Fault Indicators Allocation to Maximize the Performance of a Fault Locator Based on Artificial Intelligence

Juan Marín-Quintero[1]
Energy Department
Universidad de la Costa
Barranquilla, Colombia
jmarin14@cuc.edu.co

César Orozco-Henao[2]
Department of Electrical and Electronic Engineering
Universidad del Norte
Barranquilla, Colombia
chenaoa@uniorte.edu.co

Andrés Herrera-Orozco[3]
Department of Electrical Engineering
Universidad Tecnológica de Pereira
Pereira, Colombia
arherrera@utp.edu.co

***Abstract*— Fault location (FL) is one of the main challenges in Advanced Distribution Automation (ADA) of Active Distribution Networks (ADN). One of the commonly used strategies by utilities to deal with this challenge is the use of Fault Indicators (FIs), which indicate to the operator the path taken by the fault current. However, a good performance of this scheme depends on the number of installed devices, a high number of them could cause a high cost for the utility investment planning. In this context, this paper presents an artificial intelligence-based fault location strategy that determines the number and location of FI into ADN to maximize performance in fault section estimation. To achieve this objective, the ADN is divided into sections, and the FL problem is modeled as a classification problem to train an Artificial Neural Network (ANN). To determine the number of FIs to be installed and their location, the strategy uses the three-phase current magnitudes measured by the FI as features for an ANN model. Also, the strategy uses a feature selection and tuning scheme based on a multiverse optimization algorithm (MOA) to identify the features that maximize the accuracy of the ANN model. The strategy was validated on the IEEE123-node test feeder. The results showed accuracy close to 99.4% with a reduction of 40% of the number of FIs when compared with other method. The strategy shows its simplicity and promising prospects to apply it in the utility's investment planning.**

Keywords: Fault indicators, Fault Location, Microgrids, Artificial Neural Networks

I. Introduction

Currently, significant efforts have been made to achieve an energy transition to clean energies with the aim of reducing greenhouse gas emissions. One of the main allies of this transition is the electrification of processes that use fossil fuel-based energy resources such as transportation [1]. However, this requires the massive integration of non-conventional renewable energy sources throughout the electric power value chain, including distribution systems. Resources such as Distributed Generation (DG) and Energy Storage System (ESS) integrated into distribution networks have typically been referred to as Distributed Energy Resources (DER).

The integration of DER has driven the modernization of distribution networks, transforming them into Active Distribution Networks (ADN). These networks are characterized by integrating automation and control functionalities through an ADA infrastructure, which involve a fault location (FL), isolation and service restoration (FLIRS) functionalities to improve system reliability [2]. One of the main tasks of FLIRS is fault location. The specialized literature has dealt with this task by using methods such as fault apparent impedance estimation [3]–[5], traveling waves [6], [7] and artificial intelligence techniques [8]–[12]. However, several utilities worldwide have chosen to carry out this task by using FIs to reduce the multiple estimation presented by the previously methods [13].

Some FL methods based on FIs have been proposed in the technical literature. In [13] is presented a strategy for the allocation of FIs to eliminate the multiple estimation of the fault point presented by apparent impedance-based FL methods. The study considers the effect of DG integration on the allocation of FIs. However, the FIs do not improve the performance of the impedance-based FL method and the methodology does not consider the presence of microgrids.

On the other hand, in [14] a FL method using smart meters and FIs is proposed. The method proposes an analytical model based on Mixed Integer Linear Programming (MILP) where each hypothetical FL is modeled as decision variables. However, although the FL method considers the integration of DG, the problem is not solved when having high DG penetration or integrated microgrids. In addition, the method does not determine the number of smart meters and FIs for improve FL performance. In the same way, in a study by [15], a MILP model was introduced to determine the optimal number and placement of fault management equipment within a distribution network. This fault management equipment encompasses FIs, manual switches, and remotely controlled switches (RCS). The objective function of this method involves various factors, such as the cumulative cost associated with FIs and sectionalizing switches. However, while the formulation of the problem does incorporate considerations for locating and optimizing the number of FIs, the potential influence of DG and the integration of microgrids was not considered.

Alternatively, in a recent work by [16], a fault identification methodology was introduced, which is integrated into an Asset Management System (AMS). This AMS effectively combines data originating from FIs, a Distribution Dispatching Control System (DDCS), and a Feeder Dispatching Control System (FDCS). Their fault identification model, based on Petri-net technology, was constructed using an Automated Mapping/Facilities Management/Geographic Information System (AM/FM/GIS). The AMS draws upon information

Submitted to the 23rd Power Systems Computation Conference (PSCC 2024).

from the DDCS and the FDCS to access the statuses of FIs and the loads on feeders and laterals. This supports quick FL and isolation in conventional distribution systems. However, the proposed method, while effective for conventional systems, does not incorporate DERs or microgrids and the optimization of the number of FIs was not addressed in their FL methodology, which could potentially enhance FL efficiency. Also, in [17], the concept of a "fault zone" is introduced, referring to the part of an electrical system where FI devices activate. The research proposes a method focused on strategically placing FIs in system locations that lead to smaller fault zones. This approach highlights the importance of sitting FIs near areas with high power demand to reduce service interruption costs. However, this proposal specifically addresses radial distribution networks and does not consider the unique characteristics of ADN, including aspects related to DG and microgrid integration. Moreover, the method does not explore the optimization of the number of FIs but assumes a fixed quantity of FIs to be used.

