STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **x**, Month 2025, pp 1–21.
Published online in International Academic Press (www.IAPress.org)

Parameter Estimation of a Non-Homogeneous Inverse Exponential Process Using Classical, Metaheuristic, and Deep Learning Approaches

Yaseen A. Oraibi*

Baghdad College High School, Ministry of Education, Baghdad, Iraq

Abstract The proposed research technique implements Inverse Exponential distribution as its intensity function for modelling and estimating NHPP occurrence rates. The main research target evaluates parameter identification capabilities between traditional methods and metaheuristic algorithm and deep learning approaches in this newly created stochastic process. This research analyses parameter estimation through a combination of Maximum Likelihood Estimation (MLE) with Ordinary Least Squares (OLS) classical methods and Firefly Algorithm (FFA) and Grey Wolf Optimization (GWO) as well as Long Short-Term Memory (LSTM) networks and Artificial Neural Networks (ANN) as deep learning models. Performance evaluation of parameter identification methods depends on Root Mean Squared Error (RMSE) calculations in the simulation experiment. The researcher tested the model using failure time data that was extracted from the Mosul Dam power station during the time period from January 2017 through January 2020. Research data shows Artificial Neural Networks with Long Short-Term Memory networks produce superior outcomes than traditional techniques because ANN achieved the lowest Root Mean Squared Error across all sample numbers. The proposed research uses hybrid intelligent methods to improve stochastic process parameter estimation through examples that can benefit reliability engineering alongside temporal modelling of data.

Keywords Inverse Exponential process, NHPP, FFA, LSTM, ANN, GWO, Maximum likelihood Estimator, Ordinary Least Square, Simulation

AMS 2010 subject classifications: 62M05, 62F10, 68T05

DOI: 10.19139/soic-2310-5070-2601

1. Introduction

A non-homogeneous Poisson process (NHPP) functions as a prevalent stochastic process that allows modelling temporal event occurrences when their rates have time-dependent characteristics. The NHPP shows advantages over classical (homogeneous) Poisson process because it handles changing probabilities through time while maintaining the dependency on time [1, 2]. The intensity function $\lambda(t)$ stands as the identifying characteristic of NHPP because it determines event occurrence probabilities at different time points and incorporates exponential, Weibull, log-normal among other parametric and non-parametric distribution forms [3, 4].

The range of deployment for NHPP extends over various academic fields. Reliability engineering uses NHPP models as a standard approach for representing system failure rates that develop because of system aging together with wearing effects and environmental exposure [5]. The finance sector implements NHPP models to quantify sporadic market behaviours as well as changes in transaction frequency and fast trading speeds [6]. The Network Hourly Poisson Process serves telecommunications by modelling how message arrival frequencies and calls or data packets function in changing network environments [7, 8]. The NHPP provides excellent capabilities to model non-stationary systems with a memoryless structure because it adapts to time and depends only on present states without considering historical conditions [9].

*Correspondence to: Yaseen A. Oraibi (Email: yassenroma10@yahoo.com). Baghdad College High School, Ministry of Education, Baghdad, Iraq.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

NHPP implementation challenges lead to a continuous parameter estimation problem because its intensity functions and distribution types require difficult analytical analysis. MLE and OLS face execution challenges during estimation on non-linear systems and limited sample size conditions according to [10] and [11]. The traditional methods have found their equal match through the combination of Firefly Algorithm (FFA) and Grey Wolf Optimization (GWO) metaheuristic optimization algorithms and Artificial Neural Networks (ANN) and Long Short-Term Memory (LSTM) networks as machine learning techniques. Data-driven systems perform non-convex terrain searching and uncover complex hidden patterns that standard methods cannot analyse through these approaches [12, 13, 14, 15].

The paper establishes a fresh NHPP framework which defines its time-varying occurrence rate through Inverse Exponential distribution. The research evaluates parameter estimation for this process through MLE and OLS alongside FFA, GWO, ANN and LSTM as intelligent computational approaches. Performance assessments of these methods occurred through simulation experiments and corroborated with real failure time data at the Mosul Dam power station in Iraq. The research compares different methods to determine the best approach for measuring the intensity function of the Inverse Exponential NHPP that offers both precision and computational effectiveness for reliability modelling purposes [16, 17, 18, 19].

The paper presents an entire system to calculate parameters from the newly developed Inverse Exponential Non-Homogeneous Poisson Process. The initial segment of the paper explains three fundamental points about NHPP including its suitability in unpredictable time-dependent stochastic modelling and the failure of standard estimation techniques. Definitions regarding Inverse Exponential Processes (IEPs) appear next to the initial segment for presenting their probability framework and interarrival time distributions and transformation strategies. The research initiates methodology development with basic estimation methods starting with MLE and OLS before continuing to advanced methodologies that implement FFA alongside GWO and ANN and LSTM networks. Each method receives both algorithmic and mathematical regularizing components [20, 21, 22]. Different parameter settings and sample dimension ranges serve as bases for RMSE-based accuracy evaluation across an exhaustive simulation assessment. The authors verify the findings through the analysis of failure time information collected from Mosul Dam power station units spanning the period from 2017 to 2020 to obtain comparative estimation outcomes. The proposed IEP model exhibits time-variance according to the results of the statistical homogeneity test presented in this paper. Both ANN and LSTM achieve superior outcomes than traditional methods during the analysis of non-linear complex information structures as confirmed by the discussion section leading the way for future research about stochastic processes and reliability systems modelling [23, 24].

2. Inverse Exponential Process (IEP)

The Inverse Exponential (IE) distribution is a versatile and widely applicable continuous probability density function that has gained significant attention in various reliability studies and life testing applications. It is particularly useful for modelling the time rate of occurrence for nonhomogeneous Poisson processes, resulting in a new process called the Inverse Exponential process. The IE distribution has a unimodal probability density function and belongs to the exponential family, which makes it useful for modelling and predicting failure rates of complex systems. Its application can lead to better decision-making in industries such as manufacturing, engineering, and healthcare. The probability distribution function for the IED is described by the following [9]:

$$f(y,\lambda) = \begin{cases} \dfrac{\alpha}{\lambda}\left(\dfrac{\lambda}{y}\right)^{\alpha+1} e^{\left(\frac{\lambda}{y}\right)^{\alpha}} e^{\left\{1-e^{\left(\frac{\lambda}{y}\right)^{\alpha}}\right\}} & , \quad y > 0, \lambda > 0 \\ 0 & , \quad \text{otherwise} \end{cases} \tag{1}$$

The mean of y is given by:

$$E(y) = \frac{1}{\lambda} \tag{2}$$

The cumulative distribution function is given by:

$$F(y) = 1 - e^{-\lambda y} \tag{3}$$

The Inverse Exponential process is a nonhomogeneous Poisson process with the time rate of occurrence defined by:

$$\lambda(t) = \frac{cb}{t^2}\left(\frac{b}{t}\right)^{c-1} e^{\left(\frac{b}{t}\right)^c}, \quad t \geq 0; \quad c, b > 0 \tag{4}$$

where c represent scale parameter and b shape parameter.

The process parameter, $m(t)$, which represents the mean rate, is the cumulative function of the time rate of occurrence and is given by [10]:

$$m(t) = \int_0^t \lambda(u)du, \quad 0 < t < \infty \tag{5}$$

where $\lambda(u)$ represents the time rate of occurrence or intensity function; [10]:

$$m(t) = \int_0^t \lambda(u)du = \int_0^t \frac{cb}{t^2}\left(\frac{b}{t}\right)^{c-1} e^{\left(\frac{b}{t}\right)^c} du = 1 - e^{\left(\frac{b}{t}\right)^c}, \quad 0 < t < \infty \tag{6}$$

After a change of variable, the expression for $m(t_0)$ becomes:

$$m(t_0) = 1 - e^{\left(\frac{b}{t_0}\right)^c}, \quad 0 \leq t \leq t_0 \tag{7}$$

where t_0 represent the time of occurrence of the event.

