International Journal of Engineering & Technology, 7 (3.15) (2018) 91-94

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Development of Transmission Line Failure Rate Model using Polynomial Regression

M.K.N Arshad[1], N. Aminudin[2*], M. Marsadek[3], S.Z.M Noor[4], R.H Salimin[5], D. Johari[6]

[1,2,4-6]Universiti Teknologi MARA, Shah Alam, Malaysia
[3]Universiti Tenaga Nasional, Kajang, Malaysia
**Corresponding author E-mail: nza9669@gmail.com*

Abstract

Drastic climate change and more frequent occurrences of natural disaster which destruct power system infrastructure results in power delivery congestion at the transmission network. Heavily loaded transmission network that operates during adverse weather is very prone to outage, hence may trigger more critical problem such as voltage collapse. Research on risk of voltage collapse due to transmission line outage has been carried out by numerous researcher. Generally, this risk study involves two major parts; one is the assessment of voltage collapse impact due to the line outage and the other is the assessment of probability of line outage to occur. According to many literatures, precise probability estimation is very difficult to be evaluated since it is very unpredictable. Therefore, serious attention and studies have been focused in estimating the probability of transmission line outage prudently. The accuracy of probability assessed using Poisson distribution is very much dependent on its failure rate value. In this research, a weather-based transmission line failure rate model is developed using Ordinary Least Square (OLS) polynomial regression technique. To evaluate the effectiveness of the proposed method, comparative study with previous research which utilized robust MM-estimator technique is conducted. The results revealed that the proposed technique is more precise and the weather considered in the study has more significant impact compared to the preceding work. Thus, this finding contributes to more accurate probability estimation in risk of voltage collapse assessment.

Keywords: *OLS; Polynomial model; probability estimation; transmission line outage; transmission line failure rate model.*

1. Introduction

Nowadays, drastic climate change and natural disaster occur more frequently and very unpredictable. Occurrence of extreme weather incidents such as flash flood, landslide, storm and earthquake would seriously damage critical power system infrastructure involving outages of its components. This destruction leads the existing power operation to reroute and hence risk the power system stability and security. Therefore, research on identifying the probability of outages related with weather is crucial to be known precisely to minimize the impact of such incidents in power system operation.

To determine the probability of undesirable events to occur is indeed a challenging task which requires careful thought and insight. There are two fundamental methods to perform the probabilistic assessment, one is utilizing Monte Carlo simulation (MCS) and the other is by employing enumeration method [1], [2]. In either MCS or enumeration method, it is crucial to determine the most relevant probability distribution to be used for precise assessment. Based on the literatures, probability density distribution functions are employed depending on the type of outages considered. For instance, Normal distribution and Multivariate Normal distribution are often used to measure the uncertainty involving load level fluctuation [3]–[5], Poisson distribution, Bayesian model, Cartesian model, Fuzzy model, composite Fuzzy and Credibility theory are applied to assess transmission line and generator outages [4], [6]–[9] whereby Markov model and Weibull distribution are exploited to determine probability of equipment outages related to aging factor such as transformer and special protection system [10]–[12]. Among these outages, transmission line outage has drawn many attentions and has been reported to occur more frequent in recent years [13]. Majority of the researches employed Poisson distribution to determine the probability transmission line outages in power system operation [4], [14]–[18].

Poisson distribution model considers the counting nature of rare outage events which occur randomly at a fix time frame. From the formula, it can be seen that the accuracy of the overall probability assessed using this distribution is dependent on the failure rate value. many studies have been conducted to incorporate the effect of weather conditions in determining the failure rate of transmission line using various approaches. Chen et. al in [19] utilized fuzzy modeling to compute the failure rate contributed by wind speed and icing condition of the 39-bus New England test system whereby failure rate computation using Brostrom weather model under typhoon condition influenced by wind speed is introduced in [20]. Besides the previous two approaches, multiple linear regression method has been used in [14], [17], [18] to estimate the weather-related failure rate model. Reference [14] focused on online prediction of failure rate based on the temperature and wind speed, whereas [17] and [18] considers lightning activities and wind speed. In [14], maximum likelihood (MLE) or ordinary least square (OLS) regression technique is employed to determine the best estimates of the parameter in the model. The information of historical outage data with respect to weather condition during outage event is required to perform the failure rate estimation. Type of weather and approach considered in the previous studies may specifically applicable for some countries condition and sys-

Copyright © 2018 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

tem. To suit Malaysia climate and system, different weather type might need to be relooked.

