

文章编号: 1000-582X(2012)06-141-06

# 贝叶斯模型比较的多模型组合框架在软测量建模中的应用

韩露, 任江洪, 黄毅卿

(重庆大学 自动化学院, 重庆 400044)

**摘要:** 针对基于单一模型建立的软测量模型存在着预测精度需要进一步提高的问题, 在分析目前常用的 2 种多模型组合框架的基础上, 提出了一种基于贝叶斯模型比较的多模型组合框架。该框架以通过模糊 c-均值聚类分析获得的生产过程状态变化知识为基础, 对每种状态下各子模型的预测性能采用贝叶斯模型比较方法进行比较, 并以此为基础在不同状态下采用了不同的子模型加权策略。在进行模型比较时, 基于交叉检验分布, 使用子模型训练所得采样序列, 有效地减少了计算量。将该框架用于工程应用, 取得了较好效果。

**关键词:** 贝叶斯模型比较; 软传感器; 蒙特卡洛方法; 参数估计

**中图分类号:** TP212.6

**文献标志码:** A

## A multi-model composition framework based on bayesian model comparison and its application in soft-sensor modeling

HAN Lu, REN Jiang-hong, HUANG Yi-qing

(College of Automation, Chongqing University, Chongqing 400044, P. R. China)

**Abstract:** In order to improve the prediction performance of single model based soft sensor, the features of the current model combination frameworks by analyzing, a new multi-model combination framework based on the bayesian model comparison is proposed. In this framework, fuzzy c-means clustering to the historical data is used to analyze the production states, then the prediction performance of sub-models at different states are compared based on bayesian model comparison. The comparing results are the basis of the model combination strategy at different states. With adapting cross-validation predictive distribution, the samples got from the trained models are used to successfully reduce computation load of model comparison. The framework has obtained good results in the practical application.

**Key words:** Bayesian model comparison; soft sensor; Monte Carlo method; parameter estimation

工业过程中的许多过程变量由于测量技术或生产条件的限制, 难以直接在线进行测量。这些量常常对过程性能的评估和确保产品质量至关重要。传统的方法是采用样本实验室化验或者安装昂贵的在

线成分分析仪进行在线测量。成分分析仪购置和维护成本很高, 同时需要离线进行维护的时间也很长, 限制了其进一步推广使用。实验室化验的结果延迟很大(常常在 1 h 以上), 已经失去了作为反馈信号,

**收稿日期:** 2011-12-20

**基金项目:** 国家 863 计划资助项目(2009AA01Z310); 中加政府间科技合作基金资助项目(2009DFA12100); 重庆市科委自然科学基金资助项目(CSTC, 2011BB008)

**作者简介:** 韩露(1977-), 女, 重庆大学博士, 主要从事模式识别, 机器视觉, 统计学习等方向研究, (Tel)13320349910。

进行产品质量控制的作用。为解决这些问题, Joseph 和 Brosilow 提出了推断控制<sup>[1]</sup>的方法:通过建立软测量模型,基于易于测量并与目标变量相关的过程变量,对目标变量进行推断估计。近年来,软测量模型在工业生产中获得了越来越广泛的应用,并取得了较好的应用效果<sup>[2]</sup>。

目前在软测量建模中应用较多的是采用单一模型对目标变量整个变化过程进行建模。但是由于工业过程本身的动态性质(例如生产状态的频繁切换),以及可测变量与目标变量间的强非线性关系,单一模型在许多应用中的效果不尽理想。于是不断有研究者提出采用多个模型对软测量模型进行建模的方法。机器学习的研究也已给出证明<sup>[3]</sup>,当采用适当的模型组合策略时,组合模型预测误差的方差将小于任何一个用于组合的单个模型的预测误差的方差。

在软测量建模领域,多模型组合方法大致可以分为 2 类:1)对训练数据进行聚类分析,找出目标变量及可测变量的不同状态(对应于不同的生产过程状态),对不同状态的数据使用适合于该状态的不同模型(单个)进行建模<sup>[4]</sup>;2)对整个过程采用多个模型同时进行建模,然后对各个模型的预测输出采用适合的加权策略进行组合输出以获得组合模型的输出<sup>[5-6]</sup>。方法一在某种状态的训练数据较少时,相应模型的预测精度会显著下降,并且其在本质上仍是采用单个模型进行预测。方法二并未考虑生产过程状态变化的问题,而对生产过程的理解,是进一步优化控制的基础。