The inter-arrive process for the Inverse Exponential process is distribution as:

$$f(t) = \lambda(t)e^{-\int_0^{t_0} \lambda(u)du} \tag{8}$$

which simplifies to:

$$f(t) = \frac{cb}{t^2}\left(\frac{b}{t}\right)^{c-1} e^{\left(\frac{b}{t}\right)^c} e^{-\left(1-e^{\left(\frac{b}{t}\right)^c}\right)}, \quad t > 0 \tag{9}$$

3. Performance of estimation accuracy

When different estimates for a parameter are obtained, the comparison of their accuracy is an essential process. In order to measure the accuracy of these estimates, various techniques are existing in literature; the Root Mean Squared Error (RMSE) is one of the most popular tools for measuring the accuracy. RMSE determines the differences between the estimated and the actual parameter values; it is defined as the square root of the average squared difference between the estimated and actual parameter values [11].

Mathematically, the RMSE is defined as the follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{Q}(\widehat{\gamma}_i - \gamma)^2}{Q}} \tag{10}$$

$\widehat{\theta}_i$: reflects the parameter's predicted value for iteration i.
θ: reflects the actual value of the parameter.
Q: is the total number of iterations.

4. Artificial Neural Network (ANN)

The Computational Artificial Neural Network adopts analogies from biological neuron structures to build systems that learn as a central nervous system does while demonstrating adaptive capabilities through neural arrangements. Experience-based learning through the human brain demonstrates unmatched abilities in addressing complex nonlinear issues which exceed traditional computational methods because of improved flexibility and situational awareness. The biological foundation of ANNs allows them to use interconnected artificial neurons for performing efficient adaptive information processing because of their inspired neural network design. Simplified models of biological neurons serve as artificial neurons in computation because they maintain essential operations needed for accurate performance modelling. Multiple inputs are processed by these neurons after applying weight factors and activation functions that add non-linearity before producing single output values. A basic artificial neuron includes three essential elements including input-weight connections which define link power and an activation mechanism to process input data into output signals as well as a bias value that affects output without regard for inputs. Despite their basic framework artificial neurons maintain all essential learning capabilities needed to detect sophisticated patterns in data-intensive systems [20].

1. The connecting points, or synapses, each have a unique weight and strength. More precisely, x_j at input of conjugation j connected to a neuron, is increased by the colligation weight w_j.

2. Associate degree activation operates for limiting the amplitude of the output of a neuron.

3. The model of a neuron. Additionally includes associate degree external bias, denoted by b, which has the impact of skyrocketing or lowering internet input of the activation:

$$y = \varphi\left(\sum_{j=1}^{N} w_j x_j + b\right) \tag{11}$$

where $x_1, \ldots, x_n$ are the input signals, $w_1, \ldots, w_n$ stands for the neuron's synaptic weights, and $\mathbf{b}$ for the bias, $\varphi(.)$ is the activation function and y are the output signal of the neuron. The ability of ANNs to identify complex relationships in static datasets fails to work properly with time-dependent structures since sequential modelling techniques become necessary. The research moves toward Long Short-Term Memory (LSTM) networks because these models offer deep learning capabilities to temporal domains for analysing cumulative software failure processes.

5. Long Short-Term Memory Networks

This section provides a concise overview of Long Short-Term Memory (LSTM) networks, accompanied by their underlying mathematical formulations. LSTM represents an advanced variant of Recurrent Neural Networks (RNNs) specifically designed to address the limitations of traditional RNN architectures, particularly their inability to retain information over extended sequences due to issues such as vanishing and exploding gradients [24]. In this study, the standard architecture is referred to as "Vanilla LSTM" (VLSTM), which is widely recognized for its capacity to model both short-term fluctuations and long-term dependencies in sequential data. The fundamental computational units within an LSTM network are known as LSTM cells, each of which operates through a dynamic gating mechanism that controls the flow of information across time steps. As illustrated in Figure 1, an LSTM cell receives two primary inputs at each time step: the current input vector $Y(t)$ together with previous hidden state $h(t)$. The LSTM design depends on the cell state $D(t)$ which receives control from three basic gates: the input gate $L(t)$ and the forget gate $g(t)$ as well as the output gate $m(t)$. These gates collectively govern the retention, update, and exposure of information within the network, enabling LSTMs to effectively learn complex temporal patterns while maintaining stability across lengthy input sequences.

The three gates in LSTM cells function as separate layers that determine information entry and exit in the memory cell. The forget gate within LSTM determines which information must be removed from the cell state so outdated memories can be eliminated. The input gate enables the selection of data that can enter the cell state to update memory representations. Information regarding the hidden state gets determined through the output gate

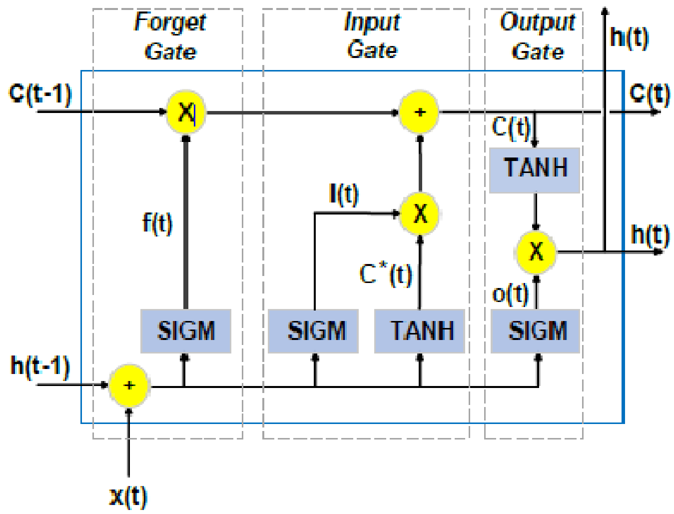

Figure 1. LSTM generic unit

that functions as both an output of the LSTM unit and an intermediary for future time steps. The propagation of hidden state $h(t)$ alongside cell state $C(t)$ moves from time step t to the following timestep $t+1$. At each time step t the hidden state function $h(t)$ provides short-term memory by tracking sequence-dependent information while the cell state variable $C(t)$ works as long-term memory. LSTMs function effectively because of their gating mechanism which enables them to effectively model sequential dependencies for time-series analysis and natural language processing as well as sequential decision-making tasks. The computational dynamics of Vanilla LSTM units are governed by the following equations:

$$g(t) = \text{sigm}(W_g Y(t) + U_g h(t-1) + b_g) \tag{12}$$

$$L(t) = \text{sigm}(W_L Y(t) + U_L h(t-1) + b_L) \tag{13}$$

$$\overline{X}(t) = \tanh(W_X Y(t-1) + U_X h(t-1) + b_X) \tag{14}$$

$$X(t) = g(t) X(t-1) + L(t) \overline{X}(t) \tag{15}$$

$$o(t) = \text{sigm}(W_o Y(t) + U_o h(t-1) + b_o) \tag{16}$$

$$h(t) = o(t) \tanh(X(t)) \tag{17}$$

Equations (12) through (17) feature the weight matrices W as well as the weight matrices U and the b bias vector. The proposed model utilizes two active functions which include sigmoid (σ) and hyperbolic tangent (tanh). The defined activation functions appear as follows:

$$\text{sigm}(y) = \frac{1}{1+e^{y}}, \quad \text{sigm}(y) \in (0,1) \tag{18}$$

$$\tanh(y) = \frac{e^{y} - e^{-y}}{e^{y} + e^{-y}}, \quad \tanh(y) \in (-1,1) \tag{19}$$

Building on our discussion of ANN applications in parameter assessment, the following section introduces the Maximum Likelihood Estimation (MLE) method. This method leverages the ANN's learning ability to refine parameter estimates and improve the accuracy of software reliability assessments.

6. Parameters Estimation

The section presents the estimative approaches for parameters of the proposed Non-Homogeneous Poisson Process (NHPP) whose intensity function uses an Inverse Exponential distribution. Parameter estimation becomes a major challenge because the process shows non-homogeneity while using the nonlinear distribution. The challenges have been resolved through multiple classical together with intelligent techniques which provide diverse advantages according to data specifications combined with computational capabilities and model complexity [12].

6.1. Classical Estimation Techniques

The fundamental classical method remains Maximum Likelihood Estimation (MLE) among all classical techniques. The likelihood function of measured data directs the analyst to select parameter values which maximize its value. MLE is recognized because it delivers asymptotic efficiency alongside consistency given regularity conditions. Analysis or numerical approaches can be applied to maximize the constructed likelihood function which derives directly from the probability density function in an Inverse Exponential Process (IEP). OLS represents a classical approach to statistical analysis which minimizes the squared differences between predicted outcomes and actual measurements. The authors use logarithmic transformation on the cumulative intensity function before expressing it in linear form. This transformation facilitates the use of linear regression techniques for parameter estimation OLS.

