Artificial Intelligence Approaches to Determine Graphite Nodularity in Ductile Iron

Maximilian Brait[a, b]*, Eduard Koppensteiner[a], Gerhard Schindelbacher[a], Jiehua Li[b], Peter Schumacher[a, b]

[a] Austrian Foundry Research Institute, Parkstraße 21, 8700 Leoben, Austria
[b] Chair of Casting Research, Montanuniversität Leoben, Franz-Josef-Straße 18, 8700 Leoben, Austria
* e-mail: maximilian.brait@ogi.at

Abstract

The complex metallurgical interrelationships in the production of ductile cast iron can lead to enormous differences in graphite formation and local microstructure by small variations during production. Artificial intelligence algorithms were used to describe graphite formation, which is influenced by a variety of metallurgical parameters. Moreover, complex physical relationships in the formation of graphite morphology are also controlled by boundary conditions of processing, the effect of which can hardly be assessed in everyday foundry operations. The influence of relevant input parameters can be predetermined using artificial intelligence based on conditions and patterns that occur simultaneously. By predicting the local graphite formation, measures to stabilise production were defined and thereby the accuracy of structure simulations improved. In course of this work, the most important dominating variables, from initial charging to final casting, were compiled and analysed with the help of statistical regression methods to predict the nodularity of graphite spheres. We compared the accuracy of the prediction by using Linear Regression, Gaussian Process Regression, Regression Trees, Boosted Trees, Support Vector Machines, Shallow Neural Networks and Deep Neural Networks. As input parameters we used 45 characteristics of the production process consisting of the basic information including the composition of the charge, the overheating time, the type of melting vessel, the type of the inoculant, the fading, and the solidification time. Additionally, the data of several thermal analysis, oxygen activity measurements and the final chemical analysis were included.
Initial programme designs using machine learning algorithms based on neural networks achieved encouraging results. To improve the degree of accuracy, this algorithm was subsequently adapted and refined for the nodularity of graphite.

1. INTRODUCTION

In recent years microstructure modelling with commercial software packages has become state of the art. The correct knowledge of the mechanical properties of a given casting is important to improve its functionalities and reduce weight, for example in the instance of design improvement using topology optimisation methods [1, 2]. However, in the case of ferritic-perlitic nodular cast iron grades, local microstructural properties could not be satisfactorily calculated. For the case of hardness, the trend across the wall thickness between the measured values and the simulation was detrimental to each other [3].

Graphite formation is strongly responsible for the resulting properties of an iron casting [4–6].

The theoretical background for the description of nucleation and graphite formation in iron-carbon alloys is well documented. In collective works by Bauer [4], Herfurth [7] and Stefanescu [8], the most important hypotheses on the formation of graphite morphologies are elucidated.

The use of machine learning and deep learning methods for regression, or clustering in material science is very common due to the high accuracy of the results using artificial neural networks [9–11]. They can be very efficiently used in the case of very complex problems or if there is no algorithmic solution available [12]. The state of the art on artificial

neural networks and deep learning methods are reviewed, e.g., in reference [13] and [14].

Recently, papers have been published which use the results of thermal analysis to predict the microstructures in ductile cast iron using artificial intelligence [15]. In work by Liliac, Yamamoto and Ogi, graphite nucleation and graphite morphology were modelled by means of neural networks [16]. However, input parameters for their prepared neural network were limited to the duration of superheating, supercooling based on stable thermal analysis, and magnesium and sulphur content. Overall, publications provide information on the most simplistic input parameters, possible methods and algorithms, which are, however, extended and adapted to the industrial environment in the present work. The complex physical relationships in the formation of graphite morphology are determined by parameters whose degree of influence is hardly detectable in everyday foundry work [8, 17, 18]. By means of artificial intelligence, the influence and their strength of various input parameters can be assessed on the basis of simultaneously occurring conditions and patterns [19]. By this means it will be possible to realistically simulate and model structural properties [20, 21], based on more precise interpretation of graphite formation. Improved knowledge on the influences and interactions of different process parameters on the microstructure formation should lead to an improvement in quality in the production of nodular cast iron and thus reduce development costs [22, 23] as well as to a deeper understanding of graphite formation in cast iron.

In the course of this work, the most important influencing factors, from melting to casting and their boundary conditions are analysed with the help of regression methods for graphite nodularity. Alongside the data from thermal analysis, as novel input parameter oxygen activity measurements, temperature measurements and expanded chemical analyses are used as input variables for the machine learning models. Additionally, a preceding thermal modelling with commercial simulation software used in the foundry industry was taken as input parameter for the prepared prediction model. A large matrix of experiments was compiled and evaluated for this purpose. A model for the prediction of nucleation in spheroidal graphite cast iron has been created using regression methods, with a particular focus on the nodularity of graphite.

