STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
*Stat., Optim. Inf. Comput.*, Vol. **x**, Month 202x, pp 0–31.
Published online in International Academic Press (www.IAPress.org)

# Survival Modelling of Breast and Brain Cancer Using Statistical Maximum Likelihood and SVM Techniques

Nidhal Q. Saadoon [1], Hiba Ali Salman [2], Adel S. Hussain [3], Emad A. Az-Zo'bi [4], Mohammad A. Tashtoush [5,6,*]

[1]*Administrative & Financial Affairs Department, Al-Iraqia University, Baghdad, Iraq*
[2]*Electrical Computer Centre, Al-Iraqia University, Baghdad, Iraq.*
[3]*IT Department, Amedi Technical Institutes, University of Duhok Polytechnic, Duhok, Iraq*
[4]*Department of Mathematics and Statistics, Faculty of Science, Mutah University, Mutah P.O.Box 7, Al-Karak 61710, Jordan*
[5]*Department of Basic Sciences, Al-Huson University College, Al-Balqa Applied University, Salt 19117, Jordan*
[6]*Faculty of Education and Arts, Sohar University, Sohar 311, Oman*

**Abstract** The research focuses on two main objectives that examine the Burr Type XII distribution through MLE parameter estimation and a comparison between MLE and SVM methods. Survival-related functions such as the survival function and hazard rate and other derived reliability measures are estimated by executing both methods on breast and brain cancer patient real-world data. The input layer of the proposed SVM framework contains distribution parameter specifications that produce output estimates for the reliability function and hazard rate function as well as probability density function, reversed hazard rate function, mills ratio, and odds function. The research data shows how the hazard function grows after diagnosis then declines toward the end of the study period which reflects the theoretical behavior patterns of Burr Type XII distributions. The survival analysis demonstrates that theoretical characteristics of the Burr Type XII distribution match experimental results thus validating its usage as cancer survival data model. This SVM method shows itself to be an accurate and stable approach for critical survival parameter prediction.

**Keywords** Survival function, Risk function, Burr Type XII distribution, maximum likelihood Estimation, Support Vector Machine

**AMS 2010 subject classifications** 60J80, 60J85, 60K10

**DOI:** 10.19139/soic-2310-5070-2619

## 1. Introduction

The disease of breast carcinoma persists as a major public health challenge due to its fatal properties when it forms inside human bodies. The identification of breast carcinoma at an early stage, along with immediate medical treatment, leads to delayed disease evolution and minimizes patient death. Breast carcinoma stands as a leading cancer among women and represents the second leading cause for cancer deaths in American Indian/Native American and Asian Pacific Islander women [1, 2]. The breast tissue origin of the disease produces uncontrollable malignant cell growth that spreads throughout the body to kill patients. Despite breast carcinoma mostly affecting female patients, it develops infrequently in men to such an extent that medical records show less than 0.05% of cases [3]. The breast cancer types, ductal carcinoma and lobular carcinoma, represent the main clinical classifications according to their location of origin between milk ducts and milk-producing lobules of the breast [4, 5, 6]. The origin of breast carcinoma occurs in extremely unusual situations in areas not pertaining to the

*Correspondence to: Mohammad A. Tashtoush (Email: tashtoushzz@su.edu.om). Depatrment of Basic Sciences, AL-Huson University College, AL-Balqa Applied University, Salt 19117, Jordan, and Faculty of Education and Arts, Sohar University, Sohar 311, Oman.

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

ducts or lobules of the breast. The primary method of preventing breast cancer fatalities is through early detection. Three main strategies provide the best chance of finding treatable breast cancer mutations: regular mammograms and clinical breast exams from healthcare professionals, along with monthly self-examinations carried out by individuals [7].

The effectiveness of artificial intelligence methodologies in modelling and forecasting complex health-related data received significant attention through artificial neural networks (ANNs). Scientific reports from [8, 9, 10] support this evidence. Different research findings show that artificial neural networks produce forecasting results equivalent to or better than standard statistical and mathematical prediction models. The researchers explored precise hazard function estimation for survival data through statistical modelling-based techniques from [11]. An analysis by their research group established a multilayer perceptron (MLP) as an extended framework of generalized linear models (GLMs) featuring multinomial error structures with a nonlinear estimator designed to support discrete-time survival distributions among competing risks scenarios. The standard weight regularization approaches were implemented to solve model complexity alongside overfitting issues. A Genetic Algorithm served as the tool to optimize independently the model parameters through adaptive optimization of neural network complexity. The proposed method obtained practical validation from its implementation among 1,793 female patients diagnosed with breast cancer and no lymph node involvement in the axilla. [12] conducted research that used ANN modelling together with maximum likelihood estimation (MLE) to determine COVID-19 mortality rates across Italian territories. The ANNs implemented by this study included nine hidden layer neurons, which produced a minimum deviation of -0.14% along with an R of 0.99836 to prove model accuracy. The analysis from both approaches showed reliable performance, which supported their research validity. The research of [13] examined how machine learning algorithms perform in survival analysis for predicting the prognosis of bladder cancer patients. Their research concentrated on analysing the forecasting capabilities of ANNs through their examination of demographic alongside clinical features to predict survival outcomes. The research evaluated Convolutional Neural Networks as deep learning models against multivariate Cox proportional hazards models to forecast 5-year survival and lifetime duration in genetic disorder patients [13, 14]. The CNN models conducted training and validation operations through 80% training and 20% validation data allocation for performing reliable assessments. Machine learning methodologies create a strong method to extract valuable patterns from complex, high-dimensional clinical information because administrative databases have substantial restrictions. The research data demonstrates that CNNs and machine learning systems show substantial potential for improving breast carcinoma risk predictions, which helps doctors create better treatment strategies [15].

A right-skewed log-logistic distribution stands as one of the key parametric models that researchers extensively use for survival analysis studies. The LL model surpasses the Weibull distribution in its ability because it features non-monotonic hazard functions that occur after $\epsilon > 1$ shape parameter, while modelling complicated hazard patterns with their initial rise followed by decline. $\epsilon > 1$ covariance values lower than or equal to one produce a hazard function with a unimodal shape and decreasing pattern in the LL distribution, which leads to wide adaptability across different survival models. The main benefit of using LL distribution is its capability to provide direct mathematical solutions for survival probability and hazard functions, which enables simple estimation and interpretation during censored data analysis. This beneficial characteristic demonstrates why the LL distribution serves as an excellent analytical instrument for theoretical work and practical survival analysis applications. The statistical characteristics, along with practical uses of the LL distribution, have been extensively researched by various studies. Literature [16, 17] provides a detailed examination of LL distribution order statistics, while [18] derived recurrence relations that determine the moments of order statistics from LL distribution samples. The LL distribution serves researchers and practitioners through its theoretical developments and model applications as described in references [19, 20].

Researchers widely recognize the Burr Type XII distribution as a skew probability model which demonstrates high flexibility during survival and reliability analysis [20]. Statisticians prefer this distribution as a survival analysis tool because it does better than conventional distribution methods in describing data patterns. The latest research on the Burr Type XII distribution introduces modified versions that demonstrate better capabilities for modelling both survival data censors and reliability analysis problems. Research on modified Burr Type XII

versions has led to successful practical validations which demonstrate its effectiveness under different hazard rate characteristics. Researchers have created Bayesian inference tools for working with the Burr Type XII distribution which proves its usefulness for modelling lifetime data [21]. Research has achieved better predictions through multiple loss functions and prior distributions which has allowed scientists to apply their models successfully to survival-related datasets [22]. New extended versions of the Burr Type XII distribution were developed to handle sophisticated hazard functions successfully when analysing medical survival data. The expanded models demonstrate excellent ability to predict various hazard rate patterns ranging from increasing to decreasing patterns or bathtub-shaped behaviors which strengthens their significance in survival analysis and risk modelling [23]. Reliability prediction receives improvements from survival modelling because scientists combined SVM and ANN with RF with parametric lifetime distributions from the Burr family. The researchers demonstrate that Burr family distributions serve practical needs when modelling time-to-failure across different real-world applications despite their non-exclusive use of Type XII distributions [24]. New discussions about survival and reliability standards reveal that Burr XII distribution models contribute importance by functioning in machine learning hybrid systems. Reliability engineering problems experience improved classification methods by using statistical model structures along with SVM algorithm learning capabilities [25].

The field of mathematical statistics relies heavily on order statistics because they form its essential foundation, which serves both theoretical and practical aspects of statistics. The ordered data analyses that rely on order statistics gain importance in multiple statistical fields, including nonparametric inference and reliability analysis and survival modelling,the because the distributional behavior can be understood through robust estimation processes. Thorough theoretical analysis of statistical inference theory at present time depends heavily on order statistics for developing distribution-free methods and studying sampling distributions. The practical application of statistical inference using ordered sampling enables the development of efficient computation methods, resulting in results that are easy to interpret, particularly when researching goodness-of-fit testing and life testing experiments under censored or ranked data evaluation conditions. Order statistics-based models provide ideal conditions for production of robust approaches in outlier detection and tail probability evaluations and statistical estimation and hypothesis testing. Order statistics continue to offer indispensable theoretical and practical value, which helps develop classical and modern statistical methodologies according to [25].

The research field of survival analysis experienced significant growth recently because this analytical method proves essential for biosciences and pharmacological investigations. The key objective within this field involves evaluating survival function estimators while employing an accurate Probability Density Function (PDF) to describe time-to-event data which involves breast carcinoma survival times from female patients. The research adopts Burr Type XII distribution as an adaptable and robust model to represent survival dynamics. The systematic assessment of the Burr Type XII survival model by using Maximum Likelihood Estimation along with Support Vector Machine remains absent from previous research. Through the integration of these three methodologies the main goal of this research becomes analysing and optimizing lifetime data prediction reliability. We utilize this research design as a means to present an innovative viewpoint in the field while resolving the significant shortage of hybrid modelling techniques for survival data research.

The presented study works thoroughly to build an advanced framework that produces precise survival and failure time predictions for breast cancer patients. Patient data from Medical City Hospital – Baghdad (2020) underwent systematic analysis through two methodology approaches which tracked survival duration from hospital entry until patient discharge due to all patients experiencing death. The successful implementation of survival analysis depended on our uses of the Burr Type XII distribution along with a predictive enhancement through SVM and MLE methods. The proposed modelling system matches current development in survival analytics and machine learning since it enhances diagnosis predictions while improving survival function interpretation. Our research design implements contemporary survival modelling approaches to ensure proper testing of these methods in breast cancer prognosis assessment.

1. The research analyses the reliability characteristics of the Burr Type XII distribution through a new set of features within SVM frameworks which were previously absent from existing literature.
2. The research aims to create predictive systems from SVM which detect and improve the closed-form properties of the Burr Type XII distribution for extensive survival analysis applications.

3. The research will assess SVM methodologies as they relate to estimating and forecasting distributional characteristics which derive from survival data of the Burr Type XII framework.
4. The proposed models will be tested against each other for maximum accuracy and robustness through Maximum Likelihood Estimation (MLE) and Support Vector Machine (SVM) modelling assessment resulting in the identification of a superior survival time prediction and distribution fitting solution.

## 2. The Burr Type XII Distribution and its Properties (Proposed Model)

The probability that the survival time remains lower than a specific time point $t$ can be calculated through the cumulative distribution function (CDF) of the random variable $T$ for the Burr Type XII distribution where $c$ and $k$ represent shape and scale parameters respectively.

$$G(t|c,k) = 1 - (1+t^c)^{-k}, t, c, k > 0, \tag{1}$$

and PDF of Burr Type XII distribution is:

$$f(t|c,k) = ckt^{c-1}(1+t^c)^{-k-1}, t, c, k > 0. \tag{2}$$

The Burr Type XII distribution functions in multiple research spaces because it both handles statistical distribution distortions and adapts to non-standard data distributions as shown in Figure 1. The distribution supports various hazard rate patterns, which enables its use in advanced theoretical as well as applied time-to-event data analysis scenarios.

### *2.1. Survival and Risk Function*

Survival analysis serves as a necessary statistical framework that experts use in medicine as well as biology and social sciences and econometrics and engineering [20]. Statistical survival analysis depends on event timing intervals in measurements spanning multiple years. The statistical discipline which examines and analyses time-to-event data through survival analysis or survival evaluation operates under the title of survival analysis [17, 21]. Survival time $T$ serves as the main variable of the study to measure the interval from an established beginning (birth or treatment start) to an end outcome (death or treatment failure). Subsequent events that theoretically could happen are excluded by this terminology since they do not fit the definition of terminal events. Survival analysis relies on $S(t)$ as its central quantitative element to represent the chances of individual survival past time $t$. The survival function $S(t)$ indicates the probability that survival time $T$ surpasses value $t$ in the nonnegative continuous random variables space.

$$S\left(t|c,k\right) = \Pr(T > t) = 1 - G(t|c,k) = (1+t^c)^{-k}, t, c, k > 0. \tag{3}$$

The survival function $S(t)$ behaves in a decreasing manner following $t = 0$ when it starts at 1 before moving toward zero as time tends toward infinity. Survival probability decreases throughout time according to the basic principle of this fundamental property. The survival function curves that show different shape-scale variations in the Burr Type XII distribution appear in Figure 2. The distribution demonstrates its modelling capacity for various survival patterns through the expected monotonic non-increasing patterns displayed on the plots. The survival function $S(t)$ along with the probability density function $f(t)$ receive a three-dimensional graphical representation in Figure 3. The graphic Visualization allows us to understand how survival probability and duration evolve together when parameters and time change within the Burr XII distribution framework.

In survival analysis, the hazard rate function acts as a fundamental principle that statisticians name the instantaneous failure rate. The hazard rate provides a measure of instantaneous failure probability, together with event occurrence risk at a specific time t for surviving up to that point. Biomedical research depends on hazard rate functions to determine how likely patients are to die during a specific time interval t when they reach that timeframe. Hazard rates represent the natural tendency of components to deteriorate throughout their life span in applications of reliability engineering. A random variable named $T$ with a non-negative continuous distribution

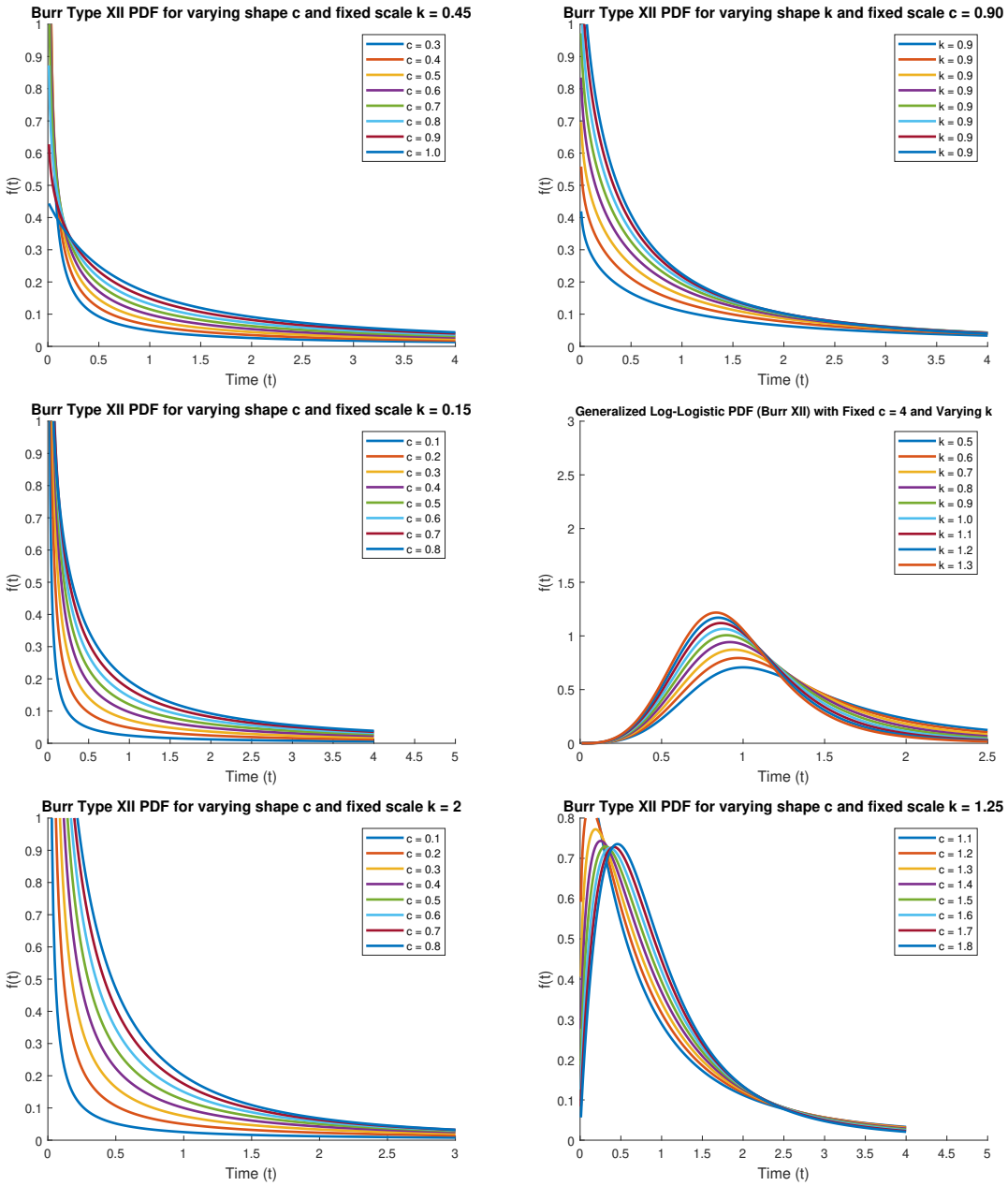


Figure 1. Burr Type XII Distribution: Behavior of PDF Under Varying Shape Parameters $c$ and $k$.

represents survival time or time-to-event. The mathematical definition of hazard rate function $h(t)$ appears as follows [26]:

$$h\left(t\right)=\lim_{\Delta t\to 0}\frac{P\left(t\leq T\leq t+\Delta t|T\geq t\right)}{\Delta t}=\frac{f\left(t\right)}{S\left(t\right)}, \tag{4}$$

where $f(t)$ is the probability density function (PDF) of the survival time, $S(t)=P(T>t)$ is the survival function, representing the probability of surviving beyond time $t$.

