# A VARIABLE SELECTION METHOD FOR PULVERIZING CAPABILITY PREDICTION OF TUMBLING MILL BASED ON IMPROVED HYBRID GENETIC ALGORITHM

## Hui Cao, Lixin Jia\*, Gangquan Si, Yanbin Zhang

State Key Laboratory of Electrical Insulation and Power Equipment, Electrical Engineering School Xi'an Jiao Tong University, Shaanxi Province, 710049, China e-mail: {huicao, lxjia, gangquan, ybzhang}@mail.xjtu.edu.cn

crossref http://dx.doi.org/10.5755/j01.itc.40.3.629

**Abstract**. Tumbling mill of thermal power plant, grinding the raw coal for the boiler, has high energy consumption, and pulverizing capability is usually used for representing the efficiency of tumbling mill. In the paper, a variable selection method for pulverizing capability prediction of tumbling mill based on improved hybrid genetic algorithm is proposed. Based on the tradition GA, the proposed method adopts the multi-population mechanism, the elites sharing mechanism and the heterogeneity mechanism for avoiding the premature convergence. The support vector machine is used for building the prediction model of the pulverizing capability with the selected variables. The proposed method is performed on the real field data. The results of the experiments verify that the proposed method has faster convergence speed and the model of pulverizing capability built with the variables selected by the proposed method has higher prediction precision. In addition, the proposed method has been put into practice and the field operation curve verifies that the pulverizing capability could be predicted successfully.

Keywords: Tumbling Mill; Pulverizing Capability; Variable Selection; Improved Hybrid Genetic Algorithm.

## 1. Introduction

Tumbling mill, pulverizing the raw coal for the boiler, is one of the major assistant equipments in a thermal power plant and has high energy consumption, which equals 15-25% of the whole consumption of the thermal power plant [1]. Therefore, ensuring the tumbling mill to be at the best efficiency is of important theoretical significance and practical motivation for the energy saving.

Pulverizing capability is usually used for representing the efficiency of tumbling mill and the accurate measurement of pulverizing capability is the premise of realizing the control and optimization of tumbling mill. Some methods of measuring the concentration of pulverized coal are proposed for estimating the pulverizing capability [2, 3]. Because tumbling mill is a multi-variable, nonlinear and strong coupling system, for overcoming the lack of single signal measurement, the modeling approaches of pulverizing capability based on the multiple process variables of tumbling mill are presented [4,5]. Nevertheless, there may be some redundant or irrelevant variables for predicting the pulverizing capability. Hence, variable selection being a common approach for solving the problem [6,7].

Variable selection approach could find the most representative process variables subset for building the prediction model of pulverizing capability, so the model complexity would be reduced and the prediction accuracy would be improved. Moreover, variable selection could be seemed as a combinatorial problem, and various metaheuristics are proposed for the variable selection problem [8]. Genetic algorithm (GA) is an optimization search methodology based on the natural selection theory. Some applications of GA for variable selection are presented in [9-13]. However, for mainly depending on the initial population, the traditional GA implementations would easily have the premature convergence, namely, fall into the local minima [14,15].

In the paper, we propose a variable selection method for pulverizing capability prediction of tumbling mill based on improved hybrid genetic algorithm. The proposed method could determine the process variables subset for representing the pulverizing capability of the tumbling mill based on the improved hybrid genetic algorithm(IHGA). The pulverizing capability would be predicted based on the support vector

<sup>\*</sup> Corresponding author

machine (SVM) with the selected variables. The main idea of IHGA is that the traditional operation of GA, such as the selection operation, the crossover operation and the mutation operation, are hybridized with the mulit-population mechanism, the elites sharing mechanism and the heterogeneity mechanism to improve the algorithm effectiveness. The multi-population mechanism lets the sub-population search the whole space in parallel to improve the performance. The elites sharing mechanism exchanges the genetic information among the sub-populations to avoid the premature convergence. The heterogeneity mechanism could increase the diversity of the sub-population. The organization of this paper is as follows: Section 2 provides some preliminaries and the problem statement. The proposed method is presented in detail in Section 3. In Section 4, the experiments are presented to verify the effectiveness of the proposed method. Finally, Section 5 concludes the paper.