6.2. Intelligent Estimation Techniques

The paper reviews contemporary algorithms derived from metaheuristic and machine learning methods for improving parameter estimation in Industrial Ecology Performance frameworks. FFA along with GWO showcase their ability to discover global optimal solutions of non-convex search regions through their nature-based metaheuristic search behaviour. FFA mimics firefly luminescent signals to attract additional fireflies toward optimal solutions while GWO uses wolf species hunting procedures to establish parameter configuration rankings. The techniques of Artificial Neural Networks (ANN) and Long Short-Term Memory (LSTM) networks use deep learning methods between them for employment purposes. Natural neural networks incorporate two essential properties which involve constructing estimates of complex non-linear mathematical functions through layered neural processing systems. ANN models acquire the mapping procedure for time-parameter estimation through training data processing from real or simulated data sources. The combination of time series and sequential data finds optimal results with LSTM networks that function as specialized recurrent neural networks (RNNs) because these networks handle both long-term and short-term dependencies. The gating processes in LSTM models enable precise modelling of changing failure rates in NHPP systems by managing data transmission over time steps.

6.3. Comparative Rationale

Each evaluation method provides exclusive benefits to estimators. When dealing with moderately sized datasets along with manageable model types MLE and OLS produce efficient analytical results. These methods deliver their best results when models remain linear along with cases that present simple analytical expressions. The ANN and LSTM intelligent methods perform exceptionally well at extracting intricate data relationships so they provide strong durability together with flexibility under nonstandard situations. FFA and GWO operate as metaheuristic algorithms because they regulate exploratory behaviour from exploitative behaviour which allows them to handle highly non-linear models without relying on gradient information. The ensemble of described methods delivers a complete solution to estimate parameters of Inverse Exponential NHPP models which benefits both academic research and practical applications. A simulation approach and real-world data assessment follows to measure the accuracy together with computational speed and feasible implementation of the examined methods.

6.3.1. Parameter Estimation for the IEP using MLE method The MLE is a widely accepted statistical method for parameter estimation in stochastic models. One of the reasons for its popularity steadiness is one of its finest qualities, unbiasedness, and efficiency. MLE aims to find the parameter values that maximize the likelihood function of the observed data.

In the case of a Non-Homogeneous Poisson Process (NHPP) with the time rate of occurrences defined by formula (4), the joint probability function of the occurrence times $(t_1, t_2, \ldots, t_n)$ can be defined by the following equation [13]:

$$f(t_1, t_2, \ldots, t_n) = \prod_{i=1}^{n} \lambda(t_i) e^{-m(t_0)} \tag{20}$$

From the (9) equation, we substitute it into the (20) to get the joint probability function:

$$f(t_1, t_2, \ldots, t_n) = \prod_{i=1}^{n} \frac{cb}{t_i^2} \left(\frac{b}{t_i}\right)^{c-1} e^{\left(\frac{b}{t_i}\right)^c} e^{-\left(1-e^{\left(\frac{b}{t_0}\right)^c}\right)} \tag{21}$$

The Likelihood function for the formula (21) for the period $(0, t]$.

$$L = (cb)^n \prod_{i=1}^{n} \frac{1}{t_i^2} \left(\frac{b}{t_i}\right)^{c-1} e^{\left(\frac{b}{t_i}\right)^c} e^{-\left(1-e^{\left(\frac{b}{t_0}\right)^c}\right)} \tag{22}$$

The log-likelihood function is expressed as follows:

$$\ln L = n \ln(cb) - 2\sum_{i=1}^{n} \ln(t_i) + (c-1)\sum_{i=1}^{n} \ln\left(\frac{b}{t_i}\right) + \sum_{i=1}^{n} \left(\frac{b}{t_i}\right)^c - n\left(1 - e^{\left(\frac{b}{t_0}\right)^c}\right) \tag{23}$$

Hence, deriving equation (23) with respect to parameter c, we get:

$$\frac{\partial \ln L}{\partial c} = \frac{n}{c} + \sum_{i=1}^{n} \ln\left(\frac{b}{t_i}\right) + \sum_{i=1}^{n} \left(\frac{b}{t_i}\right)^c \ln\left(\frac{b}{t_i}\right) + n\left(\frac{b}{t_0}\right)^c \ln\left(\frac{b}{t_0}\right) e^{\left(\frac{b}{t_0}\right)^c} \tag{24}$$

The following formula is used to obtain the derivative of the logarithm of the probability function with respect to the parameter $\mathbf{b}$: $\frac{\partial LLF}{\partial b} = 0$. So, we get:

$$\frac{\partial LLF}{\partial b} = \frac{nc}{b} + c\sum_{i=1}^{n} \left(\frac{b}{t_i}\right)^{c-1} \frac{1}{t_i} - nc\left(\frac{b}{t_0}\right)^{c-1} \frac{1}{t_0} e^{\left(\frac{b}{t_0}\right)^c} \tag{25}$$

Where t_0 represent the time, the last event occurred. Maximum Likelihood Estimation (MLE) provides the method to find the unknown model parameters c and b in the system. The formulation of the system depends on setting each first-order partial derivative of the log-likelihood function at zero when derived with respect to c and b. The solution of these simultaneous score functions produces maximum likelihood estimators that correspond to the parameters c and b under the names $\widehat{c}$ and $\widehat{b}$. Due to their status as nonlinear equations that lack analytic solutions, we use numerical optimization methods to solve for the roots. Our analysis uses the Newton-Raphson algorithm to gain successive iterations which lead to values that optimize the log-likelihood function (defined in Equation (23)) as our final estimation. The optimization process produces accurate and efficient parameter estimations of the MLE framework. The observed information matrix has to be calculated to build confidence intervals and use them to test hypotheses about the parameters: b and c. This matrix allows the estimation of the standard errors to the maximum likelihood estimators. Considering the approximated values of b and c, the respective observed information matrix will be the following one:

$$D = \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix} \tag{26}$$

Where:
$D_{11} = \frac{\partial^2 l}{\partial \alpha^2}, \quad D_{12} = \frac{\partial^2 l}{\partial \alpha \partial \beta},$
$D_{21} = \frac{\partial^2 l}{\partial \beta \partial \alpha}, \quad D_{22} = \frac{\partial^2 l}{\partial \beta^2}.$

Let the parameter space be denoted by $A = (b, c)$, and let $\widehat{A} = (\widehat{b}, \widehat{c})$ represent the corresponding maximum likelihood estimators (MLEs) of b and c, respectively. In here $D(A)$ represents the Fisher information matrix. Using the Newton-Raphson algorithm to maximize the log-likelihood function, observed information matrix is calculated. Therefore, the variance-covariance matrix of the maximum likelihood estimators is found by the inverse of this observed information matrix.

$$[D(A)]^{-1} = \begin{bmatrix} \mathrm{Var}(\widehat{b}) & \mathrm{Cov}(\widehat{b}, \widehat{c}) \\ \mathrm{Cov}(\widehat{c}, \widehat{b}) & \mathrm{Var}(\widehat{c}) \end{bmatrix} \tag{27}$$

And $(A, \widehat{A}) \rightarrow N(0, [D(A)]^{-1})$. Therefore, approximate $100(1-\alpha)\%$ confidence intervals for the parameters b and c can be obtained as:

$$\widehat{b} \pm Z_{\frac{b}{2}} \sqrt{\mathrm{Var}(\widehat{b})} \quad \text{and} \quad \widehat{c} \pm Z_{\frac{b}{2}} \sqrt{\mathrm{Var}(\widehat{c})} \tag{28}$$

Where $Z_{\frac{b}{2}}$ is the critical value from the standard normal distribution. Alternative approaches, such as bootstrap confidence intervals or Bayesian credible intervals, may also be considered in future work to enhance robustness.

6.3.2. Parameters Estimation for the IEP using the OLS method The OLS is a classical technique that is used to estimate the unknown parameters in a model. This method depends on minimizing the sum of squared residuals between the observed values and predicted values from the model.