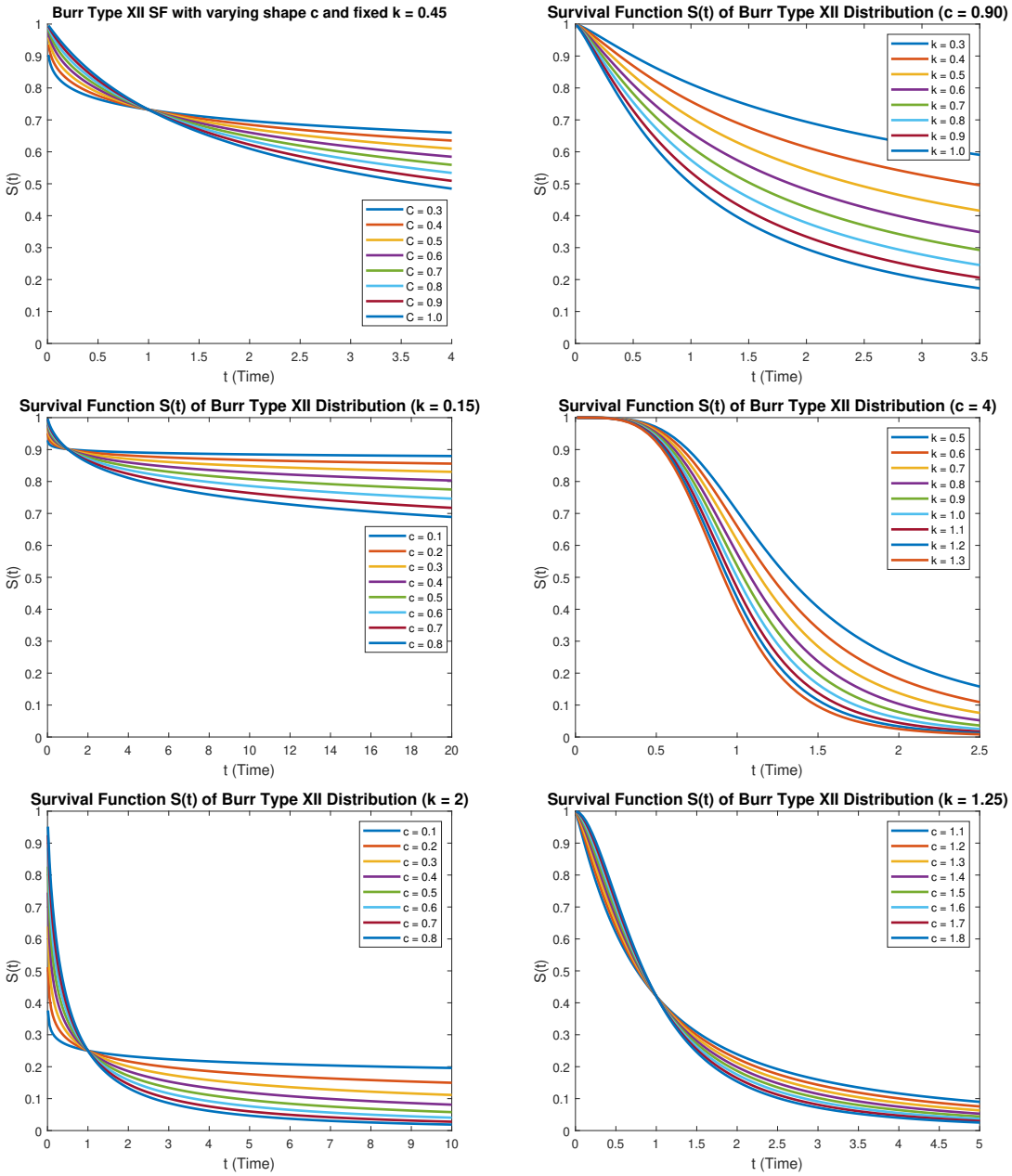


Figure 2. Burr Type XII Distribution: behavior of SF Under Varying Shape Parameters $c$ and $k$

The hazard rate function $h(t)$ stands in direct relation to the survival function $S_T(t)$ through this formula. $h\left(t\right) = \frac{S'_T(t)}{S_T(t)}$. The hazard rate represents an instantaneous failure speed at time $t$ for survivors until then, since it provides a slope measurement of survival curve variation. The hazard rate function and survival function explain the survival time distribution of $T$ through Eq. (3) and Eq. (4) in an interrelated manner.

• This method evaluates risk failure probabilities during specific time intervals while taking surviving until that period into account for precise age-related hazard analysis.

• The method enables direct comparisons between different groups including populations as well as treatment

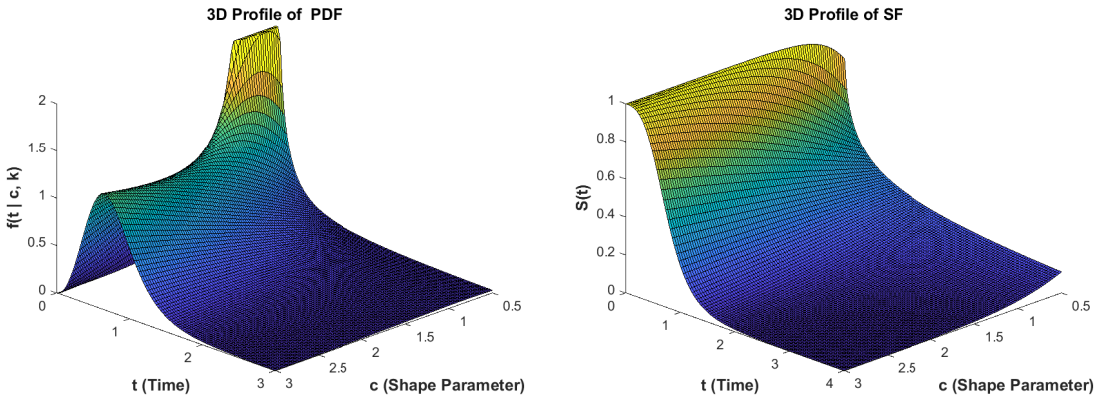


Figure 3. 3D profiles of PDF and SF

populations especially for clinical and epidemiological research.
• The proportional hazard model provides exceptional value when dealing with right censored data that contains heterogeneous failure types therefore serving as a strong tool for handling complex survival data.
• The hazard function enables straightforward comparison with the exponential distribution for modelling purposes because it allows evaluation of failure rate consistency. Time-dependent risk analysis becomes possible through deviations from the benchmark distribution since it reveals patterns that lead to improved modelling of survival scenarios.
• Reliability theory finds this model to be its essential building block because it describes the base intensity rates that affect one-component devices with a single failure type.

The hazard rate function (HRF) of Burr Type XII distribution.

$$h\left(t|c,k\right)=\frac{ckt^{c-1}}{1+t^{c}}. \tag{5}$$

The illustrations of hazard rate function (HRF) profiles appear in Figure 4 and Figure 5. Each distinctive shape of the HRF exists in decreasing patterns and unimodal and hump-shaped forms that represent beneficial characteristics for lifetime modelling. Real-world reliability analysis benefits from this model because it can handle both single-faceted and multi-faceted hazard trends, which occurs frequently in survival data systems.

The survival analysis depends heavily on the cumulative hazard rate function (CHRF) denoted by $G'(t/c,k)$, which represents the complete risk exposure of subjects from their beginning until time $t$. The CHRF enables researchers to understand risk build-up patterns for the Burr Type XII distribution while jointly providing information with the hazard and survival function for studying lifetime behavior.

$$G'(t|c,k)=\int_{0}^{t}h\left(y|c,k\right)dy=-\log\left[S\left(t|c,k\right)\right]. \tag{6}$$

Hence,

$$G'(t|c,k)=k\log\left(1+t^{c}\right). \tag{7}$$

You can see in Figure 6 how specific distributional parameters affect the cumulative hazard rate function (CHRF). The CHRF from the Burr Type XII distribution curves uniformly upwards to represent risk growth that follows patterns found in various survival and reliability situations.

A random variable that stands for lifetime duration possesses its reverse hazard rate function when the ratio between the the probability density function and cumulative distribution function exists. Mathematically, RHRF expresses failure probabilities immediately before time $t$ under the condition that the event has occurred by time

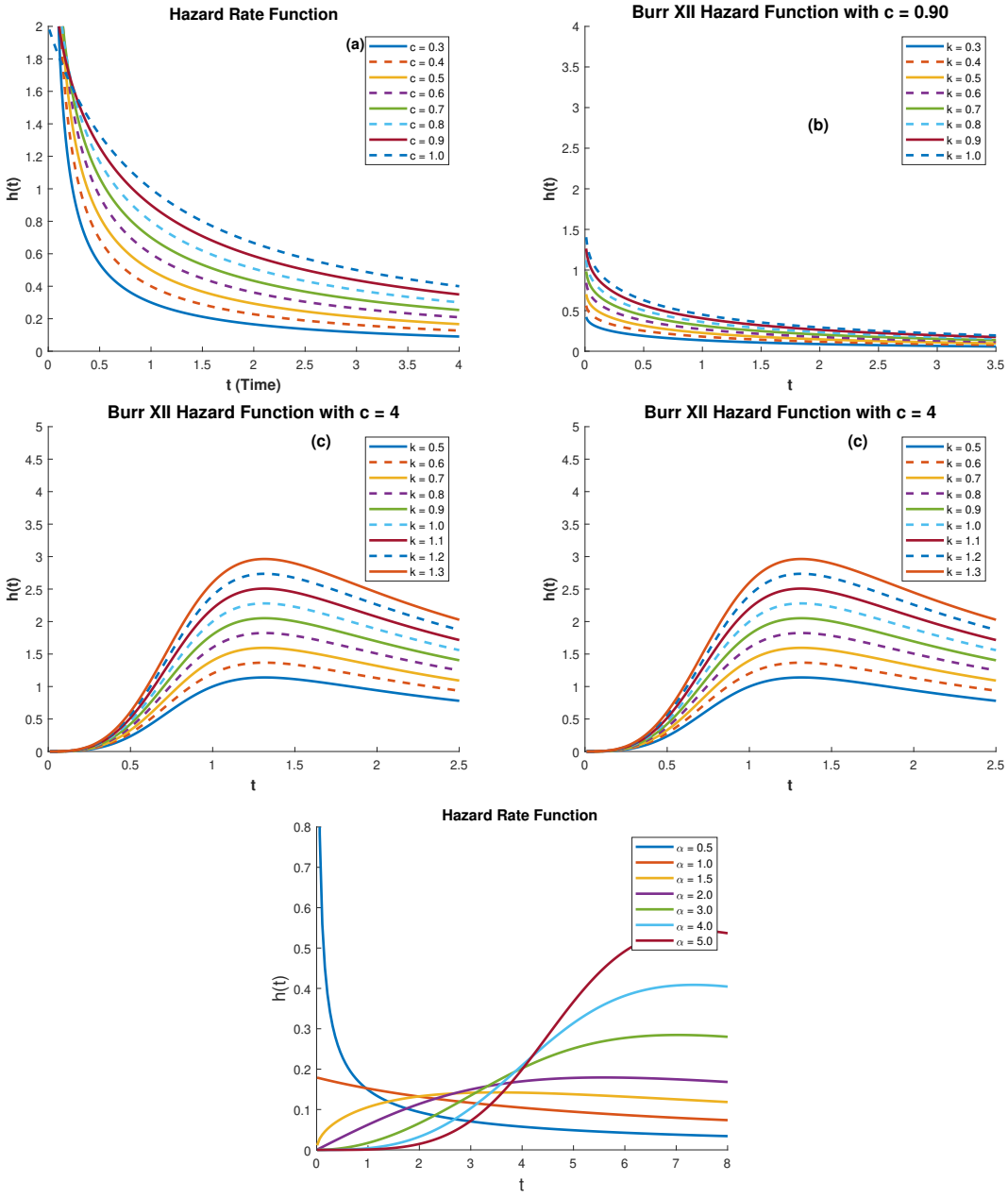


Figure 4. Features of $c$ and $k$ on HRF

$t$. The left-censored data analysis depends on RHRF for its importance in fields such as forensic science as well as actuarial studies and early failure modelling. The reliability function for survival data in terms of time $T$ can be written in this form: $h'(t|c,k) = f(t|c,k)G(t|c,k)^{-1}$, Systems that experience most failures during their initial observation period can be efficiently analysed through this function. Research in reliability theory and stochastic modelling demonstrates the essential role of RHRF for distribution characterisation, development of diagnostic checks, and new reliability model construction (see [27] for mathematical theory and [28] for utilisation and broader