## 2. Preliminaries and Problem Statement

The schematic representation of a tumbling mill system is shown in Figure 1. The raw coal is fed into the tumbling mill and pulverized to the coal powder. At the same time, the hot air and the recycle air are blown into the mill to dry and deliver the coal powder. The coal powder is transferred into the coarse classifier and fine classifier. The unqualified powder is fed back into the ball mill for further pulverizing while the accepted powder is stored in the pulverized coal bunker finally. Furthermore, the mill exhauster provides the power for the flow process of the airpowder mixture and the recycle air is from the outlet of the mill exhauster.

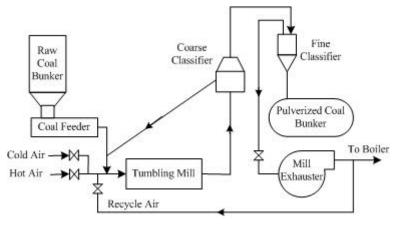


Figure 1. Tumbling mill system

The rotation speed of the tumbling mill and the mill wall are directly related to the pulverizing capability, and they are fixed by the manufacturer in the design process. Moreover, although the coal hardness would affect the pulverizing capability greatly, the coal hardness only could be calibrated by the chemistry experiment. Therefore, there are twelve process variables that could be used for building the prediction mode of pulverizing capability. The twelve process variables and their abbreviations are listed in the Table 1.

Table 1. The process variables and their abbreviations

| ID | The process variable                                     | The abbreviation |
|----|--|------------------|
| 1  | The tumbling mill load                                   | $L_m$            |
| 2  | The inlet negative pressure of the tumbling mill         | $P_i$            |
| 3  | The outlet negative pressure of the tumbling mill        | $P_0$            |
| 4  | The different inlet-outlet pressure of the tumbling mill | $P_d$            |
| 5  | The inlet temperature of the tumbling mill               | $T_i$            |
| 6  | The outlet temperature of the tumbling mill              | $T_0$            |
| 7  | The inlet negative pressure of the coarse classifier     | $P_c$            |
| 8  | The inlet negative pressure of the fine classifier       | $P_f$            |
| 9  | The inlet negative pressure of the mill exhauster        | $P_{e}$          |
| 10 | The motor current of the tumbling mill                   | $I_m$            |
| 11 | The motor current of the mill exhauster                  | $I_e$            |
| 12 | The ventilation rate                                     | F                |

The tumbling mill load is the ratio between the volume of coal powder in the mill and the interstitial volume of the static ball charge. The measuring point of the ventilation rate is in the vertical pipe of mill outlet. Because the tumbling mill system is nonlinear and strong coupling, some of the variables are redundant or weakly correlate with the pulverizing capability, namely, adopting all the variables to build the prediction model of the pulverizing capability would not only increase the model complexity but also decrease the prediction accuracy. Hence, for obtaining the accurate value of pulverizing capability, the variables used for the prediction model should be selected correctly.

#### 3. The Proposed Method

We present a variable selection method based on IHGA for building the prediction model of the pulverizing capability of the tumbling mill. Based on the traditional GA, IHGA adopts the multi-population mechanism, the elites sharing mechanism and the heterogeneity mechanism for improving the algorithm effectiveness. The flowchart of the IHGA is shown in Figure 2 and the steps of IHGA are described in the following. Moreover, we assume that the number of sub-populations is m and the size of each sub-population is n.

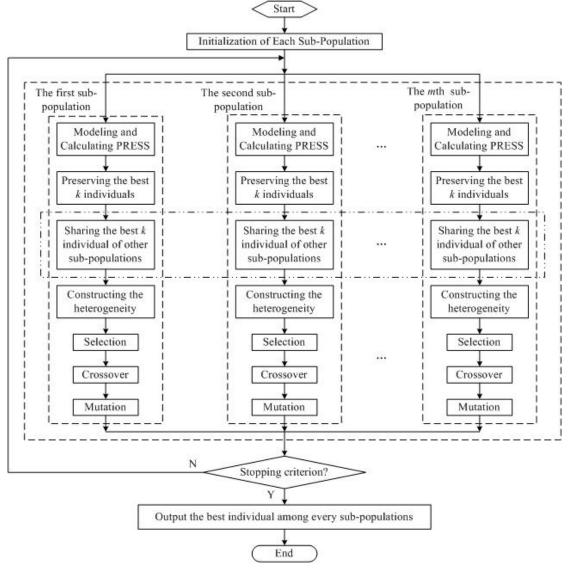


Figure 2. The flowchart of the IHGA

Step 1: Encoding and multi-population Initialization. For variable selection, IHGA uses the binary coding to encode the chromosomes and '1' or '0' indicates that the variable is selected or dropped, respectively. Each gene represents a process variable and the length of a chromosome equals the number of the variables to be selected, namely, for the tumbling mill system, the length of a chromosome is twelve.

