Journal of University of Science and Technology Beijing
Volume 12, Number 2, April 2005, Page 123

Software sensor for slab reheating furnace

Zhihua Xiong, Guohong Huang, and Huihe Shao

Institute of Automation, Shanghai Jiaotong University, Shanghai 200030, China
(Received 2004-06-08)

Abstract: It has long been thought that a reheating furnace, with its inherent measurement difficulties and complex dynamics, posed almost insurmountable problems to engineers in steel plants. A novel software sensor is proposed to make more effective use of those measurements that are already available, which has great importance both to slab quality and energy saving. The proposed method is based on the mixtures of Gaussian processes (GP) with the expectation maximization (EM) algorithm employed for parameter estimation of the mixture of models. The mixture model can alleviate the computational complexity of GP and also accords with the changes of operating condition in practical processes. It is demonstrated by on-line estimation of the furnace gas temperature in 1580 reheating furnace in Baosteel Corporation (Group).

Key words: Gaussian processes; expectation maximization; multiple models; soft sensor; reheating furnace

[This work was financially supported by the National High-Tech Research and Development Program of China (No.2002AA4120I0).]

1 Introduction

Walking-beam reheating furnace is an important device with lots of energy consumption in steel plants. During the past decade, people have made many efforts on it [1-5]. In order to improve the slab quality and reduce fuel consumption, Baosteel spent a lot of research in developing a new model for on-line optimization and control of reheating processes. Since the present furnace model, based on a complete physical description of the furnace characteristics, is too complicated for use, software sensors [6], also known for soft sensors, provide a convenient solution to meet the demand. And the key problem of soft sensors is modeling, which has different philosophy from the first principle model.

Most conventional industrial process models are global, where the industrial process is assumed to be fully characterized by a single model. However, because of multiple variables, seriously nonlinear and multiple work modes, a large number of real industrial processes are too complex to be described by a single model. Furthermore, the single model method not only reduces the estimate precision, but also increases the computational complexity [7].

Multiple model (MM) approaches [8-10] to the empirical modeling of nonlinear systems have been of interest for many years, and have been used widely in the last few years. The mixtures of Gaussian processes (GP) [11-14] have appeared in various forms.

In this paper, a novel soft sensor method is proposed on the basis of mixtures of Gaussian processes with expectation maximization algorithm employed for parameter estimation of the mixture of models.

2 Description of the reheating furnace

The structure of the walking-beam reheating furnace discussed in the following sections is shown in **figure 1.** Slabs in the furnace move from the tail zone to the soaking zone. And the flow direction of the waste gas is reverse. The control area of the furnace is divided into six zones, which are denoted from zone 1 to zone 6, respectively. The tail zone is not a control area and has no fuel input.

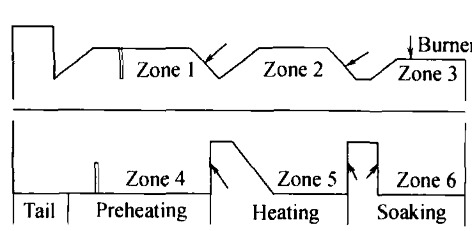

Figure 1 Structure of the walking beam reheating furnace.

Corresponding author: Zhihua Xiong, **E-mail:** zhxiong@sjtu.edu.cn

3 Mixture of GP based soft sensor

In short, multiple models involve several local models and the output of the global model is a combination of local models using an interpolation technique.

Constructing a soft sensor can be described as follows. Give the sample $\{x_k, y_k\}_{k=1}^N$, where N is the size of data samples, $x_k \in \mathbf{R}^n$ are the inputs of the soft sensor model, and $y_k \in \mathbf{R}$ is the corresponding desired targets. Suppose that the input vector for a test case is denoted x and the targets are scalar. From this training set we wish to learn a model to make accurate predictions of $\hat{y}$ for previously unseen x values.

The operating regime based model [8] of the system is formulated as

$$\hat{y} = \sum_{i=1}^{c} \phi_i(x)(a_i^T x + b_i) \quad (1)$$

where $\phi_i(x)$ is the validity function for the i-th operating regime and $\theta_i = [a^T_i\ b_i]^T$ the parameter vector of the corresponding local linear model.

