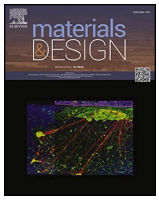
Materials & Design 248 (2024) 113449

Contents lists available at ScienceDirect

Materials & Design

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



# Machine learning powered predictive modelling of complex residual stress for nuclear fusion reactor design

Bin Zhu [a], Nathanael Leung [a], Brandon Steel [a,b], David England [a,b], Yinglong He [a,d], Andrew J. London [c], Hannah Zhang [b], Michael Gorley [c], Yiqiang Wang [c], Mark J. Whiting [a], Tan Sui [a,b,*]

[a] School of Mechanical Engineering Sciences, University of Surrey, Guildford, Surrey GU2 7XH, UK
[b] National Physical Laboratory, Hampton Road, Teddington TW11 0LW, UK
[c] United Kingdom Atomic Energy Authority, Culham Campus, Abingdon, Oxon OX14 3DB, UK
[d] Department of Mechanical Engineering, University of Birmingham, Birmingham B15 2TT, UK

ARTICLE INFO

*Keywords:*
Machine learning
Eurofer97 steel
Residual stress heterogeneity and anisotropy
Nanoindentation
Microstructures

ABSTRACT

Fusion In-vessel components, assembled and maintained using laser welding, one of the most promising techniques, exhibit complex distributions of residual stress, microstructures, and material properties. These residual stresses can compromise structural integrity and lifespan of critical components. Although using advanced experimental measurements can evaluate the residual stress for individual case, extending the measurements to massive number of components are costly and time-consuming. Traditional machine learning (ML) models struggle to account for the heterogeneity and anisotropy of these stress distributions. Here, we develop a novel ML framework based on the Eurofer97 steel, the structural material for in-vessel components. The ML framework is trained on high-resolution residual stress data derived from recently-developed evaluation techniques. Combining with microstructures, the model enables prediction of heterogenous and anisotropic residual stress distribution. It successfully predicts the compressive residual stress in fusion zone (~−200 MPa) balanced by tensile residual stress in heat affected zone (~300 MPa), aligning closely with experimental results with the *R*-squared value of 0.989 and the mean square error of $10^{-4}$. Unlike experiments that take hours, the ML model provides predictions within seconds. It offers valuable insights into residual stress prediction for various joints, enhancing the reliability and lifetime prediction of in-vessel components.

## 1. Introduction

Residual stresses, which are internal stresses that persist within a material after the completion of manufacturing processes, can originate from various sources. Eurofer97 steel is one of the most promising candidate structural materials for constructing a wide range of in-vessel components in the European Demonstration Fusion Power Plant [1], including pipes, breeding blanket and divertor cassette (Fig. 1(A)). To assemble and maintain these components, the remote laser welding is one of the most promising techniques, which induces remarkable thermal gradients, plastic deformation, phase transformations and material incompatibilities within the welded regions [2,3]. This, in turn, gives rise to substantial tensile residual stresses, often reaching levels as high as 800 MPa [4–6]. The induced residual stresses in these joints, when left unaddressed, can significantly diminish both the material performance and longevity of the engineering components. Given the high cost and long leading time of the experimental methods, there is a pressing need for a rapid and reliable computational method of evaluating residual stresses to comprehensively explore the quantitative impact of residual stress on structural integrity and to optimise welding processing.

Machine learning (ML) and deep learning (DL) models are becoming increasingly popular as the basis for data-driven models to predict the mechanical response of materials according to their manufacturing processes and microstructure [7–11]. Compared to finite element modelling, which usually relies on linear mathematical equations and simplifications, ML/DL can capture complex nonlinear relationships between inputs and outputs [12]. A recent review underscores the growing importance of models such as random forest regression (RFR), artificial neural networks (ANN), and convolutional neural networks (CNN), due to their ability to predict complex mechanical behaviour

* Corresponding author at: School of Mechanical Engineering Sciences, University of Surrey, Guildford, Surrey GU2 7XH, UK.
*E-mail address:* t.sui@surrey.ac.uk (T. Sui).

https://doi.org/10.1016/j.matdes.2024.113449
Received 26 September 2024; Received in revised form 4 November 2024; Accepted 8 November 2024
Available online 9 November 2024

[13]. Whilst the ANN and CNN approaches have proven to be powerful tools for establishing complex nonlinear relationships, such as predicting residual stress in alloys [14], they are susceptible to overfitting, especially when the training dataset is insufficient. In contrast, RFR offers a more interpretable and robust alternative, with a lower risk of overfitting. It has been successfully used to predict yield strength and elongation in reduced activation ferritic/martensitic (RAFM) steel, demonstrating strong overall performance and lower prediction errors compared to other ML/DL algorithms [13,15].

By employing experimental data, such as that from neutron diffraction, the contour method and deep hole drilling [16,17], the ML/DL models can be established to predict the low-resolution residual stress distribution in the weldment from the manufacturing conditions. However, the important reason for originating residual stress is due to the mismatch caused by microstructure changes after manufacturing. Such models, which only use manufacturing conditions, can overlook the grain-level microstructure effects on stress anisotropy, which significantly influences the mechanical properties in complex welding system. In addition, the prediction of the high-resolution heterogeneous residual stress distribution is limited by the resolution of residual stress measurement techniques (volume-averaged residual stress value). It is both timely and crucial to develop a novel machine learning predictive model, trained on data obtained from advanced characterisation techniques, that accounts for the residual stress heterogeneity and anisotropy.

A well-constructed training database, capable of capturing heterogeneous distributions at high resolution, is critical for the development of novel ML predictive models. For fast and reliable predictions, it is essential that the input data can be obtained through cost-effective methods, while the key outputs—such as the heterogeneity and anisotropy of residual stress—must be evaluated with precision. Instrumented nanoindentation, probing only a small amount of material, meets these criteria and has been widely used to assess high-resolution heterogeneity, such as variations in elastic modulus and hardness using the Oliver-Pharr method [18,19]. The load–displacement data generated from the nanoindentation has attracted increasing interest to train the ML model [5,20–22]. An example can be found through the development of ML model to address the nonlinearity based on the experimental investigation of the localised elastoplastic mechanical properties in a small region, thereby enabling a highly accurate prediction of the tensile stress [23–25].

Additionally, existing ML models for residual stress prediction are typically limited to isotropic stress states [26], given the training targets of the anisotropy of residual stress, especially in complex welded joints, are lacking. Our recent studies have revealed the anisotropic residual stress state and its relationship with microstructures and mechanical performance by employing the plasma focus ion beam and digital image correlation (PFIB-DIC) method [5,27]. A comprehensive examination of the omnidirectional residual stress fields from PFIB-DIC provides sufficient statistics data to train a ML/DL model. This approach not only overcomes the main limitation in measuring directions of the conventional nanoindentation residual stress evaluation but also yields novel ML/DL models to predict in-plane omnidirectional residual stress.

