Gust Rectification and Stacking based Methodology for Power Outage Prediction under Hurricanes

Shiwen Yu
Department of Electrical and Computer Engineering
University of Illinois at Chicago
Chicago, USA
syu65@uic.edu

Lina He
Department of Electrical and Computer Engineering
University of Illinois at Chicago
Chicago, USA
lhe@uic.edu

2021 IEEE Power & Energy Society General Meeting (PESGM) | 978-1-6654-0507-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/PESGM46819.2021.9638182

***Abstract*— The threat of hurricanes to the security and resiliency of power systems is increasing with climate change and global warming. Due to the complex relationship between the hurricane disaster and power outages, it is difficult for power system operators to predict outages accurately and allocate emergency supplies in advance. Therefore, establishing an efficient and effective approach to forecast power outage distribution is critical for hurricane mitigation and loss reduction. This paper proposes a machine-learning based prediction methodology based on the realistic historical hurricane outage data. The proposed methodology includes two parts: (1) rectifying gust via the power law, and (2) configuring stacked model based on machine learning algorithms and metrics optimization. In particular, the proposed gust rectification algorithm can improve the interpolation precision effectively. By implementing this configuration, the stacked model has a much better performance than the single model in the stacking processing. This model is trained by a comprehensive data set that consists of the information of hurricanes, power systems, and geography. In the case study, the realistic data of two historical hurricanes is applied to verify the developed model. The results demonstrate that the proposed methodology can improve the performance of power outage prediction effectively.**

Keywords—hurricanes, power outage prediction, gust rectification, power law, and stacked model

I. Introduction

Hurricanes have resulted in severe damages to America power systems and extensive economic losses for decades. For instance, hurricane Katrina in 2005 caused that 2.7 million people lost their power and the resulting total loss was up to $160 billion. Hurricane Sandy in 2012 led to a massive power outage in the Northeast and $70.2 billion loss in total. About 8.5 million people experienced power outages [1-3]. Unexpected power outages not only cause huge economic losses, but also result in casualties and political incidents. Therefore, it is significant to develop an accurate methodology to predict power outages to mitigate hurricanes and reduce economic losses.

The power loss due to outages is a critical metric in measuring power system resilience, as it indicates the ability of the electrical grid to provide adequate services under hurricanes [4-5]. Currently, some researchers have constructed the resilience models by incorporating the function of the outage prediction under hurricanes. The work of [6] applied the device failure models and power flow calculation using the Monte Carlo method to determine power outages. However, it is not applicable for complex systems due to its high computational cost. In [7], the capacity loss due to the damage of power plants was incorporated in the resilience model based on fragility cures. However, this model only considered wind speed and overlooked other information. Pre-hurricane allocation of generation resources was studied in [8] to evaluate system resilience based on the forecasted failure probabilities of devices. But the device failure is not a criterion of outages, since devices may be damaged in one place and power demand is in another. Therefore, the accurate location prediction of outages is of great importance in pre- and post-hurricane resources allocation. As known, the precise resilience measurement and practical enhancement planning strongly rely on the accurate outage prediction model.

Due to the wide variety of features involved in outage prediction, the prediction model is difficult to develop with clear causality and physical meaning. In the existing studies, outage prediction models are usually developed using statistical methods. Generalized linear models were explored solely or in combination to predict the outage risk in [9-10]. However, they mainly focused on the model fit and the applied variables were limited. The Random Forests (RF) outage model was constructed in [11] with a simplified input from [12]. Since it did not include the utility-specific data, its applications is limited to the prediction of short-term outages. The reference [13-14] discussed the two-stage approach, which first applied the grid cell with outages to predict outage quantities. However, this approach highly relied on the classifier in the first stage and can decrease the prediction accuracy in some cases. Currently, the developed outage prediction models mainly use a single algorithm, which is not sufficient to present the prediction model accurately. Compared with the single algorithm, the hybrid algorithm has shown their superiority in damage prediction under hurricanes [15]. However, it has not been applied on the multi-model configuration to improve the performance of outage prediction yet. To fill this gap, this paper proposes a new outage prediction model based on the gust rectification and stacking algorithms that can reduce prediction errors effectively [16]. The comprehensive data set that is applied to generate the model consists of the information of hurricanes, power systems, and geography.

