Evaluation of neural network attribution methods for microstructural representation of aero-engine material

Xiaopeng Zhou

Hamden Hall Country Day School; Nat Ezhou, 06517, USA

Abstract. Efficiency and accuracy of evaluation for the key to microstructure parameters in aero-engine materials is crucial for understanding the properties and performance. However, it is still a tough research topic to identify the important microstructural variables using conventional methods. In the present work, attribution (variable importance evaluation) methods based on neural networks have been systematically sorted out with a comprehensive understanding of the strengths and limitations. Microstructures of nickel base single-crystal alloys is taken as an example for discuss the variable importance evaluation methods, namely forward stepwise, backward stepwise, and partial derivative. Suggestions are provided for future application in material science and solid mechanics.

Keywords: Variable importance; attribution; microstructure representation; neural network; single crystal.

1. Introduction

Aero-engine is the supply power and the key part of aircraft. The basic components of jet engines include the compressor, combustion chamber, and turbine. The engine starts by drawing in air through the intake, where the compressor compresses the incoming air, increasing pressure and temperature to make the engine work in the proper function [1, 2]. In a jet engine, higher the import temperature of the turbine can cause better engine's performance when the engine is being operated. This is because higher temperature in turbine engine allows more efficient combustion power and expansion of gases. When thrust has been increased, the fuel efficiency will also be improved. However, the higher inlet temperatures can also bring challenges to the engine's structures and lead to some specific problems [3].

One major challenge is the engine tolerance ability of the engine components to withstand the high temperatures. The turbine engine rotor and other high temperature components will be exposed to extremely hot gases, which can reach temperatures of over 1000oC [3, 4]. The high temperature can cause thermal stress, fatigue, and lead to some material degradations such as warping, or some failure of the components. As a result, the materials being used in the construction of these components must be carefully selected to be cooperated with the elevated temperatures and maintain their structural integrity. Structural integrity is determined by material properties. As one of the commonly used areo-engine materials, nickel base single-crystal alloys are widely adopted for turbine blades. During high temperature working condition, the microstructure of single crystal materials will be rafted [5]. Thereby, it will affect the mechanical properties, characterize and evaluate the key microstructure remains to be studied.

Efficiency and accuracy of evaluation for the key to microstructure parameters in single crystal materials is crucial for understanding the properties and performance [6, 7]. To achieve this, machine learning is based on approaches offer for feasible solution that complements existing work in the field [8, 9]. Traditional methods of evaluating microstructure parameters are to rely on manual analysis, which can be time consuming, subjective, and easy to cause human error. On the other hand, machine learning has the potential to automatically and simplify the evaluation process to provide more efficient and reliable results. By training machine learning models on large datasets of microstructure images and the related parameter values has becomes possible to establish relationships between the visual features of the microstructure. These models can then be used to predict the values of these parameters for new and unseen microstructures. Different types of machine learning algorithms can

be employed for the job. For instance, Convolutional Neural Networks (CNN) or random forests. These algorithms excel at extracting intricate patterns and features from images can be enabled to capture subtle variations in microstructure and relate them to specific parameters [10, 11]. In conclusion, machine learning-based approaches offer a promising solution for efficiently and accurately evaluating key microstructure parameters in single crystal materials. By leveraging the power of algorithms to learn from data, these methods can complement existing evaluation techniques, leading to more reliable conclusions and providing a solid foundation for further analysis and exploration in material science.

The evaluation of algorithm for the microstructure characteristics of the materials needs to be studied. There are some existing works. Weng et al. [12] used the gradient-based attribution and the correlation combined analysis provided a quantitative method to evaluate the important morphological variables for the material model. They [13] subsequently developed a principal component aided version which is able to deal with the correlated input variables. Peng et al. [14] also utilized the ML model to find the dominant features in fatigue life. However, there are many evaluation methods at present contain some shortcomings. The processes are needed to systematically sort out the advantages and disadvantages and give a more suitable method. Seven methods which can give the relative contribution and/or the contribution profile of the input factors were compared in the Gevrey's work [15]. Olden et al. [16] compared the variable importance evaluation methods as well. But these works are aimed at other fields such as ecology but no one has systematically sorted out in the field of material science, which is still a research blank. A study is needed to complete the existing work and provide a comprehensive understanding of the strengths and limitations in different evaluation methods.

In the present work, taking the microstructure parameters of single crystal materials as an example, the attribution (variable importance evaluation) methods are systematically sorted out, and reliable conclusions are obtained to provide a basis for future applications.