The aim of the paper is to obtain a better understanding of graphite nodularity and its influencing factors including industrial boundary conditions in the production of nodular cast iron.

2. METHODS

2.1. Experimental methods

A suitable test specimen was constructed and designed for casting. The test specimen has different wall thicknesses and thus different solidification times (Fig. 1). The casting system is designed according to Nielsen [24] to achieve a very uniform laminar filling. The mould is vertically split and consists of two coldbox cores (Fig. 2). The cores are produced by means of a core shooter (Röpperwerk 5 l), which ensures a productive and at the same time dimensional accurate sand core moulds.
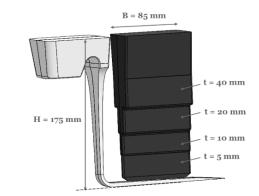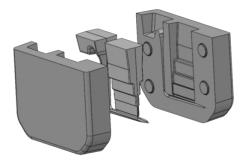

Fig. 1. Stepped test specimen

Fig. 2. Vertically split mould for the casts made from coldbox cores

Experimental casts were made with different parameters according to Table 1, which originate from the industrial environment and serve as input parameters for the prediction algorithms, resulting in an experimental matrix including 416 variations. In total we made 26 experimental batches at the test foundry of the Austrian Foundry Research Institute over a time period of one year. 16 batches with an identical basic composition out of raw-iron, pure-iron, scrap and alloying elements and variations in the melting vessel, the overheating time, and the inoculant, which effect 256 combinations. Each charge results in 16 variations, as a test specimen is poured off after 2, 4, 6 and 8 minutes and specimens are taken from each test specimen at a solidification rate of 50, 90, 180, and 360 s. The other 10 batches (160 combinations) differ from the previous 16 batches by additionally varying the combination of raw iron, pure iron, scrap and alloy elements. For each variation there is only one casting, but 160 variations only differ in the basic composition of the charge.

Table 1
Variation of the experimental parameters

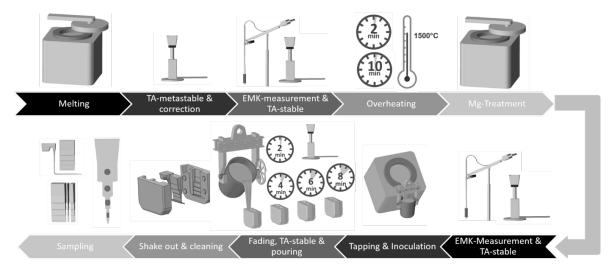| Parameter	Variation			
1 Charge	Raw iron [%]	Pure iron [%]	Scrap [%]	Alloy elements [%]
2 Melting vessel	Crucible	Lined ladle	–	–
3 Overheating time	2 min	10 min	–	–
4 Inoculant	VP216	SMW 605	Super-seed 75	Amerinoc
5 Fading	2 min	4 min	6 min	8 min
6 Solidification	50 s	90 s	180 s	360 s

Fig. 3. Schematic representation of the sample preparation process

Figure 3 shows a schematic diagram of the sample preparation process. Each test began with the melting of the basic charge. After an initial metastable thermal analysis, the melt was corrected to obtain a targeted silicon content of 2.30% and a carbon content of 3.65% in the castings. This was followed by a measurement of the oxygen activity in the melt by using a celox sonde and a stable thermal analysis was made. Then the melt was superheated to 1500°C for 2 min or 10 min respectively. After the magnesium treatment according to the bell dipping process, with 1.2% MgFeSi master alloy (Elmag: 48% Si and 6% Mg), the melt was tapped into the ladle with ladle inoculation. The inoculants listed in Table 1 were always used with 0.4% addition. This was followed by a further measurement of oxygen activity. 2 min after the melt treatment, a stable thermal analysis was performed, and the first test specimen was cast. After 4 min, 6 min and 8 min, further thermal analyses were performed and at each timestep a test specimen was cast. Finally, the castings were knocked out and samples were taken and prepared for metallographic investigations and chemical analysis.

In order to evaluate the effects of the individual changes, the following measurements were conducted:

- thermal analysis (Appromace data systems GmbH, OCC);
- oxygen activity measurement in the melt (Heraeus Celox Foundry [8]);
- temperature measurement (TMH, type-K);
- chemical analysis of the charge and the final castings (Inductively coupled plasma atomic emission spectroscopy (ICP-AES), carbon and sulphur analysis by combustion (LECO)).

