Risk Assessment of a Transmission Line Insulation Breakdown due to Lightning and Severe Weather

Tatjana Dokic[1], Payman Dehghanian[1], Po-Chen Chen[1], Mladen Kezunovic[1], Zenon Medina-Cetina[2], Jelena Stojanovic[3], Zoran Obradovic[3]

[1]Department of Electrical and Computer Engineering
Texas A&M University
College Station, TX, U.S.A.

[2]Department of Civil Engineering
Texas A&M University
College Station, TX, U.S.A.

[3]Computer and Information Sciences Department
Temple University
Philadelphia, PA, U.S.A.

Abstract

The transmission line insulation breakdown is typically assessed by performing insulator chain tests, and by conducting network modeling and simulation studies incorporating various stress conditions. This paper investigates how historical data coming from the lightning detection network and measurement stations capturing associated weather conditions can be utilized to provide a predicted assessment of risk of insulation breakdown for a given exposure and associated weather threats. The proposed analysis is enabled by the space and time correlation of the transient data recorded in the substations at the end of the lines, as well as by the assimilation of data obtained from the lightning detection network and weather stations. The proposed modeling and simulation tools are utilized to facilitate the time and space correlation analysis that leads to predication of the risk.

1. Introduction

Based on [1], 30% of blackouts in US in 2014 were weather related. In addition to weather directly causing the outage, weather conditions are one of the main factors affecting deterioration of equipment, which also may lead to failure. Lightning strikes generate overvoltages that travel along the transmission lines to substations. While they will not always cause the failure of equipment, their intensity and frequency of occurrence will affect the rate of insulation deterioration. Evaluating impact of lightning caused overvoltages on the insulators along the transmission lines is hence of utmost importance.

Georeferenced data from weather stations, lightning detection network, and utility measurements are readily available to conduct spatio-temporal analysis correlating weather threats with their corresponding impacts [2]. With such spatio-temporal framework, utility measurements coming from Intelligent Electronic Devices (IED), assets data, and weather and lightning data can be integrated to provide better decision making in outage and asset management in electric power systems [3].

Lightning studies and experiences with insulation coordination have been reported in [4-8]. For the purpose of estimating probability of a lightning strike, historical lightning data has been used in [4, 5]. Correlation of lightning data with transient measurements has been studied in literature [2, 9]. In [9], real time monitoring of transmission line transients under lightning strikes was presented, which allows for spatio-temporal correlation of lightning data and transient measurements to evaluate the impact on insulation coordination. In [2], the lightning data is correlated with traveling wave fault locator data in order to provide better accuracy and robustness of the fault location algorithm.

Correction factors for utilization of weather station data for insulation coordination have been described in [10]. In [11], statistical method for lightning-related risk analysis has been performed. An optimization procedure to determine locations of line arresters that would minimize the risk has been implemented. In [11], the weather conditions have been taken into account; however, the study has been performed based on a randomly generated data.

This paper builds on the previous study [2] by adding the risk assessment to improve transmission system asset and outage management. It is organized as follows: Section 2 provides background for the topics of insulation coordination, lightning detection network, weather stations, and Geographical Information System (GIS). In Section 3, a methodology for evaluating risk of insulation breakdown is outlined. Results of a case study are presented in Section 4, and Section 5 lists the conclusions.

1530-1605/16 $31.00 © 2016 IEEE
DOI 10.1109/HICSS.2016.311

2. Background

2.1. Typical Insulation Coordination Risk Assessment Approach

The insulators play an important role in power transmission system as the integrity of an overhead transmission line is directly governed by the electrical and mechanical performance of such equipment. According to the statistics, while the insulators account for only 5% to 8% of the direct capital cost of the transmission line, more than 70% of the line outages and up to 50% of line maintenance costs are being caused by the insulator-induced outages [12].

Insulation coordination is the study used to select insulation strength to withstand the expected stress caused by lightning and switching overvoltages. There are two approaches to tackle this problem: deterministic and probabilistic. In the former, minimum strength is set to be equal to maximum stress while in the later, the lightning flashover rate and lightning-related failures are calculated statistically and insulation strength is selected accordingly.

Insulation strength can be described using the concept of Basic Lightning Impulse Insulation Level (BIL). Statistical BIL represents a voltage level for which insulation has a 90% probability of withstand and 10% probability of failure. Standard BIL is expressed for a specific waveshape of lightning impulse and standard atmospheric conditions. A typical lightning impulse waveshape is presented in Fig. 1 [13]. In order to estimate expected stress on insulation, the failure risk is calculated statistically as (Fig. 2):

$$R = \int_0^\infty f(V) \cdot D(V) dV \tag{1}$$

where $f(V)$ is the probability of overvoltage occurrence and $D(V)$ is the probability of a disruptive discharge. Probability of overvoltage occurrence can be described with density function as follows:

$$f(V) = \frac{1}{\sigma_O \sqrt{2\pi}} e^{-\frac{(V-V_O)^2}{2\sigma^2}} \tag{2}$$

where V_O is voltage for which probability density of overvoltage occurrences has a maximum, and σ_O is a standard deviation. Probability of a disruptive discharge can be expressed with a cumulative function:

$$D(V) = \frac{1}{\sigma_D \sqrt{2\pi}} \int_{-\infty}^{V} e^{-\frac{(V-V_D)^2}{2\sigma^2}} dV \tag{3}$$

where V_D is voltage for which insulation has 50% probability of a flashover, and σ_D is a standard deviation. More details about probabilistic models of insulation flashover can be found in [11, 14-16].

