ARCHIVES of FOUNDRY ENGINEERING

ISSN (2299-2944)
Volume 2023
Issue 4/2023

22 – 33

3/4

10.24425/afe.2023.146671

Published quarterly as the organ of the Foundry Commission of the Polish Academy of Sciences

The Use of the XGBoost and Kriging Methods in the Prediction of the Microstructure of CGI Cast Iron

Ł. Sztangret [a, *], I. Olejarczyk-Wożeńska [a], K. Regulski [a], G. Gumienny [b], B. Mrzygłód [a]
[a] AGH University of Science and Technology, Poland
[b] Lodz University of Technology, Poland
* Corresponding author. E-mail address: szt@agh.edu.pl

Received 13.07.2023; accepted in revised form 16.08.2023; available online 07.11.2023

Abstract

Compacted Graphite Iron (CGI) is a unique casting material characterized by its graphite form and extensive matrix contact surface. This type of cast iron has a tendency towards direct ferritization and possesses a complex set of intriguing properties. The use of data mining methods in modern foundry material development facilitates the achievement of improved product quality parameters. When designing a new product, it is always necessary to have a comprehensive understanding of the influence of alloying elements on the microstructure and consequently on the properties of the analyzed material. Empirical studies allow for a qualitative assessment of the above-mentioned relationships, but it is the use of intelligent computational techniques that allows for the construction of an approximate model of the microstructure and, consequently, precise predictions. The formulated prognostic model supports technological decisions during the casting design phase and is considered as the first step in the selection of the appropriate material type.

Keywords: Compacted Graphite Iron, Machine Learning, Artificial Neural Networks, Kriging, XGBoost

1. Introduction

Materials engineering is a multidisciplinary field that applies knowledge from physics, chemistry, and biology to improve engineering materials. The development of new materials is critical to the creation of innovative solutions and designs. With advances in materials engineering, we now have materials that can withstand extreme conditions, enabling the design of machines that can withstand such environments. The properties of objects and devices depend on the materials used, which in turn are influenced by the manufacturing process and the resulting structure.

Predictive models of microstructure can improve and automate the design process for materials with desired properties. Understanding the rules, laws, and relationships within the field is essential, and data mining tools can support experimental research. Alloying additives affect the structure of compacted graphite iron, but these relationships are not linear and traditional regression methods may not be effective.

In cases where statistical tools fall short, artificial intelligence models have proven to be applicable. Data mining and machine learning techniques based on similar principles have gained popularity. These models use historical empirical data for calibration and capture the relationships between variables. Predictive models are used for quantitative dependent variables, while classification models handle discrete or categorical dependent variables. In this case, the focus is on determining the content of phase constituents in the microstructure based on chemical composition and wall thickness. Limited publications

exist on the influence of alloying elements on compacted graphite iron.

Compacted Graphite Iron (CGI) is a remarkable casting material due to its specific form of graphite and large contact area with the matrix. It tends to ferrite and has a number of interesting properties. Compared to gray cast iron, CGI has higher mechanical properties, improved ductility, and a less thickness sensitive matrix microstructure. Compared to ductile iron, it has a lower coefficient of thermal expansion, higher thermal conductivity, greater resistance to temperature changes, better vibration damping capacity, and improved castability. These advantages make it suitable for various applications, including brake discs, engine blocks and automotive parts. Extensive research has been conducted on the properties of CGI [1-14].

By adjusting the chemical composition, it is possible to modify the microstructure and properties of cast iron. Ausferrite, a mixture of bainitic ferrite and carbon-supersaturated austenite, is a desirable constituent with potential for strengthening through twinning-induced martensitic transformation. Heat treatment, including isothermal quenching within the austenite-bainite transformation range, is required to produce ausferrite. Alternatively, modification of the chemical composition with elements such as molybdenum, copper, or nickel can also produce ausferrite.

The authors have already addressed this issue by building models using machine learning methods such as rosette logic, neural networks, or decision trees. Although good results were obtained, the implemented models did not cope well with the extrapolation of results as well as with interpolation beyond the areas covered by experimental results. The aim of this paper is to present new research results using modern techniques such as kriging and XGBoost trees.

2. Research methodology

2.1. Material experiment

Cast iron was smelted in a medium-frequency induction furnace (Elkon, Poland) with a capacity of 30 kg. The charge consisted of special pig iron (with reduced sulfur content), ferrosilicon and ferromanganese. After superheating the cast iron to 1480°C, the slag was drawn off. In the case of alloyed cast iron, technically pure Sn, Ni, Cu, Mo and/or Cr were added. The concentration of magnesium depended on the content of elements hindering the formation of compacted graphite (e.g. Mo). Schematic layout of elements in the mold is shown in Figure 1.

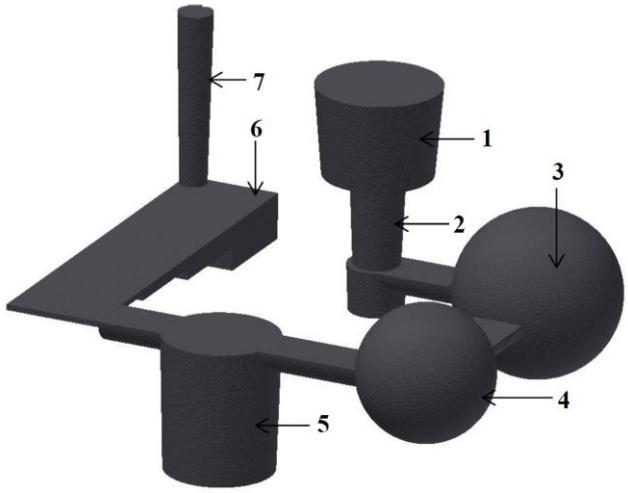

Fig. 1. Schematic layout of elements in the mold: 1 – pouring basin, 2 – sprue, 3 – reaction chamber, 4 – mixing chamber, 5 – control chamber, 6 – test casting, 7 – flow-off

The gating system incorporates the spherically-shaped reaction chamber (2) with φ85 mm diameter. where Lamet® 5504 magnesium master alloy (Elkem, Norway) was placed. It contains nodulizers (5 – 6% wt.% Mg, 0.25 – 0.40 wt.% La) inoculants (0.25 – 0.40 wt.% Al, 0.25 – 0.40 wt.% Ca) as well as graphite-forming element (44 – 48 wt.% Si). The mixing chamber (4) allowed the rest of the master alloy to dissolve and prevented it from flowing into the rest of the gating system. An S-type thermocouple (PtRh10-Pt) was placed in the thermal center of the control chamber (5) to record the temperature of the cast iron. It was connected via a compensation cable to a voltage-frequency transducer and a computer where thermal and derivation analysis curves were recorded. The test casting (5) has a stepped configuration with the wall thickness of 3, 6, 12 and 24 mm. For the study presented in this paper, 51 CGI melts were taken; the range of chemical composition is shown in Table 1.

Table 1.
The chemical composition of CGI tested

Chemical composition, wt.%							
C	Si	Mn	Mg	Mo	Cu	Ni	Cr
2.91–3.82	2.28–2.71	0.03–1.31	0.015–0.023	0–2.44	0–3.80	0–21.04	0–2,81

Such a wide range of chemical compositions made it possible to obtain a ferritic-pearlitic, pearlitic, austenitic, martensitic as well as ausferritic matrix.