The available data samples are collected in matrix Z formed by concatenating the input data matrix X and the output vector y:

$$X = \begin{bmatrix} x_1^T \\ x_2^T \\ \vdots \\ x_N^T \end{bmatrix}, \quad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}, \quad Z^T = [X\, y].$$

Each observation thus is an $n+1$ dimensional column vector $z_k = [x_{1,k}, \cdots, x_{n,k}, y_k]^T = [x_k^T\ y_k]^T$. Through clustering, the data set Z is partitioned into c clusters. The result is a fuzzy partition matrix $U=[\mu_{i,k}]_{c\times N}$, whose element $\mu_{i,k}$ represents the degree of membership of the observation z_k in cluster i.

Denote $P(\eta_i)$ the unconditional cluster probability (normalized such that $\sum_{i=1}^{c} p(\eta_i) = 1$), given by the fraction of the data that it explains. $p(z|\eta_i)$ is the domain of influence of the cluster, and will be taken to be the multivariate Gaussian function $N(v_i, F_i)$ in terms of a mean v_i and covariance matrix F_i.

$$p(z \mid \eta) = \sum_{i=1}^{c} p(z, \eta_i) = \sum_{i=1}^{c} p(z \mid \eta_i) p(\eta_i) \quad (2)$$

where the $P(z|\eta_i)$ distribution generated by the i-th cluster is represented by the Gaussian function

$$p(z \mid \eta_i) = \frac{1}{(2\pi)^{\frac{n+1}{2}} \sqrt{|F_i|}} \exp(-\frac{1}{2}(z - v_i)^T (F_i)^{-1} (z - v_i)) \quad (3)$$

In this paper, we propose to use the Gath-Geva (GG) clustering [15-18] algorithm instead of the widely used Gustafson-Kessel (GK) method, because with the GG method, the parameters of the univariate membership functions can directly be derived from the parameters of the clusters. Through GG clustering, the $p(z) = p(x, y)$ joint density of the response variable y and the regressors x is modeled as a mixture of c multivariate $n + 1$ dimensional Gaussian functions. And the conditional density $p(y|x)$ is also a mixture of Gaussian models. Therefore, the prediction of y on unseen value x can be formulated:

$$\hat{y} = E(y \mid x) = \int y p(y \mid x) dy = \frac{\int y p(y, x) dy}{p(x)} = \sum_{i=1}^{c} \frac{[[x^T\ 1]\theta_i] p(x \mid \eta_i) p(\eta_i)}{p(x)} = \sum_{i=1}^{c} p(\eta_i \mid x)[[x^T\ 1]\theta_i] \quad (4)$$

Here, θ_i is the parameter vector of local models and $p(\eta_i|x)$ is the probability that the i-th Gaussian component is generated by the regression vector x:

$$p(\eta|x) = \frac{\frac{p(\eta_i)}{(2\pi)^{n/2}\sqrt{|F_i^{xx}|}} \exp[-\frac{1}{2}(x - v_i^x)^T (F_i^{xx})^{-1} (x - v_i^x)]}{\sum_{i=1}^{c} \frac{p(\eta_i)}{(2\pi)^{n/2}\sqrt{|F_i^{xx}|}} \exp[-\frac{1}{2}(x - v_i^x)^T (F_i^{xx})^{-1} (x - v_i^x)]} \quad (5)$$

where F^{xx}_i is obtained by partitioning the covariance matrix F_i as follows.

$$F_i = \begin{bmatrix} F_i^{xx} & F_i^{xy} \\ F_i^{yx} & F_i^{yy} \end{bmatrix} \quad (6)$$

where F^{xx}_i is the $n\times n$ sub-matrix containing the first n rows and columns of F_i, F^{xy}_i is an $n \times 1$ column vector containing the first n elements of the last column of F_i, F^{yx}_i is an $1\times n$ row vector containing the first n elements of the last row of F_i, and F^{yy}_i is the last element in the last row of F_i.

The mixture of Gaussian processes defined by equations (4) and (5) is in fact a kind of operating regime based model (1) where the validity function is chosen as $\phi_i(x) = p(\eta_i \mid x)$.

4 Parameter estimation

Through a linear transformation of the input variables [19], the antecedent partition can be accurately captured and no decomposition error occurs. Unfortunately, the resulting model is not transparent as it is hard to interpret the linguistic terms defined on the linear combination of the input variables. In order to form an easy interpretable model that does not rely on transformed input variables, a new clustering algorithm is proposed based on the expectation maximization (EM) [16]. The clusters obtained by GG clustering are multivariate Gaussian functions. The alternating optimization of these clusters is identical with the EM identification of the mixture of these Gaussian models when the fuzzy weight exponent $m = 2$.

