

GENERALIZED RENEWAL PROCESS PARAMETERS ESTIMATION VIA MAXIMUM LIKELIHOOD AND GENETIC ALGORITHM

Beatriz Sales da Cunha

cunha.beatrizsales@gmail.com

Monalisa Cristina Moura dos Santos

monalisamoura24@gmail.com

Rafael Luiz Velôzo Santiago

rafa.velozo@hotmail.com

Márcio das Chagas Moura

marcio@ceerma.org

Isis Didier Lins

isis.lins@ceerma.org

Center for Risk Analysis, Reliability and Environmental Modeling (CEERMA), Department of
Production Engineering, Federal University of Pernambuco, Recife, Brazil
Avenida da Arquitetura, s/n - Cidade Universitária, Recife - PE, CEP: 50740-550

ABSTRACT

Although “as good as new” and “as bad as old” assumptions are usually adopted to describe the system state after a maintenance, they occasionally do not have a practical use. Generalized Renewal Process (GRP) is a point process capable of modeling all classes of repair, including the imperfect repair, which is a more realistic assumption. In this paper, we propose an approach to find the estimators of the Weibull-GRP parameters, by maximizing the likelihood function using Genetic Algorithm (GA). We compare the proposed solution with results in the literature, obtained through Monte Carlo Simulation (MCS). Results show that our methodology reach a smaller error through MSE criterion and GRP present a better fitting to the real data, when compared with the Renewal Process (RP) and the Non-Homogeneous Poisson Process (NHPP).

KEYWORDS. Generalized renewal process. Maximum likelihood estimators. Genetic algorithm.

Paper topics: SIM, OC.

1. Introduction

Estimating the number of failures of a given system is a task of reliability studies; as it requires collecting data from past failures in order to formulate probabilistic models to predict the next ones. When it comes to repairable systems, its state after repair may assume one of the following possibilities: ‘as good as new’, ‘as bad as old’, ‘better than old worse than new’, ‘better than new’, or ‘worse than old’. Traditional probabilistic models are generally focused on estimating the first two states, which do not properly describe a common system behavior. Hence, the generalized renewal process (GRP) is a robust probabilistic model to simulate all repair situations [Yañez et al. 2002] representing maintenance circumstances close to real.

When a large amount of data about the failure behavior is not available, the GRP’s parameters can be estimated through the Maximum Likelihood (ML) method. Yañez et al. (2002) proposed a methodology to obtain Maximum Likelihood estimators (MLE) by solving a complex system of equations obtained from the derivatives of the log-likelihood. As some of those equations did not have a closed mathematical solution form, they proposed a limited random search algorithm based on a simple Monte Carlo Simulation (MCS) to solve the problem. Wang & Yang (2012) proposed a nonlinear constrained programming to obtain MLE, where the objective function is the negative log-likelihood function, and the constrained variables are the model parameters. The solution of the optimization problem was obtained with the exterior penalty method [Yeniay 2005].

Moura et al. (2014) proposed a model for dependent and imperfect condition-based preventive and corrective maintenance. This model combined alarm repair alert, a competing risk-based model and GRP. Also, in this field, Oliveira et al. (2016) developed a simple goodness of fit test (GOFT) and applied it on a Weibull GRP. In addition they simulated a Uniform-based GRP in order to compare results with the Weibull GRP in terms of the GOFT. These models were tested in a case study of offshore, windshield and transformer facilities.

In this context, this paper uses a Genetic Algorithm (GA) to maximize the log-likelihood function and estimate the failure distribution parameters for systems. Firstly, we defined the log-likelihood function, considering that the first failure in GRP follows a Weibull distribution, and, then estimated the distribution parameters via GA. Thereby, the expected number of failures was obtained with MCS. The estimation of failure parameters was obtained to others stochastic process, e.g. the Renewal Process (RP) and the Non-Homogeneous Poisson Process (NHPP), for comparison purposes. Results will be presented for real equipment data from the literature.

The rest of this paper is organized as follows: Section 2 presents a brief theoretical background, with the main concepts adopted by this paper regarding GRP and GA. In Section 3, the proposed methodology is presented, including MLE’s equations and GA parameters. Finally, Section 4 presents the validation results and Section 5 concludes remarks.