Specimens for metallographic studies were cut from the central parts of the stepped casting. They were then ground on abrasive papers of grain sizes 180, 600 and 1200. Polishing was carried out using diamond slurries of gradations 3 and 1 μm. The metallographic sections were etched with a 4% solution of nitric acid in ethanol. Microstructure images were taken on a Nikon MA200 microscope at ×500 magnification. Phase contribution studies were carried out using NIS Elements BR software.

2.2. Machine learning

The authors' analysis is concerned with predicting the volumetric fraction of phases in the microstructure of compacted graphite iron. This problem builds on their previous research [15-17], but with an increased emphasis on regression models. Previous work has already made progress in predicting phases within a microstructure, including the identification of constituents such as ausferrite. The application of machine learning (ML) methods in metals engineering is gaining popularity [18]. Numerous studies have explored the use of supervised learning techniques such as Artificial Neural Networks (ANN) [19], Decision Trees (DTs) [20], and other methods such as XGBoost and Ridge Regression [21] to predict metal properties. Predicting material properties using ML methods is an interesting area of research. Some studies [22, 23] have used ML techniques to establish relationships between defects and mechanical properties. In many cases, ML tools have been used to improve the control of the manufacturing process [24]. However, it is more difficult to find examples of research using ML tools specifically to analyze the properties of compacted graphite iron (CGI). Existing publications on CGI either rely on qualitative analysis [25,26] or use traditional statistical approaches to identify dependencies and construct linear regression models [27, 28].

This manuscript presents an approach to modeling properties based on microstructure with the use of such methods as kriging or XGBoost, but also with the use of neural networks and traditional linear regression. The level of interpretability varies among the individual algorithms used in intelligent data analysis. Decision trees [29-31], once very popular, have given way to more precise techniques such as artificial neural networks and support vector machines [32]. These newer methods offer greater accuracy at the expense of ease of interpretation. However, there is currently a resurgence of simpler (more transparent) methods that have been adapted to increase efficiency for human comprehension [33-35].

2.2. XGBoost

XGBoost (eXtreme Gradient Boosting) is an advanced implementation of the Gradient Boosted Trees algorithm. The idea behind this algorithm is to create a series of trees that successively increase the accuracy of the prediction results. Each successive tree tries to correct the errors of the previous trees using the so-called "gradient boosting" technique. The generalized XGBoost algorithm can be represented as follows:

1. Initialize the model with constant values:

$$\hat{f}_{(0)}(x) = \underset{\theta}{\operatorname{argmin}} \sum_{i=1}^{N} L(y_i, \theta) \tag{1}$$

Where: $\{(x_i, y_i)\}_{i=1}^{N}$ – train dataset, $L\big(y, F(x)\big)$ – differentiable loss function, M – number of iterations (number of weak students - trees), α – learning rate.

2. For $m = 1$ to M:

2.1. Calculate the so-called pseudo-residues ($\hat{g}_m$ – gradients and $\hat{h}_m$ hessians):

$$\hat{g}_m(x_i) = \left[\frac{\partial\big(y_i, F(x_i)\big)}{\partial f(x_i)}\right]_{f(x)=\hat{f}_{(m-1)}(x)} \tag{2}$$

$$\hat{h}_m(x_i) = \left[\frac{\partial^2\big(y_i, F(x_i)\big)}{\partial f(x_i)^2}\right]_{f(x)=\hat{f}_{(m-1)}(x)} \tag{3}$$

2.2. Matching the "weak student" using the training set in the form $\left\{x_i, -\frac{\hat{g}_m(x_i)}{\hat{h}_m(x_i)}\right\}_{i=1}^{N}$, by solving the optimization problem:

$$\hat{\phi}_m = \underset{\phi \epsilon \Phi}{\operatorname{argmin}} \sum_{i=1}^{N} \frac{1}{2} \hat{h}_m(x_i) \left[-\frac{\hat{g}_m(x_i)}{\hat{h}_m(x_i)} - \phi(x_i)\right]^2 \tag{4}$$

$$\hat{f}_{(m)}(x) = \alpha \hat{\phi}_m(x) \tag{5}$$

2.3. Model upgrade:

$$\hat{f}_{(m)}(x) = \hat{f}_{(m-1)}(x) + \hat{f}_{(m)}(x) \tag{6}$$

3. Output:

$$\hat{f}(x) = \hat{f}_{(M)}(x) = \sum_{i=0}^{M} \hat{f}_{(m)}(x) \tag{7}$$

The operation diagram of the XGBoost algorithm is shown in Fig. 2.

The main hyperparameters of this model include [36]:

- learning_rate / eta – parameter telling after each calculated iteration what step we want to take forward. The bigger the step, the faster we get to the goal, but if it is too big, we may not reach the best result.
- max_depth – maximum depth of simple trees. The deeper the trees, the stronger the model is, but it also has a greater tendency to overfitting.
- n_estimators – the number of simple trees we want to build.
- min_child_weight – indicates the minimum number of observations in each leaf of the tree. The higher the weight, the more conservative the model - we need more weight to make a given division.
- gamma – is responsible for reducing the losses required to create another leaf node.
- seed – a seed that is used to generate random numbers.

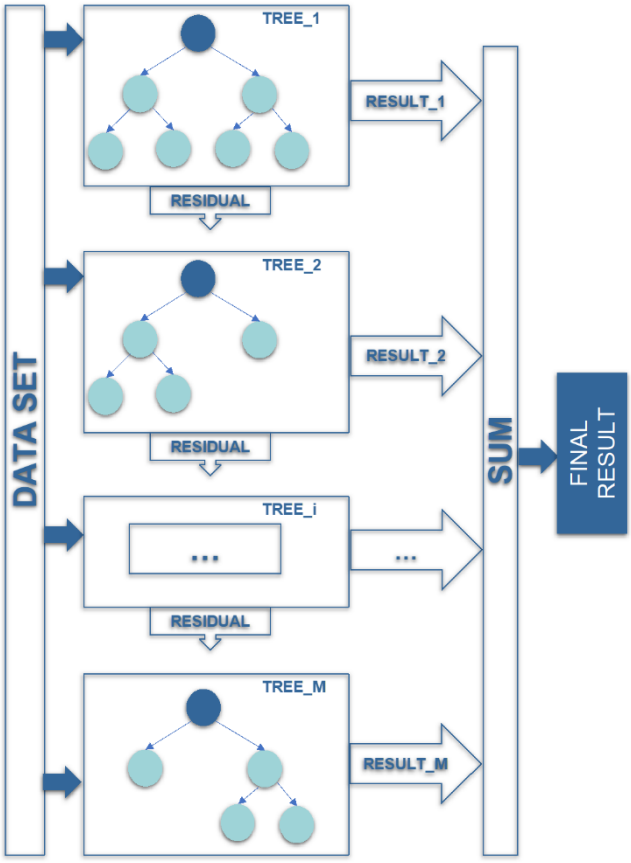

Fig. 2. Flow chart of XGBoost

2.3. Kriging method

The Kriging method, having its origins in geostatics, can be defined as modelling of unknown function through the implementation of a random process [20]. Kriging is based on the idea that the value in a given point can be estimated on the basis of an average of known values in the neighbouring points, assuming that the influences of these points are proportional to the distance to the considered point. In other words, the approximation procedure has to follow the trends of the experimental data and the surrogate function should increase when such increase is expected for an increment of the variables.