This paper developed the failure rate model to determine the probability of transmission line outage for Malaysian system. In the research, the effect of local weather such as the lightning strength and amount of rain are utilized. Generally, the data used in this research is the same with [21]. However, the major difference between these works are in terms of the technique used to develop the model. In [21], robust MM-estimator is used whereby in this study, simple OLS with higher degree of polynomial is engaged in the development of the model. It is anticipated that the proposed technique will improve the accuracy of the failure rate as compared to the previous research.

2. Research Method

2.1. Data Collection for Failure Rate Model

The probability of transmission line outage can be calculated from Poisson distribution as shown in (1):

$$\Pr(E_i) = (1 - e^{-\lambda_i}) \times \exp^{-(\sum_{j \neq i} \lambda_j)} \quad (1)$$

where E_i is the *i*th outage, *Pr(Ei)* is the probability of E_i and λ_i is the failure rate of line *i*. The above equation showed clearly that the probability of the outage is very dependent on the failure rate of transmission line. In this study, the failure rate in Poisson equation is estimated using multiple linear regression model which is developed based on the historical lightning and rain data for Malaysia's 275kV real system. Four years data of outage events as well as lightning activities during such incidents from 2004 until 2007 are obtained from the local power utility provider. As for the rain data, the information on amount of rain during outage event is collected from the Department of Irrigation and Drainage (JPS) Malaysia.

2.2. Data Management and Failure Rate Calculation

The outage and weather data need to be processed to prepare a database for failure rate model construction. This process involves pooling (combine data that has the same characteristic) and discretizing (to its discrete value to ease data handling) of historical weather data according to the outage events. After the data is pooled and discretized, the failure rate is calculated. Generally, detail procedures on data management and failure rate calculation can be referred to [14] and [21]. The algorithm for data management process is shown in Fig. 1.

In failure rate calculation, the data that has the same discretized value (according to lightning and rain weather block) is combined. The failure rate for each group of outages at the same weather condition is determine as follows:

$$\lambda' = \log\left\{ \frac{\sum_{i=1}^{n_b} z_i}{D \times \sum_{i=1}^{n_b} L_i} \right\} \quad (2)$$

where λ' is the failure rate of group outage, $\sum_{i=1}^{n_b} z_i$ is the total number of lines in the group, $\sum_{i=1}^{n_b} L_i$ is the total length of line in the group and D is the duration recorded data (number of days in four years). Equation (2), is then re-transformed to determine the failure rate of each transmission line as follows:

$$\lambda_i = L_i \times \exp(X_i \times B) \quad (3)$$

where λ_i is the failure rate for *i*th line, L_i length of the *i*th line, X_i is the matrix of forecasted discrete weather of ith line ($X_i = [1 \quad x_{1(lightning)} \quad x_{2(rain)}]$) and B is the coefficient parameters arranged in matrix form ($B = [\beta_0 \quad \beta_1 \quad \beta_2]^T$). The coefficient parameters can be identified from the linear regression method.