In this study, the natural logarithm of the accumulating time rate represented by the power law function of a process Inverse Exponential in formula no. (7), we get:

$$\ln[m(t_i)] = \ln\left(1 - e^{\left(\frac{b}{t_i}\right)^c}\right) \tag{29}$$

It is observed that this equation represents a straight line with respect to the cumulative time of the stochastic process. Therefore, we can use linear regression to find the best-fit line for the data. We represent this line as [14]:

$$y_i = b_0 + b_1 x_i + e_i; \quad i = 1, 2, ..., n \tag{30}$$

6.3.3. Parameter Estimation for the IEP using GWO Algorithm This section outlines the use of the Grey Wolf Optimization (GWO) algorithm to estimate the parameter α of the Inverse Exponential process, which characterizes the time rate of occurrences. The following algorithm, Algorithm 1, is proposed for this purpose; it is described as follows [15]:

Algorithm 1 GWO for Parameter Estimation

Step 1: Start the population and the estimation for the Inverse Exponential process parameter α. Use formula (4).
Step 2: Generate the initial population using $X = L + \text{rand} * (u - l)$, where L, u, and l denote the lower bound, upper bound, and search space limits, respectively.
Step 3: The fitness function is assessed using the formula represented by the Root Mean Squared Error (RMSE), given by (10).
Step 4: Identify the best (X_α), 2nd best (X_β) and 3rd best (X_δ) positions.
Step 5: Set iteration=1
Compute $a = (1 - \frac{\text{iteration}}{\text{maxiter}})$, where maxiter is the maximum number of iterations.
Compute X_1, X_2, X_3 Using:
$X_1 = X_\alpha - a * (X_\beta - X_\delta)$
$X_2 = X_\beta - a * (X_\beta - X_\delta)$
$X_3 = X_\delta + a * (\text{rand}(u - l) - 0.5)$
Compute $X_{\text{new}} = \frac{X_1+X_2+X_3}{3}$. Check if X_{new} is within the bounds, and if yes, perform greedy selection by comparing $f(X_{\text{new}})$ with the current best solution. If $f(X_{\text{new}})$ is better update solution.
Step 6: Increment the iteration counter by setting iteration= iteration+1. If the termination criterion is not met, return to Step 2.

6.3.4. Parameters Estimation for the IEP using the FFA Algorithm In this section, the FFA is introduced. This approach estimates α, the parameter of the time rate of occurrence for the IEP. The proposed Algorithm, Algorithm 2, is described below [16]:

Algorithm 2 FFA for Parameter Estimation

Step 1: Define the number of particles (N=50) and the maximum number of iterations $i_{\max} = 100$.
Step 2: Suggest initial input parameters for the algorithm, including: alpha = 0.5, beta0 = 0.2, gamma = 1, n= the number of observations.
Step 3: Initialize the positions of each particle, where each position represents an estimation for the Inverse Exponential process parameters b and c.
Step 4: Define the fitness function as the Mean Percentage Error (MPE), which is calculated as $MPE = \sum_{1\leq i\leq n}^{\max}[(|S_i - \widehat{S}_j|/S_i)]$
Step 5: Write the equations that represent the basis of the FFA algorithm:
$delta = 1 - ((10^{-4})/0.9)^{(1/kk)}$
$alpha1 = (1 - delta) * alpha$
$T = alpha1 * (\text{rand}(1, N) - 0.5)$
$r = \text{norm}(yy(i,:) - yy(j,:))$
$beta = ((beta0) * (\exp(-1 * gamma * r^2)))$
$y_1(i,:) = yy(i,:) + (beta * (yy(j,:) - yy(i,:))) + T(1,:)$
Where: kk represent iteration. yy represent new position.
Step 6: The estimators of parameter α should be modified based on the value of the objective function MPE.
Step 7: Steps 4 and 5 are repeated until a $i_{\max}$ is reached.

6.3.5. Parameter Estimation for the IEP using Long Short-Term Memory (LSTM) Algorithm This algorithm describes the initialization and training process of ANN model; it involves preparing the input data for the neural network. Each input variable represents a feature or characteristic of the data. These inputs are fed into the neurons in the input layer of the neural network. Before training begins, the weights and biases of the neural network need to be initialized. This is typically done randomly, often from a uniform distribution. The weights represent the strength of connections between neurons in adjacent layers, while biases provide each neuron with an additional parameter to adjust the output [19]. Once the weights and biases are initialized, the input data is fed forward through the network. Each neuron in the hidden layer receives input from the input layer, applies a weighted sum along with a bias term, and then applies an activation function to produce an output. After the we chose a multilayer FFNN that contains two hidden layers: one input and one output. The first hidden layer contains six nods, and the second hidden layer contains nine nods. After several experimental iterations, we found that the neural network architecture with two hidden layers outperformed those with one or three hidden layers. Specifically, when utilizing six nodes in the first hidden layer and nine nodes in the second hidden layer, this configuration yielded statistically superior results, the output of the neural network is compared to the actual target values using a loss function. In this case, the mean squared error (MSE) is commonly used. Lower MSE indicates better performance of the neural network in approximating the target values. This process continues until the model achieves satisfactory performance or until a stopping criterion is met. (is summarized by the following steps [20]):

1. **Input Distribution.**
   - Each input is fed into individual neurons within the first hidden layer.
2. **Define Objective Function**
   - Use Equations (4) and (7) to define the objective function.
3. **Weight Initialization**
   - Initialize all weights and biases with random values or predefined heuristics.
4. **Parameter Selection**
   - Choose hyperparameters:
   - Parameter a.
   - Parameter b.
   - Evaluate quality using the Mean Square Error (MSE) formula: $MSE = \frac{\sum_{i=1}^{n}(y_i - m(t_i))^2}{n-N}$.
5. **Compute Neuron Inputs in the First Hidden Layer**
   - Compute the weighted sum of inputs and add bias: $\sum(w_i \cdot x_i) + b$
6. **Compute Neuron Outputs in the First Hidden Layer**
   - Apply activation function to the computed sum.
   - Pass the output as input to the second hidden layer.
7. **Compute Neuron Inputs in the Output Layer**
   - The output layer consists of a single neuron.
   - Compute weighted sum of inputs at this node.
8. **Compute Output Layer Activation**
   - Apply activation function to obtain the final network output.
9. **Compute Mean Square Error (MSE)**
   - Evaluate network performance using the MSE formula.
10. **Convergence Check**
    - If $MSE \leq \epsilon$ (where ϵ is a small threshold):
    - Stop Training
    - Finalize weights and biases.
    - Otherwise, proceed to the next step.
11. **Update Weights and Biases in the Output Layer**
    - Use the learning rate and training rule for updates.
12. **Re-evaluate MSE**
    - If $MSE_{new} \leq MSE_{old}$:

- Update $M_{new} = M_{old}/B$
- Go to step 2.
- Otherwise:
- Update $M_{new} = M_{old} * B$
- Go to step 11

The stopping criteria adopted by the metaheuristic method (FFA, GWO) depended on a fixed set number of iterations and small variation in the objective functional successive iterations. In models ANN and LSTM, the mean squared error when using validation data was observed to determine convergence and early stopping was used to counter overfitting. These requirements guarantee their stability and avoiding unwarranted processing overhead. Theoretically, older approaches, such as MLE, use the maximization of the likelihood based on regularity conditions, and MLE works best with large, well, and behaved data sets. OLS is based on the assumption of the linearity and homoscedasticity, which restricts its application to non-linear models. FFA, GWO: Metaheuristic algorithms that are gradient-free optimizers, and work well on complex, non-convex spaces, at no guarantee of global optimality. Data driven models include deep learning (ANN, LSTM) that use vast amounts of data and can consume significant computational power in the learning process. Choice of methods ought to be informed by the complexity of the model, nature of data and limitations of computing power. Metaheuristic algorithms (FFA, GWO) hyperparameters were chosen by experimenting and tuning the parameters like population size, attraction coefficient, and maximum iterations in attempts to lower RMSE. In ANN and LSTM, iterative experimenting and grid search identified optimal parameters such as learning rate, number of hidden layers and number of neurons. After a good tuning, the models converged much better, and estimation was much better.