Suppose there is a dataset contains m pairs of points $\{\mathbf{x}_i, y_i\}$, where $\mathbf{x}_i \in \mathbb{R}^n$ and $y_i \in \mathbb{R}$ for $i = 1,\ldots,m$. Let the matrix $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_m]^{\mathrm{T}}$ and the vector $\mathbf{y} = [y_1, \ldots, y_m]^{\mathrm{T}}$. The data should satisfy the normalization conditions, i.e.:

$$\mu(\mathbf{X}_{:,j}) = 0, \mathrm{cov}(\mathbf{X}_{:,j}, \mathbf{X}_{:,j}) = 1 \; for \; j = 1,\ldots,n \tag{8}$$

and

$$\mu(\mathbf{X}_{:,j}) = 0, \mathrm{cov}(\mathbf{X}_{:,j}, \mathbf{X}_{:,j}) = 1 \tag{9}$$

where X_(:,j) refers to the jth column of matrix X, μ and cov are mean and covariance, respectively.

Before the Kriging model can be train using dataset, the regression and correlation models must be defined.

As regression functions zero, first and second order polynomials are usually used. In case of zero order polynomial the number of regression functions is equal to p=1 and

$$r_1(\mathbf{x}) = 1 \tag{10}$$

The number of regression functions for first order polynomial is equal to p=n+1 and they are defined by:

$$\begin{aligned} r_1(\mathbf{x}) &= 1 \\ r_2(\mathbf{x}) = x_1, \ldots, r_{n+1}(\mathbf{x}) &= x_n \end{aligned} \tag{11}$$

When the second order polynomial is chosen the number of regression functions is p=(n+1)(n+2)⁄2 and they are given by equations:

$$\begin{aligned} r_1(\mathbf{x}) &= 1 \\ r_2(\mathbf{x}) = x_1, \ldots, r_{n+1}(\mathbf{x}) &= x_n \\ r_{n+2}(\mathbf{x}) = x_1^2, \ldots, r_{2n+1}(\mathbf{x}) &= x_1 x_n \\ r_{2n+2}(\mathbf{x}) = x_2^2, \ldots, r_{3n}(\mathbf{x}) &= x_2 x_n \\ \ldots, r_p(\mathbf{x}) &= x_n^2 \end{aligned} \tag{12}$$

For further purposes, let the vector r(x) and the matrix R be defined as follows:

$$\mathbf{r}(\mathbf{x}) = \left[r_1(\mathbf{x}), r_2(\mathbf{x}), \ldots, r_p(\mathbf{x})\right]^{\mathrm{T}} \tag{13}$$

$$\mathbf{R} = [\mathbf{r}(\mathbf{x}_1), \mathbf{r}(\mathbf{x}_2), \ldots, \mathbf{r}(\mathbf{x}_m)]^{\mathrm{T}} \tag{14}$$

The correlation function is defined by the equation:

$$c(\boldsymbol{\theta}, \mathbf{a}, \mathbf{b}) = \prod_{i=1}^{n} c_i(\theta_i, a_i, b_i) \tag{15}$$

where the one-dimensional correlation $c_i(\theta_i, a_i, b_i)$ usually takes one of the following form:

- exponential function $c_i(\theta_i, a_i, b_i) = \exp(-\theta_i |a_i - b_i|)$ (16)
- general exponential function $c_i(\theta_i, a_i, b_i) = \exp(-\theta_i |a_i - b_i|^{\eta})$, where $0<\eta\leq 2$, (17)
- Gauss function $c_i(\theta_i, a_i, b_i) = \exp(-\theta_i (a_i - b_i)^2)$ (18)
- linear function $c_i(\theta_i, a_i, b_i) = \max\{0, 1 - \theta_i |a_i - b_i|\}$ (19)
- spherical function $c_i(\theta_i, a_i, b_i) = 1 - 1.5\xi_i + 0.5\xi_i^3$, where $\xi_i = \min\{1, \theta_i |a_i - b_i|\}$, (20)
- spline function

$$c_i(\theta_i, a_i, b_i) = \begin{cases} 1 - 15\xi_i^2 + 30\xi_i^3 & for\ 0 \le \xi_i \le 0.2 \\ 1.25(1-\xi_i)^3 & for\ 0.2 < \xi_i < 1 \\ 0 & for\ \xi_i \ge 1 \end{cases} \text{ where } \xi_i = \theta_i |a_i - b_i|. \quad (21)$$

Vector θ, which occurs in equations (16) – (21), is responsible for the rate of correlation descent. The higher value leads to faster decrease.

Base on the correlation function (15) the vector c(x) and the matrix C are defined as follows

$$\mathbf{c}(\mathbf{x}) = [c(\boldsymbol{\theta}, \mathbf{x}_1, \mathbf{x}), \ldots, c(\boldsymbol{\theta}, \mathbf{x}_m, \mathbf{x})]^{\mathrm{T}}, \quad (22)$$

$$\mathbf{C} = [c_{i,j}] = c(\boldsymbol{\theta}, \mathbf{x}_i, \mathbf{x}_j) \text{ for } i, j = 1, \ldots, m. \quad (23)$$

The Kriging model is created using a function g in the form:

$$\hat{y} = g(\mathbf{x}) = \mathbf{r}(\mathbf{x})^{\mathrm{T}}\boldsymbol{\beta} + \mathbf{c}(\mathbf{x})^{\mathrm{T}}\boldsymbol{\gamma}, \quad (24)$$

where vectors r(x) and c(x) are given by equations (6) and (15), respectively, whereas vectors β and γ are established during the model training using the equations:

$$\boldsymbol{\beta} = (\mathbf{R}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{R})^{\mathrm{T}}\mathbf{R}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{y}, \quad (25)$$

$$\mathbf{C}\boldsymbol{\gamma} = \mathbf{y} - \mathbf{R}\boldsymbol{\beta}, \quad (26)$$

where matrixes R and C are defined by (13) and (23), respectively.

Typically, Kriging models are fitted to the data that are obtained for larger experimental areas than the areas used in low-order polynomial regression.

Comparison of effectiveness of Kriging modelling technique with e.g. Artificial Neural Networks can be found in [19].

3. Results

3.1. Classification

As part of the research, the XGBoost algorithm was used to classify individual phases and predict the volume fraction of individual microstructure components. Calculations were made in Python using the following libraries: numpy, pandas, matplotlib, sklearn.

The data set was divided into a training set and a test set using the train_test_split function in a ratio of 80 to 20. Chemical composition and casting wall thickness were assumed as explanatory variables.

In the case of classification, the model was used with the default settings of hyperparameters. The only parameter that was set is the number of trees (n_estimators), which took the value of 20.

The results of the classification, in the form of a confusion matrix for the training and test sets, are shown in Fig. 3 and Fig. 4, respectively.

