T3.1 Definition of the algorithm for the Early Detection

Short description of the Early Detection algorithm

The project FELINES aims at designing a protection system capable of sensing electromagnetic fields that are preliminary to a lightning event, and consequently disconnect part (or all) of the electric infrastructure under its protection. These fields are generated by the so-called Preliminary Breakdown (PB) pulses, localized events taking place during the first phases of the lightning inception.

In the project the experimental part was not planned, for this reason the PB pulses are simulated according to [1]. The PB pulse is the "trigger" that starts the algorrithm for the Early Detection, which is easily described in the following points:

1. PB pulse start
2. PB pulse detection by the effect it induces on a Transmission Line (in terms of overvoltages)
3. Evaluation of the consequent Return Stroke (RS) danger level

Points 2 and 3 are carried out by using Machine Learning (ML) based procedures. More in detail:

A. A Neural Network (NN) is properly designed and trained in order to predict the overvoltage caused by PB pulses. This steps serves as PB pulse detection (point 2) and is the input to point 3.
B. A different NN is properly designed and trained in order to predict the overvoltage peak caused by the consequent RS. The knowledge of these overvoltages and the characteristics of the TL lead to a classification of the event as dangerous or not dangerous.

According to the authors this a-posteriori comparison and classification (once the threshold is defined) gives more information than a simple classification obtained by a classifier, because with the proposed approach the magnitude of the overvoltage is a fundamental additional information. In order to achieve good performances it is important to build a model that is sensitive to both dangerous and safe cases and this is achieved by working on a balanced dataset as explained before.

The two NNs above mentioned need to be designed and trained; in order to do so a proper dataset must be created.

Dataset description and generation

The dataset is referred to the following case study, depicted in Fig. 1: a two-conductor overhead distribution line consisting of phase and neutral. The overhead line is characterized by a 50-kV Critical FlashOver voltage (CFO) is considered, and the threshold to assess if a indirect lightning strike results into a dangerous flashover is 1.5 CFO, as recommended in [2]. The line geometry is the following:

- Line length: 1 km.
- Phase conductor height: 10 m .
- The neutral is two meters below it.
- Conductor radius: 1 cm.
- The neutral is periodically grounded by means of a 30 Ω grounding resistor placed every 125 m.
- The soil is characterized by a conductivity of 1 mS/m and a relative permittivity of 10.
- Matching impedance are set at both the line ends to minimize reflections.
- Induced voltages are observed at a two voltage sensors, designated as VS1 and VS2, installed in parallel with a line insulator and located at x = 500 m and x = 250 m, respectively.
- The line insulators are modeled by means of their parasitic capacitance of 1 pF .

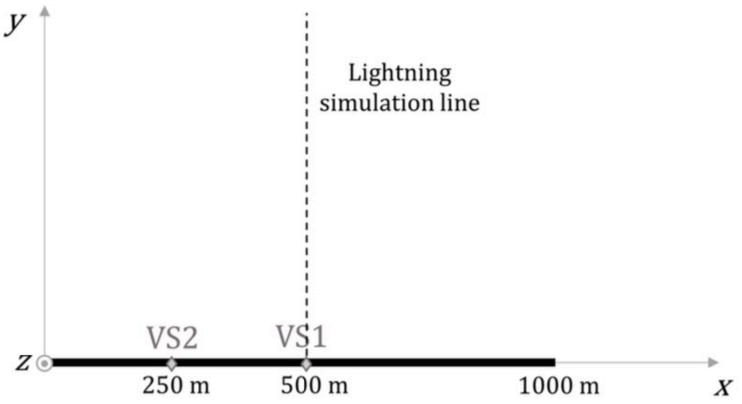

Fig. 1 Line geometry

The generation of a single event in the dataset is carried out according to the flowchart shown in Figure 2, with the stages described in the following points.

- *Strike location extraction*: PB and RS (x,y)-locations are assumed to be coincident, since the characterization of the PB current model has been done with such a hypothesis [1]. For the sake of simplicity, the x-coordinate of the strike location is assumed to be 500 m, i.e., the lightning always strikes in front of VS1. On the other hand, the y-coordinate is sampled from uniform distributions spanning from 1 m to 400 m. This upper bound was selected to ensure the generation of a sufficiently balanced dataset, since increasing the maximum distance from the line reduces the proportion of dangerous events in the dataset.