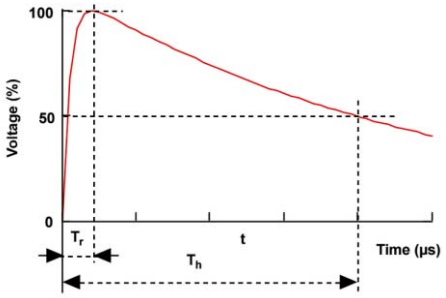

Figure 1. Standard lightning impulse, [13] (Tr = 0.1–20 µs, Th < 300 µs, where Tr is the time-to-crest value, Th is the time-to-half value)

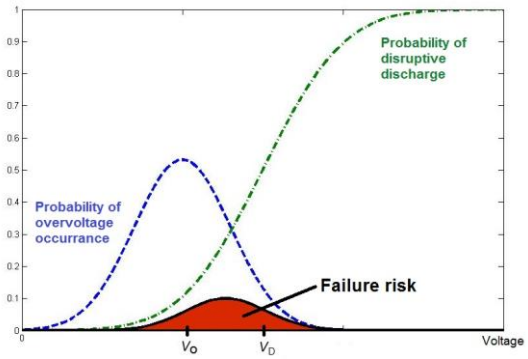

Figure 2. Risk of component failure, [14]

2.2. Weather Data and Correlation Approach

GIS and GPS together provide a framework for conducting spatial-temporal correlation between the weather threats and their corresponding impacts. As stated in [17], the spatial and temporal correlation of data plays an essential role in the process of integrating big data analytics into the electric power industry applications. Spatial correlation of data is done by integrating different data sets as layers of GIS, while GPS is used for time synchronization between events, and for synchronizing the sampling.

Two distinct categories of GIS data, spatial and attribute data can be identified [18]. Data which describes the absolute and relative context of geographic features is spatial data. For transmission towers, as an example, the exact spatial coordinates are usually accessible by the operator. In order to provide additional characteristics of spatial features, the attribute data is included. Attribute data includes characteristics that can be either quantitative or

qualitative. For example a table with the physical characteristics of a transmission tower can be described with the attribute data. Any kind of data with a spatial component can be integrated into GIS as another layer of information, [19]. As new information is gathered, these layers can be automatically updated.

In terms of spatial data representation, raster and vector data can be used. In case of vector data polygons, lines and points are used to form shapes on the map. Raster presents data as a grid where every cell is associated with one data class. Typically, different data sources will provide different data formats and types.

GPS consists of a system of 24 satellites installed by the US Department of Defense [20]. It provides location and time information for GPS receivers located on the Earth. In order to use this service, devices such as traveling wave recorders and lightning sensors are equipped with GPS receivers that supply information about longitude, latitude, and altitude, as well as a precise time tag and PPS clock.

Lightning data collected from the sensors or received from external source such as National Lightning Detection Network, [21], consist of the following information: latitude and longitude of the strike, a GPS time stamp, peak current, lightning strike polarity, and type of lightning strike (cloud-to-cloud or cloud-to-ground). Location accuracy for the ground-based lightning location system is within 0.7-1 km, [22]. GIS representation of lightning data is vector dataset including points of lightning strike locations.

Weather stations measure a set of weather parameters, such as temperature, pressure, humidity, wind speed and direction, etc. Each weather station reports measured values at the location with certain time resolution. After values from multiple stations are collected, interpolation algorithms are used to estimate the values of parameters in an area of interest (i.e. transmission lines). The results are then presented as a vector or raster maps than can be overlaid with the network georeferenced map.

Weather data is used to calculate BIL under nonstandard atmospheric conditions [23], BIL_A as:

$$BIL_A = \delta \, H_C BIL_S \tag{4}$$

where BIL_A is the BIL under nonstandard conditions, BIL_S is the standard BIL, δ is the relative air density, and H_C is the humidity correction factor. Relative air density can be calculated using:

$$\delta = \frac{PT_S}{P_S T} \tag{5}$$

where T_S and P_S are standard temperature and pressure respectively; T and P are measured temperature and pressure respectively. Humidity correction factor is equal to 1 for rainy conditions and for dry conditions can be calculated using:

$$H_C = 1 + 0.0096 \cdot \left[\frac{H}{\delta} - 11 \right] \tag{6}$$

3. Methodology

3.1. Network Modeling and Simulation

The network is modeled using the ATP version of EMTP [24]. J. Marti's frequency dependent model [25] was used for modeling of transmission line segments between towers. For representing towers, multistory transmission tower model for lightning surge analysis proposed in [26] is used. Tower model parameters are calculated using (7a-7f):

$$Z_t = 60\left(ln\frac{H}{r} - 1\right) \tag{7.a}$$

$$r = \left(\frac{r_1 h_1 + r_2 h_2 + r_3 h_3}{2H}\right) \tag{7.b}$$

$$H = \sum_{i=1}^{3} h_i \tag{7.c}$$

$$R_i = \frac{-2Z_t ln\sqrt{\gamma}}{h_1+h_2+h_3} h_i, \quad i = 1..3 \tag{7.d}$$

$$R_4 = -2Z_t ln\sqrt{\gamma} \tag{7.e}$$

$$L_i = \frac{\alpha R_i 2H}{V_t}, \quad i = 1,4 \tag{7.f}$$

where:

Z_t – tower impedance,
H – tower height,
R_i – resistances of sections,
L_i – inductances of sections,
γ – attenuation coefficient,
α – damping coefficient,
V_t – propagation velocity,
h_1, h_2, h_3 – distances between stories,
h_4 – distance between lowest story and ground,
r_i – distances between tower center and tower edge at the level of a story i.
R_f – tower footing resistance.