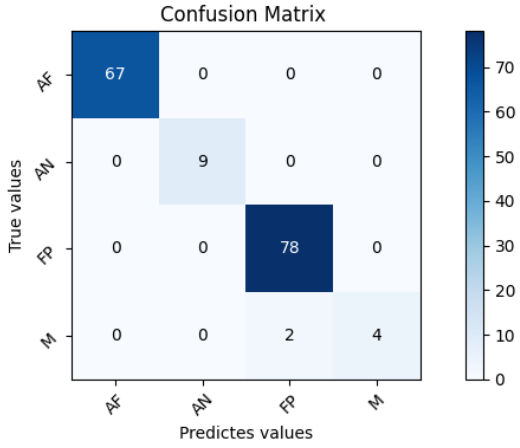

Fig. 3. XGBoost – confusion matrix, training set

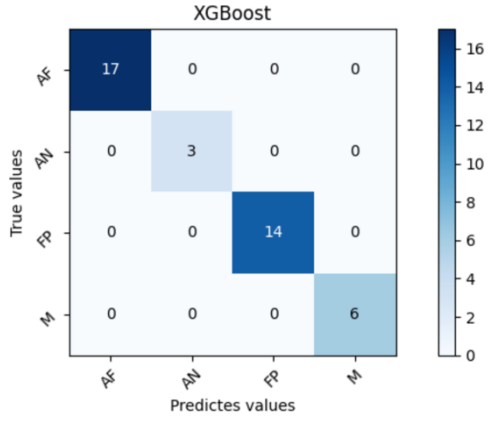

Fig. 4. XGBoost – confusion matrix, testing set

Analyzing the obtained results, it can be seen that in the case of the training set, the classifier classified correctly in most cases, only in the case of martensite there were two incorrect markings. In the case of the test set, however, the classifier achieved 100% accuracy.

The importance of the predictors for the analyzed model is shown in Figure 5. Analyzing the graph shown in Figure 5, it can be seen that molybdenum and nickel have the greatest influence on the type of microstructure. The remaining elements and the wall thickness of the casting have only a small influence on the dependent variable.

3.2. Prediction

The research conducted consisted in the development of models to predict the volume fraction of each phase (ferrite, pearlite, carbides, martensite, ausferrite, austenite) based on the chemical composition and wall thickness of the casting.

A total of 6 models were developed, one for each phase. The basic hyperparameters of the model (learning_rate, max_depth, n_estimators) were selected using the GridSearchCV function from the sklearn library. For each model, the hyperparameter search range was the same, as shown in Tab. 2.

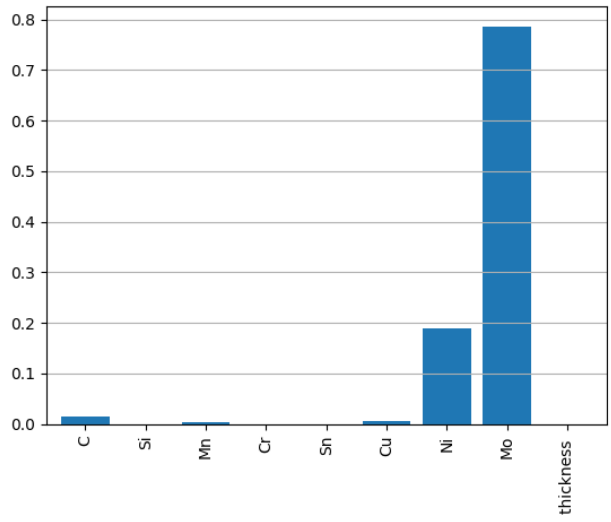

Fig. 5. Predictor importance (XGBoost)

Table 2.
Hyperparameter search range

Hiperparameter	Min	Max	Step
learning_rate	0.01	0.11	0.02
max_depth	3	10	1
n_estimators	20	300	1

The results of searching for hyperparameters with the use of GridSearchCV for each model are presented in Tab. 3.

Table 3.
Values of selected hyperparameters for individual models

Hiperparameter	model F	model P	model C	model M	model_AF	model A
learning_rate	0.09	0.09	0.09	0.05	0.9	0.7
max_depth	3	8	3	3	9	5
n_estimators	298	140	290	240	160	260

As part of the research, the evaluation of the tree learning was also analyzed for the selected learning_rate and max_deph parameters. Mean Absolute Error (MAE) was chosen as the quality of fit metric. The calculation results for ferrite, pearlite, carbides, martensite, ausferrite and atenite are shown in Fig. 6-11, respectively. In the plots, the results for the training set are marked in blue and the results for the test set are marked in orange. The gray line marks the optimal number of trees that was ultimately used in the model.

The coefficient of determination (R^2) was used to assess the quality of the models. The values of the R^2 coefficient for the trained models for the test and training set are presented in Tab. 4.

Analyzing the obtained results (Tab. 4), it can be seen that both for the training and test sets, very high determination coefficients were obtained. For most models, R^2 was obtained at the level of 0.99 for the training set, only the model predicting the share of ferrite obtained a slightly lower value. For the training set, most models predicted with an accuracy exceeding 0.95, only for the model predicting the share of carbides R^2=0.88 was obtained.

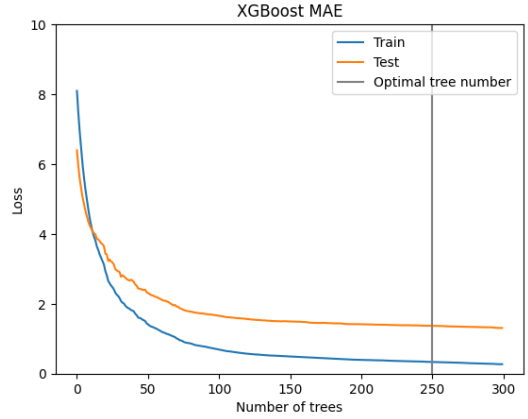

Fig. 6. Change of the MAE error for successive trees for models predicting the volume fraction of ferrite

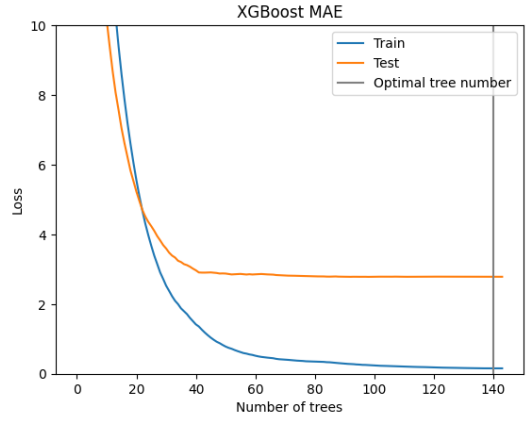

Fig. 7. Change of the MAE error for successive trees for models predicting the volume fraction of perlite

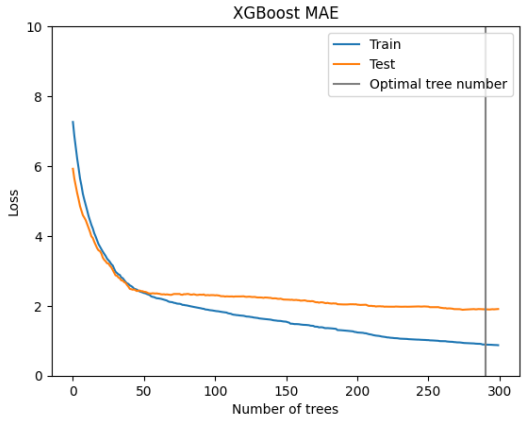

Fig. 8. Change of the MAE error for successive trees for models predicting the volume fraction of carbides

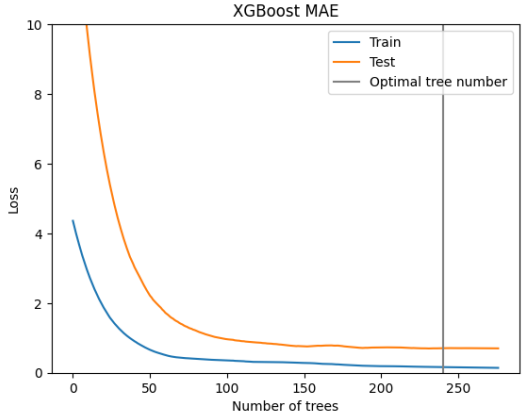

Fig. 9. Change of the MAE error for successive trees for models predicting the volume fraction of martensite

