Combined data-driven and model-driven location of lightning strikes: Application to meshed HVDC grids

Paul Verrax, Michel Kieffer, Louis Milhiet, Bertrand Raison

Combined data-driven and model-driven location of lightning strikes: Application to meshed HVDC grids

Paul Verrax, Michel Kieffer, *Senior Member, IEEE*, Louis Milhiet, and Bertrand Raison, *Senior Member, IEEE*

***Abstract*—This paper addresses the low-delay location of faults due to lightning strikes using single-ended measurements in HVDC grids. A combined data-driven and model-driven approach is used to estimate the fault location iteratively. No prior knowledge of the voltage evolution at the fault location is required. Once the first wavefront due to a fault is detected at some observation location of the HVDC grid, the measurements are fed to a model of the propagation of transient waves along the lines. The model is parameterized by the estimated fault location. As a result, a model of the waveform samples starting from the second wavefront is obtained. The measured and modeled waveform samples are then compared to update the estimate of the fault location. The performance of the proposed approach is evaluated via simulations, including field measurements of lightning currents. Typical location accuracy of 300 m is obtained by considering observations performed at 1 MHz over an observation time interval of less than 1.5 ms.**

***Index Terms*—Lightning strikes, fault location, data-driven approach, model-driven approach.**

I. Introduction

A. Context

Lightning is one of the most frequent sources of outage for power system components, in particular for overhead transmission lines (OHL) [1]. Depending on the magnitude of the lightning current and other parameters such as the grounding of the tower, lightning strikes may cause the insulator to flash, which creates a fault, or only a disturbance that propagates along the line. Faults (or disturbances) may also be due to vegetation or pollution. In what follows, the term *event* refers to any situation (lightning strike, short-circuit...) in the grid causing a fault or a disturbance.

Whatever the event causing a fault, protection algorithms [2] must be able to detect and identify the affected line, in order to isolate it with a minimum delay. Events causing only a disturbance should not lead to the triggering of protection equipment. Moreover, after the identification and clearing of the fault, the accurate location of the fault is important for line patrol as well as diagnosis and monitoring of the line [3].

In HVAC grid, the widely used distance protection operates after one to five cycles and requires observations of voltage (and current) during dozens of milliseconds after the event occurrence [4]. Over such time intervals, the impact of the lightning current vanishes and becomes negligible compared to the short-circuit behavior. In HVDC grid, however, the limited current breaking capability of HVDC breakers requires that the identification of the fault is performed in few milliseconds. This significantly limits the amount of available observations for post-mortem actions such as accurate fault location. Considering an observation time interval of few milliseconds, the behavior of the voltage and current along lines is governed by the propagation of the traveling waves (TW) throughout the grid and the direct impact of the lightning current cannot be neglected. One of the main difficulties then lies in the representation of the lightning current whose amplitude and waveform may vary considerably, due to their stochastic nature [5]. Pure model-based approach can hardly represent the variety of lightning waveforms. Alternative approaches are then necessary.

B. Overview of the proposed approach

This paper proposes a single-ended hybrid data-driven and model-driven approach to estimate the lightning strike locations, without assuming any prior knowledge of the lightning current at the strike location. The proposed algorithm is assumed to be run within a protection relay (the observation location) monitoring a transmission line and to exploit only its own measurements. A sketch of the proposed event location approach is presented in Figure 1. Its main steps are as follows.
1. Consider an initial, possibly very coarse, estimate of the event location.
2. Once an event is detected at the observation location (OL), record the samples of the waveforms starting from the event detection time (samples starting with the first wavefront of the TW generated by the event).
3. Feed the *measured* waveform samples as an input to a simplified TW propagation model of the grid. The model parameters depend on the estimated event location. Modeled waveform samples at the observation location are then obtained from the model.
4. An evaluation time interval is built starting at the time of the arrival of the second wavefront of the TW and ending at

This work was carried out at the SuperGrid Institute, an institute for the energetic transition (ITE). It is supported by the French government under the frame of "Investissements d'avenir" program with grant reference number ANE-ITE-002-01.

P. Verrax and L. Milhiet are with the Supergrid Institute, F-69100 Villeurbanne, France, e-mail: paul.verrax@supergrid-institute.com.

M. Kieffer is with Univ Paris-Saclay, CNRS, CentraleSupelec, Laboratoire des Signaux et Systèmes, F-91192 Gif-sur-Yvette, e-mail: michel.kieffer@l2s.centralesupelec.fr.

B. Raison is with the Univ. Grenoble Alpes, CNRS, Grenoble INP*, G2Elab, F-38000 Grenoble (* Institute of Engineering Univ. Grenoble Alpes), e-mail: Bertrand.Raison@univ-grenoble-alpes.fr.

the arrival time of the first wavefront of a TW not predicted by the model of the grid (both times also depend on the estimated event location).
5. The mean squared error (MSE) between the measured samples (at Step 2) and the modeled samples (at Step 3) over the evaluation time interval is computed as well as its derivative with respect to the event location to update the estimated event location.
6. As long as the estimate changes significantly, return to Step 3.

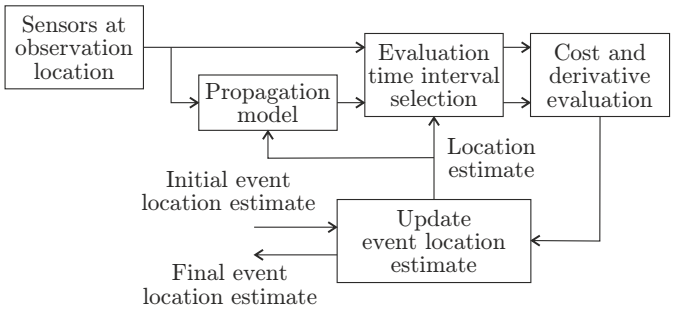


Fig. 1. Overview of the proposed event location approach involving single-ended measurements from sensors at some observation location. The measurements are fed to the input of a simplified TW *propagation model* of the grid. This model allows to simulate the evolution of the waveform at the observation location considering an estimated event location. A *cost function and its derivative* with respect to the estimated event location are evaluated by comparing the modeled waveform and the measurements over a properly chosen *evaluation time interval*, also depending on the estimated event location. An *update of the event location estimate* is then performed and the process is iterated until convergence.

This approach is partly data-driven as it uses the waveform samples acquired once an event is detected at the OL to feed a simplified TW propagation model of the grid (this is the model-driven part of the approach). The model is used to simulate the second and following wavefronts due to reflections and transmissions of the first and following wavefronts. To limit complexity, the simplified TW propagation model does not account for all nodes of the grid. Consequently it is only able to produce approximate simulations over a limited time interval. Then, the measured and modeled waveforms are compared over a time interval starting at the time of arrival of the second waveform and ending when the TW model of the grid is no more valid. The resulting cost function and its derivative with respect to the event location are used to update the estimate of the event location until convergence.

This combined data-driven and model-driven approach provides a model of the subsequent voltage waves at the observation location whatever the lightning current that stroke the monitored line. The proposed approach is also able to account for faults that are not due to lightning, for instance to pollution or vegetation.

To get more insights about the proposed approach, consider Figure 2, which provides an example of the voltage observed at a station ending a 350 km long line. Figure 2 shows the measured voltage after a fault due to a direct strike and the voltage resulting from a fault due to pollution, both occurring 70 km away from the station on the negative pole. Assuming the event happens at $t = 0$ ms, the first TW originating from

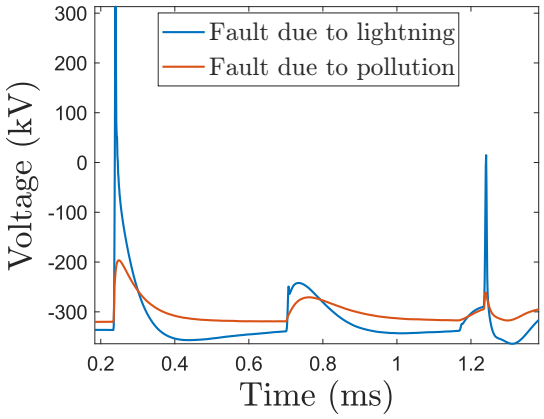


Fig. 2. EMT simulation of voltage waveforms observed at the end of a line for two events occurring 70 km away from the observation location: fault caused by a direct strike on the negative pole (in blue) and fault due pollution (in red)

the fault reaches the station at around $t = 0.2$ ms. The second wave, observed after $t = 0.7$ ms, is due to the reflection of the first wave at the station and at the fault location. Then, two different waves arrive around $t = 1.2$ ms, one after a second reflection at the fault and one due to the reflection at an adjacent station within the grid.

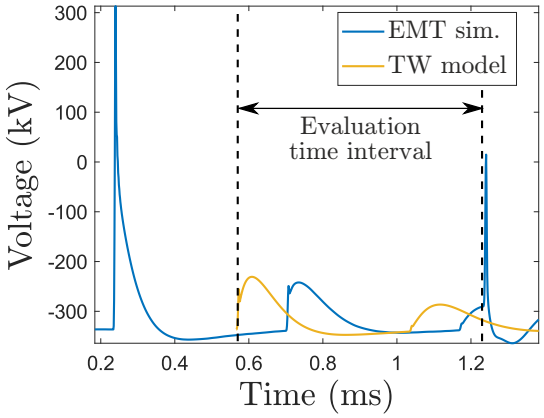


Fig. 3. Voltage waveforms observed at the end of a line: EMT simulation of a fault caused by a direct strike on the negative pole 70 km away from the observation location (in blue) Simulation obtained by from a TW propagation model fed by the first samples of the EMT simulated waveform considering an estimated event location 50 km away from the observation location (in yellow); evaluation time interval for the cost function.

A sketch of the proposed location approach is presented in Figure 3. An initial, possibly very coarse, estimate of the event location is first considered (here 50 km). The measurements following the time of arrival of the first TW observed at the station (from $t = 0.2$ s in Figure 3) are fed to the input of a direct TW propagation model to estimate the voltage of the subsequent TW. Considering the estimate of the event location, they should reach the OL at t =0.56 ms. The MSE between the voltage of the simulated TW and the voltage measured at the station is evaluated over a time interval consistent with the estimated event location (here between $t = 0.56$ s and $t = 1.26$ s in Figure 3). As the estimated event location differs from the actual strike location, the MSE will be large. The

estimated event location is updated using the MSE and its derivative with respect to the estimated event location and a new iteration may start.

The paper is structured as follows. Related works are briefly presented in Section II. The modeling approach is detailed in Section III for the initial voltage waveform at the event location based on the voltage measured at the end of the line as well as for the subsequent voltage TW. Section IV describes the iterative event location estimation algorithm. Section V evaluates the model accuracy and the performance of the proposed algorithm on simulations including measured bipolar lightning strikes.

In general, upper case letters refer to Laplace domain variables, while lower case letters are used for time domain variables. Vectors are written in bold, and matrices into brackets.

