

DIMENSIONAMIENTO DE MICRORREDES AISLADAS A PARTIR DE MODELOS
BASADOS EN REDES NEURONALES

DEICY PAOLA PLATA GUTIÉRREZ

UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FÍSICO-MECÁNICAS
ESCUELA DE INGENIERÍAS ELÉCTRICA, ELECTRÓNICA Y DE
TELECOMUNICACIONES
BUCARAMANGA

2022

DIMENSIONAMIENTO DE MICRORREDES AISLADAS A PARTIR DE MODELOS
BASADOS EN REDES NEURONALES

DEICY PAOLA PLATA GUTIÉRREZ

Trabajo de grado para optar el título de Ingeniera Electricista

Director:

JUAN MANUEL REY LOPEZ

Doctor en Ingeniería Electrónica

Codirectores:

IVÁN EDGARDO JIMÉNEZ VARGAS

Ingeniero Electricista

GERMAN ALFONSO OSMA PINTO

Doctor en Ingeniería

UNIVERSIDAD INDUSTRIAL DE SANTANDER

FACULTAD DE INGENIERÍAS FÍSICO-MECÁNICAS

ESCUELA DE INGENIERÍAS ELÉCTRICA, ELECTRÓNICA Y DE
TELECOMUNICACIONES

BUCARAMANGA

2022

DEDICATORIA

A Dios, a mi madre Chiquinquirá Gutiérrez quien me ha brindado su apoyo incondicional durante toda mi existencia e inicio de este proceso, Prospero Díaz por su labor de padre de crianza y darme a conocer mi alma mater, mis hermanos Leidy y Daniel los motores de mi vida, a Juan Carlos Pabón y Henry Juliam Pineda por su amistad y consejos en el transcurso de estos años, Juan Pablo Moreno por su apoyo incondicional, colaboración en el desarrollo de este trabajo de grado y por su compañía en los momentos más difíciles para llevar a cabo mis propósitos de vida. Finalmente, a todos los que estuvieron temporalmente y a aquellos que no están en este plano terrenal, pero tuvieron la oportunidad de acompañarme en algún momento de la vida.

-Deicy Paola Plata Gutiérrez-

TABLA DE CONTENIDO

INTRODUCCIÓN	16
1. MICRORRED: DEFINICIÓN Y ESTRUCTURA BÁSICA	18
2. DIMENSIONAMIENTO DE MICRORREDES.....	19
2.1. ANÁLISIS DE COSTOS	20
2.2. ANÁLISIS DE CONFIABILIDAD.....	22
2.2.1. LPSP (Loss of Power Supply Probability).....	23
2.2.2. LOLP (Loss of Load Probability).....	23
2.2.3. LOLH (Loss of Load Hour)	24
2.2.4. Nivel de autonomía (LA).....	24
2.2.5. Energía Total Perdida (TEL).....	25
2.2.6. Estado mínimo de carga de la batería (MSOC).....	25
2.3. ESTRATEGIA DE GESTIÓN DE ENERGÍA BASADA EN REGLAS Y DESPACHO	26
2.4. RESTRICCIONES	27
2.4.1. Balance de potencia	27
2.4.3. Límite de potencia del generador diésel.....	28
2.4.4. Carga/descarga de la batería	28
2.4.5. Estado de carga	29
2.4.6. Cantidad de componentes.....	30
2.5. MODELADO DE LOS COMPONENTES.....	31
2.5.1. Turbinas eólicas	31
2.5.2. Sistema fotovoltaico	32
2.5.3. Sistema de Almacenamiento de las Baterías (BESS)	33
2.5.4. Generador diésel.....	35
3. REDES NEURONALES.....	36
3.1. DEFINICIÓN DE RED NEURONAL ARTIFICIAL (RNA).....	36
3.2. COMPONENTES DE UNA RNA	38

3.2.1. Función de entrada.....	39
3.2.2. Función de activación.....	40
3.2.3. Función de salida.	41
3.3. BACKPROPAGATION Y ENTRENAMIENTO DE RNA.....	42
3.3.1. Propagación hacia adelante	43
3.3.2. Propagación hacia atrás.....	43
4. OBJETIVOS.....	43
4.1. OBJETIVO GENERAL	43
4.2. OBJETIVOS ESPECÍFICOS	44
5. CASO DE ESTUDIO.....	44
5.1. ESCENARIO DE ESTUDIO	44
5.2. RECURSOS ENERGÉTICOS PRIMARIOS	45
5.3. PERFIL DE CARGA	46
6. METODOLOGÍA PARA EL DIMENSIONAMIENTO DE MG BASADO EN RNA	
.....	47
6.1. ETAPA PRELIMINAR	49
6.1.1 Definición de Entradas y Salidas del algoritmo de Dimensionamiento ...	49
6.1.2 Definición del Volumen Total de los Datos.....	50
6.1.3 Definición de Criterios para la Creación del Dataset.....	51
6.1.4 Representación gráfica del comportamiento (Análisis de dispersión de	
datos).....	51
6.1.5. Técnicas o métodos para el preprocesamiento de los datos.....	54
6.1.5.1. Método convencional	54
6.1.5.2. Método de decimación o reducción de muestras	62
6.1.5.3. Restricción Crítica de Combinaciones	64
6.2 ETAPA DE ANÁLISIS PRE-IMPLEMENTACIÓN.....	68
6.2.1. Acondicionamiento de los Conjuntos de Datos de Entrenamiento	69
6.2.2. Análisis de Distribución Normal de Training Dataset.....	71

6.2.3. Criterios de diseño de la RNA	74
6.3 ETAPA DE IMPLEMENTACIÓN RNA	75
7. ETAPA DE ANÁLISIS DE RESULTADOS	76
7.1. ANÁLISIS DE LOSS Y ACCURACY.....	77
7.2. ANÁLISIS DE DISPERSIÓN DE DATOS	78
7.3. ANÁLISIS DE DISTRIBUCIÓN NORMAL DE LOS DATOS OBTENIDOS....	83
7.4. ANÁLISIS DEL COEFICIENTE DE DETERMINACIÓN (R2 SCORE)	86
CONCLUSIONES	88
RECOMENDACIONES.....	90
BIBLIOGRAFÍA.....	91
ANEXOS.....	94
ANEXO 1. SEUDOCÓDIGO DEL ALGORITMO DE ENTRENAMIENTO DE LA RNA USADA EN EL TRABAJO DE GRADO.....	94

LISTA DE FIGURAS

Figura 1. Configuración de la microrred para el caso de estudio. Tomado de jiménez vargas .	18
Figura 2. Estrategia para el dimensionamiento de una microrred.	20
Figura 3. Curva de potencia típica de una turbina eólica adaptada de bukar et al.	32
Figura 4. Modelo de la red neuronal implementada.	38
Figura 5. Estructura de una neurona que tiene una entrada y una salida adaptado de ruiz et al.	38
Figura 6. Funciones de activación comúnmente utilizadas adaptado de burgos.	41
Figura 7. Perfiles para irradiancia, temperatura y velocidad del viento para dos días cualquiera. Tomado de jimenez vargas et al.	46
Figura 8. Perfil de carga en por unidad durante 50 horas del período de evaluación. Tomado de jimenez vargas et al.	47
Figura 9. Metodología empleada para el dimensionamiento de mg a partir de modelos basados en rna.	48
Figura 10. Definición de las entradas y salidas del algoritmo de dimensionamiento.	50
Figura 11. Estructura del ds para el preprocesamiento de los datos.	51
Figura 12. Estructura del dataset para el procesamiento de los datos con indicador de posición.	51
Figura 13. Gráfica de dispersión del indicador de confiabilidad lpsp para el 100% de los datos.	52
Figura 14. Gráfica de dispersión del indicador de confiabilidad lolh para el 100% de los datos.	53
Figura 15. Gráfica de dispersión para el indicador económico cost para el 100% de los datos.	53
Figura 16. Estructura general del dataset para los datos de entrenamiento y evaluación empleado en el método convencional.	55

Figura 17. Dataset con 10% de los datos para entrenamiento y 90% de los datos para evaluación.	56
Figura 18. Gráfica de dispersión para el indicador de confiabilidad lpsp para el 10% vs 100% de los datos.	57
Figura 19. Gráfica de dispersión para el indicador de confiabilidad lolh para el 10% vs 100% de los datos.	57
Figura 20. Gráfica de dispersión para el indicador económico cost para el 10% vs 100% de los datos.	58
Figura 21. Dataset con 1% de datos para entrenamiento y 99% de datos para evaluación.	58
Figura 22. Gráfica de dispersión para el indicador de confiabilidad lpsp para el 1% vs 100% de los datos.	59
Figura 23. Gráfica de dispersión para el indicador de confiabilidad lolh para el 1% vs 100% de los datos.	59
Figura 24. Gráfica de dispersión para el indicador económico cost para el 1% vs 100% de los datos.	60
Figura 25. Dataset con 0.1% de datos para entrenamiento y 99.9% de datos para evaluación.	60
Figura 26. Gráfica de dispersión para indicador de confiabilidad lpsp para el 0.1 vs 100% de los datos.	61
<i>Figura 27. Gráfica de dispersión para indicador de confiabilidad lolh para el 0.1 vs el 100% de los datos.</i>	<i>61</i>
Figura 28. Gráfica de dispersión para indicador económico cost para el 0.1% vs 100% de los datos.	62
Figura 29. Estructura general del dataset empleado en el método de decimación o reducción de muestras (saltos de 1000).	63
Figura 30. Gráficas de dispersión para el indicador de confiabilidad lpsp de la decimación de muestras vs 100% de los datos.	63
Figura 31. Gráficas de dispersión para indicador de confiabilidad lolh de la decimación de muestras vs 100% de los datos.	64

Figura 32. Gráficas de dispersión para indicador económico cost de la decimación de muestras vs 100% de los datos.....	64
Figura 33. Estructura general del dataset empleado en el método restricción crítica de combinaciones.....	67
Figura 34. Gráfica de dispersión para el indicador de confiabilidad lp sp con restricción crítica de datos.....	67
Figura 35. Gráfica de dispersión para el indicador de confiabilidad lolh con restricción crítica de datos.....	68
Figura 36. Gráfica de dispersión para el indicador económico cost con restricción crítica de datos.....	68
Figura 37. Gráfica de distribución normal para el indicador de confiabilidad lp sp para el método convencional con diversos porcentajes de datos, método de decimación de muestras y restricción crítica de datos.	73
Figura 38. Gráfica de distribución normal para el indicador de confiabilidad lolh para el método convencional con diversos porcentajes de datos, método de decimación de muestras y restricción crítica de datos	73
Figura 39. Gráfica de distribución normal para el indicador económico cost para el método convencional con diversos porcentajes de datos, método de decimación de muestras y restricción crítica de datos.	74
Figura 40. Arquitectura de la rna implementada.....	76
Figura 41. Gráfica de dispersión para el indicador de confiabilidad lp sp de la rna con datos de entrenamiento vs referencia en el método convencional.	79
Figura 42. Gráfica de dispersión para el indicador de confiabilidad lolh de la rna con datos de entrenamiento vs referencia en el método convencional.	79
Figura 43. Gráfica de dispersión para el indicador económico cost de la rna con datos de entrenamiento vs referencia en el método convencional.....	80
Figura 44. Gráfica de dispersión para el indicador de confiabilidad lp sp de la rna con datos de entrenamiento vs referencia en el método de decimación o reducción de muestras.....	80

Figura 45. Gráfica de dispersión para el indicador de confiabilidad lolh de la rna con datos de entrenamiento vs referencia en el método de decimación o reducción de muestras.	81
Figura 46. Gráfica de dispersión para el indicador económico cost de la rna con datos de entrenamiento vs referencia en el método de decimación o reducción de muestras.....	81
Figura 47. Gráfica de dispersión para el indicador de confiabilidad lp sp de la rna con datos de entrenamiento vs referencia en el método de restricción crítica de combinaciones.	82
Figura 48. Gráfica de dispersión para el indicador de confiabilidad lolh de la rna con datos de entrenamiento vs referencia en el método de restricción crítica de combinaciones.	82
Figura 49. Gráfica de dispersión para el indicador económico cost de la rna con datos de entrenamiento vs referencia en el método de restricción crítica de combinaciones.	83
Figura 50. Gráfica de distribución normal para el indicador de confiabilidad lp sp en los 3 métodos implementados.....	84
Figura 51. Gráfica de distribución normal para el indicador de confiabilidad lolh en los 3 métodos implementados.....	84
Figura 52. Gráfica de distribución normal para el indicador económico cost en los 3 métodos implementados.	85

LISTA DE TABLAS

Tabla 1. Valores máximos y mínimos de los elementos de la mg con la cantidad posible de combinaciones.	50
Tabla 2. Parámetros de las distintas bases de datos para cada uno de los métodos seleccionados.....	71
Tabla 3. Parámetros de la rna implementada en python.	75
Tabla 4. Recopilación de resultados obtenidos en los 3 métodos implementados.	77
Tabla 5. Resultados de σ y \bar{x} obtenidos en los 3 métodos seleccionados.....	86
Tabla 6. Coeficiente de determinación (r2 score) obtenido en los 3 métodos implementados.	87

RESUMEN

TÍTULO: DIMENSIONAMIENTO DE MICRORREDES AISLADAS A PARTIR DE MODELOS BASADOS EN REDES NEURONALES*

AUTOR: DEICY PAOLA PLATA GUTIÉRREZ**

PALABRAS CLAVE: Microrred, Red Neuronal, Dataset, Función de activación, Algoritmo de Dimensionamiento.

DESCRIPCIÓN:

El siguiente Trabajo de Grado tiene como objetivo presentar una alternativa en los procesos que demanda las tareas del dimensionamiento de una Microrred (MG) aislada, basándose en el uso de conceptos de vanguardia como el Machine Learning y el diseño e implementación de Redes Neuronales Artificiales (RNA). Como aspecto novedoso, este documento presenta una propuesta para la reducción de información requerida en el entrenamiento de una RNA sin comprometer la exactitud y desempeño de las variables de salida, describiendo distintos métodos de construcción de conjuntos de datos para ser implementados y analizados posteriormente.

También se describe una metodología que condensa toda la experiencia adquirida identificando las etapas clave para la construcción de una RNA, teniendo como punto de partida un caso de estudio, la identificación de todas las variables técnicas que juegan un papel importante para el entrenamiento de la RNA, como lo son los Elementos Almacenadores de Energía (EAE), los Elementos de Conversión de Energía (ECA), el análisis de costos y los indicadores de confiabilidad.

Dentro de las etapas definidas de la metodología, se propone realizar un análisis de la información obtenida del caso de estudio, presentando 3 alternativas de construcción del conjunto de datos o Dataset requeridos por la RNA en el proceso

de entrenamiento (training) y evaluación (testing) de su desempeño para predecir una salida ante el ingreso de información desconocida fuera de los datos utilizados en la etapa de entrenamiento.

*Dimensionamiento de Microrredes Aisladas a Partir de Modelos Basados en Redes Neuronales

** Facultad de Ingenierías Fisicomecánicas. Escuela de Ingenierías Eléctrica, Electrónica y de Telecomunicaciones. Ingeniería Eléctrica. Director: Juan Manuel Rey López. Doctor en Ingeniería Electrónica. Codirector (es): German Alfonso Osma Pinto. Doctor en ingeniería, Iván Edgardo Jiménez Vargas. Ingeniero Electricista.

ABSTRACT

TITLE: SIZING OF ISOLATED MICROGRIDS USING ARTIFICIAL NEURONAL NETWORKS*

AUTHOR: DEICY PAOLA PLATA GUTIÉRREZ**

KEYWORDS: Microgrid, Neuronal Network, Dataset, Activation Function, Sizing Algorithm.

DESCRIPTION:

The following Graduation Project presents an alternative to the isolated microgrids sizing methodology, based on cutting-edge concepts such as Machine Learning and Artificial Neural Networks (ANN). As a novel aspect, this document presents a proposal for reducing information required in the training of an ANN without compromising the accuracy and performance of the output variables, describing different methods of construction of data sets.

A methodology is also described that condenses all the experience acquired by identifying the key stages for the construction of an ANN, having as a starting point a case study, the identification of all the technical variables that play an essential role in the training of the ANN, such as Energy Storage Elements (ESE), Energy Conversion Elements (ECE), cost analysis, and reliability indicators.

Within the defined stages of the methodology, it is proposed to carry out an analysis of the information obtained from the case study, presenting 3 alternatives for constructing the dataset required by the ANN in the training process and evaluating its performance.

* Sizing of Isolated Microgrids using Artificial Neuronal Networks

**Faculty of Physical-Mechanical Engineering. School of Engineering Electrical, Electronics and Telecommunications. Electrical Engineering. Director: Juan Manuel Rey López. PhD Electronic Engineering. Co-director: German Alfonso Osma Pinto. PhD Engineering, Iván Edgardo Jiménez Vargas. Electrical Engineer.

INTRODUCCIÓN

Según el Departamento de Energía Eléctrica de Estados Unidos¹, las microrredes se definen como un grupo de cargas interconectadas y de recursos energéticos distribuidos operando dentro de límites eléctricos claramente definidos, comportándose como una única entidad controlable con relación a la red; estas redes incluyen a su vez una variedad de fuentes renovables en los que se encuentran los arreglos de paneles fotovoltaicos, turbinas eólicas, y cualquier elemento o tecnología válida dentro de este grupo. De igual modo, se destaca el uso de fuentes de carácter no renovable como es el caso de los generadores de diésel y sistemas de almacenamiento de energía como las baterías, supercondensadores, entre otros.

Las microrredes son una alternativa para el abastecimiento de la demanda local en lugares remotos pertenecientes a zonas no interconectadas (**ZNI**) del territorio nacional, los cuales por su topografía y relieve imposibilitan la instalación de líneas de transmisión; siendo una solución en el aprovechamiento de las energías renovables para la generación, de una forma más ecológica en comparación con los sistemas convencionales². Las microrredes requieren de estrategias de despacho para la coordinación de los flujos de energía de sus componentes, garantizando el óptimo dimensionamiento sin ver afectados el costo y el

¹ OFFICE OF ELECTRICITY DELIVERY AND ENERGY RELIABILITY SMART GRID R&D PROGRAM. *DOE Microgrid Workshop Report*. San Diego, California: DEPARTMENT OF ENERGY U.S.A, 2011. P 1.

² UBILLA, K. JIMÉNEZ ESTÉVEZ, G. A. HERNÁNDEZ, R. REYES CHAMORRO, L. HERNÁNDEZ IRIGOYEN, C. PALMA BEHNKE, R. SEVERINO, B. «Smarts Microgrids as a Solution for Rural Electrification: Ensuring Long-Term Sustainability Through Cadastre and Business Models,» *IEEE Transactions on Sustainable Energy*, vol. 5, nº 4, pp. 1310-1318, 2014

aprovechamiento máximo de los recursos energéticos dispuestos en la zona de instalación.

El enfoque principal de este trabajo de grado es el desarrollo de una herramienta de diseño basada en los conceptos de la inteligencia artificial y machine learning para abordar de forma alternativa la etapa de dimensionamiento, específicamente en la obtención de los principales parámetros de confiabilidad que puede poseer una MG y los costos asociados a las posibles combinaciones entre sus Elementos de conversión de Energía (ECE) y Elementos Almacenadores de Energía (EAE), su mantenimiento y operación.

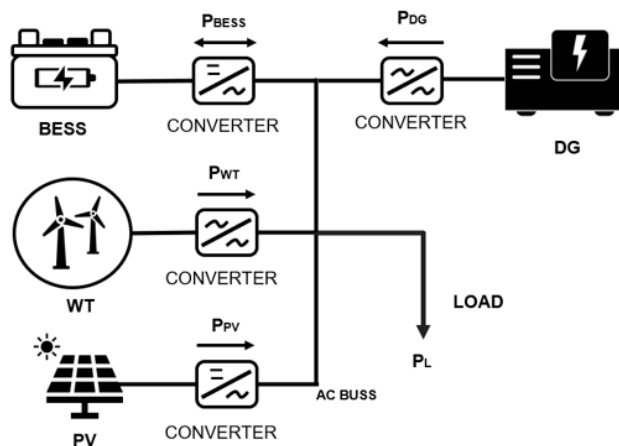
Este documento se compone de 2 grandes secciones distribuidas en 7 capítulos, que permiten describir todo el trabajo desarrollado. Como punto de partida, se realizó un enfoque de formulación del problema de dimensionamiento de microrredes aisladas para identificar las variables de entrada y salida de la red neuronal, teniendo en cuenta condiciones climáticas y de carga específicas para la creación de la base de datos de entrenamiento. En la segunda sección se define la metodología empleada para la obtención de cada uno de los modelos RNA, la descripción de los métodos para la construcción de los conjuntos de datos de entrenamiento y el análisis de los resultados obtenidos.

1. MICRORRED: DEFINICIÓN Y ESTRUCTURA BÁSICA

Las microrredes son sistemas de distribución eléctrica generalmente en BT o MT que funcionan como una sola unidad controlable, integrándose en su estructura cargas y fuentes de energía distribuidas en los que se incluyen elementos de generación, dispositivos de almacenamiento de energía o cargas controlables que pueden funcionar de forma coordinada o controlada conectados a la red o en modo aislado³.

La Figura 1 presenta los elementos que usualmente son empleados en las microrredes y que se tendrán en cuenta para el caso de estudio de este trabajo de grado en la sección del capítulo 5.

Figura 1. Configuración de la microrred para el caso de estudio. Adaptado de Jiménez Vargas ⁴.



³ GONZÁLEZ, N. CUSGÜEN, C. MOJICA NAVA, E. PAVAS, A. «Control Strategies for Power Quality in Rural Microgrids,» UIS ingenierías, vol. 16, n° 2, pp. 93-103, 2017.

⁴ JIMÉNEZ VARGAS, I. OSMA PINTO, G. REY LÓPEZ, J. M. «Sizing of autonomous microgrid considering life cycle emissions,» 2020. P 3.

2. DIMENSIONAMIENTO DE MICRORREDES

El dimensionamiento de microrredes tiene como objetivo identificar la mejor distribución y configuración de cada uno de los elementos ECE y EAE conformando un sistema de energía híbrido⁵. El dimensionamiento también implica resolver el problema que conlleva la resolución de los modelos complejos asociados a la no linealidad y la complejidad que demandan el diseño y modelado de ECE/EAE⁶.

Estos tipos de problemas pueden resolverse definiendo diversos parámetros importantes como la función objetivo, restricciones, gestión energética y el despacho de potencia de cada uno de los componentes de la MG (Ver Figura 2). Además, el dimensionamiento también se ocupa de la identificación de distintos indicadores de fiabilidad que definen el abastecimiento de las cargas al igual que las pérdidas de suministro de energía que imposibilitan la satisfacción de la demanda y el exceso de la capacidad de generación.

