

Weibull mélange pour l'analyse de la fiabilité

Dr. BENAÏCHA Halima

Université d'Oran2

Résumé

Dans cet article, Le modèle mélange de Weibull est proposé pour modéliser les temps de défaillance d'un système électrique en exploitation avec différents modes de défaillance et analyser la fiabilité. L'estimation des paramètres de la loi de fiabilité se fait par la méthode du maximum de vraisemblance. Basé sur des données de durée de vie des transformateurs de puissance de la Société nationale de l'électricité et du gaz (SONELGAZ), nous avons analysé et comparé la performance de notre modèle avec la distribution classique de Weibull à deux paramètres. En toute rigueur, Il est nécessaire de vérifier lequel des deux modèles ajustent correctement les données. Pour résoudre ce problème, une solution possible consiste à comparer les modèles en utilisant le critère d'information d'Akaike.

Mots-clés : Modèle Weibull; Mélanges de deux Weibull; Estimation des paramètres; Algorithme de la Vraisemblance; Analyse de la fiabilité; AIC.

I. Introduction

Depuis la première tentative d'analyse d'un modèle de mélange par Pearson[1], l'étude des mélanges de lois est devenue un domaine à part entière de la statistique moderne. De nombreux ouvrages de références existent sur le sujet, le plus récent est Mc Lachlan et Peel [2], faisant un "état de l'art" des différentes approches développées jusqu'à maintenant. Pour un point de vue détaillé des techniques

d'analyses non-bayésiennes, on pourra se référer à Titterton et al. [3], ainsi qu'à McLachlan et Basford [4].

Les mélanges de distributions ont fourni un ensemble d'approches mathématiques pour la modélisation statistique d'une grande variété de phénomènes aléatoires.

Vue leur utilité comme méthodes de modélisation flexibles, les modèles de mélange ont continué à avoir un succès et un intérêt croissant du point de vue pratique et théorique.

Les domaines dans lesquels les modèles de mélange ont été expliqués avec succès comprennent l'astronomie, la biologie, la génétique, la médecine, la psychiatrie, l'économie, l'ingénierie et le marketing.

Les mélanges de lois constituent une démarche puissante (une grande diversité de distributions peuvent être modélisées sous cette forme) et parcimonieuse (la distribution ainsi modélisée est décrite par un nombre réduit de paramètres) [5,6,7]. Par ailleurs, la composition de la distribution comme un mélange peut souvent être exploitée pour identifier une structure cachée dans le jeu de données, c.à.d. de déterminer des sous-populations intéressantes du point de vue de l'interprétation du jeu de données.

L'hétérogénéité dans un contexte d'exploitation distribuée se manifeste par une mixture de probabilités de défaillance $f_j(t)$ résultant du comportement de l'équipement en exploitation sur chaque environnement j . Lorsque les mêmes équipements sont exploités sur un nombre m d'environnements distincts (organisation, climatique, conditions de production) où le processus d'usure est susceptible d'être modifié par les variables d'exploitation, alors les données des durées de vie sont une mixture (mélange). Le modèle

résultant de la mixture est un mélange pondéré de probabilités de défaillance [8]. En conséquent, le modèle est une somme de lois de défaillances $f_j(t)$ pondérées par un poids p_j représentant le pourcentage des données des durées de vie provenant de chaque environnement j . En conséquent, le modèle est une somme de lois de défaillances $f_j(t)$ pondérées par un poids p_j représentant le pourcentage des données des durées de vie provenant de chaque environnement j [9,10].

La suite de cet article s'organise de la façon suivante.

Tout d'abord, nous présentons la modélisation des temps de défaillances par le modèle de Weibull à deux paramètres et modèle de Weibull mélange à 5 paramètres. Nous discutons dans la section 3, l'estimation des paramètres de ces lois effectuée en utilisant la méthode du maximum de vraisemblance. Un exemple illustratif décrivant l'analyse de la fiabilité des transformateurs de puissance, en comparant les performances du modèle simple à 2 paramètres à celle d'un modèle à 5 paramètres, basé sur un mélange additif de deux lois de Weibull. dans la section 4. La section 6 conclut l'article.

II. Modèle probabiliste des durées de vie d'un équipement

II.1 Modèle de Weibull à deux paramètres

Le principe de la distribution de Weibull est associé à l'étude de la probabilité de défaillance ou de succès. C'est la plus populaire des lois, utilisée dans plusieurs domaines (électronique, mécanique,...). Elle caractérise le comportement du système dans les trois phases de vie : période de jeunesse, période de vie utile et période d'usure ou vieillissement [11].

