Evaluation of ANN, ICA-ANN and PSO-ANN predicting ability in the prediction of CO$_2$ emissions during the calcination of cement raw material

Yakoub Boukhari $^1$

$^1$ Ziane Achour University, P.O. Box 3117 Road Moudhiba Djelfa, Algeria

Abstract
Cement industry releases large amounts of carbon dioxide CO$_2$ as by-product to the atmosphere during the calcination of cement raw material. In fact, the calcination is a complex process and not completely understood. The amount of CO$_2$ emitted varies with the grain size, chemical composition, burning temperature and time to pass through the kiln during calcination process. However, due to interaction of several parameters, it is not easy to establish accurate mathematic model to calculate the real amount of CO$_2$ emission. Moreover, using the laboratory experiments to determine the amount of CO$_2$ emissions are not usually easy, time-consuming, expensive and require good quality of reagents and equipments. To overcome the above problems, artificial neural network (ANN), ANN optimised by imperialist competitive algorithm (ICA-ANN), ANN optimised by particle swarm optimization (PSO-ANN) are applied to predict amount of CO$_2$ emissions. A comparative accuracy of these tools is evaluated based on the coefficient of determination $R^2$, $R^2$ adjusted, mean absolute percentage error (MAPE) and scatter index (SI).

The results obtained are promising and demonstrate that all proposed tools represent a good alternative for the prediction of CO$_2$ emission with adequate accuracy. PSO and ICA are capable to improve the predicting accuracy of ANN. In addition, PSO-ANN can predict slightly better than ICA-ANN. Based on testing data, the results obtained show that 98.61%, 98.18% and 97.5% of experimental data are explained by PSO-ANN, ICA-ANN and ANN, respectively with average relative error less than 1.41% and SI less than 0.1.

Keywords
CO$_2$ emissions, calcination process, artificial neural network, imperialist competitive algorithm, particle swarm optimization

1. Introduction

The cement is used extensively in a diversity of construction projects. The one of the most important step in the cement production process is clinker calcination process of raw materials. The calcination of raw materials is produced in cement kilns at high temperature. In fact, the calcination process is a complex thermo-chemical reaction, and at the same time, it is greatly influenced by heat transfer, mass transfer from inside particle to reaction interface, chemical reaction and experimental conditions [1]. It is complicated process due to complex interactions of the influencing parameters between them [2]. In fact, the calcination process is a complex process and not completely understood.

The main by-product of clinker calcination process of raw materials is CO$_2$ emitted from thermal chemical decomposition reaction of limestone [3]. The amount of CO$_2$ emissions during calcination of raw materials is very important and it has strong influence in determining cement quality. The main oxides present in in the raw materials are CaO, SiO$_2$, MgO, Al$_2$O$_3$ and Fe$_2$O$_3$ [6]. These oxides have
very significant role in determining the amount of CO$_2$ emissions during calcination. The amount of CO$_2$ emissions vary with their grain size, chemical composition and burning time. However, the influence of these parameters on an amount of CO$_2$ emissions is still not clear. Due to the complexity of calcination process, it is very difficult to account the amount of CO$_2$ emissions by traditional mathematical methods. Moreover, using the laboratory experiments to determine the amount of CO$_2$ emissions are not usually easy, time-consuming, expensive and require good quality of reagents and equipments.

For decades, intelligence methods are widely used in several domain to predict the behavior of complex phenomena [4,5]. The objective of present study is to evaluate the predicting ability of artificial neural network (ANN), ANN optimised by imperialist competitive algorithm (ICA-ANN) and ANN optimised by particle swarm optimization (PSO-ANN) in the prediction of CO$_2$ emissions during the calcination of cement raw material. These models do not need to understand the process behavior for extracting prior knowledge and have strong capability to adapt to system variation. Due to their advantages, ANN, PSO-ANN, ICA-ANN are widely used to solve a diversity of complex problems in many fields. ANN is a very effective tool for predicting the pitting corrosion [7] and velocity of sound in liquid water [8]. PSO-ANN is also successfully applied for predicting of cobalt leaching rate from waste lithium-ion batteries [9] and solar space heating system parameters [10]. ICA-ANN is successfully applied as intelligence models to predict maximum surface settlement caused by tunnelling with higher reliability [11] and oil flow rate of the reservoir [12].

The present paper is organised as follows: Section 2 describes briefly the artificial intelligence tools used to predict the target. Section 3 presents the used materials and methods. Section 4 presents and discuss predicting results. Finally, Section 5 presents our conclusions.

