POLITECNICO DI TORINO
Repository ISTITUZIONALE

Thermographic and Machine Learning approaches for a rapid estimation of gears bending fatigue strength

Original
Thermographic and Machine Learning approaches for a rapid estimation of gears bending fatigue strength / Corsaro, L., Dehghanpour Abyaneh, M., Cura, F.M., Sesana, R.. - In: FORSCHUNG IM INGENIEURWESEN-ENGINEERING RESEARCH. - ISSN 0015-7899. - ELETTRONICO. - 89:1(2025), pp. 1-13. [10.1007/s10010-025-00863-6]

Availability:
This version is available at: 11583/3002699 since: 2025-09-01T14:33:02Z

Publisher:
Springer

Published
DOI:10.1007/s10010-025-00863-6

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

08 November 2026

Forschung im Ingenieurwesen (2025)89:122
https://doi.org/10.1007/s10010-025-00863-6

ORIGINALARBEITEN/ORIGINALS

Thermographic and Machine Learning approaches for a rapid estimation of gears bending fatigue strength

Luca Corsaro[1] · Mohsen Dehghanpour Abyaneh[1] · Francesca Curà[1] · Raffaella Sesana[1]

Received: 20 March 2025 / Accepted: 23 July 2025
© The Author(s) 2025

Abstract

The estimation of the tooth root bending fatigue strength of gears is a topic of great interest in the field of mechanical engineering. The assessment of this mechanical property is generally conducted through the execution of a series of tests and, in many cases, a long-time experimental campaign is necessary for the bending fatigue strength evaluation. The present study aims at the estimation of the bending fatigue strength in gears by using the well-known Thermographic Method with integrated Machine Learning techniques implementing Gaussian process regression and artificial neural networks. This approach allows for the combination of a Non-Destructive, green technique with Artificial Intelligence algorithms, determining a rapid and reasonable estimation of the bending fatigue strength for gears. Among all methods, the statistical analyses confirm that all models have high accuracy. However, Gaussian process regression and deep neural networks may be superior in comparison with other methods, and their precision and reliability may be higher for advanced fatigue assessment. This tool could be helpful to cut down experimental workload with the help of Thermographic Method for the tooth root bending fatigue strength estimation, hence enabling very fast Non-Destructive evaluation of gear performance. Thermography approach combined with Machine Learning agrees sustainability by saving critical resource-intensive testing and leads to an advanced mechanical properties evaluation framework in gear systems, hence offering important alternative to the classical methods.

1 Introduction

The application of Passive Thermography (PT) as Non-Destructive Testing (NDT) technique presents numerous advantages, including non-contact, full field measurements and reduced testing time. In the specific context of mechanical research, several studies were conducted by using infrared (IR) detectors to investigate fatigue phenomena. The feasibility of these detectors for temperature analysis in fatigue tests was initiated around the 1970's [1] and, ten years later, the topic of the fatigue limit estimation attracted significant interest. The earliest documented application of the Thermographic Method can be found in [2, 3], where the One Curve Method (OCM) was proposed as a novel approach for a rapid fatigue limit estimation through the utilisation of PT. The papers identified an empirical linear relation between stabilisation temperatures and the applied stress amplitudes, allowing for the fatigue limit estimation. Subsequently, various papers were published with the aim of improving the OCM methodology, as an example reported in [4, 5]. As a matter of fact, an accurate analysis of the intrinsic dissipations generated during the fatigue test was found to be crucial in identifying the fatigue limit. A further analysis was proposed in [6], where the estimation was obtained from an iterative process that examined both low and high loads. More in detail, the fatigue limit was identified using the intersection of two linear curves approximating the surface temperature increment, starting from an initial guess of the fatigue limit. This procedure is known as the Two Curves Method (TCM). Furthermore, other thermal parameters were presented for the purpose of analysing the fatigue limit, which were derived from temperature profiles recorded during fatigue tests [7]. In a recent article [8], the focus of the Thermographic Method moved towards its application in the analysis of a mechanical component instead of classical samples. At the end, the results regarding the fatigue resistance by using the Thermographic approach were in good agreement with those provided from consolidated approaches. So, the Thermo-

All authors approved publication of the manuscript.

✉ Luca Corsaro
luca.corsaro@polito.it

[1] Department of Mechanical and Aerospace Engineering (DIMEAS), Politecnico Di Torino, 10129 Torino, Italy

Published online: 01 September 2025

graphic Method offers a rapid and economical alternative to conventional methods, such as the Staircase analysis, for the fatigue life estimation, although its primary application remains the analysis of classical samples.

Over the last years, the advent of Artificial Intelligence (AI) techniques, such as Machine Learning (ML), appears to be a valid approach in the mechanical field, specifically in the context of fatigue life prediction. Examples of this approach can be found in the literature, such as in [9, 10], where ML were widely used to predict the fatigue life of materials, extend fatigue datasets and to identify significant factors for improving prediction accuracy. One of the most significant advantages of this approach is the possibility of optimising computational time and cost, as demonstrated in the work proposed in [11]. More in detail, the authors proposed a simplified model for fatigue life prediction without compromising accuracy. Another significant application of ML is evidenced in the S-N curve prediction, as illustrated in [12] for aluminium alloys or in [13] for high-strength steel. The utilisation of ML for the fatigue life prediction was further explored for particular materials, such as additively manufactured metals [14], with a focus on investigating the impact of defect location, size and morphology on the fatigue life. Moreover, the possibilities offered by ML techniques in terms of the estimation of low-cycle fatigue life were also investigated for 316 stainless steel in [15]. Even though the low-cycle fatigue life of stainless steel is affected by many factors, and the relationship between these factors and fatigue life is complicated and nonlinear, a ML algorithm was proposed for this purpose. On the basis of these considerations, it is evident that the subject of fatigue life prediction is of significant interest from a variety of viewpoints. As a matter of fact, the demand for more accurate results is combined with the need for rapid and low-cost analyses.

In this work, the Thermographic Method was applied to a mechanical component such as gears with the aim of performing a rapid and economical estimation of the tooth root bending fatigue strength. The proposed fatigue life prediction was improved with ML techniques such as Gaussian Process Regression (GRP), Artificial Neural Networks (ANN) and Support Vector Regressor (SVR), which also allowed in minimizing the number of experimental tests. The proposed methodology, specifically the combination of Thermographic and ML approaches, made it possible to determine the tooth root bending fatigue strength using only one gear with a significant reduction in testing time and number of samples compared to the classic Staircase approach, providing comparable results.

2 Fatigue limit estimation with the Thermographic Method

In this activity, the Thermographic approach adopted for the evaluation of the tooth root bending fatigue strength is the "Two Curves Method" [6, 7, 16], in the specific case for the estimation of the endurable pulsating force ($F_{pn\sim}$).