Similar to [14], in [18] an algorithm is introduced for the sequential determination of outage scenarios by using data from various sensors, including smart meters and remote fault indicators (RFIs). The research presents an evidence-driven, rule-based search algorithm as a solution to address outage management challenges in ADN. This approach is capable of handling multiple simultaneous failures and, consequently, can estimate with precision the FL either along the primary feeder or on a lateral line. However, while the proposed algorithm considers the integration of DG into the network, it does not account for microgrid integration, topological changes, or the optimization of the number of FIs. The study presented in [19] introduces a novel index termed the Load Point Long-term Interruption Frequency Index (LLIFI) with a focus on enhancing the reliability of specific buses within a distribution network. This index is designed to quantify the long-term interruption statistics associated with these buses. Its application is primarily aimed at optimizing the placement of FIs and RCS throughout the distribution system. However, this method is centered on enhancing bus reliability without addressing the specifics of FL. It does not explore the relationship between the quantity of FIs and FL accuracy. Additionally, the utilization of this index increases the overall system cost and results in an elevated number of automation devices positioned near designated buses. In the study [20], an approach to identify a fault sections within a distribution system is introduced. This approach leverages ANNs to validate FIs. The ANN model is trained using data to establish correlations between fault current information and the validation outcomes of FIs. One feature of this method is its ability to identify fault sections even in scenarios where reverse fault currents originate from the downstream side of the FL. However, the proposed method does not consider aspects such as the integration of microgrids, network topological changes, or the influence of the quantity of FIs on the accuracy of the FL process.

In this way, considering the challenges previously presented, this paper addresses the complexities associated with FL into ADN by establishing an artificial intelligence-driven approach. Table I provides a comprehensive comparison between the proposed strategy and relevant aspects analyzed in referenced works. The strategy aims to optimize the performance of FIs in fault section estimation within the ADN. The key contributions of this strategy are as follows:

- Considering the FI allocation problem into the formulation of FL method based in artificial intelligence using only the current magnitudes of the involved lines section.
- Eliminating of the need for directional FIs installation, which means the proposed strategy can be implemented in ADNs without requiring additional equipment or additional communication infrastructure.
- Considering critical characteristics of the ADN, including inherent network imbalances, topological variations, integration of microgrids, and the high penetration of DER.

The paper is structured as follows. Section II describe the FL formulation. Section III introduces the fault locator based on artificial intelligence which is integrated in the optimal FI allocation. Section IV outlines the cases of studies. Section V presents the results and engages in a comprehensive discussion. Finally, Section VI highlight conclusions drawn from this study.

TABLE I. ASPECTS CONSIDERED TO DEVELOP THE PROPOSED STRATEGY.

Considered aspect	**Methods - Reference**								
	[13]	[14]	[15]	[16]	[17]	[18]	[19]	[20]	**Proposed strategy**
Uses some fault location method.	✓	✓	✓	✓	x	✓	x	✓	✓
Eliminates the multiple estimates.	✓	✓	✓	✓	✓	✓	✓	✓	✓
FIs installation considered to improve the performance of FL strategy.	x	✓	x	x	x	✓	x	x	✓
Considers the presence of microgrids	x	x	x	x	x	x	x	x	✓
No needing to use of directional FIs	x	x	x	-	-	x	x	✓	✓
Determines the number of FIs that optimize the FL accuracy.	x	x	✓	x	x	x	x	x	✓
Considers topological changes.	x	x	x	x	x	x	✓	x	✓
Considers high DERs integration.	✓	x	x	x	x	x	✓	✓	✓
Includes aspects about the fault characteristics.	x	x	x	x	x	x	x	x	x

x Not considered; ✓ Considered; - It is not necessary.

II. FAULT LOCATION FORMULATION

This section describes FL method-based on artificial intelligence and the use of FIs to determinate a fault zone. The proposed method can be shown in Fig. 1. To implement the method is necessary to split an ADN into several zones.

The fault location problem is modeled as a classification problem with N labels, where the number of labels is equal to the number of zones into which the ADN is divided. Consider also, the occurrence of P faults in the ADN, where each fault scenario $\boldsymbol{x}_{pk} = \{x_{p1}, x_{p2}, x_{p3}, \cdots x_{pk}\}$, is represented by a set of k features, defined by the current magnitudes measured by each FI. Therefore, considering the set of P fault scenarios as a training data set, with a dimensional space k, an ANN-based learning model is formulated in (1) to solve the FL problem as a classification problem.

$$y_i(x_p) = \sigma\left(\sum_{r=1}^{m} w_r * h\left(\sum_{q=1}^{k} w_{rq} * x_q\right) + W_0\right) \qquad (1)$$
$$s.t$$
$$x_p \in R^k \text{ and } y_i \in \{1, 2, \cdots, N\}$$

Where, w_r and w_{rq} are the weight parameters, m is the number of neurons in the hidden layer, σ and h are the active functions, and W_0 is the bias parameter used by the activation function. [21].

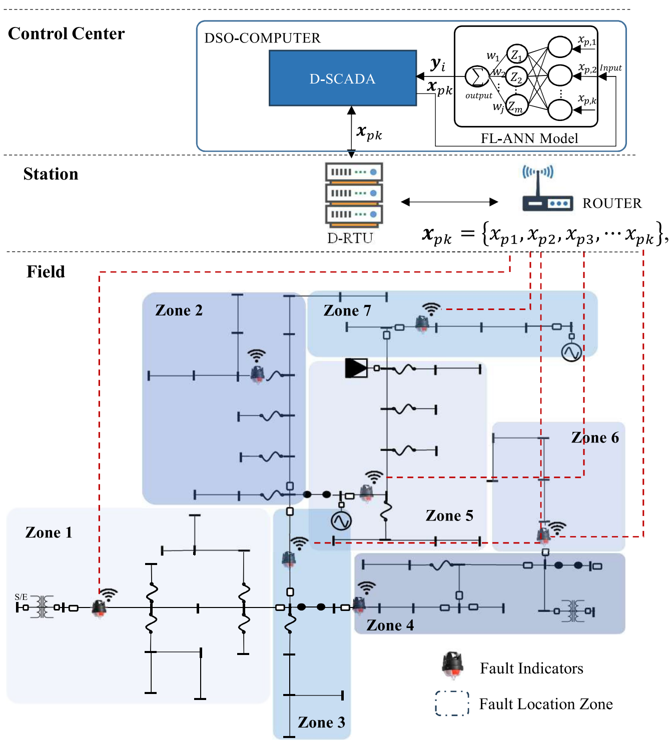

Fig 1. FL formulation as a classification problem.