NOMENCLATURE

$[Z], [Y]$	Distributed series impedance and shunt admittance matrix, respectively
$[H]$	Propagation function
$[K_{1\to2}], [T_{1\to2}]$	Reflection and transmission matrices from medium 1 to medium 2, respectively
$\mathbf{p}, \mathbf{p}^*, \widehat{\mathbf{p}}$	Current, true, and estimated value of the unkown fault parameter vector
$\mathbf{V}, \mathbf{V}^{\mathrm{m}}$	Total measured and modeled voltage at observation point, respectively
$\mathbf{V}_{\mathrm{init}}, \widehat{\mathbf{V}}_{\mathrm{init}}$	Initial voltage at the event location and its estimate, respectively
$\mathbf{V}^{\mathrm{m}}_{\mathrm{t,i}}$	Measured and modeled voltage at the observation point due to the i-th wave, output, respectively
τ	Length of the evaluation time interval
c	Cost function
c_{w}	Propagation speed of TW
d_{f}	Fault distance
d_e	Length of line e
R_{f}	Fault resistance
R_{g}	Grounding resistance of the transmission tower
$t_{\mathrm{f}}, t_{\mathrm{d}}$	Fault time and detection time, respectively
$t_{\mathrm{m}}, t_{\mathrm{nm}}$	Arrival time of the first model and first non-modeled waves, respectively

II. RELATED WORK

As HVDC breakers typically have limited current breaking capabilities [6], they must act before the fault current exceeds their maximum breaking current. Location algorithms that try to benefit from the propagation of TW are thus best suited for DC protection purposes [7]. These algorithms may be single-ended or double-ended, depending on whether they use measurements from one end or both ends of the line, see for instance [8]. Double-ended methods require properly synchronized measurement devices as well as a reliable communication channel between the two ends of the line, which may introduce additional delays [9]. One of the main challenge associated with single-ended TW based methods is the ability to detect and interpret correctly the multiple reflected and transmitted waves that travel throughout the grid. This is especially true for complex waveforms resulting, *e.g.*, from lightning strikes [3].

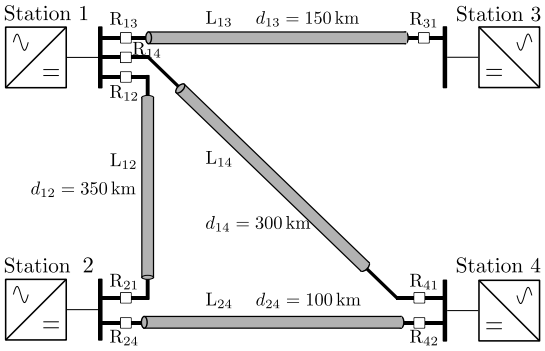


Fig. 4. Example of a four-station meshed HVDC grid: Each transmission line is protected by the two relays located at both ends of the line.

The problems raised by lightning strikes are generally considered in insulation coordination studies, see for instance [10]. Such studies involve a detailed model of the transmission lines, and of several towers. The lightning itself is generally represented using a standardized waveform, *e.g.*, the CIGRE current source [11] or the 1.2/50 waveform [12]. By contrast, faults in protection studies are generally represented as simple short circuits, regardless of the cause of the fault. Given the short time scales on which HVDC protection actions must be triggered, the TW that reach the line end after a lightning strike differs significantly from that caused by a short circuit, for instance due to pollution, as illustrated in Figure 2. Methods that make strong assumptions on the characteristic of the voltage surge at the event location, *e.g.*, considering a voltage step, are likely to malfunction in case of events that present different characteristics, such as lightning strikes.

Some papers account for different types of faults, using standardized waveforms for the lightning current. In [13], [14], several measurement-based ratios and thresholds are proposed to distinguish shielding failures with flashover from lightning disturbances. Considering the same standard waveform, [15] proposes a double-ended location method where the TW are detected from the residuals of a Kalman filter. Such approaches heavily rely on the chosen waveform, *e.g.*, a double exponential, to represent the lightning current for the definition of relevant thresholds. Methods that are designed and calibrated with such data are likely to malfunction in the presence of actual lightning measurements due to the discrepancy between standard and actual lightning current.

Machine learning approaches have been recently introduced to locate and classify faults affecting transmission lines. In the context of AC distribution grids, [16] used gradient boosting trees to detect faults and identify the faulty branch as well as to classify the type of fault. The method uses few input features such as the voltage and current recorded in all the branches of the network. The method shows robust identification accuracy with respect to the grid topology and the fault resistance. In a similar context, a convolutional neural network (CNN) is used in [17] to classify the fault, including the affected feeder and the type of fault, but not the root cause of the failure. The CNN

is fed with voltage phasor measurements over particularly long time intervals (5 to 20 s) that are first pre-processed using dynamic mode decomposition and continuous wavelet transform.

A genetic fuzzy system is proposed in [18] to distinguish internal from external faults using principal components analysis of the wavelet transform of the DC current after a fault. The method is successfully employed on a cable-based multi-terminal HVDC system and requires short observation time interval of about 1 ms. As for all machine learning based approaches, the main limit lies in the required training data set, that must account for all possible faults that will have to be identified. This leads to a high number of simulations to be performed off-line, varying all relevant parameters such as the faulty line, fault distance, fault resistance. The robustness of such approach to the wide variety of lightning induced waveforms is difficult to ensure.The Electro Magnetic Time Reversal (EMTR) approach is applied in [19] to locate lightning strikes affecting OHL. This approach assumes Telegrapher's equations, used to model the transient behavior after the inception of a fault, as being invariant through time reversal. A set of candidate fault locations (CFL) is considered. The recorded transient signals are time-reversed and back-injected from the observation point in a model of the grid. The current at each CFL is then calculated. This requires one simulation per CFL. The CFL for which the current signal energy is maximum is taken as the estimate of the fault location. The EMTR approach is also partly data-driven, as it does not require any prior knowledge of the waveform at the fault location. Nevertheless, as only the lossless approximation of the Telegrapher's equations are invariant through time reversal [20], the EMTR approach is not well suited for protection of long transmission lines where attenuation and distortion of TW cannot be neglected. This can be compensated by the use of long observation time intervals, *e.g.*, 30 times the propagation delay, which corresponds to more than 10 ms of data for lines above 100 km [21]. Such long observation time intervals may typically not be available in HVDC protection due to the stringent tripping time constraints.

To reduce the number of online simulations to be performed, [21] proposes an entirely data-driven approach. Again, a large set of CFL is considered. For each CFL, an arbitrary signal is injected into a network model at the observation location. The current waveform data at the observation location are then stored. These simulations are performed offline and do not require any time reversal. When a fault is detected, the resulting waveform is directly convoluted with the various offline simulated waveforms. The CFL for which the energy of the resulting signal is maximum is selected as the estimated fault location. Compared to the approach of [19], this algorithm is fast as it does not require any online simulation of the propagation of waves within the network. There are, however, several downsides. The observation time interval has to be longer than that of the original EMTR approach. The fault location is correctly estimated when the peaks (wavefronts) of the measured waveform and of one of the simulated waveforms coincide. The approach is very inaccurate when only two or three wavefronts are observed. In addition, EMTR approaches and their variants assume that large resistors are placed at each end of the transmission lines [22], [20]. This may not be the case in meshed grid typologies where rather small DC reactors (*e.g.,* 50 mH) are placed at the end of lines, or even omitted and located at the converter output.

By contrast, the proposed event location approach is based on a TW propagation model of the voltage at the end of the line that accounts for voltage wave distortion and attenuation along the line. This propagation model makes the use of a short observation time interval of typically less than 1 ms.

III. Traveling wave modeling

Events occurring on transmission lines generate transient phenomena that can be described with the theory of TW, briefly recalled in Section III-A. In this paper, TW are employed to obtain a model of the voltage at the OL, where sensors are available. Using an estimate of the event location, this model can be inverted to infer the voltage waveform at the event estimated location, as described in Section III-B. The estimated voltage waveform at event location can then be used to compute subsequent TW reaching the station located at the end of the line, see Section III-C. As will be seen, the estimate of the voltage at the event location is no more required: Only the voltage samples acquired at the OL are necessary to infer the subsequent waveform samples for a given value of the event parameters.

A. Traveling waves

Consider a generic transmission line e of length d_e, comprising n_c conductors, with stations q and q' at its extremities. The considered OL sensors are at the output of station q. Consider an event occurring at a distance d_f from station q and $d_e - d_\mathrm{f}$ from station q'. The evolution of the voltage $\mathbf{V}(s,x)$ at a location x along the line is described in the Laplace domain by the Telegrapher's equation [23]

$$\frac{\partial^2 \mathbf{V}(s,x)}{\partial x^2} = [Z(s)]\,[Y(s)]\,\mathbf{V}(s,x)\,, \tag{1}$$

where $[Z(s)]$ and $[Y(s)]$ are the distributed series impedance and shunt admittance matrices, respectively and s is the Laplace variable. To be solved, (1) must be decoupled through the transformation of phase voltage to modal voltage, using a matrix $[T_V]$. For the sake of simplicity, all employed quantities, including measurements, are assumed in what remains to be expressed in modal domain.

The solution of (1) consists of the superposition of two waves traveling in opposite directions. The wave traveling towards the positive x is

$$\mathbf{V}(s,x) = \underbrace{\exp\left(-\left[\Gamma(s)\right]x\right)}_{[H(s,x)]}\mathbf{V}_{\text{init}}(s) \tag{2}$$

where $\mathbf{V}_{\text{init}}(s)$ is the initial surge at the event location, $[\Gamma(s)] = \sqrt{[Z(s)]\,[Y(s)]}$ is the propagation matrix, and $[H(s,x)]$ is the propagation function. When a forward wave $\mathbf{V}_\mathrm{f}$ is traveling from a medium of surge admittance $[Y_{\mathrm{s},1}]$ to a medium of surge admittance $[Y_{\mathrm{s},2}]$, a reflected wave

$\mathbf{V}_\mathrm{r} = [K_{1\to2}]\,\mathbf{V}_\mathrm{f}$ and a transmitted wave $\mathbf{V}_\mathrm{t} = [T_{1\to2}]\,\mathbf{V}_\mathrm{f}$ appear with the reflection coefficient

$$[K_{1\to2}] = \left([Y_{\mathrm{s},2}] + [Y_{\mathrm{s},1}]\right)^{-1}\left([Y_{\mathrm{s},1}] - [Y_{\mathrm{s},2}]\right) \tag{3}$$

and the transmission coefficient

$$[T_{1\to2}] = [K_{1\to2}] + [I_{n_\mathrm{c}}]\,.$$

The voltage measured at the interface corresponds to the transmitted wave.

The voltage at the end of the line due to the arrival of the first TW after the event occurrence is

$$\mathbf{V}_{\mathrm{t},1}\,(s, d_\mathrm{f}) = \underbrace{\left([T_q(s)][H\,(s, d_\mathrm{f})]\right)}_{[H_1(s,d_\mathrm{f})]}\mathbf{V}_{\text{init}}\,(s)\,, \tag{4}$$

where $[T_q(s)]$ is the transmission coefficient from the line to the station q. $[T_q(s)]$ accounts for the impact of other lines connected to station q. As shown in [24], the attenuation and distortion of TW can be accounted for using relatively simple models such as first order filters for $[H\,(s, d_\mathrm{f})]$.

