Leakage current monitoring in high voltage insulator strings with neural network optimized with genetic algorithm

Maria Gabriely L. Silva, Jander R. S. Vieira,
Byron L. D. Bezerra, Sergio C. Oliveira
Polytechnic School of Pernambuco
University of Pernambuco
Recife, Brazil
{mgls, jrsv, byronleite, scampello}@ecomp.poli.br

Eduardo Fontana
Department of Electronic and Systems
Federal University of Pernambuco
Recife, Brazil
fontana@ufpe.br

***Abstract*—Partial discharges and flashovers in insulator strings, among others, are directly related to pollution presence on their surfaces and the ambient humidity. A partial discharge monitoring system has been developed and implemented in some locations in the Northeast Region of Brazil. This system has recorded approximately two years of information on humidity, temperature and peak leakage current activity during its operation. From the data collected by this network, it is possible to identify variations in the activity of leakage current peaks, showing more intense activity, which is related to the increased risk of flashover. This article presents an analysis of leakage current data using the prediction of current peaks with LSTM (Long Short-Term Memory) and MLP (Multilayer Perceptron) neural networks optimized with a genetic algorithm. The optimized models undergo a comparative analysis through hypothesis tests, along with experiments exploring their adaptation and generalization capabilities on different datasets. The results show that the LSTM model outperforms the MLP model in predicting leakage current peaks, achieving a significant reduction. The LSTM model demonstrated remarkable ability to anticipate activity peaks with a low error rate, enabling forecasts five days in advance, even under variable weather conditions. In interregional tests, it achieved an MSE of 0.041 and a MAE of 0.008 when predicting unobserved data. These results enable maintenance teams to act proactively, mitigating lightning risks and optimizing the scheduling of transmission line cleaning operations.**

***Index Terms*—insulators, flashover, leakage current, predict, hybrid systems**

I. Introduction

Pollution combined with high relative humidity in high-voltage insulator strings forms a conductive layer on their surface, allowing leakage current to flow. This current heats the pollution layer, causing moisture to evaporate and forming dry bands, dielectric regions surrounded by conductive areas [1]. The electric field across these bands may exceed the dielectric strength of air, triggering partial discharges. While isolated discharges are generally harmless [2][3], increased pollution and persistent humidity can lead to multiple discharges and, ultimately, a flashover [4], which interrupts power transmission and damages the insulator.

This work was supported by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)—Finance Code 001; Fundação de Amparo a Ciência e Tecnologia de PE (FACEPE); and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

To prevent flashover, maintenance actions such as washing the insulators are required. However, this process is costly and logistically complex, often requiring days of preparation [5]. Accurate prediction of flashover risk enables better decision-making and maintenance planning.

By analyzing leakage current and humidity data, it is possible to identify patterns associated with partial discharges and evaluate pollution severity [5]–[7]. A monitoring system installed in northeastern Brazil [5] records hourly data on leakage current pulses, average humidity, and temperature. This study investigates the use of such data to support preventive actions against flashover.

A neural network system with hyperparameters optimized via a genetic algorithm (GA) is proposed to predict leakage current activity. The GA tunes the number of hidden layers, neurons, and sliding window size for two widely used time series models: Multilayer Perceptron (MLP) and Long Short-Term Memory (LSTM). MLPs are known for modeling nonlinear relationships and generalizing across diverse data, while LSTMs effectively capture long-term dependencies and handle sequential data. These characteristics are crucial for modeling temporal patterns in leakage current. These networks have been widely applied in time series analysis due to their ability to generalize nonlinear patterns and capture long-term dependencies, properties essential for predictive maintenance in high-voltage [28].

Although decision tree-based models, such as Random Forests and XGBoost, have been widely employed for predictive tasks, their performance is often limited in time series forecasting scenarios that involve sequential dependencies [29]. In this study, greater emphasis was placed on neural network architectures due to their capacity to model the temporal dynamics characteristic of leakage current behavior. Nevertheless, statistical methods such as SARIMAX (Seasonal AutoRegressive Integrated Moving Average with eXogenous variables) were also evaluated to investigate their effectiveness in this context.

SARIMAX is suitable for modeling time series data and allows the inclusion of exogenous variables, such as humidity and previous current levels, in the forecasting process. Despite its capabilities, this type of statistical model may face challenges in representing nonlinear behaviors and complex variations commonly observed in insulator leakage current, which motivates the exploration of more robust techniques.

The main contributions of this work are as follows: (i) the development of a predictive model for leakage current activity based on LSTM and MLP neural networks with hyperparameters optimized via genetic algorithms; (ii) a comparative performance evaluation of both models under geographically distinct climatic conditions; (iii) the use of real-world data collected from a long-term monitoring system in northeastern Brazil; and (iv) the demonstration that LSTM models, when properly tuned, enable five-day-ahead predictions of high leakage current activity, supporting preventive maintenance planning in power transmission systems.

The following sections of this paper are organized as follows: Section II discusses the works related to flashover risk classification and prediction; Section III describes the leakage current monitoring system that provides the basis used in this work; Section IV describes the implementation of the SARIMAX model; Section V describes the proposed methodology with LSTM and MLP models and the experiments performed; Section VI presents the results and finally, the conclusions and final considerations are described in Section VII.