The complexity of the FL problem and the ANN model training depends on the number of labels or zones into which the ADN is divided. While higher are the number of zones into which the system is divided, then higher are the number of classes of the ANN model [22]. Therefore, more precise tuning of the model is necessary. Therefore, to maximize the performance of the ANN model, a features selection and tuning technique based on a Multiverse Optimization Algorithm (MOA) is implemented and explained in the following section [23].

III. Fault Locator Based on Artificial Intelligence

The performance of the ANN model as a fault zone classifier depends on the location of the FIs, since the current magnitudes recorded by these devices correspond to the features of the ANN model. Therefore, the objective function of the MOA is to determine the number of FIs devices and their location into ADN that maximizes the performance of the ANN model. Additionally, it determines the tuning of the ANN model parameters. Fig. 2 presents the flowchart of the optimal FIs allocation technique as a feature selection and tuning strategy based on the MOA. The steps involved in this technique are explained below.

Stage 1:Data Generation

A significant number of faults must be considered to obtain the ANN model as a classifier of the fault zone. However, the number of faults occurrences into ADN are few since they are designed to minimize their probability of fault. Therefore, it is necessary to generate a synthetic fault database by automatic simulation of ADN faults. The synthetic data set is obtained by cooperative work between an Electromagnetic Transient (EMT) software and a numerical computation software.

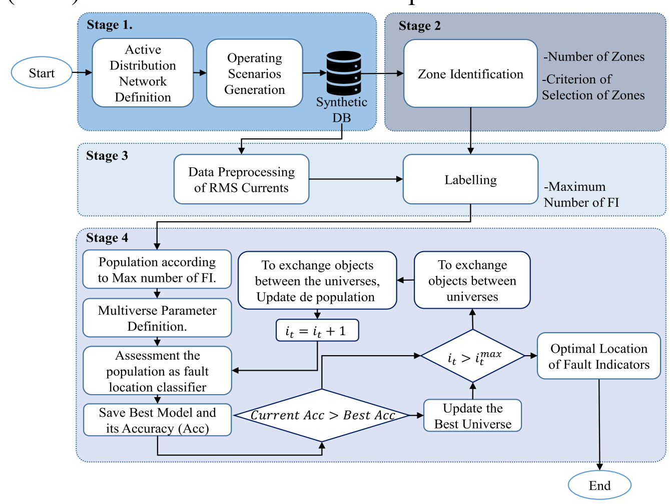

Fig 2. Features selection technique based on the multiverse optimization algorithm.

The factors that represent the behavior of ADN such as load variation, topological changes, connection, or disconnection of DERs and microgrid mode can be considered. Applying these operative conditions, the fault scenarios are simulated to then varying factors as type of fault, fault resistance and fault location. Then, the RMS values of the current into the terminals of all line sections are extracted. Table II presents the factors and levels used in the proposed formulation [8], [24].

Stage 2: Zones definition

Splitting the ADN into zones is not an easy task, it process can consider different factors. Some of these factors are related to system characteristics, such as the geographical area covered by each zone, the type of line (overhead or underground), and the failure rate of each line segment within the system. Another factor is associated with the type of user, including their capacity, priority, and the option of power supply restoration through reconfiguration [22]. Additionally, factors related to the geographical placement of the zones, such as ease of access, can also be used. An example of zoning is presented in Fig. 3.

Stage 3: Preprocessing and labelling

At this stage the fault scenario database is standardized by (2). This process is necessary for improving the performance ANN models [25].

$$x_{fm}^{st} = \frac{x_{fm} - \mu_m}{\sigma_m} \quad (2)$$

Where, x_{fm} is the value for the *f-th* fault operating scenario in the *m-th* feature, μ_m is the average for the *m-th* feature, σ_m is the standard deviation for the *m-th* feature, and x_{fm}^{est} is the standardized value for the *f-th* fault operating scenario in the *m-th* feature.

On the other hand, each fault scenario is labeled according to the number zone where the fault was simulated.

TABLE II. FACTORS AND LEVELS COMMONLY USED IN ADN OPERATING SCENARIOS.

Group	Factor	Levels	Number of levels
No-fault operation	Topology change	**T1**: S_0, S_1, S_2, S_4, S_5 close, S_3, S_6, S_7, S_8 open; **T2**: S_0, S_1, S_3, S_4, S_5 close, S_2, S_6, S_7, S_8 open; **T3**: S_0, S_1, S_2, S_3, S_5 close, S_4, S_6, S_7, S_8 open; **T4**: S_0, S_1, S_2, S_3, S_4 close, S_5, S_6, S_7, S_8 open; **T5**: S_0, S_2, S_3, S_4, S_5 close, S_1, S_3, S_6 open; **T6**: S_0, S_1, S_2, S_4, S_6 close, S_3, S_5, S_7, S_8 open, **T7**: $S_0, S_1,, S_4, S_5$ close, $S_2, S_3, S_5, S_6, S_7, S_8$ open,. **T8**: S_0, S_1, S_4, S_6 close, S_2, S_3, S_5, S_7, S_8 open. **T9**: S_0, S_1, S_2, S_4, S_7 close, S_3, S_5, S_6, S_8 open, **T10**: S_1, S_2, S_4, S_7, S_8 close, S_0, S_3, S_5, S_6 open,	10
	MG Mode	On-grid/off-grid	2
Fault operation	Fault type	Single-phase faults, double-phase faults, double-phase to ground faults and three-phase faults	10
	Fault location	30, 55 and 75% of all single-phase lines and 50% of all three-phase lines	162 60
	Fault resistance	Single-phase faults: 0Ω to 90Ω phase-phase faults: 0Ω to 40Ω	10 5

Stage 4: Optimal allocation of FIs based in MOA

In this stage, a MOA is implemented to determine the number of FIs and their location, which define the combination of features that maximizes the accuracy of the ANN model. The concept behind the MOA draws its inspiration from the multi-verse theory in physics, particularly concerning to global optimization as mentioned in references [23], [26]. Here the solutions are denominated as universes. Therefore, the modeling of the problem considers the accuracy of the ANN model as a fitness function and the universes that make up the population are represented by a mixed coding, as shown in Fig. 4.

