Swelling Of Steels Subjected To Fast Neutrons' Irradiation Using Optimized Ensemble Regressions

Ziqi Ren[1,2,3,4*], Yaping Zhao[1,3], Baoxiu Li[1,2,3], and Zixuan Zhang[1,3]

[1] *Department of Materials and Construction Engineering, Hebei Institute of Mechanical and Electrical Technology, Xingtai City, Hebei Province, 054000, China*
[2] *Hebei Science Popularization Base for Additive Manufacturing, Xingtai City, Hebei Province, 054000, China*
[3] *Hebei Technology Innovation Centre for New Energy Vehicle Integrated Die Casting, Xingtai City, Hebei Province, 054000, China*
[4] *Xingtai Green Digital Precision Casting Technology Innovation Center, Xingtai City, Hebei Province, 054000, China*
[*] *Corresponding author. E-mail: renziqi9701@yeah.net*

Received: June 22, 2026; Accepted: September 20, 2026

Radiation-induced swelling (swe) of structural steels under fast neutron irradiation is a critical degradation phenomenon that limits the service lifetime and safety of nuclear reactor components. Conventional experimental approaches for evaluating swelling behavior are costly, time-consuming, and difficult to apply under high-dose irradiation conditions. This study develops an optimized machine learning framework for predicting the swelling behavior of austenitic steels exposed to fast neutron irradiation. A dataset containing 444 experimental samples collected from published literature was established, including twelve input variables related to chemical composition (Fe, Cr, Ni, Mo, Mn, C, Ti, Si, and P), cold work degree, irradiation dose, and temperature. Two ensemble regression models, Adaptive Boosting Regression (ABR) and Extra Tree Regression (ETR), were optimized using the Tyrannosaurus Rex Optimization (TRO) algorithm to improve hyperparameter selection and prediction performance. The dataset was randomly divided into 75% training data (333 samples) and 25% testing data (111 samples), with preprocessing methods including normalization, randomization, interquartile range (IQR)-based outlier treatment, and four-fold cross-validation applied to enhance model reliability. The optimized Extra Tree Regression model (ExtraT-T) achieved superior predictive performance, obtaining coefficients of determination (R^2) of 0.9908 and 0.9563 for the training and testing datasets, respectively, compared with 0.9679 and 0.8991 for the optimized ABR model (AdaB-T). Feature analysis revealed that irradiation dose and cold work degree were the dominant factors affecting swelling behavior. The proposed optimized machine learning framework provides an efficient approach for predicting radiation-induced swelling, reducing dependence on irradiation experiments, and supporting the design of radiation-resistant nuclear materials.

Keywords: SWE; Neutrons; Irradiation; Austenitic steel; ML; Boosting algorithms
© The Author('s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are cited.

http://dx.doi.org/10.6180/jase.202701_36.046

1. Introduction

One of the main processes leading to changes in volume of structural materials irradiated with energetic particles is the creation of voids from the build-up of radiation defects, most of all nucleation and growth of voids [1]. Unacceptable losses in resistance, dimensional stability, and load-bearing capability occur when the swe level reaches about 5% of the substance volume [2]. This state is incompatible with the demanding demands of nuclear engineering applications. The levels of radiation exposure that materials used in fusion reactors are required to withstand are very high, often going beyond a few hundred displacements per atom (dpa). Reduced-activation ferritic/martensitic

(RAFM) steels, seen as attractive options for fusion structural components due to their comparatively high resistance to radiation-induced deterioration, may nonetheless incur intolerable swe under sustained high-dose neutron exposure. Because these experiments are expensive and take a long time when done with neutrons, self-ion irradiation is a very common method and a good substitute for neutron irradiation. In addition, ion and neutron-irradiated materials exhibit very different changes in void size, number density, and depth profiles because the damage regions caused by ions and neutrons are not only different in space but also in time. Besides, the mechanisms of void nucleation and growth are very rapid processes at atoms and their direct observation and quantification remain a challenge. Investigating the swe actions of nuclear building components require a theoretical model. By reason of its fast computation speed and lengthy calculable time span, the rate theory (RT) model is particularly well-suited for long-term swe action monitoring under high-dose irradiation. In the early 1970s, Brailsford, Bullough, and others laid the theoretical groundwork for reaction-rate (RT) swe models [3]. Later research expanded this base to explore the mechanisms that bring about the formation of voids and bubbles in materials exposed to radiation [4]. In particular, microstructural sinks such as dislocation loops and network dislocations have been studied because of their effect on void growth, interstitial fluxes, and vacancy and dislocation fluxes [5]. In contrast, the conventional RT model is not successful in sufficiently explaining the incubation time of swe. Assuming primary void and loop widths are finite allows us to circumvent these restrictions in the technique. Computationally expensive, sophisticated cluster-dynamics approaches keep trying to determine the nucleation level more rigorously, but they still need to include complex substance effects like irradiation-induced segregation, gas generation, and alloy chemistry [6]. The estimative reliability of these simulations also declines when they depend on several fitted factors [7].

In engineering sciences, ML has grown in importance due to data-driven methodology's quick advancements [8–11]; it allows for the extraction of complicated correlations among processing factors, performance in harsh conditions, and substance composition [12–15]. Prior studies have illustrated that methods based on ML and computational intelligence are useful in predicting intricate and expensive engineering features [16–19]. Recent studies have shown that using meta-heuristic algorithms to tune hyperparameters can improve the performance of machine learning models in engineering problems [20, 21]. As an important degrading mechanism in nuclear system structural substances, radiation-induced swe presents a formidable issue to conventional experimental and mechanistic models; ML offers a robust alternative. ML methods can quickly and quantitatively forecast crucial agents like the starting dose of steady-state swe by learning the multidimensional dependencies that control swe action from existing empirical data collections. This capability helps with the development of next-generation nuclear structural materials by speeding up the evaluation of applicant substances and facilitating the design of alloys with greater resistance to irradiation damage. In the work of Jin et al., ML a data-collection-based computer model has already demonstrated its usefulness for forecasting the Donset [22]. ML was able to rapidly compute and predict the Donset in the absence of a physical model, based on the many empirical findings concerning the radiation swe of structural substances. Unfortunately, ML is unable to offer credible estimations of swe activity following hundreds of doses since there is a dearth of empirical swe data under high-dose settings.

This investigation aims to ascertain whether or not it is possible to anticipate swe in austenitic steels when they are exposed to neutron irradiation. Among the ML models that are being researched to demonstrate how effectively they function are the ABR model and the ETR models. The consistency of the findings is highly influenced by the ABR and ETR hyperparameters, which should be chosen utilizing metaheuristic optimization approaches. This goal may be achieved by using one of the available methods, such as the TRO. It is associated with 444 software engineering data points that were gathered from publications in the academic realm. Twelve different factors were taken into consideration for this inquiry. All of these factors are included: The Dose, Temperature (Tem), Degree of Cold Work (CW), Fe, Cr, Ni, Mo, Mn, C, Ti, Si, and P. Despite the overall importance of the radiation behavior of nuclear structural materials, the scope of the present study is specifically limited to predicting the swelling number of austenitic steels under neutron irradiation. Previous ML investigations have mainly focused on forecasting the swelling onset dose or using conventional models, including neural networks; while an optimized and interpretable cumulative framework for forecasting the swelling amount by simultaneously considering chemical composition, cold work rate, irradiation dose, and temperature has not been sufficiently investigated. To address this gap, the present study uses 444 experimental observations of neutron-irradiated austenitic steels and optimizes adaptive boost regression models and super random trees with the Tyrannosaurus Rex optimization algorithm. In addition to evaluating the

prediction performance, SHAP analysis is also used to interpret the contribution of input parameters. Therefore, the innovation of the study is in providing an optimized and interpretable aggregation framework for forecasting the amount of swelling within the range of compositions and irradiation circumstances available in the dataset.