If the characteristics of the event are known, $\mathbf{V}_{\text{init}}\,(s)$ may be computed. A fault due to pollution can for instance be represented as a voltage source connected to the ground, in series with a resistance R_f that closes at time t_f onto the affected conductor [25]. The initial voltage surge is then

$$\mathbf{V}_{\text{init}}\,(s, R_\mathrm{f}) = [K_{\mathrm{e}\to\mathrm{f}}\,(R_\mathrm{f})]\,\frac{\exp\left(-t_\mathrm{f}s\right)}{s}\mathbf{v}_{\mathrm{bf}},$$

where $\mathbf{v}_{\mathrm{bf}}$ is the phase voltage at the event location just before the occurrence of the fault. The reflection coefficient from the line to the fault, $[K_{\mathrm{e}\to\mathrm{f}}\,(R_\mathrm{f})]$, depends on the type of fault (*e.g.,* pole-to-ground) and on the fault resistance R_f. The voltage due to subsequent waves can also be computed. For example, the wave, reflected at station q and then at the fault location, traveling back towards station q is

$$\begin{aligned}\mathbf{V}_{\mathrm{t},2}\,(s, \mathbf{p}) = {}& [T_q]\,[H\,(s, d_\mathrm{f})]\,[K_{\mathrm{e}\to\mathrm{f}}\,(R_\mathrm{f})]\,\times \\ & [H\,(s, d_\mathrm{f})]\,[K_q(s)]\,[H\,(s, d_\mathrm{f})]\,\mathbf{V}_{\text{init}}\,(s)\,.\end{aligned} \tag{5}$$

The vector $\mathbf{p} = (R_\mathrm{f}, d_\mathrm{f})$ gathers the parameters characterizing the fault, *i.e.*, the fault location d_f and the fault resistance R_f. Similarly the voltage wave that travels first towards the station q' at the remote end of the line is expressed as

$$\begin{aligned}\mathbf{V}_{\mathrm{t},3}\,(s, \mathbf{p}) = {}& [T_q]\,[H\,(s, d_\mathrm{f})]\,[T_{\mathrm{e}\to\mathrm{f}}\,(R_\mathrm{f})]\,\times \\ & [H\,(s, d_e - d_\mathrm{f})]\,[K_{q'}(s)]\,[H\,(s, d_e - d_\mathrm{f})]\,\mathbf{V}_{\text{init}}\,(s)\,.\end{aligned} \tag{6}$$

More generally, any wave that travels throughout the grid can be modeled to get an expression similar to that of (6), possibly with more terms to account for additional reflections, transmissions, and propagation. This requires the characteristics of the grid and the event parameters to be known, as well as the path through the grid followed by the considered TW, see Section IV for more details.

B. Data-driven estimate of the initial voltage

In the general case, the characteristics of the event are unknown *a priori*. This is especially the case for faults due to lightning strikes, as the initial fault voltage depends on the lightning current. Standard waveform models have been proposed such as the CIGRE current source or the 1.2/50 [11], [12]. Nevertheless, the amplitude and waveform of the lightning current is stochastic. Consequently, in the general case, the forward model (4) cannot be directly applied to locate the event, as the initial fault voltage $\mathbf{V}_{\text{init}}$ is unknown.

Nevertheless, assuming that the fault location d_f is known, the initial voltage $\mathbf{v}_{\text{init}}\,(t)$ can be estimated based on the *measured* voltage $\mathbf{v}_{\mathrm{t},1}\,(t)$ at station q by inverting (4), *i.e.*,

$$\widehat{\mathbf{V}}_{\text{init}}\,(s, d_\mathrm{f}) = [H_1(s, d_\mathrm{f})]^{-1}\mathbf{V}_{\mathrm{t},1}\,(s)\,. \tag{7}$$

The dependency in d_f has been removed in the time domain expression of $\mathbf{v}_{\mathrm{t},1}\,(t)$ and in its Laplace transform $\mathbf{V}_{\mathrm{t},1}\,(s)$ in (7) and in what follows to indicate that these are *measured* quantities. Evaluating the inverse of $[H_1(s, d_\mathrm{f})]$ is not straightforward as the transfer function comprises delays, making the inverse acausal. One should thus compensate for this delay to compute (7). Note that (7) assumes the event distance d_f as known, but does not involve the fault resistance. An estimate of the initial voltage $\mathbf{v}_{\text{init}}\,(t)$ can thus be computed regardless of the type of event and potential fault resistance, provided that the actual value of d_f is available, or at least an estimate of it. Taking the inverse Laplace transform of (7), one gets

$$\widehat{\mathbf{v}}_{\text{init}}(t, d_\mathrm{f}) = \mathcal{L}^{-1}\left([H_1(s, d_\mathrm{f})]^{-1}\right) \otimes \mathbf{v}_{\mathrm{t},1}\,(t)\,, \tag{8}$$

where $\otimes$ is the convolution product.

The estimate $\widehat{\mathbf{v}}_{\text{init}}(t, d_\mathrm{f})$ is useful to classify the type of event. Nevertheless, as will be seen in what follows, it is not necessary in the proposed event location approach.

C. Model for subsequent waves

In Section III-A, a model of the different waves traveling from the fault towards the station q and q' at both ends of the line has been obtained, assuming the initial voltage is available. Conversely, in Section III-B, the model of the first TW is inverted to compute the initial voltage at the event location from voltage measurements at the station. In this section, these two results are combined to estimate the TW following the first one independently of any prior knowledge of the initial voltage at event location.

Using (7) in the expression (5) of the second wave traveling back from the fault location, one gets

$$\begin{aligned}\mathbf{V}^{\mathrm{m}}_{\mathrm{t},2}\,(s, \mathbf{p}) = {}& [T_q]\,[H\,(s, d_\mathrm{f})]\,[K_{\mathrm{e}\to\mathrm{f}}\,(R_\mathrm{f})]\,[H\,(s, d_\mathrm{f})]\,\times \\ & [K_q(s)]\,[H\,(s, d_\mathrm{f})]\,\underbrace{\left([T_q(s)]\,[H\,(s, d_\mathrm{f})]\right)^{-1}\mathbf{V}_{\mathrm{t},1}\,(s)}_{\widehat{\mathbf{V}}_{\text{init}}(s,d_\mathrm{f})},\end{aligned} \tag{9}$$

which boils down to

$$\mathbf{V}^{\mathrm{m}}_{\mathrm{t},2}\,(s, \mathbf{p}) = [H_{q,\mathrm{f}}\,(s, \mathbf{p})]\,\mathbf{V}_{\mathrm{t},1}\,(s)\,, \tag{10}$$

with

$$\begin{aligned}[H_{q,\mathrm{f}}\,(s, \mathbf{p})] = {}& [T_\mathrm{q}]\,[H\,(s, d_\mathrm{f})]\,[K_{\mathrm{e}\to\mathrm{f}}\,(R_\mathrm{f})] \\ & [H\,(s, d_\mathrm{f})]\,[K_q(s)]\,[T_q(s)]^{-1}\,.\end{aligned} \tag{11}$$

In this specific case, $[H_{q,\mathrm{f}}(s,\mathbf{p})]$ is directly causal, which simplifies the evaluation. The expression (10) allows one to compute an estimate of the second TW (which has been reflected once at the fault) based on the measurements of the first TW $\mathbf{v}_{\mathrm{t},1}(t)$. The second TW depends on the fault distance d_{f} as well as on the fault resistance, via $[K_{\mathrm{e}\to\mathrm{f}}(R_{\mathrm{f}})]$, due to the reflection at the fault location. The expression (10) can be extended to the n-th wave traveling from the line end to the fault location and back to the line end, as

$$\mathbf{V}^{\mathrm{m}}_{\mathrm{t},n}(s,\mathbf{p}) = [H_{q,\mathrm{f}}(s,\mathbf{p})]\,\mathbf{V}_{\mathrm{t},n-1}(s). \tag{12}$$

The model $\mathbf{V}^{\mathrm{m}}_{\mathrm{t},n}(s,\mathbf{p})$ depends on $\mathbf{V}_{\mathrm{t},n-1}(s)$ which should be measured at station q. Unfortunately, in general, only $\mathbf{V}_{\mathrm{t},1}(s)$ may be measured, since during the first milliseconds after the detection of an event, only the first wave appears (see Figure 2). The second and following waves superpose to each other, and it is difficult to isolate $\mathbf{V}_{\mathrm{t},n-1}(s)$.

To address this issue, one observes in (12) the additional distortion and attenuation of the wave $\mathbf{V}_{\mathrm{t},n-1}(s)$ after it traveled back and forth between station q and the fault location. Considering (12) for all $n \geqslant 1$, the total voltage at the station q is obtained as

$$\begin{aligned}\mathbf{V}^{\mathrm{m}}_{\mathrm{t}}(s,\mathbf{p}) &= \sum_{n=0}^{\infty}\mathbf{V}^{\mathrm{m}}_{\mathrm{t},n+1}(s,\mathbf{p})\\ &= \mathbf{V}_{\mathrm{t},1}(s) + \sum_{n=1}^{\infty}[H_{q,\mathrm{f}}(s,\mathbf{p})]\,\mathbf{V}_{\mathrm{t},n}(s)\\ &= \mathbf{V}_{\mathrm{t},1}(s) + [H_{q,\mathrm{f}}(s,\mathbf{p})]\,\mathbf{V}_{\mathrm{t}}(s)\end{aligned} \tag{13}$$

where

$$\mathbf{V}_{\mathrm{t}}(s) = \sum_{n=1}^{\infty}\mathbf{V}_{\mathrm{t},n}(s)$$

is the voltage measured at station q gathering the contributions of all TW. Assuming that the fault parameter vector $\mathbf{p} = (R_{\mathrm{f}}, d_{\mathrm{f}})$ is known, compared to (12), (13) gives an estimate, of the total voltage at station q based on the total measured voltage $\mathbf{V}_{\mathrm{t}}(s)$.

Similar derivations can be performed for waves following a different path, in particular for TW first reflected at the remote station q' and traveling back towards the station q, by combining (6) and (7).

The voltage in the temporal domain is obtained using the inverse Laplace transform of (13) to get

$$\begin{aligned}\mathbf{v}^{\mathrm{m}}_{\mathrm{t}}(t,\mathbf{p}) &= \mathcal{L}^{-1}\left(\mathbf{V}_{\mathrm{t},1}(s) + [H_{q,\mathrm{f}}(s,\mathbf{p})]\,\mathbf{V}_{\mathrm{t}}(s)\right).\\ &= \mathbf{v}^{1}_{\mathrm{t}}(t) + \mathcal{L}^{-1}\left([H_{q,\mathrm{f}}(s,\mathbf{p})]\right)\otimes\mathcal{L}^{-1}\left(\mathbf{V}_{\mathrm{t}}(s)\right)\\ &= \mathbf{v}^{1}_{\mathrm{t}}(t) + [h_{q,\mathrm{f}}(t,\mathbf{p})]\otimes\mathbf{v}_{\mathrm{t}}(t).\end{aligned} \tag{14}$$

where the impulse response $h_{q,\mathrm{f}}(t,\mathbf{p})$ can be computed analytically for any value of $\mathbf{p}$. For measurement times at which only the first TW reaches the station, the model boils down to

$$\mathbf{v}^{\mathrm{m}}_{\mathrm{t}}(t) = \mathbf{v}^{1}_{\mathrm{t}}(t).$$

In practice, when the event is not too close to the station, the voltage induced by the first TW $\mathbf{v}^{1}_{\mathrm{t}}(t)$ tends to zero when the subsequent waves arrive at some time t_{s} and the model (14) becomes

$$\mathbf{v}^{\mathrm{m}}_{\mathrm{t}}(t,\mathbf{p}) = [h_{q,\mathrm{f}}(t,\mathbf{p})]\otimes\mathbf{v}_{\mathrm{t}}(t),\ t > t_{\mathrm{s}}. \tag{15}$$

The main idea of the proposed fault location approach is, from the first samples of the signal $\mathbf{v}_{\mathrm{t}}(t)$ acquired at the OL, to infer its next samples using (14) in Section III-C. We focus on the samples related to the subsequent TW following the first TW reaching the station. For a given estimate of the event location, the inferred waveform samples are then compared to the observed ones to potentially update the estimate of the event location, see Section IV-A.