2. Brief description of artificial intelligence tools

2.1. Artificial neural network (ANN)

Artificial neural network is inspired from the biological nervous system within the human brain [13]. It is the most popular intelligence tools because it is able to predict the output of complex nonlinear relationships among variables in a wide range of areas. It is composed of input layer (IL), output layer (OL) and at least one intermediate layers called hidden layers (HL). Each layer contain one or more nodes arranged (neurons). The neurons in each layer are fully connected with neurons in the subsequent layer. There are no links between neurons in the same layer. The neuron mainly consists of weight, bias and activation functions. The weight, bias factors are adjusted and optimised at every iteration by Levenberg Marquardt (LM) based back propagation (BP) during training process [14]. The cost function used by ANN during leaning process mean square error (MSE). It is used to measure the difference between the predicted output and the desired output.

The number of neurons in the input layer and the output layer equal the number of input and output variables in the data, respectively. Whereas, there are no general rules to determine the suitable number of hidden nodes and number of its neurons. The common way is to set a relative large number of neurons at the beginning, and then reduce it gradually until the desired error are achieved. There are some cases where the ANN tool has the disadvantages of slow learning convergence, local optima trapped instead of global optimal solution [15]. PSO and ICA are proposed to overcome the previous shortcomings of ANN and to improve its applications.

2.2. Imperialist Competitive Algorithm-Artificial Neural Network (ICA-ANN)

As mentioned previously, despite the popularity of ANN in prediction complex system, it still has the possibility to fall in a local optimum. Hence, Imperialist Competitive Algorithm is combined with ANN to find the global optimal and avoid premature convergence toward local.

The imperialist competitive algorithm ICA is a new optimization algorithm which is inspired by the imperialistic competition processes of human [16]. ICA algorithm is applied to update the weights and biases during the training process in order to improve efficiency of ANN. Recently, it is very attractive [17] and widely applied to solve discrete optimization problems due to its good convergence rate and
better global optima finding. ICA algorithm starts by initialising a randomly populated colony called countries. In reality, there are two groups of countries which are imperialists and colonies regarding their power. The strongest countries with the minimum best cost are chosen as imperialists, whereas the weakest countries are taken as colonies of these imperialists. The imperialist and their colonies are united to construct the initial empires. ICA algorithm begins an iterative process to arrive at optimal solutions after some number of decades or generations.

Three main operators of ICA are assimilation, revolution, and competition [18]. In assimilation, each colony starts to move to their corresponding imperialist in order to develop its position. During the movement, the colony can attain great power (lower cost) compared to its imperialist. This procedure is called revolution. In this case, their positions will be exchanged and the empire has a new imperialist. At the next step, imperialist competition starts and the weakest empire is eliminated from the competition. At the end, only one of these empires is remained and all the other countries are their colonies. The remaining empire presents the optimal solution. The most important ICA parameters are number of countries, number of imperialists, number of decades.

2.3. Particle Swarm Optimization-Artificial Neural Network (PSO-ANN)

Similar to ICA, Particle Swarm Optimization (PSO) is combined with ANN to form powerful tools and to adjust its setting parameters. PSO is a popular algorithm due to its competitive performance and easy implementation [19]. It is inspired by natural phenomena of birds flocking or schooling fish while searching for food sources. In the natural, birds randomly move in groups and work together by sharing information to achieve a nearest food source. Each bird tries to follow the bird which is nearest to the food. Bird searching for food updates both its speed and position. This process is repeated iteratively until the source of food is found. After a sufficient number of iterations, all birds will eventually discover the nearest path from the nest to the food source. The nearest path is the desired solution [20]. PSO is more attractive because of its quick convergence and only few parameters adjustments are required [21]. The performance of PSO is related principally to the number of particle, number of iterations.

3. Materials and methods

3.1. Materials and experiments

In the present study, CO$_2$ emission is considered as a function of chemical composition, grain size and time exposed. The raw materials are blended and preheated to around 300°C to remove water combined in the hydration products and then up to 850°C to remove impurities, which can affect the cement quality.

Four different grain size distribution (71, 125, 250 and 350 µm) of raw materials used are selected separately. The chemical composition and mix proportions of four raw materials used are summarised in Table 1. Finally, each mixture of raw materials with gain size are burned in the laboratory furnace at 1000°C for different times 5, 10, 15, 20, 30 min. The amount of CO$_2$ emissions is calculated before and after burning of each mixture of raw materials at 1000°C.

