Hassen et al. Iraqi Journal of Science, 2023, Vol. 64, No. 6, pp: 3054-3065
DOI: 10.24996/ijs.2023.64.6.33

ISSN: 0067-2904

Estimating the Rate of Occurrence of Extreme value process Using Classical and Intelligent Methods with Application: nonhomogeneous Poisson process with intelligent

Adel.S.H. [1]*, Khwazbeen Saida Fatah [2], Muthanna Subhi Sulaiman [3]

[1] *IT Department, Amedi Technical Institutes, University of Duhok Polytechnic*
[2] *Department of Mathematics, College of Science, University of Salahaddin*
[3] *Department of Statistics, College of Computer and Mathematics, University of Mosul*

Received: 16/8/2022 Accepted: 10/10/2022 Published: 30/6/2023

Abstract

In this paper, the propose is to use the xtreme value distribution as the rate of occurrence of the non-homogenous Poisson process, in order to improve the rate of occurrence of the non-homogenous process, which has been called the Extreme value Process. To estimate the parameters of this process, it is proposed to use the Maximum Likelihood method, Method of Moment and a smart method represented by the Artificial Bee Colony:(ABC) algorithm to reach an estimator for this process which represents the best data representation. The results of the three methods are compared through a simulation of the model, and it is concluded that the estimator of (ABC) is better than the estimator of the maximum likelihood method and method of moment in estimating the time rate of occurrence of the proposed Extreme value process. The research also includes a realistic application that deals with the operating periods of two successive stops for the raw materials factory from the General Company for Northern Cement / Badush Cement Factories (new) during the period from 1/4/2018 to 31/1/2019, in order to reach the time rate of factory stops.

Keywords: Extreme-value process, ABC Algorithm, Maximum likelihood Estimator, Modified Moment Estimation, Simulation

تقدير معدل حدوث عملية القيمة المتطرفة باستخدام الأساليب الكلاسيكية والذكية مع التطبيق: عملية بواسون غير متجانسة مع الذكية

عادل سفيان حسين[1]*, خوازبين سيدا فتاح[2], مثنى صبحي سليمان[3]

[1] نظم معلومات, معهد عمادية, الجامعة دهوك التقنية, دهوك, العراق
[2] رياضيات, الكلية العلوم, الجامعة صلاح الدين, أربيل, العراق
[3] القسم الاحصاء ونظم معلوماتية, كلية علوم الرياضيات والحاسبات, جامعة الموصل, الموصل, العراق

* Email: Adel_83_1983@yahoo.com

الخلاصة

في هذا البحث ، تم أقتراح استخدام توزيع القيمة المتطرفة كمعدل لحدوث عملية بواسون غير المتجانسة ، من أجل تحسين معدل حدوث العملية غير المتجانسة ، والتي تسمى عملية القيمة القصوى. من أجل تقدير معلمات هذه العملية ، تم اقتراح استخدام طريقة الاحتمالية القصوى وطريقة اللحظة وطريقة ذكية ممثلة بخوارزمية (ABC) للوصول إلى مقدر لهذه العملية التي تمثل أفضل تمثيل للبيانات. تمت مقارنة نتائج الطرق الثلاث من خلال محاكاة النموذج ، وخلصت إلى أن مقدّر (ABC) أفضل من مقدر طريقة الاحتمالية القصوى وطريقة اللحظة في تقدير المعدل الزمني لحدوث النموذج المقترح. عملية القيمة القصوى. كما تضمن البحث تطبيقا واقعيا تناول فترات التشغيل لمحطتين متتاليتين لمصنع المواد الخام من الشركة العامة للاسمنت الشمالية / مصانع اسمنت بادوش (جديد) خلال الفترة من 2018/4/1 الى 31/1 / 2019 ، للوصول إلى المعدل الزمني لتوقفات المصنع.

1. Introduction

Nonhomogeneous Poisson Process (NHPP), which is considered the best-known generalization of the Poisson process models, it is mainly used for analysing and modelling the failure data in recoverable systems. These models assume that the point process $\{N(\tau),\tau\geq 0\}$ with independent increments is distributed as the Poisson distribution and the occurrence of events in the Poisson process is random and monotonous during a certain period and with a fixed incidence rate per unit time which is denoted by the symbol λ, while the rate at which events occur in a nonhomogeneous Poisson process is time-varying τ with the ratio called the time rate of occurrence or intensity function; it is denoted by the symbol $\lambda(\tau)$. The Extreme-value process, which is introduced by Extreme-value, is a special model for NHPP that has appeared as a special case of the Weibull distribution; it plays a key role in modelling and analysing failure data accumulated over time [1]. In this paper, Extreme-value distribution is introduced and different methods for estimating their parameters are presented. In this study, both ABC algorithm and the maximum Likelihood Estimation (MLE), and the method of Moment (MM) are used to estimate Extreme value Process parameters.

