

# Data-Driven Modeling of Mechanical Properties of Cast Iron Using Fuzzy Logic

He TAN <sup>a,1</sup>, Vladimir TARASOV <sup>a</sup> and Vasileios FOURLAKIDIS <sup>b</sup> and Attila DIOSZEGI <sup>b</sup>

<sup>a</sup> *Department of Computer Science and Informatics, School of Engineering, Jönköping University, Sweden*

<sup>b</sup> *Department of Materials & Manufacturing, School of Engineering, Jönköping University, Sweden*

**Abstract.** For many industries, an understanding of the fatigue behavior of cast iron is important but this topic is still under extensive research in materials science. This paper offers fuzzy logic as a data-driven approach to address the challenge of predicting casting performance. However, data scarcity is an issue when applying a data-driven approach in this field; the presented study tackled this problem. Four fuzzy logic systems were constructed and compared in the study, two based solely upon experimental data and the others combining the same experimental data with data drawn from relevant literature. The study showed that the latter demonstrated a higher accuracy for the prediction of the ultimate tensile strength for cast iron.

**Keywords.** Fuzzy logic system, Data scarcity, Fatigue related properties prediction, Cast iron components

## 1. Introduction

Casting is essential to many industries. The ability to predict the performance of a cast component helps to produce near-net shape cast components and provides greater freedom for manufacturing design. There are two categories of approach to casting performance prediction: model-based and data-driven. The model-based approach attempts to develop mechanism models to explain observed phenomena, e.g. the Griffith and Hall-Petch equations, and their adopted versions. However, the mechanical properties of a cast component can be affected by many parameters. Melt treatment and the casting process both display stochastic behavior. The assumptions and simplifications of the mechanism models limit the performance of this approach.

Data-driven techniques utilize experimental data for prediction. Such techniques have been proposed to deal with uncertainty and imprecision when predicting the mechanical properties of cast components or when relating the mechanical properties to the parameters of the casting process, and include: regression methods, machine learning methods, neural networks, and fuzzy logic. In our previous work [1] two fuzzy logic systems (FLS) were built to predict yield strength (YS) for two Al-Si-Mg alloys. The re-

---

<sup>1</sup>Corresponding Author: He Tan, E-mail: he.tan@ju.se.

sults showed high accuracy of prediction. When compared with several machine learning methods, the prediction accuracies of the two constructed FLS were at the same level as or slightly better than the machine learning methods. The advantage of the FLS approach over machine learning methods is that the former provides a highly interpretable model. This is important for industrial applications. The interpretability of fuzzy logic method was appreciated by the industry partners in the research project where the work in [1] where the work on Al-Si-Mg alloy yield strength was conducted.

The current work investigated the use of fuzzy logic to address the problem of predicting the ultimate tensile strength (UTS) of a particular form of cast iron morphology – Lamellar Graphite Iron (LGI). LGI is the most used casting material nowadays and is especially important in the automotive industry. For example, cylinder heads and cylinder blocks for diesel truck engines are mostly manufactured with LGI. The UTS is one of the main factors that affect the performance and fuel consumption of these engines.

One challenge posed by the application of data-driven approach in material science is data scarcity. This is caused by an unusually high data acquisition cost in the field compared to other domains where any data-driven approach has proven useful. The current work examined using data published in literature to deal with this challenge. The data from literature together with the current experimental data provided a larger set of training data to construct FLS. The results showed that the constructed FLS provided a high accuracy of prediction, which was not be achieved by the FLS constructed only from the experimental data.

The paper is structured as follows. Related work is given in Section 2. Section 3 presents the data collected from experimental work and the data derived from literature. The method to design and build FLS and the evaluation results are presented in Section 4. The conclusions are drawn in Section 5.

## **2. Related Work**

Although the attempts to apply fuzzy logic in the casting industry are found to be limited, studies have shown that fuzzy logic models of mechanical properties of die casts and metallic materials can achieve desired performance. In [2] fuzzy logic was applied to estimate surface roughness of die casting alloys. A FLS was constructed to map the presence of additives and machining parameters to surface roughness. The percentage error of the FLS is 5.4%. The authors of [3] showed that fuzzy expert system was efficient to optimize the parameters in the milling process and predict the performance measures. The fuzzy logic model built in [4] predicted surface hardness of aluminium alloy with TiN (titanium nitride). The percentage error and accuracy of the model are 6.09 and 96.142%, respectively. In our previous work [1], two FLS were constructed to predict YS for Al-Si-Mg alloys. The prediction errors of the models are 3.53% for the first and 3.19% for the second alloy. In the presented work we applied fuzzy logic to predict the UTS for LGI. The purpose is to evaluate the applicability of fuzzy logic as a data-driven approach to address the problem of understanding fatigue behavior of cast iron. The latter is a challenge in the field of materials science.

