Predicting Impact of Weather Caused Blackouts on Electricity Customers Based on Risk Assessment

Qin Yan, *Student Member, IEEE*, Tatjana Dokic, *Student Member, IEEE*, Mladen Kezunovic, *Fellow, IEEE*
Department of Electrical and Computer Engineering
Texas A&M University
College Station, TX, U.S.A
upforever_yan@tamu.edu; tatjana.djokic@tamu.edu; kezunov@ece.tamu.edu

***Abstract*—This study evaluates the outage probability and electricity customer cost under the potential weather caused power blackouts. Risk assessment of the weather impact on customers is implemented and visualized in ArcGIS map. The methodology correlates the historical large power outage events with the corresponding weather condition of the time, uses weather forecast data to assess the risk for customers, and compares the results in different predicted weather conditions.**

***Keywords*—blackout; customer impact; data analysis; geographical information system; meteorology; risk assessment**

NOMENCLATURE

CAIDI	Customer Average Interruption Duration Index
CDF	Customer Damage Function
CIC	Customer Interruption Cost
CS	Customer Surveys
ECOST	Expected Interruption Cost Index
EENS	Expected Energy Not Supplied
EO/C	Ratio of Economic Output to Energy Consumption
GEFS	Global Ensemble Forecast System
GFS	Global Forecast System
GIS	Geographical Information System
NAM	North American Mesoscale
NDFD	National Digital Forecast Database
SAIDI	System Average Interruption Duration Index
SAIFI	System Average Interruption Frequency Index
SVM	Support Vector Machine

I. INTRODUCTION

Severe weather conditions can cause damage to electricity delivery system and power infrastructures, leading to power interruptions to large number of customers. Studies indicate that estimated annual cost from storm-related outages to American economy is between $20-55 billion and the trend is still growing [1]. The historical blackout data from 2012 to 2014 in Texas shows 33% of the historical outage events are caused by weather/ falling trees [2]. The Vermont study [3] analyzed 933 outage events from over 20 years and stated that about 44% of the events were related to different weather conditions. It also pointed out that some of the events are triggered by "multiple factors".

To reduce the storm-related outages, possible methods can be tree-trimming schedules, reliability-centered maintenance regulations, distributed generation support, grid redundancy improvement, underground cables construction, and mutual assistance agreement [1]. All these methods are for long-term purposes. In a short-term view, if the utilities are aware of an upcoming severe weather scenario and the estimated severity of the related customer impact, preventive measures can be deployed to mitigate the customer vulnerabilities ahead.

There are two types of indices to measure reliability [4], load point reliability indices and system reliability indices which sum up all the load points. The most commonly used indices are SAIDI, CAIDI, SAIFI, and EENS. However, these indices consider only the outage parameters (restoration time, affected number of people, event frequency) and unsupplied energy which is the only customer parameter involved, and they calculate average values for a given time period, normally on either monthly or yearly basis. For the purpose of evaluating impact from a single event, CIC can be used. Brief summaries of the CIC estimation methods are provided in [5, 6]. The most common assessment is through CS [7], which are the most proper tools for individual customers. In most cases customers cannot comprehensively estimate their loss, and it takes a long time and requires prohibitively high cost to do that. Economists prefer the EO/C in terms of national economic impact [8]. The amalgamated CS [9] and mapped CS methods [10] are used to save expense and time by utilizing the existing survey results. Post-event analysis of blackout impact that already happened is used in [11]. The issue with this is the limitation of the geographic area, duration and characteristics of the analyzed outages, which were usually in urban area with high population density. Customers' economic losses can be expressed by CDF [12, 13]. ECOST, one of the reliability indices described in [4], requires the output from CDF. CIC indices are often grouped based on outage durations and customer types. Tracking weather conditions in real-time reveals quite different influence to various customer categories not considered before such as the health impact and the inconvenient transportation impact. This paper looks at variety of weather factors and evaluates the probability of blackouts and the corresponding customer financial and health impact under the forecasted weather conditions. It then develops the risk assessment of the weather-related outage impact on different customer categories.