2. Method

2.1 Microstructures of nickel-based single-crystal alloys

Long-term thermal exposure leads to some significant changes in the microstructure of nickel-based single-crystal alloys, particularly the formation of rafting, where the γ and γ′ phases become long strips instead of their original cubic distribution [5]. The virgin microstructure of superalloys is shown in Fig. 1(a). The white phase is γ, while the black phase is γ′. The γ′ precipitates are distributed in cubic symmetrically. After e.g., 1000 hours of aging treatment at 1100◦C, the result of the microstructure is displayed in Fig. 1(b), where the γ and γ ′ phases become as the long strips, i.e., rafting. The rafting results in the degradation of mechanical properties and reduction of fatigue life in these alloys [17].

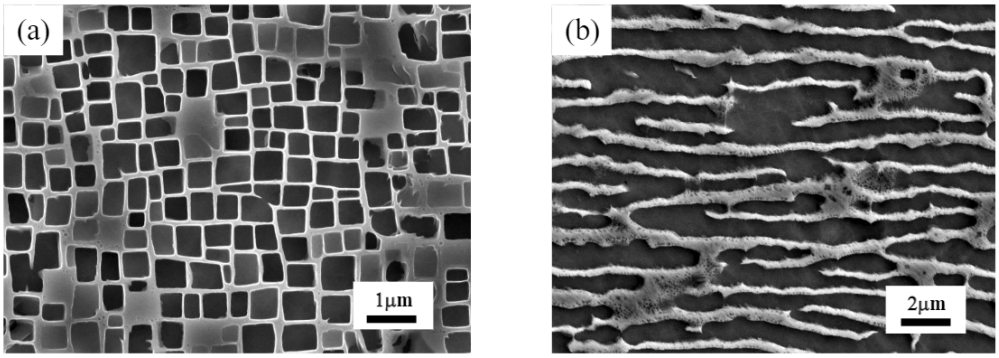

Fig.1 (a) Virgin and (b) rafted states of microstructure of single crystal alloys [13, 17].

2.2 Mean intercept length (MIL)

To describe the microstructure change during rafting, the Mean Intercept Length (MIL) method [18] was introduced to extract the microstructural information. It involves averaging the intercept length of parallel equidistant lines in a specific orientation φ, which reads as

$$d(\varphi)=\frac{L_{\text{int}}(\varphi)}{N_{\text{int}}(\varphi)} \tag{1}$$

where L_{int} is the total length of parallel lines that belongs to the family and intersect the region of interest and N_{int} is the number of intercepts. As depicted in Fig. 2, the black phase of the simple microstructure is analyzed in orientation φ. The total length of the red lines yields L_{int} and N_{int} = 6 because the parallel lines are cut off into six intercepts. The MIL data can be presented as the using of polar coordinates, creating a rose diagram that shows the length distribution of an Orientation Distribution Function (ODF) as a function of orientation. Taking the examples from [13], the microstructures are described by the rose diagrams.

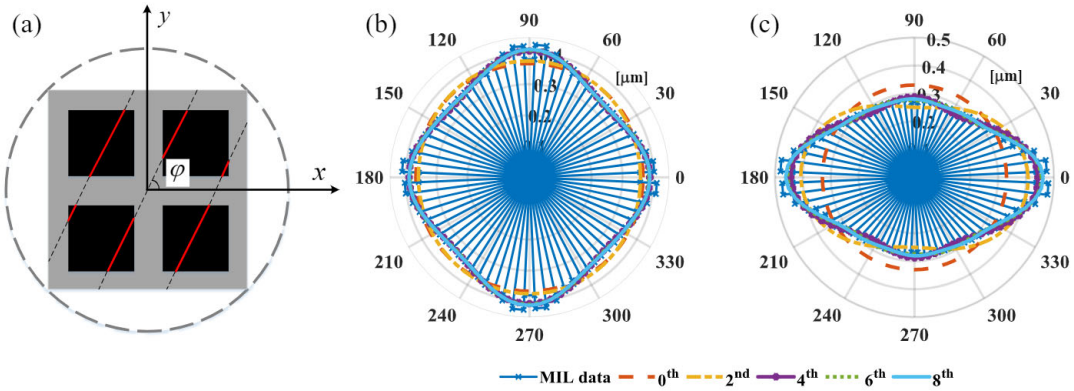

Fig.2 Representation of the microstructures of the single-crystal alloy taken from [12, 13]. (a) Schematic diagram of MIL implementation. (b) and (c) MIL discrete data and Fourier series representations of the virgin and rafted state microstructures. Different dashed curves are Fourier series of different orders.