IV. Fault location

In Section III-C, the presentation focused on waves traveling between the event location and the station. In practice, TW may follow more complex paths through the network before reaching the station in the evaluation time interval used to determine the event location. Ideally, all waves should be taken into account in the TW propagation model considered in the proposed algorithm. Nevertheless, to limit the evaluation complexity of the model, only a subset of TW is simulated. The time interval over which the comparison between the waveforms at model output and the measured waveforms is performed depends on the model complexity, as described in Section IV-B. The estimated location may then be updated if necessary, as detailed in Section IV-C.

In this section, the (unknown) true value $\mathbf{p}^* = (d^*_{\mathrm{f}}, R^*_{\mathrm{f}})$ of the fault parameters is distinguished from a generic value $\mathbf{p} = (d_{\mathrm{f}}, R_{\mathrm{f}})$ of the fault parameters.

A. Fault location approach

An event occurring at time t_{f} and at a distance d_{f} from the station where the measurements are performed will be detected around time

$$t_{\mathrm{d}} = t_{\mathrm{f}} + d_{\mathrm{f}}/c_{\mathrm{w}}, \tag{16}$$

where c_{w} is the TW propagation speed. During a short time interval starting at t_{d}, only the first TW is observed in $\mathbf{v}_{\mathrm{t}}(t)$. The corresponding samples can be used, as detailed in Section III, to infer the voltage at the event location. The second TW reaches the station at time t_{m}. This is the first wave obtained at the output of the model (15) when it is fed with the first samples of $\mathbf{v}_{\mathrm{t}}(t)$. From that time instant, the measurements may be used to evaluate the consistency of the estimated event location. Measurements can be exploited until t_{nm}, the time at which the first wave that traveled through a path not accounted for in the model reaches the station. Depending on the estimate of the event location, the role of the *Evaluation time interval selection* block in Figure 1 is to determine the time interval $[t_{\mathrm{m}}, t_{\mathrm{nm}}]$ over which the model output can be compared with the measurements. This is further explained in Section IV-B.

In the example of Figure 2, $t_{\mathrm{d}} = 0.23$ ms. The second TW arrives at $t_{\mathrm{m}} = 0.7$ ms and a third wave arrives at $t = 1.2$ ms due to the reflection at an adjacent station. If the model does not account for such reflections, only measurements between

$t_{\mathrm{m}} = 0.7\,\mathrm{ms}$ and $t_{\mathrm{nm}} = 1.2$ ms can be exploited to update the estimate of the event location.

Considering the (unknown) true value of the fault parameters $\mathbf{p}^*$, over the interval $[t_{\mathrm{m}}, t_{\mathrm{nm}}]$, the voltage measurements $\mathbf{v}_{\mathrm{t}}(t)$ are assumed to be described as

$$\mathbf{v}_{\mathrm{t}}(t) = \mathbf{v}_{\mathrm{t}}^{\mathrm{m}}(t, \mathbf{p}^*) + \varepsilon(t), \quad \forall t \in [t_{\mathrm{m}}, t_{\mathrm{nm}}], \tag{17}$$

where $\mathbf{v}_{\mathrm{t}}^{\mathrm{m}}(t, \mathbf{p}^*)$ is given by (14) when $\mathbf{p} = \mathbf{p}^*$ and $\varepsilon(t)$ represents the zero-mean white Gaussian measurement noise. The maximum likelihood estimate $\widehat{\mathbf{p}}$ [26] of $\mathbf{p}^*$ is

$$\widehat{\mathbf{p}} = \arg\min_{\mathbf{p}} c(\mathbf{p}) \tag{18}$$

with

$$c(\mathbf{p}) = \sum_{k=1}^{n} \left(\mathbf{v}_{\mathrm{t}}(t_k) - \mathbf{v}_{\mathrm{t}}^{\mathrm{m}}(t_k, \mathbf{p})\right)^2, \tag{19}$$

and where $n = f_{\mathrm{s}} \lfloor t_{\mathrm{nm}} - t_{\mathrm{m}} \rfloor$ is the number of samples in the evaluation time interval when sampling is performed at a frequency f_{s} and $t_k = \left((n-k)\,t_{\mathrm{m}} + (k-1)\,t_{\mathrm{nm}}\right)/(n-1)$. The minimization of $c(\mathbf{p})$ is then performed iteratively, *e.g.*, using the Levenberg-Marquardt [26] algorithm to evaluate an estimate $\widehat{\mathbf{p}}$ of $\mathbf{p}^*$. The determination of the appropriate observation time interval $[t_{\mathrm{m}}, t_{\mathrm{nm}}]$ over which the cost function (18) has to be evaluated is essential and is addressed in Section IV-B.

B. Evaluation time interval determination

This section describes a method to evaluate an appropriate evaluation time interval for the samples involved in the cost function (18). It consists in determining the arrival times t_{m} of the first *modeled wave* and t_{nm} of the first *non-modeled wave* of the TW propagation model. To this end, the HVDC grid is described using an undirected graph $\mathcal{G} = (\mathcal{Q}, \mathcal{E})$ composed of vertices Q and edges $\mathcal{E}$ as proposed in [27] and briefly recalled here.

Each vertex $q \in \mathcal{Q}$ of $\mathcal{G}$ represents an interconnection between two or more lines or between a line and a station. Each line is represented by an edge $e \in \mathcal{E}$ of the graph $\mathcal{G}$. The edge between the nodes q_i and q_j is denoted $e_{i,j}$. Since the graph is undirected, $e_{i,j} = e_{j,i}$. The length of the segment represented by the edge $e_{i,j}$ is $d_{i,j}$. We assume that at $t = t_{\mathrm{f}}$, an event occurs in the edge $e_{\mathrm{f}} = e_{i,j} \in \mathcal{E}$. The event leads to a modification of the graph $\mathcal{G}$. A node q_{f} is added to $\mathcal{Q}$ to get $\mathcal{Q}_{\mathrm{f}} = \mathcal{Q} \cup \{q_{\mathrm{f}}\}$ and the edge $e_{\mathrm{f}} = e_{i,j} \in \mathcal{E}$ where the event occurs is replaced by the edges $e_{i,\mathrm{f}}$ and $e_{\mathrm{f},j}$ of lengths $d_{\mathrm{f},i}$ and $d_{\mathrm{f},j}$ to get $\mathcal{E}_{\mathrm{f}} = \mathcal{E} \cup \{e_{i,\mathrm{f}}, e_{\mathrm{f},j}\} \setminus \{e_{i,j}\}$. Consider a node $q_{\mathrm{s}} \in \mathcal{Q}_{\mathrm{f}}$ at which voltage is measured. A TW is entirely determined by its *path*, *i.e.*, the sequence of nodes it has traversed. The set of all possible paths from q_{f} to q_{s} is

$$\mathcal{P}_{q_{\mathrm{f}} \to q_{\mathrm{s}}} = \{(q_{n_1}, .., q_{n_m}) \mid q_{n_1} = q_{\mathrm{f}}, q_{n_m} = q_{\mathrm{s}}, \left(q_{n_i}, q_{n_{i+1}}\right) \in \mathcal{E}_{\mathrm{f}}, m \geqslant 2\}. \tag{20}$$

A path $\pi \in \mathcal{P}_{q_{\mathrm{f}} \to q_{\mathrm{s}}}$ may comprise several times the same node, including the node q_{f} where the event occurs and the node q_{s} where the voltage is measured (the OL).

The set $\mathcal{E}_{\mathrm{f}}$ is partitioned into a set of edges $\mathcal{E}^{\mathrm{m}}$ which are accounted for in the voltage evolution model of Section III (*modeled* edges) and a set of edges $\mathcal{E}^{\mathrm{nm}} = \mathcal{E}_{\mathrm{f}} \setminus \mathcal{E}^{\mathrm{m}}$ which are not taken into account (*non-modeled* edges). The set $\mathcal{P}_{q_{\mathrm{f}} \to q_{\mathrm{s}}}$ is partitioned similarly into a subset of *modeled* paths

$$\mathcal{P}^{\mathrm{m}}_{q_{\mathrm{f}} \to q_{\mathrm{s}}} = \{(q_{n_1}, .., q_{n_{\mathrm{m}}}) \mid q_{n_1} = q_{\mathrm{f}}, q_{n_{\mathrm{m}}} = q_{\mathrm{s}}, \left(q_{n_i}, q_{n_{i+1}}\right) \in \mathcal{E}^{\mathrm{m}}, n_{\mathrm{m}} \geqslant 2\}. \tag{21}$$

and *non-modeled* paths $\mathcal{P}^{\mathrm{nm}}_{q_{\mathrm{f}} \to q_{\mathrm{s}}} = \mathcal{P}_{q_{\mathrm{f}} \to q_{\mathrm{s}}} \setminus \mathcal{P}^{\mathrm{m}}_{q_{\mathrm{f}} \to q_{\mathrm{s}}}$. The set $\mathcal{E}^{\mathrm{m}}$ has to include enough edges, especially those adjacent to the fault $\left\{e_{i,\mathrm{f}}, e_{j,\mathrm{f}}\right\}$ and to the station q_{s} to have $\mathcal{P}^{\mathrm{m}}_{q_{\mathrm{f}} \to q_{\mathrm{s}}} \neq \emptyset$. Including more edges increases the time interval over which the resulting model is valid, at the price of a larger computing complexity.

For each TW, the arrival time t_π of the wave at node q_{s} is determined by its path $\pi \in \mathcal{P}_{q_{\mathrm{f}} \to q_{\mathrm{s}}}$and the propagation speed along the path. The time t_π thus depends on the location of the event. The problem of finding t_{m} and t_{nm} can then be formulated as

$$t_{\mathrm{m}} = \min\, t_\pi + t_{\mathrm{f}},\ \pi \in \mathcal{P}^{\mathrm{m}}_{q_{\mathrm{f}} \to q_{\mathrm{s}}} \setminus \{(q_{\mathrm{f}}, q_{\mathrm{s}})\} \tag{22}$$

$$t_{\mathrm{nm}} = \min\, t_\pi + t_{\mathrm{f}},\ \pi \in \mathcal{P}^{\mathrm{nm}}_{q_{\mathrm{f}} \to q_{\mathrm{s}}}. \tag{23}$$

For an event occurring in the half of the transmission line closest to the OL, t_{m} is the time instant at which the first TW reaches the station a second time after a first reflection at the event location[1]

$$t_{\mathrm{m}}(d_{\mathrm{f}}) = t_{\mathrm{f}} + 3d_{\mathrm{f}}/c_{\mathrm{w}}. \tag{24}$$

Then using (16), one gets

$$t_{\mathrm{m}}(d_{\mathrm{f}}) = t_{\mathrm{d}} + 2d_{\mathrm{f}}/c_{\mathrm{w}}. \tag{25}$$

For events occurring in the half of the line furthest to the OL, t_{m} is the time instant at which the initial TW propagating towards the remote end of the line reaches the OL after it has been reflected at the remote station and transmitted at the event location

$$t_{\mathrm{m}}(d_{\mathrm{f}}) = t_{\mathrm{f}} + (d - d_{\mathrm{f}})/c_{\mathrm{w}} + d/c_{\mathrm{w}}. \tag{26}$$

Using again (16), one gets

$$t_{\mathrm{m}}(d_{\mathrm{f}}) = t_{\mathrm{d}} + 2(d - d_{\mathrm{f}})/c_{\mathrm{w}}. \tag{27}$$

The time instant t_{nm} can be computed by solving a multiple shortest paths problem, using a Dijkstra-like algorithm [28]. Knowing the topology of the grid, the algorithm returns all the paths taken by the waves that reach the receiving node until a non-modeled edge is taken, as well as the associated traveled distances for each paths. As for t_{m} which depends on d_{f}, an estimate of the event location is required, since all propagation delays are determined by the distance between the event location and its neighbors.

