Coal mill fault diagnosis based on Gaussian process regression

Longfei Zhu¹, Shuangbai Liu¹, Deli Zhang¹, Xiaozhi Qiu² and Weiqing Zhou³

¹State Grid Jibei Electric Power Co. Ltd. Research Institute, North China Electric Power Research Institute Co. Ltd., Beijing 100045, China
²Beijing Jingneng Power Co. Ltd., Beijing 100025, China
³Nanjing Institute of Technology School of Energy and Power Engineering, Nanjing 211167, China
wqzhou@njit.edu.cn

Abstract. A typical operating set of equipment can be obtained through cluster analysis of historical data. Two state monitoring models for HP medium speed coal mill are established based on Gaussian process regression and the similarity index calculated by this model can be used for measuring the operating status of HP mills. Finally a method for fault diagnosis of HP mill based on Gaussian regression modelling is proposed combined with fault diagnosis knowledge base of this HP mill. Taking the HP medium speed mill of a 660MW thermal power unit as an example, the real operating data is collected and used for modelling and analysis. Results shows that the equipment parameter estimation calculated by Gaussian process regression is accurate. It can be used for early-warning and diagnosed of equipment fault and also for practical engineering application.

1. Introduction
As the key link of entire thermal power unit the milling system takes role of providing fuel for boiler combustion and its operating status is related to the safe and economic operation of the unit. According to statistics on the causes of boiler failures in domestic units it reveals that the coal mill failure is one of the main reasons [1]. However, because of the complex environment where milling system of thermal power plant running the coal mill failures often show greater uncertainty and it is difficult to achieve the desired results by using traditional fault diagnosis methods.

With the rapid increase of automation level of modern power plant units diagnostic methods based on data modeling are developing rapidly and it is applied in fault diagnosis of coal powder system in thermal power plant. Literature [2] uses radial basis neural network for fault diagnosis of HP bowl medium speed mill based on historical fault data. Literature [3] proposes a fusion method of fuzzy clustering and D-S evidence theory for fault diagnosis of coal mill. Literature [4] proposes a grey correlation method based on D-S combination rules which is used for early failure diagnosis of coal mill. Recent scientific reports on the monitoring and diagnosis of coal mill can also be found in the literature [5-8].

This paper establishes a Gaussian process regression estimation model for HP medium speed coal mill based on principle of maximum likelihood estimation. The fault diagnosis method is proposed combined with fault expert knowledge database and fault diagnosis of coal mill is achieved.
2. Gaussian regression method

2.1 Gaussian regression method

Gaussian regression decomposes events into several discrete regression models based on Gaussian probability density functions. According to probability and statistics theory the triggering of an event can be approximated by multiple Gaussian distribution functions and parameters of the functions are generally solved by maximum likelihood estimation \(^9\).

Suppose \(x \in \mathbb{R}^k\) is historical data from a multi-case process and its probability density function can be expressed in Gaussian function form:

\[
p(x|\mu, \Sigma) = \sum_{k=1}^{K} \omega_k g(x|\mu_k, \Sigma_k)
\]  

(1)

In formula (1): \(K\) is the total number of observation vectors, \(\omega_k\) is the weight of the \(k\)-th observation vector, \(\mu_k\) and \(\Sigma_k\) are mean and covariance of the local Gaussian model. \(g(x|\mu_k, \Sigma_k)\) is the multivariate Gaussian density function for the \(k\)-th observation vector and its calculation formula is:

\[
g(x|\mu_k, \Sigma_k) = \frac{1}{(2\pi)^{L/2} |\Sigma_k|^{1/2}} \exp \left[ -\frac{1}{2} (x - \mu_k)^T \Sigma_k^{-1} (x - \mu_k) \right]
\]  

(2)

The parameters \(\Theta=\{\{\omega_1, \mu_1, \Sigma_1\}, \cdots, \{\omega_K, \mu_K, \Sigma_K\}\}\) in equation (2) are generally determined by the expected maximum algorithm (EM). Given training data \(X = \{x_1, x_2, \cdots, x_n\}\), mixed component number \(K\) and initial value of parameters \(\Theta=\{\{\omega_1^{(0)}, \mu_1^{(0)}, \Sigma_1^{(0)}\}, \cdots, \{\omega_K^{(0)}, \mu_K^{(0)}, \Sigma_K^{(0)}\}\}\) the EM algorithm updates parameters through two iterative steps which objective is to make likelihood of the observed data increasing monotonically.