2. Theoretical background

Typically, maintenance actions restore a repairable equipment to one of the following conditions: (i) “as good as new” - perfect repair, (ii) “as bad as old” - minimal repair, or (iii) “better than old but worse than new” - imperfect repair. The first two conditions are modelled through a Renewal Process (RP) and a Non-Homogeneous Poisson Process (NHPP), respectively [Ross 2014]. However, those situations have simplifying assumptions and occasionally do not have practical use [Lins & Drogue 2011]. So, GRP come as a method to model imperfect repair conditions and approximate the simulated problem to a realistic situation.

2.1 Generalized Renewal Process

In order to solve the limitations of RP and NHPP, [Kijima & Sumita 1986] developed a probabilistic method based on the concept of virtual age, that enables the modelling of all repair stages presented above, known as GRP. The model incorporates a parameter q , called the rejuvenation parameter, that makes the model more flexible. As seen in Santana et al. (2018), here we will consider that maintenance actions only act on the degradation over x_i , time between the $(i - 1)^{th}$ repair and the i^{th} failure (Kijima Type I). The calculation of the virtual age v_i is shown in Eq.1.

$$v_i = v_{i-1} + qx_i = q \sum_{j=1}^i x_j \quad (1)$$

As q approaches 0, the better is the quality of repair, i.e., the equipment is almost completely restored to a “as good as new” condition. Analogously, as q approaches 1, the maintenance action approaches minimal repair. At the limit we can define [Santana et al. 2018]:

- $q = 0$ represents a perfect repair;
- $q = 1$ corresponds to a minimal repair;
- $0 < q < 1$ indicates imperfect repair.

The CDF for the time between the $(i-1)^{th}$ and i^{th} failures can be determined following the definition of conditional probability (Eq. 2).

$$F(x_i|v_{i-1}) = P(X \leq x_i | X > v_{i-1}) = \frac{F(v_{i-1} + x_i) - F(v_{i-1})}{1 - F(v_{i-1})} \quad (2)$$

Due to the flexibility to fit various degradation stages, the time to first failure will follow a Weibull distribution. Thus, the conditional Weibull-GRP CDF is shown in Eq. 3, where α is the scale parameter and β is the shape parameter of the Weibull distribution.

$$F(x_i|v_{i-1}) = 1 - \exp \left[\left(\frac{v_{i-1}}{\alpha} \right)^\beta - \left(\frac{v_{i-1} + x_i}{\alpha} \right)^\beta \right] \quad (3)$$

Through the inverse transform method [Ross 2012], Eq.3 is transformed into the GRP-Weibull random generator of times until failures presented in Eq.4., where u is a uniformly distributed random number between 0 and 1.

$$x = \alpha \left[\left(\frac{v_{i-1}}{\alpha} \right)^\beta - \ln(u) \right]^{\frac{1}{\beta}} - v_{i-1} \quad (4)$$

In this paper, the Weibull parameters presented above, i.e. MLE, will be estimated via GA by maximizing the log-likelihood function, that will be further discussed in Section 3. The MLE will feed a discrete event simulation (DES) algorithm to estimate the number of failures by replicating the system behavior through the random generation of failure times (Eq. 4) and repair times.

2.2 Maximum likelihood estimators

The aim of this paper is to develop a GRP MLE for time terminated and failure terminated analysis. We define a case as time terminated, if it is intended to estimate the time of the next failure given an operation time t_{op} after the occurrence of the last failure, and before the occurrence of the next one. As an example, to illustrate the time terminated case, suppose the operation of an equipment during the time horizon of $T = 1000h$. The exact moments of each failure were: $t_1 = 100h$, $t_2 = 400h$, $t_3 = 800h$, $t_4 = 900h$, where t_i represent the accumulated time of i^{th} failure. No failure was observed between the interval $(900h;1000h]$. To predict t_5 considering all previous failures and the time interval $t_{op} = 100h$ between the last failure and T , we perform a time terminated analysis. Similarly, for failure terminated case, the prediction is done at the moment of the last failure. Next, we are going to present likelihood function for each case. As a form of simplifying the algorithm calculations, we used the parameterization made by Wang & Yang (2012) shown in Eq. 5.