2. Data gathering and pre-processing

This investigation involved data collection and preprocessing to anticipate the sw variable (Swe. $\rightarrow$ % vol) through a ML methodology, which is extensively utilized in engineering disciplines, particularly in substances engineering and the examination of substance properties under intricate conditions (such as radiation, Tem, and deformation). This method is highly appreciated for being able to capture nonlinear interactions effectively, reducing the dependence on simplifying assumptions, and increasing prediction accuracy. The data utilized were sourced from credible technical literature and prior experimental research, with selection criteria encompassing source credibility, information completeness, consistency of empirical situations, and a direct correlation of factors with the sw phenomena [23–39]. After initial data collection and examination, the set of 444 observations was divided into 333 training and 111 test examples using a fixed random partition. The test set was not used until the model selection and hyperparameter tuning procedure was completed. In the training set, the interquartile range was computed for each variable, and observations outside the range Q1-1.5IQR to Q3+1.5IQR were flagged as possible outliers. These were then matched to the values reported in the original sources and the physical limits of the variables. None of the 444 observations were discarded based solely on the IQR rule; in particular, high inflation values were retained, as they may reflect the true response of the material under severe irradiation situations. Normalization parameters were determined only from the training data and the same parameters were applied to the test data without refitting. Cross-validation was also performed only within the training collection; so that in each iteration, the data-dependent steps were fitted only on the training subset of the same fold to prevent information leakage. The input parameters encompass $Fe(\%wt)$, $Cr(\%wt)$, $Ni(\%wt)$, $Mo(\%wt)$, $Mn(\%wt)$, $C(\%wt)$, $Ti(\%wt)$, $Si(\%wt)$, $P(\%wt)$, Degree of CW ($CW \rightarrow$ %), Dose | dpa |, and Temperature $^{\circ}C$, chosen for their direct influence on chemical composition, microstructure, cold deformation rate, radiation dose, and Tem in the formation of lattice defects, voids, and resultant swe phenomena, all of which are physically correlated with the resultant parameter. To enhance the stability and generalizability of the model, four-part Cross-Validation procedures were implemented for precise assessment, Randomization to eliminate ordinal bias, Normalization to standardize data and facilitate model convergence, and IQR to detect and manage outliers; these preprocessing measures are crucial for augmenting the precision, reliability, and overall efficacy of the ML model.

Table 1 presents the central tendency and dispersion metrics of the input and output parameters for both the training and test data collections, illustrating a strong statistical concordance between the two collections. The proximity of the minimum, maximum, and mean values in most variables (such as Fe, Cr, Ni, Mo, Mn, and C) suggests that the data partitioning method was executed randomly and is typical of the full dataset. The standard deviation quantities denote the permissible variability of the chemical composition, procedure circumstances, and radiation agents crucial for training the ML model. The examination of skewness and kurtosis reveals that certain agents, including Cr, P, and swe, exhibit asymmetric distributions and elevated kurtosis, signifying data concentration within specific intervals and the influence of extreme values on the swe phenomenon. Conversely, variables such as Fe, Ni, $Dose$, and Tem display distributions that approximate normality. The Degree of CW and Dose agents exhibit significant variability, affirming their critical influence on microstructural alterations and the increase in swe. Furthermore, the data about the outcome agent swe exhibits considerable dispersion and positive skewness, indicating the complexity of this procedure and necessitating the application of ML techniques to framework nonlinear interactions between input and output variables. The statistical findings of this table demonstrate the quality of the data, the wide variety of agents, and their capability to create a dependable estimative model.

Fig. 1 illustrates the box plot of the input agents utilized for swe predict, distinctly depicting the normalized distribution of the data, the extent of dispersion, the range of variations, and the existence of outliers for each agent. Alloying components such as Fe, Cr, Ni, Mo, and Mn exhibit a considerable range of variations and distinct medians, reflecting the heterogeneity of the specimens' chemical composition and their influence on microstructural alterations and the formation of radiation defects. The agents Ti and Si have a limited range; yet, the existence of outliers suggests that minor variations in these elements can substantially influence the swe action. Conversely, procedure agents and situations such as Degree of CW, Dose, and Tem exhibit more dispersion, indicating the swe phenomenon's heightened sensitivity to the extent of cold deformation, radiation dose, and Tem. The reduced range of certain agents,

Table 1. Central Tendency and Dispersion Metrics of variables

Feature	Subsection	Category	Unit	Statistics						
				Min.	Max.	Avg.	St.D.	Var.	Skew.	Kurt.
Fe	Training	Input 1	%wt.	63.372	73.078	66.219	2.369	5.614	0.860	-0.379
	Testing			63.398	73.078	66.476	2.421	5.863	0.484	-1.294
Cr	Training	Input 2	%wt.	7.5	19.6	16.861	2.049	4.200	-2.119	7.515
	Testing			7.5	19.6	17.163	2.061	4.246	-1.831	5.900
Ni	Training	Input 3	%wt.	8.5	19.4	12.562	2.338	5.465	0.378	1.247
	Testing			8.5	19.4	12.268	2.446	5.981	0.376	0.676
Mo	Training	Input 4	%wt.	0	3.49793	1.661	1.091	1.191	-0.643	-1.350
	Testing			0	3.47167	1.527	1.148	1.317	-0.414	-1.717
Mn	Training	Input 5	%wt.	0	2	1.639	0.542	0.293	-1.540	1.681
	Testing			0	2	1.569	0.572	0.327	-1.212	0.677
C	Training	Input 6	%wt.	0.003	0.12	0.059	0.028	0.001	0.473	0.146
	Testing			0.003	0.12	0.059	0.027	0.001	0.333	0.242
Ti	Training	Input 7	%wt.	0	0.7	0.201	0.241	0.058	0.996	-0.414
	Testing			0	0.7	0.169	0.237	0.056	1.171	-0.019
Si	Training	Input 8	%wt.	0	1.5	0.769	0.286	0.082	0.475	3.188
	Testing			0	1.5	0.745	0.260	0.068	-0.044	4.242
P	Training	Input 9	%wt.	0	0.155	0.019	0.025	0.001	2.896	11.802
	Testing			0	0.155	0.015	0.024	0.001	3.700	18.639
CW	Training	Input 10	%	0	30	8.997	10.044	100.892	0.357	-1.605
	Testing			0	30	9.081	9.690	93.894	0.236	-1.795
Dose	Training	Input 11	dpa	0	118.5	44.986	24.782	614.152	0.339	-0.312
	Testing			2.5	128.85	40.428	25.331	641.642	0.466	-0.101
Tem.	Training	Input 12	°C	302	687.29	478.150	97.178	9443.615	0.069	-0.897
	Testing			304	689.079	472.533	99.397	9879.755	0.289	-0.932
Swe.	Training	Goal	%vol.	0.1	28.04	3.286	5.117	26.181	2.363	5.438
	Testing			0.1	18.0797	2.181	3.537	12.513	2.846	8.290

coupled with multiple outliers (notably for P and swe), suggests an asymmetric data distribution and nonlinear system activity. The swe plot exhibits positive skewness, with data predominantly clustered at lower numbers and a few extremely high numbers. This complexity in forecasting the agent underscores the need for ML model's adept at recognizing nonlinear links. This chart shows the variety and flexibility of the data are high enough for the data to be utilized in making a precise and reliable prediction model.

