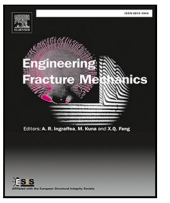
Engineering Fracture Mechanics 292 (2023) 109595

Contents lists available at ScienceDirect

Engineering Fracture Mechanics

journal homepage: www.elsevier.com/locate/engfracmech



Quantification of uncertainty in a defect-based Physics-Informed Neural Network for fatigue evaluation and insights on influencing factors

Emanuele Avoledo, Alessandro Tognan, Enrico Salvati *

Polytechnic Department of Engineering and Architecture, University of Udine, Via delle Scienze 206, Udine, 33100, Italy

ARTICLE INFO

Keywords:
Fatigue
Defects
Physics-Informed Neural Networks
Uncertainty quantification
Sensitivity analysis

ABSTRACT

Substantial advances in fatigue estimation of defective materials can be attained through the employment of a Physics-Informed Neural Network (PINN). The fundamental strength of such a framework is the ability to account for several defect descriptors while maintaining predictions physically sound. The first objective of the present work is the assessment of the PINN estimated fatigue life variability due to uncertainties carried by the inputs. Additionally, a set of sensitivity indices are employed to explore the influence of defect descriptors in fatigue life. The work suggested that some traditionally neglected defect descriptors may play a relevant role under specific circumstances.

1. Introduction

The full exploitation of Additive Manufacturing (AM) for producing near-net shape parts in high performance sectors is typically hampered by several outstanding issues. Amongst these, the difficulty in predicting the mechanical behaviour is one of the key concerns. The lack of well established protocols to control the aspects of AM makes the fabricated material highly inhomogeneous at both the macro and micro-scale, e.g. defects [1,2], residual stress [3–5], texture [6], and others [7,8]. Moreover, the onset of such inhomogeneities is considerably sensitive to the selection of the process parameters, such as laser power, scanning strategy, layer thickness and powder components [9]. The presence of defects, e.g. gas pores, lack-of-fusion (LOF) and key holes, is one of the crucial factors impacting fatigue performance. In this regard, gas droplets can be trapped while melting the powder bed, resulting in the formation of gas pores, which show a distinctive spheroidal shape. Differently, some regions may experience incomplete melting of the powder layer resulting in LOF defects. LOF defects are generally more detrimental to the mechanical properties of AM parts, due to their highly irregular shape, elongation and relatively large extension [10]. In addition, LOF defects typically have a low aspect ratio, resulting in flattened crack-like defect. The origin of this peculiar shape is due to the presence of a preferred growth direction perpendicular to the build direction. Although a defect-free AM material is hardly attainable, thermo-mechanical post-treatments can be used to limit the presence and the influence of defects in AM parts, such as hot isostatic pressing (HIP) [10] and shot-peening [11,12].

The presence of these inhomogeneities is one of the leading causes of the anisotropic behaviour in fatigue performance implying a relevant influence on the build direction, and thus on the loading conditions [1]. Particularly, when AM parts are subjected to fatigue loading, critical stress concentrations may occur in the regions surrounding defects [13,14]. Therefore, these locations represent potential initiation sites for fatigue crack formation and consequent propagation leading to the failure of the structural

* Corresponding author.
E-mail addresses: avoledo.emanuele@spes.uniud.it (E. Avoledo), tognan.alessandro@spes.uniud.it (A. Tognan), enrico.salvati@uniud.it (E. Salvati).

https://doi.org/10.1016/j.engfracmech.2023.109595
Received 30 June 2023; Received in revised form 8 August 2023; Accepted 3 September 2023
Available online 9 September 2023

component [1]. Defects can influence fatigue in many ways, and their size is not the only parameter that defines its influence. In fact, the closer is the defect to the free surface of the sample, the more its inclination to trigger fatigue crack is [15,16].

To deal with the fatigue characterisation in presence of defects in materials, Fracture Mechanics is generally employed. This is justified by the fundamental assumption that defects can be approximated to cracks under certain circumstances. Kitagawa-Takahashi's (KT) diagram is an useful approach to relate the applied cyclic load to a characteristic length of the crack [17], thus allowing to understand whether a specific loading and defect conditions can lead to fatigue failure. A more convenient manner to describe this semi-phenomenological behaviour, can be achieved through El-Haddad (EH) formulation [18], which considers the fatigue stress intensity factor (SIF) threshold and the intrinsic stress range at a given predefined fatigue endurance limit; other models also exist [19]. To deal with irregularly-shaped defects, Murakami proposed its well-renowned parameter to associate defects with their projected area on the plane orthogonal to the loading direction, i.e. $\sqrt{\text{area}}$. Such parameter defines a representative defect's length that can be promptly exploited the restate KT or EH models [20]. Furthermore, the SIF formulated in terms of $\sqrt{\text{area}}$ and distance of the defect from the material free surface has been proven to accurately characterise the severity of a defect given the applied load, as seen by Romano et al. [21]. Interestingly, this approach has been extensively used in numerous scenarios when the defect-based assessment of fatigue performance was sought [22]. A recent study by Niu et al. used extreme value statistics combined with the weakest link theory to derive a defect-based probabilistic framework to estimate the fatigue scatter of AM metallic materials.

Nevertheless, current models based on Fracture Mechanics may restrict the exploitation of other important geometrical and positional characteristics of defects. Future challenges in fatigue prediction of defective materials aim at overcoming this outstanding limitation.

Over the years, Machine Learning (ML) algorithms have been increasingly pursued to study problems related to both finite and infinite fatigue life [23–25]. Compared to semi-empirical or analytical models, ML models have the potential to effectively account for a wide range of physical properties without significantly increasing computational costs. Considerable progress has been achieved aiming to improve the prediction of fatigue behaviour. For instance, Extreme Gradient Boosting proved to be an effective algorithm for predicting the fatigue crack growth rate in Ti6Al4V alloy fabricated via Selective Laser Melting (SLM) [26]. Konda et al. trained a Neural Network (NN) and a Random Forest algorithm to forecast the fatigue life of AlSi10Mg, SS316L and Ti6Al4V additively manufactured via SLM [27]. Li et al. devised a Support Vector Regression (SVR) approach to establish a correlation between the fatigue life of SLM parts with some critical defect features. In this instance, a sensitivity analysis was performed and the correlations between defect descriptors and fatigue life were found to be consistent with both theoretical expectations and experimental observations [28]. Another study correlating defect characteristics such as location, size and morphology with fatigue life was carried out by Bao et al. In this study, a Support Vector Machine was trained with defect characteristics and showed high accuracy in predicting fatigue life [29]. Furthermore, Moon et al. utilised a set of Ti6Al4V additively manufactured with different scanning strategies and different surface conditions. In particular, the data obtained by computed tomography (CT) and surface profilometry were processed to train a drop-out NN which showed high fatigue life prediction accuracy [30]. He et al. compared a $\sqrt{\text{area}}_{\text{eff}}$-based Linear Elastic Fracture Mechanics (LEFM) model with a Random Forest model in order to predict the fatigue life of a martensitic stainless steel and KSFA90 steel alloys. The ML model provided more truthful prediction of the fatigue life than those of the reference theoretical models [31]. Peng et al. predicted the finite fatigue life of AlSi10Mg by using Extreme Gradient Boosting. Therein, fractography and fatigue characterisation data were used to train the surrogate model. Subsequently, the numerical predictions were validated against Murakami's semi-empirical model and an empirical model, referred to as X-parameter model. Further sensitivity analyses permitted $\sqrt{\text{area}}$ to be identified as the most impacting feature on fatigue life [32].

The mere data-driven nature of the ML methods just reviewed allows for neatly recognising patterns provided adequately large datasets – typically not available in the context of fatigue. For this reason, Salvati et al. have recently exploited Physics-Informed Neural Networks (PINNs) to conceive a predictive model to forecast the finite fatigue life of a small batch of AlSi10Mg AM samples [33]. Therein, the PINN was trained on an experimental dataset containing numerous defect descriptors while constraining the learning stage by a LEFM semi-empirical law. As a result, the PINN learned to make predictions balancing the experimental observation and the knowledge inherited from the LEFM model. Besides effectively handling a small-data regime, the PINN held superior predictive capabilities, and a natural tendency at preventing overfitting compared to the sole NN counterpart. It is worthwhile noting that PINNs have been applied to model the finite fatigue life behaviour of AM material using mere S-N data to accomplish the training stage [34], even within a probabilistic setting [35]. He et al. proposed a PINN-based strategy that incorporates a multiaxial fatigue critical plane model for predicting fatigue life [36]. The research showed that properly choosing the theoretical model to support of an NN can enhance the accuracy of the predictions [36]. Interestingly, Ciampaglia et al. [37] devised a PINN to account for AM parameters as well defect features to investigate their joint effect on the finite fatigue response of AlSi10Mg SLM alloy [37]. In this instance, the PINN lead to more accurate predictions compared to the equivalent NN whose physical branch was suppressed. Wang et al. tested two different physics-based ML models, i.e. SVR and NN, for predicting the fatigue life of AM metallic materials. Both models were designed to take a few defect features, such as $\sqrt{\text{area}}$, and the applied stress range as inputs. The case studies showed that the physics-based models outperformed the mere data-driven counterpart and the traditional semi-empirical approaches [24].

In order to make reliable fatigue life predictions, it is important to recognise and quantify the uncertainties associated with the model and data errors. Although diverse sources of uncertainty are typically involved in fatigue characterisation [38], this research focuses on those arising from the measurement of the applied load and defectivity characterisation. In this regard, the present study proposes a systematic method to model the uncertainties derived from the defectivity characterisation in terms of location, geometry and morphology.

The lack of analytical input–output relationship in ML models prevents one from applying the classical Error Propagation Theory (EPT), hence an alternative method must be employed. The present study aims to quantify the uncertainties in fatigue prediction due to the aleatoric errors of the input data and to attempt identifying the most influencing defect descriptors in the authors' previously proposed PINN framework [33]. The aleatoric errors affecting the experimental quantities, which serve as input to the PINN, were obtained using EPT based on carefully selected assumptions to avoid unphysical scenarios. Subsequently, a Monte Carlo (MC) simulation was employed to perform a large number of simulations, by individually perturbing the inputs of the considered PINN. This was carried out to explore the influence of each experimental parameter on the fatigue behaviour using three statistical tools: Pearson correlation coefficient (PCC), Permutation Feature Importance (PFI) and Accumulated Local Effects (ALE). Eventually, in order to assess the overall impact of inputs' uncertainty on the fatigue prediction, in terms of number of cycles to failure, all the inputs were simultaneously and appropriately perturbed. Finally, a χ^2 test was conducted on the resulting MC distributions to evaluate statistical parameters such as mean, standard deviation and the probability density function of the predicted number of cycles to failure.

2. Method and analysis

The present section is meant to cover the fundamental concepts of the PINN model for fatigue prediction and to provide details on the three statistical tools employed in the sensitivity analysis of the present study.

2.1. Physics-informed neural networks

Given their data-driven nature, the predictive performance of NNs is massively conditioned by the processed training dataset. When sufficient data is lacking, or in presence of noisy inputs, NN might learn unintended patterns, thus leading to predictions not obeying with the physics of the observed phenomenon. The realm of ML has recently seen the advent of PINNs which allows for tolerating small-data regime while ensuring the physical consistency of NN models [39,40]. Fundamentally, a PINN couples a conventional NN with a phenomenological law of the underlying – hypothesised – physical phenomenon. In this instance, the training of PINNs entails defining a bespoke loss function consisting of two terms: data loss and model loss. The former is given by the discrepancy between the prediction and the ground truth, whereas the latter quantifies to which extent the predictions differs from the physical model. As a result, PINNs seek a trade-off between these competing terms so that the predictions of given inputs will do.