1.1 Extreme-Value Process (EVP)

We will assume that the Poisson process $\{Y(\tau), \tau \geq 0\}$ represents the NHPP; Since the number of events that occur over a period $(0,\tau)$follows the Poisson distribution as a function of probability density:

$$p[Y(\tau)=y]=\frac{[\lambda(\tau)]^{y}e^{-m(\tau_0)}}{y!} \quad , \ y=1,2,3,... \tag{1}$$

$m(\tau)$represents the process parameter (mean rate), which is the cumulative function of the time-of-occurrence rate, which is determined by the following formula [2]:

$$m(\tau)=\int_0^{\tau}\lambda(u)\,du \qquad , \ 0<\tau<\infty \tag{2}$$

because $\lambda(u)$ represents the time rate of occurrence or intensity function, therefore, the Extreme-value process is a nonhomogeneous Poisson process[2], with the time rate of occurrence is defined as follows:

$$\lambda(t)=\frac{1}{\sigma}e^{(\frac{t-\mu}{\sigma})} \quad ,-\infty\leq\ t\ \leq\ \infty;\ -\infty<\mu<\infty\ ,\sigma\ >0 \tag{3}$$

$$m(t)=\int_0^{t}\lambda(u)\,du \qquad , \ 0<t<\infty \tag{4}$$

$$=\int_{-\infty}^{t_0}\lambda(u)\,du$$

$$=\int_{-\infty}^{t_0}\frac{1}{\sigma}\ e^{\frac{u-\mu}{\sigma}}\,du$$

$$=e^{\frac{t_0-\mu}{\sigma}} \qquad ,-\infty\leq\ t\ \leq\ t_0 \tag{5}$$

σ, μ are the parameters of the accident rate for the Extreme-value Process; parameters estimation for such processes has been extensively studied and a large number of techniques have been proposed.

1.2 Artificial Bee Colony (ABC).

The Artificial Bee Colony (ABC) algorithm is a swarm-based on the developed metaheuristic algorithm for numerical problem optimization. Honey bees' creative foraging activity inspired it. The method is mainly based on the model for honey bee colony foraging Behavior. There are three categories of bees in the (ABC) algorithm, namely employed bees, onlooker bees, and scout bees. The employed bees hunt for food in their memory near the food source; in the meantime, they relay the information about these food sources to the onlooker bees. Onlooker bees tend to pick good food sources from those discovered by hired bees. The food supply with greater quality (fitness) will be more likely to choose onlooker bees than those with a lower rate. The scout bees are derived from a small number of employed bees that quit their food sources in quest of new ones. A colony of artificial forager bees (agents) hunts for rich artificial food sources in (ABC) (good solutions for a given problem). To apply (ABC) algorithm, the optimization problem at hand is first transformed into the problem of determining the optimum parameter vector that minimizes an objective function. The artificial bees then discover a population of initial solution vectors at random and then iteratively enhance them by utilizing the strategies: migrating towards better solutions via a neighbor search mechanism while abandoning inferior answers [3].

Foraging Behavior of Honey Bees

The minimal model of forage selection leads to the emergence of the collective intelligence of honeybee swarms consists of three basic components: food sources, employed and unemployed foragers. The model identifies two main modes of behavior: the recruitment of the wealthy. The source of nectar and give up the poor source.[4]

- **Food Sources:** the values of a food source depend on several factors, such as its proximity to the nest, its richness or concentration of its energy, and the ease with which this energy can be extracted. For simplicity, the profitability of a food source can be represented by a single quantity.

- **Employed foragers:** They are associated with a particular food source that they are currently exploited or employed. They carry with them information about this specific source to the hive. The information can be the distance and direction from the nest, sharing this information with a certain probability.

- **Unemployed foragers:** They are continually on the lookout for a food source to exploit. There are two types of unemployed foragers: scouts, searching the nest's environment for new food sources and onlookers waiting in the nest and establishing a food source through the information shared by employed foragers. The mean number of scouts averaged over conditions is about when the above explained foraging behavior of honey bees is reexamined, it is seen that the defined principles are fully satisfied.[5]

Algorithmic Structure of (ABC)