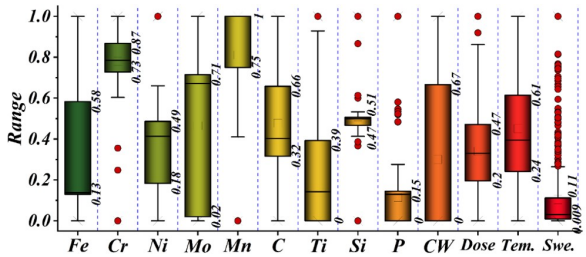


Fig. 1. Introduced input parameters' box plot to estimate Swe.

Fig. 2 is presented as an exploratory representation of the pairwise linear relationships among the features and is not to be interpreted as a formal test of multicollinearity. Radiation dose shows the strongest positive correlation with swelling ($r \approx 0.49$), followed by Ni, Cr, and irradiation temperature with moderate correlations of about 0.34, 0.31, and 0.30, respectively; while the other variables show weaker linear relationships. Since ABR and ETR are decision tree-based cumulative models, unlike linear regression models, they do not rely on the estimation of independent coefficients and can model nonlinear relationships and interactions between correlated variables. Therefore, no feature was eliminated based solely on pairwise correlations, and all physically based variables, including chemical composition, cold work, dose, and irradiation temperature, were retained in the model to preserve information relevant to swelling behavior.

3. Ml-based frameworks

3.1. ETR

Geurts et al. [41] introduced the ETR approach, a boosted variant of the Random Forest (RF) framework. ETR is a

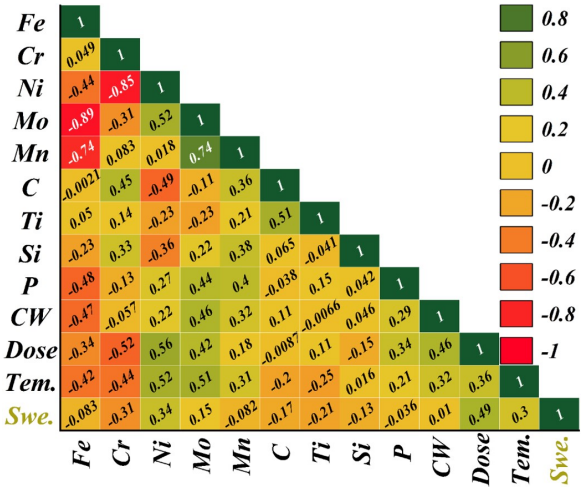


Fig. 2. Pearson correlation matrix between input variables and radiation-induced inflation [40].

procedure, adhering to the conventional top-down methodology, that constructs a collection of uncondensed decision or regression trees. During the training stage of the base predictor, it utilizes random sub-collections of attributes, much to the RF [42]. The ETR is less intricate than an RFR, as stated by Geurts [41], despite the considerable dimensions of the tree. Moreover, it outperforms RFR and other tree-bagging approaches in terms of speed. The ETR and RFR procedures fundamentally vary in two aspects. At first, the ETR employs all potential cutting points and arbitrary chooses one for node division. Secondly, to mitigate bias, it employs the entirety of the training step data for tree training. The RF model is trained with the bootstrap replica. The splitting procedure in the ETR technique is regulated by two variables: n_{min}, which denotes the least specimen size required to differentiate nodes, and k, representing the number of attributes arbitrarily selected at each node. Furthermore, n_{min} and k denote the average result, noise intensity, and feature selection intensity, accordingly. In the ETR framework, precision is boosted, and overfitting is reduced by modifying these two factors. Regression utilizes relative variance fall, and Eq. (1) presents the points. The labels T_i and T_j denote the subsets of specimens from T that relate to the result of the hypothetical divisions s.

$$\text{Score}_i(s,T) = \frac{\text{var}\{z \mid T\} - \frac{|T_i|}{|T|}\,\text{var}\{z \mid T_i\} - \frac{|T_j|}{|T|}\,\text{var}\left\{z \mid T_j\right\}}{\text{var}\{z \mid T\}} \tag{1}$$

In 1990, Schapire established the basis for the boosting approach, an estimation method that integrates multiple weak learners into a single strong learner [44]. Boosting is an ML technique that utilizes the outputs of prior model training to assign weight to every framework's training data specimen for the later framework. The results of previous learning will affect future learning, and the significance of the data will increase as one progress through the boosting period. The generic tree is eliminated, so only the most powerful tree is incorporated with every subsequent totalling of tree frameworks, as demonstrated through Eq. (2).

$$F_n(x) = F_{m-1}(x) + \text{argmin}_h \sum_{i=1}^{n} L\left(y_i, F_{m-1}\left(x_i\right) + h\left(x_i\right)\right) \tag{2}$$

$F_n|x\rangle$ represents the universal model, $F_{m-1}|x\rangle$ denotes the universal model achieved in the preceding round, $h(x_i)$ signifies the newly added tree, and y_i indicates the forecasting output for the ith tree. The model's overall performance will improve as additional repetitive computations are executed. The primary feeble regressor is derived using training specimens, after which additional training specimens are generated by amalgamating erroneous specimens with untrained data. Later, the example is utilized to derive the second feeble regressor. To obtain the third feeble regressor, the erroneous example is combined with the untutored data to produce an additional training example. The procedure is repeated continuously until the requisite value of base learners is generated. When training a fundamental tree model, it is crucial to adjust the weight spread of every example in the dataset. The outputs of the training procedure will differ with each dataset, although they will invariably be aggregated into a singular sum [45].

3.2. TRO

The Tyrannosaurus Rex Optimization Algorithm (TRO) is a population-based metaheuristic method inspired by the hunting behavior of Tyrannosaurus Rex. In the present study, this algorithm was used solely to optimize the hyperparameters of the adaptive reinforcement regression (ABR) and extra-random tree regression (ETR) models. Each candidate solution represented a specific combination of model hyperparameters within predefined ranges, and its fitness was assessed using the average RMSLE obtained from cross-validation on the training set. The initial population was randomly generated in the search space and then iteratively updated using the exploration and exploitation mechanisms of the TRO algorithm, including simulated search, pursuit, and escape of prey. Each new configuration was accepted when it reduced the cross-validation RMSLE value compared to the previous solution [43].

3.2.1. Stage-by-Stage Procedure of TOA

The sequential procedure is established to boost the picture for ascertaining the best share for picture retrieval utilizing

Procedure 1. Pseudocode of the TOA procedure [43]

```
Start
Set up Pr situations
Determine Fitness for every Pr
While (termination circumstance not met):
  Calculate fitness F1 for every Pr
  Find Pr with the lowest situations → target of T-Rex
  If Rand<E then
    Revise Pr location using Eq. (3)
  Else
    Arbitrary revise Pr situation
  End if
  Compute fitness F2 for Pr
  If F2 > F1 then
    Revise T-Rex situation, target and fitness number
  Else
    Target is zero
  End if
End while
End
```

TOA. TOA facilitates the equitable distribution of the community, enhancing lifetime reliability, implementation duration, and power consumption within the task-scheduling framework.