Data-driven approach is a new paradigm in materials science. It is extremely expensive to generate data in the field. The community also lacks a unified approach to reusing and sharing materials data [5]. To the best of the authors' knowledge, none of the previ-

ously published works utilizes published data sets for modeling of mechanical properties of cast components. The current paper presented a study to utilize data set published in literature together with experimental data during construction of fuzzy logic models to address the data scarcity.

### 3. Data Collection

This section presents the data sets that were used to design, built and evaluate the FLS in Sect. 4. The first data set was generated from the experimental work. The second data set is a combination of the experimental data and the data found in the published literature.

#### 3.1. Experimental Data

This section describes the experimental work and the data collected during the experiment. For more details the reader is referred to [6]. Four hypoeutectic lamellar graphite iron heats with varying chemical composition were produced. The four samples were labeled individually with C13-4, C13-B, C13-C and C13-D. The carbon content of the samples is 3.62, 3.34, 3.05 and 2.8, respectively. The experimental mould contained three cylindrical cavities. Each was surrounded by a different material, i.e. steel, sand or isolation, to provide the different cooling rates of 3.5°C/s (high), 0.8 °C/s (medium) and 0.2 °C/s (low), respectively.

The microstructure was examined using an optical microscope. Secondary dendrite arm spacing (SDAS) was measured by a linear intercept method. The eutectic cell (EC) size was calculated as the average diameter of the 5–10 biggest eutectic cells. The obtained microstructure is fully pearlitic (>95% pearlite) for all the cooling conditions.

Uniaxial tensile testing was performed on one sample machined from each cylindrical casting. The tests were conducted on a Lloyd EZ50 testing machine, at a strain rate of 0.035 mm/s and at room temperature. Tensile strength was measured using dog-bone shaped bars according to VOLVO standard 5C25.

Table 1 presents the data collected during the experiment. In total 12 data points were generated from the experiment. The value of ECEL was not collected during the experiment for the sample C13-D when cooling rate is high.

#### 3.2. Literature data

A number of papers are published concerning the prediction of UTS and they contain relevant experimental data (e.g. [6,7,8]). The investigation data from [7] together with the current experiment data were employed in this work (see Table 2). The experimental procedure from [7] involves the production of 11 fully LGI with varying N content and inoculation addition aimed to provide material with difference in microstructure and mechanical properties. Contrary to the current experiment, the work in [7] utilized similar carbon content but constant cooling rates for all the castings.

### 4. Fuzzy Modeling of Ultimate Tensile Strength

In the presented work, the FLS were constructed as multiple-input-single-output mamdani-type models. Each FLS has two inputs. The inputs are considered to be the

**Table 1.** Data collected during the experimental work.

Sample	Cooling rate	$\sigma_{UTS}$ [MPa]	ECEL [ $\mu\text{m}$ ]	SDAS [ $\mu\text{m}$ ]
C13-A	High	363	423	25.0
	Medium	195	1153	55.9
	low	154	1977	92.2
C13-B	High	368	330	25.0
	Medium	254	1072	63.9
	low	211	1583	87.0
C13-C	High	440	263	30.0
	Medium	286	965	56.4
	low	250	1454	95.0
C13-D	High	447	-	29.0
	Medium	337	899	63.2
	low	289	1380	101.0

**Table 2.** The combination of the data collected from [7] (see the first 11 rows) and the experimental data when the cooling rate is  $0.2^\circ\text{C/s}$ , i.e. Low (see the last 4 rows)

Sample	Carbon content	$\sigma_{UTS}$ [MPa]	ECEL [ $\mu\text{m}$ ]	SDAS [ $\mu\text{m}$ ]
W1	3.36	335	185	30
W2	3.41	288	332	28
W3	3.36	318	193	28
W4	3.41	323	181	27
W5	3.42	323	285	26
W6	3.36	345	224	26
W7	3.4	358	176	28
W8	3.28	347	307	28
W9	3.3	373	165	30
W10	3.3	393	140	27
W11	3.36	357	179	27
C13-A	3.62	154	1977	92.2
C13-B	3.34	211	1583	87.0
C13-C	3.05	250	1454	95.0
C13-D	2.8	289	1380	101.0

most important input variables to predict the output by the experts in the field of materials science. The input variables also have high correlation coefficient with the output. The first input is "Carbon content" that represents the percentage of C in the iron cast alloy. The second input concerns microstructure. It is either "SDAS", representing microstructure coarseness, or "ECEL", representing the eutectic cell size. The output is the resulting ultimate tensile strength, i.e. "UTS". The systems were implemented using Scikit-Fuzzy <sup>2</sup>.