10 micrographs were taken of each sample, and the microstructures were evaluated by using the software Amguss V1.1.60 [25, 26, 27]. The focus of this paper is on nodularity. We analysed the nodularity according to EN ISO 945-4 [28, 29] with a minimum particle size of 10 μm. It was calculated corresponding to Formula (1) [28, 29]. The roundness of a particle was calculated according to Formula (2) [28, 29], which determines whether the particle is assigned to class V or IV. Particles with a roundness between 0.6 and 0.8 are assigned

to class V and particles with a roundness between 0.8 and 1 are assigned to class VI.

$$Nodularity = \frac{\sum Area_{VI(0.8-1)} + \sum Area_{V(0.6-0.8)}}{\sum Area_{all}} \quad (1)$$

$$Roundness = \frac{4 \times Area}{\pi \times max.length\,of\,the\,graphite\,particle^2} \quad (2)$$

The mean value of the nodularity was calculated from 10 micrographs each. In this way, the 4160 × 46 matrix was reduced again to a 416 × 46 matrix. This matrix was the basis for all correlation calculations (45 input parameters, 1 label).

2.2. Modelling methods

Results were evaluated and collected in a Python data-frame. Rarely, measured values are missing from the tests due to broken thermocouples (1.9%) or due to measurement problems during oxygen activity measurement with the Celox probes (7.7%). DataWig [30], a software package written in Python, was used to generate the missing measurements. This tool allows imputation of the missing measurements through a combination of deep learning feature extractors [31] with automatic hyperparameter tuning [30, 32]. Missing readings for oxygen activity, maximum temperature, and upper and lower eutectic temperature were calculated.

Different statistical methods for regression (Linear Regression (LR) [33], Gaussian Process Regression (GPR) [34], Regression Trees (RT) [35], Boosted Trees (BT) [36], Support Vector Machines (SVM) [37], Shallow Neural Networks (SNN) [38] and Deep Neural Networks (DNN) [39, 40]) were applied to the prepared dataset. Software Matlab® [41] and the deep learning framework Keras [42] written in Python were used for data analysis. Keras is the high-level Application Programming Interface (API) of TensorFlow 2 [43].

Different statistical methods were compared based on the coefficient of determination (R^2), if there is a linear correlation between true and predicted values, and the Root Mean Square Error (RMSE) calculated from the application to a separate test data set [44].

The basis for the correlations is formed by a 416 × 46 data matrix. For evaluating the regression methods, we split the dataset in data for training (60%), validation (20%) and testing (20%) not random, but always all samples of a single casting together. The dataset for training and validation is extended to a 10,000 × 46 matrix by a bootstrap based sample augmentation by resampling [45–48]. 45 apparently relevant input parameters, listed in Table 2 except nodularity, were used to predict nodularity by statistical methods. The prepared model and analysis are valid for the input and output parameters in the range of the minimal (min.) and maximal (max.) values according to Table 2. Additionally, means ($\bar{x}$), and coefficients of variation (cv) of the used dataset are listed.

