XLI CILAMCE
XLI Ibero-Latin American Congress on Computational Methods in Engineering
16-19 November, 2020 FOZ DO IGUAÇU PARANÁ – BRAZIL

Fatigue Curves in Steel Railway Generated by Genetic Algorithm

Maria Rafaela B. de M. Ribeiro[1], José Antônio F. Santiago[1], Ritermayer M. Teixeira[2]

[1] Civil Engineering Program, *COPPE, Federal University of Rio de Janeiro*
Centro de Tecnologia, Ilha do Fundão, Zip-Code 21941-972, Rio de Janeiro/Rio de Janeiro, Brazil
maria.ribeiro@coc.ufrj.br, santiago@coc.ufrj.br
[2] Dept. of Civil Engineering, *ITEC, Federal University of Pará*
Rua Augusto Corrêa, Zip-Code 66075-900, Pará/Belém, Brazil
ritermayer@ufpa.br

Abstract. Genetic Algorithms (GA) are global optimization, based on natural selections and genetic mechanisms for which structured and random search strategies are geared by reinforcing the search for high amplitude points in which the function to be minimized (maximized) has relatively low (high) values. This work deals with a study on the Genetic Algorithm application to estimate parameters inherent to S-N-p Curves generated for typically structural details of the metal railway bridges applying together with The Maximum Likelihood Method to infer runout data, such as censored data for failure probability functions. MatLab software is employed to develop the computational program. To compare the obtained results, the Interior Point Algorithm is also presented, due to its wide use in problems involving linear and nonlinear quadratic programming.

Keywords: Fatigue, Railway Bridge, Probabilistic Analysis, Genetic Algorithm.

1 Introduction

In Brazil, the first railroads date from the second half of the 19th century, as well as the bridges builted for the circulation of these standard trains. Most of these structures remain in operation, for economic reasons, even more than a century after their construction. It is known that such bridges were designed with riveted connections, where they were not originally designed due to fatigue failure [1]. The assessment of the safety conditions of these bridges has an increasing relevance, as well as the techniques applied, as these bridges were designed for traffic conditions completely different from those that occur in the present times, which are increasingly intensified. As a consequence, it is known that certain sections whose bridges had been designed for loads of 90 kN / axis have undergone successive reinforcements, starting to support compositions with up to 200 kN / axis [2].

More robust and reliability-based techniques assist in the assessment and analysis of this type of problem in order to update the reinforcement and maintenance methodologies. Naturally, the search for the best efficiency of a process or the best performance of a structure leads us to think about the development of several routes to reach an optimal result through experimental methods. It is known that the idea of improving or optimizing a problem, implies some freedom to be able to modify it in order to obtain the best performance of it. The potential for changes is typically expressed in terms of the variations allowed in a group of parameters that can be defined as a vector $\boldsymbol{\theta} = [\theta_1, \theta_2, ..., \theta_n]^T$, where are the number of design variables. The set of variables that provide the optimal value of the evaluated problem is called the optimal point and can also be represented by a vector $\boldsymbol{\theta}^* = [\theta_1^*, \theta_2^*, ..., \theta_n^*]^T$. Optimizing implies the existence of a merit function (objective function) that can be improved and used as a measure of the effectiveness of the problem to be assessed. This function is called an objective function. They can be functions of a variable (one-dimensional) or multivariable (multidimensional).

2 Optimization Algorithms

Optimization is a branch of mathematics applicable in several areas, focusing on some of the problems commonly encountered in structural engineering. Optimization can be defined as a process of searching for the maximum or minimum values of a function (or several functions), whose variables satisfy certain restrictions in the form of equality or inequality [3].

The set of points that maximize or minimize a function (or several functions) is called optimal solutions. The maximization of a function $f(\theta)$ can be converted into a minimization of $g(\theta)$, where $g(\theta) = -f(\theta)$. Thus, maximizing a function $f(\theta)$ is the same as minimizing $-f(\theta)$. These functions are known as objective functions and represent the amount that you want to minimize or maximize in the problem. Figure 1, extracted and adapted from [4], shows the situation in which the point θ^* generates the minimum value of a function $f(\theta)$ and the maximum negative value of the function $-f(\theta)$.