In order to apply the proposed sensitivity analysis framework, this study utilises the pre-trained PINN that some of the authors of the present manuscript previously conceived [33]. Although the details of this PINN can be found elsewhere, it is worth recalling the key stages of its implementation. Specifically, the PINN was built upon a database borrowed from the literature containing n.12 AlSi10Mg specimens manufactured through SLM [21]. The specimens were also subjected to a fatigue testing with constant stress amplitude at load ratio of $R = -1$. Making use of (CT), many defect characteristics were extracted, amongst which: the distance between defects and the free surface, h, the external defect surface, A_D, the defect volume V_D. The external surface A_D was further processed to compute the projected area on the plane perpendicular to the loading direction, i.e. Murakami's parameter $\sqrt{\text{area}}$. Based on these experimental measurements, relevant descriptors of the morphology and geometry of the defects were derived. The first is sphericity, S, expressed as:

$$S = \frac{\pi^{\frac{1}{3}}(6V_D)^{\frac{2}{3}}}{A_D} \tag{1}$$

and the equivalent three-dimensional diameter, d:

$$d = 2\left(\frac{3V_D}{4\pi}\right)^{\frac{1}{3}} \tag{2}$$

It is essential to report the criterion adopted to categorise defects according to their relative position to the free surface of the corresponding sample [33]:

$$\begin{cases} \text{Surface} & \text{if } h/r \le \pi/4 \\ \text{Sub-surface} & \text{if } \pi/4 < h/r < 1.25 \\ \text{Embedded} & \text{if } h/r \ge 1.25 \end{cases} \tag{3}$$

where r is the 2D equivalent radius of the defect:

$$r = \sqrt{\frac{\text{area}}{\pi}} \tag{4}$$

Once identified and evaluated the key descriptors influencing the fatigue life of each tested sample, i.e. $\Delta\sigma$, $\sqrt{\text{area}}$, h, D, S, these will serve as inputs to the considered PINN model. Additionally, it is important to highlight that only the three most critical defects are involved in the PINN model. The ranking criterion counts the fatigue SIF range related to $\sqrt{\text{area}}$ and normalised by the fatigue SIF threshold obtained with Murakami and Endo's formulation [20], called a *normalised SIF range* δK [33]. Accordingly, the input vector associated with the i-th sample can be formalised as follows:

$$\mathbf{x}^{(i)} = \left[\sqrt{\text{area}}_1^{(i)}\ \sqrt{\text{area}}_2^{(i)}\ \sqrt{\text{area}}_3^{(i)}\ h_1^{(i)}\ h_2^{(i)}\ h_3^{(i)}\ S_1^{(i)}\ S_2^{(i)}\ S_3^{(i)}\ d_1^{(i)}\ d_2^{(i)}\ d_3^{(i)}\ \Delta\sigma^{(i)}\right] \tag{5}$$

Indicating the PINN by $\mathcal{H}$, one can succinctly states its functional input–output relationship:

$$N = \mathcal{H}(\mathbf{x}) \tag{6}$$

The PINN model was then architected by prescribing n.13 input neurons (n.12 related to the defect morphology and one referring to the loading condition), n.16 and n.8 neurons over the first and second hidden layer, respectively. Whilst, the output layer is scalar as it provides the forecast finite fatigue life. Sigmoid activation functions were globally adopted.

As far as the physical constraint of the PINN is concerned, a customised Basquin's-like curve was adopted:

$$N = A\,\delta K^{B} \tag{7}$$

where δK is the aforementioned normalised SIF, and N is the predicted number of cycles to failure. It should be highlighted that A and B in Eq. (7) were identified through Ordinary Least Squares method upon post-mortem fractographic investigation of the defect that triggered fatigue failure, i.e. the "killer defects", where identifiable. Conversely, the PINN processes defect data obtained from ante-mortem CT scans and the applied stress range. As quickly mentioned earlier, only the three most critical defects, sorted by δK, of each sample are used to train the PINN. These defects are deemed as "potential killer defects".

2.2. Uncertainty quantification of the PINN prediction

2.2.1. Error propagation theory

The EPT is briefly described here. Let y depends on $q_i\ \forall\, i = 1, 2, \dots, k$ uncertain experimental measurements:

$$y = f(q_1, \dots, q_k) \tag{8}$$

Assume that each q_i is associated with a measurement error δq_i. As common practice, one can compute the propagated uncertainty, namely δy, derived from $q_i\ \forall\, i = 1, 2, \dots, k$ taking the square root of the total derivative of y:

$$\delta y = \sqrt{\frac{\partial y(q_1, \dots, q_k)}{\partial q_1}\delta q_1^2 + \cdots + \frac{\partial y(q_1, \dots, q_k)}{\partial q_k}\delta q_k^2} \tag{9}$$

According to Eq (6), one wish to propagate the uncertainty associated with each entry of $\mathbf{x}$ (Eq. (5)) towards N. This procedure is particularly insightful to ascertain the reliability of the predictions. In order to propagate the uncertainties throughout the PINN, it is necessary to determine the errors of the experimental input data obtained via CT as well as those involved in $\Delta\sigma$. With regard to Eq (5), let $\delta\sqrt{\text{area}}$, δh, δS, and δd be the measurement error of the defect descriptors. It is important to highlight that same error holds when multiple instances of the same feature are concerned. In this way, the error on the linear dimension due to the resolution of the reconstructed defect can be extended to the error arising when estimating a three-dimensional geometrical feature. Assuming l as the length of the side of the cube-like defect, its external volume can be equivalently expressed as $V_D = l^3$. Consequently, restating Eq. (9) as $V_D = f(l)$ leads to the volume uncertainty:

$$\delta V_D = 3V_D^{\frac{2}{3}}\,\delta l \tag{10}$$

In agreement with the cube-like approximation, the external area of the defects can be modelled as $A_D = 6\,l^2$. Thus, stating $A_D = f(l)$ allows one to compute the area error according to Eq. (9):

$$\delta A_D = 12\sqrt{\frac{A_D}{6}}\delta l \tag{11}$$

Following, δV_D and δA_D were utilised to calculate the sphericity error. To do so, it suffices to write $S = f(A_D, V_D)$ and exploit again Eq. (9):

$$\delta S = \sqrt{\left[\frac{12}{A_D}\left(\frac{\pi}{6}\right)^{\frac{1}{3}}\right]^2\delta l^2 + \left[-\frac{12\pi^{\frac{1}{3}}(6V)^{\frac{2}{3}}}{\sqrt{6}A_D^{\frac{3}{2}}}\right]^2\delta l^2} \tag{12}$$

Upon plugging $V_D = l^3$ into Eq. (2) and applying Eq. (9), the error related to the equivalent diameter can be promptly retrieved:

$$\delta d = \left(\frac{6}{\pi}\right)^{\frac{1}{3}}\delta l \tag{13}$$

Whilst $\sqrt{\text{area}}$ can be interpreted as the square root of a single lateral face of the cube-like defect, hence $l = \sqrt{\text{area}}$. Therefore, the corresponding error turns out to be:

$$\delta\sqrt{\text{area}} = \delta l \tag{14}$$

and differentiating further:

$$\delta\text{area} = 2\sqrt{\text{area}}\,\delta l \tag{15}$$

The last relationship gives the error of the projected area. Furthermore, since h is obtained by a direct measurement, its error is assumed as:

$$\delta h = \delta l \tag{16}$$

One can claim certain arbitrariness when using a cubic-like defect approximation. Alternatively, it is possible to regard defect as sphere without increasing the complexity of the error estimation process. Assuming r as the radius of the sphere-like defect, using the approach just outlined, and properly restating Eq. (9), it is possible to formulate the expressions for the uncertainties of the defect characteristics. In this instance, the error related to the sphericity is given by:

$$\delta S = \sqrt{\left[\frac{4\pi^{\frac{2}{3}}(6V)^{\frac{1}{3}}}{A_D}\right]^2 \delta r^2 + \left[-\frac{12\pi^{\frac{5}{6}}(6V)^{\frac{2}{3}}}{A_D^{\frac{3}{2}}}\right]^2 \delta r^2} \tag{17}$$

For sphere-like defects, the equivalent diameter turns out to be equal to twice the radius of the defects, thus the corresponding error turns out to be:

$$\delta d = 2\delta r \tag{18}$$

Whilst the error of $\sqrt{\text{area}}$ can be express as follow:

$$\delta\sqrt{\text{area}} = \sqrt{\pi}\delta r \tag{19}$$

Lastly, to quantify the error of the load amplitude, one can compute:

$$\delta\Delta\sigma = \sqrt{\left(\frac{1}{A}\right)^2 \delta\Delta F^2 + \left(-\frac{\Delta F}{A^2}\right)^2 \delta A^2} \tag{20}$$

where $\delta\Delta F$ is the error that affected the measurement of the applied force ΔF exerted by the fatigue rig, A refers to the cross-sectional area of the specimen and δA is the corresponding uncertainty due to the manufacturing process. In particular, δA can be obtained as follows:

$$\delta A = \pi\frac{\Phi}{2}\delta\Phi \tag{21}$$

where Φ is the diameter of the specimen's cross section and $\delta\Phi$ the corresponding dimensional tolerance.

2.2.2. Monte Carlo simulation

The appraisal of the error just presented should be propagated in terms of fatigue life to ascertain the reliability of the PINN. If the explicit analytical expression of Eq. (6) were available, Eq. (9) could be applied in the following form:

$$\delta N = \sqrt{\frac{\partial N(\sqrt{\text{area}}_1, \dots, \Delta\sigma)}{\partial\sqrt{\text{area}}_1}\delta\sqrt{\text{area}}_1^2 + \cdots + \frac{\partial N(\sqrt{\text{area}}_1, \dots, \Delta\sigma)}{\partial\Delta\sigma}\delta\Delta\sigma^2} \tag{22}$$

to retrieve the uncertainty regarding the prediction N. Unfortunately, the PINN notoriously acts as a black-box, hence pure EPT cannot be directly applied. To circumvent this limitation, PINN is seen from an agnostic viewpoint and a MC approach was pursued. In this regard, an adequate number of input samples is generated to satisfactorily cover the range where the descriptors are supposed to – plausibly – vary. In any case, the distributions are centred at the nominal values of the input features, whereas the standard deviations correspond to the measured errors. The PINN is then employed to compute the corresponding output so as to realise a database for further statistical post-processing and investigations. It is worth emphasising that two different sampling method were considered, namely one-parameter-at-a-time (OPT) and multiple-parameters-at-a-time (MPT). As concerns the former the inputs are independently sampled. In regards to the latter all the input features are sampled simultaneously.

For both sampling methodologies, a convergence test is performed in order to determine the minimum number of trials required to cancel the fluctuation that the random sampling typically generates [41]. The convergence is monitored through the following parameter expressed as:

$$c_i = 100\,\frac{|\log\overline{N_i} - \log\overline{N_{i-1}}|}{\log\overline{N_i}} \tag{23}$$

which practically represents the relative absolute difference in terms of expected value of fatigue lives $\overline{N_j}$ between two consecutive simulation. The convergence can be considered as reached when c_i yields values below a predefined threshold – 0.5% in the present work.

Let $\mathcal{P}$ be the output distribution of fatigue life N having n samples. In order to probe $\mathcal{P}$ a χ^2 test was performed. Under the so-called null hypothesis H_0, this test initially conjectures a certain distribution $\mathcal{D}$ for the output population whereby the samples of $\mathcal{P}$ may have been drawn. With reference to a certain significance level α, if the observations substantiate the hypothesis, H_0 cannot be rejected. In contrast, if the observations deviate from the expectation H_0 must be refused and one should be inclined to H_1. This test is commonly stated as:

$$\begin{cases} H_0 : \mathcal{P} \sim \mathcal{D} \\ H_1 : \mathcal{P} \nsim \mathcal{D} \end{cases} \tag{24}$$

Formally, H_0 cannot be rejected if the following condition holds:

$$U < \chi^2_{1-\alpha,l} \tag{25}$$

in which U is the test statistics:

$$U = \sum_{i=1}^{k} \frac{\mathcal{O}_i - \mathcal{E}_i}{\mathcal{E}_i} \tag{26}$$

where, in turn, k is the number of bins used to partition the output distribution N, $\mathcal{O}_i$ and $\mathcal{E}_i$ are the observed and theoretical frequency of each bin, respectively. The latter is defined as:

$$\mathcal{E}_i = M \cdot \mathbb{P}[N \in \mathcal{X}_i] \tag{27}$$

where M is the sample numerosity, and $\mathcal{X}_i$ is the i-th bin. Referring back to Eq. , $\chi^2_{1-\alpha,l}$ is the $1-\alpha$ quartile of a χ^2 distribution having l degrees of freedom, which are computed by the following rule:

$$l = k - 1 - p \tag{28}$$

where p is the number of parameters of $\mathcal{N}$ estimated from the sample. As concerns OPT sampling, three different H_0 were tested $\mathcal{P} \sim \mathcal{N}$, $\mathcal{P} \sim \mathcal{S}$, and $\mathcal{P} \sim \mathcal{L}$, i.e. $\mathcal{P}$ was hypothesised as normal distribution, skew-normal distribution and log-normal distribution, respectively. Conversely, two H_0 were formulated when considering MPT, namely $\mathcal{P} \sim \mathcal{S}$, and $\mathcal{P} \sim \mathcal{L}$. In any case, $\alpha = 0.05$ was adopted as the significance level.