Due to the strong correlation between distributions of outages and gusts, the gust distribution is of great significance to accurate outage prediction. In this paper, the power law distribution reflecting the change of wind speed with the altitude is first used to rectify the gust, as the altitude impact is overlooked when generating a gust map using interpolation. After gust rectification, a case study is given to verify the overall improvement of the prediction metrics. There are two layers in the stacked model that is developed using different machine learning algorithms. The first layer composed of RF, Gradient Boosting Decision Tree (GBDT), and Adaptive Boosting (AdaBoost) is to generate new explanatory variables by 2-fold cross-validation to avoid overfitting, and the second layer based on RF is to make the final prediction. The stacked model theoretically outperforms a single algorithm because it has comprehensive learning mechanisms and the metric to be optimized in each layer can be adjusted flexibly to achieve the best prediction.

Compared with the existing methods, the proposed methodology applied the gust rectification algorithm, which ensures that the model can use both the high-resolution gust data and the gust data interpolated from low-resolution data.

This can greatly enhance the adaptability of the model and maintain a good prediction performance even when the resolution of the gust data is not high. In the stacked model, the proposed hybrid model has a better performance than the single models in all predictive metrics by configuring the stacking structure and setting the optimization metrics.

In practice, the target area is divided into multi grid-cells to capture the spatial correlation and prediction accuracy convenience [17]. In this manner, even a simple model like the logistic regression discussed in [18] can be used in a wide area by fitting each grid cell. Since the grid cells with a smaller size can capture the spatial correlation and make full use of the higher resolution of some data, this paper uses 1 km^2 grid cell to ensure the accuracy of the proposed model [5]. The effectiveness of the proposed methodology is verified via the realistic data of two historical hurricanes and power outages. This methodology can help power system operators make planning decisions.

This paper is organized as follows. Section II introduces the data and variable processing. Section III outlines the stacked model construction, metrics, and cross-validation testing. Section IV shows the testing results of outage prediction and verifies its effectiveness. Section V is the conclusion.

II. Data and Variables

A. Description of data and variables

The data used in this paper comes from hurricanes, power systems, and geography. The hurricane data refers to the maximum 3-sec wind gust. The power system data includes the number of customers, poles, pole-mounted transformers, box-type transformers, guy wires, and the length of overhead lines. The geography data includes the longitude, the latitude, the altitude, the slope aspect, the slope position, the underlying surface type, and the surface roughness. In [13], all explanatory variables were categorized into the static and dynamic types based on whether they are time-dependent. Similarly, the data used in this study is divided into 8 static and 6 dynamic explanatory variables. In this way, only dynamic explanatory variables need to be updated in actual applications.

The target area is divided into the grid cells with the size of 1 km^2 for data collection. Variables are defined based on the collected data. The response variable (Y) is the dichotomous variable, which is 1 if outages happen in the grid cell and otherwise is 0. The proposed method in this paper is to predict the response variable of the grid cell. The static explanatory variables includes the longitude ($X_{\_Lon}$), the latitude ($X_{\_Lat}$), the altitude ($X_{\_Alt}$), the slope aspect ($X_{\_SA}$), the slope ($X_{\_Slo}$), the underlying surface type ($X_{\_UST}$), and the surface roughness ($X_{\_SR}$). The dynamic explanatory variables include the maximum 3-sec wind gust ($X_{\_Gus}$), the number of customers ($X_{\_Cus}$), poles ($X_{\_Pol}$), pole-mounted transformers ($X_{\_PT}$), box-type transformers ($X_{\_BT}$), guy wires ($X_{\_GW}$), and the length of overhead lines ($X_{\_LOOL}$).