As in the minimal model of forage selection of natural honey bees, the colony ofartificial bees in (ABC) contains three groups of bees: Employed Bees associated with specific food

sources, Onlooker Bees watching the dance of employed bees within the hive to choose a food source, and Scout Bees search for the food sources randomly. Both onlookers and scouts are also called Unemployed Bees. Initially, all food source positions were discovered by scout bees [6]. After that, the nectar of the food sources is exploited by employed bees and onlooker bees, and this continual exploitation will ultimately cause them to become exhausted. Then, the employed bee becomes a scout bee which searches for additional food sources using the exhausted food source. In other words, the employed bee whose food source has been exhausted becomes a scout bee. In (ABC), the position of a food source represents a possible solution to the problem, and the nectar amount of a food source corresponds to the quality (fitness) of the associated answer. In the basic form, the number of employed bees is equal to the number of food sources (solutions) since each employed bee is associated with one and only one food source. The general algorithmic structure of the (ABC) optimization approach. [7]
In the initialization phase, the population of food sources (solutions) is initialized by artificial scout bees and set control parameters. In the employed bees phase, artificial employed bees search for new food sources having more nectar within the neighborhood of the food source in their memory. They find a neighbor food source and then evaluate its fitness. After producing the new food source, its fitness is calculated. A greedy selection is applied between it and its parent. After that, employed bees share their food source information with onlooker bees waiting in the hive by dancing in the dancing area. [8] In the onlooker bee's phase, artificial onlooker bees probabilistically choose their food sources depending on the information that is provided by the employed bees. For this purpose, a fitness-based selection technique can be used, such as the roulette wheel selection method. After a food source for an onlooker bee is probabilistically chosen, a neighborhood source is determined, and its fitness value is computed. As in the employed bee’s phase, a greedy selection is applied between two sources [9]. In the scout bee's phase, the employed bees whose solutions cannot be improved through a predetermined number of trials, which is called limit, become scouts, and their solutions are abandoned. Then, the scouts start to search for new solutions randomly. Hence, those sources which are initially poor or have been made destitute by exploitation are abandoned, and negative feedback Behavior arises to balance the positive feedback [9]. These three steps are repeated until a termination criterion is satisfied, for example, a maximum cycle number or a maximum (CPU) time. The artificial bee colony (ABC) algorithm is a recently proposed optimization technique that simulates honey bees' intelligent foraging Behavior. A set of honey bees is called a swarm that can accomplish tasks through social cooperation. In the ABC algorithm, the first half of the swarm consists of employed bees, and the second half constitutes the onlooker bees. The number of employed bees or the onlooker bees equal the number of solutions in the swarm. The (ABC) generates a randomly distributed initial population of (SN) solutions (food sources), where (SN) denotes the Swarm Size. Let $X_i = \{x_{i,1}, x_{i,2}, \dots, x_{i,D}\}$ represent the (i^{th}) solution in the swarm, where (D) is the dimension size. Each employed bee (X_i) generates a new candidate solution (V_i)[10].

1.3 Performance of estimation accuracy

In order to analyse the accuracy of the parameter estimation, different criteria were used. One of them is the Maximum Percentage Error (MPE); it is one of the measurements used to analyse the efficiency of the estimation by comparing different estimation methods. In this study, MPE was used to compare both MLE, MM methods and ABC algorithm applied to estimate the Extreme-value process parameters; this measure is considered as a fitness function for ABC algorithm. If S_i and $\hat{S}_i$ are defined as follows:

$$S_i = \sum_{j=1}^{i} Y_j \quad , \qquad and\ \hat{S}_i = \sum_{j=1}^{i} \hat{Y}_j \quad (6)$$

Then MPE is evaluated by:

$$\mathrm{MPE} = \sum_{1\le i\le n}^{\max} \left[\left|\mathrm{S_i} - \hat{\mathrm{S}}_\mathrm{j}\right| / \mathrm{S_i}\right] \quad (7)$$

2. Parameters Estimation

In this part, we discuss the estimation problem for the extreme-value process.

2.1 parameters Estimation for the Extreme-value Process using the Maximum Likelihood Estimator (MLE)

The Maximum Likelihood Estimator (MLE) is one of the most widely used methods for estimating the parameters of stochastic models due to its good properties, including stability and Minimum Variance Unbiased estimators, as the parameter values make the maximum likelihood function of the observations at its maximum limit.

Assume that $\{X(t), t \geq 0\}$ is a NHPP with the time rate of occurrences defined by the formula (3), then the joint probability function of the occurrence times $(t_1, t_2, \dots, t_n)$ in which $(0 < t_1 \leq t_2 \leq \cdots \leq t_n \leq t_0)$ is defined by the following [11]:

$$f(t_1, t_2, \dots, t_n) = \prod_{i=1}^{n} \lambda(t_i) e^{-m(t_0)} \tag{8}$$

From the (3), (5) equations, we substitute them into the (8) to get the joint probability function:

$$f(t_1, t_2, \dots, t_n) = \prod_{i=1}^{n} \frac{1}{\sigma} e^{\frac{t_i-\mu}{\sigma}} * \ e^{-e^{\frac{t_0-\mu}{\sigma}}} \tag{9}$$

The Likelihood function for the formula (8) for the period (0, t].

