Application of BP Neural Network and PSO Methods for Process Design of Cold Rolling Steel

Min-ming CHEN, Xian-hui LIU, Wei-dong ZHAO and Yu-fei CHEN

College of Electronics and Information Engineering of Tongji University, China

Keywords: Process design, Cold rolling steel, Particle swarm optimization, Neural Network.

Abstract. Exploring the relationship between composition, process and mechanical properties of cold rolling steel is very important for steel production. The most direct ways is to directly put into actual production. According to the results of actual production, we can tell the composition and process and to obtain the property. But this method is expensive for steel companies. In this paper, we proposed a prediction model for predicting the mechanical properties of cold rolling steel. And based on the mechanical properties prediction model and particle swarm optimization, a process design method is proposed. This model can predict the material composition and process of the steel given the steel mechanical properties. The two models both achieved good results finally.

Introduction

We will deduce the final mechanical properties according to the composition and process of the steel materials, which is called the prediction of the steel properties [1]. And we will calculate the proper composition and process according to the required mechanical properties, which is called process design. How to predict the mechanical properties of steel is always popular research work for steel enterprises. A good steel mechanical properties prediction model can reduce the cost of time and funds. However, no one has yet proposed a mature method to solve the problem of steel properties prediction until now. And the process design method we need is much few and more difficult.

In the past few decades, a large number of scholars have devoted to the study of steel mechanical properties prediction models. At present, the main ideas are divided into two types, one is based on the metallurgical mechanism model [3-4], the other is the mechanical properties prediction model[5-10] through statistics. In the metallurgical mechanism based model, people are mainly concerned to find out the relationship between the constituent elements of steel and the ultimate mechanical properties. [3] Such models often start from the process to explore the effects of metallurgical processes and materials on the composition of the steel material produced, and then the mechanical properties are calculated by the composition of the steel material obtained[4] . But such methods often require deep physical and chemical knowledge and fully understanding of each step, so it is often difficult to build a model through this way. [2]

The data accumulation during recent decades has provided a basis for the establishment of the prediction model of steel mechanical properties through mathematical statistics. In the last decade, a large number of scholars have invested in this research[5]. The main idea is to fit a nonlinear model with high accuracy based on a large number of data. [2, 4-10] In 2001, Perzyk proposed a mechanical properties prediction model based on ANN network, and achieved good effect [4]. After that, some scholars used support vector machines and Bayesian networks respectively to establish the mechanical properties prediction model [6, 7]. D. M. Jones compares multiple linear regression, multiple nonlinear regression and neural networks, within which neural networks has achieved the best effect. [5]. This kind of method is different from the traditional metallurgical mechanism model, and it does not need to grasp deep metallurgy knowledge to establish an excellent model.

The previous research mainly focused on how to put forward a high accuracy prediction model, but the enterprise is more likely to know what the composition and process of these steel types may need, given the mechanical properties of the required steel. L. A. Dobrzenski proposed a prediction model based on genetic algorithm, which can predict the corresponding material composition and process
reversely according to the required mechanical properties, and achieved a certain effect [10]. This provides a way for the research of process design.

In this paper, a mechanical properties prediction model applied to the cold rolling steel is proposed. On the basis of this model, a process design model is proposed, which is equivalent to the reverse model of the above model. The model can obtain the corresponding components and processes given the mechanical properties the steel mechanical properties.

Model

Mechanical Properties Prediction Model

Cold rolling steel mainly uses hot rolling steel as material, and after pickling, rolling, annealing, leveling, oiling, and finally packaging to produce, cold rolling steel with better mechanical properties is produced. Compared with hot rolling steel, cold rolling steel is thinner and better in surface quality, which can satisfy wider mechanical mechanical properties requirements. But the production process of cold rolling product is also more complicated.

The mechanical properties prediction model are mainly used to accomplishes the mechanical properties prediction in the cold rolling production process. We use 15 parameters affecting the steel mechanical properties in cold rolling production as input of the forecast model, including chemical composition such as C, S, P, N, and the main process parameters of cold rolling production such as coiling temperature and rolling speed. The output of the model is the five mechanical mechanical properties indicators of steel YP, TS, EL, R90 and N90.

BP neural network is a nonlinear model with good predictive mechanical properties. In the previous work, there has been a lot of work to prove that the prediction mechanical properties of BP neural network is somewhat better than other nonlinear prediction models. Therefore, this experiment will use BP neural network as the benchmark model of mechanical properties prediction model. We need consider the number of hidden layers, the number of hidden layer neurons, activation function, error functions and training methods.