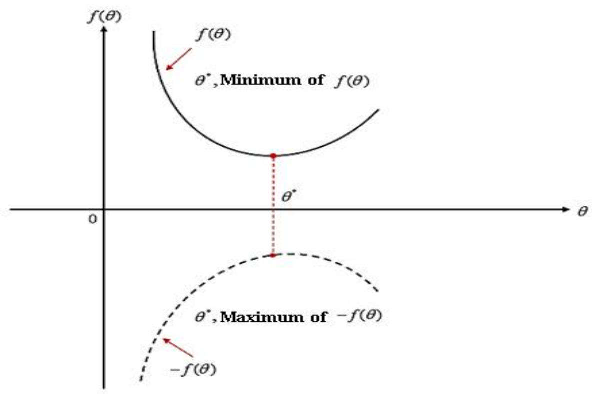

Figure 1. Minimizing $f(\theta)$ is the same as Maximizing $-f(\theta)$.

2.1 Interior Point Algorithm

The numerical method of the interior point algorithms is a more robust method, proposed by Karmarkar [5] and extremely useful in solving problems that involve linear and nonlinear quadratic programming and can be adopted as an alternative to problems that cannot always be obtained by analytical means [6]. The search for parameters is given by the interior of the feasible region, a different philosophy from the method that preceded it, also known as the Simplex method where the search for parameters is given towards the optimum of the border, that is, by the edges of the feasible region.

Several models of Interior Point were developed, among these models can be highlighted the primal-dual method of logarithmic barrier, or simply barrier function, found in the Optimization Toolbox of Matlab® [7].

2.2 Genetic Algorithm

Genetic algorithms (Genetics Algorithms - GA's) are tools for optimization, which seek to improve results for any circumstance by stochastic methods. In engineering, the aim is, in general, to minimize resources and efforts and maximize results for certain tasks. In recent years, there has been a considerable increase in the proposals for new non-traditional, or also called modern, optimization methods, which deviate from the concept of the usual optimization techniques.

These techniques can be bio-inspired, based on behaviors and characteristics of neurobiological systems, particle swarm, ant colony optimization, biology and molecules, for example. According to Rao [4], this is a technique inspired by natural genetics and the principles of natural selection, philosophically, it is based on the Darwinian theory of the evolution of species. The use of GA's to solve engineering problems becomes very attractive, since, unlike classical methods that have limited applications, they are able to identify global solutions, which, in general, can converge to the optimal solution of the problem and not to an local solution [8].

GA applications in fatigue are found in the literature. Franulovic [9] they carried out an elastic - plastic study of the material under a cyclic loading. Due to the complexity of the problem, routines with genetic operators were used to search for the parameters in order to obtain a reliable convergence for the results comparing with experimental data.

3 Modeling Fatigue Strength Curves

The fatigue strength models of structural details are traditionally presented in the form of S-N curves that relate the stress range, $\Delta\sigma$, for the percentile probability of the number of cycles to failure, N. Such curves can be obtained from fatigue tests. at constant stress amplitudes (CAFL).

The type of curve model to be analyzed is linear on a logarithmic scale. In this type of modeling the fatigue limit is fixed. In general, percentile curves are used to assess the voltage range corresponding to a specific number of cycles. In terms of international normative codes for riveted connections, there are class “D” curves for AREMA [10] and class 71 for Eurocode [11], in which curves with failure probabilities of 2.25% and 5% are assumed, respectively. In any case, the fatigue limit is assumed arbitrarily, since the model is not capable of predicting values for γ. The S-N curves present an exponential behavior in the N domain (eq. 1), requiring linearization on a logarithmic scale in order to facilitate the interpretation of the data (eq.2 and 3).