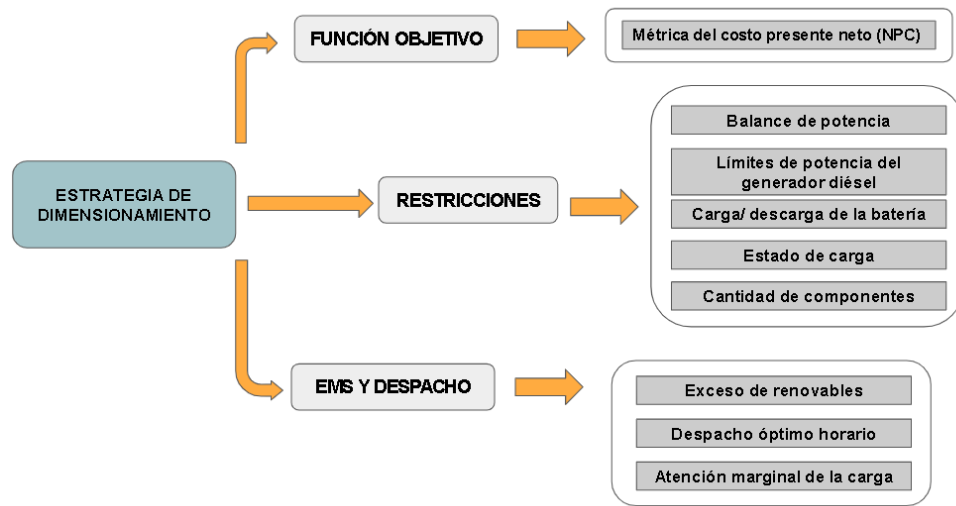
Para la obtención de los indicadores de confiabilidad se requiere definir los posibles factores externos e internos que puedan comprometer el nivel de autonomía con el que contará el sistema. Dentro de estos factores, toma importancia el comportamiento en el tiempo que pueda presentar la demanda de la carga, la distribución de los elementos que componen la MG y la gestión que realizara la red con el exceso de energía generado⁷.

⁵ LIAN, J. ZHANG, Y. MA, C. YANG, Y. CHAIMA, E. A review on recent sizing methodologies of hybrid renewable energy systems, *Energy Conversion and Management*, vol. 199, 2019. P 4-6.

⁶ REY LÓPEZ, J. M. VERGARA, P. P. SOLANO MARTÍNEZ, J. ORDOÑEZ PLATA, G. Design and optimal sizing of microgrids, Springer, pp. 337-367, 2019.

⁷ FALCON A, Andres F. Confiabilidad en Microrredes Eléctricas AC/DC usando control robusto H_{∞} y estrategias de tolerancia a fallas. Universidad Politecnica Salesiana-Sede Quito. 2019. P 2.

Figura 2. Estrategia para el dimensionamiento de una microrred.



2.1. ANÁLISIS DE COSTOS

Este tipo de análisis suele ser esencial en el proceso de dimensionamiento de una microrred, debido a que puede limitar la planeación y desarrollo de un proyecto donde los costos de inversión iniciales, infraestructura, operación y mantenimiento son relevantes para definir su viabilidad.

Para ello se define una función económica objetivo donde se tiene en cuenta un indicador económico conocido como el costo presente neto (*NPC*), el cual define los costos actuales de cada uno de los componentes del sistema, incluyendo los costos generados en su proceso de instalación, mantenimiento y reemplazo; en algunos casos los costos de instalación y mantenimiento se consideran como un porcentaje del costo total de los elementos. Teniendo en cuenta lo antes mencionado, el *NPC* de un sistema se define como:

$$NPC = C_{IC} + C_{OM} + C_{FC} + C_{RE} + C_{SA} \quad (1)$$

Donde:

C_{IC} = Capital inicial.

C_{OM} = Costo de operación y mantenimiento.

C_{FC} = Costo de consumo de combustible.

C_{RE} = Costo de reemplazo de componentes.

C_{SA} = Costo del aprovechamiento de los equipos al finalizar su vida útil.

El cálculo de los costos anuales se puede hacer a través de las siguientes ecuaciones:

$$C_{ICy} = \sum_G N_G PR_G \quad (2)$$

$$C_{OMy} = \sum_{G \neq DG} (N_G OM_G) + N_{DG} OM_{DG} OH_{DG} \quad (3)$$

$$C_{FCy} = PR_d FC_{DG} \quad (4)$$

$$C_{REy} = \sum_G 0.8 N_G PR_G \quad (5)$$

$$C_{SAy} = \sum_G 0.8(1 - DE_G) N_G PR_G \quad (6)$$

Donde:

G = Conjunto de los cuatro recursos energéticos distribuidos (DER).

N_G = Cantidad de cada recurso que componen los DER.

PR_G = Precio de los DER en USD/unidad.

OM_G = Costos de operación y mantenimiento de G en UDS / unidad.

OM_{DG} = Costos de operación y mantenimiento del generador diésel (DG) en USD/h.

OH_{DG} = Tiempo de operación del generador diésel (DG) en horas.

PR_d = Precio del Diésel en USD/L.

DE_G = Porcentaje de degradación lineal de G.

FC_{DG} = Consumo de combustible en litros del generador diésel.

$\overline{PE_{DG}}$ = Factor de penalización de las emisiones promedio en USD/kWh.

P_{DGt} = Potencia generada por el generador diésel en kWh.

Teniendo en cuenta lo anterior, se hace una conversión de valores anuales a presentes a través de la siguiente expresión:

$$NPC = \sum_{y=1}^n \frac{F_y}{(1+i_f)^y} \quad (7)$$

Donde:

NPC = Costo presente neto de un flujo de caja.

F_y = Suma de las ecuaciones para el cálculo de los costos anuales en un año y , teniendo en cuenta los costos de reemplazo de componentes y su aprovechamiento al finalizar su vida útil.

n = Número de años.

Nota: Los costos de reemplazo y de aprovechamiento se multiplicaron por 0.8 con el propósito de considerar el no reemplazo y reventa de la totalidad de componentes del sistema.

2.2. ANÁLISIS DE CONFIABILIDAD

Los indicadores de confiabilidad son utilizados en el dimensionamiento de una microrred como punto de referencia para definir y evaluar la confiabilidad de todo el sistema de distribución de energía eléctrica. Basado en lo anterior, el reto que suponen los nuevos desarrollos de microrredes implica una selección más precisa

de aquellos indicadores que permitan el análisis de confiabilidad con el propósito de asegurar una configuración de elementos óptima; entre los indicadores más utilizados se exponen los citados a continuación:

2.2.1. LPSP (Loss of Power Supply Probability)

Es la probabilidad de no satisfacer la demanda requerida por la carga cuando se presenta un exceso de carga con respecto a la capacidad de generación, expresándose como la relación entre la potencia no suministrada a la carga y su demanda en un determinado período de tiempo, calculándose a través de la expresión matemática que se presenta a continuación:

$$LPSP = \frac{\sum LPS}{\sum P_{LOAD}} \quad (8)$$

Donde:

P_{LOAD} = Demanda horaria de la carga.

LPS = Pérdidas de suministro de energía de la carga.

Siendo LPSP el criterio más empleado en el proceso de dimensionamiento de microrredes y sistemas de generación híbridos.

2.2.2. LOLP (Loss of Load Probability)

Es la probabilidad de que la demanda de un sistema exceda a su capacidad de generación en un determinado período de tiempo, expresándose como la relación entre las pérdidas de potencia suministrada a la carga (LPS) y las horas en el período de análisis N dado en horas como se aprecia a continuación:

$$LPSP = \frac{\sum LPS}{N} \quad (9)$$

Donde:

P_{LOAD} = Demanda hora a hora requerida por la carga.

N = Número de horas del período de tiempo considerado.

2.2.3. LOLH (Loss of Load Hour)

Indica el porcentaje de horas en el cual la demanda del sistema supera la capacidad de generación, calculándose a través de la expresión matemática que se aprecia a continuación:

$$LPSP = \frac{HLPS}{N} * 100\% \quad (10)$$

Donde:

$HLPS$ = Cantidad de horas totales en donde las pérdidas de potencia suministrada a la carga tienen valores positivos.

N = Número de horas del período de tiempo considerado.

2.2.4. Nivel de autonomía (LA)

Indica el porcentaje de días en el que la demanda del sistema no supera la capacidad de generación; debido a su relación con el LOLH se puede expresar a través de la siguiente ecuación:

$$LA = 100\% - LOLH \quad (11)$$

Donde:

$LOLH$ = Porcentaje de horas en donde la demanda del sistema supera su capacidad de generación.

2.2.5. Energía Total Perdida (TEL)

Es el exceso de energía generada y no consumida por la carga, este criterio generalmente es empleado en microrredes que carecen de sistema de almacenamiento, aunque en algunos casos, se utilizan con sistemas de almacenamiento siempre y cuando se considere un abastecimiento de la carga completo y los excedentes generados como pérdidas, se expresa con la siguiente ecuación:

$$TEL = \sum_{t=1}^T (P_G - P_L) = \sum_{t=1}^T P_{LOST}, \quad \text{si } P_G > P_L \quad (12)$$

Donde:

P_G =Potencia disponible para el sistema de generación y almacenamiento.

P_L =Potencia demandada por la carga.

Considerándose solamente valores positivos para las pérdidas de potencia (P_{LOST}), dadas cuando la potencia generada y disponible por el sistema sea mayor al requerido por la carga.

2.2.6. Estado mínimo de carga de la batería (MSOC)

Se garantiza una vida útil más prolongada en las baterías cuando no se presentan descargas profundas y se mantiene un valor superior al estado de carga (SOC) suministrado por el fabricante, en el caso del dimensionamiento de microrredes puede considerarse el estado mínimo de carga como una restricción en las reglas de despacho de los sistemas de gestión de energía (EMS).

2.3. ESTRATEGIA DE GESTIÓN DE ENERGÍA BASADA EN REGLAS Y DESPACHO

Las reglas de despacho en los sistemas de gestión de energía permiten aprovechar la generación de fuentes renovables, además de optimizar el despacho energético de los generadores diésel (DG) y el sistema de almacenamiento de las baterías (BESS), teniendo en cuenta para el caso de estudio que se llevará a cabo posteriormente (en la sección 5) los siguientes criterios o acciones:

1. Disponibilidad de los datos necesarios para tener una solución como es el caso de la cantidad máxima de elementos posibles de la microrred que se encontrará constituida por generadores diésel, baterías, turbinas eólicas y paneles solares.
2. Los cálculos de generación de energía renovable correspondientes a cada uno de los elementos de generación de la microrred, obtenidos a partir de su modelamiento matemático (ver sección 2.53).
3. Cálculo de la diferencia entre la generación con energía renovable y la demanda requerida por la carga.
4. Tener en cuenta para cada instante de tiempo t las siguientes condiciones:
 - Si la generación de energía renovable es mayor que la demanda, el sistema de almacenamiento de las baterías se encontrará cargado siendo la energía no admitida por el sistema de almacenamiento como energía de descarga.
 - Si la generación de energía renovable es menor a la demanda, la potencia despachada por el sistema de almacenamiento de las baterías y el generador diésel se lleva a cabo de tal modo que se minimicen los costos de energía al igual que la degradación de los componentes y la penalización por emisiones.
 - Si la capacidad máxima del sistema al combinar el almacenamiento de las baterías y el generador diésel no es suficiente para satisfacer la demanda de la carga, la potencia de salida del sistema se establecerá en su máximo y se hará cálculo de una hora de carga desatendida.

2.4. RESTRICCIONES

Las restricciones en el proceso de dimensionamiento de microrredes tienen el trabajo de limitar los valores que las variables del diseño puedan adquirir, dándole un sentido físico al problema haciéndolo más preciso y realista. En el caso de estudio abordado en este trabajo de grado (Ver sección 5) se tendrán en consideración cinco (5) tipos de restricciones:

2.4.1. Balance de potencia

En esta restricción se garantiza que la generación de cada uno de los componentes satisfaga la demanda requerida por la carga, aspecto que convierte esta restricción en la más importante de la estrategia de dimensionamiento. Para ello, se iguala la potencia generada por el generador diésel, la potencia absorbida o entregada por el sistema de almacenamiento de las baterías (BESS), la potencia generada por los paneles solares y turbinas eólicas con la potencia demandada por la carga, tal como se aprecia a continuación:

$$P_{gd}(t) + P_{bs}(t) = P_c(t) - (P_{wt}(t) + P_{fv}(t)) \quad (13)$$

Donde:

P_{gd} = Potencia generada por el generador diésel en [kW].

P_{bs} = Potencia entregada o absorbida por la batería.

P_c = Potencia demandada por la carga.

P_{wt} = Potencia generada por la turbina eólica.

P_{fv} = Potencia generada por los paneles fotovoltaicos.

Adicionalmente a la restricción anteriormente descrita, se le agrega un indicador de confiabilidad con el fin de mantener el suministro de potencia en la carga, siendo en este caso aplicado el indicador LPSP (Ver sección 2.2.1), fijándose un valor

promedio límite de 5%. El LPSP en conjunto con la primera restricción se puede expresar del siguiente modo:

$$LPSP(t) = 100 * \sum_{t=0}^N \left(\frac{P_c(t) - P_{dg}(t) + P_{bs}(t) + P_{wt}(t) + P_{fv}(t)}{P_c(t)} \right) \quad (14)$$

$$\overline{LPSP(t)} \leq 5\%$$

Donde:

N = Horas del período de análisis

2.4.2. Límite de potencia del generador diésel

El generador diésel posee un límite máximo y mínimo de potencia que puede generarse en un instante de tiempo t , siendo generalmente el límite superior la potencia nominal de la planta y el límite inferior una potencia de 0 [kW], expresándose como una desigualdad se tiene:

$$\underline{P_{gd}} \leq P_{gd}(t) \leq \overline{P_{gd}} \quad (15)$$

Donde:

$\underline{P_{gd}}$ = Límite inferior de la potencia del generador diésel.

$\overline{P_{gd}}$ = Límite superior de la potencia del generador diésel.

2.4.3. Carga/descarga de la batería

El sistema de almacenamiento de las baterías (BESS) tiene un límite máximo correspondiente a la descarga de la batería para un instante de tiempo t y un límite inferior que indica la potencia máxima de carga tomándose para este caso la

potencia de descarga como positiva y la de carga negativa. Al expresarse esta restricción como una desigualdad resulta:

$$\underline{P_{bs}} \leq P_{bs}(t) \leq \overline{P_{bs}} \quad (16)$$

Donde:

$\underline{P_{bs}}$ = Potencia máxima de carga de la batería.

$\overline{P_{bs}}$ = Potencia máxima de descarga de la batería

2.4.4. Estado de carga

El estado de carga o SOC (State of Charge) es un indicador de la energía almacenada en una batería, el cual es expresado en porcentaje y cuyo valor varía de 0 a 100⁸. Con el fin de prolongar la vida útil en las baterías los fabricantes suelen recomendar mantener su carga por encima del 30% o en algunos casos sugieren mantener la carga por debajo del 95% de su capacidad. Para el caso de estudio tratado en la sección 5, el límite inferior será de un 40% mientras que el superior estará fijado en el 100%⁹, expresándose como sigue a continuación:

$$\underline{SOC} \leq SOC(t) \leq \overline{SOC} \quad (17)$$

Donde:

\underline{SOC} = Límite inferior de carga de la batería

\overline{SOC} = Límite superior de carga de la batería

⁸ GUEMBE ZABALETA, J. SAN MARTÍN BIRRIUN, I. PASCUAL MIQUELEIZ, J. M. «Cálculo del estado de carga en baterías de plomo-ácido: diseño y validación experimental,» Pamplona, 2016, p. 9.

⁹ JIMÉNEZ VARGAS, I. OSMA PINTO, G. REY LÓPEZ, J. M. «Sizing of autonomous microgrid considering life cycle emissions,» 2020. P 5.

También se incluye una restricción en donde se considera la energía presente en la batería en un instante determinado de tiempo $E_{bs}(t)$, expresándose matemáticamente como sigue:

$$E_{bs}(t) = E_{bs}(t - 1) - P_{bs}(t - 1)\Delta t \quad (18)$$

Donde:

$E_{bs}(t - 1)$ =Energía presente en la batería en un instante de tiempo anterior.

$P_{bs}(t - 1)$ =Potencia en la batería en un intervalo de tiempo anterior.

Δt = Duración de un instante de tiempo t en horas.

2.4.5. Cantidad de componentes

En esta restricción se especifica la cantidad de componentes en donde se establece que estos sean mayores a cero para que el algoritmo de solución no arroje valores negativos¹⁰. Esta restricción se expresa como sigue:

$$C_{fv} \geq 0$$

$$C_{wt} \geq 0$$

$$C_{gd} \geq 0$$

$$C_{bs} \geq 0$$

Con el cumplimiento de esta restricción se garantiza que la cantidad de componentes dentro del conjunto de combinaciones sea un valor positivo.

¹⁰ JIMÉNEZ VARGAS, I. OSMA PINTO, G. REY LÓPEZ, J. M. «Sizing of autonomous microgrid considering life cycle emissions,» 2020. P 7.

2.5. MODELADO DE LOS COMPONENTES

En esta subsección, se apreciarán los modelos matemáticos correspondientes a cada uno de los componentes de la microrred, los cuales generan energía eléctrica a partir del aprovechamiento de los recursos energéticos de la zona de instalación, como es el caso de las turbinas eólicas y sistemas fotovoltaicos. También se hará una breve revisión del modelo utilizado para el sistema de almacenamiento de las baterías (BESS) y la relación entre la potencia suministrada y el consumo de combustible del generador diésel.

2.5.1. Turbinas eólicas

Para determinar la potencia generada por una turbina eólica (Ver **Figura 3**), se debe tener en cuenta la velocidad del viento con su modelo cúbico, el cual tiene velocidades de corte inferiores y superiores¹¹. El modelo matemático correspondiente a este elemento es el siguiente:

$$P_{WT} = \begin{cases} 0, & V < V_{c_{in}} \wedge V > V_{c_{out}} \\ N_{WT}NC_{WT} \left(\frac{V(t)^3 - V_{c_{in}}^3}{V_R^3 - V_{c_{in}}^3} \right), & V_{c_{in}} \leq V < V_R \\ N_{WT}NC_{WT}, & V_R \leq V < V_{c_{out}} \end{cases} \quad (19)$$

Donde:

NC_{WT} = Potencia nominal de la turbina eólica en [kW].

V_R = Velocidad del viento nominal en [m/s].

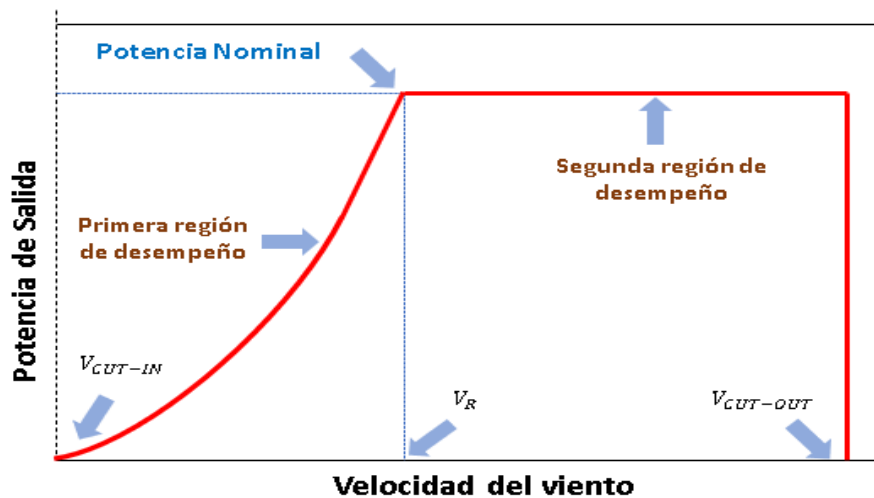
¹¹ BUKAR, A. L. TAN, C.W. LAU, K. Y. «Optimal sizing of an autonomous photovoltaic/wind/battery/diesel generator microgrid using grasshopper optimization algorithm,» *Solar Energy*, vol. 188, pp. 685-696, 2019.

$V(t)$ = Velocidad de la turbina en un instante de tiempo t .

$V_{c_{in}}$ = Velocidad de corte inferior en [m/s].

$V_{c_{out}}$ = Velocidad de corte superior en [m/s].

Figura 3. Curva de potencia típica de una turbina eólica adaptada de Bukar et al ¹².



2.5.2. Sistema fotovoltaico

Se seleccionó el modelo mostrado por los autores Bukar et al¹³, en el cual se consideran los parámetros brindados por el fabricante, la temperatura ambiente y la irradiancia para el respectivo cálculo de la potencia de salida, como sigue a continuación:

¹² *Ibíd.*, p. 685-696

¹³ BUKAR, A. L. TAN, C.W. LAU, K. Y. «Optimal sizing of an autonomous photovoltaic/wind/battery/diesel generator microgrid using grasshopper optimization algorithm,» *Solar Energy*, vol. 188, pp. 685-696, 2019.

$$P_{pv}(t) = N_{PV}\eta_{PV}A_{PV}I(t) \left(1 + \left(\frac{k_p}{100} \right) (T_c(t) - T_{STC}) \right) \quad (20)$$

Donde:

$P_{pv}(t)$ = Potencia generada en [kW] en un instante de tiempo t.

$I(t)$ = Irradiancia en [kW/m²] en un instante de tiempo de t.

$T_c(t)$ = Temperatura de la celda en °C en un instante de tiempo t.

N_{PV} = Número de paneles fotovoltaicos.

η_{PV} = Eficiencia.

A_{PV} = Área efectiva del panel fotovoltaico en m².

k_p = Constante de la temperatura en $\frac{1}{^\circ\text{C}}$.

T_{STC} = Temperatura del panel en °C.

2.5.3. Sistema de Almacenamiento de las Baterías (BESS)

Teniendo en cuenta la estrategia de gestión basada en reglas (EMS) definida anteriormente en la subsección 2.3, la potencia de entrada y/o salida en el sistema de almacenamiento de baterías depende del estado de carga, los límites de potencia y la etapa de funcionamiento de la EMS¹⁴.

Por lo tanto, el BESS funciona durante las siguientes situaciones:

Exceso de renovables: Si hay un exceso de generación producido por las fuentes de energía renovables tal que supere la demanda requerida por la carga, la energía sobrante cargará las baterías, siendo el único caso en donde el BESS almacene

¹⁴ JIMÉNEZ VARGAS, I. OSMA PINTO, G. REY LÓPEZ, J. M. «Sizing of autonomous microgrid considering life cycle emissions,» 2020. P 7.

energía. La inyección de potencia estará limitada por el estado de carga de las baterías (SOC) y la potencia máxima de carga de las baterías¹⁵.

$$P_B(t) = -(P_{WT}(t) + P_{FV}(t) - P_L(t)) \quad (21)$$

Donde:

$P_B(t)$ = Potencia de las baterías en un instante de tiempo t.

