

A concept for application of B-spline algorithm to density estimation of fatigue failures in 18Ni300 steel

Łukasz Blacha^{1*}, Adam Deptuła², Kamil Urbanowicz³

¹ Faculty of Mechanical Engineering, Opole University of Technology, ul. Mikołajczyka 5, 45-271 Opole, Poland

² Faculty of Production Engineering and Logistics, Opole University of Technology, ul. Ozimska 75, 45-370 Opole, Poland

³ Faculty of Mechanical Engineering and Mechatronics, West Pomeranian University of Technology in Szczecin, al. Piastów 19, 70-310 Szczecin, Poland

* Corresponding author, email: l.blacha@po.edu.pl

ABSTRACT

P-s-n curves detailing fatigue are basic data required to ensure high reliability design. Since fatigue testing requires significant time and resource costs, the development of reliable small-sample methods is crucial to ensure the reduction in time and costs of design process. This paper investigates the possibility for applying B-spline interpolation basis functions to approximate the probability density functions on various levels of high-cycle uniaxial fatigue using the sample data of only selected stress levels. It was found that using the proposed model an iterative approach can be used to ensure acceptable fitting accuracy and to predict fatigue life on the desired probability level. As a result, the findings show the possibility for rapid modeling of the probability density function and the resulting reliability in high-cycle fatigue by using the well-known interpolation techniques based only on the basic dispersion parameters. The proposed approach provides a robust framework for fatigue life prediction, paving the way for broader applications in engineering analysis. The proposed data augmentation method seems to demonstrate the potential to reduce fatigue testing costs and simplify testing procedures.

Keywords: B-spline, 18Ni300 steel, fatigue, fatigue probability density function.

INTRODUCTION

Fatigue failure is a type of structural damage classified as progressive and localized. The process leading to such failure is characterized as an irreversible, monotonic degradation of material properties due to the acting alternating stresses. It is a direct reason for fatigue cracking, an effect of material degradation mechanisms. This type of failure is typical for properly designed and maintained engineering structures, i.e. when the other failure factors can be excluded. In this way, fatigue analysis can be used to evaluate the service life of components and reduce the probability of accidents. It is the reason why fatigue analysis is of increasing importance and is a main issue in research and development in many engineering branches. Although the fatigue phenomena

has been intensively studied from 19th century, still the approaches applied are mostly deterministic. Such methods often rely on safety factors and fixed material properties although the inherent variability in e.g. material properties calls for a more realistic and comprehensive approach. Within the frames of analysis, fatigue life cannot be considered as deterministic due to its great variation for the same design in identical loading conditions [1–3]. Linear, bilinear or trilinear models were proposed to reproduce the s-n or ϵ -n field (stress-life or strain-life, respectively), justified on the basis of client requirements or preferences. Such justification may be viewed as unfortunate [1], targeted on one specific approach and motivation. On the other hand, assuming constant material properties often leads to overly conservative or unsafe predictions. In view of the above,

the probabilistic approach to fatigue analysis seems to be not only advantageous but necessary.

The variability in fatigue life is usually material-driven, evidenced as the effect of microscopic inhomogeneity due to manufacturing defects and microstructural inconsistencies [3–5]. It should be noted however, that the material microstructure and the resulting fracture mechanism can be affected also by the type of fatigue loading, as it was shown by e.g. Kowal and Szala [6], Rui et al. [7] or Lehner et al. [8]. The effect may become particularly significant in the case of SLM materials, as it was shown by Macek et al. [9], who evidenced strong influence of the uniaxial fatigue loading applied during the tests on the fracture surface roughness of SLM-manufactured AISI H13 steel. The correlation between the parameters of the fracture surface topography and fatigue performance under certain loadings can be evidenced using e.g. the FRASTA [10] or entire fracture surface method [11]. As a conclusion, various load levels need to be analyzed in order to properly describe the s-n field.

Probabilistic methods account for this variation by considering the material’s strength and fatigue life as linear or log random variables with specific probability distributions (e.g. normal, log-normal and Weibull (most popular) distributions [1, 12–14]), leading to more accurate predictions of fatigue behavior. This enables to assess the risk of failure at different confidence levels, which is critical for high-reliability systems. It allows for a more refined safety margin and better-informed decision-making regarding component lifespan and maintenance schedules. Within the fatigue reliability analysis, the probability density function (PDF) as an estimate of the density of random variable serves as a tool providing a framework to model and quantify uncertainty in the behavior of systems and components. Apart of the possibility to evaluate the likelihood of failure at the given time period, PDF can also be easily transformed in order to find e.g. mean time to failure, reliability function and hazard rate function - widely used in condition monitoring and planning of maintenance activities [15, 16].

The relation between the probability distribution and stress-life behavior can be most conveniently seen on a joint plot of a probability distribution and one of the typical engineering fatigue curves. One of the fatigue curves most popularly used in deterministic engineering approaches is the stress-life relation portrayed in a double-log

scale (i.e. the “s-n curve”), an example of which was plotted in Figure 1 together with the selection of corresponding PDFs.

Figure 1 illustrates the profile of a s-n curve in the medium/high cycle fatigue (MCF/HCF, respectively; in state of the art, the difference between ranges is blurred) at which the curve is assumed to be linear. Outside of this range (towards infinity) s-n curve is assumed to be horizontal (if the material has a fatigue limit) or with a different gradient (for materials that do not have a fatigue limit) [19], sometimes reaching a plateau at longer fatigue lives. The low cycle fatigue (LCF) is assumed to be reflected better in an ϵ -n approach due to the large amount of plastic strain within the load history. What needs to be pointed out as a conclusion is that a different distributions are needed to quantify the s-n field at different ranges of fatigue life and that the scatter changes along with the change of stress level.

Apart of the empirical distributions, estimating the PDF of a random variable can be performed in a (i) parametric or (ii) nonparametric approach. Parametric approach involves determining the underlying distribution of data from either a sample or observational data, when the type of the distribution is already assumed. There are several methods for estimating the PDF, each with its advantages and limitations. In case of the group (i) approaches one of the most known is the method of moments which involves estimating the parameters of a distribution by equating the sample moments to the corresponding theoretical moments of a chosen distribution. Although it has proved to be simple and fast, it may not produce a

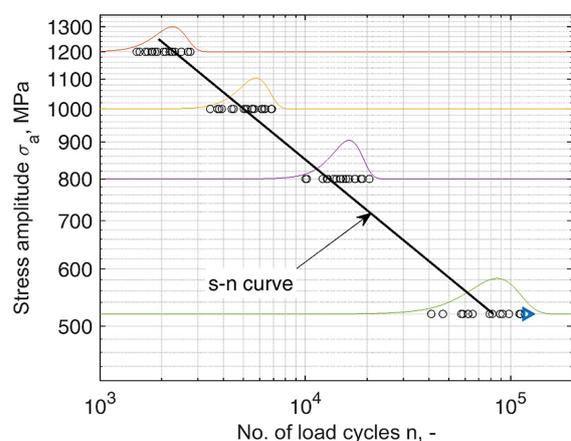


Figure 1. s-n curve for 18Ni300 maraging steel under uniaxial tension-compression, together with the simulated shape of Weibull PDFs on the tested stress levels [17]

good fit if the moments don't uniquely determine the distribution. Another method, the maximum likelihood estimation (MLE) involves numerical fitting a parametric model to the data in reference to the likelihood function obtained on the basis of the assumed distribution. The estimation algorithm determines the local maximum of the likelihood function L , which evaluates the joint probability density at the observed data sample. This type of approach is widely used in practical engineering, usually providing an optimal and unbiased estimate (if the assumed model is correct). A different approach gaining more attention nowadays is the Bayesian estimation. The technique used incorporates prior knowledge or beliefs about the distribution's parameters through a prior distribution, and updates this belief based on the observed data using Bayes' theorem. This may be successful, especially due to the natural capability of providing uncertainty estimates. Unfortunately, such approach may require significant computational effort as well as certain knowledge needed to select an appropriate prior.

