## Supplementary material

### Data Collection and Description

The data used in this study was meticulously collected from various academic papers and technical reports related to radiation embrittlement [1-17]. The collected steel types include Eurofer97, F82H, T91, OPTIFER, ORNL, and CLAM. This dataset includes detailed information on the chemical composition of reduced-activation steels, the conditions of radiation exposure, and the DBTT.

To ensure data consistency and reduce variability arising from heterogeneous experimental conditions, the following screening criteria were applied:

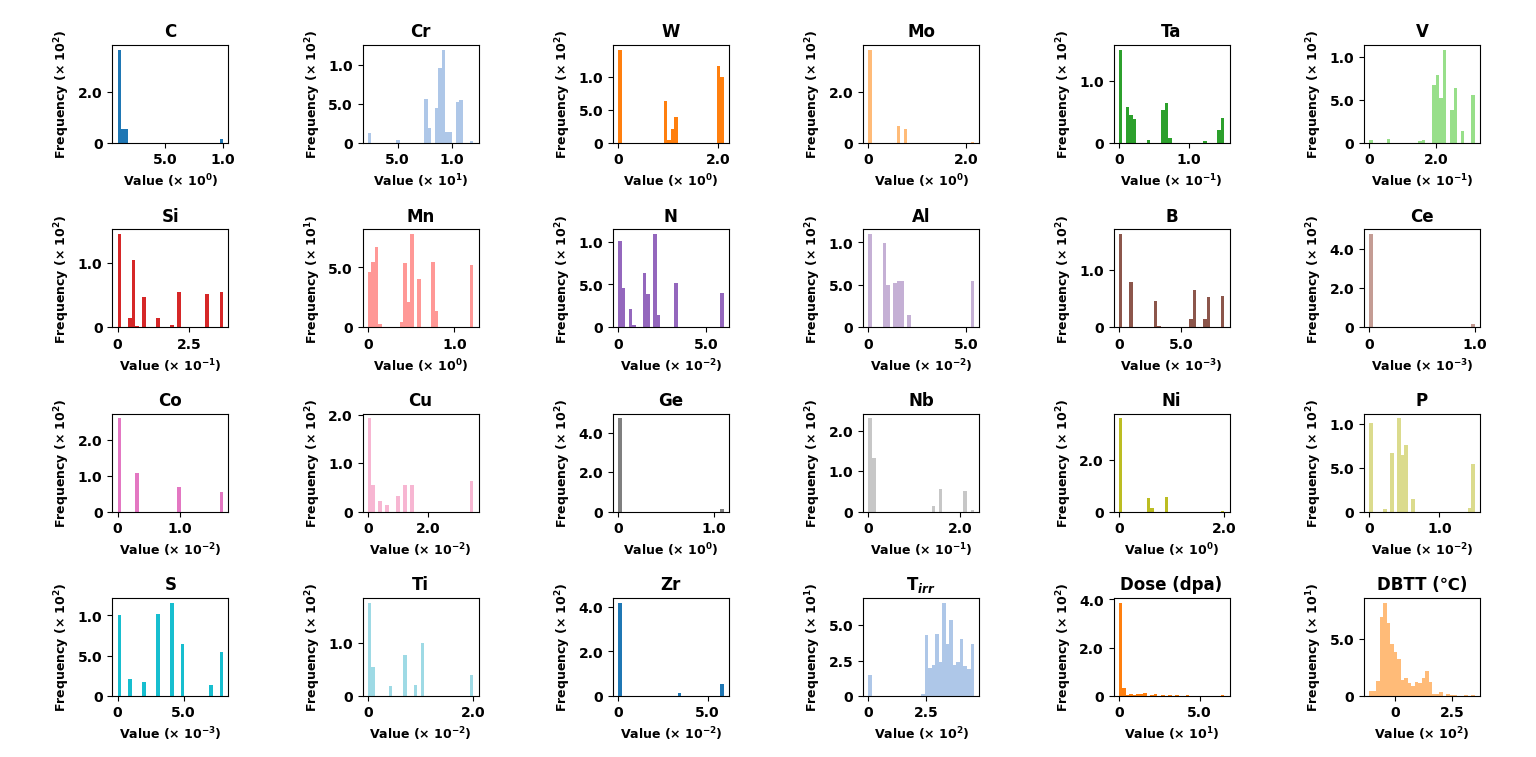
* Only neutron irradiation experiments were included; ion irradiation data were excluded.
* The irradiation dose ranged between 0.1 dpa and 70 dpa.
* The irradiation temperature covered 0°C to 500°C, with clear temperature control reported.
* Only cases with standard mechanical property evaluations (e.g., Charpy impact test for DBTT determination) were selected.
* Datasets where the neutron energy spectrum was fast neutron dominated (typical for fusion or fission test reactors) were preferred.
* Data with unspecified or highly uncertain irradiation parameters were excluded.