<table>
<thead>
<tr>
<th>Raw Materials</th>
<th>SiO$_2$</th>
<th>CaO</th>
<th>MgO</th>
<th>Fe$_2$O$_3$</th>
<th>Al$_2$O$_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Material 1</td>
<td>12.38</td>
<td>80.28</td>
<td>1.38</td>
<td>1.69</td>
<td>4.27</td>
</tr>
<tr>
<td>Material 2</td>
<td>3.96</td>
<td>92.62</td>
<td>0.99</td>
<td>0.65</td>
<td>1.78</td>
</tr>
<tr>
<td>Material 3</td>
<td>14.06</td>
<td>78.69</td>
<td>1.35</td>
<td>1.68</td>
<td>4.22</td>
</tr>
<tr>
<td>Material 4</td>
<td>14.16</td>
<td>78.04</td>
<td>1.36</td>
<td>2.21</td>
<td>4.23</td>
</tr>
</tbody>
</table>

3.2. Dataset collection

The dataset extracted from the experimentation is collected in a table of 80 rows and 8 columns. Each row in this table presents an experiment. From 1 to 7 columns are inputs where the last column is
output. The size particle, time exposed, SiO$_2$ (%), CaO (%), MgO (%), Fe$_2$O$_3$ (%), Al$_2$O$_3$ (%) are inputs and the amount of CO$_2$ emissions is the output.

The total dataset are randomly divided into two sets: training and testing. For each algorithm, 75% of dataset is used for training while the remaining 25% (unseen dataset) of dataset are kept out to evaluate the generalisation ability. The most common performance criterion used to evaluate the accuracy of each algorithm are the coefficient of determination $R^2$, adjusted $R^2$ (adj), the mean absolute percentage error (MAPE) and the scatter index (SI). The tool performance is perfect when value of $R^2$ is very close to 1, while value of MAPE are very close 0. The predictive accuracy is excellent when SI is inferior of 0.1; good if SI among 0.1 and 0.2; and bad if SI more than 0.3. [22].

4. Results and discussion

In fact, selecting an optimal settings parameters of each intelligence methods tools during training stage is a challenging task. The trial-and-error method is considered as best method to find the optimal settings parameters [23]. It is applied in this study to determine different setting parameters.

4.1. ANN results

It is well known, that ANN has ability to learn the relationship between inputs and outputs in the presence sufficient number of hidden layer and neurons with suitable transfer functions [7]. After trying various ANN parameters during the training phase, the more appropriate structure parameters of ANN are determined as listed in Table 2.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of HL</td>
<td>1</td>
</tr>
<tr>
<td>Neuron number in HL</td>
<td>12</td>
</tr>
<tr>
<td>Transfer function for HL</td>
<td>transig</td>
</tr>
<tr>
<td>Transfer function for OL</td>
<td>purelin</td>
</tr>
<tr>
<td>Training Algorithm</td>
<td>LM based BP</td>
</tr>
</tbody>
</table>

The comparison between experimental and predicted CO$_2$ emissions values obtained by ANN for training and testing phases are shown in Figure 1 and Figure 2, respectively.

![Figure 1: Comparison between predicted and experimental CO$_2$ emissions in testing phase](image-url)
Figure 2: Comparison between predicted and experimental CO$_2$ emissions in training phase

The distribution of the relative error obtained ANN during training and testing phases is illustrated in Figure 3. It is clear from Figure 3 that the relative error values are nearly spread around the zero line. The average relative error for training phase is 0.58% while it is 1.41% for testing phase. While the maximum error for training phase is 2.36% and for testing phase is 4.02%. These low values of MAPE are indicative of very small difference between experimental of CO$_2$ emissions and predicted ones.

The results illustrated in Figure 3 confirm that ANN has good generalisation capability and can predicts the amounts of CO$_2$ emissions adequately, as indicated by the low value of MAPE equal to 1.41%, high value R$^2$ equal to 0.9763.

Figure 3: Distribution of the relative error obtained ANN during training and testing phases

From Table 3, the values of R$^2$ adjusted indicate that ANN is able to predict approximately 98.94% of training dataset and 97.50% of testing dataset. The values of SI less than 0.1 obtained in both phases mean adequate predictive capability. The high accuracy of ANN is usually due to its flexible architecture and its excellent performance in solving the nonlinear mapping between the inputs and outputs.

Table 3

<table>
<thead>
<tr>
<th>Parameters</th>
<th>R$^2$ adj</th>
<th>SI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training phase</td>
<td>98.94%</td>
<td>0.0076</td>
</tr>
<tr>
<td>Testing phase</td>
<td>97.50%</td>
<td>0.0176</td>
</tr>
</tbody>
</table>
4.2. PSO-ANN results

The PSO algorithm is used to train and optimise weights and biases of previously ANN architecture to form powerful tool. The objective of using the same architecture is to evaluate the capability optimising of PSO algorithm. The optimal parameter configuration of PSO-ANN utilised are summarises in Table 4.

