IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received October 27, 2020, accepted November 4, 2020, date of publication November 9, 2020, date of current version November 19, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3036819

Machine Learning and q-Weibull Applied to Reliability Analysis in Hydropower Sector

EDILSON MACHADO DE ASSIS[1], CELSO LUIZ SANTIAGO FIGUEIRÔA FILHO[2], GABRIEL ALVES DA COSTA LIMA[1], LUIS AUGUSTO NAGASAKI COSTA[1], AND GISELE MARIA DE OLIVEIRA SALLES[3]

[1]Fundação para o Desenvolvimento de Bauru, São Paulo 17033-360, Brazil
[2]Escola Politécnica, Programa de Engenharia Industrial (PEI) Universidade Federal da Bahia–UFBA, Salvador 40210-630, Brazil
[3]Paranaense Company of Energy (COPEL), Curitiba 80420-170, Brazil

Corresponding author: Edilson Machado de Assis (edilsonassis@gmail.com)

This work was supported by Companhia Paranaense de Eletricidade (COPEL) under Project PD-06491-0391-2015.

ABSTRACT Brushes are critical components in power generation equipment. The interruptions in brush operation because they have failed in their service can cause financial losses avoidable by proper maintenance planning. Therefore, this article aims to offer a methodology for estimating the reliability parameter of brushes used in hydroelectric generators through machine learning concepts and through a statistical distribution compatible with complex phenomena. The method uses six selection patterns by graphical plotting of lifetimes and brush lengths to recognize problems of recording wear information. Then, the information is separated into three data sets according to the failure mode. The brushes reliability prediction uses an artificial neural network with assisted learning to predict a cumulative distribution function based on the operating time extracted from the equipment's hour meter. The method compares the statistical models q-Weibull, Weibull, q-exponential, and exponential with the prediction function. Three measures of goodness of fit were calculated, the logarithm of likelihood, coefficient of determination, and mean squared error. Most of the values found point to an advantage in the use of artificial neural networks over the use of the q-Weibull distribution. The method compares density and failure rate functions. The application of artificial neural networks in reliability analysis can have a significant impact on reducing maintenance costs, as it leads to results closer to reality. This article presents artificial neural networks for the first time compared to a distribution based on non-extensive statistical mechanics in the context of hydroelectric brushes.

INDEX TERMS Artificial neural network, brushes, q-Weibull, reliability, statistical distribution.

I. INTRODUCTION

ONE of the main devices designed to transfer current between stationary and moving parts of machines and electrical equipment is the electrical sliding contact. Heating of the parts in contact, sparks, and arcs are common in the transmission of electric current by sliding contact. The mechanical vibration of the sliding contacts and the wear lead to an unbalanced distribution of the electric current between the contacts (brushes), overcurrents and thermal overloads result in decreased reliability of the contact device with the brushes [1]. Excessive brush wear is a problem in many hydroelectric plants. Cleaning and replacing brushes leads to increased maintenance costs [2].

The associate editor coordinating the review of this manuscript and approving it for publication was Xiao-Sheng Si.

Sliding electrical contacts are present in some critical engine and generator components. Electric brushes are generally used in these contacts to conduct current between the stationary part and the moving part of the motor. Mechanical and electrical wear are the main factors influencing on brush wear. Temperature, material properties, sliding speed, contact force, and interfacial and environmental conditions also influence brush wear. The mechanical wear of the brushes is proportional to the brush spring pressure and the sliding speed, and the electrical wear of the brushes is associated with the current drop and contact voltage [3].

Brush wear can cause malfunctions in generators and DC electric motors. The effects of polarity and current density on the wear rate of electric brushes are well delineated since [4]. The relationship between drop in contact voltage and friction wear for sliding electrical contacts in DC motors when high currents flow into the brush was discussed by [5], specifically, how to avoid burning the brush when the high current flows for the brush.

The prediction of the service life of DC motors through brush wear can be made from the understanding of friction and wear due to the mechanical and electrical contact between a brush and a commutator. The operating voltage showed a negligible influence on the wear behavior while the electric current and the temperature have a high contribution [6].

Wear of copper-graphite brushes in small brush-type DC motors was analyzed in [3]. Variable electrical current conditions were obtained by changing the brush spring pressure and sliding speed. The results were resistance to electrical contact, voltage drop, increased brush surface temperature, and so on. The brush wear has been greatly altered by the electric current, which indicates that the high current itself not only produces more heating but also causes an increase in voltage drop.

According to [7], temperature and pressure significantly affect the wear performance of copper-graphite composite brushes. An accelerated wear test was performed to assess life characteristics. The probability of failure up to the specified time according to the lognormal distribution, reliability, and analysis of life characteristics were calculated from time to failure data using the temperature-nonthermal-accelerated life-stress model. Flaws obtained in the accelerated wear test were extrapolated to the condition of normal use.

The brushes supply current to the rotor windings in turbogenerators. Individual brushes and brush groups are installed on removable cassettes. Pressure control on the slip ring is necessary to regulate the brush operation. Otherwise, after some time, there will be an overload on some brushes and the unloaded condition on the others. This can increase brush wear, overheat, and damage the contact pairs and the entire cassette [1]. Several brushes are connected in parallel to obtain the current required for operation. Even so, the current can be distributed unevenly between the brushes. This can lead to an overload of some brushes and also to uneven wear [2].

In microscopic terms, the structure of the surfaces involved is important to understand the performance of the sliding contacts. The surface structures were analyzed by confocal microscopy, with special emphasis on the components of the anisotropic structure. The anisotropy factor can be linked to the type of microscopic surface, as well as electrical noise in a sliding system model [8].

A new type of electrical contact material is metal fiber. Generally, the metal fiber brush is made up of many metal fibers aligned along the length direction. Each metal fiber can be treated as a small independent brush [9] and this suggests the phenomenon of scale invariance that is common in complex situations.

The two interactions of competition and cooperation between brushes together ([1], [2]), the influence of the surface structure involved in wear [8], and the apparent scale invariance mentioned in [9] seem to suggest that there is a complex behavior in the brush wear.

Complex systems are made up of components that work with numerous interdependence or subordination relationships, presenting a difficult understanding. In these systems, chaos theory does not apply, nor does it presuppose the occurrence of Emergency, which is a phenomenon or process of forming complex patterns based on a multiplicity of simple interactions as occurs in social (social networks), biological systems (animal colonies) and physical (climate).

Complex systems usually have long-range spatial interactions, or long-term memory, or cooperation/competition effects, as seen in [10].

Failure of a component may have many (recent or not) multiple and interacting causes, some of them acting on a cooperative and others on a conflictive basis, so it is not surprising that complex behavior may appear. If this happens, power-law-like expressions are expected to substitute exponential in the statistical description [11].

Figure 1 shows the methodology of this article. The input data from the maintenance sector make up a matrix of brush lengths and time. These values are submitted to a consistency test and, if approved, will compose the three scenarios, all failures, wear failures, and non-wear failures. The rest of the input data is isolated from the analysis. A second pattern recognition test checks whether the replacement was performed for wear or not. Subsequently, the method performs in each scenario, three sets of modeling, Artificial Neural Networks, q-Weibull, and a package with three widely used models (Weibull, q-exponential, and exponential). The last step performs the comparison of the five models through log-likelihood, mean squared error, coefficient of determination, and Akaike information criterion (AIC) and decides on the acceptance of artificial neural network modeling.

II. RELATED WORKS

Artificial neural networks (ANNs) have been considered promising tools for analyzing reliability data since the 1990s. Identifying the appropriate probability distribution for describing a data set and estimating parameters were necessary steps in reliability studies. Chi-square tests or suitability tests were usually part of the distribution identification step along with the interpretation of the failure data histograms. These procedures may be inaccurate on small samples. The graphical approach to a fitting could be analyzed as a pattern recognition problem and parameter estimation would be a classification problem. Both cases are solved by neural networks. In [12], a neural network was used to classify distributions as normal, uniform, exponential, Weibull, and another neural network was used to calculate the parameters.

Another form of reliability forecasting has gained expressiveness in the software area. The existing models were based on assumptions about development environments, the nature of the failures, and the probability of individual failures occurring. Since all of these assumptions must be made before the project starts and since many projects are unique,

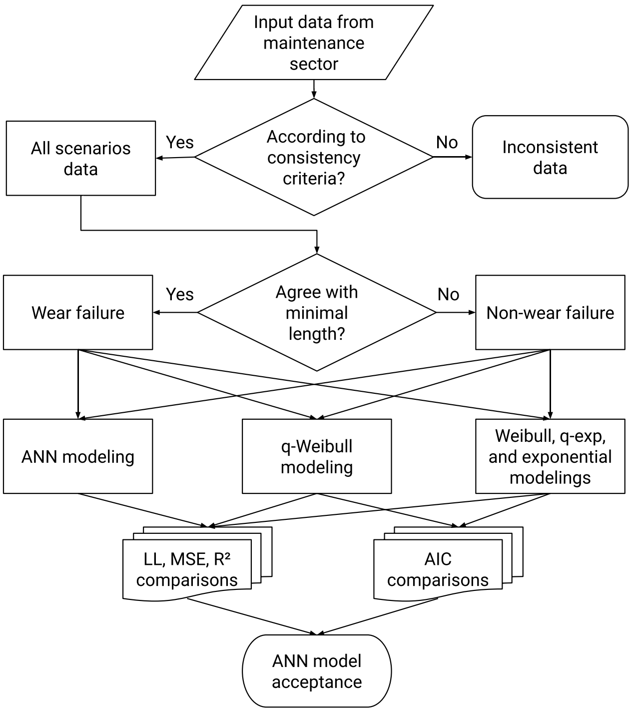


FIGURE 1. Flowchart of the methodology of the article.

failures could be predicted based on the failures of similar projects. These models should predict whether reliability has grown enough to guarantee product launch. Reference [13] used testing and debugging data to predict the number of faults at the end of each day through a neural network.

Reference [14] proposed the training of four ANNs to directly estimate some important measures of reliability analysis. The mean time to failure (MTTF), the mean time to repair (MTTR), the unavailability rate, and the estimated time for scheduled maintenance were calculated. The four ANNs were powered by a data distributor and received only specific data for their analysis. The analyzes were carried out separately.

In the early 2000s, [15] developed a simulation-based probabilistic neural network model to estimate the probability of failure of aged pipes vulnerable to corrosion. The approach consisted of calculating pipeline reliability in an adaptive connectionist representation instead of using a simulation-based probabilistic analysis structure. This ANN model used eight tube parameters as input variables. The output variable was the probability of failure. The proposed method is an evolution of the probabilistic neural network (PNN), originally proposed by [16], is generic and can be applied to several decision problems related to the maintenance of old engineering systems.

Reference [17] validated the identification of the most significant indicators to represent the reliability of a hydropower plant and also the method to represent the efficiency of the plant by a physical model. The multi-criteria decision-making ensemble methods made the selection of the most important alternative to be considered. Thus, an index represented as a weighted function of multiple independent indicators with the potential to change the plant's utilization factor represented its reliability.