Table 2
Minimal and maximal values of the gathered dataset

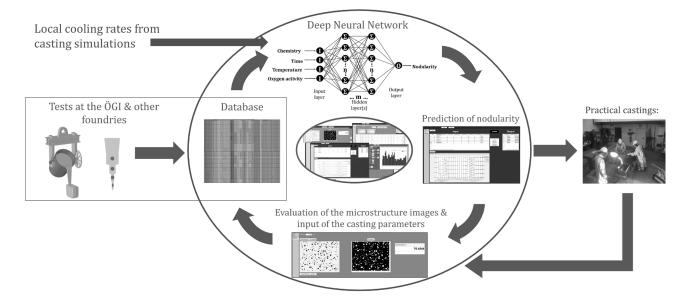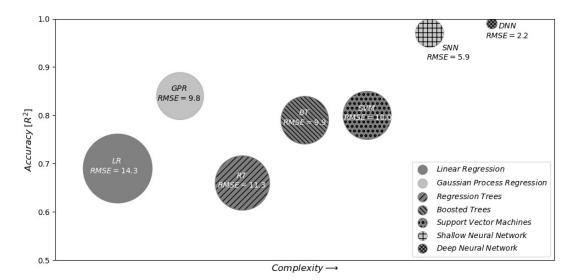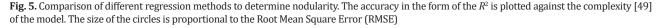| Parameter	Label	min.	max.	$\bar{x}$	cv
Different melting vessel [1] (0 = graphite crucible; 1–15 lined ladles with different degree of wear and tear)	melting-vessel	0	15	4.21	1.316
Time the melt was superheated [min]	overheating-time	2	10	3.56	0.890
Fading time of the inoculation [min]	fading	0	8	4.13	0.651
(0 = late stream inoculation)	–	–	–	–	–
Solidification time [s]	solidification	50	360	170.66	0.703
Scrap content in the charge [%]	scrap	0	58.18	26.38	0.869
Raw-iron content in the charge [%]	raw-iron	15	82	54.94	0.468
Coal content in the charge [%]	C-charge	0	1.3	0.25	1.920
Ferro-silicon content in the charge (FeSi 75) [%]	Si-charge	0	0.8	0.34	1.176
Inoculant (1 = VP216, 2 = SMW605, 3 = Superseed 75, 4 = Amerinoc)	inoculant	1	4	2.33	0.481
Liquidus temperature of the untreated melt [°C]	TLiq_TA1	1149.3	1183.7	1169.48	0.006
Upper eutectic temperature of the untreated melt [°C]	TEup_TA1	1121.0	1166.9	1150.79	0.006
Lower eutectic temperature of the untreated melt [°C]	TElow_TA1	1121.0	1166.8	1146.47	0.006
Oxygen activity of the untreated melt [ppb]	O2_EMK1	541.5	5341.3	2006.79	0.498
Superheating temperature [°C]	T-overheating	1373	1565.1	1474.33	0.025
Casting temperature [°C]	Tcast	1152.0	1354.1	1259.29	0.045
Liquidus temperature of the final treated melt [°C]	TLiq_TA2	1138.74	1163.7	1151.85	0.004
Upper eutectic temperature of the final treated melt [°C]	TEup_TA2	1137.1	1160.8	1154.89	0.003
Lower eutectic temperature of the final treated melt [°C]	TElow_TA2	1135.2	1157.6	1148.86	0.004
Temperature at final measurement of oxygen activity [°C]	T_EMK2	1355	1445	1414.31	0.014
Oxygen activity of the final treated melt [ppb]	O2_EMK2	20.7	197.5	108.49	0.364
Aluminium content [%]	Al	8.0E-3	0.019	0.00123	2.846
Wolfram content [%]	W	1.1E-3	0.017	0.0092	0.401
Cobalt content [%]	Co	1E-3	0.006	0.00238	0.639
Niobe content [%]	Nb	1E-3	0.012	0.00155	1.548
Bor content [%]	B	2.0E-4	0.0072	0.00274	0.850
Arsen content [%]	As	1E-3	0.004	0.00113	0.477
Tin content [%]	Sn	1E-3	0.01	0.00154	1.234
Zinc content [%]	Zn	1E-3	0.002	0.00102	1.422
Nitrogen content [%]	N	3.0E-3	0.012	0.00521	0.507
Cerium content [%]	Ce	1E-3	0.025	0.00377	1.281
Lantan content [%]	La	1E-3	0.002	0.00125	0.347
Selenium content [%]	Se	3.0E-3	0.008	0.00324	0.298
Tellurium content [%]	Te	1E-3	0.007	0.00428	2.255
Phosphorus content [%]	P	1.6E-2	0.05	0.0409	0.022
Sulphur content [%]	S	4.0E-3	0.015	0.00937	0.340
Copper content [%]	Cu	2.6E-2	0.59	0.445	0.045
Chrome content [%]	Cr	1E-2	0.021	0.0132	0.231
Nickel content [%]	Ni	4.0E-3	0.048	1.31	0.008
Molybdenum content [%]	Mo	2E-3	0.007	0.00309	0.392
Titan content [%]	Ti	6.0E-3	0.012	0.0109	0.150
Vanadium content [%]	V	2E-3	0.004	0.00302	0.150
Carbon content [%]	C	3.46	3.74	3.61	0.016
Silicon content [%]	Si	2.09	2.57	2.31	0.051
Manganese content [%]	Mn	4.94E-2	0.15	0.0789	0.369
Magnesium content [%]	Mg	1.5E-2	0.039	0.0266	0.185
Nodularity of graphite spheres [%]	Nodularity	12.5	93.9	62.42	1.316

Fig. 4. Graphical user interface to operate the software for prediction of nodularity based on artificial neural networks

The algorithm is supposed to learn and improve itself with each new test due to the increasing amount of data. In addition, a user-friendly interface was being developed for practical use and equipped with the necessary functions. In Figure 4 the procedure to predict the nodularity and the sphere density based on deep neural networks is visualised.

3. RESULTS AND DISCUSSION

By using Linear Regression (LR) only a low accuracy was achieved for nodularity. This simple model obtained a R^2 for predicting the nodularity of 0.69 and a Root Mean Square Error (RMSE) of 14.3. Using a rational quadratic Gaussian Process Regression (GPR) the model improved to a R^2 of 0.84 and a RMSE of 9.8. Conducting modelling of Regression Trees (RT) and Boosted Trees (BT) the accuracy could not be improved in comparison to the Gaussion Process Regression despite the complexity of these models being higher. Promising methods

for this kind of regression are Support Vector Machines (SVM). A model with a R^2 of 0.8 and a RMSE of 10.0 was obtained for SVM. Finally artificial neural networks were applied to the problem. With a Shallow Neural Network (SNN), using a Sigmoid activation function, one hidden layer and 10 nodes the accuracy was improved to predict the nodularity. The R^2 for this machine learning algorithm was 0.95 and its RMSE 5.9. Additionally a Deep Neural Network (DNN) consisting of 5 hidden layers, 40 nodes at each layer with a ReLU-activation function was used. The ReLU-function improves the learning performance of a Deep Neural Network in contrast to the Sigmoid function, which limits the outputs of the node to the value of one regardless of the size of the input [20]. This DNN achieved the highest accuracy. The RMSE for predicting the nodularity is 2.2 and the R^2 is 0.99. In Figure 5 the results of the tested regression methods are compared. The accuracy (R^2) is plotted against the model complexity according to [49]. The size of the circles is proportional to the RMSE for nodularity.