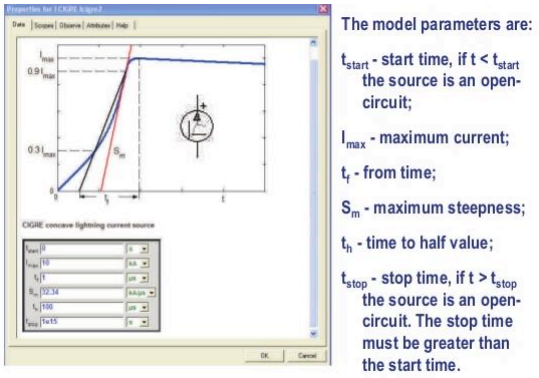

Figure 3. CIGRE concave lightning model, [27]

In the model, the lowest points of tower surge impedances are connected to the tower footing resistances.

Modeling of lightning impulse is done using current source based on CIGRE concave lightning model presented in Fig. 3 [27]. Lightning peak current is obtained from the lightning detection network. A time characteristic is synchronized with the time of a lightning strike.

The simulation process is presented in Fig. 4. First the lightning strike is selected and data from lightning detection network are sent to the ATP model in order to generate the fault. Then the simulation is run. After simulation for each node (tower) of interest, the maximum value of voltage is recorded. In parallel, the nonstandard BIL for the component is calculated using weather data. In the end, measured maximum voltage is compared to the component's nonstandard BIL and data is sent to the prediction model where it will be used as historical data for training.

3.2. Weather and Network Georeferenced Data

A complete list of data used in this study is presented in Table I. All towers and substations were geographically referenced in order to spatially correlate their locations with those of lightning strikes and weather stations.

The 100 m buffer is created around transmission lines. Only lightning strikes inside the buffer were selected and each of them was simulated as one fault scenario inside ATP. Location of lightning strike in the ATP model is determined based on lightning location in relation to the transmission network map.

Data obtained from three weather stations was used: Station NCHT2 - 8770777 - Manchester, TX, Station LYBT2 - 8770733 - Lynchburg Landing, TX, Station MGPT2 - 8770613 - Morgans Point, TX, [28]. The locations of stations are presented in Fig. 5. Time instances of interest are times of lightning strikes obtained from lightning detection network. In order to temporally correlate lightning data with data from weather stations, linear interpolation is used. For each tower and substation, weather parameters at the locations were calculated based on distance to the weather stations as:

$$P = \frac{\frac{P_1}{d_1} + \frac{P_2}{d_2} + \frac{P_3}{d_3}}{\frac{1}{d_1} + \frac{1}{d_2} + \frac{1}{d_3}} \tag{8}$$

where P is an estimated parameter value at the component location, P_i is a parameter value measured at weather station I; and d_i is a distance from the weather station i to the component.

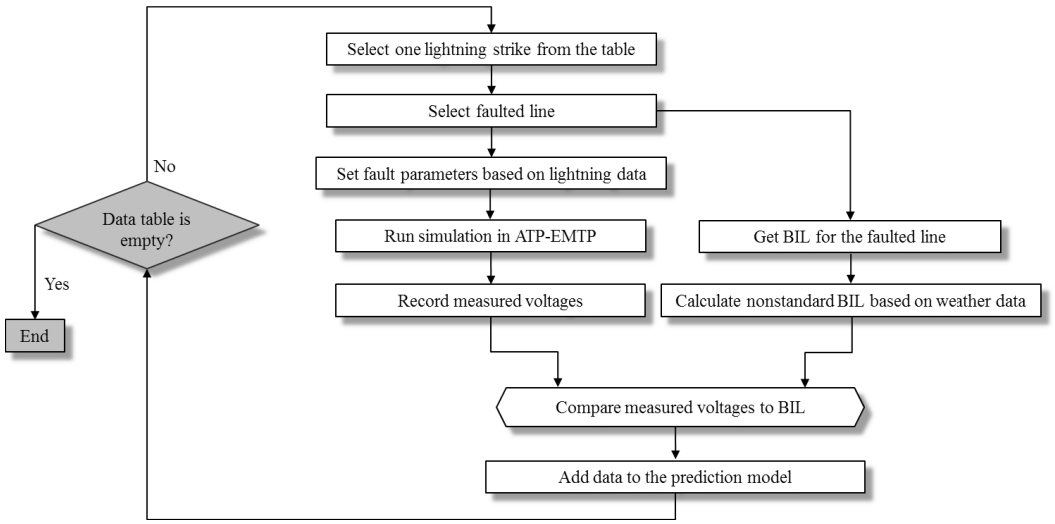

Figure 4. The simulation process

Table I. List of data

Lightning Detection Network	Weather	Insulation Studies	Geography	Traveling Wave Fault Locators
Date and time of lightning strike	Temperature	Surge impedances of towers	Location of substations	Date and time when event was recorded
Location of a strike (latitude and longitude)	Atmospheric pressure	Surge impedances of ground wire	Geographical representation of the line	Distance to the fault from the line terminals
Peak current and lightning strike polarity	Relative humidity	Footing resistance	Location of towers	Transient signals recorded at the line terminals
Type of lightning strike (cloud to cloud or cloud to ground)	Precipitation	Components BIL (Basic Lightning Impulse Insulation Level)	Location of surge arresters	