The time series modeling technique using ANNs offers a promising alternative for predicting failures and reliability. Neural network modeling via feed-forward Multilayer Perceptron (MLP) can suffer from long computing time. The radial basis function (RBF) neural network architecture is considered a viable alternative due to its shorter training time. In RBF neural network, the neuron's output decreases as the input is moved away from its centroid, at a rate determined by the radius. According to [18], the RBF's ability to recognize whether an entry is close to the training set or is in an untrained region of the entry space gives the RBF a significant advantage over the MLP structure. Further comparison studies between feed-forward MLP and RBF architecture demonstrated that modeling via MLP had a lower error and better performance than the RBF model for cycle time prediction under convoluted multivariate data sets and predicting new product success before market entry [19], [20].

An application of ANNs, addressing Feed-Forward, Back-Propagation to make performance predictions of the hydroelectric plant on the Himreen dam-Diyala lake was developed by [21]. The method used data obtained during research over a 10 year period in terms of net turbine head, a flow rate of water, and power production to predict plant performance. The correlation coefficient (R) between the predicted and observed production variables was higher than 0.96.

This article proposes an ANN following the MLP structure (see Figure 3). Backpropagation is the algorithm used for training and the network output is a proposed cumulative distribution function. This allows the calculation of all the reliability measures necessary for our analyzes, in addition to the comparison with results obtained from widely used probability distributions. Two generalizations of statistical distributions based on non-extensive statistical mechanics were compared with the ANN model

III. RELIABILITY MODELING

Technological advances have produced increasingly complicated systems, with high installation costs and the possibility of generating large financial losses or environmental damage if they do not work as designed (see [22]).

Maintenance focused on reliability has become imperative to keep the system functioning properly and to guarantee the quality of its products over time.

A set of organized data, with reliable origin and representative nature, should generate useful information for improving the system, making it more efficient, more robust, and less costly.

The consequences of incorrect design or inefficient maintenance can negatively affect safety, the environment, or the cost in many categories of industrial processes [23].

In general, the objective of maintenance is to provide increased availability, either of the component or the system as a whole.

Reliability analysis often uses the Weibull distribution, which is a simple and powerful empirical model. For more details see [24]. The probability density function (pdf) at time t, where $t < T$ and T is the time to failure, is given by:

$$f(t) = \frac{\beta}{\eta - t_0}\left(\frac{t-t_0}{\eta - t_0}\right)^{\beta-1} \exp\left[-\left(\frac{t-t_0}{\eta-t_0}\right)^{\beta}\right], \quad (1)$$

where $\beta > 0$, $\eta > t_0$, $t \geq t_0$ and $\int_0^{\infty} f(x)dx = 1$. The exponential distribution is a particular case of the Eq.(1) when $\beta = 1$.

Several generalizations of the Weibull model have been proposed: linear and nonlinear transformations of time, the use of multiple distributions, parameters as a function of time, and stochastic models, among others. In [25] there are several of these model proposals.

As seen in [26] and [27] there is no clarity as to the originality of the probability density functions used in reliability. According to [28], the use of the stretched exponential (Weibull) is registered prior to the article by [24] in a work by Kohlrausch describing the capacitor discharge.

Almost all proposals for generalizing the Weibull model have an exponential structure, be it simply exponential, nested exponentials, or exponentials of various functions. However, a generalization based on non-extensive statistical mechanics has been employed because of its flexibility.

The statistical mechanics of simple systems has a well-established structure through exponential probability distributions, for example, Boltzmann weight and Maxwellian distribution. These exponential distributions are derived from the Boltzmann-Gibbs-Shannon (BGS) entropy. The theoretical basis of the statistical description of complex systems is still the subject of research, however, there is much evidence that points in the direction of non-extensive statistical mechanics.

A. STATISTICAL NON-EXTENSIVE MODELING

A generalization of the concept of entropy in terms of a parameter q was introduced by [29]:

$$S_q = k\frac{\sum_i^W p_i^q - 1}{1-q}, \quad (2)$$

where k is a positive constant that provides dimensional consistency to the expression, the Boltzmann constant, p_i is the probability of occurrence of the i-th microstate and W is the total number of microstates in the system. This expression recovers the Boltzmann-Gibbs-Shannon (BGS) entropy, $S_1 = -k\sum_i^W p_i \ln p_i$, when $q \to 1$.

Non-extensive statistical mechanics induced generalizations in other fields, for example in mathematics. Two functions usually appear at the beginning of the formalism, the $\exp_q(x)$ and the $\ln_q(x)$, called q-exponential and q-logarithm respectively. Both functions have a control parameter q in addition to the argument x.

Non-extensive statistical mechanics introduces generalizations of the exponential and logarithm functions, using a q dimensionless parameter, these functions are called q-exponential and q-logarithm, defined by [30]:

$$\exp_q(x) = \begin{cases} [1+(1-q)x]^{\frac{1}{1-q}}, & \text{if } [1+(1-q)x] > 0 \\ 0, & \text{otherwise,} \end{cases} \quad (3)$$

and

$$\ln_q x = \frac{x^{1-q}-1}{1-q} \quad (x > 0,\ q \neq 1), \quad (4)$$

where $x, q \in \mathbb{R}$. These functions are inverse of each other, that is, $\exp_q(\ln_q x) = \ln_q(\exp_q x) = x$, and the usual functions ($\exp(x)$ and $\ln(x)$) are recovered at the limit $q \to 1$. The functions still satisfy $\ln_q 1 = 0$ and $\exp_q 0 = 1$, $\forall q$. For certain values of the parameters the q-exponential presents a cross-over between an exponential behavior and a power-law regime ($\exp_q(-ax)$ with $a > 0$ and $q > 1$ is asymptotically a power-law for large x, leading to fat-tailed distributions). Other properties of these q-functions can be found in [31]. These generalized functions have been applied in several areas in addition to statistical mechanics, for example, in mathematics and statistics. Properties of q-exponential can be found in [32].

B. q-WEIBULL MODEL

The q-Weibull model is obtained from the three-parameter Weibull model (Eq. (1)) by replacing the exponential function with the q-exponential. This model has been applied in several areas of knowledge because of its flexibility ([33]–[35]). In [36] more details about this replacement are shown. According to this model, t_0 is the location or minimum life parameter, that is, it is a minimum time value such that failures below this are not expected. The probability density function of q-Weibull is defined by:

$$f_q(t) = (2-q)\frac{\beta}{\eta - t_0}\left(\frac{t-t_0}{\eta-t_0}\right)^{\beta-1} \exp_q\left[-\left(\frac{t-t_0}{\eta-t_0}\right)^{\beta}\right], \quad (5)$$

where $\beta > 0$, $\eta - t_0 > 0$ and $t - t_0 \geq 0$. The factor $(2-q)$ and the restriction $q < 2$ are necessary to guarantee the normalization of $f_q(t)$. Making the limit $q \to 1$, the q-Weibull probability density function becomes Weibull. The q-exponential distribution is a particular case of Eq. 1 when $\beta = 1$. The value $\eta - t_0$ is known as θ, the scale parameter of the distribution.

The reliability function of q-Weibull is defined as:

$$\begin{aligned} R_q(t) &= \int_t^{\infty} f_q(t')dt' \\ &= \left[1-(1-q)\left(\frac{t-t_0}{\eta-t_0}\right)^{\beta}\right]_+^{\frac{2-q}{1-q}} \\ &= \left\{\exp_q\left[-\left(\frac{t-t_0}{\eta-t_0}\right)^{\beta}\right]\right\}^{2-q}. \end{aligned} \quad (6)$$

The cumulative function $F_q(t)$ is the complement of the reliability function and is defined by:

$$F_q(t) = 1 - \left\{ \exp_q \left[-\left(\frac{t - t_0}{\eta - t_0} \right)^\beta \right] \right\}^{2-q}. \tag{7}$$

The failure rate function is:

$$h_q(t) \equiv \frac{f_q(t)}{R_q(t)}, \tag{8}$$

and written as:

$$\begin{aligned} h_q(t) &= \frac{(2-q)\beta}{\eta - t_0} \left(\frac{t - t_0}{\eta - t_0} \right)^{\beta - 1} \left[1 - (1-q) \left(\frac{t - t_0}{\eta - t_0} \right)^\beta \right]_+^{-1} \\ &= \frac{(2-q)\beta}{\eta - t_0} \left(\frac{t - t_0}{\eta - t_0} \right)^{\beta - 1} \left[\exp_q \left[-\left(\frac{t - t_0}{\eta - t_0} \right)^\beta \right] \right]^{q-1}. \end{aligned} \tag{9}$$

Taking into account that the q-logarithm (see Eq. (4)) is the inverse function of the q-exponential, the Eq. (7) is written as $y = \beta x + b$, with $y = \ln \left\{ -\ln_{\frac{1}{2-q}} \left[1 - F_q(t) \right] \right\}$, $x = \ln(t - t_0)$, and $b = -\beta \ln \left[\frac{\theta}{(2-q)^{\frac{1}{\beta}}} \right]$.

Sample data are the time to failure arranged in ascending order and non-reliability values are estimated using Bernard's median rank approximation provided in [37]

$$\hat{F}_i = \frac{i - 0.3}{n + 0.4}, \tag{10}$$

where n is the sample size, i is the failure order number ranging from 1 to n. The median rank is the probability of failure for each experiment when n of these statistically independent experiments are carried out, the probability of i or more failures occurring is 50%, that is, it is a solution of $P(X \geq n) = 0.5$ in the variable p within the context of a variable of binomial distribution $X \sim B(n, p)$.

For each time t_i we have

$$x_i = \ln(t_i - t_0), \tag{11}$$

and

$$y_i = \ln \left[-\ln_{\frac{1}{2-q}} (1 - \hat{F}_i) \right]. \tag{12}$$

The parameter of Eq. (7) are estimated by maximizing the coefficient of determination R^2,

$$R^2 = 1 - \frac{\sum_{i=1}^{n} [y_i - \widehat{y_i}]^2}{\sum_{i=1}^{n} [y_i - \overline{y_i}]^2}, \tag{13}$$

$$\text{subject to:} \begin{cases} \beta > 0 \\ \theta > 0 \\ t_0 < t_{min} \\ q < 2, \end{cases} \tag{14}$$

where $\hat{y}_i$ ln $\left\{ -\ln_{\frac{1}{2-q}} \left[1 - F_q(t_i) \right] \right\}$, $\overline{y_i}$ is $\frac{\sum y_i}{n}$, and t_{min} is the lowest sampling time.

Equation 13 returns $R^2 \leq 1$, including negative values. Note that the parameters of the Weibull distribution can be determined by maximizing R^2 with the additional constraint $q = 1$. As additional criteria for the quality of fit, the mean squared error, MSE $= \sum \left[F_q(t_i) - \hat{F}_i \right]^2 / n$, the likelihood $\mathcal{L} = \sum \log \left[f_q(t_i) \right]$, and Akaike Information Criterion (AIC) were calculated. In general, the lower the MSE, the higher the quality of the adjustment. Lower values of $\mathcal{L}$, indicate better fittings qualities. The AIC [38] is an index that compares models with different number of parameters. AIC requires a bias correction for small number of points [39] . This index is expressed by $AIC = n \ln \left(\frac{RSS}{n} \right) + 2k + \frac{2K(K+1)}{n-K-1}$, where n is the number of data points (x_i, y_i), RSS is the residual sum of squares, and K is the number of parameters of the model. The best model is supposed to be that one with the lowest AIC. The value $\Delta AIC_i \equiv AIC_i - \min [AIC_i]$ may be used to compare the models, the best one has $\Delta AIC = 0$.