Level 1: setting up Several Prs is generated arbitrarily within the quest area using the community-dependent method referred to as TOA. Let Y be the state of the Pr, arbitrarily produced inside the quest zone defined by the upper and lower borders specified in Eq. (3).

$$Y_j = \text{sand}(ml, dim) * (va - pa) + pa \tag{3}$$

In this context, (ml, dim) represents the situation of the Pr, whereas m indicates the dimension, pa signifies the community count, dim describes the dimension of the quest zone, and va refers to the upper limit.

Level 2: random generation Haphazardly generate the input variables post-initialization. Optimal fitness numbers are calculated using explicit hyperparameter situations.

Level 3: fitness function (ff) estimation An arbitrary answer is produced from the setup process. This is assessed by parameter optimization to boost the weight variable of the P and Y classifier as per Eq. (4).

$$\text{fitness function} \;=\; \text{enhancing}\,(P\&Y) \tag{4}$$

Level 4: chasing The T-Rex pursues its Pr akin to a wolf or a lion. The T-Rex attempts to hunt a victim in proximity. A Pr sporadically protects itself from hunters or manages to evade them, while chasing arbitrarily as indicated in Eq. (5).

$$Y_{\text{new}} \;=\; \begin{cases} y_{\text{new}} \;\text{if rand}() < Ts \\ \text{Random else} \end{cases} \tag{5}$$

Here, T_s denotes the evaluation of successfully capturing the dispersed Pr. The Pr starts to disintegrate as the T-Rex initiates its pursuit, tracking the Pr's situation by the use of Eq. (6).

$$y_{new} = y + rand * rs * (epos * es - target * ls) \tag{6}$$

in which location, es represents the success rate for chasing, ranging from $[0.1, 1]$. If the hunt didn't work out, the Pr managed to escape, so the success rate is zero, and because of this, the Pr's location is upgraded. The target shows where the Pr is relative to the T-Rex.

Level 5: utilization to boost p and y The Pr's current location and its previous situation are critical factors in the selection process. The Pr position is deleted if the T-Rex doesn't pursue and instead goes away or fights off the hunters. This is evident from the FF presented in Eq. (7).

$$y_j^{q+1} = \begin{cases} P < Y \\ target = 0 \end{cases} \tag{7}$$

P denotes the FF for the starting arbitrary Pr placement, while Y signifies the FF for updating the Pr's situation.

Level 6: termination The weight agent for operator P and Y of SPGAN is augmented utilizing TOA; otherwise, repeat stage 3 iteratively til the termination criterion $Y = Y + 1$ is met. Finally, SPGAN-TOA-LDC distinguishes between liver illness and non-liver disease with improved precision while diminishing computational time without error. The pseudocode of the TOA procedure is displayed in the following section.

4. Performance indicators

We evaluated the performance of the ABR and ETR models using a variety of metrics. The framework's skill to explain the variation is displayed by the Coefficient of Determination (R^2) and adjusted R^2. In order to avoid overfitting, the second one considers the number of predictors. The average forecast bias is shown via the Mean Percentage Error (MPE). Positive values indicate that the forecast was too high, while negative numbers indicate that it was too low. The Symmetric Mean Absolute Percentage Error (SMAPE) more equitably handles over- and under-predictions and displays forecast accuracy as a percentage. The accuracy of a model is compared to a naive forecast using the Mean Absolute Scaled Error (MASE). A value less than one indicates that the model is superior. To determine relative differences, Mean Squared Logarithmic Error (MSLE) and its square root, Root MSLE (RMSLE), use a logarithmic transformation. Underestimations become more severe as a result. Last but not least, Theil's U Statistic (U) evaluates a model's predictive ability against that of a naive model. Better predictions may be made by the model when the value is less than 1.

$$R^2 = \left(\frac{\sum_{j=1}^{m} \left(g_j - \bar{g}\right)\left(y_j - \bar{y}\right)}{\sqrt{\left[\sum_{j=1}^{m} \left(g_j - g\right)^2\right]\left[\sum_{j=1}^{m} \left(y_j - \bar{y}\right)^2\right]}} \right)^2 \tag{8}$$

$$MPE = \frac{1}{m}\sum_{j=1}^{m}\left(\frac{g_j - y_j}{g_j}\right) \times 100 \tag{9}$$

$$SMAPE = \frac{1}{m}\sum_{j=1}^{m}\frac{\left|g_j - y_j\right|}{\frac{\left(|g_j| + |y_j|\right)}{2}} \times 100 \tag{10}$$

$$\text{MASE} = \frac{\frac{1}{m}\sum_{j=1}^{m}\left|g_j - y_j\right|}{\frac{1}{m-1}\sum_{j=1}^{m}\left|g_j - g_{j-1}\right|} \tag{11}$$

$$A_R^2 = 1 - \frac{\left(1 - R^2\right) \times (m-1)}{m - c - 1} \tag{12}$$

$$MSLE = \frac{1}{m}\sum_{j=1}^{m}\left(\ln\left(g_j + 1\right) - \ln\left(y_j + 1\right)\right)^2 \tag{13}$$

$$RMSLE = \sqrt{\frac{1}{m}\sum_{j=1}^{m}\left(\ln\left(g_j + 1\right) - \ln\left(y_j + 1\right)\right)^2} \tag{14}$$

$$U = \frac{\sqrt{\frac{1}{m}\sum_{j=1}^{m}\left(y_j - g\right)^2}}{\sqrt{\frac{1}{m}\sum_{j=1}^{m} g_j^2} + \sqrt{\frac{1}{m}\sum_{j=1}^{m} y_j^2}} \tag{15}$$

The following variables are described by the previously given formulas: With m representing the total number of findings, g_j representing the real swe, $\bar{g}$ representing the average of the actual swe, y_j representing the expected value of swe, and $\bar{y}$ representing the average of the predicted value of swe. With the exception of the intercept, c is the number of predictors (independent variables) in the model.

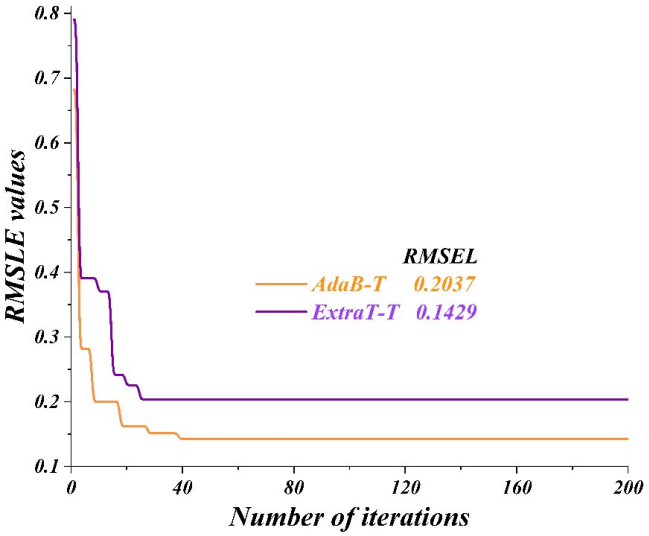


Fig. 3. The reduction in RMSLE values during optimization as objective function.