Figure 1 aims to give a brief description of the Thermographic Method and how this methodology was properly adapted in case of gears. Figure 1a shows a generic temperature profile recorded with an IR camera during a fatigue test. The thermal evolution is typically characterized by an initial heating followed by a stabilized trend. The observed stabilization trend is indicative of a steady-state thermal equilibrium, which is typically exhibited by samples due to their limited volume involved in thermal exchange with the

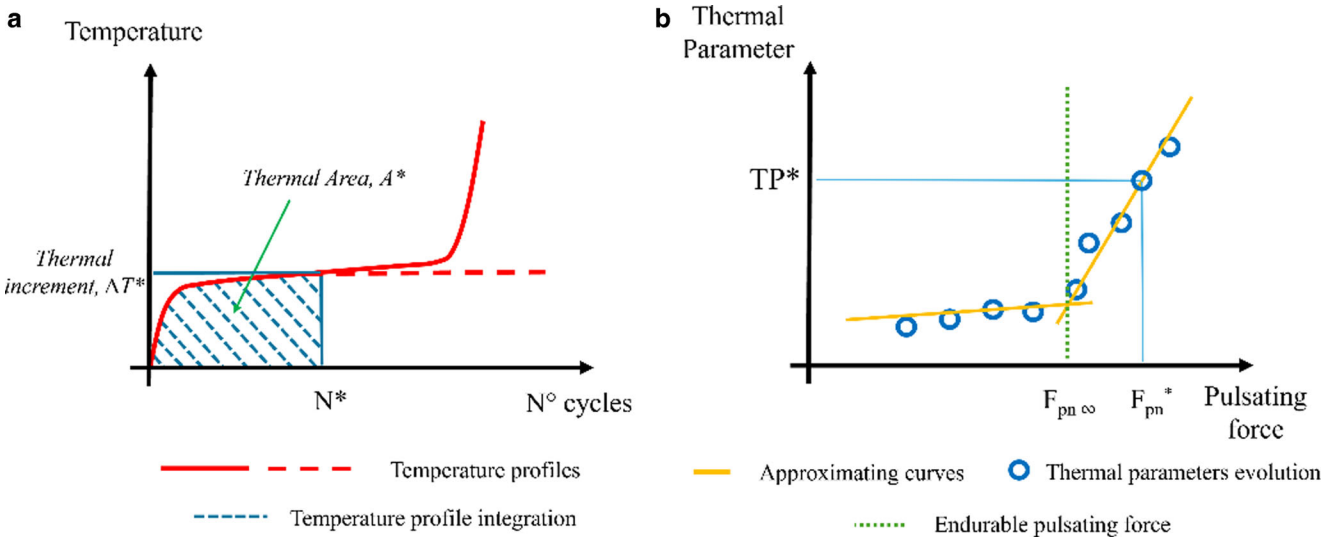

Fig. 1 Thermographic method: **a** thermal profile, **b** fatigue limit estimation

surrounding environment and the gripping mechanism of the testing apparatus. Additionally, the number of cycles at which the temperature reaches a stabilized trend is also related to the testing frequency. Then, if the load value applied during the fatigue test is higher than the corresponding endurance limit, an abrupt thermal increment is observed immediately before the failure of the material (solid red line). On the other hand, for load value below the fatigue limit, the thermal profile remains stabilized (dashed red line). So, from a temperature profile generated for a given load value, specific thermal parameters can be defined for the purpose of fatigue limit estimation. In particular, the temperature increment (thermal increment, ΔT^*), or the area subtended to the temperature profile (thermal area, A*), can be considered for a given time (the number of cycles, N*).

The fatigue limit estimation is performed according to Fig. 1b. In this process, a series of loading conditions are examined, starting with a low load level and increasing the load when a specified number of cycles is recorded with the IR camera. Before the application of another loading condition, particular attention must be paid to the necessity of cooling the sample. This is a crucial step in order to accurately detect the intrinsic dissipations that are produced during the execution of the fatigue test. The procedure concludes when an abrupt thermal parameter increment (illustrated by the blue circles in Fig. 1b) is detected for a certain number of load levels. So, a flat trend is obtained for low load values and, from a certain level, a change in the evolution of the thermal parameters is observed. The two different evolutions of the thermal parameters enable the determination of a transition zone in which the fatigue limit can be estimated. One advantage of the proposed methodology (TCM) is that the thermal parameters necessary for the intrinsic dissipation evolutions, and the corresponding fatigue limit estimation, can be evaluated directly from temperature profiles without the necessity of a steady-state evolution. This is due to the fact that the thermal parameters obtained from different stress amplitudes are investigated over the same number of loading cycles.

During the years, a number of combinations of approximating curves were proposed in literature. These combinations included linear approximating curves, as well as combinations of linear, parabola and power law approximating curves. The reason for these possibilities in the approximating curves is attributable to the intrinsic dissipations evolution that can occur during the fatigue damage process. In general, if the approximating curves provide variability in the endurance limit values, the choice of the fatigue limit is based on the most conservative estimation.

In the specific case of tooth root bending fatigue strength estimation, presented in this work, the TCM methodology was adapted on the basis of the nominal pulsating force rather than the amplitude stress (F_{pn} as reported in Fig. 1b), as specifically adopted in case of fatigue limit estimation for classical samples. Concerning the thermal parameter (TP* as reported in Fig. 1b), the thermal profiles from both the lateral faces of the tested gear were integrated over time, resulting in integrated areas of the lateral faces for each specific loading condition. Then, this approach allows for the estimation of the $F_{pn\sim}$, providing a rapid and non-destructive procedure as an alternative to that estimated by means of the Staircase methodology.

3 Materials and methods

3.1 Samples

The tooth root bending fatigue strength was estimated for two spur gears of identical geometry, manufactured from 20MnCr5 and 18NiCrMo5 steels, which were subjected to a carburizing surface treatment.

The main parameters of the tested gears are listed in Table 1. More in detail, the geometrical data are reported on the left column, where both geometrical (normal modulus, m_n, number of teeth, z, pressure angle, α, face width, b, shift profile, x) and mechanical (tooth root roughness, R_z, and tooth root hardness, HV) parameters are illustrated. The hardness analysis was conducted on the external surfaces of the gear with the objective of verifying the uniformity of the surface treatment. This way, the endurable pulsating force estimated with the proposed method by using just one gear is representative of the investigated surface treatment and gear geometry. Moreover, the ISO 6336 Standard coefficients [17, 18], properly evaluated for the tooth root bending fatigue strength computation, are presented on the right side of Table 1. The gears examined in this study were classified as Eh typology. With regard to the material quality, the ML grade was selected since not all the requirement indications provided in [18] were available for verification. These classifications permitted an assessment of the minimum load for the Thermographic Method starting from an indication of the endurable pulsating force obtained from computations (see Sect. 3.2). The obtained hardness measurements revealed values that exceeded the maximum recommended value of [18]. Consequently, the maximum value presented in [18] was adopted for the computations.

The tooth root bending fatigue strength of each gear was also evaluated with the Staircase methodology in order to provide a comparison with the proposed approach. The endurable pulsating force was found to be the same for the two tested gears, 14.60 kN with a standard deviation of 1.06 kN respectively. As a result of the Staircase, the obtained value was considered with a 50% of failure probability. The subsequent step was to calculate the corresponding stress (permissible bending stress). This was achieved by using the

Table 1 Gear parameters.

Geometrical and mechanical data		ISO 6336 coefficients	
Parameter	Value	Parameter	Value
Normal modulus, m_n	2.2 mm	Y_s	1.59
N° of teeth, z	30	Y_f	1.75
Pressure angle, α	17°	$Y_{\delta relT}$	0.99
Face width, b	20 mm	Y_{RrelT}	1.05
Shift profile, x	0.25 mm	Σ_{flim}	312 MPa
Tooth root roughness, Rz	4.3	–	–
Tooth root hardness, HV	970	–	–

equation proposed in ISO 6336—Part 3 (Method B) (see ISO 6336 coefficients in Table 1; [17]). This computation was also corrected considering 1% of failure probability and with a factor of 0.9 adopted in case of pulsator tests [19]. At the end, a result of 683 MPa was obtained for both the gears.

In order to measure the thermal emissions during the bending fatigue tests, the gear was carefully painted with black paint, which allows to set a known emissivity value of 0.95 according to the value suggested in the reference tables [20].

3.2 Thermal signals processing

The thermal profiles used to estimate the tooth root bending fatigue strength were acquired by using two IR cameras. The bending fatigue tests were carried out using the Single Tooth Bending Fatigue (STBF) configuration, and a dedicated equipment was developed to perform bending fatigue tests at high testing frequency (174 Hz) with a mechanical pulsator.

Figure 2 shows the experimental setup adopted at the mechanical pulsator. In particular, the thermal emissions detection was performed by means of two FLIR A700 IR cameras (see Fig. 2 n°1), located at the opposite site. The aim was to acquire thermal emissions produced by both lateral faces of the gear (see Fig. 2 n°2) under testing. Furthermore, the environmental temperature was monitored during the entire fatigue test using thermal emissions data acquired from the reference sample (see Fig. 2 n°3).