To find the optimal solutions MOA employs three fundamental notions: black holes, white holes, and wormholes. Furthermore, the method incorporates two dynamic coefficients aimed at enhancing performance: The Travelling Distance Rate (TDR) and the Wormhole Existence Probability (WEP). While TDR and WEP could be treated as constants, a strategic choice has been made to transition them from fixed values to variables influenced by ongoing and maximum iterations. These parameters are given by (3) and (4).

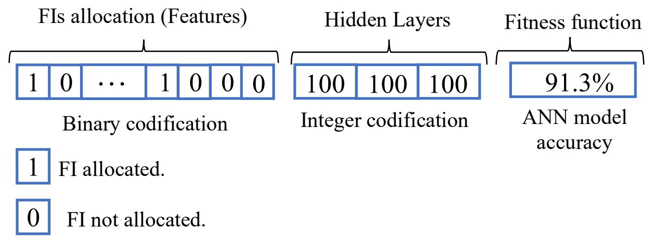

Fig. 4. Coding of the optimal allocation problem of FIs.

$$TDR = 1 - \frac{it^{0.166}}{it_{max}^{0.1666}} \quad (3)$$

$$WEP = WEP_{min} + it * \left(\frac{WEP_{max} - WEP_{min}}{it_{max}}\right) \quad (4)$$

where *it* is the current iteration and it_{max} is the maximum number of iterations.

Algorithm 1 shows the steps that carry out the MOA. Black and white holes play an essential role in enabling the transfer of objects between different universes; highly accurate universes have a significant probability of sending objects to universes with lower accuracy.

Similarly, wormholes transfer objects from the most optimal current universe to another universe. These fundamental concepts allow the MOA to perform exploration, exploitation, and search within the solution set. The algorithm of the multiverse optimizer is described below:

Algorithm 1 - Multiverse Optimizer

Input: *Database with the whole features*

Output: *ANN model. Features selection and hyperparameter*

Begin

1: *Set the initial maximum number of hidden layers (HL_{max}), maximum number of neurons for each hidden layer(HLN_{max}), TDR, WEP_{max}, WEP_{min} values; number of universes UN, it_{max}.*
2: *To obtain the initial population*
3: *it=1*
4: ***While*** $it \leq it_{max}$ ***repeat***
5: *Assessing (Accuracy) each universe*
6: ***If*** Acc_{it}>*Incumbent*
7: *Incumbent*=Acc_{it}
8: ***end***
9: *To exchange the FI allocations and hyperparameters between universes using black and white holes*
10: *Set t = t + 1. {Iteration counter increasing}*

Updating of WEP and TDR

End While

Produce the best solution for FIs allocation: ANN model, attributes (FIs allocation) and hyperparameters

End For

End

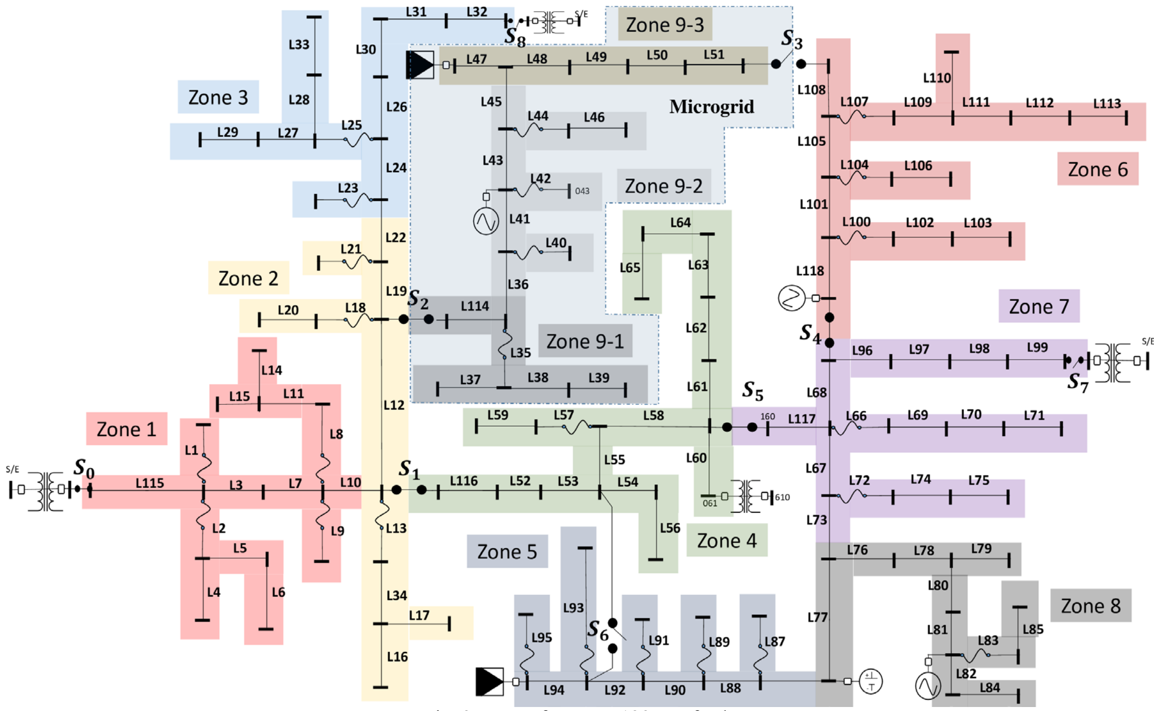

Fig. 3. Zones for IEEE 123 test feeder.

IV. CASE OF STUDY

The proposed fault locator was validated on a modified IEEE 123 node test feeder operating at a nominal voltage of 4.16 kV [27]–[29]. This network is characterized by overhead and underground lines, unbalanced load, four voltage regulators, shunt capacitor banks, and multiple switches. The test feeder was modified by adding six DERs along the grid: three non-inverters connected DERs (synchronous generator), and three inverters connected DERs (two PV systems and one battery energy storage system). In addition, this system has six switches that allow reconfiguration processes to be performed. Fig. 3 shows the modified 123-node IEEE test feeder. The FIs can be located on any line section of the ADN.