5. Results, comparison, and verification

swe is evaluated by the use of ExtraT-T and AdaB-T, which are methodologies that integrate the ABR and ETR approaches with the TRO techniques. Fig. 4 illustrates the actual values of swe as well as the values that were approximated using the ExtraT-T and AdaB-T methods. Information gathered throughout the learning and examination stages of the project. In addition, Fig. 4 illustrates the spread of the modeled swe/reported swe value among the various models. The findings of the evaluations that were conducted on the designation that were procured throughout the learning and evaluation levels of the product enhancement strategy are shown in Table 4. ExtraT-T and AdaB-T, two forecasting models, were used to estimate swe, and their success in making these predictions was evaluated based on the metrics R^2, MPE, SMAPE, MASE, A_R^2, MSLE, RMSLE, and U. All study designs were given a score rating at each level of the assessment process, in addition to a percentage of the progress that had been made. The overall performance of the models, as well as their precision and endurance, were evaluated by using previous investigations from published sources, such NN, SVM, and RF [46]. This was to enhance the accuracy of the merged

frameworks as a whole.

According to Table 2, the hyperparameters of the two models were optimized with the aim of minimizing the RMSLE index. For the AdaB-T model, 109 estimators and a decision tree with a maximum depth of 11 were selected, and the optimal learning rate was 0.078; this combination shows that the model has created a balance between learning speed and prevention of overfitting with a relatively limited number of learners and a lower learning rate. In the ExtraT-T model, 275 trees, a maximum depth of 12, at least four samples for each node division, and the squared error criterion were computed as the optimal configuration, which indicates that this model needs a larger set of random trees to reduce variance and improve the stability of the prediction. In both models, the population size was 50, the number of iterations was 200, and the number of independent runs was 5 so that the optimization and comparison situations of the models were the same. However, the reported range for the AdaB-T learning rate, [0.1-0.1], is not consistent with the optimal value of 0.078 and should be corrected to the actual search range before publication.

Fig. 3 shows that the value of the RMSLE objective function in both models decreased rapidly during the initial iterations and reached a stable state after about 40 iterations; so that AdaB-T with a final value of about 0.143 had less error and provided better performance than ExtraT-T with a value of about 0.204.

The results of the five-part cross-validation show that the ExtraT-T model has a lower RMSLE than AdaB-T in all parts. The average error of ExtraT-T is 0.14295, which is about 29.8% lower than the value of 0.2037 of the AdaB-T models. Also, the lower standard deviation of ExtraT-T (0.007977 vs. 0.011366) indicates its greater stability and generalizability across different data parts (Table 3).

AdaB-T and ExtraT-T are both capable of properly predicting swe based on the data. With recorded values of 0.9908 and 0.9563, accordingly, the ExtraT-T produced excellent R^2 metric scores during the learning and evaluation stages. When compared to the other results, the AdaB-T training and examination stages showed lower reliability, with values of 0.9679 and 0.8991, respectively. Even though the A_R^2 index findings were comparable to the R^2 values, the data showed that the $ExtraT-T$ strategy outperformed the $AdaB-T$ method of R^2 values. This sample cannot be used to examine the framework's dependability. Furthermore, in addition to the performance metrics that have been discussed, there are several error-based metrics that are calculated, including as MPE, SMAPE, MASE, MSLE, RMSLE, and U. The lower the value of these metrics, the more dependable the accuracy is assumed to be. When compared to other frameworks, the $ExtraT-T$ model achieved the lowest values with the greatest accuracy, as shown by a review of the values that were collected by the metrics. As shown by MSLE values of 0.0167 in the learning stage, which indicated an improvement rate of -52.119%, and 0.0246 in the testing stage, which indicated an improvement rate of -49.52%, the precision of the $ExtraT-T$ learning and testing stages outperformed that of the $AdaB-T$ approach. With learning and testing values of 0.0349 and 0.0487, accordingly, the $AdaB-T$ demonstrated superior MSLE performance. During the training and evaluation stages, the $ExtraT-T$ model demonstrated strong functional dependability with SMAPE values of 18.9835 and 21.9374. SMAPE readings of 20.8709 and 27.6081, obtained using the $AdaB-T$ method, were very close. Various procedures were used to examine the validity and reliability of the frameworks. Logical analysis, percentage-based improvement rates, and rating scores were some of the methods. When it comes to fulfilling its primary purpose, the $ExtraT-T$ model performs better than the $AdaB-T$ model, which is similar to it.

Table 4 provides a comparison of the findings that were acquired from the approaches that were employed in this study with the findings that were published in the previous body of research (NN, SVM, and RF [46]). This comparison considers not only the accuracy performance but also the models that were developed. For the purpose of comparing the data, Extra$T-T$ was used since it was the model that generated the most accurate conclusions for the inquiry. The outcomes of the accuracy index show a considerable difference, with a higher R^2 in the training stage leading to significant benefits in the long run. The R^2 metrics for the training phase of the Extra$T-T$ program is 0.9908, and the value of each metric is decided by its corresponding value. In all of the data, the values of 0.86 in NN, 0.83 in SVM, and 0.78 in RF [46] were found to be close to the previously reported values. However, when compared to these values, the result that was obtained was superior and more reliable.

The ratio of anticipated swe to reported swe (Fig. 4) is used in this illustration to provide a comparison of the predictive capabilities of AdaB-T and Extra$T-T$ in both training and testing conditions. When it comes to the training category, both models show comparable performance, with AdaB-T (1.5, 0.85) and Extra$T-T$ (1.6, 0.85) both exhibiting the same level of fitting ability. On the other hand, although Extra$T-T$ has a low spread as well as strong accuracy and consistency (1.6, 0.84) in the test category, AdaB-T reveals a considerable spread and poor stability, with values of 2.1 and 0.80, respectively. The outcomes

Table 2. The optimal configuration of developed algorithms.

Algorithm	Hyperparameters	Optimal Value [Search bounds]	Algorithm	Hyperparameters	Optimal Value [Search bounds]
$AdaB-T$	n_estimators	109 [10-1000]	ExtraT - T	n_estimators	275 [10-1000]
	learning_rate	0.078 [0.1-0.1]		criterion	squared_error
	base_estimator	Decision Tree		max_depth	12 [1-50]
	max_depth	11 [1-50]		min_samples_split	4 [1-10]
	Population size	50		Population size	50
	No. of iterations	200		No. of iterations	200
	No. of independent runs	5		No. of independent runs	5

Table 3. The results of cross validation considering RMSLE values.

Algorithm	**F1**	**F2**	**F3**	**F4**	**F5**	**Avg.**	**St. D.**
AdaB-T	0.2223	0.2108	0.1982	0.1893	0.2001	0.2037	0.011366
ExtraT-T	0.156003	0.147933	0.13909	0.132845	0.140424	0.14295	0.007977

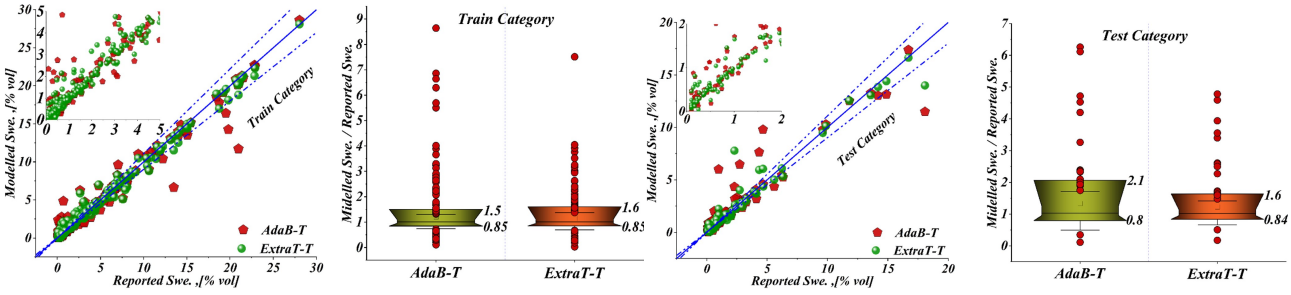
Table 4. Prediction Results from Different ML Models.