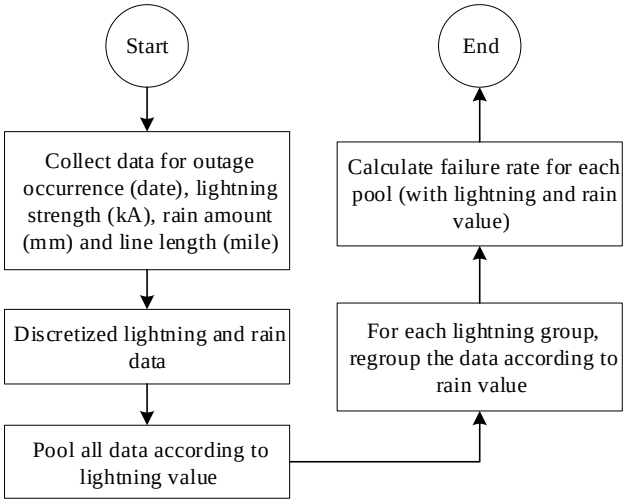


Fig. 1: Flowchart for overall failure rate data preparation

2.3. Failure Rate Model Based On Multiple Linear Regressions

Linear regression is a statistical tool that is adopted to model the correlation between two variables by fitting a linear equation to the observed or historical data. These two variables are referred to as independent and dependent variables. Independent variable is the variable that stands alone and the value is independent of changes of another variable in the model. As for the dependent variable, its value is changed when there are changes in independent variable's value. In this research, the failure rate is regarded as the dependent variable in which its value depends on the changes of lightning and rain weather (independent variables). Since there is more than one independent variable considered in the study, thus, the approach is called as multiple linear regression. Equation (4) shows one of the examples of multiple linear regression model.

$$\lambda'(x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + e \quad (4)$$

where

λ'	Failure rate
$\beta_0, \beta_1, \beta_2$	Model coefficient parameters
x_1	Discrete lightning value
x_2	Discrete rain value
e	Error

To determine failure rate value from (3), one must first construct the model as depicted by (4). If the values of model coefficient parameters, lightning and rain are known, then, the estimated fail-

ure rate ($\hat{\lambda}'$) can be calculated directly from the model. To estimate the value of model coefficient parameter, historical data which consists of the information on calculated failure rate and the respective weather conditions as discussed in sub-section 2.1 and 2.2 is required. This data will be utilized by the regression technique to obtain the model coefficient. OLS is one of the most popular techniques used in multiple linear regression modeling due to its simplicity.

The OLS estimation approach optimizes the fit of the model by minimizing the sum of the squared deviations between the actual failure rate calculated from historical data (λ') and estimated failure rate ($\hat{\lambda}'$) value obtained from the model. The OLS estimation method in residual form can be represented by [22]:

$$Min \sum_{i=1}^{n} \hat{e}_i^2 \tag{5}$$

where

$i = 1,2,3...n$ Sample size (n is the total sample size)

$\hat{e}_i$ Error between λ' and $\hat{\lambda}'$ ($\hat{e}_i = \lambda' - \hat{\lambda}'$)

$\hat{\lambda}'(x_1, x_2) = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2$

There are many combinations of regression model available for weather-related failure rate estimation. Equation 2.4 showed one of the examples of failure rate model used in the study. Nevertheless, there is no general consensus highlighting the advantage of one model over another [17]. Hence, the performances of several regression models need to be examined to identify the best model. Appropriate selection of modeling type must be chosen prudently because each model results in different estimation error and relationship between the variables. Generally, each of the models differs from one to another in terms of how the combination of lightning and rain is presented. Research in [21] only considers first and second-degree polynomials for the possible failure rate models. However, in this research higher degree of polynomials (3rd and 4th degree) are tested to improve the findings in [21]. As such, 216 possible failure rate models are developed and tested. Table 1 tabulates some of the sample of failure rate models used in the assessment.

The performance of failure rate model developed in this research will be compared with the model in [21] based on a few criteria such as its p-value, residual standard error and adjusted R^2 value. The p-value of a model shows the association between dependent (failure rate) and independent variables (weather). The correlation of these variables exists if the p-value is less than 0.05. Residual standard error represents the difference between actual (from raw data) and predicted (from regression model) failure rate value. The adjusted R^2 value denotes the impact of independent variable on the dependent variable. In this research, high value of adjusted R^2 signifies that weather has significant impact to the failure rate value. Generally, the procedures in determining the best failure rate model and comparative study performed in this work is shown in Figure 2. *m* in the flowchart represents total regression models (216) being tested.