$P_{WT}(t)$ = Potencia de las turbinas eólicas en instante de tiempo t.

$P_{FV}(t)$ = Potencia generada por los paneles fotovoltaicos en un instante de tiempo t.

$P_L(t)$ = Potencia demandada por la carga en un instante de tiempo t.

Defecto de renovables: Cuando la generación de energía renovable sea insuficiente para satisfacer la demanda requerida por la carga, se despachará energía del BESS en conjunto con el generador diésel con el fin de minimizar los costos de generación en ese instante de tiempo. El costo de generar con baterías se encontrará asociado con su degradación¹⁶, por lo tanto:

$$C_{BESS}(t) = \frac{PR_{BESS}}{ET_{BESS}} P_B(t), P_B(t) > 0 \quad (22)$$

Donde:

$C_{BESS}(t)$ = Costo de generación con el sistema de almacenamiento de las baterías.

¹⁵ *Ibíd.*, p. 7.

¹⁶ JIMÉNEZ VARGAS, I. OSMA PINTO, G. REY LÓPEZ, J. M. «Sizing of autonomous microgrid considering life cycle emissions,» 2020. P 7.

PR_{BESS} = Precio del capital en USD.

ET_{BESS} = Rendimiento energético del sistema de almacenamiento de las baterías en [kWh].

La ecuación descrita anteriormente, asigna un precio o tarifa a los kWh entregados por las baterías teniendo en cuenta su vida útil en términos de energía¹⁷. Los costos de generación con el sistema de baterías no consideran los costos de operación y mantenimiento (C_{OM}) o los costos del combustible (C_{FC}) generados por el consumo del generador diésel (C_{FC}) debido a que se encuentran incluidos en los cálculos de los costos de reemplazo de componentes (C_{RE}) y los costos del aprovechamiento de los equipos al finalizar su vida útil (C_{SA}).

2.5.4. Generador diésel

Al igual que el BESS su despacho depende de las EMS¹⁸; el cálculo del consumo de combustible es utilizado para obtener el despacho y el costo del consumo presentados en el generador diésel. La función lineal del costo del consumo de combustible para un instante de tiempo t está dado por la siguiente expresión:

$$FC_{DG}(t) = 0.246P_{DG}(t) + 0.08415CN_{DG} \quad (23)$$

Donde:

$P_{DG}(t)$ =Potencia del generador diésel en un instante de tiempo t .

¹⁷ JIMÉNEZ VARGAS, I. OSMA PINTO, G. REY LÓPEZ, J. M. «Sizing of autonomous microgrid considering life cycle emissions,» 2020. P 7 – 8.

¹⁸ PASQUALINO, J., CABRERA, C., CHAMORRO, M.V. The environmental impacts of folic and solar energy implementation in the Colombian Caribe. 13, 68–75 (2015). <https://doi.org/10.15665/rp.v13i1.361>.

CN_{DG} = Capacidad nominal del generador diésel en [kW].

En el cálculo del despacho del generador diésel también se tiene en cuenta un factor adicional que incluye la degradación presentada por su uso y las penalizaciones impuestas por emisiones de GEI¹⁹. La función de costos del despacho del generador diésel está dada por la siguiente ecuación:

$$C_{DG}(t) = PR_{DG}FC_{DG} + \overline{PE}_{DG}P_{DG}(t) + \frac{PR_{DG}}{ET_{DG}}P_{DG}(t) \quad (24)$$

Donde:

ET_{DG} = Energía total que el generador diésel puede despachar durante toda su vida útil.

PR_{DG} = Precio del generador diésel en USD.

$P_{DG}(t)$ = Potencia del generador diésel en un instante de tiempo t.

FC_{DG} = Consumo de combustible en litros del generador diésel.

\overline{PE}_{DG} = Factor de penalización de las emisiones promedio en USD/kWh.

3. REDES NEURONALES

3.1. DEFINICIÓN DE RED NEURONAL ARTIFICIAL (RNA)

El concepto de Redes Neuronales es un término que lleva desarrollándose por varios años y que actualmente se ha convertido en una parte fundamental que

¹⁹ JIMÉNEZ VARGAS, I. OSMA PINTO, G. REY LÓPEZ, J. M. «Sizing of autonomous microgrid considering life cycle emissions,» 2020.p 7.

abarca una gran cantidad de servicios y productos que son utilizados en el día a día²⁰ (aplicaciones móviles, smartphone, motores de búsqueda en internet, etc.). Ya es casi imposible encontrar una aplicación o campo de la ciencia que no pueda ser aprovechada y explotada por una de ellas (medicina, economía, educación, historia, entre otras). Como definición básica, una red neuronal se compone principalmente de la interconexión de elementos simples (Neuronas Artificiales) que responden a una expresión matemática, los cuales son asociados a un nivel de inteligencia o experticia diferente a los programas basados en una lógica procedimental o lineal²¹.

La diferencia radica en un proceso denominado entrenamiento o *training*, en donde es puesto a prueba el algoritmo que lo compone con información previamente estudiada o seleccionada, que luego será evaluado usando distintos datos que la RNA considera como nueva y que tiene una correlación con el comportamiento de los datos aprendidos anteriormente²².

A continuación, se muestra el esquema básico que posee la RNA (Figura 4), y la estructura la compone una neurona (Figura 5) que será ampliada más adelante en este capítulo.

²⁰ BASOGAIN OLABE, Xabier. Redes neuronales artificiales y sus aplicaciones. Escuela Superior de Ingeniería de Bilbao, EHU. p 1-2.

²¹ *Ibíd.*, p 2.

²² F Izaurieta, C Saavedra. Redes neuronales artificiales. Departamento de Física, Universidad de Concepción Chile. 2000

Figura 4. Modelo de la red neuronal implementada.

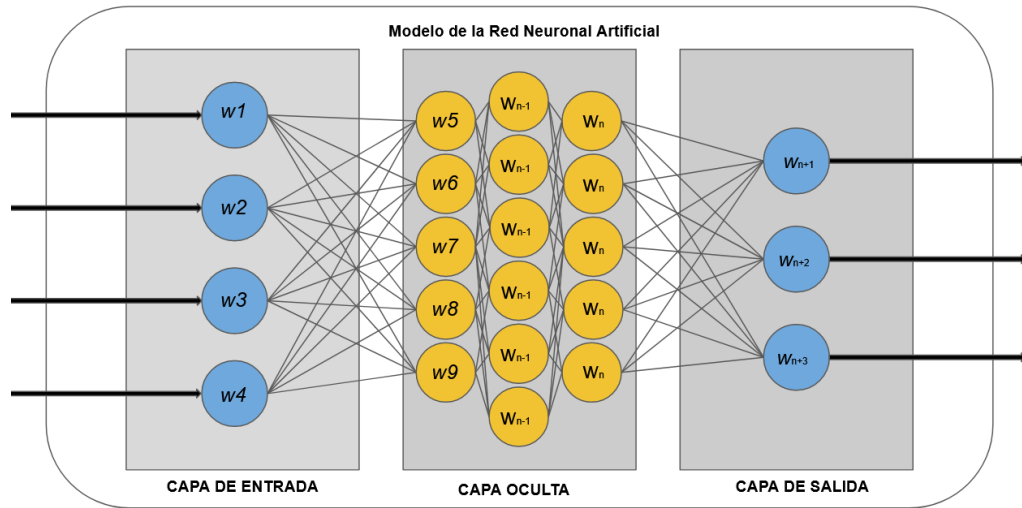
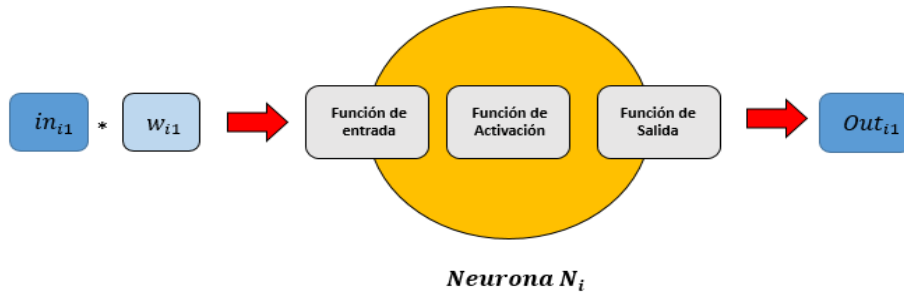


Figura 5. Estructura de una neurona que tiene una entrada y una salida adaptado de RUIZ et al²³.



3.2. COMPONENTES DE UNA RNA

La RNA se conforma por la interconexión y distribución de neuronas a través de 3 capas principales: capa de entrada, capa oculta y capa de salida (Figura 4); la

²³ RUIZ, C. A. BASUALDO, M. S. Redes neuronales: conceptos básicos y aplicaciones, Rosario, México, 2001. p 6.

función de este elemento radica en permitir el paso de los datos por la capa de entrada para ser multiplicados por los pesos ajustados anteriormente durante el proceso de entrenamiento, para finalmente arrojar una respuesta deseada usando las funciones de activación y de salida.²⁴

Los pesos son constantes que se caracterizan por no estar restringidos y por estar sujetos a cambios dependiendo de la influencia presentada en los valores de entrada; esto quiere decir que entre mayor sea la influencia de la entrada, el peso será proporcional en tamaño. (Ver Figura 5).

Las neuronas suelen globalizar los valores de entrada como uno solo a través de una función de entrada, siendo calculada a partir de un vector como se expresa a continuación:

$$input_i = (in_{i1} \cdot w_{i1}) * (in_{i2} \cdot w_{i2}) * \dots (in_{in} \cdot w_{in}) \quad (25)$$

Donde:

* = operador de la función de entrada (sumatoria, máximo o producto)

n = Número de entradas a la neurona N_i

w_i = peso de la neurona N_i

3.2.1. Función de entrada

Las funciones de entrada que se utilizan a menudo en la arquitectura de RNA son las siguientes:

²⁴ MATICH, Damián J. Redes Neuronales: Conceptos Básicos y Aplicaciones. Informática Aplicada a la Ingeniería de Procesos – Orientación I. Universidad Tecnológica Nacional – Facultad Regional Rosario. Marzo 2001. p.4.

- **Sumatoria de las entradas pesadas:** Suma de los valores de entrada a la neurona multiplicado con sus respectivos pesos.
- **Productoria de las entradas pesadas:** Producto de los valores de entrada a la neurona multiplicados con sus respectivos pesos.
- **Máximo de las entradas pesadas:** Toma en consideración el valor de entrada más predominante multiplicado previamente con su peso correspondiente.

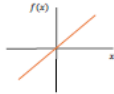
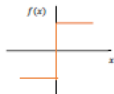
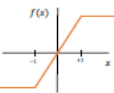
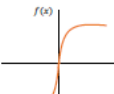
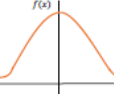
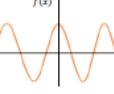
La red neuronal empleada para este trabajo de grado utiliza una función de entrada correspondiente a la **sumatoria de las entradas pesadas**, expresada con la siguiente ecuación:

$$\sum_j(n_{ij}w_{ij}) , \quad \text{con } j = 1,2,3, \dots n \quad (26)$$

3.2.2. Función de activación

Las funciones de activación se encargan del cálculo de la actividad de una neurona, estas se seleccionan dependiendo del objetivo de entrenamiento de la RNA. En el diseño de RNA se utilizan en mayor proporción las siguientes funciones de activación (Ver Figura 6):

Figura 6. Funciones de activación comúnmente utilizadas adaptado de BURGOS²⁵.

	Función	Rango	Gráfica
Identidad	$y = x$	$[-\infty, +\infty]$	
Escalón	$y = \text{sign}(x)$ $y = H(x)$	$\{-1, +1\}$ $\{0, +1\}$	
Lineal a tramos	$y = \begin{cases} -1, \text{si } x < -1 \\ x, \text{si } -1 \leq x \leq 1 \\ +1, \text{si } x > 1 \end{cases}$	$[-1, +1]$	
Sigmoidea	$y = \frac{1}{1 + e^{-x}}$ $y = \text{tgh}(x)$	$[0, +1]$ $[-1, +1]$	
Gaussiana	$y = Ae^{-Bx^2}$	$[0, +1]$	
Sinusoidal	$y = A \text{sen}(\omega x + \varphi)$	$[-1, +1]$	

3.2.3. Función de salida.

La función de salida representa la pieza final de todo el conjunto en una neurona simple, siendo esta necesaria para su funcionamiento al permitir determinar el valor transferido a las neuronas vinculadas. Si la función de activación no se encuentra sobre un umbral determinado proporcionado con anterioridad en la definición de dicha función, la salida no podrá pasar a la neurona subsiguiente perteneciente a la arquitectura de la RNA²⁶.

²⁵ BURGOS, F. J. P. Herramientas en GNU/LINUX para estudiantes universitarios/-Redes neuronales con GNU/LINUX, The Public's Library and Digital Archive, 2003.

²⁶ BURGOS, F. J. P. Herramientas en GNU/LINUX para estudiantes universitarios/-Redes neuronales con GNU/LINUX, The Public's Library and Digital Archive, 2003.

Comúnmente se cuenta con dos funciones de salida:

- **Ninguna:** Considerada como la función más sencilla debido a que en este caso la salida es igual a la entrada conocida también como función identidad.
- **Binaria:** Este tipo de función tiene dos valores: donde el valor de la función de salida es 1 si la función de activación es mayor o igual al umbral y 0 en el caso contrario.

3.3. Backpropagation y entrenamiento de RNA

Las RNA se caracterizan por su capacidad de aprendizaje y adaptabilidad como resultado de un proceso de entrenamiento, cuyo objetivo es conseguir que un conjunto de entradas genere un conjunto de salidas deseadas con un mínimo grado de consistencia; para llevar a cabo este proceso se aplican diversos conjuntos de vectores de entrada para ser ajustados a los pesos que convergen posteriormente hacia un valor deseado reflejado en el vector de salida.²⁷

Un concepto importante que surge en la tarea de entrenar con éxito una RNA es el método de entrenamiento de *backpropagation*, muy usado para redes multicapas de neuronas. Se caracteriza por ser un algoritmo enfocado en el mapeo e identificación de patrones, en la flexibilidad para seleccionar distintos números de capas ocultas y la gran variedad de aplicaciones debido a que no requiere relacionar matemáticamente los patrones de la entrada y salida²⁸. Una red de backpropagation se caracteriza por presentar un método de entrenamiento supervisado, donde cada

²⁷ ANDRADE TEPÁN, E.C. Estudio de los principales tipos de redes neuronales y las herramientas para su aplicación. Universidad Politécnica Salesiana Sede Cuenca. 2013.p 49.

²⁸ BASOGAIN OLABE, Xabier. Redes neuronales artificiales y sus aplicaciones. Escuela Superior de Ingeniería de Bilbao, EHU.p.28

conjunto de datos de entrada está emparejado a un conjunto de datos de salida deseado, disminuyendo en cada iteración la diferencia entre los resultados arrojados por la RNA y los de referencia.

El algoritmo de backpropagation se compone de una fase de propagación hacia adelante y otra hacia atrás, repitiendo cada proceso por cada pareja de entrada y salida que llega a la red²⁹.

3.3.1. Propagación hacia adelante

En esta fase cada conjunto de entrada se propaga desde la capa de entrada hasta la capa de salida, generando una respuesta por parte de la RNA y el error asociado entre la predicción de la red y los datos de referencia³⁰.

3.3.2. Propagación hacia atrás

De acuerdo con los errores cometidos en la iteración del conjunto de datos de entrada de la RNA y los datos de referencia, el algoritmo de backpropagation se encarga de corregir el valor de los pesos entre las conexiones de las neuronas mediante la retropropagación del error desde la capa de salida hacia la capa de entrada a través de las capas ocultas³¹.

4. OBJETIVOS

4.1. OBJETIVO GENERAL

Diseñar un modelo basado en redes neuronales para el dimensionamiento de microrredes aisladas para condiciones climáticas y de carga específicas.

²⁹ *Ibíd.*, p. 28.

³⁰ ANDRADE TEPÁN, E.C. Estudio de los principales tipos de redes neuronales y las herramientas para su aplicación. Universidad Politécnica Salesiana Sede Cuenca. 2013. p 16.

³¹ *Ibíd.*, p.16.

4.2. OBJETIVOS ESPECÍFICOS

1. Seleccionar un enfoque de formulación del problema de dimensionamiento de microrredes aisladas como punto de referencia para identificar las variables de entrada y salida de la red neuronal.
2. Dimensionar un conjunto diverso de microrredes con unas condiciones climáticas y de carga específicas para la creación de la base de datos de entrenamiento.
3. Entrenar el modelo implementado basado en redes neuronales con la información de entrada y salida del conjunto de microrredes dimensionadas.
4. Cuantificar el impacto de la variación en el tamaño del set de datos para entrenamiento en la precisión del modelo.
5. Evaluar diferentes estrategias de selección de datos para el entrenamiento del modelo.

5. CASO DE ESTUDIO

5.1. ESCENARIO DE ESTUDIO

Se definió como caso de estudio una solución de dimensionamiento de una MG aislada (Ver Figura 1) instalada en el Norte del departamento de la Guajira, en las coordenadas geográficas (12.154N, 72.063W), el cual es un sector caracterizado por la abundancia de energía solar y eólica, aspecto que lo convierte en una

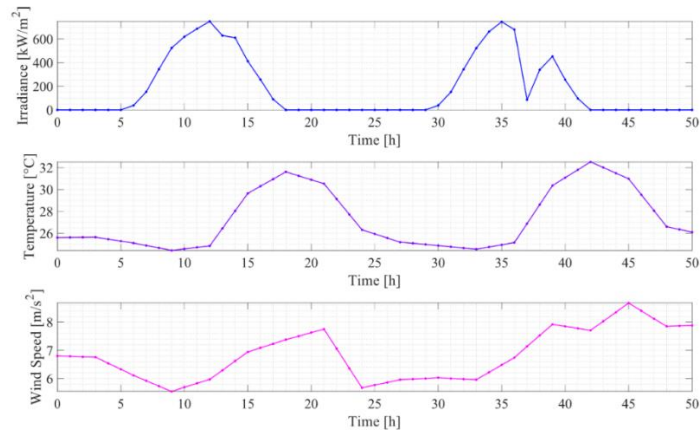
ubicación óptima para el aprovechamiento de estos recursos por parte de los elementos de generación que posee la MG³².

5.2. RECURSOS ENERGÉTICOS PRIMARIOS

En el modelado de los elementos de generación vistos en la sección 2.5, para el caso de las turbinas eólicas y los paneles fotovoltaicos, se hizo uso de información meteorológica obtenida de EU Science Hub, donde se empleó la irradiancia, velocidad del viento y temperatura ambiente correspondientes a las 24 horas del día en un lapso equivalente a 5 años. En la Figura 7 se presenta el perfil de irradiancia, velocidad del viento y temperatura, correspondientes a dos días seleccionados aleatoriamente.

³² JIMÉNEZ VARGAS, I. OSMA PINTO, G. REY LÓPEZ, J. M. «Sizing of autonomous microgrid considering life cycle emissions,» 2020. P 8.

Figura 7. Perfiles para irradiancia, temperatura y velocidad del viento para dos días cualquiera. Tomado de JIMENEZ VARGAS ET AL³³.



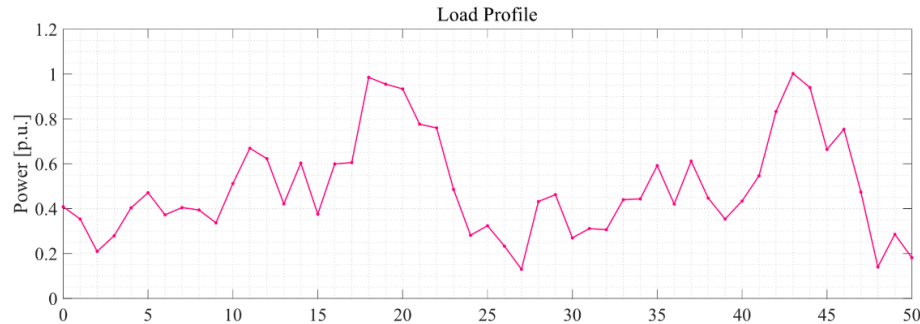
5.3. PERFIL DE CARGA

El perfil de carga se generó a partir de los perfiles típicos residenciales en Colombia. La demanda más alta se da al inicio de la noche (18:00 a 20:00) y es igual 20 kWp, en promedio. El perfil contiene una variabilidad aleatoria del 15%. La Figura 8 presenta el perfil de potencia en por unidad para dos días aleatorios ³⁴.

³³ Pasqualino, J., Cabrera, C., Chamorro, M.V.: The environmental impacts of folic and solar energy implementation in the Colombian Caribe. 13, 68–75 (2015).

³⁴ JIMÉNEZ VARGAS, I. OSMÁ PINTO, G. REY LÓPEZ, J. M. «Sizing of autonomous microgrid considering life cycle emissions,» 2020. p 8

Figura 8. Perfil de carga en por unidad durante 50 horas del período de evaluación. Tomado de JIMENEZ VARGAS ET AL³⁵.



6. METODOLOGÍA PARA EL DIMENSIONAMIENTO DE MG BASADO EN RNA

Para el desarrollo de este trabajo de grado fue de gran importancia resumir en una metodología todas las etapas ejecutadas describiendo de manera precisa y procedimental la solución final, lo que permitió responder satisfactoriamente a cada uno de los objetivos definidos anteriormente.

Dicha metodología se basó fundamentalmente en el uso de los denominados *métodos ágiles*³⁶, cuya finalidad es la de facilitar y mejorar el proceso de planeación, gestión y ejecución de proyectos basados en software. Dentro de los métodos ágiles consultados se tiene la técnica de SCRUM y KANBAN³⁷.

³⁵ *Ibíd.*, p. 8.

³⁶ BHAVSAR, Krunal; SHAH, Vrutik; GOPALAN, Samir. Scrumbanfall: an agile integration of scrum and kanban with waterfall in software engineering. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, 2020, vol. 9, no 4, p. 2075-2084.

³⁷ ANDERSON, David J., et al. A comparative study of Scrum and Kanban approaches on a real case study using simulation. En *International Conference on Agile Software Development*. Springer, Berlin, Heidelberg, 2012. p. 123-137.

Gracias a la experiencia obtenida durante cada subdesarrollo o tarea, fue posible identificar las etapas macro que agrupan a su vez, otros conjuntos de actividades complementarias. Todas las actividades se definen a partir de un objetivo clave y se le asigna una función única en la metodología, facilitando así, la divulgación y la réplica de resultados que demanda el uso de RNA y el dimensionamiento de MG.