3.3. Risk Framework

The risk assessment framework used for this research is defined as follows:

$$R = P[T] \cdot P[C|T] \cdot u(C) \tag{9}$$

where R is the *State of Risk* for the system (or component), T is the *Threat* intensity (i.e. lightning peak current), *Hazard* $P[T]$ is a probability of a lightning strike with intensity T, $P[C|T]$ is the *Vulnerability* or probability of an insulation total failure if lightning strike with intensity T occurred, and the *Worth of Loss*, $u(C)$, is an estimate of financial losses in case of insulation total failure. While typical risk assessment approach described in (1) does capture the hazard and vulnerability part, it does not take into account economical losses due to insulation breakdown and does not enable application of prediction algorithms to the parts of risk assessment model.

The proposed risk measure can be defined as a stochastic process referenced in time and space as follows [29]:

$$R(X,t) = P[T(X,t)] \cdot P[C(X,t)|T(X,t)] \cdot u(C(X,t)) \tag{10}$$

where X represents the spatial parameter (longitude and latitude) and t represents the time parameter obtained using GPS. As an example, the impact of lightning is associated with certain time and location. The impact that lightning will have on a component depends on the component's distance from the lightning strike. With the stochastic risk maps, the early warning system can be spatially and temporally mapped accordingly.

Figure 5. Location of three weather stations

Fig. 6 shows the relationship model for risk assessment. *Lightning* data are indicating the probability of a *lightning strike* that is impacting probability of a *backflashover*. Probability of a *backflashover* is also under impact of weather conditions (*temperature*, *pressure*, *humidity* and *precipitation*). If there was a *backflashover*, the probability of a *component total failure* (situation where insulation is significantly damaged and needs to be replaced) is examined. Not every flashover will cause insulation total failure so probability of a flashover and probability of insulation total failure are expressed separately and then combined within overall risk framework. Due to *component failure*, some *losses* are expected to be imposed. The chain of events will impact the final *risk* as discussed in [29]. The three components of risk analysis (hazard, vulnerability and worth of loss) can be identified in Fig. 6 and are explained in the following sections.

3.3.1. Hazard

Probability of a lightning strike is estimated based on historical lightning data in the radius around the affected components. Historical data for a period of 10 years were used. For each node, the lightning frequency is calculated as:

$$LD_i = \frac{L_A}{L_T} \tag{11}$$

where L_A is the number of lightning strikes in the area with radius of 100 m around the node and L_T is the number of lightning strikes in the total area of the network. Fig. 7 shows the final lightning frequency map for the transmission network under study from which the value of Hazard is selected for each component of transmission network.

3.3.2. Vulnerability

In order to estimate a new *BIL* as time progresses (BIL_{new}), data described above are represented here in

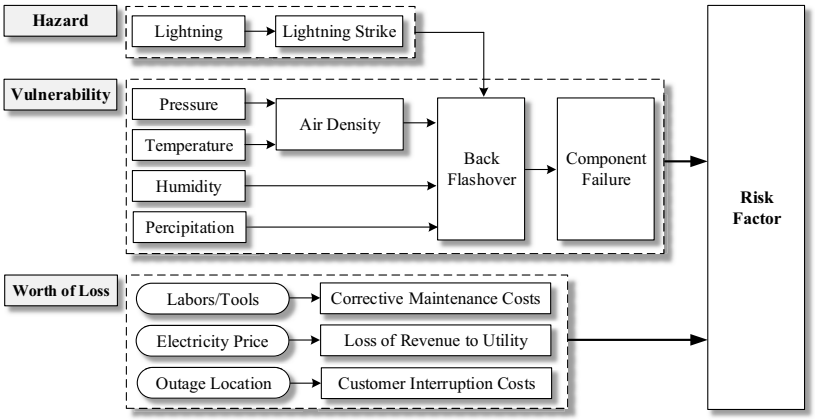

Figure 6. Risk analysis model

form of a power system network where each node represents a substation or a tower and links between nodes are calculated using impedance matrix as illustrated in Fig. 8. For each node in the graph, there are several input attribute values (x): temperature, atmospheric pressure, relative humidity, precipitation from the weather stations; Peak current and lightning strike polarity and the values of components *BIL* (Basic Lightning Impulse Insulation Level) prior to lightning strike (BIL_{old}). The output of interest (y) is the *BIL* after occurrence of lightning strike (BIL_{new}).