II. Related Works

Several studies have proposed different approaches for predicting failures in insulator strings using leakage current data. Sierra et al. [12] applied the Exponentially Weighted Moving Average (EWMA) for determining optimal cleaning times. Lima et al. [13] developed a fuzzy system with an MLP network using two days of risk input to predict flashover risk on the tenth day, though with reduced precision for higher peaks.

C.-T. Yeh et al. [14] proposed a BiLSTM-based model for real-time classification of leakage current in 15kV and 25kV insulators, enabling rapid anomaly detection and improved operational reliability. In a classification context, Jander et al. [18] applied a Bayesian classifier to estimate flashover risk based on activity rates and humidity, successfully identifying the progression from tolerable to critical pollution levels using real-world sensor data [5].

Gao et al. [19] used a backpropagation neural network to predict leakage current, performing well with high current values but showing deviations at low values. Thanh et al. [20] employed a Gated Recurrent Unit (GRU) model using meteorological variables, achieving approximately 45% accuracy.

Other deep learning approaches include a custom CNN for classifying RGB images from partial discharge signals, outperforming models like AlexNet and ResNet50 with over 98% accuracy [25]; an LSTM-CNN hybrid for categorizing current levels from 0 to 3 and capturing trends and cyclic behaviors [26]; and an ANN trained with features selected via Random Forest to predict pollution severity and asymmetric aging [27].

In light of these limitations — including low precision for peak events, lack of temporal modeling, restricted use of exogenous variables, and focus on classification instead of prediction — this study adopts recurrent neural network architectures (MLP and LSTM) for multivariate time series forecasting. The goal is to capture nonlinear behaviors and temporal dependencies in leakage current signals, enabling more accurate prediction of severe activity levels.

III. Leakage Current Monitoring System

There are several methods for measuring the deposition of pollutants on insulators [6]-[10]. Among them is the monitoring of the leakage current flowing over the surface of insulators [5][6]. Fig. 1 shows the leakage current monitoring system [5], consisting of three parts: the optical sensor module, the processing module, and the satellite transmission module.

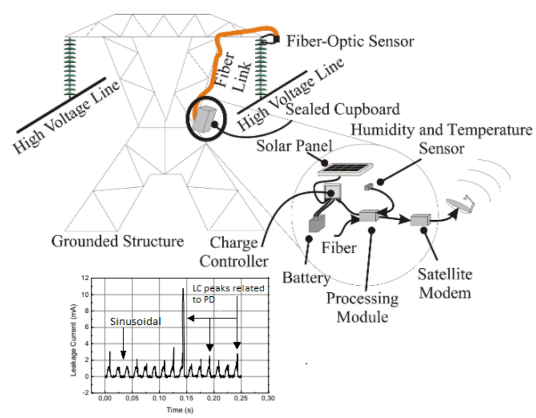

Fig. 1. Schematic of the leakage current monitoring system and the waveform captured [5].

The optical sensor module captures the waveform and peaks of the leakage current. Only the positive cycle of the sinusoid is captured by the sensor module, but it has similar information to the negative cycle of the leakage current. From the waveform captured, it is possible to observe in Fig. 1 that the leakage current is composed of a sinusoidal component that accompanies the voltage signal being applied, and a component characterized by rapid leakage current (LC) peaks, which are directly related to partial discharges (PD) on the insulator surfaces.

The processing module has the purpose of detecting, amplifying and storing the current peaks captured by the optical sensor module. To analyze the amplitude and frequency of the current, it was decided to classify the peaks into amplitude ranges and count the occurrence of peaks in these ranges per time unit. The peaks were classified into four ranges with amplitudes above: 5mA, 10mA, 20mA and 40mA, represented, respectively, by N1, N2, N3 and N4 [5].

Finally, the transmission module has the objective of transmitting, every hour, the processed information, N1 to N4, including average temperature and humidity data, via satellite to the ADECI system (Assessment of Electrical Performance in Insulator Strings) of CHESF (Companhia Hidro Elétrica do São Francisco) which is the company responsible for generating and transmitting electricity in the Northeast of Brazil, mainly from hydroelectric plants. This leakage current monitoring system was installed in six high voltage monitoring stations in Northeast Brazil, namely Sobral (500kV), Pitaguary (500kV), and Fortaleza (500kV) in the state of Ceará, Mossoró (230kV) in the state of Rio Grande do Norte; Angelim (500kV) in the state of Pernambuco; and São Miguel dos Campos (230kV) in the state of Alagoas.

IV. Forecasting Leakage Current Using SARIMAX

Given the temporal nature of leakage current data and the need to predict the evolution of these events over time, the SARIMAX model was considered. This model was chosen for its ability to capture patterns in time series data with seasonality, while incorporating exogenous variables, such as humidity and previous current levels [21].

The motivation for testing SARIMAX lies in the fact that statistical time series models are widely used for forecasting in various industrial and engineering applications. Unlike purely statistical models like ARIMA, SARIMAX allows the incorporation of external variables into the forecasting process, potentially improving the accuracy of predictions by including relevant environmental and electrical information.

Our objective with this approach was to assess whether SARIMAX could effectively predict leakage current peaks based on past values and environmental conditions, enabling the anticipation of potential failures in the electrical system.

Thus, the SARIMAX model was implemented with the aim of forecasting the leakage current (N3), considering exogenous variables such as N1, N2, and Humidity.