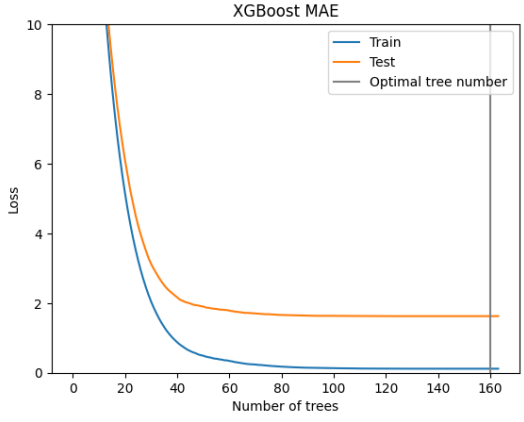

Fig. 10. Change of the MAE error for successive trees for models predicting the volume fraction of ausferrite

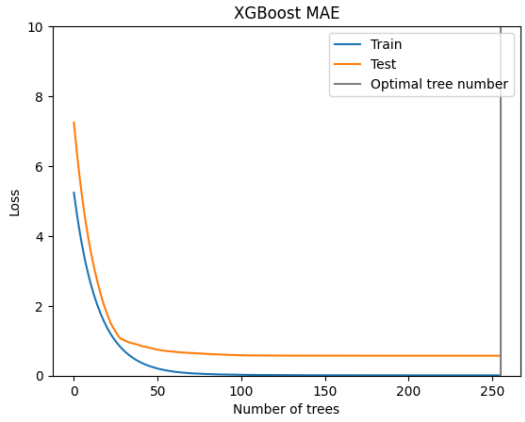

Fig. 11. Change of the MAE error for successive trees for models predicting the volume fraction of austenite.

Table 4.
Coefficients of determination of the developed models for the training and test sets

Model	R^2 train dataset	R^2 test dataset
model_F	0.99	0.95
model_P	0.99	0.98
model_C	0.99	0.98
model_M	0.99	0.99
model_AF	0.99	0.99
model_A	0.99	0.98

Table 5.
Mean absolute error (MAE) of the developed models for the training and test sets

Model	MAE train dataset	MAE test dataset
model_F	0.64	1.89
model_P	0.06	2.50
model_C	1.87	3.63
model_M	0.16	0.70
model_AF	0.12	1.62
model_A	0.01	0.57

As part of the research, an attempt was also made to use one hot encoding to encode the CLASS variable and check the impact of this information on the quality of the models. The results of the coefficient of determination after this modification are presented in table 6.

Table 6.
Coefficients of determination of the developed models for the training and test sets after one hot encoding the CLASS variable

Model	R^2 train dataset	R^2 test dataset
model_F	0.999	0.969
model_P	0.999	0.988
model_C	0.996	0.831
model_M	0.999	0.997
model_AF	0.999	0.992
model_A	0.999	0.993

3.3. Linear regression

The linear regression model is the simplest one among the models which was built within the scope of this paper. The predicted value is the inner product of two vectors:

$$h_\theta(\mathbf{x}) = \boldsymbol{\theta}^T\mathbf{x} \tag{27}$$

where: $\boldsymbol{\theta}$ – vector of model parameters, $\mathbf{x}$ – vector of features (input of the model with added element $x_0 = 1$).

Due to the simplicity of equation 27 it is important to select the appropriate features from all accessible inputs. The modelled process has 9 inputs and 6 outputs. Five sets of the linear correlation coefficient was computed for all combination of inputs and outputs.

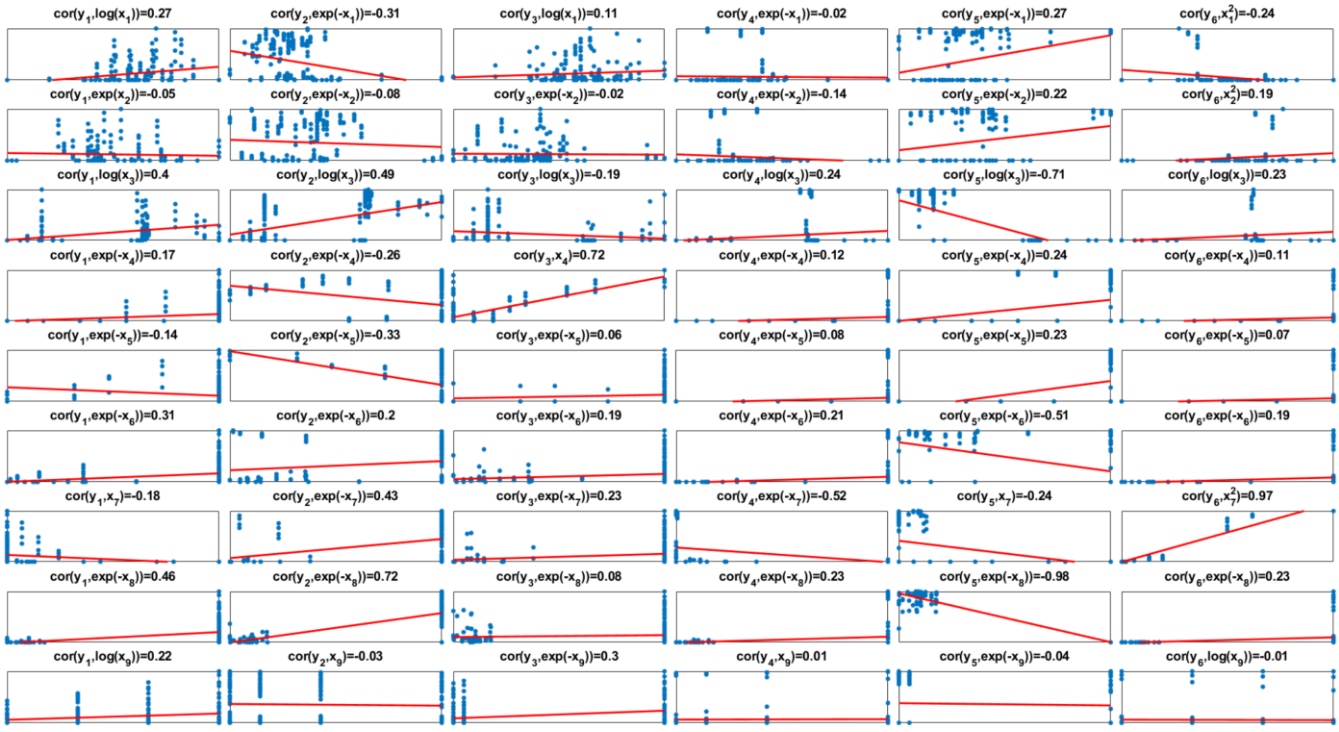

Fig. 12. The correlation between inputs and outputs.

