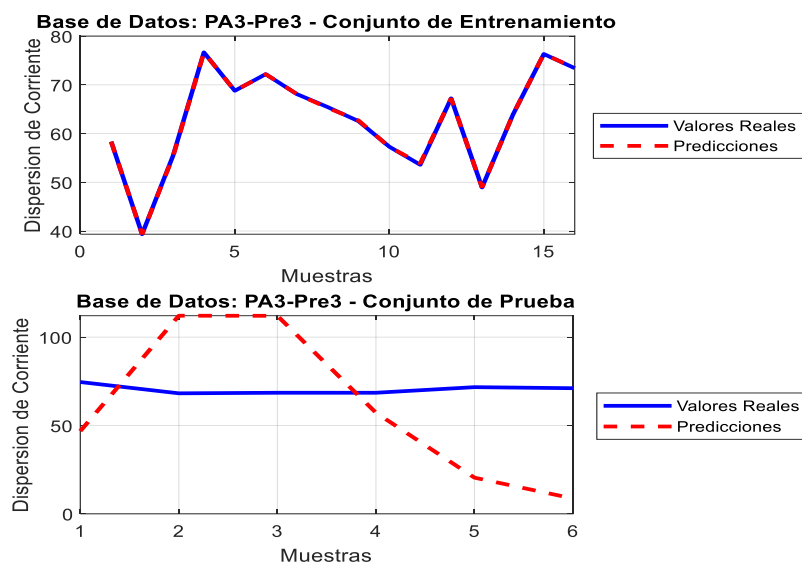


Fig. 5-33: Grafico prediccion v/s valores reales (4 funciones de membresia) PA3-Pre3

Base 4:

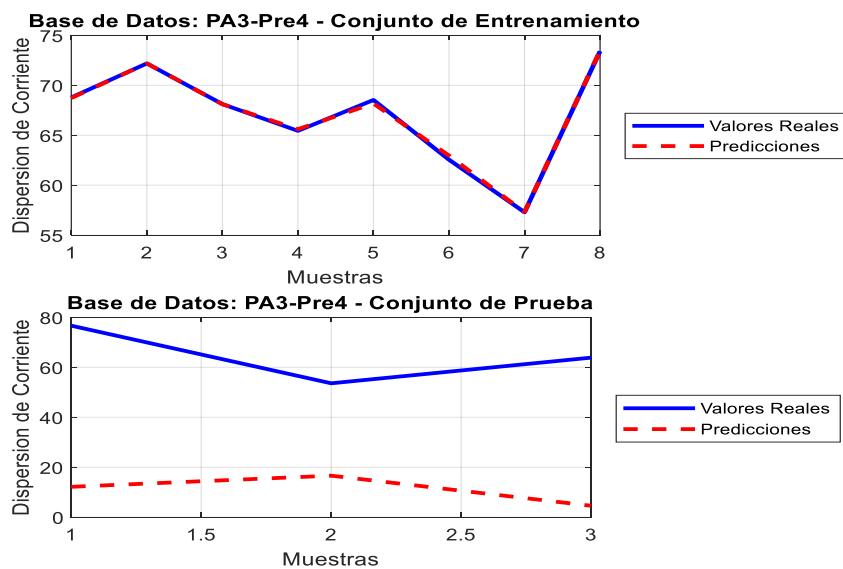


Fig. 5-34: Grafico prediccion v/s valores reales (4 funciones de membresia) PA3-Pre4

Comentarios:

Se puede ver que los resultados obtenidos tanto en 6,5 y 4 funciones de membresía es un resultado muy variable y con demasiados errores en el sistema se ve que los conjuntos de entrenamiento aun no logran acoplarse bien a los valores reales.

