Modeling the Chemical Composition of Ferritic Stainless Steels with the Use of Artificial Neural Networks

Rafal Honysz

Department of Engineering Materials and Biomaterials, Faculty of Mechanical Engineering, Silesian University of Technology, 44-100 Gliwice, Poland; rafal.honysz@polsl.pl; Tel.: +48-(32)-237-18-81

Abstract: The aim of this paper is an attempt to answer the question of whether, on the basis of the values of the mechanical properties of ferritic stainless steels, it is possible to predict the chemical concentration of carbon and nine of the other most common alloying elements in these steels. The author believes that the relationships between the properties are more complicated and depend on a greater number of factors, such as heat and mechanical treatment conditions, but in this paper, they were not taken into account due to the uniform treatment of the tested steels. The modeling results proved to be very promising and indicate that for some elements, this is possible with high accuracy.

Artificial neural networks with radial basis functions (RBF), multilayer perceptron with one and two hidden layers (MLP) and generalized regression neural networks (GRNN) were used for modeling. In order to minimize the manufacturing cost of products, developed artificial neural networks can be used in industry. They may also simplify the selection of materials if the engineer has to correctly select chemical components and appropriate plastic and/or heat treatments of stainless steel with the necessary mechanical properties.

Keywords: analysis and modeling; numerical techniques; computational material science; artificial neural networks; ferritic stainless steel

1. Introduction

Developments in material engineering have resulted in increased market competition, especially for corrosion-resistant steels. These materials’ properties are strictly dependent on their chemical composition and processing type. It is therefore important that the chemical composition, as well as the appropriate heat and mechanical treatment conditions, should be selected according to the customer’s requirements, in order to obtain the required mechanical properties and relatively low production costs. The classical approach, i.e., the execution of a series of experiments with the development of the required number of samples to determine the characteristics of each of these steel grades, is a breakneck undertaking that requires an extremely large amount of time and financial expenditure. Artificial intelligence techniques, together with experimental data, enable the creation of a model that enables the chemical composition of stain-ferritic steels to be predicted with high precision in a very short time. The main objective of designing such a model is to minimize the costs associated with the material testing of these steels and to improve access to the measurement results more quickly. The use of artificial intelligence allows stainless steel technology to be advanced in many respects, even though only a limited number of definition vectors are available [1–12].

In recent years, the issue of the application of artificial intelligence algorithms in material engineering has been dealt with by many scientists from around the world. Several computer models have been developed that explain the relationships between steel phenomena, their properties, chemical composition, and processing conditions. The models can be implemented in the manufacturing sector in order to minimize the production expenses of goods. The choice of materials can also be simplified if the engineer has...
to correctly choose chemical elements and suitable plastics and/or stainless-steel heat processing with the appropriate mechanical characteristics [13–21].

2. Materials and Methods

Data for the construction of computation models for predicting steel properties were obtained by laboratory testing of certain grades of ferritic stainless steels, following PN-EN 10088-1: 2014. The main criteria for selecting steel grades were carbon concentration from 0.3 to 0.8%, chromium concentration from 10 to 16% and nickel concentration from 0.1 to 2%, together with other alloying elements [1–5,22–25]. Steel was smelted in electric arc furnaces equipped with vacuum arc degassing (VAD) devices. The material was delivered in the form of round rolled rods with a diameter of 150 mm after normalization treatment at the temperature of 660 °C for 180 min. As the heat and plastic treatment of steel were uniform, these values were not included in the training vectors. From metallurgical approvals, chemical element concentration values were read and used as output variables in the process of teaching artificial neural networks:

- Carbon (C);
- Molybdenum (Mn);
- Silicon (Si);
- Phosphorus (P);
- Sulphur (S);
- Chrome (Cr);
- Nickel (Ni);
- Molybdenum (Mo);
- Copper (Cu);
- Aluminum (Al).

After the analysis of the chemical composition of the tested steels, it was found that the concentration of carbon and alloying elements is appropriate for the correct teaching of artificial neural networks. The concentration of the elements that are impurities in steel is very small. This is obvious from the point of view of the quality of steel and the products made from it. Unfortunately, most likely the concentration values of these elements are too small to teach artificial neural networks and in the case of modeling these elements, it would not be possible to obtain satisfactory results.

The results of laboratory tests were used to build a dataset with 3272 training vectors. Input variables were:

- Yield strength (Rp0.2);
- Tensile strength (Rm);
- Relative elongation (A);
- Relative area reduction (Z);
- Impact strength (KcU2);
- Brinell hardness (HB).

The determination of strength properties consisted of carrying out a tensile test for steel samples following [26]. Hardness tests were carried out using the Brinell method following [27].