Although efforts were made to ensure data consistency, some degree of heterogeneity remains due to differences in neutron energy spectra, dose rates, and irradiation facilities (e.g., HFIR, BOR-60, BR2, SINQ). However, the adoption of ensemble learning models (Stacking of XGBoost, RF, GBDT, and MLP) is particularly suited for handling heterogeneous datasets, as ensemble methods average over multiple learners and are more robust against noise and variability. Additionally, feature importance analysis and SHAP interaction plots were used to interpret and verify model behaviour, ensuring that physically meaningful trends were captured despite experimental heterogeneities.

Fig. 1 presents the distribution of various input variables within the dataset. These variables encompass the chemical elements, Niobium (Nb), Nickel (Ni), Phosphorus (P), Carbon (C), Chromium (Cr), Tungsten (W), Molybdenum (Mo), Vanadium (V), Silicon (Si), Manganese (Mn), Nitrogen (N), Aluminium (Al), Arsenic (As), Boron (B), Cobalt (Co), Copper (Cu), Oxygen (O), Titanium (Ti), Zirconium (Zr), and Molybdenum (Mo). Additionally, it includes two irradiation-related parameters: Irradiation Dose (Dose) and Irradiation Temperature (Tirr), with DBTT as the dependent variable.

Table 1 offers a statistical summary of these variables, displaying their minimum, maximum, mean, and standard deviation values. The elements commonly used in reduced-activation steels include C, Cr, W, Mo, Ta, V, Si, and Mn. Other elements such as N, B, As, Co, Cu, Ni, Al, P, Pb, S, Ti, and Zr are present in relatively low amounts in the samples. The mean and standard deviation values of these elements show their distribution across different samples. Although these elements are present in low concentrations, they can significantly impact the properties of the steel in trace amounts. The data also includes radiation conditions and DBTT.

The collected data covers a wide range of RAFM steel samples and their performance under various irradiation conditions. This comprehensive statistical summary not only provides the fundamental characteristics of the materials but also lays a solid foundation for subsequent analysis and modelling. This study aim to gain valuable, data-driven insights into the behaviour of these materials under irradiation exposure.



**Fig. 1.** Histograms displaying the distribution of variables, including element content (wt.%), irradiation Dose (dpa), irradiation temperature (°C), and DBTT (°C), with the x-axis representing the value range and the y-axis indicating frequency.