In the experiment, we first tested the network of three and four layers, with 30 to 70 neurons. We found that under a single network structure, the result of the prediction on the five mechanical mechanical properties was not good, and there was a large error in some dimensions. We tried to use five different networks to predict the five properties, and we find that the results are much better than single network, so our model uses five networks to predict five outputs respectively. Although some people have considered using multiple networks to predict properties, many networks have the same structure. [2] we consider that whether we can get better results if we use different structure networks. Finally, our network structure is shown in Figure 1.

![Figure 1. Network structure.](image)

We design the input layer of the five prediction networks with 15 input neurons corresponding to the 15 main parameters affecting the steel mechanical properties mentioned above and each input will not have a large difference by normalization process. The output layer contains four output neurons corresponding to the 5 main mechanical mechanical properties we need. We mainly use two
structures to define these five neural networks. The first structure is applied to the YP network and the EL network. This network structure selects two hidden layers, including 31 and 33 neurons, and the activation function is RELU. The second kind of neural network is applied to TS network, R90 network and N90 network. This kind of network structure is a single hidden layer with 61 neurons. We choose the Relu as our activation function. In both two types of networks, we use RMSE as our cost function and we train it through gradient descent. Eq. 1 express the function of RMSE. In this function $o_i, i = 1,2,...,5$ represents the predicted value of network $i$ and $y_i$ is the output to the test set.

$$
RMSE = \sqrt{\frac{\sum_{i=1}^{5} (o_i - y_i)^2}{5}}
$$

(1)

**Process Design Model**

The process design model can predict the 15 dimension input parameters of the mechanical properties when given the five mechanical mechanical properties. we can regard it as the reverse model of the mechanical properties prediction model. Some people considered using genetic algorithms to solve the problem of process design[10], but the effect is not good. The reason is that the inputs and outputs of genetic algorithm is discrete variables, but variables in the steel process design problem are continuous. Particle Swarm Optimization (PSO) is a new evolutionary algorithm developed by J. Kennedy and RC Eberhart et al. [11]. Starting from a random solution, it finds the optimal solution by iteration. Compared with genetic algorithms, particle swarm optimization is more suitable for solving continuous problem.

Our process design model is mainly based on the PSO algorithm and the mechanical properties prediction model we built above. The main idea is to produce a set of random particle populations as our initial solution. We calculate the mechanical properties indicators of the particle populations by the above-mentioned mechanical properties prediction models, and optimize the particle populations by certain rules. Then calculate the mechanical properties indicators until the resulting mechanical properties indicators have a small difference from our required mechanical properties indicators.

In the first iteration, we randomly generate 20 sets of 15 dimensional random numbers as the particles. We denote them as $x_i = (x_{i1}, x_{i2},...,x_{i15}), i = 1,2,...,20$ and randomly generated the velocity of these particles $v_i = (v_{i1}, v_{i2},...,v_{i15})$, then we calculated the fitness $f_i, i = 1,2,...,20$ of each particle by Eq. 2.

$$
f_i = \sum_{n=1}^{5} (\text{net}_n(x_i) - y_n)^2
$$

(2)

$\text{net}_n(x), n = 1,2,...,5$ denote the five mechanical properties prediction networks established by us, and $y_n$ is the target mechanical mechanical properties. The particles with smaller fitness are more consistent with our needs. In the first iteration, we use $f_i$ as the best solution for particle $i$ denote as $\text{bf}_i$, and the corresponding $x_i$ of the $\text{bf}_i$ denote as $\text{pb}_i$, and the corresponding $x_i$ to the best $\text{bf}_i$ of all particles denote as $\text{gb}$. Then we update the $x_i$ values and $v_i$ values of the 20 particles by Eq. 3 and Eq. 4.

$$
v_i = v_i + c_1 * r() * (\text{pb}_i - x_i) + c_2 * r() * (\text{gb} - x_i)
$$

(3)

$$
c^2 = a^2 + b^2
$$

(4)
N is a function that generates random numbers between 0 and 1. \( c_1, c_2 \) are two constant values and we set both of them to 1.4995.

We update \( x_i \) and recalculate \( f_i \) and record the smaller one in \( f_i \) and \( bf_i \) as new \( bf_i \). Then we update \( pb_i \) and \( gb \). After above options we update \( y_i \) and \( x_i \) again. We repeat the above steps until we reach a certain number of iterations or the fitness of a particle to meet our requirements. Here we think that the iteration can be stopped when the iteration reaches 100 or a particle's \( f_i < 0.05 \), and \( pb_i \) is our prediction result.