Table 1. Sample of regression models

Degree of Polynomial	Model	Equation
1st	1_1	$\lambda' = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + e$
	1_2	$\lambda' = \beta_0 + \beta_1 x_1 + \beta_{12} x_{12} + e$
2nd	2_1	$\lambda' = \beta_0 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + e$
	2_2	$\lambda' = \beta_0 + \beta_{11} x_1^2 + \beta_2 x_2 + e$
	2_3	$\lambda' = \beta_0 + \beta_{22} x_2^2 + \beta_1 x_1 + e$
3rd	3_1	$\lambda' = \beta_0 + \beta_1 x_1 + \beta_{32} x_{12}^2 + \beta_{13} x_1^3 + e$
	3_2	$\lambda' = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{32} x_{12}^2 + \beta_{33} x_{12}^3 + e$
	3_3	$\lambda' = \beta_0 + \beta_2 x_2 + \beta_3 x_{12} + \beta_{32} x_{12}^2 + \beta_{13} x_1^3 + e$
4th	4_1	$\lambda' = \beta_0 + \beta_2 x_2 + \beta_{32} x_{12}^2 + \beta_{13} x_1^3 + \beta_{34} x_{12}^4 + e$
	4_2	$\lambda' = \beta_0 + \beta_2 x_2 + \beta_{32} x_{12}^2 + \beta_{33} x_{12}^3 + \beta_{14} x_1^4 + e$
	4_3	$\lambda' = \beta_0 + \beta_3 x_{12} + \beta_{22} x_2^2 + \beta_{13} x_1^3 + \beta_{34} x_3^4 + e$

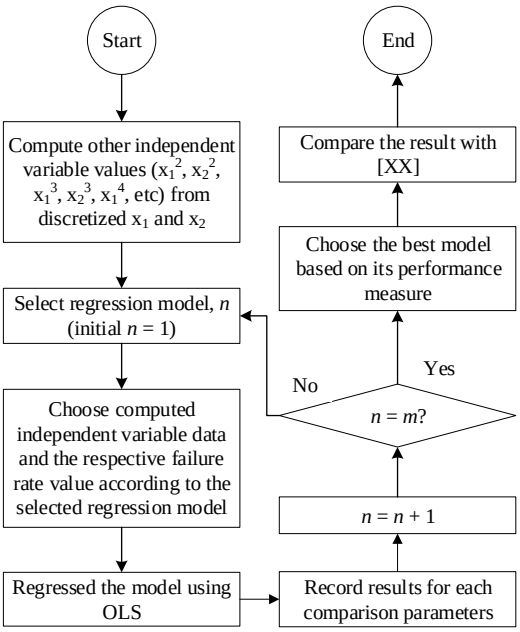


Fig. 2: Flowchart to determine the best failure rate model

3. Results and Analysis

In this research, the proposed polynomial failure rate models are developed using R software and data as explained in the previous section. The p-value, residual standard error as well as the adjusted R^2 of failure rate model developed using proposed method (OLS) are compared with the model in [21]. Table 2 shows the best failure rate model obtained for third and fourth degree polynomial using OLS technique and the best model in [21]. The performance for each model are tabulated in Table 3.