In the considered iterative estimation approach, the estimate of the event location is regularly updated to improve the cost function (19). The instants $t_{\mathrm{m}}(d_{\mathrm{f}})$ and $t_{\mathrm{nm}}(d_{\mathrm{f}})$ as well as the evaluation time interval $[t_{\mathrm{m}}(d_{\mathrm{f}}), t_{\mathrm{nm}}(d_{\mathrm{f}})]$ are thus likely to

[1]For simplicity reasons, the unusual case where a wave traveling throughout multiple stations of the grid reaches the OL before the first reflected wave is ignored. Formally, this assumes that the triangle inequality holds for the distances between any node triplet of the graph.

change from one iteration to the other. It may be convenient to perform the estimation on a time interval of constant duration τ. By choosing the length of the time interval

$$\tau = \min_{d_\mathrm{f} \in [0, d_{ij}]} \left(t_\mathrm{nm}\left(d_\mathrm{f}\right) - t_\mathrm{m}\left(d_\mathrm{f}\right) \right) \tag{28}$$

as the minimum length for all possible distances, it is ensured that non-modeled waves are never present in the evaluation time interval, whatever $d_\mathrm{f} \in [0, d_{ij}]$. It may happen that for some event location, the evaluation interval becomes too small (or even empty, *i.e.*, $\tau = 0$) to perform an accurate estimation of the event parameters. If so, the set of modeled edges $\mathcal{E}^\mathrm{m}$ has to be extended to account for more previously non-modeled waves and thus increase t_nm and consequently the size of the available evaluation time interval.

C. Global optimization

The estimation of the event parameters requires the minimization of the cost function (19). Preliminary observations showed that the sensitivity of the cost function with respect to the fault resistance is very small. For a direct lightning strike causing a negative pole-to-ground fault at a distance $d_\mathrm{f}^* = 50\,\mathrm{km}$ with a fault resistance $R_\mathrm{f}^* = 65\,\Omega$, Figure 5 (left) shows the contour plot of the cost function (19), normalized such that the minimum cost is 1. The vertical "valley shape" around $(d_\mathrm{f}^*, R_\mathrm{f}^*)$ indicates that the sensitivity of the cost function with respect to the fault resistance is very small, making it difficult to estimate. Conversely, an erroneous value $R_\mathrm{f} \neq R_\mathrm{f}^*$ does not affect significantly the accuracy of the fault distance estimate, as observed in Figure 5 (right) where $R_\mathrm{f} = 40\,\Omega \neq R_\mathrm{f}^* = 65\,\Omega$ is used. Consequently, in what

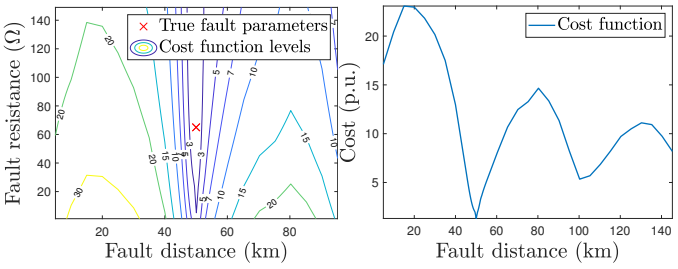


Fig. 5. Contour plot of the cost function (19) as function of the fault distance and fault resistance (left) and for a fixed resistance $R_\mathrm{f} = 40\,\Omega$ different from $R_\mathrm{f}^* = 65\,\Omega$ (right). The cost function is normalized such that its minimum is 1.

follows, the estimation of the fault parameters is limited to that of the fault distance d_f^*, considering an average value for the fault resistance set to $R_\mathrm{f} = 10\,\Omega$.

Figure 5 (right) shows that the shape of the cost function as a function of the fault distance includes several local minima. Although global optimization techniques exist, local approaches were chosen for their better computation performance. To avoid convergence towards a local minimum, several optimizations are performed in parallel with different initial values of the fault distance. The estimated event location minimizing $c\left(\mathbf{p}\right)$ is then selected as the final estimate of the event location.

Algorithm 1 summarizes the proposed approach. A list $\mathcal{D}$ of initial estimates of the event location is considered. The length τ of the evaluation time interval is evaluated offline. $\mathcal{D}$ and τ are provided as input to the algorithm (Line 1). When an unusual behavior is detected (Line 4), several optimization algorithms are launched in parallel (Line 7), for each of the initial estimate in $\mathcal{D}$. Each algorithm is run until a stopping condition (maximum number of iteration, small updates, or small gradient) is reached (Lines 8-14). Upon convergence of the optimization algorithms, the final estimate of the event location is chosen as the one providing the minimum value of the cost function (Line 18).

Algorithm 1 Event location algorithm

1: Input: List of initial estimated fault locations $\mathcal{D}$, Lenght of evaluation time interval τ
2: Output: Estimated fault location $\widehat{d_\mathrm{f}}$
3: **if** Unusual behavior detected **then**
4: Get arrival time t_d of first TW
5: Record measurements $v(t)$ from t_d
6: $k = 0$
7: **for** $d_\mathrm{f} \in \mathcal{D}$ **do**
8: **while** stopping_conditions = false **do**
9: Evaluate $t_\mathrm{m}(d_\mathrm{f})$ using (25) or (27)
10: Simulate $v^m(t, d_\mathrm{f})$ using TW propagation model (13)
11: Evaluate cost function (19) and derivative over $[t_\mathrm{m}, t_\mathrm{m} + \tau]$
12: Update d_f^k
13: Evaluate stopping_conditions
14: **end while**
15: Assign $\widehat{d}_{\mathrm{f},k} = d_\mathrm{f}$
16: $k = k + 1$
17: **end for**
18: Return $\arg\min_k c(\widehat{d}_{\mathrm{f},k})$
19: **end if**

V. Results

This section presents simulation results for the proposed modeling and fault location approaches. The EMT simulation set-up is first presented in Section V-A, including the fault and lightning representation. The accuracy of the models to evaluate the voltage at the event location and of the subsequent voltage waves at the OL is compared against EMT simulations in Section V-B. The proposed fault location algorithm is then evaluated using standard lightning waveforms as well as experimental lightning currents in Section V-C.

A. Simulation set-up

The test grid implemented in EMTP-RV [29] used for the simulation comprises 4 converter stations with half-bridge MMC as presented in Figure 4. The grid topology is taken from the benchmark grid used in the PROMOTIoN project [30]. The overhead transmission lines are composed of a positive and a negative pole connected as a rigid bipole [31] with 2 sky wires for lightning strike protection. The

configuration of the transmission lines obtained from [32] is listed in Table I. The soil resistivity is considered uniform with a constant value of $\rho = 100\,\Omega$m. The parameters of the MMC stations are listed in Table II.

TABLE I
CHARACTERISTICS OF THE OVERHEAD TRANSMISSION LINES

	Poles	Sky wires
DC resistance (Ω/km)	0.024	1.62
Height at tower (m)	37.2	41.7
Height at mid-span (m)	22.2	26.7
Horizontal distance (m)	±4.465	±3.66
Outside diameter (cm)	4.775	0.98

TABLE II
CHARACTERISTICS OF THE MMC STATIONS

Rated power (MW)	1000
DC rated voltage (kV)	320
Arm inductance (p.u.)	0.15
Transformer resistance (p.u.)	0.001
Capacitor energy in each submodule (kJ/MVA)	40
Conduction losses of each IGBT/diode (Ω)	0.001
Number of sub-modules per arm	400
Grounding impedance (Ω)	0.5

The lightning is simulated as a current source connected to the stroke point, here a conductor. Insulation coordination studies usually represent a tower with a distributed parameter model, which requires a very short simulation period to account for the propagation time along the tower. Here, a tower model based on lumped elements [33] is used, see Figure 6. The body and top of the tower are represented as series resistance and inductance, while the cross-arms are modeled using only inductances. The insulator string are represented by air gaps. The grounding of the tower, R_g, may change from one tower to another, and typically depends on the nature of the soil. As the fault resistance accounts for all the current path between the faulty conductor and the ground, it should at least comprise the tower grounding resistance, *i.e.*, $R_\mathrm{f}^* \geq R_\mathrm{g}$.

For simulation involving a standard waveform, the CIGRE current source is employed [11]. For a strike occurring at time $t = 0$, it is defined as

$$i_{\mathrm{CIGRE}}(t) = \begin{cases} At + Bt^n & \text{for } 0 \leqslant t < t_\mathrm{n} \\ i_1 e^{\frac{-(t-t_\mathrm{n})}{t_1}} - i_2 e^{\frac{-(t-t_\mathrm{n})}{t_2}} & \text{for } t \geqslant t_\mathrm{n}. \end{cases} \quad (29)$$

The two parts represent the current front and the current tail. The parameters $t_\mathrm{n}, i_1, i_2, A, B$, and n of the model (29) are deduced from four parameters: the maximum amplitude i_max, the maximum steepness S_m, the front time t_ft, and the time to half value t_h. The front time typically amounts to few microseconds and is kept constant at $t_\mathrm{ft} = 2\,\mu$s. A correlation between the maximum steepness and the amplitude is generally observed and S_m is obtained from i_max by $S_\mathrm{m} = 3.9 i_\mathrm{max}^{0.55}$ [5]. Statistical distributions have been proposed to account for the variability of the remaining parameters i_max and t_h.

Simulations involving measured lightning currents have also been conducted. The measurements have been performed at the telecom base station Miluccia in Corsica, see [34] for more details.

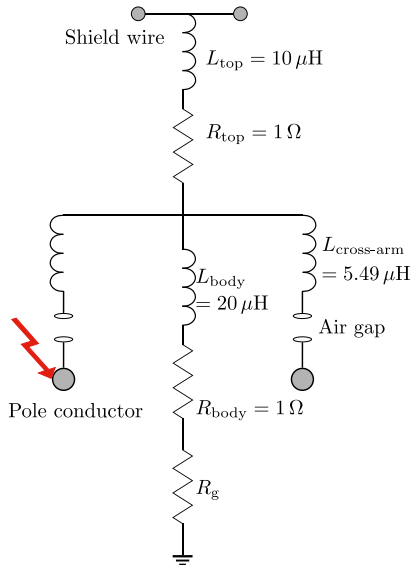


Fig. 6. Representation of the stroke tower in EMT software during a direct strike

As detailed in Section V, the typical observation duration is less than 1.5 ms. Sometimes, lightning may comprise subsequent strikes. As the typical interstroke time interval is about 60 ms [5] in average, the proposed algorithm would handle the different strikes as independent events. The evaluation time interval to locate the first strike would end before the arrival of the subsequent strikes. Lightning comprising only a single strike are thus considered.