2.3 Fourier series for 2D microstructures

The basic idea is to construct a condensed representation for the microstructure in some different orientations. From the MIL method, the rose diagram can be described by the function d(φ), with φ as the polar angle that illustrated in Fig. 2. d(φ) is a representation of the given microstructure and can be expanded in Fourier series [12], as

$$d(\varphi)=D^{(0)}+\sum_{m=1}^{\infty}\left(a_m\cos m\varphi+b_m\sin m\varphi\right) \tag{2}$$

where Re is an operator that takes the real part, while $D^{(0)}$, a_m and b_m are the constant coefficients. These parameters of different orders reduce the MIL data dimension, which are easier representations of microstructures.

2.4 Methods for quantifying variable importance in ANNs

2.4.1 Forward stepwise

Assesses the change in the mean square error, which means the smaller the error, the better the model of the network by sequentially adding input neurons. The more input parameters, the more model information, the better the model to the neural network (rebuilding the neural network at each step). The resulting change in mean square error for each variable addition illustrates the relative importance of the predictor variables [15].

Take three variables as an example to judge the order of importance of these three variables. First, select one of the variables in turn, and use the relationship between the neural network variable (microstructure) and the output (mechanical properties), and find that the model established by the three variables and the mechanical properties will have different prediction errors, and there will be a best variable. Leave it on. For the remaining two variables, repeat the above operation, and keep the better variable among them, so far the importance ranking is obtained.

2.4.2 Backward stepwise

Assesses the change in the mean square error of the network by sequentially removing input neurons from the neural network (rebuilding the neural network at each step). The resulting change in mean square error for each variable removal illustrates the relative importance of the predictor variables [15].

Take three variables as an example to judge the order of importance of these three variables. First, build a model with all variables, then remove each variable in turn, and use the relationship between neural network variables (microstructure) and output (mechanical properties), and find that the models established by the three variables and mechanical properties will have different prediction errors. There will be a best variable, keep it. For the remaining two variables, repeat the above operation, and keep the better variable among them, so far the importance ranking is obtained.

2.4.3 Partial derivatives

There is a certain relationship between partial derivative and importance. The partial derivative is computed from the ANN output with respect to the input neurons [19]. To illustrate that the derivative can reflect the importance, examples are taken:

Example 1: $Y = 2x1 + x2$,

Example 2: $Y = x12 + 2x2$.

The partial derivative of Example 1 can be written as $\varphi Y/\varphi x1 = 2$ and $\varphi Y/\varphi x2 = 1$. The partial derivatives represent the sensitivity of the output Y with respect to each input variable. In this case, the derivative values indicate that a unit change in x1 has twice the impact on the output compared to a unit change in x2. Therefore, x1 can be considered more important than x2 in determining the output Y. The partial derivative of Example 2 can be written as $\varphi Y/\varphi x1 = 2x1$ and $\varphi Y/\varphi x2 = 2$, if it is a high-order expression, there are variables in the derivative, so we integrate the derivative in an interval to judge the importance, and finally expression: In this case, the derivatives contain variables, which can complicate the assessment of importance. To overcome this, we integrate the derivatives over a specific interval to evaluate the overall importance. The resulting expression will help us determine the relative importance of the input variables [12]:

$$s_{ij} = \frac{\sum_{x\in\chi}\left|p_{i,j}\right|}{\sum_{j=1}^{n_a}\sum_{x\in\chi}\left|p_{i,j}\right|} \tag{3}$$

where n_a is the number of the evaluated variables. $p_{i,j}$ is the partial derivative of the i^{th} input on the j^{th} output. s_{ij} indicates the effect of the i^{th} input on the j^{th} output.

3. Results and discussion

3.1 Implementation detail

Regular microstructures were generated to assess the ability of microstructure Fourier series. Representative volume elements (RVEs) were created according to the microstructures with periodic boundary condition assigned. The stress was prescribed and the strain is the output. The stress and strain were set as the input and output of the neural network as

$$\sigma = f\left(F, \theta, \varepsilon\right) \tag{4}$$

where F denote the Fourier series consists of different order components. After the key stress and strain points were extracted from the curves, there are 3353, 418 and 418 samples in training, validation and test datasets. The dataset generation process can be found in [12] in detail.