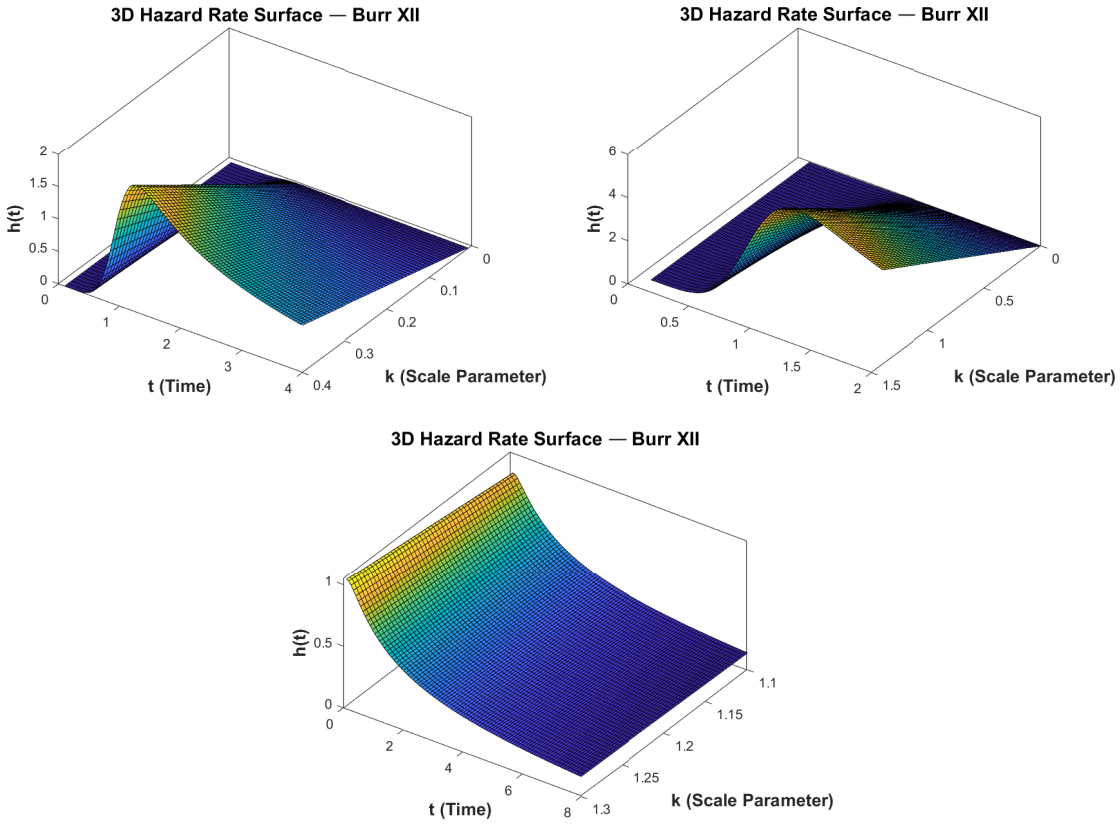


Figure 5. 3D profile of HRF of Burr Type XII distribution

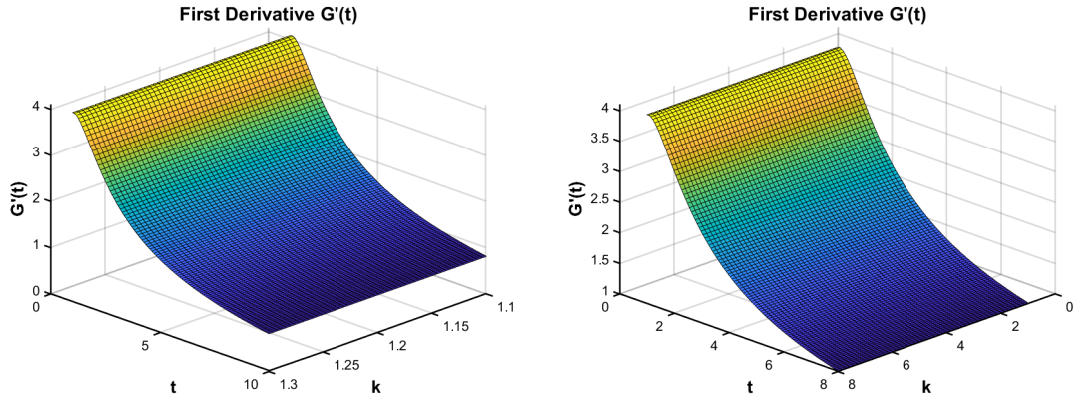


Figure 6. Features of $c$ and $k$ on CHRF

applications) in Figure 7.

$$h'(t|c,k) = \frac{ck}{t(1+t^c)} \tag{8}$$

Figure 8 shows the Mills ratio, due to its intrinsic relationship with the failure (hazard) rate, serves as a valuable analytical tool in the assessment of reliability and durability. The distribution tail behavior comes into focus through

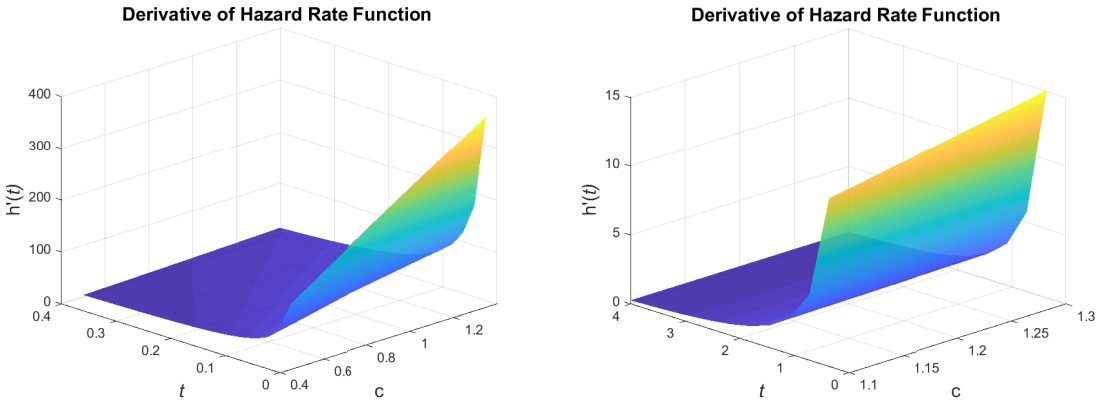


Figure 7. Features of $c$ and $k$ on RHRF

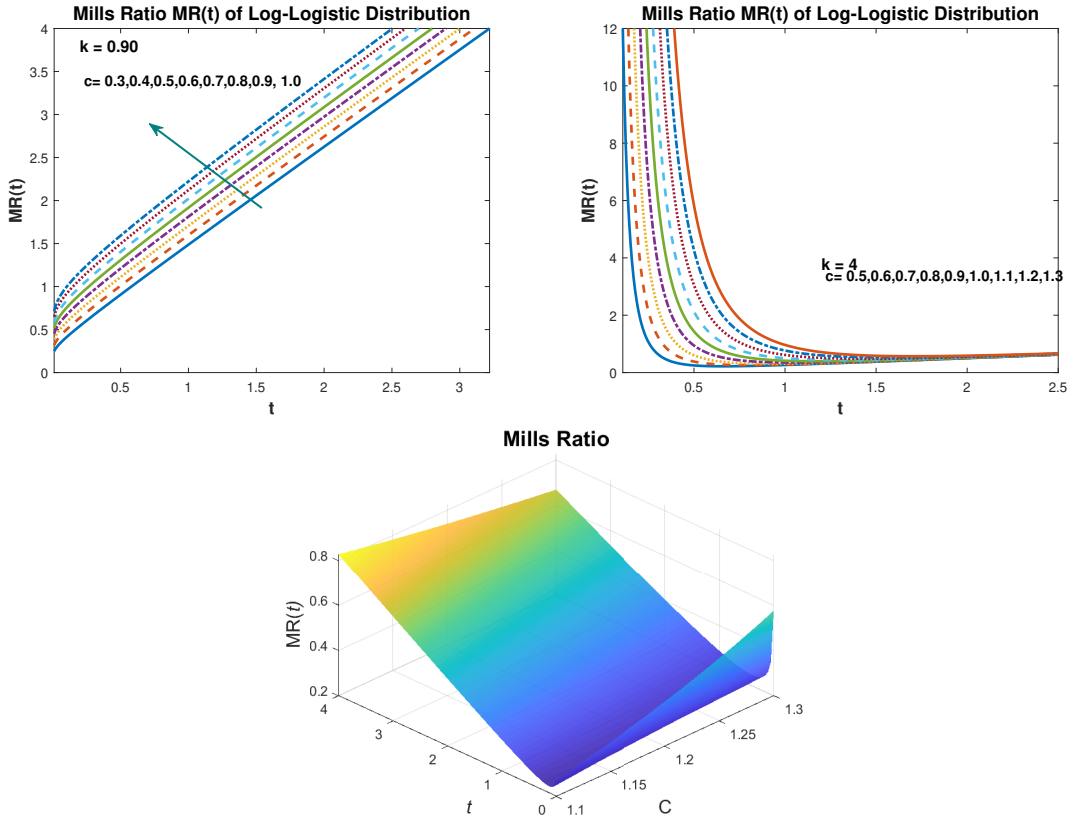


Figure 8. Features of $c$ and $k$ on MR

this metric which proves especially valuable for survival analysis of systems throughout their operational periods.

$$MR(t|c,k) = \frac{t}{ck}\left(1+t^{-c}\right) \tag{9}$$

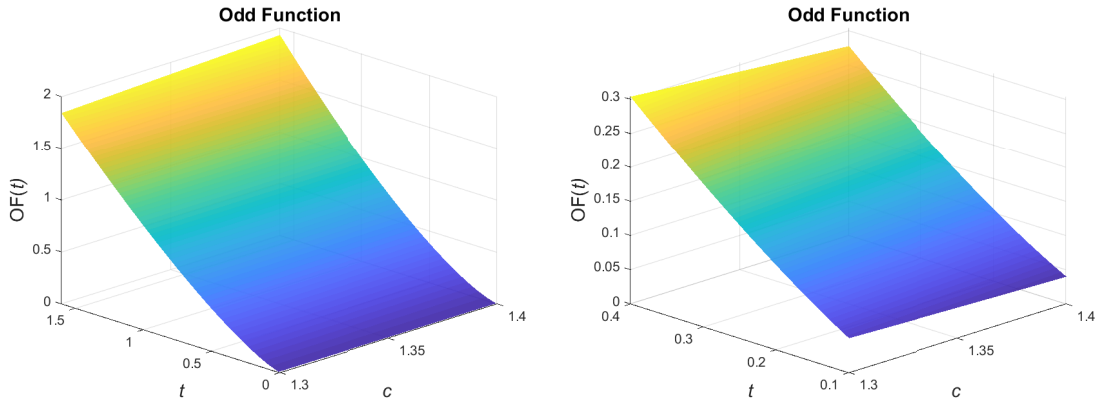


Figure 9. Features of $c$ and $k$ on OF

The odd function of $T\ OF\ (t|c,k) = \frac{G(t|c,k)}{S(t|c,k)}$ is:

$$OF\ (t|c,k) = \frac{1}{(1+t^c)^{-k}} - 1. \tag{10}$$

The function is plotted as shown in Figure 9.

## 3. Estimation Methods

### *3.1. Maximum Likelihood Estimation (MLE)*

This estimation method is recognized as one of the most effective approaches due to its desirable statistical properties, including stability, asymptotic efficiency, and consistency under certain conditions. Let us assume a random sample of size $n$ drawn from a log-logistic distribution, denoted by $t_1, t_2, \ldots, t_n$. The corresponding log-likelihood function for the parameters c and k is expressed. And the maximum likelihood estimation (MLE) function contains the probability density function $t$ that has been substituted into the general log-likelihood expression as presented in this statement:

$$f_n\left(t_1, t_2, \ldots, t_n\right) = \mathrm{n}! \prod_{\mathrm{i}=1}^{\mathrm{n}} f\left(t_i\right). \tag{11}$$

Then,

$$L(c,k/t) = n! \prod_{i=1}^{n} f(t_i) = n! \prod_{i=1}^{n} ckt_i^{c-1}(1+t_i^c)^{-k-1}, \tag{12}$$

Hence,

$$L(c,k/t) = n!c^n k^n \prod_{i=1}^{n} t_i^{c-1} \prod_{i=1}^{n} (1+t_i^c)^{-k-1}. \tag{13}$$

The natural logarithm of the likelihood function replaces the likelihood function in analysis due to ease of operation through Eq. (13). Taking the exponential of the natural logarithm of the likelihood function produces this expression for the log-likelihood:

$$\ln L(c,k|t) = \ln n! + n\ln c + n\ln k + (c-1)\sum_{i=1}^{n} \ln\left(t_i\right) - (k+1)\sum_{i=1}^{n} \ln\left(1+{t_i}^c\right). \tag{14}$$

The derivative $\frac{\partial \ln L(.)}{\partial(.)}$ w.r.t. $c$ and $k$ are:

$$\frac{\partial \ln L(c,k|t)}{\partial c} = \frac{n}{c} + \sum_{i=1}^{n} \ln(t_i) - c(k+1) \sum_{i=1}^{n} \frac{{t_i}^{c-1}}{1+{t_i}^{c}}, \tag{15}$$

$$\frac{\partial \ln L(c,k|t)}{\partial k} = \frac{n}{k} - \sum_{i=1}^{n} \ln(1+{t_i}^{c}). \tag{16}$$

Appropriate numerical optimization techniques are used to solve the system derived when the partial derivatives of log-likelihood functions in Eqs. (15) and (16) become zero for parameter estimation of $c$ and $k$. The survival analysis applies parameter estimates for the Burr Type XII distribution through subsequent investigations reported.

**Survival Function**

Estimating the survival function that has been substituted into the general log-likelihood expression as presented in this statement:

$$L(t|c,k) = n! \prod_{i=1}^{n} S(t_i|c,k) = n! \prod_{i=1}^{n} (1+{t_i}^{c})^{-k}. \tag{17}$$

The natural logarithm of the likelihood function replaces the likelihood function in analysis due to ease of operation through Eq. (17). Taking the exponential of the natural logarithm of the likelihood function produces this expression for the log-likelihood:

$$\ln L(\mathrm{t}|\mathrm{c},\mathrm{k}) = \ln n! - k \sum_{i=1}^{n} \ln(1+{t_i}^{c}). \tag{18}$$

The derivative $\frac{\partial \ln L(.)}{\partial(.)}$ w.r.t. $c$ and $k$ are:

$$\frac{\partial \ln L(c,k|t)}{\partial c} = -ck \sum_{i=1}^{n} \frac{{t_i}^{c-1}}{1+{t_i}^{c}}, \tag{19}$$

$$\frac{\partial \ln L(c,k|t)}{\partial k} = -\sum_{i=1}^{n} \ln(1+{t_i}^{c}). \tag{20}$$

### *3.2. Modified Support Vector Machine*

The goal is to use SVM regression to approximate the relationship between the input data y and the output of the function $f(t;c,k)$ and $S(t;c,k)$, and then extract the parameter estimates $\hat{c}$ and $\hat{k}$. The following algorithm:

BEGIN
Step 1: Data Preparation:
- Normalize dataset $D$ to improve numerical stability.
- Split $D$ into training set $D_{train}$ and validation set $D_{val}$.
Step 2: Train SVM Regression Model:
- Initialize the SVM model using the specified Kernel Type
- Train SVM using $D_{train}$ to learn mapping: $t_i \rightarrow f_i$ .
Step 3: Parameter Initialization:
- Set $c \leftarrow c_0$
- Set $k \leftarrow k_0$.
Step 4: Optimization Loop:
-Define objective function: $MSE = \frac{1}{n} \sum_{i=1}^{n} (f_i - f(t_i;c,k))^2$.
- Use Optimization Method to minimize MSE.

For $iter = 1$ to Max Iterations DO
Update $(c, k)$ based on optimization rule.
Compute MSE.
End for.
Step 5: Parameter Extraction:
- Set $\hat{c} \leftarrow$ optimized value of $c$.
- Set $\hat{k} \leftarrow$ optimized value of $k$.
Step 6: Validation:
For each $t_j$ in $D_{val}$ DO
Compute predicted output $\hat{f}_j$ by using $SVM$.
Compute theoretical output $f\left(t_j; \hat{c}, \hat{k}\right)$.
End for
Compute $RMSE = \sqrt{\frac{\sum_{i=1}^{Q}\left(\widehat{f}_i - f\left(t_j;\hat{c},\hat{k}\right)\right)^2}{Q}}$.
Output $\hat{c}$, $\hat{k}$ and $RMSE$.
End.

In order to overcome the issue of interpretability of models, we offer the study of the inner behavior of SVM model. The action of the kernel function in changing the feature space, the effect of the support vectors on the regression function and decision boundaries based on margins are described in detail to explain the predictive mechanism. Moreover, post-hoc explainability methods like SHAP values were utilized in order to determine feature contributions to the model outputs. These observations provide a clear understanding of how the model works and make it more useful in the context of clinical decision making.

