Artificial neural network study of the electrical conductivity of mould flux

How to cite:

© 2021 Rongshan Qin

https://creativecommons.org/licenses/by-nc-nd/4.0/

Version: Accepted Manuscript

Link(s) to article on publisher’s website:

oro.open.ac.uk
Artificial neural network study of the electrical conductivity of mould flux

Rongshan Qin*

School of Engineering and Innovation, The Open University, Walton Hall, Milton Keynes MK7 6AA, United Kingdom

Abstract

The electrical conductivity of mould flux with chemical constitution of CaO-SiO$_2$-Al$_2$O$_3$-NaO-K$_2$O-MgO-CaF$_2$-Cr$_2$O$_3$-FeO-MnO has been investigated. The assessed database contains one unitary, five binary, nine ternary, four quaternary, two quinary, two senary and one octonary subsystems. Each constitutional component is in connection to another via some direct or indirect links. A multilayer artificial neural network method was developed and implemented to the database. The work provides a method to calculate the relationships between the composition, temperature and electrical property of the mould flux within the defined parameter ranges. The results have been validated against those experimental data that are not included in the training of the neural networks.

Keywords: Electrical conductivity; Mould flux; Multilayer neural network; Continuous casting
Introduction

Many steelmaking companies are using a mould flux with chemical constitution of CaO-SiO$_2$-Al$_2$O$_3$-NaO-K$_2$O-MgO-CaF$_2$-Cr$_2$O$_3$-FeO-MnO to cast stainless steels in continuous casting mould. The mould slag contains high concentration of fluorine to reduce viscosity and solidification temperature [1]. Cr$_2$O$_3$-FeO-MnO can either pre-exist in mould powder or enter to liquid slag from oxidation of alloying elements in liquid stainless steel at casting mould and tundish [2]. The electrical conductivity of this system has never been assessed systematically although the data for its subsystem are notably rich.

The primary driving force to study the electrical properties of mould powder is for electroslag remelting processing [3-4]. This has, therefore, attracted considerable experimental measurement activities [3-6], theoretical modelling [7-8] and data assessment [9] for several subsystems. The recent environment regulation on carbon neutral steelmaking promotes the electrification of continuous casting. Electric field affects materials segregation [10], viscosity [11], distribution of oxide inclusions [12] and surface roughness in the cast mould [13]. This demands an analytical expression to represent the relationship between the chemical constitution, temperature and electrical conductivity of the system. The aim of the present work was to provide a mean to calculate the constitution and temperature dependent electrical conductivity of the mould flux materials.

This work uses artificial neural network to approach the target. The method relies on available data to extrapolate values in the unknown parameters’ range [14]. Ideally, a relationship between the electrical conductivity, constitution and temperature should be derived from micro-mechanisms [15]. The previous theoretical modelling for the electrical conductivity of mould slag has been based largely on an assumption that electrical conduction is carried out by the moving ionized atoms. Obviously, viscosity affects the mobility of ionized atoms and hence plays an important role in electrical conduction. The basicity of mould slag affects the amount and length of silicate chains. This lays a foundation for the optical basicity model to calculate the electrical conductivity of mould slag [7]. On the other side, the interactions between constitutional components influence the mobility of atoms. This forms a basis to use particle interaction to calculate electrical conductivity [8]. For the oxides which can conduct electricity by electronic means, such as FeO, the above-mentioned theoretical models stray away [2].
The data-based artificial neural network has proved to be an effective solution to provide alternative solution [9].

The artificial neural network is different from data fitting such as the least square method. The former prevents overfitting but the latter seeks best fitting to data. The recent development of artificial intelligent learning enables the method to reduce its dimension according to the conservative laws that are hidden in the data [16-17]. Artificial neural network has potential to indicate some physics natures buried in the big data. This work intends to provide a method to calculate the constitution and temperature dependent electrical conductivity of the mould flux.