Se definieron 4 etapas macro que permiten clasificar todo el conjunto de actividades y procedimientos (Ver Figura 9). La metodología define como etapa preliminar el estudio de la naturaleza de los datos que se usarán para conformar la Base de Datos (**BS**) de la *RNA*. Una segunda etapa se encarga de la preparación y preprocesamiento de la **BS** que usará la *RNA* en el entrenamiento. La siguiente etapa comprende la implementación de cada una de las versiones predefinidas de la *RNA* con sus **BS** asociadas; y como último paso, se requiere que se realice su respectivo análisis de los resultados obtenidos mediante indicadores que permitan evaluar su desempeño.

Figura 9. Metodología empleada para el dimensionamiento de MG a partir de modelos basados en RNA.



6.1. ETAPA PRELIMINAR

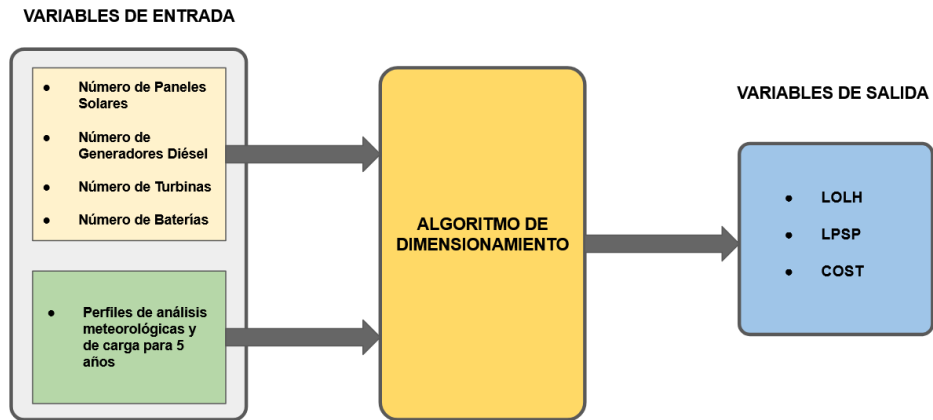
Todo proceso de diseño e implementación de una solución basada en RNA requiere como punto de partida un estudio previo y un análisis de la información o Base de Datos que ejecutará la RNA para su etapa de entrenamiento (*training*) y su etapa de evaluación (*testing*). La RNA tiene como objetivo sustituir el arduo proceso de calcular los modelos matemáticos que definirán la MG y a sus ECE/EAE, así como todas las funciones y ecuaciones para el dimensionamiento. Esta tarea demanda una gran cantidad de recursos computacionales para su procesamiento, resultando en el aumento del tiempo que requiere el equipo de cómputo para llegar a un resultado.

En esta etapa preliminar se agrupan 4 grandes procesos fundamentales que servirán de base para las etapas posteriores. Para la primera parte, se requiere estudiar la naturaleza de los datos recolectados e identificar cuáles son las variables asociadas a la entrada y cuáles a la salida. Luego es necesario acotar el volumen total de datos que se usará en la etapa *training*, e identificar hasta dónde se puede reducir el cálculo de las variables de salida para obtener una RNA confiable sin comprometer la precisión de los resultados.

6.1.1 Definición de Entradas y Salidas del algoritmo de Dimensionamiento

La solución propuesta en el caso de estudio presenta un sistema definido por los componentes de entrada y salida, como se muestra en la Figura 10. El algoritmo está implementado en Matlab, usando como entrada la cantidad máxima de elementos ECE/EAE. También tiene en cuenta los perfiles de análisis meteorológicos y de carga para un determinado tiempo (en este caso para 5 años), necesarios para que el algoritmo de dimensionamiento pueda entregar las variables de salida. Las salidas seleccionadas para la solución comprenden el cálculo de los valores de LPSP, LOLH y COST.

Figura 10. Definición de las entradas y salidas del algoritmo de dimensionamiento.



6.1.2 Definición del Volumen Total de los Datos

Las entradas se definen con las variables N_d (Número máximo de motores diésel), N_w (Número máximo de turbinas eólicas), N_p (Número máximo de paneles solares) y N_b (Número máximo de baterías). En la tabla 1 se muestran los valores máximos y mínimos para cada uno de estos elementos, así como la cantidad de combinaciones posibles entre sí. (Ver Tabla 1)

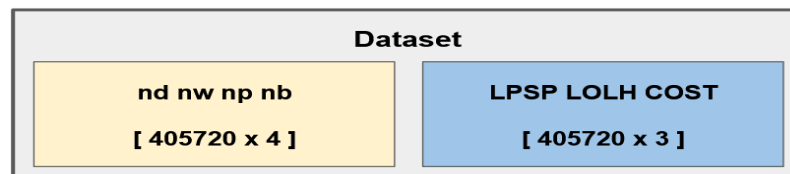
Tabla 1. Valores máximos y mínimos de los elementos de la MG con la cantidad posible de combinaciones.

	Elemento EAE y ECE de la MG			
	N_d	N_w	N_p	N_b
Cantidad Mínima	1	0	0	1
Cantidad Máxima	4	20	160	30
Combinaciones Posibles	405720			

6.1.3 Definición de Criterios para la Creación del Dataset

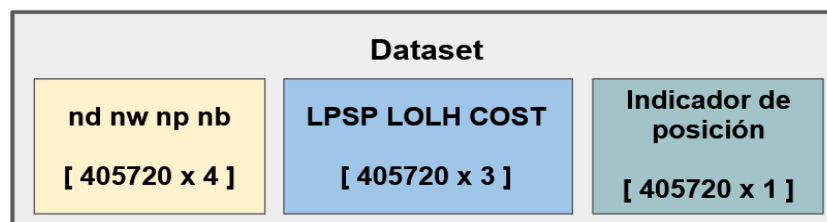
Para el preprocesamiento o tratamiento previo de los datos que se utilizarán para la etapa de entrenamiento y test de la RNA, fue necesario partir de la construcción básica de un *Dataset (DS)*, conformado por todas las entradas y salidas posibles (Ver Figura 11). Estos conjuntos de datos serán un punto de referencia durante todo el proceso de ajuste y análisis de cada una de las técnicas o métodos que se presentarán en las secciones siguientes.

Figura 11. Estructura del DS para el preprocesamiento de los datos.



También fue necesaria la construcción de un segundo Dataset, el cual presenta una nueva columna que tiene asignado un indicador de posición, lo que facilita identificar una combinación en especial para el análisis gráfico de cada uno de los métodos implementados (Ver Figura 12).

Figura 12. Estructura del Dataset para el procesamiento de los datos con indicador de posición.



6.1.4 Representación gráfica del comportamiento (Análisis de dispersión de datos)

El uso de gráficas es muy importante para estudiar en gran medida el comportamiento que lleva consigo cada conjunto de datos. La gráfica de dispersión de datos permite ese tipo de análisis siendo a su vez una herramienta que facilita conocer la posición de cada uno de los conjuntos de datos que conforman los métodos implementados en el proceso de training.

A continuación, se muestran las distintas gráficas obtenidas (Figura 13, Figura 14 y Figura 15) utilizando la totalidad de los datos que luego serán asignados para los conjuntos de entrenamiento y evaluación de cada uno de los métodos propuestos. El resultado obtenido en esta etapa de la metodología será utilizado como punto de referencia para poder analizar los cambios que presentan los métodos propuestos en capítulos posteriores.

Figura 13. Gráfica de dispersión del indicador de confiabilidad LPSP para el 100% de los datos.

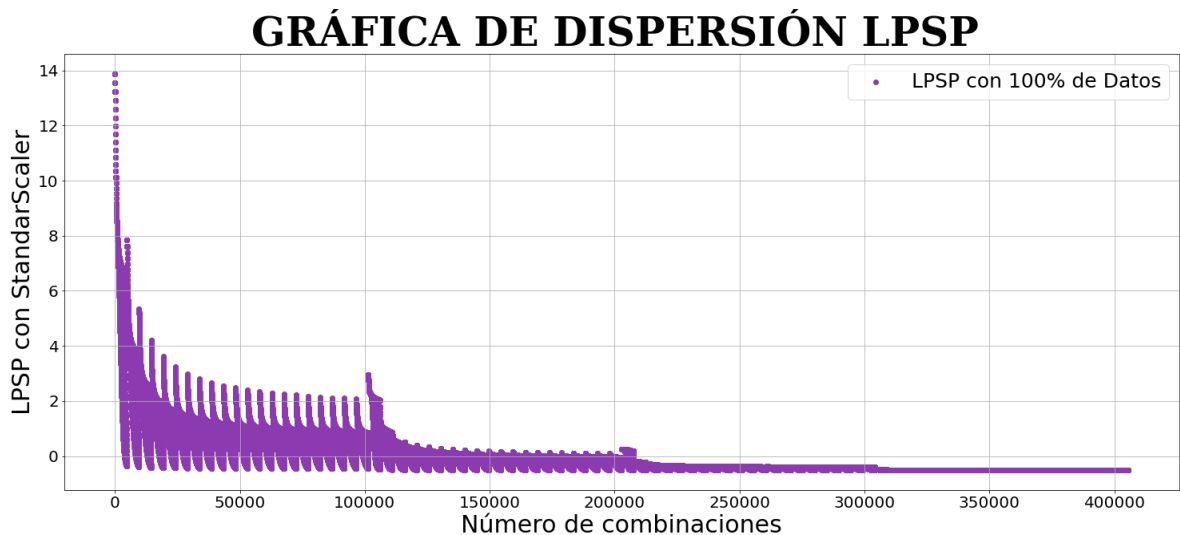


Figura 14. Gráfica de dispersión del indicador de confiabilidad LOLH para el 100% de los datos.

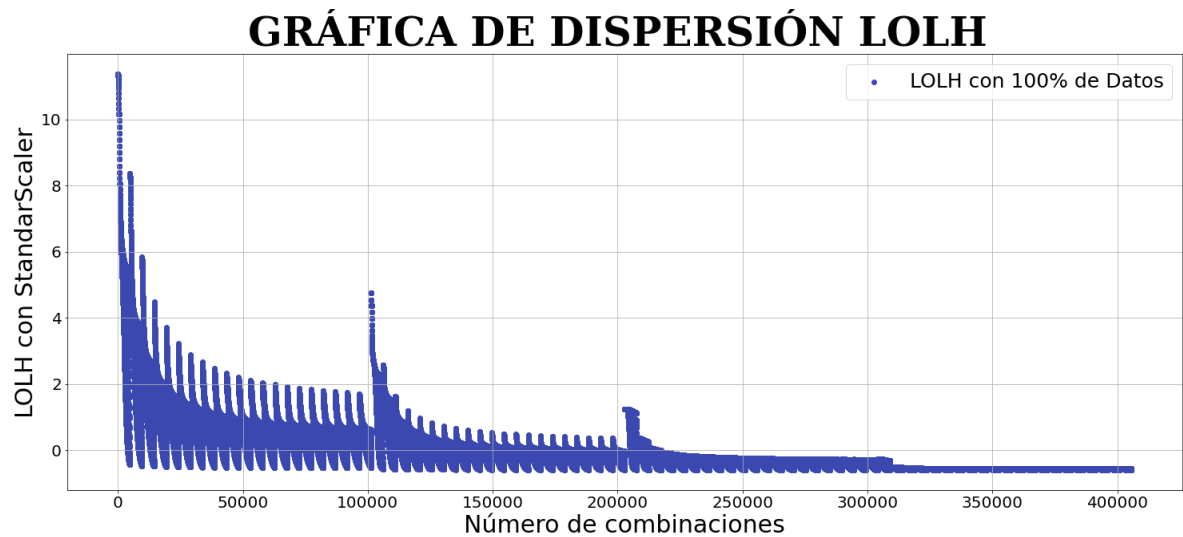
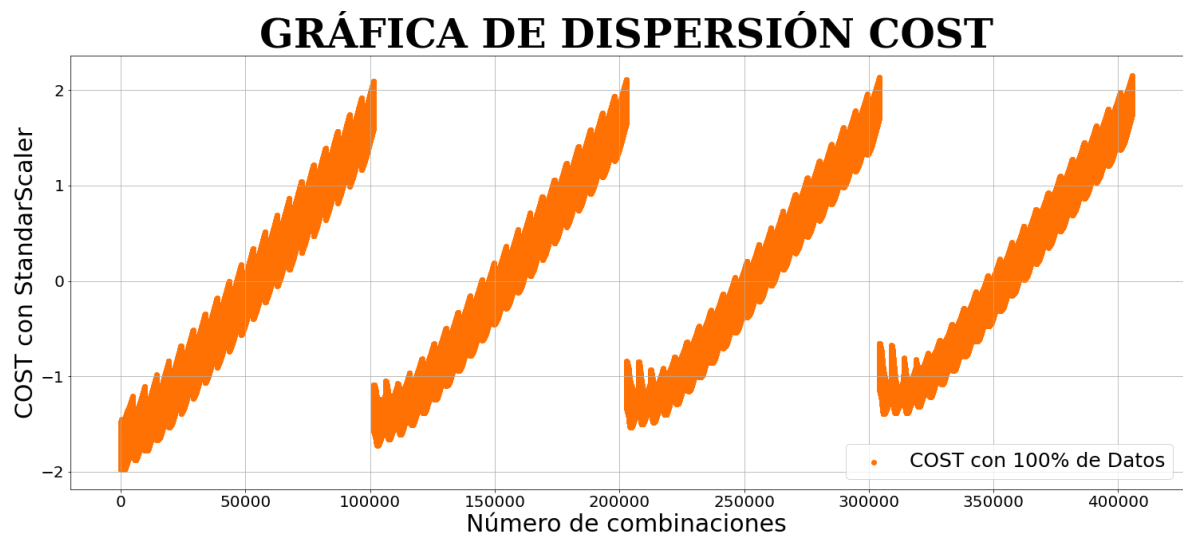


Figura 15. Gráfica de dispersión para el indicador económico COST para el 100% de los datos.



6.1.5. Técnicas o métodos para el preprocesamiento de los datos

En esta sección se presentará a profundidad los tres métodos seleccionados para la construcción de los conjuntos de datos propuestos para identificar la mejor forma de obtener una RNA entrenada para emplear la menor cantidad de información y a su vez, que asegure una gran precisión en la predicción de resultados bajo la asignación de nuevas entradas.

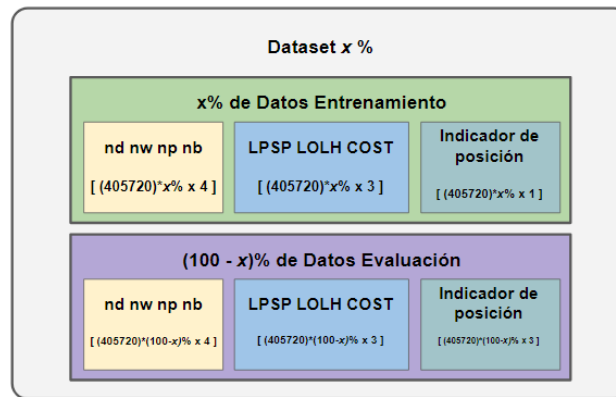
6.1.5.1. Método convencional (MC)

El primer método seleccionado e implementado fue el método que se utiliza comúnmente en casi todos los procesos de entrenamiento de RNA. El uso de separación de datos usando la herramienta o librería *SKLEARN*³⁸. Este módulo permite separar mediante una función *train_test_split*, los datos para la evaluación y test de la RNA. Estos datos son separados aleatoriamente asegurando que la información utilizada para el entrenamiento pueda distribuirse en todo el rango total del Dataset original.

En la Figura 16 se puede observar la representación de la distribución que presenta la función de *train_test_split*. Por una parte, se asigna un valor de *x%* porcentual para el conjunto de datos asignados para el entrenamiento. Los datos restantes son asignados para conformar los datos de evaluación.

³⁸ GÉRON, Aurélien. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow. Concepts, Tools, and Techniques to Build Intelligent Systems. Second Edition. O'Reilly Media Inc. 2019.

Figura 16. Estructura general del Dataset para los datos de entrenamiento y evaluación empleado en el método convencional.



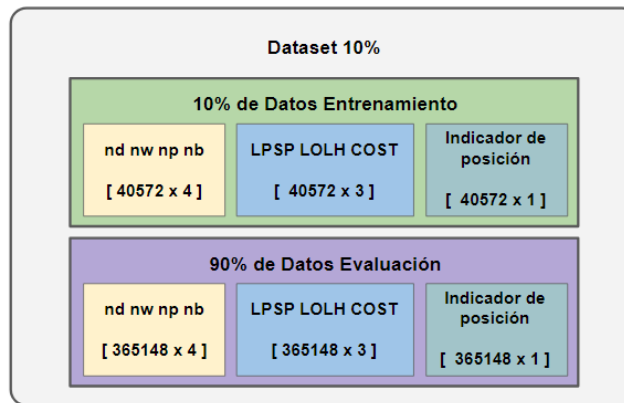
Este método fue seleccionado como punto de partida durante el proceso de la búsqueda de posibles soluciones, permitiendo ahondar más en el desarrollo de una RNA adaptada adecuadamente en el campo de las microrredes³⁹. Basado en la premisa anterior se propuso la creación de 3 Dataset diferentes, asignando una cantidad porcentual que representa el 10%, el 1% y el 0.1% de los datos totales, destinando los datos restantes para la evaluación del desempeño.

I. DATASET 10%

En la Figura 17 se muestra el Dataset obtenido a partir del porcentaje definido en 10%, dejando el 90% restante para la etapa de evaluación de la RNA.

³⁹ SALEH, Hyatt. Machine Learning Fundamentals. Packt Publishing Editorial. 2018.

Figura 17. Dataset con 10% de los datos para entrenamiento y 90% de los datos para evaluación.



En la Figura 18, Figura 19 y Figura 20 se muestran los comportamientos del conjunto Dataset para el 10% de los datos totales en comparación al Dataset original. (Los puntos a color representan los datos de referencia y en negro los resultados obtenidos para cada porcentaje de datos asignados aleatoriamente)

A continuación, se presentan los resultados mediante la gráfica de dispersión de las variables LPSP (*Figura 18*), LOLH (*Figura 19*) y el valor COST (*Figura 20*) obtenidas por este método:

Figura 18. Gráfica de dispersión para el indicador de confiabilidad LPSP para el 10% vs 100% de los datos.

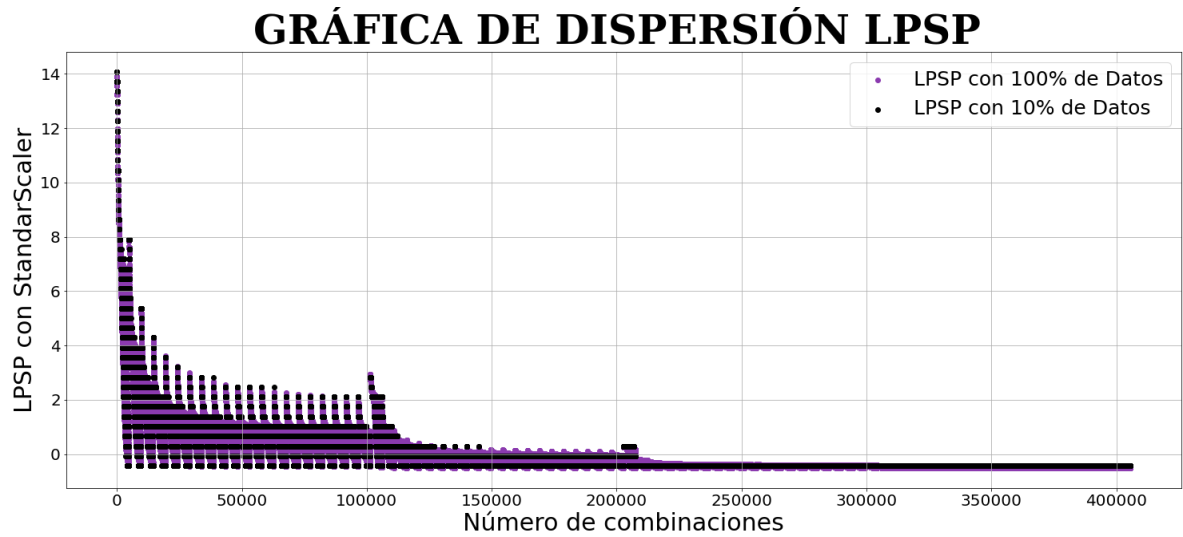


Figura 19. Gráfica de dispersión para el indicador de confiabilidad LOLH para el 10% vs 100% de los datos.