$$L = \prod_{i=1}^{n} \lambda(t_i) e^{-m(t_0)} \tag{10}$$

$$L = \frac{1}{\sigma^n} e^{\sum_{i=1}^{n} \frac{t_i-\mu}{\sigma}} * e^{-e^{\frac{t_0-\mu}{\sigma}}} \tag{11}$$

To simplify the calculations, the natural logarithm of the maximum function is taken instead of the maximum function itself represented by the equation (11). Hence, the following formula is obtained:

$$\ln L = \ln(1) - n\ln(\sigma) + \frac{1}{\sigma}\sum_{i=1}^{n} \tau_i - \frac{n\mu}{\sigma} - \ e^{\frac{\tau_0-\mu}{\sigma}} \tag{12}$$

Since the parameters μ and σ of the function $\lambda(\tau)$ both are known. Then the maximum likelihood estimator for them is found by finding the first derivative of the equation (12) the equations are equal to zero and my comparisons[12]:

$$\frac{\partial \ln L}{\partial \mu} = \frac{n}{\mu} - \frac{1}{\sigma} e^{\frac{\tau_0-\mu}{\sigma}}$$

$$\left.\frac{\partial \ln L}{\partial \mu}\right|_{\mu=\hat{\mu}} = 0$$

$$\frac{n}{\hat{\mu}} - \frac{1}{\hat{\sigma}} e^{\frac{\tau_0-\hat{\mu}}{\hat{\sigma}}} = 0 \tag{13}$$

Therefore, the maximum likelihood estimator for the parameter (μ) in an extreme-value model is it:

$$\hat{\mu}_{MLE} = \ln(n)\,\hat{\sigma} + \ \tau_0 \tag{14}$$

The distribution of parameter b can be inferred through the conditional distribution of the variable $S = \sum_{i=1}^{n} t_i$ conditional on the number of incidents n, the reason is that the observations in the potential function of the nonhomogeneous Poisson process only comes from $n, \sum_{i=1}^{n} t_i$. In order to find the maximum likelihood estimator for parameter b, we need to find the probability distribution for it, which represents the conditional distribution of the variable$S = \sum_{i=1}^{n} t_i$, conditional on the number of events n. To get the probability distribution of the variable S conditioned by the number of incidents n, this is done by dividing the potential function of the nonhomogeneous Poisson process in the formula (8) as follows [2]:

$$L[S|N(t) = n] = \frac{n!\,\frac{1}{\sigma^n}\, e^{-\frac{1}{\sigma}\sum_{i=1}^{n} t_i}}{\left(e^{\frac{t_0}{\sigma}}\right)^n} \ , \ \sigma \neq 0 \tag{15}$$

The log-likelihood function for formula (15) is expressed as follows:

$$\ln \mathrm{L} = \ln(\mathrm{n}!) - \mathrm{n} \ln \sigma + \frac{1}{\sigma}\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{t_i} - \frac{\mathrm{n}t_0}{\sigma} \ , \ \sigma \neq 0 \quad (16)$$

The derivative of the logarithm of the possible function concerning the parameter σ is found σ as follows:

$$\frac{\partial \ln \mathrm{L}}{\partial \sigma} = -\frac{\mathrm{n}}{\sigma} + \frac{1}{\sigma^2}\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{t_i} + \frac{nt_0}{\sigma^2} \qquad \sigma \neq 0 \quad (17)$$

As much as possible for the parameter σ it can be found by solving the following equation:

$$\left.\frac{\partial \ln \mathrm{L}}{\partial \sigma}\right|_{\sigma=\hat{\sigma}} = 0 \quad (18)$$

$$-\frac{\mathrm{n}}{\sigma} + \frac{1}{\hat{\sigma}^2}\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{t_i} + \frac{nt_0}{\sigma^2} = 0$$

$$\frac{1}{\sigma^2}\left(nt_0 - \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{t_i}\right) = \frac{\mathrm{n}}{\sigma}$$

Therefore,

$$\frac{1}{\sigma}\left(nt_0 - \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{t_i}\right) = n \ .$$

Therefore, the estimator of the maximum likelihood for parameter σ using the method of maximum likelihood is:

$$\hat{\sigma}_{MLE} = \frac{1}{n}\left(nt_0 - \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{t_i}\right) \quad (19)$$

From formulas (14) and (19), we will get:

$$\hat{\mu}_{MLE} = t_0 - \ln(n)\frac{1}{n}\left(nt_0 - \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{t_i}\right) \quad (20)$$

A program has been prepared in the programming language MATLAB/R2019b to find the maximum possible estimators for the parameters σ , μ of the time rate of occurrences of the Extreme-value stochastic process.