An essential requirement for conducting a meaningful χ^2 test is the meticulous choice of a binning scheme for the output fatigue population. Since this operation is often non-trivial, the guidelines laid out in Refs. [42,43] were followed. Accordingly, the bin is recommended to possess at least an absolute frequency equal to 1, 80% bins are expected to have at least a frequency as high as 5, and $k \in [2n^{2/5}, 4n^{2/5}]$. Conservatively, $k = 4n^{2/5}$ is adopted in this study.

Each input feature has a specific physical meaning, so the corresponding perturbations must respect physical constraints. Each input cannot assume negative values, and the sphericity cannot exceed one (i.e. perfect spherical shape). As a result of this constraints, some of the input distributions resulted to be truncated. The following procedure was used to determine the statistical parameters of the theoretical distributions.

1. MC simulations were performed without removing the non-physical values of the input distributions, only with the aim of establishing the correct statistical parameters.
2. The statistical parameters of the theoretical distributions were determined using the Maximum Likelihood Estimation (MLE) and then tested through the χ^2 test.
3. The input distributions were filtered to obtain the actual fatigue life distributions.
4. The χ^2 test was performed to check whether the parameters of latter distributions could be approximated by those previously obtained.

2.3. Sensitivity analysis

2.3.1. Pearson's correlation coefficient

Pearson's correlation coefficient (PCC) is used to determine whether a certain feature of $\mathbf{x}$ (Eq. (5)), i.e. x_h is correlated with the predicted fatigue life N. The evaluation of PCC is pursued within the OPT framework, for each i-th fatigue samples. In order to lighten the notation, the superscript (i) will be omitted compared to Eq. (5). The evaluation of PCC is separately conducted for each features x_h. Accordingly, a dataset $D_{x_h} = \{x_{h,j} : j = 1, 2, \dots, n\}$ was prepared by drawing $n = 10\,000$ sorted and equally spaced trials in $[m_h, M_h]$ such that:

$$m_h = \min_{i=1,2,\dots,12} x_h^{(i)} \tag{29}$$

$$M_h = \max_{i=1,2,\dots,12} x_h^{(i)} \tag{30}$$

The last relationships seek the minimum and the maximum of the feature x_h among the fatigued specimens. Following, the PINN provided the set of predictions $D_{N,x_h} = \{N_{h,j} : N_{h,j} = \mathcal{H}(\mathbf{x}_{h\mapsto j}), x_{h,j} \in D_{x_h}, j = 1, 2, \dots, n\}$. In this instance, $\mathbf{x}_{h\mapsto j}$ is used to indicate that the h-th element of $\mathbf{x}$ is replaced with the j-th instance $x_{h,j} \in D_{x_h}$, whereas the remaining features are unchanged. As concerns the generic fatigued specimen, let $\hat{x_h}$ and $\hat{N}$ the expected value of x_h and N, respectively. Hence, PCC is defined as:

$$\mathrm{PCC} = \frac{\sum_{j=1}^{n}(x_{h,j} - \hat{x_h})(N_{h,j} - \hat{N})}{\sqrt{\sum_{j=1}^{n}(x_{h,j} - \hat{x_h})^2 \sum_{j}^{n}(N_{h,j} - \hat{N})^2}} \tag{31}$$

This coefficient can be seen as a normalised covariance taking values in $(-1, 1)$. As concerns its extreme instances, if $\mathrm{PCC} = 1$, then there exist a positive linear correlation between x_h and N. Conversely, if $\mathrm{PCC} = -1$ the variables are negatively correlation holds. Alongside, $\mathrm{PCC} = 0$ implies that no correlation holds between the variables can be inferred.

2.3.2. Permutation feature importance

One of the drawbacks of ML is the difficulty in assessing whether their outcomes are physically sound. Physics-Informed approaches can limit such a shortcoming, however the influence of not physically-informed parameters might not have a physical significance. The sensitivity of each input on the predicted fatigue life was investigated by means of the Permutation Feature Importance (PFI) [44]. In essence, this index allows one to quantify how far the predicted output depends upon a certain input feature. From a practical perspective, the evaluation of this index is performed by shuffling the input features and measuring the associated variation in terms of the output.

In regards to the present research, let $\mathbf{X}$ be the matrix whose rows are the representative vectors of the specimens Eq. (5):

$$\mathbf{X} = \begin{bmatrix} \sqrt{\text{area}}_1^{(1)} & \sqrt{\text{area}}_2^{(1)} & \dots & \Delta\sigma^{(1)} \\ \vdots & \vdots & \dots & \vdots \\ \sqrt{\text{area}}_1^{(r)} & \sqrt{\text{area}}_2^{(r)} & \dots & \Delta\sigma^{(r)} \\ \sqrt{\text{area}}_1^{(s)} & \sqrt{\text{area}}_2^{(s)} & \dots & \Delta\sigma^{(s)} \\ \vdots & \vdots & \dots & \vdots \\ \sqrt{\text{area}}_1^{(12)} & \sqrt{\text{area}}_2^{(12)} & \dots & \Delta\sigma^{(12)} \end{bmatrix} = \begin{bmatrix} \mathbf{x}^{(1)} \\ \vdots \\ \mathbf{x}^{(r)} \\ \mathbf{x}^{(s)} \\ \vdots \\ \mathbf{x}^{(12)} \end{bmatrix} \tag{32}$$

hence, each element of $\mathbf{X}$, i.e. $X_j^{(i)}$ corresponds to the j-th feature of the (i)-th fatigued sample. It is essential to note that $X_j^{(i)}$ are the measured input descriptors. The feature of interest is then selected, i.e. the i-th column, and the elements of such column are randomly shuffled T times, thus originating the permuted input matrix $\mathbf{X}'$ whose i-th row is $\mathbf{x}^{(i)'}$. Subsequently, both $\mathbf{X}$ and $\mathbf{X}'$ are processed row-wise according to Eq (8). This step provides the prediction of the finite fatigue life $N_j = \mathcal{H}(\mathbf{x}^{(i)})$, and $N_j' = \mathcal{H}(\mathbf{x}^{(i)'})$.

The mean absolute percentage error (MAPE) was chosen to quantify the deviation of N_j' from the original prediction N_j. With reference to the i-th feature and t-th permutation, MAPE is:

$$\text{MAPE}_{i,t} = 100\,\frac{1}{n}\sum_{j=1}^{n}\frac{|N_j - N_j'|}{N_j} \quad \forall\, t = 1, 2, \dots, T \tag{33}$$

where n is the number of samples – $n = 12$ for the present study. Finally, the minimum, average, and maximum $\text{MAPE}_{i,t}$ are taken in order to facilitate the interpretation of the results:

$$\text{MAPE}_{i,\min} = \min_{t=1,2,\dots,T} \text{MAPE}_{i,t} \tag{34}$$

$$\text{MAPE}_{i,\text{mean}} = \frac{1}{T}\sum_{t=1}^{T}\text{MAPE}_{i,t} \tag{35}$$

$$\text{MAPE}_{i,\max} = \max_{t=1,2,\dots,T} \text{MAPE}_{i,t} \tag{36}$$

Since the full permutation analysis would have required n. 12! A stopping criterion was adopted. Specifically, the analysis was stopped when the second decimal figure of $\text{MAPE}_{i,\min}$ and $\text{MAPE}_{i,\max}$ $\forall\, i = 1, 2, \dots, 13$ stabilised.

2.3.3. Accumulated local effect

The Accumulated Local Effects (ALE) is a statistic, yet handy index to ascertain whether potential correlation might exists among the input features and the predicted output. Differently from PFI, ALE allows one to continuously visualise the detected correlation throughout a given spectrum of values instead of providing a single scalar value [45].

It is worth recalling how ALE is typically computed. In order not to overload the main content of the present manuscript, the following explanation will refer to a predefined input feature x_j selected amongst those in $\mathbf{x}^{(i)}$ (Eq. (5)). The determination of ALE commence with gathering the values of x_j for each (i)-th specimen into the set $A_{x_j} = \{x_j^{(i)} : i = 1, 2, \dots, 12\}$, thus defining the interval $I_{x_j} = [\min A_{x_j}, \max A_{x_j}]$. Next, I_{x_j} is evenly partitioned into K sub-intervals $I_{j,k} = (x_{j,k-1}, x_{j,k}]$:

$$I_{x_j} = \bigcup_{k=1}^{K} I_{j,k} = \bigcup_{k=1}^{K} (x_{j,k-1}, x_{j,k}]$$

Spanning $I_{j,k}$ $\forall\, k = 1, 2, \dots, K$, if $x_j^{(i)} \in A_{x_j}$ belongs to $I_{j,k}$, the whole vector $\mathbf{x}^{(i)}$ is assumed to belong to $I_{j,k}$. Consequently, the prediction $N_j = \mathcal{H}(\mathbf{x}^{(i)})$ is replaced with those in correspondence of the lower and upper bound of $I_{j,k}$, i.e. $x_{j,k-1}$ and $x_{j,k}$. In this regard, let $\mathbf{x}_{j\mapsto k}^{(i)}$ be the input vector where x_j is replaced by $x_{j,k}$ – analogously for $x_{j,k-1}$. Therefore, the related predictions are $N_{j,k-1} = \mathcal{H}(\mathbf{x}_{j\mapsto k-1}^{(i)})$, and $N_{j,k} = \mathcal{H}(\mathbf{x}_{j\mapsto k}^{(i)})$. This operation is repeated for each $\mathbf{x}^{(i)} \in I_{j,k}$, and arithmetic average of the set $\{N_{j,k-1} - N_{j,k}\}_{\mathbf{x}^{(i)}\in I_{j,k}}$ is taken. Such an average is then *accumulated* up to an arbitrary threshold $x_{j,k}$ $\forall\, k = 1, 2, \dots K$:

$$\text{ALE}_{x_j}'(x_{j,k}) = \sum_{x_{j,0},\dots,x_{j,k}} \frac{1}{\nu_{j,k}} \sum_{\mathbf{x}^{(i)}\in I_{j,k}} N_{j,k-1} - N_{j,k} \tag{37}$$

Table 1
Absolute errors of PINN inputs.

	$\delta\sqrt{\text{area}}$ [μm]	δh [μm]	δS [−]	δd [μm]	$\delta\Delta\sigma$ [MPa]
Cube-like defect	30	30	0.1	40	4
Sphere-like defect	50	30	0.1	60	4

Table 2
MC input distributions. $\mathcal{N}(\mu,\sigma)$ refers to the normal distribution where μ and σ are the expected value and the standard deviation, respectively.

Input	Distribution
$\sqrt{\text{area}}_j^{(i)}$	$\mathcal{N}(\sqrt{\text{area}}_j^{(i)}, 30)$ μm
$h_j^{(i)}$	$\mathcal{N}(h_j^{(i)}, 30)$ μm
$S_j^{(i)}$	$\mathcal{N}(S_j^{(i)}, 0.1)$
$d_j^{(i)}$	$\mathcal{N}(d_j^{(i)}, 40)$ μm
$\Delta\sigma_j^{(i)}$	$\mathcal{N}(\Delta\sigma_j^{(i)}, 4)$ MPa

where $\nu_{j,k}$ denotes the number of $\mathbf{x}^{(i)} \in I_{j,k}$. Since ALE is typically centred at zero, i.e. mean null effect, Eq. (37) is translated through:

$$\text{ALE}_{x_j}(x_{j,k}) = \text{ALE}'_{x_j}(x_{j,k}) - \frac{1}{n}\sum_{x_{j,0},\ldots,x_{j,K}} \nu_{j,k}\,\text{ALE}'_{x_j}(x_{j,k}) \tag{38}$$

where n is the number of samples, i.e. $n = 12$.

Since 12 samples were available, it was therefore decided to divide each input domain in six parts, each containing two samples.