$$\lambda = \alpha^{-\beta} \quad (5)$$

2.2.1 Time terminated case

For a single system, it is applied the parametrization described in Eq. 5. To define the likelihood function of time terminated case we should include terms related to the conditional reliability of the system at the time T where t_n is the accumulated time of the last failure observed at the time horizon. The definition of likelihood is shown in Eq. 6.

$$L = f(x_1) \prod_{i=2}^n f(x_i|v_{i-1}) R(T|t_n) \quad (6)$$

Then taking the log-likelihood from the CDF of Eq.3, we arrive at the expression in Eq.7, which is the function to be optimized through GA in order to find the time terminated MLE's $\hat{\lambda}, \hat{\beta}, \hat{q}$. To access the complete algebraic procedure, see Wang & Yang (2012).

$$\ln L = n(\ln \lambda + \ln \beta) - \lambda[(T - t_n + q \times t_n)^\beta - (qt_n)^\beta] - \lambda \sum_{i=1}^n [(x_i + qt_{n-1})^\beta - (qt_{n-1})^\beta] + (\beta - 1) \sum_{i=1}^n [\ln(x_i + qt_{n-1})] \quad (7)$$

2.2.2 Failure terminated case

In this case, the estimate is made at the exact moment of the last failure, so there is no additional time t_{op} after the occurrence of the last fault. Thus, through the definition of likelihood Eq. 6 is reduced to Eq. 8.

$$L = \prod_{i=1}^n f(x_i | v_{i-1}) \quad (8)$$

Similarly, applying the parametrization contained in Eq. 5 then taking the log-likelihood from the CDF in Eq.3, we arrive at the expression in Eq. 9. Note that the expression $T - t_n$ disappear because the estimation is made at the exact moment of the last fault, where $T - t_n = 0$. Eq.9. presents the function to be optimized through GA to obtain failure terminated MLE's $\hat{\lambda}, \hat{\beta}, \hat{q}$.

$$\ln L = n(\ln \lambda + \ln \beta) - \lambda \sum_{i=1}^n [(x_i + qt_{n-1})^\beta - (qt_{n-1})^\beta] + (\beta - 1) \sum_{i=1}^n [\ln(x_i + qt_{n-1})] \quad (9)$$

2.3 Genetic algorithm

Optimization problems typically entails (i) a wide search space, (ii) the presence of different data types, and (iii) computational complexity of objective functions [Seidian et al. 2016]. Solving these kinds of problems using exhaustive procedures, such as Monte Carlo Simulation, is time consuming and computationally intensive. In this context, heuristic methods can overcome such complications by finding possible optimal solutions in a reasonable time, although globally optimal solutions are not ensured since they are probabilistic methods.

Among the heuristic methods, Genetic Algorithms (GAs) are inspired by Darwin's theory of evolution and attempt to mimic the natural evolution process by the use of genetic operators, such as: selection, crossover, and mutation. The idea of GA is that an individual is a possible solution for the problem in question, represented by its phenotype, containing the values for each variable of the problem. Individuals are randomly initialized, forming a population, and the genetic operators are used to explore new solution spaces and prevent premature convergence of results [Altıparmak et al. 2009].

In this paper, we are interested in estimating the failure distribution parameters for a given equipment by maximizing the likelihood function, which will be detailed in the next section, using GA. In our case, the individuals' phenotype is composed by the estimates for the distribution parameters. As a form of validation of the MLE obtained with GA, we intend to simulate the system and compare with real failure data. Additionally, we intend to compare our results with the ones obtained by Yañez et al. (2002) through the criteria of Mean Squared Error (MSE).

3. Methodology

Unlike the methodology proposed by [Yañez et al. 2002], that had to deal with a complex system of equations, this paper intend to maximize the likelihood directly using GA. The steps to obtain the MLE $\hat{\lambda}, \hat{\beta}, \hat{q}$ of GRP and validate the methodology are in Fig. 1. Firstly, the log-likelihood function for time terminated and failure terminated cases will be obtained. Posteriorly, each function will be optimized via GA with respect GRP parameters. The output, i.e. $\hat{\lambda}, \hat{\beta}$, and $\hat{q}$, will feed a DES algorithm to reproduce system's operation and obtain important measures, as the expected number of failures and times between failures, which will be compared with the real

observed data. Finally, our methodology will be compared with the results of the MLE through MCS by the criteria of Mean Squared Error (MSE).