The thermal signals processing, which was crucial for the thermal parameters computation, was performed from both lateral faces of the upper tooth. The temperature evolution during the bending fatigue tests was extracted from the thermogram in a Region Of Interest (ROI, see Fig. 2 n°4) located at the tensile tooth root near the 30° tangent lines [17]. The temperature evolution produced from the reference sample (see Fig. 2 n°3) was also processed in order to consider the relative temperature of the tensile tooth root as a temperature profile.

The thermal emissions of the tested gears were acquired for different loads, and a total number of cycles equal to 100,000 was considered during each fatigue test. In the specific case of application of the Thermographic Method to a mechanical component, it is difficult to achieve a steady-state thermal profile for a low number of cycles, as is generally the case for samples (see Sect. 2). As a matter of fact, the gear has a significant volume which results in an increased exchange of heat with the environment, and the presence of the equipment influences the thermal equilibrium with the grip of the testing machine. In any case, the proposed thermal parameters (the integrated area calculated from both lateral faces of the gears, see Sect. 2) enabled the analysis of the intrinsic dissipation evolution and the estimation of the endurable pulsating force by using the Two Curves Method, taking into account solely the thermal emissions produced for 100,000 cycles. In order to define the nominal pulsating forces to be used in the proposed Thermographic approach, the lowest value of the loading had to be determined. This was based on an estimation of the endurable pulsating force obtained from ISO 6336 Part 3 (Method B) and Part 5 computations (see ISO 6336 coefficients in Table 1; [17, 18]). In particular, the obtained value was subjected to further correction, which involved

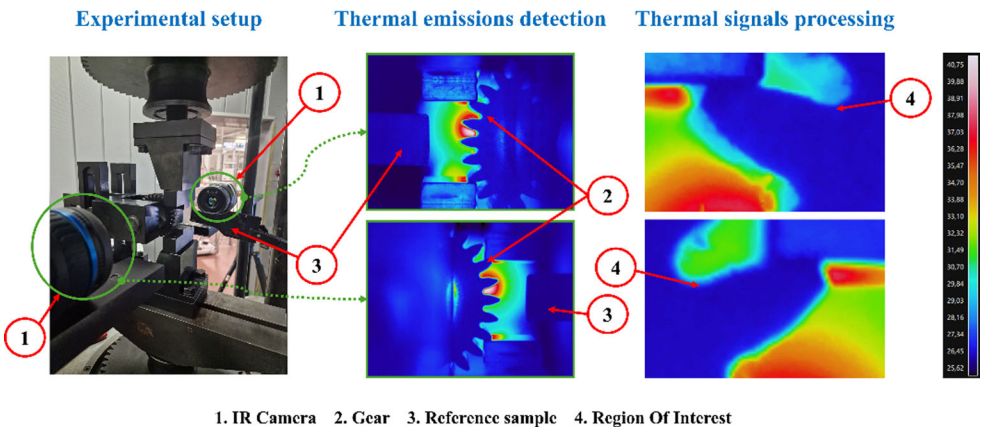

Fig. 2 Experimental setup at the mechanical pulsator

the transformation of the result to the real testing conditions of the experimental activity, in terms of 50% of failure probability, and with consideration of the STBF testing conditions by using the 0.9 coefficient [19]. The final result of this calculation resulted in an indicative value of 13.86 kN, which was adopted for the experimental plan definition at the mechanical pulsator. It is important to note that this value is intended as an indication, selected for the purpose of identifying the lowest nominal pulsating force for the application of the proposed Thermographic Method starting from an indication of the endurable pulsating force evaluated from computations. So, the experimental campaign for the endurable pulsating force estimation was carried out as follows. An initial load below the aforementioned estimated force (13.86 kN from ISO 6336 Standards computations [17, 18] and corrections [19]) was considered. Then, the loads were increased until an abrupt increment in the thermal parameters evolution was detected. In conclusion, a total of seven nominal pulsator forces were considered for each gear, according to the possible gear tooth combination available from the gear geometry. In this activity, the endurable pulsating forces were estimated from the thermal parameters evolution by using three possible combinations of approximating curves: linear-linear, linear-power law and parabola-power law.

At the end, the corresponding tooth root bending fatigue strength was calculated, starting from the endurable pulsating force and by using ISO 6336—Part 3 (Method B) computations [17]. The value was also corrected considering the pulsator test and the failure probability [19]. The computation was performed for the endurable pulsating forces estimated from each combination of approximating curves adopted in this work, and a mean value was adopted as an indication of the tooth root bending fatigue strength of the tested gear. The stress calculation was also performed from the endurable pulsating force obtained from the Staircase methodology in order to compare the results obtained with the proposed approach.

3.3 Machine Learning implementation

Essential elements of material science are material analysis and design; without any of these, long-term success is scarcely possible. Fortunately, with current development, experimental data is starting to flow in by volume, and major characteristics of materials are found in enormous databases. When utilized sensibly with the ML solution, such massive datasets show enormous potential for enhanced material manufacture and general efficiency [21].

3.3.1 Supported Vector Regression (SVR)

Support Vector Machines (SVMs) have changed categorization in areas related to data analysis, although their initial architecture was inadequate for regression problems. Support Vector Regression (SVR) was designed to expand the SVM framework for continuous output evaluations, enabling the resolution of intricate regression problems. SVR has essential hyperparameters, including the epsilon-insensitive zone, which permits minor deviations, function approximation for mapping inputs to higher-dimensional spaces, and the application of kernel functions (such as linear, polynomial, and radial basis function kernels). These hyperparameters are used to find both linear and non-linear relationships. The technique also works based on support vectors and optimizes the regression function to minimize error, which is used by the other hyperparameters such as C and epsilon adjusted to balance model complexity and error tolerance (Fig. 3; [22, 23]).

SVR is a nonlinear issue handling extension of linear regression. It exploits kernel functions to convert data into a higher-dimensional space, therefore facilitating linear separation. Aiming to obtain the flattest feasible function in the feature space for best regression performance, Fig. 3 shows the SVR interpolation method (Eq. 1; [24]).

$$y\left(x\right) = f\left(x\right) = \sum_{i=1}^{N} \left(\alpha_i^* - \alpha_i\right) k\left(x_i, x\right) + b \tag{1}$$

α_i^* and α_i represent Lagrange multipliers. The kernel function is defined by a linear dot product of the nonlinear mapping $k(x_i, x)$. The coefficients α_i^* and α_i in Eq. 4 are produced by minimizing the following regularized risk functional.

$$R_{reg}\left[f\right] = \frac{1}{2}\left|\left|\omega\right|\right|^2 + C\sum_{i=1}^{l} L_{\varepsilon}\left(y\right) \tag{2}$$

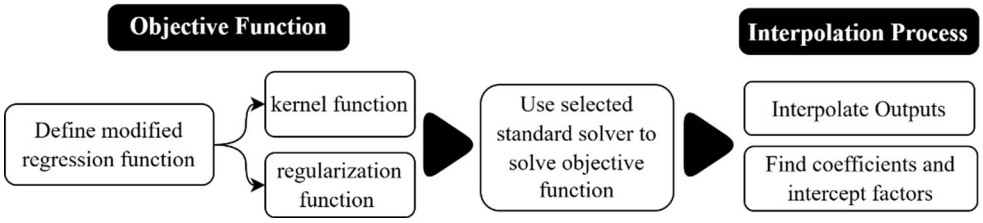

Fig. 3 Interpolation process using an SVR method to produce interpolated outputs

Table 2 A common SVR kernel functions.

