Journal of Sustainable Development of Energy, Water and Environment Systems

http://www.sdewes.org/jsdewes

Year 2024, Volume 13, Issue 1, 1120531

Original Research Article

Stochastic Characterization of Faults in Electrical Transmission Networks: Case Study of the Electrical Community of Benin

***Mohamed Barate**[*1], **Eyouléki T. G. Palanga**[1], **Ayité S. A. Ajavon**[1], **Kodjo Agbossou**[2]*

[1]Centre d'Excellence Régional pour la Maîtrise de l'Electricité, Université de Lomé, 01BP1515 Lomé, Togo

e-mail: secretariat@cerme-togo.org

[2]Electrical and Computer Engineering Department, Université Du Québec à Trois-Rivières, Quebec, Canada

Cite as: Barate, M., Palanga, E. T. G., Ajavon, A. S. A., Agbossou, K., Stochastic Characterization of Faults in Electrical Transmission Networks: Case Study of the Electrical Community of Benin, J.sustain. dev. energy water environ. syst., 13(1), 1120531, 2025, DOI: https://doi.org/10.13044/j.sdewes.d12.0531

ABSTRACT

Electrical energy is a key factor in the development of any nation. Demand has been rising in recent years. This is putting existing power networks in difficulty, as they have not anticipated this meteoric rise. Blackouts occur daily and hurt the socio-economic development of nations, especially the most vulnerable. A predictive network outage solution would contribute effectively to better transmission network planning. Unfortunately, outage data for conventional networks are based solely on dispatcher reports and are difficult to exploit. This is the case for the Electrical Community of Benin transmission network. Understanding the predictive nature of this data would help implement fault prediction algorithms in this network. This paper aims to model outages and production in the Electrical Community of Benin power grid using probability laws. The objective is to contribute to the security of the electricity transmission network. Predicting the number of outages, their duration and the overall power lost will allow dispatchers to adjust electrical energy sources to avoid blackouts and save on electrical energy to impact the cost of producing electrical energy. The Kolmogorov Smirnov test, the error estimation using Akaike's information criterion and Bayesian information criterion on the one hand, and the Chi-2 test and the error estimation using the Root Mean Square Error on the other, were used to fit Benin Electrical Community network outages and accumulated sources using Weibull's law, outage duration using Erlang's law and energy lost using the Exponential law.

KEYWORDS

Transmission grids, Fault prediction algorithms, Outage durations, Stochastic characterization.

INTRODUCTION

Electrical energy is a development vector for every nation. It contributes to their economic and social development. But power cuts are a major brake on development. A study carried out in Indonesia shows that a 1% increase in the frequency of power cuts leads to a 0.055% drop in productivity per worker, and a 0.061% drop in value added per worker. Similarly, power outages reduce labor productivity and losses, which amount to around 4.91 million USD per year [1]. The impacts of power grid failures are well known in the literature, especially the major worldwide failures described in [2] and [3]. Most of these outages are caused by line tripping, protection system malfunctions, power oscillations, and voltage instability, as well as system splitting and collapse [4]. Among these, frequency drop and phase shift cannot be

[*] Corresponding author

prevented by system operating actions, whereas overloads and transmission voltage drops can. Given that inaccurate identification of a fault point can delay network recovery time, resulting in economic losses and customer dissatisfaction [5], It seems necessary to consider structural and reliable mechanisms to predict network failures. For example, predicting cascading faults would effectively manage voltage instability during cascading [6].

As the study carried out by [7] shows a weak predictive character of SCADA data in the transmission power system, an analysis based on event report data coupled with operating data of a conventional SCADA-supervised power system is required. The aim is to set up a predictive system that would enable players in the field to anticipate outages and minimize their impact, to contribute effectively to the development of any nation, in line with one of the criteria of smart grids [8].

A survey carried out by [9] showed that the breakdowns by region in the world in 2011 put sub-Saharan Africa in fourth place, with 210 major outages lasting 7.5 hours per outage. The transmission networks of Togo and Benin are no exception. A study carried out on the Electrical Community of Benin (ECB) network showed that between 2015 and 2018 there were 136,487 outages, or 3.89 outages per hour [10]. Given these facts about the ECB transmission network, an outage prediction solution would enable stakeholders in the field to prioritize certain power transmission lines in both countries (health, the army, government services, and certain essential neighborhoods, etc.), to minimize the impact of outages and prevent blackouts of the entire network. Because a minor interruption would affect a country's critical infrastructure [11].

For the characterization of electric power operating data, there are bottom-up methods, which combine statistical and engineering methods, and top-down methods, which combine economic and technological methods [4]. Statistical analysis effectively contributes to highlighting the predictive nature of a system or the data of a system [12]. Failure Modes, Effects, and Criticality Analysis is a widely used qualitative risk analysis method across various industrial and service applications [13]. For this analysis, probabilistic risk assessment (PRA) methods are gaining increasing attention in the context of high-voltage transmission network operations, as they offer a new approach to assessing safety in power system operations. This is due to the diversity of renewable energy sources connected to the grid, which tends to render deterministic methods for evaluating power system security obsolete [14]. In this context, [15] proposed a new PRA methodology aimed at assessing operational risk, measured by the probability of tripping a set of lines within a 10 to 15-minute interval. In [16], the practical method of risk probability is studied based on the failure probability of key equipment by defining the equipment factor and the failure factor. The method for evaluating risks and hazards in network planning is studied by defining the social influence factor and the loss-of-load factor. The methods applied to the risk assessment of the Shenzhen network are reasonable and practical. In [17], authors analyzes the interactions between protection system components and the power network under extreme events involving simultaneous faults and cascading failures. The proposed risk assessment considers detailed reliability models of the protection system components, including circuit breakers and protective relays. The effectiveness of the proposed risk assessment method is demonstrated using a modified 9-bus system and the IEEE 68-bus system. Risk-based electricity dispatching has been proposed as a viable alternative to security-constrained dispatching to reduce power grid costs and help better understand significant hazards [18]. A high penetration of renewable energy has caused stochastic power injection at the interface between the transmission and distribution systems [19]. As a result, stochastic analysis remains one of the most reliable methods for characterizing electrical energy data. Stochastic processes are mathematical models describing the behavior of evolving random variables [20]. Electrical data has these behaviors, which can be modeled by these processes. For statistical methods, probability density functions (PDFs) are commonly used. Mathematical models for characterizing these PDFs are numerous and varied.

For sources, [21] proved that among the set of probability distribution functions such as Gaussian, Skewnormal, Rayleigh, and Exponential, the Skewnormal function best describes BEC source data. Similarly, of the Gaussian, Skewnormal, and Weibull distributions used to characterize imports, Weibull holds sway for the TCN source and Skewnormal for the VRA/CIE source. The Kolmogorov-Smirnov test and performance indicators such as root Mean Square Error (RMSE) and coefficient of determination (R^2) were used. For the prediction of wind power production, [22] have shown that the mixture of Weibull and extreme value distributions (mixture of Gumbel, Freche, and Weibull) describes unimodal and bimodal wind behavior, while the mixture of extreme value and Lognormal distributions describes unimodal bell-shaped wind behavior.