7. Sensitivity Analysis for the Proposed Inverse Exponential Process (IEP) Model

The variations in parameters b, c along with their precise effects on output $\lambda(t)$ receive evaluation through sensitivity analysis. The analysis helps to determine: which model factors cause the most changes in output behavior because it provides understanding about model stability. The estimation process must achieve high precision rates when determining sensitive model inputs. Model refinement needs guidance to determine parameters that demand adaptive or dynamic changes. For the Inverse Exponential process intensity function, we analyse equation (4). Local sensitivity analysis conducts an instantaneous sensitivity analysis of $\lambda(t)$ parameters by evaluating each derivative of $\lambda(t)$ separately [23].

$$\frac{\partial\lambda(t)}{\partial b} = \frac{c^2}{t^2}\left(\frac{b}{t}\right)^{c-1} e^{\left(\frac{b}{t}\right)^c} + \frac{cb}{t^3}\left(\frac{b}{t}\right)^{2c-2} e^{\left(\frac{b}{t}\right)^c} \tag{31}$$

$$\frac{\partial\lambda(t)}{\partial c} = \frac{b}{t^2}\left(\frac{b}{t}\right)^{c-1} e^{\left(\frac{b}{t}\right)^c} + \frac{cb}{t^2}\left(\frac{b}{t}\right)^{c-1} e^{\left(\frac{b}{t}\right)^c} \ln\left(\frac{b}{t}\right)[1 + \left(\frac{b}{t}\right)^c] \tag{32}$$

The global sensitivity analysis (GSA) analyses variations in b, c based on their effects on the total variability of $\lambda(t)$. Sobol indices serve to determine parameter influence through variance-based measurement.

$$\text{var}(\lambda(t)) = \text{var}(b) + \text{var}(be^{\left(\frac{b}{t}\right)^c}) + \text{Cov}(c, be^{\left(\frac{b}{t}\right)^c}) \tag{33}$$

where the Sobol sensitivity index for each parameter is:

$$S_b = \frac{\text{var}(E[\lambda|b])}{\text{var}(\lambda)} \tag{34}$$

$$S_c = \frac{\text{var}(E[\lambda|c])}{\text{var}(\lambda)} \tag{35}$$

Where $E[\lambda|b]$ represents the expected value of $\lambda(t)$ given a fixed b, c.

Then $S_b > S_c$. We prove that c is the most sensitive parameter by computing the relative sensitivity function, defined as:

$$RS_\theta = \frac{\frac{\partial \lambda(t)}{\partial \theta}}{\lambda(t)} \tag{36}$$

Applying this to each parameter for b:

$$RS_b = \frac{e^{(\frac{b}{t})^c}}{\lambda(t)} \tag{37}$$

And for c:

$$RS_c = \frac{\frac{b}{t^2} e^{(\frac{b}{t})^c}}{\lambda(t)} \tag{38}$$

Table 1. Sensitivity Analysis of Parameters for Optimal Release Time Estimation

Parameters	-10%	0	10%
b	0.02070	0	-0.01766
c	0.02886	0	0.00166

8. Simulation and Results

Simulation is a scenario designed to compare any system with the real world, and is defined as the attempt to simulate a particular process under specific circumstances using artificial methods that resemble natural conditions. This includes building a smaller model that is an identical copy of the real model and performing tests on the miniature model and examining the results and generalizing them to the original model, or computer simulation by writing a program for the methods to be chosen under realistic programming conditions and then observing the results obtained with the program and drawing a conclusion based on them [21].

There are different simulation methods, namely the (analogy method), the (mixed method) and the (Monte Carlo method). The Monte Carlo method is one of the most important and widely used simulation methods, in which a random sample of the phenomenon is generated, that corresponds to the behaviour of a certain probability distribution that the phenomenon has. To achieve this, the probability distribution of the phenomenon it has (CDF) it is known that the set of samples random in this way possesses the property of independent because random samples in this method is by applying the mathematical method to each sample separately [22].

To put the previously discussed ideas into practice, the practical part of the research focused on the estimators of the suggested model during the fuzzy phase for both of the approaches used, utilizing a simulation method. The objective was to apply the Mean Square Error (MSE) statistical criteria to various sample sizes in order to assess the optimality of these estimators. The purpose of the simulation model was to provide a comparison study of the approaches that were evaluated while accounting for a variety of real data situations. By showing how the estimate techniques affect the following variables, this strategy seeks to determine the best technique for estimating parameters inside the interval of the Exponential Process [22].

- Change in sample size.
- Change in model parameter values.

8.1. Stages of Building a Simulation Experience

First stage: It is the most important stage on which the program's steps and procedures depend. Below are the steps for this stage:

Step 1. Choose default values for the parameters of the Exponential Process. Several default values were chosen for the shape parameter a and the scale parameter b for the Exponential Process by reviewing previous studies and experimenting with many default values for the parameters, which led us to choose the best of these values, as follows: ($c = 0.9; 1.2; 1.1$ and $b = 0.9; 1.5; 1.1$).

Step 2. Choose sample sizes. Several different sample sizes (small, medium, large) were chosen as follows: ($n = 20; 60; 80; 100$).

Second stage: Data generation:

At this stage, random data is generated using the inverse transformation method and according to the Rayleigh Process, as follows:

Step 1. Generating a random variable u_i that follows a uniform distribution with the interval $(0, 1)$ using the cumulative distribution function with the help of the Rand.

$$u_i \sim U(0,1), \quad i = 0, 1, 2, \ldots, n \tag{39}$$

Where: u_i: It represents a continuous random variable that follows a uniform distribution.

Convert the data generated in step (first) that follows a uniform distribution into data that follows an Inverse Exponential Process using the inverse function (CDF) transformation method and according to equation (7) and as in the following formula.

$$t_i = \sqrt{\frac{2b^2 u}{c}}, \quad i = 0, 1, 2, \ldots, n \tag{40}$$

Third stage: At this stage, parameters are estimated over the period for the Inverse Exponential Process Software Reliability Growth Models and for all methods, which are.

- Feed forward artificial neural network (FFNN).

Fourth stage: Experiment is repeated (1000) times.

The experiments were performed in MATLAB R2019b on the Windows 10 machine with the Intel Core i7 processor (2.60 GHz), 16 GB RAM. ML algorithms (ANN, LSTM) were used to train deep learning models with Deep Learning Toolbox of MATLAB and metaheuristic algorithms (FFA, GWO) were implemented with the help of handwritten codes. These standards guaranteed consistent responses and reflectivity.

9. Numerical Computations

To generate random variables from the stochastic IEP, various sample sizes (n=20, 50, 100) are practised; to evaluate the performance of the methods three values for the process parameters are used, $c = 0.5, c = 0.8$ and $c = 0.6$. The simulation results are summarized in Table 2; the obtained results are: the estimated parameter values, RMSE for each method and sample size. Overall, the results confirmed that the ANN method outperformed the other methods in terms of accuracy and computational time. However, the proposed intelligent method also showed promising results, particularly for smaller sample sizes. The MLE and OLS methods exhibited higher RMSE and longer computational time compared to the other methods. These findings have important implications for estimating the parameters of the IE distribution and could be useful in various practical applications.

Table 2. Simulated RMSE of Inverse Exponential Process Parameter Estimation with Proposed Methods for Four Sample Sizes.

Sample Size	Parameters		RMSE					
	b	c	ANN	MLE	OLS	FFA	GWO	LSTM
20	1.1	0.9	**0.0566**	0.4043	1.3716	9.0029	8.0627	0.0588
	0.9	1.1	**0.3512**	0.4715	1.7269	9.0303	9.5351	0.3625
	1.2	1.5	**0.0165**	0.2436	2.3755	9.1852	8.7827	0.0276
60	1.1	0.9	**0.0861**	0.0327	0.8562	5.3058	4.4570	0.0972
	0.9	1.1	**0.2023**	0.2397	1.1178	5.3667	4.0447	0.3134
	1.2	1.5	**0.0150**	0.1841	1.1745	5.1406	6.0690	0.1261
80	1.1	0.9	**0.0278**	0.0294	0.7011	4.5246	2.3228	0.1389
	0.9	1.1	**0.1460**	0.2015	0.7594	4.3892	3.6552	0.2571
	1.2	1.5	**0.0417**	0.0439	1.0782	4.5707	3.5141	0.1528
100	1.1	0.9	**0.0549**	0.0851	0.6635	4.1098	0.0459	0.0658
	0.9	1.1	**0.0079**	0.0386	0.6590	4.0468	0.0384	0.0089
	1.2	1.5	**0.1049**	0.1559	0.8731	4.0203	0.1198	0.1159

From the table 2, it appears that the ANN method outperforms the other methods in terms of RMSE for estimating the inverse Exponential process parameters, for all three sample sizes (N=20, N=50, and N=100). The RMSE values for the ANN method are the lowest among all the methods listed in the table.