IV. MACHINE LEARNING

Artificial intelligence has been a topic of public and private interest for decades. Since the 1950s, there was great hope that classic artificial intelligence techniques based on logic, knowledge representation, reasoning, and planning would result in revolutionary software that could, among other things, understand language, control robots, and provide expert advice. Although the advances based on such techniques will be real in the future, many researchers have begun to doubt these classic approaches, choosing to focus their efforts on the design of systems based on statistical techniques and machine learning [40].

Machine learning (ML) and the intelligent systems that emerged from it, such as search engines, recommendation platforms, and speech and image recognition software, have become an indispensable part of modern society. Rooted in statistics and relying heavily on the efficiency of numerical algorithms, machine learning techniques take advantage of the world's most powerful computing platforms and the availability of immensely large data sets. Also, as the fruits of their efforts become so easily accessible to the public through various modalities - such as the cloud, interest in machine learning continues to increase dramatically, generating more social, economic, and scientific impacts. One of the pillars of machine learning is mathematical optimization, which, in this context, involves the numerical calculation of parameters for a system designed to make decisions based on data not yet seen. That is, based on the data currently available, these parameters are chosen to be optimal concerning a particular learning problem. The success of certain optimization methods for machine learning has inspired a large number of researchers in various communities to face even more challenging machine learning problems and to design new methods that are widely applicable [40].

Machine learning has a huge variety of possible applications. Image recognition, automatic image caption generation, document text analysis, recommendation systems (the link between users and products) are just a few examples [41].

Using machine learning requires several different skills. One is the necessary programming skill. The other skills have to do with the development and implementation of an appropriate model [42]. The domain of models requires knowledge of specific mathematical tools [43].

Several programming languages can be used in machine learning, each has its advantages and disadvantages over others (see [44]–[48] for details). Currently, hardware and technology have evolved in a specialized way to meet the implementation of learning machines [49], [50].

Machine learning is about making computer programs automatically improve with experience. Machine learning is one of the most successful technical fields today. It lies at the intersection of computer science and statistics and the core of artificial intelligence and data science. Recent progress in machine learning has been driven by the development of new algorithms and the continued explosion in the availability of online data and low-cost computing. In virtually all of science, technology, and commerce you can see the adoption of data-intensive machine learning methods [41].

The machine learning (ML) approach requires the definition of a flexible program whose behavior is determined by several parameters. Then, a data set is used to determine the best possible set of parameters. These values improve the performance of our program in relation to some measure of performance. The program is called a model, once the parameters have been set. The set of all different programs (input and output mappings) that can be produced only by manipulating the parameters is called a model family. And the program that uses the data set to choose parameters is called a learning algorithm. It is necessary to define the problem precisely before addressing the learning algorithm. identify the exact nature of the inputs and outputs and choose an appropriate model family [51].

In this case, our model will receive a numeric vector of times until failure and will output another vector that will represent probabilities.

If the right model family is chosen, there must be a parameter setting, so that the model returns a value close to the probability of failure when it receives a time until failure.

The training process can take place as follows [52]:

1. Start the program with random parameters

2. Analyze the program outputs by comparing them with the expected values

3. Adjust the parameters to reduce the outputs closer to the expected values

4. Repeat steps 2 and 3 until the error is below the desired

Learning can be supervised or unsupervised. The supervised learning algorithm is a function that takes as input a set of input data and output results and generates a function that is the model learned. One of the simplest supervised learning tasks is regression. Unsupervised learning is carried out without comparison with the expected result [51].

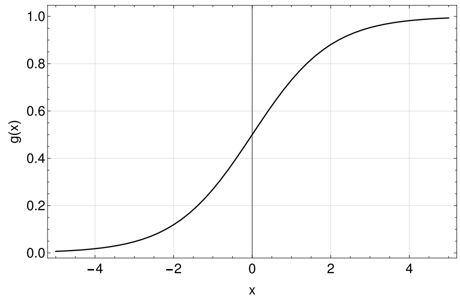


FIGURE 2. Sigmoid $g(x) = 1/(1 + \exp(-x))$.

A. NEURAL NETWORKS

ANNs are computational implementations of the neuronal structure of human brains. The brain contains neurons that are like organic switches. They can change their output state, depending on their electrochemical input. The neural network in the brain is strongly interconnected through neurons. The output of one neuron can be the input for thousands of other neurons in these networks. Learning occurs through the repetitive activation of certain neural connections. This reinforces such connections making them more likely to produce the desired result, given a specified input. This learning involves feedback - when the desired outcome occurs, the neural connections that cause that outcome are strengthened.

Artificial neural networks try to simplify and mimic brain behavior. They can be trained in a supervised or unsupervised manner. In a supervised ANN, the network is trained in providing samples of corresponding input and output data, with the intention of having the ANN provide the desired output for a given input. Learning takes place by adjusting the weights of ANN connections. Unsupervised learning in an ANN is an attempt to get ANN to understand a little about the behavior of the input data even though there are no outputs provided for comparison.

The biological neuron is simulated in an ANN by an activation function. A function commonly used in ANN is the sigmoide $g(x) = 1/(1 + \exp(-x))$, as shown in Figure 2.

Note that there is a similarity between the format of the activation function and a typical accumulated distribution function such as that represented by Figure 11. It is worth remembering that, each node represents only part of the network modeling process and it is not necessary that the ANN output function behaves the same as the activation function.

The function is activated, that is, it moves from 0 to 1 when input x is greater than a certain value, in this case, 0. The analytical formulation of the derivative of the activation function is important for the ANN training algorithm. This is a feedback algorithm and is called backpropagation. The sigmoid function does not change from 0 to 1 instantly as a step function. This means that the derivative of this function is continuous.

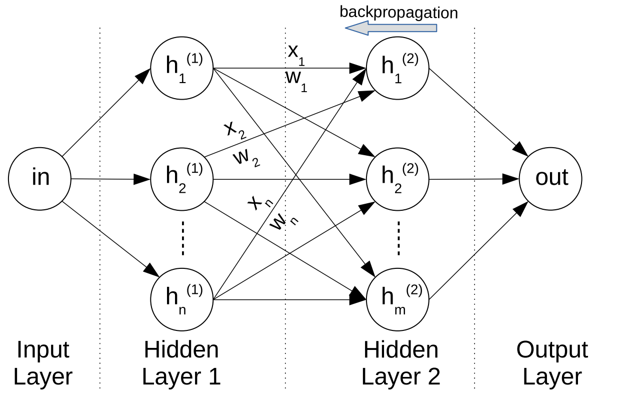


FIGURE 3. Example of artificial neural network.

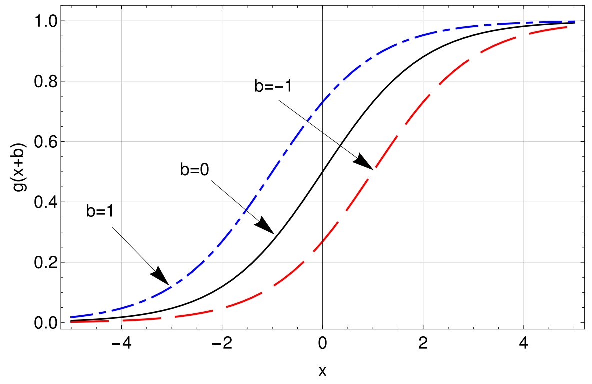


FIGURE 5. Sigmoids with different bias values (*b*).

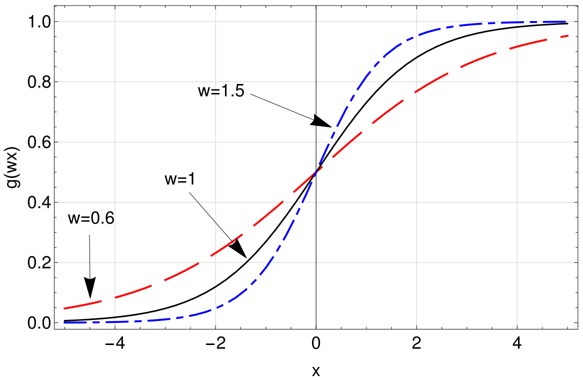


FIGURE 4. Sigmoids plotted with different values of weights (*w*).

ANNs can be represented as connected layers of nodes. Each node receives several weighted inputs, applies the activation function to the sum of these inputs, and generates an output. Figure 3 represents the ANN arrangement used.

The circles in Figure 3 represent the nodes and are also called perceptrons. The activation function is located in the node and takes the x values of inputs weighted by the w values, adds them up and inserts them in the activation function. The output of the activation function is shown as h in Figure 3. The activation function $h_1^{(2)}$ corresponds to the first function of the second Hidden Layer and can be written as:

$$h_1^{(2)} = g\left(x_1w_1 + x_2w_2 + \cdots + x_nw_n + b_1^{(2)}\right). \quad (15)$$

The weights modify the inclination of the activation function, as shown in Figure 4. When $w > 1$ the activation function becomes steeper, whereas for values $w < 1$ the slope becomes smaller.

Although Figure 3 does not show, besides the real number entries x_i and w_i with $i = 1, 2, \cdots, n$, there is a real bias entry (b) in each node. The bias values shift the entire curve to the left if they are negative or to the right, if they are positive (see Figure 5).

The goal of supervised learning is to reduce the error between the desired input and output. Ideal for supervised learning is to provide many known data input and output data pairs and vary the weights and bias based on these samples so that error expression is minimized. This optimization process is performed with a gradient (derivatives). Backpropagation is processed in this way.

V. METHOD

The maintenance sectors of a set of hydroelectric plants sent data in a tabular form containing approximately 7500 rows. This information concerns two hydroelectric power generation plants (PCH and UHE). The lengths of the turbine brushes and the respective hour meter values were noted for the various brush positions. Each pair of these values define a point in Figure 6.

The information collection process is not automated and inconsistent values can be reported. A set of six validation criteria was established to filter the observed inconsistencies (See Figure 6).

- **Too accelerated wear**. The brush length information shows a decreasing trend with an angle above the maximum allowed.
- **Too slowed wear**. The lengths of a brush form a very small slope.
- **Negative operating time**. There are negative time interval values. This is probably due to errors in the annotation of the hour meter.
- **The install length is greater than the maximum**. The annotation shows a brush length that exceeds the limit and would make installation impossible.
- **The Install length is less than the minimum**. The length of the newly installed brush is less than the minimum allowed.
- **Inconsistent length increase**. The annotation shows an increase in brush length over some time.

Classification is a technique that aims to identify a cluster for a given data set. In this way, depending on the value of the target or output attribute, the entire data set can be qualified to belong to a class. This technique helps to identify the behavior patterns of the data. This is, in short, a discrimination mechanism [53], [54]

The application of filters based on the geometric arrangement of plotted data has been used in learning machines. These filters allow you to quickly validate the data set.