In this study, we establish new ML/DL models that associate the nanoindentation mechanical response, experimentally obtained residual stress and microstructure distribution. A laser-welded Eurofer97 joint

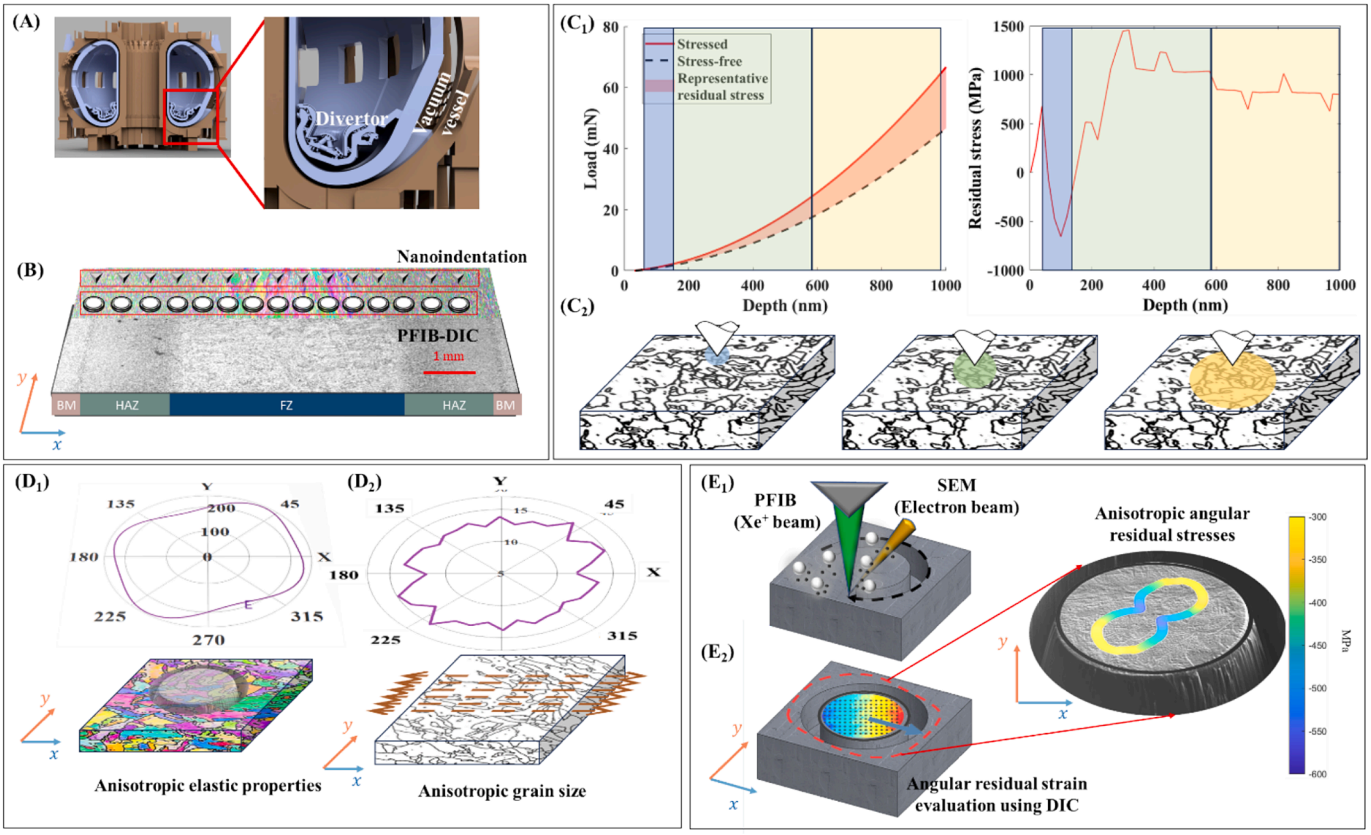



**Fig. 1.** Experimental data acquisition, feature selection and data processing workflow. (A) The in-vessel components containing laser-welded Eurofer97 are highlighted in purple. (B) Illustration of the laser-welded Eurofer97 sample that is being used to establish ML/DL residual stress predictive models. (C) Nanoindentation data acquisition and processing. The selection range of the input features for the ML/DL model, which were highlighted in the yellow box region in ($C_1$), was determined by ($C_2$) the initial estimation of number of grains involved during indentation. (D) EBSD data processing for angular distribution of material anisotropic elastic properties ($D_1$) and grain size ($D_2$). (E) The PFIB-DIC measurement for evaluating the residual stress. ($E_1$) demonstrates the PFIB incremental milling procedures schematically and ($E_2$) shows the deformation of the ring-core pillars after milling and analysis of the angular residual strain and stress.

was examined to evaluate its microstructure, mechanical properties and residual stress, providing essential training data for the ML/DL models. The experimental data was transformed to an angular basis using a modified algorithm based on previous studies [27]. The material anisotropy, mechanical responses from nanoindentation, and angular residual stress from the PFIB-DIC method were obtained. Predictive models, including RFR, ANN and CNN, were used to forecast the local heterogeneous residual stress in different directions. Successful implementation of the model would contribute to developing a framework for predicting experimental trends in materials with complex heterogeneous microstructures, ultimately promoting the welding or other joining applications in nuclear fusion reactors.

## 2. Training database collection from advanced characterisations

The two main parts of establishing the ML/DL models include the training data preparation and determination of training data architecture and ML/DL hyperparameters. In this model, the database was collected from the laser-welded Eurofer97 joint, which was butt welded with a 5 kW Yb-fibre single laser source and a welding speed of 1.2 m/min by the TWI. The SEM images of the joint region can be found in Fig. 1(B) The advanced characterisation techniques were used for database collection, including electron backscatter diffraction (EBSD), nanoindentation, and PFIB-DIC. For clarity regarding experimental and model design, an overview of the experiments is shown in Fig. 1(B). The EBSD was used to generate an orientation map for analysis of material microstructures and anisotropies across an area covering both the fusion zone (FZ) and heat-affected zone (HAZ) regions of the weldment. Additionally, in the area encompassed by the EBSD maps, the heterogeneous residual stress distribution was evaluated by PFIB-DIC, with a point-to-point spacing of 200 μm across the weldment. The nanoindentation measurements were conducted to collect loads at different depths beside the PFIB-DIC measurement with a 30 μm offset to avoid interference. The detailed experiment and data interpretation methods can be found in the following sections.

### 2.1. Nanoindentation data collection and interpretation

Nanoindentation was conducted using a Berkovich tip within an Agilent G200 located at the Materials Research Facility, UK Atomic Energy Authority (UKAEA). The collection of data was completed under displacement control utilising continuous stiffness measurement to a depth of 1000 nm at 2 nm/s [28]. For comparative results, the indents were gathered in similar manner across the stressed joint at a spacing of 200 μm steps, and on two stress-free pillars in the FZ and HAZ respectively.

The selection of specific nanoindentation loads and depths is pivotal in establishing the workflow for ML/DL models. Given that the region of deformation caused by nanoindentation extends beyond the size of the indent itself, it can be representative of different length scales of residual stress [29]. Beneath the nanoindentation tip, deformation occurs in several zones: core, plastic, and elastic. The core and plastic zones typically extend to the spherical regions with a radius of approximately 10 times the penetration depth for a Berkovich tip [30], whilst the elastic zone is even larger, though not quantitatively measured. Located within this affected volume, for the laser-welded Eurofer97 joint, tens of surrounding grains are likely involved in responding to the indent load when the depth reaches 1 μm (the radius of the spherical plastic zone can be up to 10 μm). Quantitatively identifying the transition boundary from micro- to macro-scale deformation is challenging, as the size of the deformation zone varies based on material hardness and the level of macroscale residual stress, as demonstrated in Fe material [31]. In this study, such selection was determined from equi-biaxial principal residual stress values calculated from the load-depth curve in Fig. 1($C_1$) [20]. Although this conventional measurement is susceptible to errors in contact area, it provides valuable initial evidence for feature selection. As shown in the residual stress versus depth curve within Fig. 1($C_1$), the results via the conventional model, initially exhibited noticeable fluctuations at shallower depths but gradually stabilised as the depth increased. During the analysis of the profiles, three distinct depth-related domains were revealed and categorised as follows: The noise domain marked in blue, the microscale domain highlighted in green, and the macroscale domain in yellow. The noise domain has the largest variability, and the mechanical response is influenced by numerous factors, for example, the indentation size, surface roughness, surface precipitates, and tip blunting [32]. At greater depths, their impact is less as the microdomain is entered. As shown schematically in Fig. 1($C_2$), only a limited number of grains contributed to accommodating the indent-induced deformation in this domain, signifying microscale stress levels. The indentation depth then increases further and reaches a steady-state residual stress region, known as the macro domain, explained by a statistical convergence resulting from the increased number of deforming grains. Therefore, the assumption can be made that the loads at various depths would yield identical residual stress values, given that the individual load values within the macroscale domain are indicative of the same macroscale residual stress value. To merge data from various sources for model training, the load intervals must match the number of angle intervals. In the context of this ML/DL model, the load for both measured and reference data (from stress-free pillars) acquired at multiple depths (highlighted in yellow within Fig. 1($C_1$)) was employed as the inputs for the ML/DL models.