On the other hand, the ADN has the possibility of connecting and disconnecting a Microgrid (MG), which is delimited by the switches S2 and S3. The microgrid consists of 18 nodes corresponding to zone 9 of the ADN. This zone is divided into three sub-zones called Zone 9-1, 9-2, and 9-3 respectively.

The methodology is validated by considering two scenarios. The first scenario is associated with the whole ADN (including the MG) and the second scenario just considers the MG. In this way, we will obtain the best combination of FIs that improve the performance of fault locator in both scenarios.

The fault locator was validated considering the operational scenarios presented in Table II.

Finally, the total number of fault scenarios are determined for each line configuration. For single-phase lines, 14580 fault scenarios were simulated, for double-phase lines 720 fault scenarios were simulated and for three-phase lines 14400 faults were simulated. Additionally, the criterion used for establishing the zones was the distance covered by each zone, which is reported in Table III.

TABLE III. LINE LENGTH

Zone	Length (m)	Number of lines
1	3775	14
2	3450	10
3	3525	11
4	4950	15
5	5075	17
6	5950	15
7	5100	15
8	3325	9
9	3350	11

V. RESULT AND DISCUSSION

The results obtained by validating the methodology for the described scenarios in section IV are presented in sections A, B and C. All the results were obtained by using a PC with the following characteristics: CPU AMD Ryzen 7 5800H with Radeon Graphics 3.20 GHz-RAM 16GB and hard drive SSD 500GB.

A. Performance in the main network (ADN)

To show the advantages of the optimal allocation of FIs to maximize the performance of the ANN model as a fault locator, the maximum number of FIs to be considered was varied from 10 to 55 in steps of 5 as shown in Fig. 5.

The results presented in Fig. 5 show that for locating faults in the main network (ADN), the number of FIs that maximizes the performance of the fault locator is approximately 40 FIs and its accuracy reaches 99.4%. This finding underscores the existence of an optimal number of FIs that effectively maximizes the accuracy of the FL method. Moreover, the use of the MOA for feature selection within the ANN demonstrates a highly satisfactory outcome in the placement of FIs.

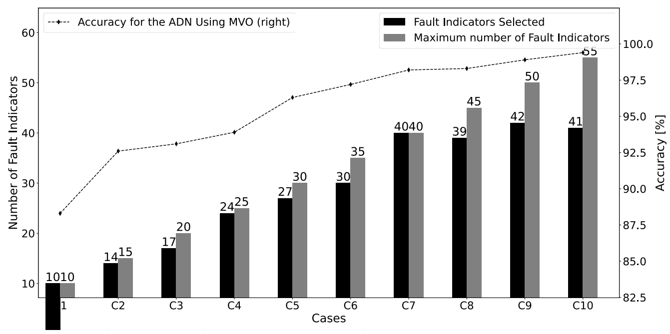

Fig.5. Performance of the methodology for the optimal FIs location into the ADN.

Additionally, Table IV presents the confusion matrix of the ANN model within the evaluated cases. This matrix provides a visual representation of the FL method's robustness, even in the face of dynamic features in an ADN, such as topology changes and the operation modes of the microgrid. Fault identification in zone 3 emerges as the most challenging aspect, achieving an accuracy of 96.7% among 568 faults analyzed within this zone. Specifically, of the 568 faults analyzed in Zone 3, 549 were correctly identified within the same zone, while the remaining 19 were misplaced in Zone 1, i.e., outside their designated zone. This analysis can be extrapolated to other areas.

TABLE IV. CONFUSION MATRIX FOR THE TENTH CASE

99.4%	Z1	Z2	Z3	Z4	Z5	Z6	Z7	Z8	Z9
Z1	877	0	0	0	0	0	1	0	0
Z2	0	612	0	0	2	0	0	0	0
Z3	19	0	549	0	0	0	0	0	0
Z4	0	0	0	1023	0	0	0	0	0
Z5	0	0	0	0	962	1	3	0	0
Z6	0	0	0	0	0	874	0	0	0
Z7	0	0	0	0	0	0	922	0	0
Z8	0	0	0	0	0	1	12	672	1
Z9	0	0	0	0	0	0	0	0	637

Also, Table V shows in which lines the 41 FIs were placed for the tenth case of the ANN model, also it shows the number of neurons per each hidden layer obtained by the MOA. Some commercial solutions such as that presented in [30] install one FI per bifurcation. If we consider this criterion to locate the FIs and obtain the ANN model as a fault locator, the number of FIs installed and the accuracy obtained could be 79 and 97.9%, respectively, as shown in Table VI.

If we take this solution as our reference case and compare it with the best solution obtained (the tenth case) we observe that the accuracy of both ANN models is comparable and higher than 97%. However, the number of FIs installed are approximately half for the best solution when compared to the number of FIs installed in the reference solution, which implies a reduction in the total cost of the solution. This shows the advantages of formulating the FIs placement problem as a features selection problem for obtaining the ANN model as a fault locator.