The data was divided into training (70%) and testing (30%) sets, ensuring that the model was trained with a representative portion of the time series before being evaluated. To capture the dynamics of leakage current over time, a sliding window approach of size 5 was employed, allowing the SARIMAX model to learn short-term patterns within the series.

The models were configured with (1,1,1) for the SARIMAX hyperparameters, which represent:

- AutoRegressive (AR) component: Indicates that the model considers one past value of the series to predict the next value. This means that the leakage current at a given moment is influenced by the immediately preceding value.
- Integration (I) component: Represents the number of differencing operations applied to make the series stationary. Since leakage current data may exhibit trends over time, a single differencing step was applied to remove potential trends and make the data more suitable for statistical modeling.
- Moving Average (MA) component: Allows the model to capture dependencies between past forecast errors. With an MA term of order 1, SARIMAX attempts to correct recent errors when predicting the next value.

In addition to these fundamental components, the model incorporated exogenous variables, including N1, N2, and Humidity, to enhance the prediction of N3. The inclusion of these variables aims to capture relationships between environmental conditions and leakage current behavior over time.

Finally, a small random noise was added to the data before training. This procedure, common in statistical modeling, helps stabilize numerical calculations by preventing singularity issues and improving the model's robustness. The noise was generated from a normal distribution with zero mean and minimal variance $\sigma^2 = 10^{-4}$, ensuring that its influence was negligible on actual values but sufficient to prevent computational problems during model fitting.

V. Methodology with LSTM and MLP models

A. Hybrid system architecture

The hybrid system proposed in this paper uses GA to optimize key hyperparameters of the LSTM and MLP networks. Genetic algorithms efficiently explore the search space to find configurations that improve neural network performance, especially important in this research, where hyperparameter selection impacts the model's ability to capture complex leakage current patterns. The tuned parameters are the number of hidden layers (1 to 3), neurons per layer (32 to 128), and sliding window size (5 to 15).

The activation function used for training was the hyperbolic tangent [16]. The optimization algorithm applied to the gradient descent during the training process is Adam, with a learning rate of 0.001 and with a predefined number of epochs at 100. A dropout rate of 0.5 was applied to each layer during training, while the cost function was used the binary cross entropy loss because this system requires only one neuron as output to predict a variable. In addition, the batch size presented during training is 32 packets, in order to avoid turbulence during training and improve convergence.

With the optimized values, the LSTM and MLP models are compared. The mean square error (MSE) and mean absolute error (MAE) in the output were used as evaluation metrics, which presents consolidated results when used in quantized continuous variables.

B. Data selection and training

To obtain the optimized hyperparameters for the neural networks, training was carried out with 35 runs on the São Miguel dos Campos database. The decision to select a specific database was based on an analysis of the correlation between the variables N1-N4, humidity and temperature in each of the six databases obtained by the monitoring system. The importance of this choice lies in the search for a database that not only accurately represents the relationships between the variables, but also demonstrates significant patterns for predicting the level of leakage current in the insulators. The base with the best correlations between activity levels was São Miguel dos Campos. These values can be seen in Fig. 2.

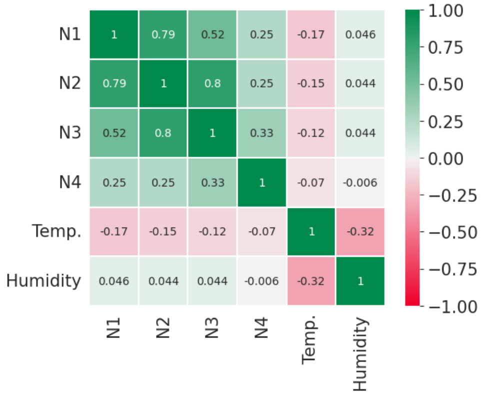

Fig. 2. Correlation of the variables of the São Miguel dos Campos database.

The activity levels N1-N4 are related to the pollution degree of the insulator strings [17], so the more peaks of higher amplitude occur the more likely the string is dirty and the more likely a flashover will occur when the humidity is high. The N3 level of 20 mA can already be considered an indication of the degree of pollution [17] and in this work it was chosen as the output variable of the network. The records were transformed into days based on the sum of the hourly peak values. Predicting the leakage current on a daily basis allows potential problems to be anticipated days in advance. This long-term forecast daily is crucial for scheduling power line maintenance, providing enough time to plan, organize resources and carry out the necessary maintenance without compromising the operation of the network. The Min-Max normalization technique was also applied to the daily data according to the equation 1.

$$Z = \frac{X - Min(X)}{Max(X) - Min(X)}. \quad (1)$$

A sliding window technique was employed in the training. This technique is commonly used in the training of neural networks and other machine learning models on time series or sequential data [22]. This approach allows the model to learn from multiple subsets of the training data instead of using the complete set at once. The prediction approach used in this work was multistep prediction, which involves forecasting multiple steps ahead in a time series.

Starting from the initial training window, the model is used to predict the next point, representing the day immediately following the last point of the utilized window. After predicting the next point, the time series is updated. The last predicted point is added to the window, and the oldest point is removed. This keeps the window size constant, with the most recent points incorporating the predictions. The model is then used to predict the next point in the updated time series, representing the second day following the original window's last point. This process is repeated until predicting the fifth day beyond the window size.