3.2 Comparison of different attribution algorithm

3.2.1 Variable importance

Forward stepwise and Backward stepwise, which are likely referring to statistical variable selection techniques used to identify important variables in a regression model. These methods iteratively add or remove variables based on certain criteria, such as p-values or adjusted R-squared, to determine the most influential predictors. Partial derivative can be automatically computed using pyTorch framework. Ranked variable importance obtained from three methods is plotted in Fig. 3. Dm(0) and Dp0 are related to isotropic properties of the material, while Dm2 and Dp2 describe the anisotropy of the material. As for directional microstructures, Dm2 and Dp2 will be more crucial. The most important component identified by the methods is Dp2, which agrees with the expectation.

However, when evaluating the partial derivatives, the importance of Dm2 is relatively small compared to what was determined by the Forward and Backward stepwise methods. This could be due to the high correlation between Dm2 and Dp2. When variables are highly correlated, their individual contributions may become less distinguishable, leading to differences in their importance rankings between variable selection methods.

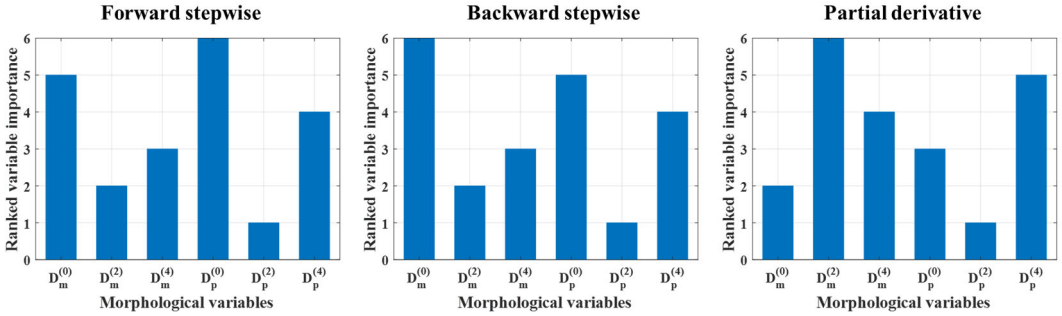

Fig. 3 Ranked variable importance obtained by three different evaluation methods. 1-most important, 6-least important.

3.2.2 Robustness and computational expense

Firstly, in terms of algorithm robustness, each training is random. It's important to note that random initialization can indeed lead to variations in the results obtained from iterative methods like Forward stepwise and Backward stepwise. This randomness is inherent to the algorithms, and as a result, the training errors may vary across different runs. However, this does not affect the evaluation of variables whose importance are significantly different. The randomness problem occurred when two variables have similar importance is acceptable. On the other hand, the Partial Derivative method has been known as poor robustness when dealing with highly correlated variables. Since the Partial Derivative method can be depended on any of the highly correlated variables, the results may vary depending on which variable is chosen, leading to poor robustness. However, if there are no highly correlated variables present, the Partial Derivative method can still be considered a good option.

Additionally, the calculation involved in the Partial Derivative method is much smaller compared to Forward stepwise and Backward stepwise. This computational advantage can make the Partial Derivative method attractive in situations where efficiency is a priority.

Overall, the choice of variable selection method depends on the specific requirements and characteristics problems. If robustness is crucial and contain with highly correlated variables, it might be advisable to consider other methods or techniques that can handle such situations more effectively. Similarly, if computational efficiency is a priority and highly correlated variables are not a concern, the Partial Derivative method can be a suitable option.

3.2.3 Discussion of forward and backward stepwise

In this section, the difference between forward and backward stepwise is discussed. Here is the computational experiment of three microstructural variables.

```
# Forward stepwise
# x_train = x_train[:,[1,10,11]] # Input: microstructure variable 1; MSE= 3%;
# x_train = x_train[:,[2,10,11]] # Input: microstructure variable 2; MSE= 6%;
# x_train = x_train[:,[3,10,11]] # Input: microstructure variable 3; MSE= 3.3%;

# x_train = x_train[:,[1,2,10,11]] # Input: microstructure variable 1 & 2; MSE= 2.8%;
# x_train = x_train[:,[1,3,10,11]] # Input: microstructure variable 1 & 3; MSE= 3%;
# Importance rank: 1>2>3

# Backward stepwise
# x_train = x_train[:,[1,2,10,11]] # Delete: microstructure variable 3; MSE= 2.8%;
# x_train = x_train[:,[1,3,10,11]] # Delete: microstructure variable 3; MSE= 2.95%;
# x_train = x_train[:,[2,3,10,11]] # Delete: microstructure variable 3; MSE= 2.98%;

# x_train = x_train[:,[2,10,11]] # Delete: microstructure variable 1 & 3; MSE= 6%;
# x_train = x_train[:,[3,10,11]] # Delete: microstructure variable 1 & 2; MSE= 3.3%;
# Importance rank: 1>2>3
```