TABLE V. FAULT INDICATOR SOLUTION FOR THE TENTH CASE

Fault Indicator 1-20		Fault Indicator 21-40		Fault Indicator 41-60		Fault Indicator 61-80		Fault Indicator 81-100		Fault Indicator 101-118		Hidden Layers	
L1	0	**L21**	0	**L41**	0	**L61**	0	**L81**	0	**L101**	0	**H1**	80
L2	0	**L22**	1	**L42**	1	**L62**	0	**L82**	0	**L102**	0	**H2**	130
L3	0	**L23**	1	**L43**	0	**L63**	0	**L83**	1	**L103**	1	**H3**	82
L4	0	**L24**	1	**L44**	0	**L64**	1	**L84**	1	**L104**	1	**H4**	55
L5	0	**L25**	0	**L45**	0	**L65**	0	**L85**	1	**L105**	0		
L6	1	**L26**	0	**L46**	0	**L66**	0	**L86**	1	**L106**	0		
L7	1	**L27**	0	**L47**	0	**L67**	0	**L87**	0	**L107**	1		
L8	0	**L28**	1	**L48**	0	**L68**	1	**L88**	1	**L108**	1		
L9	0	**L29**	0	**L49**	0	**L69**	0	**L89**	0	**L109**	0		
L10	1	**L30**	0	**L50**	0	**L70**	0	**L90**	0	**L110**	0		
L11	0	**L31**	0	**L51**	0	**L71**	1	**L91**	0	**L111**	1		
L12	1	**L32**	0	**L52**	0	**L72**	1	**L92**	1	**L112**	0		
L13	1	**L33**	0	**L53**	0	**L73**	1	**L93**	1	**L113**	0		
L14	1	**L34**	0	**L54**	0	**L74**	0	**L94**	1	**L114**	0		
L15	0	**L35**	0	**L55**	1	**L75**	1	**L95**	0	**L115**	0		
L16	0	**L36**	0	**L56**	0	**L76**	0	**L96**	1	**L116**	1		
L17	1	**L37**	0	**L57**	0	**L77**	0	**L97**	0	**L117**	1		
L18	1	**L38**	0	**L58**	1	**L78**	1	**L98**	0	**L118**	0		
L19	1	**L39**	0	**L59**	0	**L79**	0	**L99**	1				
L20	0	**L40**	0	**L60**	0	**L80**	0	**L100**	1				

B.Performance into the Microgrid

Fig. 6 presents the accuracy of the proposed strategy implemented for the second scenario. Similar to section A, the number of maximum FIs is placed in a range between 4 to 14 and progressively increased by 2. The results show that from the initial case, the ANN-based FL method achieves an accuracy ranging from 90% to 93.6% with varying placements of FIs.

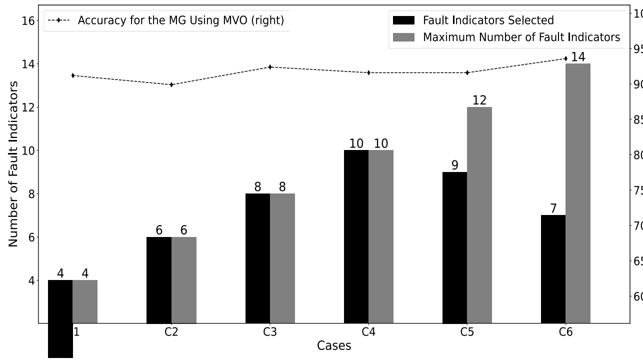

Fig. 6. Performance of the methodology for the optimal FIs location into the MG.

TABLE VI. PERFOMANCE FOR THE BASE CASE.

Number of Fault Indicators	Accuracy [%]	Hidden Layers	Neurons per each Hidden Layer
79	97.9	4	100

From Fig. 6, the ANN model with the highest accuracy is the one obtained in sixth case, achieving a FL accuracy of 93.6% with a total of 7 FIs. Table VII shows the confusion matrix for this case.

TABLE VII. CONFUSION MATRIX FOR THE SIXTH CASE

93.63%	**Z0**	**Z9-1**	**Z9-2**	**Z9-3**
Z0	1182	29	13	1
Z9-1	11	200	14	0
Z9-2	10	14	426	47
Z9-3	1	0	7	353

Zone 9-1 presented the lowest accuracy in zone identification. Of the 225 scenarios evaluated for this zone, only 88.88% of the cases were identified in the correct zone. This can be attributed to the variability of the fault currents recorded by the FIs since the short-circuit currents vary significantly by the ongrid/offgrid modes of operation of the MG. Nevertheless, the FL method showed adequate performance for the cases evaluated. The zone 0 represents all the fault scenarios outside of the MG, here, the methodology obtained the better results for all the classes of the ANN model. Table VII shows the optimal location of the FIs for the sixth case.

TABLE VIII. FAULT INDICATOR SOLUTION FOR THE SIXTH CASE

Fault Indicator 1-6		Fault Indicator 7-12		Fault Indicator 13-18		Hidden Layers	
L35	1	**L41**	1	**L47**	1	**H1**	87
L36	1	**L42**	1	**L48**	0	**H2**	137
L37	0	**L43**	1	**L49**	1	**H3**	108
L38	1	**L44**	1	**L50**	0	**H4**	97
L39	0	**L45**	1	**L51**	1		
L40	0	**L46**	1	**L114**	1		

Finally, Fig. 7 shows the evolution of the performance of the MOA for the tenth ADN case and the sixth case of MG. For the tenth case (Dotted line with rhombuses in the Fig. 7), it is observed that beyond the fourteenth iteration, the algorithm stabilizes its performance, converging towards a fitness function value of 99.4% considering all the evaluated scenarios. This behaviour evidence that there is an optimal number of FIs and placement that maximize the performance of the ANN model as a fault locator.

Similarly, for the microgrid case (Dotted line with squares), we observe that in the sixth iteration, the fitness function reaches a stable value of 93.6%. This outcome affirms the efficacy of the MOA in fine-tuning the ANN.

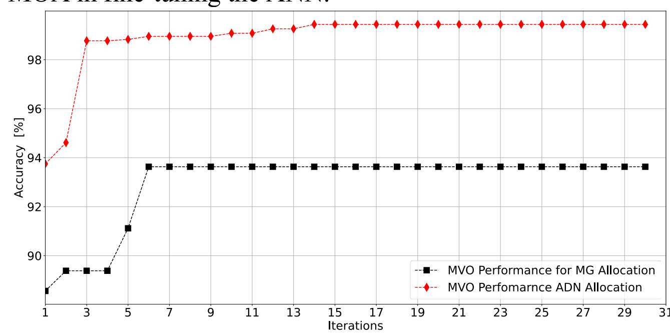

Fig.7. Performance for multiverse optimization algorithm for the tenth case (for main network ADN) and sixth case (for the microgrid).

C. Comparison between optimal location algorithms

With the objective to compare the performance of the proposed method, another metaheuristic optimization technique was used to determinate the optimal number of FIs that improve the fault location of the MG. This optimization technique is well known in technical literature such as Cuckoo Search Algorithm (CSA). Both techniques used the same values for the maximum number of iterations (30) and the size of population (25). Fig 8 shows the performance of the method using CSA and MOA for the MG, additionally, it shows the number of FIs obtained by each technique.