In contrast to the above, the group (ii) approaches are statistical methods that do not assume any specific parametric form for the distribution from which the data are drawn. One of these is a histogram-based estimation in which the data is divided into intervals and the frequency of data points in each interval is computed. Here, the weakest point is on the selection of the length of such interval – a poor choice can lead to a considerable under- or over-smoothing. Kernel density estimation is a more sophisticated method that estimates the probability density function by smoothing the data with a one of a number of kernel functions combined with a specific smoothing parameter. The overall density estimate is the sum of kernels placed at data points and normalized by their number. It is more flexible than histograms, but a critical attention must be paid to appropriate selection of the bandwidth scaling the smoothing parameter which poorly chosen can over-smooth or under-smooth the estimate. Finally, another approach can be also be formulated through a spline-based estimation. Very wide applicability of piecewise polynomial functions can be used also in order to estimate the probability density function. Smoothness of the estimate can be controlled by the choice of the spline degree and the smoothness parameter controls. As a result, a possibility exists that a very smooth density estimates can be obtained with fewer assumptions

about the underlying distribution than the parametric methods.

In the paper a possibility to model the probability density function at different stress levels of high cycle fatigue via a B-spline algorithm is explored. Apart of the typical spline estimation, the B-spline algorithm used herein is applied only up to the procedure of finding the basis functions, without the interpolation part itself which allowed to shorten the procedure. In this way, the assumed model allowed to use an iterative approach to ensure acceptable fitting accuracy and to predict high cycle fatigue life on the desired probability level, with prior calibration using the sample data of only selected stress levels. Such approach could simplify fatigue-reliability component design by plugging only the boundary conditions at some stress levels and - as a result of the iterative algorithm - obtaining the PDFs of fatigue life at the desired, wide range of HCF stress level in an automated manner. The approach was verified through experimental validation on a large sample of fatigue specimens, additively manufactured from 18Ni300/MS1 steel powder. Although having the advantages like design flexibility, material efficiency or sustainability, their probabilistic fatigue behavior and the assumption of material stability is still inadequately explored.

The remainder of this paper is organized as follows. Description of the fatigue tests and specimen preparation process are presented, followed by the methodology and background of the research on the PDF estimation procedure. The PDFs are estimated and compared in the following section, with the results discussed to conclude the article.

METHODOLOGY

This study was validated on the log-fatigue lives of 18Ni300 maraging steel, a random variable assumed to follow a Weibull distribution. Weibull originally formulated the parameters of his distribution without a formulation applicable to a wide range of stress levels (although there is a natural difference between their values at different stress levels) [20]. Here, initially the parameters were investigated and estimated using the maximum likelihood estimation. This investigation was aimed on finding the PDFs corresponding to the experimentally verified fatigue data on selected stress levels. In this way a data basis was developed, to which the following B-spline

estimated PDFs were compared to in a validation process. The procedure served as a tool for verification of the assumptions underlying the concept for a semi-nonparametric estimation of PDF for fatigue life, a rapid way of analyzing probabilistic fatigue behavior of an engineering material.

Fatigue tests

The tested specimens were fabricated by direct metal laser sintering (DMLS/SLM) on an EOSINT M280 machine at Opole University of Technology. Such manufacturing technique uses metallic powder materials which are melted to melting temperature by a laser beam in such a way that the layer of metallic powder is fully molten throughout. The powder was supplied by EOS GmbH, Germany; its composition is presented in Table 1. Geometry of the specimens reflected the ASTM guidelines [15], as shown in Figure 2. The specimens were manufactured with laser powers of 285 W, 138 W and 60 W for stripe, contour and edge scanning at scanning speeds of 960 mm/s, 300 mm/s and 700 mm/s, respectively. The laser spot size was approximately 0.1 mm. In the process, layers were added on the base plate heated to 40 °C with an oxygen concentration less than 0.25% in the process gas atmosphere. After the printing process the specimens were cut to the desired lengths using an electrical bandsaw. At the end surfaces were smoothly finished in a bead blasting process using glass beads with diameters $\in <90, 150> \mu\text{m}$. The material

microstructure was normalized using a Nabertherm N41/H industrial furnace. During this heat treatment process, the material was heated to 490 °C, maintained at this temperature for 4 h and cooled in the furnace for 48 h. In the tests, the load applied to the specimen was constant in amplitude, sinusoidally alternating between tension and compression with a load ratio $R = -1$. The specimen was axially loaded in an Instron Electropuls E10000 testing machine, realizing a constant-frequency force-controlled system following the ASTM standard [22]. The 19–20 specimens were tested at four stress amplitude levels from 520–1200 MPa until reaching the failure criteria, which was assumed as total rupture or a 50% decrease in force. Number of reached load cycles was assumed to be the number of cycles to failure n_f - the durability. Amplitude of the stress applied at the gauge section was calculated as

$$\sigma_a = F_a/A \tag{1}$$

where: F_a is the amplitude of the uniaxial force (N) and A is the area of the test section (mm^2). The obtained $\sigma_a - n_f$ results are presented in Table 2. The processed data can also be found in [17].

After the fatigue tests, the obtained fractures were subjected to fracture surface fractography using the Tescan Vega scanning electron microscope (SEM). In overall, in such material configuration a soft martensite microstructure that includes hard nanoprecipitates is typical [18]. After the separation of outliers, the investigated

Table 1. Chemical composition of powder used in specimen manufacturing process (wt.%) [21]

Element	Fe	Ni	Co	Mo	Ti	Al	Cr
Min		17.00	8.50	4.50	0.60	0.05	-
Max		19.00	9.50	5.20	0.80	0.15	0.50
Element	Cu	C	Mn	Si	P	S	
Max	0.50	0.03	0.10	0.10	0.01	0.01	

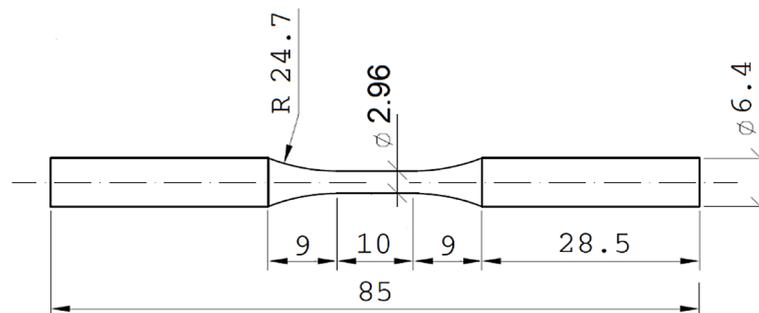


Figure 2. Geometry and dimensions of tested specimens

Table 2. The stress-life data obtained through fatigue tests [17]