吸取以上 2 种方法的优点,并为解决以上方法存在的问题,提出了一种基于贝叶斯模型比较的模型组合框架。其基本思想是通过模糊 C-均值聚类方法对训练数据进行聚类分析,获得生产过程状态切换的知识;用全局数据对多个模型进行训练;对各个过程状态,计算每个模型对该状态数据的模型证据(model evidence)  $p(D|M_i)$  作为模型对该状态数据的预测精度的评估,该评估是对各子模型的预测进行组合的依据。通过采用交叉检验分布,利用马尔科夫链蒙特卡罗方法(Markov chain Monte Carlo, MCMC)实现了该方法的有效计算,使其适合于子模型采用贝叶斯统计模型(例如贝叶斯神经网络、高斯过程回归、相关向量机)进行软测量建模的情况。

## 1 相关概念

### 1.1 贝叶斯模型组合

最简单的模型组合方法是 Breiman 提出的 bagging 方法<sup>[6-7]</sup>。多个模型组合输出的结果为

$$y_{\text{COM}}(x) = \frac{1}{K} \sum_{k=1}^K y_k(x). \quad (1)$$

也即对  $K$  个模型的输出进行算术平均。在此基础上更为复杂、自适应性更好的模型组合方法是 Jacobs 等提出的混合专家模型(mixture of expert)<sup>[8]</sup>

$$p(t|x) = \sum_{k=1}^K \pi_k(x) p_k(t|x). \quad (2)$$

其对多个子模型的预测分布,根据相应的输入值进行加权组合,得到组合模型的预测分布。其中  $0 \leq \pi_k(x) \leq 1$ ,  $\sum_{k=1}^K \pi_k(x) = 1$ 。由于是对模型的预测分布进行加权组合,因此适应于子模型采用概率预测模型进行建模的情况。这也是所提出的模型框架的基础。

### 1.2 贝叶斯模型比较

由公式(2)可知,混合专家模型的关键是如何确定系数  $\pi_k(x)$ 。最直观的想法是根据模型的预测精度来确定组合系数:在输入值  $x$  处预测精度高的模型获得较大的组合权值,相应预测精度低的模型组合权值也较低。在贝叶斯统计建模研究中比较模型预测性能的基本方法是贝叶斯模型比较。其基本思想为:在训练数据  $D$  已知的情况下,可以评价模型  $M_k$  的后验分布如下

$$p(M_k|D) \propto p(M_k)p(D|M_k), \quad (3)$$

其中: $p(M_k)$ 是模型  $k(k=1, \dots, K)$  的先验概率; $p(D|M_k)$ 是模型  $k$  在模型空间的似然函数,又被称作模型证据,描述了由数据  $D$  表明的对模型  $k$  的偏好程度。当所有模型的先验概率都取相同概率的情况下,可通过只比较  $p(D|M_k)$  来进行模型比较。

$$p(D|M_k) = \prod_{i=1}^n p(y_i|M_k), (x_i, y_i) \in D, \quad (4)$$

其中

$$p(y_i|M_k) = \int p(y_i|\theta_k, M_k) p(\theta_k|M_k) d\theta_k. \quad (5)$$

计算  $p(y_i|M_k)$  意味着要在模型  $k$  的整个先验概率空间进行积分。在模型  $k$  比较复杂的情况下,以上积分无法以解析的方式进行,只有采用马尔科夫链

蒙特卡诺方法<sup>[9]</sup>来近似处理。

### 1.3 马尔科夫链蒙特卡诺方法

蒙特卡诺方法的基本思想是通过抽样从 $\theta_k$ 的分布 $p(\theta_k | M_k)$ 获得 $\theta_k$ 的一系列独立样本 $\theta_k^{(t)}, t=1, \dots, T$ ,那么式(5)可以用以下的求和方法来近似

$$\hat{p}(y_i | M_k) = \frac{1}{T} \sum_{t=1}^T p(y_i | \theta_k^{(t)}, M_k). \quad (6)$$

上式又被成为蒙特卡诺积分。蒙特卡洛方法的关键是找到独立的样本序列。马尔科夫链是一种寻找独立样本序列的方法。通过它获得随机变量序列 $\theta_0, \theta_1, \dots$ 。序列的下一个状态只依赖于上一个状态。当满足一定的条件时<sup>[9]</sup>,无论序列的初始值取什么,序列在迭代 $m$ 次后,将收敛于目标静态分布 $p(\theta)$ ,那么随后的迭代生成的样本 $\theta^{(t)}, t=m+1, \dots, n$ 就可作为蒙特卡诺积分所需的独立样本。有关马尔科夫链蒙特卡诺方法的更详细内容请参见文献[10]。