The forward stepwise method involved iteratively adding one microstructure variable at a time to the model based on its individual contribution to model fit. The observation shows that the first step yielded the lowest error (3%) when incorporating microstructure variable 1. Subsequent steps showed that the inclusion of microstructure variables 2 and 3 further reduced the error, with microstructure variable 1 being the most important predictor overall. Similarly, the backward stepwise method started with a model containing all microstructure variables and iteratively removed one variable at a time based on its individual impact on the model fit. We found that the first step, which involved removing microstructure variable 3, resulted in the lowest error (2.8%). Subsequent steps demonstrated that the removal of microstructure variables 2 and 1 increased the error, indicating their importance in the predictive model. Based on findings, microstructure variable 1 holds the highest predictive value among the variables considered, followed by microstructure variable 2 and then microstructure variable 3. These conclusions were consistent across multiple iterations, indicating the accuracy and stability of the forward and backward stepwise methods. The acknowledgement that the stability of the algorithm was assessed by examining the standard deviation of errors across iterations. This analysis revealed the consistency of the algorithm's performance in variable selection. However, it is important to note that the specific results may vary depending on the dataset and problem complexity.

The similarities and differences can be summarized as follows.

Similarities:

Both forward stepwise and backward stepwise are repeated procedures. They involve some sequences of steps where the variables can be added or be removed from the equation. Based on some predefined standard. Moreover, both methods require a criterion for evaluating the quality of the model at each step.

Differences:

The main difference between forward stepwise and backward stepwise methods are based on the direction of variable selection. Forward stepwise procedure starts with empty model and gradually adds variables that contribute the most to the changes in the equation. In contrast, backward stepwise starts with a full equation that contains all of the variables that have the least impact on the model fit. At the same time, because of the different direction of calculation, forward stepwise tends to be computationally more efficient compared to backward stepwise. This is because forward stepwise

starts with a smaller model and gradually expands it due to evaluate fewer variables in the equation. In contrast, backward stepwise starts with a larger model and progressively reduces it, potentially requiring more computations.

Both methods can be prone to selection bias. Forward stepwise may include variables early on that have a high correlation with the response variable but might not be truly significant. Backward stepwise may remove variables that, although individually insignificant, might have a significant effect in combination with other predictors. This bias can lead to overestimation or underestimation to the true effects of predictors. While variable selection aims to improve model performance, there is a risk of overfitting when the selection process is based solely on the fit to the training data. Including too many predictors can lead to a model that performs well on the training data but poorly on new, unseen data. Backward stepwise can be computationally expensive, especially with a large number of predictors. As it involves evaluating models with different combinations of variables, the computational time can increase substantially.

Overall, the computer self-studying demonstrates the usefulness of forward and backward stepwise methods for variable selection, providing insights into the importance of microstructure variables in predicting the response variable. The findings contribute to the field by highlighting the key predictors and guiding future research in developing more accurate and interpretable models.

3.2.4 Summary

The Forward Stepwise and Backward Stepwise methods share certain similarities and have some distinct differences. In terms of advantages, both Forward Stepwise and Backward Stepwise provide a systematic approach to variable selection, allowing for the identification of important predictors in a regression model. They can be particularly useful when dealing with a large number of potential variables. In summary, the Forward Stepwise and Backward Stepwise methods offer a systematic approach to variable selection, providing reliable results and comparable performance. They differ in terms of computational requirements, with Backward Stepwise involving slightly more calculations. The Partial Derivative method offers computational efficiency but suffers from poor robustness when highly correlated variables are present. In conclusion, to choose the most suitable method, it is essential to consider the specific characteristics of the dataset, such as the presence of correlations, the desired level of computational efficiency, and the importance of stability in the variable selection process.

4. Conclusion

In the present study, three variable importance evaluation methods based on ANN are fully discussed. The microstructural variables of single-crystal alloys are taken as a test example. The study shows several conclusions for the future application of those methods, as

- Forward Stepwise and Backward Stepwise methods offer a systematic approach to variable selection, providing more reliable and stable results than partial derivative.
- In terms of computational expense, Backward stepwise involving slightly more calculations than Forward stepwise. Partial derivative method is the most efficient one.
- Partial Derivative method has poor robustness when dealing with highly correlated variables. However, if there are no highly correlated variables present, the Partial Derivative method can still be considered a good option.
- The results in this study can provide a basis for future application in microstructural parameter selection.