Values of these properties are input values for respective artificial neural networks. The ranges of selected input variables are shown in Table 1.

<table>
<thead>
<tr>
<th>Range</th>
<th>Mechanical Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rp0.2 (MPa)</td>
</tr>
<tr>
<td>minimum</td>
<td>208</td>
</tr>
<tr>
<td>maximum</td>
<td>920</td>
</tr>
</tbody>
</table>

Table 1. The range of selected input variable values.
Material tests were conducted in such a way as to obtain an even distribution of values in the range of variability of the given input value without excessive data clusters or empty spaces. Data uniformity was confirmed using the histogram tool. These vectors were randomly divided into three sets. A training set with 1635 vectors and a validation set with 818 vectors was used in network learning processes. The remaining vectors were included in the test file and were used to check the correctness of the network operation. Before the learning process, the input values of all training vectors were normalized. The process of assigning cases to individual sets was repeated many times. After each new draw, the process of teaching artificial neural networks was repeated several times to obtain the best regression statistics. Research for the best artificial neural network for regression issues was narrowed to structures such as:
- Radial basis functions (RBF);
- General regression neural network (GRNN);
- Multi-layer perceptron (MLP).

Radial basis functions (RBF) artificial neural networks use radial basis functions as activation functions. The output of the network is a linear combination of radial basis functions of the inputs and neuron parameters. These are commonly used types of artificial neural networks for function approximation problems. Radial basis function networks are distinguished from other neural networks due to their universal approximation and faster learning speed. An RBF network typically has only one hidden layer containing radial neurons, each of which models a Gaussian response surface. Due to the strongly non-linear nature of these functions, usually one hidden layer is enough to model functions of any shape. However, the condition for the creation of an effective model of any function by the RBF network is ensuring a sufficient number of radial neurons in the structure of the network. If there are enough of them, an appropriate radial neuron can be attached to every important detail of the modeled function, which guarantees that the obtained solution will reproduce the given function with completely satisfactory fidelity.

Multi-layer perceptron (MLP) is the most popular type of artificial neural network. This type of network usually consists of one input layer, several hidden layers and one output layer. Hidden layers usually consist of McCulloch–Pitts neurons. It is defined by its weights and threshold value, which together give the equation of a specific line and determine the rate of change of the function value along with its distance from the designated line. Transfer functions in hidden and output layers are often hyperbolic. Training MLP networks is possible thanks to the method of backpropagation of errors. MLPs are designed to approximate any continuous function and can solve problems that are not linearly separable. The major use cases of MLP are pattern classification, recognition, prediction and approximation.

General regression neural networks (GRNN) are networks that combine the advantages of a radial basis function network and a multi-layer perceptron network. In the radial layer, which is the equivalent of the first hidden layer, radial neurons are used to group the input data. This layer may consist of a very large number of neurons, which corresponds to detecting a large number of data clusters in the input data set. The second layer consists of only two summing neurons and is called the regression layer. The output neuron performs only one action, which produces the quotient of the scores of both summation neurons. It can be shown that the GRNN provides the best estimate of the required output value in regression networks.

The appropriate selection of the network structure is one of the most important tasks necessary to build an optimal ANN model. While the number of input and output neurons is determined by the number of input and output variables, the selection of the number of hidden layers and the number of neurons in these layers is an extremely complicated task. There are no universal criteria for the selection of ANN structure [28–41].

The training process was conducted with use of these algorithms:
- Error backpropagation;
- Conjugate gradient;
Quasi-Newton;
Levenberg–Marquardt;
Fast propagation;
Delta-bar-delta.

To verify the neural network usability for the prediction and modeling purposes, the following regression statistics were used:
• Mean absolute error (MAE);
• Mean absolute percentage error (MAPE);
• Pearson R correlation [28–41].

The mean absolute error (MAE) is defined as the difference between the measured value and the value computed at the output for the output variable (1):

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} (|x_i^z - x_i^o|)$$

where:
$n$—size of the set
$x_i^z$—$i$-th measured value
$x_i^o$—$i$-th computed value

The mean absolute percentage error (MAPE) is defined as the difference between the measured and computed absolute value divided by the measured value and multiplied by 100% (2):

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{|x_i^z - x_i^o|}{x_i^z} \times 100\% \right)$$

Correlation is determined by the standard Pearson correlation coefficient $R$ for the measured value and the value obtained at the output (3):

$$R = \frac{n(\sum_{i=1}^{n} x_i^z x_i^o) - (\sum_{i=1}^{n} x_i^z)(\sum_{i=1}^{n} x_i^o)}{\sqrt{(n \sum_{i=1}^{n} x_i^z)^2 - (\sum_{i=1}^{n} x_i^z)^2} \sqrt{(n \sum_{i=1}^{n} x_i^o)^2 - (\sum_{i=1}^{n} x_i^o)^2}}$$

All calculations were made in the Statistica 13 package developed by Statsoft [42] on desktop computer with an i5-3450 processor with 8 GB ram.