This paper is organized as follows: Section II describes the weather data and its impacts. Section III describes the weather-driven risk model. In Section IV, data processing and integration is described. Case study is presented in Section V and Section VI summarizes the conclusions.

This effort was funded by NSF CURENT ERC and contracted through University of Tennessee under the project titled "The Use of Big Data for Outage Management in Distribution Systems".

978-1-5090-4168-8/16/$31.00

II. WEATHER IMPACTS

A. Weather Data

Two types of weather data are of interest in this paper: i) historical weather data used for training of prediction model, and ii) weather forecast data used for real time decision-making.

Historical weather data can be extracted from land-based station data, radar or satellite data, [14]. In this paper the land-based station data has been used. In case of land-based stations the interpolation techniques need to be applied in order to provide wide-area weather conditions that can be mapped using GIS software such as ESRI ArcGIS, [15].

When it comes to weather forecast data, there are several services that provide data such as NDFD [16] that provides short- and long-term weather prediction for variety of weather parameters. These services make weather predictions using one of numerical models such as GEFS, GFS, NAM, etc. [17].

Weather parameters that are of great interest for analyzing power system outages are: temperature; precipitation; humidity; wind characteristics; storm, hail, and severe wind probabilities; lightning characteristics and frequency maps, etc.

B. Weather Impact on Outages

Weather impact on outages in power systems can be classified into two groups:

- *Direct impact to utility assets*: This type of impact includes all the situations where severe weather conditions directly caused the component to fail. Examples are: lightning strikes to the utility assets, wind impact making trees or tree branches come in contact with lines, etc. In this case the outage occurs during the time of impact. When post fault analysis is performed these types of outages are marked as weather caused outages.
- *Indirect impact to utility assets*: This type of impact accrues when weather didn't directly cause the outage but instead created the situation in the network that indirectly caused the component to fail. Examples are: hot weather conditions increasing the demand thus causing the overload of the lines, exposure of assets to long term weather impacts causing component deterioration, etc. In this case the outage can occur after the time of impact. In post-fault analysis this type of outages are marked as equipment failure.

Focus of this paper is direct impact of weather to utility assets. Thus, when analyzing outage data only group of data marked as weather causing outages has been used.

C. Weather Impact on Customers Under Outage

Power supply interruption can bring extensive cost to customers in the way of spoiling the perishable materials/food, damaging equipment, causing production loss, income loss, health impact, and extra expenses [5]. Severe weather conditions can deteriorate the impact by the change of temperature (either heatwave or cold front), unusual humidity, heavy precipitation, high wind, poor visibilities, etc. For indoor customers, temperature and humidity will be the main factors. Once the weather condition is out of human's comfort range, it may lead to extensive health issues, especially for people with injury or illness. For the patients who are extremely dependent on the health appliances such as infusions or respirators which require electric power to operate, the outage may cause very serious health problems and it can be deadly even with only a few hours of outage in many cases. Appliances as simple as heaters or fans can also result in health issues when they are not working. Strokes can be triggered in a freezing house in winter or a sweltering house at peak summer without any air condition. The same applies to those requiring certain condition to survive, such as tropical plants or pets. Besides the personal health, safety issues may also arise by the increased robbery rate due to the non-operating street lights and security system. [18]

The extent of the weather impact on different customer categories varies. Residential customers may not be affected by wind or rain if they stay inside, while the thunderstorm may ruin the uncovered crops or damage the outdoor equipment for some industries. Commercial stores may lose business due to the inconvenient transportation caused by poor weather. The same impact also applies to the product transport for industries and patient transfer for health centers, etc. Property loss and business interruptions are usually considered when estimating the customer impact, because such losses are obvious and easy to be quantified. While it is difficult to evaluate the health impact on the customers, it cannot be neglected. It would be helpful to design a methodology for calculation and estimate the possible customer financial and health loss under outage caused by unusual weather factors and quantify it as an impending event risk value [19]. The results can be used by the utilities as a reference of how serious the customer may have to suffer from the upcoming event and whether certain mitigation measures are necessary to avoid the loss.