<sup>2</sup><https://pythonhosted.org/scikit-fuzzy/>

As described in Sect. 3, two data sets were collected. Each data set was divided into a training data set and a test data set. The training data sets were used to build the FLS, i.e. to define the parameters for membership functions (Sect. 4.1) and to learn fuzzy inference rules (Sect. 4.2). The first data set has 12 data points. 8 data points constituted the training data. The rest were used for test (4 data points when the second input is SDAS, or 3 data points when the second input is ECEL). The second data set contains 15 data points. 11 data points constitutes the training data set, the rest were used for test.

#### 4.1. Linguistic variables and membership functions

The membership functions for both input and output variables are Gaussian functions, because Gaussian function was considered to be the most adequate by the experts in materials science. For each membership function the standard deviation  $\sigma$  and mean  $c$  were determined by manual examination of the training data sets.

The examination were carried out as follows: 1) compiling the values of an input/output into a list, sorting it and determining the range of the variable; 2) partitioning the values in the list into segments to determine the number of memberships functions (linguistic labels) for an input/output variable and the  $\sigma$  and  $c$  for each membership function. The size of each partition was chosen to encompass several values. Finally, *max* was used for the aggregation and *centroid* was the defuzzification method.

#### 4.2. Fuzzy inference rules

A set of two-input one-output inference rules were produced for each fuzzy logic system. The rules were first determined to be "and" rules (i.e., logical conjunction) by a knowledge engineer. Then the rules were learned from the training data using the method adapted from [9]. The method consists of three steps:

1) Generate a rule for each input-output data pair  $(x_1, x_2; y)$  in training data:

IF  $x_1^i$  is  $A$  and  $x_2^i$  is  $B$ , THEN  $y^i$  is  $C$ , where  $x_1$  and  $x_2$  are the two inputs,  $y$  is the output,  $\mu_A(x_1^i)$ ,  $\mu_B(x_2^i)$ , and  $\mu_C(y^i)$  are membership functions with the maximum membership degree for each respective value.

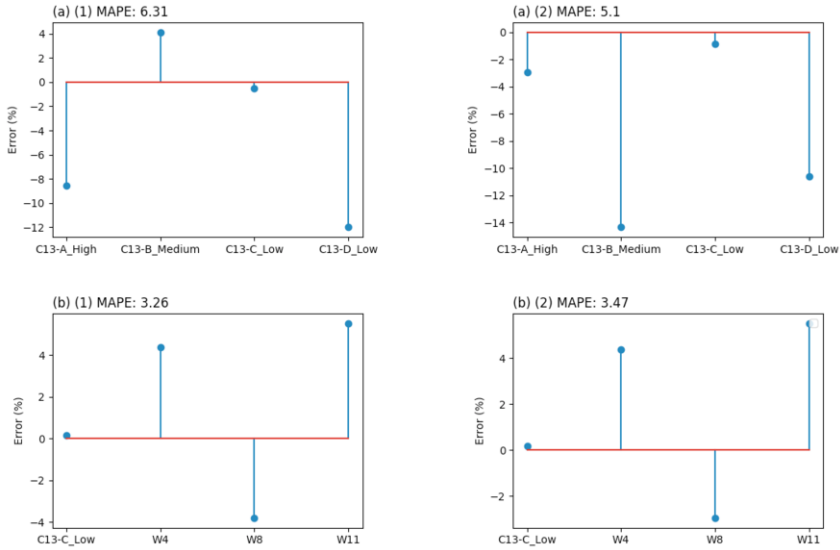
2) Calculate a degree for each rule:  $D(rule^i) = \mu_A(x_1^i) \times \mu_B(x_2^i) \times \mu_C(y^i)^i$

3) Generate the final list of rules. If there is more than one rule with the same antecedent, the rule that has the maximum degree is selected for the final list.

This is a general and straightforward method for learning rules from data. The method was implemented in Python. The performance of the method was evaluated in [10] which showed that the rules generated using the presented method were identical to the rules manually created by a knowledge engineer.

#### 4.3. Evaluation Results

The prediction accuracy of the constructed FLS was measured using Mean Absolute Percentage Error (MAPE) [11] (see Fig. 1). The percentage prediction errors are also depicted in Fig. 1. The results show that the prediction accuracy of the two fuzzy logic systems for the second data set is satisfactory and also at the same level as reported in the related work described in Sect 2. The prediction accuracies of the two FLS are very similar, despite different sets of input variables were used.