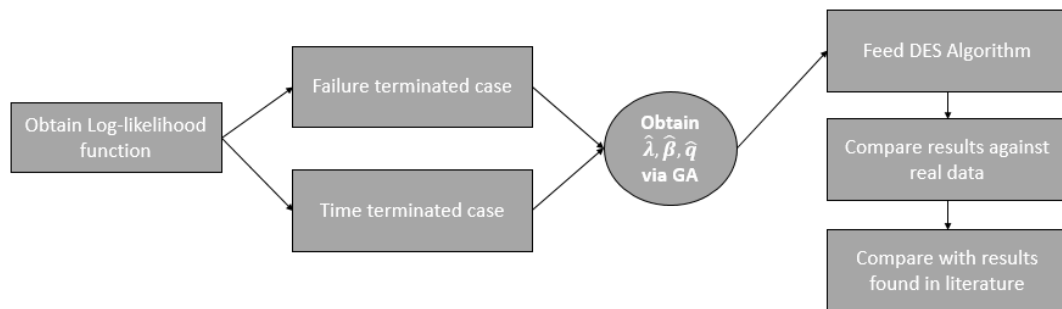


Figure 1: Proposed methodology. Source: this paper.

As mentioned before, the optimization objective is to maximize the log-likelihood function. In this case, the individuals' phenotypes are composed by $\hat{\lambda}, \hat{\beta}, \hat{q}$ and individuals' fitness are defined by Eq. 7 and Eq. 9 for the time terminated and the failure terminated cases, respectively. The specific GA parameters are defined in Table 1, where the selection criterion indicates how individuals are selected in each generation to participate in the next one. The elitist criterion indicates that the individuals with greater fitness remain for the next generation.

Table 1: Example of specific GA parameters.

Parameter	Value
Number of generations	2000
Population (individuals per generation)	2000
Crossover probability	0.95
Mutation probability	0.1
Selection criterion	Elitist

Source: this paper.

4. Results

To attest the effectiveness of the proposed solution method we will perform both a time and a failure terminated analysis using two datasets of real equipment and we will make a comparison with previously obtained results in literature. Times between failures (X_i) for the time terminated and the failure terminated case, obtained from Yañez et al. (2002), are shown in Tables 2 and 3, respectively.

Table 2: Time terminated data

Table 3: Failure terminated data

i	X_i	i	X_i
1	860	13	367
2	1,608	14	2,758
3	1,134	15	355
4	2,703	16	1,084
5	645	17	855
6	95	18	280
7	1,278	19	490
8	605	20	945
9	344	21	105
10	1,054	22	127
11	680	23	61
12	405	24	326

Source: [Yañez et al. 2002]

i	X_i	i	X_i
1	3,456	13	360
2	1,584	14	998
3	236	15	656
4	516	16	180
5	1,820	17	244
6	452	18	1,528
7	432	19	44
8	1,264	20	3,064
9	3,072	21	324
10	384	22	1,528
11	2,448	23	348
12	32	24	336

Source: [Yañez et al. 2002]

These data were incorporated at the likelihood functions in Eq.7 and Eq. 9 and optimized through GA with parameters described in Section 3.2. The MLE for each stochastic process – i.e., GRP, NHPP, and RP – are showed in Table 4 (time terminated) and 5 (failure terminated).