**Experiment**

**Mechanical Properties Prediction Model**

We first verify the effect of the mechanical properties prediction model. This experiment uses the data collected in the actual production. Finally, 4500 sets of data left after removing the partial data containing human interference. The fifteen dimension parameters include the content of C, S, P, N, Mn, O2, Al and other technological parameters such as finishing rolling temperature, coiling temperature, thickness, annealing temperature, rolling speed, entry and exit tension, annealing times. In addition to these 15 dimensional parameters, some parameters are not used as input parameters of this experiment due to their small deviation. We randomly select 3000 groups of data as our training set, and use the remaining 1500 sets of data as our test set. We use EP to express the error rate of a sample and calculate by the Eq. 5:

\[
EP = \frac{O - y}{y} \times 100
\]

(5)
o represents the predicted value of the network, and \( y \) represents the actual value. We tested the error rates of five networks separately, and the results are shown in the Fig. 2-6.

![Figure 2. YP Error rate distribution.](image1)

![Figure 3. TS error rate distribution.](image2)

![Figure 4. EL error rate distribution.](image3)

![Figure 5. R90 error rate distribution.](image4)

![Figure 6. N90 error rate distribution.](image5)
We use 1-ErrorPercent as the accuracy rate and the result is shown in Table 1. It can be seen that although some networks have poor prediction results on some test data, such as YP network and R90 network. But the prediction accuracy of each network is above 90 and it is a very good result.

<table>
<thead>
<tr>
<th></th>
<th>YP</th>
<th>TS</th>
<th>EL</th>
<th>R90</th>
<th>N90</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>75.88</td>
<td>94.25</td>
<td>84.63</td>
<td>70.71</td>
<td>90.96</td>
</tr>
<tr>
<td>Average</td>
<td>95.15</td>
<td>98.77</td>
<td>96.37</td>
<td>93.54</td>
<td>98.49</td>
</tr>
<tr>
<td>Maximum</td>
<td>99.99</td>
<td>99.99</td>
<td>100</td>
<td>99.99</td>
<td>100</td>
</tr>
<tr>
<td>Variance</td>
<td>5.36</td>
<td>1.54</td>
<td>4.51</td>
<td>8.07</td>
<td>1.83</td>
</tr>
</tbody>
</table>

We use NEC to express error rate of the model and is calculated by Eq. 6.

\[ NEC = 1 - \frac{1}{5N} \sum_{n=1}^{N} \sum_{i=1}^{5} \left| \frac{o_{n,i} - y_{n,i}}{y_{n,i}} \right| \times 100 \]  

(6)

Among them, \( o_{n,i} \) represents the predicted value of the network \( i \) to the \( n \) sample, and \( y_{n,i} \) represents the actual value. Through the above formula, we calculate the average accuracy of the whole model reaches 96.464.

**Process Design Model**

The test of the process design model also selects the above 1500 sets of test data. We use the error rate to represent the PEP single test data and calculate by the following formula. We use PEP to represent the error rate of the single test data and calculate by the Eq. 7.

\[ PEP = \frac{\sum_{i=1}^{5} \left| \text{net}_i (\text{pb}_i) - y_i \right|}{5} \times 100 \]  

(7)

\( y_{i,1,2,3,4,5} \) is the actual value of \( i \) mechanical property. \( \text{pb}_i \) is the 15 dimensional prediction process parameters we calculated. net, express the five neural networks we have established. Finally, we get the distribution of error values for all test samples as shown in Figure 7:

![Figure 7. Process design model error rate distribution.](image)

The statistical results of the above data are shown in Table 2. Although the error distribution of the output results is wide and has a large variance, the prediction error of the 87.9% sample is less than 20%. So generally speaking, the process design model has good mechanical properties in the process design of cold rolling steel.
Summary

In this article, we applied the BP neural network to the mechanical properties prediction of cold rolling steel successfully. And the model reflects the relationship between 15 input parameters and 5 steel mechanical properties with a high accuracy. At the same time, we put forward a process design model successfully, and have a good prediction effect. The model calculates the required process parameters accurately given the required properties of the steel. We believe the process design model will provide great help to steel enterprises in the future.

Acknowledgement

The National Science Foundation. National Key Research and Development Program 2017YFB030410. Construction Specific Project of Shanghai Engineering Research Center 17DZ2251600 Central University Fund.

References