Metrics	Metrics value		Difference (%)	Metrics rank		[46]		
	AdaB - T	ExtraT - T		AdaB - T	ExtraT - T	NN	SVM	RF
	Training section					All data		
R^2 (Best is 1.0)	0.9679	0.9908	2.364	2	1	0.86	0.83	0.87
MPE (Best is 0.0)	-28.0786	-15.2840	-45.566	2	1			
SMAPE (Best is 0.0)	20.8709	18.9835	-9.0434	2	1			
MASE (Best is 0.0)	0.0802	0.0591	-26.321	2	1			
A_R^2 (Best is 1.0)	0.9666	0.9899	2.4166	2	1			
MSLE (Best is 0.0)	0.0349	0.0167	-52.119	2	1			
RMSLE (Best is 0.0)	0.1868	0.1292	-30.805	2	1			
U (Best is 0.0)	0.0783	0.0426	-45.514	2	1			
	Testing section							
R^2 (Best is 1.0)	0.8991	0.9563	6.3656	2	1			
MPE (Best is 0.0)	-33.2182	-22.6555	-31.798	2	1			
SMAPE (Best is 0.0)	27.6081	21.9374	-20.54	2	1			
MASE (Best is 0.0)	0.1569	0.0983	-37.365	2	1			
A_R^2 (Best is 1.0)	0.8965	0.9553	6.564	2	1			
MSLE (Best is 0.0)	0.0487	0.0246	-49.52	2	1			
RMSLE (Best is 0.0)	0.2206	0.1567	-28.95	2	1			
U (Best is 0.0)	0.1374	0.0905	-34.114	2	1			

The values for the NN, SVM, and RF models are taken from [46] and are presented only as benchmarks reported in previous studies. Due to differences in dataset, data partitioning, and validation protocol, these values should not be interpreted as a direct, concurrent comparison with the models in the present study.

indicate that although both models exhibit comparable performance when it comes to training data, Extra$T-T$ is more effective at generalizing to new data, making it the more advantageous model overall.

The enhanced efficacy of the ExtraT-T model relative to AdaB-T may be ascribed to the attributes of the Extra Tree regression method and the proficiency of TRO-based hyperparameter optimisation. ExtraT-T produces numerous stochastic decision trees and diminishes model variance via ensemble averaging, thereby enhancing its ability to elucidate the intricate nonlinear associations among irradiation conditions, alloy composition, and swelling behaviour. Conversely, while AdaBoost enhances weak learners in an iterative manner, it exhibits heightened sensitivity to noise and outliers, thus diminishing its generalisation capacity for experimental irradiation datasets. The forecast outcomes suggest that swelling is regulated by the interplay of several factors instead than only by individual variables. Irradiation dosage affects the formation of defects, whereas temperature and cold working alter defect mobility, recombination, and void development. Consequently, ExtraT-T's capacity to simulate nonlinear interactions enhances its



(a) Train category **(b)** Test category

Fig. 4. Prediction Results from Different ML Models for Train and Test category.

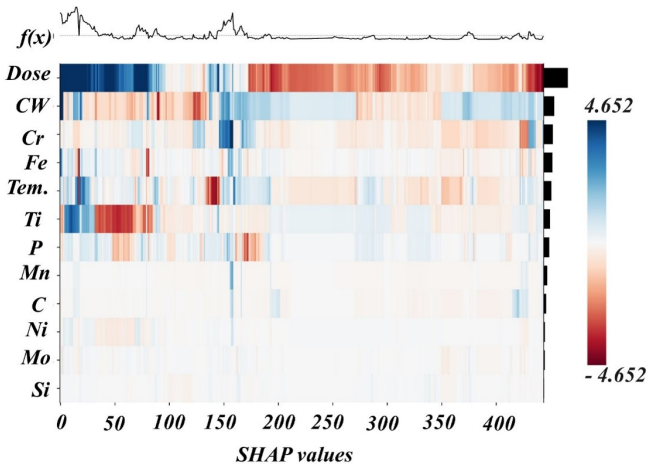


Fig. 5. SHAP values to assess the input's impact on target [47].

predictive precision.

From an engineering standpoint, the suggested ExtraT-T model serves as an expedited assessment instrument for determining the swelling resistance of austenitic steels. It has the potential to diminish reliance on costly irradiation tests, facilitate material selection, and aid in the creation of radiation-resistant nuclear structural materials. Nonetheless, empirical verification is essential prior to utilizing the model for conclusive design choices.

To interpret the model output and determine the contribution of each input variable to inflation prediction, the Shapley incremental value method (SHAP) was used; this method, based on game theory, calculates the contribution of each feature in increasing or decreasing the predicted value [35]. SHAP analysis was conducted to enhance the interpretability of the ExtraT-T model and ascertain the influence of each input variable on swelling prediction. The SHAP summary plot indicates that the irradiation dose is the predominant factor, succeeded by cold work (CW) and temperature. The beneficial effect of dosage aligns with radiation damage processes, as heightened neutron exposure leads to greater displacement damage, vacancy buildup, and void expansion, culminating in increased swelling. The SHAP dependence plots demonstrate nonlinear associations between critical variables and swelling behavior. CW demonstrates a multifaceted influence as cold deformation generates dislocations that may function as defect sinks, mitigating swelling to some extent, whereas excessive CW could elevate defect density and hasten swelling. Temperature influences swelling by regulating defect mobility, recombination, and the processes of void evolution. Among the alloying constituents, Cr and Ni demonstrate moderate influences owing to their impact on phase stability and radiation-induced segregation, while Mo and Si present comparatively lower SHAP values in the analyzed dataset. These results do not indicate insignificant physical impacts but reflect lesser influences under the observed conditions.

The values for the NN, SVM, and RF models are taken from [46] and are presented only as benchmarks reported in previous studies. Due to differences in dataset, data partitioning, and validation protocol, these values should not be interpreted as a direct, concurrent comparison with the models in the present study.

6. Conclusions

This investigation reveals the practicality of anticipating radiation-induced swelling (swe) in neutron-irradiated austenitic steels via optimized machine learning frameworks. A dataset comprising 444 experimental samples derived from published studies was employed to formulate and evaluate Adaptive Boosting Regression (ABR) and Extra Tree Regression (ETR) models refined by the Tyrannosaurus Rex Optimisation (TRO) algorithm. The findings indicate that the enhanced ExtraT-T model delivers better predictive accuracy than AdaB-T, attaining R^2 values of 0.9908 for the training set and 0.9563 for the testing set. The

enhanced performance is ascribed to ExtraT-T's capability to capture nonlinear interactions between irradiation circumstances, material composition, and microstructural elements. SHAP analysis reveals that irradiation dosage and cold work are the primary determinants influencing swelling behaviour, whereas alloying elements like Mo and Si contribute to a lesser extent within the examined dataset. The established model offers an effective method for assessing swelling behaviour while minimizing reliance on expensive irradiation experiments. Nonetheless, the forecasts are constrained to the material compositions, irradiation doses, and temperatures encompassed in the existing dataset. Caution is advised when extrapolating beyond this applicability domain, and further experimental data are necessary to diminish uncertainty and enhance model reliability. Subsequent research ought to emphasize the augmentation of the database through high-dose irradiation studies, the inclusion of supplementary microstructural characteristics, and the execution of independent experimental validation. These initiatives will augment the dependability of machine learning-facilitated forecasting for the formulation and evaluation of radiation-resistant nuclear structural materials.