- *PB pulse simulation*: The PB pulse parameters which are not assumed constant are randomly extracted from the distribution appearing in [1] (Table 2), where a statistical characterization of the model has been performed by fitting experimental electric field data from more than 3000 high intensity negative Cloud to Ground lightning flashes measured in Florida, U.S.A. [3]. Moreover, the correlation coefficients between such parameters are taken into account too (Table 3).

- *RS simulation*. The only free parameter of the RS model is the channel-base peak current. This is randomly extracted starting from the PB peak current, by using the PB/RS peak current ratio distribution provided in [1], whose arithmetic mean, geometric mean, median, and standard deviation are 0.27, 0.19, 0.20, and 0.25, respectively.

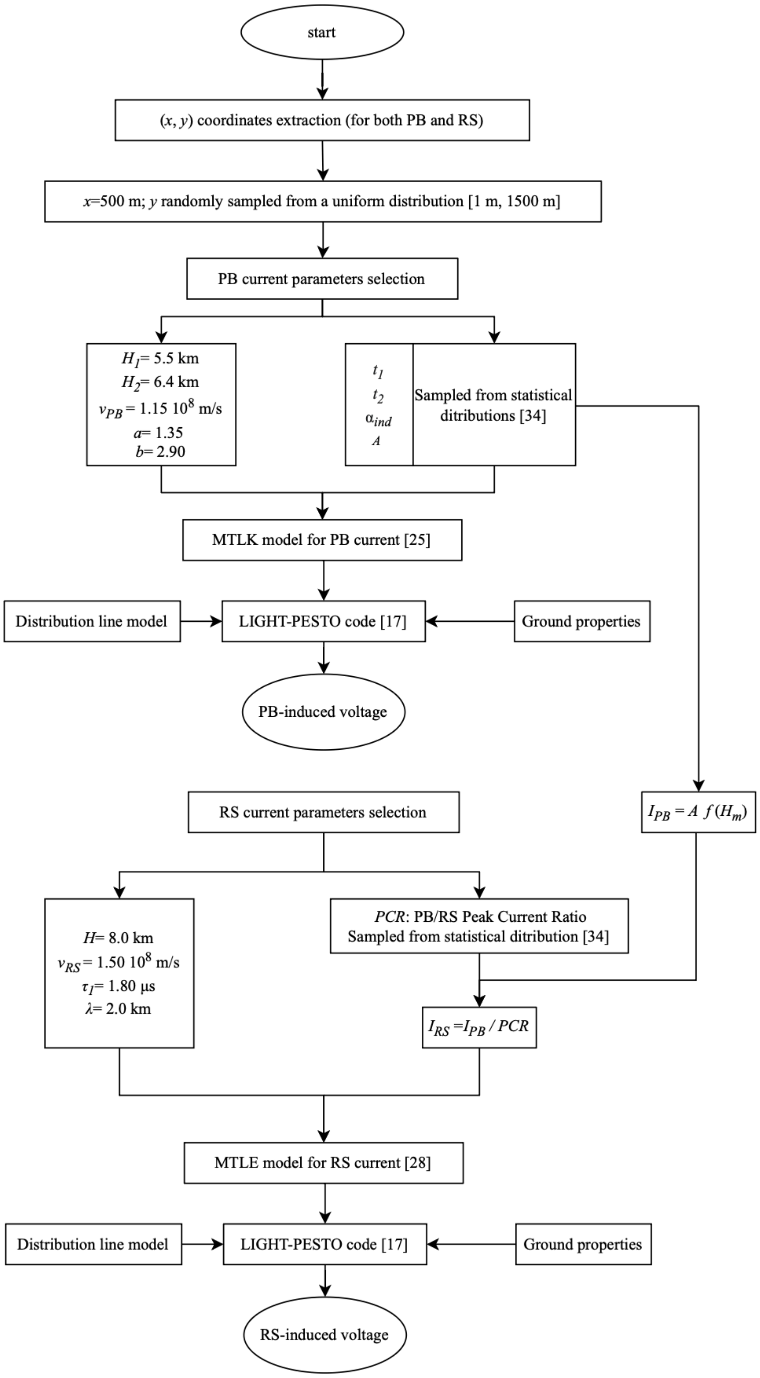

Fig. 2 Simulation of PB-induced and RS-induced voltages for a single lightning flash.