We adopt the multi-population mechanism to improve the performance. The whole population is divided into several sub-populations and the evolution process of each sub-population is independent. Since each sub-population searches for the optimal solution

of the whole search space in parallel, the computational cost would be reduced. Nevertheless, for the same search space, the smaller population would lead to worse precision. For the proposed method, the size of sub-population is smaller than that of whole population, and we present the elites sharing mechanism to enhance the interaction between sub-populations for improving their cooperation for optimization. Moreover, all sub-populations are randomly initialized with every chromosome in each sub-population encoding a candidate variables set.

**Step 2: Fitness function.** For a chromosome in any sub-population, according to the coding of the chromosome, the selected variables are determined and the prediction model of the pulverizing capability is built by SVM, which is one of the most competitive regression analysis methods for small samples [16]. The fitness function is determined as the prediction residual error sum of squares (PRESS) of the built prediction model, namely, the lower the PRESS, the better the prediction model.

Step 3: Elites sharing. The elites of a sub-population are defined as the chromosomes whose PRESS is top-k, where the PRESS of each chromosome of the sub-population being in ascending order and k is an integer. For each sub-population, the elites of the sub-population and the elites of other sub-populations are put into the next generation of the sub-population directly without performing any operations. By the elites sharing mechanism, the genetic information of various sub-populations could be exchanged and the search scope of the sub-population is guided in the direction of the global optimum. Meanwhile, the sub-populations through continuous exchange of genetic information could avoid the limitation of the searching process.

**Step 4: Constructing heterogeneity.** We construct a chromosome to be farthest away from all chromosomes of a sub-population and the chromosome is named the heterogeneity for the sub-population. The heterogeneity is also put into the next generation of the sub-population. The heterogeneity mechanism would prevent the phenomenon of inbreeding, increase the diversity of the sub-population, and reduce the possibility of the sub-population finding the local optimum. In addition, for the heterogeneity mechanism, the Hamming distance is used as the distance measurement between two chromosomes.

**Step 5: Selection.** The selection operation determines which parent individuals would be put into the next generation. For the proposed method, we use the roulette wheel strategy to perform proportional selection for determining the  $(n - m \cdot k - 1)$  parent chromosomes to be put into the offspring, where  $n > m \cdot k$ .

**Step 6: Crossover.** The crossover operation combines two chromosomes to create a new chromosome, which may be better than both of the two chromosomes if the new chromosome inherits the best characteristics from each of the two chromosomes.

For IHGA, according to a user-definable crossover probability, the chromosome would exchange the genes with its next neighbor chromosome. Furthermore, the number of the exchanged genes is determined randomly.

**Step 7: Mutation.** The mutation is an operator which realizes the diversity of each sub-population. IHGA choices some genes of each chromosome of a sub-population to be flipped based on the user-definable probability of mutation.

**Step 8: Termination and output.** For IHGA, when the number of generation is larger than the user-definable threshold, the evolution of each sub-population terminates and the chromosome with the lowest fitness value among all sub-populations, namely, the whole population, is output.

According to the best individual obtained by the proposed method, we could determine the process variables for the prediction model of the pulverizing capability and the model built with the selected variables would have higher accuracy. Since we present the multi-population mechanism, the elites sharing mechanism and the heterogeneity mechanism, the proposed method could have faster convergence speed and avoid the premature convergence under a certain extent. In the next section, we would use the experiment results to further verify the effectiveness of the proposed algorithm.