Firstly, the kernel function that was used in SVM model was Radial Basis Function (RBF) kernel which is flexible in handling non-linear patterns of data as well documented in its application in survival analysis. Second, grid search strategy and 10-fold cross-validation were used to optimize the hyper-parameters, in which the regularization parameter $C$ and the kernel scale $\gamma$ were optimized. These values were the best because they presented the least Root Mean Square Error (RMSE) throughout validation folds. Third, regarding the characteristic of inputs, the current SVM model has just used time-to-event variable following Burr Type XII distribution in modelling learning. Even though this approach enables us to compare our estimation with those estimated by MLE directly, we recognize that since there are no clinical covariates (e.g., age, tumour stage) the model may have little clinical utility. We plan to enlarge the SVM input space in future to include such covariates in order to achieve a more interpretable input space that will be of greater practical value. In particular, we used the Radial Basis Function (RBF) kernel that is generally useful when there is a need to model non-linear relationships that are prevalent in survival analysis. The important hyper-parameters of SVM were:
- Regularization parameter $c = 1.0$.
- Kernel coefficient $\gamma =' scale'$, $\gamma = \frac{1}{n_{features} * Var(X)}$.

They were optimized in order to provide the best performance in terms of the grid search method with 10-fold cross-validation and the best configuration was the one that provided the minimum average Root Mean Square Error (RMSE) across folds. Regarding input features, SVM model in this paper assumed time to event as the sole input feature according to the Burr Type XII distribution structure form. In order to enhance the numerical stability and the convergence of the model, all of the input data were modified to a standard scale (zero mean, unit variance). To further assist reproducibility, we have also provided a flowchart of the SVM pipeline, i.e., the data pre-processing steps, training and tuning of the model and prediction steps.

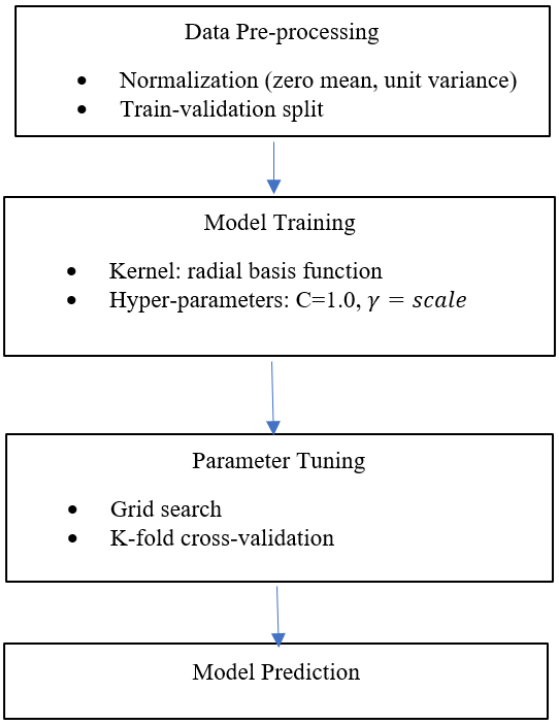


Figure 10. The SVM pipeline flowchart

## 4. Simulation

Simulation operates as a methodological instrument which duplicates real-world systems during controlled experimental analysis for the purpose of behavior investigation. The basic notion consists of creating an artificial reproduction of a process which reproduces actual system actions under precise conditions. The process requires building a reduced or easy-to-understand model which effectively represents both the essential structures and operational principles of the complete system. The surrogate model undergoes experimental testing which leads to analysis results that provide implications for the full-scale system. Simulation conducted through computers relates to the development and execution of mathematical algorithms which generate models that reflect real programming scenarios. The system performs iterative simulations which allow for output observation and analysis to draw statistically valid conclusions regarding performance and behavior in conditions of uncertainty [29].

The most well-known simulation approaches consist of the analogy method along with the mixed method and the widely-used Monte Carlo method. The Monte Carlo simulation stands out because it uses probabilistic sampling approaches to duplicate complex phenomena subject to known probability distributions. The simulation technique needs a specified cumulative distribution function (CDF) to create Independent Identically Distributed (IID) random samples. The system treats observations one by one by performing appropriate mathematical treatments that maintain statistical validity and replication potential [30]. This research applied simulation to evaluate estimators of the proposed model under fuzziness through the practical conceptualization of these theoretical constructs. The proposed estimation approaches underwent experimental comparison through a simulation testing scheme. The Root Mean Square Error criterion served as a performance evaluation tool for estimators while they operated with different sample size conditions thus determining their precision level [31]. The simulation model served to evaluate the suggested estimation methods comparatively under different data environments which represented actual real-world scenarios. This research analyzes various estimation methods to establish the most

reliable and efficient method for determining parameters within the Burr Type XII distribution. An approach like this helps decision making through better evaluation of lifetime data and other domains featuring heavy-tailed behavior and extreme value phenomena [31, 32].
• Change in sample size.
• Change in model parameter values.

**Stage I: Model Initialization and Parameter Specification**

The first step plays a pivotal role by creating a base for all following simulation processes. The first step includes all operations that establish core hypothesis along with parameter value selection while defining process behavior. This phase contains three sequential elements for completion:

**Step 1: Default Parameter Values get Selected During this First Step of the Procedure**
The simulation process starts by setting initial default values to the parameters used in Burr Type XII Process. The chosen parameter settings draw from past experimental studies together with comprehensive testing work to maintain robustness and applicability of configured parameters. Two specified parameter configurations showed the best results from evaluating different simulation parameter options. Set 1: $c = 0.2; k = 2.0$, Set 2: $c = 2.5; k = 2.5$, Set 3: $c = 2.7; k = 4.6$ and Set 4: $c = 3.0; k = 5.3$.

These parameters respectively define the shape, scale, location, and additional distributional characteristics necessary for generating synthetic data that closely resemble the theoretical behavior of the Burr Type XII.

**Step 2: Determination of Sample Sizes**

Different sample sizes of small medium and large datasets successfully measure the stability and performance of the estimators during the simulation. $n = 50; 100; 200; 400$. This stratification allows for rigorous analysis of estimator sensitivity and efficiency under varying data volumes.

**Stage II: Random Data Generation via Inverse Transformation**

This stage involves the generation of pseudo-random data points that follow the probability distribution function of the Burr Type XII Process, utilizing the Inverse Transform Sampling Method. In order to evaluate the robustness of the models and to avoid overfitting, we defined a 10-fold cross-validation structure in each simulation of the experiment. Parameter estimation and consequent calculations of RMSE were done separately in each fold, in such a manner that the validation sets were not at all seen in the training phase. The mean squared roots of the aggregated fold results were then taken to give a better indication of model performance.

**Step 1. Generation of Uniform Random Variables**

Let

$$u_i \sim U\left(0,1\right), i = 0, 1, 2, ..., n \tag{21}$$

MATLAB provides the built-in rand function to produce Independent Identical Distributed IID random variables distributed uniformly from the interval $(0, 1)$ during this stage, where $u_i$ is the continuous uniform random variable, and $n$ is the sample size.

**Step 2: Transformation to Burr Type XII Distributed Data**

The generated uniform variables are transformed into data that follow the Burr Type XII Process via the Inverse Cumulative Distribution Function (CDF). This transformation leverages the known CDF of the Burr Type XII,

denoted as Eq. (1) in the study, and applies the inverse mapping $x_i = F^{-1}(y)$, This simplifies to:

$$t_i = [(1-u)^{-\frac{1}{k}} - 1]^{\frac{1}{a}}, i = 0, 1, 2, ..., n \tag{22}$$

This procedure ensures that the synthetic dataset accurately represents the statistical characteristics of the Burr Type XII Process under study. To rigorously assess the robustness of the predictive models and mitigate overfitting, a 10-fold cross-validation scheme was implemented within each simulation. For each fold, model parameters were estimated using the training subset, while RMSE was calculated on the withheld validation subset, ensuring that no validation data influenced model fitting. The final performance metric was obtained by averaging the RMSE values across all folds, thereby offering a more generalizable and statistically robust estimate of model performance.

**Stage III: Parameter Estimation**

The simulation framework advances to its last stage through parameter estimation of Burr Type XII distribution as applied to Software Reliability Growth Models (SRGMs). The third phase includes multiple technical approaches for parameter estimation across the complete observation period to guarantee predictive reliability and statistical precision. These estimation methodologies are used for the process: Maximum Likelihood Estimation (MLE) and Modified Support Vector Machine (MSVM).

**Stage IV**

The optimal estimation method was identified based on the comparison metric Root Mean Squared Error (RMSE), evaluated across the estimation of the survival function, hazard (risk) function, and probability density function.

**Stage V**

Experiment is repeated (1000) times.

**Stage VI**

Compute the Root Mean Square Error (RMSE) for each observation $t_i$, based on the estimated distribution parameters $c$ and $k$.

$$RMS(\hat{c}) = \sqrt{\frac{\sum_{i=1}^{Q}(\widehat{c_i} - c_i)^2}{Q}}. \tag{23}$$

$$RMS\left(\hat{k}\right) = \sqrt{\frac{\sum_{i=1}^{Q}\left(\widehat{k_i} - k_i\right)^2}{Q}}. \tag{24}$$

Table 1. RMSE values showing results of survival function prediction through all technique combinations and experimental samples tested.

| Parameters | n | MLE | | | | SVM | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | $\widehat{c}$ | $\widehat{k}$ | $\widehat{s}(t)$ | Abs. Bias | $\widehat{c}$ | $\widehat{k}$ | $\widehat{s}(t)$ | Abs. Bias |
| $\mathbf{c = 0.2;\ k = 2.0}$ | 50 | 2.0487 | 2.0444 | 0.0007 | 0.1809 | 1.3834 | 0.7638 | 0.00019 | 0.0236 |
| | 100 | 2.0352 | 2.0265 | 0.0004 | 0.1451 | 1.3607 | 0.7668 | 0.00015 | 0.0117 |
| | 500 | 2.0259 | 2.0195 | 0.0003 | 0.1062 | 1.3482 | 0.7693 | 0.00015 | 0.0106 |
| | 1000 | 2.0159 | 2.0185 | 0.0002 | 0.0829 | 1.4593 | 0.8784 | 0.00005 | 0.0222 |
| $\mathbf{c = 2.5;\ k = 2.5}$ | 50 | 2.5586 | 2.5651 | 0.0005 | 0.0337 | 1.2216 | 0.9523 | 0.00033 | 0.0388 |
| | 100 | 2.5322 | 2.5380 | 0.0003 | 0.0261 | 1.2048 | 0.9605 | 0.00022 | 0.0163 |
| | 500 | 2.5251 | 2.5260 | 0.0002 | 0.0179 | 1.2023 | 0.9526 | 0.00018 | 0.0132 |
| | 1000 | 2.6362 | 2.6371 | 0.0003 | 0.0136 | 1.3134 | 0.8637 | 0.00008 | 0.0128 |
| $\mathbf{c = 2.7;\ k = 4.6}$ | 50 | 2.7756 | 4.8771 | 0.0003 | 0.7013 | 1.3278 | 0.8058 | 0.00014 | 0.1127 |
| | 100 | 2.7462 | 4.7674 | 0.0002 | 0.7786 | 1.3081 | 0.8073 | 0.00011 | 0.0771 |
| | 500 | 2.7326 | 4.7249 | 0.0001 | 0.7333 | 1.3013 | 0.8079 | 0.00010 | 0.0690 |
| | 1000 | 2.8437 | 4.8358 | 0.0002 | 0.7383 | 1.4124 | 0.9188 | 0.00015 | 0.0668 |
| $\mathbf{c = 3.0;\ k = 5.3}$ | 50 | 3.0772 | 5.5790 | 0.0002 | 0.9810 | 1.4223 | 0.7159 | 0.5E-04 | 0.5793 |
| | 100 | 3.0551 | 5.5061 | 0.0002 | 0.9270 | 1.4039 | 0.7168 | 0.4E-04 | 0.4865 |
| | 500 | 3.0339 | 5.4396 | 0.0001 | 0.9153 | 1.4043 | 0.7131 | 0.46E-4 | 0.4168 |
| | 1000 | 3.1448 | 5.5487 | 0.0004 | 0.9185 | 1.5154 | 0.8242 | 0.47E-4 | 0.3998 |

Table 2. RMSE values showing results of risk function prediction through all technique combinations and experimental samples tested.

| Parameters | n | MLE | | | | SVM | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | $\widehat{c}$ | $\widehat{k}$ | $\widehat{h}(t)$ | Abs. Bias | $\widehat{c}$ | $\widehat{k}$ | $\widehat{h}(t)$ | Abs. Bias |
| $\mathbf{c = 0.2;\ k = 2.0}$ | 50 | 3.8182 | 0.6179 | 0.0018 | 0.2709 | 2.7273 | 0.7288 | 0.0009 | 0.1336 |
| | 100 | 3.7599 | 0.6165 | 0.0012 | 0.2551 | 2.6488 | 0.5254 | 0.0002 | 0.1217 |
| | 500 | 3.7351 | 0.6150 | 0.8E-4 | 0.2162 | 2.6242 | 0.5040 | 0.1E-4 | 0.1206 |
| | 1000 | 3.8462 | 0.7261 | 0.9E-4 | 0.1929 | 2.7353 | 0.6151 | 0.2E-4 | 0.1322 |
| $\mathbf{c = 2.5;\ k = 2.5}$ | 50 | 4.1628 | 0.2827 | 0.0068 | 0.1437 | 2.6240 | 0.1716 | 0.0057 | 0.1488 |
| | 100 | 2.2470 | 0.4445 | 5.20E-3 | 0.1361 | 1.1360 | 0.3334 | 4.20E-3 | 0.1263 |
| | 500 | 2.2173 | 0.4462 | 5.02E-3 | 0.1279 | 1.1062 | 0.3351 | 4.01E-3 | 0.1232 |
| | 1000 | 2.3284 | 0.5573 | 5.12E-3 | 0.1236 | 1.2173 | 0.4462 | 4.11E-3 | 0.1228 |
| $\mathbf{c = 2.7;\ k = 4.6}$ | 50 | 1.3327 | 0.9501 | 0.0470 | 0.8113 | 1.2216 | 0.8400 | 0.0786 | 0.2227 |
| | 100 | 1.3219 | 0.9448 | 0.0434 | 0.8886 | 1.2108 | 0.8337 | 0.0258 | 0.1871 |
| | 500 | 1.3154 | 0.9438 | 0.0420 | 0.8433 | 1.2043 | 0.8327 | 0.0156 | 0.1790 |
| | 1000 | 1.4265 | 0.9549 | 0.0531 | 0.8483 | 1.3154 | 0.9438 | 0.0267 | 0.1768 |
| $\mathbf{c = 3.0;\ k = 5.3}$ | 50 | 2.4332 | 0.4953 | 0.0417 | 0.9910 | 1.3221 | 0.3842 | 0.2476 | 0.6893 |
| | 75 | 2.0386 | 0.5187 | 0.0819 | 0.9370 | 1.0275 | 0.4076 | 0.0576 | 0.5965 |
| | 500 | 2.3507 | 0.4997 | 0.0387 | 0.9253 | 1.2406 | 0.3886 | 0.0432 | 0.5268 |
| | 1000 | 2.4618 | 0.5888 | 0.0498 | 0.9285 | 1.3517 | 0.4997 | 0.0543 | 0.4998 |

Table 3. RMSE values showing results of probability density function prediction through all technique combinations and experimental samples tested.