Table 2. The best failure rate model for each degree of polynomial

Degree of Polynomial	Model	Equation
2nd [21]	2_2	$\lambda' = -5.145000 + 0.000811 x_1^2 + 0.047334 x_2$
3rd	3_2	$\lambda' = -5.271000 + 0.031660 x_1 + 0.083080 x_2 - 0.000387 x_{12}^2 + 0.0000031 x_{12}^3$
4th	4_1	$\lambda' = -5.113000 + 0.065350 x_2 - 0.000177 x_{12}^2 + 0.000038 x_1^3 + 0.00000001 x_{12}^4$

Table 3: Parameter measured for every model

		Model		
		2_2	3_2	4_1
Residual standard error		0.315	0.291	0.290
Adjusted R^2 (%)		19.210	30.910	31.570
P-value	β_0	2.000×10^{-16}	2.000×10^{-16}	2.000×10^{-16}
	β_1	-	4.084×10^{-3}	-
	β_2	4.000×10^{-4}	2.780×10^{-5}	7.900×10^{-5}
	β_{11}	8.050×10^{-5}	-	-
	β_{13}	-	-	3.390×10^{-3}
	β_{32}	-	4.460×10^{-4}	3.090×10^{-3}
	β_{33}	-	5.060×10^{-4}	-
	β_{34}	-	-	4.000×10^{-4}

The above table shows the parameter measured for each model as outlined in Table 2. Model 2_2 is the best model developed in previous study [21] using robust MM-estimator method, whereby model 3_2 and 4_1 are the models developed in this research using OLS technique for third and fourth degree polynomial. The performance of Model 2_2 in determining failure rate value is compared with the models constructed in this research.

Overall, it is observed that the p-value of all coefficients obtained for every model is less than 0.05. This result means that the weather considered (lightning and rain) has significant impact in influencing the value of failure rate. Among these three models, model 4_1 has the least residual standard error as well as highest adjusted R^2 value as compared to others. The error produced by this model is reduced about 8% and the adjusted R^2 value is improved about 64% compared to preceding research.

Based on the above result, model 4_1 is chosen as the best model to estimates failure rate value in probability of transmission line outage assessment for Malaysia condition. The constructed mathematical model is given as follows:

$$\lambda' = -5.113000 + 0.065350x_2 - 0.000177x_{12}^2 + 0.000038x_1^3 + 0.00000001x \quad (6)$$

4. Conclusion

This research proposed a development of new weather-related failure rate model for Malaysia condition using OLS with higher degree of polynomial. Precise estimation of failure rate value is vital to ensure reliable probability of transmission line outage assessment is conducted for precise risk assessment. Based on the results, it can be concluded that the failure rate model developed using OLS technique with higher degree of polynomial has better estimation on failure rate value as compared to robust MM-estimator technique. Model 4_1 which was constructed using the proposed method has more accurate estimation since the residual standard error is 8% less than Model 2_2. Besides, high adjusted R^2 value of Model 4_1 also shows that the lightning and rain have more significant impact in the model, in contrast to Model 2_2.

Acknowledgement

The authors would like to thank the Research Management Institute (RMI), Universiti Teknologi MARA through research grant 600-IRMI/MyRA 5/3/LESTARI(0160/2016) for the financial support of this research.

References

[1] R. Billinton and L. Wenyuan, "Hybrid approach for reliability evaluation of composite generation and transmission systems using Monte-Carlo simulation and enumeration technique," *IEE Proc. C Gener. Transm. Distrib.*, vol. 138, no. 3, p. 233, 1991.

[2] W. Li, *Risk assessment of power systems models, methods and applications*. Canada: John Wiley & Sons Inc., 2005.

[3] J. McCalley *et al.*, "Probabilistic security assessment for power system operations," in *IEEE Power Engineering Society General Meeting, 2004.*, 2004, p. 212–220 Vol.1.

[4] N. Ming, J. D. McCalley, V. Vittal, and T. Tayyib, "Online risk-based security assessment," *IEEE Trans. Power Syst.*, vol. 18, no. 1, pp. 258–265, 2003.

[5] W. Hua, J. D. McCalley, and V. Vittal, "Risk based voltage security assessment," *IEEE Trans. Power Syst.*, vol. 15, no. 4, pp. 1247–1254, 2000.

[6] J. D. McCalley, V. Vittal, H. Wan, Y. Dai, and N. Abi-Samra, "Voltage risk assessment," in *IEEE Power Engineering Society Summer Meeting, 1999.* , 1999, vol. 1, pp. 179–184 vol.1.