**Table 1**

Statistical summary of variable measurements (Composition elements in wt.%).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variables | Mean | Standard Deviation | Maximum | Minimum |
| C | 0.139528 | 0.14934 | 0.09 | 1 |
| Cr | 9.015787 | 1.521929 | 2.25 | 12 |
| W | 1.187104 | 0.863852 | 0 | 2.12 |
| Mo | 0.191607 | 0.347249 | 0 | 2.17 |
| Ta | 0.040403 | 0.047781 | 0 | 0.15 |
| V | 0.229967 | 0.045311 | 0 | 0.314 |
| Si | 0.125278 | 0.131151 | 0 | 0.37 |
| Mn | 0.456356 | 0.368688 | 0 | 1.22 |
| N | 0.017792 | 0.016164 | 0 | 0.06 |
| Al | 0.01438 | 0.01531 | 0 | 0.054 |
| B | 0.003349 | 0.003262 | 0 | 0.0085 |
| Ce | 2.86E-05 | 0.000167 | 0 | 0.001 |
| Co | 0.003945 | 0.005701 | 0 | 0.017 |
| Cu | 0.008822 | 0.011533 | 0 | 0.035 |
| Ge | 0.032048 | 0.183541 | 0 | 1.1 |
| Nb | 0.049676 | 0.078346 | 0 | 0.23 |
| Ni | 0.209292 | 0.368047 | 0 | 2.01 |
| P | 0.004635 | 0.004195 | 0 | 0.015 |
| S | 0.003435 | 0.002423 | 0 | 0.008 |
| Ti | 0.00549 | 0.005992 | 0 | 0.02 |
| Zr | 0.00772 | 0.019128 | 0 | 0.059 |
| Tirr | 336.6876 | 83.12703 | 0 | 456.9514 |
| Dose (dpa) | 3.961391 | 9.871974 | 0 | 65.29018 |
| DBTT (℃) | 8.173883 | 74.13604 | -113 | 349 |

**Data preprocessing**

Generally, the original data obtained is incomplete and may contain a significant number of missing values, outliers, and duplicate entries. If these unprocessed data are directly used to build a prediction model, the prediction accuracy of the model would be adversely affected [18, 19]. Therefore, normalisation was applied as a data preprocessing method in the current study as shown in Eq. (1).

(1)

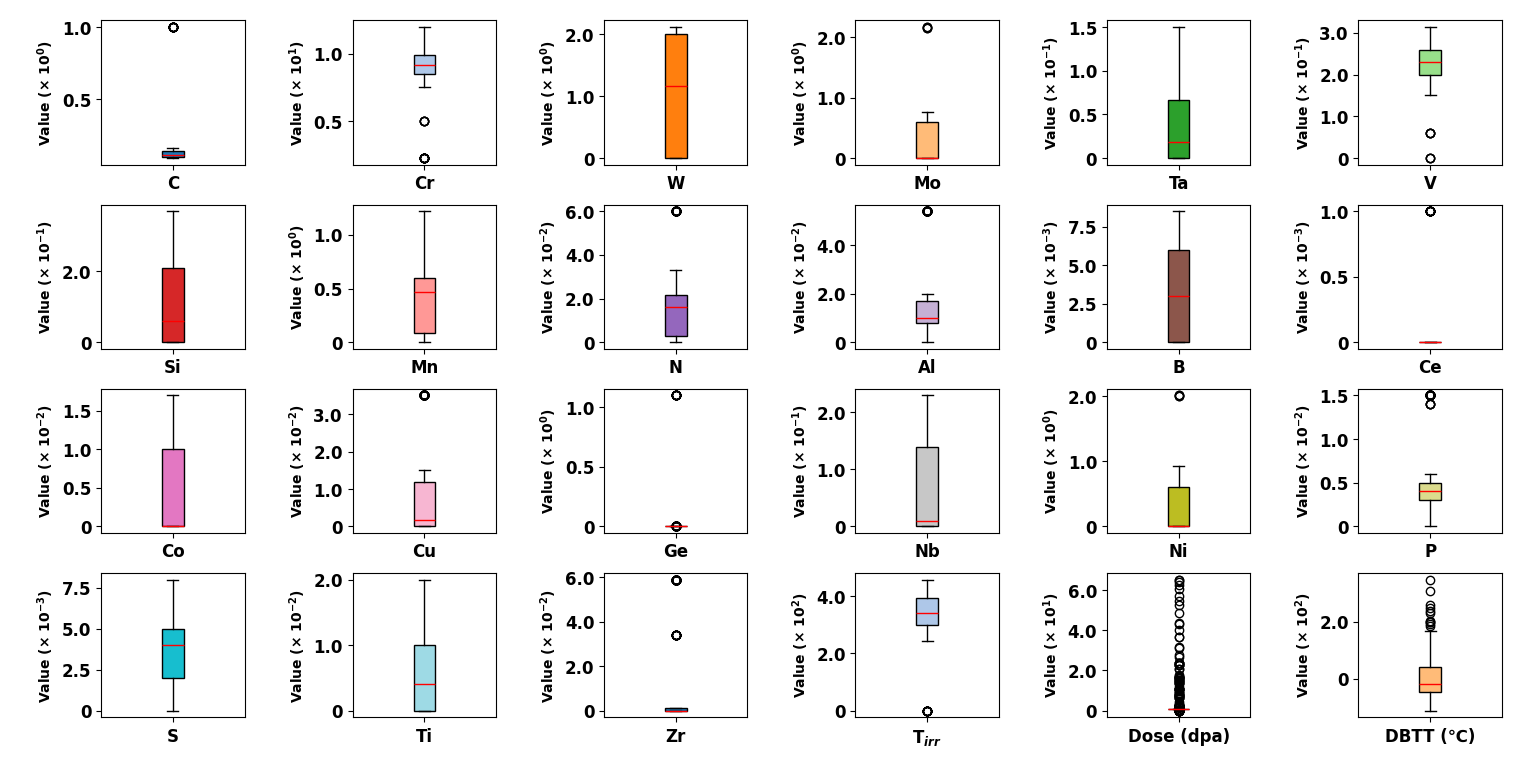
where is the normalised value, is the original value, and and are the minimum and maximum values of the dataset respectively.