3. Results and discussion

3.1. Error propagation in experimental data

In order to evaluate the error of the experimental inputs of the PINN, it was essential to establish the length error and the error of the applied force. The length error was assumed to be 30 μm, based on the resolution of CT and the process required to measure the experimental quantities involved, so, both δl and δr were considered to be equal to 30 μm. The load cell used to perform the fatigue test was assumed to have an applied force uncertainty of $\delta\Delta F = 5$ N. The value chosen for $\delta\Delta F$ might appear conservative in some instances, as the controller of servo-hydraulic machines can allegedly introduce larger uncertainties. Nevertheless, the proposed framework is sufficiently versatile such that δN can readily be updated. In any case, the adopted δN is deemed to cover a vast variety of instances and sufficient to conduct a meaningful uncertainty quantification. All the specimens were machined before undergoing the fatigue test. Since the dimensional tolerance of machined parts is typically equal to 0.02 mm, then $\delta\Phi = 0.02$ mm. As it can be seen from Eqs. (14)–(16), only the sphericity error depends on other experimentally measured quantities, i.e. A_l and V. For the sake of simplicity, the sphericity error was assumed to be equal to the average of the sphericity errors calculated for all the three considered defects of all the specimens. The absolute errors of the inputs, are reported in Table 1.

In terms of the error associated with $\sqrt{\text{area}}$ and d, the sphere-like approximation provided more variability on the inputs, thus resulting in a more conservative approach than the cubic-like one.

3.2. Uncertainty quantification of fatigue life predictions

In order to perform the MC simulations, a Normal distribution was prescribed over each input of the PINN. In this instance, the Normal distribution simulates the acquisition of the experimental data affected by random noise, which is frequently encountered in practice. Although this choice is arbitrary, the assumption is not restrictive as one can hypothesise any other distributions deemed as relevant to the problem. The adopted distributions are shown in Table 2 which refers to the cube-like approximation for the shape of the defects. The same analysis was repeated by setting the standard deviation of the Normal distributions equal to the errors obtained with the spherical approximation. It should be noted that the input distributions are centred at the experimental values, whereas the standard deviations are taken from Table 1. The peculiar choice of the normal distribution ensures that for a sufficiently large number of trials the expected value converges to the prescribed mean, i.e. the experimental value in this case. This holds according to the Central Limit Theorem. For the sake of illustrating the sampling strategy, the resulting distribution of d is given in Fig. 1. Additionally, the leftmost tail of the distribution is truncated as a result of the removal of non-physical values.

Fig. 2 shows the results of the convergence tests applied to the outcomes of the MC simulations obtained for the MPT case. This case had the higher variability so it was also considered as a reference case for the OPT case. As it can be seen in Fig. 2, the minimum number of trials required to achieve convergence test was 7000, referring to samples n.1, under the sphere-like approximation. Conservatively, the number of trials was set to 10 000 both for OPT and MPT and both for the shape approximations.

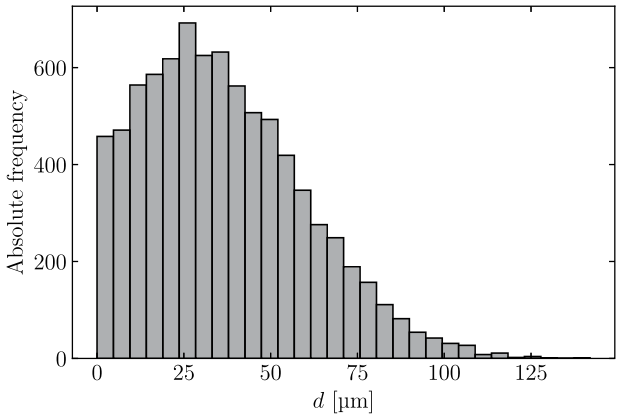


Fig. 1. Example of a truncated input distribution.

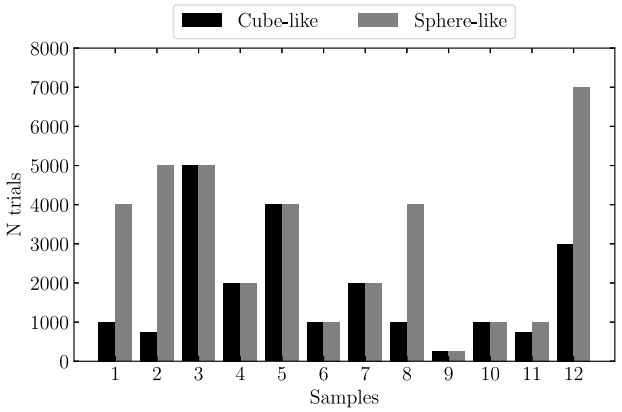


Fig. 2. Number of necessary trials required to attain MC test convergence for each considered tested sample and for both the shape defect approximations.

Table 3
Chi-square test results for MPT case. $\mathcal{L}$ and $\mathcal{S}$ refers to log-normal and skew-normal distribution respectively.

Defect shape	Sample											
	1	2	3	4	5	6	7	8	9	10	11	12
Cube-like	$\mathcal{S}$	$\mathcal{S}$	$\mathcal{L}$	$\mathcal{L}$	$\mathcal{S}$	$\mathcal{L}$	$\mathcal{L}$	$\mathcal{L}$	$\mathcal{L}$	$\mathcal{L}$	$\mathcal{L}$	$\mathcal{L}$
Sphere-like	$\mathcal{S}$	$\mathcal{S}$	$\mathcal{L}$	$\mathcal{L}$	$\mathcal{L}$	$\mathcal{L}$	$\mathcal{L}$	$\mathcal{L}$	$\mathcal{S}$	$\mathcal{L}$	$\mathcal{L}$	$\mathcal{L}$

MC simulations allowed the output distribution of fatigue to be investigated for both OPT and MPT. Appendix B reports the distributions under H_0 which was not rejected in any examined case of OPT.

Table 3 illustrates the results of the χ^2 test for MPT, in particular $\mathcal{L}$ and $\mathcal{S}$ refers to the log-normal and skew-normal distributions respectively. It should be mentioned that, differently from the OPT case, the MPT did not involved a normal distribution for H_0. A preliminary critical inspection of the resulting output distributions revealed asymmetric distributions only. As such, the normal distribution was excluded from H_0. As it can be seen from the results of the χ^2 test reported in Table 3, there is no significant difference between the results obtained when the defect shape is approximated by a cube or a sphere.

Regarding MPT, the output variability is a measure of the uncertainty in the fatigue life prediction. Fig. 3 shows the box plots of the output of the MPT analysis for both the defect shape approximation. Therein each box is representative of a unique sample whose input features were perturbed simultaneously. Specifically, each box is used to visually and concisely display the resulting output fatigue distribution of each sample. In particular, the box spans the interquartile range, whereas the whiskers cover the range from 5% to 95% of the predicted output fatigue distribution. Additionally, each box plot show a red line corresponding to the median value. The position of the median in the box plots shows that, even if the fatigue life has a skew-normal distribution, the asymmetry is barely perceptible. Presumably, this effect can be attributed to truncating the input distributions. Additionally, one can raise concerns about the asymmetry of the output distribution, even though symmetric (normal) distribution were used. This phenomenon can be interpreted as being a result of the affine operations and non-linearities that the PINN intrinsically holds. Hence,

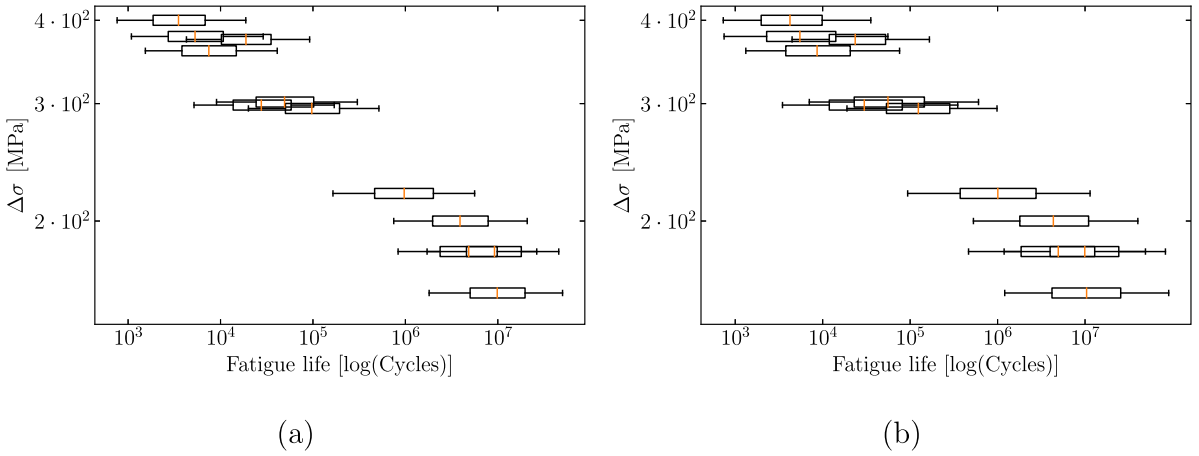


Fig. 3. Box plots of fatigue life predictions obtained for MPT under (a) the cubic approximation and (b) the spherical approximation. Each box plot refers to a different sample. The red line represents the median of the distributions, the sides of the box correspond to the interquartile range and the whiskers cover the range from 5% to 95% of the predicted output fatigue distribution. The samples are sorted by the number of cycles to failure predicted by the PINN as per Tables A.4 and A.5.

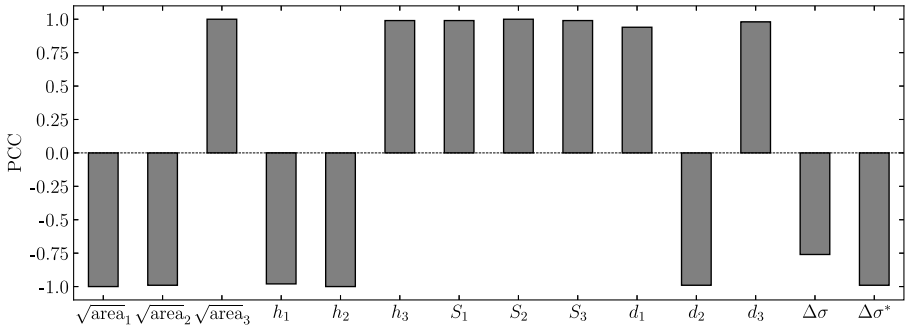


Fig. 4. Comparison of the PCC averaged result of each input. Note that $\Delta\sigma^*$ refers to the PCC of the applied stress range upon taking the logarithm of both $\Delta\sigma$ and N prior the application of Eq. (31).

in principle, the PINN might not constitute an even function. Lastly, Fig. 3 also shows that as the median fatigue life increases, the corresponding variability range becomes higher.

As it should be expected, the input Normal distributions based on the sphere-like approximation result in a more pronounced variability in the fatigue life prediction than those based on the cube-like approximation.

3.3. Sensitivity analysis

3.3.1. Pearson's correlation coefficient

The obtained results for PCC are reported in Appendix C. Fig. 4 shows the average PCC results grouped by defect characteristics.

The PCC results show that the PINN established high linear correlations between the inputs related to defect features and the fatigue life. Only the relationship between the fatigue load amplitude and the number of cycles appears to be farther from this scenario. The last result is consistent with the theoretical model used in the PINN: the theoretical fatigue life was predicted by Eq. (7), which is a power law that express fatigue life depending on the fatigue load amplitude. This relationship becomes linear when plotted on a log–log scale. In fact, the calculation of the average PCC of the applied stress range upon taking the logarithm of N and $\Delta\sigma$ turned out to be close to 1. In this instance, the PCC is shown in Fig. 4 in correspondence of $\Delta\sigma^*$.

As concerns $\sqrt{\text{area}}$, Fig. 4 shows that the first and the second inputs have a negative correlation and have the greatest influence on the fatigue life. The third input shows a different trend but also has the smallest influence on the PINN result. According to LEFM model, the higher the $\sqrt{\text{area}}$, the higher the SIF is, thus leading to a decrease in terms of fatigue life. Apart from $\sqrt{\text{area}}_3$, the obtained correlation corroborates the LEFM model. It is important to recall that, $\sqrt{\text{area}}_3$ is very unlikely to trigger fatigue failure as it holds the smallest normalised SIF, therefore its influence may have been inaccurately modelled by the analysed PINN framework. Analogous considerations also applies to the analysis of the remaining input features.

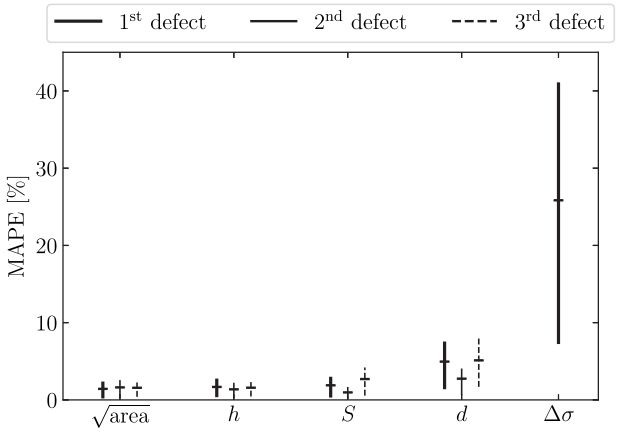


Fig. 5. Visual representation of PFI results.