**Artificial neural network**

A supervised multilayer artificial neural network with back propagation learning algorithm has been coded for the present purpose. The network has an input layer containing 11 units to record the constitutional compositions and temperature. The output layer has 1 unit to provide the computational result for electrical conductivity. There are T-1 hidden layers each containing numerous units. The architecture of 2-layer network is illustrated schematically in Fig. 1, where the hidden layer has n-units. The mapping function for 2-layer neural network in the present work is defined as

$$x_j = \sum_{i=1}^{10} \omega_{j,i}^{(1)} c_i + \omega_{j,11}^{(1)} T + b_j^{(1)}$$  \hspace{1cm} (1.1)

$$u_j = 1/(1 + e^{-x_j})$$  \hspace{1cm} (1.2)

$$y = \sum_{j=1}^{n} \omega_{1,j}^{(2)} u_j + b_1^{(2)}$$  \hspace{1cm} (1.3)

$$\sigma = 1/(1 + e^{-y})$$  \hspace{1cm} (1.4)

The mapping function for >2 layers neural network can be manipulated in the same way. A code package to calculate up to 4-layer neural network has been developed by the author. $\omega_{j,i}^{(k)}$ is the weight factor between $j^{th}$ unit in $k^{th}$ layer and $i^{th}$ unit in $(k-1)^{th}$ layer, where 0$^{th}$ layer is the input layer. $b_j^{(k)}$ is the bias for $j^{th}$ unit in $k^{th}$ layer. $x_j$ is the activation of $j^{th}$ unit in 1$^{st}$ layer $u_j$ is its activation function. $y$ is the activation of output layer. The electrical conductivity $\sigma$ is the activation function of $y$. 
Fig. 1 Schematic diagram illustrates 2-layer neural network.

The activation functions in Eqs. (1.2) and (1.4) are nonlinear and with output values between 0 and 1. Many other neural network models choose other functions such as $f(x) = \tanh(x)$ [14], which has value between -1 and 1, or Gaussian distribution [16]. It is important to normalize the data according to the format of activation functions. In the present model, all the electrical conductivity data are normalized to between 0 and 1. The numerical results are denormalized to get the true value.

Following the standard method in neural network calculation [18], the total difference is defined as

$$E_D = \frac{1}{m} \sum_{l=0}^{m} \left( \sigma_c^{(l)} - \sigma_e^{(l)} \right)^2$$

where $m$ is the total number of sets of training data. $\sigma_e^{(l)}$ and $\sigma_c^{(l)}$ are the $l^{th}$ training value and calculated value, respectively. To prevent neural network optimization from overfitting the noise in training data,
a regularization is defined to minimize the total value of weight factor square as \( E_W = \sum_{i,j,k} (w_{j,i}^{(k)})^2 \).

The overall target function is defined as [18]

\[ E = \frac{1}{2} \beta E_D + \frac{1}{2} \alpha E_W \]  

(3)

where \( \beta \) and \( \alpha \) are coefficient. Their ratio reflects a balance between accuracy and simplicity. The intelligent learning is to perform a gradient descent minimization of the total difference via [19]

\[ \Delta w_{j,i}^{(k)} = -\eta \frac{\partial E}{\partial w_{j,i}^{(k)}} \]  

(4)

where \( \eta \) is the learning rate. Eq. (4) defines one of the learning methods that always works but unnecessarily the most effective method. In the following section, other learning methods will be used and compared with Eq. (4). The analytical expression for each weight factor and bias can be obtained by back propagation of the partial difference of target function according to Eq. (4).

**Data assessment**

The composition in the database is defined to be atomic percentage. The weight percentage data from literature has been converted according to their molar weight [3, 5-6, 21]. Renormalization has been performed for those data that the original total composition doesn’t come to 100% [22]. This is not rare in steel company become the initial constitutions of C and CO\(_2\) in mould powder are not counted in liquid slag. However, some data in literature has a total composition in excess of 100%. Those data are ignored entirely. For the data contains tiny fraction (3.00 wt.% ) of NaO and K\(_2\)O but without detail information about their ratio [21]. An assumption of 50:50 is applied. This assumption does not affect the specification of the individual contribution from NaO and K\(_2\)O to the electrical conductivity because significant amount of NaO and K\(_2\)O data from other literature have been applied to justify the assumption [23-24]. For the data that have been assessed previously [8], the assessed data instead of large amount of raw experimental data has been adopted to minimize noise. Some data are reported as scattered points in plotted figure. To get the values from the plotted figure with high accuracy, a java-based code package has been developed by the author to convert the figure to data. Some figures have been plotted artificially with scales in coordinates not proportional [22]. Only those data with clear indication of their values have been adopted. For the plotted continuous curves, only those points
accompanied with experimental values are adopted. After the critical assessment, a database contains 752 sets of data with one unitary, five binary, nine ternary, four quaternary, two quinary, two senary and one octonary subsystems has been built up. The detailed subsystems are listed in Table 1. The parameter ranges are shown in Table 2. The unit of temperature is Kelvin. The unit of electrical conductivity is $\Omega^{-1}\cdot\text{cm}^{-1}$.