Table 4: Time terminated MLE's

	GRP	NHPP	RP
$\hat{\lambda}$	0.00000025	0.00000854	0.00032028
$\hat{\beta}$	2.02603741	1.50342690	1.19098285
$\hat{q}$	0.14582259	1	0

Source: this paper

Table 5: Failure terminated MLE's

	GRP	NHPP	RP
$\hat{\lambda}$	0.00000435	0.00192509	0.00293258
$\hat{\beta}$	1.54603724	0.93140092	0.84863035
$\hat{q}$	0.72728872	1	0

Source: this paper

Following the procedure described by Yañez et al. (2002), estimations of the expected number of failures $E[N]$ were obtained for each case through DES algorithm with 100,000 replications, considering that failure intervals were generated with Eq.4. To attest the quality of the estimators we made a graphical comparison between the simulation and the real data by plotting the i_{th} failure number versus the i_{th} accumulated time t_i . Results are shown in Figures 2 (a) and (b) for the time terminated and the failure terminated case, respectively.

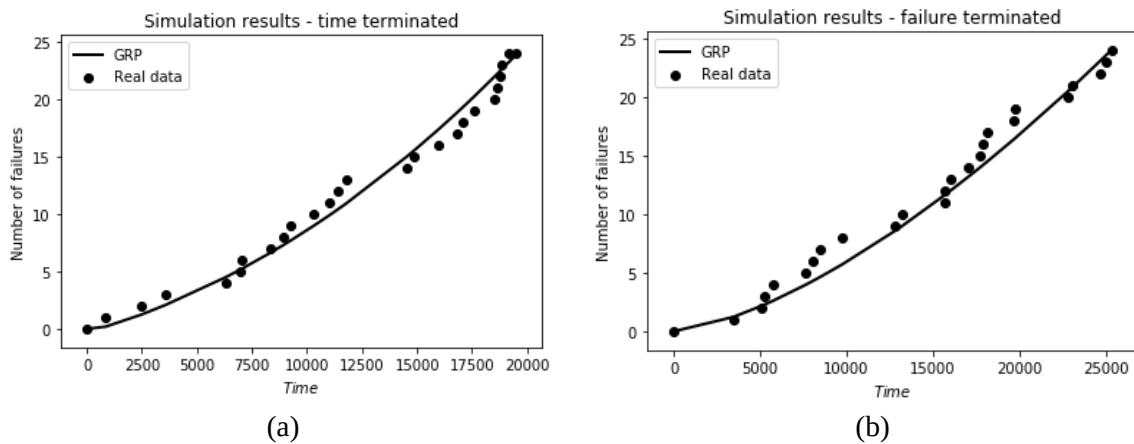


Figure 2: Simulation results x real data for (a) time terminated and (b) failure terminated cases.
Source: this paper

To compare our methodology with the one proposed by Yañez et al. (2002) – a GRP-Weibull simulation performed via MCS estimation - the mean squared error (MSE) (Tables 6 and 7) between simulation results and real data was computed. By MSE comparison, it is seen a better adjustment through GRP (imperfect repair) assumption. Furthermore, our methodology, through a simpler way, achieved better results than those found by Yañez et al. (2002).

Table 6: MSE for time terminated results

NHPP	1.218
RP	2.447
GRP	1.086
Yañez et al. (GRP)	1.095

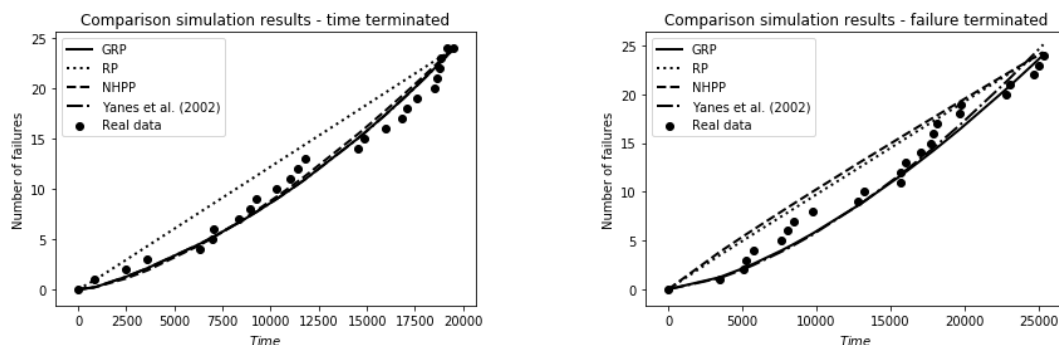
Source: this paper

Table 7: MSE for failure terminated results

NHPP	2.412
RP	2.047
GRP	1.282
Yañez et al. (GRP)	1.317

Source: this paper

A visual comparison between the processes and methodologies described in this paper with the real data is seen in Figure 3 (a) and (b). Since the errors found by our methodology and Yañez et al. (2002) are close, the green dots and solid line almost overlap in both figures. It is seen also that RP and NHPP for failure terminated case overestimate the expected number of failures, this overestimation occurs at RP model for time terminated case.