3. Results

Table 2 contains architecture and regression statistics for the validation set, respectively, for the best RBF, GRNN and MLP networks developed for investigating stainless steels. An automatic network designer was used to estimate the number of neurons in hidden layers for artificial neural networks of the RBF and MLP type. In the case of GRNN, the number of neurons in the radial layer is defined by the number of training vectors. Multi-layer perceptron architecture is described by three or four values, which are the number of input neurons, the number of neurons in one or two hidden layers and single output neuron. For example, the MLP network architecture used for carbon concentration prediction is 6-7-1. This means 6 neurons in the input layer, 7 neurons in one hidden layer and 1 neuron in the output layer. The same network for manganese concentration has the architecture 5-15-5-1, which means 5 neurons in the input layer, 5 neurons in the first hidden layer, 5 neurons in the second hidden layer and 1 neuron in the output layer. In the case of GRNN, the number of neurons in the first layer, called the regression layer, is always equal to the number of learning points. The second layer consists of only two summing neurons. This is why all developed GRNNs developed on the same data set differ only in the number of input neurons.
Table 2. Architecture and regression parameters of developed artificial neural networks (testing set).

<table>
<thead>
<tr>
<th>Chemical Element</th>
<th>RBF Network</th>
<th>MLP Network</th>
<th>GRNN Network</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AR</td>
<td>MAE</td>
<td>MAPE</td>
</tr>
<tr>
<td>C</td>
<td>5-33-1</td>
<td>0.50</td>
<td>23.3%</td>
</tr>
<tr>
<td>Mn</td>
<td>6-54-1</td>
<td>0.11</td>
<td>14.7%</td>
</tr>
<tr>
<td>Si</td>
<td>5-24-1</td>
<td>0.03</td>
<td>11.5%</td>
</tr>
<tr>
<td>Cr</td>
<td>5-12-1</td>
<td>0.78</td>
<td>2.7%</td>
</tr>
<tr>
<td>Ni</td>
<td>6-30-1</td>
<td>0.33</td>
<td>22.9%</td>
</tr>
<tr>
<td>Mo</td>
<td>6-28-1</td>
<td>0.33</td>
<td>39.0%</td>
</tr>
<tr>
<td>Cu</td>
<td>6-57-1</td>
<td>0.09</td>
<td>20.8%</td>
</tr>
<tr>
<td>Al</td>
<td>3-18-1</td>
<td>0.05</td>
<td>36.6%</td>
</tr>
<tr>
<td>P</td>
<td>5-9-1</td>
<td>0.04</td>
<td>35.4%</td>
</tr>
<tr>
<td>S</td>
<td>2-10-1</td>
<td>0.06</td>
<td>39.4%</td>
</tr>
</tbody>
</table>

Where: AR—neural network architecture; MAE—mean absolute error; MAPE—mean absolute percentage error; R—Pearson R correlation.

Figure 1 introduces a comparison of the testing set Pearson R correlation of the best artificial neural network of all types, red for RBF, green for MLP and red for GRNN. Greater values are better. Figure 2 introduces a comparison of testing the set mean absolute error of the best artificial neural network of all types; colors are the same, smaller values are better. Figure 3 introduces a comparison of testing the set mean absolute percentage error of the best artificial neural network of all types. Mean absolute percentage error is more readable than mean absolute error because it shows what percentage of the predicted variable is in the prediction error. Again, colors are the same, and smaller values are better.

![Figure 1](image-url)

**Figure 1.** Comparison of Pearson correlation for the best achieved artificial neural networks (testing set), built for examine steels.
The Pearson R correlation graphs were developed to demonstrate the prediction efficiency in a graphical way. The distinction between the values computed using the artificial neural network and those experimentally tested in the actual laboratory is provided. In all three subsets, the distribution of mechanical property vectors for each of the approximate steel vectors is very similar, confirming the correctness of the learning processes of the
networks. Major variations between groups in the distribution of vectors would suggest the probability of errors and, thus, a network of poor quality. Sample graphs of Person R correlation for a testing subset are shown in Figure 4.

![Figure 4](image.png)

Figure 4. Examples of comparative graphs for (a) carbon concentration, (b) chromium concentration and (c) phosphorus concentration, calculated with use of the artificial neural networks and determined experimentally.