B. Data preprocessing

When the gust data from monitoring stations is available, the interpolation is often used to generate high-resolution gust maps. However, the interpolation only considers the distance difference and ignores the impact of the altitude on wind speed that can be depicted by the power law. Hence, the generated wind gust is just the wind gust at the average altitude of the monitoring stations, rather than that at the actual altitude of the grid cells. Therefore, the collected maximum 3-sec wind gust in each grid cell is rectified by using the power law shown in (1).

$$X_{i,Gus} = V_{i,Gus}\left(\frac{Z}{Z_i}\right)^{\alpha_i} \tag{1}$$

where the symbol $X_{i,Gus}$ is the rectified gust in the grid cell i, the symbol $V_{i,Gus}$ is the original gust, the symbol Z is the averaged altitude of all the monitoring stations, the symbol Z_i is the altitude, and the symbol α_i is the surface roughness coefficient [17]. Based on the power law, the wind speed increases with the increase of the height. Thus, the maximum 3-sec wind gust is rectified to the actual altitude of grid cells.

For grid cells that do not have customers, there is no outage under any circumstances. These grid cells are less meaningful for the model training while increasing the computing cost. Therefore, the grid cells that have no customers are deleted from the dataset. For the nominal variable, this paper uses the one-hot encoding to convert it to 0/1 bits, which enables to enlarge the variable space and speed up the modeling [19]. One-hot encoding is illustrated in Fig. 1.

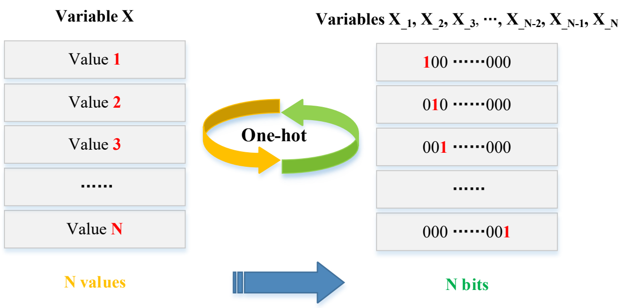

Fig. 1. One-hot encoding

As shown in Fig. 1, before encoding, the nominal variable X has N types of values, representing N categories. After the one-hot encoding, it is encoded as a bit combination of one '1' and N-1 '0'. Thus, the variable X is extended to N variables according to its N categories. In this study, the underlying surface type ($X_{\_UST}$) is classified into 10 categories, which is extended to 10 explanatory variables (from $X_{\_U1}$ to $X_{\_U10}$) by using the one-hot encoding.

The variables with a large absolute value are prone to be recognized important in modeling. They can lead to large prediction errors. Therefore, the variables need to be scaled to eliminate such effect. Since the value ranges of variables are different and change with data updates, it is complicated to define the base line values. Thus, this paper uses the min-max normalization as shown in (2) to convert the variable in the range from 0 to 1 [17].

$$X^* = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{2}$$

where the symbol X is the original variable and the symbol X^* is the scaled variable ranging in [0,1]. the symbols X_{max} and X_{min} are the maximum and minimum values of X. After data preprocessing, the summary of variables used in this paper are shown in TABLE I.

TABLE I. SUMMARY OF THE VARIABLES USED IN THIS PAPER

Explanatory Variable	Static		Dynamic	
	Notation	***Meaning***	***Notation***	***Meaning***
1	$X_{\_Lon}$	longitude	$X_{\_Gus}$	the maximum 3-sec wind gust
2	$X_{\_Lat}$	latitude	$X_{\_Cus}$	number of customers
3	$X_{\_Alt}$	altitude	$X_{\_Pol}$	number of poles
4	$X_{\_SA}$	slope aspect	$X_{\_PT}$	number of pole-mounted transformers
5	$X_{\_Slo}$	slope	$X_{\_BT}$	number of box-type transformers
6	$X_{\_USI}$ – $X_{\_USI0}$	underlying surface type	$X_{\_GW}$	numbers of poles without guy wires.
7	$X_{\_SR}$	surface roughness	$X_{\_LOOL}$	length of overhead lines