$$N = \Delta\sigma^{-m} \cdot k, \quad \Delta\sigma > \gamma \tag{1}$$

$$\ln(N) = -m \cdot \ln(\Delta\sigma) + \ln(k), \quad \ln(\Delta\sigma) > \ln(\gamma) \tag{2}$$

$$x = -m \cdot y + \ln(k), \quad y > v \tag{3}$$

To write these curves within a probabilistic field, the median curve is added to an Inverse Cumulative Probability Distribution, assumed for a given probability, as well as the standard deviation. As shown in eq. 4 below.

$$x = \bar{x}(y) + \left[F_x^{-1}(p)\right] \cdot S_x \tag{4}$$

In this way, S-N-p curves are generated that belong to an assumed probabilistic field. Sarkani *et al.*[12], for example, made analyzes assuming these correspondences of distributions. The probability function to be used for life fatigue is of the Lognornal type, in which it is equivalent to assume the Normal distribution for the logarithmic domain.

The pdf and cdf of can be determined according to eq. 5 and 6, respectively.

$$f_x(x; y, \boldsymbol{\theta}) = \frac{1}{S_x} \cdot \phi\left[\frac{x - \bar{x}(y; m, k)}{S_x}\right], \quad N > 0 \tag{5}$$

$$F_x(x; y, \boldsymbol{\theta}) = p = \Phi\left[\frac{x - \bar{x}(y; m, k)}{S_x}\right] \tag{6}$$

4 Maximum Likelihood Estimation

The principle of maximum likelihood was a technique initially developed by R. A. Fisher in 1920 [13]. It is a methodology for estimating parameters consolidated within statistics, as its estimators have characteristics that show a good performance. This method consists of finding the set of parameters that maximizes the likelihood function of the population samples. The likelihood function consists of the joint probability density function of the random variables that represent the samples.

In the study of fatigue models, Pascual and Meeker [14] proposed a 5-parameter probabilistic model using the MLE (Maximum Likelihood Estimation). Pascual [15] Proposed an MLE model considering experimental multi-factors. The method determines the S-N curve that describes the most likely location of each test result, including runout. Sarkani *et al.* [12] it presented a generalized methodology based on MLE for inclusion of runouts, where it is possible to consider different probability distributions for life to fatigue. An adopted model was based on the Weibull distribution, was demonstrated and applied to a set of real data to estimate the median curves and confidence intervals.

The MLE method consists of maximizing the likelihood of a set of parameters calculated to fit the model. For a given set of independent x_i observations obtained at different levels y_i, the likelihood function is given by eq. 7:

$$\Gamma(\boldsymbol{\theta}) = \prod_{i=1}^{n} \left[f_x(x_i; y_i, \boldsymbol{\theta}) \right]^{\delta_i} \cdot \left[1 - F_x(x_i; y_i, \boldsymbol{\theta}) \right]^{1-\delta_i} \tag{7}$$

Where $\delta_i = 1$ to fail and $\delta_i = 0$ runout; $f_x(x_i; y_i, \boldsymbol{\theta})$ and $F_x(x_i; y_i, \boldsymbol{\theta})$ are, respectively, the probability density function and the cumulative probability distribution of x_i ; $\boldsymbol{\theta}$ is the vector of model parameters.

In general, it is easier to work with the likelihood function in a logarithmic domain, because in this form the maximization is done by a sum instead of a product. Thus, the log-likelihood function $L_i(z_i; \boldsymbol{\theta})$ for a single test point in the sample is given by eq. 8:

$$L_i(z_i; \boldsymbol{\theta}) = \delta_i \cdot \left\{ \ln\left[f_x(x_i; y_i, \boldsymbol{\theta}) \right] \right\} + (1-\delta_i) \cdot \ln\left[1 - F_x(x_i; y_i, \boldsymbol{\theta}) \right] \tag{8}$$

Thus, the log-likelihood function for the entire data set, $\Lambda(\boldsymbol{\theta}) = \ln\left[\Gamma(\boldsymbol{\theta})\right]$, is given by the contribution of each point in the sample, according to eq. 9.

$$\Lambda(\boldsymbol{\theta}) = \sum_{i=1}^{n} L_i(z_i; \boldsymbol{\theta}) \tag{9}$$

The vector with the parameters $\boldsymbol{\theta}$ can be solved by applying optimization algorithms in order to maximize the function $\Lambda(\boldsymbol{\theta})$. For this, a subroutine was developed using a native function known as “ga” available in MATLAB Global Optimization Toolbox ™ [16] and the “fmincon” available in MATLAB Optimization Toolbox ™ [7] in which it presents the Interior Point Algorithm.