### 2.2. EBSD data collection and interpretation

The material microstructures and anisotropies were characterised using a Jeol-7100F scanning electron microscope equipped with an EBSD detector (a Thermo Fisher Lumis system), at the University of Surrey. EBSD data was acquired with an accelerating voltage of 20 kV, a beam current of 12 nA, and an exposure time of ~20 ms. Pixel binning of 2 × 2 and step size of 0.5 μm was used to collect the EBSD orientation maps (512 × 384 pixels). The grain orientation and boundaries were effectively reconstructed by interpreting the orientation maps, employing the MATLAB package MTEX version 5.2.8 [33]. To enhance precision and reduce random errors, noise reduction techniques based on nearest neighbour algorithms were applied. The EBSD data for the area of 60 × 60μm$^2$ surrounding the PFIB-DIC measurement region was extracted to investigate the localised material anisotropic elastic properties and grain size.

The texture determined by EBSD was further interpreted to calculate the Young's modulus and Poisson's coefficients in a certain direction. Utilising the stiffness matrix associated with Eurofer97 material [34], they were rigorously computed in an iterative manner, with the EBSD maps being systematically rotated at angular intervals of 1° within the angle range from 1° to 360°. The Chill tensor average method, which computes the arithmetic mean between Voigt upper and Reuss lower bounds, was integrated into MTEX for this purpose [35]. Fig. 1($D_1$) shows the example of Young's modulus for a ring-core measurement position in the FZ region. This method was extended to all positions where the PFIB-DIC ring-core measurements were applied.

The grain size was determined using the mean linear intercept method on the EBSD orientation maps [36,37]. The automation of grain size determination was achieved via MATLAB, which involved the identification of grain boundaries and the subsequent quantification of their occurrences in binary images. During identification, a thickness offset was introduced to reduce the error from the width of the grain boundaries. The automated counting mechanism from previous studies was adapted to avoid overestimating the number of grain boundaries, including the identification of junction and end intercepts [38]. Fig. 1 ($D_2$) schematically shows that the intercepts were systematically identified along ten lines. The mean grain size was then evaluated according to the method in the ASTM E112-13 [37]. To calculate the angular grain size, the automated detection was achieved with angular intervals of 1°.

An example of an automated measurement result shows that the angular distribution of grain sizes for the region around a ring-core in the FZ region was in a range from 12 to 16 μm. The grain size obtained here using the automated detection processes in the *x*-direction was consistent with previous findings [3].

### 2.3. PFIB-DIC residual stress data collection and interpretation

The residual stress in the target of the ML/DL model was evaluated by the PFIB-DIC ring-core technique performed on the TESCAN Fera III $Xe^+$ PFIB-SEM at the University of Surrey. The sample was polished and etched with Vilella's reagent, whereby the features of grain boundaries or precipitates on the surface after echting are sufficient to achieve high-quality DIC correlations [27]. As shown in Fig. 1($E_1$), with a 30 keV ion beam at 20 nA, a 30 μm pillar was generated by milling an annulus in 50 increments until complete strain relief was achieved in the pillar, and 10 high-resolution images were collected at each increment [39]. For the residual strain evaluation, an array of ring-cores was collected across the joint with a spacing of 200 μm shown in Fig. 1(B).

While the PFIB removes the materials and fabricates the ring-core geometry, the results from this method yield a two-dimensional displacement field, shown in Fig. 1($E_2$). This is caused by the stress relief deformation arising from the difference between the original ring-core design, indicated by the black circle, and the final shape of the pillar, shown by the dashed red line. A prior processing of coordinate transformation was applied to the correlation data following the strain calculation using the MATLAB-based DIC [40]. This expands the strain calculation to encompass all in-plane angular orientations from a single direction. Generalised Hooke's Law was applied to convert the relaxed strains into residual stresses, employing the elastic constants of the Eurofer97 steel to calculate the anisotropic Young's modulus (E) and Poisson's coefficients ($\nu$) for all in-plane angular directions from the EBSD within the same region. An example of in-plane omnidirectional residual stress result in the FZ region is shown in Fig. 1($E_2$), where a compressive residual stress, up to -600 MPa, was observed close to the *x*-direction. Given the inherent characteristics of this measurement method, the PFIB-DIC with EBSD gives a dependable result for assessing residual stresses in specimens that contain a high degree of heterogeneity in their mechanical properties.

## 3. ML/DL model establishments

### 3.1. Construction of ML/DL database

In this section, detailed architecture of the database and the selection of hyperparameters for the ML/DL model is shown. Fig. 2 shows the maps of experimentally measured distribution using the motheds above regarding material anisotropies, grain size and residual strain and stress versus angles at all the measurement positions. As shown in Fig. 2(A), the database preparation was extended across the welding region by using the method developed above, where the material anisotropies (Young's modules and Possion's ratios) that were derived from the EBSD orientation maps and the residual strain obtained from the PFIB-DIC measurements were used to calculate the omnidirectional residual stress (Fig. 2(B)). To leverage ML/DL for predicting residual stress, the RFR and two neural networks, ANN and CNN, served as predictive models (Fig. 2(B)). As shown in Fig. 2(C), eight features, including distance to the welding centre ($C$), grain size ($GS$), anisotropic Young's modulus ($E$), nanoindentation load at measured and stress-free position ($L$ and $L_R$), angle ($\theta$), depth ($d$) and anisotropic Poisson's ratio ($v$), were selected as input and organised in a tabular format. The demonstration of converting from the results of data interpretation (the nano-indentation load maps and grain size maps) to the input training data was shown in Fig. 2(C). The relationship between the input data and the target residual stress can be approximated by a predictive model. To make full use of the power of neural networks, a sub-window, which combines data in multiple directions forming a grid-like input, was introduced and assigned to a single target to improve the predicting accuracy (Fig. 2(C)).

### 3.2. Construction of ML/DL models

To improve predictive accuracy and reliability, the normalising and randomising processes were applied across all models. The datasets were subjected to normalisation to eliminate scale invariance and enhance numerical stability. Subsequently, the datasets underwent randomisation and were divided into training, validation, and testing sets with ratios of 0.7, 0.15, and 0.15, respectively. Backpropagation was employed in each model, whereas weights associated with nodes were iteratively updated based on an error function. A batch size of 32 was employed, indicating the number of samples processed before backpropagation-initiated weight updates. The mean squared error (MSE) and R-square ($R^2$) were then used to examine the models' performance. Both metrics provide single values that enable direct comparison of different algorithms, offering a clear indication of their predictive accuracy. MSE quantifies the average squared difference between the experimental and predicted values, with a lower MSE indicating a better fit. It is calculated using Eq. (1) [41]:

$$MSE = \frac{\sum_{i=1}^{n}(y_i - \widehat{y_i})^2}{n} \tag{1}$$

where $y_i$ is the $i$-th experimental value, $\widehat{y_i}$ is the $i$-th predicted value and $n$ is the total number of data points.