C. Variable selection

Not every variable is indispensable for modeling, since too many variables with low importance will increase the complexity and computational cost of the model. In addition, some variables introduce noises to reduce the prediction effect of the model. Therefore, the variables need to be selected and the method is introduced as follows. First, the Pearson correlation analysis is performed between the explanatory variables to help check and understand variables. This analysis is helpful to discover potential redundant variables. For example, if the linear correlation between two variables is strong, it indicates that one of them is redundant. Second, the importance of explanatory variables is identified and ranked based on the RF classifier. In this way, variables with high importance are retained, while variables with low importance are deleted.

III. MODELING

A. Models and metrics

This study uses three machine learning algorithms as the basis to build a stacked model, including the RF, the GBDT , and the AdaBoost. The RF is an ensemble learning method constructed by the decision tree and bagging [20]. It uses the subsets of data and features to train each decision tree, and the trees are ensemble to build the forest. It performs well in outage prediction under hurricanes and can ensure a high accuracy with small generalization errors [10-13]. The GBDT builds the tree by the boosting method [21]. It uses the negative gradient of the loss function to train the boosting tree, thus have a high accuracy and strong robustness. The AdaBoost trains a series of weak classifiers by updating the weight of the training data, and ensembles the weak classifiers to obtain a strong classifier [22-23]. It is the first boosting algorithm and fast in training. It can be less susceptible to overfitting.

It is noted that their respective advantages are represented in the stacked model. For example, the RF and GBDT algorithms have a high prediction accuracy, and the AdaBoost algorithm can speed up the computational speed. The configuration of the stacking structure is introduced in the next section. The performance of outage prediction is measured by three metrics, including the precision, the recall, and the F_β score [24]. They are given by

$$P_r = \frac{T_P}{T_P+F_P} \tag{3}$$

$$R_e = \frac{T_P}{T_P+F_N} \tag{4}$$

$$F_\beta = (1+\beta^2)\cdot\frac{P_r \cdot R_e}{\beta^2 \cdot P_r + R_e} \tag{5}$$

where the symbol P_r is the precision metric that represents the proportion of the grid cells, which actually have outages among the grid cells that are predicted to have outages, the symbol R_e is the recall metric that represents the prediction to find all grid cells that actually have outages. The symbol T_P is the number of grid cells identified as outage in both actual and predicting conditions. The symbol F_P is the number of grid cells identified as outage in prediction but as not in actual condition. The symbol F_N is the number of grid cells identified as no outage in prediction but as outage in actual condition. The symbol F_β score is a comprehensive measure of precision and recall, and the symbol β is a positive real factor which means how much times recall is as important as precision.

B. Stacking

As shown in Fig.2, the stacked model is divided into two layers. The first layer is to generate new variables based on the original dataset and add them to form a new dataset. To avoid overfitting, this paper uses the 2-fold cross-validation when producing the new variables. Specifically, for each model in the first layer, one fold is used to train the model and the other is used to produce new variables. The second layer is to use the new dataset generated by the first layer to build a prediction model.