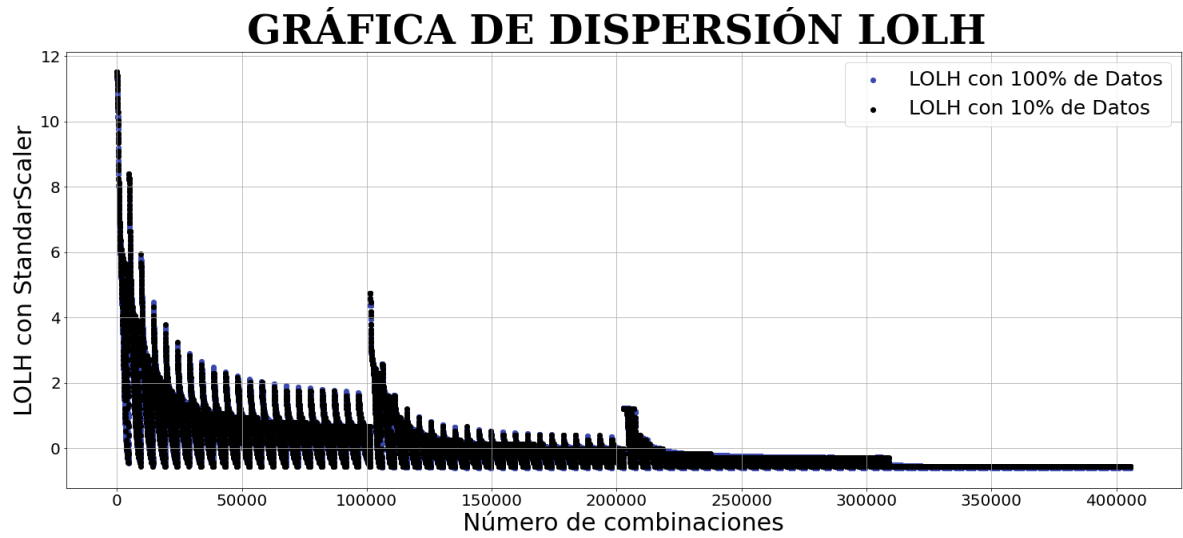
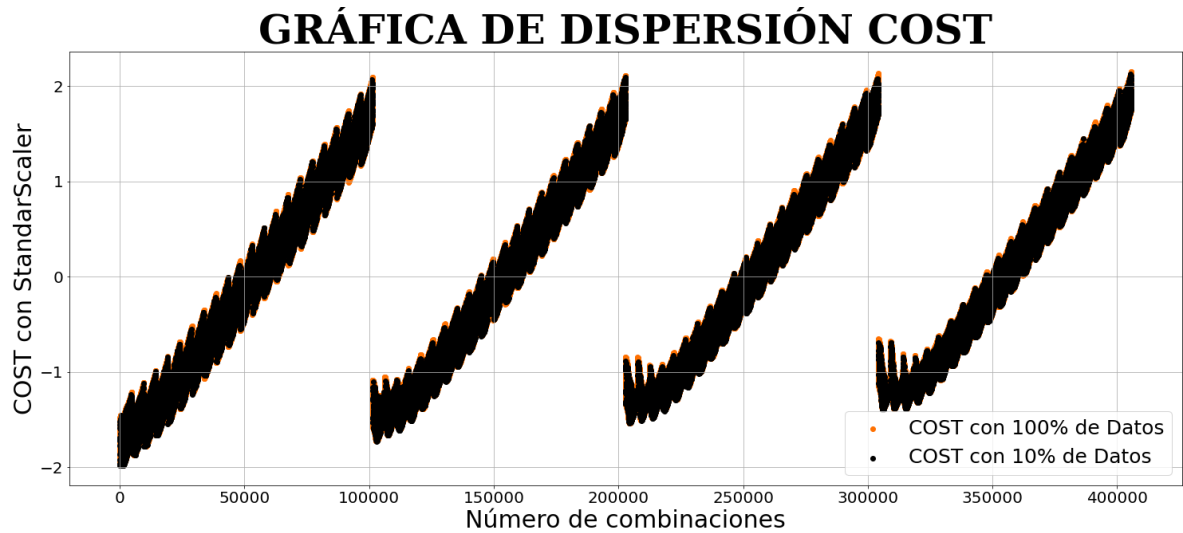


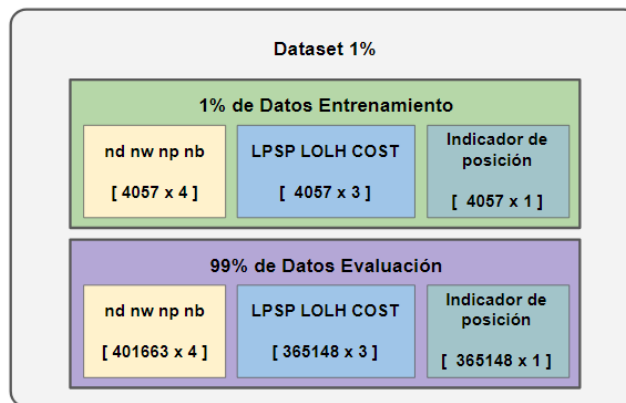
Figura 20. Gráfica de dispersión para el indicador económico COST para el 10% vs 100% de los datos.



II. DATASET 1%

Se presentan en esta sección los resultados obtenidos al repetir el procedimiento anterior, esta vez con el 1% de los datos totales.

Figura 21. Dataset con 1% de datos para entrenamiento y 99% de datos para evaluación.



A continuación, se presentan los resultados mediante la gráfica de dispersión de las variables LPSP (Figura 22), LOLH (Figura 23) y el valor COST (Figura 24) obtenidas por este método:

Figura 22. Gráfica de dispersión para el indicador de confiabilidad LPSP para el 1% vs 100% de los datos.

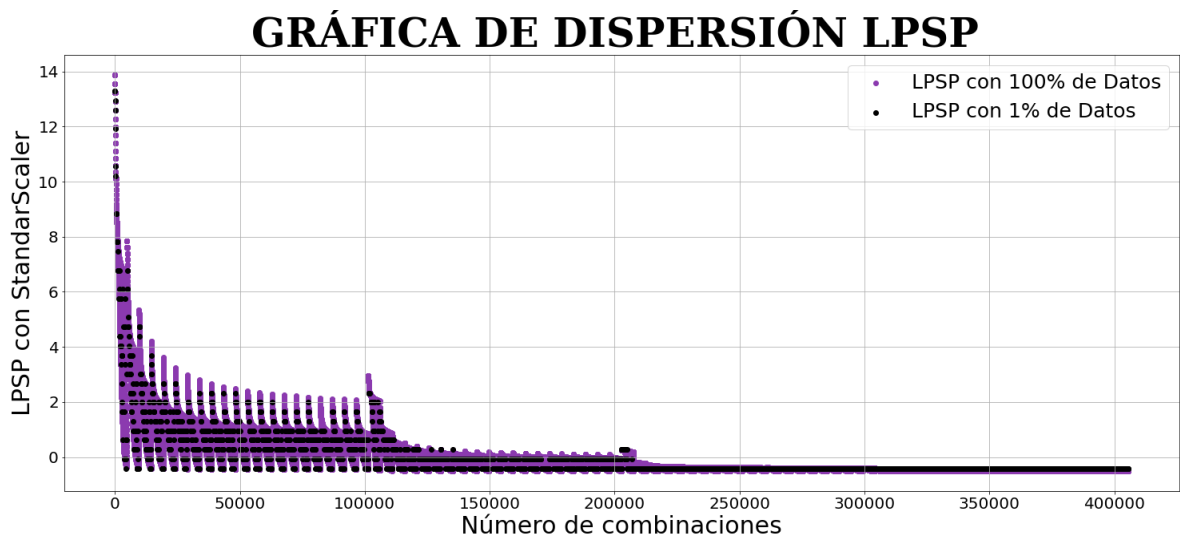


Figura 23. Gráfica de dispersión para el indicador de confiabilidad LOLH para el 1% vs 100% de los datos.

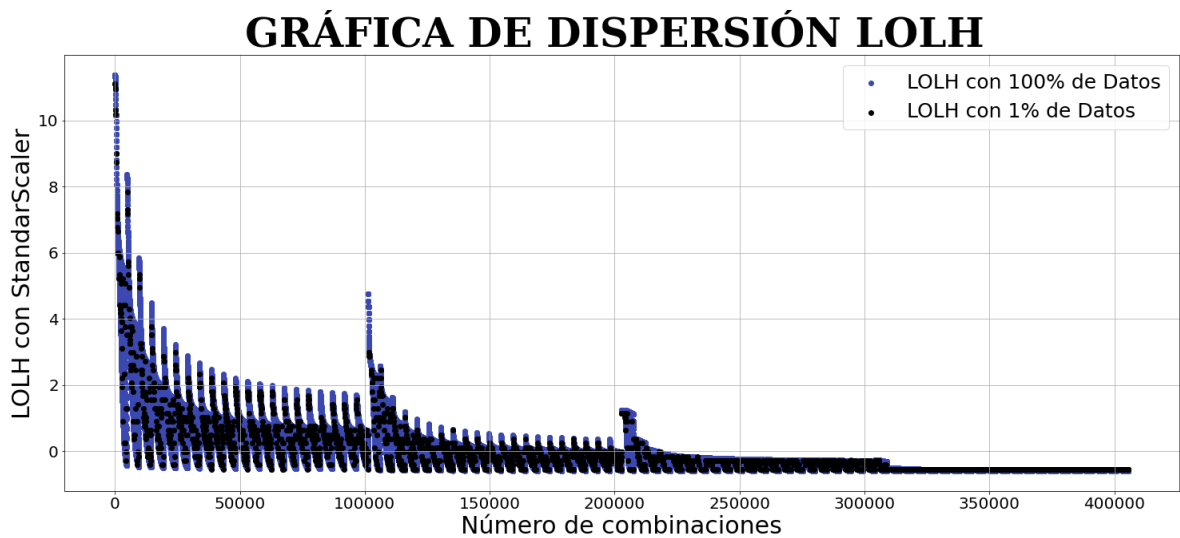
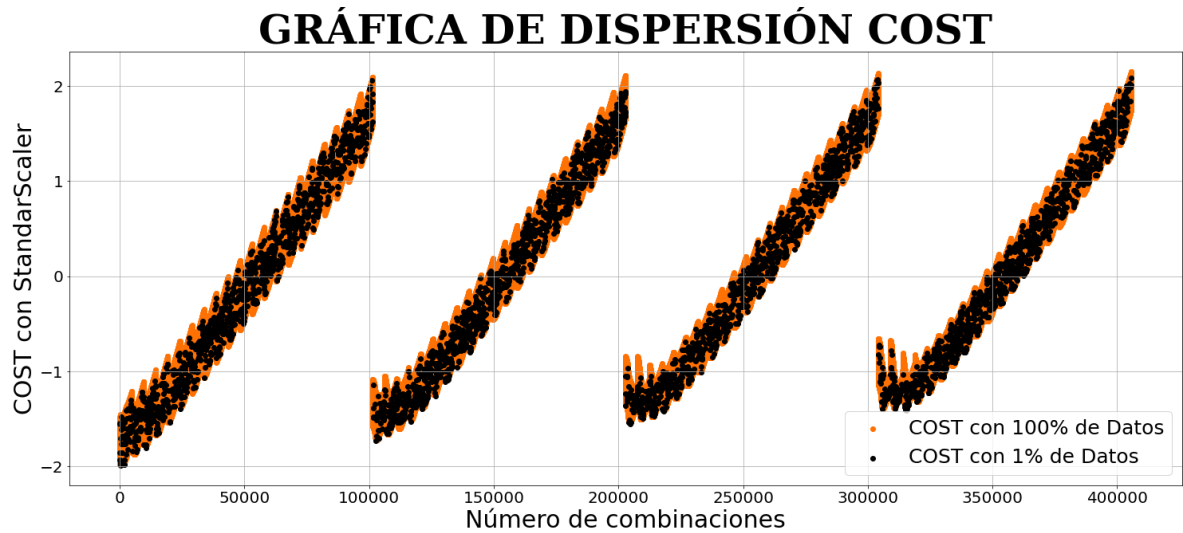


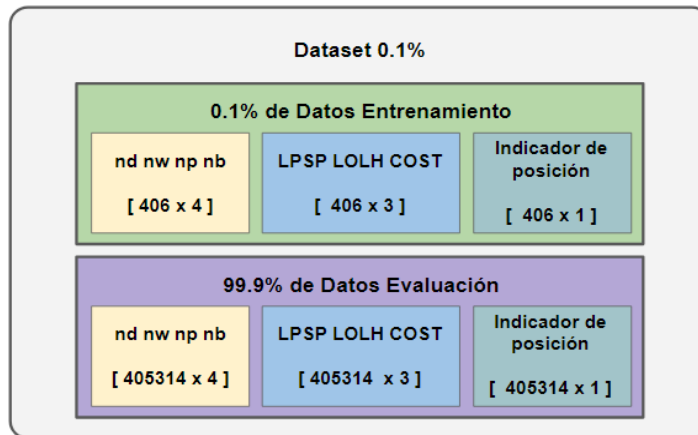
Figura 24. Gráfica de dispersión para el indicador económico COST para el 1% vs 100% de los datos.



III. DATASET 0.1%

Para el caso del 0.1% se obtuvo los siguientes resultados:

Figura 25. Dataset con 0.1% de datos para entrenamiento y 99.9% de datos para evaluación.



A continuación, se presentan los resultados mediante la gráfica de dispersión de las variables LPSP (Figura 26), LOLH (Figura 27) y el valor COST (Figura 28) obtenidas por este método:

Figura 26. Gráfica de dispersión para indicador de confiabilidad LPSP para el 0.1 Vs 100% de los datos.

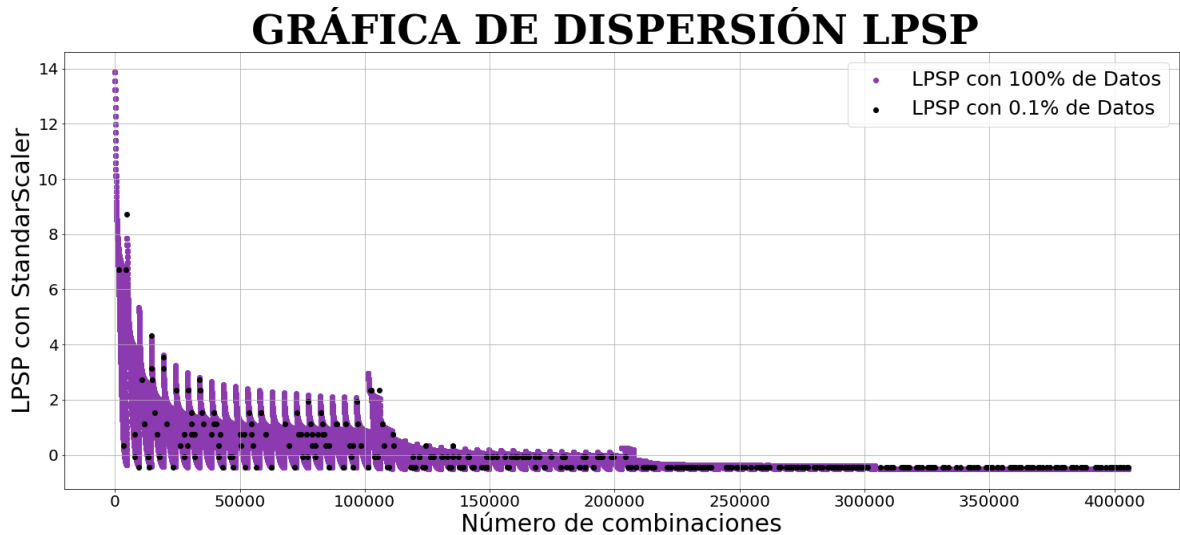


Figura 27. Gráfica de dispersión para indicador de confiabilidad LOLH para el 0.1 Vs el 100% de los datos.

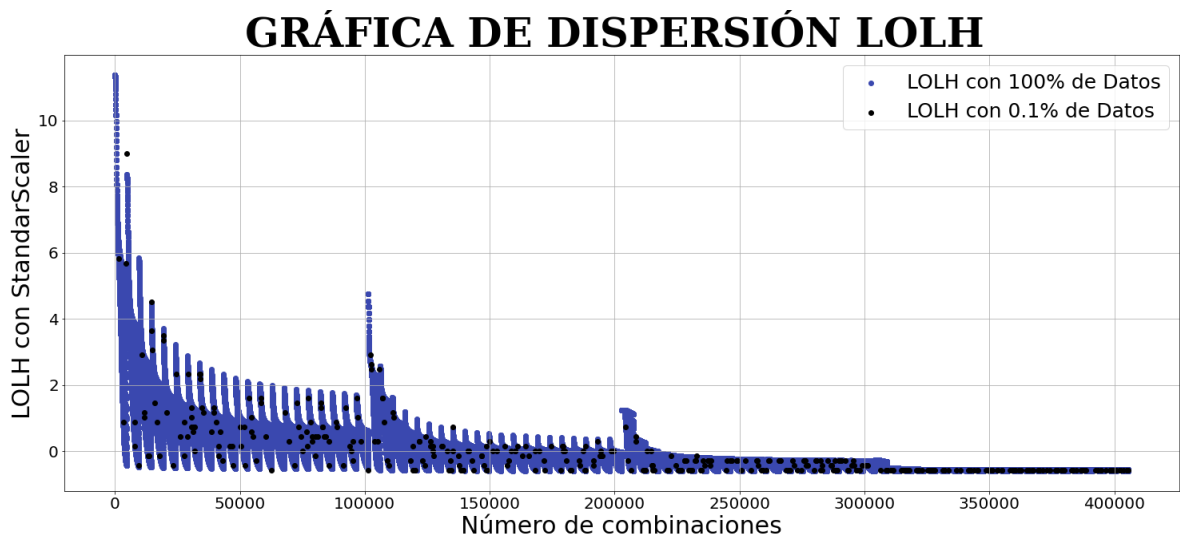
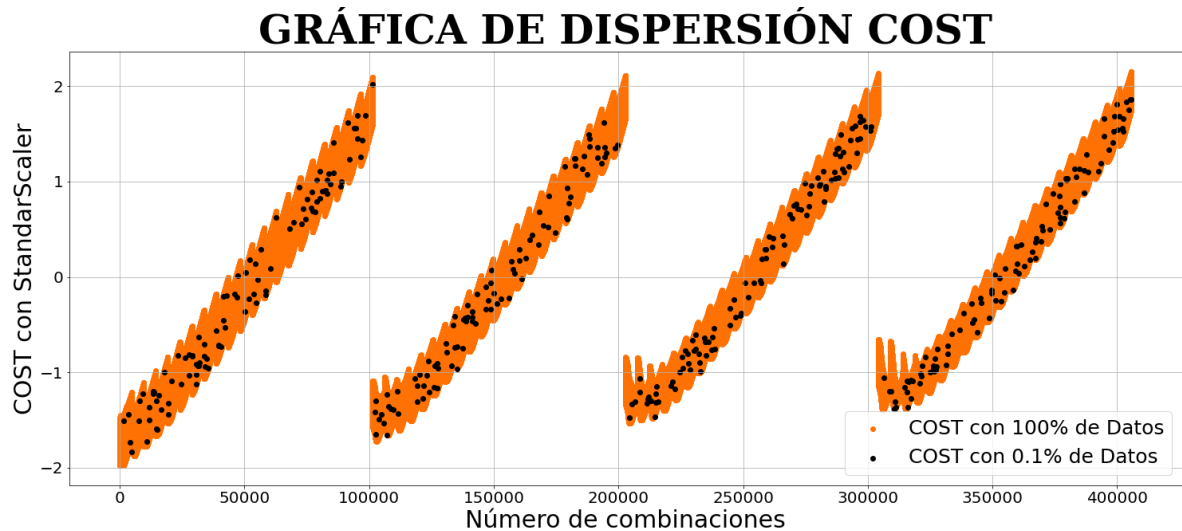


Figura 28. Gráfica de dispersión para indicador económico COST para el 0.1% Vs 100% de los datos.



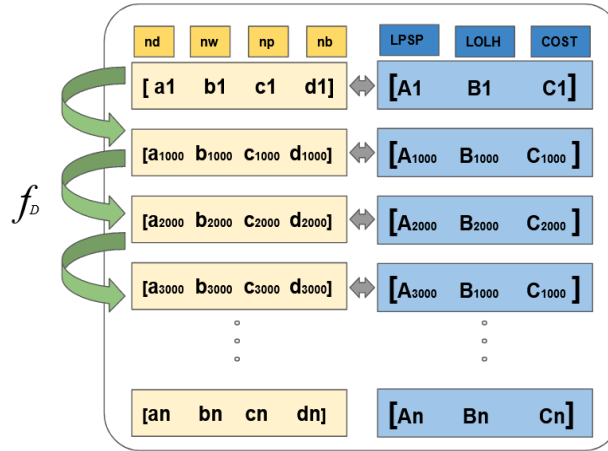
6.1.5.2. Método de decimación o reducción de muestras (MDoRM)

Como segundo método se ha seleccionado una técnica como la decimación de muestras, usada comúnmente en técnicas de procesamiento de señales digitales⁴⁰, lo cual permite extraer una secuencia reducida de muestras a partir de la señal original, descartando las restantes, dando como resultado una nueva señal con un tamaño reducido.

Para el caso de interés, la frecuencia de decimación usada tiene como función garantizar la toma de una muestra cada mil datos, lo que resulta en la reducción considerable en el volumen de datos resultantes que se ejecutará posteriormente en el entrenamiento de la RNA (Ver figura 29).

⁴⁰ OPPENHEIM, Alan V. SCHAFFER, Ronald W. BUCK, John R. Tratamiento de señales en tiempo discreto. Pearson Educación, 2000. p 193.

Figura 29. Estructura general del Dataset empleado en el método de decimación o reducción de muestras (saltos de 1000).



A continuación, se presentan los resultados mediante la gráfica de dispersión de las variables LPSP (Figura 30), LOLH (Figura 31) y el valor COST (Figura 32) obtenidas por este método:

Figura 30. Gráficas de dispersión para el indicador de confiabilidad LPSP de la decimación de muestras Vs 100% de los datos.

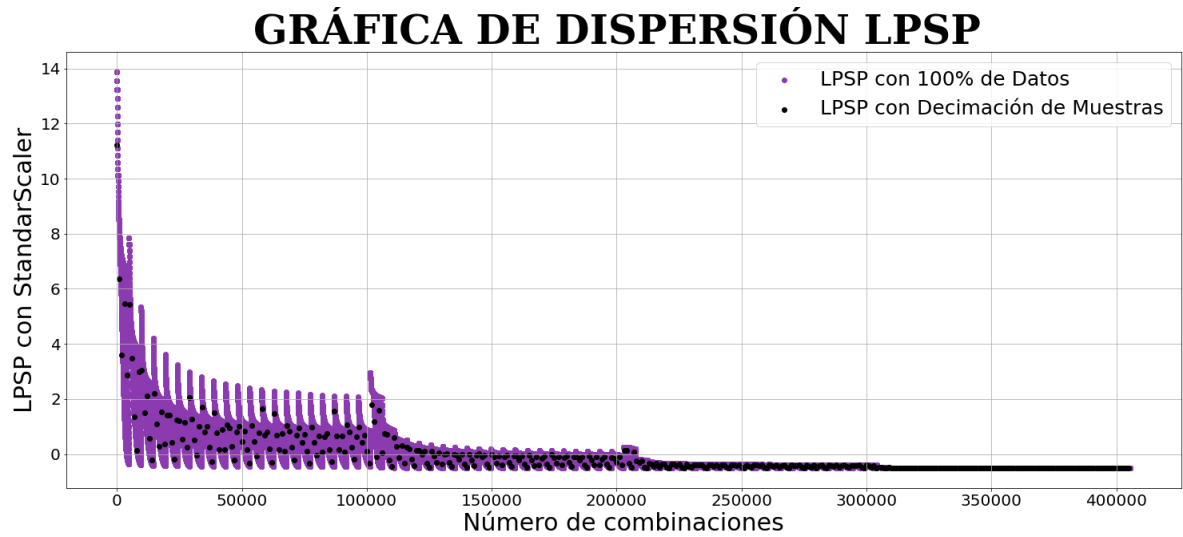


Figura 31. Gráficas de dispersión para indicador de confiabilidad LOLH de la decimación de muestras Vs 100% de los datos.