Stress amplitude σ_a , MPa	Number of cycles to failure n_f -
1200	2189, 1899, 2683, 2241, 2688, 1663, 2042, 2258, 2686, 1902, 1710, 2757, 1510, 2320, 1561, 2089, 2493, 1852, 1784, 1799
1000	6350, 6826, 5557, 3745, 4481, 6332, 6366, 6875, 6199, 5498, 3923, 5030, 5649, 6200, 3813, 4384, 5129, 3443, 6111, 5228
800	14173, 12725, 10027, 17423, 10175, 14777, 15276, 12158, 15792, 18678, 15793, 14905, 18678, 18990, 12856, 17504, 13711, 16331, 20542, 14875
520	57521, 58672, 65341, 41105, 62215, 111748, 46798, 91366, 79041, 110653, 98261, 88955, 117488, 81423, 61968, 4x runout ($n_f > 1.2 \cdot 10^6$: 717254, 198394, 135495, 1184961)

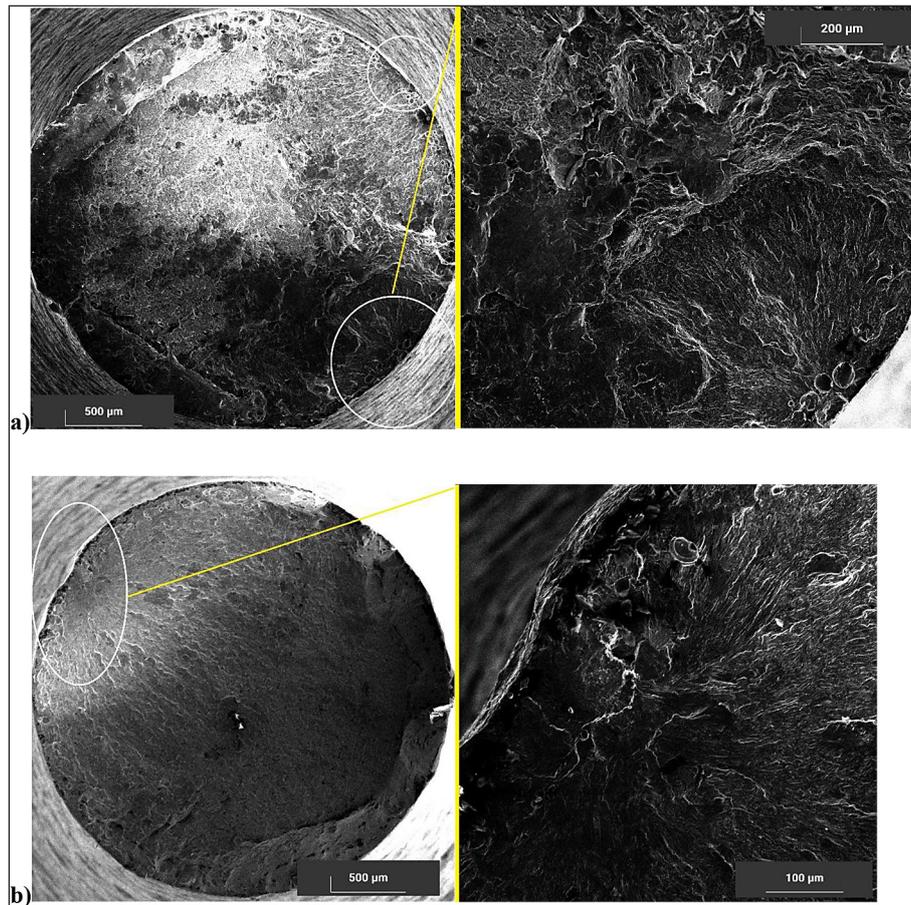


Figure 3. Fatigue fracture surfaces of the specimen failed at: (a) 18990 cycles of load at $\sigma_a = 800$ MPa, and (b) 61968 cycles of load at $\sigma_a = 520$ MPa, as an example of fracture surface topography

surfaces evidenced that cracks were initiated at the surface, with multiple or single initiation sites. Within the tested specimens, the fatigue fracture surfaces presented in Fig. 3 were observed along the plane of maximum normal stress.

Distribution fitting

A two-parameter variant of the Weibull distribution was used to model the fatigue life data, described by the log fatigue lives random variable ($N_f \sim W(\alpha, \beta)$, where: $N_f = \log(N)$). The Weibull

distribution (in different variants) is usually used in fatigue analysis and is expected to agree as well or better with the test data than the log-normal distribution [6.23–25]. This distribution type was chosen as the origin in the process of formulation of the B-spline PDF estimation model.

Maximum likelihood estimation

The Weibull PDF in a two-parameter variant is represented in the following relation:

$$f(n_f) = (\alpha/\beta)(n_f/\beta)^{\alpha-1} \exp(-(n_f/\beta)^\alpha) \quad (2)$$

where: α and β are the shape and scale parameters, respectively. Value of these parameters at the tested stress levels were determined through maximum likelihood estimation.

MLE is an efficient tool for estimating the parameters of a probability distribution based on sample data. The estimation algorithm determines the local maximum of the likelihood function L , which evaluates the joint probability density at the observed data sample. The parameters were chosen by maximizing the probability to obtain a data sample that most likely resembles the actual observed sample. Typically, the likelihood function is parameterized by a multivariate parameter θ . If we assume that the probability density of the analyzed random variable N_i is a function of n_i with θ fixed (it depends on parameter θ), the likelihood function should be viewed as a function of θ with n_i fixed. If so, the function

$$L(\theta | n_i) = f_{\theta}(n_i) \tag{3}$$

is the likelihood function. A detailed description of the applied algorithm can be described as follows:

- Create a column vector F for the shape parameter and another column vector G for the scale parameter. Both should have the same length j , with points spaced in a specified interval Θ (the parameter space):

$$F = [\alpha_1 \dots \alpha_j], G = [\beta_1 \dots \beta_j] \tag{4}$$

The quality of the results monotonically depends upon the length j (assuming an unchanged interval Θ).

- Generate the matrix M for the maximum number of possible different combinations of the parameters in the likelihood function. Vectors F and G should be reshaped, reorganized, and merged in the following manner:

$$M = \begin{bmatrix} \alpha_1 & \beta_1 \\ \vdots & \vdots \\ \alpha_1 & \beta_j \\ \vdots & \vdots \\ \alpha_j & \beta_1 \\ \vdots & \vdots \\ \alpha_j & \beta_j \end{bmatrix} \tag{5}$$

the resulting matrix M has the dimension of $(j^2, 2)$.

- Calculate the likelihood function L for each vector $\theta_i = [\alpha_i, \beta_i]$ represented in the i th row of matrix M , where $i = \langle 1 \dots j^2 \rangle$.

Table 3. Parameters of Weibull distribution obtained according to MLE

Stress amplitude σ_s , MPa	Shape parameter $\hat{\alpha}$	Scale parameter $\hat{\beta}$
1200	45.606	3.355
1000	53.047	3.761
800	59.985	4.215
520	41.868	4.937

- Find the value of i for which the θ_i vector maximises the likelihood function.
- The found vector is considered as the maximum likelihood estimate $\hat{\theta}$, $\hat{\theta} = \arg \max_{\theta \in \Theta} L(\theta; n)$ a vector whose entries are the values of α and β that make the observed data most probable within the distribution.

The obtained results are presented in Table 3 with respect to the tested load levels and without considering runouts.