Table 4
PSO-ANN parameters

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of particle</td>
<td>14</td>
</tr>
<tr>
<td>Number of iteration</td>
<td>16</td>
</tr>
<tr>
<td>Acceleration constant ((C_1=C_2))</td>
<td>1.5</td>
</tr>
<tr>
<td>Number of HL</td>
<td>1</td>
</tr>
<tr>
<td>Neuron number in HL</td>
<td>12</td>
</tr>
<tr>
<td>Transfer function for HL</td>
<td>transig</td>
</tr>
<tr>
<td>Transfer function for OL</td>
<td>purelin</td>
</tr>
<tr>
<td>Training Algorithm</td>
<td>LM based BP</td>
</tr>
</tbody>
</table>

The predicted values of CO\(_2\) emissions obtained from PSO-ANN are compared with the experimental ones for training phase and testing phase as shown in Figure 4 et Figure 5, respectively. It is clear that almost points are very closely clustered around the line of equality \((y=x)\) and lied exactly linear fit. Moreover, the linear fit is sloped with an angle close to 45° that means strong linear relationships between to predicted and experimental dataset. PSO-ANN can achieve \(R^2\) of 0.9870 in the training phase and 0.9868 in the testing phase. The \(R^2\) values close to 1 mean that the predicted CO\(_2\) emissions are very close to the real experimental values. The \(R^2\) mean that less than 1.4% of testing dataset and training dataset can not explain by PSO-ANN.

The distribution of the relative error obtained during training and testing phases are plotted in Figure 6. It is observed that almost of points are tightly concentrated near to line zero. The relative errors are relatively less in both phases, where the maximum error not exceed 3.15% in training phase and 3.17% in testing phase. The average relative error for training and testing are 0.6% and 0.97%, respectively.

![Figure 4: Comparison between predicted and experimental CO\(_2\) emissions in testing phase](image-url)
Figure 5: Comparison between predicted and experimental CO\textsubscript{2} emissions in training phase

Figure 6: Distribution of the relative error obtained PSO-ANN during training and testing phases

The prediction results approve the feasibility of the PSO-ANN and show the good generalization capability. As reported in Table 5, the values of $R^2$ adjusted mean that PSO-ANN can predicted 98.6\% of total dataset correctly. In addition, the SI values less than 0.1 reflect the excellent predicting ability of amount of CO\textsubscript{2} emissions.

Table 5

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$R^2$ adj</th>
<th>SI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training phase</td>
<td>98.70%</td>
<td>0.0088</td>
</tr>
<tr>
<td>Testing phase</td>
<td>97.50%</td>
<td>0.0114</td>
</tr>
</tbody>
</table>

The high performance of PSO-ANN is explained by the capability of PSO to find the global optimum solution and optimum structure of ANN and high capability of ANN to learn by example during training process. In summary, PSO-ANN tool is very useful in predicting the amount of CO\textsubscript{2} emissions with very high value $R^2$ and very low value of MAPE.

4.3. ICA-ANN results

Similar to previous case, ICA is also used for optimising the weights and bias values in ANN. The best parameters values of ICA utilised during training process to optimise and to improve the prediction performance accuracy of ANN are shown in Table 6.
### Table 6
ICA-ANN parameters

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of countries</td>
<td>25</td>
</tr>
<tr>
<td>Number of initial imperialists</td>
<td>10</td>
</tr>
<tr>
<td>Number of decades</td>
<td>3</td>
</tr>
<tr>
<td>Number of HL</td>
<td>1</td>
</tr>
<tr>
<td>Neuron number in HL</td>
<td>12</td>
</tr>
<tr>
<td>Transfer function for HL</td>
<td>transig</td>
</tr>
<tr>
<td>Transfer function for OL</td>
<td>purelin</td>
</tr>
<tr>
<td>Training Algorithm</td>
<td>LM based BP</td>
</tr>
</tbody>
</table>

The capability of ICA-ANN to predict amount of CO\(_2\) emissions during calcination process is shown in Figure 7 and Figure 8.

**Figure 7:** Comparison between predicted and experimental CO\(_2\) emissions in testing phase

**Figure 8:** Comparison between predicted and experimental CO\(_2\) emissions in training phase

The plots clearly illustrate that almost of dataset in training and testing phases fall on a linear fit which is mostly overlapped with line of line of equality (y=x). The R\(^2\) value are high for both phases and are near to one, reflecting strong linear relationships between predicted amount of CO\(_2\) emissions...
and experimental ones. The values of $R^2$ reveal that more than 98% of testing and training dataset are predicted perfectly by ICA-ANN.