2.2 Estimation of Extreme-value Process by using ABC

In this section, extreme-value process parameters σ,μ that represent the parameter for the time rate of occurrences is estimated using ABC algorithm. For the estimation process, an algorithm, Algorithm (1), is proposed; it is described as follows:

Algorithm (1)

Step1: input parameter (N=50) and the number of iterations with $i_{max} = 100$.

Step2: defining objective & Fitness function is the MPE, in which $\mathrm{MPE} = \sum_{1\le \mathrm{i} \le \mathrm{n}}^{\max}\left[\left|\mathrm{S_i} - \hat{\mathrm{S}}_\mathrm{j}\right|/\mathrm{S_i}\right]$

Step3: Generate an initial population

Step4: Perform Employed and Onlooker phases

- Select variable and partner
- Generate New Solution $X_{new} = X + \emptyset\left(X - X_p\right), \emptyset \in [-1,1]$;
- Calculate new fitness.
- Perform Greedy Selection

Step5: Memorize the best solution

Step6: Perform the Scout phase

Step7: Plot the result

2.3 Parameters Estimation of Extreme value Process using Moment of Method (MM)

The process {x(t); t ≥ 0} represents the extreme value process rate of time in the equation (9). The first sample moment $m_1$is calculated by[13]:

$$m_1 = \frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{t_i} \quad (21)$$

To calculate the first population moment for an extreme process, we need to find the $E(X)$which is denoted by the $\mu_r$that can be easily written as:

$$\mu_r = E(X^r) = \sum_{k=0}^{r} \binom{k}{r} \sigma^{r-k} M^k \int y^{r-k} e^{-(y+e^{-y})} dy \quad (22)$$

In particular cases, namely r=1 and r=2, we get after computation and simplification:

$$\mu_n = E(X^n) = n\mu + n^2\sigma\gamma \quad (23)$$

The mean of the Ext can be obtained by putting n=1 in formula (23), we get:

$$\mu_1 = E(X^1) = \mu + \sigma\gamma \quad (24)$$

By equating the first sample moment with the first population moment, we obtain the parameter of the Ext process:

$$\mu_1 = m_1$$

$$\mu + \sigma\gamma = \frac{1}{n}\sum_{i=1}^{n} t_i$$

$$\mu = \frac{1}{n}\sum_{i=0}^{n} t_i - \sigma\gamma \quad (25)$$

The second sample moment $m_2$is calculated by:

$$m_2 = \frac{1}{n}\sum_{i=0}^{n} {t_i}^2 \quad (26)$$

If n=2 substitute in the formula (23), we get:

$$\mu_2 = 2\mu + 4\sigma\gamma$$

$$2\mu + 4\sigma\gamma = \frac{1}{n}\sum_{i=0}^{n} {t_i}^2$$

$$\mu = \frac{1}{2n}\sum_{i=0}^{n} {t_i}^2 - 2\sigma\gamma \quad (27)$$

By equating the second sample moment with the second population moment, we obtain the parameter of the extreme process. We equal two equations (25) and (27) we will get:

$$\frac{1}{n}\sum_{i=0}^{n} t_i - \sigma\gamma = \frac{1}{2n}\sum_{i=0}^{n} {t_i}^2 - 2\sigma\gamma$$

$$\sigma_{MM} = \frac{1}{\gamma}\left\{\frac{1}{2n}\sum_{i=0}^{n} {t_i}^2 - \frac{1}{n}\sum_{i=0}^{n} t_i\right\} \quad (28)$$

And substitute formula (28) in formula (25), we will get:

$$\mu_{MM} = \frac{1}{n}\sum_{i=0}^{n} t_i - \left\{\frac{1}{2n}\sum_{i=0}^{n} {t_i}^2 - \frac{1}{n}\sum_{i=0}^{n} t_i\right\} \quad (29)$$

Where, $\gamma = 0.577215$.

3. Simulation

Simulation, which is a computer-based methodology that enables experimentation on a valid digital representation, is considered one of the best methods used for generating random variables from specific distribution functions; it is a flexible methodology characterized by showing the ability to conduct experiments and tests via repeating the process many times by easily changing the inputs of the estimation processes. The importance of the simulation lies in using simulated random numbers to estimate some parameters. The random numbers used in the first experiment are independent of those used in the second experiment and so on. The most common simulation method used in statistical analysis is the Monte Carlo simulation method used to estimate parameters or statistical measures, and examine the properties of the estimates. In this section, a comprehensive simulation study was conducted to compare the two estimation methods to reach the best estimate for the EVP parameters. The following four stages represent a description of the design simulation experiments [14].

3.1 Stage of Assigning Virtual Values

This stage is considered as a basis for the other stages, in which the required virtual values are specified for generating random numbers; they are:

1- Sample Size (n)

To obtain efficient and accurate results, the size of the sample plays an important role. For this study, for generating random numbers, different sizes were practiced; a small-sized sample with (N=20), a middle-sized sample with (N=50), and a large-sized sample with (N=100).