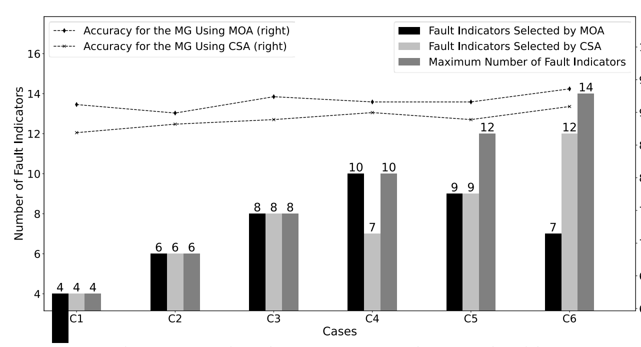

Fig 8 Comparison between CSA and MOA algorithms.

As presented in Fig. 8, the performance of CSA is below of the values obtained by the MOA in different aspects such as number of FIs selected, accuracy and simulation time. For instance, the maximum accuracy obtained by the CSA was 91.2%, contrary to the MOA that was 93.46%. On the other hand, the average simulation time was obtained for MOA and CSA techniques, with 85 and 119 minutes, respectively. The last evidence the potential, robustness, and efficiency of MOA as technique for searching better solutions in the allocation of fault indicators. Also, the accuracy for the fault location algorithm obtained by MOA is 2.7% above of the value obtained by CSA.

VI. Conclusion

In this paper, an artificial intelligence-based FL strategy was presented, focusing on maximizing ADN fault section estimation performance by allocating FIs. The strategy shows that dividing the ADN into sections (zones or labels) to model the FL problem as a classification problem and considering the RMS currents measured by each FI as features for the ANN model, allows determining the optimal number of FIs and their location that maximizes the accuracy of the fault zone estimation. Then, this strategy not only allows obtaining an ANN model with high accuracy for the estimation of the fault zone but also reduces the investment of the assignment of FIs in a distribution network by the Distribution System Operator (DSO), since it determines the optimal number of IFs that should be installed. The proposed strategy was tested on the modified IEEE 123-node test feeder, considering the dynamic characteristic of ADN, through the presence of topology changes and on-grid/off-grid operation modes of microgrids. The results obtained showed that the number of FIs that maximize the precision of the ANN model trained as an estimator of the fault zone is 41 for the main network (ADN), achieving an accuracy of 99.4%. For the microgrid, with at least 4 FIs an accuracy greater than 92% is achieved. Additionally, the MOA evidence its robustness and efficiency when the results were compared with other optimization technique. The proposed strategy showed its simplicity, and convenient to implement a model to locate fault into a real ADN.

Acknowledgements

The authors would like to thank to the program of Electrical Engineering, the master's program in electrical engineering and the academic vice-rector of the Universidad Tecnologica de Pereira for the financial support received to present the results of this research. Likewise, this work was supported by Universidad del Norte, and Departamento Administrativo de Ciencia, Tecnología e Innovación - COLCIENCIAS (now Colombian Ministry of Science, Technology, and Innovation - Minciencias) by call contest ‘‘Convocatoria 852 - Conectando conocimiento de 2019’’ – Project Integra2023, code 111085271060, contract 80740-774-2020.

References

[1] International Renewable Energy Agency (IRENA), “World energy transitions outlook 2023,” 2023. [Online]. Available: https://irena.org/Digital-Report/World-Energy-Transitions-Outlook-2022%0Ahttps://irena.org/publications/2021/March/World-Energy-Transitions-Outlook

[2] Cisco Systems, “Distribution Automation - Secondary Substation Design Guide,” 2019. [Online]. Available: https://www.cisco.com/c/en/us/td/docs/solutions/Verticals/Distributed-Automation/Secondary-Substation/DG/DA-SS-DG.pdf

[3] A. S. Bretas, C. Orozco-Henao, J. Marín-Quintero, O. D. Montoya, W. Gil-González, and N. G. Bretas, “Microgrids physics model-based fault location formulation: Analytic-based distributed energy resources effect compensation,” Electr. Power Syst. Res., vol. 195, no. October 2020, 2021, doi: 10.1016/j.epsr.2021.107178.

[4] C. G. Arsoniadis and V. C. Nikolaidis, “Precise Fault Location in Active Distribution Systems Using Unsynchronized Source Measurements,” IEEE Syst. J., vol. 17, no. 3, pp. 4114–4125, 2023, doi: 10.1109/JSYST.2023.3279232.

[5] D. Lu and Y. Liu, “Fault location for general AC/DC transmission lines: Multi-phase, non-homogeneous, partially mutually coupled and multi-terminal lines,” Electr. Power Syst. Res., vol. 222, no. April, p. 109484, 2023, doi: 10.1016/j.epsr.2023.109484.

[6] Y. Xu, C. Zhao, S. Xie, and M. Lu, “Novel Fault Location for High Permeability Active Distribution Networks Based on Improved VMD and S-transform,” IEEE Access, vol. 9, pp. 17662–17671, 2021, doi: 10.1109/ACCESS.2021.3052349.

[7] X. Wang et al., “Fault location based on variable mode decomposition and kurtosis calibration in distribution networks,” Int. J. Electr. Power Energy Syst., vol. 154, no. August, p. 109463, 2023, doi: 10.1016/j.ijepes.2023.109463.

[8] J. Atencia-De la Ossa, C. Orozco-Henao, and J. Marín-Quintero, “Master-slave strategy based in artificial intelligence for the fault section estimation in active distribution networks and microgrids,” Int. J. Electr. Power Energy Syst., vol. 148, no. June 2022, p. 108923, 2023, doi: 10.1016/j.ijepes.2022.108923.

[9] A. Farughian, L. Kumpulainen, and K. Kauhaniemi, “Review of methodologies for earth fault indication and location in compensated and unearthed MV distribution networks,” Electr. Power Syst. Res., vol. 154, pp. 373–380, 2018, doi: 10.1016/j.epsr.2017.09.006.