B. Evaluation of the accuracy of the model

Results are first presented to assess the accuracy of the TW model compared to EMT simulations. The performance of the estimator of the voltage at the event location (see Section III-B) is evaluated as well as the model of the subsequent TW when fed by voltage measurements obtained at the observation location (Section III-C).

Consider a direct lightning strike occurring on the negative pole of line L_{12}, at a distance $d_\mathrm{f} = 70\,$km. Two different sets of lightning parameters have been tested, corresponding to two different events. The first lightning current has a peak value $i_\mathrm{max} = 10\,$kA and half-time value $t_\mathrm{h} = 75\,\mu$s while the stroke tower has a ground resistance of $R_\mathrm{g} = 50\,\Omega$. The second lightning current has a peak value $i_\mathrm{max} = 5\,$kA and half-time value $t_\mathrm{h} = 21\,\mu$s while the stroke tower has a ground resistance of $R_\mathrm{g} = 20\,\Omega$. In the first event, the over-voltage at the stroke tower causes a flashover of the insulator string, inducing a fault. In the second event, the magnitude of the lightning current is smaller and the insulation withstands the overvoltage, causing the lightning current to propagate as a disturbance throughout the stroke conductor without causing any fault.

To evaluate the accuracy of the estimator of the voltage at the event location alone, the fault distance d_f^* is assumed to be known. The estimated voltage at event location is compared to the output of EMT simulation for both events in Figure 7. In both cases, the overvoltage on the negative pole is well

represented as well as the tail of the waveform. The knowledge of the type of event (pole-to-ground fault or disturbance) is not required to obtain the initial voltage at the event location. This is expected as the reflection coefficient at the fault location is not involved in (4). For the model of the subsequent waves, the knowledge (or a guess) of the type of event is important, since it impacts the value of the reflection coefficient.

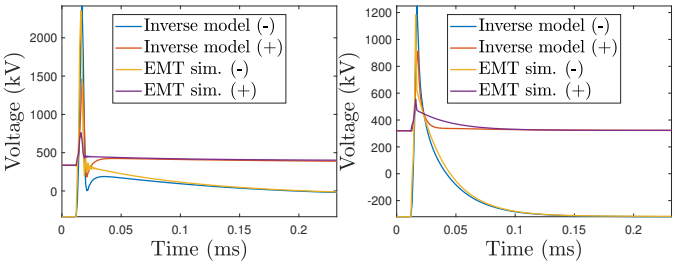


Fig. 7. Direct lightning strike (CIGRE model) on line L_{12} leading to a negative pole-to-ground fault (left) and to a disturbance (right): Voltage at the event location obtained at the output of the inverse model (8) and provided by the EMT simulation.

For the first event (resulting in a negative pole-to-ground fault), the output of the model of the subsequent waves arriving at Station 1 is depicted in Figure 8 and compared with EMT simulation. The first sample of the measured signal (provided by the EMT simulation)are used to determine the waveform at fault location as described in Section III-B. They are fed to the subsequent TW propagation model to get an estimate of the voltage at the OL after time $t = 0.7\,\text{s}$. The TW arriving at $t = 0.7\,\text{ms}$ is due to the reflection at the fault location while the surge at about $1.2\,\text{ms}$ is due to the reflection of the first wave at Station 3.

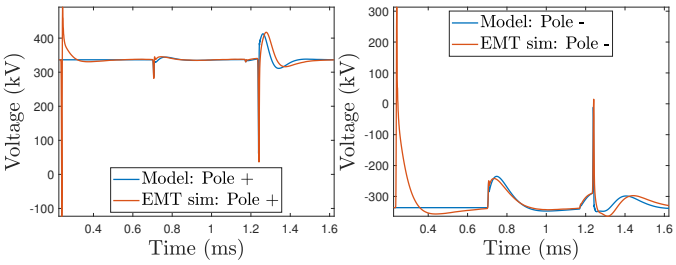


Fig. 8. Direct lightning strike (CIGRE model) on line L_{12} leading to a negative pole-to-ground fault: Voltage at Station 1 obtained from the output of the subsequent wave model (15) introduced in Section III-C and provided by the EMT simulation.

The model for subsequent waves (15) accurately represents the voltage at the station after the occurrence of a simulated lightning strike with unknown parameters on a transmission line, assuming the type of fault and the fault parameters are known. The characteristic of the lightning strike (29), however, is not required to simulate the subsequent waves.

C. Fault location

The obtained model for the fault voltage at the end of a line is applied to the location of the fault using the algorithm introduced in Section IV. Illustrative examples are first provided considering the CIGRE lightning current model in Section V-C1, and then using experimental lightning current measurements, in Section V-C2. All events occur in line L_{14}, monitored by the event location algorithm placed at Station 1. The set of modeled edges $\mathcal{E}^{\text{m}}$ introduced in Section IV-B consists of the lines L_{14} and L_{13}, which gives a minimum length of the evaluation time interval (28) of $\tau = 0.5\,\text{ms}$. Though L_{14} is the monitored line, L_{13} is also included in the modeled edges to provide a sufficiently long evaluation time interval to the algorithm.

1) CIGRE lightning waveforms (29) *:* The first event consists of a negative pole-to-ground fault occurring on line L_{14} at a distance $d_{\text{f}}^* = 50\,\text{km}$ from Station 1. The parameters of the lightning current are $i_{\max} = 10\,\text{kA}$, and $t_{\text{h}} = 34\,\mu\text{s}$. Moreover, the tower grounding resistance is taken as $R_{\text{g}} = 65\,\Omega$. Though the grounding resistance is unknown, the estimate of the fault resistance is set to $R_{\text{f}} = 10\,\Omega$ and is not estimated along with the fault distance.

Figure 9 shows the cost function (19) as a function of the candidate event distance. One observes the global minimum at $d_{\text{f}} = d_{\text{f}}^* = 50\,\text{km}$, as well as several local minima. This justifies the need for multiple initialization points. The fault location algorithm considers a set $\mathcal{D} = \{5, 25, 45, 65, 85, 105, 125, 145, 165, 185\}$ km of $n = 10$ different initial estimates of the event location. The algorithm converges towards the global minimum for the initial estimates $d_{\text{f}} = 25\,\text{km}$ and $d_{\text{f}} = 45\,\text{km}$ and the estimated fault distance is $\widehat{d}_{\text{f}} = 50\,\text{km}$. The model of the subsequent waves is compared with the EMT simulation in Figure 9 (right). The mismatch between the EMT simulated and TW propagation model output voltages of about $20\,\text{kV}$ is related to the value of the fault resistance $R_{\text{f}} = 10\,\Omega$ employed in the model that does not account for the value of the tower grounding $R_{\text{g}} = 65\,\Omega$. This results illustrates the robustness of the approach to an uncertainty in the knowledge of R_{f}. This is due to the small sensitivity of the cost function to the fault resistance, as shown in Figure 5.

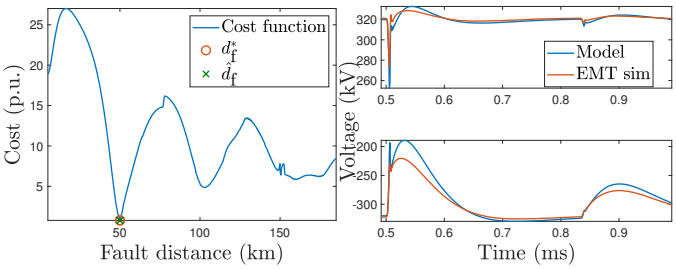


Fig. 9. Direct lightning strike (CIGRE model) on the negative pole on line L_{14} at $d_{\text{f}}^* = 50\,\text{km}$: value of the cost function as a function of the estimated event location (left), as well as output at Station 1 of the subsequent voltage waves model (considering $\widehat{d}_{\text{f}} = 50\,\text{km}$) and of the EMT simulation, (top-right: positive pole; bottom-right: negative pole).

A second event is considered with a lightning strike occurring on the negative pole of line L_{14} at a distance $d_{\text{f}}^* = 160\,\text{km}$ from Station 1. The parameters of the lightning current are $i_{\max} = 33\,\text{kA}$, $t_{\text{h}} = 16\,\mu\text{s}$, and $R_{\text{g}} = 80\,\Omega$. Figure 10 (left) shows the values of the cost function (19) as a function of the estimated event distance. A global minimum is present at $d_{\text{f}}^* = 160\,\text{km}$. The iterative algorithm provides an estimated

fault distance $\widehat{d}_{\rm f} = 159.2\,$km, slightly different from $d_{\rm f}^* = 160\,$km, as the algorithm gets trapped in a local minimum close to $d_{\rm f}^*$. As for the events presented in Section V-B, the type of event (*e.g.* pole-to-ground fault or disturbance) is needed in the model of the subsequent waves, for instance through the reflection coefficients in (12). The type of event can be determined for instance by comparing the voltage of the two poles. Alternatively, several event identification algorithms can be run in parallel with different assumptions on the type of fault. Figure 10 (right) shows the subsequent waves, obtained from the propagation model and the estimate at event location and the wave obtained from the EMT simulation.

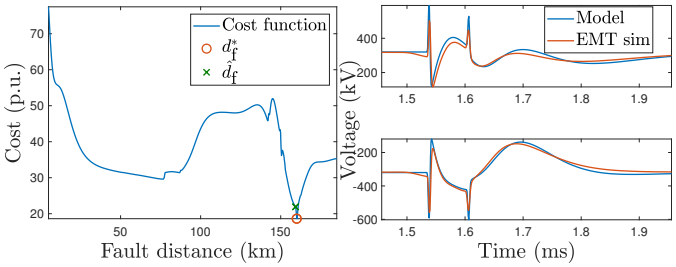


Fig. 10. Direct strike (CIGRE model) on the negative pole on line L_{14} at $d_{\rm f}^* = 160\,$km: Cost function as a function of the estimated event location (left) as well as output at Station 1 of the subsequent voltage waves model (considering $\widehat{d}_{\rm f} = 159.2\,$km) and of the EMT simulation (right); the representation is over the evaluation time interval.

2) Experimental lightning current waveforms : A first event is considered where a lightning strikes the positive pole of line L_{14} at a distance $d_{\rm f} = 120\,$km from Station 1. To evaluate the behavior of the fault location algorithm, the considered OL is relay R_{14}. As can be seen on the initial voltage waveform at the fault location in Figure 11 (right), the flashover only occurs about $0.1\,$ms after the beginning of the strike, which differs significantly from the behavior observed with CIGRE lightning waveforms. The fault location algorithm is applied with the same tuning parameters as in Section V-C1. The value of the cost function (18) is represented in Figure 11 (left) as a function of the estimated fault distance. A global minimum is observed at $d_{\rm f} = d_{\rm f}^*$. There are again local minima as for the CIGRE waveform of Section V-C1. Moreover, the attraction basin of the cost function around the global minimum is reduced, compared to that of Figure 9. Nevertheless, the estimated fault distance is $\widehat{d}_{\rm f} = 119.7\,$km, only $300\,$mfrom $d_{\rm f}^* = 120\,$km. The modeled voltage at Station 1 is compared with the EMT simulation in Figure 12. The estimated voltage at fault location is compared with the EMT simulation in Figure 11 (right). Despite the rather erratic waveform, the model fits relatively well the voltage at fault location.