Regarding the distance between the surface and the defect, the first two inputs are negatively correlated with fatigue life predictions, whilst the third is positively correlated. Experimentally, surface-defects cause a greater reduction of fatigue life as compared with those in the bulk. Therefore, one would expect a positive correlation with the fatigue life instead. Nevertheless, it is important to note that, according to the criterion expressed in Eq. (3), the considered dataset showed that around 80% considered defects are surface defects. This implies that the sole PINN model cannot fully capture the correlation between h and the fatigue life since the variability of this descriptor is limited and bulk defects are under represented, i.e. much fewer bulk defects are involved in the calculations. This could reasonably explain the unexpected behaviour of the h related to the first two defects individuated by the PINN.

As far as defect's sphericity is concerned, S, a linear positive correlation is found with the number of cycles to failure for all the considered defects. Such a result is convincing as it is well accepted that the most dangerous defects are those showing crack-like traits such as flatness and high geometrical irregularity (e.g. LOFs).

The last defect feature considered by the PINN model is the equivalent diameter d. This parameter is related to the size of the defect. As d increases, a decrease in fatigue life should be observed due to the fact that a larger defect is more detrimental than a small one when they are approximately at the same distance from the surface. Unexpectedly, the first and the third inputs show a positive correlation. Such an unexpected result may arise as a consequence of an hidden interplay between the defect sphericity, apart from the limited number of experimental data used to train the model. For instance, defects with large d can intrinsically be associated with a LOF defect or the other way around. This interesting observation provides some clues on the appropriateness of using such specific defect descriptors to well represent the influencing defect characteristics.

Lastly, the positive correlation of $\Delta\sigma$ seem to concur well with the hypothesised phenomenon, i.e. for increasing values of $\Delta\sigma$ the fatigue life progressively decreases.

3.3.2. Permutation feature importance

In Appendix D are reported the results of PFI analysis in terms of MAPE, and number of permutation used for each input feature. Alongside, Fig. 5, offers a succinct representation of the data in Appendix D.

Conversely, $\sqrt{\text{area}}$ which was supposed to show the second most influencing effect, in fact, it yields to the smallest MAPE, as well as the distance to the free-surface, h. Whilst, S and d appear to be the most influencing factor, after the applied load range. Again, due to the very limited dataset used for the training process, uniquely unveiling the role of each descriptor remains utopistic. Nonetheless, some useful insights can be obtained herein. The large influence attributed to the equivalent diameter and the sphericity is a factual data, meaning that they certainly play a relevant role, comparable with what is well-known and accepted $\sqrt{\text{area}}$. It should not be excluded that the use of different defect's descriptors can lead to better insights. These results are consistent with [28], which showed that the role of a specific defect feature can vary significantly with the variation of other parameters. Thus, the choice of leading descriptors may vary as a function of experimental conditions and material.

3.3.3. Accumulated local effects

The ALE plot showing the influence of the fatigue load amplitude on the fatigue life predicted by the PINN model is depicted in Fig. 6. Therein, the trend of $\Delta\sigma$ resembles the classic Basquin's law widely used in fatigue. In particular, the variation of predicted fatigue life increase for decreasing values of $\Delta\sigma$.

With regard to the influence of the defect characteristics, Fig. 7 shows the corresponding ALE plots. The trends depicted in Fig. 7 reflect relation shown by PCC in Table C.12, qualitatively. All features show straight vertical trends in certain areas, indicating low variability. Whilst, milder gradient indicate a more pronounced dependency. In particular, $\sqrt{\text{area}}$, h, and d result to have a significant effect on the fatigue life only above a certain threshold. Interestingly, the sphericity seems to play a role only if below a certain value: more flattened defects — which appears to be reasonable. This very interesting result highlight these defect's descriptors may play a different role depending on their magnitudes.

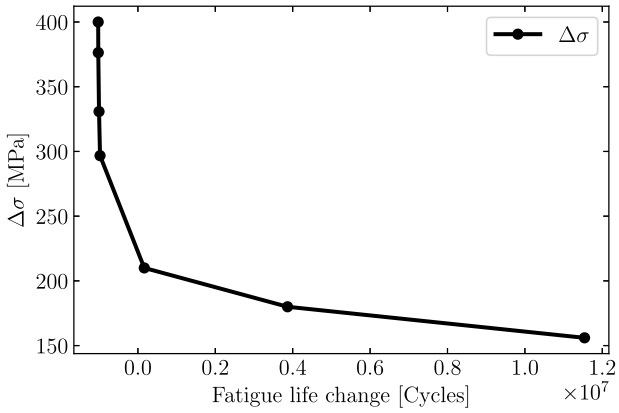


Fig. 6. ALE plot for the amplitude fatigue load.

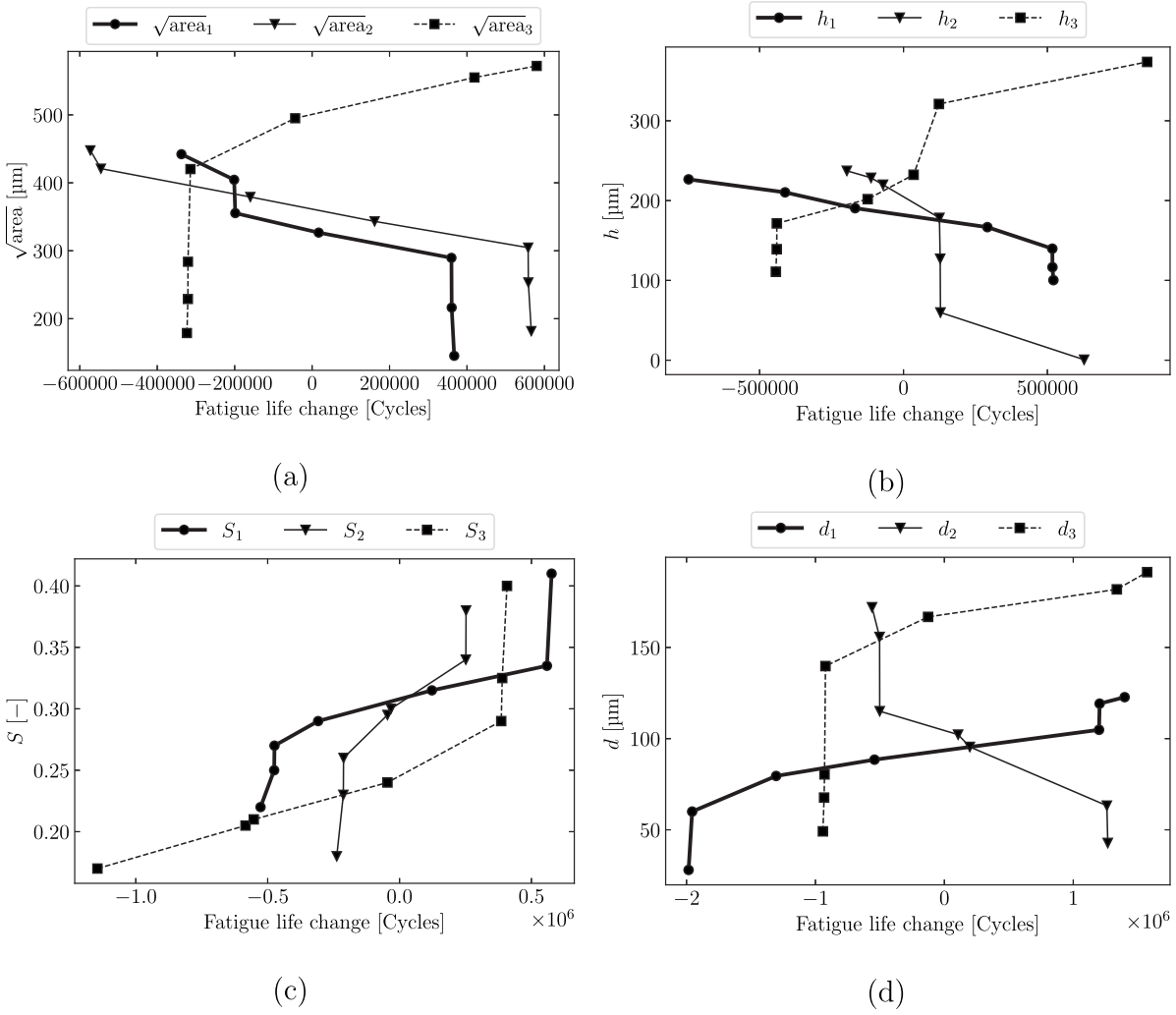


Fig. 7. ALE plots for defect features.

4. Conclusions

The objective of this study was twofold. On one side, it addressed the uncertainty quantification of a defect-based pre-trained PINN model used to predict the fatigue life of a small-sized batch of SLM AlSi10Mg tested samples. On the other, it involved a sensitivity analysis to verify whether the PINN model was able to identify relationships between input quantities and the fatigue life prediction, despite the limited experimental data set.

The uncertainty quantification was conducted by performing a series of MC simulations upon the prescription of realistic uncertainties in the considered inputs. This resulted in a rigorous quantification of uncertainties over the predicted fatigue life, which are of great interest when designing materials against fatigue failure. The sensitivity analysis involved the assessment of three different statistical indices leading to the following results:

- PCC showed that all the considered defect features correlated with the fatigue life through the PINN. Unexpected relations were found regarding the distance to the free surface and the equivalent diameter.
- PFI allowed the influence of each input on the PINN prediction to be quantified. The fatigue load amplitude had the largest influence, whilst Murakami's $\sqrt{\text{area}}$ and the distance from the free surface had the smallest.
- ALE plots allowed for the observation of input–output correlations, and highlighted which values of each descriptor show the strongest influence on the fatigue life. Moreover, this analysis showed a limited role of the third-ranked defect (in terms of δK)

Overall, the PINN framework appeared to be a powerful tool to evaluate fatigue and to provide insights into fatigue influencing features. Indeed, the present study provided interesting clues on how the employed defect's descriptors may play different roles depending on their magnitude and considered fatigue regimes. In any case, it is imperative to reiterate that the prerequisite for a PINN training dataset is its coverage over a wide range of features' magnitudes, rather than simply relying on very large datasets that do not enclose important regimes. For instance, the small influence of the $\sqrt{\text{area}}$ is certainly due to the narrow range of values employed during the PINN training process. This issue can be coped by either introducing additional physical constraints or by adding more experimental tests at substantially different killer defects dimensions. Same applies to the distance to the free surface to the killer defects.

In conclusion, it is necessary to remark that results of the present research may not be universally applicable to any metallic material, similarly to what recently unveiled A. Li using a pure ML method [28]. Therefore, further extensive fundamental research is required to gain a deep understanding of the underlying microstructural phenomena involved, both from the defect morphology side and from the material susceptibility to complex-shaped defects, including effects of residual stress.

CRediT authorship contribution statement

Emanuele Avoledo: Writing – original draft, Visualization, Methodology, Formal analysis, Data curation, Conceptualization. **Alessandro Tognan:** Writing – review & editing, Validation, Methodology, Conceptualization. **Enrico Salvati:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This work has been supported by the project "CONCERTO – Multiscale modelling/characterisation and fabrication of nanocomposite ceramics with improved toughness" funded by the MIUR Progetti di Ricerca di Rilevante Interesse Nazionale (PRIN) Bando 2020, Italy – grant 2020BN5ZW9.

Appendix A. Experimental dataset

Table A.4 reports the experimental CT dataset involved in the present study. Each row of the table corresponds to a different input vector (Eq. (5)) and includes the geometric characteristics of the three defects with the highest δK and the load amplitude associated with the corresponding fatigue test.

Table A.5 shows the results of the fatigue tests in terms of number of cycles and applied stress range. Alongside, the table indicates whether a killer defect was detected inside the related sample.

Table A.4
Experimental CT dataset used to train the PINN in [33].