[10] A. Ghaemi, A. Safari, H. Afsharirad, and H. Shayeghi, “Accuracy enhance of fault classification and location in a smart distribution network based on stacked ensemble learning,” Electr. Power Syst. Res., vol. 205, no. December 2021, p. 107766, 2022, doi: 10.1016/j.epsr.2021.107766.

[11] Z. Li et al., “Cognitive knowledge graph generation for grid fault handling based on attention mechanism combined with multi-modal factor fusion,” Comput. Electr. Eng., vol. 111, no. PA, p. 108855, 2023, doi: 10.1016/j.compeleceng.2023.108855.

[12] P. Stefanidou-Voziki, N. Sapountzoglou, B. Raison, and J. L. Dominguez-Garcia, “A review of fault location and classification methods in distribution grids,” Electr. Power Syst. Res., vol. 209, no. April, p. 108031, 2022, doi: 10.1016/j.epsr.2022.108031.

[13] R. F. G. Sau, V. P. Dardengo, and M. C. de Almeida, “Allocation of fault indicators in distribution feeders containing distributed generation,” Electr. Power Syst. Res., vol. 179, no. October 2019, p. 106060, 2020, doi: 10.1016/j.epsr.2019.106060.

[14] Y. Jiang, “Data-driven fault location of electric power distribution systems with distributed generation,” IEEE Trans. Smart Grid, vol. 11, no. 1, pp. 129–137, 2020, doi: 10.1109/TSG.2019.2918195.

[15] B. Li, J. Wei, Y. Liang, and B. Chen, “Optimal Placement of Fault Indicator and Sectionalizing Switch in Distribution Networks,” IEEE Access, vol. 8, pp. 17619–17631, 2020, doi: 10.1109/ACCESS.2020.2968092.

[16] T. T. Ku, C. S. Li, C. H. Lin, C. S. Chen, and C. T. Hsu, “Faulty Line-Section Identification Method for Distribution Systems Based on Fault

Indicators," IEEE Trans. Ind. Appl., vol. 57, no. 2, pp. 1335–1343, 2021, doi: 10.1109/TIA.2020.3045672.

[17] G. G. Santos and J. C. M. Vieira, "Optimal Placement of Fault Indicators to Identify Fault Zones in Distribution Systems," IEEE Trans. Power Deliv., vol. 36, no. 5, pp. 3282–3285, 2021, doi: 10.1109/TPWRD.2021.3101671.

[18] Y. Jiang, "Outage Management of Active Distribution Systems With Data Fusion From Multiple Sensors Given Sensor Failures," IEEE Trans. Power Deliv., vol. 38, no. 3, pp. 1891–1903, 2023, doi: 10.1109/TPWRD.2022.3227184.

[19] M. Gholami, I. Ahmadi, and M. Pouriani, "Optimal placement of fault indicator and remote-controlled switches for predetermined reliability of selected buses," IET Gener. Transm. Distrib., vol. 17, no. 12, pp. 2799–2810, 2023, doi: 10.1049/gtd2.12854.

[20] J.-U. S. Myong-Soo Kim, Jae-Guk An, Yun-Sik Oh, Seong-Il Lim, Dong-Hee Kwak, "A Method for Fault Section Identification of Distribution Networks Based on Validation of Fault Indicators Using Artificial Neural Network," Energies, vol. 16, pp. 1–14, 2023, doi: 10.3390/ en16145397 Academic.

[21] C. M. Bishop, Pattern Recognition and Machine Learning (Information Science and Statistics). Berlin, Heidelberg: Springer-Verlag, 2006.

[22] E. Correa-Tapasco, J. Mora-Flórez, and S. Perez-Londoño, "Performance analysis of a learning structured fault locator for distribution systems in the case of polluted inputs," Electr. Power Syst. Res., vol. 166, no. August 2018, pp. 1–8, 2019, doi: 10.1016/j.epsr.2018.09.016.

[23] E. Hosseini, K. Z. Ghafoor, A. Emrouznejad, A. S. Sadiq, and D. B. Rawat, "Novel metaheuristic based on multiverse theory for optimization problems in emerging systems," Appl. Intell., vol. 51, no. 6, pp. 3275–3292, Jun. 2021, doi: 10.1007/s10489-020-01920-z.

[24] J. Marín-Quintero, C. Orozco-Henao, J. C. Velez, and A. S. Bretas, "Micro grids decentralized hybrid data-driven cuckoo search based adaptive protection model," Int. J. Electr. Power Energy Syst., vol. 130, no. February, p. 106960, 2021, doi: 10.1016/j.ijepes.2021.106960.

[25] Y. Bengio. Practical recommendations for gradient-based training of deep architectures. In Neural Networks: Tricks of the Trade: Second Edition pp. 437-478. Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/978-3-642-35289-8_26.

[26] S. Mirjalili, S. M. Mirjalili, and A. Hatamlou, "Multi-Verse Optimizer: a nature-inspired algorithm for global optimization," Neural Comput. Appl., vol. 27, no. 2, pp. 495–513, Feb. 2016, doi: 10.1007/s00521-015-1870-7.

[27] K. P. Schneider et al., "Analytic Considerations and Design Basis for the IEEE Distribution Test Feeders," IEEE Trans. Power Syst., vol. 33, no. 3, pp. 3181–3188, May 2018, doi: 10.1109/TPWRS.2017.2760011.

[28] W. Kersting, Distribution System Distribution Syste Modeling and Analysis. New Mexico Boca: CRC Press, 2012.

[29] W. H. Kersting, "Radial Distribution Test Feeders," Trans. Power Syst., vol. 6, no. 3, pp. 975–985, 1991, [Online]. Available: http://ewh.ieee.org/soc/pes/dsacom/testfeeders.html.

[30] Y. Gong and A. Guzmán, "Distribution feeder fault location using IED and FCI information," 2011 64th Annu. Conf. Prot. Relay Eng., pp. 168–177, 2011, doi: 10.1109/CPRE.2011.6035617.