Fig. 5. Comparison of different regression methods to determine nodularity. The accuracy in the form of the R^2 is plotted against the complexity [49] of the model. The size of the circles is proportional to the Root Mean Square Error (RMSE)

In Figure 6a the density distribution of the errors (true nodularity minus calculated nodularity) by the DNN is visualised as histogram. Additionally, the area between the 10% percentile and the 90% percentile are highlighted by a lighter grey background. In the diagram on the right in Figure 6b the predicted values are plotted over the true values for nodularity.

Virtual tests and an investigation of the influence of a single parameter on nodularity were performed by the trained DNN. The range of input parameters for the DNN to predict graphite nodularity are shown for one of the castings in Figure 7. Each parameter was right in the range of min and max according to Table 2. Figure 7 shows the calculated influence of individual parameters on nodularity using the DNN respectively, while changing individual parameters separately within the according parameter range (min. – max. in Table 2). The empty bars are values whose range is smaller than the RMSE (RMSE = 2.2) of the prepared DNN, and the shaded bars are between the single RMSE and the double RMSE. The cross shaded bars visualise the significant influence parameters, which are larger than the double RMSE and are of interest for the investigated case.

In the test sample in Figure 7 magnesium and oxygen activity have a high influence on the graphite nodularity according to the extended bars. The model also calculates a very strong dependence of the nodularity, e.g., on the solidification time and also on overheating temperature as well as the casting temperature.
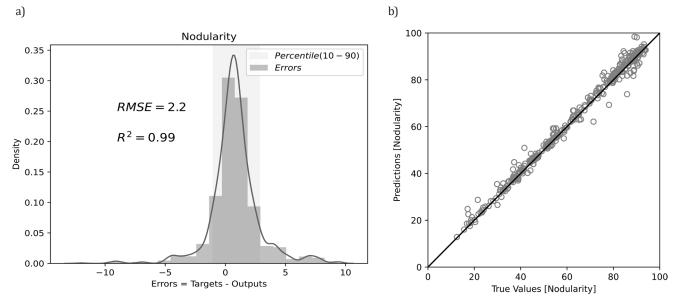
a)

b)

Fig. 6. Accuracy of the Deep Neural Network to predict nodularity of graphite spheres: a) histogram to visualize the distribution of prediction errors; b) predicted values plotted against true nodularity
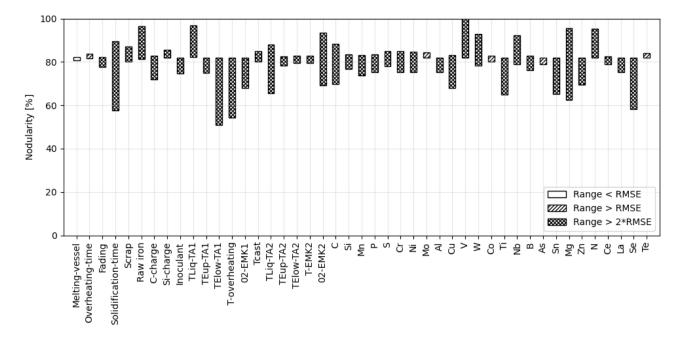

Fig. 7. Illustration of the influence of the different parameters on the nodularity calculated with the previously trained deep neural network for one casting. Note the input parameters are parameter of one specific test sample

Figure 8 shows the influence of fading time, solidification time, cast temperature, oxygen activity measurement and Mg-content on nodularity graphically for the given test sample.

As a first validation, the calculated dependence of nodularity on the solidification time can be seen of improved nodularity in real casting. Figure 9 shows metallographic micrographs after a solidification time of a) 50 s, b) 90 s, c) 180 and d) 360 s as a comparison to the calculated values in Figure 8b. The nodularity increases with the increasing solidification time. It is obvious that the nodularity in Figure 9a is lower (solidification time is 50 s), than in Figure 9d (solidification time is 360 s). However, the models created are based on correlations alone and do not necessarily indicate physical or metallurgical causalities.
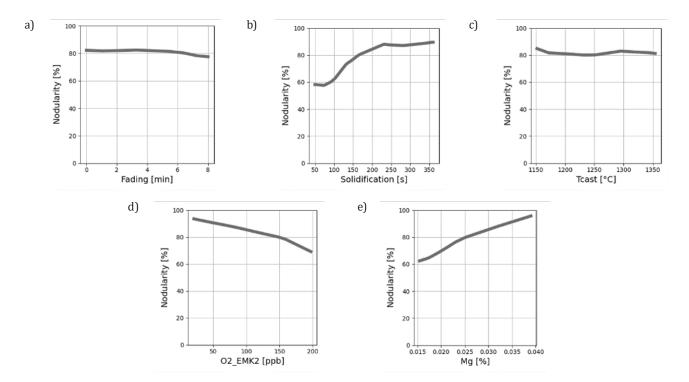