References

[1] Norman, H. H.; Parziale, E. A.; Saluja, J. K.; Schenz, R. F. (1972). Nerva Flight Engine Control System Design. Nuclear Technology, 15(3), 447–454. doi:10.13182/NT72-A16041

[2] Dooley, J. L.; Whittlesey, J. W. (1966). SAE Technical Paper Series [SAE International National Aeronautic Meeting and Production Forum - (APR. 25, 1966)] SAE Technical Paper Series - A Highly Compounded Two-Cycle Radial Light-Plane Engine. , 1(), –. doi:10.4271/660173

[3] Lin J, Zhang J, Yang S, et al. Reliability analysis of aero-engine blades considering nonlinear strength degeneration[J]. Chinese Journal of Aeronautics, 2013, 26(3): 631-637.

[4] Franz-Josef Ebert (2007). An Overview of Performance Characteristics, Experiences and Trends of Aerospace Engine Bearings Technologies. , 20(4), 378–384. doi:10.1016/s1000-9361(07)60058-2

[5] Kamaraj, M. (2003). Rafting in single crystal nickel-base superalloys—An overview. Sadhana, 28, 115-128.

[6] Edward J. Garboczi (1990). Permeability, diffusivity, and microstructural parameters: A critical review. , 20(4), 591–601. doi:10.1016/0008-8846(90)90101-3

[7] D.A. Hughes; Q. Liu; D.C. Chrzan; N. Hansen (1997). Scaling of microstructural parameters: Misorientations of deformation induced boundaries. , 45(1), 105–112. doi:10.1016/s1359-6454(96)00153-x

[8] Kumar, S., & Kochmann, D. M. (2022). What machine learning can do for computational solid mechanics. In Current trends and open problems in computational mechanics (pp. 275-285). Cham: Springer International Publishing.

[9] Weng, H., Xu, Y., Chen, J., Ruan, H., Nan, C., Liang, L., & Ju, X. (2021). An enhanced greedy algorithm for failure resistant material design with application to composite delamination. Composite Structures, 278, 114681.

[10] Li, Z., Liu, F., Yang, W., Peng, S., & Zhou, J. (2021). A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects. IEEE Transactions on Neural Networks and Learning Systems, 1–21. doi:10.1109/tnnls.2021.3084827

[11] Xu, Y., Weng, H., Ju, X., Ruan, H., Chen, J., Nan, C., ... & Liang, L. (2021). A method for predicting mechanical properties of composite microstructure with reduced dataset based on transfer learning. Composite Structures, 275, 114444.

[12] Weng, H., Luo, C., & Yuan, H. (2022). ANN-aided evaluation of dual-phase microstructural fabric tensors for continuum plasticity representation. International Journal of Mechanical Sciences, 231, 107560.

[13] Weng, H., & Yuan, H. (2023). Quantitative representation of directional microstructures of single-crystal superalloys in cyclic crystal plasticity based on neural networks. International journal of plasticity, under review.

[14] Peng, X., Wu, S., Qian, W., Bao, J., Hu, Y., Zhan, Z., ... & Withers, P. J. (2022). The potency of defects on fatigue of additively manufactured metals. International Journal of Mechanical Sciences, 221, 107185.

[15] Gevrey, M., Dimopoulos, I., & Lek, S. (2003). Review and comparison of methods to study the contribution of variables in artificial neural network models. Ecological modelling, 160(3), 249-264.

[16] Olden, J. D., Joy, M. K., & Death, R. G. (2004). An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data. Ecological modelling, 178(3-4), 389-397.

[17] Yang, X., Cui, X., & Yuan, H. (2020). Correlations between microstructure evolution and mechanical behavior of a nickel-based single crystal superalloy with long-term aging effects. Materials Characterization, 169, 110652.

[18] Moreno, R., Borga, M., & Smedby, Ö. (2014). Techniques for computing fabric tensors: a review. Visualization and processing of tensors and higher order descriptors for multi-valued data, 271-292.

[19] Dimopoulos, Y., Bourret, P., & Lek, S. (1995). Use of some sensitivity criteria for choosing networks with good generalization ability. Neural Processing Let