JEAS,Vol. 31 No. 147 (2025) pp (145-157)
P-ISSN 2518-5764 - E-ISSN 2227-703X
DOI: https://doi.org/10.33095/pfny6g72

Journal of Economics and Administrative Sciences

A Nonparametric Estimator of the Reliability Function for a Carbon Fiber-Reinforced Polymer under Cyclical Stress

Haider K. Raheem*
Department of Administrative and Financial Affairs, AL- Karkh University of Science
***Corresponding author**

Alaa H. Jalob
Department of Studies and Planning, AL- Karkh University of Science

Received: 2/3/2025 **Accepted:21/4/2025** **Published: 1/6/2025**

Abstract:
This paper introduces a method for measuring the Reliability Function of Carbon Fiber Reinforced Polymer (CFRP) under cyclic stress using nonparametric estimators. The study's goal is to gain a better understanding of material deterioration by combining microscopic and macroscopic approaches while accounting for the uncertainty associated with the data and models. Several statistical techniques were used to estimate the reliability function from failure data generated under cyclic stress, such as Nadaraya-Watson, spline estimation, kernel density estimation, Bayesian estimation, and Gaussian process regression.
Real data will be simulated as a result of experiments conducted in Stanford structures and the vehicle laboratory (SACL) in partnership with NASA's AMES Research Center for Predictive Excellence (PCoE). Thus, developing the method of generating data based on failure rate and the number of cycles that allow calculation of possible failure periods, the parametric models will accurately predict physical behavior under periodic pressure, as evidenced by data analysis using performance measures such as Mean Squared Error (MSE) and the Determination Factor (R 2). The study also confirms the need to include uncertainty in forecasts to increase accuracy of results-level estimation stresses the use of statistical models that explain uncertainty in particular Bayesian Kernel Estimator and Gaussian Process Regression to estimate the reliability function of carbon fiber reinforced polymers CFRP maintenance costs that are important in the development of vehicle safety methodologies.
Keywords: Nonparametric Estimation of Reliability Function, Carbon Fiber-Reinforced Polymer (CFRP), Cyclic Stress, Kernel Density Estimation (KDE), Gaussian Process Regression (GPR), Bayesian Estimation.

1. Introduction:

Reliability functionality is an essential statistical tool used in the study of performance assessment and durability of engineering materials and systems (Härdle, 1990), which means that the component can remain in operation or functioning without interruption or failure during a given period. This function is often used to determine material resistance to repeated loads, especially in applications involving dynamic pressures such as space structures and aircraft, where the accuracy of prediction of the reliability function is an essential subject for improving engineering designs.

(Taki et al., 2018). Due to its unique properties, CFRP (carbon fibre-reinforced polymer) is one of the most essential composite materials used in engineering applications where its properties include resistance, lightweight and high strength. Resistance is widely used in space structures and aviation due to its high power-to-weight ratio. When polymer material is under repeated periodic pressure, it decomposes over time, creating precise cracks at the end. These cracks gather, as a result of which the system fails. (Polymer), periodic stress analyses and weakens the performance of CFRP by causing cumulative damage to the internal structure, which reduces its ability to maintain future loading as a result. The study of the reliability function under periodic stress contributes to understanding the behaviour and patterns of failure to give a clear idea of the possibility of improving the design of these composite materials to withstand these harsh operational conditions.

Calculating the reliability function of CFRP composite materials is important to ensure long-term performance consistent with various conditions of operational modes to predict anticipated failure times and identify factors leading to accelerated degradation and damage of composite substances under stress, we assess the reliability function and help to make the best decisions regarding maintenance based on accurate data, which reduces operational costs and, as a result of improving the integrity of the structure of systems that rely on these substances for their work. So, the main problem is to build a reliable estimation model for the reliability that should take into account data errors and the effects of other factors, such as load differences, temperature fluctuations.

To focus light on the relationship between stress cycles and failure rates, NASA provided failure data for CFRP under periodic stress that included detailed records of failure times for a sample set tested in various operational settings, resulting in the creation of statistical models based on real data that are concerned with predicting physical behaviour and estimating reliability functions with high accuracy, which are an important and reliable source as they represent the results of tests obtained from strongly controlled laboratories. This strong control thus leads to accurate statistical analysis and the ability to anticipate physical behaviour in cases of actual use.

For the purpose of providing a model for estimating R(t), the simulation method was used based on failure data generated using different methods to distribute non-standard probabilities. This approach seeks to consider the subject of the impact of periodic stress on the composite material and how failure escalates and develops over time. The reliability function was estimated using methods including Nadaraya-Watson, kernel regression, spline estimation, KDE, Bayesian estimators, and Gaussian process regression.

The reason why these models are suitable for data analysis without the need for rigorous assumptions about basic data distributions has given them reliability in predicting failure rates. Therefore, non-pivotal methods have been used to analyse CFRP failure data and then compare their performance by relying on measures such as MSE and R^2 to assess estimation accuracy.

In addition, to verify the different effects of carbon fibre-reinforced polymer such as frequency, stress and load levels, on material reliability. Results revealed that reliance on non-standard models gives accurate and flexible estimates of reliability function compared to standard methods, especially when we deal with data with uncertain distributions. As a result, the study showed that the integration of Bayesian techniques increases the accuracy of the estimate by calculating the uncertainty in the evidence and adding new methods of analysis to assess the reliability, which will help to improve polymer design and thus reduce the risk of failure.