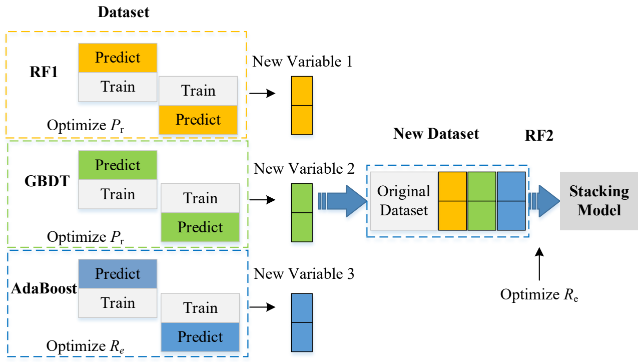

Fig. 2. Stacking construction

At each layer, the single models are assigned to optimize different metrics based on the aforementioned advantages. In the first layer, the RF and GBDT are optimized to improve the precision due to their high accuracy, and the AdaBoost is optimized to improve the recall due to its high speed [11-14, 17]. In the second layer, the RF is optimized to improve the recall as the power system operator cares more about the comprehensiveness of forecast results. Therefore, all the outage grid cells are included in the prediction results.

The proposed stacked model integrates different learning mechanisms and advantages of single algorithms. It has the same input and output with the single model, but achieves better prediction performance by assigning metrics and generating new variables.

C. Cross-validation testing

When generating new variables as shown in Fig. 2, the model is trained on one fold first. On this fold, the target metric is optimized in the hyper parameter space by randomly selecting 80% data of the fold to train the model and the 20% to test. The optimization is repeated for 20 times until the hyper-parameters are found. Similarly, when measuring the performance of the stacked and single models, this paper also utilizes the cross-validation testing. This paper first optimizes the hyper-parameters on a random sample with 80% of the whole dataset, and then predicts the power outages and measures the metrics for the rest of 20% dataset. The test is conducted for 20 times and the average of each metric is used to evaluate the performance of the model.

IV. Case Study

A case study is conducted to verify the effectiveness and feasibility of the proposed method. The applied data is the realistic outage information of a coastal county under 2 historical hurricanes with identical geographic coordinates. In particular, since there is no high-resolution gust data available, the gust data is from the interpolation based on the gust of monitoring stations. As introduced before, there are 8 static and 6 dynamic explanatory variables. The effectiveness of the gust rectification algorithm and stacked model is verified, and the results are contrasted and discussed in this section.

A. Data preprocessing

The wind gust is rectified by using (1) as the gust map is generated using the interpolation. To verify the effectiveness of the gust rectification, this paper compares the performance based on the stacked model before and after the rectification. Due to the nominal characteristic, the underlying surface type is converted to 0/1 bits based on one-hot encoding as shown in Fig. 1, thus 23 explanatory variables are prepared. Finally, the variables are scaled by the min-max normalization as introduced in (2) to eliminate the negative impact of variables with large absolute values.

B. Variable selection

To check the rationality and understand the dataset, the Pearson correlation between variables are shown in Fig. 3. The stronger positive correlation between two variables is represented with the redder color. On the contrary, the stronger negative correlation between two variables is represented with the bluer color.

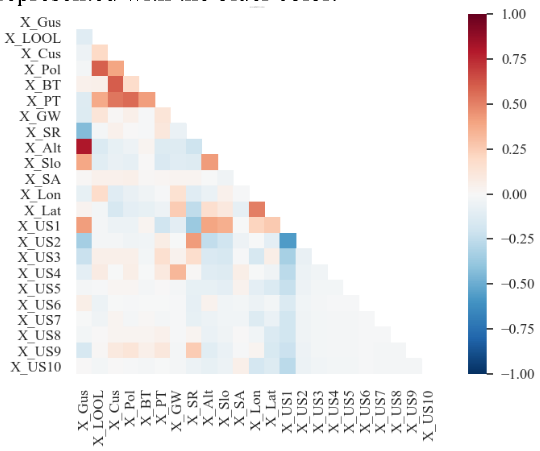

Fig. 3. Pearson analysis for variables

In Fig. 3, some correlations are in alliance with practical experience. For example, the maximum 3-sec wind gust is stronger with the altitude increases, and is weaker with the surface roughness increases. In addition, the grid cells with more customers need more poles, transformers, and the length of overhead lines. Although this indicates that the dataset is reasonable, the strong positive correlations implies that some variables are redundant and need to be deleted to simplify the model. It is noted that the correlations between the one-hot encoded underlying surface types are 0, since they are orthogonal bases within one group in this variable space. The variable importance is identified and ranked using the RF classifier. And Pearson correlation coefficients between important variables and Y are also calculated [25]. The results are shown in Fig. 4 and TABLE II.