4. Discussion

The greatest efficiency in modeling the chemical composition of ferritic stainless steels is shown by general regression neural networks (GRNN). For nine out of ten elements, they have the best regression statistics. The best results were achieved by modeling the concentration of carbon, manganese, chromium, nickel and molybdenum. For these elements, the Pearson correlation exceeded the level of 0.9. Moderately good results were obtained for silicon and copper, where the correlation values are in the range from 0.8 to 0.9. None of the developed networks was able to model the concentrations of aluminum, phosphorus or sulfur.

In the case of chromium and molybdenum, we see a clear advantage of the GRNN. The greatest efficiency in modeling the chemical composition of ferritic stainless steels is shown by general regression neural networks (GRNN). For nine out of ten elements, they have the best regression statistics. The best results were achieved by modeling the concentration of carbon, manganese, chromium, nickel and molybdenum. For these elements, the Pearson correlation exceeded the level of 0.9. Moderately good results were obtained for silicon and copper, where the correlation values are in the range from 0.8 to 0.9. None of the developed networks was able to model the concentrations of aluminum, phosphorus or sulfur.

In the case of carbon, when comparing the Pearson R correlation, it can be seen that all networks were equally good at modeling. The values for the RBF and MLP networks, which are 0.92, differ only slightly from the GRNN networks with a correlation of 0.95. The difference in the mean absolute error is only 0.04 percent of the concentration. The comparisons of the values measured in the laboratory and those obtained computationally, shown in Figure 4a, also show a very good concentration of points. The RBF network rejected one input field which was yield strength (Rp0.2).

In the case of carbon, when comparing the Pearson R correlation, it can be seen that all networks were equally good at modeling. The values for the RBF and MLP networks, which are 0.92, differ only slightly from the GRNN networks with a correlation of 0.95. The difference in the mean absolute error is only 0.04 percent of the concentration. The comparisons of the values measured in the laboratory and those obtained computationally, shown in Figure 4a, also show a very good concentration of points. The RBF network rejected one input field which was yield strength (Rp0.2).

In the case of manganese, the GRNN also has the best correlation and the smallest error. The parameters of the RBF and MLP networks are generally lower. Although the correlation is lower by 0.06 in the case of the RBF network and by 0.09 in the case of the MLP network, and the errors of these networks are twice as large as the GRNN network, these values can be considered acceptable. This time, it was the MLP network that rejected the relative elongation A5 as a negligible value in the modeling process.

A similar situation occurred in the case of modeling the nickel concentration. Here also the correlation values between the networks are small and range from 0.89 to 0.92, but the errors are generally larger and amount to 0.18 (9.9%) for GRNN and from 0.31 (18.8%) to 0.33 (22.9%) for other networks. This is related to a higher concentration of this element in the tested steels. All three types of networks can be used for modeling; however, due to the smallest error value, GRNN is recommended.

In the case of chromium and molybdenum, we see a clear advantage of the GRNN network over the other two networks. The difference in correlation is huge, at 0.29 compared to the RBF network. For MLP networks, this difference is not much better. The RBF absolute error difference is almost three times the GRNN error of 0.3 (1.4%). For steels with chromium concentrations greater than ten percent, this is a good result. The variable
The best results were achieved for modeling the concentration of carbon, manganese, chromium, nickel and molybdenum. For these elements, the Pearson’s R correlation exceeded the level of 0.9 with a relatively low value of the mean absolute percentage error, which ranges from 1.4% for chromium to 9.9% for nickel. Equally good artificial neural networks, although with slightly lower Pearson’s R correlation values (which were below 0.9 and with mean average percentage errors below 9%) were obtained for silicon and copper.

The regression statistics of these networks indicate that the developed artificial neural networks can be successfully used to predict the chemical composition of ferritic stainless steels.

None of the developed networks were successful in modeling the concentration of aluminum, phosphorus or sulfur. Unfortunately, previous concerns about impurities modeling in ferritic stainless steels have been confirmed. The too low chemical concentration of these elements, and thus too little variability of the values in the training vectors, was the main reason for the failure in properly training any type of artificial neural network and thus in properly modeling these elements.

General regression networks (GRNN) showed the best efficiency in modeling the chemical composition of ferritic stainless steels.

The aim of this paper was an attempt to answer the question of whether, on the basis of the values of the mechanical properties of ferritic stainless steels, it is possible to predict the chemical concentration of carbon and nine other most common alloying elements in these steels. The answer is yes. Since we already know that it is possible, the developed
base model may in the future be expanded with new steel grades with different chemical compositions and processed in different ways. This will certainly expand the possibilities of using this model in the industry.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data sharing is not possible due to privacy issues.

**Conflicts of Interest:** The author declares no conflict of interest.

**References**