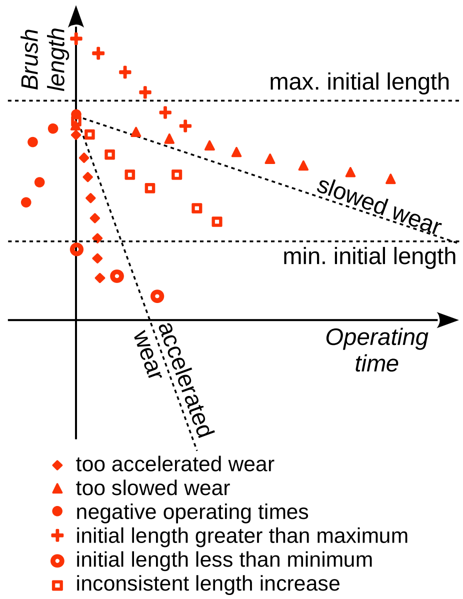


FIGURE 6. Validation criteria.

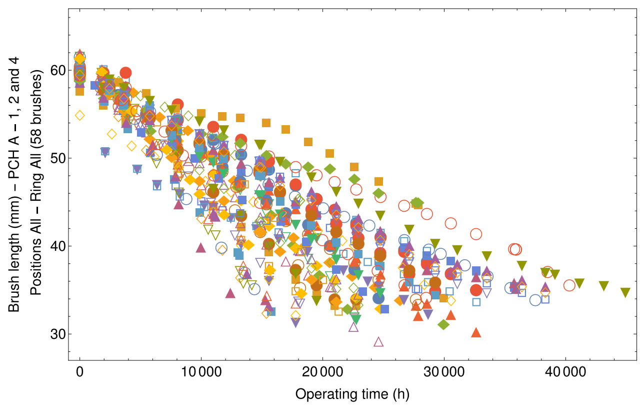


FIGURE 7. Brush lengths and operation times of PCH A usine for all failure modes.

The lengths of the brushes in all positions of units 1, 2, and 4 of the PCH A plant are plotted in Figure 7. Each series of values corresponds to a brush. Plotting was performed with different markers only for better visual comfort.

The values of the last readings of each brush length were used to infer which failure mode occurred. Brushes that have been replaced by wear remain in operation until a minimum length is reached. If the last length reading reached a value close to this minimum, it is assumed that the replacement occurred by wear failure mode, otherwise, the failure mode is characterized as non-wear. Figure 8 shows the readings of the 58 brushes replaced by wear failure mode.

The time values (abscissa axis) corresponding to the last readings of each brush shown in Figure 8 are the lifetime for the reliability analysis. These times until failure are ordered in ascending order and then each median rank is calculated according to Eq. (10). Each ordered pair is plotted in Fig. 9. The curve values were calculated as a supervised learning

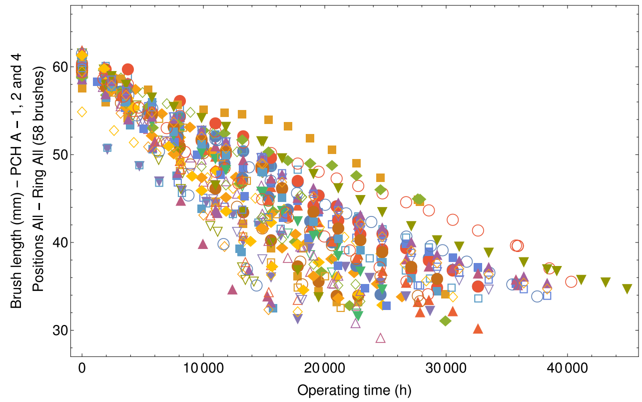


FIGURE 8. Brush lengths and operation times of PCH A usine for wear failure mode.

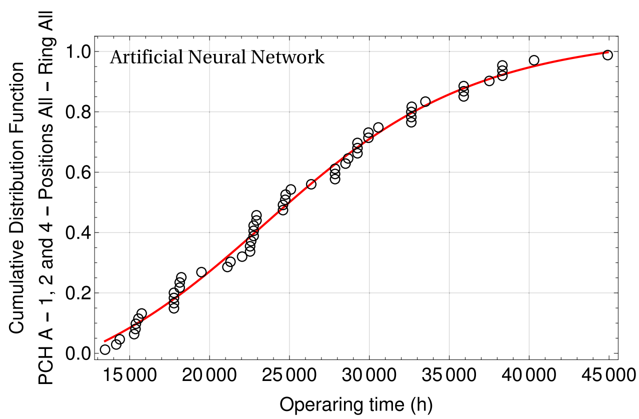


FIGURE 9. Time to failure (circles) and cumulative distribution function (line) for PCH A usine and wear failure mode.

neural network. All data were used to train the neural network to allow a comparison with the adjustment of the statistical distribution. A large number of values were used as ANN test data. These values are the abscissa of the cumulative distribution function curves. The adopted method trains ANN with all the lifetimes. This will allow a direct comparison between the application of ANN with q-Weibull distribution fitness. Overfitting disturbances can be visually identified in the graph if they occur.

Some precautions were added because the model represents a cumulative distribution function. The function predicted by the artificial neural network must always be increasing, this will guarantee non-negative values for its derivative (probability density function). The function must also return non-negative values and less than or equal to 1. Once these conditions are ensured, the values returned by the artificial neural network can be interpreted as failure probabilities up to a specified time.

The same time-to-failure values were used to fit a q-Weibull distribution. This model is represented by a straight line whenever the variable changes described in the equations Eq. (11) and Eq. (12) are made. This property is not verified for the neural network model.

The Figure 10 shows black circles representing the times until failure, the continuous line representing the adjusted

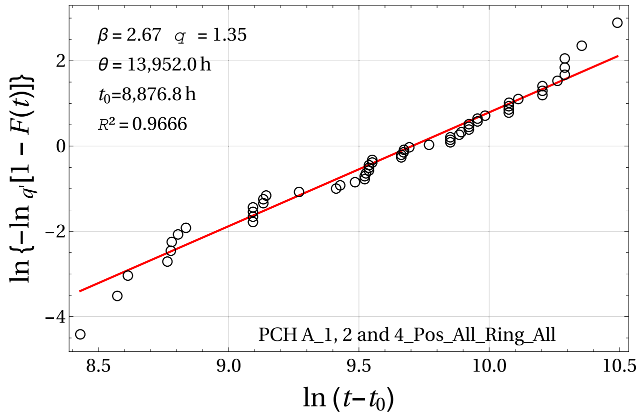


FIGURE 10. ***q*-Weibull fitness of the PCH A usine time to failure data (wear failure mode).**

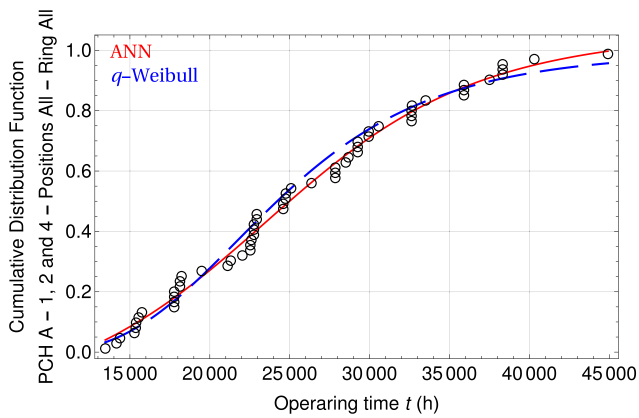


FIGURE 11. **Artificial neural network and *q*-Weibull cumulative distribution functions of the PCH A usine (wear failure mode).**

model, the distribution parameters, and the value of the determination coefficient (R^2). The coefficient of determination of the adjusted q-Weibull distribution is $R^2 = 0.9666$. This means that the quality of this fitting is very good.

The coefficient of determination for the artificial neural network model can be calculated by:

$$R^2_{ANN} = 1 - \frac{\sum_{i=1}^{n}\left[\hat{F}_i - \widehat{F}_{ANN}\right]^2}{\sum_{i=1}^{n}\left[\hat{F}_i - \bar{\hat{F}}_i\right]^2}, \tag{16}$$

where n is the number of lifetimes, $\hat{F}_i$ is the i-th median rank value obtained from Eq. (10) and F_{ANN} is the value of cumulative distribution function calculated according to the machine learning method for the i-th lifetime. The coefficient of determination for the machine learning model is $R^2_{ANN} = 0.9921$. This value is higher than that found in the q-Weibull model. In fact, the curve of the machine learning model is closer to the sample data as shown in 11.

The likelihood and mean squared error values for the q-Weibull model are $\mathcal{L}_{qW} = -600.82$ and $\mathrm{MSE}_{qW} = 0.00131$, and for the ANN model they are $\mathcal{L}_{ANN} = -598.01$ e $\mathrm{MSE}_{ANN} = 0.00065$ (see Table 3). These likelihood values

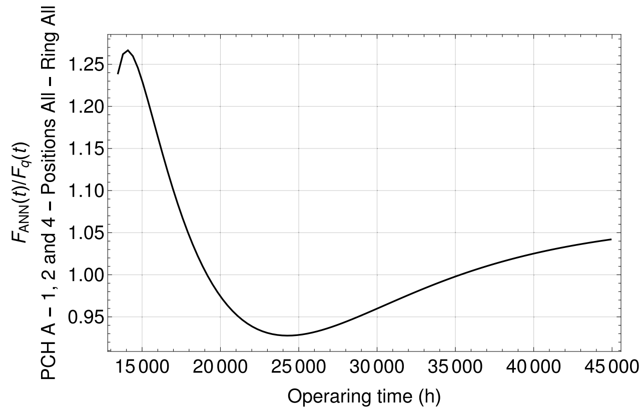


FIGURE 12. **Artificial neural network cumulative distribution functions values divided by *q*-Weibull cumulative distribution function of the PCH A usine (wear failure mode).**

are very close and differ by approximately 0.5%, with little advantage for the ANN model. Although the mean square error values are small, the ANN model is about 50% smaller than the q-Weibull model.

In addition to the q-Weibull model, three other statistical distributions were adjusted, Weibull, exponential, and q-exponential (Table 1). The likelihood values are all close with little advantage for the q-exponential model. All models were calculated by maximizing R^2. The q-Weibull model presents the highest R^2 among the statistical distributions (7% higher than the exponential model). The value MSE_{qW} corresponds to 46% of the second-best fit according to this criterion. AIC is the only criterion in which the q-Weibull model is in second place, losing to the q-exponential model. Note that AIC takes into account the computational effort required by the model to handle a larger number of parameters.

The comparative curve of the values of failure probabilities up to time t of the two models was plotted with the expression $\alpha(t) = F_{ANN}(t)/F_q(t)$ (See Figure 12). The results show that $1 \leq \alpha(t) \leq 1.4$ for times below 19,000h. The values of $\alpha(t)$ are between 1 and 0.93 when $19,000\text{h} \leq t \leq 35,500\text{h}$. The values are slightly increasing from 35,500h reaching 1.04 in the final times. Note that the biggest differences were found for the initial time instants.

The time derivative of the failure probability function $F_{ANN}(t)$ is equivalent to the probability density function ($f_{ANN}(t)$). The probability density function of the q-Weibull model is expressed in Eq. (5). The curves for both functions present very similar shapes as shown in Figure 13.

The results of $f_{ANN}(t)/f_q(t)$ show that the probability density function of the q-Weibull distribution returns values greater than the ANN model only within the range of values from $t = 15,300\text{h}$ to $t = 26,000\text{h}$. Figure 14 shows that curve $f_{ANN}(t)/f_q(t)$ assumes values between 0.83 and 1.55 and the minimum being approximately at $t = 19,500\text{h}$

The previous analyzes show proximity between the values of the functions obtained by the artificial neural network and by the q-Weibull distribution (see Figure 15). However, the

TABLE 1. The parameters and goodness of fits measures of the *q*-Weibull, Weibull, *q*-exponential, and exponential distributions.