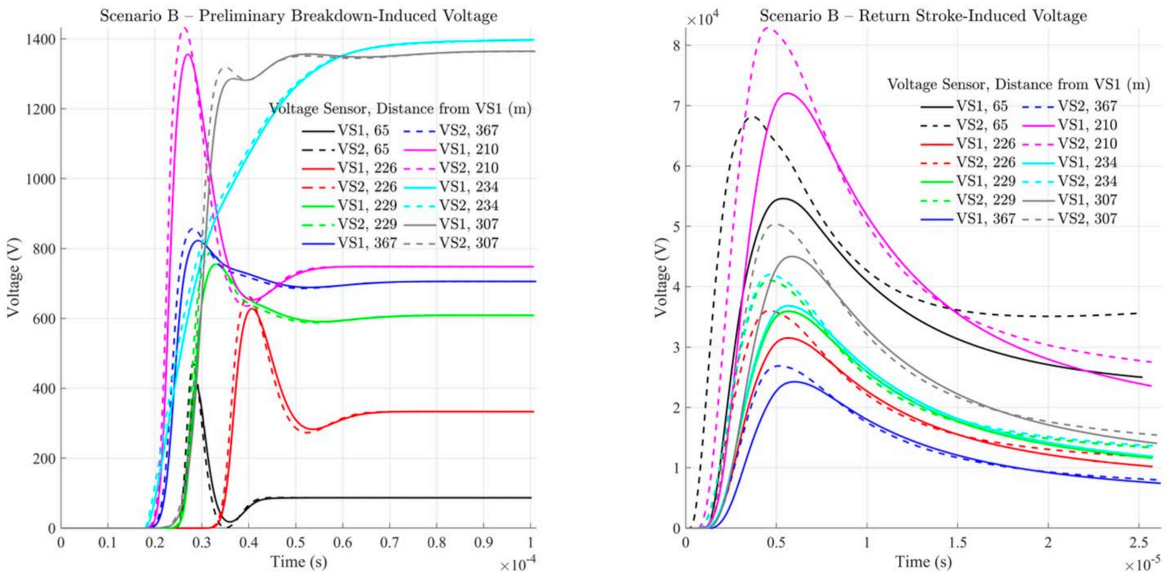

Figure 3. Examples of waveforms

The dataset is created by using a softrare named LIGHT-PESTO, previously developed by the researchers at UniGe. Each simulation produces two voltage waveforms for each sensor, associated with PB and RS, respectively. It should be mentioned that the flashes with RS leading to direct strikes are identified by applying the well known electro-geometric criterion [4] and removed from the datasets. Indeed, this type of phenomenon produces completely different voltage pulses on overhead conductors with respect to those induced by lightning EM fields.

In the dataset, among the 13095 events with RS producing an indirect flashover, 6925 are classified as dangerous, thus the dataset balance is assessed. An excerpt from each induced voltage dataset is displayed in Figure 3.

Algorithm for early detection

The Deep Learning based algorithm for the Early Detection is here presented. As explained at the beginning of this deliverable, it can be defined as a regression tool instead of a simple classificator, since it estimates the overvoltage peak an not simply a true/false outpur relative to the danger level.

The model takes as input the PB voltage waveform acquired from two sensors (VS1 and VS2) and predict the maximum of the RS-induced voltage on the line.

Signal pre-processing and feature extraction

Before feeding the PB voltage waveforms into the DL model to predict the RS maximum voltage, the time-domain PB signals are preprocessed and subjected to dimensionality reduction. Each time-domain waveform consists of $n = n_{PB}$ samples, which should be reduced and transformed into a lower-dimensional feature space to mitigate the curse of dimensionality during training.

A first step in dimensionality reduction comes from the consideration that, in practical applications, the system can detect a voltage exceeding a specified threshold, and his would be the origin of the time coordinates; on the contrary, the time $t = 0$ at which the PB pulse ignites (coincident with $t = 0$ in the simulations) is not known. Consequently, in each pair of voltage waveforms, the initial leading-zero values (Figure 3) corresponding to the delay between the initiation of the PB pulse and the arrival time of the induced voltage at the closer sensor are removed.

As a result this action creates vectors of different dimensions; on the contrary a specific Neural Network receives as input a vector of a defined dimension. For this reason, in order to have signals of the same length as input to the algorithm, an additional action on the time domain series must be performed. Since the tail section of each signal (vector) is the less informative, we decided to remove samples from the tail part of each signal up to the size of the shortest waveform.