BIL_{new} in our experiments is predicted using Gaussian Conditional Random Fields (GCRF) based on structured regression [30, 31]. The model captures both the network structure of variables of interest (y) and attribute values of the nodes (x). It is a model based on a general graph structure and can represent the structure as a function of time, space, or any other user-defined structure. It models the structured regression problem as estimation of a joint continuous distribution over all nodes:

$$P(y|x) = \frac{1}{Z} exp\left(-\sum_{i=1}^{N}\sum_{k=1}^{K} \alpha_k \left(y_i - R_k(x)\right)^2 - \right.$$
$$\left. -\sum_{i,j}\sum_{l=1}^{L} \beta_l e_{ij}^{(l)} S_{ij}^{(l)}(x)\left(y_i - y_j\right)^2 \right) \quad (12)$$

where the dependence of target variables (y) on input measurements (x) based on k unstructured predictors $R_1, \ldots, R_k$ is modeled by the "association potential" (the first double sum of the exponent in the previous equation). The structure between outputs based on multiple layers of node inter-dependence is modeled by the "interaction potential" (the second double sum of the exponent in the equation). With such feature functions, the distribution can be expressed in Gaussian form that makes inference and learning of the model more feasible. The inference problem is then formulated as the mean of the Gaussian distribution that maximizes $P(y|x)$. Learning the parameters ($\alpha_1, \ldots, \alpha_k; \beta_1, \ldots, \beta_l$) is done by convex optimization of the log likelihood.

Outputs in terms of predictions of BIL_{new} variable are then used to calculate the probability of a flashover in case of a lightning strike $P[F|T]$.

The next step is calculating the probability of an insulation total failure in case there was a flashover with a cumulative function:

$$\mathrm{P[C|F]} = \frac{1}{\sigma_F\sqrt{2\pi}} \int_{-\infty}^{V} e^{-\frac{(V-V_F)^2}{2\sigma^2}} dV \quad (13)$$

where V_F is voltage for which insulation has 50% to exhibit total failure, and σ_F is a standard deviation.

Vulnerability is calculated as a combination of

Figure 7. Lightning frequency map

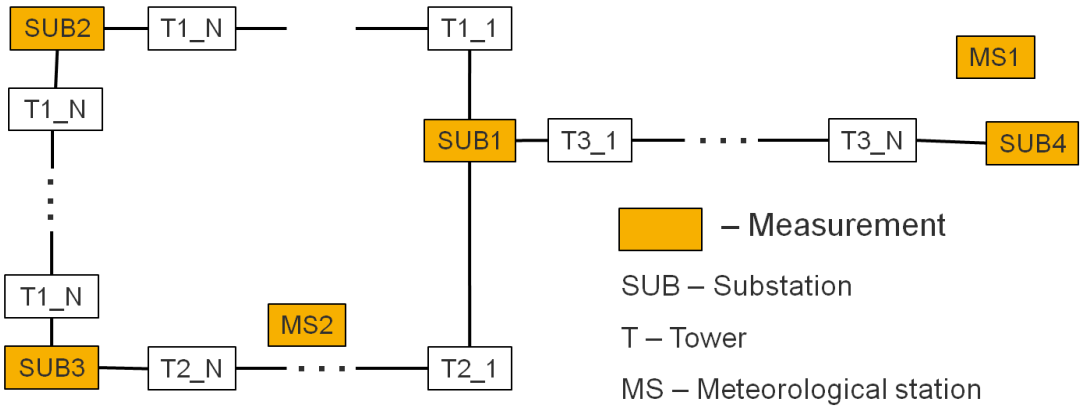

Figure 8. Illustration of a network data X = (Lightning Current, Temperature, Pressure, Humidity, Precipitation, Measured Voltage, Old BIL); Y = (New BIL); Links: impedance matrix.

probability of a flashover in case of a lightning strike $P[F|T]$ and a probability of an insulation total failure after a flashover has occurred, $P[C|F]$ as:

$$Vulnerability = P[C/T] = P[C/F] \cdot P[F|T] \tag{14}$$

As a result of Vulnerability analysis, the probability of insulation total failure is expressed in terms of lightning peak current.

3.3.3. Worth of Loss Assessment

In case where the failure of an insulator ends up in a transmission line outage, the imposed outage cost can be quantified. The total imposed costs corresponding to the failure of insulator k and accordingly outage of transmission line i at time t, $\Phi_{k,i}^{t}$, are quantified in (15) comprising of three monetary indices.

$$\Phi_{k,i}^{t} = C_{\mathrm{CM},k,i}^{t} + \sum_{\substack{d=1 \\ d\in LP}}^{D} \left(C_{\mathrm{LR},k,i}^{t} + C_{\mathrm{CIC},k,i}^{t} \right) \tag{15}$$

The first monetary term in (15) is fixed and highlights the corrective maintenance activities to fix the damaged insulator. This cost index, which in some cases can be regarded as the replacement cost of the insulator, also includes the cost of required labor, regular tools, and maintenance materials. The variable costs (second term) include the lost revenue cost imposed to the utility ($C_{\mathrm{LR},k,i}^{t}$) as well as the interruption costs imposed to the affected customers ($C_{\mathrm{CIC},k,i}^{t}$). The cost function $C_{\mathrm{LR},k,i}^{t}$ is associated with the cost imposed due to the utility's inability to sell power and hence the lost revenue when the insulator (and the associated transmission line) is out of service during the maintenance or replacement interval. This monetary term can be calculated using (16) [29].

$$C_{\mathrm{LR},k,i}^{t} = \sum_{\substack{d=1 \\ d\in \mathrm{LP}}}^{D} \left(\lambda_{d}^{t} . \mathrm{EENS}_{d,k,i}^{t} \right) \tag{16}$$

where, λ_{d}^{t} is the electricity price ($/MWh.) at load point d and $\mathrm{EENS}_{d,k,i}^{t}$ is the expected energy not supplied (MWh.) at load point d due to the failure of insulator k and outage of line i accordingly at time t. Here, the EENS index of reliability is calculated by solving the following optimization problem [30]:

$$\min_{\theta,V,P,Q} \quad \sum_{g\in \mathrm{G}} C_g(P_g^{t}) + \sum_{g\in \mathrm{G_R}} C_g^{R}(r_g^{t}) \tag{17}$$

s.t.