The input to the neural networks consists of a time window ranging from 5, 10, and 15 days in the past, a number of layers ranging from 1, 2, and 3 layers, and a number of neurons ranging from 32, 64, and 128 neurons, all determined by the genetic algorithm. The input variables include normalized leakage current activity levels (N1, N2, N3). The model predicts leakage current activity (N3) for the fifth day ahead. This multistep prediction procedure relies on the sequential iteration of the model to forecast each future point in the time series, updating the series as new predictions are generated. It is a continuous process until the desired number of steps ahead is predicted. The days ahead of prediction are essential in this work, enabling maintenance teams for power transmission lines to plan and travel to critical pollution points before complete discharges occur.

The complete dataset was split into 70% for training and 30% for testing. Within the training set, 80% was used for model fitting and 20% for validation during the optimization phase. To ensure statistical robustness, all experiments were repeated 30 times with different random seeds controlling initialization and sample shuffling. The reported metrics correspond to the average across these 30 independent runs. Table I shows the parameters used in the net.

TABLE I
NEURAL NETWORK PARAMETERS AND HYPERPARAMETERS

Training Optimizer	Adam
Dropout	0.5 in each layer
Loss function	Cross Entropy
GA optimized parameters	Number of layers: [1, 2, 3]
	Number of neurons: [32, 64, 128]
	Sliding window size: [5, 10, 15]
Epochs size	100
Batchsize	32
Learning Rate	0.001

C. Hyperparameter Optimization

Hyperparameter optimization evaluates the performance of each combined model by varying some critical parameters. The number of neural nodes in each layer defines the complexity of the neural network, which significantly affects the training and response time. The number of neurons per layer controls the ability to search for critical patterns in massive data. Also, the addition of hidden layers could improve the accuracy and performance of the training process due to the complex data representation [11].

The search engine used through the Genetic Algorithm does a scan in order to find an optimal combination. As it is three parameters to be optimized, each with three possible values, the algorithm tries to combine such parameters to save effort. The GA was set with an initial population of 10 individuals for three generations, in order to find the best combination without extrapolating the equivalent of brute force scanning, which would be at least 27 combinations. Reproduction was carried out using tournament selection, with genetic diversity introduced through a uniform crossover (rate = 0.25) and a mutation rate of 0.1.

Although the total number of possible configurations (27) is relatively small, an exhaustive grid search would require testing all combinations, which can be computationally expensive given that each training run consists of 100 epochs.

The choice of genetic algorithm (GA) as the hyperparameter optimization strategy was motivated by its proven balance between search space exploration and convergence speed. While other approaches such as Particle Swarm Optimization (PSO), Bayesian Optimization, or random search could also be applied, GA stands out for its simplicity, adaptability, and ease of parallelization. Moreover, comparative studies [23], [24] have demonstrated that GA often yields competitive or superior results in neural network tuning, particularly in problems with small to moderate search spaces, as is the case in this study.

In this scenario, after executing all configurations generated by the optimization algorithm each with 100 epochs, the hyperparameter configuration that presented the lowest mean square error for the LSTM was [1 0 2], which has topology of 2 hidden layers, composed of 32 neurons each, and with window size of 15 events (Fig. 3). For the MLP network, the configuration with the lowest error was [2 1 2], which has a topology of 3 hidden layers, composed of 64 neurons each, and a window size of 15 events (Fig. 4). It is possible to observe that during the runs the error in the MLP network does not present large variations, which may indicate that the network is unable to learn. Analyzing the best topology obtained by the two networks in this experiment, the LSTM network is already superior with an error of 0.003898, while the MLP has an error of 0.02903.

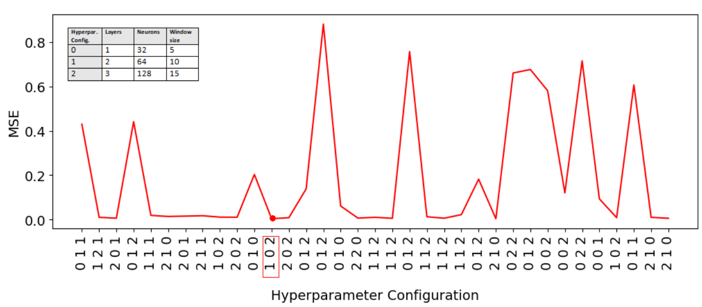

Fig. 3. Performance of optimized hyperparameters in the LSTM model with data São Miguel dos Campos.

Two different tests were carried out to assess the performance and adaptability of the neural network under different climatic and geographical conditions.

In the first test, the aim was to analyze the neural network's generalization power under the influence of different climates. To do this, the network was trained using only the climate database of São Miguel dos Campos. After training, the network's performance was evaluated using databases from two different regions: Fortaleza and Angelim. This procedure allowed us to see how the network, trained in a specific environment, would behave when applied to different climates.

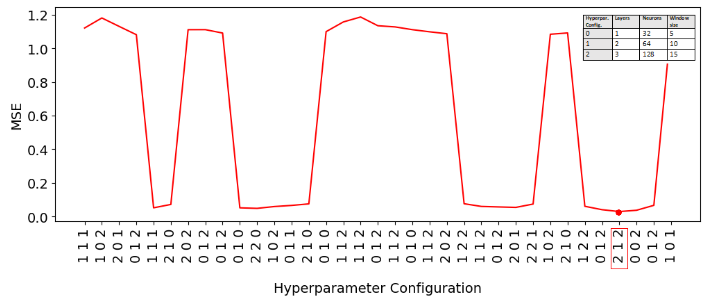

Fig. 4. Performance of optimized hyperparameters in the MLP model with data São Miguel dos Campos.