As a result the waveforms from VS1 are truncated to $n_1 < n$ time samples while the waveforms from VS2 are truncated to $n_2 < n$ time samples (it is worth noticing that n_1 and $n_2$may differ).

To further reduce the waveform size, each time-domain voltage signal is divided into two equal sections: the front and the tail. A compression procedureis then applied to both parts, in which every group of N consecutive samples is replaced by their average value. Different averaging windows are applied to the two parts based on their relative importance. The front section, which includes more information, is compressed using a smaller averaging window of five samples. In contrast, the tail section, which holds less essential information, is compressed more aggressively by averaging every six consecutive samples. Several tests have been performed, with different subdivision between front section and tail section; after careful performance evaluation, and also taking into account the simplicity of the this pre-processing part, the time domain signal is split into the initial and tail section in its mid point.

After the compression procedure, both compressed waveforms from VS1 and VS2 are stacked and then processed by Principal Component Analysis (PCA) , a linear dimensionality-reduction technique that projects the input data onto an orthogonal basis defined by the directions of maximum variance, further reducing the feature dimensionality to nPCA components, retaining Γ% of the total variance.

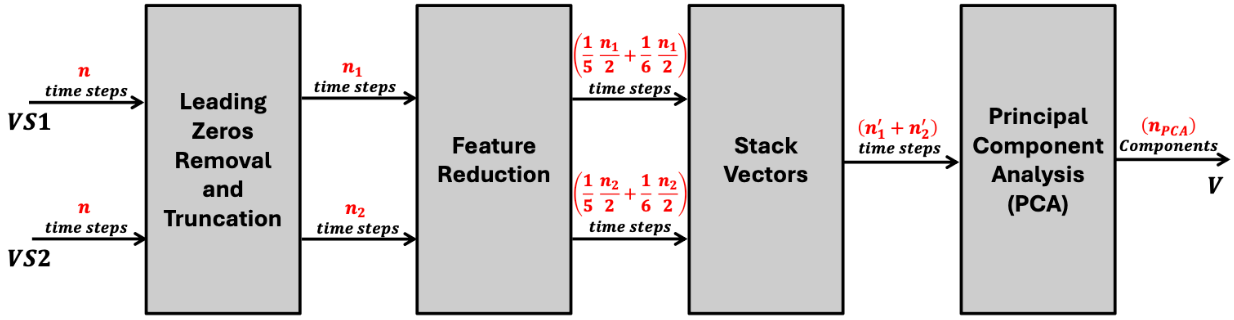

Fig. 4. Pre-processing process

Model Architecture and Training

The preprocessed feature vector V is used as the input to the residual neural network, whose architecture is illustrated in Figure 5, for the regression task of predicting the RS peak voltage.

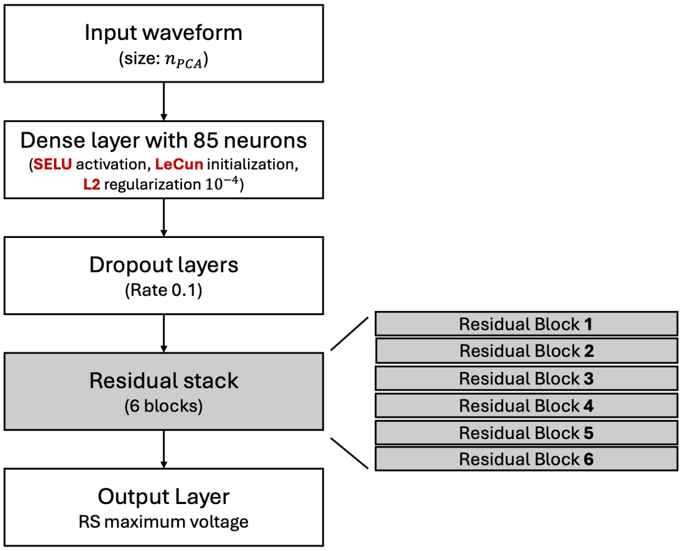

Figure 5: Architecture of the residual neural network used for predicting the RS maximum voltage.

The network begins with a fully connected dense layer comprising 85 neurons, employing the Scaled Exponential Linear Unit (SELU) activation function and L2 regularization [5].

As described in Figure 5, the residual network is very dense and includes many layers that cause it to be more prone to vanishing gradients during training. Therfore, the LeCun initialization [6] is applied to sets the initial weight variance to $1/n_{input}$ (i.e., inversely proportional to the number of input neurons), and normalize pre-activation values $z = Wx + b$ (denoting the weights, inputs, and bias, respectively). Then the SELU activation function.