This normalisation process scales the data to a range between 0 and 1, ensuring that all variables contribute equally to the analysis and model training, thus improving the overall prediction accuracy.

Outliers can significantly affect the accuracy and robustness of prediction models [19]. Therefore, it is crucial to detect and treat outliers appropriately. In this study, boxplot analysis was conducted to identify outliers in the dataset. Fig. 2 illustrates the boxplots for various input variables, including chemical elements, irradiation conditions and Yield Stress. Each boxplot provides a visual representation of the distribution of values for a particular variable, highlighting the central tendency and spread of the data, as well as identifying any outliers.

From the boxplots, it was evident that several variables contained outliers. These outliers appear as individual points that lie significantly outside the range of the majority of the data. To address the presence of these outliers, the Interquartile Range (IQR) method was employed. The IQR method involves calculating the range between the first quartile (Q1) and the third quartile (Q3) of the data. Data points that lie below Q1 - 1.5IQR or above Q3 + 1.5IQR are considered outliers.

By applying the IQR method, outliers were systematically identified and removed from the dataset. This process ensured that the remaining data is more representative of typical conditions and reduced the likelihood of skewed analysis or model performance. The removal of outliers contributes to the overall robustness and accuracy of the predictive models developed in this study.



**Fig. 2.** Boxplots representing the spread and central tendency of variables, with values on the vertical axis. Each plot corresponds to a different variable, such as elements or conditions. The red line within each box indicates the median of the data.

**Stacking-Based Model for DBTT Prediction**

In this study, a stacking ensemble learning approach was employed to enhance the prediction accuracy of DBTT under irradiation conditions. The dataset, extracted from an Excel file, contained multiple independent variables, including chemical composition and irradiation conditions, with DBTT as the dependent variable. To ensure consistency and robustness, StandardScaler was applied to normalise the dataset, and the data was split into a training set (80%) and a testing set (20%) with a fixed random seed (42) for reproducibility.