In the second test, the goal was to enhance the neural network's adaptability to geographical differences. For this, it was retrained using a mixed and individually normalized dataset with climate data from Fortaleza, Angelim, Mossoró, and São Miguel dos Campos. This ensured balanced contributions from each location and helped the model learn generalizable patterns. For each input, data from the same day across all cities were included (e.g., day 1 combined N1–N3 values from all cities). Performance was again evaluated using the Fortaleza and Angelim datasets to assess whether the broader dataset improved adaptability to regional and climatic variations.

D. Hypothesis Testing

To compare the performance of the LSTM and MLP networks, an experiment was performed with 30 runs on both networks with the best topology found by the genetic algorithm. To compare the two models the T-Test hypothesis test [15] was executed, but to execute it is necessary to verify that the samples come from a normally distributed population.

The Shapiro-Wilk normality test [15] was used to verify that the samples come from a normal population. The null and alternative hypotheses are:

- H0: The evaluation metrics come from a normal population.
- H1: The evaluation metrics do not come from a normal population.

If the p-value is less than the statistical significance level α (adopted 5%), the H0 should be rejected. Tables II and III present the calculated p-values for the MSE and MAE metrics, respectively. Since no p-value was smaller than the adopted significance (0.05), the H0 was not rejected. Therefore, the analyzed metrics come from a normally distributed population.

TABLE II
P-VALUE FOR SHAPIRO-WILK TEST.

Network	MSE	P-Value
LSTM	0.00830476	0.1587
MLP	0.01279	0.2981

TABLE III
P-VALUE FOR SHAPIRO-WILK TEST.

Network	MAE	P-Value
LSTM	0.073453	0.1336
MLP	0.092819	0.3775

With the normality criterion met, the T-Test was run in order to verify if there is a statistically significant difference between the tested topologies. The hypotheses defined, were:

$$H_0 : MetricMean(LSTM) - MetricMean(MLP) < 0, \tag{2}$$

$$H_1 : MetricMean(LSTM) - MetricMean(MLP) > 0. \tag{3}$$

Since the p-value was less than the adopted significance (0.05) for the MSE (0.000586) and MAE (0.007562) metrics, the H0 was rejected. Therefore, the LSTM model is superior to the MLP.

VI. Results

Figure 5 shows the actual leakage current values and the predictions generated by the SARIMAX model. It is observed that the model was unable to adequately capture the leakage current peaks, often underestimating or ignoring these extreme values.

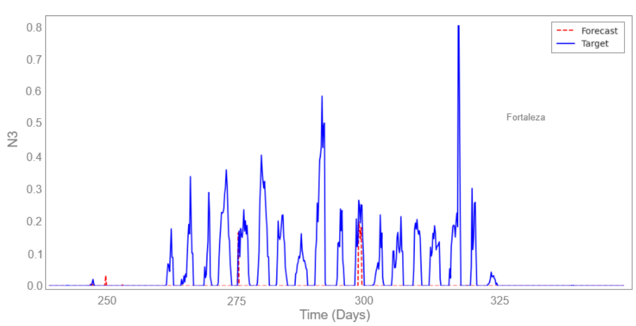

Fig. 5. Prediction results with SARIMAX model trained with the mixed database and tested with the Fortaleza.

The results indicate that SARIMAX was not effective in forecasting leakage current peaks in this specific application. This limitation can be attributed to the nature of current peaks, which often represent rare and abrupt events, posing a challenge for traditional autoregressive models. Additionally, SARIMAX's sensitivity to highly variable and non-stationary data, even after normalization, contributed to this performance. The absence of mechanisms within the model to learn complex nonlinear patterns present in the peaks is another intrinsic limitation of methods based on moving averages and autoregressive approaches.

These results underscore the importance of exploring more advanced techniques, such as neural networks (LSTM or MLP), which have greater capacity to model nonlinear relationships and complex dynamics.

Fig. 6(a) shows the results of testing the LSTM net trained on São Miguel dos Campos and tested on Fortaleza. Fig. 6(b) shows the results with the same net tested with the Angelim base. With these results, it is possible to observe that the network identifies the behavior of activities but with less accuracy during peak periods. This outcome is likely associated with the dependence on the phenomenon of pollutant deposition on insulators, influenced by climatic factors such as wind direction, rainfall, pollutant type, and annual cycles of pollutant deposition based on both rainfall patterns and local economic activities.

However, these less accurate peaks, near days 1, 50, and 75 in Fig. 6(a), correspond to rapid, short-duration phenomena, such as sudden increases in humidity or transient pollution events. These events tend to generate sharp but brief activity spikes that are not consistently captured by the model. Although the impact on the overall predictive performance is minor, future work should investigate higher-frequency data collection and specialized architectures to improve responsiveness to short-duration peaks. On the other hand, when observing longer-lasting activities (wider peaks), which are the more critical cases requiring maintenance, the network accurately predicted with greater precision during peak periods.