Regarding network load, the generalized form of the Weibull distribution model provides a global characterization of billing data and household consumption [23]. Whereas, the Weibull and Log-normal distributions fit individual consumption in dwellings [24]. Each bus of the 95 IEEE network was well-fitted with a Gaussian mixture [25]. The estimation of the maximum demand in the low-voltage electrical network is successfully simulated using the Monte Carlo model, taking into account the statistical deviation of the demand of each half-hour from a gamma distribution [26].

Electrical network faults are also characterized in the literature. In [27], it proved that the Weibull distribution makes it possible to specify in a probabilistic manner the faults of equipment such as circuit breakers and current transformers during stochastic short-circuit events generated in the electrical network. The results indicate that the defects thus modeled are short-term. This does not generally characterize the defects in the transmission networks of developing countries. In [28], it is shown the non-Gaussian nature of the errors in PMUs. With the adjustment criteria of AIC, BIC and that of the modules. The errors in the PMUs were adjusted with a semiparametric Gaussian mixture. In [29], to model the resistance of electrical distribution networks with photovoltaics as backup against network failures, the Weibull function was selected as the best model. The probability laws are used as a performance criterion by taking the parameters of long-term failures. The data is not consistent enough to reflect the reality on the ground.

The transmission network is the backbone of the electricity network. For better planning, it is necessary to have the overall profile of the source, load faults, and its characters. To minimize the error rate in the choice of the model that best describes the quantities thus cited, it is clear that a single test with one or two adjustment criteria remains insufficient. This article proposes an analysis of transmission network failures based on all the data of sources, loads, and general parameters of ECB network failures with two tests and several adjustment criteria as selection criteria. Knowing that an imbalance between the load and the source can cause a frequency imbalance and lead to a total blackout, predicting holistically, the failures of the electrical network, the duration of these failures, and the lost power would contribute to quantifying the reliability of the electrical transmission network and save on the price of electrical energy.

The data modeling method will be presented in section two, followed by the study results in section three.

METHODS

Several methods are used today to characterize electrical data: the graphical method, the power density method, etc. [12], which determine the constants of the chosen distribution function. The machine learning approach is being increasingly studied, which allows several probability density models to be proposed so that the machine can propose the models that best describe the data under study. This study is based on this approach. The data to be modeled will be presented, followed by the method for modeling these data. It should be noted that the probability density will be calculated with the SciPy module in Python. The calculated probability density functions are in normalized form. To scale, you must use the *loc* and *scale* parameters according to eq. (1):

$$pdf(x) = \frac{pdf(y)}{scale} \quad \text{with} \quad y = \frac{x - loc}{scale} \tag{1}$$

Presentation of Electrical Community of Benin

The Electrical Community of Benin (ECB) manages the Togo and Benin electricity transmission networks. The network comprises 1,288.3 km of high-voltage lines, with transformer stations with a total capacity of 397.16 MVA. The network is supplied by renewable sources (hydroelectric and solar), thermal sources, and two import lines from the Volta River Authority (VRA) from Ghana and the Transmission Company of Nigeria (TCN). The network is managed using the N-1 method and a SCADA/EMS (Supervisory control and data acquisition/Energy Management System). A section of the ECB network in Togo is presented by [30]. It is represented by Figure 1.

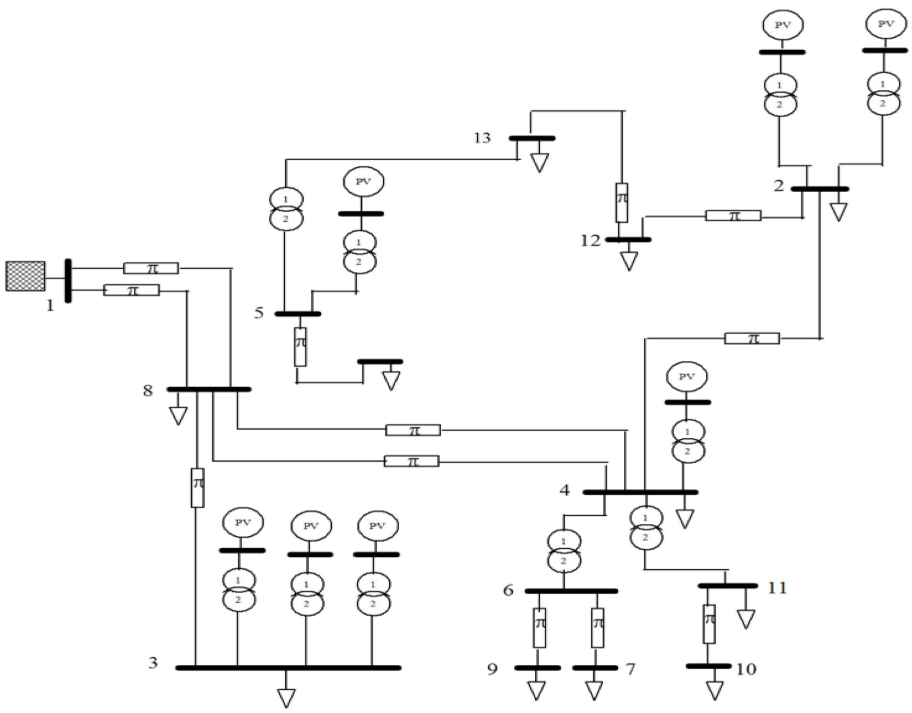


Figure 1. ECB network of Togo [30]

Data modeling method

The data to be modeled are ECB data from February 2014 to December 2018. These data come from SCADA measurement points and dispatcher incident reports. These data have undergone preprocessing to obtain the study data. Data from the SCADA database is hourly. And the incident reports are weekly. The preprocessing consisted of summing the hourly data (sources and consumption) to obtain the weekly data. The two types of data were then concatenated in the date column to find the study data. This processing was carried out using Anaconda's Jupiter 6.4.8. The study data consists of 204 rows and 4 columns. The statistical description of the data used is shown in Table 1.

The probability densities of the data were run through Kolmogorov Smirnov tests with Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) error estimators compared to existing models. This is followed by a test of cumulative Chi-2 densities and the root mean square error (RMSE) estimator to select the model that best characterizes the data studied. This method is described by the flow chart in Figure 2.

Table 1. Descriptive statistics

	Total power from sources (MW)	Number of triggers	Duration of outages (min)	Non distribute power (MW)
mean	76293,77	40,11	2035,16	159,03
std	24238,59	17,65	1815,60	110,38
min	6935,29	6	46	22,67
25%	54921,68	27,75	826	78,49
50%	74261,75	40	1645,5	122,39
75%	98405,06	51	2644	209,05
max	119392,98	110	12172	545,84
Coefficient of Skewness	0,02	0.56	2.35	1.26
Coefficient of kurtosis	-1,06	0,51	7.56	0.995