The EM algorithm is widely used for parameter estimation of the mixture of models, in particular the mixture of Gaussian model. The basics of EM are the followings. Suppose we know the observed values of a random variable z and wish to model the density of z using a model parameterized by 77. EM obtains parameter estimates $\hat{\eta}$ which maximize the likelihood $L(\eta) = p(z|\eta)$ of the data. The EM assumes that this estimation is intractable and the values of a missing or hidden random variable h would make the problem more tractable. Let $p(z,h|\eta)$ denote the joint probability of z and h parameterized by η. It is assumed that z and h are such that maximizing the complete data likelihood $L_C(\eta) = p(z,h|\eta)$ is more tractable than maximizing $L(\eta)$. However, the values of h are unknown. The EM algorithm tackles this problem by iteratively generating a probability over the values h and estimating the parameters that maximize the expected value of $L_c(\eta)$ with respect to h.

The optimal θ_i parameter vector of the local model can be obtained as:

$$\min_{\theta_i} \frac{1}{N}(y - X_e\theta_i)^T \Phi_i (y - X_e\theta_i) \quad (7)$$

where .X_c=[X1] denotes the extended regression matrix obtained by adding a unitary column to X, and Φ_i is a matrix having the membership degrees on its main diagonal:

$$\Phi_i = \begin{bmatrix} \mu_{i,1} & 0 & \cdots & 0 \\ 0 & \mu_{i,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mu_{i,N} \end{bmatrix} \quad (8)$$

The weighted least-squares estimate of consequent parameters is given by

$$\theta_i = (X_e^T \Phi_i X_e)^{-1} X_e^T \Phi_i y \quad (9)$$

5 Experimental results

Furnace gas temperature is required to be computed or monitored for real-time use in modeling and control of the reheating furnace. The simulation experiment presented in this paper is derived from the 1580 reheating furnace in Baosteel Corporation of China. We collected 300 samples in 5 h, and the ratio of training samples to test samples is 2:1. According to technics analysis, the fuel flow at the current moment and the temperature of furnace gas at previous moment are used to the input of the software sensor. After training, the soft sensor model can estimate the gas temperature at the current moment online.

The test results for both the estimate output of the soft sensor and the actual value are shown in **figures 2-7** (suppose the cluster $c = 4$). The actual output is the dotted line and the estimate output is the solid one.

Through a series of experimental simulations, we can find that if we increase the value of c, we will get the less generalization error with the cost of more computational time as a balance and *vice versa.* With the change of c, the results for both generalization MSE and computational time in different zones are shown in **table 1.** The computer used for all these simulations is PIV 1.5 GHz PC with 256 MB RAM and MATLAB version is 6.5 (R13) running under the operating system Windows 2000 professional.

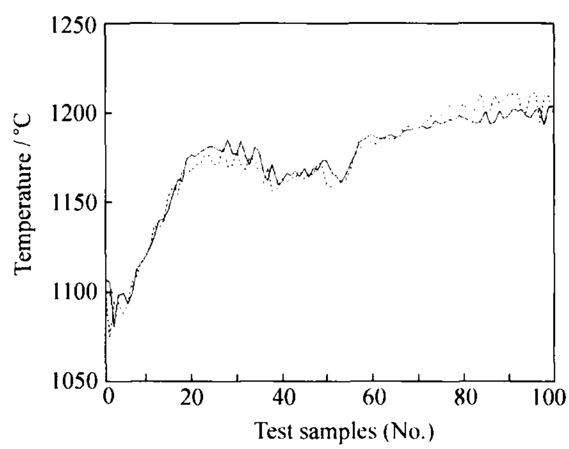

Figure 2 Test results for the estimated output and the actual output in zone 1.

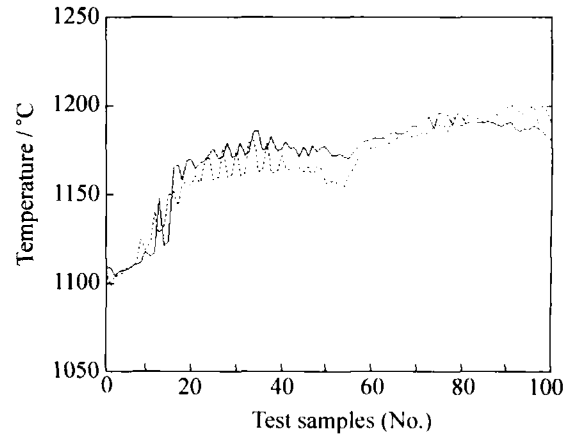

Figure 3 Test results for the estimated output and the actual output in zone 4.