1.1. Literature Review and Hypothesis:

More than one statistical method has been reviewed and examined to estimate reliability, as Nadaraya-Watson is a core-based smoothing statistical method that is a desirable option for non-standard regression (Ali, 2022). This method is effectively good at capturing basic data trends and patterns without assuming a specific parameter format as described in (Adam et al., 2025) which uses the Bayesian method to characterize the reliability of fatigue. Similarly, KDE estimates the calculation and determination of probability intensity functions that are important for characterizing the behaviour of failure distributions.

Reliability modelling options are many of them, Spline techniques such as supervised Spline and cubic B-shaped Spline (Härdle, 1990). The light focused on splinters and their use in predictive modelling by demonstrating their efficiency in smoothing longitudinal data and that Bayesian assessors improve the estimation of reliability functions by relying on previous information before subsequent data analysis, as well as using Bayesian neural networks to identify uncertainty in predicting fatigue life and highlighting the value of potential modelling.

Gaussian Process Regression (GPR) has been used as an effective method of reliability analysis, especially in calculating uncertainty for failure predictions and knowing GPR's ability to represent complex interactions during random processes in research such as (Norkin & Pichler, 2025), where adaptive KD learning was used to improve the parametric model that GPR conforms to today's practical reality, and recent developments in probable machine learning and its use in structural health monitoring make it a reliable and good choice.

The NASA Ames PCoE Research centre for Predictive Excellence, in collaboration with the Stanford Laboratory of SACL Structures and Compounds, collected experimental data on CFRP failure under periodic load, which has helped to create a failure data generation system that likely allows periodic calculation. Performance measures, including average square error MSE and the R2 determination coefficient, were used to verify the accuracy of the model to verify the effectiveness of the non-standard method of estimating reliability functionality.

Assessing reliability to improve predictive models requires the inclusion of uncertainty, in which the estimate of the Bayesian nucleus and the regression of the Gaussian process works very effectively while estimating the reliability function of the CFRP study subject polymer. These methods provide clear insights into the remaining life of composite materials. They help improve the safety of polymer and reduce the costs for its maintenance. Future studies should consider the approach of mixed modelling using many statistical techniques for refining estimates and improving their accuracy in analysing the reliability of composite materials.

2. Carbon Fiber-Reinforced Polymer under Cyclic Stress:

A probability that the composite material will remain with a reliable function after a certain time has passed:

$$\begin{aligned} R(t) &= P(T > t) \\ &= 1 - F(t) \quad \dots \quad (1) \end{aligned}$$

(Larrosa & Chang, 2012) clarify F (t) is the cumulative distribution function for complex material failure times. The failure rate can be described using the PDF function with t failure time as follows:

$$f(t) = -\frac{d}{dt}R(t) \qquad \dots \; (2)$$

The CFRP carbon fibre reinforced polymer behaviour operates under periodic stress conditions and has a significant impact on the reliability and durability of the structure of those composites because the periodic stress on the composite substances analyses these substances by scattering small and precise cracks resulting in the end exposure of these composites to failure. The failure rate of CFRP polymer can therefore be studied using models provided by this study to analyse these composite substances. (Taki et al., 2020) explained that the function of λ failure rate (N) grows exponentially with periodic stress:

$$\varphi(N) = \varphi_0 e^{aw} \quad \ldots \quad (3)$$

Since φ_0 represents the initial failure rate and a degradation or corrosion inhibitor and based on equation No. 3, CFRP is subject to the law (Nsanzumuhire et al., 2025):

$$\frac{da}{dw} = a(\Delta M)^K \quad = \quad \ldots \quad (4)$$

ΔM represents the periodic stress intensity coefficient.

3. Research Methodology:

In this study, a nonparametric statistical modeling approach was used to calculate the reliability function of carbon fiber-reinforced polymer (CFRP) under cyclic stress. The study is based on real NASA data, which includes recorded failure times and cycle counts for a group of samples that underwent cyclic stress tests under various operating settings.

Statistical simulation techniques were utilized to create analytical models that replicated the material's behavior using real data from experiments carried out at the Stanford Structures and Composites Laboratory (SACL) in partnership with NASA Ames Research Center's Prognostic Center of Excellence (PCoE). Several nonparametric approaches, including Nadaraya-Watson Kernel Regression, Spline Estimation, Kernel Density Estimation (KDE), Bayesian Estimator, and Gaussian Process Regression (GPR), were used to investigate the link between the number of cycles and failure time. Furthermore, modern statistical approaches were used to examine the effect of different operating circumstances on material performance. This methodology attempts to offer reliable reliability function estimates, which will aid in material design optimization and failure risk reduction in engineering applications.