This combination of both SELU activation function and LeCun normal initializer enables the network to automatically preserve the mean and variance of neuron activations across layers, a feature known as self-normalization that causes activations remain normalized and prevents exploding or vanishing gradients during training.

An L2 regularization term with a coefficient of 10^{-4} is applied to the kernel weights to add a penalty that discourages large weights and helps prevent overfitting by penalizing large parameter values. Following this layer, a dropout layer with a rate of 0.1 randomly deactivates a portion of neurons during training, further improving model generalization.

The central part of the network is a stack of six residual blocks, as described in Figure 6, and each of them learns a nonlinear transformation of its input and then adds this transformation back to the original signal through a skip connection. This design allows information and gradients to flow more smoothly through the network, reducing the risk of vanishing gradients and helping the model converge faster and more reliably during training.

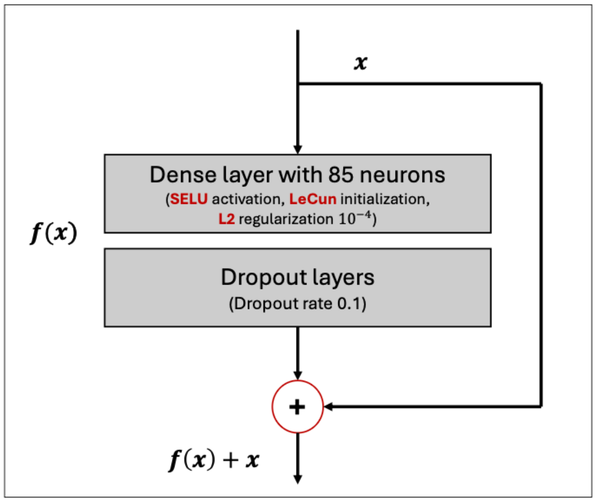

Figure 6: Structure of a single residual block used in the residual stack.

The network comprises a single neuron in both the input and output layers, while all hidden layers contain 85 neurons each. It is designed to predict the RS maximum voltage directly from the learned feature representation.

The residual neural network is implemented using Python's Keras library and trained with the Adam optimizer, using a learning rate of 10^{-4} and the Huber loss function [8].

Final Comments

The designed Machine Learning based algorithm is capable of predicting the RS maximum votage strarting from PB voltage waveforms acquired from two sensors (VS1 and VS2) and predict the maximum of the RS-induced voltage on the line. The Early Detection algorithm here proposed has been tested on several distribution lines models, single and three phase and characterized by different complexity, and in all cases it has proven to be efficient and reliable.

[1] D. Mestriner, M. Nicora, M. Brignone, R. Procopio, Y. Zhu and V. A. Rakov, "Modeling and Statistical Characterization of Preliminary Breakdown Pulses in Negative Cloud-to-Ground Lightning Flashes," in *IEEE Transactions on Electromagnetic Compatibility*, doi: 10.1109/TEMC.2026.3666274.

[2] IEEE, Guide for improving the lightning performance of electric power overhead distribution lines, IEEE Std. 1410-2010 (Revision of IEEE Std 1410-2004) (2011) 1–73doi:10.1109/IEEESTD.2011.5706451.

[3] Y. Zhu, V. A. Rakov, M. D. Tran, A study of preliminary breakdown and return stroke processes in high-intensity negative lightning discharges, Atmosphere 7 (10) (2016). doi:10.3390/atmos7100130.

[4] A. Mousa, K. Srivastava, The implications of the electrogeometric model regarding effect of height of structure on the median amplitude of collected lightning strokes, IEEE Transactions on Power Delivery 4 (2) (1989) 1450–1460. doi:10.1109/61.25632.

[5] G. Klambauer, T. Unterthiner, A. Mayr, S. Hochreiter, Self-normalizing neural networks, Advances in neural information processing systems 30 (2017).

[6] Y. LeCun, L. Bottou, G. B. Orr, K.-R. Muller, Efficient backprop, in: Neural Networks: Tricks of the Trade, Springer, 2012, originally published 1998; revised edition.

[7] G. Klambauer, T. Unterthiner, A. Mayr, S. Hochreiter, Self-normalizing neural networks, in: Advances in Neural Information Processing Systems, 2017.

[8] I. Goodfellow, Y. Bengio, A. Courville, Deep Learning, MIT Press, 2016.