Fig. 8. Correlation of: a) fading; b) solidification time; c) cast temperature; d) oxygen activity; e) magnesium content on nodularity
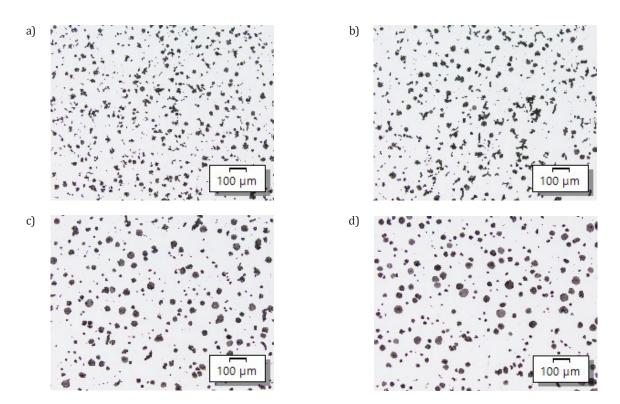

Fig. 9. Metallographic images with different solidification times and similar parameters tocalculated values in Figure 8b. Solidification time is: a) 50 s; b) 90 s; c) 180 s; d) 360 s

4. CONCLUSION

In this work, a previously reported AI to describe nodularity was expanded to 45 relevant input parameters (listed in Table 2) and described by linear regression, Gaussian process regression, regression trees, boosted trees, support vector machines, shallow neural networks and deep neural networks. The highest accuracy was achieved by DNN with 5 hidden layers, 40 nodes at each layer with ReLU-activation functions.

According to the model based on DNN, for a particular sample graphite nodularity is strongly dependent on solidification time, overheating temperature as well as on the Mg content and thus also on oxygen activity, as these are interdependent. While these influencing factors are well known to the metallurgist [50], other industrial relevant parameters which determine the boundary conditions can be analysed using this model.

The uniqueness of this paper is that it describes regression methods to determine nodularity using 45 input parameters, which include information about the basic composition of the charge, overheating time, types of melting vessel, type of inoculant, fading and solidification time. Furthermore, the database is well documented by several measurements including thermal analysis, oxygen activity measurements and chemical analysis.

A limitation of the current research is that the investigations carried out so far have only been performed in the foundry of the Austrian Foundry Research Institute with batch weights between 50 and 120 kg. The limited range of tests only allows the nodularity to be predicted with high accuracy for castings in this foundry. In the near future, further test casts will be conducted with other foundries to obtain a higher range of parameters and to assess the influence of larger batches. Further regression analysis and hypothesis testing will be performed to study the individual parameters and to analyse the nodule count.

REFERENCES

[1] Bendsoe M.P. & Sigmund O. (2003). *Topology optimization: theory, methods, and applications*. Springer Science & Business Media, Heidelberg.

[2] Eschenauer H., Koski J. & Osyczka A. (Eds.) (2012). *Multicriteria design optimization: procedures and applications.* Springer Science & Business Media, Heidelberg.

[3] Cziegler A. (2020). *Optimization of Green Sand Thermophysical Properties for Simulation of Cast Iron Processes.* International MAGMA User Meeting 2020. 6–8 October 2020.

[4] Bauer W. (2017) *Graphitmorphologie: Entstehung, Einflussgrößen, Kontrolle.* Österr.eichisches Gießerei-Institut, Leoben.

[5] Benedetti M., Fontanari V. & Lusuardi D. (2019). Effect of graphite morphology on the fatigue and fracture resistance of ferritic ductile cast iron. *Engineering Fracture Mechanics*, 206, 427–441. Doi: https://doi.org/10.1016/j.engfracmech.2018.12.019.

[6] Liu Y., Xing J., Li Y., Wang S. & Tao D. (2018). Tomographical study of the effect of graphite on properties of cast iron. *Steel research international*, 89(8), 1800086. Doi: https://doi.org/10.1002/srin.201800086.

[7] Herfurth K. (2016). Wachstum der Graphitkristalle im Gusseisen unter besonderer Berücksichtigung der Grenzflächenspannung. *Gießerei-Praxis*, 6, 254–260.

[8] Stefanescu D., Alonso G., Larrañaga P., De la Fuente E. & Suarez R. (2018). A comparative study of graphite growth in cast iron and in analogous systems. *International Journal of Metalcasting*, 12(4), 722–752. Doi: https://doi.org/10.1007/s40962-017-0204-1.