2- Parameter Values for EVP

To determine an estimate for EVP parameters, the probability distribution function for Extreme-value function is used to generate the extreme-value random variables using different uniform random numbers assuming values for $\hat{\sigma} = 0.5$, 0.8 ;$\hat{\mu}$ =0.8.

3- Sample Repetition Size (i)

To get high homogenous, the experiments were repeated ($i = 50$) once for each experiment.

3.2 Generating random variables

Random variables are generated from the extreme-value distribution function based on each value of the virtual parametric values and the assumed sample size N as follows:

1- Generate random numbers that follow continuous uniform distribution over [0,1].

2- Change the generated uniform data into data that follow the extreme-value probability distribution by using the cumulative density function and according to the inverse transform method; this method is one of the simplest simulation techniques and the most important methods to get random variables from continuous and discrete distributions. The algorithm of the inverse transform method for generating random variables from the extreme-value distribution.

3- From the cumulative probability function for the extreme-value distribution, which is defined as:

$$F(Y) = 1 - e^{-e^{-\left(\frac{y-\mu}{\sigma}\right)}} \tag{30}$$

Since $U = m(\tau)$

$$u = 1 - e^{-e^{-\left(\frac{\tau-\mu}{\sigma}\right)}}$$

$$e^{-e^{-\left(\frac{\tau-\mu}{\sigma}\right)}} = 1 - u$$

By taking the natural logarithm of both sides, the following is obtained:

$$-e^{-\left(\frac{\tau-\mu}{\sigma}\right)} = \ln(1-u)$$

Again, we will take the natural logarithm of both sides, the following is obtained:

$$\left(\frac{\tau-\mu}{\sigma}\right) = \ln(-\ln(1-u))$$

The random generator for the extreme-value process is represented by the following equation:

$$\frac{1}{\sigma}\tau = \frac{\mu}{\sigma} + \ln(\ln(1-u)) \tag{31}$$

Hence, by implementing the inverse transform Method, various random variables from the Extreme-value function are obtained using the MATHLAB program:

Therefore:

$$\tau(i) = \mu + \sigma \log(\log(1-u(i))) \quad ; \text{for } i = 1,2,\ldots N \tag{32}$$

Where $U(i) \sim U(0,1)$

3.3 Comparison Stage

After finding the model parameters estimator using different methods, a comparison is made between two different methods by using the Root Mean Squares Error (RMSE), according to the following form [15] :

$$RMSE = \sqrt{\frac{\sum_{i=1}^{Q}(\hat{\gamma}_i-\gamma)^2}{Q}} \tag{33}$$

$\hat{\gamma}_i$: Represents the value of the parameter estimated in iteration i.
γ : Represents the real parameter value.
Q : Represents the number of iterations.

3.4 **Numerical Computations**

The following results are shown in Table 1, which represents the results of performing a simulation to generate different random variables from the stochastic extreme-value process using different sample sizes (n=20,50,100), for two values for extreme-value process parameters μ, σ (σ =0.5 ,0.8 and μ=0.8) for the three methods, MLE, MM and the proposed intelligent ABC method.

Table 1: The Simulated RMSE for the ML, MM and ABC Estimator for the parameters μ, σ for Extreme-value process.

μ	σ	n	Methods	$RMSE(\hat{\mu})$	$RMSE(\hat{\sigma})$
0.5	0.8	20	MLE	4.2036	1.9352
			ABC	0.3648*	1.2411*
			MM	5.4824	5.0873
		50	MLE	2.6586	1.2239
			ABC	0.2307*	07749*
			MM	3.4674	3.2175
			MLE	1.8799	0.8654
		100	MM	2.4518	2.2751
			ABC	0.1631*	0.5550*
			MLE	4.0770	1.8817
		20	ABC	0.7467*	1.3864*
			MM	5.7726	5.2153
	0.8		MLE	2.5785	1.1901
0.8		50	ABC	0.4723*	0.8768*
			MM	3.6509	3.2984
			MLE	1.8233	0.8415
		100	MM	2.5816	2.3324
			ABC	0.3340*	0.6200*

From the table above, the numerical results show that the ABC method is better than the MLE, MM methods for estimating the Extreme-value process parameters.

4. Application

To evaluate the applicability of the two methods, real data from Badush Cement Factory is used. The new Badush Cement Factory in Nineveh Governorate is one of the most important factories for the General Cement Company in the north of Iraq; it is the main source for cement production for the governorates of Iraq in general and Nineveh Governorate in particular. The data represent the successive operating periods in days between two successive stops for the cement production during the period from 1/4/2018 to 31/1/2019. To ensure the adequacy fit of data, a test of goodness of fit is needed; it is explained below.