A second event is considered with a different lightning current striking the negative pole of line L_{14} at a distance $d_{\rm f}^* = 75\,$km from Station 1, and a grounding resistance of $R_{\rm g} = 60\,\Omega$. The voltage at the fault location is depicted in Figure 13 (right) and compared to the output of the inverse model (8). The flashover of the positive pole occurs about $0.08\,$ms after the inception of the strike. The cost function (19) as a function of the estimated event location is depicted in Figure 13 (left), showing a global minimum at $d_{\rm f} = d_{\rm f}^* = 75\,$km.

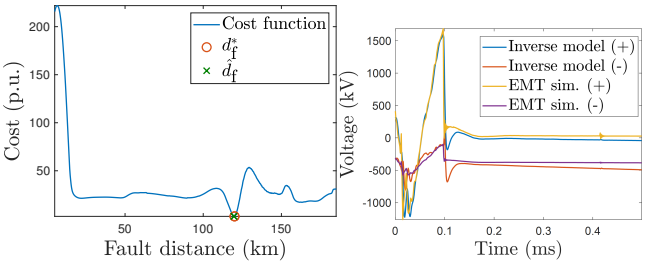


Fig. 11. Direct strike (experimental lightning current) on the positive pole on line L_{14} at $d_{\rm f}^* = 120\,$km: Cost function as a function of the estimated event location (left) as well as estimate of the voltage waveform at the fault location (considering $\widehat{d}_{\rm f} = 119.7\,$km) and of the EMT simulation involving the experimental waveform samples (right).

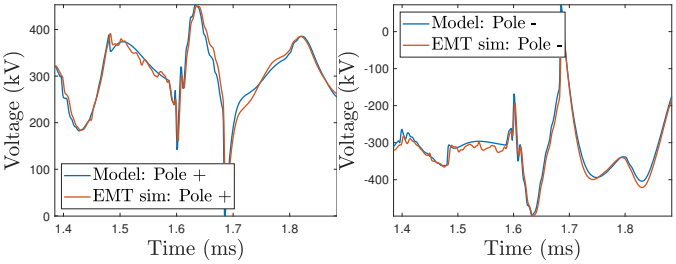


Fig. 12. Direct strike (experimental lightning current) on the positive pole on line L_{14} at $d_{\rm f}^* = 120\,$km: Subsequent voltage waves for the positive (left) and negative pole (right) at Station 1 at the output of the TW propagation model (considering $\widehat{d}_{\rm f} = 119.7\,$km) and of the EMT simulation; the representation is over the evaluation time interval.

The estimated event distance is $\widehat{d}_{\rm f} = 74.9\,$km. The modeled voltage at Station 1 is compared with the EMT simulation in Figure 12. As in the previous cases, the voltage waveform is affected by the mismatched fault resistance $R_{\rm f}$, without degrading significantly the accuracy of the estimated event location.

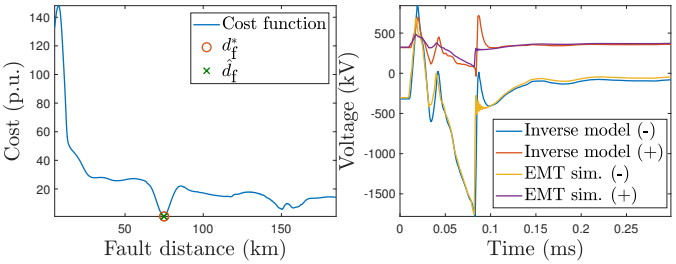


Fig. 13. Direct strike (experimental lightning current) on the positive pole on line L_{14} at $d_{\rm f}^* = 75\,$km: Left: cost function as a function of the estimated event location and estimated fault location $\widehat{d}_{\rm f}$ after a direct strike on the negative pole on line L_{14} at $d_{\rm f}^* = 75\,$km. Right: estimated fault voltage using $\widehat{d}_{\rm f} = 74.9\,$km compared with EMT simulation.

D. Sensitivity analysis

The proposed approach is based on a TW propagation model of the transmission line (2), which requires the knowledge of some parameters, such as the value of the distributed impedance and shunt admittance (1). Moreover, the value of the grounding resistance of the transmission towers is also needed in the reflection matrix $[K_{\rm e\to f}(R_{\rm f})]$, see (5).

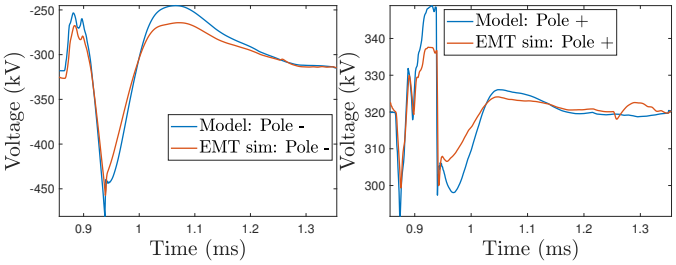


Fig. 14. Direct strike (experimental lightning current) on the positive pole on line L_{14} at $d_{\mathrm{f}}^{*} = 75\,\mathrm{km}$: Comparison of the modeled subsequent voltage waves for negative (left) and positive pole (right) at Station 1 with the EMT simulation following a lightning strike on the negative pole using the estimated fault distance $\widehat{d}_{\mathrm{f}} = 74.9\,\mathrm{km}$; the representation is over the evaluation time interval.

In what follows, the sensitivity of the algorithm to a mismatch between the parameter values used in the TW propagation model and in the EMT simulation is investigated for the distributed inductance and the ground resistance. In both cases, extensive simulations are carried out on a specific fault scenario. A lightning strike (simulated wthe CIGRE model) is simulated on line L_{13} at a distance $d_{\mathrm{f}}^{*} = 85\,\mathrm{km}$ from station 1 with a tower grounding resistance of $R_{\mathrm{g}} = 25\,\Omega$ and lightning parameters $i_{\max} = 14\,\mathrm{kA}$ and $t_{\mathrm{h}} = 25\,\mu\mathrm{s}$. The lightning strike causes a negative pole-to-ground fault.

Different values of the grounding resistance in the TW propagation model are taken in the interval $[1, 100]\,\Omega$. The distributed inductance depends on the conductor temperature, and hence both on weather and load conditions. Conductor temperature can typically vary between $-10°\mathrm{C}$ and $50°\mathrm{C}$, which amounts in inductance variation of $\pm 15\%$ around its nominal value [35]. Different values of the distributed inductance are thus considered in the interval $[0.85, 1.15]$ of its nominal value.

Figure 15 illustrates the effects of the erroneous distributed inductance (left) and grounding resistance (right) on the event location error. The estimated event location error remains less than 330 m in all cases.

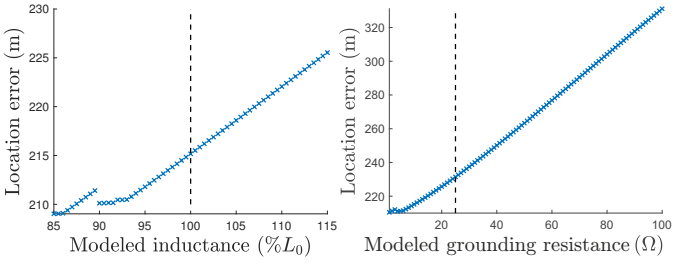


Fig. 15. Sensitivity analysis with respect to the value of the distributed inductance L (left) and tower grounding resistance R_{g} (right) considered in the TW propagation model; The dashed black lines indicate the true parameter value used in the EMT simulation.

E. Implementation issues

The proposed algorithm requires measurements to be collected during a sufficiently long observation time interval starting when an event is detected at the OL. The measurements are fed to the TW propagation model. Measured samples belonging to the evaluation time window are also compared to the TW model output. The observation time interval has thus to be such that at least two wave fronts are observed. For an event situated at the midpoint of a line of 300 km, this would need 1.5 ms of observed data.

In the event location algorithm, the most time-consuming part is the construction at each step of the transfer function $[h_{q,\mathrm{f}}(t, \mathbf{p})]$ involved in the TW propagation model (15). In the current Matlab implementation, $[h_{q,\mathrm{f}}(t, \mathbf{p})]$ is obtained in 10 ms at each iteration. Between 5 and 15 steps are required by the local optimization algorithm to converge. When 10 initial estimates of the event location are considered, the total computing time is about 1 s.

This computing time may be dramatically reduced. First, the local minimization starting from different initial estimates of the event location can be run in parallel. Moreover, an offline evaluation of $[h_{q,\mathrm{f}}(t, \mathbf{p})]$ can be performed for many candidate event locations separated, *e.g.*, by 500 or 1000 m. Since the proposed approach is robust to the fault resistance R_{f}, only the event locations have to be changed. The resulting transfer function have then to be stored, so as to be reused within the event location algorithm.Comparison with alternative approaches

The proposed method is finally compared to alternative approaches able to locate faults due to lightning on transmission lines. The performance of the different methods are borrowed from the reference publications and focus on several criteria: the maximum error in the estimated fault distance relative to the total line length, the type of lightning waveform used for performance evaluation, and the required length of the observation time interval. For the double-ended schemes, the observation time interval needed is set to, at least, twice the propagation time along the entire line, as it corresponds to the worst case propagation time for faults occurring at one extremity of the line.

The performance of the various algorithms are presented in Table III using information available in the references. The proposed approach provides the most accurate results, yet with a single-ended scheme using a limited observation time interval. Foremost, it has been evaluated with experimental lightning waveform samples, showing similar results than with standard lightning waveform models.

TABLE III
COMPARISON OF LOCATION ALGORITHMS FOR LIGHTNING-INDUCED FAULTS

Ref	Method	Ended	Observation interval	Sampling rate	Relative location error	Lightning model required
[15]	Adaptive Kalman	Double	> 0.7 ms	2 MHz	0.3 %	Yes
[19]	EMTR	Single	> 3 ms	10 Mhz	2.5 %	No
[36]	Data mining	Double	> 0.2 ms	100 Mhz	0.3 %	During training
Proposed		Single	$[0.5, 1.5]$ ms	1 MHz	0.1 %	No

VI. CONCLUSIONS

This paper addresses the problem of low-delay location of faults due to lightning strikes from single-ended measurements in HVDC grids. A combined data-driven and model-driven approach is proposed. An estimate of the fault location is

iteratively evaluated. No prior knowledge of the evolution of the voltage at the strike location is required. Data are acquired at some observation location of the HVDC grid. An initial estimate of the strike location is considered. The acquired data are fed to a model of the TW following the first wave. The subsequent waves are then compared to the measured data to update the estimate of the strike location. The performance of the proposed approach is evaluated via simulations including field measurements of faults due to bipolar lightning strikes. A typical location accuracy of 300 m is obtained considering observations performed at 1 MHz over a observation time interval of 1 ms.

REFERENCES

[1] S. A. De Almeida, R. Pestana, and F. P. Barbosa, "The main causes of incidents in the portuguese transmission system - Their characterization and how they can be used for risk assessment," in *Proc. 6th International Conference on the European Energy Market, EEM*, Leuven, 2009.

[2] E. O. Schweitzer, B. Kasztenny, M. V. Mynam, A. Guzmán, N. Fischer, and V. Skendzic, "Defining and Measuring the Performance of Line Protective Relays," *70th Annual Georgia Tech Protective Relaying Conference*, no. October 2016, 2016.