These algorithms were used in order to minimize $-\Lambda(\boldsymbol{\theta})$ (equivalent to maximize $\Lambda(\boldsymbol{\theta})$) when initial restrictions are given to $\boldsymbol{\theta}$, in order to limit the search space within a doable region. The constraints (constraints) of the parameters can be introduced directly into the algorithm through a linear inequality, according to the optimization eq. 10 expressed below:

$$\min_{\boldsymbol{\theta}} \left\{ -\Lambda(\boldsymbol{\theta}) \text{ tal que } \mathbf{A} \cdot \boldsymbol{\theta} \leq \mathbf{b} \right\} \tag{10}$$

Where $\mathbf{A}$ is a matrix and $\mathbf{b}$ is a vector created to establish restrictions, based on the doable domain of each parameter. The generalization of the constraint matrix is shown in the equation being applied to both algorithms in eq. 11:

$$\begin{cases} \theta_1^{\inf} \leq \theta_1 \leq \theta_1^{\sup} \\ \theta_2^{\inf} \leq \theta_2 \leq \theta_2^{\sup} \\ \vdots \\ \theta_k^{\inf} \leq \theta_k \leq \theta_k^{\sup} \end{cases} \Leftrightarrow \begin{cases} \theta_1 \geq \theta_1^{\inf} \\ \theta_1 \leq \theta_1^{\sup} \\ \theta_2 \geq \theta_2^{\inf} \\ \theta_2 \leq \theta_2^{\sup} \\ \vdots \\ \theta_k \geq \theta_k^{\inf} \\ \theta_k \leq \theta_k^{\sup} \end{cases} \Leftrightarrow \begin{cases} -\theta_1 \leq -\theta_1^{\inf} \\ \theta_1 \leq \theta_1^{\sup} \\ -\theta_2 \leq -\theta_2^{\inf} \\ \theta_2 \leq \theta_2^{\sup} \\ \vdots \\ -\theta_k \leq -\theta_k^{\inf} \\ \theta_k \leq \theta_k^{\sup} \end{cases}$$

$$\begin{cases} -1\cdot\theta_1+0\cdot\theta_2+0\cdot\theta_3+\ldots+0\cdot\theta_k \le -\theta_1^{\inf} \\ 1\cdot\theta_1+0\cdot\theta_2+0\cdot\theta_3+\ldots+0\cdot\theta_k \le \theta_1^{\sup} \\ 0\cdot\theta_1-1\cdot\theta_2+0\cdot\theta_3+\ldots+0\cdot\theta_k \le -\theta_2^{\inf} \\ 0\cdot\theta_1+1\cdot\theta_2+0\cdot\theta_3+\ldots+0\cdot\theta_k \le \theta_2^{\sup} \\ \vdots \\ 0\cdot\theta_1+0\cdot\theta_2+0\cdot\theta_3+\ldots+1\cdot\theta_k \le -\theta_k^{\inf} \\ 0\cdot\theta_1+0\cdot\theta_2+0\cdot\theta_3+\ldots+1\cdot\theta_k \le \theta_k^{\sup} \end{cases} \Leftrightarrow \begin{bmatrix} -1 & 0 & 0 & \ldots & 0 \\ 1 & 0 & 0 & \ldots & 0 \\ 0 & -1 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & 0 & \cdots & -1 \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix} \cdot \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \vdots \\ \theta_k \end{bmatrix} \le \begin{bmatrix} -\theta_1^{\inf} \\ \theta_1^{\sup} \\ -\theta_2^{\inf} \\ \theta_2^{\sup} \\ \vdots \\ -\theta_k^{\inf} \\ \theta_k^{\sup} \end{bmatrix}$$