## 4. Experiment Results

In this section, some experiments are presented to evaluate the effectiveness of the proposed algorithm. We focus on comparing IHGA with GA for selecting the process variables of the tumbling mill system. The SVM is adopted to build the prediction model of the pulverizing capability. IHGA and GA use the PRESS of the built prediction model as the fitness functions. For IHGA, we assume that the number of subpopulations is *m* and the size of each sub-population is *n*. To ensure the fairness of the experiment, the size of population of GA is  $n \cdot m$ , and the initialization conditions of IHGA and GA are kept consistent. Moreover, IHGA and GA also adopt the same crossover probability and the same mutation probability. Following the same value used in [13], the crossover probability and the mutation probability are set to be 0.8 and 0.2, respectively, and are not changed in the following experiments. Two field datasets are used in the experiments and they are obtained from the field database of QinLing Thermal Power Plant under different work conditions. The two field datasets both include fifteen samples and are shown in Table 2 and Table 3 respectively, where pc represents the pulverizing capability and the values of pc are measured based on the field experiments in the steady state. Experiments are performed in MATLAB 7.0.4 and the running environment is an Athlon64 X2 3600+ machine with 2 GB of RAM and running Windows XP Professional. In addition, the proposed method has been put into practice in QinLing Thermal Power Plant and the running curve of pulverizing capability would be shown.

| Table 2 | <b>2.</b> The | field | dataset | 1 |
|---------|---------------|-------|---------|---|
|         |               |       |         |   |

| ID | $L_m(\%)$  | $P_i(\text{Pa})$ | $P_0(\mathrm{Pa})$ | $P_d$ (Pa) | $T_i(^{\circ}\mathrm{C})$ | $T_0(^{\circ}\mathrm{C})$ | $P_c$ (Pa) |
|----|------------|------------------|--------------------|------------|---------------------------|---------------------------|------------|
| 1  | 48.76      | -704.95          | -3182.97           | 2520.54    | 195.9                     | 100.8                     | -4184.96   |
| 2  | 48.90      | -721.97          | -3196.35           | 2505.19    | 195.9                     | 100.8                     | -4188.38   |
| 3  | 49.07      | -600.54          | -3193.73           | 2527.47    | 195.9                     | 100.9                     | -4189.45   |
| 4  | 56.44      | -656.66          | -3368.04           | 2719.29    | 176.0                     | 92.6                      | -4548.5    |
| 5  | 57.51      | -635.56          | -3486.76           | 2878.13    | 174.8                     | 91.3                      | -4739.07   |
| 6  | 48.64      | -712.54          | -3181.78           | 2516.47    | 195.0                     | 100.1                     | -4207.47   |
| 7  | 48.54      | -616.85          | -3203.48           | 2529.38    | 195.9                     | 100.9                     | -4174.48   |
| 8  | 48.4       | -700.37          | -3202.31           | 2513.26    | 196.3                     | 101.2                     | -4197.76   |
| 9  | 48.51      | -569.09          | -3193.39           | 2529.04    | 195.9                     | 100.9                     | -4190.52   |
| 10 | 51.05      | -715.35          | -3403.53           | 2691.82    | 213.7                     | 85.4                      | -4532.85   |
| 11 | 48.81      | -804.68          | -3215.03           | 2539.71    | 196.0                     | 101.0                     | -4199.02   |
| 12 | 51.34      | -813.48          | -3353.22           | 2623.96    | 213.7                     | 86.8                      | -4483.04   |
| 13 | 52.00      | -771.35          | -3358.11           | 2612.38    | 213.7                     | 86.1                      | -4469.42   |
| 14 | 53.27      | -795.44          | -3430.40           | 2729.22    | 213.9                     | 84.8                      | -4576.07   |
| 15 | 57.51      | -635.56          | -3486.76           | 2878.13    | 174.8                     | 91.3                      | -4739.07   |
| ID | $P_f$ (Pa) | $P_e$ (Pa)       | $I_m$ (A)          | $I_e$ (A)  | F (Km3/h)                 | <i>pc</i> (ton/h)         | _          |
| 1  | -6310.99   | 3764.82          | 116.31             | 66.22      | 151.67                    | 21.46                     |            |
| 2  | -6320.28   | 3764.82          | 116.32             | 66.18      | 151.67                    | 23.46                     |            |
| 3  | -6316.35   | 2285.57          | 116.52             | 66.08      | 152.48                    | 29.82                     |            |
| 4  | -6727.65   | 3113.13          | 118.56             | 65.75      | 150.78                    | 48.49                     |            |
| 5  | -6932.61   | 2871.12          | 119.19             | 66.07      | 142.08                    | 51.65                     |            |
| 6  | -6340.08   | 2832.74          | 116.45             | 66.09      | 150.21                    | 19.46                     |            |
| 7  | -6301.29   | 2341.67          | 116.36             | 65.87      | 152.64                    | 32.64                     |            |
| 8  | -6332.67   | 2227.59          | 116.21             | 65.77      | 155.85                    | 36.18                     |            |
| 9  | -6319.09   | 2276.46          | 116.16             | 66.16      | 152.26                    | 26.21                     |            |
| 10 | -6818.75   | 1361.86          | 116.1              | 66.58      | 156.29                    | 41.97                     |            |
| 11 | -6323.55   | 3469.58          | 116.48             | 65.97      | 152.76                    | 35.12                     |            |
| 12 | -6742.18   | 1663.85          | 115.48             | 66.17      | 159.18                    | 38.59                     |            |
| 13 | -6733.18   | 1433.58          | 116.23             | 66.24      | 156.56                    | 40.24                     |            |
| 14 | -6850.81   | 1528.54          | 116.39             | 66.36      | 156.46                    | 43.08                     |            |
| 15 | -6932.61   | 2871.12          | 119.19             | 66.07      | 142.08                    | 51.65                     |            |