$$\mathbf{g_P}(\boldsymbol{\theta},\mathbf{V},\mathbf{P}) = 0 \tag{18.a}$$

$$\mathbf{g_Q}(\boldsymbol{\theta},\mathbf{V},\mathbf{Q}) = 0 \tag{18.b}$$

$$\mathbf{h_F}(\boldsymbol{\theta},\mathbf{V}) \le 0 \tag{18.c}$$

$$\mathbf{h_T}(\boldsymbol{\theta},\mathbf{V}) \le 0 \tag{18.d}$$

$$\delta_n^{\min} \le \delta_n \le \delta_n^{\max}, \quad \forall n \in \mathrm{N} \tag{18.e}$$

$$V_n^{\min} \le V_n \le V_n^{\max}, \quad \forall n \in \mathrm{N} \tag{18.f}$$

$$P_g^{\min} \le P_g^{t} \le P_g^{\max}, \quad \forall g \in \mathrm{G} \tag{18.g}$$

$$Q_g^{\min} \le Q_g^{t} \le Q_g^{\max}, \quad \forall g \in \mathrm{G} \tag{18.h}$$

$$0 \le r_g^{t} \le \min(r_g^{\max}, \Delta_g), \quad \forall g \in \mathrm{G_R} \tag{18.i}$$

$$P_g^{t} + r_g^{t} \le P_g^{\max}, \quad \forall g \in \mathrm{G_R} \tag{18.j}$$

$$\sum_{g\in G, Z_m} r_g^{t} \ge R_{Z_m}^{t}, \quad \forall m \tag{18.k}$$

$$P_d^{\min} \le P_d \le P_d^{\max} \qquad \forall d \in \mathrm{LP} \tag{18.l}$$

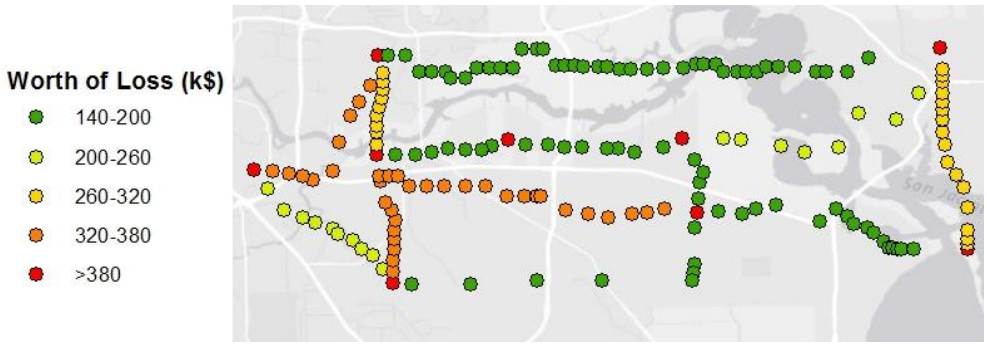

Figure 9. Worth of Loss Map

$$\mathrm{IL}_d^t = P_{d_j} - P_{d_j}^{\text{supplied}} \quad \forall d \in \mathrm{LP}, \forall j \in \mathrm{N} \tag{18.m}$$

$$\mathrm{EENS}_{d,k,i}^t = \sum_{i \in \mathrm{N}} \sum_{h \in \mathrm{H}} P_h^t . \mathrm{IL}_{d,i,k}^t . RT_{d,i,k}^t \tag{18.n}$$

The optimization problem in (17) and (18) dispatches the energy and reserve to optimize the social welfare by minimizing the total cost of energy and reserves while satisfying AC power flow equations, ancillary service requirements, and transmission and operating constraints. Constraints (18.a)-(18.b) represent the non-linear nodal active and reactive power balance equations. Network constraints (18.c)-(18.d) represent the branch flow limits for the "to" and "from" ends of each branch, respectively. Constraints (18.e)-(18.f) present the equality upper and lower limits on all bus voltage phase angles and magnitudes. Supply constraints are presented in (18.g)-(18.h) and (18.i)-(18.k) are capacity reserve constraints. Constraint (18.i) reflects the reserve for each generating unit that must be positive and limited above by a reserve offer quantity as well as the physical ramp rate (Δg) of the unit. Constraint (18.j) enforces that the total amount of energy plus reserve of the generating unit does not exceed its capacity. Constraint (18.k) is enforced to ensure that the right amount of capacity is procured according to the reserve requirements in each region. Constraint (18.l) restricts the demand at each load point between the lower and upper bounds. Constraint (18.m) calculates the interrupted load at each load point and constraint (18.n) evaluates the EENS corresponding to the outage condition.