$$\mathbf{A} = \begin{bmatrix} -1 & 0 & 0 & \ldots & 0 \\ 1 & 0 & 0 & \ldots & 0 \\ 0 & -1 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & 0 & \cdots & -1 \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix}; \quad \boldsymbol{\theta} = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \vdots \\ \theta_k \end{bmatrix}; \quad \mathbf{b} = \begin{bmatrix} -\theta_1^{\inf} \\ \theta_1^{\sup} \\ -\theta_2^{\inf} \\ \theta_2^{\sup} \\ \vdots \\ -\theta_k^{\inf} \\ \theta_k^{\sup} \end{bmatrix} \tag{11}$$

5 Results Obtained

The input dates used in this paper was obtained by Taras and Greiner [17], Pedrosa *et al.* [18] and Mayorga *et al.* [19] in a experimental campaign in some bridges in the workdwide. In this work, tests by different authors were compiled, in which the failure points and runout points of structural details in real scale of bridges with riveted connections are specified. With the application of the MLE method in relation to the optimization algorithms, new parameters were obtained for m, k e S_x.

Figure 2 presents the median curves and the S-N-p curves of 2.25% and 97.5% of probability of failure, highlighting the generation of the curves the parameters found by the estimator according to the IP optimization algorithm. Figure 3 shows the results generated by GA, which was obtained for 167 generations, as shown in figure 4.

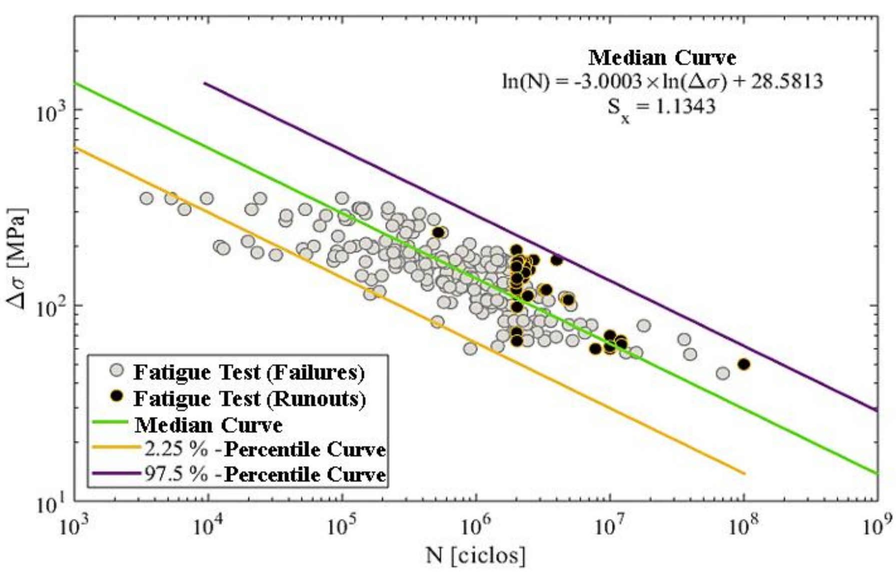

Figure 2. S-N-p Curves generated by IP.