References

[1] Y. V. Konobeev, A. M. Dvoriashin, S. I. Porollo, and F. A. Garner, (2006) *"Swelling and microstructure of pure Fe and Fe–Cr alloys after neutron irradiation to 26 dpa at 400 C"* **Journal of Nuclear Materials *355***: 124–130. DOI: 10.1016/j.jnucmat.2006.04.011.

[2] F. A. Garner. "Void swelling and irradiation creep in light water reactor (LWR) environments". In: *Understanding and Mitigating Ageing of Nuclear Power Plants*. Elsevier, 2010, 308–356. DOI: 10.1533/9781845699956.2.308.

[3] R. Bullough, B. L. Eyre, and K. Krishan, (1975) *"Cascade damage effects on the swelling of irradiated materials"* **Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences *346***: 81–102. DOI: 10.1098/rspa.1975.0167.

[4] G. A. Cottrell, (2003) *"Void migration, coalescence and swelling in fusion materials"* **Fusion Engineering and Design *66***: 253–257. DOI: 10.1016/S0920-3796(03)00299-0.

[5] M. P. Surh, J. B. Sturgeon, and W. G. Wolfer, (2005) *"Radiation swelling behavior and its dependence on temperature, dose rate, and dislocation structure evolution"* **Journal of Nuclear Materials *336***: 217–224. DOI: 10.1016/j.jnucmat.2004.09.016.

[6] A. A. Kohnert, B. D. Wirth, and L. Capolungo, (2018) *"Modeling microstructural evolution in irradiated materials with cluster dynamics methods: A review"* **Computational Materials Science *149***: 442–459. DOI: 10.1016/j.commatsci.2018.02.049.

[7] S. I. Golubov, A. M. Ovcharenko, A. V. Barashev, and B. N. Singh, (2001) *"Grouping method for the approximate solution of a kinetic equation describing the evolution of point-defect clusters"* **Philosophical Magazine A *81***: 643–658. DOI: 10.1080/01418610108212164.

[8] M. A. Golsefidi, M. Esmaeili-Falak, and H. Sarbaz, (2026) *"An interpretable AI framework using XGB-POA for micropile compressive stiffness prediction"* **Scientific Reports**: DOI: 10.1038/s41598-026-60145-1.

[9] R. S. Benemaran, E. Khajavi, and A. R. T. Khanghah, (2026) *"Construction of multi-parameter intelligent comprehensive estimation algorithms for soil compression modulus"* **Scientific Reports**: DOI: 10.1038/s41598-026-43812-1.

[10] B. Razzaghzadeh, (2025) *"Data-Driven Forecasting of Plastic Viscosity in Stressed Fresh Concrete Using Metaheuristic Boosted XGBoost Models"* **Novel Approaches in Civil and Geotechnical Engineering *1***: 69–84.

[11] M. Esmaeili-Falak, (2026) *"Optimized Gradient Boosting Model for Estimating Unconfined Compressive Strength of Pozzolanic Geopolymer-stabilized Granular Materials Using Pufferfish Algorithm"* **Transportation Infrastructure Geotechnology *13***: 106. DOI: 10.1007/s40515-026-00834-6.

[12] M. S. Borchelooei and M. Esmaeili-Falak, (2026) *"Predicting Water Absorption in Hardened Concrete Using Machine Learning: A Comparative Analysis of XGBoost Optimization Methods"* **International Journal of Pavement Research and Technology**: DOI: 10.1007/s42947-026-00707-8.

[13] H. S. Milani and R. S. Benemaran, (2026) *"Estimation of bond strength of steel-concrete composites subjected to high temperature using tuned tree-based algorithms"* **Multiscale and Multidisciplinary Modeling, Experiments and Design *9***: 81. DOI: 10.1007/s41939-026-01177-y.

[14] M. Esmaeili-Falak and B. Sojoudi, (2026) *"Enhancing dimensional stability in cement pastes: a machine learning approach with metaheuristic-optimized categorical boosting"* **Multiscale and Multidisciplinary Modeling, Experiments and Design *9***: 94. DOI: 10.1007/s41939-026-01183-0.

[15] H. Kou, J. Quan, S. Guo, and E. Hassankhani, (2024) *"Light and normal weight concretes shear strength estimation using tree-based tunned frameworks"* **Construction and Building Materials *452***: 138955. DOI: 10.1016/j.conbuildmat.2024.138955.

[16] M. Esmaeili-Falak and A. Letafat, (2025) *"Prediction of Unconfined Compressive Strength in Pozzolanic Geopolymer-Stabilized Granular Materials Using Tree-Based Models"* **Transportation Infrastructure Geotechnology *12***: 247. DOI: 10.1007/s40515-025-00701-w.

[17] X. Sun, X. Dong, W. Teng, L. Wang, and E. Hassankhani, (2024) *"Creation of regression analysis for estimation of carbon fiber reinforced polymer-steel bond strength"* **Steel and Composite Structures *51***: 509–527. DOI: 10.12989/scs.2024.51.5.509.

[18] N. K. Oghli and E. F. Mahzad, (2025) *"Predicting Liquefaction Triggering Potential Using Metaheuristic GMDH Approaches"* **Geotechnical and Geological Engineering *43***: 1–24. DOI: 10.1007/s10706-025-03275-z.

[19] R. S. Benemaran and H. S. Milani, (2026) *"Post peak flexural strength assessment of concrete reinforced with steel fiber employing categorical boosting frameworks"* **Innovative Infrastructure Solutions *11***: DOI: 10.1007/s41062-026-02645-7.

[20] H. Zheng, (2025) *"Predicting Splitting Tensile Strength of Basalt Fiber-Reinforced Concrete Using Hybrid Decision Tree Algorithms Optimized by Metaheuristic Techniques"* **Novel Approaches in Civil and Geotechnical Engineering *1***: 52–70.

[21] K. Zhang, Y. Zhang, and B. Razzaghzadeh, (2024) *"Application of the optimal fuzzy-based system on bearing capacity of concrete pile"* **Steel and Composite Structures *51***: 25–41. DOI: 10.12989/scs.2024.51.1.025.

[22] M. Jin, P. Cao, and M. P. Short, (2019) *"Predicting the onset of void swelling in irradiated metals with machine learning"* **Journal of Nuclear Materials *523***: 189–197. DOI: 10.1016/j.jnucmat.2019.05.054.

[23] H. R. Brager and F. A. Garner, (1983) *"Microstructural and Microchemical Comparisons of AISI 316 Irradiated in HFIR and EBR-II"* **Journal of Nuclear Materials *117***: 159–176. DOI: 10.1016/0022-3115(83)90022-3.