The last variable term of the cost function in (15) reflects the customer interruption costs due to the failure of insulator k and corresponding outage of transmission line i at time t which can be calculated through (19). As it can be seen, $C_{\mathrm{CIC},k,i}^t$ is a function of the EENS index and the value of lost load (VOLL_d) which is governed by various load types being affected at each load point. The value of lost load ($/MWh.) is commonly far higher than the electricity price as obtained through customer surveys [31].

$$C_{\mathrm{CIC},k,i}^t = \sum_{\substack{d=1 \\ d \in LP}}^{D} \left(\mathrm{VOLL}_d . EENS_{d,k,i}^t \right) \tag{19}$$

The cost function in (15), which is actually the failure consequence of an insulator, can be calculated for each insulator failure in the network making it possible to differentiate the impact of different outages on the system overall economic performance. The worth of loss map is presented in Fig. 9.

4. Results

4.1. Studied Network

The network segment contains 170 locations of interest (10 substations and 160 towers). Based on the geographical representation and network connectivity, the prediction graph is constructed as described in Fig. 8. All towers were modeled using model described in section 3.1.

4.2 Test Scenarios

Historical data is prepared for the period of 10 years, starting from January 1st 2005, and ending with December 31st 2014. 1000 lightning strikes were assumed in the area of interest for the period of 10 years and used for Hazard calculation in the proposed risk framework. Out of 1000 strikes 100 strikes caused a flashover and were considered as an input data for prediction of vulnerability. For each instance of lightning strike, the weather parameters were obtained as described in section 3.2. The separate weather parameters were calculated for each

(a) (b) (c)

Figure 10. Total risk calculated on (a) January 1st 2009; (b) December 31st 2014; (c) January 5th 2015 (prediction after the next lightning strike)

component location. An example of weather data table is presented in Table II.

Before the first lightning strike, all components are assumed to have a BIL provided from the manufacturer. In each time step, new BIL is calculated based on the old BIL and collected weather data for each lightning strike.

4.3 Study Results

For each network component, risk value was calculated, assigned, and presented on a map as shown in Fig. 10. The value of risk is presented as a percentage, where 100% is assigned to the component with highest risk for the future lightning-caused failures. In part (a) of Fig. 10, the risk map on January 1st 2009 is presented, while in part (b), the risk map after the last recorded event is presented. After several years of lightning impact, some of the zones that have experienced high rates of lightning activities have an increased value of risk. It is of utmost importance to observe the probability of future lightning strikes for such vulnerable zones.

With the use of weather forecast, the prediction of future Risk values can be accomplished. In Fig. 10 (c), the prediction for the next time step is demonstrated. For the time step of interest, the lightning location is predicted to be close to the line 11 (marked with red box in Fig. 10 (c). Thus, risk values assigned to the line 11 will have the highest change compared to that of the previous step. The highest risk change on line 11 happens for node 72 with changed from 22.8% to 43.5%. The Mean Squared Error (MSE) of prediction of GCRF algorithm on all 170 test nodes is 0.0637 ± 0.0301 volts when predicting the new value of *BIL* (BIL_{new}).

5. Conclusion

This paper introduces a new framework for predictive insulation breakdown risk assessment and maintenance plan. More specifically:

- An automated risk-based early warning system (EWS) for prediction of insulation breakdown and associated economic impacts is developed and integrated with Geographical Information System (GIS).
- Accumulated impacts of past disturbances to the lightning protection components are taken into account to assess unfolding component's vulnerability after exposure to continued weather threats. Wide range of weather conditions surrounding lightning strikes is used.
- Insulation breakdown risk is assessed by analyzing time/space correlation between historical lightning and weather data, and network measurements.
- The algorithm is capable of predicting risk in case of future lightning strikes using Gaussian Conditional Random Fields (GCRF) structured regression model. With this model components geographical configuration is taken into account for a prediction.
- The worth of loss assessment that effectively differentiates the impact of different outages on the overall system economic performance has been performed.

Acknowledgement

This work has been funded by NSF I/UCRC PSerc under the project titled "Systematic Integration of Large Data Sets for Improved Decision-Making", DARPA grant FA9550- 12-1-0406 negotiated by AFOSR, and National Science Foundation Big Data grant NSF-14476570.

References

[1] Eaton, "Power Outage Annual Report – United States Annual Report 2014," [Online], Available: http://electricalsector.eaton.com/forms/BlackoutTrackerAnnualReport

[2] M. Kezunovic, et al., "Improved Transmission Line Fault Location Using Automated Correlation of Big Data from Lightning Strikes and Fault-induced Traveling Waves," 48th Hawaii Int. Conf. Syst. Sciences (HICSS), Jan. 2015.

[3] P.-C. Chen, *et al*., "Utilizing Geographical Information System for Transmission and Distribution Outage Management," IEEE/PES Innovative Smart Grid Technologies Latin America (ISGT-LA), October 2015.

[4] I. M. Rawi, et al., "Lightning study and experience on the first 500kV transmission line arrester in Malaysia," 2014 International Conference on Lightning Protection (ICLP), Shanghai, China, 2014.

[5] W. Sones, S. M. Wong, "Overview on Transient Overvoltages and Insulation Design For a High Voltage Transmission System," High Voltage Engineering and Application (ICHVE), 2010 International Conference on, New Orleans, LA, 2010.

[6] Z. G. Datsios, et al., "Estimation of the minimum shielding failure current causing flashover in overhead lines of the hellenic transmission system through ATP-EMTP simulations." 2012 International Conference on Lightning Protection (ICLP), Vienna, Austria, 2012.

[7] S. T. Mobarakei, T. Sami, B. Porkar, "Back flashover phenomenon analysis in power transmission substation for insulation coordination," 11th Int. Conf. on Environment and Electrical Engineering (EEEIC), Venice, May 2012.