**Table 1.** The subsystems in the electrical conductivity database

<table>
<thead>
<tr>
<th>No.</th>
<th>Component</th>
<th>System</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>CaF$_2$</td>
<td>[3, 6]</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>CaO+SiO$_2$</td>
<td>[8, 23]</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>CaO+Al$_2$O$_3$</td>
<td>[4, 8]</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>CaO+CaF$_2$</td>
<td>[3, 4, 8]</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>Al$_2$O$_3$+CaF$_2$</td>
<td>[3, 4]</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>MgO+CaF$_2$</td>
<td>[4]</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>CaO+SiO$_2$+Al$_2$O$_3$</td>
<td>[5, 7]</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>CaO+SiO$_2$+NaO</td>
<td>[23, 24]</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>CaO+SiO$_2$+K$_2$O</td>
<td>[23]</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>CaO+SiO$_2$+CaF$_2$</td>
<td>[6]</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>CaO+Al$_2$O$_3$+CaF$_2$</td>
<td>[3]</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>SiO$_2$+Al$_2$O$_3$+NaO</td>
<td>[24]</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>SiO$_2$+NaO+FeO</td>
<td>[24]</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>SiO$_2$+NaO+MgO</td>
<td>[24]</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>SiO$_2$+NaO+MnO</td>
<td>[24]</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>CaO+SiO$_2$+Al$_2$O$_3$+NaO</td>
<td>[24]</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>CaO+SiO$_2$+NaO+K$_2$O</td>
<td>[23]</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>CaO+SiO$_2$+Al$_2$O$_3$+K$_2$O</td>
<td>[23]</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>CaO+SiO$_2$+Al$_2$O$_3$+MgO</td>
<td>[23]</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>CaO+SiO$_2$+Al$_2$O$_3$+NaO+K$_2$O</td>
<td>[4, 5, 7]</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>CaO+SiO$_2$+Al$_2$O$_3$+CaF$_2$+FeO</td>
<td>[23]</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>CaO+SiO$_2$+Al$_2$O$_3$+MgO+CaF$_2$+FeO</td>
<td>[22]</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>CaO+SiO$_2$+Al$_2$O$_3$+MgO+FeO+MnO</td>
<td>[4, 22]</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>CaO+SiO$_2$+Al$_2$O$_3$+NaO+K$_2$O+MgO+CaF$_2$+Cr$_2$O$_3$</td>
<td>[21]</td>
</tr>
</tbody>
</table>

**Table 2.** The range of parameters’ values in the database.

<table>
<thead>
<tr>
<th>Value</th>
<th>CaO</th>
<th>SiO$_2$</th>
<th>Al$_2$O$_3$</th>
<th>NaO</th>
<th>K$_2$O</th>
<th>MgO</th>
<th>CaF$_2$</th>
<th>Cr$_2$O$_3$</th>
<th>FeO</th>
<th>MnO</th>
<th>T (K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>471</td>
</tr>
<tr>
<td>Max</td>
<td>72.24</td>
<td>60.81</td>
<td>43.24</td>
<td>29.21</td>
<td>6.00</td>
<td>20.00</td>
<td>100.00</td>
<td>0.523</td>
<td>34.80</td>
<td>20.00</td>
<td>2039</td>
</tr>
</tbody>
</table>