The estimates shown in Table 3 were used to investigate if the θ parameter could be considered as an unbiased estimator of the population distribution parameters. To do so, sample mean μ was compared with the estimated population mean \bar{n}_i , calculated according to the following relation:

$$\bar{n}_i = \beta \cdot \Gamma(1 + 1/\alpha) \tag{6}$$

where: Γ is the Euler’s function (Euler integral of the second kind). The sample mean μ was calculated at each stress level for the log fatigue lives in Table 2. The obtained values and the corresponding probability of failure P_f are shown in Table 4.

$$P_f = 1 - \exp(-(n_i/\beta)^\alpha) \tag{7}$$

B-spline interpolation algorithm

A B-spline is a generalization of the Bézier curve and has its place in data analysis. It allows to derive smooth curves and surfaces on the basis of a sequence of limited data which values are known as control points. The mathematical model behind this approach allows to characterize this method as an elastic tool which can be applied in modelling of different types of curves and surfaces.

B-spline curve interpolation algorithm is based on generation of a smooth curve which goes through a set of control points P , $P = \{P_0, P_1, \dots, P_n\}$ where $P_i = (x_i, y_i)$ ($i = \{0, 1, \dots, k\}$) are coordinates. The curve is fitted with regards to these points and the *basis functions* N_{ij} . The

Table 4. Sample and population mean μ and \bar{n}_l , respectively

Stress amplitude σ_a , MPa	Mean value	Probability of failure
	\bar{n}_l	$P_f(\bar{n}_l)$
1200	3.314	0.435
1000	3.721	0.435
800	4.176	0.434
520	4.872	0.435

applied algorithm can be described within the following points:

- Define the degree of a B-spline curve, denoted by p .
- Generate *knots* u , $U = \{u_0, u_1, \dots, u_m\}$ where $m = k + p + 1$. Knots could be defined as uniform (i.e. equally spaced) or in a different manner.
- Calculate basis functions, $N_{i,j}(u)$ for each knot separately, according to the following recurrence relation:

$$N_{i,j}(u) = \frac{u - u_i}{u_{i+j} - u_i} N_{i,j-1}(u) + \frac{u_{i+j+1} - u}{u_{i+j+1} - u_{i+1}} N_{i+1,j-1}(u) \tag{8}$$

where: j is the degree of polynomial in a current iteration, $j = \{0, 1, \dots, p\}$ and i is the ordinal, $i \in [0, j]$ (i.e. value dependent on the current iteration number).

The algorithm for calculating Equation 8 could be illustrated through scheme shown in Figure 4.

Normally, when the basis functions are derived the algorithm follows on with the aim to determine the coordinates of points lying on the fitted curve with respect to each knot. However, this is not the case here since the basis function of a specific degree could already provide a sigmoidal shape of the PDF by itself, which – when utilized properly – could significantly reduce

the effort needed to achieve a solution. As it can be seen in view of Figure 4, the more knots are defined, the higher degree of the basis function is assumed and the shape of such function becomes more sigmoidal which becomes clear in view of Figure 5.

When defining the knot locations a special care must be taken as it has a major impact on the shape of this function, especially the first and the last one (u_0 and u_m , respectively) [26]. When knots are equally spaced, the basis function is symmetrical; however, when the knot vector is characterized by changing knot interval then the basis function becomes non-symmetrical with the left- or right asymmetry (depending on the chosen interval). Additional advantage is that the non-uniform B-splines achieve better approximating results [27]. Within this manner it is worth to investigate whether an introduction of knot locations defined with respect to the distribution moments (at only several test levels) could be capable of providing a solution applicable to a wider range of fatigue lives. Such an approach will disable the necessity of e.g. estimating the feature functions like in [28]. With regards to the above, the steps in the procedure were defined as:

- fatigue tests of a representative samples on different $\sigma_a^{(i)}$ stress levels, where $i \in (1, \dots, 4)$
- definition of the border knots, u_0 and u_m ,
- definition of the *inside* knots u_1, \dots, u_{m-1} ,
- introduction and optimization of the $u_i(\sigma_a)$ function, aimed on providing the solution for knot localization on stress levels between and different than $\sigma_a^{(i)}$.

By having the results of previously described fatigue tests the corresponding u_0 and u_m locations were assumed as possible to find by kernel density estimation (KDE) with sigmoidal kernel. Derived results were validated in comparison

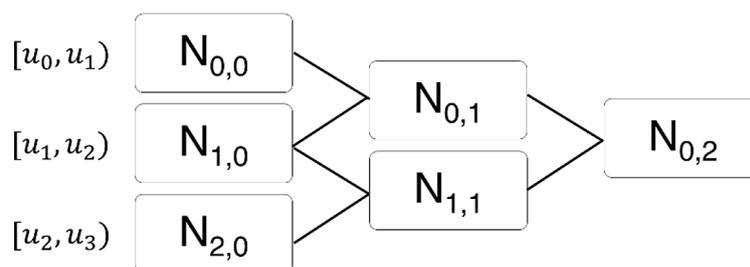


Figure 4. Iteration scheme illustrating the algorithm for calculating the basis function for a given point u , on the example of a degree two basis function

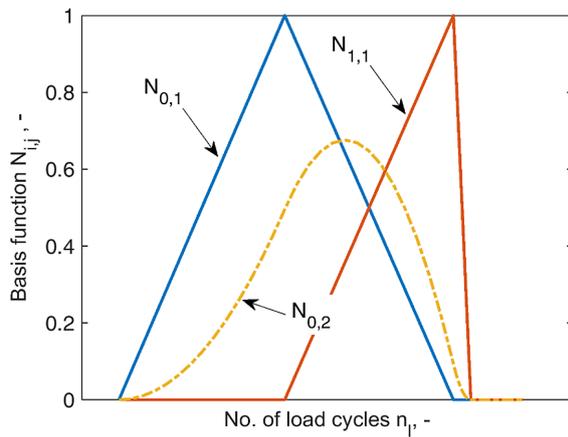


Figure 5. Simulation of selected non-uniform B-spline bases of a different degrees corresponding to a 7-element knot vector U defined in the domain of log number of load cycles n_l at specific stress level

with the distribution portrayed in PDFs estimated in a MLE approach. As a result, sixth degree basis function was found as capable to estimate the shape of PDF of fatigue lives. These findings would be difficult to derive without some initial assumption concerning the rest of the unknown knots, i.e. u_1, \dots, u_{m-1} . Simulation studies [29] allowed to derive a suggestion providing certain fitting accuracy:

$$u_1 = \bar{n}_l - 2std, u_2 = \bar{n}_l + std, u_3 = u_2 + \Delta u, \dots, u_6 = u_2 + 4\Delta u \quad (9)$$

where: \bar{n}_l is the first raw moment (i.e. mean log fatigue life) and std is the square root of the second central moment (i.e. standard deviation in a sample of N_l random variable).

Based on such assumption concerning the location of those nodes, regression analysis was performed in order to find the $u_i(\sigma_a)$

relation in order to implement the applicability in a wide range of fatigue lives. The derived regression was then optimized in an iterative approach aimed on minimizing the error function in order to: (i) eliminate imperfections resulting from simple least squares fitting and (ii) to fit the model more accurately to the real data, especially if the data is irregular or has deviations. Based on the conducted analysis, a generalized block diagram of the algorithm for estimating the probability density function has been developed, as shown in Figure 6. The solution and obtained results are presented in the next section.

RESULTS AND DISCUSSION

To estimate the parameters of the probability distribution of the analyzed 18Ni300 steel, the results of fatigue tests described in section 2.1 were used in the maximum likelihood estimation algorithm, which resulted in the values shown in Table 3. Then, using Table 3 and equation (2), PDF plots were generated at the analyzed stress amplitude levels σ_a , as illustrated in Figure 7.