[24] D. S. Gelles, (1984) *"Swelling in several commercial alloys irradiated to very high neutron fluence"* **Journal of Nuclear Materials *122***: 207–213. DOI: 10.1016/0022-3115(84)90597-X.

[25] A. V. Kozlov and I. A. Portnykh, (2009) *"Dependence of steady-state radiation swelling rate of 0.1C–16Cr–15Ni–2Mo–2Mn–Ti–Si austenitic steel on dpa rate and irradiation temperature"* **Journal of Nuclear Materials *386–388***: 147–151. DOI: 10.1016/j.jnucmat.2008.12.072.

[26] E. N. Shcherbakov, A. V. Kozlov, P. I. Yagovitin, M. V. Evseev, E. A. Kinev, V. L. Panchenko, I. Isobe, M. Sagisaka, T. Okita, and N. Sekimura, (2009) *"Influence of damage rate on physical and mechanical properties and swelling of 18Cr–9Ni austenitic steel in the range of* $3x10^{-}9$ *to* $4\ x\ 10^{-}8$ *dpa/s"* **Journal of Nuclear Materials *386–388***: 152–156. DOI: 10.1016/j.jnucmat.2008.12.080.

[27] S. I. Porollo, Y. V. Konobeev, and F. A. Garner, (2009) *"Swelling and microstructure of austenitic stainless steel ChS-68 CW after high dose neutron irradiation"* **Journal of Nuclear Materials *393***: 61–66. DOI: 10.1016/j.jnucmat.2009.05.005.

[28] F. A. Garner, M. B. Toloczko, and B. H. Sencer, (2000) *"Comparison of swelling and irradiation creep behavior of fcc-austenitic and bcc-ferritic/martensitic alloys at high neutron exposure"* **Journal of Nuclear Materials *276***: 123–142. DOI: 10.1016/S0022-3115(99)00225-1.

[29] T. R. Allen, J. I. Cole, C. L. Trybus, D. L. Porter, H. Tsai, F. Garner, E. A. Kenik, T. Yoshitake, and J. Ohta, (2006) *"The effect of dose rate on the response of austenitic stainless steels to neutron radiation"* **Journal of Nuclear Materials *348***: 148–164. DOI: 10.1016/j.jnucmat.2005.09.011.

[30] I. A. Portnykh, A. V. Kozlov, V. L. Panchenko, and N. M. Mitrofanova, (2012) *"Characteristics of radiation porosity formed upon irradiation in a BN-600 reactor in the fuel-element cans of cold-deformed steel EK-164 (06Kh16N20M2G2BTFR)-ID cd"* **Physics of Metals and Metallography *113***: 520–531.

[31] I. A. Portnykh and V. L. Panchenko, (2016) *"Characteristics of radiation porosity and structural phase state of reactor austenitic 07C–16Cr–19Ni–2Mo–2Mn–Ti–Si–V–P–B steel after neutron irradiation at a temperature of 440–600 C to damaging doses of 36–94 dpa"* **Physics of Metals and Metallography *117***: 611–623. DOI: 10.1134/S0031918X16060090.

[32] F. A. Garner and H. R. Brager, (1985) *"The role of phosphorus in the swelling and creep of irradiated austenitic alloys"* **Journal of Nuclear Materials *133–134***: 511–514. DOI: 10.1016/0022-3115(85)90200-4.

[33] F. A. Garner, H. R. Brager, and R. J. Puigh, (1985) *"Swelling behavior of titanium-modified alloys in EBR-II"* **Journal of Nuclear Materials *133–134***: 535–539. DOI: 10.1016/0022-3115(85)90205-3.

[34] P. Dubuisson, A. Maillard, C. Delalande, D. Gilbon, and J.-L. Seran. *The effect of phosphorus on the radiation induced microstructure of stabilized austenitic stainless steels*. Tech. rep. Gif-sur-Yvette, France: CEA Centre d'Etudes Nucleaires de Saclay, 1990.

[35] F. A. Garner and A. S. Kumar. "The influence of both major and minor element composition on void swelling in austenitic steels". In: *ASTM STP 955*. 1987, 289–314. DOI: 10.1520/STP33826S.

[36] F. A. Garner and H. R. Brager, (1988) *"The influence of Mo, Si, P, C, Ti, Cr, Zr and various trace elements on the neutron-induced swelling of AISI 316 stainless steel"* **Journal of Nuclear Materials *155–157***: 833–837. DOI: 10.1016/0022-3115(88)90425-4.

[37] F. A. Garner and M. A. Mitchell, (1992) *"The complex role of phosphorus in the neutron-induced swelling of titanium-modified austenitic stainless steels"* **Journal of Nuclear Materials *187***: 223–229. DOI: 10.1016/0022-3115(92)90501-B.

[38] K. L. Murty and I. Charit, (2008) *"Structural materials for Gen-IV nuclear reactors: Challenges and opportunities"* **Journal of Nuclear Materials *383***: 189–195. DOI: 10.1016/j.jnucmat.2008.08.044.

[39] G. M. Bond, B. H. Sencer, F. A. Garner, M. L. Hamilton, T. R. Allen, and D. L. Porter. "Void swelling of annealed 304 stainless steel at 370–385 C and PWR-relevant displacement rates". In: *Proceedings of the Ninth International Symposium on Environmental Degradation of Materials in Nuclear Power Systems–Water Reactors*. Wiley, 1999, 1045–1050. DOI: 10.1002/9781118787618.ch109.

[40] J. Benesty, J. Chen, Y. Huang, and I. Cohen. "Pearson Correlation Coefficient". In: *Noise Reduction in Speech Processing*. Springer, 2009, 1–4. DOI: 10.1007/978-3-642-00296-0_5.

[41] P. Geurts, D. Ernst, and L. Wehenkel, (2006) *"Extremely randomized trees"* **Machine Learning *63***: 3–42. DOI: 10.1007/s10994-006-6226-1.

[42] V. John, Z. Liu, C. Guo, S. Mita, and K. Kidono. "Real-time lane estimation using deep features and extra trees regression". In: *Image and Video Technology: 7th Pacific-Rim Symposium, PSIVT 2015*. Springer, 2016, 721–733. DOI: 10.1007/978-3-319-29451-3_57.

[43] V. S. D. M. Sahu, P. Samal, and C. K. Panigrahi, (2023) *"Tyrannosaurus optimization algorithm: a new nature-inspired meta-heuristic algorithm for solving optimal control problems"* **e-Prime - Advances in Electrical Engineering, Electronics and Energy *5***: 100243. DOI: 10.1016/j.prime.2023.100243.

[44] R. E. Schapire, (1990) *"The strength of weak learnability"* **Machine Learning *5***: 197–227. DOI: 10.1007/BF00116037.

[45] R. E. Schapire. "Explaining adaboost". In: *Empirical Inference: Festschrift in Honor of Vladimir N. Vapnik*. Springer, 2013, 37–52. DOI: 10.1007/978-3-642-41136-6_5.

[46] P. Korotaev and A. Yanilkin, (2025) *"Neural networks for prediction of swelling in austenitic steels irradiated by fast neutrons"* **Computational Materials Science *246***: 113408. DOI: 10.1016/j.commatsci.2024.113408.

[47] S. M. Lundberg, G. Erion, H. Chen, A. DeGrave, J. M. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal, and S.-I. Lee, (2020) *"From local explanations to global understanding with explainable AI for trees"* **Nature Machine Intelligence *2***: 56–67. DOI: 10.1038/s42256-019-0138-9.