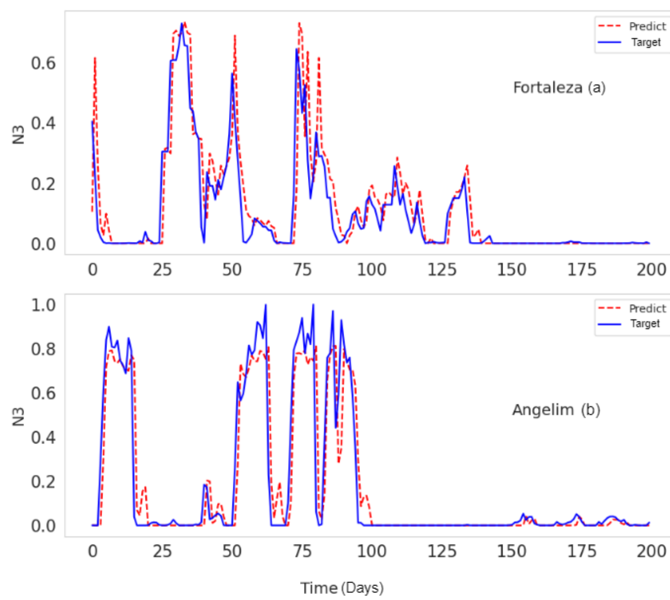

Fig. 6. Prediction results with the optimized LSTM network trained with the data São Miguel dos Campos and tested with the data Fortaleza (a) and Angelim (b).

The same scenario was evaluated for the MLP network, in Fig. 7(a) test in Fortaleza and Fig. 7(b) test in Angelim. However, the network could not identify the behavior of the activities, obtaining higher errors as can be seen in the Table IV.

Fig. 8 shows the results of the training conducted with the mixed network and evaluated in Fortaleza and Angelim. With the addition of data from other bases in the training process, it can be observed that there was a reduction in the peaks of

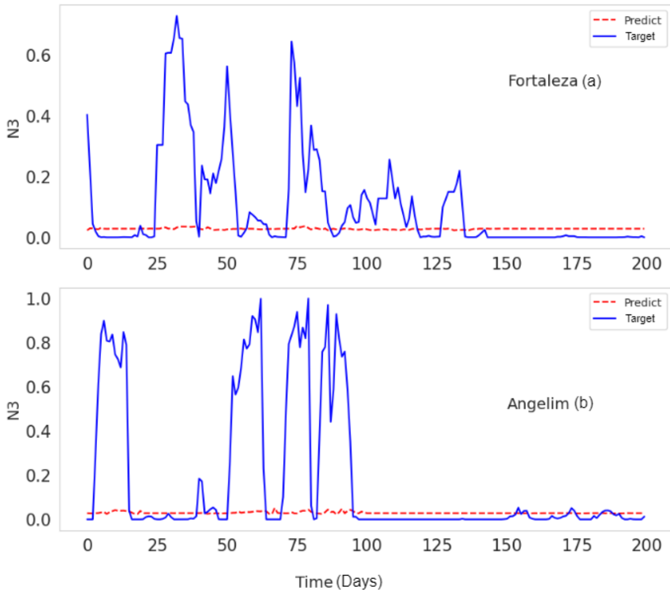

Fig. 7. Prediction results with the optimized MLP network trained with the data São Miguel dos Campos and tested with the data Fortaleza (a) and Angelim (b).

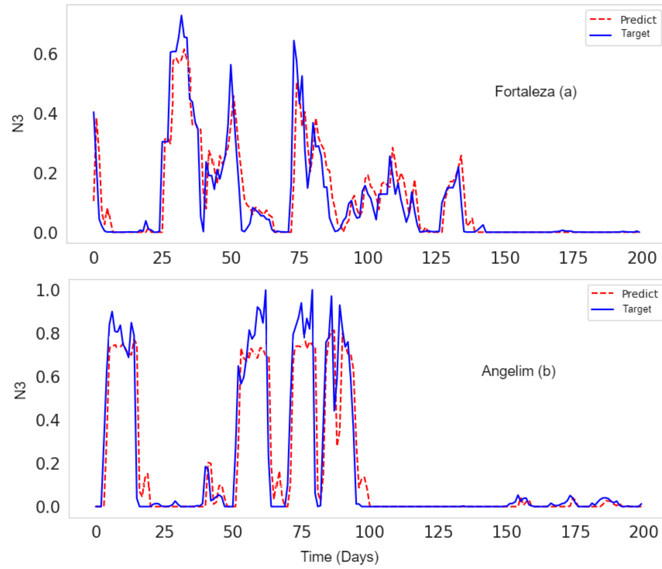

Fig. 8. Prediction results with the optimized LSTM network trained with the mixed database and tested with the Fortaleza (a) and Angelim (b)

TABLE IV
TEST RESULTS WITH THE MLP AND LSTM TRAINED WITH SÃO MIGUEL DOS CAMPOS DATA SET AND TESTED WITH ANGELIM AND FORTALEZA DATA SETS.

Network	Error (MSE)	Error (MAE)	Test Data Set
LSTM	0.029713	0.02033	Angelim
LSTM	0.04136	0.008843	Fortaleza
MLP	0.2743	0.1872	Angelim
MLP	0.2731	0.1542	Fortaleza

larger amplitudes compared to Fig. 6. The peak near day 80 in Fig. 6(a) experienced a reduction of approximately 33.4%.