4. Materials and Methods:

4.1. Nadaraya-Watson Kernel Estimator:

The Nadaraya-Watson Kernel Estimator is a nonparametric approach for estimating a smooth function using observed data, we apply kernel regression to estimate it and get R(t) and The failure periods T_i and their empirical cumulative probabilities $F(T_i)$ fulfill the following:

$$F(T_i) = E[Y_i \mid T_i]$$

Y_i is an indicator of failure occurrence, our goal is to estimate F(t), which is Y's conditional expectation given T Using kernel regression, (Ali, 2022) say the conditional expectation is:

$$\hat{F}(t) = \frac{\sum_{i=1}^{n} K_h(T_i - t) F(T_i)}{\sum_{i=1}^{n} K_h(T_i - t)} \quad \ldots \quad (5)$$

$K_h(u) = \frac{1}{h} K\left(\frac{u}{h}\right)$ is a kernel function with bandwidth h, where K(u) is a symmetric probability density function (e.g., Gaussian kernel), and h determines the smoothing level Therefore the estimated reliability function is :

$$\hat{R}(t) = 1 - \hat{F}(t)$$

Common possibilities for K(u) include the Epanechnikov Kernel, the bandwidth h is commonly determined via cross-validation. via the kernel-based estimate of F(t) from Equ.5 (Hussein et al., 2012) refers to the Nadaraya-Watson Kernel Estimator for the reliability function is:

$$\hat{R}(t) = 1 - \frac{\sum_{i=1}^{n} K_h(T_i - t) F(T_i)}{\sum_{i=1}^{n} K_h(T_i - t)} \quad \ldots \quad (6)$$

This yields a smooth, nonparametric estimate of dependability that is tailored to the underlying failure time distribution.

4.2. Spline-Based Estimation:

Spline-based estimation is an effective nonparametric technique for estimating smooth functions, including the reliability function R(t) Splines approximate functions using piecewise polynomials assuring smoothness and continuity at predetermined knots, a spline function is a piecewise polynomial function of degree that satisfies smoothness constraints at certain breakpoints called knots, a cubic spline (widely used due to its smoothness) F(t) is unknown we estimate it using splines and then deduce R(t), The spline function S(t) is a piecewise polynomial function of degree k that meets smoothness constraints at specific breakpoints known as knots, a cubic spline (often utilized for its smoothness) is defined as n say (Hasan et al., 2018) :

$$S(t) = \begin{cases} a_1 + b_1 t + c_1 t^2 + d_1 t^3, & t \in [\tau_0, \tau_1] \\ a_2 + b_2 t + c_2 t^2 + d_2 t^3, & t \in [\tau_1, \tau_2] \\ \vdots \\ a_m + b_m t + c_m t^2 + d_m t^3, & t \in [\tau_{m-1}, \tau_m] \end{cases} \quad \dots \quad (7)$$

Each section is a cubic polynomial to estimate the dependability function (Hasan et al., 2018) we approximate the F(t) using S(t), where. So, our reliability estimate becomes:

$$\hat{R}(t) = 1 - S(t)$$

For a valid spline, we impose continuity and smoothness criteria S(t) should be continuous at :

$$S(\tau_i^-) = S(\tau_i^+), i = 1,2,\dots,m-1 \quad \dots \quad (8)$$

These constraints result in a system of linear equations, which we solve to find the spline coefficients to fit the spline, we minimize the residual sum of squares (RSS) over the recorded failure times (T_i and F_i):

$$\min \sum_{i=1}^{n} \left(F_i - S(T_i)\right)^2$$

where is the empirical cumulative failure probability (Seyala et al.,2024) :

$$F_i = \frac{\sum_{j=1}^{i} \delta_j}{n} 1 \quad \dots \quad (9)$$

Where $\delta_j = 1$. If failure happens, $\delta_j = 0$ is else, the solution of minimizing the problem produces the spline coefficients $((a_i, b_i, c_i, d_i))$, which define S (t) After solving for S(t), the estimated reliability function is $\hat{R}(t) = 1 - S(t)$ which is a smooth and flexible approximation of the true reliability function.

4.3. Kernel Density Estimator:

The Kernel Density Estimation method is a nonparametric approach to estimating the probability density function f(t) (Mraoui et al., 2024) , we can integrate it to estimate the cumulative distribution function F(t), and then use Equ2 to create the reliability function R(t) The kernel density estimator of the probability density function f(t) is defined as:

$$\hat{f}(t) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{t - T_i}{h}\right) \quad \dots \quad (10)$$

K(x) is the kernel function, often the Gaussian kernel, defined as $\frac{1}{\sqrt{2\pi}} e^{-x^2/2}$. The kernel function to derive the cumulative distribution function F(t) integrate the estimated density function $\hat{f}(t)$:

$$\hat{F}(t) = \int_{-\infty}^{t} \hat{f}(u) du$$

Substituting $\hat{f}(t)$ from Equ.8, then:

$$\hat{F}(t) = \int_{-\infty}^{t} \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{u - T_i}{h}\right) du$$

By interchanging the summation and integral $\hat{F}(t) = \frac{1}{n}\sum_{i=1}^{n}\int_{-\infty}^{t}\frac{1}{h}K\left(\frac{u-T_i}{h}\right)du$ as a probability density function and the cumulative kernel function is defined as:

$$K^*(x) = \int_{-\infty}^{x} K(u)du$$

Thus, we rewrite the estimated CDF is $\hat{F}(t) = \frac{1}{n}\sum_{i=1}^{n} K^*\left(\frac{t-T_i}{h}\right)$ For a Gaussian kernel:

$$K^*(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}du = \Phi(x)$$

Where $\Phi(x)$ is the standard normal CDF, Thus:

$$\hat{F}(t) = \frac{1}{n}\sum_{i=1}^{n}\Phi\left(\frac{t-T_i}{h}\right) \quad \ldots \quad (11)$$

(Sadek et al., 2024), The KDE estimator for R(t) is, where from Equ.9 substituted.

$$\hat{R}(t) = 1 - \frac{1}{n}\sum_{i=1}^{n}\Phi\left(\frac{t-T_i}{h}\right) \quad \ldots \quad (12)$$

This offers a smoothed nonparametric estimate of the reliability function using kernel methods, also refers the bandwidth h is critical in determining the smoothness of $\hat{f}(t)$and subsequently $\hat{R}(t)$ A small h results in an overfitted (noisy) approximation, while a big h oversmoothes the function. Silverman's Rule of Thumb is a popular method for determining h:

$$h = 1.06\sigma n^{-1/5} \quad \ldots \quad (13)$$

where σ is the sample standard deviation.