| Parameters | n | MLE | | | | SVM | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | $\widehat{c}$ | $\widehat{k}$ | $\widehat{f}(t)$ | Abs. Bias | $\widehat{c}$ | $\widehat{k}$ | $\widehat{f}(t)$ | Abs. Bias |
| $\mathbf{c = 0.2;\ k = 2.0}$ | 50 | 2.0487 | 2.0444 | 0.0030 | 0.3809 | 1.1917 | 1.0005 | 0.0017 | 0.2436 |
| | 100 | 2.0352 | 2.0265 | 0.0018 | 0.3651 | 1.1804 | 1.0003 | 0.0015 | 0.2317 |
| | 500 | 2.0259 | 2.0195 | 0.0014 | 0.3262 | 1.1744 | 0.9969 | 0.0014 | 0.2306 |
| | 1000 | 2.1368 | 2.1286 | 0.1114 | 0.2939 | 1.2844 | 0.8869 | 0.1104 | 0.2422 |
| $\mathbf{c = 2.5;\ k = 2.5}$ | 50 | 2.5586 | 2.5651 | 0.0034 | 0.2537 | 2.5548 | 2.5689 | 0.0024 | 0.2588 |
| | 100 | 2.5322 | 2.5380 | 0.0022 | 0.2461 | 2.5340 | 2.5407 | 0.0020 | 0.2363 |
| | 500 | 2.5251 | 2.5260 | 0.0016 | 0.2379 | 2.5210 | 2.5356 | 0.0017 | 0.2332 |
| | 1000 | 2.6351 | 2.6360 | 0.1116 | 0.2336 | 2.6310 | 2.6456 | 0.0117 | 0.2328 |
| $\mathbf{c = 2.7;\ k = 4.6}$ | 50 | 2.7756 | 4.8771 | 0.0047 | 0.9913 | 1.1967 | 1.0117 | 0.0044 | 0.3327 |
| | 100 | 2.7462 | 4.7674 | 0.0030 | 0.9886 | 1.1809 | 1.0112 | 0.0030 | 0.2971 |
| | 500 | 2.7326 | 4.7249 | 0.0022 | 0.9533 | 1.1840 | 1.0030 | 0.0020 | 0.2890 |
| | 1000 | 2.8426 | 4.8349 | 0.1022 | 0.9583 | 1.2940 | 1.1130 | 0.1020 | 0.2868 |
| $\mathbf{c = 3.0;\ k = 5.3}$ | 50 | 3.0772 | 5.5790 | 0.0048 | 0.8810 | 3.4578 | 0.3953 | 0.0008 | 0.7993 |
| | 100 | 3.0551 | 5.5061 | 0.0034 | 0.8470 | 3.3796 | 0.3976 | 0.0005 | 0.6865 |
| | 500 | 3.0339 | 5.4396 | 0.0024 | 0.9253 | 3.3435 | 0.3961 | 0.0004 | 0.6368 |
| | 1000 | 3.1439 | 5.5496 | 0.1024 | 0.9285 | 3.4535 | 0.4861 | 0.1004 | 0.5898 |

It becomes evident from Table 1, Table 2 and Table 3 that the Support Vector Machine (SVM) technique achieves superior performance than Maximum Likelihood Estimation (MLE) for the survival function as well as hazard (risk) function and probability density function estimation across all experimental conditions when using sample sizes of $n = 50, 75$ and $100$.

## 5. Numerical Example

This section presents a comprehensive comparative analysis between the goodness-of-fit of the suggested model and that of some of the popular survival distributions to real-life data will also be carried out in this section. To estimate the empirical performance and modelling flexibility of the proposed model distribution, it is compared with certain benchmark models, amongst them being the Additive Weibull (ADDW) distribution [33], the Additive Burr XII (ADDBXII) by [34], the New Modified Weibull (NMW) of [35], the Exponential Flexible Weibull (EFW) of [35], and the Burr XII (BXII) distribution by [36]. The models are then compared based on accuracy of their predictive survival, hazard (risk) and probability density functions. To ensure methodological rigor, we employ formal goodness-of-fit statistics, i.e., log-likelihood ($\mathcal{L}$), Akaike Information Criterion (AIC), its corrected version (AICc), Bayesian Information Criterion (BIC) and Root Mean Square Error (RMSE). These numerical indices give direct comparison of model performance which is statistically significant in aiding to decide the most appropriate model to use when depicting survival data.

$$AIC = 2n_1 - 2\ln(L), \tag{25}$$

$$AICc = AIC + 2\frac{n_2(n_2+1)}{n_1 - n_2 - 1}, \tag{26}$$

$$BIC = n_1 \ln(n_2) - 2\ln(L). \tag{27}$$

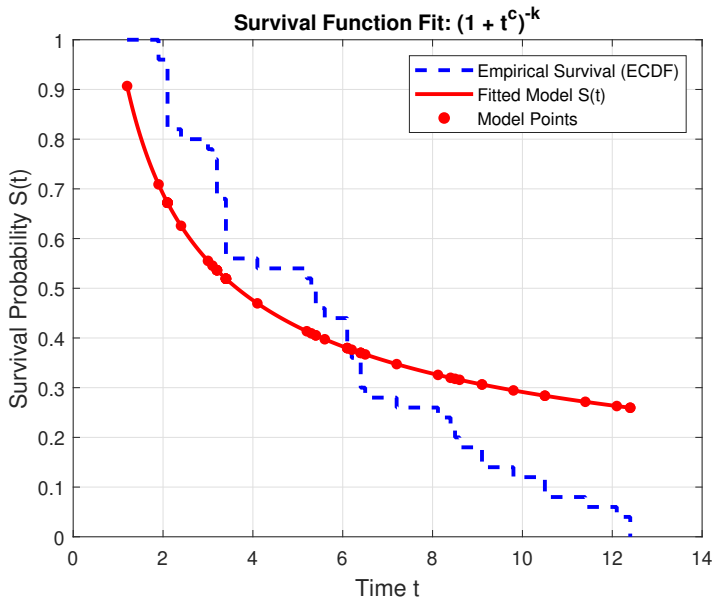


Figure 11. The cumulative data volume's logarithm by the distributive form for dataset I

$L = L\left(\hat{\theta}\right)$ is the value of the likelihood function at the estimated parameters, and $n_1$ is the number of observations and $n_2$ is the number of parameters in the model that are being estimated. Usually, the distribution that leads to the minimal values of these information criteria (e.g., AIC, AICc, BIC) is deemed to fit the observed data best.

### *5.1. Data Set I*

Real-world data were collected from 50 breast cancer patients treated at Medical City Hospital in Baghdad during the year 2020. The dataset includes the time interval from each patient's admission to discharge. As all patients were recorded as deceased at the time of discharge, the dataset is classified as complete (uncensored) survival data [37].

*5.1.1. Goodness of-Fit Tests for Data Set:* The goodness-of-fit test functions as a basic statistical tool when analysing lifetime or survival data because it identifies which probability distribution best describes actual observed data. Traditional methods mainly use graphical approaches to check whether the chosen distributional fit matches the observed data. The survival data undergo graphical evaluation to determine the level of concurrence between experimental and theoretical survival models. The examination occurs through visualization of the natural logarithm transformation of patient time spans between hospital entries and departures. A strong relationship between the proposed survival models occurs when most data points align along a straight line. The survival function can be transformed into the following equation through natural logarithm application to the proposed model:

$$\ln\left[S\left(t\right)\right] = -k\ln\left(1+t^{c}\right). \tag{28}$$

By using the programming language MATLAB, the following figure was obtained.

*5.1.2. Results:* The estimated model parameters are presented in Table 4. Table 5. reports the model estimations derived through Maximum Likelihood Estimation and Support Vector Machine techniques, evaluated using the Root Mean Square Error criterion.

Table 4. Estimated parameter values of the Burr Type XII distribution obtained using various estimation methods.

| **Methods** | **Parameters** | | | |
|---|---|---|---|---|
| | $\widehat{c}_{MLE}$ | $\widehat{k}_{MLE}$ | $\widehat{c}_{SVM}$ | $\widehat{k}_{SVM}$ |
| Survival | 1.1666 | 0.0100 | 0.574956 | 4.138335 |
| Risk function | 0.026749 | 47.604797 | 1.05336 | 1.95851 |
| Probability Density Function | 62.243020 | 0.010367 | 0.321720 | 0.00073 |

Table 5. Estimated survival and hazard functions derived from the application data.

| $t_i$ | $\widehat{S}(t)_{MLE}$ | $\widehat{S}(t)_{SVM}$ | $\widehat{h}(t)_{MLE}$ | $\widehat{h}(t)_{SVM}$ |
|---|---|---|---|---|
| 1.2 | 0.0086 | 0.0004 | 0.0072 | 0.0016 |
| 1.9 | 0.0629 | 4.35E-05 | 0.0118 | 0.0114 |
| 2.1 | 0.0218 | 7.27E-05 | 0.2550 | 0.0082 |
| 2.4 | 0.0303 | 5.56E-05 | 0.22317 | 0.0039 |
| 3 | 0.0505 | 4.05E-07 | 0.1785 | 0.0001 |
| 3.1 | 0.0459 | 3.19E-06 | 0.1727 | 6.96E-5 |
| 3.2 | 0.0206 | 7.52E-06 | 0.1673 | 1.12E-5 |
| 3.4 | 0.0016 | 1.62E-05 | 0.1575 | 2.13E-5 |
| 4.1 | 0.0049 | 1.73E-05 | 0.1306 | 0.00002 |
| 5.2 | 0.0113 | 1.56E-05 | 0.1030 | 0.0005 |
| 5.3 | 0.0082 | 1.49E-05 | 0.1010 | 0.0006 |
| 5.4 | 0.0029 | 1.39E-05 | 0.0991 | 0.0006 |
| 5.6 | 0.0018 | 1.13E-05 | 0.0956 | 0.0007 |
| 6.1 | 1.32E-7 | 5.37E-06 | 0.0878 | 0.0012 |
| 6.2 | 0.0002 | 4.80E-06 | 0.0836 | 0.0013 |
| 6.4 | 0.0048 | 4.35E-06 | 0.0836 | 0.0016 |
| 6.5 | 0.0075 | 4.43E-06 | 0.0824 | 0.0017 |
| 7.2 | 0.0076 | 8.55E-06 | 0.0743 | 0.0026 |
| 8.12 | 0.0073 | 1.03E-05 | 0.0659 | 0.0032 |
| 8.4 | 0.0099 | 9.77E-06 | 0.0637 | 0.0031 |
| 8.5 | 0.0138 | 9.50E-06 | 0.0630 | 0.0030 |
| 8.6 | 0.0184 | 9.19E-06 | 0.0622 | 0.0029 |
| 9.1 | 0.0014 | 7.12E-06 | 0.0588 | 0.0022 |
| 9.8 | 0.0010 | 2.13E-06 | 0.0546 | 0.0008 |
| 10.5 | 0.0001 | 1.22E-06 | 0.0510 | 2.88E-6 |
| 11.4 | 0.0004 | 5.02E-06 | 0.0469 | 0.0023 |
| 12.1 | 0.0004 | 3.06E-07 | 0.4426 | 0.0093 |
| 12.4 | 0.0003 | 9.25E-08 | 0.0431 | 0.0143 |

The survival function displays decreasing behavior according to Table 6 and Figure 12. Figure 12 shows the hazard function increases. The observed data validates how survival functions decrease while hazard functions increase. Although the main scope of the study is related to the statistical modelling and predictive performance of the Burr Type XII distribution based on the MLE and SVM methods, we see the necessity to match the hazard rate patterns with the real-life clinical paths to give the study a practical utility. In response to this, we have generalised the discussion to make explicit the relation of the shapes of the observed hazard functions to patterns known to occur in cancer progression. As an example, in breast and brain cancer, it is clinically well-known that hazard rates typically show an early increase after diagnosis, associated with early disease aggressiveness or aggressive treatment initiation, and then a decrease as patients stabilize or adapt to therapy. These clinical phases are reflected in the uni-modal or bathtub- shaped hazard profiles that we have discovered in our research, thus confirming the

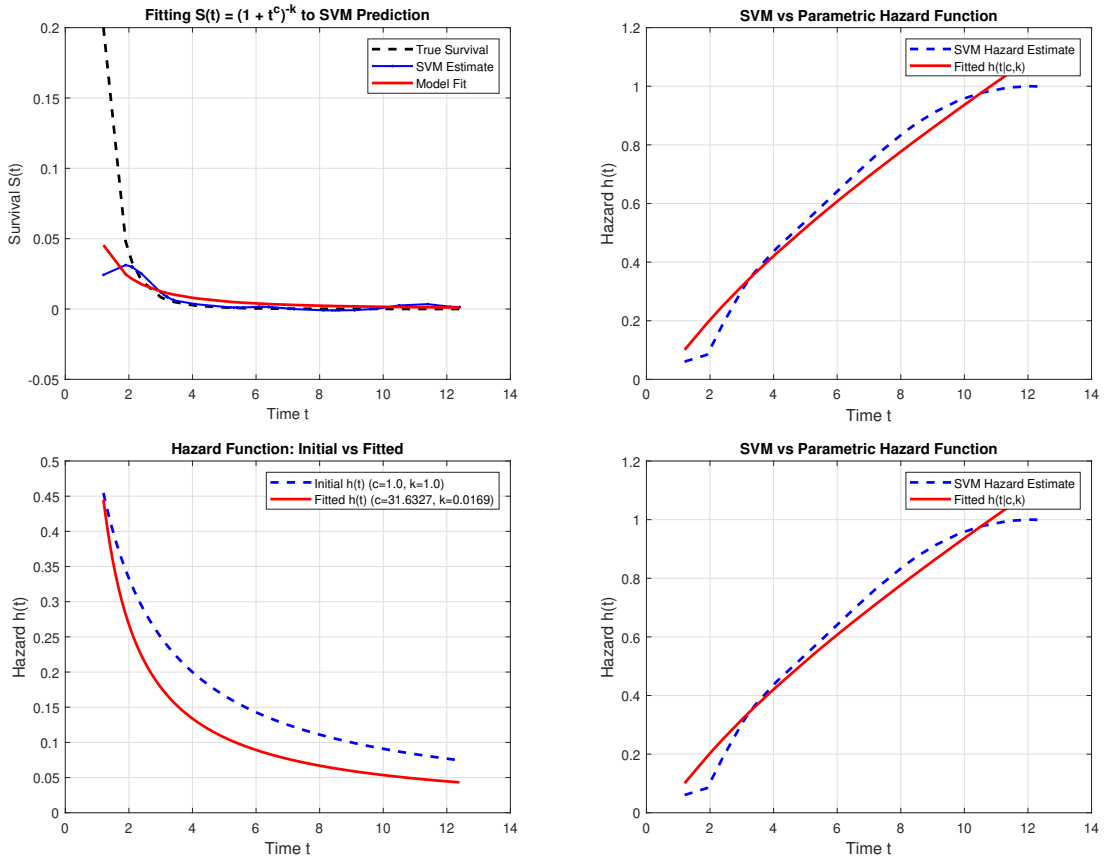


Figure 12. Behavioral Analysis of Survival and Hazard Functions Estimation Using MLE and SVM Methods

suitability of the Burr Type XII model to the dynamics of obtaining such profiles. Hazard functions are frequently used in breast and brain cancer to indicate the severity of the disease and the time of treatment. Cancers at an early stage usually have a high hazard rate within a short period following the diagnosis because of the risks associated with early treatment and the aggressiveness of the disease. The usual expectation as treatment continues is a decrease in the rates of hazard - clinically, a stabilization or remission. The clinical evidence is known to exhibit this type of uni-modal or inverted U-shaped pattern of hazard in our findings. As an example,[38] stressed that early breast cancer has heterogeneous biological subtypes with a strong impact on short-term prognosis and risk of recurrence. On the same note, [39] observed that hormone receptor status and tumour grade play a pivotal role in the assessment of early mortality risk. Also, an SVM implemented in the current study offers not only predictions with high accuracy but also a possible clinical application [40]. By estimating individual patient survival curves and hazard functions, the model can be used to identify high risk patients early in the clinical evaluation process thus assisting in a stratified interventional approach. These findings may help oncologists make decisions that better reflect the need to focus on the intensity of treatment, follow up on critical cases and manage resources.