MCMC方法已被成功地用于贝叶斯建模,例如Neal在文献[10-11]中将MCMC方法用于贝叶斯神经网络和高斯过程回归。这也是在对子模型进行建模时所用的方法。

## 2 基于贝叶斯模型比较的组合预测模型框架

流程工业的生产过程是一个动态过程,常常在不同的状态间切换。在不同的状态下,影响过程变量的因素不尽相同<sup>[12]</sup>。采用单一的模型对整个生产过程进行建模,模型的输入变量和模型结构都是确定不变的,不能保证在所有的状态下都获得理想的预测性能。因此需要对生产状态进行辨识,针对不同的状态采用不同的建模策略。

### 2.1 生产过程状态的辨识

采用聚类分析的方法,基于历史生产数据对生产过程状态进行状态辨识是可行的方法<sup>[13-14]</sup>。采用了模糊C均值聚类算法<sup>[15]</sup>进行过程状态的辨识。模糊聚类的目标是通过迭代过程,寻找使以下的目标函数取得最小值的聚类中心

$$J_m = \sum_{i=1}^N \sum_{c=1}^C u_{ij}^m \|x_i - c_j\|^2, 1 \leq m \leq \infty, \quad (7)$$

其中: $c_j$ 是第 $j$ 个聚类中心的位置; $x_i$ 是第 $i$ 个训练数据; $u_{ij}$ 表示数据 $x_i$ 对于第 $j$ 个聚类的隶属度,具体定义如下

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left( \frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{1}{m-1}}}. \quad (8)$$

模糊c-均值聚类的基本步骤见文献[15]。传统的模糊c-均值算法中的聚类中心数量必须事先确定。对于一些较简单或研究比较透彻的过程,聚类中心的数量比较容易确定,例如PVC的生产过程。但对于较复杂的生产过程,由于对其机理特性缺乏更深入的认识,常常很难事先确定聚类中心的数量,例如我们面对的硫的回收过程。因此在实际应用中,采用了文献[16]提出的 $I$ 准则,结合生产专家的意见,对聚类中心数量的取值基于生产数据进行自适应的选择。通过后面的数据分析表明,该方法在实践中效果较好。

### 2.2 基于交叉检验分布的贝叶斯模型比较

由前讨论知直接计算式(5)很困难,需要寻找高效准确的计算方法。

交叉检验是一种常用的模型训练方法,在概率建模中,使用交叉检验方法即是对其中一个训练数据 $(x_i, y_i)$ 使用除 $(x_i, y_i)$ 以外的训练数据 $D_{[-i]}$ 来建立模型(获得模型参数 $\theta$ 的后验分布),并用此模型来估计 $y_i$ 的交叉检验分布,用公式表示如下(为了公式的清晰,在分布的条件中省略了表示模型的 $M_k$ )

$$p(y_i | D_{[-i]}) = \int p(y_i | \theta, D_{[-i]}) p(\theta | D_{[-i]}) d\theta. \quad (9)$$

当使用参数回归模型时,例如贝叶斯神经网络, $y_i$ 的分布与 $D_{[-i]}$ 并不直接相关,上式也可简化为

$$p(y_i | D_{[-i]}) = \int p(y_i | \theta) p(\theta | D_{[-i]}) d\theta. \quad (10)$$

Geisser和Eddy<sup>[17]</sup>的研究指出可以用交叉检验分布来有效地估计模型证据 $p(D | M_k)$

$$\hat{p}(D | M_k) = \prod_{i=1}^n p(y_i | M_k), (x_i, y_i) \in D. \quad (11)$$

在实际应用中,一般对上式取对数,以更利于计算和比较,如下式所示

$$\delta = \log(\hat{p}(D | M_k)) = \sum_{i=1}^N \log(p(y_i | D_{[-i]}, M_k)). \quad (12)$$

如果按照MCMC方法来训练模型,那么用上式来估计 $p(D | M_k)$ ,意味着要进行 $N$ 次MCMC。如果有 $K$ 个子模型,则共要进行 $K \times N$ 次MCMC,这在目前的计算能力下是没有实用价值的。