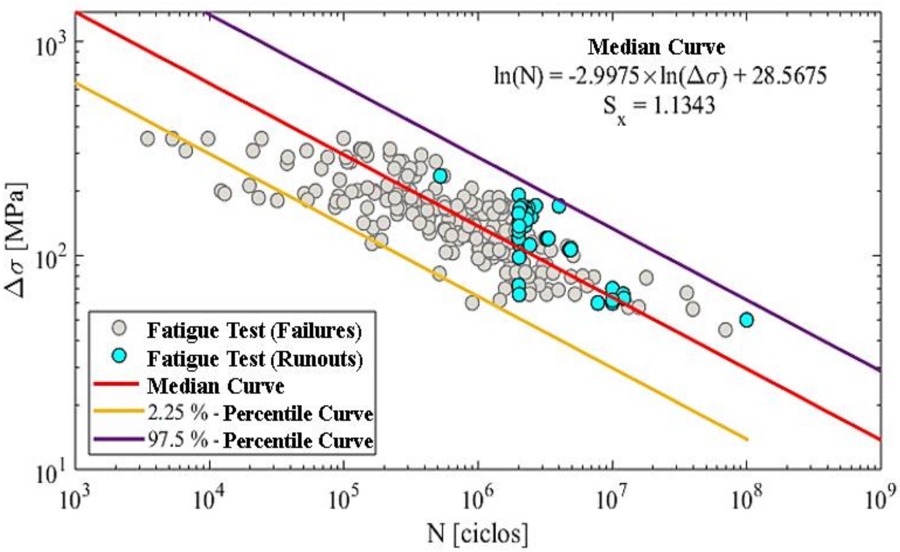

Figure 3. S-N-p Curves generated by GA.

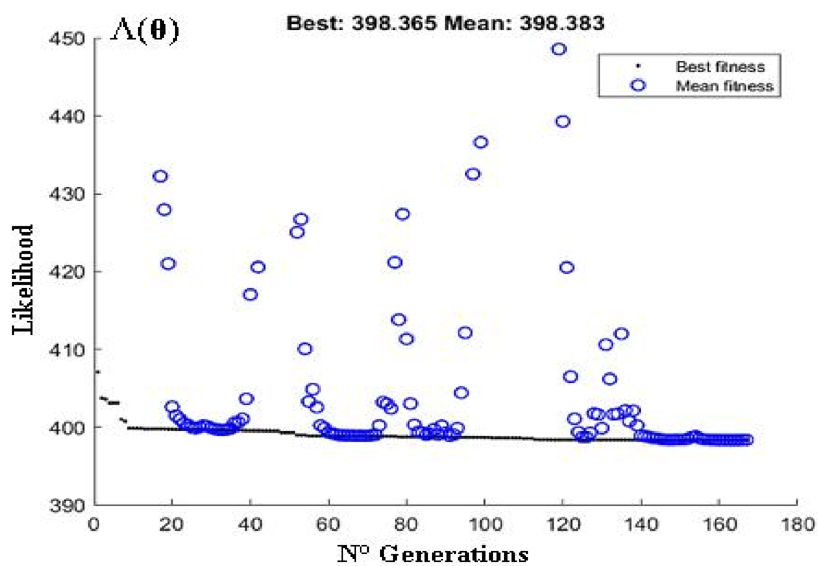

Figure 4. 167 generations carried out by GA.

Table 1 shows the comparison of the optimal parameters obtained between the optimization algorithms. There is a very small difference between the results, showing a good convergence of the MLE models.

Table 1. Comparison of the obtained parameters.

	k	m	S_x	Exp(k)
MLE - IP	28.5813	-3.0003	1.1343	$2{,}58.10^{12}$
MLE - GA	28.5675	-2.9975	1.1343	$2{,}55.10^{12}$
Difference (%)	0.048	0.093	0	1.162

6 Conclusions

This was an initial approach to the application of a genetic algorithm to determine S-N-p curves for riveted details. GA showed good results in estimating the parameters found. The S_x parameter coincided in both algorithms. Note that the biggest difference in relation to the search for the results obtained by the interior point algorithm is in relation to the transformation of k for the N domain with 1.162%. Although in absolute terms this value is small, this greater difference in relation to the other parameters is due to the great variability that exists in the N domain, and how sensitive this parameter is for determining fatigue life.

Acknowledgements. The authors of this work are greatful with the Federal University of Rio de Janeiro (UFRJ) and Federal of University of Pará (UFPA) to encourage this research and the CNPq by the finnancial support given.

Authorship statement. The authors hereby confirm that they are the sole liable persons responsible for the authorship of this work, and that all material that has been herein included as part of the present paper is either the property (and authorship) of the authors, or has the permission of the owners to be included here.