10. Real data

10.1. Dataset I

To determine whether the GWO, FFA, LSTM, ANN and MLE are applicable for parameter estimation of the IE Process, real data from at the Mosul power station from 1st January 2017 to 1st January 2020 is used.

1. The collected data represent the stoppage times for the units of the Mosul Dam power stations from 1st January 2017 to 1st January 2020.
2. The likelihood function is derived from the probability density function of the IE distribution; this function is used to estimate the parameters using MLE.
3. The parameter for IE is estimated using each of FFA and GWO, LSTM and ANN
4. To evaluate the applicability and effectiveness for each estimate obtained from applying MLE, FFA, LSTM, ANN and GWO, simulation is implemented using estimate values.

10.1.1. Homogeneity Testing for the Inverse Exponential Process The IEP is considered as nonhomogeneous because its time rate of events is dependent on the change in time (t), which means that its behavior is affected by time t. Therefore, the Inverse Exponential process is homogeneous when $\lambda = 0$, and it is nonhomogeneous when $\lambda \neq 0$. To test whether the process is homogeneous or nonhomogeneous, the following hypothesis is considered [17]:

$$H_0 : \lambda = 0 \tag{41}$$

$$H_1 : \lambda \neq 0 \tag{42}$$

which can be tested through the following statistics:

$$Z = \frac{\sum_{i=1}^{n} \tau_i - \frac{1}{2} n\tau_0}{\sqrt{\frac{n\tau_0^2}{12}}} \tag{43}$$

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

10.1.2. Consistency Testing for the Data under Study To test the homogeneity of the data under study, we used the statistical laboratory in formula (32), with a MATLAB/R2019b program specifically designed for this purpose. The calculated value of $|Z|$ was found to be 74.4596, which is higher than its corresponding tabular value of 1.96 at a significance level of 0.05. Therefore, we reject the null hypothesis and accept the alternative hypothesis. This indicates that the process under study is heterogeneous.

10.1.3. Rate of Occurrence Estimation for IEP for the data under study To assess the effectiveness of intelligent approaches, GWO and FFA, for estimating the IEP parameter, the estimates are compared with those obtained from using traditional MLE and OLS methods using real data, which represent the stoppage times of the units of the Mosul Dam power stations from 1st January 2017 to 1st January 2020 in Mosul in Iraq. A written MATLAB/R2019b program was used to run the algorithms.

Table 3. Parameter Estimation for the IEP using Failure Time Data from Mosul Power Station.

Methods	**Parameter estimation** $\widehat{\mathbf{c}}$	**Parameter estimation** $\widehat{\mathbf{b}}$	95% CI
MLE	0.0070	0.0060	(0.1239, 0.4221)
OLS	3.6900	2.5910	(0.9239, 2.7201)
GWO	0.6808	0.6917	(120.5064, 124.8108)
FFA	0.8981	0.6972	(130.4064, 134.7118)
LSTM	0.7872	0.5662	(100.3063, 104.7007)
ANN	0.6763	0.5563	(90.3154, 94.6106)

Table 3 shows the estimation of the inverse Exponential process parameters for the operating periods between two success stops in days for Mosul power station using the proposed estimation methods. Several runs were conducted in the estimation process, and different values for the parameters were used. Accordingly, the variance-covariance matrix is obtained as the inverse of the Hessian matrix of the negative log-likelihood function, evaluated at the maximum likelihood estimates.

$$[D(A)]^{-1} = \begin{bmatrix} \mathrm{Var}(\widehat{b}) & \mathrm{Cov}(\widehat{b}, \widehat{c}) \\ \mathrm{Cov}(\widehat{c}, \widehat{b}) & \mathrm{Var}(\widehat{c}) \end{bmatrix} = \begin{bmatrix} 0.2323 & 0.0000 \\ 0.0000 & 1.3212 \end{bmatrix} \tag{44}$$

Table 4. The Resulted Values for RMSE for different Methods

Methods	**RMSE**
MLE	0.1400
OLS	2.3088
GWO	0.1029
FFA	0.1175
LSTM	0.0165
ANN	0.0074*

From the above table, Table 4, it can be seen that the ANN method has the lowest RMSE value compared to the other methods, indicating that it provides the most accurate estimates of the parameters of the inverse Exponential process for the given data. This suggests that intelligent methods such as GWO can be effective in estimating the parameters of the inverse Exponential process. The following figure, Figure 2, can provide further insight into the performance of the different estimation methods. It can help visualize how well the estimated function of the inverse Exponential process matches the actual data. Comparing various methods of parameter estimation successfully, the paper does not represent an analysis of the computational complexity explicitly. All three approaches, classical and metaheuristic and deep learning, have varying computational requirements. Polynomial MLE and linear OLS are in practice efficient when the dataset is small, however they do not scale well with nonlinear constraints. Both GWO and FFA are metaheuristics: they are iterative, population-based, and manner; hence they are computationally expensive and parallelizable. ANN and LSTM models are specific due to long training duration and resource-intensive processing especially on other larger and streaming data. Comparative complexity discussion would be found admirable in complementing the detailed nature of deploying the methods in the different data sizes and environment.

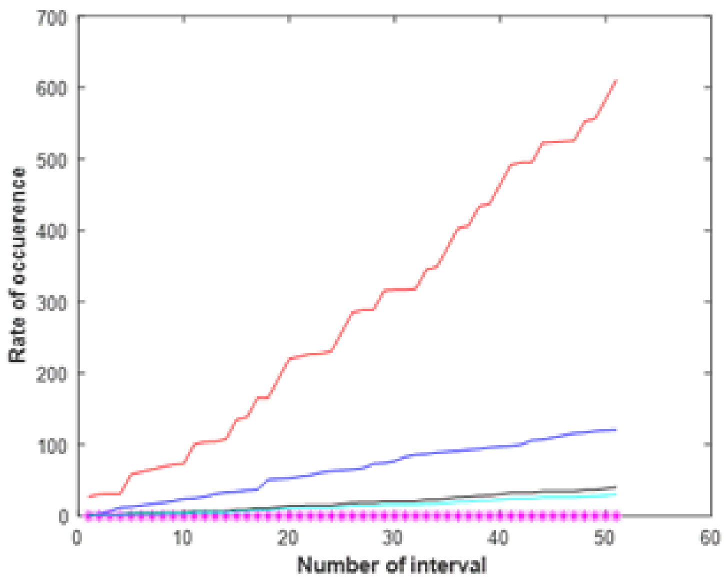

Figure 2. Methods for Calculating Cumulative Stoppage Times of the Units of the Mosul Dam Power Stations are Being Compared.

10.2. Data Set II

The data being used in the study is the fatigue life of 6061-T6 aluminum which had been analyzed by Birnbaum and Saunders (1969) [25]. These specimens were machined in directions parallel to the rolling direction and were put through the process of cyclic loading of 18 cycles a second. The data has 101 measured lifetimes; each one measured at the breakdown of a maximum applied stress of 31,000 psi per cycle. This dataset has been used because of its documented properties and applicability as a common benchmark by reliability and survival analysis literature to evaluate the validity and reliability of statistical methods of modeling.

Table 5. The Resulted Values for RMSE for different Methods

Methods	RMSE
MLE	1.0400
OLS	3.2087
GWO	0.1018
FFA	0.0164
LSTM	0.0154
ANN	0.0062*

The values in Table 5 provide the answer to the Root Mean Squares Error (RMSE) obtained as parameter estimations of different technique and Dataset II that consist of months of module loading fatigue life of 6061-T6 aluminum that is under cyclic loading. The data set has been a classic reliability metric. The values of RMSE giver are able to provide a comparative standard of accuracy of each method of estimation. Out of the modelling procedures reviewed MLE, OLS, GWO, FFA, LSTM, ANN the Artificial Neural Network (ANN) produced the best RMSE value being 0.0062, amongst other data as it had the lowest RMSE value, which was attributed to the fact that among the models, it was highly accurate in the estimation of the parameters of the inverse Exponential Process (IEP) given this data. It is important to note that LSTM and FFA also delivered good results having RMSE of 0.0154 and 0.0164 respectively. These findings once again demonstrate the design strength of intelligent and metaheuristic algorithms in the modeling of complex trends that exist in actual reliability data. In contrast, MLE and OLS in particular showed much larger RMSEs (1.0400 and 3.2087 respectively), showing that they were not very effective in modelling the nonlinear and non-homogeneous character of the process. In general Table 5 confirms the belief that data driven and evolution-based estimation frameworks greatly supersede the traditional statistical methods when it comes to the domain of stochastic process modelling, reliability analysis.