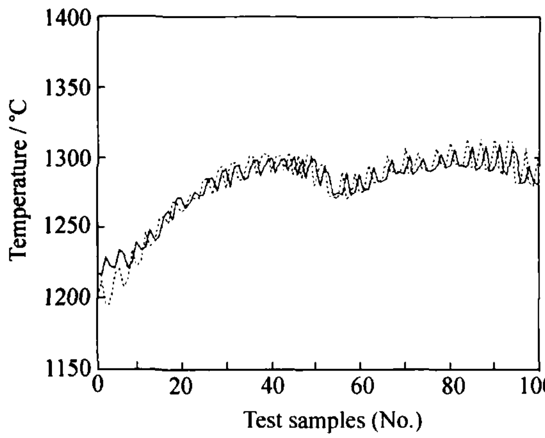

Figure 4 Test results for the estimated output and the actual output in zone 2.

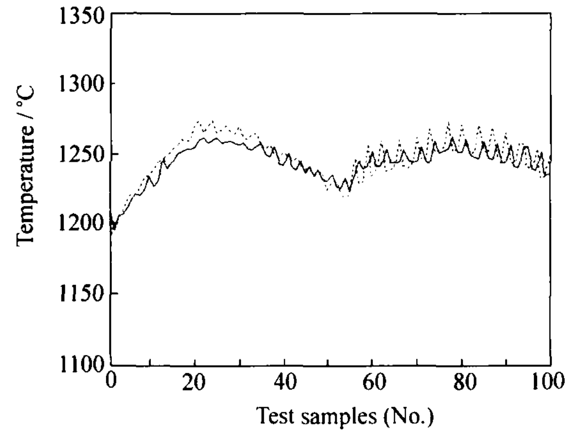

Figure 6 Test results for the estimated output and the actual output in zone 3.

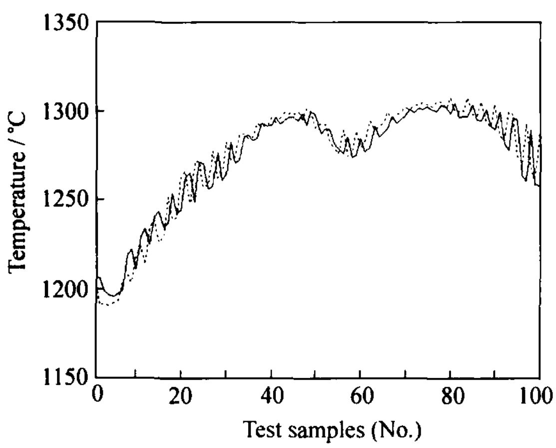

Figure 5 Test results for the estimated output and the actual output in zone 5.

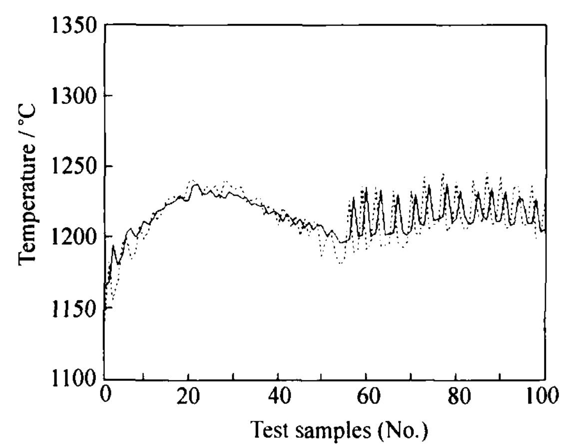

Figure 7 Test results for the estimated output and the actual output in zone 6.

Table 1 A series of results for the change of c in different zones

c	Zone 1		Zone 2		Zone 3		Zone 4		Zone 5		Zone 6	
	MSE/°C	Time/s	MSE/°C	Time/s	MSE/°C	Time/s	MSE/°C	Time/s	MSE/°C	Time/s	MSE/°C	Time/s
2	7.6286	0.1500	9.2553	0.1400	8.8070	0.2110	9.9654	0.1600	10.0879	0.1600	9.5922	0.1600
3	7.6239	0.2110	8.9754	0.2300	8.6147	0.2300	9.5958	0.1800	9.9512	0.1900	9.4719	0.2400
4	7.4682	0.2900	8.9726	0.2400	8.3426	0.2410	9.4082	0.3200	9.7934	0.3600	9.3734	0.2700
5	7.2985	0.3500	8.7488	0.4610	8.2780	0.2600	9.3928	0.3510	9.6912	0.3900	9.3454	0.3500
6	7.1695	0.4300	8.2479	0.5410	8.2115	0.2800	9.1368	0.3710	9.6394	0.4010	9.3095	0.4400

6 Conclusions

(1) Software sensors, which provide the on-line estimation of process variables that are difficult or costly to measure from available sensors, are powerful tools in system modeling and control. Their performance depends on both the measurement quality delivered by available sensors and the associated estimation algorithm.