Next, using kernel density estimation on a sample without outliers and referring to the estimated PDFs, locations of the extreme nodes, u_0 and u_m were determined as the extreme points of a PDF. Kernel density estimation is a common nonparametric estimation approach using each data point, a smoothing parameter and kernel function in order to estimate the PDF. Nevertheless it has to be remembered that actual performance of the kernel density estimator

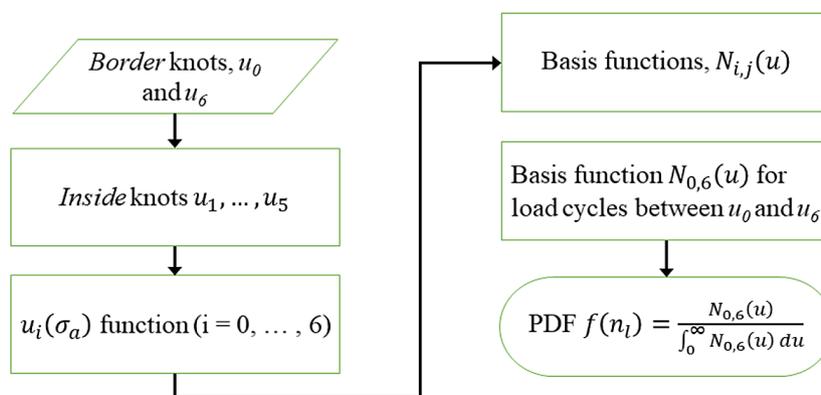


Figure 6. Block diagram for the algorithm for estimating the probability density function at the desired stress level of high cycle fatigue loading history

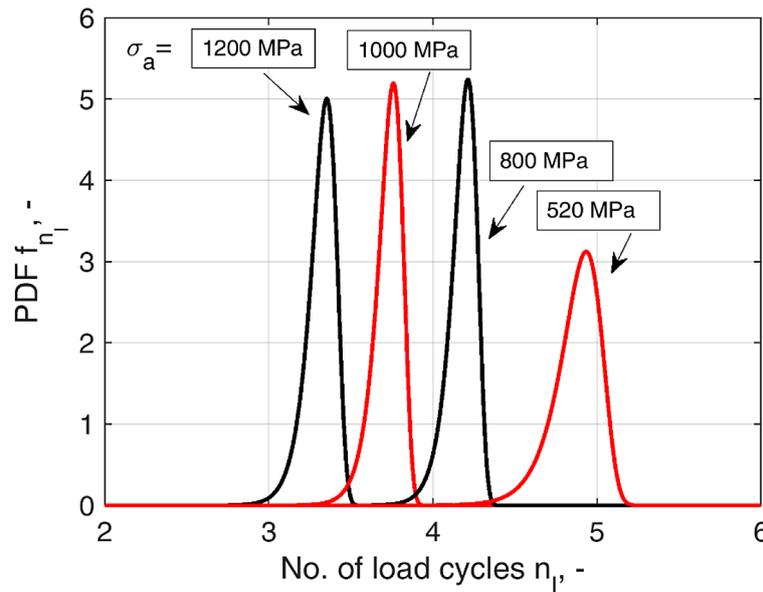


Figure 7. Probability density functions according to the MLE estimates of the Weibull parameters

still has many shortcomings like the significant impact of the smoothing parameter. Here, the sigmoidal kernel was used:

$$K(r) = 2/[\pi(\exp(r) + \exp(-r))] \quad (10)$$

$$r = (n_l - n_{l,i})/h \quad (11)$$

as providing the estimations which were most reliable when compared with the MLE obtained PDFs. Within this procedure the smoothing parameter h was assumed according to the Silverman’s optimum estimate:

$$h = 0.9 \cdot s_c \cdot n^{-1/5} \quad (12)$$

where: $s_c = \min(std, IQR/1.34)$ (IQR is the interquartile range).

In reference to the above and combined with the results obtained after the application of Equation 9 to fatigue data, a knot vector U was obtained at the four tested stress levels and summarized in Table 5.

Table 5. Knot vectors obtained according to Kernel Density Estimation and Equation 9

Stress amplitude σ_a , MPa	Knot vector U
1200	[2.860, 3.131, 3.406, 3.434, 3.463, 3.491, 3.520]
1000	[3.020, 3.544, 3.810, 3.850, 3.890, 3.930, 3.970]
800	[3.600, 3.999, 4.264, 4.295, 4.327, 4.358, 4.390]
520	[3.900, 4.578, 5.018, 5.076, 5.134, 5.192, 5.250]

Analysis and discussion of knot placement and its impact

The knot vectors used in B-spline basis functions play a critical role in shaping the resulting PDF. By strategically placing knots, one can effectively control the flexibility and accuracy of the B-spline estimation and, as a result – the smoothness of the resulting PDF, ensured by recursively defined $N_{i,j}(u)$ B-spline basis functions (Equation 8). The inherent smoothing effect ensures that the estimated PDF remains continuous and differentiable, which is beneficial in density integration. Following Table 5, this section delves deeper into the practical implications and optimization of these vectors for estimating the Probability Density Function (PDF) of fatigue lives in 18Ni300 steel. Figure 8 illustrates the impact of knot placement on the smoothness and accuracy of PDF estimation using B-spline basis functions. Two PDFs were compared:

- PDF with original (non-uniform) knots: this curve demonstrates higher precision in areas with greater data concentration, such as the intervals between u_4 and u_5 . The smaller spacing between knots in these regions allows for better representation of local variations in fatigue life data.
- PDF with uniform knot spacing: the curve for uniformly spaced knots is smoother but less precise in capturing the irregularities of the data distribution. This approach better reflects general trends but lacks the adaptability needed for complex data sets.

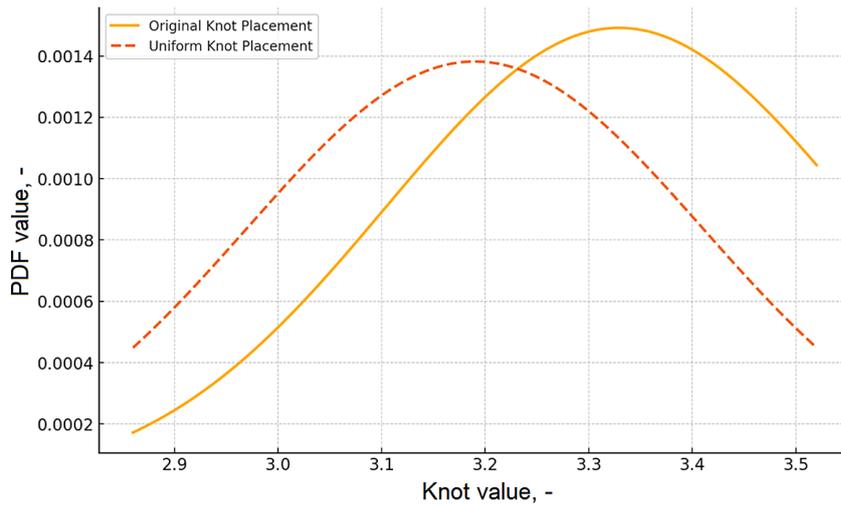


Figure 8. Impact of knot spacing on PDF smoothness

As shown in Figure 8, using non-uniform knots enables the PDF model to better adapt to the actual data distribution, which is crucial for analyzing the variability of fatigue life of 18Ni300 steel. In the study non-uniform spacing was adopted due to the asymmetry observed in Figure 7, additionally leveraging the flexibility of KDE and regression-based adjustments to better capture the stochastic nature of fatigue life distributions. It has to be remembered, that non-uniform spacing of knots accommodates varying data scatter across stress levels. For example: at $\sigma_a = 1200$ MPa, the smaller intervals between knots (u_4 and u_5) allow finer resolution in regions of high data concentration; at $\sigma_a = 520$ MPa, the wider spacing (u_5 to u_6)

reflects the greater scatter in fatigue life at lower stress level. Based on such assumption concerning the location of those nodes, regression analysis was performed in order to find and optimize the $u_i(\sigma_a)$ relation for HCF range.