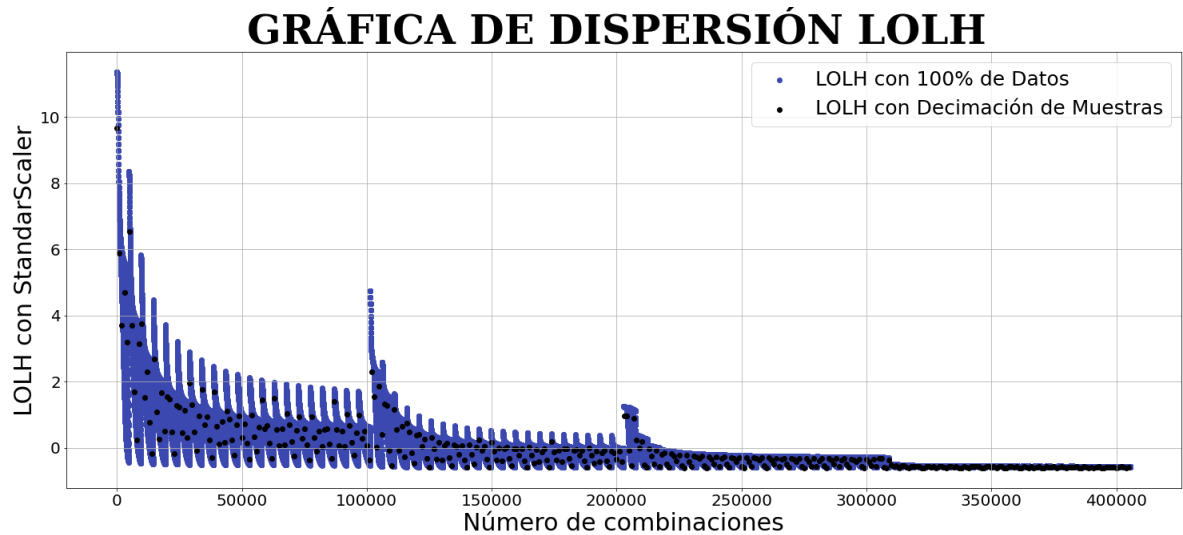
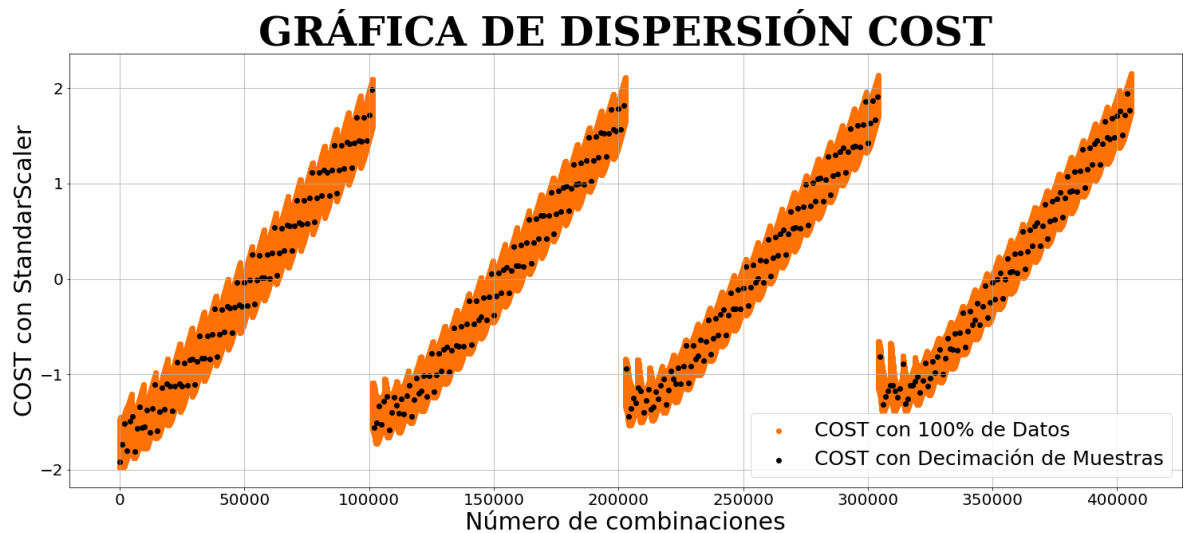


Figura 32. Gráficas de dispersión para indicador económico COST de la decimación de muestras Vs 100% de los datos.



6.1.5.3. Método de Restricción Crítica de Combinaciones (MRCC)

El tercer método implementado se basa en un análisis profundo en el comportamiento que presentan cada una de las salidas y el impacto que tiene las 4 entradas sobre ellas. Las 3 salidas LOLH, LPSP y COST tienen una tendencia a

cambiar significativamente con las variaciones de los parámetros **nd** y **nw**. En cambio, para los parámetros **np** y **nb**, existe un comportamiento con muy poco impacto en la respuesta y que sus aportes representan pequeñas distorsiones sobre la curva global.

El proceso de reducción de muestras parte de la identificación de los parámetros ECE/EAE con menos impacto en el comportamiento de las variables de salida (Análisis de dispersión gráfica), designando una cantidad reducida de elementos en la combinatoria, sin modificar los valores de aquellas variables que si presentan un mayor impacto. En otras palabras, la matriz de combinaciones resultante no estará reducida en su totalidad para todas las variables asignadas, sino que se reducirán aquellas muestras redundantes y consecutivas, utilizando el diezmado únicamente en el vector de valores que define el rango del parámetro ECE/EAE antes de conformar la matriz de combinaciones que se utilizará más adelante como Dataset.

Este método asegura un mayor control de la cantidad de datos que se quiere emplear para el entrenamiento de la RNA y también un mayor control en la selección crítica de los parámetros de entrada. Evitar usar la matriz de combinaciones totales y seleccionar una nueva ajustada al criterio de reducción y precisión requeridas.

La restricción se puede realizar teniendo en cuenta varios aspectos técnicos y de dimensionamiento al identificar cuáles elementos de la microrred tendrán o no un mayor impacto en las ecuaciones según sus modelos y parámetros intrínsecos y no de funciones aleatorias como las usadas en el método convencional o semicontroladas como la decimación de muestras, que, aunque permite reducir periódicamente las muestras, no se puede determinar el valor significativo de aquellas seleccionadas.

Para obtener las salidas de la nueva matriz de combinaciones, basta simplemente con repetir el cálculo empleando nuevamente los algoritmos de dimensionamiento

para consolidar una nueva base de datos, tal como se desarrolló para el Dataset original.

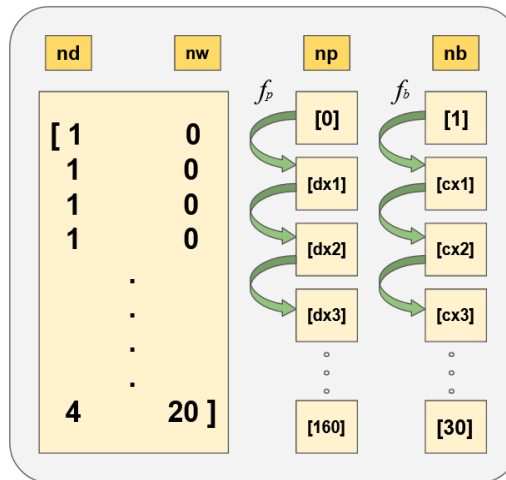
Este nuevo conjunto de datos no deriva del Dataset consolidado con la totalidad de las combinaciones posibles, lo que no permite la posibilidad de construir un vector para la posición dificultando así la comparación en sus gráficas de dispersión. Por tal motivo, en esta sección solo se muestra su comportamiento presentado sin la presencia del Dataset usado como referencia.

El nuevo conjunto está conformado por la totalidad de los datos de las variables **nd** y **nw**; para la entrada **np** se reduce su cantidad de muestras, al asignar una frecuencia de muestreo, recorriendo todo el rango desde la cifra mínima hasta su valor máximo. El mismo proceso se repite para la variable **nb**. Como resultado se tiene que el nuevo conjunto de combinaciones se presenta de la siguiente manera:

$$combi_{m3} = nd_{[]} \times nw_{[]} \times \widehat{np}_{[]} \times \widehat{nb}_{[]} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 4 & 20 & 160 & 30 \end{pmatrix} \quad (27)$$

El tamaño de la nueva matriz $combi_{m3}$ se determina al realizar la multiplicación del número que compone cada vector de datos, $\widehat{np}_{[]}$ y $\widehat{nb}_{[]}$ son los nuevos conjuntos diezmados que contiene el mismo rango en información, con la diferencia que sus elementos no son secuenciales y están reducidos a comparación con la matriz de entrada original.

Figura 33. Estructura general del Dataset empleado en el método Restricción crítica de combinaciones.



A continuación, se presentan los resultados mediante la gráfica de dispersión de las variables LPSP (Figura 34), LOLH (Figura 35) y el valor COST (Figura 36) obtenidas por este método:

Figura 34. Gráfica de dispersión para el indicador de confiabilidad LPSP con restricción crítica de datos.

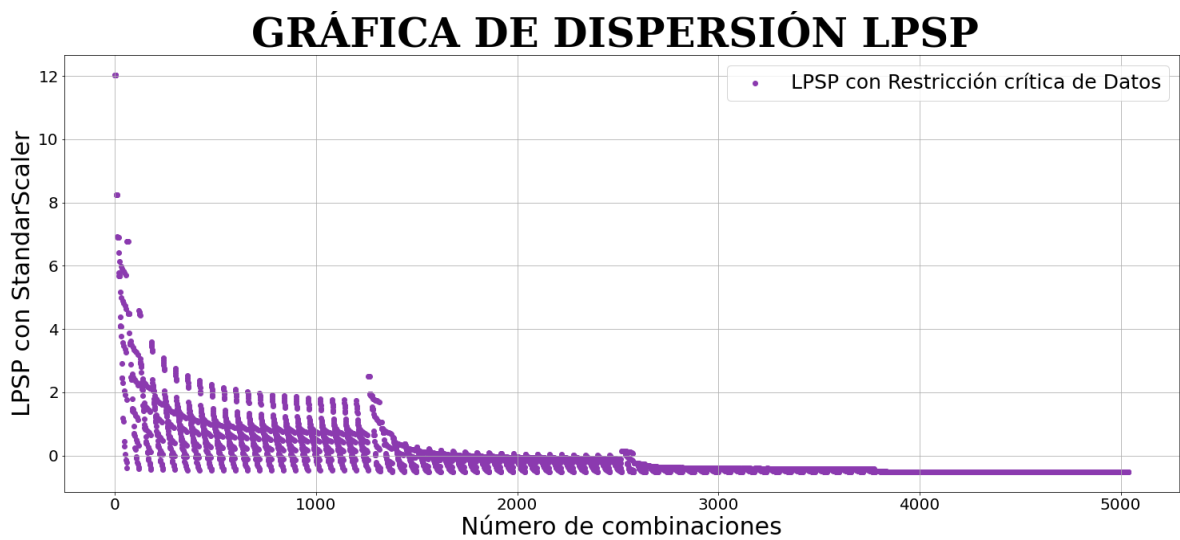


Figura 35. Gráfica de dispersión para el indicador de confiabilidad LOLH con restricción crítica de datos.

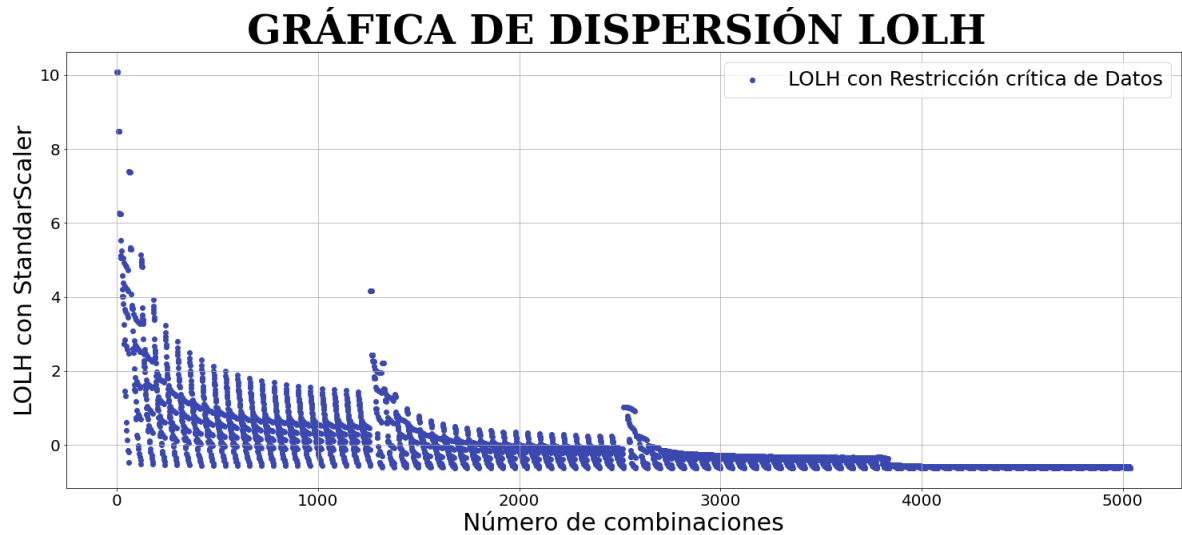
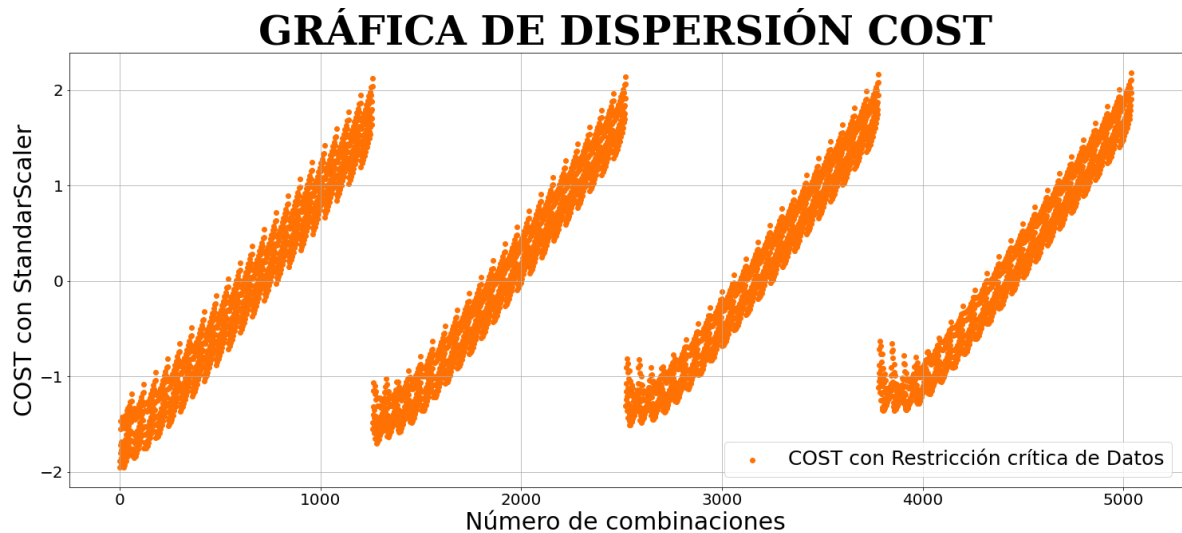


Figura 36. Gráfica de dispersión para el indicador económico COST con restricción crítica de datos.



6.2 ETAPA DE ANÁLISIS PRE-IMPLEMENTACIÓN

Antes de ejecutar la etapa de implementación de la RNA, se requiere realizar un estudio o análisis previo de las características estadísticas, así como una

adecuación de todos los datos que serán utilizados por cada RNA a entrenar. Los modelos de RNA que emplea la herramienta Tensorflow requieren una normalización en sus datos de entrada, con el fin de facilitar el proceso de corrección de los valores de los pesos que componen el algoritmo de cada una de las capas implementadas. La red se ve afectada por las diferencias en la magnitud y rango de los datos que ingresan en su capa de entrada, comprometiendo la fiabilidad que posee el sistema para identificar patrones de esta. En esta sección se presenta un método comúnmente usado para corregir este problema que ofrece el kit de herramientas de Tensorflow como lo es *StandardScaler*, de la librería *Sklearn*.

6.2.1. Acondicionamiento de los Conjuntos de Datos de Entrenamiento

Todo proceso de entrenamiento de la RNA requiere un acondicionamiento en su conjunto de datos destinados para dicha tarea. El acondicionamiento comprende asignar un formato válido de los datos, incluyendo la normalización, paso necesario que demanda la librería *Tensorflow*⁴¹.

La normalización es una etapa previa al *training* y se efectúa utilizando la librería *Sklearn*, dentro de la cual se encuentran diferentes tipos de transformaciones y funciones usadas en el *machine learning*. La transformación seleccionada ha sido la normalización *StandardScaler*, función encargada de escalar de manera que la RNA pueda ajustar el valor de sus pesos a medida que se adapta a una nueva entrada y se adapta al comportamiento que el Dataset presente⁴².

Debido a que, por lo general los datos pueden presentar valores con escalas muy grandes entre ellos, pasar de unidades pequeñas a cantidades cientos de veces

⁴¹ SALEH, Hyatt. Machine Learning Fundamentals. Packt Publishing Editorial. 2018.

⁴² *Ibíd.*

mayores; el objetivo de usar esta herramienta es mantener la misma relación entre sus elementos, pero a una escala que le sea mucho más fácil ajustarse a los cambios de los pesos de la RNA.

La función de normalización de StandardScaler⁴³ se define hallando la media y la desviación estándar del conjunto de datos que lo conforma, mostrado en la siguiente ecuación:

$$x_s[i] = \frac{x[i] - \bar{x}}{\sigma} \quad (28)$$

Donde:

$x_s[i]$ corresponde al nuevo dato normalizado.

$x[i]$ corresponde al valor real.

σ corresponde a la desviación estándar de la totalidad de la base de datos.

\bar{x} corresponde a la media aritmética de la base de datos.

En la Tabla 2 se presentan los valores σ y \bar{x} calculados para los Datasets de los métodos a estudiar.

⁴³ PEDREGOSA, F. VAROQUAUX, G. GRAMFORT, A. MICHEL, V. THIRION, B. GRISEL, O. BLONDEL, M. PRETTENHOFER, P. WEISS, R. DUBOURG, V. VANDERPLAS, J. AND PASSOS, A. COURNAPEAU, D. BRUCHER, M. PERROT, M. DUCHESNAY, E. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research. Vol 12. P. 2825—2830. 2011.

Tabla 2. Parámetros de las distintas bases de datos para cada uno de los métodos seleccionados.

Método Convencional con el 10% del total de los datos							
Datos Training	Datos Testing	LPSP		LOLH		COST	
40572	365148	\bar{x}	σ	\bar{x}	Σ	\bar{x}	Σ
		1.1702	2.7480	4.0905	7.3744	287006.5310	114242.7603
Método Convencional con el 1% del total de los datos							
Datos Training	Datos Testing	LPSP		LOLH		COST	
4057	401663	\bar{x}	σ	\bar{x}	Σ	\bar{x}	σ
		1.2072	2.9187	4.1816	7.6476	287971.3017	113427.3064
Método Convencional con el 0.1% del total de los datos							
Datos Training	Datos Testing	LPSP		LOLH		COST	
405	405315	\bar{x}	σ	\bar{x}	Σ	\bar{x}	σ
		1.1555	2.5080	3.9703	6.8991	288340.1703	113245.6494
Método Decimación o Reducción de Muestras							
Datos Train	Datos Test	LPSP		LOLH		COST	
406	405314	\bar{x}	σ	\bar{x}	Σ	\bar{x}	σ
		1.7643	3.5020	5.3755	8.6933	278652.2810	114041.2927
Método de Restricción Crítica de Combinaciones							
Datos Train	Datos Test	LPSP		LOLH		COST	
5040	400680	\bar{x}	σ	\bar{x}	Σ	\bar{x}	σ
		1.6989	3.2654	5.2099	8.3471	283562.2697	114207.1169

6.2.2. Análisis de Distribución Normal de Training Dataset

Las gráficas de distribución normal permiten visualizar el comportamiento en un conjunto de datos aleatorios al mostrar la distribución que presenta alrededor de su

media aritmética⁴⁴. La función de distribución normal se emplea en esta sección y en secciones siguientes como una herramienta para observar la varianza y la tendencia a la que se ve afectada cada Dataset al momento de modificar la forma en que se construyen.

Para generar cada gráfica, se sustrajo los valores de σ y \bar{x} obtenidos en la sección 6.2.1 para la matriz de datos del Dataset de referencia y los conjuntos de datos del Método Convencional con el 10%, 1% y 0.1% de la totalidad de los datos de estudio, Método Decimación o Reducción de Muestras y el Método de Restricción Crítica de Combinaciones.

A continuación, se presentan los resultados obtenidos en las Figura 37, Figura 38 y Figura 39, para los indicadores de confiabilidad LPSP, LOLH y el indicador económico de costos.

⁴⁴ SALAZAR, Cecilia. DEL CASTILLO, Santiago. Fundamentos básicos de estadística. Primera edición. 2018.

Figura 37. Gráfica de distribución normal para el indicador de confiabilidad LPSP para el método convencional con diversos porcentajes de datos, método de decimación de muestras y restricción crítica de datos.