The difference between the sets consisted in transforming the input values with a nonlinear function:

- $c_{i,j} = cor(y_i, x_j), i = 1, \ldots, 6, j = 1, \ldots, 9$
- $c_{i,j} = cor(y_i, x_j^2), i = 1, \ldots, 6, j = 1, \ldots, 9,$
- $c_{i,j} = cor\left(y_i, log(x_j)\right), i = 1, \ldots, 6, j = 1, \ldots, 9,$
- $c_{i,j} = cor\left(y_i, exp(x_j)\right), i = 1, \ldots, 6, j = 1, \ldots, 9,$
- $c_{i,j} = cor\left(y_i, exp(-x_j)\right), i = 1, \ldots, 6, j = 1, \ldots, 9.$

The values of coefficients $c_{i,j}$ from sets were compared and, for each combination of $i = 1, \ldots, 6, j = 1, \ldots, 9$ the one with the higher absolute value was selected as the feature for linear regression model. The results are presented in the figure 12.

After analyzing the first column in the figure 12, the feature vector for the first model (for prediction of the volume fraction of ferrite) was selected as follows:

$$\mathbf{x} = \begin{bmatrix} log(x_1) \\ exp(x_2) \\ log(x_3) \\ exp(-x_4) \\ exp(-x_5) \\ exp(-x_6) \\ x_7 \\ exp(-x_8) \\ log(x_9) \end{bmatrix} \tag{28}$$

The features vectors for the rest of the models were selected in the similar way.

To take into account the bias and the interaction between values in vector $\mathbf{x}$ (28) it was enlarged by value $x_0 = 1$ and all combination of its elements, i.e. $log(x_1) \cdot exp(x_2), log(x_1) \cdot log(x_3)$, ... and to This results in the vector length equals to 46.

To avoid overfitting problem, the regularization term was added to cost function which was used during the models training:

$$J(\boldsymbol{\theta}) = \frac{1}{2m}\sum_{i=1}^{m}(h_\theta(\mathbf{x}_i) - y_i)^2 + \frac{\lambda}{2m}\sum_{i=1}^{n}\theta_i^2 \tag{29}$$

where: m is number of training data, n is the length of feature vector, λ is regularization coefficient.

The training was performed for different values of the regularization parameter λ. The obtained learning curves are presented in the figures 13-18.

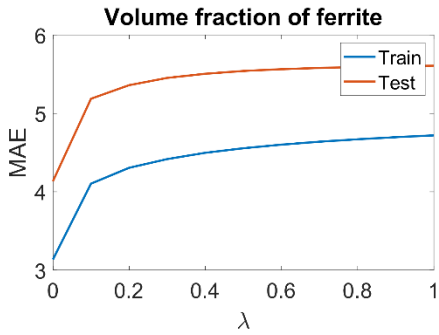

Fig. 13. Learning curves obtained for the model_F.

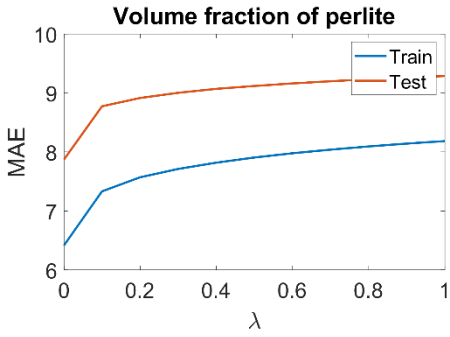

Fig. 14. Learning curves obtained for the model_P.

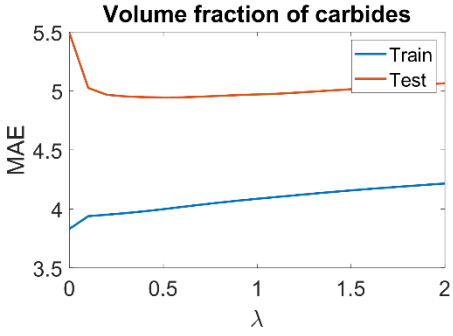

Fig. 15. Learning curves obtained for the model_C.

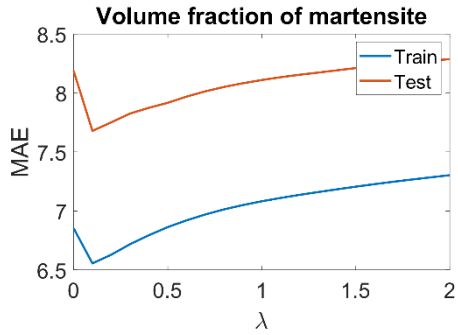

Fig. 16. Learning curves obtained for the model_M.

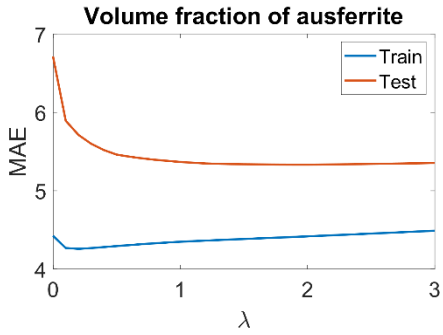

Fig. 17. Learning curves obtained for the model_AF.

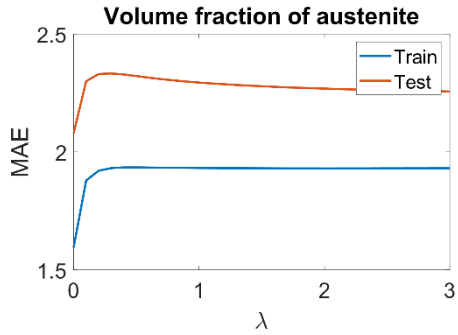

Fig. 18. Learning curves obtained for the model_A.

The obtained errors and coefficients of determination for all linear regression models are presented in tables 7 and 8.

Table 7.
Mean absolute error (MAE) of the developed linear regression models for the training and test sets

Model	MAE train dataset	MAE test dataset
model_F	3.13	4.13
model_P	6.41	7.87
model_C	3.83	4.95
model_M	6.56	7.68
model_AF	4.26	5.37
model_A	1.59	2.08

Table 8.
Coefficients of determination of the developed linear regression models for the training and test sets

Model	R^2 train dataset	R^2 test dataset
model_F	0.916	0.855
model_P	0.931	0.903
model_C	0.824	0.640
model_M	0.684	0.537
model_AF	0.975	0.957
model_A	0.979	0.962

3.4. Kriging model

The Kriging model is a combination of linear regression functions and correlation functions (equation 23). Therefore, the input vector for this model was in the form presented by equation (28). It took into account the linear correlation analysis described in section 3.3, but was not enlarged by combination of its elements. Before training the set of regression and correlation function must be chosen. In case of regression function the decision was made to use the first order polynomials (equation 11). Selecting the zero order polynomials resulted in higher approximation error, while choosing the second order polynomials caused model to be too complicated to train using 200 training records (calculation became too ill conditioned). The training was performed using all correlation functions (equations 16-21). Tables 9 and 10 present the chosen correlation functions as well as mean absolute error and coefficient of determination. The error was calculated only for the test set, because kriging is an interpolation algorithm.

Table 9.
Mean absolute error (MAE) of the developed Kriging models for the test set

Model	Correlation function	MAE test dataset
model_F	spherical function (20)	1.31
model_P	linear function (19)	3.84
model_C	spherical function (20)	4.63
model_M	general exponential function (17)	0.97
model_AF	general exponential function (17)	4.67
model_A	spline function (21)	0.24

Table 10.
Coefficients of determination of the developed Kriging models for test set

Model	Correlation function	R^2 test dataset
model_F	spherical function (20)	0.99
model_P	linear function (19)	0.988
model_C	spherical function (20)	0.839
model_M	general exponential function (17)	0.997
model_AF	general exponential function (17)	0.984
model_A	spline function (21)	0.999

3.5. Artificial neural networks

The last model was built using the artificial neural networks (ANN). ANNs are built with a given number of artificial neurons which are arranged in three layers: input, hidden and output layers. The number of layers and number of neurons define the network topology. The number of hidden layers is usually not higher than 2. The number of neurons in input and output layer depends on the dimension of training data, while the number of neurons in hidden layer(s) is chosen before training. There is no method which would be able to determine the best networks topology. Therefore, the training of each network was performed 200 times. Each time the numbed of hidden layers (1 or 2) and the number of neurons (5-20) was selected randomly. The activation function in all neurons in input and hidden layers was sigmoid, while the activation function in neuron in output layer was linear.