Optimized fitting of the knot vector in a wide range of fatigue lives

At the beginning stress amplitudes were selected from a set $\sigma_a = \{520; 800; 1000; 1200\}$ MPa. Corresponding knot values (u_i) were derived from experimental observations as shown in Table 5. Linear regression was employed to explore potential trends in knot placement across the stress amplitudes. Each knot position at i -th

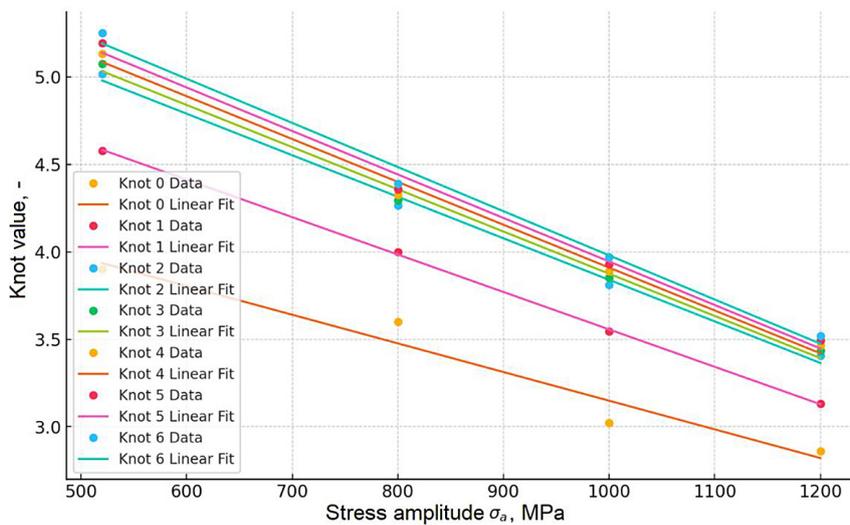


Figure 9. Linear fit of knot vectors for stress amplitudes

load level (u_i) was plotted against its corresponding stress amplitude ($\sigma_a^{(i)}$). The following equation described the linear fit for each knot:

$$u_i = a_i \cdot \sigma_a^{(i)} + b_i \quad (13)$$

where: a_i and b_i are the regression coefficients. Figure 9 illustrates the fitted lines for each knot position, demonstrating how knot locations evolve with varying stress amplitudes.

While the linear regression model, as depicted in Equation 13, provides an initial approximation of the relationship between stress amplitude (σ_a) and knot locations (u_i), its accuracy is limited by the assumption of a strictly linear dependency. Given the observed variability in the knot positions across the stress amplitudes, it was necessary to explore more flexible models to better capture the nonlinear nature of these relationships. Here, a quadratic regression model was employed for intermediate knots (u_1, u_2, \dots, u_{m-1}). The quadratic model extends the linear formulation by

introducing a second-order term, allowing for a more accurate representation of the data trends:

$$u_i = a_i \cdot (\sigma_a^{(i)})^2 + b_i \cdot \sigma_a^{(i)} + c_i \quad (14)$$

where: a_i, b_i and c_i are the regression coefficients.

Such approach ensures that the predicted knot values closely match the target knot values derived from experimental observations and Equation 9. The regression coefficients (a_i, b_i, c_i) for each quadratic model were optimized iteratively to minimize the prediction error, as detailed in Equation 15:

$$errf = \sum_{i=0}^m (u_{pred,i} - u_{target,i})^2 \quad (15)$$

The error function, as defined in Equation 15, quantifies the differences between the predicted knot positions (u_{pred}) and the experimentally derived values (u_{target}) estimated according to Equation 9 for each of the stress amplitude. To further illustrate the effectiveness of the quadratic regression model, a heatmap of the magnitude

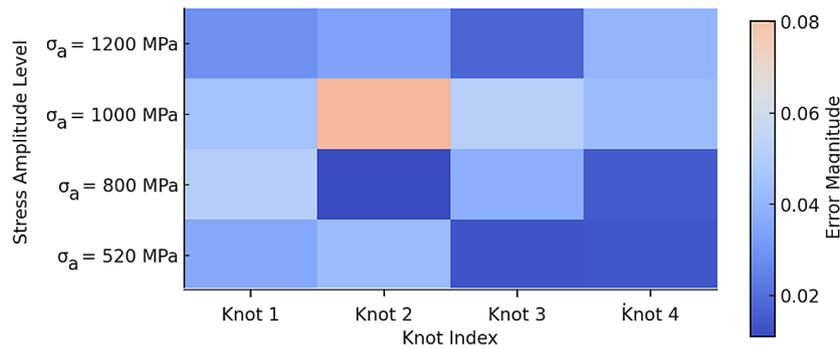


Figure 10. Heatmap of knot placement prediction errors

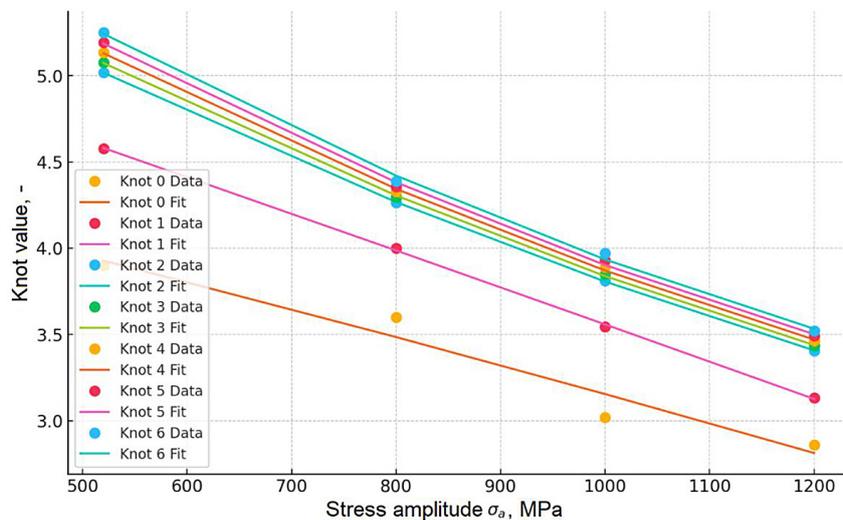


Figure 11. Quadratic regression fitting for knot placement across stress amplitudes

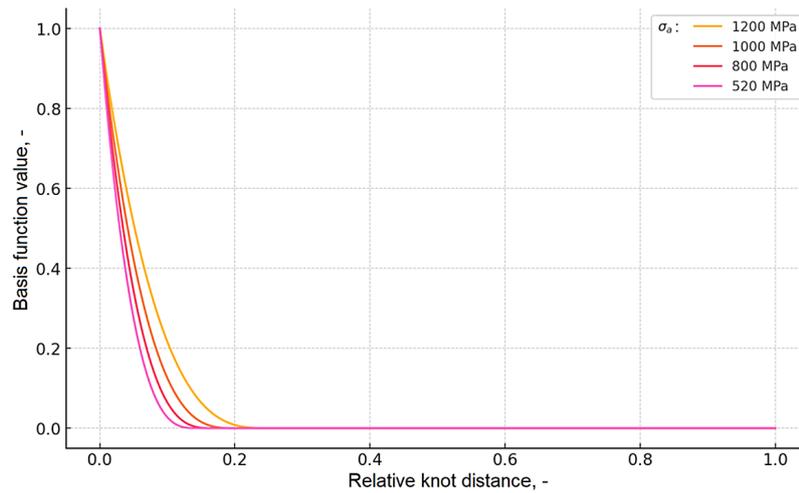


Figure 12. B-spline basis functions for stress amplitudes

of prediction errors (*errf*) is presented in Figure 10. The heatmap highlights the regions with the highest and lowest errors, providing a clear indication of the model’s performance across the dataset. The darker regions correspond to smaller errors, demonstrating the stability of the optimized quadratic regression model. In view of Figure 10 it can be seen that the quadratic regression model effectively minimized errors in intermediate knot placement, which ensures stability across the dataset.