Dans sa forme la plus générale, la distribution de Weibull dépend de ses deux paramètres suivants : α et β . La densité de probabilité d'une loi de Weibull a pour expression [12]:

$$f(t; \psi) = \frac{\beta}{\alpha} \left(\frac{t}{\alpha}\right)^{\beta-1} e^{-\left(\frac{t}{\alpha}\right)^\beta} \quad (1)$$

Où $t \geq 0, \alpha > 0, \beta > 0$.

Le paramètre de forme β (sans unité) qui peut orienter un diagnostic, sa valeur étant caractéristique de certains modes de défaillance.

Le paramètre d'échelle α (unité de temps) qui, comme son nom l'indique, nous renseigne sur l'étendue de la distribution. La fonction fiabilité s'écrit :

$$R(t; \psi) = e^{-\left(\frac{t}{\alpha}\right)^\beta} \quad (2)$$

II.2 Mélanges de deux Weibull à cinq paramètres

La densité de probabilité d'une loi de mélanges de deux Weibull a pour expression [13].

$$f(t) = pf_1(t) + (1-p)f_2(t) \quad (3)$$

$$f(t; \Theta) = p \frac{\beta_1}{\alpha_1} \left(\frac{t}{\alpha_1}\right)^{\beta_1-1} e^{-\left(\frac{t}{\alpha_1}\right)^{\beta_1}} + (1-p) \frac{\beta_2}{\alpha_2} \left(\frac{t}{\alpha_2}\right)^{\beta_2-1} e^{-\left(\frac{t}{\alpha_2}\right)^{\beta_2}} \quad (4)$$

Où $\Theta = (p, \psi_1, \psi_2)$

avec $\psi_i = (\alpha_i, \beta_i) \quad i=1,2 \quad \alpha_1 > 0, \beta_1 > 0, p > 0$

La fonction fiabilité s'écrit:

$$R(t) = pe^{-\left(\frac{t}{\alpha_1}\right)^{\beta_1}} + (1-p)e^{-\left(\frac{t}{\alpha_2}\right)^{\beta_2}} \quad (5)$$

II. Méthode du Maximum de vraisemblance

L'idée fondamentale de l'estimation par maximum de vraisemblance est, comme son nom l'indique, de trouver un ensemble d'estimations des paramètres, pour que la vraisemblance de l'échantillon utilisé, soit maximum.

III.1 Maximum de vraisemblance de la loi de Weibull à deux paramètres

L'une des méthodes les plus importantes d'estimation des paramètres d'une fonction de densité de probabilité est la méthode de maximum de vraisemblance. Le principe de la méthode est de maximiser la fonction de la densité de probabilité pour une fonction de densité de probabilité [14,15,16].

Partant d'une série $\{t_j : j=1,2,\dots,n\}$ d'une variable aléatoire positive T qui représente les temps de défaillance d'un système.

La loi de vraisemblance des données relativement au modèle de paramètre Ψ s'écrit :

$$L(t_1, t_2 \dots t_n; \psi) = \prod_{j=1}^n f(t_j; \psi) \quad (6)$$

Il est plus facile de maximiser la log-vraisemblance au lieu de la fonction de vraisemblance elle-même. Si ψ maximise $\ln(L(\psi))$, alors il maximise également $L(\psi)$. Cela est dû à la monotonie de la fonction logarithme. Son logarithme est [17]:

$$L(t_1, t_2 \dots t_n ; \psi) = \sum_{j=1}^n (f(t_j; \psi)) \quad (7)$$

Où $f(t_j; \psi)$ est donnée en (3.1) avec $\psi = (\alpha; \beta)$

La log-vraisemblance vaut :

$$\ln(L(t; \psi)) = n \ln \beta - n \beta \ln \alpha + (\beta - 1) \sum_{j=1}^n \ln(t_j) - \alpha^{-\beta} \sum_{j=1}^n \ln(t_j)^\beta \quad (8)$$

Une dérivée partielle nulle étant un moyen éprouvé pour trouver un extremum, on pose l'équation de vraisemblance et on dérive la fonction par rapport au paramètre à estimer [18,19].