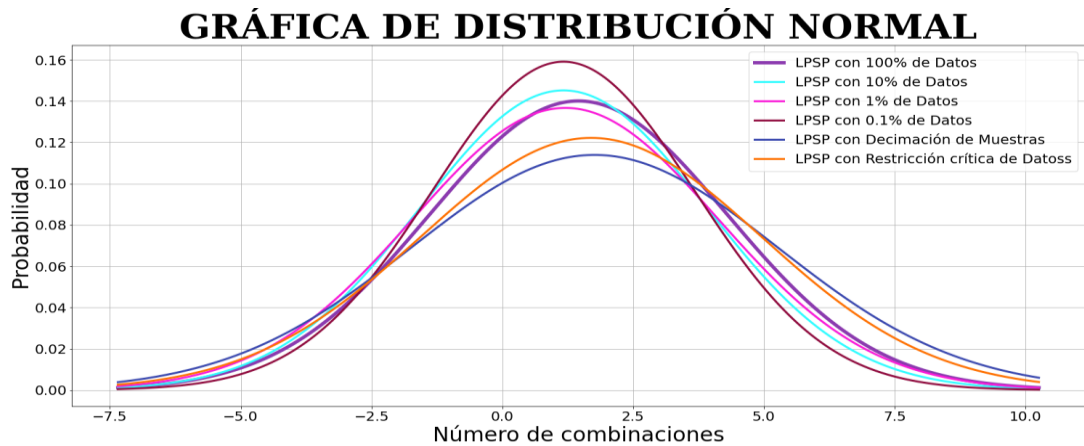


Figura 38. Gráfica de distribución normal para el indicador de confiabilidad LOLH para el método convencional con diversos porcentajes de datos, método de decimación de muestras y restricción crítica de datos

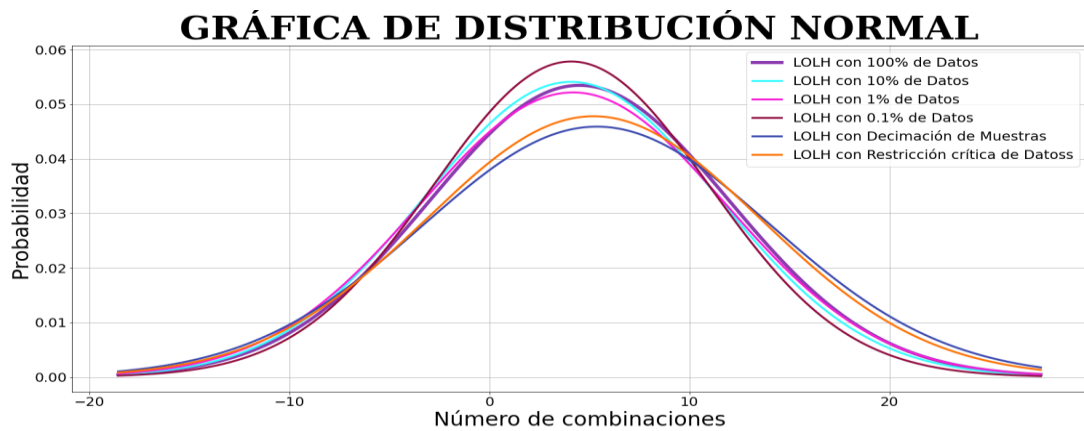
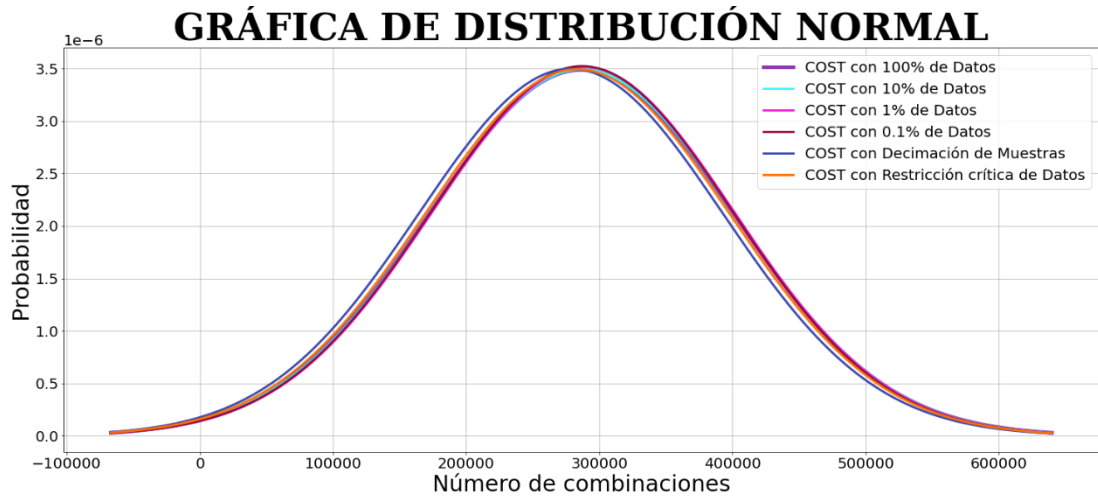


Figura 39. Gráfica de distribución normal para el indicador económico COST para el método convencional con diversos porcentajes de datos, método de decimación de muestras y restricción crítica de datos.



6.2.3. Criterios de diseño de la RNA

Se requiere implementar una RNA que tenga la capacidad de interpretar los comportamientos de los datos obtenidos de la MG para la predicción de las combinaciones que no hacen parte de la base de datos de entrenamiento. El mejor modelo para la predicción de comportamiento lo presenta la estructura *sequential* de la librería de *Tensorflow*⁴⁵. La RNA presenta una estructura donde sus interconexiones se propagan desde la capa de entrada, pasando por sus capas ocultas hasta la capa de salida. Otro aspecto importante para tener en cuenta es la función de activación, que depende mucho del tipo de información que recibe y desea transmitir la neurona.

⁴⁵ GÉRON, Aurélien. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow. Concepts, Tools, and Techniques to Build Intelligent Systems. Second Edition. O'Reilly Media Inc. 2019.

A continuación, se muestran los parámetros que componen la RNA implementada para la predicción de los indicadores de confiabilidad y económicos obtenidos en el proceso de dimensionamiento de MG aisladas.

Tabla 3. parámetros de la RNA implementada en Python.

	NO DE NEURONAS	MODELO DE RNA	FUNCIÓN DE ACTIVACIÓN
CAPA DE ENTRADA	4	Sequential	Relu
CAPAS OCULTAS	996		Relu
CAPA DE SALIDA	3		Linear

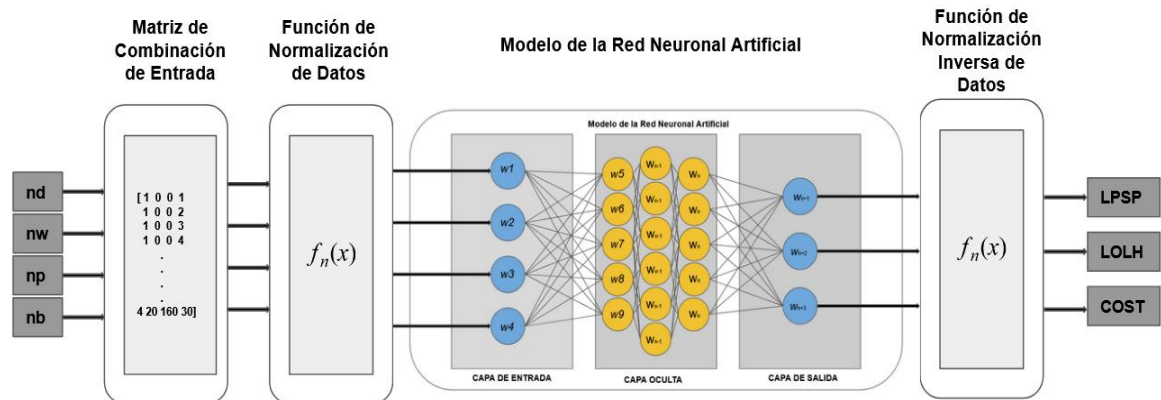
6.3 ETAPA DE IMPLEMENTACIÓN RNA

La etapa de implementación de la RNA descrita en la etapa de diseño se realiza bajo el uso de la librería de Tensorflow en el lenguaje de Python, usando como ejemplos otras implementaciones de distintas aplicaciones que se estudiaron como punto de referencia⁴⁶. El código implementado se compone de diferentes secciones que se encargan de realizar una tarea específica, definidos en 4 grandes módulos: módulo de preprocesamiento de datos que genera la matriz de combinación de entrada, módulo de normalización de datos, módulo de la RNA entrenada, y

⁴⁶ GÉRON, Aurélien. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow. Concepts, Tools, and Techniques to Build Intelligent Systems. Second Edition. O'Reilly Media Inc. 2019.

finalmente, el módulo de normalización inversa de la salida. La arquitectura del sistema se describe en el diagrama presentado en la Figura 40.

Figura 40. Arquitectura de la RNA implementada.



7. ETAPA DE ANÁLISIS DE RESULTADOS

Dentro del marco de la metodología, la presentación y análisis de resultados requiere de mayor atención debido al impacto que tienen los métodos experimentales que se proponen como una alternativa a la forma que comúnmente se realiza en el entrenamiento de Machine Learning y las aplicaciones de RNA.

Se seleccionaron indicadores como el valor de Loss y Accuracy, valores que normalmente arroja la herramienta Tensorflow durante el proceso de entrenamiento de la Red Neuronal. Además, se revisó el comportamiento de los conjuntos de datos obtenidos para cada RNA entrenada con los 3 métodos de Dataset, mediante un análisis gráfico de dispersión y distribución normal comparados con los datos de referencia.

En conjunto con los indicadores mencionados anteriormente, se propone la revisión del método del Coeficiente de Determinación (R2 Score), un concepto usado en el ambiente de la Inteligencia Artificial, para medir el desempeño de la RNA y la

eficiencia en la predicción de resultados a la presencia de nueva información en la entrada de la Red.

7.1. Análisis de Loss y Accuracy

En la siguiente sección se presentan los resultados obtenidos mediante el uso de la herramienta Tensorflow para determinar el nivel de exactitud (*Accuracy*) que presenta la salida de la RNA al compararse con los datos asignados para la etapa de testing. El valor de *Loss* permite determinar el nivel de desempeño de los pesos calculados, correspondiendo a la variación que presenta el cálculo de los nuevos pesos con relación a la última entrada en el proceso de training.

Tabla 4. Recopilación de resultados obtenidos en los 3 métodos implementados.

RNA Entrenada	LOSS	ACCURACY
Método Convencional	0,0016	97,31%
Método Decimación o Reducción de Muestras	0,0213	91,76%
Método de Restricción Crítica de Combinaciones	0,0108	93,27%

Según los resultados mostrados en la Tabla 4, los 3 métodos implementados dictan una respuesta satisfactoria para los 3 casos analizados. Las 3 Redes Neuronales entrenadas proporcionan un indicador de precisión (*Accuracy*) satisfactorio por encima del umbral del 90%, y un valor de *Loss* muy por debajo del 0.03. El método

que muestra un resultado sobresaliente es el *Método Convencional*, con un 97.31% de exactitud en los datos obtenidos en su tarea de predicción al ingreso de nuevas entradas. El método de Restricción Crítica de Combinaciones (RCC), presenta una mejoría en la precisión al compararse con el método de Decimación o Reducción de muestras (DoRM), pero no lo suficientemente cerca del primer método.

La diferencia mostrada en estos resultados permite inferir que el algoritmo de la RNA responde mejor a una variación aleatoria en la presentación de los datos del Dataset, a diferencia de una matriz de entrada con secuencia progresiva y ordenada. La etapa de preprocesamiento y adecuación de datos antes del proceso de entrenamiento de la RNA suele generar un mayor impacto en su salida de lo esperado, no obstante, por la naturaleza descriptiva del origen de la información proveniente de la MG, se obtuvieron resultados satisfactorios cumpliendo con el objetivo general de este trabajo.

7.2. Análisis de Dispersión de Datos

Esta sección presenta los resultados gráficos del comportamiento en dispersión de los 3 métodos implementados anteriormente. Se aplica el mismo método de análisis presentado en la sección 6.1.4, se toman las muestras obtenidas de la predicción de cada RNA y se compara con los datos de referencia. A continuación, se presentan las gráficas resultantes de cada uno de los métodos descritos e implementados a partir de la Figura 41 hasta la Figura 49.

Figura 41. Gráfica de dispersión para el indicador de confiabilidad LPSP de la RNA con datos de entrenamiento vs referencia en el método convencional.

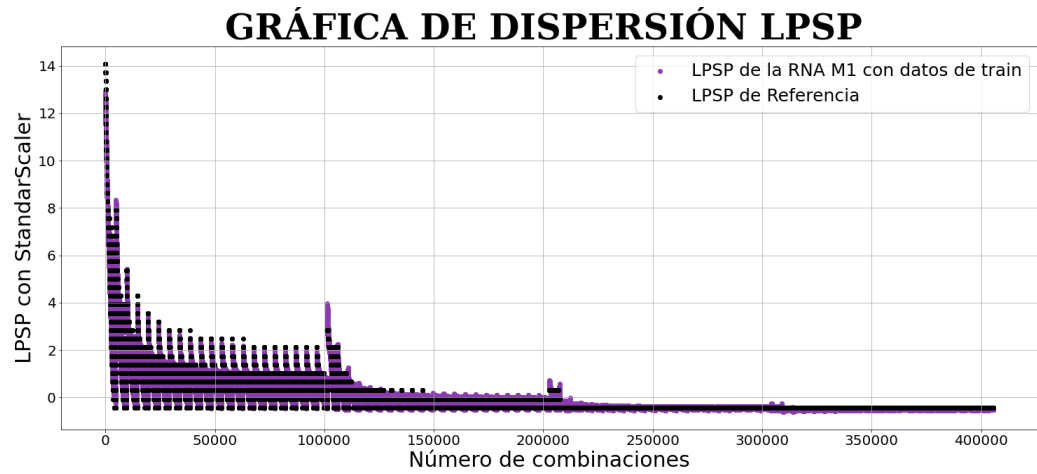


Figura 42. Gráfica de dispersión para el indicador de confiabilidad LOLH de la RNA con datos de entrenamiento vs referencia en el método convencional.

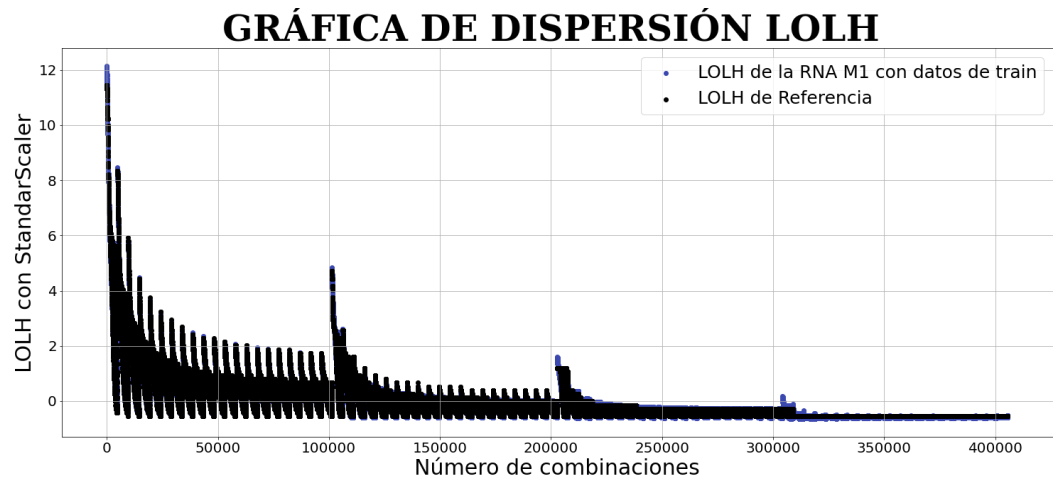


Figura 43. Gráfica de dispersión para el indicador económico COST de la RNA con datos de entrenamiento vs referencia en el método convencional.

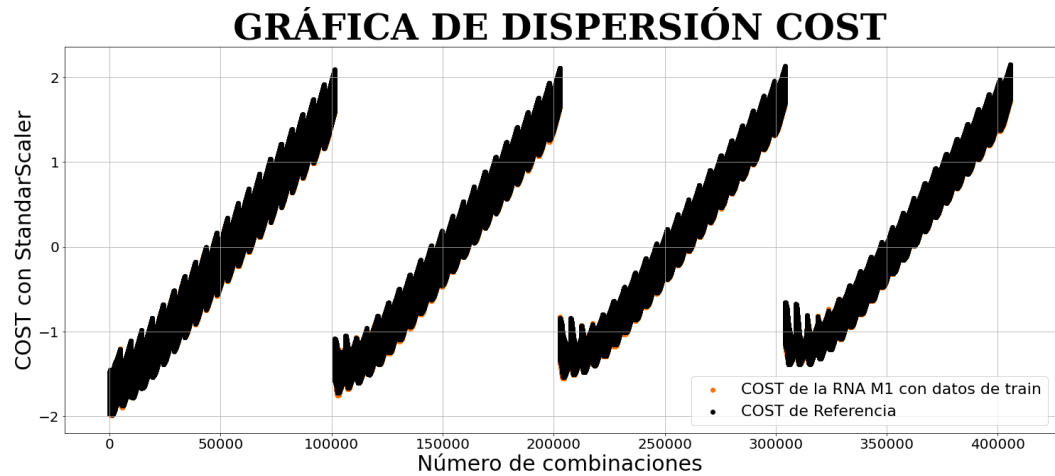


Figura 44. Gráfica de dispersión para el indicador de confiabilidad LPSP de la RNA con datos de entrenamiento vs referencia en el método de decimación o reducción de muestras.

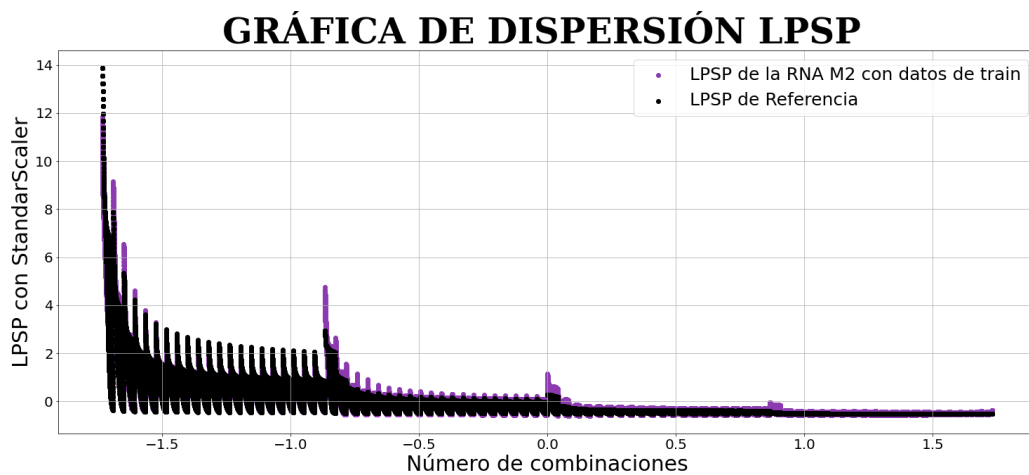


Figura 45. Gráfica de dispersión para el indicador de confiabilidad LOLH de la RNA con datos de entrenamiento vs referencia en el método de decimación o reducción de muestras.

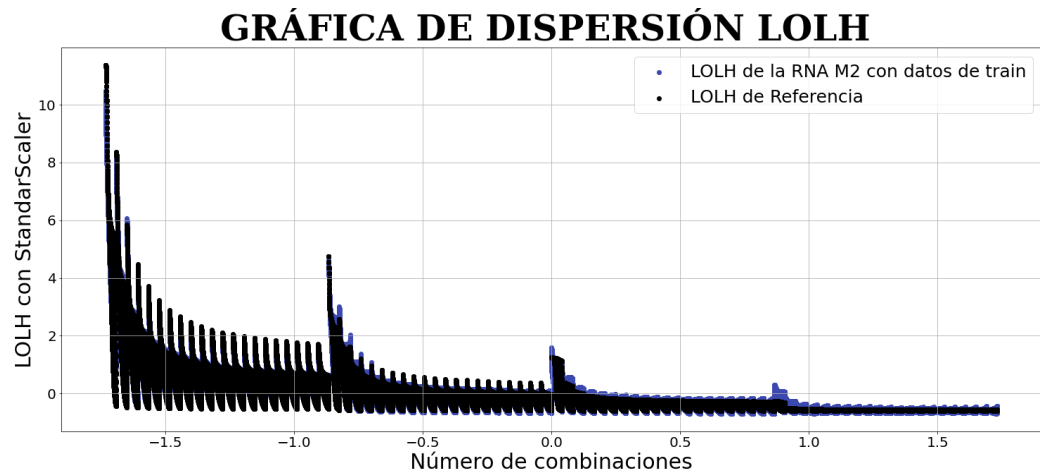


Figura 46. Gráfica de dispersión para el indicador económico COST de la RNA con datos de entrenamiento vs referencia en el método de decimación o reducción de muestras.

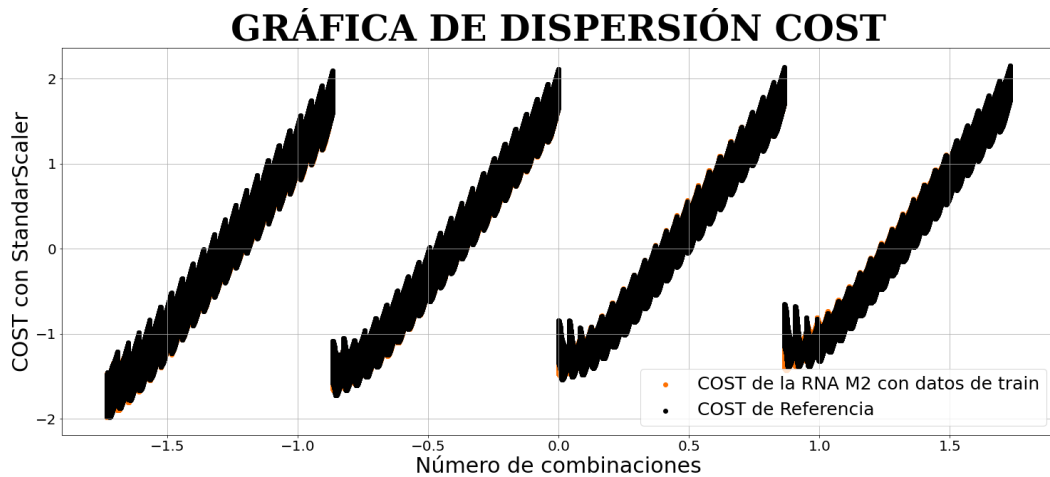


Figura 47. Gráfica de dispersión para el indicador de confiabilidad LPSP de la RNA con datos de entrenamiento vs referencia en el método de restricción crítica de combinaciones.

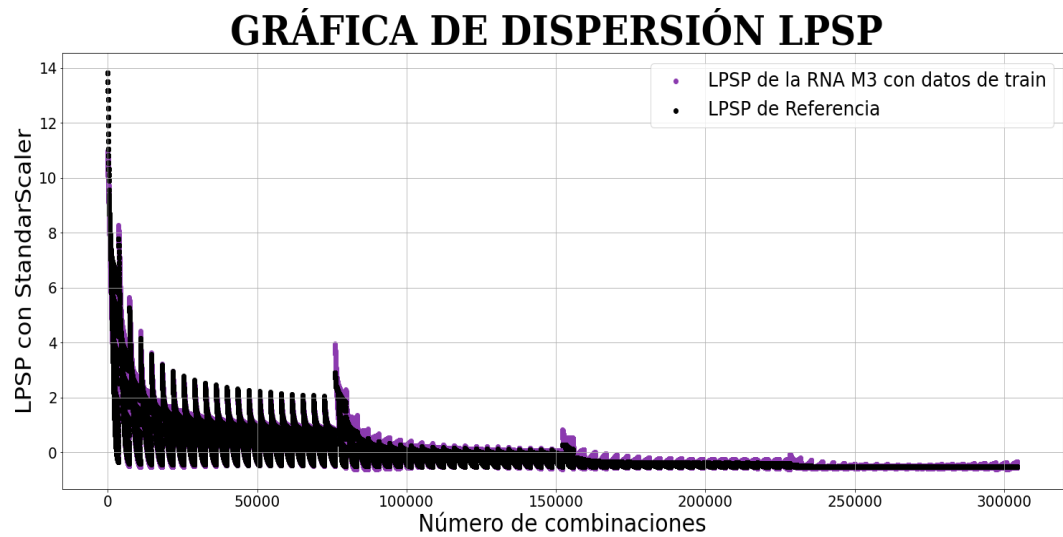


Figura 48. Gráfica de dispersión para el indicador de confiabilidad LOLH de la RNA con datos de entrenamiento vs referencia en el método de restricción crítica de combinaciones.