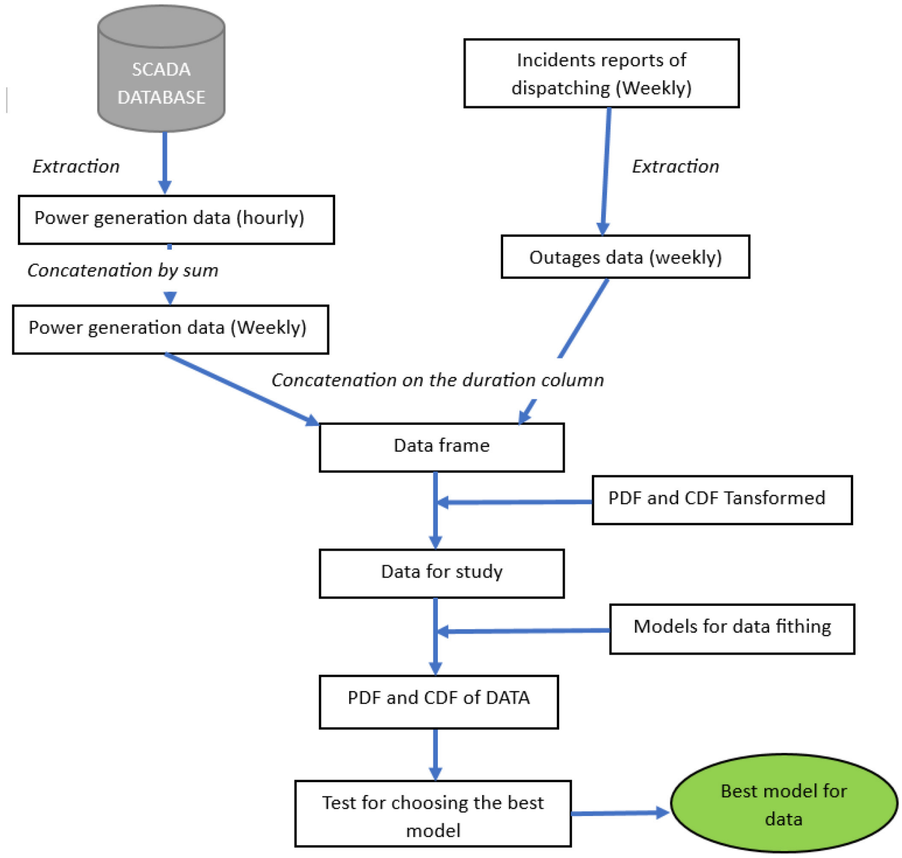


Figure 2. Methodology diagram

Theoretical foundations of the tools used

The following describes the tools used in this paper to model the data.

Kolmogorov-Smirnov hypothesis test. The Kolmogorov-Smirnov test is a hypothetical test used to determine whether a sample follows a given distribution known by its continuous distribution function, or whether two samples follow the same distribution. The reference hypothesis *H*0 and the opposite hypothesis follow eq. (2):

$$\begin{cases} H0 \rightarrow \ F^{*}(x) = F(x) \\ H1 \rightarrow F^{*}(x) \neq F(x) \end{cases} \tag{2}$$

where: $F^{*}(x)$, and $F(x)$ are the empirical and theoretical functions respectively.

In the case of curve fitting of given functions or models, the Kolmogorov-Smirnov (K-S) statistic is used as a relative indicator. The K-S test is best at estimating errors in curve-fitting models. It represents the level of rejection of the null hypothesis (*H*0) of the Kolmogorov-Smirnov test. The lower the KS value, the higher the acceptance rate of the null hypothesis. When the K-S value is less than 0.05, you are informed that the mismatch is significant. The *P*_value determines whether the *H*0 hypothesis is accepted or not. If the *P*_value is less than 0.05, the *H*0 hypothesis is rejected; otherwise, it is accepted.

<u>Akaike Information Criterion and Bayesian Information Criterion.</u> When it comes to choosing a model to describe data, it's difficult to choose from the ever-growing pool of models described in the literature. The Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are criteria for choosing the right model, striking a balance between the adequacy and complexity of the model that best describes the data under study. They are powerful measures of model selection in regression analysis. They enable the selection of simple models that best describe the data under study. AIC finds a model that maximizes the likelihood of the data while taking into account the number of parameters used. It is described by eq. (3):

$$\mathrm{AIC} = -\log(L) + 2k \tag{3}$$

where: *L* represents the maximum likelihood and *k* the number of parameters including the intercept and any additional predictors.

Like the AIC, the Bayesian Information Criterion (BIC) is another model selection criterion that takes into account both model fit and complexity. The BIC is based on Bayesian principles and provides a higher penalty for model complexity than the AIC. The BIC is given by eq. (4):

$$\mathrm{BIC} = \ -2\log(L) + kLog(n) \tag{4}$$

where: *L* represents the maximum likelihood, *k* is the number of parameters including the intercept and any additional predictors, and *n* is the sample size. The lower their values, the better the model describes the data under study.

<u>Sum of squared errors.</u> The sum of squared errors (SSE) is a measure of the deviation between the data and an estimation model. It is commonly referred to as deviance. The lower the SES, the tighter the fit of the model to the data used. It is used as an optimality criterion in parameter and model selection when fitting data. It is defined by eq. (5):

$$\mathrm{SSE} = \sum\nolimits_{\mathrm{i}=1}^{n}(X_{\mathrm{i}} - \bar{X}) \tag{5}$$

<u>Chi-2 goodness-of-fit test.</u> The chi-2 test is a goodness-of-fit test that checks whether a sample of a random variable *F*(*x*) gives observations comparable to those of a defined probability law *P*. The null hypothesis (*H*0) is that the random variable *F*(*x*) follows the probability law P. The null hypothesis here is that the observation is sufficiently close to the theory. The null hypothesis (*H*0) is that the random variable *F*(*x*) follows the probability law *P*. The null

hypothesis here is that the observation is sufficiently close to the theory and is generally rejected when $p \leq 0.05$.

Root mean square error. The root-mean-square error (RMSE) in statistics is an indicator that best measures the difference between the actual observed distributions and the predicted probabilities for each observation. The lower the RMSE value, the more effective the model is at modeling the data. The value of the RMSE is given by eq. (6) [12]:

$$\mathrm{RMSE} = \left[\frac{1}{n}\sum_{\mathrm{i}=1}^{n}(F_{\mathrm{i}}^{*} - F_{\mathrm{i}})^{2}\right]^{1/2} \tag{6}$$

where: F_i* and F_i represent the empirical and theoretical functions of the observed models, respectively. In this case, the CDF.

RESULTS AND DISCUSSION

This section describes the data modelling results in PDF and CDF and presents the test results for the best model choice.

Best models selecting

PDF functions have been chosen for each of the quantities studied (see **Table 2**). The literature models passed the Kolmogorov-Smirnov tests, the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) convenience test, and the sum-of-squares-errors estimation test. Considering the Kolmogorov-Smirnov test, only the Erlang and Beta distributions apply to the outage duration data. For the number of outages, all the distributions found are acceptable. As for the non-distributed power data due to outages, the Weibull and Lognormal distributions are accepted by the Kolmogorov test, while the last three are rejected. For the sources, even if the Kolmogorov test passes for the Weibull distribution with a rate of 90.1%, the *P*-value is less than 0.05, which allows us to reject. According to this test, no distribution can be used to model total data from ECB sources. Therefore, to use the data from the ECB sources, they will have to be taken independently, according to the studies carried out by [21]. All these results are shown in **Table 2**. The choice of results was made by first comparing the statistics of the Kolmogorov test. The lower the value of the statistic, the higher the success rate. Followed by SES, AIC, and BIC in succession. That said, in the case of outage durations for example, the Exponential distribution has the lowest AIC value, but its statistic value is 0.102, i.e. a Kolmogorov test success rate of 89.8% against the Erlang and Beta functions with successive AIC values of 2205.59, 2207.31 for a statistic of 0.045, i.e. a success rate of 95.5%, putting these two models ahead of the Exponential model.