The accuracy of amount of CO$_2$ emissions prediction of ICA-ANN is shown in Figure 9. It is clear, the dispersion of points dataset is quite close to the line zero. The relative error is almost low in the slip range of 0 to 2.51% in both phases. ICA-ANN is capable of providing average relative error values equal to 0.6% and 1.12% for training and testing phases, respectively. These values illustrate that the CO$_2$ emissions predicted are very close to the real experimental ones.

The distribution of the relative error obtained by ICA-ANN during training and testing phases is shown in Figure 9. It is clear, the dispersion of points dataset is quite close to the line zero. The relative error is almost low in the slip range of 0 to 2.51% in both phases. ICA-ANN is capable of providing average relative error values equal to 0.6% and 1.12% for training and testing phases, respectively. These values illustrate that the amount of CO$_2$ emissions predicted by ICA-ANN are very close to the real experimental ones.

![Figure 9: Distribution of the relative error obtained ICA-ANN during training and testing phases](image)

The performance criteria for both phases are illustrated in Table 7. The adjusted $R^2$ adjusted indicate that only 1.08% of training and 1.72% of testing dataset are not explained by this model. The values of SI that are less than 0.1 signify excellent capability of predicting.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$R^2$ adj</th>
<th>SI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training phase</td>
<td>98.92%</td>
<td>0.0080</td>
</tr>
<tr>
<td>Testing phase</td>
<td>98.28%</td>
<td>0.0118</td>
</tr>
</tbody>
</table>

Results obtained reveal that ICA-ANN can produce excellent predicting results with high values of $R^2$ and low values of MAPE. The high accuracy of ICA-ANN is mostly due to its great capability of optimizing and the self-adaptive learning ability of ANN.

4.4. Comparison between different tools

Based on testing dataset, ANN non-optimised is compared to PSO-ANN, ICA-ANN to evaluate predicting ability of each tool in the prediction of CO2 emissions during the calcination of cement raw material and the capacity of PSO and ICA in optimising of parameters of ANN. This comparison is presented in Figures 10 and Figure 11.

Firstly, PSO-ANN, ICA-ANN and ANN act as robust and powerful tools in predicting amount of CO$_2$ emissions and can generate good accuracy. As can be seen from Figure 10 and Figure 11, using PSO and ICA for optimising weight and bias can lead to a good predicting ability on result compared to simple ANN. For PSO-ANN, the values of $R^2$ and MAPE are 0.9763 and 1.14%, respectively whereas after combining ANN with PSO the value of $R^2$ and MAPE become 0.9868 and 1.01%, respectively. A similar improvement is observed with ICA-ANN. The results reveal the highest prediction capacity of PSO-ANN compared to ICA-ANN and ANN. Furthermore, the ANN efficiency is less than ICA-ANN by according to the results obtained via R2 and MAPE. The predicting ability of
ANN, ICA-ANN and PSO-ANN are as excellent as expected and they can reveal the real relationship between the influencing parameters and target. Based on testing dataset, the results obtained show that 1.39%, 1.82% and 2.50% of experimental dataset are not explained by PSO-ANN, ICA-ANN and ANN, respectively with average relative error less than 1.41% and SI less than 0.1.

![Figure 10: Comparison between PSO-ANN, ICA-ANN and ANN in term of $R^2$ and MAPE](image)

![Figure 11: Comparison between PSO-ANN, ICA-ANN and ANN in term of SI and $R^2$ adjusted](image)

5. Conclusion

In present paper, PSO-ANN, ICA-ANN and ANN tools are proposed, and their prediction performances of amount of CO$_2$ emissions is evaluated through a comparison with the experimental ones. Based on testing dataset, the results obtained demonstrate that all tools proposed are very useful tools for fast prediction of amount of CO$_2$ emissions with high generalization performance. Using PSO and ICA for optimising weight and bias can lead to a good predicting ability on result compared to simple ANN. Based on the same neural network architecture, PSO-ANN has highest predicting ability with comparative high value of $R^2$ and less value of MAPE, followed ICA-ANN, while slightly less performance is seen in the case of ANN non optimised.

Based on testing dataset, the results obtained show that 98.61%, 98.18% and 97.5% of experimental dataset are explained by PSO-ANN, ICA-ANN and ANN, respectively with average relative error less than 1.41% and SI less than 0.1. Finally, the results obtained are promising and demonstrate that all proposed tools represent a good alternative for the prediction CO$_2$ emission during the calcination of cement raw material with excellent accuracy.
6. References