4.4. Bayesian Nonparametric Estimation

Bayesian nonparametric approaches offer a versatile framework for estimating the reliability function R(t) without requiring a specific parametric distribution for failure times instead, we apply a Bayesian prior to the reliability function and update it with observed failure data. In this paradigm, we use stochastic processes, specifically the Dirichlet Process (DP) to explain the uncertainty in the reliability function. The cumulative distribution function F(t) and the reliability function R(t) are connected from Equ.2, with regard F(t) as an unknown function and assign a prior distribution over the potential functions, (Adam et al., 2025):

$$F(t) \sim \text{Stochastic Process Prior} \quad \ldots \quad (14)$$

The Dirichlet Process (DP) is widely utilized prior. The Dirichlet Process (DP) is a distribution over a distribution that is frequently used as a prior for unknown CDF from Equ.12 We defined:

$$F \sim \text{DP}(\alpha, F_0) \quad \ldots \quad (15)$$

Where α is the concentration parameter controlling variability,F_0 is the base distribution and often taken as a parametric family like an exponential or weibull distribution given failure times, where α is the concentration parameter governing variability, and F_0 is the basis distribution, generally considered as a parametric family like Weibull distribution, for failure times $T = \{t_1, t_2, \ldots, t_n\}$, the posterior estimate of F(t) is as follows:

$$F(t) \mid T \sim \frac{\alpha}{\alpha+n}F_0(t) + \frac{n}{\alpha+n}\hat{F}(t)$$

Where $\hat{F}(t)$ is the empirical CDF, $\hat{F}(t) = \frac{1}{n}\sum_{i=1}^{n} I\,(t_i \le t)$, the Bayesian reliability estimator is

$$\hat{R}(t) = 1 - F(t) = 1 - \left(\frac{\alpha}{\alpha+n}F_0(t) + \frac{n}{\alpha+n}\hat{F}(t)\right) \quad \ldots \quad (16)$$

(Gorgees et al., 2018) Prove that it represents a Bayesian, nonparametric estimate of the R (t).

4.5. Gaussian Process Regression Estimation:

Gaussian Process Regression is an effective nonparametric Bayesian method for function estimation, especially when dealing with noisy data. It offers a probabilistic framework for estimating a function and quantifying uncertainty, In the framework of reliability analysis we intend to estimate the reliability function R(t), we estimate it using Gaussian Process Regression (GPR) a Gaussian Process (GP) is a distribution over functions in which any finite set of function values follows a multivariate normal distribution A function is modeled as:

$$f(t) \sim GP\big(\mu(t), k(t,t')\big)$$

Where μ(t) is the mean function (commonly assumed to be zero: μ(t)=0 for simplicity), $k(t,t')$ is the covariance function that defines smoothness, The Squared Exponential (SE) kernel is a commonly used choice for $k(t,t')$:

$$k(t,t') = \sigma_f^2 \exp\left(-\frac{(t-t')^2}{2\ell^2}\right)$$

Where controls the variance is the length scale, determining how quickly correlations decay gave a set of observed failure times and their corresponding empirical failure probabilities. We aim to predict any new time, we assume the observed values are generated from a noisy process the variance is controlled by whereas the length scale (l) determines how rapidly correlations fade, given a series of observed failure times and their corresponding empirical failure probabilities, suppose the observed values come from a noisy process:

$$Y_i = F(t_i) + \epsilon, \epsilon \sim N(0, \sigma_n^2)$$

Where σ_n^2 is the noise variance, the joint prior distribution for the observed data and the new test point follows a multivariate normal distribution.

$$\begin{bmatrix} Y \\ F(t_*) \end{bmatrix} \sim N\left(0, \begin{bmatrix} K(T,T) + \sigma_n^2 I & K(T,t_*) \\ K(t_*,T) & K(t_*,t_*) \end{bmatrix}\right)$$

Where $K(T,T)$ is the covariance matrix between training points $K(T,t_*)$is the covariance vector between training and test points, and $K(t_*,t_*)$ is the test point's covariance according to Bayes' theorem, the posterior distribution of $F(t_*)$ given observations is also Gaussian:

$$F(t_*) \mid T, Y, t_* \sim N\left(\hat{F}(t_*), \sigma^2(t_*)\right) \quad \ldots \quad (17)$$

$$\hat{F}(t_*) = K(t_*,T)[K(T,T) + \sigma_n^2 I]^{-1} Y$$

This calculates the mean estimate of $\hat{F}(t_*)$ and the uncertainty $\sigma^2(t_*)$, using R(t), the estimated reliability function is $\hat{R}(t) = 1 - \hat{F}(t_*)$ then Expression for Estimation using GPR:

$$\hat{R}(t) = 1 - K(t,T)[K(T,T) + \sigma_n^2 I]^{-1} Y \quad \ldots \quad (18)$$

The variance equation is $\sigma^2(t) = K(t,t) - K(t,T)[K(T,T) + \sigma_n^2 I]^{-1} K(T,t)$this provides a probabilistic nonparametric estimate of the reliability function.