III. CUSTOMER IMPACT RISK ANALYSIS

The risk assessment framework is formulated and defined as follows, [20]:

$$R = P[T] \cdot P[C|T] \cdot u(C) \quad (1)$$

where R is the *State of Risk* of the customer impact from the upcoming weather conditions, *Hazard* P[T] is the probability of a *Threat* intensity T (i.e. a certain weather scenario that may cause power outage), *Vulnerability* P[C|T] is the probability of blackout in case hazardous weather conditions have occurred, and Worth of Loss u(C) is an estimate of customer interruption losses in case of the power outage [21]. Simplified description of Risk assessment is presented in Fig. 1.

A. Hazard

Hazard assessment is done solely based on weather data. This factor describes how likely certain weather conditions will occur in the area. The following weather parameters are

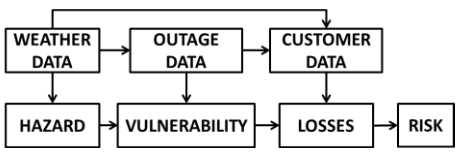


Fig. 1: Risk analysis for customer impact

observed: temperature, humidity, precipitation, and wind speed. The weather conditions have been classified in several groups based on values of the listed parameters. Then hazard probability has been assigned to each of these groups based on weather forecast data.

B. Vulnerability

Vulnerability analysis uses historical outage and weather data to predict what is the probability of a blackout in case of predicted weather. The SVM has been used for prediction. The SVM function classifies the data into two classes, while minimizing the classification error. The training data D can be described as:

$$D = \{(x_i, y_i) | x_i \in \mathbb{R}^p, y_i \in \{-1,1\}\}_{i=1}^{n} \qquad (2)$$

where labels are classified into two classes: one for situations where weather conditions did cause a blackout ($y_i = 1$) and one for situation where they did not cause a blackout ($y_i = -1$). The x_i is the *p*-dimensional real vector containing observed weather parameters, and *n* is the number of training points.

The decision boundary is:

$$g(x) = \langle a, x \rangle + \mathrm{b} \qquad (3)$$

Problem of maximizing the margin of separation between two classes (selecting *a* and *b*) can be described as quadratic programming (QP) optimization:

$$\Phi(a, \xi) = \min \left\{ \frac{1}{2} \|a\|^2 + C \sum_{i=1}^{n} \xi_i \right\} \qquad (4)$$

where ξ_i are non-negative slack variables. All samples are placed on or outside the margin as:

$$y_i \left(\langle a, x_i \rangle + b \right) \geq 1 - \xi_i, \;\; i = 1, \ldots, n$$

C. Loss

In terms of the duration and cause of the failure, power interruptions can be classified into three types: momentary, sporadic and chronic interruptions [22]. The cause of chronic interruptions is generally load shedding, while momentary ones are due to switching actions. Sporadic interruptions are mainly caused by severe weather condition which tends to last longer, affect larger number of customers, cause longer restoration time due to the affected transportation or unavailable resources and finally lead to high financial impacts.

Among the existing CIC estimation methods, most of the results give the loss indices reflecting unsupplied energy, which are classified by types of customers (residential, small/large industrial & commercial, agricultural, etc.). Some methods provide indices based on different outage duration [10]. Some of the customer filling surveys differentiated the questions for winter days and summer days, but the statistical final indices only show the average values since they are evaluating the reliability on a yearly basis [7, 23-24]. This study considers important customer categories, i.e. health care centers, schools, fire stations.