$R^2$ measures the proportion of variance in the experimental data that is explained by the predictions. It ranges from 0 to 1, with 0 indicating no correlation and 1 indicating perfect correlation. It is defined using Eq. (2) [41]:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \widehat{y_i})^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2} \tag{2}$$

where $\overline{y}$ is the mean of the experimental values.

To build the DL models, ANN and CNN algorithms were employed via TensorFlow, a Python-based open-source ML library. Each of the hidden layers containing nodes, is equipped with an activation function that operates on incoming data to determine the output. The ANN in this study incorporated six layers, comprising 32, 64, 128, 64, 32, and 16 nodes, as empirical findings indicated stability with these parameters. The Rectified Linear Unit (ReLU) function was chosen as the activation function for all nodes to introduce non-linearity into the model, effectively capturing the complex relationship between the inputs and the target outputs. Furthermore, to explore any spatial associations between the inputs and target, various feature combinations were investigated. Data sub-window functions were employed, enabling the combination of 1–9 input data sets (indicated by red dash box), which were then assigned to a single target data set (blue dash box), as shown in Fig. 2(C).

Another commonly used DL model, the CNN, was applied, which shares similarities with the ANN but introduces an additional convolutional layer at the beginning of the model to further investigate any spatial relations. This layer treats the input angle as an image, denoted as the feature vector. Two additional hyperparameters are associated with this convolutional layer, specifically the kernel size and the number of filters. The kernel size corresponds to the length of the 1D convolutional window, equal to the previously discussed data window size, whereas the number of filters determines the dimension of the convolutional output, set at 32, aligning with the batch size. The activation function chosen for this layer was also ReLU. Following this convolutional layer, the architecture was identical to the ANN.

The RFR was also employed in this study via Scikit, another Python-based ML library. To remove bias, the predictions from each decision tree were averaged. The sampling procedure led to the exclusion of some

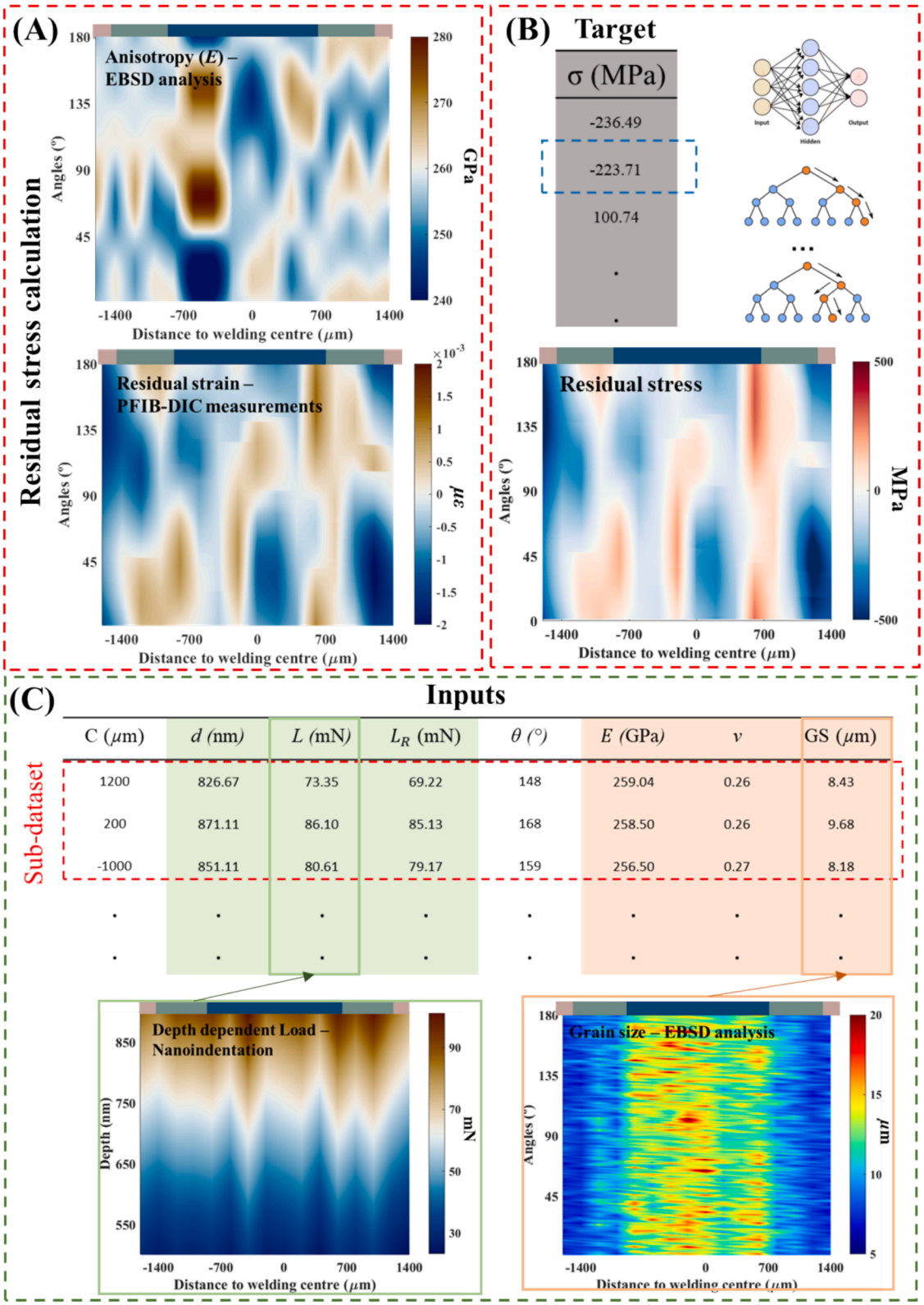



**Fig. 2.** Training data and workflow for training and validating the ML/DL models. (A) Distribution of anisotropic material properties obtained from EBSD analysis and non-equi-biaxial residual strain obtained from PFIB-DIC evaluation. (B) The target of the training data, omnidirectional residual stress calculated using the data from (A), and the architectures of two neural networks (NNs) and random forest regression (RFR). (C) Input features were selected from the experimental data and organised in a tabular format. The nanoindentation load and grain size are mapped versus angles and distance to the welding centre. The sub-window, chopping the input data into segments of consecutive samples, as indicated by the red dash box. The blue dashed box in (B) indicates the assigned target corresponding to the input data selected by the sub-window.

training data in each subset, referred to as out-of-bag (OOB) data. Leveraging this OOB data, RFR conducts error analysis in parallel and assesses the importance of each input variable without the necessity for additional validation data. The MSE and $R^2$ were then calculated across an increasing number of trees (0–200), with the integration of the grid search method to optimise other hyperparameters. After the training phase, the RFR and best-performed ANN model underwent validation. The validation of angular predictions was subsequently conducted by employing a measurement corresponding to an individual testing position within the FZ region, whilst the validation of distribution prediction was achieved using all the data without randomisation.

## 4. Results

### 4.1. Performance of the ML/DL models

The performance of the models was evaluated using MSE and $R^2$ values and summarised in Fig. 3. As shown in Fig. 3(A), notably the results exhibit improved performance with an increase in $R^2$ and a concomitant reduction in MSE, up to a threshold of ~100 trees. Beyond this point, the values stabilise at slightly above 0.99 for $R^2$ and approximately $3 \times 10^{-4}$ for MSE. Further expansion of the number of trees would increase computational costs with diminishing returns. In addition, a distinctive attribute of the RFR model is its capacity to ascertain the significance of different features employed in the analysis, as illustrated in Fig. 3(B). In this context, the distance to the welding centre (*C*) emerges as most influential, attaining a value of approximately 0.45. The grain size (*GS*) has the second ranking of importance with a value of approximately 0.2, whereas the third important factor, anisotropic Young's modulus (*E*), has a value of 0.1. The values of the remaining features, nanoindentation load at measured and stress-free position ($L$ and $L_R$), angle ($\theta$), depth ($d$) and anisotropic Poisson's ratio ($\nu$) exhibit comparable importance values, suggesting equal predictive utility for the model.