[7] C. Ma, X. Y. Xiao, C. S. Li, Y. Zhang, and H. Q. Li, "Uncertain risk assessment model for catastrophic accidents in power system," *Int. J. Electr. Power Energy Syst.*, vol. 62, pp. 374–382, 2014.

[8] X. Fei and J. D. McCalley, "Power System Risk Assessment and Control in a Multiobjective Framework," *IEEE Trans. Power Syst.*, vol. 24, no. 1, pp. 78–85, 2009.

[9] F. Yongqing, W. Wenchuan, Z. Boming, and L. Wenyuan, "Power System Operation Risk Assessment Using Credibility Theory," *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 1309–1318, 2008.

[10] F. Weihui, Z. Sanyi, J. D. McCalley, V. Vittal, and N. Abi-Samra, "Risk assessment for special protection systems," *IEEE Trans. Power Syst.*, vol. 17, no. 1, pp. 63–72, 2002.

[11] F. Weihui, J. D. McCalley, and V. Vittal, "Risk assessment for transformer loading," *IEEE Trans. Power Syst.*, vol. 16, no. 3, pp. 346–353, 2001.

[12] L. Ning, W. Wu, B. Zhang, and P. Zhang, "A time-varying transformer outage model for on-line operational risk assessment," *Int. J. Electr. Power Energy Syst.*, vol. 33, no. 3, pp. 600–607, 2011.

[13] P. Henneaux, P. E. Labeau, and J. C. Maun, "Blackout probabilistic risk assessment and thermal effects: Impacts of changes in generation," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4722–4731, 2013.

[14] X. Fei, J. D. McCalley, Y. Ou, J. Adams, and S. Myers, "Contingency Probability Estimation Using Weather and Geographical Data for On-Line Security Assessment," in *International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), 2006.*, 2006, pp. 1–7.

[15] A. M. Leite da Silva, I. P. Coutinho, A. C. Zambroni de Souza, R. B. Prada, and A. M. Rei, "Voltage collapse risk assessment," *Electr. Power Syst. Res.*, vol. 54, no. 3, pp. 221–227, 2000.

[16] V. Krishnan and J. D. McCalley, "Contingency assessment under uncertainty for voltage collapse and its application in risk based contingency ranking," *Int. J. Electr. Power Energy Syst.*, vol. 43, no. 1, pp. 1025–1033, 2012.

[17] P. Kankanala, A. Pahwa, and S. Das, "Regression models for outages due to wind and lightning on overhead distribution feeders," in *Power and Energy Society General Meeting, 2011 IEEE*, 2011, pp. 1–4.

[18] Z. Yujia, A. Pahwa, and Y. Shie-Shien, "Modeling Weather-Related Failures of Overhead Distribution Lines," *Power Syst. IEEE Trans.*, vol. 21, no. 4, pp. 1683–1690, 2006.

[19] X. Chen, Y. Wu, and S. Lou, "Risk assessment for power system static security based on fuzzy modeling of weather conditions," in *2011 IEEE Power Engineering and Automation Conference*, 2011, vol. 1, no. 1, pp. 367–371.

[20] X. Song, Z. Wang, H. Xin, and D. Gan, "Risk-based dynamic security assessment under typhoon weather for power transmission system," *Asia-Pacific Power Energy Eng. Conf. APPEEC*, no. 1, pp. 1–6, 2013.

[21] N. Aminudin, M. Marsadek, N. M. Ramli, T. K. A. Rahman, and N. Razali, "Robust model for weather-related contingency probability estimation used for risk based security assessment," *Int. Rev. Model. Simulations*, vol. 7, no. 5, 2014.

[22] R. E. Schumacker, M. P. Monahan, and R. E. Mount, "A Comparison of OLS and Robust Regression using S-Plus," in *Multiple Linear Regression Viewpoints*, 2002, vol. 28(2).