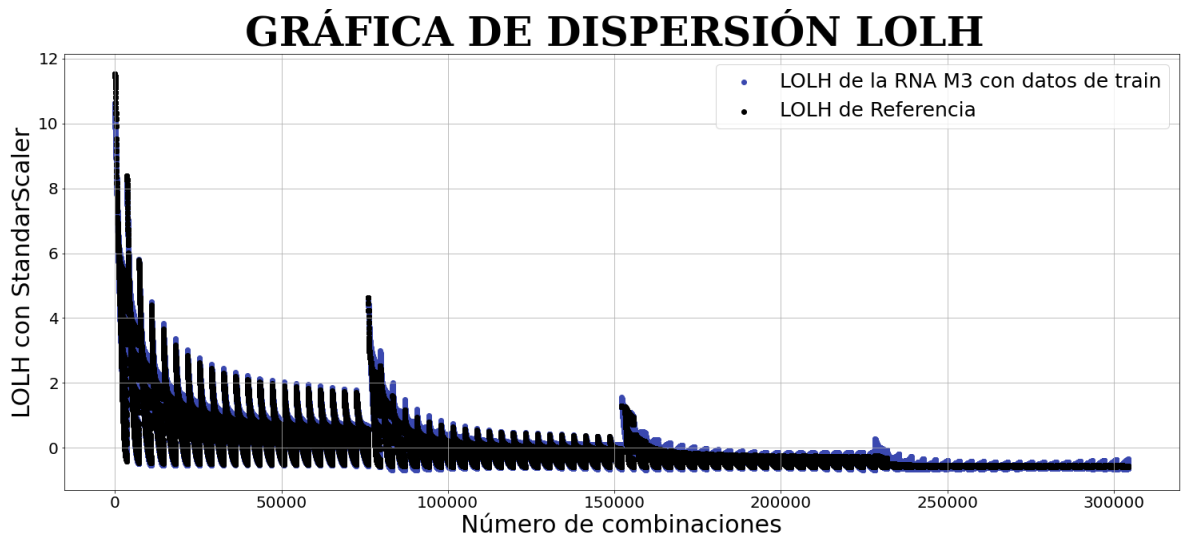
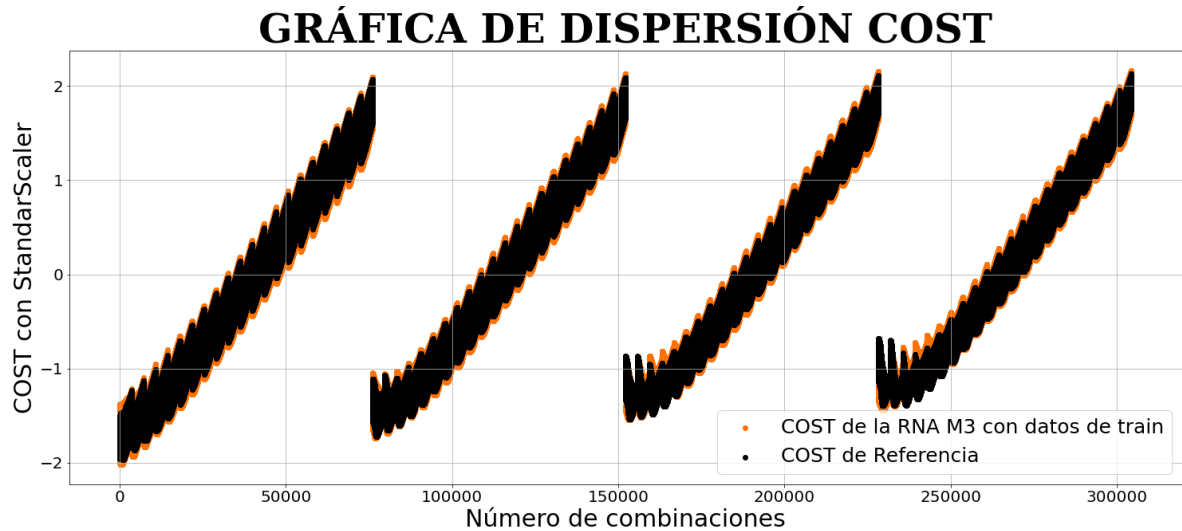


Figura 49. Gráfica de dispersión para el indicador económico COST de la RNA con datos de entrenamiento vs referencia en el método de restricción crítica de combinaciones.



7.3. Análisis de Distribución Normal de los Datos obtenidos

En la Figura 50, Figura 51 y Figura 52 se aprecian las gráficas de distribución normal para los indicadores de confiabilidad LPSP, LOLH y el indicador económico COST correspondientes a cada uno de los 3 métodos implementados. Los resultados obtenidos en esta sección indican un comportamiento congruente con respecto a los análisis presentados en secciones anteriores. El MC y MRCC muestran el mismo comportamiento en su distribución, en comparación al MDoRM que si tiene una variación notable en el valor pico para los indicadores de LPSP y LOLH. Para el valor del indicador económico COST, los 3 métodos tienen mayor convergencia en su gráfica y comportamiento en general.

Figura 50. Gráfica de distribución normal para el indicador de confiabilidad LPSP en los 3 métodos implementados.

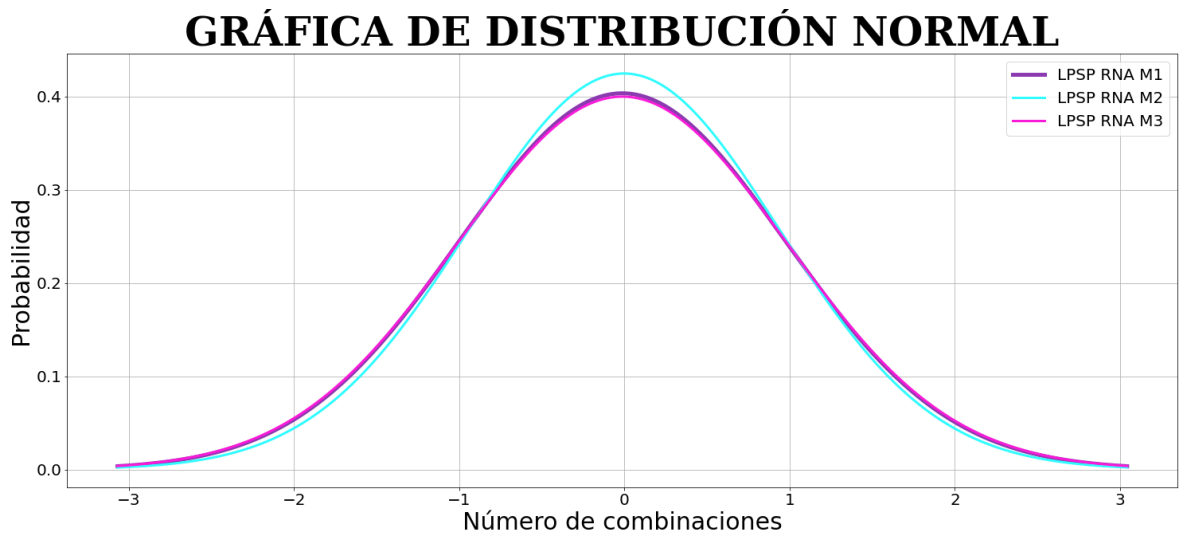


Figura 51. Gráfica de distribución normal para el indicador de confiabilidad LOLH en los 3 métodos implementados.

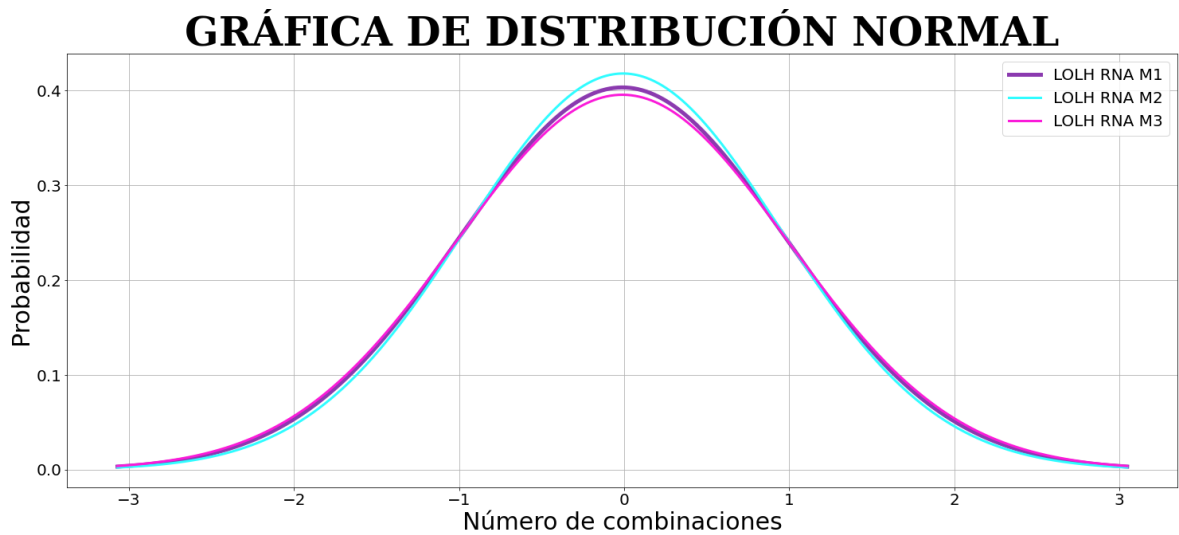
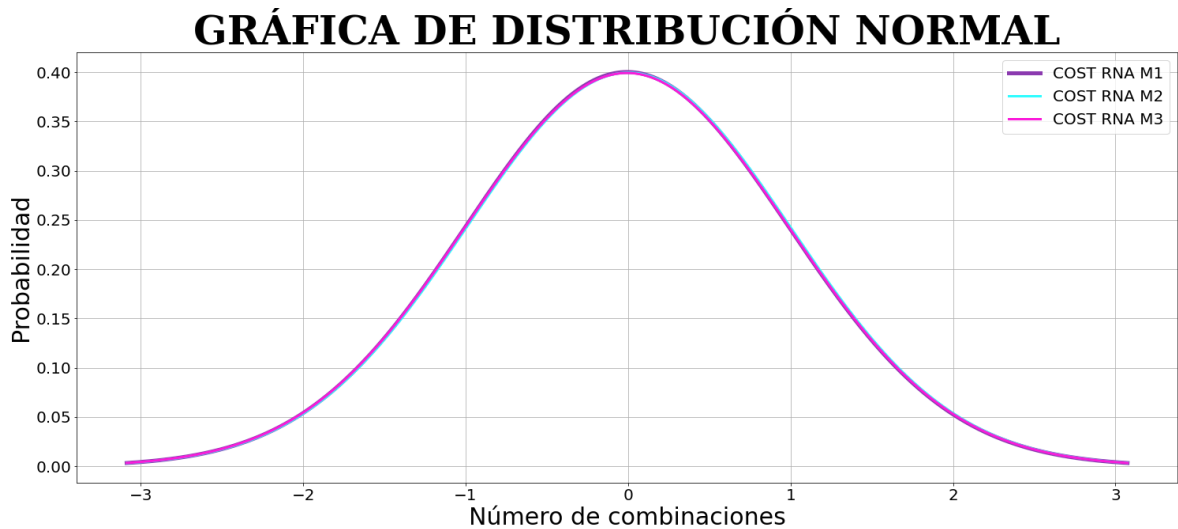


Figura 52. Gráfica de distribución normal para el indicador económico COST en los 3 métodos implementados.



Basado en las gráficas de dispersión anteriores, se extraen los valores de la media (\bar{x}) y la desviación estándar (σ) de cada método implementado, como otro indicador de desempeño cuantificable para determinar la veracidad de los resultados obtenidos. Estos dos indicadores reflejan el comportamiento compacto y global de la información presente en cada conjunto de Dataset, indicando la tendencia y el comportamiento de los datos al agruparse a su valor medio.

A continuación, se presentan los resultados obtenidos de cada RNA al evaluar su respectivo Dataset de testing. (Tabla 5)

Tabla 5. Resultados de σ y \bar{x} obtenidos en los 3 métodos seleccionados.

Método Convencional con el 1% del total de los datos					
LPSP		LOLH		COST	
\bar{x}	σ	\bar{x}	σ	\bar{x}	σ
-0.0127	0.9887	-0.0117	0.9896	-0.0064	0.9968

Método Decimación o Reducción de Muestras					
LPSP		LOLH		COST	
\bar{x}	σ	\bar{x}	σ	\bar{x}	σ
-0.0019	0.9393	-0.0071	0.9548	0.0025	0.9982

Método de Restricción Crítica de Combinaciones					
LPSP		LOLH		COST	
\bar{x}	σ	\bar{x}	σ	\bar{x}	σ
0.0045	0.9975	0.0045	1.0089	0.0024	0.9991

7.4. Análisis del Coeficiente de Determinación (R2 Score)

El Coeficiente de Determinación (R2) representa un indicador de desempeño importante en la evaluación de los sistemas basados en modelos *Machine learning* (ML) y es comúnmente empleada para determinar la precisión que posee un modelo matemático bajo condiciones no relacionadas con entradas conocidas y que no hicieron parte de la etapa de entrenamiento.

La expresión que define el R2 se muestra a continuación:

$$R^2 = \frac{\sum_{t=1}^T (\hat{Y}_t - \bar{Y})^2}{\sum_{t=1}^T (Y_t - \bar{Y})^2} \quad (29)$$

Tabla 6. Coeficiente de Determinación (R2 Score) obtenido en los 3 métodos implementados.

RNA Entrenada	R2
Método Convencional	99,84
Método Decimación o Reducción de Muestras	97,87
Método de Restricción Crítica de Combinaciones	98,92

Los resultados obtenidos para cada uno de los 3 métodos propuestos permiten visualizar que, sin importar los datos seleccionados para el entrenamiento de la RNA, su desempeño se posiciona por encima del 95%, ratificando la robustez de la arquitectura seleccionada e implementada. El método convencional se posiciona en un primer lugar por encima de los otros métodos propuestos, por una diferencia poco relevante, pero permite identificar el impacto que tiene años de mejoramiento y optimización de las herramientas Keras y Tensorflow.

CONCLUSIONES

Con base en los resultados obtenidos en la Sección 7 de análisis de resultados, se presentaron distintos métodos e indicadores de evaluación para determinar la calidad de la respuesta obtenida para los 3 métodos propuestos en este Trabajo de Grado. En el análisis presentado en la Sección 7.1, el método que presenta mejor desempeño en aspectos como exactitud (Accuracy) al momento de evaluar su predicción con datos distintos a los empleados para su entrenamiento, así como un mejor desempeño en la asignación de sus pesos de neurona (Loss), fue el Método Convencional. También se presenta como mejor opción al revisar los resultados del coeficiente de determinación (R2 Score) en la Sección 7.4, lo que permite concluir que la arquitectura y parámetros de diseño seleccionado para la RNA implementada se ajusta mucho mejor a la entrada de valores aleatorios que a las entradas ordenadas y consecutivas que presentan los métodos de Decimación o Reducción de Muestras y el de Restricción Crítica de Combinaciones.

El MDoRM y MRCC propuestos e implementados como una alternativa al MC, demostraron un desempeño satisfactorio en los indicadores propuestos en la Sección 7.1 y 7.3, con respecto al método de referencia MC. Este resultado permite aseverar que la arquitectura y parámetros de diseño seleccionados para la RNA propuesta en la Sección 6.2.3, fue acertada para el trabajo de predicción de comportamientos en los análisis de confiabilidad presentados en la Sección 2.2 y las restricciones en la cantidad de componentes de la Sección 2.4.6.

El presente Trabajo de Grado demostró la viabilidad de implementar una solución alterna como son las aplicaciones de Redes Neuronales y el Machine Learning en el campo del dimensionamiento de Microrredes aisladas, al reducir los volúmenes de datos requeridos en el proceso de entrenamiento para el MC (10%, 1%, 0.1%), el MDoRM (0.1%) y el MRCC (1.24%) para la predicción de los indicadores económicos y de confiabilidad de la MG (LPSP, LOLH, COST), determinando un

valor mínimo global de combinaciones (405 elementos) con una predicción final de 97.31% de accuracy.

RECOMENDACIONES

El presente TG tuvo como objetivo realizar un primer acercamiento en el uso de herramientas como las RNA y el Dimensionamiento de Microrredes aisladas, evaluando una alternativa de reducir el tamaño de la muestra de datos requeridos para las etapas de training y test. La selección de la arquitectura de la RNA se basó en un solo tipo de aplicación para la detección de patrones, la cual fue lo suficientemente robusta al momento de procesar los métodos propuestos de Dataset. Se recomienda para trabajos futuros explorar otros tipos de arquitecturas, parámetros y configuraciones de RNA, con el fin de profundizar más en la optimización del código y el algoritmo de entrenamiento.

También se propone evaluar el desempeño de la RNA implementada en diferentes casos de estudio, contemplando la posibilidad de aumentar la complejidad en la arquitectura de la MG, así como en el estudio de más indicadores de confiabilidad presentes en el dimensionamiento que no se tuvieron en cuenta en este trabajo.

BIBLIOGRAFÍA

OFFICE OF ELECTRICITY DELIVERY AND ENERGY RELIABILITY SMART GRID R&D PROGRAM. *DOE Microgrid Workshop Report*. San Diego, California: DEPARTMENT OF ENERGY U.S.A, 2011.

UBILLA, K. JIMÉNEZ ESTÉVEZ, G. A. HERNÁNDEZ, R. REYES CHAMORRO, L. HERNÁNDEZ IRIGOYEN, C. PALMA BEHNKE, R. SEVERINO, B. «Smarts Microgrids as a Solution for Rural Electrification: Ensuring Long-Term Sustainability Through Cadastre and Business Models,» *IEEE Transactions on Sustainable Energy*, vol. 5, nº 4, pp. 1310-1318, 2014.

GONZÁLEZ, N. CUSGÜEN, C. MOJICA NAVA, E. PAVAS, A. «Control Strategies for Power Quality in Rural Microgrids,» *UIS ingenierías*, vol. 16, nº 2, pp. 93-103, 2017.

JIMÉNEZ VARGAS, I. OSMA PINTO, G. REY LÓPEZ, J. M. «Sizing of autonomous microgrid considering life cycle emissions,» 2020.

LIAN, J. ZHANG, Y. MA, C. YANG, Y. CHAIMA, E. «A review on recent sizing methodologies of hybrid renewable energy systems,» *Energy Conversion and Management*, vol. 199, 2019.

REY LÓPEZ, J. M. VERGARA, P. P. SOLANO MARTÍNEZ, J. ORDOÑEZ PLATA, G. «Design and optimal sizing of microgrids,» Springer, pp. 337-367, 2019.

GUEMBE ZABALETA, J. SAN MARTÍN BIRRIUN, I. PASCUAL MIQUELEIZ, J. M. «Cálculo del estado de carga en baterías de plomo-ácido: diseño y validación experimental,» Pamplona, 2016, p. 9.

BUKAR, A. L. TAN, C.W. LAU, K. Y. «Optimal sizing of an autonomous photovoltaic/wind/battery/diesel generator microgrid using grasshopper optimization algorithm,» *Solar Energy*, vol. 188, pp. 685-696, 2019.

BURGOS, F. J. P. Herramientas en GNU/LINUX para estudiantes universitarios/- Redes neuronales con GNU/LINUX, The Public's Library and Digital Archive, 2003.

ANDRADE TEPÁN, E.C. Estudio de los principales tipos de redes neuronales y las herramientas para su aplicación. Universidad Politécnica Salesiana Sede Cuenca. 2013.

BASOGAIN OLABE, Xabier. Redes neuronales artificiales y sus aplicaciones. Escuela Superior de Ingeniería de Bilbao, EHU.

BHAVSAR, Krunal; SHAH, Vrutik; GOPALAN, Samir. Scrumbanfall: an agile integration of scrum and kanban with waterfall in software engineering. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, 2020, vol. 9, no 4, p. 2075-2084.

IZAURIETA, F. SAAVEDRA, C. Redes neuronales artificiales. Departamento de Física, Universidad de Concepción Chile. 2000

RUIZ, C. A. BASUALDO, M. S. Redes neuronales: conceptos básicos y aplicaciones, Rosario, México, 2001.

MATICH, Damián J. Redes Neuronales: Conceptos Básicos y Aplicaciones. Informática Aplicada a la Ingeniería de Procesos – Orientación I. Universidad Tecnológica Nacional – Facultad Regional Rosario. Marzo 2001.

GÉRON, Aurélien. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow. Concepts, Tools, and Techniques to Build Intelligent Systems. Second Edition. O'Reilly Media Inc. 2019.

SALEH, Hyatt. Machine Learning Fundamentals. Packt Publishing Editorial. 2018.

OPPENHEIM, Alan V. SCHAFFER, Ronald W. BUCK, John R. Tratamiento de señales en tiempo discreto. Pearson Educación, 2000.

ANEXOS

Anexo 1. Seudocódigo del algoritmo de entrenamiento de la RNA usada en el Trabajo de Grado.

Programa: Algoritmo de entrenamiento

Módulo: carga del dataset

```
import numpy as nup
trainingdata = nup.loadtxt ("datos_training.csv", delimiter=",")
```

Módulo: normalización de datos

```
from sklearn.preprocessing import StandardScaler
trainingdata _x= trainingdata[:,0:4]
trainingdata _y= trainingdata[:,4:7]
ssc = StandardScaler()
trainingdata _x_sd=ssc.fit_transform(trainingdata _x)
trainingdata _y_sd=ssc.fit_transform(trainingdata _y)
X_train = trainingdata _x_sd
Y_train = trainingdata _y_sd
```

Módulo: descripción de RNA

```
import tensorflow as tf
model = tf.keras.Sequential()
model.add(tf.keras.layers.Dense(1000,activation='relu', input_shape = (4,)))
model.add(tf.keras.layers.Dense(3,activation='linear'))
model.compile(loss='mean_squared_error', optimizer='sgd', metrics=['accuracy'])
model.fit(X_train, Y_train, batch_size=10,epochs=150)
scores = model.evaluate(X_test, Y_test)
model.save('model.h5')
```

Módulo: entrenamiento RNA

```
testingdata = nup.loadtxt ("datos_testing.csv", delimiter=",")
testingdata _x= testingdata[:,0:4]
testingdata _x_sd=ssc.fit_transform(testingdata _x)
X_test = testingdata _x_sd
prediction=model.predict(X_test)
```