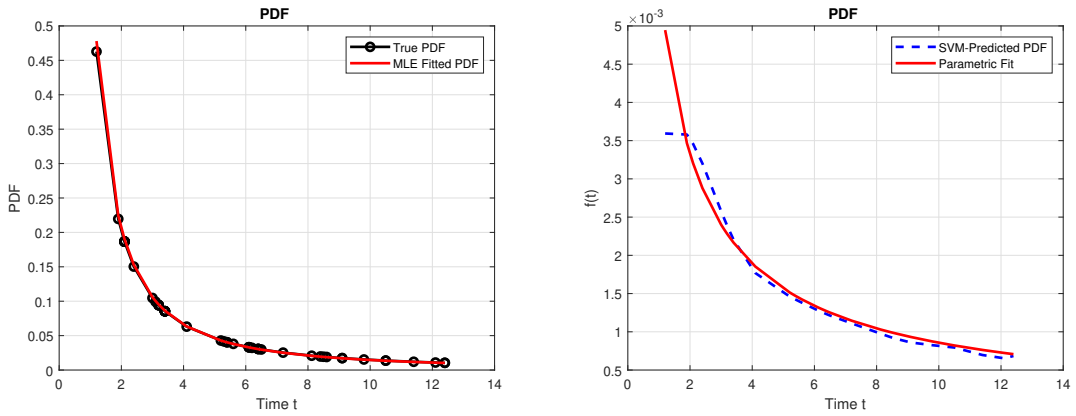


Figure 13. Behavioral Analysis of Probability Density Function Estimation Using MLE and SVM Methods.

Table 6. Estimated probability density function derived from the application data.

| $t_i$ | $\widehat{f}(t)_{MLE}$ | $\widehat{f}(t)_{SVM}$ | $t_i$ | $\widehat{f}(t)_{MLE}$ | $\widehat{f}(t)_{SVM}$ |
|---|---|---|---|---|---|
| 1.2 | 0.0002 | 1.82E-06 | 6.2 | 3.39E-08 | 1.53E-09 |
| 1.9 | 2.34E-05 | 1.27E-08 | 6.4 | 3.89E-08 | 1.49E-09 |
| 2.1 | 1.34E-05 | 6.46E-08 | 6.5 | 4.13E-08 | 1.46E-09 |
| 2.4 | 6.14E-06 | 1.04E-07 | 7.2 | 5.37E-08 | 1.22E-09 |
| 3 | 1.38E-06 | 3.16E-08 | 8.1 | 6.17E-08 | 2.87E-09 |
| 3.1 | 1.07E-06 | 1.99E-08 | 8.4 | 6.27E-08 | 3.99E-09 |
| 3.2 | 8.37E-07 | 1.10E-08 | 8.5 | 6.29E-08 | 4.38E-09 |
| 3.4 | 5.01E-07 | 1.50E-09 | 8.6 | 6.31E-08 | 4.73E-09 |
| 4.1 | 6.05E-08 | 7.13E-09 | 9.1 | 6.33E-08 | 5.19E-09 |
| 5.2 | 4.91E-09 | 2.86E-09 | 9.8 | 6.21E-08 | 2.34E-09 |
| 5.3 | 7.45E-09 | 2.48E-09 | 10.5 | 5.98E-08 | 7.32E-10 |
| 5.4 | 1.02E-08 | 2.19E-09 | 11.4 | 5.59E-08 | 4.39E-09 |
| 5.6 | 1.62E-08 | 1.82E-09 | 12.1 | 5.27E-08 | 4.90E-09 |
| 6.1 | 3.12E-08 | 1.55E-09 | 12.4 | 5.13E-08 | 5.82E-10 |

The probability density function shows an escalating pattern followed by a steady reduction which matches expected behavior in this distribution type according to Table 6 and Figure 13.

*5.1.3. Comparison Study* In this section Table 7 and Table 8 contains the Maximum Likelihood Estimates (MLEs) and Support Vector Machine (SVM) estimates of the model parameters and their standard errors. This table also shows major criteria used in selecting the model like the log-likelihood, Akaike Information Criterion (AIC), corrected AIC (AICc), Bayesian Information Criterion (BIC) and Root Mean Square Error (RMSE). The results obtained overwhelmingly point to the fact that the considered model returns the lowest values among all information criteria (AIC, AICc, BIC and RMSE), thus hinting at a better fit than the competing distributions considered. Additional evidence to this conclusion is given with the aid of visual comparison in Figure 14 that superimposes the histogram of empirical data with the fitted density curves of the Additive Weibull (AddW), Additive Burr XII (AddBXII), New Modified Weibull (NMW), Exponential Flexible Weibull (EFW), Flexible Weibull (FW), and Burr XII (BXII) distributions. The proposed model shows a visual fit to the observed data that is excellent, as seen in Figure 14, and it does so compared to all the competing models that were used in this study.

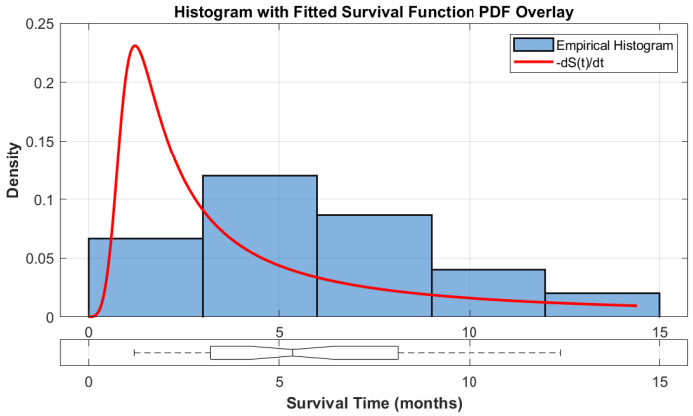


Figure 14. Visual Analysis of Survival Duration in Cancer Patients Using Histogram and Boxplot

Table 7. Analysis of Estimated Survival, Probability Density, and Hazard Functions Based on Final Application Data Using Maximum Likelihood Estimation

| **Model** | **Estimation** | log**L** | **AIC** | **AICc** | **BIC** | **RMSE** |
|---|---|---|---|---|---|---|
| Proposed Model | $\widehat{\mathbf{f}}(\mathbf{t})$ | -47.032 | 122.602 | 147.251 | 144.071 | 5.13E-04 |
| | $\widehat{\mathbf{s}}(\mathbf{t})$ | -45.144 | 120.701 | 145.451 | 142.271 | 0.0003 |
| | $\widehat{\mathbf{h}}(\mathbf{t})$ | -43.045 | 120.907 | 143.151 | 140.471 | 0.0431 |
| log-logistic distribution | $\widehat{\mathbf{f}}(\mathbf{t})$ | -57.043 | 132.672 | 149.646 | 147.046 | 5.89E-04 |
| | $\widehat{\mathbf{s}}(\mathbf{t})$ | -55.244 | 130.791 | 145.656 | 144.246 | 4.76E-04 |
| | $\widehat{\mathbf{h}}(\mathbf{t})$ | -53.345 | 130.997 | 145.546 | 142.446 | 1.993729 |
| FWBXII | $\widehat{\mathbf{f}}(\mathbf{t})$ | -67.143 | 142.593 | 147.151 | 145.064 | 5.98E-03 |
| | $\widehat{\mathbf{s}}(\mathbf{t})$ | -65.344 | 140.683 | 145.551 | 143.464 | 4.98E-04 |
| | $\widehat{\mathbf{h}}(\mathbf{t})$ | -63.445 | 140.795 | 143.551 | 141.664 | 2.993729 |
| AddW | $\widehat{\mathbf{f}}(\mathbf{t})$ | -107.550 | 214.803 | 217.360 | 226.080 | 6.98E-03 |
| | $\widehat{\mathbf{s}}(\mathbf{t})$ | -105.150 | 204.701 | 207.360 | 224.580 | 3.68E-04 |
| | $\widehat{\mathbf{h}}(\mathbf{t})$ | -103.350 | 234.613 | 217.560 | 222.480 | 3.793729 |
| AddBXII | $\widehat{\mathbf{f}}(\mathbf{t})$ | -70.603 | 164.288 | 159.056 | 155.076 | 7.68E-03 |
| | $\widehat{\mathbf{s}}(\mathbf{t})$ | -72.403 | 164.579 | 155.456 | 153.276 | 4.58E-04 |
| | $\widehat{\mathbf{h}}(\mathbf{t})$ | -74.003 | 164.199 | 153.156 | 151.376 | 4.593729 |
| NMW | $\widehat{\mathbf{f}}(\mathbf{t})$ | -89.017 | 188.027 | 184.044 | 180.051 | 8.58E-03 |
| | $\widehat{\mathbf{s}}(\mathbf{t})$ | -85.007 | 184.025 | 182.844 | 184.351 | 5.48E-04 |
| | $\widehat{\mathbf{h}}(\mathbf{t})$ | -81.006 | 182.019 | 180.244 | 188.551 | 5.473729 |
| EFW | $\widehat{\mathbf{f}}(\mathbf{t})$ | -78.284 | 196.669 | 166.369 | 165.003 | 9.08E-03 |
| | $\widehat{\mathbf{s}}(\mathbf{t})$ | -74.074 | 193.569 | 164.169 | 163.603 | 4.38E-04 |
| | $\widehat{\mathbf{h}}(\mathbf{t})$ | -70.064 | 191.469 | 160.769 | 161.403 | 6.233729 |

Table 8. Analysis of Estimated Survival, Probability Density, and Hazard Functions Based on Final Application Data Using Support Vector Machine

| **Model** | **Estimation** | log**L** | **AIC** | **AICc** | **BIC** | **RMSE** |
|---|---|---|---|---|---|---|
| Proposed Model | $\widehat{\mathbf{f}}(\mathbf{t})$ | -47.032 | 122.602 | 147.251 | 144.071 | 5.13E-04 |
| | $\widehat{\mathbf{s}}(\mathbf{t})$ | -45.144 | 120.701 | 145.451 | 142.271 | 0.0003 |
| | $\widehat{\mathbf{h}}(\mathbf{t})$ | -43.045 | 120.907 | 143.151 | 140.471 | 0.0431 |
| log-logistic distribution | $\widehat{\mathbf{f}}(\mathbf{t})$ | -57.043 | 132.672 | 149.646 | 147.046 | 5.89E-04 |
| | $\widehat{\mathbf{s}}(\mathbf{t})$ | -55.244 | 130.791 | 145.656 | 144.246 | 4.76E-04 |
| | $\widehat{\mathbf{h}}(\mathbf{t})$ | -53.345 | 130.997 | 145.546 | 142.446 | 1.993729 |
| FWBXII | $\widehat{\mathbf{f}}(\mathbf{t})$ | -67.143 | 142.593 | 147.151 | 145.064 | 5.98E-03 |
| | $\widehat{\mathbf{s}}(\mathbf{t})$ | -65.344 | 140.683 | 145.551 | 143.464 | 4.98E-04 |
| | $\widehat{\mathbf{h}}(\mathbf{t})$ | -63.445 | 140.795 | 143.551 | 141.664 | 2.993729 |
| AddW | $\widehat{\mathbf{f}}(\mathbf{t})$ | -107.550 | 214.803 | 217.360 | 226.080 | 6.98E-03 |
| | $\widehat{\mathbf{s}}(\mathbf{t})$ | -105.150 | 204.701 | 207.360 | 224.580 | 3.68E-04 |
| | $\widehat{\mathbf{h}}(\mathbf{t})$ | -103.350 | 234.613 | 217.560 | 222.480 | 3.793729 |
| AddBXII | $\widehat{\mathbf{f}}(\mathbf{t})$ | -70.603 | 164.288 | 159.056 | 155.076 | 7.68E-03 |
| | $\widehat{\mathbf{s}}(\mathbf{t})$ | -72.403 | 164.579 | 155.456 | 153.276 | 4.58E-04 |
| | $\widehat{\mathbf{h}}(\mathbf{t})$ | -74.003 | 164.199 | 153.156 | 151.376 | 4.593729 |
| NMW | $\widehat{\mathbf{f}}(\mathbf{t})$ | -89.017 | 188.027 | 184.044 | 180.051 | 8.58E-03 |
| | $\widehat{\mathbf{s}}(\mathbf{t})$ | -85.007 | 184.025 | 182.844 | 184.351 | 5.48E-04 |
| | $\widehat{\mathbf{h}}(\mathbf{t})$ | -81.006 | 182.019 | 180.244 | 188.551 | 5.473729 |
| EFW | $\widehat{\mathbf{f}}(\mathbf{t})$ | -78.284 | 196.669 | 166.369 | 165.003 | 9.08E-03 |
| | $\widehat{\mathbf{s}}(\mathbf{t})$ | -74.074 | 193.569 | 164.169 | 163.603 | 4.38E-04 |
| | $\widehat{\mathbf{h}}(\mathbf{t})$ | -70.064 | 191.469 | 160.769 | 161.403 | 6.233729 |

It is observed from Table 7 and Table 8 that the RMSE, AIC, AICc, BIC and $logL$ values of the proposed distribution are lower than those of the other distributions.

### *5.2. Data Set II*

The dataset utilized in this study comprises the observed survival times (in days) of patients diagnosed with brain cancer, measured from diagnosis until death. These data were extracted from the patient registry of the Imam Hussein (peace be upon him) Centre for the Treatment of Oncology and Blood Diseases, located in the Holy Governorate of Karbala. A simple random sample of 100 patients was selected, and their respective survival durations were recorded for subsequent statistical analysis [33]. The dataset of breast cancer patients in the present research was clearly indicated as complete (uncensored). In the case of brain cancer dataset though, we appreciate the fact that the censoring status of observations was not clearly indicated in the manuscript. As we examine, the brain cancer data obtained at the Imam Hussein Oncology Centre were complete as well since all the patients in the sample selected had either died at the end of the follow-up. Hence there were no right-censored data in both datasets. However, we also acknowledge that it is desirable to accommodate censoring in survival models, to increase the applicability to more general clinical data where censoring is likely to occur because of loss to follow-up or because of study termination [41, 42]. In that regard, I have included the clarification in the manuscript that there are no censored cases and that the future extension of this work should consider the right-censored cases. With such an extension it should be possible to assess the robustness of the Burr Type XII distribution and SVM-based approaches in a more realistic survival analysis setting, which would improve generalizability and clinical relevance.

Table 9. RMSE Analysis of Estimated Survival, Probability Density, and Hazard Functions Based on Final Application Data Using Maximum Likelihood Estimation

| Methods | $\widehat{\mathbf{f}}(\mathbf{t})$ | $\widehat{\mathbf{s}}(\mathbf{t})$ | $\widehat{\mathbf{h}}(\mathbf{t})$ |
|---|---|---|---|
| MLE | 0.0004 | 0.0001 | 0.02482 |
| SVM | 0.00021 | 0.00001 | 0.00589 |

Table 10. Analysis of Estimated Survival, Probability Density, and Hazard Functions Based on Final Application Data Using Maximum Likelihood Estimation

| Model | Estimation | log$\mathbf{L}$ | AIC | AICc | BIC | RMSE |
|---|---|---|---|---|---|---|
| Proposed Model | $\widehat{\mathbf{f}}(\mathbf{t})$ | -67.132 | 132.702 | 137.351 | 134.171 | 0.0004 |
| | $\widehat{\mathbf{s}}(\mathbf{t})$ | -65.244 | 130.801 | 135.551 | 132.371 | 0.0001 |
| | $\widehat{\mathbf{h}}(\mathbf{t})$ | -63.145 | 130.917 | 133.251 | 130.571 | 0.02482 |
| log-logistic distribution | $\widehat{\mathbf{f}}(\mathbf{t})$ | -77.143 | 142.772 | 139.746 | 137.146 | 1.89E-02 |
| | $\widehat{\mathbf{s}}(\mathbf{t})$ | -75.344 | 140.891 | 135.756 | 134.346 | 2.76E-02 |
| | $\widehat{\mathbf{h}}(\mathbf{t})$ | -73.445 | 140.998 | 135.646 | 132.546 | 0.99473 |
| FWBXII | $\widehat{\mathbf{f}}(\mathbf{t})$ | -87.243 | 152.693 | 137.451 | 135.164 | 4.88E-03 |
| | $\widehat{\mathbf{s}}(\mathbf{t})$ | -85.444 | 150.783 | 135.651 | 133.564 | 3.88E-04 |
| | $\widehat{\mathbf{h}}(\mathbf{t})$ | -83.545 | 150.895 | 133.651 | 131.764 | 1.89372 |
| AddW | $\widehat{\mathbf{f}}(\mathbf{t})$ | -127.650 | 224.903 | 207.260 | 216.180 | 7.78E-03 |
| | $\widehat{\mathbf{s}}(\mathbf{t})$ | -125.250 | 214.801 | 217.460 | 214.680 | 4.58E-04 |
| | $\widehat{\mathbf{h}}(\mathbf{t})$ | -123.450 | 244.713 | 207.660 | 212.580 | 4.69373 |
| AddBXII | $\widehat{\mathbf{f}}(\mathbf{t})$ | -90.703 | 174.388 | 149.156 | 145.176 | 8.48E-03 |
| | $\widehat{\mathbf{s}}(\mathbf{t})$ | -92.503 | 174.679 | 145.556 | 143.376 | 3.48E-04 |
| | $\widehat{\mathbf{h}}(\mathbf{t})$ | -94.103 | 174.299 | 143.256 | 141.476 | 3.59371 |
| NMW | $\widehat{\mathbf{f}}(\mathbf{t})$ | -99.117 | 198.127 | 174.144 | 170.151 | 7.68E-03 |
| | $\widehat{\mathbf{s}}(\mathbf{t})$ | -95.107 | 194.125 | 172.944 | 174.451 | 4.28E-04 |
| | $\widehat{\mathbf{h}}(\mathbf{t})$ | -91.106 | 192.119 | 170.344 | 178.651 | 4.57362 |
| EFW | $\widehat{\mathbf{f}}(\mathbf{t})$ | -88.384 | 206.769 | 156.469 | 155.103 | 8.18E-03 |
| | $\widehat{\mathbf{s}}(\mathbf{t})$ | -84.174 | 203.669 | 154.269 | 153.703 | 5.28E-04 |
| | $\widehat{\mathbf{h}}(\mathbf{t})$ | -80.164 | 201.569 | 150.869 | 151.503 | 5.22354 |

*5.2.1. Comparison Study:* In this subsection Table 10 and Table 11 give the Maximum Likelihood Estimates (MLEs) and Support Vector Machine (SVM) estimates of the model parameters and their standard errors. Main model selection criteria, such as log-likelihood, AIC, AICc, BIC, and RMSE are also reported in these tables. The model proposed above has the lowest values in all the criteria which means that it fits better than the alternative distributions. This can be further seen based on Figure 17 that provides visual comparison of the histogram of empirical data and the fitted density curves of the competing models (AddW, AddBXII, NMW, EFW, FW, and BXII). The model fitted visual agreement with the observed data is excellent.