5. Carbon Fiber Reinforced Polymer Reliability Simulation:

Simulation is an important technique for studying how cyclic stress affects carbon fiber-reinforced polymer (CFRP). Failure data was created using statistical simulation approaches based on cycle count and stress rate allowing for an accurate examination of failure time distribution and estimate of the reliability function. Statistical model was created to provide simulated data that depicted the effect of cyclic stress on the material, this model uses probabilistic equations to calculate the failure probability for each sample based on cycle count and stress level A total of 2,660 failure times were generated using a random distribution within a predetermined range, such as 1 to 1,000 cycles and Sample sizes ranged from 50 to 2,500. These sizes cover a wide range of data volumes allowing for an examination of the impact of sample size on the accuracy of various estimators (Talreja, & Varna,2023) say, the chance of failure was determined using an equation based on the failure and stress level:

$$\text{Probability of Failure} = 1 - \exp(-\text{Failure Rate} \times \text{Number of Cycles} \times \text{Stress Level})$$

After computing the failure probability for each sample, the sample was determined to have failed using a random distribution. If the estimated probability exceeded a random number between 0,1, the sample was judged unsuccessful. The failure times were then recorded and classified for cumulative distribution analysis (ECDF). The cumulative distribution function (Empirical Cumulative Distribution Function - ECDF) was used to depict the fraction of samples that failed before a certain time. This function provides the foundation for comparing the performance of various estimators.

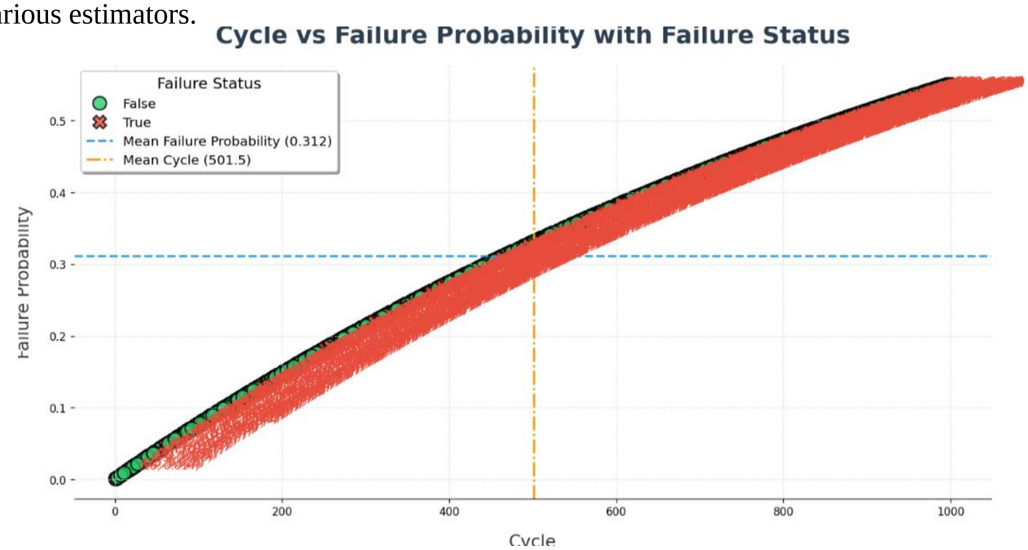

Figure 1: shows the probability of failure over 8651 stress cycles, based on recorded failures status.

Source: researcher's preparation

The carbon fiber-reinforced polymer was subjected to a total of 8,650 loading stress cycles with the severity of cyclic stress changing each time. This variation in stress intensity represents the real-world conditions that the material faces in engineering applications, which helps to improve the accuracy of the models used to mimic the material's behavior.

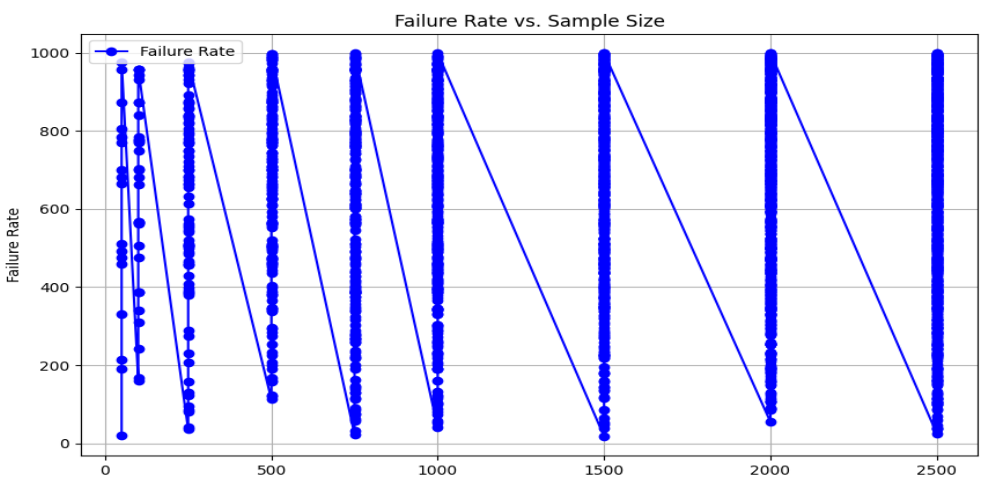

Figure 2: represents the generation of failure rate (2659 observations) for all sample sizes.