$$\frac{dL(t; \psi)}{d\alpha} = \frac{-n\beta}{\alpha} + \frac{\beta}{\alpha^{\beta+1}} \sum_{j=1}^n \ln(t_j) = 0 \quad (9)$$

$$\frac{dL(t; \psi)}{d\beta} = \frac{n}{\beta} - n \ln \alpha + \sum_{j=1}^n \left(\frac{t_j}{\alpha} \right)^\beta \ln \frac{t_j}{\alpha} = 0 \quad (10)$$

En remplaçant α par son expression dans l'Équation (9), nous obtenons :

$$\alpha = \left(\frac{1}{n} \sum_{j=1}^n (t_j)^\beta \right)^{1/\beta} \quad (11)$$

$$\frac{dL(t; \psi)}{d\beta} = \frac{1}{\beta} + \frac{1}{n} \sum_{j=1}^n \ln(t_j) \cdot 1(12) - \frac{\sum_{j=1}^n (t_j)^\beta \ln(t_j)}{\sum_{j=1}^n (t_j)^\beta} = 0$$

Pour estimer le paramètre β , nous utilisons la méthode de Newton-Raphson avec une bonne condition initiale $\beta^{(0)}$ [20,21].

Les étapes de ce procédé sont les suivantes:

$$\beta^{(k+1)} = \beta^k - \frac{g(\beta)}{g'(\beta)} \quad (13)$$

Où

$$g(\beta) = \frac{\beta}{n} \sum_{j=1}^n \ln(t_j) \sum_{j=1}^n (t_j)^\beta - \beta \sum_{j=1}^n (t_j)^\beta \ln(t_j) + \sum_{j=1}^n (t_j)^\beta \quad (14)$$

$$g'(\beta) = \frac{1}{n} \sum_{j=1}^n \ln(t_j) \left(\sum_{j=1}^n (t_j)^\beta + \beta \sum_{j=1}^n (t_j)^\beta \ln(t_j) \right) - \beta \sum_{j=1}^n (t_j)^\beta \ln^2(t_j) \quad (15)$$

III.2 Maximum de vraisemblance de mélange de deux Weibull

L'estimation des paramètres de Weibull mélange se fait par la méthode du maximum de vraisemblance.

Partant d'une série $\{t_j : j=1,2,\dots,n\}$ de temps défaillance d'une variable aléatoire T relative à la durée de fonctionnement avant défaillance issue de densité de probabilité du modèle Weibull mélange.

La dérivée de la fonction log-vraisemblance par rapport à chacun des cinq paramètres est la suivante [22] :

$$\frac{dL(t; \psi)}{d\alpha_1} = \sum_{j=1}^n \frac{p}{f(t_j)} \frac{df_1}{d\alpha_1} \quad (16)$$

$$\frac{dL(t; \psi)}{d\beta_1} = \sum_{j=1}^n \frac{p}{f(t_j)} \frac{df_1}{d\beta_1} \quad (17)$$

$$\frac{dL(t; \psi)}{d\alpha_2} = \sum_{j=1}^n \frac{(1-p)}{f(t_j)} \frac{df_2}{d\alpha_2} \quad (18)$$

$$\frac{dL(t; \psi)}{d\beta_2} = \sum_{j=1}^n \frac{(1-p)}{f(t_j)} \frac{df_2}{d\beta_2} \quad (19)$$

$$\frac{dL(t; \psi)}{dp} = \sum_{j=1}^n \frac{f_1(t_j; \psi_1) - f_2(t_j; \psi_2)}{pf_1(t_j; \psi_1) + (1-p)f_2(t_j; \psi_2)} \quad (20)$$

La suite des opérations est bien sûr l'annulation des dérivées partielles.

Si l'on considère souvent la méthode du maximum de vraisemblance comme une méthode pratique et efficace, il existe d'autres méthodes d'estimation qui peuvent être employées.

IV.Exemple illustratif

Afin de se rapprocher le plus possible de la réalité, nous considérons des données de défaillances de 109 transformateurs de puissance de moyenne tension 30kV et 10 kV de puissance 160 MVA à 630 MVA. Ces données sont fournies par la Société de Distribution de l'Electricité et du Gaz de l'Ouest (SONELGAZ) au cours de la période 2008 à 2013.

Les états du transformateur de puissance sont enregistrés depuis sa première mise en service jusqu'à sa panne. La figure 1 schématise l'âge de la panne de chaque équipement de l'échantillon.