As shown in Fig. 3(C), the impact of varying sub-window sizes on the determination of an optimal model architecture for predicting residual stress was investigated, as well as the choice of neural network types. Although the design of the sub-window aims to improve the performance of the framework, it is seen that increasing the sub-window size leads to an escalation in the MSE values for both ANN and CNN models. This might be attributed to the anisotropic properties, on which the residual stress distribution is directionally dependent. Specifically, the model showed the best training performance without using the sub-window setup, and the MSE values for ANN and CNN are just below 0.02 and 0.03, respectively. The $R^2$ values obtained are 0.989 for ANN and 0.978 for CNN. Looking at the comparison between the two neural networks (Fig. 3(C)), both MSE and $R^2$ indicate that the ANN outperformed CNN. This outcome was observed even though the multiple datasets were organised into sub-windows, indicating the grid-like data format did not fully leverage the strengths of CNN. It is noteworthy that the inclusion of sub-windows had a relatively minor impact on the ANN model. Fig. 3(D) presents the evolution of the MSE for training, validation, and testing datasets for an increasing number of epochs for the best performance model in Fig. 3(C) (ANN model without sub-window). It is observed that the MSE converges to a value of 0.024 after 32 epochs, and the highest $R^2$ of 0.971 for the testing datasets. An example of validation for the ANN and RFR models was given at an individual residual stress measurement using a comparative analysis between

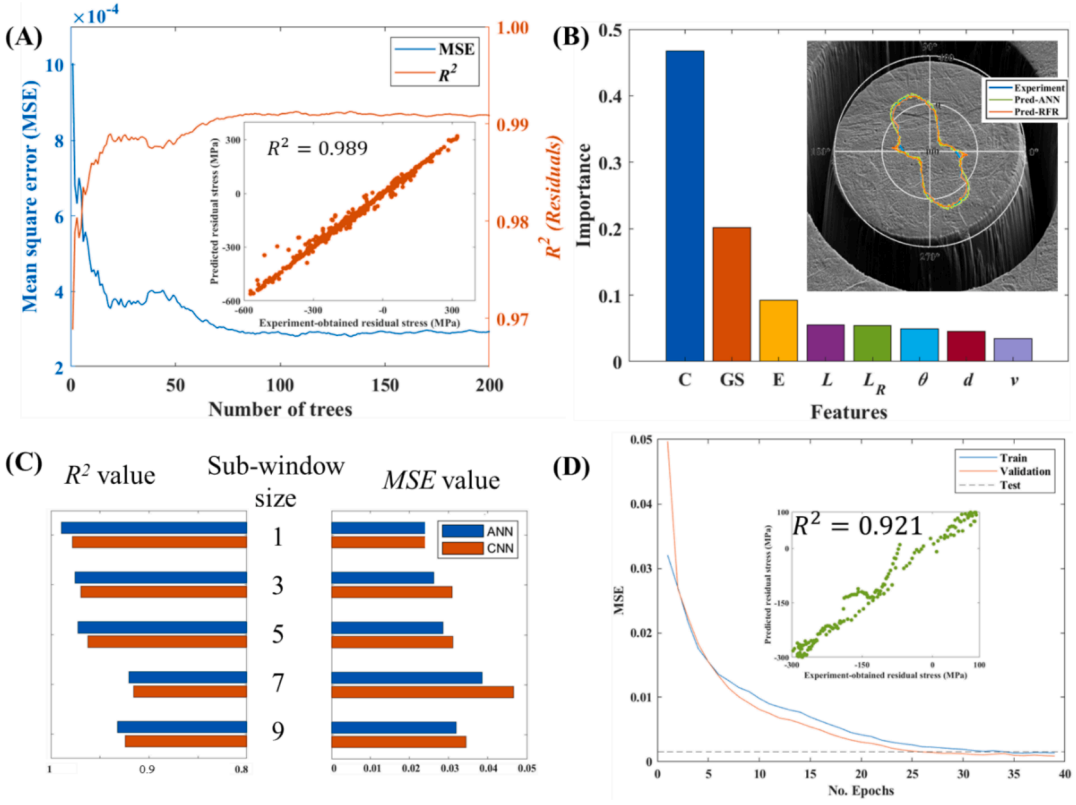



**Fig. 3.** Results of ML/DL models' performance. (A) The learning curve of the RFR model, showing the evolution of MSE as a function of the number of trees and the R-square value at the optimised hyperparameters. (B) Ranking of importance for input features, i.e., welding centre (*C*), grain size (*GS*), anisotropic Young's modulus (*E*), nanoindentation load at measured and stress-free position ($L$ and $L_R$), angle ($\theta$), depth ($d$) and anisotropic Poisson's coefficient ($\nu$), in the RFR model. Insert figure shows the angular residual stress at individual measurements predicted by RFR and ANN. (C) Comparison between two neural networks, ANN and CNN, regarding MSE and R-square values at different sub-window sizes. (D) The learning curve of the best performance model in (C), ANN model without sub-window, showing the evolution of MSE for the training, validation and testing dataset for the sub-window size of 1.

experimental results and model predictions, as shown in the insert image in Fig. 3(B). Both models demonstrate excellent agreement with the experimental data, displaying only slight discrepancies at approximately 0 and 180 degrees.

### 4.2. Validation of ML/DL model using experiment-obtained results

The predicted results were validated using the experimental results gathered from the laser-welded Eurofer97 joint. In Fig. 4(A) and (B), the maps indicate the predicted residual stresses obtained from the RFR and ANN models. This evaluation was performed with respect to the distance from the welding centre and for multiple angles. Similar residual stress values and distributions can be seen from the two predictions, whilst the map of the ANN prediction (Fig. 4(B)) shows a higher noise level. This contrast in predictive accuracy is further corroborated by comparing with the experiment-obtained results from the laser-welded Eurofer97 joint along the *x*- and *y*-directions and presented in Fig. 4(C) and (D). Although the ANN model demonstrates commendable performance in predicting residual stress, the RFR model surpasses this with superior performance.

## 5. Discussion

Machine learning and deep learning predictive models were developed for estimating heterogeneous residual stress distribution based on material microstructures and anisotropies, which were acquired via advanced characterisation techniques. The model has yielded remarkably accurate predictions, even with a relatively small dataset for training. Unlike previous ML/DL models, which were established via the residual stress data from the joints with different manufacturing parameters, the proposed model has the adaptability that stems from the inclusion of microstructure distribution originating from the welding setups. The elimination of the dependence on welding parameters reduces the cost of acquiring training data experimentally for every welds and promotes adaptability to different welding scenarios according to their microstructure distributions. The advanced techniques allow for a finer scale characterisation, enabling the heterogeneity and anisotropy of the residual stress evaluation and providing a unique insight into investigating crack initiation and propagation. Compared with finite element modelling, which has been a longstanding method for predicting residual stress distribution, the ML/DL model excels in terms of efficiency. The model is capable of predicting residual stress distributions with remarkable accuracy in less than a second. This stark contrast in prediction time demonstrates a significant advantage of the ML/DL approach, creating a high throughput approach for resolving residual stress. The training methodology can also be extended to different materials and types of joints to accommodate diverse joining methods without significant additional efforts.