Model	β	$\theta(h)$	$t_0(h)$	q	Shape	$\mathcal{L}$	R^2	MSE	AIC	ΔAIC
q-Weibull	2.69	1.40×10^4	8,785	1.35	Unimodal	-600.82	0.9666	0.00131	-144.89	35.94
Weibull	1.53	1.57×10^4	11,305	1	Increasing	-600.5	0.9347	0.00243	-129.11	51.72
q-exponential	1	1.32×10^5	13,212	-2.9	Increasing	-599.82	0.9585	0.00808	-180.82	0
exponential	1	1.76×10^4	13,000	1	Constant	-608.75	0.9050	0.00756	-109.58	71.24

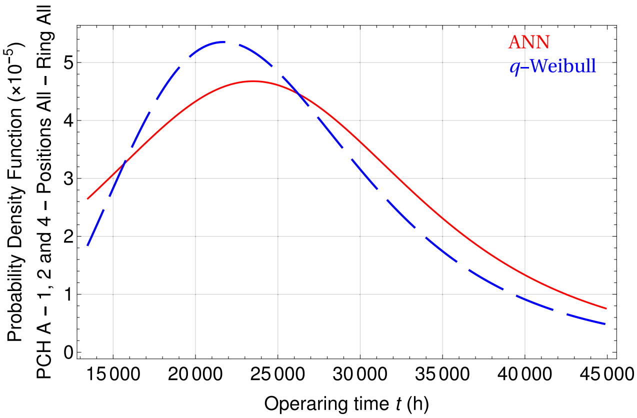


FIGURE 13. Artificial neural network (solid line, red online) and *q*-Weibull probability density (dashed line, blue online) functions of the PCH A usine (wear failure mode).

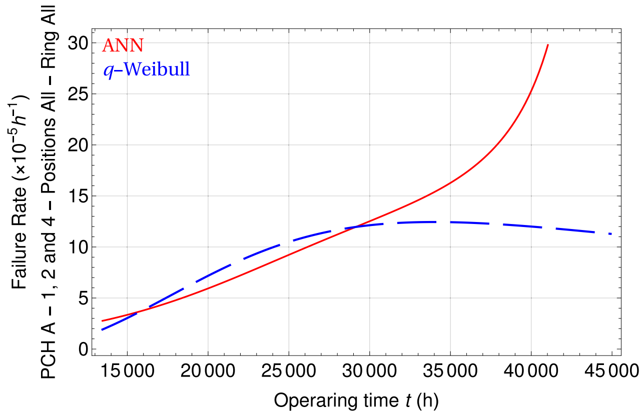


FIGURE 15. Failure rate values for artificial neural network (solid line, red online) and *q*-Weibull (dashed line, blue online) models of the PCH A usine (wear failure mode).

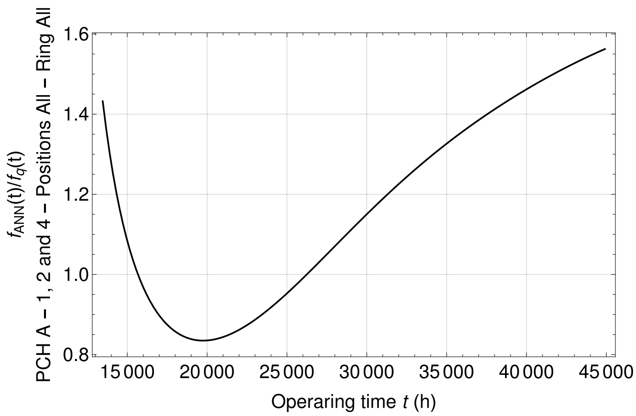


FIGURE 14. The function equivalent of the probability density function for Artificial neural network model divided by -Weibull probability density function of the PCH A usine (wear failure mode).

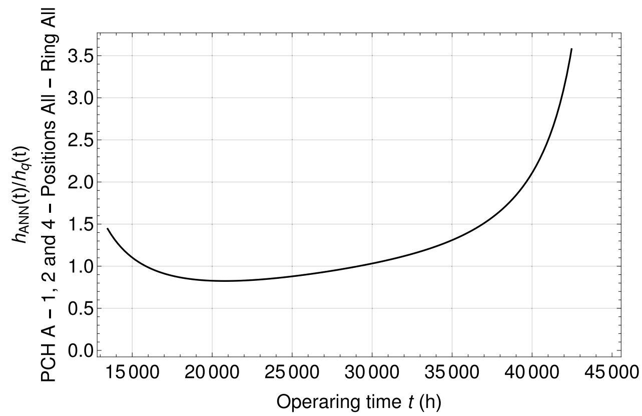


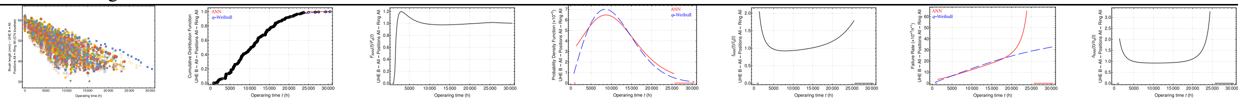
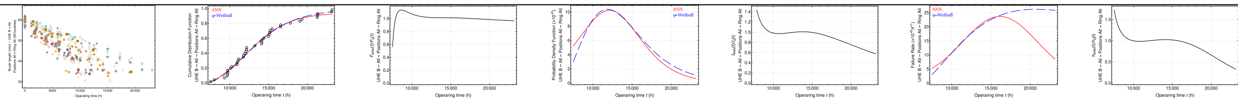
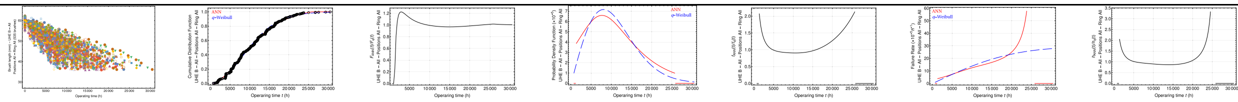
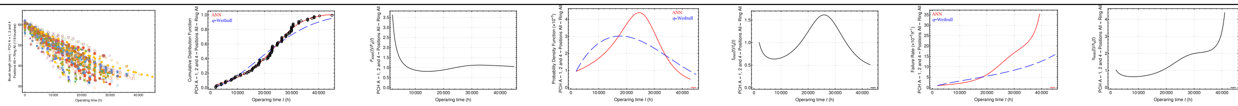
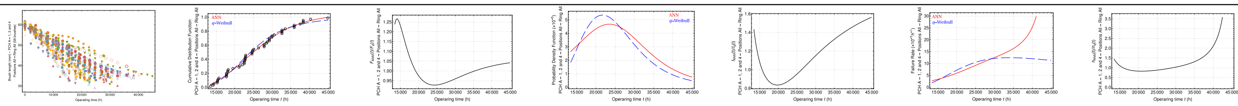
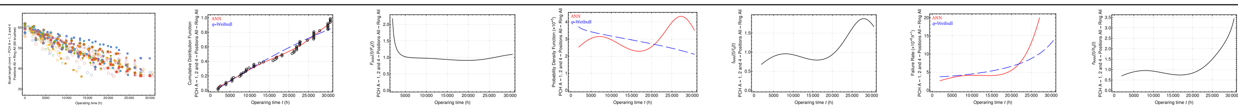
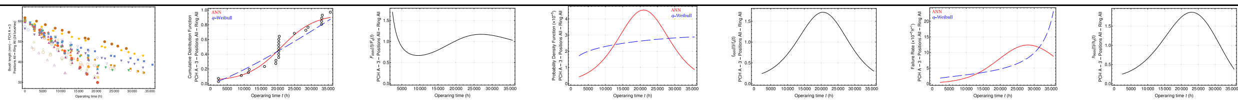
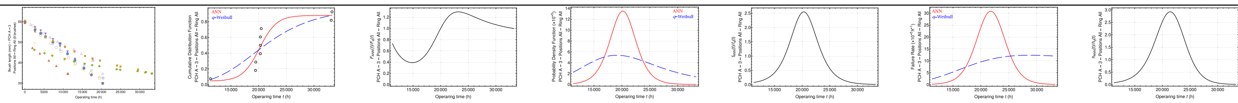
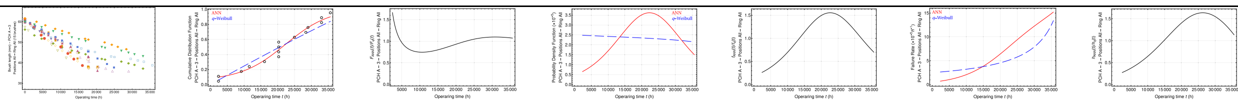
FIGURE 16. Failure rate values of artificial neural network model divided by -Weibull failure rate of the PCH A usine (wear failure mode).

failure rate shows very different results. For times below $t = 30,000$h, there is a variation close to the previous analyzes. But for time values greater than this the curves deviate and the expression $f_{ANN}(t)/f_q(t)$ returns values around 3.5 for the last time values as shown in Figure 16.

Table 2 shows a comparative summary of the fittings made by the q-Weibull distribution and the artificial neural network model. The first column shows the index i which is used to identify the numerical results of the Table 3. The "Operating time and brush length" column shows a scatter plot of the brush length values and their time intervals measured by the hour meter. These values are the result of applying the filters shown in Figure 6. The third, fifth, and seventh columns show the graph of the cumulative distribution function, the probability density function, and the failure rate function with the ANN model plotted in solid line, red online, and q-Weibull model plotted in dashed line, blue online. The fourth, sixth, and eighth columns show the graphs of $F_{ANN}(t)/F_q(t)$, $f_{ANN}(t)/f_q(t)$, and $h_{ANN}(t)/h_q(t)$.

The first three rows of the Table 2 are related to the UHE plant. The first row refers to all the failure modes involved (574 cases - see Table 3), the second row only refers to wear failure mode (39 replacements), that is, only the brushes that were replaced because they did not have enough length for correct operation. Row three shows the replacements made for a reason other than wear (535 cases). Note that 6.8% of UHE replacements were due to wear, while at PCH units 1, 2, and 4 this proportion is 49% and at unit 3 it is around 37%. The behaviors of the plots of the three analyzes are similar, except failure rates. The failure rate values for high

TABLE 2. Comparative summary of ANN and *q*-Weibull plots.

i	Operating time and brush length	Cumulative distribution function	$F_{ANN}(t)/F_q(t)$	Probability density function	$f_{ANN}(t)/f_q(t)$	Failure rate function	$h_{ANN}(t)/h_q(t)$
1. UHE All							
2. UHE Wear							
3. UHE Other							
4. PCH 124 All							
5. PCH 124 Wear							
6. PCH 124 Other							
7. PCH 3 All							
8. PCH 3 Wear							
9. PCH 3 Other							

times-to-failure are higher in the q-Weibull model when we analyze the wear failure mode. The behavior is the reverse for analyzes with all failure modes and also for modes that exclude wearing. Discontinuity is observed above $t = 25,000$h in the probability density and failure rate plots of the analyzes with all failure modes and non-wear failure modes. This range of high time values was not observed in the hour meter readings in the wear failure mode, so the discontinuity does not seem to have interfered in the divergence of the behavior models.