To further illustrate the effectiveness of the quadratic regression model, Figure 11 depicts the fitted knot locations (u_i) against the experimental data for each stress amplitude (σ_a). The close alignment of the fitted lines with the observed knot positions demonstrates the robustness of the regression model in capturing trends across different stress levels.

The skewness of the generated basis functions was investigated according to Figure 12. In this figure each colored line represents a single basis function for a specific stress amplitude level.

These functions were constructed using a set of “knots” that divide the parametric range u (x-axis).

The basis functions are local, meaning their influence is limited to a specific segment of the u -range. The sum of the basis function values at any point u equals 1 (which is a fundamental property of B-splines). Y- values are restricted to the range from 0 to 1, reflecting the relative influence of the basis function at different points. It can be seen that each basis function reaches its peak value near the center of its corresponding knot and decreases as it moves away. Increasing the number of knots or the degree of the spline alters the shape and width of the basis functions.

To evaluate the differences between local and global models, PDFs generated using B-spline basis functions were compared with those obtained from a fitted Weibull model across the experimentally tested stress levels. The differences between both PDF types were quantified through residual analysis, where the obtained residuals highlight the local areas where the

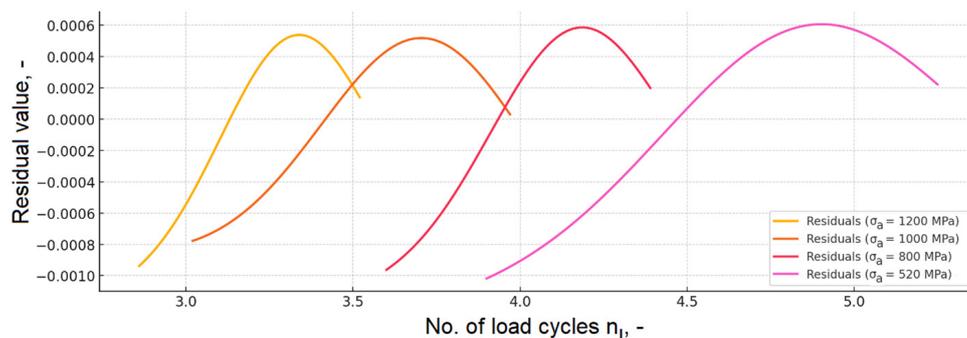


Figure 13. Residuals obtained from comparison between B-spline basis functions and Weibull PDFs

Table 6. Root mean square error between the PDFs estimated according to the B-spline and Weibull formulation

Stress amplitude σ_a , MPa	RMSE, -
1200	4.89e-4
1000	4.40e-4
800	5.24e-4
520	5.47e-4

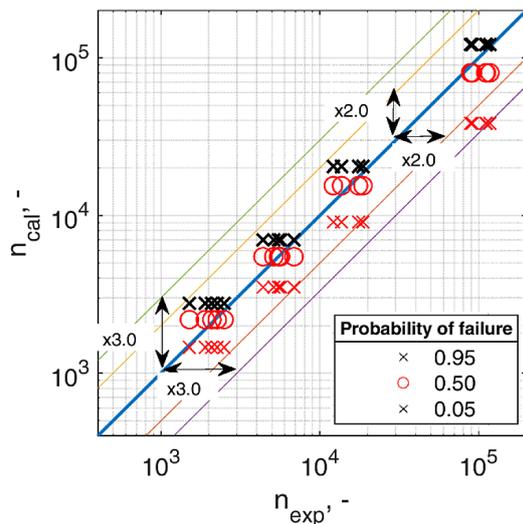


Figure 14. Comparison between fatigue lives: experimental (n_{exp}) and calculated at selected probability levels (n_{cal})

models differ. Although some local differences could be seen, the B-spline PDFs aligned closely with the Weibull PDFs, validating the proposed methodology. Figure 13 illustrates the comparison across different stress amplitudes. The close alignment validates the stability of the proposed algorithm in capturing the stochastic nature of fatigue data. Additionally, Table 6 presents the root mean square error (RMSE) for each stress level, indicating generally good agreement between Weibull and B-spline PDFs, despite some noticeable local discrepancies.

Practical application

Finally, the developed optimized quadratic regression model was applied to generate normalized PDFs at selected HCF stress levels. The PDFs were used to estimate fatigue lives at selected levels of failure probability $P_f \in \{0.95, 0.50, 0.05\}$. Estimated lives were compared to uniformly distributed six-element sample of experimental fatigue lives and plotted in Figure 14.

CONCLUSIONS

In the paper a concept for B-spline based estimation of fatigue life PDFs was investigated. The B-spline interpolation algorithm was applied in order to find a basis function with the corresponding knot vector which would be suitable for deriving the fatigue PDF. Based on the results of fatigue tests of 18Ni300 SLM manufactured maraging steel, knot vectors were found for the tested stress levels and verified with respect to MLE. Based on the obtained results, the following conclusions can be drawn:

1. The b-spline basis functions, with their localized influence, allowed for smooth and precise pdf representation, making them well-suited for fatigue life analysis.
2. Focusing on basis functions instead of direct interpolation allowed to significantly reduce the computational complexity, making the algorithm scalable for larger datasets or additional stress levels.
3. By controlling the number of knots, knot placement and the degree of the basis functions, the pdf of fatigue life random variable can be estimated with respect to different ranges of high-cycle fatigue regime.
4. The use of non-uniform knots allowed to evidence a close alignment with observed mle fatigue life distributions and to capture the variability across stress levels, especially at higher stress amplitudes, where data concentration was higher.
5. By leveraging a combination of kde and optimized quadratic regression, the method could successfully accommodate the stochastic nature of fatigue data, providing precise knot placement and accurate pdf estimation.
6. Satisfying level of control over the final pdf was retained after employing the introduced knot location optimization model.
7. The knot location model made the proposed probability density estimation applicable to stress levels different than the experimental ones.
8. Computational efficiency of the proposed approach could make it a valuable addition to the suite of probabilistic modeling techniques.
9. The proposed approach provides a robust framework for fatigue life prediction which validates the model as having the potential for application in engineering component design.

While the proposed approach demonstrates stability and accuracy for the analyzed stress levels, several limitations constitutes the field for further investigation:

- the analysis is confined to four stress levels. Extending the methodology to a broader range of stress amplitudes or additional materials could further validate the approach.
- the impact of outliers in fatigue data on knot placement and PDF estimation requires additional study.
- the current approach assumes unimodal distributions. Adapting the algorithm to handle multi-modal fatigue data could enhance its applicability.

Acknowledgements

The authors wish to thank Mariusz Prażmowski for performing fractography. This work was financially supported by the Opole University of Technology under GAMMA project no. 155/22.