CIC is closely related to the degree of customers' dependence on electricity supply [5], thus it was formulated to be a function of outage parameters and customer features. To estimate the customer loss from a weather-related power failure, the formula needs to be improved to consider the additional financial loss and health effect caused by the weather elements, in addition to the previous economic value for EENS. The improved formula is as follows.

$$Loss^{t,HA} = CDF^{t,HA}(dr, s, d, w, cf) + EL^{t,HA}(EF, cf, dr) + HL^{t,HA}(EF, cf) \qquad (5)$$

where the estimated *Loss* at target time *t* for each hazard area *HA* includes three monetary terms: existing *CDF*, additional economic loss *EL* caused by unusual environmental features *EF*, and health loss *HL* under outage. *EF* includes all the weather elements that may influence the customer cost, including temperature *tp*, humidity *hu*, storm type *st*, wind speed *ws*, precipitation *pr*, and others *ot*. Once *t* is targeted, season *s*, time of a day *d*, day of a week *w* are given. *HA* is determined by the forecasted weather. The interruption duration *dr* can be statistically estimated based on the historical blackout event data and the forecasted weather condition. Customer features *cf* comprise customer type, number of people, time schedule at *t*, presence of interruption-sensitive equipment and back-up equipment or generators, etc [25-26].

Specifically, the three terms in (5) are defined respectively:

$$CDF^{t,HA} = \sum_{\substack{lp=1 \\ lp \epsilon HA}}^{L} CL^{lp}(dr, s, d, w, cf) \cdot EENS_t^{lp} \qquad (6)$$

$$or \qquad CDF^{t,HA} = \sum_{\substack{lp=1 \\ lp \epsilon HA}}^{L} CE_{d,s,dy,w,cc}^{lp} \qquad (7)$$

At load point *lp*, *CL* represents the interruption cost per kWh ($/kWh) while *CE* represents the interruption cost per event. *L* is the total customer number based on the size of each hazard area. *EENS* can be reliably predicted if based on smart meter measurement. According to the data availability and accuracy, either (6) or (7) can be chosen to estimate *CDF*.

$$EL^{t,HA} = \sum_{\substack{lp=1 \\ lp \epsilon HA}}^{L} \{ACL^{lp}(dr, s, d, w, cf) \cdot AE_{EF,t}^{lp} + O_{EF,t}^{lp}\} \qquad (8)$$

where *ACL* is the interruption cost for additional unsupplied energy *AE* for each *lp* and *O* represents other financial losses caused by weather, i.e. spoiled food/material, damaged equipment, production loss. These can be estimated based on the type and severity of the forecasted bad weather condition.

$$HL^{t,HA} = \sum_{\substack{lp=1 \\ lp \epsilon HA}}^{L} f_{cf}^{lp}(EF) \qquad (9)$$

where *f* is the function to calculate the health cost for each *lp* in terms of the *cf*. The function can be expressed as follows.

$$f_{cf}^{lp} = cs^{lp} \cdot ec^{ct} \cdot [\alpha^{ct} \cdot |tp - tp^s| + \beta^{ct} \cdot |hu - hu^s| + \gamma^{ct} \cdot g(st) + \delta^{ct} \cdot |ws - ws^s| + \varepsilon^{ct} \cdot |pr - pr^s| + \theta^{ct} \cdot k(ot)] \qquad (10)$$

where $\alpha^{ct}, \beta^{ct}, \gamma^{ct}, \delta^{ct}, \varepsilon^{ct}, \theta^{ct}$ are the weight coefficients to express the impact of each weather element on the customer type *ct*. *cs* is the customer size at *lp*, and *ec* is the equivalent economic cost of health impact based on the medical cost standard. The value with a superscript *s* indicate the threshold of the parameter's normal range there. Since storm type is not a