As elucidated in Fig. 3(B), the importance diagram generated by the RFR model illustrates the critical roles of key parameters in precisely predicting residual stress in the weldment. Notably, the coordinates, which indicate the distance to the welding centre, are of paramount significance, given the distinct variations in thermal input in relation to the proximity to the welding centre. Moreover, the constraints imposed by cooling materials exhibit variable effects on the residual stress magnitude in different directions [42,43]. In conjunction with coordinates, microstructure emerges as the second most critical determinant for predicting residual stress. Grain size and material anisotropies inherently represent the outcomes of welding configurations, encapsulating crucial information as input to replace traditional welding parameters. Grain elongation during welding is primarily a result of thermal gradient and plastic deformation via interactions with welding tools, changing the aspect ratio of grain size significantly [44,45]. Columnar grains interact with the surrounding grains, increasing mismatch [46]. A greater degree of elongation in specific directions is indicative of a higher likelihood of anisotropy, leading to anisotropic

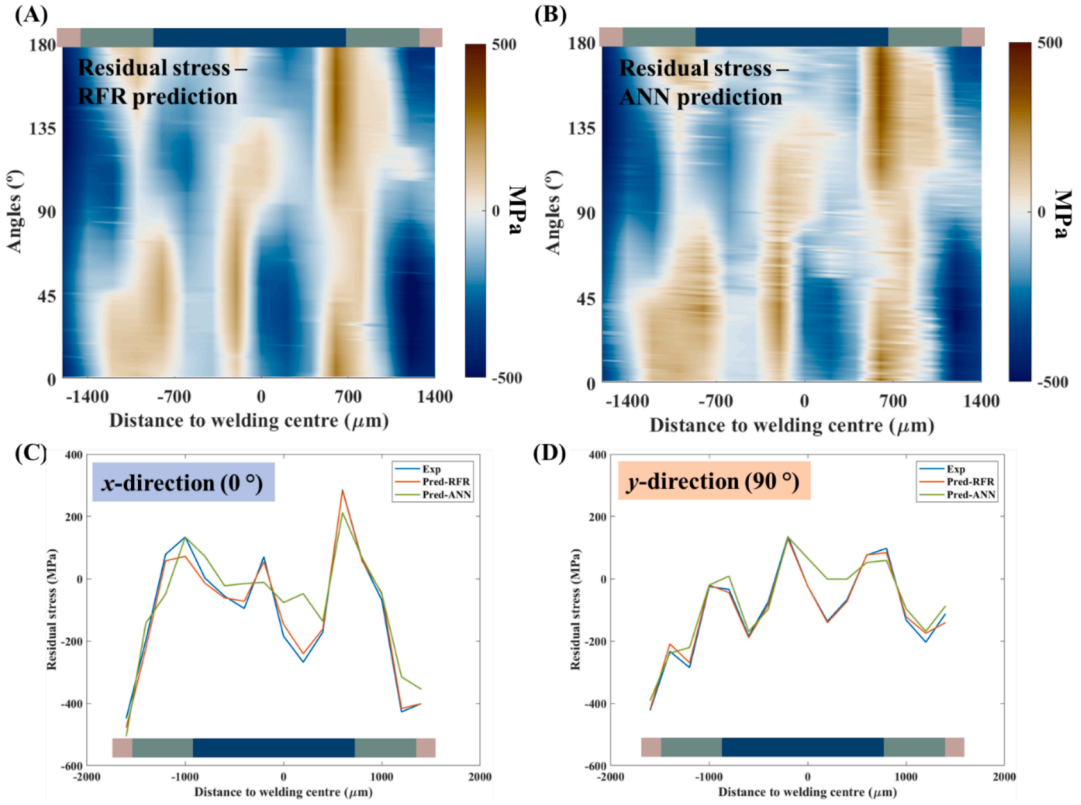



**Fig. 4.** Predictions of angular residual stress distribution across the weldment using (A) the RFR model and (B) the ANN model without sub-window. The predicted residual stress at (C) *x*- and (D) *y*-direction were extracted and compared with experimentally obtained results using PFIB-DIC.

residual stress distribution [43,47]. The calculated average grain size (Fig. 2(C)) in different directions serves as a valuable indicator of elongation severity. In addition, the heat input determines the temperature field which leads to different degrees of texturing. Normally the combined effects of grain size distribution and texture cannot be separated [46,47]

. However, in this case, with laser-welded Eurofer97, the anisotropic properties are less obvious than grain size (see the importance features in Fig. 3(B)). Despite the intricacies of the quantitative relationship, ML/DL models adeptly discern these associations, offering optimal regression fitting.

The use of conventional residual stress measurement through nanoindentation, by comparing measured data with stress-free surfaces, provides a fundamental basis for constructing ML/DL models. Employing well-established calculation procedures ensures that these models are beyond "black boxes" and become interpretable and insightful models. A valuable observation, illustrated in Fig. 1(C), reveals that even as the loads vary when the indenter transitions into the macroscale domain, the resultant residual stress values remain consistent across different load scenarios. This reduces the importance of load selection in the macroscale regime, reaffirming the initial assumptions made during model development. The successful development of the ML/DL model effectively eliminates the constraints inherent in nanoindentation-based residual stress measurements. This achievement prevents the necessity for measuring contact area and enables the evaluation of non-equi-biaxial residual stress in all directions.

The selection of ML/DL algorithms is critical in establishing the predictive model in this study. Here, the RFR, ANN and CNN models, which have been broadly used to develop predictive tools for estimating stress/strain under tension based on microstructures and experimental data, were used in building the framework [10,48]. The RFR stands out for its interpretability, which provides feature importance scores, enabling a more comprehensive understanding of each feature's contribution to the predictive outcome. Furthermore, the results in the insert figure of Fig. 3(B) suggest that RFR offers superior predictive capabilities, particularly for measurements within pillars, as indicated by the slightly higher $R^2$ value. In the domain of prediction, RFR also emerges as the more effective choice owing to its ability to produce predicted residual stress maps with a notably reduced level of noise (Fig. 4(A) and (B)). Moreover, it demonstrates a higher degree of consistency with the experimental results (Fig. 4(C) and (D)). This indicates a greater degree of variation in ANN predictions. This divergence directly stems from the increased complexity introduced by neural network algorithms, which are data-hungry [49]. The RFR, however, is better suited for smaller datasets as the shallow learning algorithm is less susceptible to overfitting [50]. It should be noted that ANN remains a valuable tool, especially when dealing with extensive datasets encompassing diverse input types – a common scenario in applications across various industries and materials.

## 6. Conclusion

In summary, this study developed a novel framework to establish a ML/DL powered model to predict residual stress at multiple directions. Leveraging the cost-effectiveness and sophistication of PFIB-DIC, instrumented indentation, and microstructure characterisation techniques, the training data encompasses a comprehensive spectrum of high-resolution measurements. This dataset captures intricate details of material behaviour and the interplay between residual stresses, microstructural features, and mechanical properties. This lays a robust foundation for training the predictive model, enabling accurate predictions and providing insights into localised residual stresses at a finer scale in the laser-welded Eurofer97 joint. The framework can also be further improved by its nature of being highly adaptable and dynamic in refinements with the particular ML/DL algorithms and features. This enables the present framework to be readily extended to a wide variety of manufacturing parameters for fast and reliable residual stress prediction.