Due to nonlinearity of an activation functions of neurons in input and hidden layers, ANNs are able to learn any nonlinear relation. Therefore, there is no need to perform any correlation analysis (like it was done in case of linear regression model). The input values for all ANN models was original vector x (composed of chemical composition of the alloy and the casting wall thickness. The best topology of each network is presented in table 11, while the mean absolute error and coefficient of determination in tables 12 and 13.

Table 11.
Topologies of the best networks

Model	Number of neurons is each layer	Number of weight elements
model_F	9-14-6-1	237
model_P	9-11-1	122
model_C	9-7-10-1	161
model_M	9-14-10-1	301
model_AF	9-13-12-1	311
model_A	9-9-14-1	245

Table 12.
Mean absolute error (MAE) of the developed ANN models for the training and test sets

Model	MAE train dataset	MAE test dataset
model_F	1.09	1.86
model_P	2.48	4.04
model_C	1.65	2.11
model_M	0.21	0.28
model_AF	2.71	2.69
model_A	0.08	0.09

Table 13.
Coefficients of determination of the developed ANN models for the training and test sets

Model	R^2 train dataset	R^2 test dataset
model_F	0.985	0.956
model_P	0.986	0.959
model_C	0.944	0.923
model_M	0.998	0.999
model_AF	0.972	0.995
model_A	1	1

4. Discussion and summary

The discussion focuses on the prediction of phase volume fractions in the microstructure of compacted graphite iron (CGI). Previous studies by the authors [15-18] have developed prediction models, particularly for ausferrite. In this study, the authors proposed a solution based on kirging and boosted trees to solve the problem of overfitting and extrapolation error in property modeling. In the second phase, the appropriate model of an Artificial Neural Network and linear regression has been developed to compare the results. This approach allows the use of relatively simple and interpretable models, resulting in an efficient architecture that allows fast and accurate assessment (especially the classification tree with XGBoost) of the future phase composition in the CGI microstructure. Figure 19 presents the comparison of the MEA (mean absolute error). As expected, linear regression produced the worst results. The other algorithms oscillated around the error rate. None of the algorithms outperformed the others in any of the phases. While XGBoost was the most accurate in predicting pearlite and ausferrite, neural networks were the best at predicting carbides, martensite, and austenite. Kriging, on the other hand, was best for ferrite and very good for austenite. This juxtaposition shows another feature of the presented models, each of them has different advantages and their combined performance provides the analyst with the most new knowledge about the behavior and dependencies in the studied phenomena.

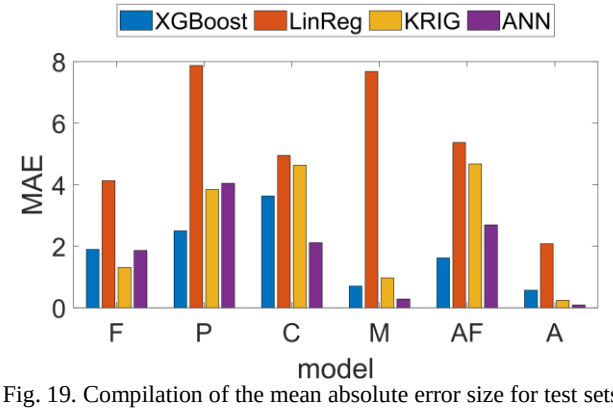

Fig. 19. Compilation of the mean absolute error size for test sets for developed models for each phase (F- ferrite, P - pearlite, C - carbides, M - martensite, AF - ausferrite, A - austenite).

The proposed solution combines the advantages of all those techniques, allowing easy interpretation of dependencies and accurate prediction. The presented methodology of data-driven modeling for the prediction of compacted graphite iron microstructure proved to be highly effective for the experimental data set. It improved the overall prediction of microstructure composition and facilitated chemical composition selection. The authors believe that as the training database grows with subsequent material experiments, this approach can be successfully applied to the design of new chemical compositions, potentially including other alloying additives.

Acknowledgements

This study was carried out as part of the fundamental research financed by the Ministry of Science and Higher Education, grant no. 16.16.110.663

References

[1] König, M. (2010). Literature review of microstructure formation in compacted graphite iron. *International Journal of Cast Metals Research*. 23(3), 185-192. https://doi.org/10.1179/136404609X12535244328378.

[2] Dawson, S. & Hang, F. (2009). Compacted graphite iron-a material solution for modern diesel engine cylinder blocks and heads. *China Foundry*. 6(3), 241-246.

[3] Chen, Y., Pang, J. C., Li, S. X., Zou, C. L. & Zhang, Z. F. (2022). Damage mechanism and fatigue strength prediction of compacted graphite iron with different microstructures. *International Journal of Fatigue*. 164, 107126, 1-14. https://doi.org/10.1016/j.ijfatigue.2022.107126.

[4] Sandoval, J., Ali, A., Kwon, P., Stephenson, D. & Guo, Y. (2023). Wear reduction mechanisms in modulated turning of compacted graphite iron with coated carbide tool. *Tribology International*. 178, 108062, 1-13. https://doi.org/10.1016/j.triboint.2022.108062.

[5] Hosadyna-Kondracka, M., Major-Gabryś, K., Warmuzek, M. & Brůna, M. (2022). Quality assessment of castings manufactured in the technology of moulding sand with furfuryl-resole resin modified with PCL additive. *Archives of Metallurgy and Materials*. 67(2), 753-758. https://doi.org/10.24425/amm.2022.137814.

[6] Mrzygłód, B., Łukaszek-Sołek, A., Olejarczyk-Wożeńska, I. & Pasierbiewicz, K. (2022). Modelling of plastic flow behaviour of metals in the hot deformation process using artificial intelligence methods. *Archives of Foundry Engineering*. 22(3), 41-52. DOI: 10.24425/afe.2022.140235.

[7] Palkanoglou, E.N., Baxevanakis, K.P. & Silberschmidt, V.V. (2022). Thermal debonding of inclusions in compacted graphite iron: Effect of matrix phases. *Engineering Failure Analysis*. 139, 106476, 1-13. https://doi.org/10.1016/j.engfailanal.2022.106476.

[8] Patel, M., Dave, K. (2022). An insight of compacted graphite iron (CGI) characteristics and its production: a review. Recent Advances in Manufacturing Processes and Systems: Select Proceedings of RAM 2021, 131-148.

[9] Górny, M., Lelito, J., Kawalec, M. & Sikora, G. (2015). Influence of structure on the thermophisical properties of thin walled castings. *Archives of Foundry Engineering*. 15(2), 23-26.

[10] Górny, M., Kawalec, M., Witek, G. & Rejek, A. (2019). The influence of wall thickness and mould temperature on structure and properties of thin wall ductile iron castings. *Archives of Foundry Engineering*. 19(2), 55-59. DOI: 10.24425/afe.2019.127116.