REFERENCES

1. Castillo, E., Fernández-Canteli, A. A unified statistical methodology for modeling fatigue damage. Springer Netherlands, Dordrecht; 2009. <https://doi.org/10.1007/978-1-4020-9182-7>
2. Schijve, J. (Ed.). Fatigue of structures and materials. Springer Netherlands, Dordrecht; 2009. <https://doi.org/10.1007/978-1-4020-6808-9>
3. Pook, L. Metal fatigue, solid mechanics and its applications. Springer Netherlands, Dordrecht; 2007. <https://doi.org/10.1007/978-1-4020-5597-3>
4. Kufoin, E., Susmel, L. Quantitative review of probabilistic approaches to fatigue design in the medium cycle fatigue regime. Probabilistic Engineering Mechanics. 2024; 75: 103589. <https://doi.org/10.1016/j.probengmech.2024.103589>
5. Tomaszewski, T., Skibicki, A. Probabilistic prediction of the size effect on the fatigue strength for variable length specimens of selective laser melted 316L stainless steel. International Journal of Fatigue. 2024; 182: 108225. <https://doi.org/10.1016/j.ijfatigue.2024.108225>
6. Kowal M., Szala M. Diagnosis of the microstructural and mechanical properties of over century-old steel railway bridge components. Engineering Failure Analysis. 2020 Mar 1; 110: 104447. <https://doi.org/10.1016/j.engfailanal.2020.104447>
7. Rui S.S., Wei S., Sun C. Microstructure evolution, crack initiation and early growth of high-strength martensitic steels subjected to fatigue loading. International Journal of Fatigue. 2024 Nov 1; 188: 108534. <https://doi.org/10.1016/j.ijfatigue.2024.108534>
8. Lehner P., Blinn B., Beck T. Changes in microstructure and mechanical properties of ferritic high chromium steel and P91 induced by isothermal fatigue. Materials Science and Engineering: A. 2025 Feb 1; 923: 147713. <https://doi.org/10.1016/j.msea.2024.147713>
9. Macek W., Martins R.F., Branco R., Marciniak Z., Szala M., Wroński S. Fatigue fracture morphology of AISI H13 steel obtained by additive manufacturing. Int J Fract. 2022 May 1; 235(1): 79–98. <https://doi.org/10.1007/s10704-022-00615-5>
10. Macek W., Sampath D., Pejkowski Ł., Żak K. A brief note on monotonic and fatigue fracture events investigation of thin-walled tubular austenitic steel specimens via fracture surface topography analysis (FRASTA). Engineering Failure Analysis. 2022 Apr 1; 134: 106048. <https://doi.org/10.1016/j.engfailanal.2022.106048>
11. Macek W., Branco R., de Jesus J., Costa J.D., Zhu S.P., Masoudi Nejad R., et al. Strain energy density and entire fracture surface parameters relationship for LCF life prediction of additively manufactured 18Ni300 steel. International Journal of Damage Mechanics. 2024 Sep 1; 33(9): 725–47. <https://doi.org/10.1177/10567895241245879>
12. Schijve, J. A normal distribution or a weibull distribution for fatigue lives. Fatigue & Fracture of Engineering Materials & Structures. 1993; 16: 851–859. <https://doi.org/10.1111/j.1460-2695.1993.tb00124.x>
13. D’Antuono, P. An analytical relation between the Weibull and Basquin laws for smooth and notched specimens and application to constant amplitude fatigue. Fatigue & Fracture of Engineering Materials & Structures. 2024; 43: 991–1004. <https://doi.org/10.1111/ffe.13175>
14. Blacha, Ł., Karolczuk, A. Validation of the weakest link approach and the proposed Weibull based probability distribution of failure for fatigue design of steel welded joints. Engineering Failure Analysis. 2016; 67: 46–62. <https://doi.org/10.1016/j.engfailanal.2016.05.022>
15. Little, R. Mechanical reliability improvement: Probability and statistics for experimental testing. CRC Press, Boca Raton, USA; 2002. <https://doi.org/10.1201/9780203910993>
16. Banjevic, D. Remaining useful life in theory and practice. Metrika. 2009; 69: 337–349. <https://doi.org/10.1007/s00184-008-0220-5>
17. Blacha, Ł., Derda, Sz. Fatigue life for 18Ni300 steel under uniaxial fully reversed loading, Mendeley Data, V1. 2024. <https://doi.org/10.17632/r8pwh9ygn5.2>

18. Tan, C., Zhou, K., Ma, W., Zhang, P., Liu, M., Kuang, T. Microstructural evolution, nanoprecipitation behavior and mechanical properties of selective laser melted high-performance grade 300 maraging steel. *Materials & Design*. 2017 Nov 15; 134: 23–34. <https://doi.org/10.1016/j.matdes.2017.08.026>
19. Haibach, E. Modified linear damage accumulation hypothesis accounting for a decreasing fatigue strength during increasing fatigue damage (in German). Lab. für Betriebsfestigkeit, LBF, Darmstadt, TM Nr.50; 1970.
20. Weibull, W. A statistical representation of fatigue failures in solids, vol. 27. Transaction of The Royal Institute of Technology, Stockholm; 1949.
21. EOS Maraging Steel MS1 Material Data Sheet. EU - EOS Store. <https://www.eos.info/> website (Accessed: 23.04.2024).
22. ASTM E 466-15. Standard Practice for Conducting Force Controlled Constant Amplitude Axial Fatigue Tests of Metallic Materials; ASTM International.
23. Yu, Y., Lee, B., Cho, Y. Multiaxial fatigue reliability assessment using a differential ant-stigmergy algorithm. *Journal of Mechanical Science and Technology*. 2019; 33: 2093–2099. <https://doi.org/10.1007/s12206-019-0413-z>
24. Olsson, E., Olander, A., Öberg, M. Fatigue of gears in the finite life regime – Experiments and probabilistic modelling. *Engineering Failure Analysis*. 2016; 62: 276–286. <https://doi.org/10.1016/j.engfailanal.2016.01.012>
25. Zhao, Y.-X., Gao, Q., Wang, J.-N. An approach for determining an appropriate assumed distribution of fatigue life under limited data. *Reliability Engineering & System Safety*. 2000; 67: 1–7. [https://doi.org/10.1016/S0951-8320\(99\)00036-8](https://doi.org/10.1016/S0951-8320(99)00036-8)
26. Gofuku, S., Tamura, S., Maekawa, T. Point-tangent/point-normal B-spline curve interpolation by geometric algorithms. *Computer-Aided Design*. 2009; 41: 412–422. <https://doi.org/10.1016/j.cad.2009.02.005>
27. Zhao, Y., Zhang, M., Ni, Q., Wang, X. Adaptive Nonparametric Density Estimation with B-Spline Bases. *Mathematics*. 2023; 11. <https://doi.org/10.3390/math11020291>
28. Wang, X., Zhao, Y., Ni, Q., Tang, S. Nonparametric density estimation with nonuniform B-spline bases. *Journal of Computational and Applied Mathematics*. 2024; 440: 115648. <https://doi.org/10.1016/j.cam.2023.115648>
29. Blacha, Ł. A B-spline based density estimation of fatigue failures in SLM 18Ni300 steel. In: *Engineering Mechanics*, 2024; 30, Fuis, V., Hajek, P. (ed.); Brno University of Technology, Institute of Solid Mechanics, Mechatronics and Biomechanics. <https://doi.org/10.21495/em2024-050>