(2) The proposed soft sensor can be implemented *via* online computation which can be employed in reheating processes for optimal running and control. Furthermore, we believe the proposed method sheds some light on the potential application to industrial fields.

References

[1] L.M. Pedersen and B. Wittenmark, On the reheat furnace control problem, [in] *Proceedings of the American Control Conference*, Philadelphia, 1998, p.3811.

[2] Y. Kim, K.C. Moon, B.S. Kang, C. Han, and K.S. Chang, Application of neural network to the supervisory control of a reheating furnace in the steel industry, *Control Eng. Pract.*, 6(1998), No.8, p.1009.

[3] H.S. Ko, J.S. Kim, T.W. Yoon, M. Lim, D.R. Yang, and I.S. Jun, Modeling and predictive control of a reheating furnace, [in] *Proceedings of the American Control Conference*, Chicago, 2000, p.2725.

[4] G Ditzhuijzen, D. Staalman, and A. Koorn, Identification and model predictive control of a slab reheating furnace, [in] *Proceedings of the 2002 International Conference on Control Applications*, Glasgow, 2002, p.361.

[5] Z. Chen, C. Xu, B. Zhang, H. Shao, and J. Zhang, Advanced control of walking-beam reheating furnace, *J. Univ. Sci. Technol. Beijing,* 10(2003), No.4, p.69.

[6] G Martin, Consider soft sensors, *Chem. Eng. Prog.,* 7(1997), p.66.

[7] N. Li, S. Li, and Y. Xi, A multiple model approach to modeling based on LPF algorithm, *J. Syst. Eng. Electron.,* 12(2001),No.3,p.64.

[8] R. Murray-Smith and T.A. Johansen, *Multiple Model Approaches to Nonlinear Modeling and Control,* Taylor & Francis, London, 1997.

[9] R. Babuska, *Fuzzy Modeling for Control,* Kluwer Academic Publishers, Boston, 1998

[10] T.A. Johansen and R. Babuska, On multi-objective identification of Takagi-Sugeno fuzzy model parameters, [in] *Proceedings of IFAC World Congress,* Barcelona, 2002, p.847.

[11] V. Tresp, Mixtures of Gaussian processes, *Adv. Neural Inf. Process. Syst.,* 13(2001), p.654.

[12] J.Q. Shi, R. Murray-Smith and D.M. Titterington, *Hierarchical Gaussian Process Mixtures for Regression,* Technical Report TR-2002-107, Department of Computing Science, University of Glasgow, Scotland, UK, 2002.

[13] C.E. Rasmussen and Z. Ghahramani, Infinite mixtures of Gaussian process experts, *Adv. Neural Inf. Process. Syst,* 14(2002),p.881.

[14] J.Q. Shi, R. Murray-Smith, and D.M. Titterington, Bayesian regression and classification using mixtures of Gaussian processes, *Int. J. Adapt. Control Signal Process.,* 17(2003), p.1.

[15] I. Gath and A.B. Geva, Unsupervised optimal fuzzy clustering, *IEEE Trans. Pattern Anal. Mach. Intell.,* 7(1989), p.773.

[16] F. Hoppner, F. Klawonn, R. Kruse, and T. Runkler, *Fuzzy Cluster Analysis—Methods For Classification, Data Analysis And Image Recognition,* John Wiley and Sons, 1999.

[17] J. Abonyi, R. Babuska, and F. Szeifert, Modified Gath-Geva fuzzy clustering for identification of Takagi-Sugeno fuzzy models, *IEEE Trans. Syst. Man Cybern.,* **Part** B, 2002, p.612.

[18] E. Kim, M. Park, S. Kim, and M. Park, A transformed input-domain approach to fuzzy modeling, *IEEE Trans. Fuzzy Syst.,* 6(1998), No.4, p.596.

[19] J.A. Bilmes, A *gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models,* ICSI TR-97-021, U.C. Berkeley, 1998.