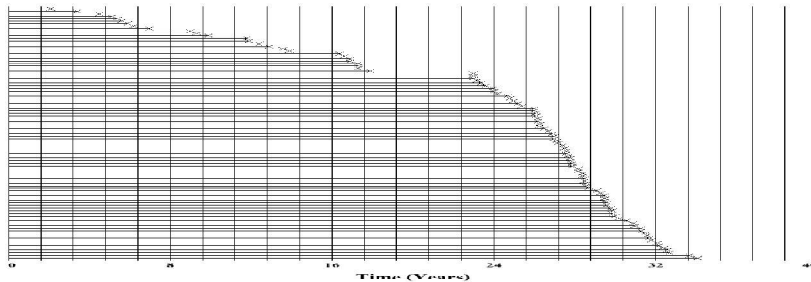


FIG.1 –Données statistiques durant un intervalle de temps (2008–2013)

IV.1.Estimation des lois de fiabilité

IV.1.1loi de weibull

Selon la fonction de distribution cumulative illustrée à la figure2, les données ne sont pas alignées ; ils forment plutôt une courbe. Nous pouvons conclure que la distribution de Weibull à 2- paramètres n'est pas en conformité avec ses données.

Cependant, la modélisation par le modèle de Weibull de la durée de vie des transformateurs de puissance n'est pas fine. Ainsi, d'autres modèles devraient être étudiés.

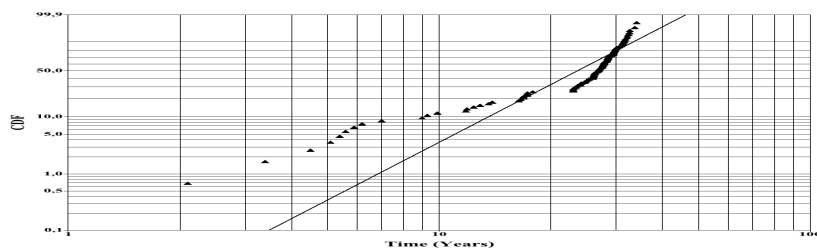


FIG. 2 – la fonction de distribution cumulative de Weibull à 2-paramètres

Nous estimons les paramètres de la loi de Weibull en maximisant la vraisemblance du modèle.

On trouve :

Le paramètre d'échelle α (Années) = 26.9 ;

le paramètre de forme β = 2.38 ;

Le paramètre de forme est une indication de l'âge des transformateurs de puissance.

le temps moyen du bon fonctionnement : $MTTF(\text{Années}) = 23.5$

IV.1.2. Mélange de deux Weibull

Dans la figure 3, nous montrons la courbe de mélange de deux lois de Weibull qui adapte beaucoup mieux les données. En effet, elle est capable de prendre en compte le mélange, la concurrence ainsi que le changement de mode de défaillance ; en fonction de l'âge de fonctionnement de l'équipement contrairement au modèle de Weibull à deux paramètres. On parle alors de l'hétérogénéité des données des durées de vie.

L'aplatissement ultérieur du début de la courbe reflète la performance de la sous-population. Par contre, la fonction ajustée commence à courber lorsque la sous-population plus vulnérable est épuisée par des défaillances sur le terrain.

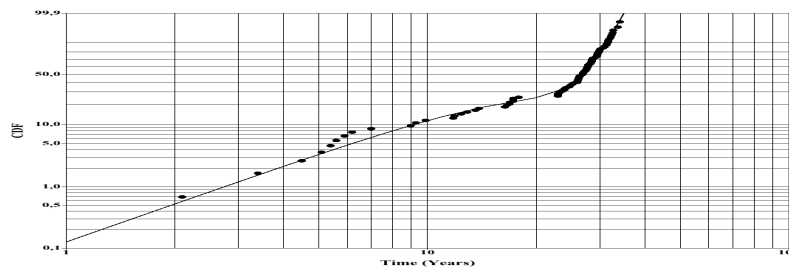


FIG.3 – la fonction de distribution cumulative de mélange de deux Weibull

Le tableau 1, présente les estimations du maximum de vraisemblance à cinq paramètres basé sur le mélange de deux lois de Weibull. Deux paramètres pour chaque composant de distribution de Weibull et une proportion p_j représentant le pourcentage des données des durées de vie provenant de chaque environnement j .

Selon le paramètre de forme et les modes de défaillances des transformateurs de puissance. On a 22% de la population échoue en

raison des défaillances précoces ou de jeunesse, (exemple erreurs de fabrication) avec un MTTF égal à 11,03 années et 78% de la population défaille en raison du deuxième mode de défaillance avec un MTTF égal à 28,04 ans.