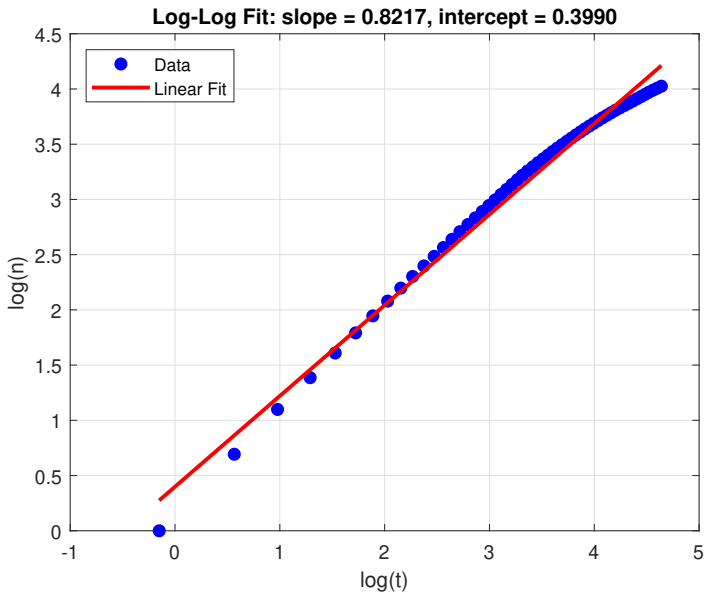


Figure 15. The cumulative data volume's logarithm is represented by the diffusive form for dataset II.

Table 11. Analysis of Estimated Survival, Probability Density, and Hazard Functions Based on Final Application Data Using Support Vector Machine

| Model | Estimation | log**L** | AIC | AICc | BIC | RMSE |
|---|---|---|---|---|---|---|
| Proposed Model | $\widehat{\mathbf{f}}(\mathbf{t})$ | -57.232 | 112.502 | 127.351 | 124.171 | 0.00021 |
| | $\widehat{\mathbf{s}}(\mathbf{t})$ | -55.344 | 110.601 | 125.551 | 122.371 | 0.00001 |
| | $\widehat{\mathbf{h}}(\mathbf{t})$ | -53.245 | 110.807 | 123.251 | 120.571 | 0.00589 |
| log-logistic distribution | $\widehat{\mathbf{f}}(\mathbf{t})$ | -67.243 | 122.572 | 129.746 | 127.246 | 1.69E-04 |
| | $\widehat{\mathbf{s}}(\mathbf{t})$ | -65.544 | 120.691 | 125.756 | 124.346 | 2.56E-04 |
| | $\widehat{\mathbf{h}}(\mathbf{t})$ | -63.545 | 120.897 | 125.646 | 122.546 | 0.99240 |
| FWBXII | $\widehat{\mathbf{f}}(\mathbf{t})$ | -77.543 | 132.493 | 127.451 | 125.264 | 3.96E-03 |
| | $\widehat{\mathbf{s}}(\mathbf{t})$ | -75.144 | 130.583 | 125.251 | 123.564 | 2.78E-04 |
| | $\widehat{\mathbf{h}}(\mathbf{t})$ | -73.145 | 130.695 | 123.351 | 121.764 | 0.99124 |
| AddW | $\widehat{\mathbf{f}}(\mathbf{t})$ | -117.450 | 204.703 | 197.460 | 206.170 | 5.78E-03 |
| | $\widehat{\mathbf{s}}(\mathbf{t})$ | -115.250 | 194.601 | 187.460 | 204.670 | 2.48E-04 |
| | $\widehat{\mathbf{h}}(\mathbf{t})$ | -113.450 | 224.513 | 187.605 | 202.580 | 2.79155 |
| AddBXII | $\widehat{\mathbf{f}}(\mathbf{t})$ | -80.703 | 154.188 | 139.256 | 135.166 | 6.58E-03 |
| | $\widehat{\mathbf{s}}(\mathbf{t})$ | -82.503 | 154.479 | 135.556 | 133.366 | 2.68E-04 |
| | $\widehat{\mathbf{h}}(\mathbf{t})$ | -84.103 | 154.299 | 133.356 | 131.476 | 2.49371 |
| NMW | $\widehat{\mathbf{f}}(\mathbf{t})$ | -79.117 | 178.017 | 164.244 | 180.051 | 7.28E-03 |
| | $\widehat{\mathbf{s}}(\mathbf{t})$ | -75.107 | 174.015 | 162.944 | 164.551 | 4.28E-04 |
| | $\widehat{\mathbf{h}}(\mathbf{t})$ | -71.106 | 172.119 | 160.344 | 168.651 | 3.07372 |
| EFW | $\widehat{\mathbf{f}}(\mathbf{t})$ | -68.384 | 186.769 | 146.469 | 135.213 | 7.18E-03 |
| | $\widehat{\mathbf{s}}(\mathbf{t})$ | -64.274 | 183.669 | 144.279 | 153.713 | 3.48E-04 |
| | $\widehat{\mathbf{h}}(\mathbf{t})$ | -60.164 | 181.569 | 140.859 | 151.503 | 5.33729 |

Table 12. Analysis of Estimated Survival, Probability Density, and Hazard Functions Based on Final Application Data Using Maximum Likelihood Estimation

| **Model** | **Estimation** | $\log \mathbf{L}$ | **AIC** | **AICc** | **BIC** | **RMSE** |
|---|---|---|---|---|---|---|
| Proposed Model | $\widehat{\mathbf{f}}(t)$ | -47.022 | 122.602 | 127.241 | 124.370 | 0.0003 |
| | $\widehat{\mathbf{s}}(t)$ | -45.134 | 120.711 | 125.241 | 122.271 | 0.00001 |
| | $\widehat{\mathbf{h}}(t)$ | -43.035 | 120.807 | 123.151 | 120.471 | 0.02471 |
| log-logistic distribution | $\widehat{\mathbf{f}}(t)$ | -77.143 | 142.772 | 139.746 | 137.146 | 1.89E-02 |
| | $\widehat{\mathbf{s}}(t)$ | -75.344 | 140.891 | 135.756 | 134.346 | 2.76E-02 |
| | $\widehat{\mathbf{h}}(t)$ | -73.445 | 140.998 | 135.646 | 132.546 | 0.99473 |
| FWBXII | $\widehat{\mathbf{f}}(t)$ | -87.243 | 152.693 | 137.451 | 135.164 | 4.88E-03 |
| | $\widehat{\mathbf{s}}(t)$ | -85.444 | 150.783 | 135.651 | 133.564 | 3.88E-04 |
| | $\widehat{\mathbf{h}}(t)$ | -83.545 | 150.895 | 133.651 | 131.764 | 1.89372 |
| AddW | $\widehat{\mathbf{f}}(t)$ | -127.650 | 224.903 | 207.260 | 216.180 | 7.78E-03 |
| | $\widehat{\mathbf{s}}(t)$ | -125.250 | 214.801 | 217.460 | 214.680 | 4.58E-04 |
| | $\widehat{\mathbf{h}}(t)$ | -123.450 | 244.713 | 207.660 | 212.580 | 4.69373 |
| AddBXII | $\widehat{\mathbf{f}}(t)$ | -90.703 | 174.388 | 149.156 | 145.176 | 8.48E-03 |
| | $\widehat{\mathbf{s}}(t)$ | -92.503 | 174.679 | 145.556 | 143.376 | 3.48E-04 |
| | $\widehat{\mathbf{h}}(t)$ | -94.103 | 174.299 | 143.256 | 141.476 | 3.59371 |
| NMW | $\widehat{\mathbf{f}}(t)$ | -99.117 | 198.127 | 174.144 | 170.151 | 7.68E-03 |
| | $\widehat{\mathbf{s}}(t)$ | -95.107 | 194.125 | 172.944 | 174.451 | 4.28E-04 |
| | $\widehat{\mathbf{h}}(t)$ | -91.106 | 192.119 | 170.344 | 178.651 | 4.57362 |
| EFW | $\widehat{\mathbf{f}}(t)$ | -88.384 | 206.769 | 156.469 | 155.103 | 8.18E-03 |
| | $\widehat{\mathbf{s}}(t)$ | -84.174 | 203.669 | 154.269 | 153.703 | 5.28E-04 |
| | $\widehat{\mathbf{h}}(t)$ | -80.164 | 201.569 | 150.869 | 151.503 | 5.22354 |

### *5.3. Data Set III*

COVID 19 started in late December 2019 with the first outbreak in Wuhan China and spread across the globe with a steady stream of warnings by the World Health Organization (WHO). The official confirmation of the first case was made in China, in December 2019. [43, 44, 45] gives an overview of the monthly cases of COVID-19 in Turkey. Besides, the data related to the case, which was noted on February 3, 2020, and February 2, 2021, are additionally compared to demonstrate the dynamics of the pandemic in one year.

## 6. Results and Discussion

A summary of the comparative project between Maximum Likelihood Estimation (MLE) and Support Vector Machine (SVM) regression appears in this section regarding their survival time modeling techniques. The techniques were used to determine probability density function (PDF) together with survival function and hazard function through performance assessment using MSE, RMSE, and IMSE metrics.

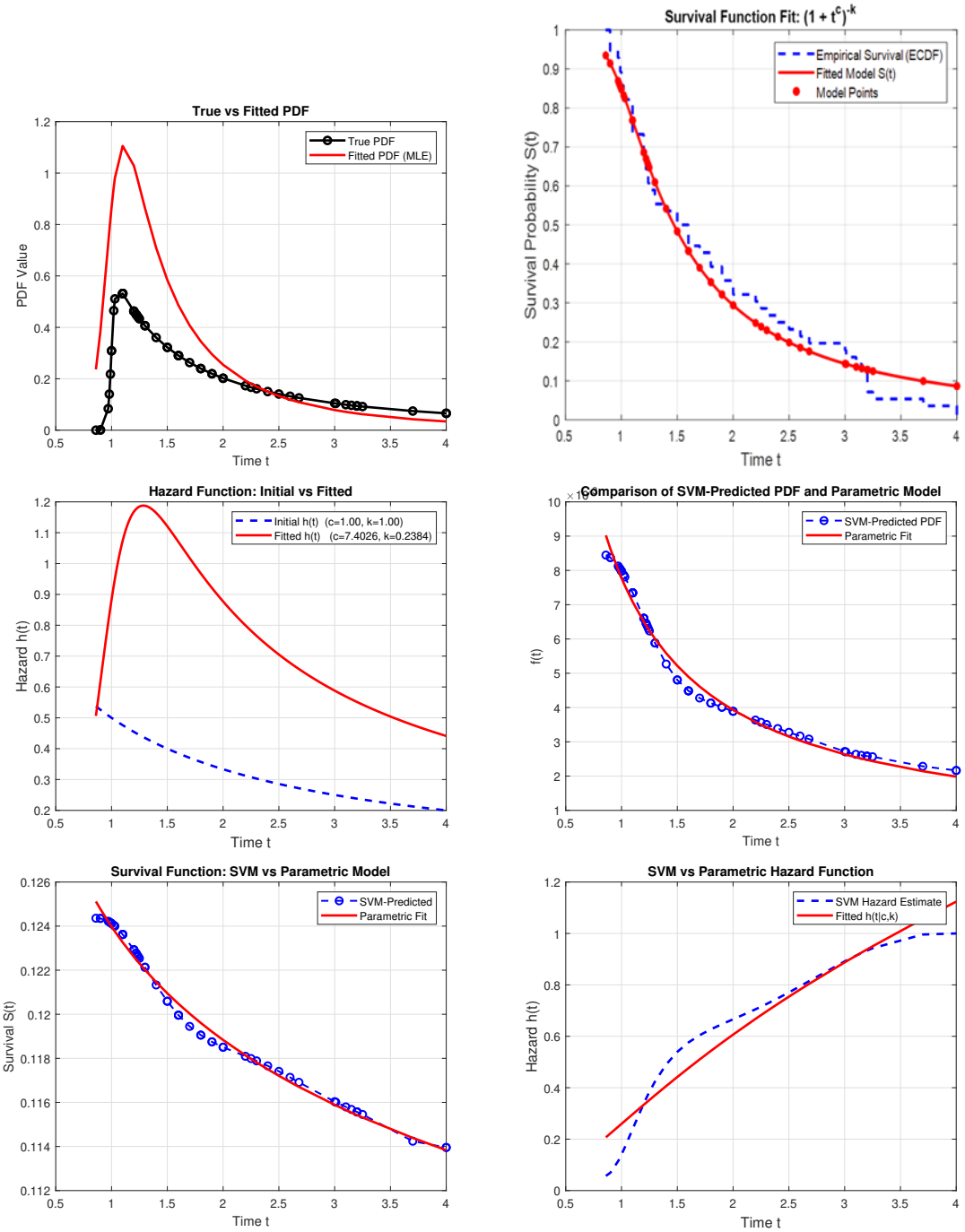


Figure 16. Estimation and behavioral Evaluation of PDF, Survival, and Hazard Functions via MLE and SVM Approaches

### *6.1. Parameter Estimation and Functional Accuracy*

The SVM approach maintained lower RMSE metrics than MLE in estimating survival and hazard functions when using sample sizes from $n = 50$ to $n = 1000$. With 1000 observation points the survival function from SVM showed an improvement in RMSE measurement that reached nearly ten times better than MLE.