Kernel Function Type	Typical Formula	Description
Linear	$K\left(x_i, x_j\right) = \left(x_i^T x_j\right)$	–
Polynomial degree 3 (Cubic)	$K\left(x_i, x_j\right) = \left[\left(x_i^T x_j\right) + c\right]^3$	c Is set to 1 in most implementations
Gaussian	$K\left(x_i, x_j\right) = \exp\left(-\frac{\left\|x_i - x_j\right\|^2}{2\sigma^2}\right)$	The definition of σ is kernel width. It regulates the Gaussian function's width, influencing the decision boundary's flexibility and smoothness

The trade-off is defined by the constant (C), whereas the quantity ω describes the model complexity. The following definition relates to the ε-insensitive loss function, denoted as $L_\varepsilon(y)$ [25].

$$L_\varepsilon\left(y\right) = \begin{cases} 0, & \left|f\left(x\right) - y\right| - \varepsilon < 0 \\ \left|f\left(x\right) - y\right| - \varepsilon, & \left|f\left(x\right) - y\right| - \varepsilon \geq 0 \end{cases} \quad (3)$$

SVR is particularly useful at interpolation areas, which enables it to forecast values within the confines of existing data points, especially when the data displays complex patterns. SVR's kernel technique makes it a flexible and reliable way to interpolate, modelling data well even when it is scattered in an uneven or sparse way. It is very important to choose the right kernel code. As shown in Table 2, the linear, polynomial, and Gaussian (RBF) kernels are the most common ones used for SVR. However, some materials list up to 12 different kernel functions [26–35].

3.3.2 Gaussian Process Regression (GPR)

Gaussian Process Regression (GPR) is a powerful, non-parametric Bayesian method for regression problems. It is reported that this method is more useful for both exploration and exploitation. When combined with Bayesian optimization and hyperparameter tuning, GPR provides probabilistic interpretations and guide the search for the best solution. Function for GPR is defined as the weighted sum of the basic functions [36].

$$y\left(x\right) = f\left(x\right) = \sum_{i=1}^{N} w_i \phi_i\left(x\right) + \sigma_f \epsilon = W^T \phi\left(x\right) + \sigma_f \epsilon \quad (4)$$

It has three parts: W is the weight matrix for the output, $\phi(x)$ is the set of values for the N basis functions at x and ϵ is the white noise model with correlation $\phi(x)$ over noise models.

The kernel function, which is also called the covariance function, is a way to figure out how linear or nonlinear raw data maps to a feature space. The function values for any two inputs are shown below, and two common kernel functions are shown in Table 3. In the end, the probability of a given function ($f(x)$) is supplied, and the optimal weight value is obtained by solving it while accounting for system noise and uncertainties. This model, in contrast to the SVR model, also has a robust solution [37].

$$cov\left(f\left(x\right), f\left(x\prime\right)\right) = \phi\left(x\right)^T \sum W \left(x\prime\right) = k\left(x, x\prime\right) \quad (5)$$

The kernel futures of GPR are flexible and could fit many kinds of data in various situations. Even in cases of com-

Table 3 Frequently used kernel functions for GPR.

Kernel Function Type	Typical Formula	Description
Squared Exponential	$K(x_i, x_j) = \sigma_f^2 \exp\left[-\frac{1}{2}\frac{\left(x_i - x_j\right)^T\left(x_i - x_j\right)}{\sigma_l^2}\right]$	–
Rational Quadratic	$K(x_i, x_j) = \sigma_f^2\left[1 + \frac{\left\|\left(x_i - x_j\right)^T\left(x_i - x_j\right)\right\|}{2\alpha\sigma_l^2}\right]^\alpha$	α Is a scale-mixture parameter with a positive value

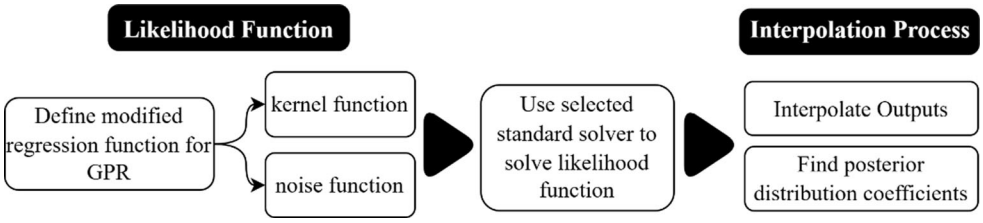

Fig. 4 Interpolation process using an GPR method to produce interpolated outputs

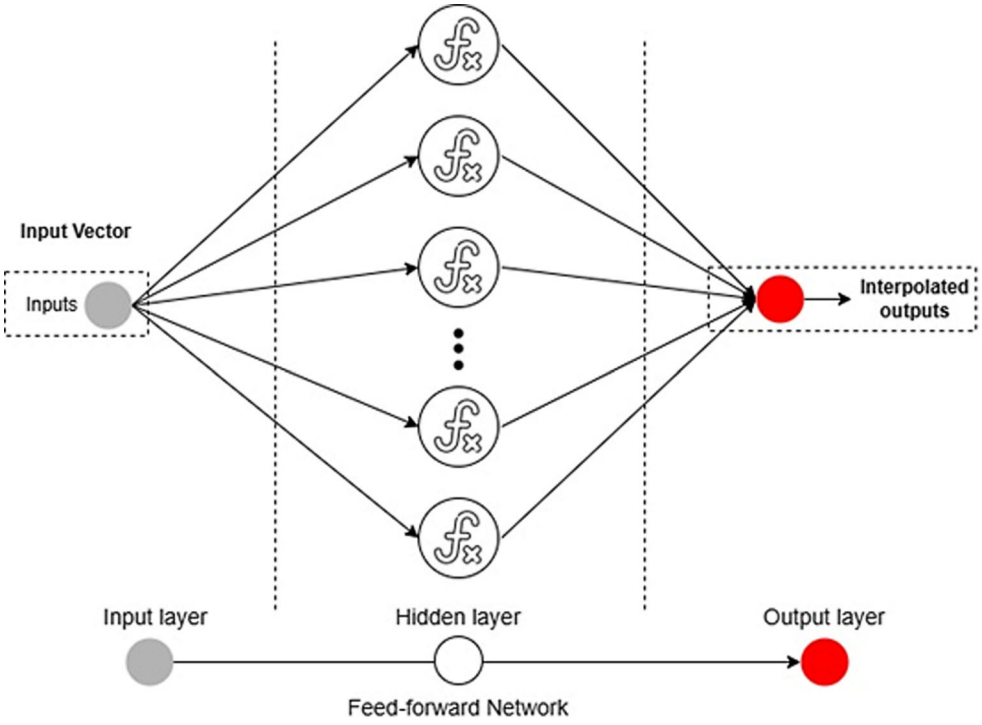

Fig. 5 Feed-forward ANN architecture used for interpolation

plicated or non-linear patterns, it is particularly useful for interpolation tasks, and value prediction within known data points ([37, 38];. Fig. 4) also shows the GPR interpolation process [37].

3.3.3 Artificial Neural Network (ANN)

Training lets a neural network (NN) change the weights and strengths of its connections as it learns. Predicting future results from input data by means of an artificial neural network (ANN) is a formidable instrument. It establishes links between input and output parameters according to prior patterns, as much as human intelligence does [39].

ANNs include many hidden layers and linked neurons to solve problems accurately. This directly depends on the complexity of the problems. Basic ANNs are appropriate for less challenging tasks, which usually have one hidden layer. On the other hand, deep ANNs have many hidden layers (deep architectures) in order to solve complex tasks and interpolations. Therefore, they are very successful for challenging uses such as image and natural language processing. Since ANNs interpolate values within the range of available data points, they are particularly useful for interpolation. This ability is especially important when handling complex or non-linear patterns, as the network can learn underlying trends and provide accurate estimations even in challenging situations ([40];. Fig. 5) shows the architecture of the ANN method for interpolation in this study.

3.3.4 Accuracy metrics

Essential tests in approaches such as regression analysis and ML interpolation models are performance measures, also referred to as evaluation or error metrics. These are the kinds of logical and mathematical instruments. These instruments are meant to evaluate the correctness of expected results by means of comparison with the actual outcomes. The models' interpolations, whether they are ANN, GPR, SVR, or comparable models, cannot always be precise. The purpose of this little error is to avoid overfitting. These models need to be measured in order to compare the performance of various ones and assess their usefulness since each interpolation has some mistakes [41].