Neural network computation and results
The activation functions, as illustrated in Eqs. (1.2) and (1.4), have an output value between 0 and 1 for the activation between \(-\infty\) and \(+\infty\). However, \(\sigma = 0.006693\) at \(y = -5\) and \(\sigma = 0.993307\) at \(y = 5\), which indicate the extremely slow approximation to either 0 or 1. Based on this consideration, the training data for electrical conductivity is not normalized by the true minimum and maximum values in the database but multiplied by 0.7 to the minimum value and 1.3 to the maximum value. The input parameters for composition and temperature are all normalized to a value between 0 and 1 to ensure every input parameter has the same weight of contribution. The denormalization and normalization procedure followed the following equations

\[
\sigma = \frac{\sigma - 0.7\sigma_{\min}}{1.3\sigma_{\max} - 0.7\sigma_{\min}} \tag{5.1}
\]

\[
\hat{c}_i = \frac{c_i - c_i_{\min}}{c_i_{\max} - c_i_{\min}} \tag{5.2}
\]

where \(\sigma\) and \(\hat{c}_i\) are the normalized electrical conductivity and composition for \(i < 11\) and temperature when \(i = 11\). The initial values for all the weight factors and biases are assigned a random float value between -5 and 5. A high quality random number generator was coded according to a probability theory developed by Marsaglia et al. [25].

In neural network calculation, it has been noted that the artificial learning by means of Eq. (4) in every time iteration does not help to find the weight factors and biases to achieve minimum total differences between the target values and calculated value. The weight factors soon to adjust their values to minimum the overall target function \(E\) rather than the total difference \(E_D\). To overcome this problem, the regulation term \(E_W\) is not included in each time iteration but replaced at the final step assessment. It is also found that the convergence rate is almost doubled by the following artificial learning mechanism, which agrees with the suggestion from Rumelhart et al. [19]

\[
\Delta w_{ji}^{(k)}(t + 1) = -\eta \frac{\partial E}{\partial w_{ji}^{(k)}} + \delta \Delta w_{ji}^{(k)}(t) \tag{6}
\]

where \(\delta\) is a coefficient. For the 2-layer neural network calculation with \(n = 16\) and \(\alpha = \beta = \eta = \delta = 0.5\), the evolution of total difference, regulation term and target function vs iteration steps are demonstrated in figure 2. It shows that that total difference drops sharply in the early stage (labelled by A), followed by slow drops (labelled by B) until a flat stage (labelled by C) to fluctuate around a
minimum value. However, the regularization term was increased slowly but monotonically until stage C. This is due to the early mentioned decision of not to include the minimization of regulation term in the time iteration. The target function has been reduced monotonically until the flat stage.

Fig. 2 The evolution of total difference ($E_D$), regulation term ($E_W$) and target function ($E$) vs. iteration steps for 2-layer neural network with hidden layer containing 16 units.

To determine the optimum number of units in the hidden layer in 2-layer neural network calculation, one has calculated the change of differences until $1.6 \times 10^7$ iteration steps for various number of units. The results are plotted in figure 3. Although $E_D$ decreases when the number of units increases, $E_W$ demonstrates some optimized values. $E_W$ increases sharply when the number of units is away from the optimized one. The target function, $E$, reveals an optimized value for the number of units in the hidden layer of the 2-layer neural network. Based on the results, $n = 16$ is chosen. It is worth mentioning that the local minimum at $n = 16$ for the curve of $E_D$ is out of expectation. To double check whether it is a numerical coincidence, the code and parameters were run at 3 difference workstations.
but the results were very similar, given the fact that the initialization of weight factors involves a random number generator which should be different at different computers. The smallest $E_D$ appeared at the 9/02/00 iteration step. The values for the weight factors and biases at this optimized condition are listed in Table 3. These values can be used to calculate the electrical conductivity of the system at any composition and temperature in the parameters' range.

### Table 3: The optimized weight factors and biases obtained by artificial neural network calculations

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.38</td>
<td>6.34</td>
<td>9.30</td>
<td>12.36</td>
<td>15.32</td>
<td>18.28</td>
<td>21.24</td>
<td>24.20</td>
<td>27.16</td>
<td>30.12</td>
<td>33.08</td>
<td>36.04</td>
<td>39.00</td>
<td>41.96</td>
</tr>
</tbody>
</table>

These values are used to calculate the electrical conductivity of the system at any composition and temperature in the parameters' range.
Fig. 3 The evolution of total difference ($E_D$), regulation term ($E_W$) and target function ($E$) vs. number of units in hidden layer after $1.6 \times 10^7$ iteration steps.