Sample	$\sqrt{area}_1$ [mm]	$\sqrt{area}_2$ [mm]	$\sqrt{area}_3$ [mm]	h_1 [mm]	h_2 [mm]	h_3 [mm]	S_1 [–]	S_2 –]	S_3 [–]	d_1 [mm]	d_2 [mm]	d_3 [mm]
1	0.223	0.296	0.253	0.131	0.120	0.165	0.31	0.36	0.30	0.069	0.012	0.082
2	0.442	0.429	0.492	0.170	0.223	0.217	0.26	0.23	0.21	0.118	0.141	0.159
3	0.145	0.182	0.179	0.102	0.116	0.122	0.41	0.38	0.40	0.028	0.050	0.049
4	0.331	0.334	0.541	0.220	0.234	0.187	0.29	0.32	0.21	0.090	0.077	0.175
5	0.374	0.313	0.315	0.148	0.174	0.178	0.28	0.30	0.30	0.101	0.101	0.078
6	0.271	0.283	0.243	0.131	0.134	0.156	0.24	0.25	0.28	0.088	0.090	0.079
7	0.363	0.413	0.370	0.227	0.237	0.242	0.26	0.23	0.23	0.109	0.172	0.138
8	0.210	0.223	0.214	0.100	0.004	0.111	0.34	0.29	0.35	0.051	0.042	0.057
9	0.347	0.448	0.500	0.200	0.183	0.310	0.22	0.18	0.17	0.121	0.171	0.191
10	0.322	0.352	0.572	0.163	0.217	0.374	0.32	0.30	0.21	0.074	0.101	0.186
11	0.435	0.365	0.470	0.200	0.223	0.222	0.29	0.30	0.25	0.123	0.107	0.142
12	0.309	0.394	0.569	0.182	0.001	0.332	0.33	0.27	0.20	0.085	0.103	0.178

Table A.5
Load amplitude, and experimentally recorded fatigue life of each considered specimen. The last column indicates whether the fractographic investigation disclosed the presence of killer defects. In particular, `I` and `NI` stands for identifiable and non-identifiable, respectively.

Sample	$\Delta\sigma$ [MPa]	Fatigue life [Cycles]	Killer defect
1	400.0	474	I
2	378.3	11 465	NI
3	360.0	3432	I
4	374.4	28 201	NI
5	298.5	19 806	NI
6	301.6	39 538	NI
7	295.0	46 255	I
8	220.0	2 622 640	I
9	200.0	15 242 310	I
10	180.0	11 352 768	I
11	180.0	237 485	NI
12	156.0	3 795 336	NI

Appendix B. Results of the Monte Carlo simulation for OPT sampling

Tables B.6–B.8 reports the results of the χ^2 test applied to the fatigue life distributions obtained for the OPT case under the cube-like assumption for the defects' shape. Whilst Tables B.9–B.11 reports the results of the χ^2 test applied to the fatigue life distributions obtained for the OPT case under the cube-like assumption for the defects' shape. In particular, $\mathcal{N}(\mu,\sigma)$, $\mathcal{L}(\mu,\sigma)$ and $\mathcal{S}(\mu,\sigma,a)$ refers to the normal, log-normal and skew-normal distributions respectively. Additionally, μ is the expected value, σ is the standard deviation, and a is the skewness parameter of the distribution. To obtain the parameters of the log-normal and skew-normal distributions as a function of μ, σ and a, the guidelines described in [46] and in [47] could be followed.

Table B.6
Results of χ^2 test applied to the output distributions of OPT, under the cube-like assumption for the defect shape, for samples 1–4.

Sample	1	2	3	4
$\sqrt{area}_1$	$\mathcal{S}(3231, 165, -0.57)$	$\mathcal{L}(5040, 230)$	$\mathcal{N}(6944, 336)$	$\mathcal{N}(14315, 763)$
$\sqrt{area}_2$	$\mathcal{L}(3228, 182)$	$\mathcal{L}(5045, 299)$	$\mathcal{L}(6938, 415)$	$\mathcal{L}(14294, 805)$
$\sqrt{area}_3$	$\mathcal{L}(3227, 50)$	$\mathcal{L}(5044, 173)$	$\mathcal{N}(6932, 178)$	$\mathcal{L}(14280, 211)$
h_1	$\mathcal{L}(3251, 347)$	$\mathcal{S}(5073, 595, 0.62)$	$\mathcal{L}(6981, 826)$	$\mathcal{L}(14327, 1589)$
h_2	$\mathcal{L}(3229, 154)$	$\mathcal{L}(5046, 268)$	$\mathcal{L}(6941, 366)$	$\mathcal{L}(14299, 671)$
h_3	$\mathcal{L}(3234, 173)$	$\mathcal{L}(5047, 298)$	$\mathcal{L}(6944, 404)$	$\mathcal{L}(14307, 789)$
S_1	$\mathcal{L}(3530, 1511)$	$\mathcal{S}(5412, 2068, 1.18)$	$\mathcal{L}(7489, 3036)$	$\mathcal{L}(15590, 6940)$
S_2	$\mathcal{S}(3373, 876, 1.31)$	$\mathcal{S}(5170, 836, 2.59)$	$\mathcal{S}(7164, 1483, 1.81)$	$\mathcal{S}(14899, 4047, 1.09)$
S_3	$\mathcal{S}(3623, 1867, 0.99)$	$\mathcal{S}(5527, 2270, 1.41)$	$\mathcal{S}(7804, 3743, 1.10)$	$\mathcal{L}(16389, 8830)$
d_1	$\mathcal{N}(4001, 2898)$	$\mathcal{S}(6524, 5429, 1.02)$	$\mathcal{N}(9008, 7268)$	$\mathcal{L}(17709, 12515)$
d_2	$\mathcal{L}(3427, 1179)$	$\mathcal{L}(5250, 1461)$	$\mathcal{L}(7182, 2002)$	$\mathcal{L}(14777, 4080)$
d_3	$\mathcal{L}(3402, 1164)$	$\mathcal{L}(5436, 2200)$	$\mathcal{L}(7477, 2900)$	$\mathcal{L}(15165, 5170)$
$\Delta\sigma$	$\mathcal{N}(3226, 30)$	$\mathcal{N}(5038, 58)$	$\mathcal{N}(6929, 75)$	$\mathcal{N}(14279, 133)$

Table B.7
Results of χ^2 test applied to the output distributions of OPT, under the cube-like assumption for the defect shape, for samples 5–8.

Sample	5	6	7	8
$\sqrt{\text{area}}_1$	$\mathcal{N}(26776, 1265)$	$\mathcal{L}(46422, 2302)$	$\mathcal{N}(915817, 40864)$	$\mathcal{N}(89464, 4520)$
$\sqrt{\text{area}}_2$	$\mathcal{L}(26809, 1689)$	$\mathcal{L}(46443, 2956)$	$\mathcal{L}(815373, 58620)$	$\mathcal{L}(89371, 5491)$
$\sqrt{\text{area}}_3$	$\mathcal{L}(26808, 1083)$	$\mathcal{L}(46410, 1646)$	$\mathcal{N}(89311, 2808)$	$\mathcal{N}(915956, 44112)$
h_1	$\mathcal{S}(26977, 3406, 0.58)$	$\mathcal{L}(46761, 5915)$	$\mathcal{L}(89986, 11236)$	$\mathcal{L}(922426, 123293)$
h_2	$\mathcal{L}(26817, 1553)$	$\mathcal{N}(46446, 2652)$	$\mathcal{N}(89426, 4924)$	$\mathcal{L}(916066, 56388)$
h_3	$\mathcal{L}(26821, 1721)$	$\mathcal{L}(46485, 2963)$	$\mathcal{L}(89480, 5653)$	$\mathcal{L}(916558, 62262)$
S_1	$\mathcal{S}(28821, 11390, 1.02)$	$\mathcal{L}(50340, 21066)$	$\mathcal{L}(96529, 40418)$	$\mathcal{L}(969416, 350621)$
S_2	$\mathcal{S}(27343, 4056, 2.27)$	$\mathcal{S}(47553, 8660, 1.48)$	$\mathcal{S}(91659, 18059, 1.31)$	$\mathcal{S}(922622, 97323, 1.20)$
S_3	$\mathcal{S}(29215, 11790, 1.20)$	$\mathcal{L}(51513, 23310, 1.00)$	$\mathcal{S}(99392, 46245, 0.51)$	$\mathcal{S}(978161, 345049, 0.73)$
d_1	$\mathcal{L}(35980, 32692)$	$\mathcal{N}(62034, 55972)$	$\mathcal{S}(118419, 100215, 0.72)$	$\mathcal{L}(1267664, 1176561)$
d_2	$\mathcal{L}(28002, 8257)$	$\mathcal{L}(48300, 14274)$	$\mathcal{L}(92726, 27222)$	$\mathcal{L}(949427, 279790)$
d_3	$\mathcal{L}(29201, 12834)$	$\mathcal{L}(51041, 21807)$	$\mathcal{L}(96914, 39704)$	$\mathcal{L}(1009217, 468271)$
$\Delta\sigma$	$\mathcal{N}(26770, 338)$	$\mathcal{N}(46369, 559)$	$\mathcal{N}(89262, 1031)$	$\mathcal{N}(914142, 12307)$

Table B.8
Results of χ^2 test applied to the output distributions of OPT, under the cube-like assumption for the defect shape, for samples 9–12.

Sample	9	10	11	12
$\sqrt{\text{area}}_1$	$\mathcal{L}(3702963, 161903)$	$\mathcal{S}(4846924, 214318, -0.66)$	$\mathcal{S}(9782630, 413674, -0.68)$	$\mathcal{L}(9153476, 383796)$
$\sqrt{\text{area}}_2$	$\mathcal{L}(3705568, 234172)$	$\mathcal{L}(4846088, 305440)$	$\mathcal{S}(9794632, 593757)$	$\mathcal{L}(9159714, 562426)$
$\sqrt{\text{area}}_3$	$\mathcal{L}(3703787, 169468)$	$\mathcal{N}(4847296, 221004)$	$\mathcal{L}(9155473, 408507)$	$\mathcal{N}(9795480, 427108)$
h_1	$\mathcal{L}(3731246, 481646)$	$\mathcal{L}(4887827, 629448)$	$\mathcal{L}(9235726, 1139652)$	$\mathcal{L}(9851748, 1227947)$
h_2	$\mathcal{N}(3706137, 221190)$	$\mathcal{L}(4849147, 287052)$	$\mathcal{N}(9160716, 529094)$	$\mathcal{L}(9798370, 572566)$
h_3	$\mathcal{L}(3708983, 241700)$	$\mathcal{L}(4858359, 319294)$	$\mathcal{L}(9167384, 575908)$	$\mathcal{N}(9798994, 622628)$
S_1	$\mathcal{L}(3919340, 1413311)$	$\mathcal{S}(5132122, 1815204, -0.67)$	$\mathcal{L}(9618561, 3296626)$	$\mathcal{L}(10279181, 3563896)$
S_2	$\mathcal{L}(3724144, 409040)$	$\mathcal{L}(4873220, 527840)$	$\mathcal{S}(9170917, 930730, -1.17)$	$\mathcal{L}(9806713, 1007671)$
S_3	$\mathcal{L}(3933467, 1361876)$	$\mathcal{L}(5070991, 1736128)$	$\mathcal{L}(9651630, 3175433)$	$\mathcal{L}(10273902, 3362221)$
d_1	$\mathcal{N}(4942568, 4396306)$	$\mathcal{L}(6507407, 5763946)$	$\mathcal{L}(11957044, 10105944)$	$\mathcal{L}(12665685, 10518454)$
d_2	$\mathcal{L}(3835457, 1095108)$	$\mathcal{L}(5026145, 1457808)$	$\mathcal{L}(9456502, 2610832)$	$\mathcal{L}(10153115, 2819081)$
d_3	$\mathcal{L}(4075112, 1795173)$	$\mathcal{L}(5287743, 2345301)$	$\mathcal{L}(10012830, 4272034)$	$\mathcal{L}(10531701, 4418256)$
$\Delta\sigma$	$\mathcal{N}(3700198, 48050)$	$\mathcal{N}(4842237, 62918)$	$\mathcal{N}(9147580, 115601)$	$\mathcal{N}(9783599, 124426)$

Table B.9
Results of χ^2 test applied to the output distributions of OPT, under the sphere-like assumption for the defect shape, for samples 1–4.