Table 2. Test results for model selection

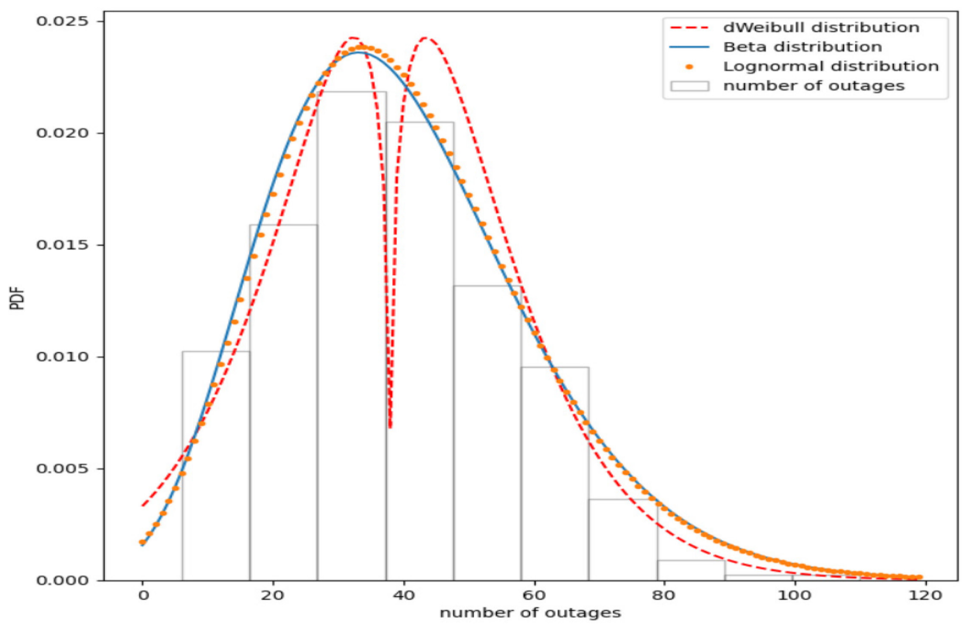
Data	Probability density functions		Model chose estimators			Kolgoromorov Smirnov test		
	Models	Distributions parameters	SSE	AIC	BIC	Stati-stic	*p*-value	Decision

Data	Probability density functions		Model chose estimators			Kolgoromorov Smirnov test		
	Models	Distributions parameters	SSE	AIC	BIC	Statistic	*p*-value	Decision
Duration of outages	Erlang	*k*=1.4718, *loc*= 35.4686 *scale*= 1358.59	2.71e-07	2205.59	-4153.67	0.05	0.782	Accepted
	Beta	*a*=1.47096 *b*=3.35 e+9 *loc*= 35.50642 *scale*=4.55e+12	2.71e-07	2207.31	-4148.31	0.05	0.779	Accepted
	Exponatial	*loc*=46.0 *scale*=1989.158	4.79e-07	2132.69	-4042.78	0.10	0.0276	Rejected
	Rayleigh	*loc*= -653.349 *scale*= 2.29 e+3	5.05e-07	2477.04	-4032.14	0.14	0.0004	Rejected
	dWeibull	*c*=1.021 *loc*= 1642.74 *scale*= 1206.65	5.33e-07	2357.95	-4015.53	0.14	0.0005	Rejected
Number of outages	dWeibull	*c*=1.2997 *loc*=38.4972 *scale*=15.1662	0.0043	1139.39	-2180.32	0.04	0.83	Accepted
	Lognormal	*s*=0.2119 *loc*=-42.515 *scale*=80.7935	0.0044	1084.17	-2176.56	0.05	0.56	Accepted
	Beta	*a*=9.20259 *b*=1.076 e+9 *loc*=-13.70973 *scale*=6.29 e+9	0.0044	1084.61	-2171.01	0.06	0.52	Accepted
	Gamma	*a*=9.20 *loc*=-13.709 *scale*=5.848	0.0044	1082.61	-2176.32	0.06	0.52	Accepted
	Skew normal	*a*= 2.529 *loc*= 20.461 *scale*= 26.379	0.0044	1088.15	-2175.14	0.06	0.47	Accepted
Power lost	Expo-normal	*K*=7.28039 *loc*=44.6098 *scale*=15.71597	0.0002	1388.36	-2860.94	0.04	0.88	Accepted
	Skewed Cauchy	*a*=0.799352 *loc*=54.232 *scale*=42.6779	0.0002	1412.03	-2.857.96	0.09	0.051	Accepted
	Skew normal	*a*=16.44380 *loc*=32.909 *scale*=167.422	0.0002	1388.53	-2799.49	0.11	0.010	Rejected
Power	dWeibull	*c*=2.1837 *loc*=75405.89 *scale*=24581.95	7.71e-09	2489.80	-4879.89	0.09	0.032	Rejected

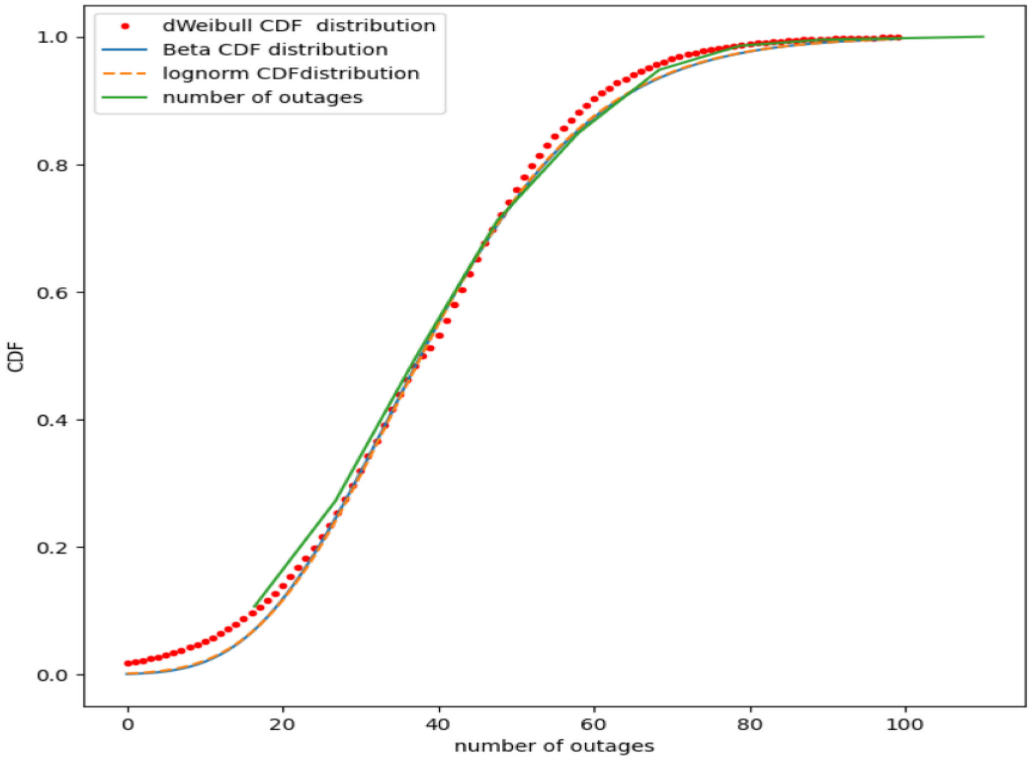
Data	Probability density functions		Model chose estimators			Kolgoromorov Smirnov test		
	Models	Distributions parameters	SSE	AIC	BIC	Stati-stic	p-value	Decision
	Beta	a=2.8054 b=1.59529 loc= 904.0497 $scale$= 1.19 e+5	1.09e-08	2390.83	-4802.13	0.17	0.010	Rejected
	Rayleigh	loc=3.60 e+4 $scale$=3.32 e+4	1.14e-08	∞	-4804.96	0.19	0.000	Rejected
	Lognormal	s= 0.037 loc=-5.605 e+5 $scale$=636437.6	1.17e-08	2425.41	-4794.25	0.16	0.000	Rejected
	Gamma	a=9618.08 loc=-2.294 e+6 $scale$=246.5	1.18e-08	2419.20	-4793.52	0.17	0.000	Rejected