The graphical results of the failure modes (all, wear, and non-wear) of units 1, 2, and 4 of the PCH plant are very similar (see rows 4, 5, and 6 in Table 2). The only easily observed difference is that the probability density function of the q-Weibull model for the other failure modes (non-wear) is decreasing, whereas for analyzes with all failure modes and wear failure mode the curves are unimodal.

The equipment of units 1, 2, and 4 of the PCH plant have similar characteristics, so they were analyzed together. Unit 3 of this plant has very different characteristics and operating conditions. There is a concentration of sample items at the center of the graphs of the accumulated distribution function. This concentration is more adequately represented by the ANN, whereas the q-Weibull distribution cannot bend to approach the sample data. The consequence of this phenomenon is that the density functions represented by the ANN are peaked and this is reflected in the failure rate curves (see rows 7, 8, and 9 of the Table 2).

VI. RESULTS

The Table 3 shows the acronym of the hydroelectric plant, failure modes, generation units, quantities of sample data, values of the logarithms of the likelihoods of the neural network and the q-Weibull distribution, values of the coefficients of determination (R^2) of the two analyzes and the mean squared error (MSE) values.

The likelihood values calculated by the ANN model ($\mathcal{L}_{\text{ANN}}$) and by the q-Weibull model ($\mathcal{L}_{q\text{W}}$) are very close. The percentage differences are less than 2%, except for row 8, in which case the difference is 7%. The likelihood values

TABLE 3. Acronym of the hydroelectric plant, failure modes, generation units, quantities of sample data, values of the logarithms of the likelihoods, values of the coefficients of determination (R^2), mean squared error (MSE) values for nine analyzes.

i	Hydroelectric plant	Failure modes	Units	n	$\mathcal{L}_{\text{ANN}}$	$\mathcal{L}_{q\text{W}}$	R^2_{ANN}	$R^2_{q\text{W}}$	MSE_{ANN}	$MSE_{q\text{W}}$
1	UHE B	All	All	574	-5638.99	-5754.02	0.9981	0.9911	0.00016	0.00030
2	UHE B	Wear	All	39	-379.87	-377.54	0.9879	0.9807	0.00099	0.00097
3	UHE B	Non wear	All	535	-5333.12	-5362.78	0.9985	0.9900	0.00013	0.00042
4	PCH A	All	1, 2, and 4	118	-1232.05	-1256.07	0.9951	0.9630	0.00041	0.00366
5	PCH A	Wear	1, 2, and 4	58	-598.01	-600.83	0.9921	0.9666	0.00065	0.00131
6	PCH A	Non wear	1, 2, and 4	60	-634.90	-629.62	0.9746	0.9777	0.00208	0.00257
7	PCH A	All	3	24	-255.51	-253.73	0.9301	0.9039	0.00563	0.00851
8	PCH A	Wear	3	9	-98.74	-92.07	0.8312	0.8076	0.01274	0.01813
9	PCH A	Non wear	3	15	-161.13	-160.02	0.9726	0.9268	0.00216	0.00469

TABLE 4. Reliability models, hydroelectric plants, shapes of failure rate, and count of shapes.

i	Model	Hydroelectric plant	Shape	Count
1	ANN	PCH A	↑	3
2	ANN	PCH A	∩	2
3	ANN	PCH A	∩,↑	1
4	ANN	UHE B	↑	2
5	ANN	UHE B	∩	1
6	q-Weibull	PCH A	↑	2
7	q-Weibull	PCH A	∪	2
8	q-Weibull	PCH A	∩	2
9	q-Weibull	UHE B	∩	3
10	Weibull	PCH A	↑	6
11	Weibull	UHE B	↑	3
12	q-exponential	PCH A	↑	6
13	q-exponential	UHE B	↑	3
14	exponential	PCH A	–	6
15	exponential	UHE B	–	3

↓= decreasing, ↑= increasing, – = constant, ∪=u-shaped, and ∩=unimodal

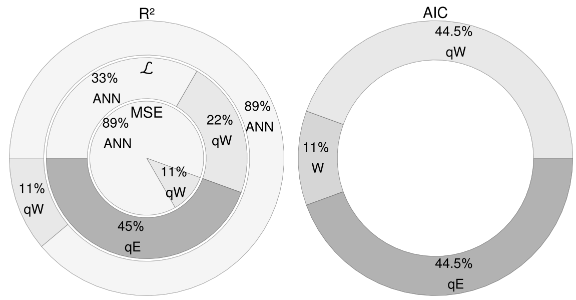


FIGURE 17. The percentages of the best models according to the fit quality criteria. Left panel: proportions of ANN, *q*-Weibull (qW), Weibull (W), *q*-exponential (qE), and exponential (E) models according to the criteria R^2 (outer ring), $\mathcal{L}$ (intermediate ring), and MSE (internal disk). The exponential model was not the best according to any criteria. Right panel: performance comparison of the *q*-Weibull (qW), Weibull (W), *q*-exponential (qE), and exponential (E) models according to the AIC criterion.

of rows 2, 6, 7, 8, and 9 are lower for the ANN model. In the other rows, the model with lower likelihood values is q-Weibull. The results of the two models are balanced.

In general, the values of the coefficients of determination for the ANN model (R^2_{ANN}) are slightly higher than the values of the q-Weibull model ($R^2_{q\text{W}}$). Although the difference is small (below 5%), this is a trend observed along almost all the rows of the Table 3. The coefficient of determination of the q-Weibull fitting ($R^2_{q\text{W}}$) of the PCH A plant for times until failure of the non-wear failure mode (row 6) is the only one greater than the coefficient calculated with an artificial neural network (only 0.32%). The mean squared error of the artificial neural network (MSE_{ANN}) is smaller than that calculated with the q-Weibull distribution in most cases, only for the UHE B plant in the wear failure mode the reverse occurs (row 2).

The ANN model presents a higher value of R^2 for 89% of the analyzed cases, in only 11% of the cases (one out of nine), the q-Weibull model presents a better performance (see Figure 17). The quality analysis according to likelihood points to the q-Weibull model as better in 22% of cases (two occurrences) and the ANN model in 33% of cases (three occurrences). All of these five cases would point to ANN if the criterion was the highest value of R^2. According to the likelihood criterion ($\mathcal{L}$), the q-exponential model was the best in the remaining four cases (45%). The MSE criterion points to ANN as the best model in 89% of cases and q-Weibull in 11%. The proportions are the same as those found when the criterion is higher R^2. It is worth remembering that the parameters of the statistical distributions were calculated with the maximization of R^2 and that the values R^2 of the Weibull, q-exponential, and exponential models are smaller than those found by the q-Weibull distribution in all nine cases. This is expected because q-Weibull is a generalization that encompasses the others. However, it is interesting to compare it with the most widely used statistical distributions. All results are summarized in Appendix A (Table 5).

The AIC criterion was used to compare the performance of statistical distributions. The q-Weibull and q-exponential models showed the same level of performance (44.5%) while the Weibull model was better in 11% of the cases.

The shape of the failure rate function is crucial for maintenance scheduling, especially preventive and corrective maintenance. The way in which the failure rate grows over time directly influences the scheduling of preventive maintenance time intervals and consequently the total costs to keep the systems running.

According to Table 4, the ANN model can identify three failure rate shapes in PCH analysis and two in UHE plants. The q-Weibull model is also capable of perceiving three formats in PCH, but only one in UHE plants. The Weibull and q-exponential models only detected a single type of

TABLE 5. Parameters, failure rate shapes, and values of the measures of goodness of fit of the ANN, *q*-Weibull, Weibull, *q*-exponential, and exponential models.

i	Plant	Unit	Failure mode	Model	β	θ	t_0	q	Shape	$\mathcal{L}$	R^2	MSE	ΔAIC
1	UHE B	All	All	ANN	-	-	-	-	↑	-5,639	0.9981	0.00016	-
2	UHE B	All	All	qW	2	1.18×10^4	93	1.04	∩	-5,754	0.9911	0.00030	0
3	UHE B	All	All	W	1.83	1.13×10^4	221	1	↑	-5,758	0.9769	0.00192	532
4	UHE B	All	All	qE	1	2.14×10^{10}	736	-964,337	↑	-5,575	0.9306	0.00475	869
5	UHE B	All	All	E	1	1.45×10^4	747	1	-	-5,904	0.8367	0.01089	1,652
6	UHE B	All	Wear	ANN	-	-	-	-	∩	-380	0.9879	0.00099	-
7	UHE B	All	Wear	qW	2.19	6.95×10^3	6,249	1.24	∩	-378	0.9807	0.00097	6
8	UHE B	All	Wear	W	1.8	8.29×10^3	6,644	1	↑	-377	0.9783	0.00118	0
9	UHE B	All	Wear	qE	1	1.49×10^5	7,297	-9.35	↑	-354	0.9442	0.00411	17
10	UHE B	All	Wear	E	1	8.68×10^3	7,378	1	-	-383	0.8865	0.00682	62
11	UHE B	All	Non wear	ANN	-	-	-	-	↑	-5,333	0.9985	0.00013	-
12	UHE B	All	Non wear	qW	2	1.11×10^4	92	1.08	∩	-5,363	0.9900	0.00042	0
13	UHE B	All	Non wear	W	1.77	1.16×10^4	235	1	↑	-5,362	0.9868	0.00042	117
14	UHE B	All	Non wear	qE	1	3.53×10^4	742	-0.03	↑	-5,421	0.9063	0.00571	1,014
15	UHE B	All	Non wear	E	1	1.41×10^4	746	1	-	-5,490	0.8459	0.00985	1,429
16	PCH A	1,2,4	All	ANN	-	-	-	-	↑	-1,232	0.9951	0.00041	-
17	PCH A	1,2,4	All	qW	1.69	3.2×10^4	16	0.74	↑	-1,256	0.9630	0.00366	42
18	PCH A	1,2,4	All	W	2.02	2.64×10^4	-1,188	1	↑	-1,255	0.9610	0.00322	61
19	PCH A	1,2,4	All	qE	1	1.02×10^{10}	1,547	-237,139	↑	-1,248	0.9610	0.00780	0
20	PCH A	1,2,4	All	E	1	2.74×10^4	1,730	1	-	-1,291	0.8456	0.01587	222
21	PCH A	1,2,4	Wear	ANN	-	-	-	-	↑	-598	0.9921	0.00065	-
22	PCH A	1,2,4	Wear	qW	2.69	1.4×10^4	8,785	1.35	∩	-601	0.9666	0.00131	36
23	PCH A	1,2,4	Wear	W	1.53	1.57×10^4	11,305	1	↑	-601	0.9347	0.00243	52
24	PCH A	1,2,4	Wear	qE	1	1.32×10^5	13,212	-2.9	↑	-600	0.9585	0.00808	0
25	PCH A	1,2,4	Wear	E	1	1.76×10^4	13,001	1	-	-609	0.9050	0.00756	71
26	PCH A	1,2,4	Non wear	ANN	-	-	-	-	∩,↑	-635	0.9746	0.00208	-
27	PCH A	1,2,4	Non wear	qW	0.97	1.23×10^5	1,381	-2.07	∪	-630	0.9777	0.00257	4
28	PCH A	1,2,4	Non wear	W	1.49	2.00×10^4	153	1	↑	-636	0.9502	0.00413	75
29	PCH A	1,2,4	Non wear	qE	1	1.68×10^5	1,266	-3.64	↑	-628	0.9776	0.00215	0
30	PCH A	1,2,4	Non wear	E	1	1.96×10^4	1,584	1	-	-642	0.9260	0.0068	97
31	PCH A	3	All	ANN	-	-	-	-	∩	-256	0.9301	0.00563	-
32	PCH A	3	All	qW	1.2	1.18×10^6	1	-57.71	↑	-254	0.9039	0.00851	0
33	PCH A	3	All	W	2	2.85×10^4	-2,993	1	↑	-256	0.8955	0.00856	12
34	PCH A	3	All	qE	1	3.47×10^{10}	596	-782,995	↑	-257	0.8873	0.01325	1
35	PCH A	3	All	E	1	2.72×10^4	1,304	1	-	-262	0.7839	0.01979	27
36	PCH A	3	Wear	ANN	-	-	-	-	∩	-99	0.8312	0.01274	-
37	PCH A	3	Wear	qW	4.22	1.83×10^4	1	1.47	∩	-92	0.8076	0.01813	17
38	PCH A	3	Wear	W	2.21	1.96×10^4	5,211	1	↑	-92	0.7992	0.01918	5
39	PCH A	3	Wear	qE	1	8.19×10^6	9,339	-305.93	↑	-92	0.7923	0.02073	0
40	PCH A	3	Wear	E	1	1.58×10^4	10,155	1	-	-94	0.7478	0.02232	2
41	PCH A	3	Non wear	ANN	-	-	-	-	↑	-161	0.9726	0.00216	-
42	PCH A	3	Non wear	qW	0.99	5.52×10^5	1	-11.43	∪	-160	0.9268	0.00469	0
43	PCH A	3	Non wear	W	1.7	2.82×10^4	-3,240	1	↑	-162	0.9193	0.00545	6
44	PCH A	3	Non wear	qE	1	1.11×10^5	-94	-1	↑	-162	0.9178	0.00741	0
45	PCH A	3	Non wear	E	1	2.56×10^4	677	1	-	-164	0.8547	0.01206	11