Méthode	p	α (année)	β	MTTF(année)
Sous	0.22	12.56	1.75	11.03
Sous	0.78	27.75	10.15	28.04

Tableau1 - Paramètres obtenue par la distribution mélange de Weibull

En se basant sur le tableau 5.4, la fonction fiabilité des transformateurs de puissance peut s'écrire alors comme suit :

$$R(t) = 0.22 \exp \left[- \left(\frac{t}{12.56} \right)^{1.75} \right] + 0.78 \exp \left[- \left(\frac{t}{27.75} \right)^{10.15} \right]$$

La formule de la fiabilité $R(t)$ peut être utilisée pour déterminer la périodicité de maintenance préventive optimale pour les transformateurs de puissance.

Le Tableau 2, représente les différentes périodicités de maintenance préventive basée sur les valeurs de la fiabilité calculées par les deux lois respectives.

Fiabilité	Périodicité de maintenance preventive (Année)	
	Loi de Weibull	Loi de mélange de
90%	9,43	10,46
70%	16,28	20,27
50%	23,09	25,52

Tableau 2- Périodicités de maintenance préventive basée sur la Fiabilité

Les résultats du tableau 2 et la figure 4 montrent que la distribution mélange de Weibull tend à surestimer la fiabilité des transformateurs de puissance donc à surestimer la périodicité de maintenance préventive ce qui réduit les coûts de maintenance et la probabilité de défaillance.

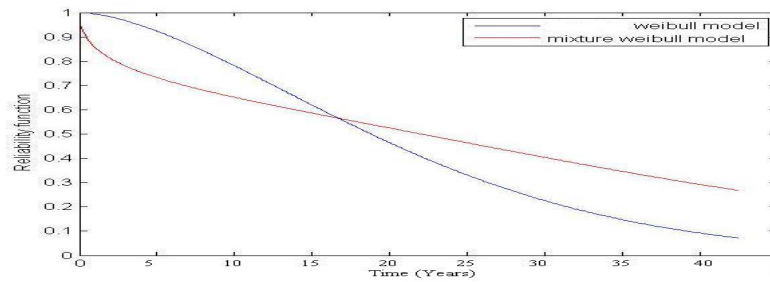


FIG.4 – Représentation de la fiabilité des deux modèles

En toute rigueur, Il est nécessaire de vérifier lequel des deux modèles ajustent correctement les données. Pour résoudre ce problème, une solution possible consiste à comparer les modèles en utilisant le critère d'information d'Akaike [23].

$$AIC = -2Ln \left(L(t, \hat{\psi}) \right) + 2k$$

où L est la vraisemblance maximisée et k le nombre de paramètres dans le modèle. Avec ce critère, la déviance du modèle ($-2 * \log(L)$) est pénalisée par 2 fois le nombre de paramètres. L'AIC représente donc un compromis entre le biais (qui diminue avec le nombre de paramètres) et la parcimonie (nécessité de décrire les données avec le plus petit nombre de paramètres possible).

Le meilleur modèle est celui possédant l'AIC le plus faible (Tableau 5.6).

Critère d'information	Modèle Weibull	Loi de mélange de Weibull
AIC	721,42	707,97

Tableau 3 : -Résultat de l'AIC

Conclusion

Le modèle mélange de Weibull est proposé pour modéliser les données de temps de défaillance des transformateurs d'un système électriques avec différents modes de défaillance et d'analyser la fiabilité. Le maximum de vraisemblance a été utilisé pour modéliser le système à l'étude. Basé sur des données de durée de vie des transformateurs de puissance de la Société nationale de l'électricité et du gaz (SONELGAZ), nous avons analysé et comparé la performance des deux modèles respectifs . L'utilisation de la fonction ajustée montre que la distribution mélange de Weibull est plus approprié pour la modélisation des transformateurs de puissance. Le résultat de l'analyse montre aussi que l'utilisation du mélange de Weibull surestime la fiabilité. En effet, elle est en mesure de spécifier les différents modes de défaillance. Notre analyse montre que 22% des transformateurs ont défailit avant 13 ans. En outre, la loi de weibull n'est pas en mesure de déterminer les défaillances précoce du transformateur .Cependant, une maintenance appropriée peut minimiser ces défaillances précoces et d'étendre de manière significative la durée de vie du transformateur.