Probability and cumulative densities functions of the best models selected

The graph shows the best visibility in terms of mathematical model selection for data modeling. The three best models that best model data have been presented in the form of PDF and CDF. According to Figure 4 for CDF greater than 0.78, the Skewed Cauchy (skewcauchy) distribution largely deviates from the CDF of the lost power data so cannot be considered a better model for the PDF distribution of undistributed power due to transmission network outages. **Figure 3**, **Figure 4**, **Figure 5**, and **Figure 6** show the respective PDF and CDF of the number of outages, the duration of outages, the power lost due to outages, and the cumulative powers of the sources and those of the models that best characterize them.

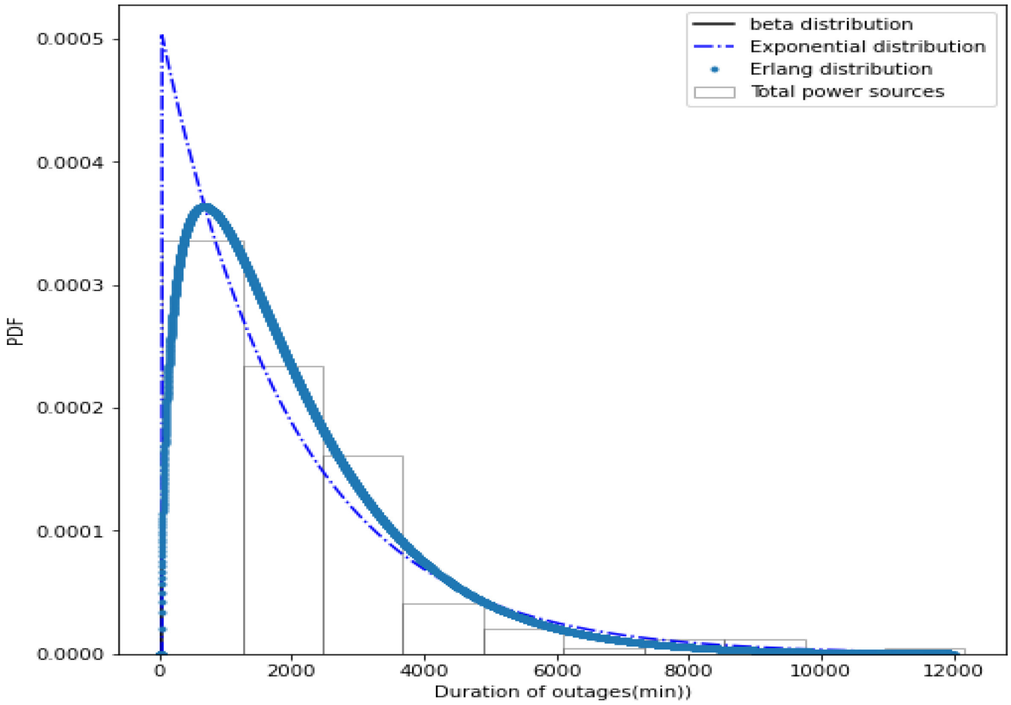


(a)

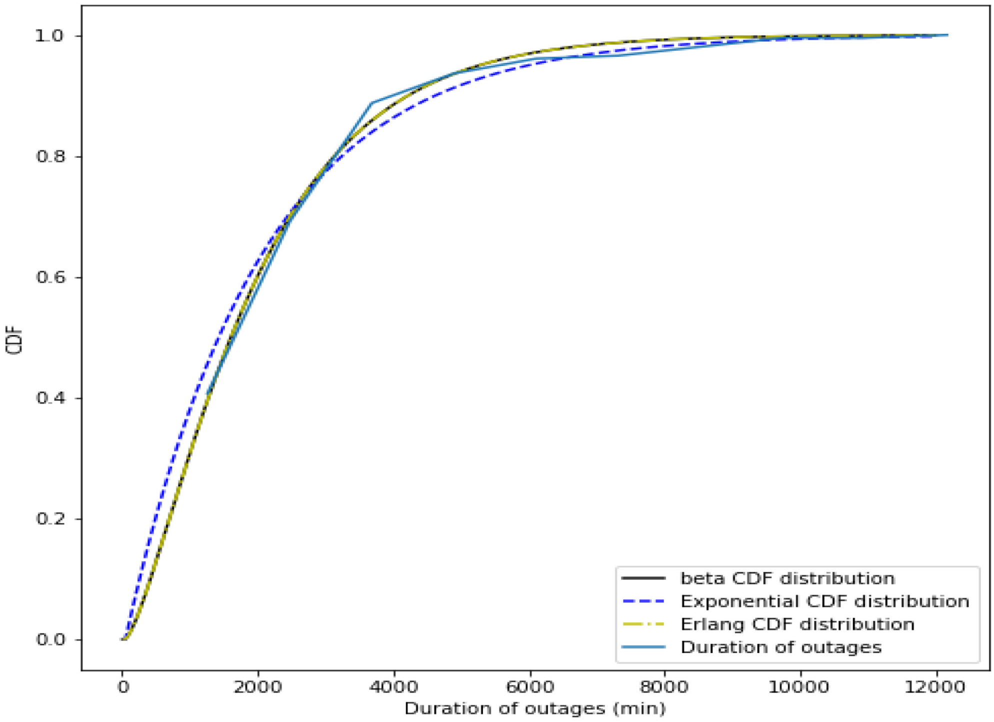


(b)