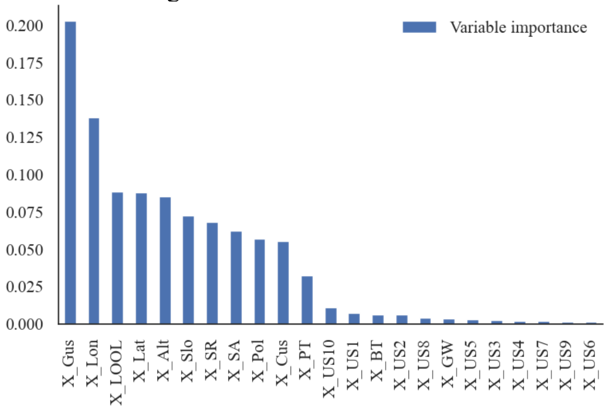

Fig. 4. Importance ranking of variables

TABLE II. Importance of the Variable

Variable	I	P	Variable	I	P
X_{Lon}	0.141	0.064	X_{Gus}	0.200	0.005
X_{Lat}	0.089	0.151	X_{Cus}	0.056	0.003
X_{Alt}	0.081	-0.024	X_{Pol}	0.057	0.067
X_{SA}	0.064	-0.048	X_{PT}	0.033	0.114
X_{Slo}	0.070	0.008	X_{BT}	0.006	0.084
X_{US1}	0.007	0.062	X_{GW}	0.003	0.058
X_{US2}- X_{US9}	<0.005	/	X_{SR}	0.068	-0.045
X_{US10}	0.012	-0.149	X_{LOOL}	0.090	0.065

As shown in Fig. 4, the importance of variables is descending ranked. It can be seen that the maximum 3-sec wind gust (X_{Gus}) is the dominant variable, and then the longitude for the hurricane track is from east to west in this case and the gust changes with the roughness. The importance (I) of each variable and the Pearson correlation coefficient (P) of important variable are provided in TABLE II. This paper selects variables with the importance more than 0.006 to feed into the model. Therefore, 13 variables including X_{Gus}, X_{Lon}, X_{Lat}, X_{Alt}, X_{LOOL}, X_{Slo}, X_{SR}, X_{SA}, X_{Cus}, X_{Pol}, X_{PT}, X_{US1}, and X_{US10} are utilized to train the stacked model. Pearson correlation coefficients provide the positive or negative correlations between variables and outage areas.

C. Modeling and predicting

The stacked model is developed based on Fig. 2 and the performance is measured by metrics introduced in section III. To emphasize the recall metric, this paper selects $\beta = 3$, then the results are shown in TABLE III.

TABLE III. Comparison of Model Prediction Results

Model	RF	GBDT	AdaBoost	Stacked model	Percent improved
P_r	0.753	0.795	0.570	0.889	11.82%
R_e	0.747	0.829	0.886	0.972	9.71%
F_3	0.748	0.826	0.840	0.963	14.64%

It can be seen that the presented model improves the P_r by 11.82%, the R_e by 9.71%, and the F_3 by 14.64%, compared with the best metrics of single models. Additionally, the performance before and after the gust rectification are compared based on the stacked model, which is shown in TABLE IV.

TABLE IV. EFFECTIVENESS OF GUST RECTIFICATION

Model	No-rectification	Rectification	Promoted percentage
P_r	0.884	0.889	0.6%
R_e	0.950	0.972	2.3%
F_3	0.943	0.963	2.12%

It shows that the gust rectification algorithm improves the P_r by 0.6%, the R_e by 2.3%, and the F_3 by 2.12% compared with the no-rectification model. This proves that in practical applications, even if there is no high-resolution gust data, the interpolation and gust rectification can be used to improve the prediction performance of the model. In TABLE III and IV, the effectiveness of the proposed stacked model and gust rectification is validated.