## CRediT authorship contribution statement

**Bin Zhu:** Writing – original draft, Visualization, Validation, Project administration, Methodology, Investigation, Formal analysis. **Nathanael Leung:** Writing – review & editing, Investigation. **Brandon Steel:** Writing – review & editing, Visualization, Investigation. **David England:** Writing – review & editing, Investigation. **Yinglong He:** Writing – review & editing, Validation, Methodology, Investigation. **Andrew J. London:** Writing – review & editing, Investigation. **Hannah Zhang:** Writing – review & editing, Investigation. **Michael Gorley:** Writing – review & editing, Supervision. **Yiqiang Wang:** Writing – review & editing, Supervision, Funding acquisition. **Mark J. Whiting:** Writing – review & editing, Supervision, Funding acquisition. **Tan Sui:** Writing – review & editing, Supervision, 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.

## Acknowledgements

The authors give thanks to the Karlsruhe Institute of Technology for providing the Eurofer97 plate for this study and Dr Simon Kirk (UKAEA) for his advice on laser welding. The authors would like to thank Dr David Cox and Mr David Jones, from the University of Surrey, for their support and assistance during the sample preparation, experiments, and data analysis. The authors are also grateful to Dr Qibin Zhao from the Shanghai Jiaotong University and Mr Tony Fry from the NPL for their review and comments. The research used UKAEA's Materials Research Facility, which has been funded by and is part of the UK's National Nuclear User Facility and Henry Royce Institute for Advanced Materials. Dr Wang, Dr Gorley and Dr London would also like to acknowledge the UK's EPSRC grants [EP/W006839/1] and the Department for Energy Security and Net Zero. Dr Sui would like to acknowledge funding from the UK's Royal Academy of Engineering under the Industrial Fellowships programme. Part of this work utilised equipment purchased by the UK's EPSRC project (EP/P001521/1).

## Data availability

Data will be made available on request.

## References

[1] M. Richardson, M. Gorley, Y. Wang, G. Aiello, G. Pintsuk, E. Gaganidze, M. Richou, J. Henry, R. Vila, M. Rieth, Technology readiness assessment of materials for DEMO in-vessel applications, J. Nucl. Mater. 550 (2021) 152906, https://doi.org/10.1016/j.jnucmat.2021.152906.

[2] S. Kirk, W. Suder, K. Keogh, T. Tremethick, A. Loving, Laser welding of fusion relevant steels for the European DEMO, Fusion Eng. Des. (2018), https://doi.org/10.1016/j.fusengdes.2018.03.039.

[3] B. Zhu, N. Leung, W. Kockelmann, S. Kabra, A.J. London, M. Gorley, M.J. Whiting, Y. Wang, T. Sui, Revealing the residual stress distribution in laser welded Eurofer97 steel by neutron diffraction and Bragg edge imaging, J. Mater. Sci. Technol. 114 (2022) 249–260, https://doi.org/10.1016/j.jmst.2021.12.004.

[4] D.J. Hughes, E. Koukovini-Platia, E.L. Heeley, Residual stress in a laser welded EUROFER blanket module assembly using non-destructive neutron diffraction techniques, Fusion Eng. Des. 89 (2014) 104–108, https://doi.org/10.1016/j.fusengdes.2013.12.041.

[5] B. Zhu, N. Leung, Y. Wang, H. Zhang, J. Dluhoš, T. Pirling, M. Gorley, M.J. Whiting, T. Sui, Investigation of the residual strain and deformation mechanisms in laser-welded Eurofer97 steel for fusion reactors, Mater. Sci. Eng. A 877 (2023) 145147, https://doi.org/10.1016/j.msea.2023.145147.

[6] B. Zhu, N. Leung, W. Kockelmann, M. Gorley, M.J. Whiting, Y. Wang, T. Sui, Neutron Bragg edge tomography characterisation of residual strain in a laser-