为解决这问题,采用了近似的蒙特卡罗估计方法:Gelfand<sup>[18]</sup>的研究指出可以通过对所有训练数据

进行 MCMC 采样,并用下式来对式(10)进行蒙特卡诺估计

$$\hat{p}(y_i | D_{[i]}) = \frac{1}{T^{-1} \sum_{t=1}^T [p(y_i | \theta^{(t)})]^{-1}} \quad (13)$$

$T$  是马尔科夫链收敛后迭代的次数。则利用式(13)和式(12)便可直接利用子模型训练所得的抽样样本估计子模型对相应状态的模型证据。

但对于高斯过程回归这样的非参数回归模型(详细介绍请参见文献[19]),并不能利用式(13)进行模型参数的 MCMC 采样。在交叉检验的情况下, $y_i$  的预测输出依赖于训练数据  $y_{[i]}$  和协方差函数。 $y_i$  的预测均值为

$$E[y_i | \mathbf{X}_i, q, D_{[i]}] = \mathbf{K}^T C_{[i]}^{-1} y_{[i]} \quad (14)$$

其中  $C_{[i]}^{-1}$  为训练数据  $D_{[i]}$  构成的  $(N-1) \times (N-1)$  的协方差矩阵,  $y_{[i]} = (y_1, \dots, y_{i-1}, y_{i+1}, \dots, y_N)$ 。如果基于上式通过式[12]来计算模型证据  $p(D | M_k)$  需要构建  $N$  次  $C_{[i]}$  并求逆,计算量太大。因此,利用了 Sundararajan 等的研究成果<sup>[20]</sup>,采用下式来近似计算高斯过程回归模型的交叉检验分布

$$\log p(y_i | x_i, \theta, D_{[i]}) = \frac{1}{2} \log 2\pi - \log c_{ii} + \frac{1}{2} \frac{q_i^2}{c_{ii}} \quad (15)$$

其中:  $c_{ii}$  是  $C^{-1}$  的第  $i$  个对角线元素;  $q_i$  是  $q = C^{-1} y$  的第  $i$  个元素;  $C$  是采用超越参数  $\theta$  计算的协方差矩阵。采用此方法只需构建一次协方差矩阵,大大降低了计算量。

### 2.3 模型的组合策略

由前已知,在不同状态下,不同的子模型的预测性能各不相同。辨识清楚生产过程的状态后,需要对各个子模型对不同状态数据的预测性能,按照以上所述的方法进行评估(实践中,选择对相应状态的隶属度值大于 0.6 的数据作为该状态的典型特征数据来作性能评估),评估的结果作为不同状态下子模型组合输出的依据。那么子模型  $k$  在状态  $m$  下,进行子模型的预测加权输出时的加权系数如下式所示

$$\omega_{km} = \frac{\delta_{km}}{\sum_{i=1}^K \delta_{im}} \quad (16)$$

其中  $\delta_{km}$  为子模型  $k$  对状态  $m$  的数据,依照公式(12)计算所得的预测性能评估。那么对属于状态  $m$  的数据  $x^*$  的组合预测均值为

$$\hat{y}_m^* = \sum_{k=1}^K \omega_{km} E[p(y^* | x^*, M_k)] \quad (17)$$

在对生产过程中的输入数据  $x^*$  作预测时,首先对输

入数据作聚类分析,获取其所属状态信息。取其中隶属度值最大的 2 个状态作为输入数据的所属状态,根据上式分别计算 2 种状态下组合模型的预测输出。最后根据输入数据对状态的隶属度大小,对 2 种状态下的预测进行组合,得出预测的均值如下

$$\overline{y^*} = \frac{a_1}{a_1 + a_2} \widehat{y}^{*(1)} + \frac{a_2}{a_1 + a_2} \widehat{y}^{*(2)} \quad (18)$$

其中  $\widehat{y}^{*(1)}, \widehat{y}^{*(2)}$  分别为输入  $x^*$  在隶属度最大的 2 个状态下的预测输出。

## 3 实验分析

将所提出的方法用于硫回收装置最后尾气中硫含量的软测量模型建模。酸性气体是大气的主要污染物,其主要危害成分是硫化氢( $H_2S$ )和二氧化硫( $SO_2$ )。由图 1 所示,回收装置输入 2 种酸性气体:第一种富含  $H_2S$ ,主要来自于洗气厂;第二种是来自于处理含硫废水产生的气体,主要含有  $H_2S$  和氨气( $NH_3$ )。两种酸性气体被送入 2 个不同的反应炉中燃烧:1 号反应炉中控制空气的输入量,使第一种酸性气体中的  $H_2S$  在空气作用下进行不完全氧化反应,生成纯硫;2 号反应炉中需输入过量的空气,以避免第二种酸性气体的燃烧生成氮和氮氧化物。