Mean Squared Error (MSE), or Mean Squared Deviation (Eq. 6), quantifies the squared discrepancies between actual and projected values. The line of best fit corresponds with the data points. A reduced MSE, approaching zero, signifies enhanced predictive accuracy. A reduced MSE, approaching 0, indicates enhanced predictive accuracy. Root Mean Squared Error (RMSE), also known as Root Mean Squared Deviation (Eq. 7), is the square root of the average of the squared errors. RMSE is the standard deviation of the errors, indicating the degree of alignment between the line of best fit and the data points. Mean Absolute Error (MAE) measures the average absolute deviation between expected and actual values. While MSE indicates the deviation of residuals, MAE offers a clear representation of the average error in the same units as the target variable. This facilitates comprehension and communication (refer to Eq. 8). The

mean absolute percentage error (MAPE) is a performance metric that shows relative error for regression models. It is helpful when sensitivity to relative changes is more important than sensitivity to absolute changes. However, MAPE has several limitations. Two of them are that it can only be used for trustworthy data and often generates lower projections. Because of these limitations, they are less effective for models that are likely to make serious errors [42, 43].

The Coefficient of Determination (R^2) is more complex when applied to nonlinear regression models or nonlinear methods like ANN-based interpolations and ML methods such as GPR and SVR. In certain cases, R^2 can result in the normal 0 to 100% range and even assume negative values (Eq. 9). A negative R^2 value suggests poor interpolation ability since the model does not match the data correctly. This often is when the model fails to find underlying patterns in the data. This would result in significant inaccuracies in interpolation tasks. On the other hand, highly flexible models sometimes overfit training data. Overfitting would result in a low R^2 score on unseen data. In such cases, other evaluation metrics such as MSE, MAE, and RMSE could provide more reliable criteria. This happens frequently when interpolating data does not follow a straightforward trend. Therefore, the models struggle to find the rule behind the input trend, leading to poor R^2 [44].

$$MSE = \frac{1}{n}\sum_{1}^{n}\left(Y_{pre} - Y_{act}\right)^2 \tag{6}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{1}^{n}\left(Y_{pre} - Y_{act}\right)^2} \tag{7}$$

$$MAE = \frac{1}{n}\sum_{1}^{n}\left|Y_{pre} - Y_{act}\right| \tag{8}$$

$$R^2 = 1 - \frac{\sum_{1}^{n}\left(Y_{pre} - \overline{Y}_{pre}\right)^2}{\sum_{1}^{n}\left(Y_{act} - \overline{Y}_{act}\right)^2} \tag{9}$$

In the above equations, Y_{act} represents the actual output value from the dataset, Y_{pre} denotes the predicted value generated by any machine learning algorithm, $\overline{Y}_{act}$ signifies the mean of the actual outputs, and $\overline{Y}_{pre}$ indicates the mean of the anticipated outputs [45–47].

4 Results

The results of the proposed approach are presented in this section.

The experimental data is optimized and augmented by AI-driven interpolation techniques to enhance precision and provide a more comprehensive analysis. More specifically, the experimental data set adopted is the evolution of the thermal parameters obtained after testing the gear at seven nominal pulsating forces. The calculation of the thermal parameters (integrated area) was performed from the thermal profiles acquired from both the lateral faces of the gear (see Sect. 2) during each testing condition. Regarding the thermal profiles, the maximum temperature increment that was observed for both gears was approximately 1–1.5 °C in the maximum loading conditions. By interpolating between actual data points, additional points are generated, allowing for deeper analysis. The purpose of this attempt is that the more points available, the more detailed the analysis can be, making it easier to identify critical points and trends in the dataset. Then, the accuracy metrics of ML-based interpolation will be discussed the impact of additional interpolated points, and how the TCM method benefits from the enhanced dataset in improving overall analysis and prediction accuracy.

Figures 6 and 7 show the interpolation results for 20MnCr5 and 18NiCrMo5 gears using three different ML techniques. More specifically, the experimental data reported in the graphs refer to the thermal parameters evolution extracted from both lateral faces of the tested gear.

Table 4 shows the average accuracy metrics of three ML interpolation methods for 20MnCr5 and 18NiCrMo5. The GPR model has an R^2 value of 1 for both gears, meaning

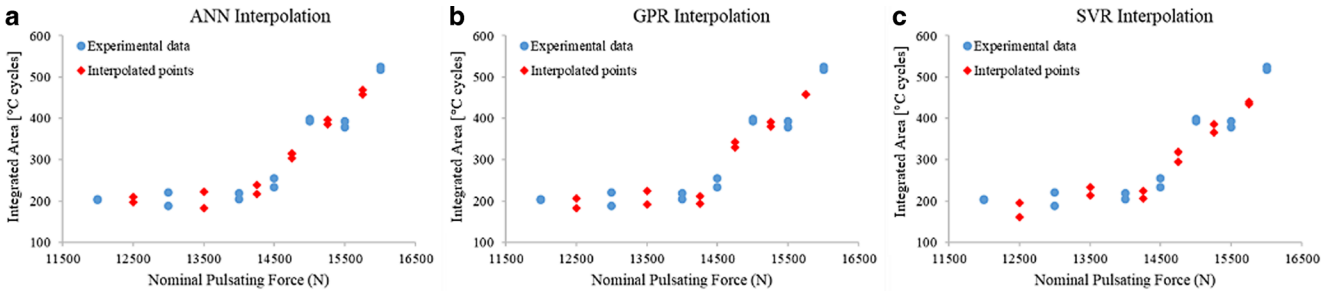

Fig. 6 20MnCr5 interpolation with ML methodology: **a** ANN, **b** GPR and **c** SVR

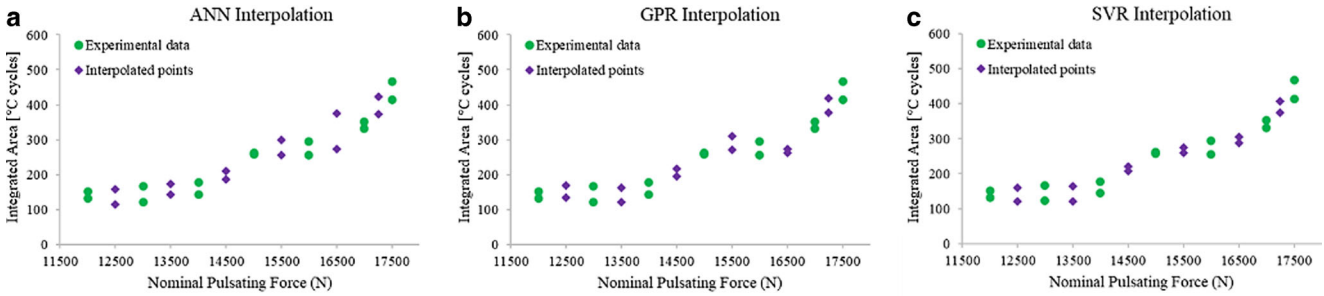

Fig. 7 18NiCrMo5 interpolation with ML methodology: **a** ANN, **b** GPR and **c** SVR

it perfectly fits the data, though this could indicate overfitting. However, based on the figures, GPR still provides reliable interpolation. ANN does better than SVR compared to the other methods, especially for 18NiCrMo5, because it has a higher R^2 and lower error values (MSE, RMSE, and MAE). From this, it is accepted that ANN is the more accurate method for interpolation, especially for 18NiCrMo5, while all of the methods show higher error values for 20MnCr5.