welded Eurofer97 joint, Nucl. Mater. Energy (2023) 101462, https://doi.org/10.1016/j.nme.2023.101462.
[7] E. Tsymbalov, Z. Shi, M. Dao, S. Suresh, J. Li, A. Shapeev, Machine learning for deep elastic strain engineering of semiconductor electronic band structure and effective mass, NPJ Comput. Mater. 7 (2021) 76, https://doi.org/10.1038/s41524-021-00538-0.
[8] F. Font-Clos, M. Zanchi, S. Hiemer, S. Bonfanti, R. Guerra, M. Zaiser, S. Zapperi, Predicting the failure of two-dimensional silica glasses, Nat. Commun. 13 (2022) 2820, https://doi.org/10.1038/s41467-022-30530-1.
[9] H. Salmenjoki, M.J. Alava, L. Laurson, Machine learning plastic deformation of crystals, Nat. Commun. 9 (2018) 5307, https://doi.org/10.1038/s41467-018-07737-2.
[10] Z. Yang, C.H. Yu, M.J. Buehler, Deep learning model to predict complex stress and strain fields in hierarchical composites, Sci. Adv. 7 (2021), https://doi.org/10.1126/SCIADV.ABD7416.
[11] W. Muhammad, A.P. Brahme, O. Ibragimova, J. Kang, K. Inal, A machine learning framework to predict local strain distribution and the evolution of plastic anisotropy & fracture in additively manufactured alloys, Int. J. Plast. 136 (2021) 102867, https://doi.org/10.1016/j.ijplas.2020.102867.
[12] E. Zhang, M. Dao, G.E. Karniadakis, S. Suresh, Analyses of internal structures and defects in materials using physics-informed neural networks, Sci. Adv. 8 (2022), https://doi.org/10.1126/sciadv.abk0644.
[13] W. Fang, J.X. Huang, T.X. Peng, Y. Long, F.X. Yin, Machine learning-based performance predictions for steels considering manufacturing process parameters: a review, J. Iron Steel Res. Int. 31 (2024), https://doi.org/10.1007/s42243-024-01179-5.
[14] B. Li, J. Zhu, X. Zhao, H. Yao, Residual stress prediction in laser shock peening induced LD-TC4 alloy by data-driven ensemble learning methods, Opt. Laser Technol. 176 (2024) 110946, https://doi.org/10.1016/j.optlastec.2024.110946.
[15] C. Wang, C. Shen, Q. Cui, C. Zhang, W. Xu, Tensile property prediction by feature engineering guided machine learning in reduced activation ferritic/martensitic steels, J. Nucl. Mater. 529 (2020), https://doi.org/10.1016/j.jnucmat.2019.151823.
[16] J. Mathew, J. Griffin, M. Alamaniotis, S. Kanarachos, M.E. Fitzpatrick, Prediction of welding residual stresses using machine learning: Comparison between neural networks and neuro-fuzzy systems, Appl, Soft. Comput. 70 (2018) 131–146, https://doi.org/10.1016/j.asoc.2018.05.017.
[17] Q. Xiong, M.C. Smith, O. Muransky, J. Mathew, Validated prediction of weld residual stresses in austenitic steel pipe girth welds before and after thermal ageing, part 2: modelling and validation, Int. J. Press. Vessel. Pip. 172 (2019), https://doi.org/10.1016/j.ijpvp.2019.02.002.
[18] G.M. Pharr, An improved technique for determining hardness and elastic modulus using load and displacement sensing indentation experiments, J. Mater. Res. 7 (1992) 1564–1583, https://doi.org/10.1557/JMR.1992.1564.
[19] S. Park, K.P. Marimuthu, G. Han, H. Lee, Deep learning based nanoindentation method for evaluating mechanical properties of polymers, Int. J. Mech. Sci. 246 (2023), https://doi.org/10.1016/j.ijmecsci.2023.108162.
[20] Y.H. Lee, D. Kwon, Measurement of residual-stress effect by nanoindentation on elastically strained (1 0 0) W, Scr. Mater. 49 (2003) 459–465, https://doi.org/10.1016/S1359-6462(03)00290-2.
[21] S. Suresh, A.E. Giannakopoulos, A new method for estimating residual stresses by instrumented sharp indentation, Acta Mater. 46 (1998) 5755–5767, https://doi.org/10.1016/S1359-6454(98)00226-2.
[22] J.E. Campbell, H. Zhang, M. Burley, M. Gee, A.T. Fry, J. Dean, T.W. Clyne, A critical appraisal of the instrumented indentation technique and profilometry-based inverse finite element method indentation plastometry for obtaining stress–strain curves, Adv. Eng. Mater. 23 (2021), https://doi.org/10.1002/adem.202001496.
[23] A. Ruiz-Moreno, P. Hähner, F. Fumagalli, V. Haiblikova, M. Conte, N. Randall, Stress−strain curves and derived mechanical parameters of P91 steel from spherical nanoindentation at a range of temperatures, Mater. Des. 194 (2020) 108950, https://doi.org/10.1016/j.matdes.2020.108950.
[24] K. Jeong, H. Lee, O.M. Kwon, J. Jung, D. Kwon, H.N. Han, Prediction of uniaxial tensile flow using finite element-based indentation and optimized artificial neural networks, Mater. Des. 196 (2020), https://doi.org/10.1016/j.matdes.2020.109104.
[25] L. Lu, M. Dao, P. Kumar, U. Ramamurty, G.E. Karniadakis, S. Suresh, Extraction of mechanical properties of materials through deep learning from instrumented indentation, Proc. Natl. Acad. Sci. 117 (2020) 7052–7062, https://doi.org/10.1073/pnas.1922210117.
[26] S. Salmani Ghanbari, A.H. Mahmoudi, An improvement in data interpretation to estimate residual stresses and mechanical properties using instrumented indentation: a comparison between machine learning and Kriging model, Eng. Appl. Artif. Intell. 114 (2022), https://doi.org/10.1016/j.engappai.2022.105186.
[27] B. Zhu, Y. Wang, J. Dluhoš, A.J. London, M. Gorley, M.J. Whiting, T. Sui, A novel pathway for multiscale high-resolution time-resolved residual stress evaluation of laser-welded Eurofer97, Sci. Adv. 8 (2022), https://doi.org/10.1126/sciadv.abl4592.
[28] W.C. Oliver, J.B. Pethica, Method for continuous determination of the elastic stiffness of contact between two bodies, US4848141A (1989).
[29] P.J. Withers, H.K.D.H. Bhadeshia, Residual stress. Part 2 – nature and origins, Mater. Sci. Technol. 17 (2001) 366–375, https://doi.org/10.1179/026708301101510087.
[30] C.K. Dolph, D.J. da Silva, M.J. Swenson, J.P. Wharry, Plastic zone size for nanoindentation of irradiated Fe—9%Cr ODS, J. Nucl. Mater. 481 (2016), https://doi.org/10.1016/j.jnucmat.2016.08.033.
[31] Y. Ding, R.R. Chromik, Relationship between indentation plastic zone size and residual stresses in plastically deformed Fe, Mater. Sci. Eng. A 696 (2017), https://doi.org/10.1016/j.msea.2017.04.017.
[32] W.D. Nix, H. Gao, Indentation size effects in crystalline materials: a law for strain gradient plasticity, J. Mech. Phys. Solids 46 (1998) 411–425, https://doi.org/10.1016/S0022-5096(97)00086-0.
[33] F. Bachmann, R. Hielscher, H. Schaeben, Texture analysis with MTEX–free and open source software toolbox, Solid State Phenom. 160 (2010) 63–68.
[34] X.X. Li, X.X. Li, S. Schönecker, R. Li, J. Zhao, L. Vitos, Understanding the mechanical properties of reduced activation steels, Elsevier Ltd (2018), https://doi.org/10.1016/j.matdes.2018.03.009.
[35] D. Mainprice, R. Hielscher, H. Schaeben, Calculating anisotropic physical properties from texture data using the MTEX open-source package, Geol. Soc. Spec. Publ. 360 (2011) 175–192, https://doi.org/10.1144/SP360.10.
[36] H. Abrams, Grain size measurement by the intercept method, Metallography 4 (1971) 59–78, https://doi.org/10.1016/0026-0800(71)90005-X.
[37] ASTM, ASTM E112-13: Standard test methods for determining average grain size, ASTM International, 2013.
[38] X. Li, L. Cui, J. Li, Y. Chen, W. Han, S. Shonkwiler, S. McMains, Automation of intercept method for grain size measurement: a topological skeleton approach, Mater. Des. 224 (2022), https://doi.org/10.1016/j.matdes.2022.111358.
[39] M. Sebastiani, E. Bemporad, F. Carassiti, N. Schwarzer, Residual stress measurement at the micrometer scale: focused ion beam (FIB) milling and nanoindentation testing, 91 (2010) 1121–1136. doi:10.1080/14786431003800883.
[40] Digital Image Correlation and Tracking - File Exchange - MATLAB Central, (n.d.). https://uk.mathworks.com/matlabcentral/fileexchange/50994-digital-image-correlation-and-tracking (accessed June 18, 2020).
[41] T. Hastie, R. Tibshirani, J. Friedman, Springer Series in Statistics The Elements of Statistical Learning – Data Mining, Inference, and Prediction, 2009.
[42] S. Kumar, A. Kundu, K.A. Venkata, A. Evans, C.E. Truman, J.A. Francis, K. Bhanumurthy, P.J. Bouchard, G.K. Dey, Residual stresses in laser welded ASTM A387 Grade 91 steel plates, Mater. Sci. Eng. A 575 (2013) 160–168, https://doi.org/10.1016/j.msea.2013.03.046.
[43] H.L. Wei, J.W. Elmer, T. DebRoy, Crystal growth during keyhole mode laser welding, Acta Mater. 133 (2017) 10–20, https://doi.org/10.1016/j.actamat.2017.04.074.
[44] P.B. Prangnell, C.P. Heason, Grain structure formation during friction stir welding observed by the “stop action technique,, Acta Mater. 53 (2005), https://doi.org/10.1016/j.actamat.2005.03.044.
[45] S. Geng, P. Jiang, L. Guo, X. Gao, G. Mi, Multi-scale simulation of grain/sub-grain structure evolution during solidification in laser welding of aluminum alloys, Int. J. Heat Mass Transf. 149 (2020), https://doi.org/10.1016/j.ijheatmasstransfer.2019.119252.
[46] H.L. Wei, J.W. Elmer, T. DebRoy, Three-dimensional modeling of grain structure evolution during welding of an aluminum alloy, Acta Mater. 126 (2017), https://doi.org/10.1016/j.actamat.2016.12.073.
[47] C.E. Seow, H.E. Coules, G. Wu, R.H.U. Khan, X. Xu, S. Williams, Wire+Arc Additively Manufactured Inconel 718: Effect of post-deposition heat treatments on microstructure and tensile properties, Mater. Des. 183 (2019), https://doi.org/10.1016/j.matdes.2019.108157.
[48] C. Yang, Y. Kim, S. Ryu, G.X. Gu, Prediction of composite microstructure stress-strain curves using convolutional neural networks, Mater. Des. 189 (2020) 108509, https://doi.org/10.1016/j.matdes.2020.108509.
[49] G.C. Silva, V.C. Beber, D.B. Pitz, Machine learning and finite element analysis: An integrated approach for fatigue lifetime prediction of adhesively bonded joints, Fatigue Fract. Eng. Mater. Struct. 44 (2021) 3334–3348, https://doi.org/10.1111/ffe.13559.
[50] J. Wei, X. Chu, X. Sun, K. Xu, H. Deng, J. Chen, Z. Wei, M. Lei, Machine learning in materials science, InfoMat 1 (2019) 338–358, https://doi.org/10.1002/inf2.12028.