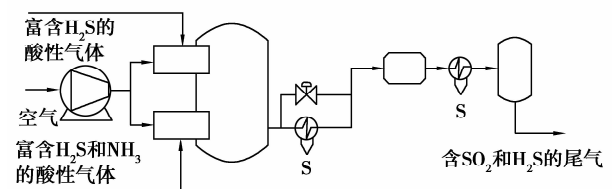


图 1 硫回收装置简单示意图

燃烧后的气体被送出反应炉进行冷却,就可得到液态硫。液态硫经过高温的传输带传送,其中所含的水将以水蒸气的形式被除去。最后的尾气包含  $H_2S$  和  $SO_2$ ,准确测量其值是准确控制空气供给量,确保回收装置性能的基础。目前能够使用在线分析仪来直接测量  $H_2S$  和  $SO_2$  的浓度。但因为与酸性气体长期接触,在线分析仪需要停机进行检修的时间很多。因此就需要软传感器作为在线分析仪的备份装置。

经过参考生产专家的意见,以及对输入变量和预测变量的相关性分析,选择了以下 5 个量作为输入变量:  $x_1$  (第一种酸性气体输入量),  $x_2$  (1 号反应炉的空气输入量),  $x_3$  (1 号反应炉的调节空气输入量),  $x_4$  (2 号反应炉的酸性气体输入量),  $x_5$  (2 号反

应炉的空气输入量)。预测目标  $y$  为尾气中  $H_2S$  的浓度。训练数据共有 1 000 组,采样时间间隔为 1 min。软测量模型的基本结构为如下的非线性移动平均模型

$$y(k) = f(x_1(k), x_1(k-5), x_1(k-7), \dots, x_5(k), x_5(k-5), x_5(k-7)))$$

分别采用贝叶斯神经网络和高斯过程回归来实现了以上的非线性移动平均模型。在每种生产状态下,输入变量对预测输出的影响不尽相同,采样数据的误差也不同,因此采用了 3 种不同的噪声模型

1) 高斯噪声模型

$$e \sim N(0, \sigma^2)$$

相应方差的先验分布为

$$\sigma^2 \sim \text{Inv-gamma}(\sigma_0^2, \nu_0)$$

采用了和 Neal<sup>[10-11]</sup>相同的超越参数设置

$$\sigma_0 = 0.05, \nu_0 = 0.5$$

2)  $t$  分布噪声模型

$$e \sim t_\nu(0, \sigma^2)$$

其层次先验分布为<sup>[21]</sup>

$$\sigma^2 \sim \text{Inv-gamma}(\sigma_0, \nu_0)$$

$$\nu = V[i], i \sim U_d(1, K)$$

$$V[1:K] = [2, 2.3, 2.6, 3, 3.5, 4, 4.5, 5 : 10, \dots, 12 : 2 : 20, 25 : 5 : 50]$$

其中  $U_d(a, b)$  为均匀分布。 $t$  分布噪声模型能够通过控制自由度  $\nu$  来控制尾部长度,对一些边界值有更好的鲁棒性。

某些输入量对噪声有较大的影响,因此采用了与输入有关的噪声模型。实验进一步发现,和输入相关的噪声模型在某些状态下有相对更好的性能。

3) 和输入相关的噪声模型如下

$$e^{(i)} \sim N(0, (\sigma^2)^{(i)})$$

$$(\sigma^2)^{(i)} = F(x^{(i)}, \theta_{\text{noise}}) + \epsilon$$

$$\epsilon \sim \text{Inv-gamma}(\sigma_0^2, \nu_0)$$

$$x^{(i)} = \{x_1(k)^{(i)}, x_1(k-5)^{(i)}, x_2(k)^{(i)}, \dots, x_2(k-5)^{(i)}, x_3(k)^{(i)}, x_3(k-5)^{(i)}\}$$

由于第一种气体在输入的酸性气体中占主导作用,因此选择和第一种酸性气体相关的输入量作为误差模型的依赖变量。 $F(\cdot)$  采用 MLP 来建立映射。

所有建模工作均在 Matlab 平台上完成,其中在贝叶斯神经网络和高斯过程回归建模上使用了 MCMCstuff 工具包。MCMC 采样使用了 OpenBUGS 软件包。