The stacking framework incorporated four base learners, namely XGBoost, Gradient GBDT, RF, and MLP. XGBoost and GBDT were selected due to their strong performance in handling non-linear relationships, RF for its ensemble stability, and MLP to capture deep feature interactions. These base models were trained independently, and their predictions were combined and fed into a Ridge Regression meta-learner, which served as the final estimator to mitigate overfitting and enhance generalisation.

To optimise the hyperparameters of each base learner, Optuna was utilised, with the objective of maximising the coefficient of determination (R²) through 5-fold cross-validation on the training dataset. The hyperparameter search space for each model is summarised in Table 2.

**Table 2**

Hyperparameter ranges for different models.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Hyperparameter | XGBoost | GBDT | Random Forest | MLP |
| n\_estimators | 100-2000 | 100-2000 | 100-2000 | N/A |
| learning\_rate | 0.001-0.3 | 0.001-0.3 | N/A | 0.001-0.1 |
| max\_depth | 3-15 | 3-15 | 3-15 | N/A |
| subsample | 0.5-1.0 | 0.5-1.0 | N/A | N/A |
| colsample\_bytree | 0.5-1.0 | N/A | N/A | N/A |
| min\_samples\_split | N/A | N/A | 2-10 | N/A |
| min\_samples\_leaf | N/A | N/A | 1-10 | N/A |
| hidden\_layer\_sizes | N/A | N/A | N/A | (50,), (100,), (50,50), (100,100) |
| activation | N/A | N/A | N/A | relu, tanh |
| alpha | N/A | N/A | N/A | 0.0001-0.1 |
| learning\_rate\_init | N/A | N/A | N/A | 0.001-0.1 |
| max\_iter | N/A | N/A | N/A | 200-1000 |

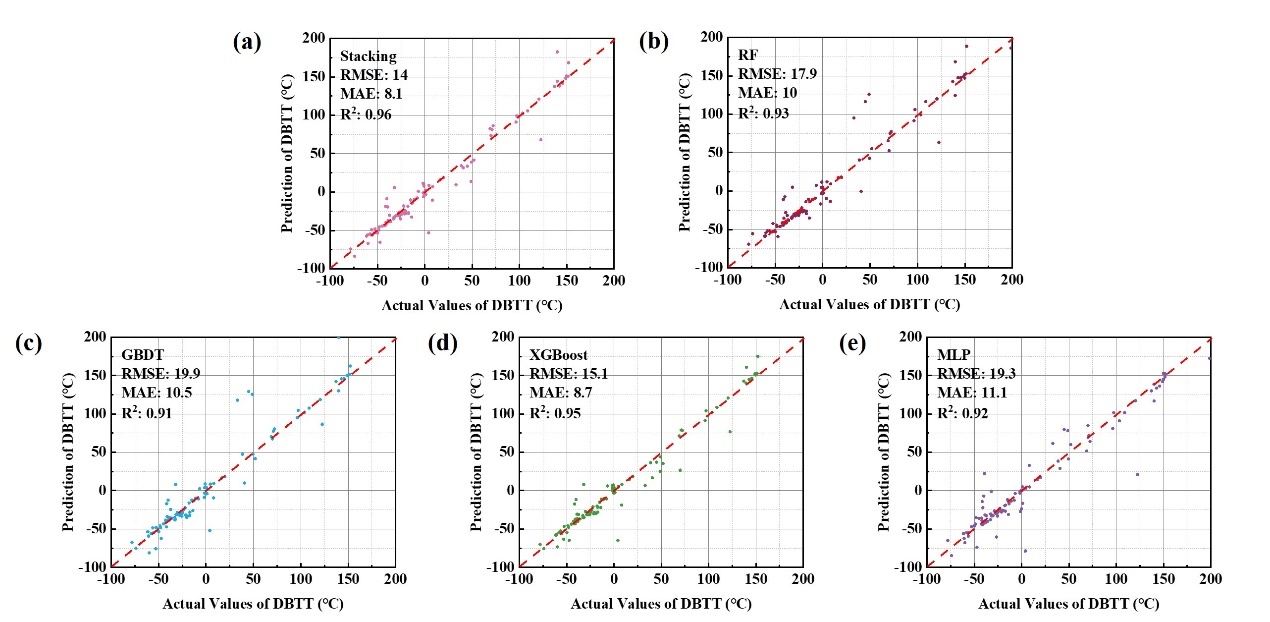
The optimised models were then assembled into the stacking framework and trained on the full training dataset. The performance of the final stacking model was evaluated on the test set using multiple metrics, including mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), and R². The formulas for these metrics are as shown in Eqs. (2-4).