Sample	5	6	7	8
$\sqrt{\text{area}}_1$	$\mathcal{L}(3243, 272)$	$\mathcal{L}(5062, 381)$	$\mathcal{L}(6965, 563)$	$\mathcal{L}(14362, 1277)$
$\sqrt{\text{area}}_2$	$\mathcal{L}(3237, 303)$	$\mathcal{L}(5059, 503)$	$\mathcal{L}(6958, 695)$	$\mathcal{L}(14329, 1347)$
$\sqrt{\text{area}}_3$	$\mathcal{S}(3229, 84, 1.04)$	$\mathcal{L}(5049, 288)$	$\mathcal{L}(6940, 298)$	$\mathcal{L}(14291, 357)$
h_1	$\mathcal{L}(3247, 356)$	$\mathcal{L}(5077, 606)$	$\mathcal{L}(6981, 826)$	$\mathcal{L}(14372, 1589)$
h_2	$\mathcal{L}(3236, 155)$	$\mathcal{L}(5047, 268)$	$\mathcal{L}(6941, 366)$	$\mathcal{L}(14299, 671)$
h_3	$\mathcal{L}(3232, 172)$	$\mathcal{L}(5049, 298)$	$\mathcal{L}(6944, 404)$	$\mathcal{L}(14307, 789)$
S_1	$\mathcal{L}(3505, 1483)$	$\mathcal{S}(5398, 2014, 1.08)$	$\mathcal{L}(7489, 3035)$	$\mathcal{L}(15590, 6940)$
S_2	$\mathcal{S}(3369, 876, 1.44)$	$\mathcal{S}(5178, 835, 2.53)$	$\mathcal{S}(7164, 1483, 1.81)$	$\mathcal{S}(14899, 4047, 1.09)$
S_3	$\mathcal{S}(3617, 1850, 1.10)$	$\mathcal{S}(5477, 2255, 1.53)$	$\mathcal{S}(7805, 3743, 1.10)$	$\mathcal{L}(16389, 8830)$
d_1	$\mathcal{S}(5314, 7103, 1.20)$	$\mathcal{S}(9549, 15932, 1.21)$	$\mathcal{S}(12586, 19594, 1.16)$	$\mathcal{S}(23156, 29533, 1.00)$
d_2	$\mathcal{L}(3485, 1433)$	$\mathcal{L}(5472, 2328)$	$\mathcal{L}(7532, 3238)$	$\mathcal{L}(15465, 6554)$
d_3	$\mathcal{S}(3711, 2042, 1.00)$	$\mathcal{N}(6093, 4074)$	$\mathcal{S}(8226, 5151, 0.89)$	$\mathcal{L}(16344, 8840)$
$\Delta\sigma$	$\mathcal{L}(3246, 302)$	$\mathcal{L}(5076, 581)$	$\mathcal{L}(6963, 743)$	$\mathcal{L}(14339, 1343)$

As it can be seen, the output distributions are mainly asymmetrical. It should be mentioned, again, that this results were obtained upon prescribing Normal distributions over the input features, see Table 2.

The standard deviations of the output distributions are also indicators of the propagation of input uncertainty to the output. However, varying one input at a time does not represent all possible interplay between input features.

Table B.10
Results of χ^2 test applied to the output distributions of OPT, under the sphere-like assumption for the defect shape, for samples 5–8.

Sample	5	6	7	8
$\sqrt{\text{area}}_1$	$\mathcal{L}(89731, 7562)$	$\mathcal{L}(89731, 7562)$	$\mathcal{L}(46609, 3841)$	$\mathcal{L}(917970, 68295)$
$\sqrt{\text{area}}_2$	$\mathcal{L}(89634, 9190)$	$\mathcal{L}(89634, 9190)$	$\mathcal{L}(46579, 4903)$	$\mathcal{L}(918305, 98149)$
$\sqrt{\text{area}}_3$	$\mathcal{L}(89415, 4691)$	$\mathcal{L}(89415, 4691)$	$\mathcal{L}(46465, 2714)$	$\mathcal{L}(916700, 73682)$
h_1	$\mathcal{L}(89986, 11236)$	$\mathcal{L}(89986, 11236)$	$\mathcal{L}(46767, 5986)$	$\mathcal{L}(922426, 123293)$
h_2	$\mathcal{N}(89426, 4924)$	$\mathcal{N}(89426, 4924)$	$\mathcal{L}(46461, 2649)$	$\mathcal{L}(916066, 56388)$
h_3	$\mathcal{L}(89480, 5653)$	$\mathcal{L}(89480, 5653)$	$\mathcal{L}(46485, 2972)$	$\mathcal{L}(916558, 62261)$
S_1	$\mathcal{L}(96529, 40418)$	$\mathcal{L}(96529, 40418)$	$\mathcal{L}(50104, 20569)$	$\mathcal{L}(969416, 350621)$
S_2	$\mathcal{S}(91659, 18059, 1.31)$	$\mathcal{S}(91659, 18059, 1.31)$	$\mathcal{S}(47620, 8715, 1.67)$	$\mathcal{S}(922622, 97323, 1.20)$
S_3	$\mathcal{S}(99392, 46245, 0.51)$	$\mathcal{S}(99392, 46245, 0.51)$	$\mathcal{S}(51485, 23151, 0.89)$	$\mathcal{S}(978160, 345049, 0.73)$
d_1	$\mathcal{S}(167507, 262375, 0.82)$	$\mathcal{S}(167507, 262375, 0.82)$	$\mathcal{S}(91716, 155967, 0.92)$	$\mathcal{L}(1855235, 3048763)$
d_2	$\mathcal{L}(97509, 43933)$	$\mathcal{L}(97509, 43933)$	$\mathcal{L}(50805, 23163)$	$\mathcal{L}(998035, 450412)$
d_3	$\mathcal{N}(107323, 70595)$	$\mathcal{N}(107323, 70595)$	$\mathcal{L}(56744, 39517)$	$\mathcal{L}(1137322, 835332)$
$\Delta\sigma$	$\mathcal{L}(89830, 10408)$	$\mathcal{L}(89830, 10408)$	$\mathcal{L}(46693, 5664)$	$\mathcal{L}(921605, 124453)$

Table B.11
Results of χ^2 test applied to the output distributions of OPT, under the cube-like assumption for the defect shape, for samples 9–12.

Sample	5	6	7	8
$\sqrt{\text{area}}_1$	$\mathcal{L}(3714433, 269787)$	$\mathcal{L}(4860658, 352518)$	$\mathcal{L}(9179447, 639256)$	$\mathcal{L}(9816999, 683230)$
$\sqrt{\text{area}}_2$	$\mathcal{L}(3715497, 388040)$	$\mathcal{L}(4861882, 507368)$	$\mathcal{L}(9182185, 931626)$	$\mathcal{L}(9819872, 997673)$
$\sqrt{\text{area}}_3$	$\mathcal{L}(3708490, 279134)$	$\mathcal{L}(4853075, 367050)$	$\mathcal{L}(9165463, 672612)$	$\mathcal{L}(9801079, 708177)$
h_1	$\mathcal{L}(3731636, 487294)$	$\mathcal{L}(4882832, 635725)$	$\mathcal{L}(9219949, 1168439)$	$\mathcal{L}(9860314, 1250710)$
h_2	$\mathcal{L}(3707330, 220889)$	$\mathcal{L}(4851365, 288199)$	$\mathcal{L}(9163574, 528256)$	$\mathcal{L}(9800531, 572701)$
h_3	$\mathcal{L}(3708988, 242475)$	$\mathcal{L}(4853612, 316869)$	$\mathcal{L}(9167389, 577764)$	$\mathcal{L}(9804419, 621557)$
S_1	$\mathcal{L}(3904929, 1385828)$	$\mathcal{L}(5102616, 3236595)$	$\mathcal{L}(10260840, 3483936)$	$\mathcal{L}(10260840, 3483935)$
S_2	$\mathcal{L}(3727515, 408211)$	$\mathcal{L}(4871166, 523182)$	$\mathcal{L}(9178867, 925965)$	$\mathcal{L}(9819668, 999473)$
S_3	$\mathcal{N}(3931764, 1360258)$	$\mathcal{N}(5132880, 1759424)$	$\mathcal{S}(9647538, 23216605, -0.95)$	$\mathcal{S}(10318600, 3419419, -0.89)$
d_1	$\mathcal{L}(7094206, 10695985)$	$\mathcal{L}(9197771, 13640010)$	$\mathcal{L}(17762207, 24392188)$	$\mathcal{L}(17762207, 24689935)$
d_2	$\mathcal{L}(4007700, 1745336)$	$\mathcal{L}(5241768, 2276211)$	$\mathcal{L}(10519372, 4392188)$	$\mathcal{L}(10519372, 4392188)$
d_3	$\mathcal{L}(4513029, 3134230)$	$\mathcal{L}(5905142, 4099392)$	$\mathcal{L}(11718965, 7719968)$	$\mathcal{L}(11718965, 7719967)$
$\Delta\sigma$	$\mathcal{L}(3728028, 486746)$	$\mathcal{L}(4878313, 636347)$	$\mathcal{L}(9850367, 1246696)$	$\mathcal{L}(9850367, 1246696)$

Appendix C. Pearson's correlation coefficient results

Table C.12 shows the results of PCC investigation obtained for each input quantity of each sample.

Table C.12
Results of Pearson correlation coefficient. Note that extreme values (i.e., −1 and 1) are rounded to the second decimal figure. $\Delta\sigma^*$ refers to the logarithm of $\Delta\sigma$.

Sample	$\sqrt{\text{area}}_1$	$\sqrt{\text{area}}_2$	$\sqrt{\text{area}}_3$	h_1	h_2	h_3	S_1	S_2	S_3	d_1	d_2	d_3	$\Delta\sigma$	$\Delta\sigma^*$
1	−0.99	−0.99	1.00	−0.98	−1.00	1.00	0.99	0.99	0.99	0.94	−0.99	0.98	−0.76	−1.00
2	−1.00	−0.99	1.00	−0.98	−1.00	0.99	0.99	0.99	0.99	0.93	−0.99	0.98	−0.75	−0.99
3	−1.00	−0.99	1.00	−0.98	−1.00	0.99	0.99	1.00	0.99	0.94	−0.99	0.98	−0.75	−0.99
4	−0.99	−0.99	1.00	−0.98	−1.00	1.00	0.99	1.00	0.99	0.95	−0.99	0.98	−0.78	−1.00
5	−1.00	−0.99	1.00	−0.98	−0.99	0.99	0.99	1.00	0.99	0.93	−0.99	0.97	−0.75	−1.00
6	−1.00	−0.99	1.00	−0.98	−1.00	0.99	0.99	1.00	0.99	0.93	−0.99	0.98	−0.75	−0.99
7	−1.00	−0.99	1.00	−0.98	−1.00	0.99	0.99	1.00	0.99	0.93	−0.99	0.98	−0.76	−1.00
8	−1.00	−0.99	1.00	−0.98	−1.00	0.99	0.99	1.00	0.99	0.93	−0.99	0.97	−0.75	−0.99
9	−1.00	−0.99	1.00	−0.98	−1.00	0.99	0.99	1.00	1.00	0.94	−0.99	0.98	−0.76	−1.00
10	−1.00	−0.99	1.00	−0.98	−1.00	0.99	0.99	1.00	1.00	0.93	−0.99	0.98	−0.75	−0.99
11	−1.00	−0.99	1.00	−0.98	−1.00	1.00	1.00	1.00	1.00	0.94	−0.99	0.98	−0.76	−1.00
12	−1.00	−0.99	1.00	−0.98	−1.00	1.00	1.00	1.00	1.00	0.94	−0.99	0.98	−0.75	−0.99

Appendix D. Permutation feature importance results

The present appendix summarises the outcomes of the PFI analysis. Specifically, Table D.13 gathers the computed range of MAPE (Eqs. (34)–(36)) for each input descriptor.

Table D.13
Range of MAPE for each input descriptor.

	$MAPE_{min}$	$MAPE_{mean}$	$MAPE_{max}$	Permutations
$\sqrt{area}_1$	0.40	1.43	2.17	2500
$\sqrt{area}_2$	0.36	1.62	2.46	3000
$\sqrt{area}_3$	0.40	1.57	2.57	5000
h_1	0.56	1.68	2.55	2500
h_2	0.35	1.37	2.12	5500
h_3	0.44	1.58	2.38	3000
S_1	0.51	1.89	2.79	1500
S_2	0.25	0.97	1.57	3000
S_3	0.56	2.70	4.2	4500
d_1	1.59	4.96	7.35	2500
d_2	0.73	2.75	3.96	3500
d_3	1.69	5.12	8.06	1500
$\Delta\sigma$	7.44	25.84	40.9	2500

References

[1] Sanaei N, Fatemi A. Defects in additive manufactured metals and their effect on fatigue performance: A state-of-the-art review. Prog Mater Sci 2021;117:100724. http://dx.doi.org/10.1016/j.pmatsci.2020.100724.