[3] ENTSO-E, "Use of traveling waves principle in protection systems and related automations," Bruxelles, Tech. Rep. April, 2021.

[4] WG B5.52, "Analysis and comparison of fault location systems in AC power networks," CIGRE, Tech. Rep. December, 2021.

[5] WG C4.407, "Lightning parameters for engineering applications," Tech. Rep. August, 2013.

[6] D. Jovcic, G. Tang, and H. Pang, "Adopting Circuit Breakers for High-Voltage dc Networks: Appropriating the Vast Advantages of dc Transmission Grids," *IEEE Power and Energy Magazine*, vol. 17, no. 3, pp. 82–93, 2019.

[7] J. Wang, B. Berggren, K. Linden, J. Pan, and R. Nuqui, "Multi-Terminal DC System line Protection Requirement and High Speed Protection Solutions," in *Proc. CIGRE Symposium*, vol. 2630, Cape Town, South Africa, 2015, pp. 1–9.

[8] A. Guzman, B. Kasztenny, Y. Tong, and M. V. Mynam, "Accurate and economical traveling-wave fault locating without communications," *71st Annual Conference for Protective Relay Engineers, CPRE 2018*, vol. 2018-Janua, pp. 1–18, 2018.

[9] N. Johannesson and S. Norrga, "Estimation of travelling wave arrival time in longitudinal differential protections for multi-terminal HVDC systems," in *Proc. 14th International Conference on Developments in Power System Protection (DPSP)*, Belfast, 2018, pp. 1007–1011.

[10] N. Manduley, S. Pack, A. Xémard, S. Touré, B. Raison, and S. Poullain, "Techniques for the improvement of the lightning back-flashover performance of double circuit HVDC lines," *Electric Power Systems Research*, vol. 200, no. June, 2021.

[11] CIGRE WG01 SC33, "Guide to procedures for estimating the lightning performance of transmission lines," Tech. Rep. October, 1991.

[12] IEC, "High-voltage test techniques - Part 1: General definitions and test requirements," International Electrotechnical Commission, Tech. Rep., 2010.

[13] G. Zou, H. Gao, W. Su, and D. Wang, "Identification of Lightning Stroke and Fault in the Travelling Wave Protection," *Journal of Electromagnetic Analysis and Applications*, vol. 01, no. 01, pp. 31–35, 2009.

[14] L. Tang, X. Dong, S. Shi, and Y. Qiu, "A high-speed protection scheme for the DC transmission line of a MMC-HVDC grid," *Electric Power Systems Research*, vol. 168, no. October 2018, pp. 81–91, 2019. [Online]. Available: https://doi.org/10.1016/j.epsr.2018.11.008

[15] Y. Xi, Y. Cui, X. Tang, Z. Li, and X. Zeng, "Fault Location of Lightning Strikes Using Residual Analysis Based on an Adaptive Kalman Filter," *IEEE Access*, vol. 7, 2019.

[16] N. Sapountzoglou, J. Lago, and B. Raison, "Fault diagnosis in low voltage smart distribution grids using gradient boosting trees," *Electric Power Systems Research*, vol. 182, no. September 2019, p. 106254, 2020. [Online]. Available: https://doi.org/10.1016/j.epsr.2020.106254

[17] V. Rizeakos, A. Bachoumis, N. Andriopoulos, M. Birbas, and A. Birbas, "Deep learning-based application for fault location identification and type classification in active distribution grids," *Applied Energy*, vol. 338, no. July 2022, p. 120932, 2023. [Online]. Available: https://doi.org/10.1016/j.apenergy.2023.120932

[18] R. Bertho, V. A. Lacerda, R. M. Monaro, J. C. Vieira, and D. V. Coury, "Selective nonunit protection technique for multiterminal VSC-HVDC grids," *IEEE Transactions on Power Delivery*, vol. 33, no. 5, pp. 2106–2114, 2018.

[19] R. Razzaghi, M. Scatena, K. Sheshyekani, M. Paolone, F. Rachidi, and G. Antonini, "Locating lightning strikes and flashovers along overhead power transmission lines using electromagnetic time reversal," *Electric Power Systems Research*, vol. 160, pp. 282–291, 2018. [Online]. Available: https://doi.org/10.1016/j.epsr.2018.03.012

[20] A. Cozza, S. Y. He, and Y. Z. Xie, "Impact of propagation losses on fault location accuracy in full transient-based methods," *IEEE Transactions on Power Delivery*, vol. 36, no. 1, pp. 383–396, 2021.

[21] G. Wang and C. Zhuang, "A Fault Location Method Using Direct Convolution: Electromagnetic Time Reversal or Not Reversal," *IEEE Transactions on Electromagnetic Compatibility*, vol. 64, no. 4, pp. 1112–1116, 2022.

[22] R. Razzaghi, M. Paolone, F. Rachidi, J. Descloux, B. Raison, and N. Retière, "Fault location in multi-terminal HVDC networks based on Electromagnetic Time Reversal with limited time reversal window," in *Proc. Power Systems Computation Conference, PSCC*, Wroclaw, 2014, pp. 1–7.

[23] L. Wedepohl, "Application of matrix methods to the solution of travelling-wave phenomena in polyphase systems," in *Proceedings of the Institution of Electrical Engineers*, vol. 110, no. 12, 1963, p. 2200.

[24] P. Verrax, A. Bertinato, M. Kieffer, and B. Raison, "Fast fault identification in bipolar HVDC grids: a fault parameter estimation approach," *IEEE Transactions on Power Delivery*, 2021.

[25] L. Tang, X. Dong, S. Shi, and B. Wang, "Analysis of the characteristics of fault-induced travelling waves in MMC-HVDC grid," in *Proc. 14th International Conference on Developments in Power System Protection (DPSP)*, vol. 2018, 2018, pp. 1349–1353.

[26] E. Walter and L. Pronzato, *Identification of parametric models from experimental data*. Springer-Verlag London, 1997.

[27] P. Verrax, N. Alglave, A. Bertinato, M. Kieffer, and B. Raison, "Low-complexity graph-based traveling wave models for HVDC grids with hybrid transmission lines: Application to fault identification," *Electric Power Systems Research*, vol. 205, 2022.

[28] E. W. Dijkstra, "A note on Two Problems in Connexion with Graphs," *Numerische Mathematik*, vol. 1, pp. 269–271, 1959. [Online]. Available: https://doi.org/10.1145/3544585.3544600

[29] J. Mahseredjian, S. Dennetière, L. Dubé, B. Khodabakhchian, and L. Gérin-Lajoie, "On a new approach for the simulation of transients in power systems," *Electric Power Systems Research*, vol. 77, no. 11, pp. 1514–1520, 2007.

[30] C. Brantl, P. Ruffing, P. Tünnerhoff, and R. Puffer, "Impact of the HVDC system configuration on DC line protection," in *CIGRE symposium*, Aalborg, 2019. [Online]. Available: http://journal.um-surabaya.ac.id/index.php/JKM/article/view/2203

[31] M. Haeusler and S. Biswas, "HVDC Solutions for Integration of the Renewable Energy Resources Comparison of Technical Alternatives and System Configurations," in *Proc. 1st International Conference on Large-Scale Grid Integration of Renewable Energy in India*, 2017.

[32] CIGRE WG B2/B4/C1.17, "Impacts of HVDC lines in the economics of HVDC projects," CIGRE, Tech. Rep., 2009.

[33] M. Ishii, T. Kawamura, T. Kouno, E. Ohsaki, K. Murotani, and T. Higuchi, "Multistory transmission tower model for lightnging surge analysis," *IEEE Transactions on Power Delivery*, vol. 6, no. 3, pp. 1327–1335, 1991.

[34] A. Xemard, E. Sellin, R. Tarafi, A. Bertinato, and P. Verrax, "Lightning overvoltages on a DC transmission line, calculated based on measured bipolar lightning strokes," *Electric Power Systems Research*, vol. 197, no. May, p. 107331, 2021. [Online]. Available: https://doi.org/10.1016/j.epsr.2021.107331

[35] Y. Wang, Y. Mo, M. Wang, X. Zhou, L. Liang, and P. Zhang, "Impact of conductor temperature time-space variation on the power system operational state," *Energies*, vol. 11, no. 4, pp. 1–15, 2018.

[36] J. A. Morales, Z. Anane, and R. J. Cabral, "Automatic lightning stroke location on transmission lines using data mining and synchronized initial travelling," *Electric Power Systems Research*, vol. 163, pp. 547–558, 2018. [Online]. Available: https://doi.org/10.1016/j.epsr.2018.01.025

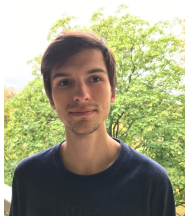
Paul Verrax was born in Sète, France, in 1995. He received his electrical engineering degree from Supélec (now CentraleSupélec), Gif-sur-Yvette, and the MSc degree in Systems, Control and Robotics from the Royal Institute of Technology, KTH, Stockholm in 2018. After working as a PhD candidate in collaboration with the Supergrid Institute and the L2S lab, he received his PhD degree in Signal and Image Processing from the University Paris-Saclay in 2021. Since then, he is a research engineer at the Supergrid Institute, Villeurbanne, France. His main area of interests are the protection, control, and design of power systems, in particular multi-terminal HVDC grids.

Michel Kieffer (M'02–SM'07) is a full professor in signal processing for communications at the University Paris-Saclay and a researcher at the Laboratoire des Signaux et Syst\'emes, Gif-sur-Yvette. From 2009 to 2016, he was part-time invited professor at the Laboratoire Traitement et Communication de l'Information, Telecom Paris, Paris.

His research interests are in signal processing and control techniques for multimedia, communications, and networking, distributed source coding, network coding, joint source-channel coding and decoding, joint source-network coding. Applications are mainly in the reliable delivery of multimedia contents over wireless channels. He is also interested in guaranteed and robust state bounding and distributed control for systems described by non-linear models in a bounded-error context.

Michel Kieffer is co-author of more than 65 contributions in journals and 140 papers published in conference proceedings. He has filed 11 patents. He is one of the co-authors of the book \emph{Applied Interval Analysis} published by Springer-Verlag in 2001 (this book was translated in Russian in 2005) and of the book \emph{Joint source-channel decoding: A crosslayer perspective with applications in video broadcasting} published by Academic Press in 2009.

He serves as associate editor of Signal Processing since 2008 and of the IEEE Transactions on Communications from 2012 to 2016. From 2011 to 2016, Michel Kieffer was junior member of the Institut Universitaire de France.



Louis Milhiet was born in Mont-Saint-Aignan, France, in 1999 and is currently completing a double degree program at the University of New South Wales (UNSW), Sydney. He will receive his engineering degree from Centrale Lille and his MSc degree in Information Technology with a specialization in Artificial Intelligence and Data Science from UNSW in 2023. During a 6-month internship at Supergrid Institute, he studied optimization methods and machine learning models for fault identification in HVDC lines. As part of his master research project, he is currently investigating the use of chromatic derivatives as a preprocessing method for speech recognition.

Bertrand Raison (M'01–SM'16) received his M.S. and Ph.D. degrees in electrical engineering from the Grenoble INP, France, in 1996 and 2000. He has joined since 2009 the Université Grenoble Alpes -G2Elab as professor. His general research interests are protection, fault detection and localization in electrical systems.