The iterative steps of the EM algorithm are divided into two major steps which includes E-step and M-step and the calculation formula is:

1. E-step

\[
p^{(s)}(C_k|x_i) = \frac{\omega_k^{(s)} g(x_i|\mu_k^{(s)}, \Sigma_k^{(s)})}{\sum_{j=1}^{K} \omega_j^{(s)} g(x_i|\mu_j^{(s)}, \Sigma_j^{(s)})}
\]  

(3)

In equation (3) \(p^{(s)}(C_k|x_i)\) is the posterior probability that the \(i\)-th training sample belongs to the \(k\)-th Gaussian component after the \(s\)-th iteration.

2. M-step

\[
\mu_k^{(s+1)} = \frac{\sum_{i=1}^{N} p^{(s)}(C_k|x_i) x_i}{\sum_{i=1}^{N} p^{(s)}(C_k|x_i)}
\]  

(4)

\[
\Sigma_k^{(s+1)} = \frac{\sum_{i=1}^{N} p^{(s)}(C_k|x_i) (x_i - \mu_k^{(s+1)})(x_i - \mu_k^{(s+1)})^T}{\sum_{i=1}^{N} p^{(s)}(C_k|x_i)}
\]  

(5)
\[ \omega_k^{(s+1)} = \frac{\sum_{i=1}^{N} p^{(s)}(C_k| x_i)}{N} \]

In equation (4)-(6) \( \mu_k^{(s+1)} \), \( \Sigma_k^{(s+1)} \) and \( \omega_k^{(s+1)} \) are the mean, covariance and prior probability of the \( k \)-th Gaussian component after the \( s \)-th iteration.

2.2 Gaussian regression modeling
Gaussian process modeling is a stochastic process modeling and its statistical properties is determined by mean and covariance. The mathematical description of Gaussian process is:

\[
f(x) \sim GP\left(m(x), k(x,x')\right)
\]

The GP in equation (7) stands for Gaussian process while actual process should include noise interference and the Gaussian regression model of the actual process is:

\[
y = f(x) + \epsilon
\]

In equation (8) \( \epsilon \) is the independent white noise which obeys a Gaussian distribution with a mean of 0 and a variance of \( \sigma_n^2 \). Suppose \( X = (x_1, \ldots, x_N) \) and \( y = (y_1, \ldots, y_N) \) are \( N \) observation samples of the Gaussian process and the distribution of \( y \) can be obtained as:

\[
y \sim GP(m(x), k(x,x') + \sigma_n^2 \delta_{ij})
\]

\[
\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}
\]

In equation (10) \( \delta_{ij} \) is Kronecker function.

3. Monitoring model based on Gaussian kernel regression

3.1 Model parameter
Taking a domestic 660MW unit milling system as an example which equipped with bowl type medium speed mill (HP type mill). According to the working principle and characteristics of the coal mill two parameter monitoring models are established including coal mill mechanical model and coal mill performance model. The parameters of these two models are shown in Table 1.

| Mechanical model                                      | Performance model                                      |
|------------------------------------------------------|-------------------------------------------------------|
| Unit load                                            | Unit load                                              |
| Coal mill motor current                              | Coal feeder total coal                                 |
| Coal mill motor winding temperature A                | Coal mill motor current                                |
| Coal mill motor winding temperature B                | Differential pressure between seal air and primary air |
| Coal mill motor winding temperature C                | Cold air adjustment valve position feedback            |
| Coal mill motor winding temperature D                | Hot air adjustment valve position feedback             |
| Coal mill motor winding temperature E                | Primary air volume                                     |
| Coal mill motor winding temperature F                | Export powder pressure                                 |
| Motor non-drive end bearing temperature               | Primary air pressure                                   |
| Motor drive end bearing temperature                   | Export powder temperature                             |
| Coal mill lubricating oil temperature                | Primary air temperature                                |
| Coal mill gearbox box temperature                     | Seal air pressure                                      |
| Coal mill A lubricating oil pressure                 |                                                        |

Taking the mechanical model of coal mill as example and historical data of the model parameters is collected. A total of 72005 groups of data in consecutive time periods is obtained and the collection
period is 30S that means data collection once in 30 seconds. After elimination working conditions of unit and coal mill shut down the final number of historical conditions is 61493 that means the historical data used by the mechanical model is a matrix of 61493× 16.

Cluster analysis of the 61493 sets of operating conditions collected are then carried out by using density-based clustering and a working condition library including 234 sets of state vectors representing the typical operating state of the coal mill is obtained which constitute the typical working conditions of the coal mill under normal operating conditions.