[8] S. Bedoui, et al., "Analysis of lightning protection with transmission line arrester using ATP/EMTP: Case of an HV 220kV double circuit line." Universities Power Engineering Conference (UPEC), 2010 45th International. IEEE, 2010.

[9] T. Sadovic, et al., "Expert System for Transmission Line Lightning Performance Determination", CIGRE Int. Colloq. on Power Quality and Lightning, Sarajevo, Jun. 2012.

[10] Zhang, J., et al. "Application of hourly meteorological records to atmospheric correction factors in insulation coordination under switching impulse voltage." High Voltage Engineering and Application, Int. Conf. on. IEEE, 2008.

[11] R. Shariatinasab, et al., "Probabilistic evaluation of optimal location of surge arresters on EHV and UHV networks due to switching and lightning surges." Power Delivery, IEEE Transactions on, vol. 24, no. 4, pp. 1903-1911, 2009.

[12] R. S. Gorur, et al., "Utilities Share Their Insulator Field Experience," T&D World Magazine, Apr. 2005, [Online] Available: http://tdworld.com/overhead-transmission/utilities-share-their-insulator-field-experience

[13] IEEE Power and Energy Society, "IEEE Std. C62.82.1-2010 - IEEE Standard for Insulation Coordination –Definitions, Principles, and Rules," IEEE Standards Association, 2010.

[14] Á. L Orille-Fernández, N. Khalil., and S. B. Rodríguez, "Failure risk prediction using artificial neural networks for lightning surge protection of underground MV cables," IEEE Trans. Power Del., Vol. 21, No. 3, pp. 1278-1282, 2006.

[15] A. P. Sakis Meliopoulos, „Lightning and Overvoltage Protection," In: D. G. Fink, H. W. Beaty, editors. Standard handbook of electrical engineering. 15 th ed. New York: McGraw-Hill; 2007.

[16] IEEE Standards, "International Standard IEC 62539 – IEEE 930: Guide for the statistical analysis of electrical insulation breakdown data," 2007.

[17] M. Kezunovic, L. Xie, and S. Grijalva, "The role of big data in improving power system operation and protection," 2013 IREP Symp. Bulk Power Syst. Dynamics and Control - IX Optimization, Security and Control of the Emerging Power Grid (IREP), pp. 1-9, Aug. 2013.

[18] David J. Buckey, Bgis Introduction to GIS, 11/10/2013. [Online]. Available: http://bgis.sanbi.org/gis-primer/

[19] ArcGIS, Esri. [Online] Available: https://www.arcgis.com

[20] GARMIN, "What is GPS?" [Online] Available: http://www8.garmin.com/aboutGPS/

[21] Vaisala Inc., "Thunderstorm and Lightning Detection Systems," [Online] Available: http://www.vaisala.com/en/products/thunderstormandlightningdetectionsystems/Pages/default.aspx

[22] U. Finke, et al., "Lightning Detection and Location from Geostationary Satellite Observations," Institut fur Meteorologie und Klimatologie, University Hannover.

[23] A. R. Hileman, "Insulation Coordination for Power Systems," CRC Taylor and Francis Group, LLC, 1999.

[24] Alternative Transients Program, ATP-EMTP, 2010. [Online]. Available: http://www.emtp.org

[25] J. R. Marti "Accurate modeling of frequency-dependent transmission lines in electromagnetic transient simulations," *IEEE Trans. Power Apparatus and Syst.*, vol. PAS-101, no. 1, pp. 147-157, Jan. 1982.

[26] M. Ishii, et al. "Multistory transmission tower model for lightning surge analysis." *Power Delivery, IEEE Transactions on* vol. 6, no. 3, pp. 1327-1335, 1991.

[27] I. Uglesic, "Modeling of Transmission Line and Substation for Insulation Coordination Studies," Simulation & Analysis of Power Syst. Transients with EMTP-RV, Dubrovnik, Apr. 2009.

[28] National Oceanic and Atmospheric Administration's national Data Buoy Center, "Station MGPT2 – Historical Data," [Online] Available: http://www.ndbc.noaa.gov/station_history.php?station=mgpt2

[29] Z. Medina-Cetina and F. Nadim F, "Stochastic Design of an Early Warning System", Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards, vol. 2, no. 4, pp. 223 – 236, 2008.

[30] J. Stojanovic, et al,. "Semi-supervised learning for structured regression on partially observed attributed graphs," SIAM International Conference on Data Mining (SDM 2015) Vancouver, Canada, April 30 - May 02, 2015, (in press).

[31] K. Ristovski, et al., "Continuous Conditional Random Fields for Efficient Regression in Large Fully Connected Graphs," Proc. The Twenty-Seventh AAAI Conference on Artificial Intelligence (AAAI-13), Bellevue, Washington, July 2013.

[32] P. Dehghanian, et al., "A Comprehensive Scheme for Reliability Centered Maintenance Implementation in Power Distribution Systems- Part I: Methodology", IEEE Trans. on Power Del., vol.28, no.2, pp.761-770, April 2013.

[33] W. Li, *Risk assessment of power systems: models, methods, and applications*, John Wiley, New York, 2005.

[34] R. Billinton and R. N. Allan, *Reliability Evaluation of Engineering Systems: Concepts and Techniques*, 2nd ed. New York: Plenum, 1992.