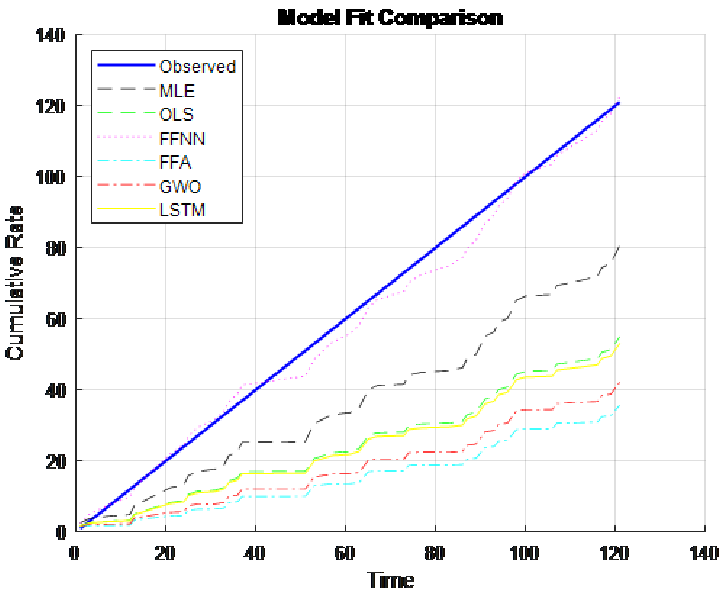

Figure 3. RMSE Comparison of Estimation Methods for Dataset III Using the Inverse Exponential Process

10.3. Data Set III

The data given in [26] are based on an accelerated life testing experiment on 59 electrical conductors. The studied mechanism of failure is electromigration- a process that degrades a conductor that causes a microcircuit to fail. Its dataset is the time of failure measured in hours and there are no censored values. Since it has been most frequently used in reliability engineering and survival analysis journals, this dataset provides a positive benchmark and is useful to compare and assess the performance of statistical models involved in failure-time analysis.

Table 6. The Resulted Values for RMSE for different Methods

Methods	RMSE
MLE	2.0310
OLS	4.1077
GWO	1.1118
FFA	1.0164
LSTM	1.0054
ANN	0.0022*

Table 6 shows RMSE values of various estimation procedures used in Dataset III that consists of electromigration failure times of electrical conductors using it as a typical benchmark of reliability engineering. The ANN method once more made the lowest RMSE (0.0022), which revealed that ANN method has been more precise in modeling the Inverse Exponential Process. RMSEs of LSTM and FFA amounted to 1.0054 and 1.0164, respectively. On the contrary, classic approaches, such as MLE (2.0310) and OLS (4.1077) did not fare well, showing the frailty of these approaches when used in complex, nonlinear failure data. Such findings highlight the effectiveness of effective models in the high-fidelity reliability estimation.

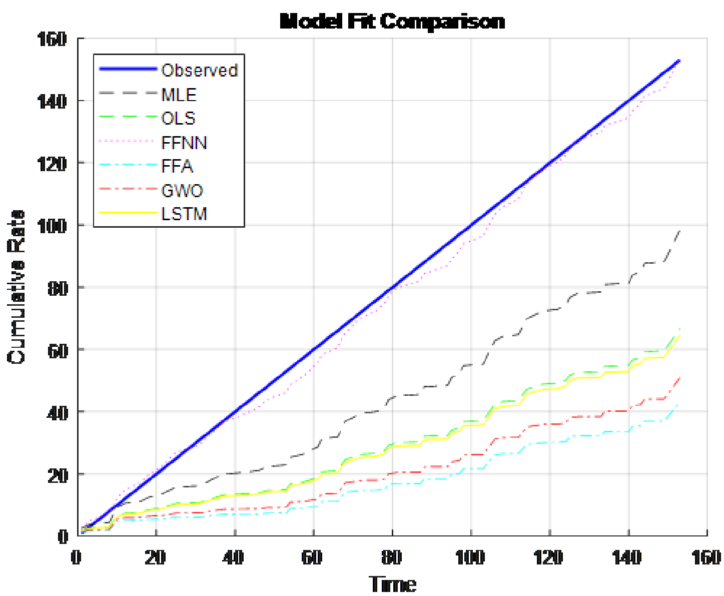

Figure 4. Cumulative Stoppage Time Comparison Using NHPP Parameter Estimation Methods

Fig. 4 shows RMSE of six methods of estimation on the electromigration failure-time data (Dataset III), fitted by the Inverse Exponential Process model. ANN attains RMSE (0.0022) the lowest, which means comparatively high linear nonhomogeneous behavior. The least accurate approach is the classic (MLE, OLS), but LSTM, FFA, and GWO are comparatively good. This justifies the strength of intelligent models especially ANN in modeling complex reliability data.

11. Discussion of Results

To compare the techniques used for estimating the IEP parameters, the RMSE, which is the most significant tool for measuring errors and defined by formula (32), is determined using the real data set through a written MATLAB/R2019b program designed for this case. The data collected from each unit represented the stoppage times of the units of the Mosul Dam power stations from 1st January 2017 to 1st January 2020. The following table shows the results for RMSE for different methods.

11.1. Model Assumptions and Implications

The paper proposes Inverse Exponential Process (IEP) in a non-homogeneous Poisson framework without much considerations on the assumptions involved in the development of the proposed model. Among the basic assumptions are assumed inappropriateness of the IEP to describe the rate of the events as time elapses, inter-arrival times independence and finally, the conditions of stationarity needed to have confidence in the parameters estimation. Assumptions of non-monotonic failure behaviour, incorrect distributional assumptions, etc., equally may lead to biassed or inconsistent estimates of the parameters such that the results of both the classical and intelligent methods lose their validity. I would suggest that it would be important to discuss these assumptions on more detail, what consequences could be expected in case they are violated, and diagnostic instruments or goodness of fit tests should be used to check whether the model is suitable.

Even though the comparative outcomes are condensed using RMSE in different approaches, the discussion can be improved as complex interpretation on behalf of why particular approaches, including ANN and LSTM repeatedly provided better results than that of the conventional and metaheuristic approaches is given. An instance is how ANN and LSTM are at a high performance in dealing with non-linear time series and time-based data structures; therefore, complex temporal dynamics can be performed with ANN and LSTM but not the MLE and OLS because simpler forms of statistics are required. Similarly, the issues that led to moderate success of FFA and GWO was due to their performance of global search with limited local accuracy of convergence. Data complexity or sample size should be additionally broken down to add to the methodological validity and feasibility of the results. Even though this paper reveals the comparative advantages of each estimation technique, it is critical to reflect on the weaknesses of the estimation options. Such classical techniques as MLE and OLS, albeit analytically efficient, are susceptible to indirectness and statistical sparsity. The metaheuristics that are globally search robust, including FFA and GWO, can experience convergence problems and they cost more. ANN LSTM Deep learning Deep learning techniques have high accuracy but expensive in data quantity, a training time, and subject to overfitting unless well regularized. Consequently, the choice of methods ought to be technique-specific in line with the structure of data, sample size, and available resources to achieve the computation.

A careful implementation of ANN, LSTM, and metaheuristic techniques into safety-critical systems (e.g., dam failure prediction) should be done with caution given that the black-box models can lead to an increase in hidden causes of decisions. Ethical implications are transparency, interpretability, as well as a possible impact of the estimation error on society. Such issues should inform the choice of methods in practice.