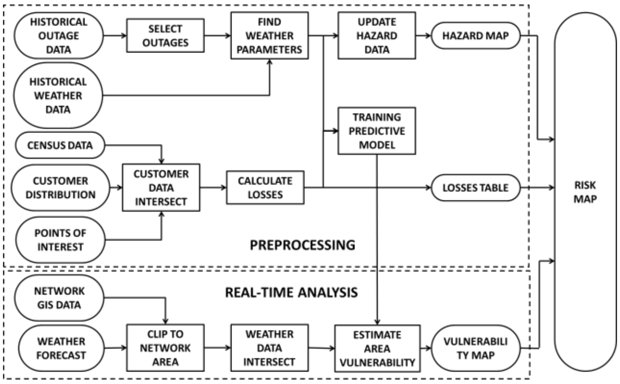


Fig. 2: Spatial integration – processing steps

quantitative value, functions *g* and *k* are used to present the corresponding impact. All the coefficients can be set by utilities in their geographic and climate circumstances.

IV. Spatio-Temporal Data Integration

In Fig. 2 processing steps for risk analysis are shown. The following data layers have been used as input data: electric network GIS data, weather data, customer type data, population count, and historical outage data.

As part of preprocessing in Fig. 2, all historical and static data are analyzed in order to provide training datasets for the prediction model. Here historical outage data is used to select time instances of interest. Then weather parameters are selected for those times. Static maps such as population count, customer distribution and points of interest are used to calculate what the associated losses for observed outages are. It is to be understood that this data is changing over time, but since the change and data availability comes once in every few years, it is considered static in this study, compared to weather data which may change every few minutes to every few hours.

As part of the real-time analysis in Fig. 2, weather forecast data obtained from NDFD is downloaded every three hours. This data is then overlaid with network data and static customer data. The risk model uses training data and weather forecast data to evaluate the risk for customers in case of weather conditions predicted by weather forecast.

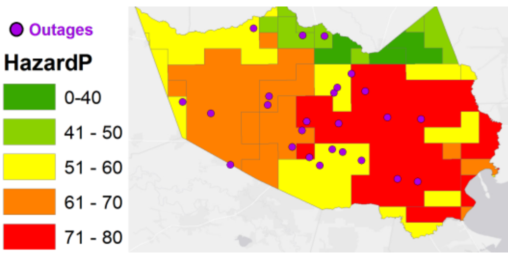


Fig. 3: The hazard distribution map

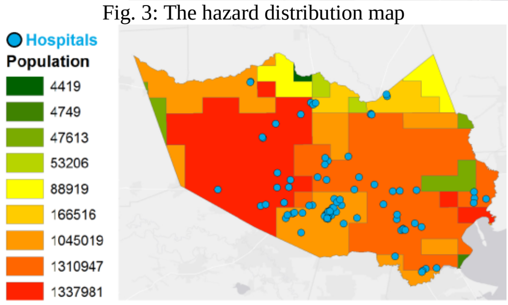


Fig. 4: Population distribution map

TABLE I. Partial Historical Outage Events in Area 4

Date	*Begin Time*	*Duration*	*Affected people*	*Event*
1/9/2012	10:07	70 min	19716	Flash Flood
8/16/2013	16:57	110 min	103000	Thunderstorm Wind
8/1/2014	4:00	55 min	26000	Flash Flood