[9] Thankachan T., Prakash K.S. & Jothi S. (2021). Artificial neural network modeling to evaluate and predict the mechanical strength of duplex stainless steel during casting. *Sādhanā*, 46(4), 197. Doi: https://doi.org/10.1007/s12046-021-01742-w.

[10] Agrawal A. & Choudhary A. (2016). Perspective: Materials informatics and big data: Realization of the "fourth paradigm" of science in materials science. *APL Materials*, 4(5), 053208. Doi: https://doi.org/10.1063/1.4946894.

[11] Bührig-Polaczek A., Rudack M., Salentin F.G., Sandt M., Hartmann D., Brait M., Koppensteiner E. & Schumacher P. (2021). Digitalisierung in der Gießereitechnik. Teil 1: Herausforderungen und Chancen. *Gießerei*, 6/21, 37–43.

[12] Çöl M., Ertunç H.M. & Yılmaz M. (2007). An artificial neural network model for toughness properties in microalloyed steel in consideration of industrial production conditions. *Materials & Design*, 28(2), 488–495. Doi: https://doi.org/10.1016/j.matdes.2005.09.001.

[13] Bashar A. (2019). Survey on evolving deep learning neural network architectures. *Journal of Artificial Intelligence and Capsule Networks*, 1(2), 73–82. Doi: https://doi.org/10.36548/jaicn.2019.2.003.

[14] LeCun Y., Bengio Y. & Hinton G. (2015). Deep learning. *Nature,* 521(7553), 436–444. Doi: https://doi.org/10.1038/nature14539.

[15] Glavas Z., Lisjak D. & Unkic F. (2007). The application of artificial neural network in the prediction of the as-cast impact toughness of spheroidal graphite cast iron. *Kovové Materiály*, 45(1), 41–49.

[16] Liliac M.M., Yamamoto K. & Ogi K. (2001). A neural network prediction for the graphite nucleation and morphology in modified cast irons. *International Journal of Cast Metals Research*, 14(3), 157–164.

[17] Amini S. & Abbaschian R. (2013). Nucleation and growth kinetics of graphene layers from a molten phase. *Carbon*, 51, 110–123. Doi: https://doi.org/10.1016/j.carbon.2012.08.019.

[18] Stefanescu D.M., Alonso G., Larrañaga P., De la Fuente E. & Suarez R. (2016). On the crystallization of graphite from liquid iron-carbon-silicon melts. *Acta Materialia*, 107, 102–126. Doi: https://doi.org/10.1016/j.actamat.2016.01.047.

[19] Ghahremani nahr J., Nozari H. & Sadeghi M.E. (2021). Artificial intelligence and Machine Learning for Real-world problems (A survey). *International Journal of Innovation in Engineering*, 1(3), 38–47.

[20] Fragassa C., Babic M., Bergmann C.P. & Minak G. (2019). Predicting the tensile behaviour of cast alloys by a pattern recognition analysis on experimental data. *Metals*, 9(5), 557. Doi: https://doi.org/10.3390/met9050557.

[21] Perzyk M. & Kochański A.W. (2001). Prediction of ductile cast iron quality by artificial neural networks. *Journal of Materials Processing Technology*, 109(3), 305–307.

[22] Swain S.K. (2008). Effect of Chemistry and Processing Variables on the Mechanical Properties of Thin Wall Ductile Iron Castings. National Institute of Technology: Rourkela [Doctoral dissertation].

[23] Lim K.Y.H., Zheng P. & Chen C.H. (2020). A state-of-the-art survey of Digital Twin: techniques, engineering product lifecycle management and business innovation perspectives. *Journal of Intelligent Manufacturing*, 31(6), 1313–1337. Doi: https://doi.org/10.1007/s10845-019-01512-w.

[24] Nielsen F. (1993). Gieß- und Anschnitttechnik Grundlagen und Anwendung einer Methode. Giesserei-Verlag GmbH, Dusseldorf.

[25] AMGuss. Professionelle Analyse der Mikrostruktur von Gusseisen. GFaI Gesellschaft zur Förderung angewandter Informatik, 30 November 2021, Retrieved from https://www.gfai.de/entwicklungen/bildverarbeitung/amguss [accessed 7.10.2021].

[26] GFaI. Gesellschaft zur Förderung Angewandter Informatik, Am Guss, Version 1.1.60.

[27] ISO/DIS 945-4:2016(E) working draft (2016). Microstructure of cast irons – Part 4: test method for determining nodularity in spheroidal graphite cast irons.

[28] Friess J., Sonntag U., Steller I. & Bührig-Polaczek A. (2020). From Individual Graphite Assignment to an Improved Digital Image Analysis of Ductile Iron. *International Journal of Metalcasting*, 14(4), 1090–1104.