Bibliographie

- [1]. McLachlan G. et Peel D. (2000). Finite mixture models. Wiley-Interscience, New York.
- [2]. Titterton D. M., Smith A. F. M. et Makov U. E. Statistical analysis of finite mixture distributions. John Wiley & Sons Ltd, Chichester.
- [3]. McLachlan G. J. et Basford K. E. (1988). Mixture models, volume 84 of Statistics : Textbooks and Monographs. Marcel Dekker Inc
- [4]. DNP. Murthy, M. Xie, R. Jiang , Weibull Models, (John Wiley and Sons, Hoboken, New Jersey, 2004).
- [5]. P. Chapouille, Fiabilité et maintenabilité. Les techniques de l'ingénieur ; traité de l'entreprise industrielle, Edition Eyrolles1999.
- [6]. Pauline Beaumont, Optimisation des plans d'essais accélérés Application à la tenue en fatigue de pièces métalliques de liaison au sol, Thèse de Doctorat, Université de Nantes Angers Le Mans.
- [7]. P. Bruneau, M. Gelgon , F. Picarougne , Parsimonious reduction of gaussian mixture models with a variational-bayes approach. Pattern Recognition, 43(3):850–858, 2010.
- [8]. P. Bruneau, M. Gelgon , F. Picarougne , Aggregation of probabilistic pca mixtures with a variational-bayes technique over parameters. In 20th International Conference on Pattern Recognition, p. 702–705, 2010.
- [9]. F. Nielsen, V. Garcia, R. Nock , Gaussian mixture models via entropic quantization. In: European Signal Processing Conference (Eusipco'2009), p. 2012–2016, 2009.
- [10]. J.F. Lawless, Statistical Models and Methods for Lifetime Data, second ed., Wiley, 2002.
- [11]. T. Hasumi, T. Akimoto, and Y. Aizawa, The Weibull-Log Weibull transition of the interoccurrence time statistics in the two-dimensional Burridge-Knopoff earthquake model,Physica A, vol. 388, pp. 483-490, 2009.
- [12]. D. N. P. Murthy, M. Xie, and R. Jiang, Weibull Models. New York: Wiley, 2004.
- [13]. Titterton DM, Smith AFM, Makov UE. Statistical analysis of finite mixture distribution. Chichester: Wiley; 1985.

- [14]. R. Jiang, M.J. Zuo, H.X. Li, Weibull and inverse Weibull mixture models allowing negative weights, *Reliability Engineering & System Safety*, vol. 66 n. 3, December 1999, pp. 227 – 234.
- [15]. Kececioglu D. *Reliability and Life Testing Handbook*. Vol. 1, Englewoods Cliffs, N.J., Prentice Hall, 1993, 917 p.
- [16]. D. Cousineau, Fitting the three-parameter weibull distribution: review and evaluation of existing and new methods, *Dielectrics and Electrical Insulation*, IEEE Transactions on, vol. 16 n. 1, August 2009, pp. 281 – 288.
- [17]. P. Norman, A. Archer, computational technique for maximum likelihood estimation with weibull models, *IEEE Transactions on Reliability*, vol. 29 n. 1, April 1980, pp. 57-62.
- [18]. S.Y. Jiang, D. Kececioglu, Maximum Likelihood Estimation from Censored-Data Mixed Weibull Distribution, *IEEE Transactions on Reliability*, vol. 41 n. 2, June 1992, pp. 248 – 255.
- [19]. Yincai Tang, Xiaoling Wei, Existence of maximum likelihood estimation for three-parameter log-normal distribution, *IEEE Reliability, Maintainability and safety, ICRMS 2009. 8th international conference*, 2009, pp. 305 – 307.
- [20]. S. W. Ng, Y. S. Lee, Variable Dimension Newton–Raphson Method, *IEEE Transactions on circuits and systems*, Vol. 47, N. 6, June 2000, pp. 809 – 817.
- [21]. M.R. Herving & al, Newton-Raphson algorithm for load-flow calculation in transmission and distribution networks, *Generation, Transmission and Distribution*, IEE Proceedings, vol. 134 n. 5, September 1987, pp. 325-330.
- [22]. Donald A. Mumford, Robust Parameter Estimation for the Mixed Weibull (seven Parameter) Including the Method of Maximum Likelihood and the Method of Minimum Distance, Thesis, Air Force Institute of Technology, 1997.
- [23]. Akaike, H., 1974. A new look at statistical model identification. *IEEE Transactions on Automatic Control* AU-19: 716-722.