Figure 3. PDF and CDF modeling of the number of outages

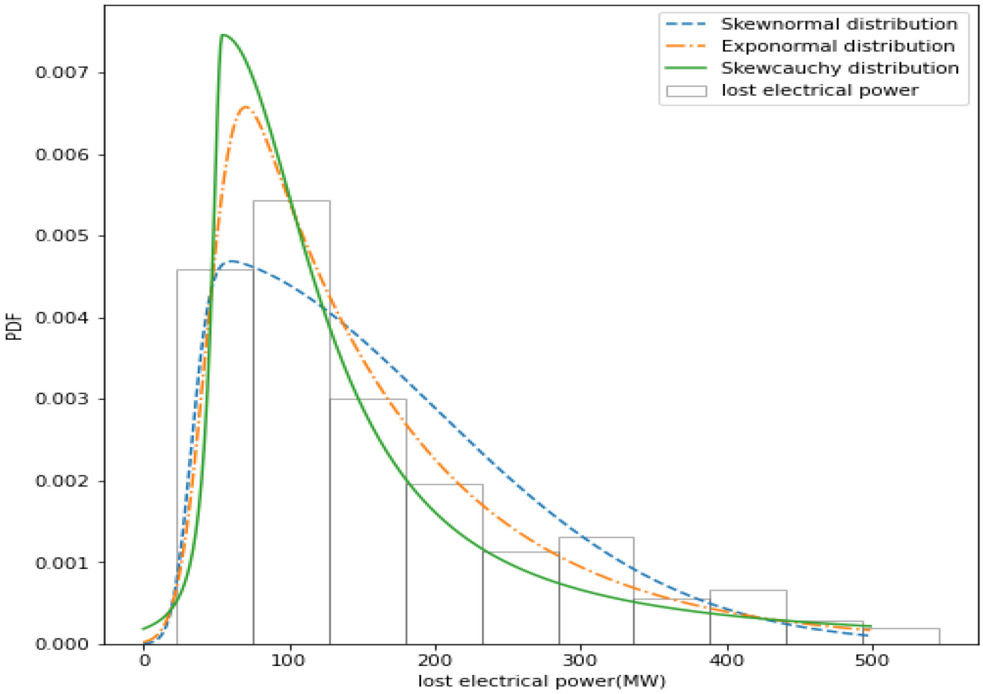


(a)

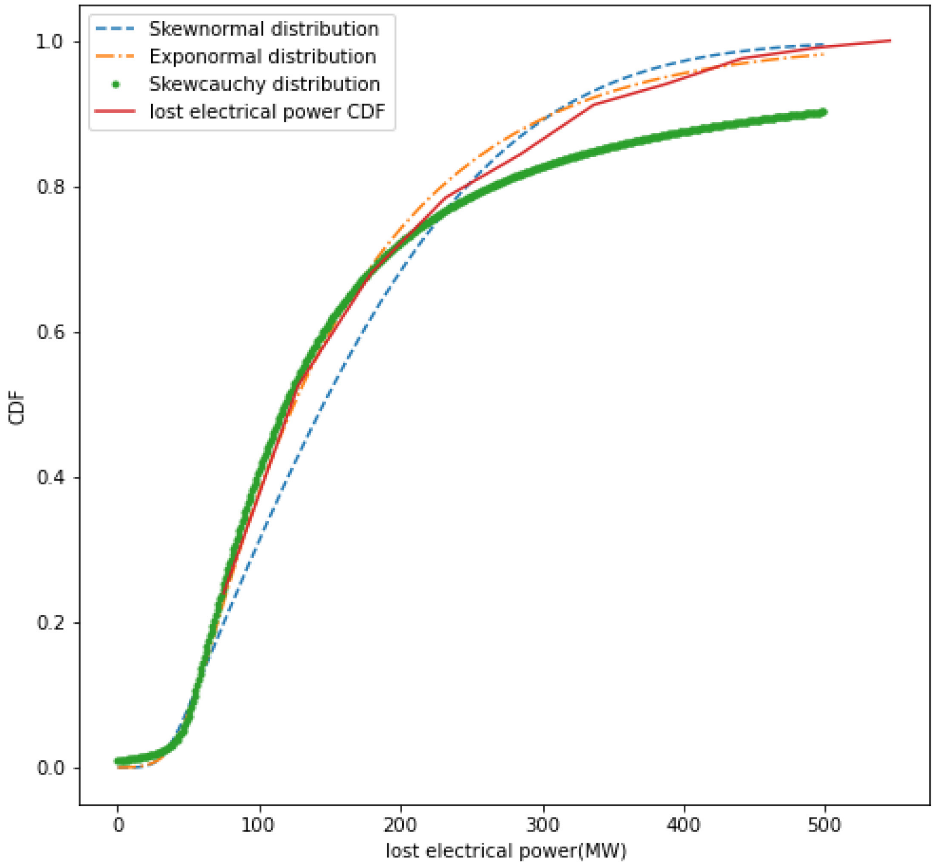


(b)

Figure 4. PDF and DCF modelling of outage duration

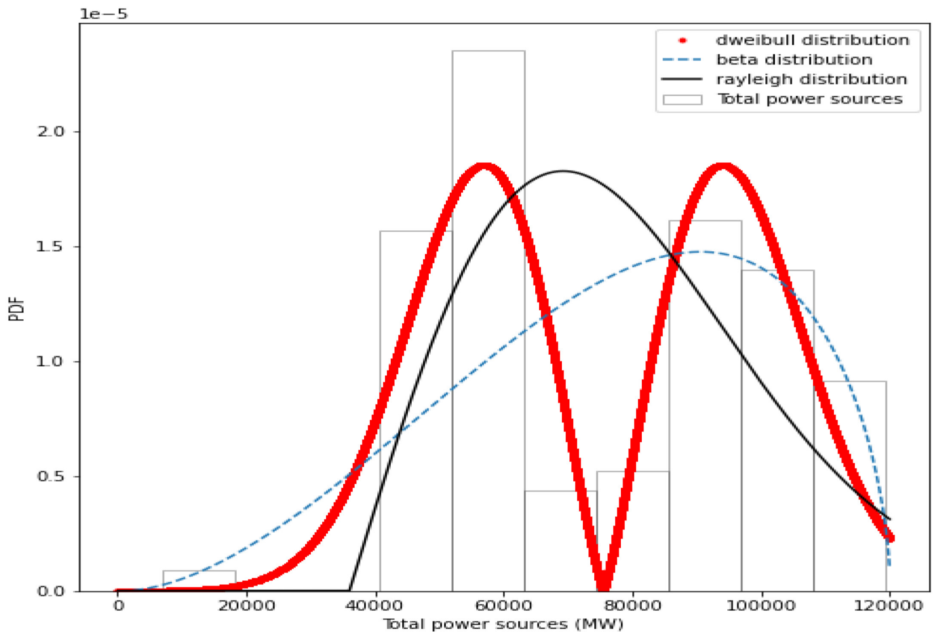


(a)

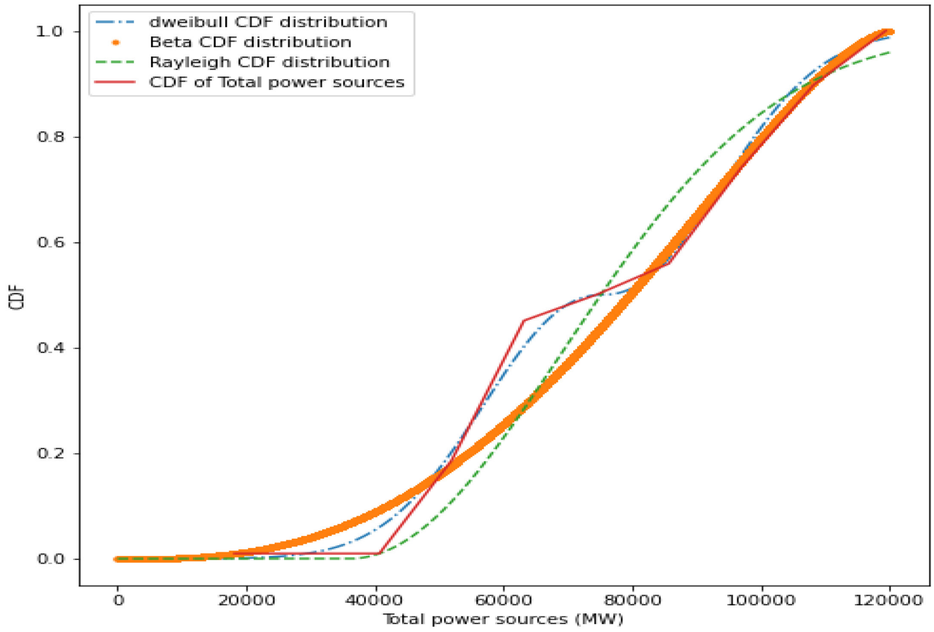


(b)