The TCM estimations (the endurable pulsating force) for both the tested gears are reported in Fig. 8. In particular, the results estimated with the original data sets are compared with the results obtained by adopting the enhanced datasets provided by ML techniques (ANN, GPR and SVR). Furthermore, the results obtained by using different approximating curves (linear-linear, linear-power and parabola-power law), adopted for both original and enhanced datasets, are also illustrated. The results derived from the Staircase methodology [40] (illustrated as dashed lines, with mean value and upper and lower limits) are also reported, with the objective of comparing the endurable pulsating forces obtained with the Thermographic Method. On the basis of the obtained results, it can be seen that the TCM with the original data set (white bar charts) is capable of providing comparable results with the Staircase methodology. A variability in the obtained estimations is appreciable on the basis of the approximating curves adopted for the endurable pulsating force estimation. In contrast, the enhanced datasets derived using different ML techniques (grey—ANN, red—GPR and black—SVR bar charts) exhibited similar results in the estimation of endurable pulsating forces, regardless of the approximating curves adopted in TCM. However, it should be noted that the investigated ML technique led to different enhanced data sets (see Fig. 6 and 7), and in one case, the enhanced thermal parameters evolution was found to be unsatisfactory in the estimation of endurable pulsating force (Fig. 8b, Linear-Power approximating curves with dataset optimised with SVR technique). It is necessary to emphasize that the ANN method had the most precise results for both 20MnCr5 (Fig. 6) and 18NiCrMo5 (Fig. 7), slightly better than GPR. However, GPR showed superior accuracy metrics, which suggested a more robust fit to the data, whereas ANN exhibited superior interpolation behaviour. SVR, in contrast, exhibited the lowest accuracy, indicating that it was less effective at interpolation than ANN and GPR respectively. So, the variance in the TCM results (on the basis of the approximating curves adopted for the endurable pulsating force estimation) is a critical factor in this comparison, and it will be further examined in the subsequent section.

Figure 9 presents an example of the procedure for the endurable pulsating force estimation performed with the Thermographic Method. More in detail, the thermal parameters evolution (Integrated Area 100,000 cycles) on the basis of the nominal pulsating force adopted during the fatigue test is reported for the worst-case scenario of variability obtained from the different approximating curves (18NiCrMo5 gear), referring the analysis to the parabola-power law approximating curves. A comparison of the intersection of the approximating curves derived from both the original (red colours) and the enhanced (green colours) data sets using artificial neural networks (ANN) is reported. The enhanced data sets (green cross) are comparable for certain loads with the original ones (red circles) due to the high quality of the ML regression (see Table 4). Then, the intersection obtained from the approximating curves pro-

Table 4 Accuracy metrics for comparing 20MnCr5 and 18NiCrMo5.

	20MnCr5				18NiCrMo5			
	MSE	RMSE	R^2	MAE	MSE	RMSE	R^2	MAE
GPR	6E–06	0.002	1	0.001	5.7 E–06	0.002	1	0.001
ANN	352.554	18.776	0.973	12.746	51.220	7.156	0.995	6.287
SVR	406.315	20.157	0.969	15.741	95.204	9.757	0.991	9.406

duced by the enhanced dataset (green colour) allows for an improvement in the endurable pulsating force estimation as a result of well-defined thermal parameters evolution. In conclusion, the lowest variability in the endurable pulsating force, among the tested approximating curves, was obtained by implementing an ANN as ML technique.

Focusing the attention on the stress (permissible bending stress [17]), the computation was carried out from the endurable pulsating forces estimated with the proposed approach. Box and whisker charts, which show the variance within the dataset related to 20MnCr5 and 18NiCrMo5 gears, are shown in Fig. 10. Results from the Thermographic Method is compared in these charts using both experimental and ML techniques. The computation was performed in accordance with ISO 6336 Part 3 [17] with the endurable pulsating forces reported in Fig. 8. The values were then transformed to take into account the 1% failure probability and the STFB tests [19]. Subsequently, a mean value was calculated between the results produced by applying the different approximating curves (linear-linear, linear-power and parabola-power). The lines in the chart show the distribution and general range of the data, the dot () denotes individual data points, and the cross (×) represents the mean. ML produces more reliable results by lowering variance and limiting the range of outcomes. When choosing the best technique, this decrease in variance is crucial. Ultimately, the benefit of ANN is obvious, compared to other methods.

The optimal values in terms of permissible bending stress that leads to a 1% failure probability results and considering the correction coefficient for STBF tests were obtained with the ANN ML technique. The 20MnCr5 gear resulted in a $\sigma_{FP1\%} = 656$ MPa with a standard deviation of 5 MPa among the three approximate curves combinations. On the other hand, the 18NiCrMo5 gear showed a value of $\sigma_{FP1\%} = 644$ MPa with a standard deviation of 13 MPa. The obtained estimated results were in good agreement with that obtained considering the endurable pulsating force evaluated with the Staircase methodology, 683 MPa for both the gears respectively (see Sect. 3.1). Furthermore, the enhanced data sets were able to reduce the variability of the estimated results compared to those obtained without the ML techniques implementations, providing similar results with different approximation curves.

5 Conclusions

The present study investigates the implementation of ML techniques to improve the estimation of the tooth root bending fatigue strength by using the Thermographic Method, the TCM respectively. The investigation was carried out on a very common mechanical component such as gears.

The TCM results obtained from the original data set were found to be in good agreement with the Staircase methodology. However, it was also observed that results obtained by using different approximation curves can lead to variability in the endurable pulsating force. In this case, in order to estimate the life prediction of the material or component, a conservative value was chosen.

The implementation of ML techniques resulted in the enhancement of the original data sets (thermal parameters evolution), with respect to those obtained from the original experimental campaign. In particular, the optimal accuracy metrics were obtained from the GPR and the ANN ML techniques.

The enhanced thermal parameter data sets were adopted in the TCM routine, which resulted in an enhancement of

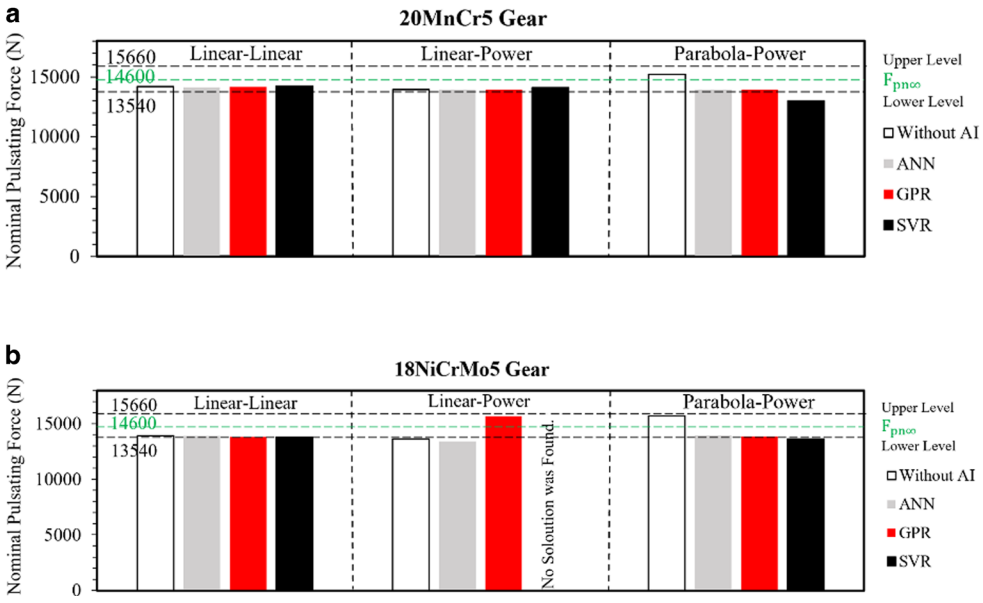

Fig. 8 Comparison of three TCM method data obtained from experiments with interpolation results using three ML methods: **a** 20MnCr5 and **b** 18CrNiMo5 Gears