Although the accuracy of such ANN or LSTM models is high, there is low interpretability of these advanced models. In practice, to make an informed decision, it is important to know the impact of the parameter estimates on the system performance in terms of failure patterns of the system or maintenance timetables. To make the interpretation easier and such that it can make differences, visualizations and sensitivity metrics can be implemented

12. Conclusions and Future work

This research presented an innovative framework to determine parameters of Inverse Exponential distribution based Non-Homogeneous Poisson Process models. Using a combined analysis of classical Maximum Likelihood Estimation and intelligent estimation methods with Ordinary Least Squares and Firefly Algorithm and Grey Wolf Optimization and Long Short-Term Memory networks and Artificial Neural Networks the study provides an in-depth comparison of estimations for time-varying stochastic processes. Simulation data showed intelligent approaches excel when estimating and are resilient at different sample levels through the use of ANN and LSTM. The real-world check using failure time information from Mosul Dam power station demonstrated that ANN generated the minimum Root Mean Squared Error (RMSE) exceeding classical and metaheuristic strategies. The study shows that data-oriented learning methods succeed in tracking changing processes of non-homogeneous systems particularly when these systems exhibit both non-linear structures and uncertain patterns. Additional research opportunities exist even though the existing results demonstrate success. The proposed framework requires expansion to handle multidimensional NHPPs since this will extend its practical application range to include predictive maintenance systems and multi-component reliability networks. Additional performance improvements concerning accuracy levels and uncertainty estimation capabilities could be obtained by including Bayesian estimation techniques together with ensemble learning models. Research should develop combination methods between neural network and evolutionary optimization to allow automatic improvements in learning rate settings and search parameter values. The deployment of proposed detection methods to real-time operations through online algorithms enables the system to adapt its fault detection capabilities in critical infrastructure monitoring applications. Scientists should conduct in-depth mathematical analysis of the theoretical characteristics which describe how estimators perform aspiring from asymptotic behavior to computational expense. Such directions enhance the current methodology while delivering substantial benefits for the entire field of stochastic process modeling alongside intelligent reliability engineering. Future studies may apply this framework to multidimensional NHPPs so that it may be used with multi-component reliability systems and networked environments. Besides, the use of Bayesian estimation methods to use prior information and measure uncertainty would also be able to improve interpretability and robustness when using limited data. Even though in the context of individual estimation techniques certain strengths were observed, considering the hybrid methods might bring even more improvements into performance. To give an example, the metaheuristic algorithms (e.g., GWO, FFA) and neural networks (ANN, LSTM) can be used together to ensure both effectiveness of the research and efficacy of the learning. This level of integration would allow autoigniting and fiddling of parameter settings leading to better convergence and robustness when applied to a variety of datasets.

Acknowledgements

The author is very grateful to the for providing access which allowed for more accurate data collection, improved the quality of this work.

REFERENCES

1. W. Hu, P. Westerlund, P. Hilber, C. Chen, and Z. Yang, A general model, estimation, and procedure for modeling recurrent failure process of high-voltage circuit breakers considering multivariate impacts, Reliab. Eng. Syst. Saf., vol. 220, p. 108276, 2022.
2. C. Li, Sequential Modelling and Inference of High-frequency Limit Order Book with State-space Models and Monte Carlo Algorithms. University of Cambridge, 2021.
3. S. Ahamad, Some studies on performability analysis of safety critical systems, Comput. Sci. Rev., vol. 39, p. 100319, 2021.
4. V. Chetlapalli, H. Agrawal, K. S. S. Iyer, M. A. Gregory, V. Potdar, and R. Nejabati, Performance evaluation of IoT networks: A product density approach, Comput. Commun., vol. 186, pp. 65–79, 2022.
5. Y. Wu, S. Tait, A. Nichols, and J. Raja, Simulation of railway drainage asset service condition degradation in the UK using a Markov chain–based approach, J. Infrastruct. Syst., vol. 27, no. 3, p. 4021023, 2021.
6. Hussain AS, Pati KD, Atiyah AK, Tashtoush MA. Rate of Occurrence Estimation in Geometric Processes with Maxwell Distribution: A Comparative Study between Artificial Intelligence and Classical Methods. Int. J. Advance Soft Compu. Appl. 2025 Mar;17(1).

7. S. Suman, S. Z. Khan, S. K. Das, and S. K. Chand, Slope stability analysis using artificial intelligence techniques, Nat. Hazards, vol. 84, pp. 727–748, 2016.
8. D. Wang et al., System impairment compensation in coherent optical communications by using a bio-inspired detector based on artificial neural network and genetic algorithm, Opt. Commun., vol. 399, pp. 1–12, 2017.
9. S. Ali, Mixture of the inverse Rayleigh distribution: Properties and estimation in a Bayesian framework, Appl. Math. Model., vol. 39, no. 2, pp. 515–530, 2015.
10. D. R. Cox and P. A. W. Lewis, The statistical analysis of series of events, 1966.
11. 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.
12. Hussain AS, Mahmood KB, Ibrahim IM, Jameel AF, Nawaz S, Tashtoush MA. Parameters Estimation of the Gompertz-Makeham Process in Non-Homogeneous Poisson Processes: Using Modified Maximum Likelihood Estimation and Artificial Intelligence Methods. 2025.
13. G. Srinivasa Rao, S. Mbwambo, and A. Pak, Estimation of multicomponent stress-strength reliability from exponentiated inverse Rayleigh distribution, J. Stat. Manag. Syst., vol. 24, no. 3, pp. 499–519, 2021.
14. Hussain AS, Sulaiman MS, Hussein SM, Az-Zo'bi EA, Tashtoush M. Advanced Parameter Estimation for the Gompertz-Makeham Process: A Comparative Study of MMLE, PSO, CS, and Bayesian Methods. Statistics, Optimization & Information Computing. 2025 Mar 6;13(6):2316-38.
15. R. Ghosh, N. Sinha, S. K. Biswas, and S. Phadikar, A modified grey wolf optimization based feature selection method from EEG for silent speech classification, J. Inf. Optim. Sci., vol. 40, no. 8, pp. 1639–1652, Nov. 2019, doi: 10.1080/02522667.2019.1703262.
16. A. Yelghi and C. Köse, A modified firefly algorithm for global minimum optimization, Appl. Soft Comput., vol. 62, pp. 29–44, 2018.
17. M. Stehlík, Exact likelihood ratio scale and homogeneity testing of some loss processes, Stat. Probab. Lett., vol. 76, no. 1, pp. 19–26, 2006.
18. Kang, Q., Yu, D., Cheong, K. H., & Wang, Z. (2024). Deterministic convergence analysis for regularized long short-term memory and its application to regression and multi-classification problems. Engineering Applications of Artificial Intelligence, 133, 108444. https://doi.org/10.1016/j.engappai.2024.108444.
19. Shafiq, A. (2023). Modeling and survival exploration of breast carcinoma: a statistical, maximum likelihood estimation, and artificial neural network perspective. Artificial Intelligence in the Life Sciences, 4, 100082.
20. Shafiq, A. (2022). Reliability analysis based on mixture of lindley distributions with artificial neural network. Advanced Theory and Simulations, 5(8), 2200100.
21. Wu, C., Wang, Q., Wang, X., Sun, S., Bai, J., Cui, D., & Sheng, H. (2024). Effect of Al2O3 nanoparticle dispersion on the thermal properties of a eutectic salt for solar power applications: Experimental and molecular simulation studies. Energy, 288, 129785.
22. Yue, Z., Shen, X., Chang, Z., & Zou, G. (2024). Experimental and simulation study on quasi-static and dynamic fracture toughness of 2A12-T4 aluminum alloy using CTS specimen. Theoretical and Applied Fracture Mechanics, 133, 104556.
23. Chupradit, S., Tashtoush, M., Ali, M., AL-Muttar, M., Sutarto, D., Chaudhary, P., Mahmudiono, T., Dwijendra, N., Alkhayyat, A. (2022). A Multi-Objective Mathematical Model for the Population-Base Transportation Network Planning. Industrial Engineering & Management Systems, 21(2), 322-331.
24. Hussain, A., Oraibi, Y., Mashikhin, Z., Jameel, A., Tashtoush, M., Az-Zo'bi, E. (2025). New Software Reliability Growth Model: Piratical Swarm Optimization -Base Parameter Estimation in Environments with Uncertainty and Dependent Failures, Statistics, Optimization & Information Computing, 13(1), 209-221.
25. Birnbaum, Z.W., & Saunders, S.C. (1969). Estimation for a family of life distributions with applications to fatigue, Journal of Applied Probability, 6, 328-347.
26. Nelson, W., & Doganaksoy, N. (1995). Statistical analysis of life or strength data from specimens of various sizes using the power-(log) normal model. Recent Advances in Life-Testing and Reliability, 377-408.