For IHGA, m, n and k are set to be 3, 10 and 1, respectively. The population number of GA is set to be 30 and the number of iterations is 30. The experiment results of dataset 1 are shown in Figure 3, where 30.72 is the PRESS value of the prediction model with all twelve variables for dataset 1. After the searching process is finished, the PRESS value of IHGA is 17.92 and the corresponding chromosome is 111101110111, namely,  $T_i$  and  $P_e$  are weakly related with the pulverizing capability. Moreover, the PRESS value of GA is 28.51 and the corresponding chromosome is 111110111111, namely,  $T_0$  is irrelevant variable for pulverizing capability. However, according to the knowledge of experts,  $T_0$  would represent the drying capability of tumbling system, namely, if the outlet temperature is lower, the drying may not be sufficient with the coal feed increasing and the pulverizing capability would be decreased [17]. Hence, GA may fall into the local minima, and IHGA has higher effectiveness for dataset 1.

The experiment results of dataset 2 are shown in Figure 4, where 3.66 is the PRESS value of the prediction model with all twelve variables for dataset 2. For IHGA, the PRESS value is 3.63 and the corresponding chromosome is 101111110111, namely,  $P_i$  and  $P_e$  are not be selected under the work condition. Although  $P_i$  affects the delivery of coal powder, the distribution of  $P_i$  values in dataset 2 is more concentrated, namely, under the work condition,  $P_i$  may contribute less for the pulverizing capability and determining  $P_i$  to be a redundant variable would be reasonable since other variables, for example, F, could represent the air draft capability under a certain extent. Furthermore, for GA, the PRESS value is 3.77, which is lager than 3.66, and the corresponding chromosome is 001111110111, namely,  $L_m$ ,  $P_i$  and  $P_e$  could not be used for modeling. However,  $L_m$  represents the coal storage of the tumbling mill. If the coal storage is lower, most of the energy would be wasted for collisions among the steel balls and the pulverizing

capability would become lower. If the coal storage is larger, the falling height of steel balls would be lower and the pulverizing capability would be lower also. Therefore,  $L_m$  should not be seemed as the irrelevant variable for pulverizing capability and GA may fall into the local minima for dataset 2. Since the hetero-

**Table 3.** The field dataset 2

geneity mechanism would let the searching scope of IHGA be larger sometimes, the PRESS values of IHGA are still relatively larger before the third iteration, and IHGA begins convergence after the 4th iteration. The results of the experiments show that IHGA has faster convergence speed and higher effectiveness.