V. Case Study

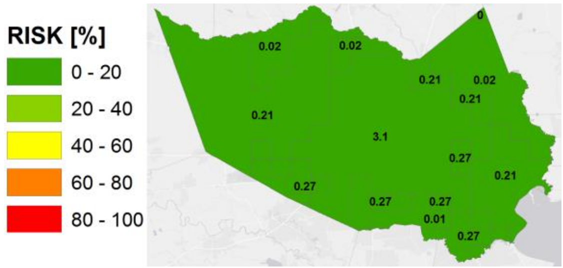
A. Test Scenario

The case study is implemented in Harris county using part of the network that is operated by CenterPoint Energy [27]. The historical blackout data has been collected from [2] (data from 2012/1/1 to 2014/12/31), correlated with historical weather conditions obtained from [14] and geo-located in the GIS map. The customer information such as population distribution [25], customer category, facility locations [26], is gathered and visualized in the same map. The forecasted weather data from NDFD [16] has been used. For comparison purpose, two case studies are implemented, one with normal weather scenario (scenario 1) and one with severe weather scenario (scenario 2). The network is split into small polygon areas grouped with the same hazard value based on the forecasted weather distribution scenario. The customer data in each area is analyzed and imported to the customer cost model, together with the historical outage data and forecasted weather data. Fig. 3 shows the hazard distribution map in scenario 2 with the points indicating the historical outage locations. Partial historical outage data is stated in Table. 1. In Fig. 4, the population distribution is presented as colors and health care locations are shown as points. They are both under scenario 2. Based on the risk assessment theory, the corresponding risk index values are calculated and demonstrated in the GIS map.

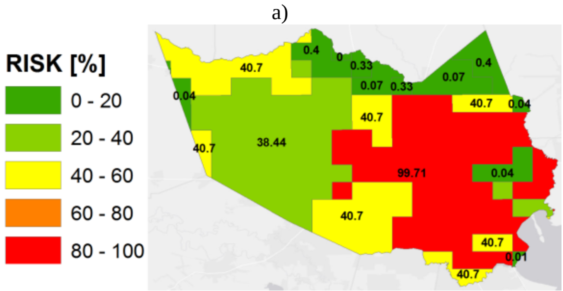
B. Study Results

The severity of the risk is indicated by using successive colors from green to red in Fig. 5 where a) shows the results under scenario 1 and b) under scenario 2. Percentages are used to present the value of risk indices, where the maximum value in scenario 2 is regarded as the denominator. The utilities can define the denominator based on their standard of risk acceptance. According to the resulting map, utilities can make decisions on whether it is necessary to send pre-warning notifications to their customers or if just mitigation actions can be taken to avoid such potential customer loss. In most cases, the customer impact can be tremendously decreased if a possible power outage is pre-notified, instead of a sudden unexpected power interruption.

The SVM had a success rate of 72%. For the remaining 28% of test cases: 25% were classified into the outage group, however there was no significant blackout during these events; and 3% of total test cases were wrongfully classified as not leading to a blackout. The prediction model used in the study seems to be well suitable for this application. It is to be noted that main concern should be with cases where blackout



a)



b)

Fig 5: Risk maps for: a) normal weather conditions, b) severe weather conditions

happened but was not predicted by the model, and these cases were only 3% of total test cases.

VI. Conclusion

The contribution of this study lies in the risk assessment methodology that estimates the impact of weather parameters on hazard, weather-caused outage vulnerability and weather-related customer losses. More specifically:

- The historical blackout events and the corresponding historical weather conditions are used to train a Support Vector Machine that predicts if future weather conditions will cause an outage.
- Weather factors are considered in the evaluation of Customer Interruption Cost for different customer categories and it is shown that the impact of different weather factors varies based on customer categories.
- It is shown that the distribution system operators can benefit from the customer risk assessment results by being aware of the impending risk allowing them to take preventive countermeasures to avoid potential customer losses.
- It is also shown that the utility customers can benefit from the customer risk assessment results by being pre-warned of the potential blackout and being provided the time to make preparedness plans to mitigate their loss.

References

[1] R. J. Campbell, "Weather-related power outages and electric system resiliency." Congressional Research Service, Library of Congress, 2012.

[2] Blackout Tracker, Eaton corporation. (January 24, 2016) [Online] Available: http://powerquality.eaton.com/blackouttracker

[3] P. Hines, et al., "Trends in the History of Large Blackouts in the United States," University of Vermont, 2008.

[4] B. Llavall, F. Xavier. "Reliability worth assessment of radial systems with distributed generation." Illinois Institute of Technology, 2010.

[5] R. Billinton et al., "Methods to consider customer interruption costs in power system analysis," CIGRE, TF 38.06.01, 2001.