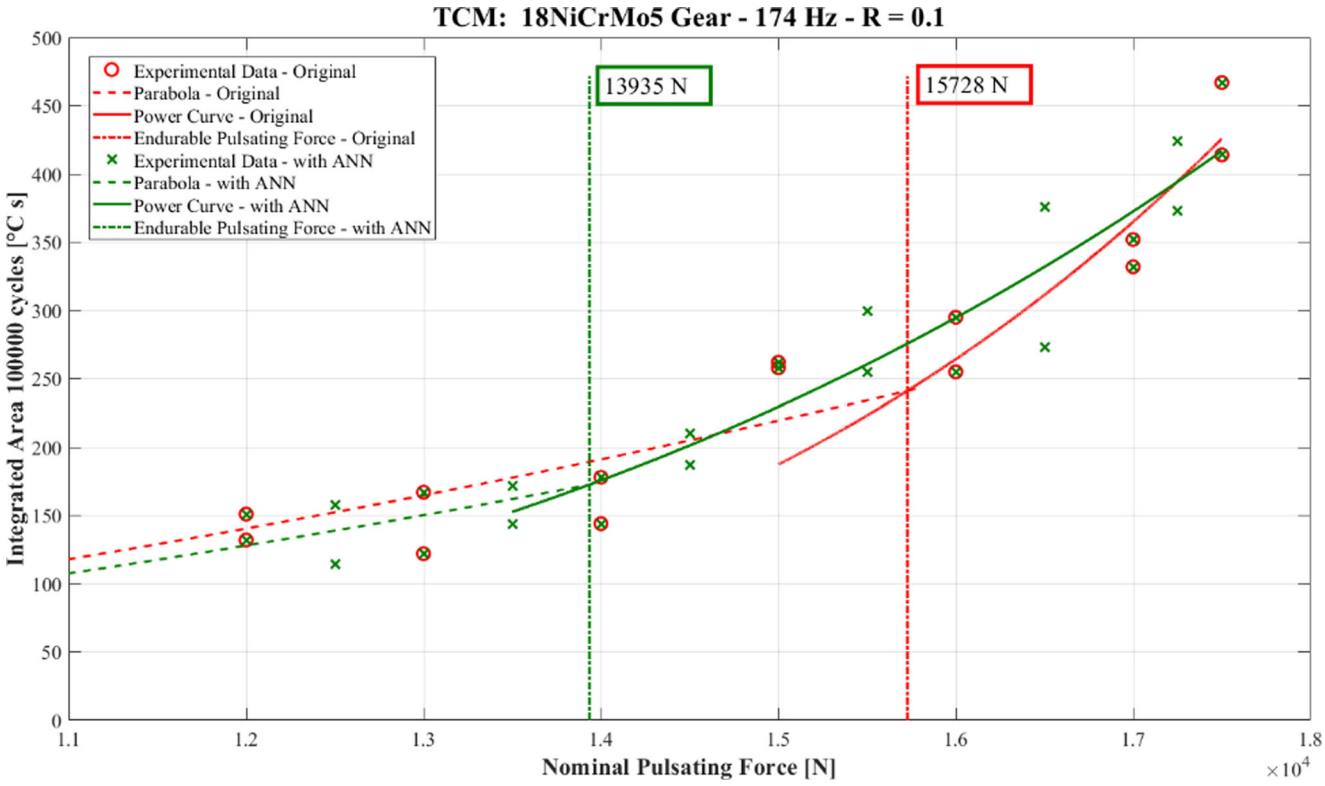

Fig. 9 Endurable pulsating force estimation with the Thermographic Method

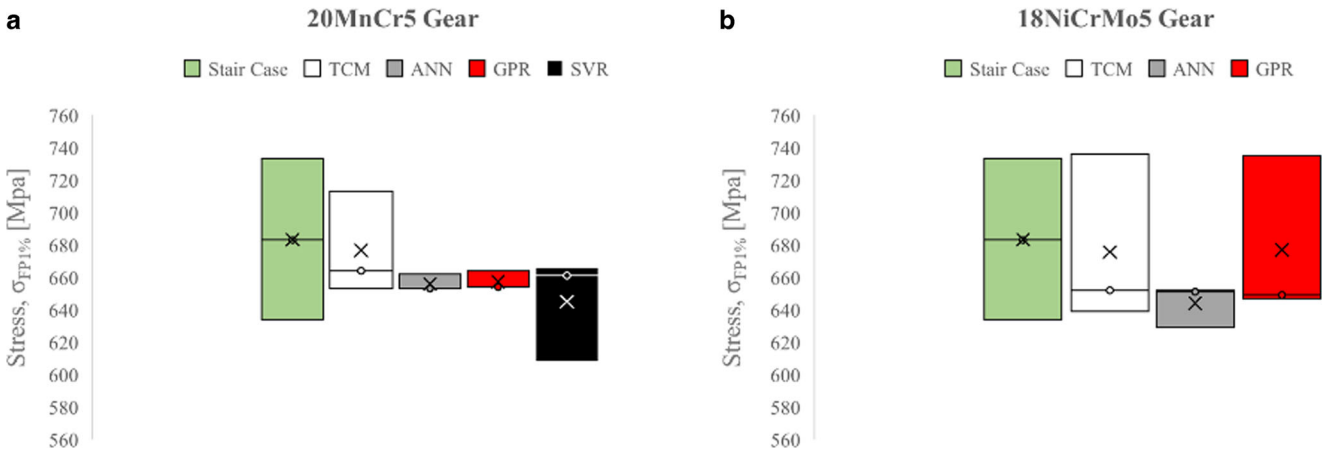

Fig. 10 Box and whisker plot across experimental and ML-based interpolation methods: **a** 20MnCr5 gear and **b** 18NiCrMo5 gear

the obtained results. As a matter of fact, a reduction in the variability of the tooth root bending fatigue strength values was observed among the approximating curves adopted for the analysis.

The Thermographic approach, which facilitates rapid and non-invasive analysis, in conjunction with ML techniques, which reduce testing time, enables for an alternative tooth root bending fatigue strength assessment. Furthermore, the enhanced data sets reduce variability in results among the approximating curves adopted for analysis, leading to more consistent estimations.

Author Contribution All authors contributed to the study. Material preparation, data collection and analysis were performed by Luca Corsaro and Mohsen Dehghanpour Abyaneh. The first draft of the manuscript was written by Luca Corsaro, Mohsen Dehghanpour Abyaneh, Francesca Curà and Raffaella Sesana. All authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding Open access funding provided by Politecnico di Torino within the CRUI-CARE Agreement.

Data availability The data are available on request.

Conflict of interest L. Corsaro, M. Dehghanpour Abyaneh, F. Curà and R. Sesana declarethat they have no competing interests.

Open Access Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden. Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen. Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf http://creativecommons.org/licenses/by/4.0/deed.de.