Table V shows the results of the tests carried out on each base for training with the mixed base. The error results show that the LSTM network is superior to the MLP network. Although the model trained on the mixed dataset exhibited slight higher overall error values, it demonstrated an improved ability to mitigate extreme peak values in leakage current predictions. This suggests that training on a more diverse dataset enhances the model's ability to generalize across different environmental conditions, reducing the sensitivity to localized noise while maintaining a reasonable predictive accuracy.

TABLE V
TEST RESULTS WITH THE MLP AND LSTM WITH TRAINING MIXED DATABASE

Network	Error (MSE)	Error (MAE)	Test Data Set
LSTM	0.03185	0.02113	Angelim
LSTM	0.04299	0.00934	Fortaleza
MLP	0.2877	0.1911	Angelim
MLP	0.2964	0.1736	Fortaleza

In both experiments, the LSTM network demonstrated superiority in predicting the peak activity N3. The LSTM showcases remarkable abilities in capturing long-term temporal patterns, standing out for its capacity to maintain and update memory, enabling the learning of more complex temporal dependencies. This differential is particularly crucial in scenarios where past events significantly impact future predictions.

On the other hand, the MLP network proves limited when confronted with long-term temporal relationships. Peaks of leakage current activity exhibit important temporal correlations, in which the LSTM excels by learning and leveraging these relationships to enhance predictions. In this context, the MLP may not be as effective.

Although the model trained with the mixed dataset showed improved generalization, some peak predictions exhibited slight delays relative to the actual events. This effect may be attributed to the daily granularity of input data and the sequential dependency structure of the LSTM. The model may require multiple confirming patterns across time steps before predicting a peak. Future work could explore the use of attention mechanisms or Transformer-based architectures to mitigate temporal lags in peak prediction.

VII. CONCLUSION

This work presented an approach for predicting peak leakage current activity in high-voltage insulators using LSTM and MLP neural networks, with hyperparameters optimized via genetic algorithms. The proposed models were evaluated using real-world datasets from multiple substations in Brazil, enabling five-day-ahead predictions of the N3 activity level.

The limitation of the MLP network in handling long-term temporal relationships of leakage current activity peaks may explain these results. Meanwhile, the LSTM was able to learn and leverage these temporal relationships to enhance predictions. The LSTM model consistently outperformed the

MLP model in both accuracy and adaptability across different climatic regions, achieving lower error rates and better generalization.

Despite satisfactory results, there is a need for more precise evaluation and new experiments with different network topologies. The pre-selection of data and enhanced exploration of columns, such as temperature, humidity, dew point, and precipitation, present opportunities. Regarding hyperparameter optimization, additional experiments are proposed with parameters like dropout, batch size, and learning rate. The application of complementary optimization techniques is also considered to further enhance the model's performance.

Even with the superior ability to predict in longer-lasting activities, there is a tendency to overestimate results. While unfavorable, it is preferable to detect a false activity than to overlook a true one. Avoiding underestimation is crucial to prevent misguided decisions about the safety of high-voltage transmission lines. Prediction errors, when they occur, encourage safer maintenance actions than optimistic forecasts. It is noteworthy that in extensive activities, the network demonstrates higher precision, reducing errors and promoting additional reliability.

Thus, predictions generated by the LSTM model optimized with genetic algorithms are effective for anticipating maintenance needs, contributing to improved planning and risk mitigation. Despite some limitations, such as prediction lags in short-duration peaks, the approach demonstrates potential for real-world application. Future work includes exploring Transformer-based architectures, alternative loss functions, and incorporating additional input variables such as dew point and precipitation to enhance the robustness and precision of the predictions.

References

[1] D. L. Williams et al., "Formation and characterization of dry bands in clean fog on polluted insulators," IEEE Transactions on Dielectrics and Electrical Insulation, vol. 6, no. 5, pp. 724–731, 1999.

[2] M. Sforzini, "Testing of polluted insulators—the present situation and problems of the future," Journal of the Franklin Institute, vol. 294, no. 6, pp. 437–468, Dec. 1972.

[3] E. Fontana et al., "Novel Sensor System for Leakage Current Detection on Insulator Strings of Overhead Transmission Lines," IEEE Transactions on Power Delivery, vol. 21, no. 4, pp. 2064–2070, Oct. 2006.

[4] B. F. Hampton, "Flashover mechanism of polluted insulation," Proceedings of the Institution of Electrical Engineers, vol. 111, no. 5, p. 985, 1964.

[5] S. C. Oliveira, "Sistema de Detecção Óptica de Descaragas Parciais em Cadeias de Isoladores de Linhas de Transmissão de Alta Tensão," PhD Thesis, PPGEE-UFPE, 2008.

[6] J. L. Fierro-Chavez, I. Ramirez-Vazquez e G. Montoya-Tena, "On-line leakage current monitoring of 400 kV insulator strings in polluted areas", IEEE Proceedings of the Generation Transmission and Distributions Conference, Vol. 143, No 6, pp 560-564, November 1996.

[7] G. Montoya-Tena, R. Hernández-Corona e I Ramírez-Vázquez, "Experiences on Pollution level measurement in Mexico", Electric Power Systems Research, Volume 76, pp 58-66, Elsevier, 2005.