ANN= Artificial neural network, qW=*q*-Weibull, W=Weibull, qE=*q*-exponential, and E=exponential.
↓= decreasing, ↑= increasing, – = constant, ∪=u-shaped, and ∩=unimodal.

format and finally, the exponential model is not able to model increasing failure rates, only constant.

Management decisions based on the behavior of the failure rate tend to be more assertive when the ANN model is used because this model can express more failure rate shapes and get closer to real behavior.

The comparison of the plots of the ANN and *q*-Weibull models shows an advantage of ANN modeling. Especially in situations where the cumulative distribution function needs more pronounced slopes. The processing speed of an ANN can also be advantageous. Many languages and programs have libraries optimized for calculating ANNs. The *q*-Weibull distribution does not have libraries in most programming languages.

A disadvantage of the ANN model is the need for additional validations for the model to represent a statistical distribution. The generated function cannot be decreasing, assuming values greater than 1 or less than 0. These controls can generate discontinuities in the accumulated distribution function. Another disadvantage is the loss of physical interpretation based on the distribution parameters.

The *q*-Weibull distribution has the advantage of relating the shape of your curves directly to the values of your parameter (see [11] for details). In this way, the nature of the modeled phenomena can be interpreted by the values of its parameter, which does not occur in an ANN.

VII. CONCLUSION

Brushes are critical components in power generation equipment. The interruptions in brush operation because they have failed in their service or for scheduled inspection tasks can cause financial losses avoidable by proper maintenance planning.

The data provided by the operation and maintenance sectors were treated using selection tools of machine learning. The extracted information was classified as valid or rejected because it appeared to be errors in registration. The valid information was re-classified to infer which failure mode caused the brush replacement.

Data from 716 brushes running at two hydroelectric power plants were used. At UHE B plant, 574 brushes were analyzed, and at PCH A 142 brushes. The information from the PCH A plant was divided into two groups. The first group comprises units 1, 2, and 4. The second group contains data only for unit 3 because its characteristics are different from other units.

The combination of reliability analysis tools and the learning machine is used to create a model of prediction of failure probability based on an artificial neural network. The q-Weibull model was used to compare the reliability analysis quantities. The cumulative distribution function curves, probability density function, failure rate function, and comparative graphs of the two models for each of these three functions were plotted.

In addition to comparing ANN's performance with the generalized q-Weibull model, comparisons with statistical distributions commonly used in reliability analysis, Weibull, exponential, and q-exponential were performed. The predictions of the ANN model are better than the predictions of models with exponential characteristics (Weibull and exponential) and also superior to models that are asymptotically power laws (q-Weibull and q-exponential). This is a strong indication that the use of ANN for predicting reliability in systems of a complex or non-complex nature can be very promising.

Three measures of goodness of fit were calculated. The coefficient of determination, the logarithm of likelihood, and the mean squared error. Most of the values found point to an advantage in the use of artificial neural networks over the use of the q-Weibull distribution.

ANN modeling was the best in 89% of the analyzes in two of the three criteria (R^2 and MSE). The comparison was performed with the q-Weibull, Weibull, q-exponential, and exponential models. It is worth remembering that all statistical distributions were calculated by maximizing R^2. ANN is the second-best model according to the likelihood criterion ($\mathcal{L}$). The reasons for this success seem to be the theoretically unlimited amount of possible failure rate shapes in ANN modeling.

The models based on artificial neural networks required additional care to meet the requirements of a probabilistic model. Even so, they were successfully used to model the cumulative distribution function in all nine cases in which it was employed.

The modeling of the cumulative distribution function by ANN allowed the replacement of the statistical model by the machine learning model without losing the requirements of a probability distribution and with the advantage of its probability density function being able to express a better concentration of values as shown in rows seven, eight and nine from Table 2.

The use of selection tools in conjunction with artificial neural networks opens a path for reliability analysis allowing views of patterns that would hardly be observed in reliability analyzes not assisted by machine learning methods.

Many programming languages have libraries designed for modeling by artificial neural networks. The level of maturity of the codes used in these tools seems to facilitate the implementation of programs to apply artificial neural networks in reliability analyzes, allowing rapid code execution.

In summary, this method of applying ANN in reliability analysis can have a significant impact on reducing maintenance costs, as it leads to results closer to reality. The ANN algorithms have reached a high level of maturity and are executed with adequate speed. Currently, such computer programs can run on hardware specially dedicated to machine learning.

APPENDIX

Parameters, failure rate shapes, and values of the measures of goodness of fit of the ANN, q-Weibull, Weibull, q-exponential, and exponential models (See Table 5).

ACKNOWLEDGMENT

- ANEEL - Agência Nacional de Energia Elétrica
- FUNDEB - Fundação para o Desenvolvimento de Bauru
- COPEL - Companhia Paranaense de Eletricidade - Project PD-06491-0391-2015

REFERENCES

[1] I. Plokhov, I. Savraev, A. Markov, A. Ilyin, O. Kozyreva, and N. Kotkov, "Industrial tests of current distribution dynamics in the brush-contact apparatus of the turbo-generator," *Environ. Technol. Resour. Int. Sci. Practical Conf.*, vol. 3, Jun. 2017, p. 258.

[2] J. K. Skjølberg, H. F. Ohma, and M. Runde, "Wear rates and current distribution of carbon brushes on steel slip rings," *IEEE Trans. Energy Convers.*, vol. 24, no. 4, pp. 835–840, Dec. 2009.

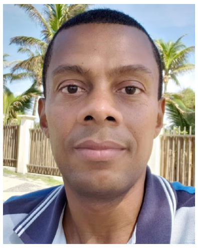
[3] W.-G. Shin, Y.-S. Song, and Y.-K. Seo, "Correlation analysis of brush temperature in brush-type DC motor for predicting motor life," *J. Mech. Sci. Technol.*, vol. 26, no. 7, pp. 2151–2154, Jul. 2012.

[4] V. P. Hessler, "Electrical brush wear," *Electr. Eng.*, vol. 54, no. 10, pp. 1050–1054, Oct. 1935.

[5] T. Ueno, K. Kadono, S. Yamaguchi, M. Aoyagi, A. Tanaka, and N. Morita, "Relationship between contact voltage drop and frictional coefficient under high-current sliding contact," *IEEJ Trans. Electr. Electron. Eng.*, vol. 5, no. 4, pp. 486–492, Jun. 2010.

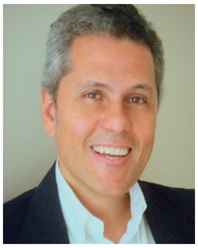
[6] W.-G. Shin and S.-H. Lee, "Determination of accelerated condition for brush wear of small brush-type DC motor in using design of experiment (DOE) based on the taguchi method," *J. Mech. Sci. Technol.*, vol. 25, no. 2, pp. 317–322, Feb. 2011.