- 3 funciones de membresía y 4000 épocas

Base1:

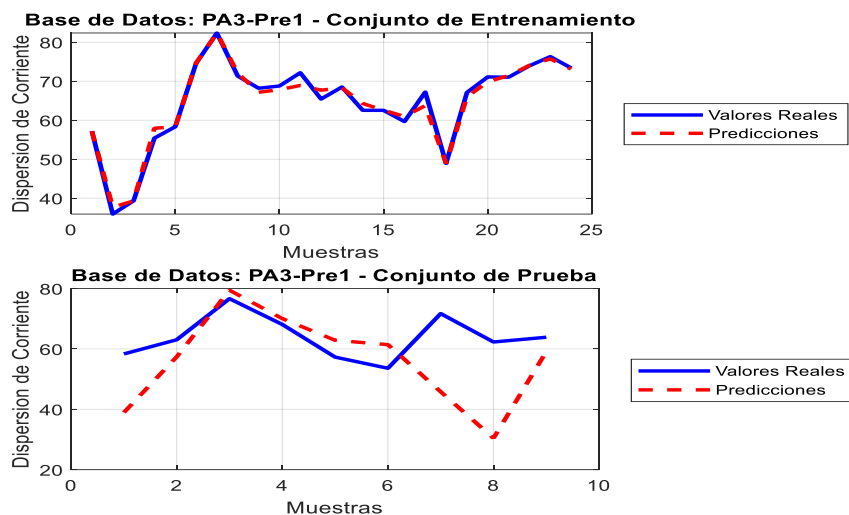


Fig. 5-35: Grafico prediccion v/s valores reales (3 funciones de membresía) PA3-Pre1

Base 2:

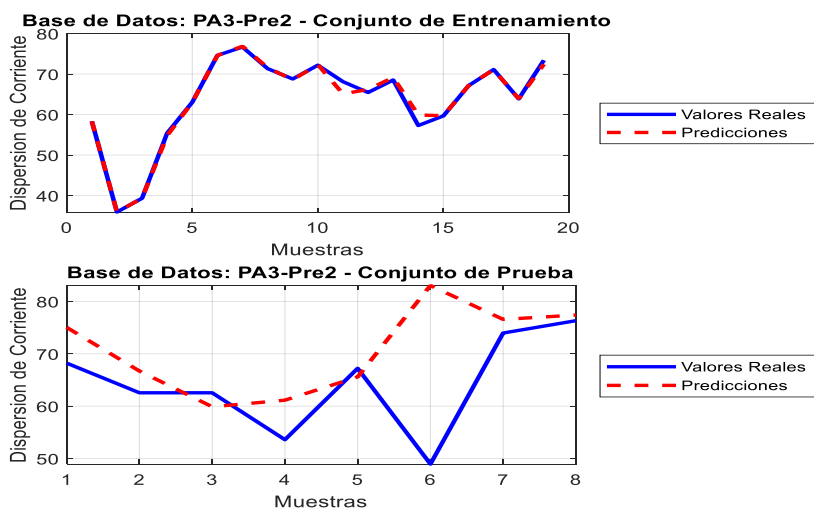


Fig. 5-36: Grafico prediccion v/s valores reales (3 funciones de membresía) PA3-Pre2

Base 3:

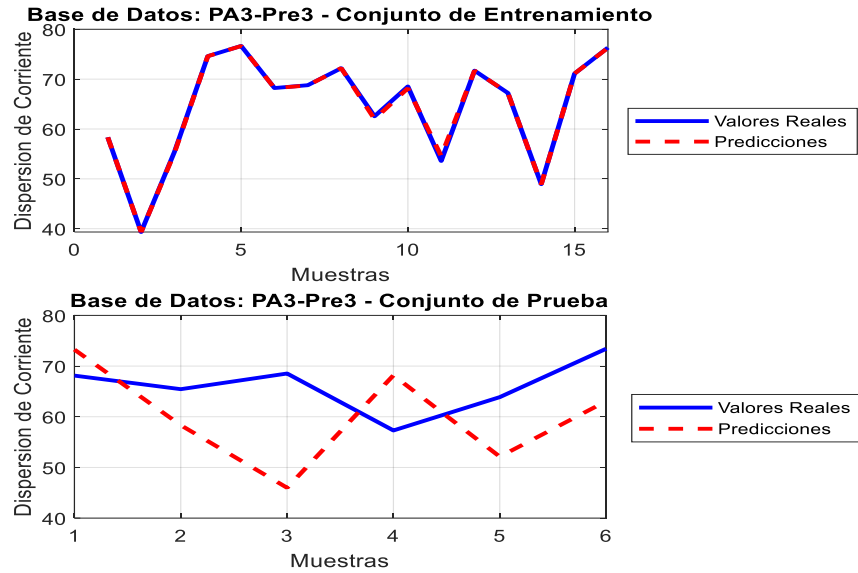


Fig. 5-37: Grafico prediccion v/s valores reales (3 funciones de membresia) PA3-Pre3

Base 4:

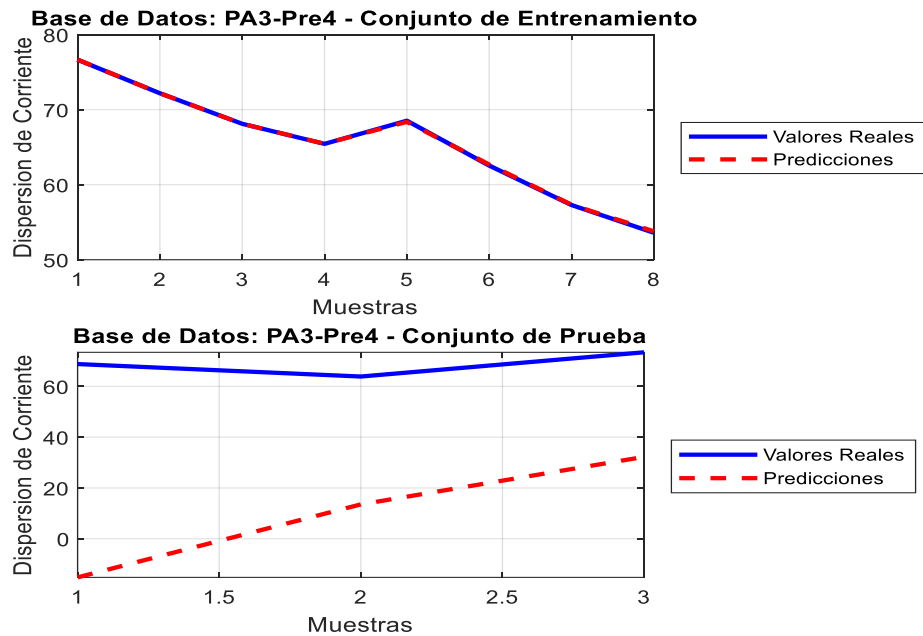


Fig. 5-38: Grafico prediccion v/s valores reales (3 funciones de membresia) PA3-Pre4

- 2 funciones de membresia y 4000 epocas

Base1:

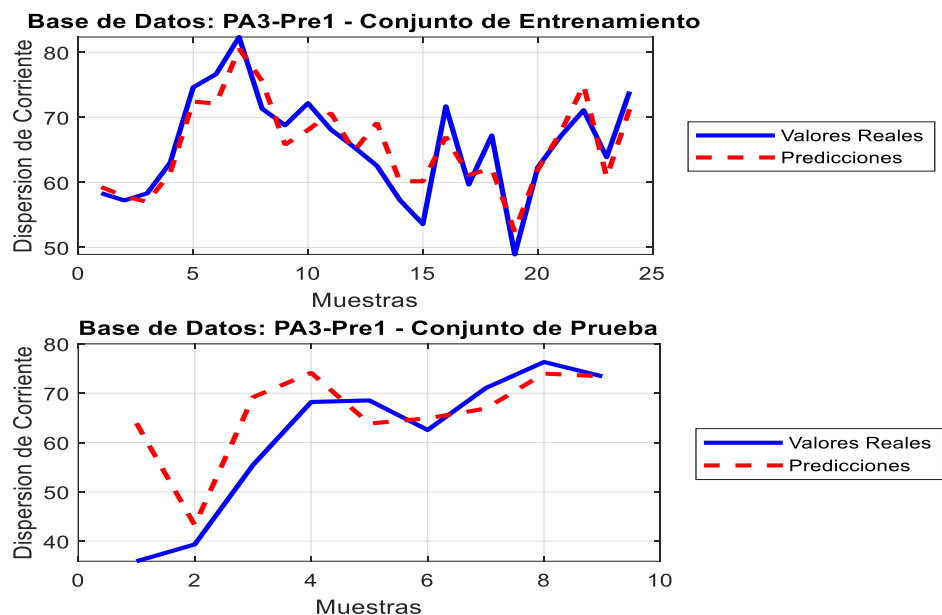


Fig. 5-39: Grafico prediccion v/s valores reales (2 funciones de membresia) PA3-Pre1

Base2:

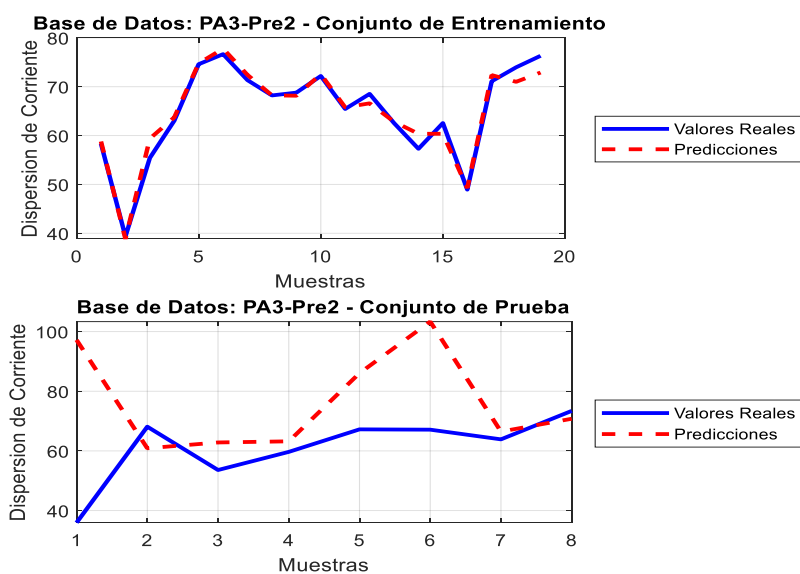


Fig. 5-40: Grafico prediccion v/s valores reales (2 funciones de membresia) PA3-Pre2

Base3:

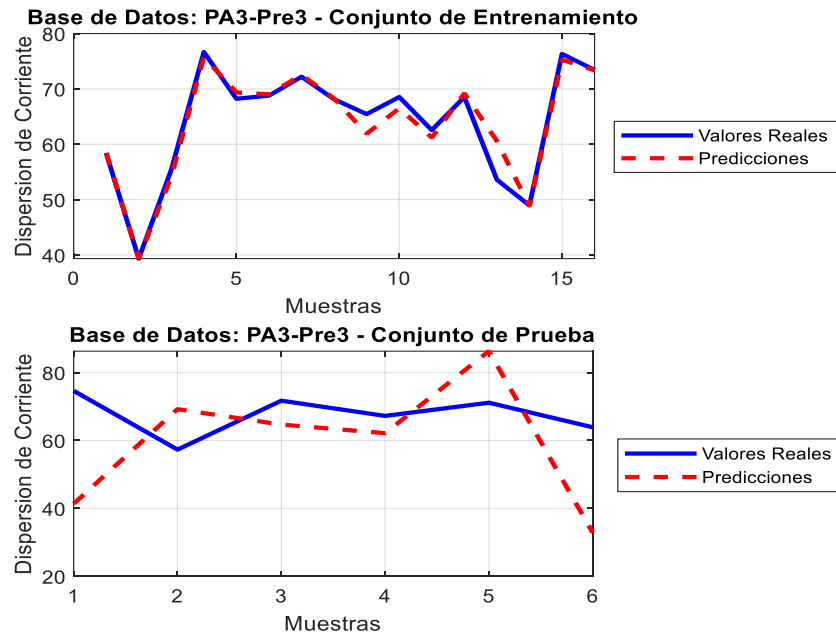


Fig. 5-41: Grafico prediccion v/s valores reales (2 funciones de membresia) PA3-Pre3

Base4:

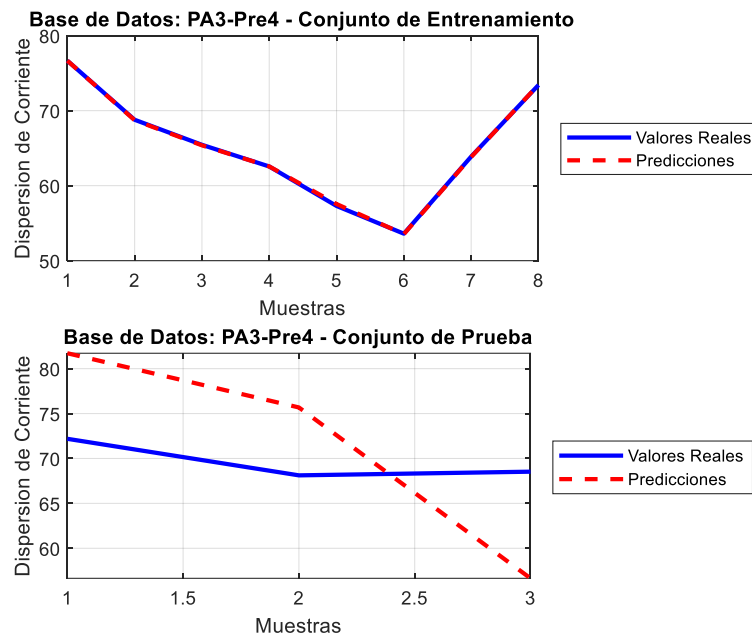


Fig. 5-42: Grafico prediccion v/s valores reales (2 funciones de membresia) PA3-Pre4

Análisis:

Se observó una mejora del 45% en la disminución del RMSE en los conjuntos de prueba a medida que se disminuía la cantidad de Funciones de membresía. Sin embargo, la mejora en el conjunto de entrenamiento es limitada debido a la escasa cantidad de datos disponibles en las bases. Utilizando redes ANFIS y sus funciones de membresía, se logró un ajuste más preciso en los datos.

En MATLAB, se utilizó un objetivo de error (error goal) de 4. Este valor establece el nivel de precisión deseado para el modelo. Es posible reducir el error tanto en el conjunto de prueba como en el de entrenamiento disminuyendo la cantidad de funciones de membresía y aumentando la cantidad de datos.

Errores Obtenidos del Modelo

Para el siguiente modelo se obtuvieron los siguientes errores de entrenamiento y Test
Conjunto de datos 1:

Tabla 7: Comparación error de Entrenamiento y Prueba Red ANFIS para Conjunto 1

Funciones de Membresia	Base de dato	PA4-Pre1	PA4-Pre2	PA4-Pre3	PA4-Pre4
2	RMSE de Entrenamiento	3,995	4,026	4,065	2,418
	RMSE de Prueba	8,892	10,820	12,920	20,770
3	RMSE de Entrenamiento	4,350	4,398	2,571	2,580
	RMSE de Prueba	12,555	15,800	13,230	21,407
4	RMSE de Entrenamiento	3,514	3,049	2,629	1,277
	RMSE de Prueba	13,550	23,466	29,465	64,092
5	RMSE de Entrenamiento	2,590	2,226	2,216	0,787
	RMSE de Prueba	31,403	37,214	54,330	61,264
6	RMSE de Entrenamiento	3,041	2,372	2,553	1,856
	RMSE de Prueba	38,894	34,306	42,327	76,811