4.1 Test of the Homogeneity of Extreme-value Process

The extreme-value process is A nonhomogeneous process because the time rate of accidents varies with change time (t), That means, it is affected by time t in its behavior. It is noted that the parameter μ is coupled to time t, and so the Extreme-value process is homogeneous in the case $\mu = 0$, and nonhomogeneous in the case $\mu \neq 0$,To conduct the test process, whether the process is homogeneous or nonhomogeneous, the following hypothesis is tested [16]:

$H_0: \mu = 0$

$H_1: \mu \neq 0$

which can be tested through the following statistics:

$$Z = \frac{\sum_{i=1}^{n} \tau_i - \frac{1}{2} n\tau_0}{\sqrt{\frac{n\tau_0{}^2}{12}}} \quad . \tag{34}$$

Where, $\sum_{i=1}^{n} \tau_i$ is the sum of the accident times for a period $(0, \tau_0]$,
n represents the number of accidents that occur in a period $(0, \tau_0]$.

4.2 The consistency test of the data under study

The homogeneity of the data under study was tested using the statistical laboratory in the formula (32. We use a program prepared for this purpose in MATLAB/R2019b. The calculated value has been obtained |Z|=74.4596. It is more than its corresponding tabular value of 1.96 at a morale level of 0.05 Therefore, the null hypothesis is rejected and the alternative hypothesis is accepted. This means the process under study is heterogeneous.

4.3 Estimation of the rate of occurrence of Extreme-value process: for Raw Material Mill Runtimes

To evaluate the performance of the ABC method for estimating Extreme-value Process Parameters, the parameters of the process under study, then we compare it with the traditional MLE and MM the real data, which represent the number of days for operating periods between two successive stops and the times of occurrence, is used. The data is raw materials for a laboratory of the new Badush cement factory in Mosul in Iraq. The programming language MATLAB/R2019b is used to run the algorithm for the estimation process.

Table 2: Estimating of Extreme-value process parameters applied to real data representing Operating periods in days for raw materials mill.

Methods	Parameter estimation $\hat{\mu}$	Parameter estimation $\hat{\sigma}$
MLE	-117.8811	62.4340
ABC	-13.5889	6.4639
MM	-2.8524	1.4539

Table 2 shows the estimation of the extreme-value process parameters for the operating periods between two consecutive stops in days for the mill for raw materials using the proposed estimation methods ABC, MOM and MLE; several runs were conducted in the estimation process and different values for the parameters were used. Then, the best estimated values for the $\hat{\mu}$, $\hat{\sigma}$ were obtained based on the following values: Consecutive operating periods between two successive stops of the raw materials mill in days during the extended period of time from 1/4/2018 to 31/1/2019Which represent 53 runs / day are:
t = [3 8 2 4 1 1 2 3 1 1 1 1 3 2 3 1 1 1 2 3 5 6 5 2 1 1 4 1 4 3 1 3 1 1 7 2 5 1 2 1 1 3 3 1 6 1 2 3 3 1 3 2 1]

4.4 Discussion of Results

To compare the used methods for estimating the parameters of the Extreme-value process, the criterion for the error of the greatest proportion was used MPE by formula (30), By using the program that is prepared for this purpose in the programming language MATLAB\R2019b, the expected number of consecutive operating periods between two successive stops was obtained for the raw materials mill for the new Badush Cement Factory during the period under study. Standard was calculated MPE between the real and the estimated values of the average plant shutdown time as in the following table:

Table 2: MPE values for the method used to estimate Extreme-value process parameters

Methods	MPE
MLE	0.9988
MM	0.9923
ABC	0.9886

It is noted from Table 3 that the value of the MPE method capabilities ABC is less than the value of the maximum likelihood method and MM in appreciation, this indicates the efficiency of the smart method in the estimation of Extreme-value Process Parameters. The following figure represents the Extreme-value process function estimated using conventional and intelligent estimation methods used in the research, compared with the real cumulative values that represent the successive operating periods between two successive stops of the raw materials mill for the new Badush Cement Factory:

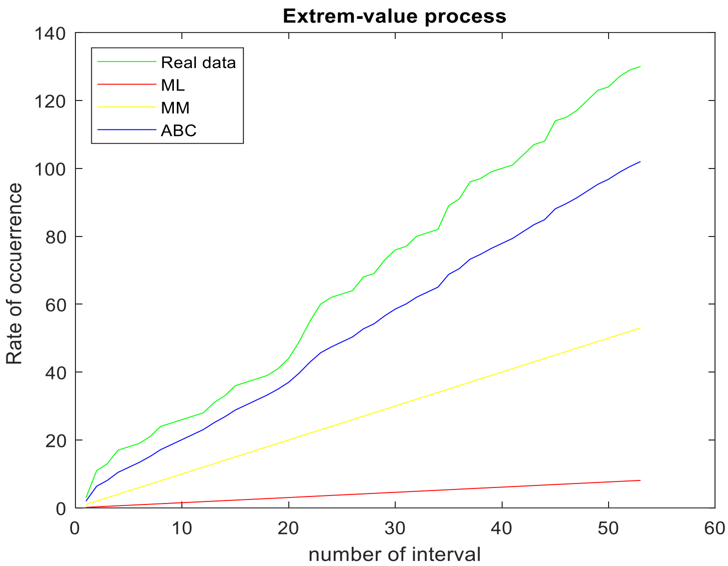

Figure 1: Estimated functions for the cumulative number of successive operating periods between two successive mill stops using different methods.