Source: researcher's preparation

From fig.2, this simulation model uses probabilistic equations to evaluate each sample's failure probability, taking into account both cycle count and stress level. The model creates 2,660 failure times based on a random distribution within a preset range, usually between 1 and 1,000 cycles. The sample sizes range from 50 to 2,500, allowing for a thorough investigation of failure behavior under various settings. This method allows for more accurate representations of reliability and durability in real-world applications.

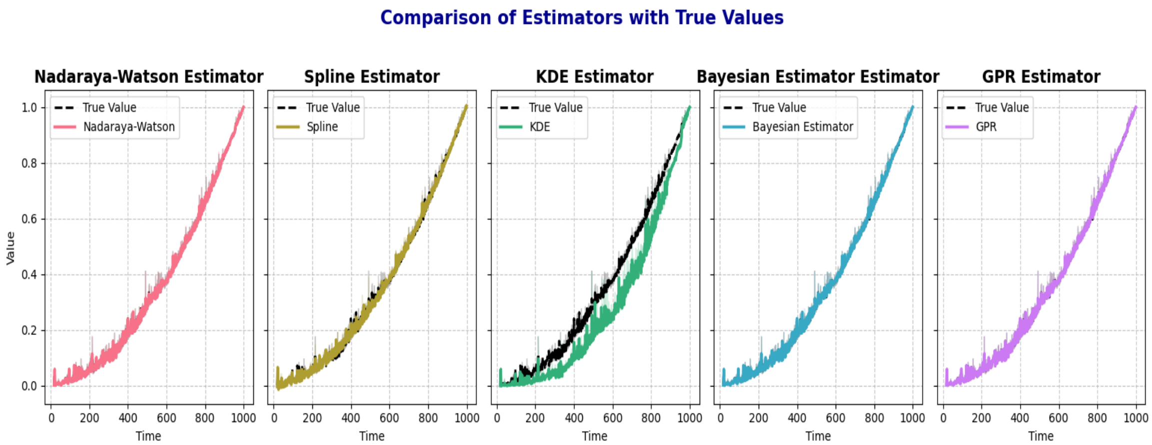

Figure 3: Represents all the nonparametric estimators curve for the reliability of carbon fiber-reinforced polymer under cyclic stress with true value.

Source: researchers' preparation

Table (1) : Represents the Comparison Metrics Between Estimators for all Different Samples.

Sample Size	Model	MSE	RMSE	MAE	R²	EVS
50	N-W	5.41975E-07	0.000736189	0.000258807	0.999993474	0.999993957
50	Spline	0.001104387	0.033232313	0.027301471	0.986701344	0.986701344
50	KDE	1.47579E-06	0.001214822	0.001054512	0.999982229	0.999988754
50	BKE	3.31021E-07	0.000575344	0.000358246	0.999996014	0.999996014
50	GPR	3.31021E-07	0.000575344	0.000358246	0.999996014	0.999996014
100	N-W	0.000116737	0.0108045	0.006337984	0.998613742	0.998629008
100	Spline	0.000978922	0.031287725	0.026501859	0.988375276	0.988375276
100	KDE	0.002524899	0.050248371	0.041713663	0.970016753	0.989602516
100	Bayesian	8.96656E-06	0.002994422	0.001217304	0.999893522	0.999893522
100	GPR	8.96656E-06	0.002994422	0.001217304	0.999893522	0.999893522
250	N-W	2.03844E-05	0.004514904	0.003269342	0.999756088	0.999756121
250	Spline	0.000405231	0.020130345	0.016519436	0.995151161	0.995151161
250	KDE	0.00182326	0.04269965	0.038132156	0.978183555	0.99547327
250	Bayesian	1.97637E-05	0.004445637	0.0033073	0.999763515	0.999763515
250	GPR	1.97637E-05	0.004445637	0.0033073	0.999763515	0.999763515
500	N-W	1.18962E-05	0.003449087	0.002587792	0.999857494	0.999857497
500	Spline	0.000120129	0.010960351	0.009051527	0.998560953	0.998560953
500	KDE	0.003750248	0.061239267	0.057062986	0.955075224	0.9940764
500	Bayesian	1.00826E-05	0.00317531	0.002338348	0.999879219	0.999879219
500	GPR	1.00826E-05	0.00317531	0.002338348	0.999879219	0.999879219
750	N-W	7.71052E-06	0.002776782	0.002034981	0.99990765	0.999907816