Figure 5. PDF and DCF modelling of lost power



(a)



(b)

Figure 6. PDF and DCF modelling of total sources

Choosing the model that best fits the data

Chi-2 tests and RMSE calculations on the CDF were used to select the best model to characterize our data. The results of these tests are shown in **Table 3**. For total source power, all functions were rejected by the Kolmogorov test. But the Chi-2 test and the RMSE value put the double Weibull (dWeibull) function in the first place. As for Power not supplied due to outages, the Chi-2 test and the RMSE value put the Exponentially modified Gaussian (Exponormal) distribution in the first place, followed by Skewcauchy. Although the RMSE value of the Exponential distribution was low for outage duration, it was rejected by the Kolmogorov test, enabling us to rank the other two. The Erlang function won out. As for the number of triggers, it's difficult to decide with both tests The Kolmogorov test should be used. For Khi-2, the Lognormal law wins out over the Kolmogorov test, which puts double Weibull ahead with a 95.8% success rate for a *P*-value of 0.83 against Lognormal with a 94.6% success rate for a *P*-value of 0.56. Considering the *P*-values of the Kolmogorov test, the Weibull distribution outweighs the Lognormal distribution. Considering the RMSE, that of the Weibull distribution is lower than that of the Lognormal distribution. Assuming the difference in chi-square between the two distributions, the Weibull model can be placed first, followed by Lognormal.

Table 3. Chi-2 test results

Data	distribution	chi_square	RMSE	Rang
Number of outages	Lognormal	4.382924	0.128426	2nd
	Beta	4.435108	0.127705	3rd
	dWeibull	4.644737	0.117384	1st
Duration of outages	Erlang	4.267320	0.171638	1st
	Beta	4.282475	0.17166	2nd
	Exponential	67.684226	0.163108	3rd
Power lost	Exponnormal	2.129430	0.13666	1st
	Skewcauchy	8.960681	0.14507	2nd
	Skewnormal	30.814606	0.16197	3rd
Power sources	dWeibull	36.928641	0.02587	1st
	Beta	73.006323	0.063119	2nd
	Rayleigh	87.413172	0.07388	3rd

Probability density functions retained after the study

According to the study, cumulative source and import data are not predictive, whereas outage data (duration of outage, number of outages, and power lost) are, as shown in **Table 4**.

Table 4. Fitted PDF for data

DATA	PDF theorical function	Functions parameters
Number of outages described by dWeibull distribution	$f(x,c) = \frac{c}{2}\lvert x\rvert^{c-1}e^{-\lvert x\rvert^{c}}$	*c*=1.2997, *loc*=38.4972 *scale*=15.1662
Duration of outages described by Erlang distribution	$f(x,k,\lambda) = \frac{\lambda^k x^{k-1} e^{(-\lambda x)}}{(k-1)!}$ with $\lambda = 1/scale$	*k*=1.4718, *loc*= 35.4686, *scale*= 1358.59
Power lost described by Exponormal distribution	$f(x,K) = \frac{1}{2K} e^{\left(\frac{1}{2K^2} - x/K\right)} erfc(-\frac{x-1/K}{\sqrt{2}})$	*K*=7.28039, *loc*=44.6098 *scale*=15.71597

This result confirms the non-Gaussian character of power system failures shown by **[28]** in the case of PMUs. As the field of study is not the same, then the models that describe power grid failures can vary depending on the systems and the source of the data. The prediction of sources can only be based on the study of data from these sources taken individually, as shown by **[21]**. The Weibull model for characterizing fault data in a network, proposed by **[29]** , is still with the results found. Because the transmission network has hybrid sources. The results show that data from network dispatcher reports can help predict outages number, and the energy a network can lose as a result of a power grid failure. They can also help predict the duration of outages in the transmission network. So the non-predictive nature of SCADA data in the power grid shown by [7] is no longer verified if static methods are used. SCADA data and data from protective device readings can therefore make an effective contribution to the prediction of outages in the transmission grid.

CONCLUSION

In this paper, to avoid blackouts in the transmission network and to save on energy losses due to repetitive breakdowns, a stochastic statistical analysis has been performed on ECB network outage data. The Kolmogorov-Smirnov test and model choice estimators such as AIC, BIC, and SSE were used to select the best models describing the PDFs of the data studied. The three best of these models are retained. The CDFs passed the Chi-2 test and the RMSE error estimation, and with an analysis of the results, the best models that characterize the data studied are retained. A rejection of the total power of ECB sources in the Kolmogorov-Smirnov test indicates that the use of separate source powers would be the best choice in characterizing ECB source data. The number of failures is well modeled with the double Weibull function, followed by the duration of failures and the power lost by these failures respectively using the Erlang function and the Exponormal function.

These models will enable us to determine predictive models for the number and duration of outages in the Togo and Benin transmission networks. This will make the network increasingly intelligent. The question that may arise is the accuracy of the reports, which can influence data quality. This work, although interesting for the security of the entire network, does not allow to locate the breakdowns of the transmission network. The continuation of this work would focus on modeling the data by taking into account the topology of the ECB network.

ACKNOWLEDGMENT

This study was carried out with the technical support of CERME and the World Bank's financial support.

NOMENCLATURE

Abbreviations

ECB	Electrical Community of Benin
PDF	Probability Density Function
CDF	Cumulative Density Function
AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
SSE	Sum of Squared Errors
RMSE	Root Mean Square Error
CERME	Centre d'Excellence Régional pour la Maîtrise de l'Electricité
SCADA	Supervisory Control and Data Acquisition
PMU	Phasor Measurement Unit

REFERENCES

1. T. Falentina and B. P. Resosudarmo, 'The impact of blackouts on the performance of micro and small enterprises: Evidence from Indonesia', World Development, vol. 124, p. 104635, 2019, https://doi.org/10.1016/j.worlddev.2019.104635.
2. Tunisian Electricity and Gas Company, 'Final report of the commission of inquiry into the blackout of August 31, 2014', Ministry of Industry, Energy and Mines, Tunisia, Technical Report, 2014.
3. N. Anandan, S. Sivanesan, S. Rama, and T. Bhuvaneswari, 'Wide area monitoring system for an electrical grid', Energy Procedia, vol. 160, pp. 381-388, 2019, https://doi.org/10.1016/j.egypro.2019.02.171.
4. K. Yamashita, J. Li, P. Zhang, and C.-C. Liu, 'Analysis and control of major blackout events', in 2009 IEEE/PES Power Systems Conference and Exposition, IEEE, 2009, pp. 1-4, https://doi.org/10.1109/PSCE.2009.4840091.
5. H. Mirshekali, R. Dashti, A. Keshavarz, and H. R. Shaker, 'Machine Learning-Based Fault Location for Smart Distribution Networks Equipped with Micro-PMU', Sensors, vol. 22, no. 3, p. 945, 2022, https://doi.org/10.3390/s22030945.
6. M. Adeen and F. Milano, 'On the negative correlation of stochastic voltage dependent loads', Electric Power Systems Research, vol. 235, p. 110663, 2024, https://doi.org/10.1016/j.epsr.2024.110663.
7. D. Renga et al., 'Data-driven exploratory models of an electric distribution network for fault prediction and diagnosis', Computing, vol. 102, no. 5, pp. 1199-1211, 2020. https://doi.org/10.1007/s00607-019-00781-w.
8. K. Demertzis et al., 'Communication network standards for smart grid infrastructures', Network, vol. 1, no. 2, pp. 132-145, 2021, https://doi.org/10.3390/network1020009.
9. H. Haes Alhelou, M. E. Hamedani-Golshan, T. C. Njenda, and P. Siano, 'A survey on power system blackout and cascading events: Research motivations and challenges', Energies, vol. 12, no. 4, p. 682, 2019, https://doi.org/10.3390/en12040682.
10. M. barate, E. T. G. Palanga, A. S. A. Ajavon, and K. M. Kodjo, 'Statistical analysis of the impact of power grid outages on sources and load: a case study of the power grid of the