[29] Mampaey F., Habets D., Plessers J. & Seutens F. (2008). The use of oxygen activity measurement to determine optimal properties of ductile iron during production. *Giessereiforschung*, 60(1), 2–19.

[30] Biessmann F., Rukat T., Schmidt P., Naidu P., Schelter S., Taptunov A., Lange D. & Salinas D. (2019). DataWig: Missing Value Imputation for Tables. *Journal of Machine Learning Research*, 20 (175), 1–6.

[31] Dara S. & Tumma P. (2018). Feature Extraction by Using Deep Learning: A Survey. In: 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), 29–31 March 2018, RVS Technical Campus, Coimbatore, India, 1795–1801.

[32] Yogatama D. & Mann G. (2014). Efficient transfer learning method for automatic hyperparameter tuning. In: Artificial Intelligence and Statistics. Proceedings of the Seventeenth International Conference on Artificial Intelligence and Statistics, PMLR 33, 1077–1085.

[33] Weisberg S. (2005). *Applied linear regression,* Vol. 528, John Wiley & Sons, Hoboken.

[34] Schulz E., Speekenbrink M. & Krause A. (2018). A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions. *Journal of Mathematical Psychology*, 85, 1–16. Doi: https://doi.org/10.1016/j.jmp.2018.03.001.

[35] Li Y. (2006). Predicting materials properties and behavior using classification and regression trees. *Materials Science and Engineering: A*, 433(1–2), 261–268.

[36] Ponomareva N., Radpour S., Hendry G., Haykal S., Colthurst T., Mitrichev P. & Grushetsky A. (2017). TF boosted trees: A scalable tensorflow based framework for gradient boosting. In: Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Springer, Cham, 423–427.

[37] Ding S., Qi B. & Tan H. (2011). An overview on theory and algorithm of support vector machines. *Journal of University of Electronic Science and Technology of China*, 40(1), 2–10. Doi: dx.chinadoi.cn/10.3969/j.issn.1001-0548.2011.01.001.

[38] Bhadeshia H., Dimitriu R., Forsik S., Pak J. & Ryu J. (2009). Performance of neural networks in materials science. *Materials Science and Technology*, 25(4), 504–510. Doi: https://doi.org/10.1179/174328408X311053.

[39] Goodfellow I., Bengio Y. & Courville A. (2016). Deep learning. MIT Press.

[40] Schmidhuber J. (2015). Deep learning in neural networks: An overview. *Neural networks*, 61, 85–117. Doi: https://doi.org/10.1016/j.neunet.2014.09.003.

[41] Kim P. (2017). Convolutional neural network. In MATLAB deep learning. Apress, Berkeley, CA, 121–147.

[42] Gulli A. & Pal S. (2017) Deep learning with Keras. Packt Publishing Ltd, Birmingham.

[43] Dillon J.V., Langmore I., Tran D., Brevdo E., Vasudevan S., Moore D., Patton B., Alemi A., Hoffman M. & Saurous R.A. (2017). TensorFlow distributions. Retrieved from: https://arxiv.org/pdf/1711.10604.pdf [accessed 7.10.2021].

[44] Chicco D., Warrens M.J. & Jurman G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, 7, e623. Doi: https://doi.org/10.7717/peerj-cs.623.

[45] Singh K. & Xie M. (2008). Bootstrap: a statistical method. Rutgers, The State University of New Jersey, USA. Retrieved from: http://www.stat.rutgers.edu/home/mxie/RCPapers/bootstrap.pdf, 1–14 [unpublished manuscript]

[46] Wang F., Sahana M., Pahlevanzadeh B., Pal S.C., Shit P.K., Piran M.J., Janizadeh S., Bandh S.S. & Mosavi A. (2021). Applying different resampling strategies in machine learning models to predict head-cut gully erosion susceptibility. *Alexandria Engineering Journal*, 60(6), 5813–5829. Doi: https://doi.org/10.1016/j.aej.2021.04.026.

[47] Eck D.J. (2018). Bootstrapping for multivariate linear regression models. *Statistics & Probability Letters*, 134, 141–149. Doi: https://doi.org/10.1016/j.spl.2017.11.001.

[48] Su Y. & Mwanakatwe P.K. (2021). The comparison study of the model selection criteria on the Tobit regression model based on the bootstrap sample augmentation mechanisms. *Journal of Statistical Computation and Simulation,* 91(7), 1415–1440. Doi: https://doi.org/10.1080/00949655.2020.1856848.

[49] Lundberg S.M. & Lee S.-I. (2017). A unified approach to interpreting model predictions. In: Proceedings of the 31st International Conference on Neural Information Processing Systems, 4768–4777.

[50] Lacaze J. (2017). Trace elements and graphite shape degeneracy in nodular graphite cast irons. *International Journal of Metalcasting*, 11(1), 44–51. Doi: https://doi.org/10.1007/s40962-016-0115-6.