[11] Saka, S.O., Seidu, S.O., Akinwekomi, A.D. & Oyetunji, A. (2021). Alloying elements variant on the development of antimony modified compacted graphite iron using rotary

furnace. *Annals of the Faculty of Engineering Hunedoara*. 19(2), 13-22.
[12] Soiński, M.S., Jakubus, A., Borowiecki, B. & Mierzwa, P. (2021). Initial assessment of graphite precipitates in vermicular cast iron in the as-cast state and after thermal treatments. *Archives of Foundry Engineering*. 21(4), 131-136.
[13] Domeij, B., Elfsberg, J. & Diószegi, A. (2023). Evolution of dendritic austenite in parallel with eutectic in compacted graphite iron under three cooling conditions. *Metallurgical and Materials Transactions* B. 1-16.
[14] Ren, Z., Jiang, H., Long, S. & Zou, Z. (2023). On the mechanical properties and thermal conductivity of compacted graphite cast iron with different pearlite contents. *Journal of Materials Engineering and Performance*. 1-9. https://doi.org/10.1007/s11665-023-07823-7.
[15] Gumienny, G., Kacprzyk, B., Mrzygłód, B. & Regulski, K. (2022). Data-driven model selection for compacted graphite iron microstructure prediction. *Coatings.* 12(11), 1676, 1-18. DOI: 10.3390/coatings12111676.
[16] Mrzygłód, B., Gumienny, G., Wilk-Kołodziejczyk, D. & Regulski, K., (2019). Application of selected artificial intelligence methods in a system predicting the microstructure of compacted graphite iron. *Journal of Materials Engineering and Performance.* 28, 3894-3904. DOI: 10.1007/s11665-019-03932-4.
[17] Wilk-Kołodziejczyk, D., Regulski, K., Gumienny, G. & Kacprzyk, B. (2018). Data mining tools in identifying the components of the microstructure of compacted graphite iron based on the content of alloying elements. *International Journal of Advanced Manufacturing Technology*. 95(9-12), 3127-3139. DOI 10.1007/s00170-017-1430-7.
[18] Wilk-Kołodziejczyk, D., Kacprzyk, B., Gumienny, G., Regulski, K., Rojek, G. & Mrzygłód, B., (2017). Approximation of ausferrite content in the compacted graphite iron with the use of combined techniques of data mining, *Archives of Foundry Engineering*. 17(3), 117-122. DOI 10.1515/afe-2017-0102.
[19] Kusiak, J., Sztangret, Ł. & Pietrzyk, M. (2015). Effective strategies of metamodelling of industrial metallurgical processes. *Advances in Engineering Software*. 89, 90-97. DOI: 10.1016/j.advengsoft.2015.02.002.
[20] Sacks, J., Welch, W.J., Mitchel, T. & Wynn, H.P., (1989) Design and analysis of computer experiments. *Stat Sci*. 4, 409-435. DOI: 10.1214/ss/1177012413.
[21] Fragassa, C. (2022) Investigating the material properties of nodular cast iron from a data mining perspective. *Metals.* 12(9), 1493, 1-26. DOI: 10.3390/met12091493.
[22] Huang, W., Lyu, Y., Du, M., Gao, S-D., Xu, R-J., Xia, Q-K. & Zhangzhou, J. (2022). Estimating ferric iron content in clinopy-roxene using machine learning models. *American Mineralogist.* 107, 1886-1900. DOI: 10.2138/am-2022-8189.
[23] Sika, R., Szajewski, D., Hajkowski, J. & Popielarski, P. (2019). Application of instance-based learning for cast iron casting defects prediction. *Management and Production Engineering Review.* 10(4), 101-107. DOI: 10.24425/mper.2019.131450.
[24] Chen, S. & Kaufmann, T. (2022). Development of data-driven machine learning models for the prediction of casting surface defects. *Metals.* 12(1), 1-15. DOI: 10.3390/met12010001
[25] Alrfou, K., Kordijazi, A., Rohatgi, P. & Zhao, T. (2022). Synergy of unsupervised and supervised machine learning methods for the segmentation of the graphite particles in the microstructure of ductile iron. *Materials Today Communications.* 30. 103174. DOI: 10.1016/j.mtcomm.2022.103174.
[26] Vantadori, S., Ronchei, C., Scorza, D., Zanichelli, A. & Luciano, R. (2022). Effect of the porosity on the fatigue strength of metals. *Fatigue & Fracture of Engineering Materials & Structures.* 45(9), 2734-2747. https://doi.org/10.1111/ffe.13783.
[27] Dučić, N., Jovičić, A., Manasijević, S., Radiša, R., Ćojbašić, Z. & Savković, B. (2020). Application of machine learning in the control of metal melting production process. *Applied Sciences.* 10(17), 6048, 1-15. DOI: 10.3390/app10176048
[28] Kihlberg, E., Norman, V., Skoglund, P., Schmidt, P. & Moverare, J. (2021). On the correlation between microstructural pa-rameters and the thermo-mechanical fatigue performance of cast iron. *International Journal of Fatigue.* 145, 106112, 1-10. DOI: 10.1016/j.ijfatigue.2020.106112.
[29] Hernando, J.C., Elfsberg, J., Ghassemali, E., Dahle, A.K. & Diószegi, A. (2020). The role of primary austenite morphology in hypoeutectic compacted graphite iron alloys. *International of Metalcasting.* 14, 745-754. DOI: 10.1007/s40962-020-00410-9.
[30] Regordosa, A., de la Torre, U., Loizaga, A., Sertucha, J. & Lacaze, J. (2020). Microstructure Changes During Solidification of Cast Irons: Effect of Chemical Composition and Inoculation on Competitive Spheroidal and Compacted Graphite Growth. *International of Metalcasting*. 14, 681-688. DOI: 10.1007/s40962-019-00389-y.
[31] Ribeiro B.C.M., Rocha F.M., Andrade B.M., Lopes W., Corrêa E.C.S., (2020). Influence of different concentrations of silicon, copper and tin in the microstructure and in the mechanical properties of compacted graphite iron, *Materials Research.* 23(2), e2019-0678, 1-10. DOI: 10.1590/1980-5373-MR-2019-0678.
[32] Tan, P.-N., Steinbach, M. & Kumar, V. (2005). *Introduction to Data Mining*. Boston: Pearson Addison-Wesley.
[33] Rokach, L. & Maimon, O. (2005). Top-down induction of decision trees classifiers-a survey. *IEEE Transactions on Systems Man and Cybernetics Part C (Applications and Reviews).* 35(4), 476-487.
[34] Barros, R.C., de Carvalho, A. & Freitas, A.A. (2015). *Automatic Design of Decision-Tree Induction Algorithms*, Springer International Publishing.
[35] Regulski, K., Wilk-Kołodziejczyk, D. & Gumienny, G. (2016). Comparative analysis of the properties of the Nodular Cast Iron with Carbides and the Austempered Ductile Iron with use of the machine learning and the support vector machine. *The In-ternational Journal of Advanced Manufacturing Technology*. 87(1), 1077-1093. DOI: 10.1007/s00170-016-8510-y.
[36] Rui, G., Zhiqian, Z., Tao, W., Guangheng, L., Jingyi, Z. & Dianrong, G., (2020) Degradation state recognition of piston pump based on ICEEMDAN and XGBoost, *Applied Sciences.* 10(18), 6593, 1-17. DOI:10.3390/app10186593