**Figure 1.** Percentage prediction errors plotted as stems for the UTS value in each data point: (a) present the results for the first data set and (b) are the result for the second data set. (1) show the results when the input variables are "Carbon content" and "ECEL"; (2) give the results when the input variables are "Carbon content" and "SDAS".

The two FLS for the second data set showed better MAPE than the systems for the first data set. The accuracy of the prediction of the UTS of the alloy samples with the low cooling rate is much better and stable for the FLS for the second data set than for the FLS for the first data set. The prediction accuracy of the two FLS for the first data set is not satisfactory. Moreover, the two FLS performed very differently when the different input variables were applied.

### 5. Conclusion

The presented work contributes a data-driven approach using fuzzy logic to address the lack of understanding of the fatigue behavior of cast irons. In addition, the study investigated the utilization of data published in literature to address data scarcity in the field of materials science when applying a data-driven approach. The study showed that the data derived from the related literature together with experimental data provided a larger set of data to support the construction of a FLS. This FLS constructed using combined data presented a high accuracy of prediction. Compared to the FLS constructed only from the experimental data, the former FLS performed better and was more stable.

For future work, we are considering three lines of activity. Firstly, we would like to apply fractional fuzzy inference systems [12] to the UTS problem addressed in this paper. We expect an FLS of this new type to deliver even better predictive performance. We will also broaden our FLS approach to investigate the mapping of other fatigue-related properties to chemical composition and the various microstructure properties of cast components. Finally, to address the data scarcity in this field, we are considering the establishment of an online infrastructure to collect, host and provide such materials data in support of further research and production-focused investigations.

## References

- [1] Vladimir Tarasov, He Tan, Anders EW Jarfors, and Salem Seifeddine. Fuzzy logic-based modelling of yield strength of as-cast A356 alloy. *Neural Computing and Applications*, pages 1–12, 2019.
- [2] Mohsen Marani Barzani, Erfan Zalnezhad, Ahmed A.D. Sarhan, Saeed Farahany, and Singh Ramesh. Fuzzy logic based model for predicting surface roughness of machined Al-Si-Cu-Fe die casting alloy using different additives-turning. *Measurement*, 61(Supplement C):150–161, 2015.
- [3] Asif Iqbal, Ning He, Liang Li, and Naeem Ullah Dar. A fuzzy expert system for optimizing parameters and predicting performance measures in hard-milling process. *Expert Systems with Applications*, 32(4):1020–1027, 2007.
- [4] E. Zalnezhad, Ahmed A. D. Sarhan, and M. Hamdi. A fuzzy logic based model to predict surface hardness of thin film TiN coating on aerospace AL7075-T6 alloy. *The International Journal of Advanced Manufacturing Technology*, 68(1):415–423, Sep 2013.
- [5] Lauri Himanen, Amber Geurts, Adam Stuart Foster, and Patrick Rinke. Data-driven materials science: Status, challenges, and perspectives. *Advanced Science*, 6(21):1900808, 2019.
- [6] Vasilios Fourlakis and Attila Diószegi. A generic model to predict the ultimate tensile strength in pearlitic lamellar graphite iron. *Materials Science and Engineering: A*, 618:161–167, 2014.
- [7] Fredrik Wilberfors and Ingvar L Svensson. The effect of nitrogen and inoculation on the tensile properties and microstructure of cast iron with lamellar graphite. In *Key Engineering Materials*, volume 457, pages 114–119. Trans Tech Publ, 2011.
- [8] Torsten Sjögren and Henrik Svensson. Study of the eutectoid transformation in grey cast irons and its effect on mechanical properties. In *Key Engineering Materials*, volume 457, pages 157–162. Trans Tech Publ, 2011.
- [9] L-X Wang and Jerry M Mendel. Generating fuzzy rules by learning from examples. *IEEE Transactions on systems, man, and cybernetics*, 22(6):1414–1427, 1992.
- [10] He Tan, Vladimir Tarasov, Anders EW Jarfors, and Salem Seifeddine. Fuzzy logic based modelling of cast component properties. *IFAC-PapersOnLine*, 52(13):1132–1137, 2019.
- [11] Arnaud de Myttenaere, Boris Golden, Bénédicte Le Grand, and Fabrice Rossi. Mean absolute percentage error for regression models. *Neurocomputing*, 192:38–48, 2016. Advances in artificial neural networks, machine learning and computational intelligence.
- [12] M. Mazandarani and X. Li. Fractional fuzzy inference system: The new generation of fuzzy inference systems. *IEEE Access*, 8:126066–126082, 2020.