| ID             | $L_m(\%)$                      | $P_i(\text{Pa})$              | $P_0$ (Pa)                 | $P_d$ (Pa)              | $T_i(^{\circ}\mathrm{C})$  | $T_0(^{\circ}\mathrm{C})$ | $P_c(\operatorname{Pa})$ |
|----------------|--------------------------------|-------------------------------|----------------------------|-------------------------|----------------------------|---------------------------|--------------------------|
| 1              | 52.64                          | -735.40                       | -3379.19                   | 2670.01                 | 213.8                      | 85.2                      | -4524.53                 |
| 2              | 52.53                          | -790.26                       | -3466.4                    | 2757.42                 | 213.9                      | 84.6                      | -4604.41                 |
| 3              | 55.50                          | -705.34                       | -3387.48                   | 2686.55                 | 214.0                      | 84.5                      | -4524.87                 |
| 4              | 55.71                          | -667.98                       | -3364.94                   | 2725.12                 | 176.4                      | 93.2                      | -4532.72                 |
| 5              | 56.76                          | -663.42                       | -3402.94                   | 2763.47                 | 176.2                      | 93.0                      | -4571.83                 |
| 6              | 55.71                          | -665.22                       | -3348.77                   | 2715.77                 | 176.1                      | 92.7                      | -4542.53                 |
| 7              | 56.17                          | -660.64                       | -3503.63                   | 2850.09                 | 176.0                      | 92.7                      | -4575.11                 |
| 8              | 56.64                          | -645.01                       | -3395.21                   | 2779.42                 | 175.9                      | 92.5                      | -4573.02                 |
| 9              | 57.67                          | -650.53                       | -3416.08                   | 2794.83                 | 175.6                      | 92.2                      | -4620.68                 |
| 10             | 57.84                          | -647.62                       | -3434.46                   | 2824.31                 | 175.0                      | 91.6                      | -4676.26                 |
| 11             | 59.13                          | -610.80                       | -3451.84                   | 2871.01                 | 174.1                      | 90.3                      | -4732.93                 |
| 12             | 63.45                          | -699.00                       | -3552.72                   | 2838.16                 | 159.6                      | 97.7                      | -4738.09                 |
| 13             | 62.14                          | -697.71                       | -3571.05                   | 2860.30                 | 159.7                      | 97.3                      | -4805.25                 |
| 14             | 61.84                          | -702.46                       | -3578.46                   | 2865.58                 | 159.6                      | 96.8                      | -4837.25                 |
| 15             | 64.68                          | -670.31                       | -3608.38                   | 2915.23                 | 159.0                      | 95.0                      | -4917.03                 |
| ID             | $P_f$ (Pa)                     | $P_e$ (Pa)                    | $I_m$ (A)                  | $I_e$ (A)               | F (Km3/h)                  | <i>pc</i> (ton/h)         | _                        |
| 1              | -6805.88                       | 1794.11                       | 116.45                     | 66.59                   | 155.17                     | 42.29                     |                          |
| 2              | -6891.13                       | 2609.04                       | 116.61                     | 66.75                   | 156.48                     | 43.47                     |                          |
| 3              | -6711.81                       | 3451.81                       | 117.55                     | 67.31                   | 142.95                     | 44.44                     |                          |
| 4              | -6704.66                       | 3179.77                       | 118.32                     | 66.04                   | 152.01                     | 45.15                     |                          |
| 5              | -6739.28                       | 3109.62                       | 118.55                     | 66.32                   | 151.40                     | 46.72                     |                          |
| 6              | -6703.17                       | 1741.48                       | 118.35                     | 65.87                   | 150.18                     | 47.69                     |                          |
| 7              | -6734.43                       | 2903.89                       | 118.72                     | 65.77                   | 150.73                     | 48.03                     |                          |
| 8              | -6737.30                       | 2242.65                       | 118.40                     | 65.55                   | 149.77                     | 49.18                     |                          |
| 9              | (00( 01                        | 2202.06                       | 110.01                     | (5.05                   | 149.90                     | 50.23                     |                          |
| ,              | -6806.24                       | 3382.86                       | 118.81                     | 65.95                   | 149.90                     | 50.25                     |                          |
| 10             | -6806.24<br>-6868.42           | 2106.70                       | 118.81                     | 65.95<br>65.91          | 149.90                     | 51.17                     |                          |
|                |                                |                               |                            |                         |                            |                           |                          |
| 10             | -6868.42                       | 2106.70                       | 119.34                     | 65.91                   | 140.50                     | 51.17                     |                          |
| 10<br>11       | -6868.42<br>-6931.5            | 2106.70<br>2805.90            | 119.34<br>119.63           | 65.91<br>66.21          | 140.50<br>141.61           | 51.17<br>52.19            |                          |
| 10<br>11<br>12 | -6868.42<br>-6931.5<br>-6876.9 | 2106.70<br>2805.90<br>3762.29 | 119.34<br>119.63<br>120.93 | 65.91<br>66.21<br>64.69 | 140.50<br>141.61<br>145.34 | 51.17<br>52.19<br>53.41   |                          |