References

1. Attermo R, Östberg G (1971) Measurements of the temperature rise ahead of a fatigue crack. Int J Fract Mech 7:122–124
2. Curti G, La Rosa G, Orlando M, Risitano A (1986) Analisi tramite infrarosso termico della temperatura limite in prove di fatica. In: Atti XIV Convegno Nazionale AIAS. Catania, pp 211–220
3. Curti G, Geraci A, Risitano A (1989) Un nuovo metodo per la determinazione rapida del limite di fatica. ATA—ingegneria Automob 42(10):634–636
4. La Rosa G, Risitano A (2000) Thermographic methodology for rapid determination of the fatigue limit of materials and mechanical components. Int J Fatigue 22(1):65–73
5. Fargione G, Geraci A, La Rosa G, Risitano A (2002) Rapid Determination of the Fatigue Curve by the Thermographic Method. Int J Fatigue 224(24):11–19
6. Curà F, Curti G, Sesana R (2005) A new iteration method for the Thermographic determination of fatigue limit in steels. Int J Fatigue 27(4):453–459
7. Curà F, Gallinatti AE, Sesana R (2012) Dissipative aspects in Thermographic methods. Fatigue Fract Eng Mater Struct 35(12):1133–1147
8. Faria JJR, Fonseca LGA, De Faria AR, Cantisano A, Cunha TN, Jahed H, Montesano J (2022) Determination of the fatigue behavior of mechanical components through infrared thermography. Eng Fail Anal 134:106018. https://doi.org/10.1016/j.engfailanal.2021.106018
9. Tefera GA, Koricho EG (2023) Prediction of fatigue life of materials using machine learning approach: a review. In: International conference on advances of science and technology. Springer Nature, Switzerland, pp 147–172
10. Karakaş Ö, Berto F, Hong Y (2023) Data science and machine learning for fatigue and fracture assessment. Fatigue Fract Eng Mater Struct 46(6):2031–2033
11. Zhou K, Sun X, Shi S, Song K, Chen X (2021) Machine learning-based genetic feature identification and fatigue life prediction. Fatigue Fract Eng Mater Struct 44(9):2524–2537
12. Ani AS, Deoghare AB (2024) Leveraging machine learning for enhanced fatigue life prediction in aluminum alloys. In: International conference on recent advancements in mechanical engineering. Springer Nature, Singapore, pp 733–743
13. Liu X, Zhang S, Cong T, Zeng F, Wang X, Wang W (2024) Very high-cycle fatigue life prediction of high-strength steel based on machine learning. Fatigue Fract Eng. Mater Struct. https://doi.org/10.1111/ffe.14213
14. Bao H, Wu S, Wu Z, Kang G, Peng X, Withers PJ (2021) A machine-learning fatigue life prediction approach of additively manufactured metals. Eng Fract Mech 242:107508. https://doi.org/10.1016/j.engfracmech.2020.107508
15. Duan H, Cao M, Liu L, Yue S, He H, Zhao Y, Zhang Z, liu Y (2023) Prediction of 316 stainless steel low-cycle fatigue life based on machine learning. Sci Rep 13(1):6753–6753
16. Gallinatti AE (2009) Theoretical and experimental study of the thermal emission in metallic materials subjected to fatigue loads. PhD thesis. Politecnico di Torino
17. ISO 6336 (2019) Calculation of load capacity of spur and helical gears. Part 3: calculation of tooth bending strength. International Standard Organization, Geneva (Switzerland)
18. ISO 6336 (2019) Calculation of load capacity of spur and helical gears. Part 5: calculation of tooth bending strength. International Standard Organization, Geneva (Switzerland)
19. Stahl K (1999) Statistische Methoden zur Beurteilung von Bauteillebensdauer und Zuverlassigkeit und ihre beispielhafte Anwendung auf Zahnrader. FVA-Forschungsvorhaben Nr. 304, Forschungsvereinigung Antriebstechnik e. V., FVA-Heft, Frankfurt, p 580
20. FLIR (2015) A6700sc/A6750sc user's manual. Document number: 29249–000, version: 6
21. Stergiou K, Ntakolia C, Varytis P, Koumoulos E, Karlsson P, Moustakidis S (2023) Enhancing property prediction and process optimization in building materials through machine learning: a review. Comput Mater Sci 220:112031
22. Carrasco M, López J, Maldonado S (2019) Epsilon-nonparallel support vector regression. Appl Intell 49:4223–4236
23. Ghorbani B, Arulrajah A, Narsilio G, Horpibulsuk S (2020) Experimental investigation and modelling the deformation properties of demolition wastes subjected to freeze-thaw cycles using ANN and SVR. Constr Build Mater 258:119688
24. Smola AJ, Schölkopf B (2004) A tutorial on support vector regression. Stat Comput 14:199–222
25. Drucker H, Burges CJ, Kaufman L, Smola A, Vapnik V (1996) Support vector regression machines. Adv Neural Inf Process Syst 9:1–7
26. Sun J, Zhang J, Gu Y, Huang Y, Sun Y, Ma G (2019) Prediction of permeability and unconfined compressive strength of pervious concrete using evolved support vector regression. Constr Build Mater 207:440–449
27. Ma X, Zhang Y, Wang Y (2015) Performance evaluation of kernel functions based on grid search for support vector regression. In: Proceedings of the 2015 IEEE 7th International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM). IEEE, pp 283–288
28. Boolchandani D, Ahmed A, Sahula V (2011) Efficient kernel functions for support vector machine regression model for analog circuits' performance evaluation. Analog Integr Circuits Signal Process 66:117–128
29. Camps-Valls G, Martín-Guerrero JD, Rojo-Álvarez JL, Soria-Olivas E (2004) Fuzzy sigmoid kernel for support vector classifiers. Neurocomputing 62:501–506
30. Hitam NA, Ismail AR, Samsudin R, Alkhammash EH (2022) The effect of kernel functions on cryptocurrency prediction using support vector machines. In: advances on intelligent Informatics and computing. Lect Notes Data Eng Commun Technol 127:319–332
31. Li Y, Wang S, Chen W, Han W (2020) Holistic comparison of different kernel functions for support vector regression based on state-of-health prediction of lithium-ion battery. In: Proc. 11th int. Conf. on Prognostics and system health management (PHM-2020), pp 40–46
32. Lin HT, Lin CJ (2003) A study on sigmoid kernels for SVM and the training of non-PSD kernels by SMO-type methods. Neural Comput 15:1–32

33. Liu H, Liu D (2006) Support vector regression based on fuzzy sigmoid kernel. Control Theory Appl 23:204–208
34. Shao Y, Jiang Q (2024) A new class of Bessel kernel functions for support vector machine. IEEE Access 12:5357–5364
35. Su H, Li X, Yang B, Wen Z (2018) Wavelet support vector machine-based prediction model of dam deformation. Mech Syst Signal Process 110:412–427
36. Kuss M (2006) Gaussian process models for robust regression, classification, and reinforcement learning. Technische Universität Darmstadt, Darmstadt, Germany
37. Schulz E, Speekenbrink M, Krause A (2018) A tutorial on Gaussian process regression: modelling, exploring, and exploiting functions. J Math Psychol 85:1–16
38. Farid M (2022) Data-driven method for real-time prediction and uncertainty quantification of fatigue failure under stochastic loading using artificial neural networks and Gaussian process regression. Int J Fatigue 155:106415
39. Basheer IA, Hajmeer M (2000) Artificial neural networks: fundamentals, computing, design, and application. J Microbiol Methods 43:3–31
40. Opěla P, Schindler I, Kawulok P, Kawulok R, Rusz S, Navratil H (2021) On various multi-layer perceptron and radial basis function based artificial neural networks in the process of a hot flow curve description. J Mater Res Technol 14:1837–1847
41. Javadi MS, Ehteshamfar MV, Adibi H (2023) A comprehensive analysis and prediction of the effect of groove shape and volume fraction of multi-walled carbon nanotubes on the polymer 3D-printed parts in the friction stir welding process. Polym Test 117:107844
42. Ehteshamfar MV, Javadi MS, Adibi H (2022) Surface modification of prototypes in fused deposition modelling using lapping process. RPJ 28:1382–1393
43. Wang N, Wang Z (2025) Integrating an interpolation technique and AI models using Bayesian model averaging to enhance groundwater level monitoring. Earth Sci Inform 18:1–15
44. Kebede H, Demissie Z, Tadesse H, Eshetu A (2024) Spatial interpolation techniques comparison and evaluation: The case of ground-based gravity and elevation datasets of the central Main Ethiopian rift. Heliyon. https://doi.org/10.1016/j.heliyon.2024.e32806
45. Draper NR, Smith H (1998) Applied Regression Analysis, 2nd edn. John Wiley & Sons, New York
46. Chai T, Draxler RR (2014) Root mean square error (RMSE) or mean absolute error (MAE)? – arguments against avoiding RMSE in the literature. Geosci Model Dev 7:1247–1250
47. Dehghanpour Abyaneh M, Narimani P, Javadi MS, Golabchi M, Attarsharghi S, Hadad M (2024) Predicting surface roughness and grinding forces in UNS S34700 steel grinding: A machine learning and genetic algorithm approach to coolant effects. Physchem 4:495–523

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.