Conjunto de datos 2:

Tabla 8: Comparación error de Entrenamiento y Prueba Red ANFIS para Conjunto 2

Funciones de Membresía	Base de dato	PA3-Pre1	PA3-Pre2	PA3-Pre3	PA3-Pre4
2	RMSE de Entrenamiento	3,999	4,000	3,992	3,692
	RMSE de Prueba	3,616	14,807	7,375	10,060
3	RMSE de Entrenamiento	3,999	4,216	3,714	2,948
	RMSE de Prueba	24,817	6,792	26,528	45,754
4	RMSE de Entrenamiento	3,779	3,711	3,562	3,872
	RMSE de Prueba	35,786	32,780	57,700	67,887
5	RMSE de Entrenamiento	4,105	3,487	3,478	1,767
	RMSE de Prueba	37,616	52,801	50,959	70,157
6	RMSE de Entrenamiento	2,174	2,820	2,687	1,953
	RMSE de Prueba	42,524	60,823	56,822	67,237

Analisis:

Claro, aquí tienes una versión mejorada y más profunda del análisis: Para el análisis de los errores en el Conjunto 1 de Redes ANFIS, se observa que el RMSE de entrenamiento es bastante bajo en todas las bases de datos, independientemente de sus funciones de membresía. Esto se debe al objetivo de error (error goal) utilizado en el modelo ANFIS, que permitió que el modelo buscara un RMSE cercano a 4. Además, se nota que el RMSE de entrenamiento disminuye para las bases PA4-Pre1, Pre2 y Pre3. Sin embargo, en PA4-Pre4, al aumentar el número de funciones de membresía, el RMSE también aumenta.

En cuanto al RMSE de prueba, se incrementa aproximadamente un 50% a medida que se aumenta la cantidad de funciones de membresía. No fue posible reducir el RMSE de prueba debido a la falta de un objetivo que pudiera controlarlo efectivamente. Al analizar tanto el Conjunto 1 como

el Conjunto 2, se observa que un mayor número de funciones de membresía en el sistema está asociado a un mayor error de prueba. Esto se debe a que la cantidad de datos de entrenamiento es insuficiente en comparación con los datos de prueba.

Este desequilibrio conduce a un problema de sobreajuste, donde el sistema muestra un bajo error en el entrenamiento pero un alto error en los datos de prueba. El sobreajuste, o overfitting, es evidente en todos los errores reportados. Este fenómeno se manifiesta cuando el modelo se ajusta demasiado a los datos de entrenamiento, capturando tanto las relaciones válidas como el ruido. Como resultado, el modelo pierde capacidad de generalización, lo que se refleja en un rendimiento pobre en los datos de prueba.

Tabla 9: Comparación error promedio de Entrenamiento .

Funciones de Membresía	2	3	4	5	6
Tipo de error	RMSE de Entrenamiento	RMSE de Entrenamiento	RMSE de Entrenamiento	RMSE de Entrenamiento	RMSE de Entrenamiento
Promedios conjuntos de datos 1	3,92075	3,719	3,731	3,209	2,409
Promedios conjuntos de datos 2	3,626	3,475	2,617	1,955	2,456

Tabla 10: Comparación error promedio de Prueba.