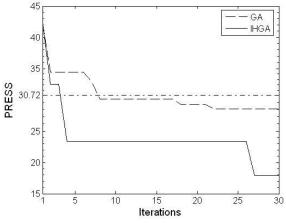


Figure 3. The experiment results of dataset 1

For IHGA, *m*, *n* and *k* are changed to be 5, 20 and 3, respectively. The size of the population of GA is set to be 100 and the number of iterations is still 30. The experiment results of dataset 1 and dataset 2 are shown in Figure 5 and Figure 6, respectively. Since the parameters of IHGA and GA are changed, the iteration processes of IHGA and GA are different from what are shown in Figure 3 and Figure 4, respectively. For the field dataset 1, the final result of IHGA is 17.92 and the corresponding chromosome is 111101110111. The final result of GA is 28.51 and the corresponding chromosome is 111110111111. Moreover, for the field dataset 2, the final result of IHGA is 3.63 and the corresponding chromosome is 101111110111. The final result of GA is 3.77 and the corresponding chromosome is 001111110111. The IHGA also has faster convergence speed and higher effectiveness. Therefore, the searching results of IHGA are not affected by the parameters setting greatly, namely, determining the parameters of IHGA would not be difficult.

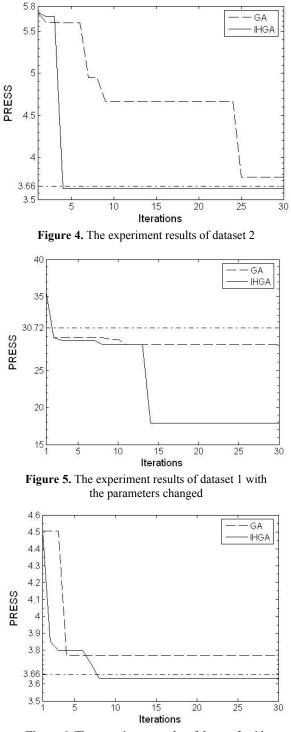


Figure 6. The experiment results of dataset 2 with the parameters changed

The results of the experiments verify that the proposed method not only could realize the variables selection correctly but also has higher effectiveness. Moreover, the proposed method has been put into practice in QinLing Thermal Power Plant. The running curve of the prediction value of pulverizing capability for half an hour is shown in Figure 7. During 1200~1400sec., since the prediction value of the pulverizing capability is almost not changed, the tumbling mill system could be seemed to be in the steady state, which is marked with a dotted rectangle in Figure 7. Although the real value of the pulverizing capability could not be measured directly, it is usually represented by the quantity of pulverized coal powder per unit of time and the tumbling mill system fulfills the rules of indestructibility of matter for steady-state operation, namely, the coal feed per unit of time approximately equals the quantity of pulverized coal powder per unit of time in the steady state. During 1200~1400sec., the coal feed per unit of time is 44.6ton/h and the average of the prediction values of the pulverizing capability is 44.46ton/h. Hence, it shows that the pulverizing capability could be predicted successfully.

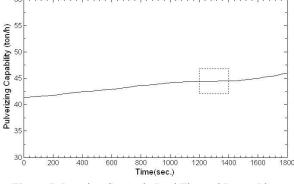


Figure 7. Running Curves in Real Thermal Power Plant

### 5. Conclusion

In the paper, we propose a variable selection method for pulverizing capability prediction of tumbling mill based on an improved hybrid genetic algorithm. The proposed method has some advantages as follows. First, the prediction model of pulverizing capability built based on the variables selected by the proposed method is more accurate. Second, the improved hybrid genetic algorithm has faster convergence speed and avoids the premature convergence under a certain extent. Third, the proposed method adopts the multipopulation mechanism for searching whole space in parallel to ensure the algorithm performance. Fourth, the proposed method uses the elites sharing mechanism to ensure the accuracy of the result. Fifth, the proposed method adopts the heterogeneity mechanism to ensure the diversity of population. Sixth, the parameters setting of the proposed method is not difficult. The results of the experiments also verify the effectiveness of the proposed method. Our method has been put into practice successfully. Moreover, the proposed method could be applied in other industrial processes including tumbling mill, for instance, the grinding circuit of mineral process. Since the performance of the proposed method may be affected by the initial population, in the future research work, we

will use some advanced implementation schemes to further improve the effectiveness of the proposed method.