(a) (b)

Figure 3: Comparison between the simulated results and real data for the (a) time terminated and the (b) failure terminated cases. Source: This paper.

5. Concluding Remarks

In this paper we proposed a method to estimate the MLE's of a Weibull-GRP, a flexible stochastic process that allows the modeling of imperfect repairs, through a GA. The use of GA enabled the direct optimization of the likelihood function instead of solving a complex system of equations. Additionally, we also estimated the parameters and extend the analysis to NHPP and a RP assumption (more simplifying assumptions), two stochastic processes widely used in the literature.

Simplifying assumptions with respect to the state of the system after a maintenance intervention can lead to bad estimations, erroneous managerial decisions and losses for the companies. At this paper, is seen that the GRP presented better adjustment of the data if compared to NHPP and RP. Also, comparing our results with Yañez et al. (2002) who used MCS to obtain MLE's, we obtained better results for Weibull-GRP model, which represents a higher level of accuracy achieved by the estimation of the MLE's parameters through the GA.

6. Acknowledgements

The authors would like to thank the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), and the Fundação de Amparo a Ciência e Tecnologia de Pernambuco (FACEPE) for the financial support through research grants. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.]

References

- Altiparmak, F., Gen, M., Lin, L., & Karaoglan, I. (2009). A steady-state genetic algorithm for multi-product supply chain network design. *Computers & Industrial Engineering*, 56(2): 521–537.
- Kijima, M., & Sumita, U. (1986). A Useful Generalization of Renewal Theory: Counting Processes Governed by Non-Negative Markovian Increments. *Journal of Applied Probability*.
- Lins, I. D., & Droguett, E. L. (2011). Redundancy allocation problems considering systems with imperfect repairs using multi-objective genetic algorithms and discrete event simulation. *Simulation Modelling Practice and Theory*, 19(1): 362–381.
- Moura, M.C., Droguett, E.L. Firmino, P.R.A.Q. (2014). A competing risk model for dependent and imperfect condition-based preventive and corrective maintenances. *Journal of Risk and Reliability*, 6: 590-605.
- Oliveira, C.C.F., Cristino, C.T., Firmino, P.R.A.Q. (2016). In kernel of modeling repairable systems: a goodness of fit for Weibull-based generalized renewal process. *Journal of Cleaner Production*, 133: 358-367.
- Ross, S. M. (2012). *Simulation* (5th ed.). Academic Press.
- Ross, S. M. (2014). *Introduction to Probability Models: Eleventh Edition*. Introduction to Probability Models: Eleventh Edition.

Saeidian, B., Mesgari, M. S., & Ghodousi, M. (2016). Evaluation and comparison of Genetic Algorithm and Bees Algorithm for location-allocation of earthquake relief centers. *International Journal of Disaster Risk Reduction*, 15: 94–107.

Santana, J. M., Santiago, R. L. V., Moura, M. D. C., & Lins, I. D. (2018). Extended warranty of medical equipment subject to imperfect repairs: an approach based on generalized renewal process and Stackelberg game. *Eksploracja I Niezawodnosc*, 20(4): 567–578.

Wang, Z. M., & Yang, J. G. (2012). Numerical method for Weibull generalized renewal process and its applications in reliability analysis of NC machine tools. *Computers and Industrial Engineering*, 63(4): 1128–1134.

Yaez, M., Joglar, F., & Modarres, M. (2002). Generalized renewal process for analysis of repairable systems with limited failure experience. *Reliability Engineering and System Safety*, 77(2): 167–180.

Yeniay, O. (2005). Penalty function methods for constrained optimization with genetic algorithms. *Mathematical and Computational Applications*, 10(1): 45–56.