Sample Size	Model	MSE	RMSE	MAE	R^2	EVS
750	Spline	0.000156221	0.012498823	0.010003174	0.998128925	0.998128925
750	KDE	0.008220539	0.090667187	0.081727688	0.901541498	0.981207176
750	Bayesian	1.26376E-05	0.003554944	0.002641178	0.999848637	0.999848637
750	GPR	1.26376E-05	0.003554943	0.002641178	0.999848637	0.999848637
1000	N-W	2.52566E-06	0.001589233	0.001153397	0.99996976	0.999969786
1000	Spline	5.78279E-05	0.007604466	0.005888106	0.999307612	0.999307612
1000	KDE	0.006395869	0.07997418	0.074490548	0.923420702	0.989192594
1000	Bayesian	1.08235E-05	0.003289912	0.002473727	0.999870407	0.999870407
1000	GPR	1.89349E-05	0.004351426	0.003391352	0.999773288	0.999773288
1500	N-W	2.05239E-06	0.001432615	0.001072178	0.999975417	0.999975417
1500	Spline	7.13706E-05	0.008448114	0.006662878	0.999145134	0.999145134
1500	KDE	0.008029635	0.089608233	0.081881435	0.903822359	0.984121772
1500	Bayesian	7.96777E-06	0.002822725	0.002141234	0.999904563	0.999904563
1500	GPR	7.96778E-06	0.002822725	0.002141235	0.999904563	0.999904563
2000	N-W	2.62349E-07	0.0005122	0.000342496	0.999996856	0.999996856
2000	Spline	3.30199E-05	0.005746295	0.004545606	0.999604293	0.999604293
2000	KDE	0.00896663	0.094692294	0.085798951	0.892544793	0.980711852
2000	Bayesian	8.4446E-06	0.002905959	0.00232522	0.999898801	0.999898801
2000	GPR	8.4446E-06	0.002905959	0.00232522	0.999898801	0.999898801
2500	N-W	3.98836E-07	0.000631535	0.000459862	0.999995222	0.999995224
2500	Spline	4.63816E-05	0.006810402	0.005696769	0.999444343	0.999444343
2500	KDE	0.011309913	0.106348073	0.095070397	0.864505942	0.972770494
2500	Bayesian	5.38965E-06	0.002321563	0.001752426	0.999935431	0.999935431
2500	GPR	5.38966E-06	0.002321564	0.001752427	0.999935431	0.999935431

Source: researchers' preparation

6. Discussion of Results:

1- The Nadaraya-Watson's estimate shows that, if data are dense Nadaraya-Watson's estimate has good accuracy, although he may experience problems in data areas that are unevenly distributed, while noting that it is consistent with the form of real value in many regions, with only minor variations in the middle and final parts.

Very low MSE, RMSE, and MAE, indicating high prediction accuracy, high R^2 and EVS values indicate a nearly perfect match MAPE values are low, indicating small percentage error STD E values are consistently minimal across sample sizes, indicating stable error dispersion, bandwidth varies slightly across sample sizes, which influences smoothing.

2- The Spline Estimator shows a high smooth curve but at the beginning and end of the period we notice it shows a significant deviation, this means it performs well in the regions or middle parts, but fails to capture the sharp changes effectively, from figure 1 a spline regression higher MSE, RMSE, and MAE than N-W, indicating more substantial prediction mistakes. R^2 values remain high, although lower compared to kernel-based models. MAPE grows with sample size, implying greater proportional errors. The Max Error is extremely high, indicating probable outlier sensitivity. STD E values are higher, indicating increased error variability.

3- The Kernel Density Estimate gives an unstable estimate, this is visibly seen in the last part of the time period leading to estimates closer to the overall average rather than disclosing fine details, limiting its effectiveness to accurate estimates. The (KDE) has a higher MSE, RMSE, and MAE than N-W and Bayesian Kernel Estimator (B-KE). R^2 and EVS values are still relatively high, but lower than other kernel-based models. MAPE values rise dramatically with sample size, indicating scalability difficulties. Higher Max Error and STD E values indicate instability in forecasts.

4- We note that Bayesian Estimator estimates are close to real value and have a stable and accurate characteristic, and remain with real values, as these estimates can handle sharp data well, although it requires advance information to avoid Bias.

The Bayesian estimator achieves exceptionally low MSE, RMSE, and MAE, rivaling N-W. R^2 and EVS values are nearly perfect, indicating high predictive power. MAPE remains low, giving it a suitable model for percentage-based forecasts. Error readings are near zero, while STD E is constantly low compared with others estimates.

5- Once the estimated curve is observed using the GPR method, its output is accurate and stable, and it is characterized by superiority in detecting sharp differences, especially in the first half of the schedule, generally performs better than KDE, with less error dispersion. Gaussian process regression (GPR) has nearly comparable performance to Bayesian Estimator across all metrics. It has extraordinarily low MSE, RMSE, and MAE, showing strong prediction skill. R^2 and EVS stay around ideal levels. MAPE readings are consistently low.

7. Conclusions:

First, based on the results of the reliability function estimate in Table 1, the Bayesian Kernel Estimator models are higher (B-KE) and Gaussian Process Regression (GPR) compared to other models tested for accuracy, as these estimates had the lowest mean error boxes (MSE, RMSE, MAE, and MAPE metrics) each has the greatest R^2 value, which shows the ability to interpret a large amount or proportion of differences in data. In addition, the above models have shown excellent stability during all different volumes of data giving them a better advantage and choice for industrial structures and materials that require special high accuracy in the issue of dynamic and complex data.

Second, an estimated performance or Nadaraya-Watson is competitive compared with the accuracy of the estimation of the Bayesian Kernel Estimator Models and Gaussian Process Regression (GPR) In addition that the bandwidth coefficient has a significant impact on performance, the results show us that suboptimal values of this coefficient affect the accuracy of prediction while emphasizing the tuning of this package leads to the highest possible performance where such technologies as Grid Search and Bayesian Optimization can be used.

Third, table (1) shows us that the KDE estimate gives the weakest results compared to the rest of the models with a clear increase in the average line during the growth of the sample size indicating that KDE is estimated to be an unsatisfactory choice while dealing with large or complex data. In addition, it has the largest absolute error values and STD E error differences which means it is less stable than the rest of the models and so when we need models that give high prediction accuracy, this option is inappropriate.