[6] D. Cheng, B. Venkatesh, "Literature survey and comparison of consumer interruption costs in North America and Europe," Electrical and Computer Engineering (CCECE), 2014 IEEE 27th Canadian Conference on, pp. 1-7, Toronto, ON, May 2014.

[7] M. Lehtonen, B. Lemstrom, "Comparison of the methods for assessing the customers' outage costs," Energy Management and Power Delivery, 1995 International Conference on, Vol. 1, pp. 1-6, Nov 1995

[8] E. Leahy, R. S. J. Tol, "An estimate of the value of lost load for Ireland." Energy Policy, 39 (3), 1514-1520, March 2011.

[9] R.K. Subramaniam, et al., "Understanding Commercial Losses Resulting from Electric Service Interruptions", IEEE Trans. Industry Applications, Vol. 29, No. 1, January-February 1993.

[10] P.J. Balducci, J. et al., "Electrical power interruption cost estimates for individual industries, sectors and U.S. economy." PNNL-13797, Pacific Northwest National Laboratory, Richland, Washington, February 2002.

[11] Systems Control, Inc. "Impact Assessment of the 1977 New York City Blackout." Prepared for United States Department of Energy by Systems Control Inc., Palo Alto, California, July 1978.

[12] L. Goel, R. Billinton, "Prediction of Customer Load Point Service Reliability Worth Estimates in an Electric Power System," IEEE Proc.-Gener, Tans, Dist, Vol.141, No. 4, July 1994.

[13] M. Sullivan, D. Keane, "Outage Cost Estimation Guidebook", Electric Power Research Institute, Palo Alto, CA: December , 1995.

[14] National Centers for Enviromental Information – National Oceanic and Atmospheric Sdministration, "Data Access," (January 24, 2016) [Online] Available: https://www.ncdc.noaa.gov/data-access

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

[16] National Digital Forecast Database (NDFD) Tkdegrib and GRIB2 DataDownload and ImgGen Tool Tutorial, NWS, NOAA. [Online] Available: http://www.nws.noaa.gov/ndfd/gis/ndfd_tutorial.pdf

[17] National Centers for Enviromental Information – National Oceanic and Atmospheric Sdministration, "Numerical Weather Prediction," [Online] Available: https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/numerical-weather-prediction

[18] M. Cerny, "Economic and Social Costs of Power Outages: The Case of Pakistan." B.Sc. Charles University in Prague, 2013.

[19] K. Alvehag, "Impact of Dependencies in Risk Assessments of Power Distribution Systems," Thesis, KTH, Stockholm, Sweden, 2008.

[20] 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.

[21] Q. Yan, T. Dokic, M. Kezunovic, "GIS-Based Risk Assessment for Electric Power Consumers Under Severe Weather Conditions", 18th Mediterranean Electrotechnical Conference (Melecon 2016), Limassol, Cyprus, April, 2016.

[22] R. Herman, C. T. Gaunt, "Direct and indirect measurement of residential and commercial CIC: preliminary findings from South Africa Surveys," Proceedings of the 10th international conference on probabilistic methods applied to power systems, 2008.

[23] O. Dzobo, C.T. Gaunt, R. Herman, Investigating the use of probability distribution functions in reliability-worth analysis of electric power systems, International Journal of Electrical Power & Energy Systems, Volume 37, Issue 1, pp. 110-116, May 2012.

[24] Herman, R.; Gaunt, T., "Probabilistic interpretation of customer interruption cost (CIC) applied to South African systems," Probabilistic Methods Applied to Power Systems (PMAPS), 2010 IEEE 11th International Conference on , vol., no., pp.564-568, 14-17 June 2010

[25] Texas Natural Resources Information System, "Census 2010," [Online] Available: http://tnris.org/data-catalog/demographics/census-2010/

[26] Harris County Appraisal District, "Public Data," [Online] Available: http://pdata.hcad.org/GIS/

[27] CenterPoint Energy – Electric Transmission & Distribution, "Outage Tracker," [Online] Available: http://gis.centerpointenergy.com/outage tracker/