V. CONCLUSION

The Gust rectification and stacking based method are proposed to improve the outage prediction accuracy. Though correcting gust data by using the power law and configuring stacking structure properly, the precision, the recall and the F_3 score of outage prediction are improved compared to single models. In the stacking process, this paper properly assigns metrics for single models to optimize. The results of the case study verify that the gust rectification and presented stacked model can effectively promote the outage prediction performance. The proposed methodology is able to reduce the outage loss and enhance power system resilience.

REFERENCES

[1] Nat. Ocean. Atmos. Admin. (2018, Jan.), Costliest U.S. tropical cyclo nes tables updated. Nat. Hurricane Center, Miami, FL, USA. [Online] . Available: https://www.nhc.noaa.gov/news/UpdatedCostliest.pdf.

[2] Office of Electricity Delivery and Energy Reliability (OE). (2005, Sep .). Hurricane Katrina Situation Report #42. U.S. Department of Energ y, Washington, DC, USA. [Avalaible]. https://www.oe.netl.doe.gov/d ocs/katrina/katrina_092305%20_1500.pdf

[3] Office of Electricity Delivery and Energy Reliability (OE). (2012, No v.). Hurricane Sandy Situation Report # 20. U.S. Department of Energ y, Washington, DC, USA. [Avalaible]. https://www.oe.netl.doe.gov/n amed_event.aspx?ID=67.

[4] A. R. Berkeley, III, M. Wallace, "A framework for establishing critical infrastructure resilience goals," *Nat. Infrastruct. Advisory Council*, Washington, DC, USA, Oct. 2010. [Online]. Available: https://www.dhs.gov/xlibrary/assets/niac/niac-a-framework-for-establishing-critical-infrastructure-resilience-goals-2010-10-19.pdf.

[5] E. Vugrin, A. Castillo and C. Silva-Monroy. (2017, Feb.). Resilience metrics for the electric power system: A performance-based approach . Sandia Nat. Laboratories, Albuquerque, NM, USA. [Online]. https:// prod-ng.sandia.gov/techlib-noauth/access-control.cgi/2017/171493.pd f

[6] H. Zhang, P. Wang, S. Yao, et al., "Resilience Assessment of Interdependent Energy Systems Under Hurricanes," *IEEE Trans. Power Syst.*, vol. 35, no. 5, pp. 3682–3694, Feb. 2020.

[7] H. Zhang, P. Wang, S. Yao, et al., "Modeling Electrical Grid Resilience Under Hurricane Wind Conditions With Increased Solar and Wind Power Generation," *IEEE Trans. Power Syst.*, vol. 35, no. 2, pp. 929–937. Sep. 2019.

[8] H. Gao, Y. Chen, S. Mei, et al., "Resilience-Oriented Pre-Hurricane Resource Allocation in Distribution Systems Considering Electric Buses," *Proc. IEEE*, vol. 105, no. 7, pp. 1214–1233, Mar. 2017.

[9] H. Liu, R. A. Davidson, D. V. Rosowsky and J. R. Stedinger, "Negative binomial regression of electric power outages in hurricanes," *Journal of Infr. Syst.*, vol. 11, no. 4, pp. 258–267, Dec. 2005.

[10] H. Liu, R. A. Davidson, and T. V. Apanasovich. "Spatial generalized linear mixed models of electric power outages due to hurricanes and ice storms," *Reliab. Eng. Syst. Saf.*, vol. 93, no.6, pp. 897–912, Jun. 2008.