Fourth, the performance of the Spline-Based Estimation model is well assessed in special circumstances and is influenced by extreme values compared to other models which can be observed through Max Error readings and standard deviations STD E which records this estimate's instability in some circumstances.

8. Recommendations:

First, We recommend the use of Bayesian Kernel Estimator (B-KE) and Gaussian Process Regression in such a study which requires the calculation of polymer reliability with high accuracy due to its industrial importance due to the fact that these two models have the lowest line rates and thus give the largest explanatory percentage of total variation in R 2 data making its performance consistent during different sample sizes.

Secondly, the possibility of using the KDE estimator should be reviewed especially in applications of composite materials used in highly important and vital industrial fields that require high efficiency and precision because it possesses the highest error rates and is characterized by instability in the size of large samples.

Third, we recommend paying more attention to the extent of results from applying the Split Regression model by addressing extreme values as it is influenced by these values resulting in higher error rates as is the case with standard deviations of errors.

Fourth, the impact of Nadaraya-Watson's package presentation on the accuracy of estimates and benchmarks with the rest of the models makes it necessary to recommend for the purpose of adjusting the bandwidth parameters of this model.

Finally, we recommend that a hybrid method combining B-KE and GPR capabilities be applied with the aim of improving their performance for the purpose of increasing the stability of capabilities in terms of efficiency in prediction accuracy.

Authors Declaration:

Conflicts of Interest: None

-We Hereby Confirm That All The Figures and Tables In The Manuscript Are Mine and Ours. Besides, The Figures and Images, which are Not Mine, Have Been Permitted Republication and Attached to The Manuscript.

- Ethical Clearance: The Research Was Approved by The Local Ethical Committee in The University.

9. References:

Adam, M. A., Singh, S. S. K., Abdullah, S., & Jedi, M. A. M. (2025). Bayesian approach in fatigue reliability characterization for stochastic-induced road load conditions. *Results in Engineering, 25,104505*. https://www.sciencedirect.com/science/article/pii/S2590123025005821

Ali, T. H. (2022). Modification of the adaptive Nadaraya-Watson kernel method for nonparametric regression (simulation study). *Communications in Statistics - Simulation and Computation*, *51*(2), 391–403. https://doi.org/10.1080/03610918.2019.1652319

Härdle, W. (1990). *Applied nonparametric regression*. Cambridge University Press. https://doi.org/10.1017/CCOL0521382483

Hasan, N. N., & Hussien, D. A. (2018a). Generalized spline approach for solving system of linear fractional Volterra integro-differential equations. *Ibn al-Haitham Journal for Pure and Applied Sciences*, *31*(1), 222–230. https://doi.org/10.30526/31.1.1842

Hasan, N. N., & Hussien, D. A. (2018b). Generalized spline method for integro-differential equations of fractional order. *Iraqi Journal of Science*, 1093–1099. https://doi.org/10.24996/ijs.2018.59.2C.13

Hussein, D., & Youssef, M. (2012). Nadaraya-Watson Estimator a Smoothing Technique for Estimating Regression Function. *Journal of Economics and Administrative Sciences*, *18*(65), 283. https://doi.org/10.33095/jeas.v18i65.1112

Larrosa, C., & Chang, F. (2012, July). Real time in-situ damage classification, quantification and diagnosis for composite structures. *In Proceedings of the 19th International Congress on Sound and Vibration* (pp. 1–8). Vilnius, Lithuania.

Mraoui, H., El-Alaoui, A., Bechrouri, S., Mohaoui, N., & Monir, A. (2024). Two-stage regression spline modeling based on local polynomial kernel regression. *Computational Statistics*. https://doi.org/10.1007/s00180-024-01498-x

Norkin, V., & Pichler, A. (2025). Models parametric analysis via adaptive kernel learning. *arXiv*. http://arxiv.org/abs/2501.14485

Nsanzumuhire, C., Daramola, O. O., Oladele, I. O., & Akinwekomi, A. D. (2025). Review of current progress on additive manufacturing of medical implants and natural/synthetic fibre reinforced composites. *Materialwissenschaft und Werkstofftechnik, 56*(1), 17–42. https://doi.org/10.1002/mawe.202400070

Sadek, A. M., & Mohammed, L. A. (2024). Evaluation of the performance of kernel non-parametric regression and ordinary least squares regression. *JOIV: International Journal on Informatics Visualization*, *8*(3), 1352–1360. https://joiv.org/index.php/joiv/article/view/2430

Seyala, N., & Abdullah, S. N. (2024). Cluster analysis on longitudinal data of patients with kidney dialysis using a smoothing cubic B-spline model. *International Journal of Mathematics, Statistics, and Computer Science, 2*, 85–95. https://ijmscs.org/index.php/ijmscs/article/view/8337

Taki, A., Firouzi, A., & Mohammadzadeh, S. (2018). Life cycle reliability assessment of reinforced concrete beams shear-strengthened with carbon fiber reinforced polymer strips in accordance with fib bulletin 14. *Structural Concrete, 19*(6), 2017–2028. https://doi.org/10.1002/suco.201700289

Taki, A., Firouzi, A., & Mohammadzadeh, S. (2020). Reliability analysis of carbon fiber-reinforced polymer shear-strengthened reinforced concrete beams subjected to deterioration. *ACI Structural Journal, 117*(4), 25–38.