Benin Electric Community', IJAR, vol. 11, no. 07, pp. 984–1000, Jul. 2023, https://doi.org/10.21474/IJAR01/17308.
11. N. Taimoor et al., 'Power outage estimation: The study of revenue-led top affected states of US', IEEE Access, vol. 8, pp. 223271-223286, 2020, https://doi.org/10.1109/ACCESS.2020.3043630.
12. H. Saleh, A. A. E.-A. Aly, and S. Abdel-Hady, 'Assessment of different methods used to estimate Weibull distribution parameters for wind speed in Zafarana wind farm, Suez Gulf, Egypt', Energy, vol. 44, no. 1, pp. 710-719, 2012, https://doi.org/10.1016/j.energy.2012.05.021.
13. A. A. Zúñiga, J. F. Fernandes, and P. J. Branco, 'Fuzzy-Based Failure Modes, Effects, and Criticality Analysis Applied to Cyber-Power Grids', Energies, vol. 16, no. 8, p. 3346, 2023, https://doi.org/10.3390/en16083346.
14. U. Shahzad, 'Probabilistic security assessment in power transmission systems: a review', Journal of Electrical Engineering, Electronics, Control and Computer Science, vol. 7, no. 4, pp. 25-32, 2021.
15. E. Ciapessoni, D. Cirio, S. Massucco, A. Pitto, and F. Silvestro, 'A Probabilistic Risk Assessment Approach to support the operation of large electric power systems', in 2009 IEEE/PES Power Systems Conference and Exposition, IEEE, 2009, pp. 1-8, https://doi.org/10.1109/PSCE.2009.4840260.
16. S. Liu, X. Shi, T. Wang, Y. Zhang, and Y. Cao, 'Power grid risk assessment method based on risk probability engineering and its application', in 2016 IEEE International Conference on Power System Technology (POWERCON), IEEE, 2016, pp. 1-5, https://doi.org/10.1109/POWERCON.2016.7753980.
17. X. Liu, M. Shahidehpour, Y. Cao, Z. Li, and W. Tian, 'Risk assessment in extreme events considering the reliability of protection systems', IEEE Transactions on Smart Grid, vol. 6, no. 2, pp. 1073-1081, 2015, https://doi.org/10.1109/TSG.2015.2393254.
18. R. Rocchetta and E. Patelli, 'A post-contingency power flow emulator for generalized probabilistic risks assessment of power grids', Reliability Engineering & System Safety, vol. 197, p. 106817, 2020, https://doi.org/10.1016/j.ress.2020.106817.
19. A. Nawaz, H. Wang, H. Yang, H. Armghan, and J. Gao, 'Risk-constrained probabilistic coordination in coupled transmission and distribution system', Electric Power Systems Research, vol. 228, p. 110005, 2024, https://doi.org/10.1016/j.epsr.2023.110005.
20. N. A. Musakkir, N. Sunusi, and S. A. Thamrin, 'Stochastic model of the annual maximum rainfall series using probability distributions', Malaysian Journal of Fundamental and Applied Sciences, vol. 19, no. 5, pp. 827-839, 2023, https://doi.org/10.11113/mjfas.v19n5.2945.
21. A. Guenoupkati, A. A. Salami, M. K. Kodjo, and K. Napo, 'Statistical Characterization of Electric Power Production and Importation: Case Study of Benin Electricity Community (CEB)', in 2019 II International Conference on High Technology for Sustainable Development (HiTech), IEEE, 2019, pp. 1-6, https://doi.org/10.1109/HiTech48507.2019.9128240.
22. R. Kollu, S. R. Rayapudi, S. V. L. Narasimham, and K. M. Pakkurthi, 'Mixture probability distribution functions to model wind speed distributions', International Journal of energy and environmental engineering, vol. 3, pp. 1-10, 2012, https://doi.org/10.1186/2251-6832-3-27.
23. G. W. Irwin, W. Monteith, and W. C. Beattie, 'Statistical electricity demand modelling from consumer billing data', in IEE Proceedings C (Generation, Transmission and Distribution), IET, 1986, pp. 328-335, https://doi.org/10.1049/ip-c.1986.0048.
24. J. Munkhammar, J. Widén, and J. Rydén, 'On a probability distribution model combining household power consumption, electric vehicle home-charging and photovoltaic power production', Applied Energy, vol. 142, pp. 135-143, 2015, https://doi.org/10.1016/j.apenergy.2014.12.031.

25. R. Singh, B. C. Pal, and R. A. Jabr, 'Statistical Representation of Distribution System Loads Using Gaussian Mixture Model', IEEE Transactions on Power Systems, vol. 25, no. 1, pp. 29-37, Feb. 2010.
26. D. H. McQueen, P. R. Hyland, and S. J. Watson, 'Monte Carlo simulation of residential electricity demand for forecasting maximum demand on distribution networks', IEEE Transactions on power systems, vol. 19, no. 3, pp. 1685-1689, 2004, https://doi.org/10.1109/TPWRS.2004.826800.
27. A. dos Santos and M. C. de Barros, 'Stochastic modeling of power system faults', Electric Power Systems Research, vol. 126, pp. 29-37, 2015, https://doi.org/10.1016/j.epsr.2015.04.015.
28. T. Ahmad and N. Senroy, 'Statistical characterization of PMU error for robust WAMS based analytics', IEEE Transactions on Power Systems, vol. 35, no. 2, pp. 920-928, 2019, https://doi.org/10.1109/TPWRS.2019.2939098.
29. F. S. Kebede, J.-C. Olivier, S. Bourguet, and M. Machmoum, 'Reliability evaluation of renewable power systems through distribution network power outage modelling', Energies, vol. 14, no. 11, p. 3225, 2021, https://doi.org/10.3390/en14113225.
30. Y. Bokovi, C. Adjamagbo, A. A. Salami, and A. S. A. Ajavon, 'Comparative Study of the Voltage Stability of an Hight Voltage Power Grid: Case of the Power Grid of the Electric Community of Benin', Science Journal of Energy Engineering, vol. 8, no. 2, pp. 15-24, 2020, https://doi.org/10.11648/j.sjee.20200802.11.

Paper submitted: 23.02.2024
Paper revised: 13.10.2024
Paper accepted: 13.10.2024