(2)

(3)

(4)

To assess the effectiveness of the stacking model, its performance was compared against the individually trained base models, as shown in Fig. 3. The scatter plots illustrate the comparison between actual vs. predicted DBTT values for each model. The results demonstrate that the stacking ensemble model achieved the best performance, with an RMSE of 14, an MAE of 8.1, and an R² of 0.96, outperforming all base learners. Among the individual models, XGBoost exhibited the highest accuracy (R² = 0.95), followed by Random Forest (R² = 0.93), MLP (R² = 0.92), and GBDT (R² = 0.91). The superior performance of the stacking model highlights the advantage of integrating multiple learning algorithms to improve DBTT prediction under irradiation conditions.



**Fig. 3.** Performance comparison of stacking and individual models for DBTT prediction. Each chart displays the relationship between actual and predicted DBTT (°C), with key evaluation metrics including RMSE (°C), MAE (°C), and R2.

## References

[1] V.V. Rybin, I.P. Kursevich, and A.N. Lapin, Effect of neutron irradiation at low temperature on the embrittlement of the reduced-activation ferritic steels, J. Nucl. Mater. 258-263 (1998) 1324-1328. <https://doi.org/10.1016/S0022-3115(98)00141-X>.

[2] H. Kayano, A. Kimura, M. Narui, Y. Sasaki, Y. Suzuki, and S. Ohta, Irradiation embrittlement of neutron-irradiated low activation ferritic steels, J. Nucl. Mater. 155-157 (1988) 978-981. <https://doi.org/10.1016/0022-3115(88)90452-7>.

[3] R.L. Klueh, M.A. Sokolov, K. Shiba, Y. Miwa, and J.P. Robertson, Embrittlement of reduced-activation ferritic/martensitic steels irradiated in HFIR at 300°C and 400°C, J. Nucl. Mater. 283-287 (2000) 478-482. <https://doi.org/10.1016/S0022-3115(00)00086-6>.

[4] R.L. Klueh, K. Shiba, and M.A. Sokolov, Embrittlement of irradiated ferritic/martensitic steels in the absence of irradiation hardening, J. Nucl. Mater. 377 (2008) 427-437. <https://doi.org/10.1016/j.jnucmat.2008.04.002>.

[5] R.L. Klueh, K. Shiba, and M.A. Sokolov, Embrittlement of irradiated F82H in the absence of irradiation hardening, J. Nucl. Mater. 386-388 (2009) 191-194. <https://doi.org/10.1016/j.jnucmat.2008.12.090>.

[6] A. Alamo, M. Horsten, X. Averty, E.I. Materna-Morris, M. Rieth, and J.C. Brachet, Mechanical behavior of reduced-activation and conventional martensitic steels after neutron irradiation in the range 250–450°C, J. Nucl. Mater. 283-287 (2000) 353-357. <https://doi.org/10.1016/S0022-3115(00)00076-3>.

[7] E. Gaganidze and J. Aktaa, Assessment of neutron irradiation effects on RAFM steels, Fusion Eng. Des. 88 (2013) 118-128. <https://doi.org/10.1016/j.fusengdes.2012.11.020>.