Figure 1 shows the estimated functions of the cumulative number of successive operating periods between two successive mill stops raw materials for the new Badush cement plant using estimation methods. Using estimation methods for the operating periods between two stops and been noticed the ABC method was the closest to the real data, which indicates the efficiency of this method of estimation compared with the maximum likelihood method and MM for search data.

5. Conclusions

In this paper, ABC was used as a tool for estimating the EVP parameters and compared with MLE and MM. The results show that the ABC is a powerful technique that performs which is better than the MLE and MM for estimating a value for the parameter of the distribution. In addition, this model was applied to real data representing the operating periods of two successive stops for the raw materials factory from the General Company for Northern Cement; it was examined graphically that the extreme value function fits the data. Finally, it is recommended to use ABC approach to estimate parameters for other distribution functions that represent NHPP such as Weibull, exponential or many others.

Acknowledgements

The authors are very grateful to the University of Mosul/ College of Computer Sciences and Mathematics for their provided facilities that help to improve the quality of this work.

References

[1] W. A. Massey, G. A. Parker, and W. Whitt, "Estimating the parameters of a nonhomogeneous Poisson process with linear rate," *Telecommun. Syst.*, vol. 5, no. 2, pp. 361–388, 1996.

[2] D. R. Cox and P. A. W. Lewis, "The statistical analysis of series of events," 1966.

[3] G. Zhu and S. Kwong, "Gbest-guided artificial bee colony algorithm for numerical function optimization," *Appl. Math. Comput.*, vol. 217, no. 7, pp. 3166–3173, 2010.

[4] D. Karaboga and B. Basturk, "On the performance of artificial bee colony (ABC) algorithm," *Appl. Soft Comput.*, vol. 8, no. 1, pp. 687–697, 2008.

[5] W. Gao and S. Liu, "Improved artificial bee colony algorithm for global optimization," *Inf. Process. Lett.*, vol. 111, no. 17, pp. 871–882, 2011.

[6] D. Karaboga, "An idea based on honey bee swarm for numerical optimization," Technical report-tr06, Erciyes university, engineering faculty, computer …, 2005.

[7] B. Akay and D. Karaboga, "A modified artificial bee colony algorithm for real-parameter optimization," *Inf. Sci. (Ny).*, vol. 192, pp. 120–142, 2012.

[8] D. Karaboga and B. Akay, "A comparative study of artificial bee colony algorithm," *Appl. Math. Comput.*, vol. 214, no. 1, pp. 108–132, 2009.

[9] K. Dervis and O. Celal, "Fuzzy clustering with artificial bee colony algorithm," *Sci. Res. Essays*, vol. 5, no. 14, pp. 1899–1902, 2010.

[10] A. Karim, A. Hassan, and M. J. Hadi, "Automatic Query Expansion for Arabic Text Retrieval," *Iraqi J. Sci.*, vol. 58, no. 4, pp. 2447–2457, 2017.

[11] Y. Aït-Sahalia and R. Kimmel, "Maximum likelihood estimation of stochastic volatility models," *J. financ. econ.*, vol. 83, no. 2, pp. 413–452, 2007.

[12] S. A. Jabr and N. S. Karam, "Gompertz Fréchet stress-strength Reliability Estimation," *Iraqi J. Sci.*, vol. 62, no. 12, pp. 4892–4902, 2021, doi: 10.24996/ijs.2021.62.12.27.

[13] L. P. Hansen, "Generalized method of moments estimation," in *Macroeconometrics and time series analysis*, Springer, 2010, pp. 105–118.

[14] P. A. W. Lewis and G. S. Shedler, "Simulation of nonhomogeneous Poisson processes with log linear rate function," *Biometrika*, vol. 63, no. 3, pp. 501–505, 1976.

[15] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)," *Geosci. Model Dev. Discuss.*, vol. 7, no. 1, pp. 1525–1534, 2014.

[16] J. F. Lawless and N. R. Mann, "Tests for homogeneity of extreme value scale parameters," *Commun. Stat. Methods*, vol. 5, no. 5, pp. 389–405, 1976.