[11] S. D. Guikema, R. Nateghi R, S. M. Quiring, A. Reilly, et al., "Predicting hurricane power outages to support storm response planning," *IEEE Access*,vol. 2, pp. 1364–1373, Nov. 2014.

[12] R. Nateghi, S. D. Guikema and S. M. Quiring. "Power outage estimation for tropical cyclones: Improved accuracy with simpler models," *Risk Anal.*, vol. 34, no.6: 1069–1078, Jun. 2014.

[13] D. B. McRoberts, S.M. Quiring and S. D. Guikema, "Improving Hurricane Power Outage Prediction Models Through the Inclusion of Local Environmental Factors," *Risk Anal.*, vol. 38, no. 12, pp. 2722–2737, Oct. 2016.

[14] H. Hui, G. Hao, X. Xiang, et al., "Research on Prediction and Evaluation of User Power Outage Area Under Typhoon Disaster," *Power Syst. Tech.*, vol. 43, no. 6, pp. 1948–1954, Jun. 2019.

[15] H. Hou, S. Yu, H. Wang, et al., "A Hybrid Prediction Model for Damage Warning of Power Transmission Line Under Typhoon Disaster," *IEEE Access*, vol. 8, pp. 85038–85050, May. 2020.

[16] P. Shrivastava, and M. Shukla, "Comparative analysis of Bagging, Stacking and Random Subspace algorithms," In *Proc. Int. Conf. Green Computing and Internet of Things (ICGCIoT)*, Noida, India, 2015, pp. 511–516.

[17] H. Hou, S. Yu, H. Wang, et al., "Risk Assessment and Its Visualization of Power Tower under Typhoon Disaster Based on Machine Learning Algorithms," *Energies*, vol. 12, no. 2, pp. 1–23, Jan. 2019.

[18] R. Eskandarpour, A. Khodaei, "Machine Learning Based Power Grid Outage Prediction in Response to Extreme Events," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 3315 – 3316, Jul. 2017.

[19] H. David, and H. Sarah, *Digital design and computer architecture*, 2nd ed. San Francisco, Calif.: Morgan Kaufmann, 2012, pp. 129.

[20] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5-32, Oct. 2001.

[21] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," *Annals of Statistics*, vol. 29, no.5, pp.1189–1232, Oct. 2001.

[22] Y. Freund, and R. E. Schapipe, "A Decision-theoretic generalization of on-line learning and an application to boosting," *J. Comput. Syst. Sci.*, vol. 55, no. 1, pp.119-139, Aug. 1997.

[23] Y. Freund, and R. E. Schapipe, "Experiments with a new Boosting algorithm," In *Proc. 13th Conf. Machine Learnin*, San Francisco, USA, 1996, pp. 148-156.

[24] D. L. Olson, and D. Delen, *Advanced Data Mining Techniques*, 1st ed. Berlin: Springer, 2008, pp. 138.

[25] L. Jiazheng, L. Yu, X. Xunjian, et al. "Prediction and early warning technology of wildfire nearby overhead transmission lines," *High Voltage Engineering*, vol.43, no.1, pp.314-320, Jan. 2017.

Shiwen Yu received the B.S degree in Electrical Engineering in 2017 and M.S. degree of electric engineering in 2020 both from Wuhan University of Technology. He is currently pursuing Ph.D. degree in electrical engineering at the University of Illinois at Chicago. His research interest includes power systems stability, machine learning, and disaster prevention.

Lina He received the Ph.D. degree in electrical engineering at University College Dublin, Ireland in 2014, and the B.S. and M.S. degrees from Huazhong University of Science and Technology, China, in 2007 and 2009, respectively. She is currently an Assistant Professor in department of electrical and computer engineering, University of Illinois at Chicago, Chicago, USA. She was a Project Manager and Senior Consultant at Siemens headquarter in Germany and Siemens US from 2014 to 2017. Her research interests include modeling, control, and protection of power electronics based power system, renewable energy integration, and machine learning applications in power systems.