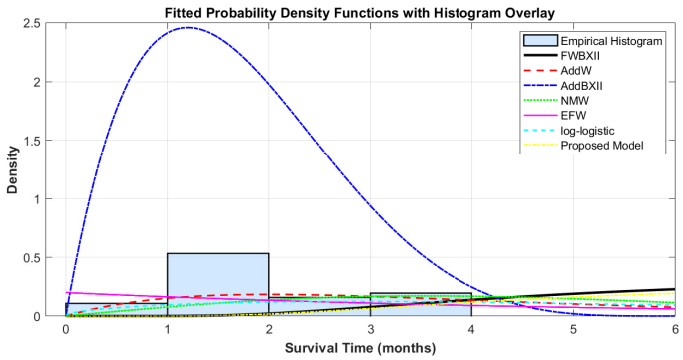


Figure 17. Visual Analysis of Survival Duration in Cancer Patients Using Histogram and Boxplot

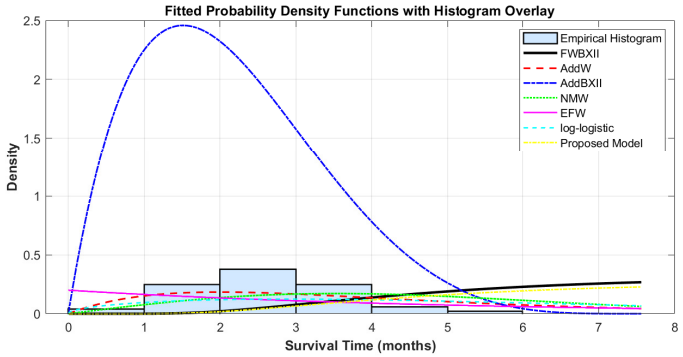


Figure 18. Visual Analysis of Survival Duration in Cancer Patients Using Histogram and Boxplot

### *6.2. PDF Estimation*

The findings demonstrate that SVM benefited PDF estimation because it produced progressively lower RMSE and IMSE outcomes with rising sample numbers. The capability of SVM to approximate cause distribution becomes noticeable through this evaluation.

### *6.3. Comparison with Previous Studies*

The Burr-type model developed here demonstrated better performance than the log-logistic model in previous research by lowering both RMSE and IMSE values which are presented in Table 8. The better fit highlights that this model works well with difficult survival data.

### *6.4. Graphical Interpretation*

Figure 12 to Figure 18 validate the numerical results. The survival function shows a plausible declining pattern which matches the inverted-U shape of the hazard function that SVM effectively detects. The SVM predictions match most closely with empirical PDF values in Figure 13. The SVM model exhibits accurate performance according to Figure 16 by showing minimal discrepancies between predicted and observed data.

### *6.5. Clinical Implications*

Initial patient survival decrease rapidly but eventually reaches a plateau according to the survival models. The prediction capabilities of SVM match this evolving pattern better than other methods so their value lies in aiding healthcare assessments and patient surveillance at an early stage.

### *6.6. Summary*

MLE showed effective results but SVM proved to have superior performance in every estimation task. Non-parametric flexibility renders SVM the ideal choice for complex or nonlinear survival data and therefore it should be the preferred method for future applications.

## 7. Conclusions

In this study applied the Maximum Likelihood Estimation (MLE) together with Support Vector Machine (SVM) through Burr Type XII distribution for their comparative analysis. The established modeling system served within survival analysis to assess remission duration and patient survival patterns among cancer patients. Modern cancer prognosis improved through the integration of computational techniques with traditional statistical methods throughout the last few decades. Support Vector Machines have emerged as strong nonlinear learning models to achieve effective results in medical diagnostic tasks and prognostic forecasting and survival analysis applications. Numerous essential results emerge from this analysis as described below:

- Survival-related characteristics obtained from the Burr Type XII distribution received more accurate predictions through the SVM-based modeling process.
- During the time period MLE and SVM methods produced equivalent behavioral trends which validated their shared trend recognition capabilities.
- An inverted bathtub shape formed by the estimated hazard function showed a time-dependent rising then declining pattern which matches known failure and risk patterns studied in survival analysis. Studies on breast and brain cancer patient cohorts demonstrated the same pattern that matches theoretical properties of the Burr Type XII distribution.
- A monotonic decrease occurred in the survival function when the time variable t rose toward higher values according to survival analysis theoretical expectations.
- The probability density function followed an escalating then declining pattern matching what survival distributions typically display.
- Support Vector Machines (SVMs) proved to be a very successful method according to performance metrics analysis for modeling the survival patterns of cancer patients.
- When survival time $t$ extended throughout duration the survival function values diminished in accordance with theoretical survival function behavior.

The research findings obtained from Support Vector Machine (SVM) modeling matched well with those produced by Maximum Likelihood Estimation (MLE) in this investigation because of widespread agreement across both methods. The experimental outcomes establish the solid structure of the proposed framework which can function as foundation work for upcoming research. The importance of extending this research stems from the opportunity to model lifetime distributions through Artificial Neural Network (ANN)-based approaches under different conditions and scenario conditions which would contribute to survival analysis methodology advancement.

**Acknowledgements**

The authors express their deep gratitude to Duhok Polytechnic of University for the facilities provided, which contributed to improving the quality of this work.

**Availability of Data and Materials**

The data used to support the findings of this study are included in the article.

**Ethical Approval**

All the authors demonstrating that they have adhered to the accepted ethical standards of a genuine research study.

**Author contributions**

All authors have sufficiently contributed to the study and agreed with the results and conclusions.

REFERENCES

1. Achieve Clinical Research, *The anxiety over breast cancer recurrence*, Retrieved from http://www.achieveclinical.com/news/the-anxiety-over-breastcancer-recurrence/, 2012.
2. American Cancer Society, *Cancer facts & figures*, Retrieved from http:// www.cancer.org/Research/CancerFactsFigures/, 2010.
3. American Cancer Society, *Learn about cancer: breast cancer*, Retrieved from http://www.cancer.org/cancer/breastcancer/detailedguide/breast-cancer-keystatistics/, 2012.
4. N. Shirawia, A. Kherd, S. Bamsaoud, M. Tashtoush, A. Jassar, E. Az-Zo'bi, *Dejdumrong Collocation Approach and Operational Matrix for a Class of Second-Order Delay IVPs: Error analysis and Applications*, WSEAS Transactions on Mathematics, vol. 23, pp. 467-479, 2024.
5. W. Garcia-Closas, *Heterogeneity of breast cancer associations with five susceptibility loci by clinical and pathological characteristics*, PLos Genet, vol. 4, no. 4, e1000054, 2008.
6. G.A. Colditz, B. A. Rosner, W. Y. Chen, M. D. Holmes, S. E. Hankinson, *Risk factors for breast cancer according to estrogen and progesterone receptor status*, J. Natl. Cancer Inst., vol. 96, no. 3, pp. 218-228, 2004.
7. W. A. Berg, *Detection of breast cancer with addition of annual screening ultrasound or a single screening MRI to mammography in women with elevated breast cancer risk*, JAMA, vol. 307, no. 13, pp. 1394–1404, 2012.
8. K. Kourou, T. P. Exarchos, K. P. Exarchos, M. V. Karamouzis, D. Fotiadis, *Machine learning applications in cancer prognosis and prediction*, Comput Struct Biotechnol J., vol. 13, pp. 8–17, 2015.
9. H. Hussein, S. Hussein, A. Hussain, M. Tashtoush, *Estimating Parameters of Software Reliability Growth Models Using Artificial Neural Networks Optimized by the Artificial Bee Colony Algorithm Based on a Novel NHPP*, Mathematical Modelling of Engineering Problems, vol. 12, no. 1, pp. 25-36, 2025. https://doi.org/10.18280/mmep.120104.
10. J. E. Bibault, P. Giraud, A. Burgun, *Big data and machine learning in radiation oncology: state of the art and future prospects*, Cancer Lett, vol. 382, no. 1, pp. 110-117, 2016.
11. F. Ambrogi, N. Lama, P. Boracchi, E. Biganzoli, *Selection of artificial neural network models for survival analysis with genetic algorithms*, Comput Stat Data Anal, vol. 52, no. 1, pp. 30-42, 2007.
12. A. Shafiq, A. B. Çolak, T. N. Sindhu, S. Lone, A. Alsubie, F. Jarad, *Comparative study of artificial neural network versus parametric method in COVID-19 data analysis*, Results Phys, vol. 53, 105613, 2020.
13. A. SH, K. S. Fatah, M. Sulaiman, *Estimating the Rate of Occurrence of Exponential Process Using Intelligence and Classical Methods with Application*, Palestine Journal of Mathematics, vol. 2, no. 12, pp. 256-269, 2023.
14. A. Ragab, J. Green, *On order statistics from the log-logistic distribution and their properties*, Commun Theory Methods, vol. 13, no. 21, pp. 2713-2729, 1984.
15. F. Ashkar, S. Mahdi, *Fitting the log-logistic distribution by generalized moments*, J. Hydrol, vol. 328, no. 4, pp. 694-703, 2006.
16. S. Bennett, *Log-logistic regression models for survival data*, Journal of the Royal Statistical Society: Series C, vol. 32, no. 2, pp. 165-171, 1983.
17. S. Jagasia, E. Tasci, Y. Zhuge, K. Camphausen, A. Krauze, *dentifying patients suitable for targeted adjuvant therapy: advances in the field of developing biomarkers for tumor recurrence following irradiation*, Expert review of precision medicine and drug development, vol. 31, no. 1, pp. 33-42, 2023.
18. M. M. Shoukri, I. U. Mian, D. S. Tracy, *Sampling properties of estimators of the log-logistic distribution with application to Canadian precipitation data*, Can J Stat, vol. 16, no. 3, pp. 223-236, 1988.

19. M. M. Ali, A. H. Khan, *On order statistics from the log-logistic distribution*, J Stat Plan Inference, vol. 17, pp. 103-107, 1987.
20. M. S. Mousa, & M. El-Morshedy, *Modified Burr Type XII distribution with applications to reliability and survival analysis*, Journal of Applied Statistics, vol. 49, no. 1, pp. 104-124, 2022.
21. A. M. Sarhan, & J. Apaloo, *Bayesian estimation of the Burr Type XII model with applications to survival data*, Communications in Statistics: Simulation and Computation, vol. 50, no. 2, pp. 447–465, 2021.
22. S. J. Almalki, & G. A. Badr, *Exponentiated Burr Type XII Distribution: Properties and Applications*, Journal of Statistical Theory and Applications, vol. 20, no. 2, pp. 163–180, 2021.
23. A. A. Ewees, et al., *Hybrid prediction model combining statistical lifetime distributions and machine learning classifiers for reliability analysis*, Expert Systems with Applications, vol. 215, 119186, 2023.
24. A. R. Nasiru, et al., *A review of recent applications of Burr distributions in machine learning and reliability engineering*, Reliability Engineering & System Safety, vol. 231, 108970, 2023. https://doi.org/10.1016/j.ress.2022.108970
25. A. Hussain, K. D. Pati, A. K. Atiyah, M. A. Tashtoush, *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, vol. 17, no. 1, pp. 1-15, 2025. https://doi.org/10.15849/IJASCA.250330.01
26. C. Cox C, H. Chu, M. F. Schneider, A. Munoz, *Parametric survival analysis and taxonomy of hazard functions for the generalized gamma distribution*, Stat Med, vol. 26, no. 23, pp. 4352–4374, 2007.
27. H. W. Block, T. H. Savits, H. Singh, *The reversed hazard rate function*, Probab Eng Inf Sci, vol. 12, no. 1, pp. 69-90, 1998.
28. A. Hussain, Y. Oraibi, Z. Mashikhin, A. Jameel, M. Tashtoush, E. Az-Zo'Bi, *New Software Reliability Growth Model: Piratical Swarm Optimization-Based Parameter Estimation in Environments with Uncertainty and Dependent Failures*, Statistics, Optimization & Information Computing, vol. 13, no. 1, pp. 209-221, 2025.
29. Y. Wang, H. Dai, Z. Chen, S. He, W. Wang, & M. Gao, *Simulation study on a novel solid–gas coupling hydrogen storage method for photovoltaic hydrogen production systems*, Energy Conversion and Management, vol. 299, 117866, 2024.
30. C. Wu, Q. Wang, X. Wang, S. Sun, J. Bai, D. Cui, & H. Sheng, *Effect of Al2O3 nanoparticle dispersion on the thermal properties of a eutectic salt for solar power applications: Experimental and molecular simulation studies*, Energy, vol. 288, 129785, 2024.
31. Z. Yue, X. Shen, Z. Chang, & G. Zou, *Experimental and simulation study on quasi-static and dynamic fracture toughness of 2A12-T4 aluminum alloy using CTS specimen*, Theoretical and Applied Fracture Mechanics, vol. 133, 104556, 2024.
32. S. S. Hassan, F. A. Doori, E. Arebe, *Estimating the survival and risk functions of a loglogistic distribution by using order statistics with practical application*, Int J Nonlinear Anal Appl, vol. 13, no. 1, pp. 2483–2502, 2022.
33. X. Hu, & H. Ren, *Reliability estimation of inverse Weibull distribution based on intuitionistic fuzzy lifetime data*, Axioms, vol. 12, no. 9, 838, 2023.
34. A. Akhgar, D. Toghraie, N. Sina, M. Afrand, *Developing dissimilar artificial neural networks (ANNs) to prediction the thermal conductivity of MWCNT-TiO2 Water-ethylene glycol hybrid nanofluid*, Powder Technol, vol. 355, pp 602–610, 2019.
35. A. B. Çolak, T. N. Sindhu,S. A. Lone, M. T. Akhtar, A. Shafiq, *A comparative analysis of maximum likelihood estimation and artificial neural network modeling to assess electrical component reliability*, Qual Reliab Eng Int, vol. 23, pp 325-338, 2023.
36. E. Az-Zo'bi et al., *Novel topological, non-topological, and more solitons of the generalized cubic p-system describing isothermal flux*, Optical and Quantum Electronics, vol. 56, no. 1, pp. 1-16, 2024.
37. A. Hussain, M. Sulaiman, S. Hussein, E. Az-Zo'bi, M. Tashtoush, *Advanced Parameter Estimation for the Gompertz-Makeham Process: A Comparative Study of MMLE, PSO, CS, and Bayesian Methods*, Statistics, Optimization & Information Computing, vol. 13, no. 6, pp. 2316-2338, 2025.
38. A. Hussain, K. Mahmood, I. Ibrahim, A. Jameel, S. Nawaz, M. Tashtoush, *Parameters Estimation of the Gompertz-Makeham Process in Non-Homogeneous Poisson Processes: Using Modified Maximum Likelihood Estimation and Artificial Intelligence Methods*, Mathematics and Statistics, vol. 13, no. 1, pp. 1-11, 2025. http://doi.org/10.13189/ms.2025.130101.
39. E., Az-Zo'bi, *An approximate analytic solution for isentropic flow by an inviscid gas model*, Archives of Mechanics, vol. 66, no. 3, pp. 203–212, 2014.
40. I. Ibrahim, W., Taha, M., Dawi, A., Jameel, M., Tashtoush, E., Az-Zo'bi, *Various Closed-Form Solitonic Wave Solutions of Conformable Higher-Dimensional Fokas Model in Fluids and Plasma Physics*, Iraqi Journal For Computer Science and Mathematics, vol. 5, no. 3, pp. 401-417, 2024.
41. A. Shafiq, A. B. Çolak, T. N. Sindhu, S. A. Lone, A. Alsubie, F. Jarad, *Comparative study of artificial neural network versus parametric method in COVID-19 data analysis*, Results in Physics, vol. 1, no. 38, 105613, 2022.
42. K. Kourou, T. Exarchos, K. Exarchos, M. Karamouzis, D. Fotiadis, *Machine learning applications in cancer prognosis and prediction*, Computational and structural biotechnology journal, vol. 1, no. 13, pp. 8-17, 2015.
43. S. Chupradit et al., *A Multi-Objective Mathematical Model for the Population-Based Transportation Network Planning*, Industrial Engineering & Management Systems, vol. 21, no. 2, pp. 322-331, 2022.
44. F. K. Wang, *A new model with bathtub-shaped failure rate using an additive Burr XII distribution*, Reliability Engineering & System Safety, vol. 70, no. 3, pp. 305-312, 2000.
45. S. Chupradit et al., *Modeling and Optimizing the Charge of Electric Vehicles with Genetic Algorithm in the Presence of Renewable Energy Sources*, Journal of Operation and Automation in Power Engineering, vol. 11, no. 1, pp. 33-38, 2023.