#### Acknowledgement

This work is supported by the National Natural Science Foundation of China (61005058). The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

## References

- [1] Heng Wang, Min-ping Jia, Peng Huang, Zuo-liang Chen. A study on a new algorithm to optimize ball mill system based on modeling and GA. *Energy Con*version and Management, Vol.51, No.4, 2010, 846-850.
- [2] Quansheng Duan, Jizhen Liu, Zhifang Wu. Design and experimental study on the pulverized coal concentration sensor based on γ-ray absorption method. *Proceedings of* 2008 IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application (PACIIA 2008), Wuhan, China, December 19-20, 2008, 901-906.
- [3] Chen, Xia, Hu, Hongli, Liu, Zhihong. A pulverized coal concentration measurement system based on capacitance sensor. *Proceedings of* 2009 *International Technology and Innovation Conference* (ITIC 2009), *Xi'an, China, October* 12-14, 2009, 1-4.
- [4] Su Zhi-gang, Wang Pei-hong, Yu Xiang-jun, Lü Zhen-zhong. Soft sensor modeling for on-line monitoring the capacity of coal pulverizing system. *Proceedings of the CSEE, Vol.*27, *No.*29, 2007, 90-95.
- [5] Hao Yongsheng, Yu Xiangjun, Zhao Gang, Lü Zhenzhong. Optimization for ball mill operation based on improved particle swarm optimization algorithm. *Journal of Southeast University (Natural Science Edition)*, Vol.38, No.3, 2008, 419-423.
- [6] Vita Špečkauskienė, Arūnas Lukoševičius. A data mining methodology with preprocessing steps. *Infor*mation Technology and Control, Vol.38, No.4, 2009, 319-324.
- [7] A. Unler, A. Murat. A discrete particle swarm optimization method for feature selection in binary classification problems. *European Journal of Operational Research, Vol.*206, *No.*3, 2010, 528-539.

- [8] Yumin Chen, Duoqian Miao, Ruizhi Wang. A rough set approach to feature selection based on ant colony optimization. *Pattern Recognition Letters*, *Vol.*31, *No.*3, 2010, 226-233.
- [9] Ran Li, Jianjiang Lu, Yafei Zhang, Tianzhong Zhao. Dynamic Adaboost learning with feature selection based on parallel genetic algorithm for image annotation. *Knowledge-Based Systems*, Vol.23, No.3, 2010, 195-201.
- [10] Xinping Song, Yongsheng Ding, Jingwen Huang, Yan Ge. Feature selection for support vector machine in financial crisis prediction-a case study in China. *Expert Systems, Vol.*27, *No.*4, 2010, 299-310.
- [11] A.L.I. Oliveira, P.L. Braga, R.M.F. Lima, M.L. Cornélio. GA-based method for feature selection and parameters optimization for machine learning regression applied to software effort estimation. *Information* and Software Technology, Vol.52, No.11, 2010, 1155-1166.
- [12] M. Bhattacharya, A. Das. Genetic algorithm based feature selection in a recognition scheme using adaptive neuro fuzzy techniques. *Int. J. of Computers, Communications & Control, Vol.5, No.4, 2010, 458-*468.
- [13] Shijin Li, Hao Wu, Dingsheng Wan, Jiali Zhu. An effective feature selection method for hyperspectral image classification based on genetic algorithm and support vector machine. *Knowledge-Based Systems, Vol.*24, *No.*1, 2010, 40-48.
- [14] Yongming Li, Sujuan Zhang, Xiaoping Zeng. Research of multi-population agent genetic algorithm for feature selection. *Expert Systems with Applications, Vol.*36, *No.*9, 2009, 11570-11581
- [15] A.M. Kuczapski, M.V. Micea, L.A. Maniu, V.I. Cretu. Efficient generation of near optimal initial populations to enhance genetic algorithms for jobshop scheduling. *Information Technology and Control*, *Vol.*39, *No.*1, 2010, 32-37.
- [16] J. Han, M. Kamber. Data mining, Concepts and Techniques (Secend Edition). *Morgan Kaufmann, San Francisco, USA*, 2006.
- [17] Tianyou Chai, Heng Yue. Multivariable intelligent decoupling control system and its application, *Acta Automatica Sinica*, *Vol.31*, *No.1*, 2005, 123-131.

Received June 2011.