[8] E. Gaganidze, H. Schneider, C. Petersen, J. Aktaa, A. Povstyanko, V. Prokhorov, R. Lindau, E. Materna-Morris, A. Möslang, and E. Diegele, Mechanical properties of reduced activation ferritic/martensitic steels after high dose neutron irradiation. in Proc. of 22st IAEA Fusion Energy Conference. 2008.

[9] V.K. Shamardin, V.N. Golovanov, T.M. Bulanova, A.V. Povstianko, A.E. Fedoseev, Y.D. Goncharenko, and Z.E. Ostrovsky, Mechanical properties and microstructure of advanced ferritic–martensitic steels used under high dose neutron irradiation, J. Nucl. Mater. 271-272 (1999) 155-161. <https://doi.org/10.1016/S0022-3115(98)00774-0>.

[10] I.V. Gorynin, V.V. Rybin, I.P. Kursevich, A.N. Lapin, E.V. Nesterova, and E.Y. Klepikov, Effect of heat treatment and irradiation temperature on mechanical properties and structure of reduced-activation Cr–W–V steels of bainitic, martensitic, and martensitic–ferritic classes, J. Nucl. Mater. 283-287 (2000) 465-469. <https://doi.org/10.1016/S0022-3115(00)00210-5>.

[11] M.A. Sokolov, H. Tanigawa, G.R. Odette, K. Shiba, and R.L. Klueh, Fracture toughness and Charpy impact properties of several RAFMS before and after irradiation in HFIR, J. Nucl. Mater. 367-370 (2007) 68-73. <https://doi.org/10.1016/j.jnucmat.2007.03.165>.

[12] B.v.d. Schaaf, C. Petersen, Y. De Carlan, J.W. Rensman, E. Gaganidze, and X. Averty, High dose, up to 80 dpa, mechanical properties of Eurofer 97, J. Nucl. Mater. 386-388 (2009) 236-240. <https://doi.org/10.1016/j.jnucmat.2008.12.329>.

[13] E. Gaganidze, C. Petersen, E. Materna-Morris, C. Dethloff, O.J. Weiß, J. Aktaa, A. Povstyanko, A. Fedoseev, O. Makarov, and V. Prokhorov, Mechanical properties and TEM examination of RAFM steels irradiated up to 70dpa in BOR-60, J. Nucl. Mater. 417 (2011) 93-98. <https://doi.org/10.1016/j.jnucmat.2010.12.047>.

[14] H.C. Schneider, B. Dafferner, and J. Aktaa, Embrittlement behaviour of low-activation alloys with reduced boron content after neutron irradiation, J. Nucl. Mater. 321 (2003) 135-140. <https://doi.org/10.1016/S0022-3115(03)00241-1>.

[15] O. Anderoglu, T.S. Byun, M. Toloczko, and S.A. Maloy, Mechanical Performance of Ferritic Martensitic Steels for High Dose Applications in Advanced Nuclear Reactors, Metallurgical and Materials Transactions A. 44 (2013) 70-83. 10.1007/s11661-012-1565-y.

[16] Y. Dai and P. Marmy, Charpy impact tests on martensitic/ferritic steels after irradiation in SINQ target-3, J. Nucl. Mater. 343 (2005) 247-252. <https://doi.org/10.1016/j.jnucmat.2004.12.020>.

[17] C. Cabet, F. Dalle, E. Gaganidze, J. Henry, and H. Tanigawa, Ferritic-martensitic steels for fission and fusion applications, J. Nucl. Mater. 523 (2019) 510-537. <https://doi.org/10.1016/j.jnucmat.2019.05.058>.

[18] K. Lakshminarayan, S.A. Harp, and T. Samad, Imputation of Missing Data in Industrial Databases, Applied Intelligence. 11 (1999) 259-275. <https://doi.org/10.1023/A:1008334909089>.

[19] R. Blessing, Outlier Treatment in Data Merging, J. Appl. Crystallogr. 30 (1997) 421-426. <https://doi.org/10.1107/S0021889896014628>.