References

[1] A. L. L Silva. “Advanced Methodologies for the Fatigue Analysis of Representative Details of Metallic Bridges”. PhD Thesis. University of Porto, 2015.
[2] R. O. Almeida, “Ruptura de Longarinas Metálicas nas Pontes da Superintendência Regional Curitiba no trecho Paranaguá-Curitiba – Um caso de Colapso Devido à fadiga”. Joinville, XXXIIº Reunião dos Residentes – RFFSA, 1990.
[3] P. Zörnig, “Nonlinear Programming”. De Gruyter Textbook. De Gruyter, 2014.
[4] S. S. Rao, “Engineering Optimization: Theory and Practice”. Roboken, New Jersey:John Wiley & Sons., 2009.
[5] N. Karmarkar, “A new polynomial-time algorithm for linear programming”. Combinatorica . n. 4 , vol.4, pp. 373--395 , 1984.
[6] R. C. L. F. Oliveira, “Otimização Linear – Método de Pontos Interiores em Programção Linear”. Notas de Aula. Faculdade de Engenharia Elétrica e Computação. UNICAMP, Campinas, SP, Brasil, 2019.
[7] Mathworks, “Optmization Toolbox User’s Guide (Release 2010b)”. Mathworks Inc., Natick, MA, USA, 2010.
[8] T. L Fonseca; “Algoritmo Genético com Regressão: Busca Direcionada Através de Aprendizado de Máquina”. Dissertação de Mestrado, Programa de Pós-Graduação em Mecânica Computacional, UFJF, Juiz de Fora, MG, Brasil, 2014.
[9] M. Franulovic, P. Basan and I. Prebil, “Genetic Algorithm in Material Model Parameters Identification for Low-Cycle Fatigue”. Computational Materials Science, vol. 45, pp. 505 – 510., 2009. http://doi:10.1016/j.commatsci.2008.11.012.
[10] AREMA, “American Railway Engineering and Maintenance-of-Way Association”. Manual for Railway Engineering. Lanham, vol. 2., 2013.
[11] EN1993-1-9. Eurocode 3: Design of Steel Structures — Part 1–9: Fatigue. CEN, Bruxelas, Bélgica, 2005.
[12] S. Sarkani, T. A. Mazzuchi, D. Lewandowski, and D. P. Kihl, “Runout Analysis in Fatigue Investigation. Engineering Fracture Fechanics”, 74 (18), pp. 2971 – 2980., 2007.
[13] M. F. Rosário, “120 Anos do Nascimento do Cientista R. A. Fisher”. Rev. Bras. Biom., vol. 27, n.4, pp. 659 – 672, 2009.
[14] F. G. Pascual and W. Q. Meeker, “Estimating Fatigue Curves with the Random Fatigue – Limit Model”. Technometrics., 1999.
[15] F. G. Pascual, “The Random Fatigue – Limit Model in Multi Factors Experiments”. Journal of Statistical Computation and Simulation., 2010. https:// doi: 10.1080/0094965031000078882.
[16] Mathworks, “Global Optimization Toolbox User’s Guide (Release 2019b)”. Mathworks Inc., Natick, MA, USA., 2019.
[17] A. Taras and R.Greiner, “Development and Application of a Fatigue Class Catalogue for Riveted Bridge Components”. Institute for Steel Structures and Shell Structures, Graz University of Technology, Graz, Austria. (Report ECCS TC6 – 2010) Riveted Components, pp. 1–27., 2010.
[18] B. Pedrosa, J. A.F.O. Correia, C. Rebelo, G. Lesiuk, A. M.P. de Jesus, A. A. Fernandes, M. Duda, R. Calçada and M.Veljkovic, “Fatigue resistance curves for single and double shear riveted joints from old portuguese metallic bridges”. Engineering Failure Analysis, n. 96, pp. 255 – 273, 2019.
[19] L. G. Mayorga, S. Sire, J. A.F.O. Correia, Abilio, M.P. de Jesus, C. Rebelo, A. F. Canteli, M. Ragueneau and B. Plu, “Statistical evaluation of fatigue strength of double shear riveted connections and crack growth rates of materials from old bridges”. Engineering Fracture Mechanics, vol. 185, pp. 241 – 257, 2017.