The optimized weight factors and bias values have been implemented to calculate the electrical conductivity for 752 sets of compositions and temperature. The results are plotted in figure 4 and compared with the values in the database. Figure 4(a) shows the comparison at linear scale coordinates. The $45^\circ$ line indicates the perfect agreement between the artificial neural network computational results and the value in database, where majority data are from experimental measurement and the rest from assessment based on experimental values. The figure shows good agreement. Due to the wide range distribution of the electrical conductivity values from 0.016 to 23.771, which across three orders of magnitude, the comparison in logarithmic scale is shown in figure 4(b). The data shows some almost evenly distribution around the $45^\circ$ line, majority with absolute error below 5%. The largest absolute discrepancy appears in the lowest electrical conductivity end, as is circled in figure 4(b). Those data are
found all belong to CaO-SiO₂-Al₂O₃ subsystem at a temperature either in 1623 K or 1673 K, and was reported in one paper.

Fig. 4 The electrical conductivity from numerical results by artificial neural network calculation vs. the value in the database: (a) in linear scale plotting, and (b) in logarithmic scale plotting.
To validate the artificial neural network calculations, the optimized weight factors and bias values have been implemented to calculate two binary systems CaF$_2$-Al$_2$O$_3$ and Cao-CaF$_2$ at different temperature. The results have been compared with the experimental results reported in various literature [6, 26-27]. Figure 5 presents the results for (a) CaF$_2$-Al$_2$O$_3$ at 1773 K, (b) CaF$_2$-Al$_2$O$_3$ at 1973 K, (c) Cao-CaF$_2$ at 1773 K and (d) Cao-CaF$_2$ at 1873 K. These experimental data are not in database during the training of neural network. Figure 5 shows that the electrical conductivity obtained in the neural network calculations are within the fluctuation of various experimental measurements. It proves that the artificial neural network prediction for the electrical conductivity can be used to predict the change of electrical conductivity at various subsystems in different compositions and temperature.

![Graph showing electrical conductivity vs. molar composition of CaF$_2$ in Al$_2$O$_3$](image-url)
Fig. 5 Validation of neural network calculation by experimental data reported in literature [6, 26-27].
(a) CaF$_2$-Al$_2$O$_3$ at 1773 K, (b) CaF$_2$-Al$_2$O$_3$ at 1973 K, (c) CaO-CaF$_2$ at 1773 K, and (d) CaO-CaF$_2$ at 1873 K.

The artificial neural network and machine learning has many potential applications in steel metallurgy [28]. In future, more works will be done to include other components to the system, such as NiO, TiO$_2$, MgF$_2$, BaF$_2$, BaO, ZrO, CaS. The availability of the new experimental measurement method for electrical conductivity enables to get more accurate data in other systems [29], which will help to build up database for training and validation of the neural networks. The future work can, hopefully, also include the effort to use the data and machine learning method to identify the main oxides that control the electrical conductivity of mould flux and the influence of temperature on the electrical properties, and to compare the results with the theoretical predictions [7-8, 15].

Conclusions

- An electrical conductivity database for CaO-SiO$_2$-Al$_2$O$_3$-NaO-K$_2$O-MgO-CaF$_2$-Cr$_2$O$_3$-FeO-MnO liquid mould slag system has been built up.
• The database has been implemented to train an artificial neural network. It is found that the two-layer neural network with 16 units in hidden layer provides the minimum difference in target function. The optimized weight factors and bias values can be used to calculate the electrical conductivity of the system in a wide range of constitutional compositions and temperature.

• The numerical prediction has been validated by the experimental results reported in literature. Excellent performance of artificial neural network derivation has been proved.

Acknowledgements

This work was financially supported by European Commission RFCS (No. 847269) and Engineering and Physical Sciences Research Council in UK (No. EP/R029598/1).

Disclosure statement

The author declares that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

References