Funciones de Membresía	2	3	4	5	6
Tipo de error	RMSE de Prueba	RMSE de Prueba	RMSE de Prueba	RMSE de Prueba	RMSE de Prueba
Promedios conjunto de datos 1	8,965	25,973	48,538	52,883	34,0896875
Promedios conjunto de datos 2	13,350	15,748	32,643	46,053	48,084

Análisis del Error promedio de Entrenamiento

Las diferencias son relativamente pequeñas con 2 y 3 funciones de membresía, sugiriendo que ambos conjuntos de datos se comportan de manera similar con estas configuraciones. Con 4 y 5 funciones de membresía, las diferencias porcentuales son mucho mayores (cerca del 30% y 39% respectivamente), indicando que el conjunto de datos 2 se beneficia significativamente más del aumento de funciones de membresía en términos de RMSE de entrenamiento. Con 6 funciones de membresía, la diferencia porcentual es mínima (cerca del 2%), sugiriendo que ambos conjuntos de datos alcanzan un punto de equilibrio en cuanto a la efectividad de las funciones de membresía adicionales.

Análisis del Error promedio de Prueba

Con 2 funciones de membresía, el RMSE de prueba del conjunto de datos 2 es aproximadamente 49% mayor que el del conjunto de datos 1, sugiriendo que el modelo ANFIS se desempeña mejor con el conjunto de datos 1 bajo esta configuración. Con 3 y 4 funciones de membresía, las diferencias porcentuales disminuyen (cerca de 39% y 33% respectivamente), indicando que el conjunto de datos 2 comienza a beneficiarse más de un mayor número de funciones de membresía en términos de RMSE de prueba. Con 5 funciones de membresía, la diferencia se reduce aún más a cerca del 13%, lo que indica una convergencia en el desempeño entre ambos conjuntos de datos. Con 6 funciones de membresía, la diferencia porcentual vuelve a aumentar a aproximadamente 41%, sugiriendo una vez más que el conjunto de datos 1 se beneficia más en términos de RMSE de prueba con esta configuración.

Capítulo 6. Conclusión

6.1. Sumario

En el presente informe de habilitación profesional, se realizó una revisión del proceso de electroobtención de cátodos de cobre para luego estudiar de manera estadística sus KPI. Por otro lado, se llevó a cabo un estudio estadístico con el fin de analizar el KPI más predominante en el proceso de electroobtención de cátodos de cobre.

Posteriormente, se generaron modelos de predicción utilizando Regresión Lineal, los cuales fueron analizados utilizando únicamente dos bases de datos diferentes. Además, se desarrollaron 10 modelos diferentes en Redes ANFIS, probando diferentes variables de entrada en el sistema y configurándolos con diferentes funciones de membresía para evaluar el error y el comportamiento de las salidas del sistema. Con el objetivo de comparar ambos modelos de Red ANFIS entre sí, se procedió a comparar el mejor modelo de Red ANFIS con la regresión lineal.

6.2. Conclusiones

La literatura siempre destacó tres indicadores como predominantes en las características de una celda de electroobtención de cobre: consumo específico, eficiencia de corriente y voltaje de celda. Sin embargo, estos valores se obtienen al final del proceso, por lo que no podían conocerse inicialmente. En la primera etapa del proyecto, se logro crear una base de datos solida la cual permitió ordenar una gran cantidad de datos; pudiendo así ordenar condiciones normales y anormales de un centro de electroobtención de cobre, adjuntando variables como voltaje de celda, eficiencia de corriente, consumo especifico, corriente, entre otros.

Evalutando mediante estadística descriptiva; análisis de correlaciones específicamente se determinó que la variable predominante es la dispersión de corriente ya que esta afecta de manera directa en el rendimiento debido a que la misma esta correlacionada con el resto de las variables.

Posteriormente, se desarrollaron dos propuestas de modelos que trabajaban con las dispersiones de corriente presentes en una nave de electroobtención de cobre. Estos modelos fueron diseñados en Matlab y se presentan como dos modelos basados en regresión lineal y diez modelos de redes ANFIS, divididos en dos grupos de cinco, cada uno representado por diferentes datos. Los modelos de redes ANFIS se compararon utilizando diferentes funciones de membresía (2, 3, 4, 5 y 6), lo que permitió obtener una perspectiva amplia de resultados.