[8] A. S. Ahmad, P. S. Ghosh, S. S. Ahmed e S. A. K. Aljunid, "Assessment of ESDD on high-voltage insulators using artificial neural network", Electric Power Systems Research, Volume 72, pp 131-136, 2004.

[9] M. Youssef, A. A. El-Alayly, "Remote monitoring of Surface Current Over Insulators," Proceedings of the IEEE Transmission and Distribution Conference, vol. 01, pp. 113-116, November 2001.

[10] R. O. Silva et al., "Isoladores tipo disco de vidro temperado – Análise de desempenho e vida útil no sistema CHESF", Procedimentos internos CHESF.

[11] Stathakis, Dimitris, How many hidden layers and nodes, International Journal of Remote Sensing, volume 30, number 8, pp. 2133–2147, year 2009

[12] R. C. Sierra et al., "A novel method for prediction of washing cycles of electrical insulators in high pollution environments", International Journal of Electrical Power and Energy Systems, Elsevier, v. 130, p. 107026, 2021.

[13] H. O. de Lima, S. C. Oliveira, and E. Fontana, "Flashover risk prediction on polluted insulators strings of high voltage transmission lines," in 2011 11th International Conference on Intelligent Systems Design and Applications, 2011, pp. 397–401.

[14] Yeh, Chao-Tsung and Thanh, Phuong Nguyen and Cho, Ming-Yuan, "Real-Time Leakage Current Classification of 15kV and 25kV Distribution Insulators Based on Bidirectional Long Short-Term Memory Networks With Deep Learning Machine," IEEE Access, 2022, pp. 7128-7140.

[15] G. K. Kanji, 100 Statistical Tests, 3rd editio. London: SAGE Publications Ltd, 2006, p.256.

[16] A. H. Namin et al., "Efficient hardware implementation of the hyperbolic tangent sigmoid function," IEEE International Symposium on Circuits and Systems, Taipei, Taiwan, 2009, pp. 2117-2120.

[17] L. H. S. Silva, S. C. Oliveira, R. A. de Lima and E. Fontana, "Long term analysis of leakage current pulses registered by an optical sensor network," 2013 SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference, Rio de Janeiro, Brazil, 2013, pp. 1-5.

[18] J. R. S. Vieira, M. G. L. Silva, S. C. Oliveira, B. L. D. Bezerra, "A Bayesian Risk Classification System Based on Two Stage Clustering For Flashover Occurrence In High Voltage Insulator String", in 9th IEEE Latin American Conference on Computational Intelligence, 2023.

[19] S. Gao et al., "Prediction method of leakage current of insulators on the transmission line based on BP neural network," 2018 IEEE 2nd International Electrical and Energy Conference (CIEEC), Beijing, China, 2018, pp. 569-572.

[20] P. N. Thanh et al., "Predicting Leakage Current of Distribution Insulators Based Deep Learning Gated Recurrent Unit," 2022 5th Asia Conference on Energy and Electrical Engineering (ACEEE), Kuala Lumpur, Malaysia, 2022, pp. 46-50.

[21] R. J. Hyndman and G. Athanasopoulos, Forecasting: Principles and Practice, 3rd ed. Melbourne, Australia: OTexts, 2021.

[22] V. Uroševic and S. Dimitrijevic, "Optimum input sequence size for a sliding window-based LSTM neural network used in short-term electrical load forecasting," 2021 29th Telecommunications Forum (TELFOR), Belgrade, Serbia, 2021, pp. 1-4.

[23] H. Alibrahim and S. A. Ludwig, "Hyperparameter Optimization: Comparing Genetic Algorithm against Grid Search and Bayesian Optimization," 2021 IEEE Congress on Evolutionary Computation (CEC), Krakow, Poland, 2021, pp. 1551-1559.

[24] K. Kiatkarun and P. Phunchongharn, "Automatic Hyper-Parameter Tuning for Gradient Boosting Machine," 2020 1st International Conference on Big Data Analytics and Practices (IBDAP), Bangkok, Thailand, 2020, pp. 1-6.

[25] P. Das, S. Chatterjee and C. Koley, "Recurrence Plot-Aided Partial Discharge Detection Framework Employing HFCT Sensor and Customized Convolutional Neural Network," in IEEE Transactions on Dielectrics and Electrical Insulation, vol. 31, no. 6, pp. 2947-2955, Dec. 2024

[26] DA, Thao Nguyen et al., A cloud 15kV-HDPE insulator leakage current classification based improved particle swarm optimization and LSTM-CNN deep learning approach. Swarm and Evolutionary Computation, v. 91, p. 101755, 2024.

[27] KHODSUZ, Masume; ZAMANI, Seyed Alireza. Novel criteria for silicone rubber insulators condition monitoring based on leakage current analysis: Considering asymmetric aging and pollution. Engineering Applications of Artificial Intelligence, v. 133, p. 108175, 2024.

[28] Yeh, Chao-Tsung et al., "Real-Time Leakage Current Classification of 15kV and 25kV Distribution Insulators Based on Bidirectional Long Short-Term Memory Networks With Deep Learning Machine," IEEE Access, pp. 7128-7140, 2022.

[29] Paliari, A. Karanikola, and S. Kotsiantis, "A comparison of the optimized LSTM, XGBOOST and ARIMA in Time Series forecasting," in Proc. 2021 12th Int. Conf. Inf., Intell., Syst. Appl. (IISA), pp. 1–7, 2021.