[7] K. Rajkumar, S. Aravindan, and M. S. Kulkarni, "Wear and life characteristics of microwave-sintered copper-graphite composite," *J. Mater. Eng. Perform.*, vol. 21, no. 11, pp. 2389–2397, Nov. 2012.
[8] C. Holzapfel, "Contact surfaces for sliding electrical contacts," in *Proc. IEEE 60th Holm Conf. Electr. Contacts (Holm)*, Oct. 2014, pp. 1–7.
[9] J.-K. Xiao, L.-M. Liu, C. Zhang, L. Zhang, and K.-C. Zhou, "Sliding electrical contact behavior of brass fiber brush against coin-silver and Au plating," *Wear*, vols. 368–369, pp. 461–469, Dec. 2016.
[10] P. Bak, *How Nature Works*. New York, NY, USA: Springer, 1996.
[11] E. M. Assis, E. P. Borges, and S. A. B. V. de Melo, "Generalizaed q-Weibull model and the bathtub curve," *Int. J. Qual. Rel. Manage.*, vol. 30, no. 7, pp. 720–736, 2013.
[12] M. C. Liu, W. Kuo, and T. Sastri, "An exploratory study of a neural network approach for reliability data analysis," *Qual. Rel. Eng. Int.*, vol. 11, no. 2, pp. 107–112, 1995.
[13] N. Karunanithi, D. Whitley, and Y. K. Malaiya, "Using neural networks in reliability prediction," *IEEE Softw.*, vol. 9, no. 4, pp. 53–59, Jul. 1992.
[14] N. Amjady and M. Ehsan, "Evaluation of power systems reliability by an artificial neural network," *IEEE Trans. Power Syst.*, vol. 14, no. 1, pp. 287–292, Feb. 1999.
[15] S. K. Sinha and M. D. Pandey, "Probabilistic neural network for reliability assessment of oil and gas pipelines," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 17, no. 5, pp. 320–329, Sep. 2002.
[16] D. F. Specht and P. D. Shapiro, "Generalization accuracy of probabilistic neural networks compared with backpropagation networks," in *Proc. Seattle Int. Joint Conf. Neural Netw. (IJCNN)*, vol. 1, 1991, pp. 887–892.
[17] P. Majumder, M. Majumder, A. K. Saha, and S. Nath, "Selection of features for analysis of reliability of performance in hydropower plants: A multi-criteria decision making approach," *Environ., Develop. Sustainability*, vol. 22, no. 4, pp. 3239–3265, Apr. 2020.
[18] K. Xu, M. Xie, L. C. Tang, and S. L. Ho, "Application of neural networks in forecasting engine systems reliability," *Appl. Soft Comput.*, vol. 2, no. 4, pp. 255–268, Feb. 2003.
[19] J. P. Giri, V. H. Tatwawadi, G. S. Mahakalkar, and P. J. Modak, "Comparative analysis of multilayer perceptron and radial basis function ANN for prediction of cycle time of structural subassembly manufacturing," *Int. J. Sci., Spirituality, Bus. Technol.*, vol. 3, no. 1, pp. 71–82, 2014.
[20] G. S. Fesghandis, A. Pooya, M. Kazemi, and Z. N. Azimi, "Comparison of multilayer perceptron and radial basis function neural networks in predicting the success of new product development," *Eng., Technol. Appl. Sci. Res.*, vol. 7, no. 1, pp. 1425–1428, 2017.
[21] A. T. Hammid, M. H. B. Sulaiman, and A. N. Abdalla, "Prediction of small hydropower plant power production in himreen lake dam (HLD) using artificial neural network," *Alexandria Eng. J.*, vol. 57, no. 1, pp. 211–221, Mar. 2018.
[22] R. Ramakumar, *Engineering Reliability: Fundamentals and Applications*. Upper Saddle River, NJ, USA: Prentice-Hall, 1993.
[23] H. A. Sandtorv, P. Hokstad, and D. W. Thompson, "Practical experience with a data collection project: The OREDA project," *Rel. Eng. Syst. Saf.*, vol. 51, no. 2, pp. 159–167, Feb. 1996.
[24] W. Weibull, "A statistical distribution function of wide applicability," *J. Appl. Mech.*, pp. 293–297, Sep. 1951.
[25] D. N. P. Murthy, M. Xie, and R. Jiang, *Weibull Models*. New York, NY, USA: Wiley, 2004.
[26] H. Pham and C.-D. Lai, "On recent generalizations of the Weibull distribution," *IEEE Trans. Rel.*, vol. 56, no. 3, pp. 454–458, Sep. 2007.
[27] S. Nadarajah and S. Kotz, "On some recent modifications of weibull distribution," *IEEE Trans. Rel.*, vol. 54, no. 4, pp. 561–562, Dec. 2005.
[28] M. Berberan-Santos, E. N. Bodunov, and B. Valeur, "History of the kohlrausch (stretched exponential) function: Focus on uncited pioneering work in luminescence," 2008, *arXiv:0804.1814*. [Online]. Available: http://arxiv.org/abs/0804.1814
[29] C. Tsallis, "Possible generalization of Boltzmann-Gibbs statistics," *J. Stat. Phys.*, vol. 52, nos. 1–2, pp. 479–487, Jul. 1988.
[30] C. Tsallis, "Extensive versus nonextensive physics," in *New Trends in Magnetism, Magnetic Materials, and Their Applications*, J. L. Moran-Lopez and J. M. Sanchez, Eds. New York, NY, USA: Plenum Press, 1994, pp. 451–463.
[31] C. Tsallis, *Introduction to Nonextensive Statistical Mechanics*. New York, NY, USA: Springer, 2009.
[32] T. Yamano, "Some properties of q-logarithm and q-exponential functions in tsallis statistics," *Phys. A, Stat. Mech. Appl.*, vol. 305, nos. 3–4, pp. 486–496, Mar. 2002.
[33] J. C. Reboredo, M. A. Rivera-Castro, and E. M. de Assis, "Power-law behaviour in time durations between extreme returns," *Quant. Finance*, vol. 14, no. 12, pp. 2171–2183, Dec. 2014.
[34] E. M. de Assis, G. A. C. Lima, A. Prestes, F. Marinho, and L. A. N. Costa, "*q*-Weibull applied to Brazilian hydropower equipment," *IEEE Trans. Rel.*, vol. 68, no. 1, pp. 122–132, Mar. 2019.
[35] E. M. Assis, E. P. Borges, S. A. B. V. de Melo, and L. Schnitman, "Modeling failure rate of a robotic welding station using generalized q-distributions," *Int. J. Qual. Rel. Manage.*, vol. 32, no. 2, pp. 156–166, Feb. 2015.
[36] U. M. S. Costa, V. N. Freire, L. C. Malacarne, R. S. Mendes, S. Picoli, Jr., E. A. de Vasconcelos, and E. F. da Silva, Jr., "An improved description of the dielectric breakdown in oxides based on a generalized weibull distribution," *Phys. A, Stat. Mech. Appl.*, vol. 361, no. 1, pp. 209–215, Feb. 2006.
[37] L. G. Johnson, "The median ranks of sample values in their population with an application to certain fatigue studies," *Ind. Math.*, vol. 2, pp. 1–9, Jul. 1951.
[38] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Autom. Control*, vol. AC-19, no. 6, pp. 716–723, Dec. 1974.
[39] C. M. Hurvich and T. Chih-Ling, "Regression and time series model selection in small samples," *Biometrika*, vol. 76, no. 2, pp. 297–307, 1989.
[40] L. Bottou, F. E. Curtis, and J. Nocedal, "Optimization methods for large-scale machine learning," *SIAM Rev.*, vol. 60, no. 2, pp. 223–311, Jan. 2018.
[41] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," *Science*, vol. 349, no. 6245, pp. 255–260, Jul. 2015.
[42] M. Bowles, *Machine Learning in Python: Essential Techniques for Predictive Analysis*, vol. 1542. Indianapolis, IN, USA: Wiley, 2015.
[43] M. P. Deisenroth, A. A. Faisal, and C. S. Ong, *Mathematics for Machine Learning*. Cambridge, U.K.: Cambridge Univ. Press, 2020.
[44] J. G. S. León, *Mathematica Beyond Mathematics: The Wolfram Language in the Real World*. Boca Raton, FL, USA: CRC Press, 2017.
[45] M. Mureşan, *Introduction to Mathematica With Applications*. Cham, Switzerland: Springer, 2017.
[46] M. Paluszek and S. Thomas, *MATLAB Machine Learning*. New York, NY, USA: Apress, 2017.
[47] A. Ghatak, *Machine Learning With R*. Singapore: Springer, 2017.
[48] M. D. Bloice and A. Holzinger, *A Tutorial on Machine Learning and Data Science Tools With Python*, vol. 1. Cham, Switzerland: Springer, 2016, pp. 435–480.
[49] K. K. Berggren *et al.*, "Roadmap on emerging hardware and technology for machine learning," *Nanotechnology*, vol. 32, no. 1, pp. 1–46, 2020.
[50] S. Cass, "Taking AI to the edge: Google's TPU now comes in a maker-friendly package," *IEEE Spectr.*, vol. 56, no. 5, pp. 16–17, May 2019.
[51] A. Zhang, Z. C. Lipton, M. Li, and A. J. Smola. (2019). *Dive Into Deep Learning Release 0.7*. [Online]. Available: https://www.en.d2l.ai
[52] L. Francis, "Neural networks demystified," *Casualty Actuarial Soc. Forum*, vol. Winter, pp. 253–320, 2001. [Online]. Available: https://www.casact.org/pubssearch/index.cfm?fa=adv_search_rs&keywords=&author=&pubYear=2001&category=&search=Search)
[53] S. Gollapudi, *Practical Machine Learning. Tackle the Real-World Complexities of Modern Machine Learning With Innovative and Cutting-Edge Techniques*. Birmingham, U.K.: Packt Publishing Ltd., 2016.
[54] W. Richert and L. P. Coelho, *Building Machine Learning System With Python*. Birmingham, U.K.: Packt Publishing Ltd, 2013.



EDILSON MACHADO DE ASSIS received the B.S. degree in civil engineering from the Catholic University of Salvador (UCSal), in 1994, the M.S. degree in production engineering from the Federal University of Bahia (UFBa / MEP), in 1999, and the Ph.D. degree in industrial engineering from the Federal University of Bahia (UFBa / PEI), in 2013.

He is currently a Reliability Researcher of FUNDEB—Bauru Development Foundation for ANEEL—National Electric Energy Agency Research and Development projects and a participant with the Graduate Program in Industrial Engineering, UFBA (PEI), as a Co-Supervisor of master's and Ph.D. degrees. He had experience in civil engineering, in the area of design of structures in reinforced concrete, also worked in the design of water supply and distribution networks. He develops consulting in the area of reliability engineering by performing risk analysis of non-tolerable events and works with methodological and product innovations.



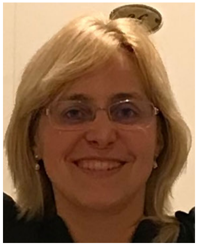
CELSO LUIZ SANTIAGO FIGUEIRÔA FILHO received the B.S. degree in mechanical engineering from the Federal University of Minas Gerais, in 1991, and the M.Sc. degree in production engineering from the Federal University of Bahia, in 1999. He is currently working as a Researcher and a Consultant in human reliability and industrial maintenance management and also holds a position as a Lecturer of the Catholic University of Salvador. His research interests include human reliability and maintenance. Since 2010, he has been a Human Reliability Researcher of FUNDEB—Bauru Development Foundation for the ANEEL—National Electric Energy Agency Research and Development projects. He is also a Consultant in the Oil and Gas, Iron and Steel, Metal Mining, Hydropower Generation, Water and Sanitation, and Transportation sectors (air, road, and rail).

GABRIEL ALVES DA COSTA LIMA received the master's degree in mines engineering from the University of Campinas, in 2004, the Ph.D. degree in industrial engineering from the University of Campinas, in 2006, and the Ph.D. degree in economy from the University of Campinas, in 2008.

He was a Professor of UNICAMP from 2008 till 2014. Since 2008, he has been a Researcher of FUNDEB—Bauru Development Foundation for the ANEEL—National Electric Energy Agency Research and Development projects. He is a Researcher projects in systems reliability, economic models applied to oil and gas, econometrics models for predictions, and uncertainties models. His research projects, such as develop of reliability, risk, and asset management quantitative methodology (UNICAMP/Comgas), E&P Portfolio optimization methodologies using 3D de components visualization techniques strategy (UNICAMP/Petrobras), methodology and computer platform for assessing human reliability in operational and maintenance activities in power generation companies (FUNDEB), development of a state-of-the-art computing platform for optimization of operational improvement project groups, electric power (FUNDEB), and development of a serial head computational platform for the economic quantification of assets employed in activities of the power generation sector based on non-conventional models (FUNDEB).

LUIS AUGUSTO NAGASAKI COSTA received the M.S. degree from the University of Campinas, Campinas, Brazil, in 2012.

Since 2012, he has been a Researcher of FUNDEB—Bauru Development Foundation for the ANEEL—National Electric Energy Agency Research and Development projects. He is currently an Electrical Engineer from the UNESP–State University of São Paulo His research interests include data analysis, optimization, risk modeling, and simulation.



GISELE MARIA DE OLIVEIRA SALLES received the B.S. degree in electrical engineering from the Federal University of Technology at Paraná, in 2006.

She has been working as a Plant Maintenance Management with Copel, Paraná Energy Company, since 2008.