Los modelos que obtenían mejores resultados eran aquellos configurados con la menor cantidad de funciones de membresía, evaluados mediante el error RMSE en los datos de entrenamiento y prueba de cada modelo. Sin embargo, los modelos utilizados sufrieron de sobreajuste debido a la falta de una cantidad mínima de datos para un buen funcionamiento del sistema. Como se evidencia en el capítulo 5 de los resultados, el error de prueba es significativamente alto debido a la escasez de datos para realizar predicciones, aproximadamente entre 5 y 6 datos.

Los modelos de Red ANFIS obtuvieron mejores resultados debido a que su RMSE de entrenamiento es hasta 4 veces más bajo, lo cual demuestra un mejor ajuste de datos en esa parte. En cambio, el modelo de regresión lineal tenía RMSE de prueba similares al modelo de Red ANFIS. Por lo tanto, desde esa perspectiva general, los modelos de Red ANFIS se ajustan mejor a las necesidades de este proyecto.

6.3. Trabajos Futuros

Entre los trabajos futuros que podrían continuar con este trabajo son:

- Generar un análisis estadístico con una base de datos mucho mayor a la presente en este Proyecto ya que al hacer la limpieza de datos se genera una disminución drástica de los datos pertenecientes a las variables de estudio.
- Este punto está relacionado con el punto anterior; es trabajar un modelo con mayor cantidad de base de datos, debido a que en ambos modelos se vio el error muy grande en los datos de prueba, lo cual no permitió realizar un limpio para los modelos, generando así modelos sin las salidas esperadas.
- Crear modelos basados en el tiempo, con el fin de poder analizar el comportamiento los datos operacionales de una celda de electro obtención de cobte.

Capítulo 7. Referencia

- [1] M. Tucker, C. Dorfling, M. Tadie. (2021). *Investigating an approach to parameter fitting for the development of a semi-empirical electrowinning model*. Stellenbosch University.
- [2] Mark E. Sclesinger Matthew J.King Kathryn C.Sole William G. Davenport. (2011). *Extractive Metallurgy of Copper 5ta Edition*. ELSERVIER.
- [3] Eduardo Wiechmann F.(2010) Evaluación Industrial Tecnología Optibar (IF)
- [4] Anibal S. Morales¹ and Leonardo G. Aburto².(2015) "Modeling and Optimization of the Specific Energy Consumption of Copper EW Tankhouse Cells employing Finite Elements Method Approach"
- [5] Jang, J. S. (1993) "ANFIS Adaptive Network based Fuzzy Inference Systems". IEEE Transactions on Systems, Man & Cybernetics Vol. 23, pp 665-685
- [6] Aros, N (2005) "Control Predictivo no lineal de un horno rotatorio para la producción de carbon activado mediante ANFIS" (UFRO).
- [7] Mora Flórez, J. J., Pérez Hernández, L. P., & Pérez Londoño, S. M. (2006). Utilización de redes ANFIS y señales de corriente para localización de la zona de falla en sistemas de distribución de energía eléctrica. *Ingeniería e Investigación*, 26(3), 93-99.
- [8] Montero Granados. R (2016): Modelos de regresión lineal múltiple. Documentos de Trabajo en Economía Aplicada. Universidad de Granada. España.
- [9] Bastos Guerrero, D, Illera, M y Sepúlveda, S. (2020). Sistema adaptativo de inferencia neuro-difusa (ANFIS) para la estimación de la radiación solar global. *Investigación e Innovación en Ingenierías*.

- [10] Cuba, C. B. (2016). Electroobtencion de Cobre Empleando el Metodo Electroanalitico Reactivo. Huancayo - Perú.
- [11] J., M. E.-M.-K.-W. (2011). Extractive Metallurgy of Copper 5ta edition. The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, UK: Elsevier.