3.2 Coal mill Gaussian regression model estimation
A total of 2000 sets of actual running data for a continuous period of time is collected as a test for the Gaussian regression model and the sampling period is 30S. Gaussian regression is used for parameter estimation and the results obtained are shown in Figure 1.

![Figure 1: Model estimation and real value](image)

Figure 1 shows the measured values and the model estimation of three measuring points including the coal mill current, the lubricating oil temperature and the non-driven end bearing temperature. It can be seen that the parameter estimation of Gaussian regression model has a high agreement with the actual measured values which means the model has a good predictive effect.

3.3 Coal mill fault diagnosis
In order to verify the fault monitoring capability of the model the last 1000 sets of data are used to simulate the failure of the coal mill lubricating oil cooling system by superimposing fault noise. The related parameter similarity trend changes as shown in Figure 2.

![Figure 2: Fault diagnosis by model similarity index](image)

(a) Similarity index of lubricating oil temperature
(b) Similarity index of gearbox box temperature

Figure 2 Fault diagnosis by model similarity index
As shown from Figure 2 the predicted similarity trends of the two measuring points have dropped significantly after the 1000-th sample. It shows that the operating status of the coal mill equipment does not match any state vector of the normal working conditions which means that some abnormality may have occurred inside the equipment.

According to parameter similarity monitoring based on Gaussian kernel regression model the measuring points where the obvious trend changes mainly include coal mill A motor current, non-drive end bearing temperature of coal mill motor, coal mill lubricating oil temperature and so on. These fault parameters with significant changes are matched to the symptoms in the fault knowledge base according to coal mill fault diagnosis expert system knowledge base and it is finally determined that coal mill lubricating oil cooling system failure has occurred. As can be seen the diagnostic result is consistent with the simulation which shows this model can be used for fault diagnosis.

4. Conclusion
This paper establishes a Gaussian regression estimation model for medium speed coal mill with that operation condition monitoring of medium speed coal mill can be realized since it has high estimation ability. Then medium speed coal mill fault diagnosis method based on Gaussian regression is proposed by combined with fault diagnosis knowledge base. Take a domestic coal mill of 660MW unit as an example the real operating data is collected and used for modelling and calculation and results show that the proposed method is more accurate for estimating equipment parameters and can be used for equipment fault diagnosis. The overall process of the proposed model is clear and intuitive and it is suitable for engineering application.

Acknowledgment
This work was supported by Science and Technology Project of SGCC (52018K170028).

References
[1] ZHANG, J.Z. (2007) Analysis and Prediction of the Jam of ZGM Medium-speed Coal Pulverizer. Journal of Engineering for Thermal Energy and Power, 27: 56-61.
[2] MA, Y.G., YIN, Z., MA, L.Y. (2011) Fault Diagnosis of Medium-speed Mills Based on RBF-Neural Network. Power Equipment, 34: 5-8.
[3] LU, X.Y., ZHAO, Z. (2011) Fault Diagnosis of Mill Based on Fuzzy Clustering Analysis and D-S Evidence Theory. Electric Power Science and Engineering, 27: 41-44.
[4] ZENG, D.L., CUI, Z.P., TIAN, L., ZHAO, Z. (2007) Fault Analysis of Coal Mills by Using Gray Correlation and D-S Combination Rules. Journal of Power Engineering, 27: 207-210.
[5] LIU, D.P., YE, X.R., CHEN, B.Y., TANG, M.Y. (2009) Fault Diagnosis of Medium Speed Mill Based on KPCA and LSSVM. Journal of Power Engineering, 29: 155-158.
[6] GAO, K., CHANG, L., CHEN, Z.G., ZHAO, J.T. (2011) Analysis of Faults in Operation of MPS Type Medium Speed Coal Pulverizers and Retrofit of its Control Functions. Thermal Power Generation, 40: 73-77.
[7] JIAO, S.M. (2018) Health Warning and Fault Diagnosis of Pulverizing System Based on Syncretic Similarity. Journal of System Simulation, 30: 595-604.
[8] REN, M.Y., JIAO, S.M. (2015) Coal milling system fault diagnosis based on improved NSET. Thermal Power Generation, 44: 87-92.
[9] YU, J., QIN, S.J. (2008) Multimode process monitoring with Bayesian inference-based finite Gaussian mixture models. AIChE Journal, 54: 1811-1829.
[10] GUO, P., WANG, Z.G. (2018) Wind turbine spindle state monitoring based on Gaussian process regression and double moving window residual processing. Electric Power Automation Equipment , 38: 34-40.