[2] Shao S, Poudel A, Shamsaei N. A linear elastic finite element approach to fatigue life estimation for defect laden materials. Eng Fract Mech 2023;285:109298. http://dx.doi.org/10.1016/j.engfracmech.2023.109298.

[3] Song X, Feih S, Zhai W, Sun CN, Li F, Maiti R, Wei J, Yang Y, Oancea V, Brandt LR, et al. Advances in additive manufacturing process simulation: Residual stresses and distortion predictions in complex metallic components. Mater Des 2020;193:108779. http://dx.doi.org/10.1016/j.matdes.2020.108779.

[4] Salvati E, Lunt AJG, Ying S, Sui T, Zhang HJ, Heason C, Baxter G, Korsunsky AM. Eigenstrain reconstruction of residual strains in an additively manufactured and shot peened nickel superalloy compressor blade. Comput Methods Appl Mech Engrg 2017;320:335–51. http://dx.doi.org/10.1016/j.cma.2017.03.005.

[5] Salvati E, Lunt AJG, Heason CP, Baxter GJ, Korsunsky AM. An analysis of fatigue failure mechanisms in an additively manufactured and shot peened IN 718 nickel superalloy. Mater Des 2020;191:108605. http://dx.doi.org/10.1016/j.matdes.2020.108605.

[6] Gordon JV, Haden CV, Nied HF, Vinci RP, Harlow DG. Fatigue crack growth anisotropy, texture and residual stress in austenitic steel made by wire and arc additive manufacturing. Mater Sci Eng A 2018;724:431–8. http://dx.doi.org/10.1016/j.msea.2018.03.075.

[7] Gockel J, Sheridan L, Koerper B, Whip B. The influence of additive manufacturing processing parameters on surface roughness and fatigue life. Int J Fatigue 2019;124:380–8. http://dx.doi.org/10.1016/j.ijfatigue.2019.03.025.

[8] Maleki E, Bagherifard S, Guagliano M. Correlation of residual stress, hardness and surface roughness with crack initiation and fatigue strength of surface treated additive manufactured AlSi10Mg: Experimental and machine learning approaches. J Mater Res Technol 2023;24:3265–83. http://dx.doi.org/10.1007/s11340-013-9828-0.

[9] Kruth JP, Froyen L, Vaerenbergh J Van, Mercelis P, Rombouts M, Lauwers B. Selective laser melting of iron-based powder. J Mater Process Technol 2004;149(1–3):616–22. http://dx.doi.org/10.1016/j.jmatprotec.2003.11.051.

[10] Zhang B, Li Y, Bai Q. Defect formation mechanisms in selective laser melting: a review. Chin J Mech Eng 2017;30:515–27. http://dx.doi.org/10.1007/s10033-017-0121-5.

[11] Maleki E, Bagherifard S, Unal O, Jam A, Shao S, Guagliano M, Shamsaei n. Superior effects of hybrid laser shock peening and ultrasonic nanocrystalline surface modification on fatigue behavior of additive manufactured AlSi10Mg. Surf Coat Technol 2023;463:129512. http://dx.doi.org/10.1016/j.surfcoat.2023.129512.

[12] Laleh M, Sadeghi E, Revilla RI, Chao Q, Haghdadi N, Hughes AE, Xu W, Graeve I De, Qian M, Gibson I, et al. Heat treatment for metal additive manufacturing. Prog Mater Sci 2022;101051. http://dx.doi.org/10.1016/j.pmatsci.2022.101051.

[13] Yang G, Xie Y, Zhao S, Qin L, Wang X, Wu B. Quality control: Internal defects formation mechanism of selective laser melting based on laser-powder-melt pool interaction: A review. Chin J Mech Eng: Addit Manuf Front 2022;100037. http://dx.doi.org/10.1016/j.cjmeam.2022.100037.

[14] Zerbst U, Madia M, Klinger C, Bettge D, Murakami Y. Defects as a root cause of fatigue failure of metallic components. I: Basic aspects. Eng Fail Anal 2019;97:777–92. http://dx.doi.org/10.1016/j.engfailanal.2019.01.055.

[15] Murakami Y. Material defects as the basis of fatigue design. Int J Fatigue 2012;41:2–10. http://dx.doi.org/10.1016/j.ijfatigue.2011.12.001.

[16] Tridello A, Fiocchi J, Biffi CA, Chiandussi G, Rossetto M, Tuissi A, Paolino DS. Effect of microstructure, residual stresses and building orientation on the fatigue response up to 109 cycles of an SLM AlSi10Mg alloy. Int J Fatigue 2020;137:105659. http://dx.doi.org/10.1016/j.ijfatigue.2020.105659.

[17] Kitagawa H. Applicability of fracture mechanics to very small cracks or the cracks in the early stage. In: Proc. 2nd int. conf. on mechanical behaviour of materials. 1976, p. 627–31.

[18] Haddad MH El, Smith KN, Topper TH. Fatigue crack propagation of short cracks. J Eng Mater Technol 1979. http://dx.doi.org/10.1115/1.3443647.

[19] Chapetti MD. Fatigue propagation threshold of short cracks under constant amplitude loading. Int J Fatigue 2003;25(12):1319–26. http://dx.doi.org/10.1016/S0142-1123(03)00065-3.

[20] Murakami Y. Metal fatigue: Effects of small defects and nonmetallic inclusions. Academic Press; 2019.

[21] Romano S, Brückner-Foit A, Brandão A, Gumpinger J, Ghidini T, Beretta S. Fatigue properties of AlSi10Mg obtained by additive manufacturing: Defect-based modelling and prediction of fatigue strength. Eng Fract Mech 2018;187:165–89. http://dx.doi.org/10.1016/j.engfracmech.2017.11.002.

[22] Zerbst U, Bruno G, Buffiere JY, Wegener T, Niendorf T, Wu T, Zhang X, Kashaev N, Meneghetti G, Hrabe N, et al. Damage tolerant design of additively manufactured metallic components subjected to cyclic loading: State of the art and challenges. Prog Mater Sci 2021;121:100786. http://dx.doi.org/10.1016/j.pmatsci.2021.100786.

[23] Chen J, Liu Y. Fatigue modeling using neural networks: A comprehensive review. Fatigue Fract Eng Mater Struct 2022;45(4):945–79. http://dx.doi.org/10.1111/ffe.13640.

[24] Wang L, Zhu SP, Luo C, Liao D, Wang Q. Physics-guided machine learning frameworks for fatigue life prediction of AM materials. Int J Fatigue 2023;172:107658. http://dx.doi.org/10.1016/j.ijfatigue.2023.107658.

[25] Tognan A, Salvati E. Probabilistic defect-based modelling of fatigue strength for incomplete datasets assisted by literature data. Int J Fatigue 2023;173:107665. http://dx.doi.org/10.1016/j.ijfatigue.2023.107665.

[26] Zhan Z, Li H. A novel approach based on the elastoplastic fatigue damage and machine learning models for life prediction of aerospace alloy parts fabricated by additive manufacturing. Int J Fatigue 2021;145:106089. http://dx.doi.org/10.1016/j.ijfatigue.2020.106089.

[27] Konda N, Verma R, Jayaganthan R. Machine Learning based predictions of fatigue crack growth rate of additively manufactured Ti6Al4V. Metals 2021;12(1):50. http://dx.doi.org/10.3390/met12010050.
[28] Li A, Baig S, Liu J, Shao S, Shamsaei N. Defect criticality analysis on fatigue life of L-PBF 17-4 PH stainless steel via machine learning. Int J Fatigue 2022;163:107018. http://dx.doi.org/10.1016/j.ijfatigue.2022.107018.
[29] Bao H, Wu S, Wu Z, Kang G, Peng X, Withers PJ. A machine-learning fatigue life prediction approach of additively manufactured metals. Eng Fract Mech 2021;242:107508. http://dx.doi.org/10.1016/j.engfracmech.2020.107508.
[30] Moon S, Ma R, Attardo R, Tomonto C, Nordin M, Wheelock P, Glavicic M, Layman M, Billo R, Luo T, Tengfei. Impact of surface and pore characteristics on fatigue life of laser powder bed fusion Ti–6Al–4V alloy described by neural network models. Sci Rep 2021;11(1):20424. http://dx.doi.org/10.1038/s41598-021-99959-6.
[31] He L, Wang Z, Ogawa Y, Akebono H, Sugeta A, Hayashi Y. Machine-learning-based investigation into the effect of defect/inclusion on fatigue behavior in steels. Int J Fatigue 2022;155:106597. http://dx.doi.org/10.1016/j.ijfatigue.2021.106597.
[32] Peng X, Wu s, Qian W, Bao J, Hu Y, Zhan Z, Guo G, Withers PJ. The potency of defects on fatigue of additively manufactured metals. Int J Mech Sci 2022;221:107185. http://dx.doi.org/10.1016/j.ijmecsci.2022.107185.
[33] Salvati E, Tognan A, Laurenti L, Pelegatti M, Bona F De. A defect-based physics-informed machine learning framework for fatigue finite life prediction in additive manufacturing. Mater Des 2022;222:111089. http://dx.doi.org/10.1016/j.matdes.2022.111089.
[34] Zhou T, Jiang S, Han T, Zhu SP, Cai Y. A physically consistent framework for fatigue life prediction using probabilistic physics-informed neural network. Int J Fatigue 2023;166:107234. http://dx.doi.org/10.1016/j.ijfatigue.2022.107234.
[35] Chen J, Liu Y. Probabilistic physics-guided machine learning for fatigue data analysis. Expert Syst Appl 2021;168:114316. http://dx.doi.org/10.1016/j.eswa.2020.114316.
[36] He GY, Zhao YX, Yan CL. MFLP-PINN: A physics-informed neural network for multiaxial fatigue life prediction. Eur J Mech A Solids 2023;98:104889. http://dx.doi.org/10.1016/j.euromechsol.2022.104889.
[37] Ciampaglia A, Tridello A, Paolino DS. Data driven method for predicting the effect of process parameters on the fatigue response of additive manufactured AlSi10Mg parts. Int J Fatigue 2023;107500. http://dx.doi.org/10.1016/j.ijfatigue.2023.107500.
[38] Zhu SP, Liu Q, Peng W, Zhang XC. Computational-experimental approaches for fatigue reliability assessment of turbine bladed disks. Int J Mech Sci 2018;142:502–17. http://dx.doi.org/10.1016/j.ijmecsci.2018.04.050.
[39] Raissi M, Perdikaris P, Karniadakis GE. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J Comput Phys 2019;378:686–707. http://dx.doi.org/10.1016/j.jcp.2018.10.045.
[40] Karniadakis GE, Kevrekidis IG, Lu L, Perdikaris P, Wang S, Yang L. Physics-informed machine learning. Nat Rev Phys 2021;3(6):422–40. http://dx.doi.org/10.1038/s42254-021-00314-5.
[41] Tognan A, Laurenti L, Salvati E. Contour method with uncertainty quantification: a robust and optimised framework via gaussian process regression. Exp Mech 2022;62(8):1305–17. http://dx.doi.org/10.1007/s11340-022-00842-w.
[42] Cochran WG. Some methods for strengthening the common χ 2 tests. Biometrics 1954;10(4):417–51. http://dx.doi.org/10.2307/3001616.
[43] Schorr B. On the choice of the class intervals in the application of the chi-square test. Math Operforsch Stat 1974;5(4–5):357–77. http://dx.doi.org/10.1080/02331887408801174.
[44] Hooker G, Mentch L, Zhou S. Unrestricted permutation forces extrapolation: variable importance requires at least one more model, or there is no free variable importance. Stat Comput 2021;31:1–16. http://dx.doi.org/10.1007/s11222-021-10057-z.
[45] Apley DW, Zhu J. Visualizing the effects of predictor variables in black box supervised learning models. J R Stat Soc Ser B Stat Methodol 2020;82(4):1059–86. http://dx.doi.org/10.1111/rssb.12377.
[46] Smith DL, Naberejnev DG. Confidence intervals for the lognormal probability distribution. Nucl Instrum Methods Phys Res A 2004;518(3):754–63. http://dx.doi.org/10.1016/j.nima.2003.09.068.
[47] Ashour SK, Abdel-hameed MA. Approximate skew normal distribution. J Adv Res 2010;1(4):341–50. http://dx.doi.org/10.1016/j.jare.2010.06.004.