基于以上模型设置,共采用 6 个不同的子模型对生产过程进行建模,不区分生产状态的变化,单独

使用这些模型对整个过程进行预测的性能(采用数据间的相关性系数来量度)如表 1 所示。

表 1 单一模型的预测性能

| 参量   | MLP-     | MLP-t    | MLP-       | GPR-     | GPR-t    | GPR-       |
|--|----------|----------|------------|----------|----------|------------|
|  | 高斯<br>噪声 | 分布<br>噪声 | 输入依<br>赖噪声 | 高斯<br>噪声 | 分布<br>噪声 | 输入依<br>赖噪声 |
| H <sub>2</sub> S 浓<br>度训练<br>数据<br>1 000 组 | 0.856    | 0.863    | 0.847      | 0.862    | 0.865    | 0.849      |
| H <sub>2</sub> S 浓<br>度检验<br>数据<br>500 组   | 0.845    | 0.850    | 0.836      | 0.847    | 0.852    | 0.839      |

从以上数据可以看出采用  $t$  分布噪声的模型预测性能较好,采用输入依赖噪声的模型性能最差。

在此基础上,作了第二个实验:在各个不同的生产状态内分析各个子模型的预测性能。采用前文所述的方法,共辨识出 5 个比较明显的生产状态。在这 5 个状态中,各模型的预测性能如表 2 所示。其中每个单元格的上下部分为预测数据与训练数据和检验数据的相关性系数。

表 2 各模型在每个生产状态的预测性能

| 参量             | 状态 I           | 状态 II          | 状态 III         | 状态 IV          | 状态 V           |
|----------------|----------------|----------------|----------------|----------------|----------------|
| MLP-<br>高斯噪声   | 0.858<br>0.849 | 0.859<br>0.846 | 0.845<br>0.831 | 0.857<br>0.841 | 0.855<br>0.836 |
| MLP-t<br>分布噪声  | 0.868<br>0.855 | 0.857<br>0.842 | 0.852<br>0.840 | 0.865<br>0.852 | 0.869<br>0.850 |
| MLP-输入<br>依赖噪声 | 0.845<br>0.826 | 0.843<br>0.835 | 0.856<br>0.848 | 0.851<br>0.842 | 0.845<br>0.829 |
| GPR-<br>高斯噪声   | 0.862<br>0.847 | 0.861<br>0.849 | 0.853<br>0.841 | 0.864<br>0.846 | 0.867<br>0.852 |
| GPR-t<br>分布噪声  | 0.871<br>0.859 | 0.862<br>0.851 | 0.860<br>0.845 | 0.867<br>0.849 | 0.866<br>0.850 |
| GPR-输入<br>依赖噪声 | 0.851<br>0.831 | 0.847<br>0.837 | 0.857<br>0.847 | 0.849<br>0.840 | 0.847<br>0.828 |

从上表看出,并没有某一个预测模型在每一个

生产状态都能作出最优的预测。采用提出的组合模型来进行预测,预测结果表 3 所示。对比表 2 和表 3,发现除了状态 I 以外,在生产过程的其他状态,组合模型的预测性能都优于最优单一模型在该状态的预测性能。

表 3 组合模型在各状态的预测性能

| 参量                  | 状态 I  | 状态 II | 状态 III | 状态 IV | 状态 V  |
|---------------------|-------|-------|--------|-------|-------|
| H <sub>2</sub> S 浓度 |       |       |        |       |       |
| 训练数                 | 0.869 | 0.865 | 0.862  | 0.871 | 0.873 |
| 1 000 组             |       |       |        |       |       |
| H <sub>2</sub> S 浓度 |       |       |        |       |       |
| 检验数据                | 0.858 | 0.855 | 0.852  | 0.856 | 0.854 |
| 500 组               |       |       |        |       |       |

#### 4 结 语

针对采用单一模型建立的软测量模型,存在着模型预测精度有待进一步提高,模型鲁棒性也有待加强的问题,提出了一种组合的预测模型框架。该框架首先采用 c-均值模糊聚类算法辨识生产过程状态,针对各生产状态的生产数据,基于贝叶斯模型比较评估各子模型对各状态数据的预测性能,将该性能评估作为组合子模型的依据。通过实验分析,证明了该方法在实验所用实例中的有效性,可以为采用组合模型来实现软测量模型的研究人员提供一定的参考。

#### 参考文献:

- [ 1 ] JOSEPH B, BROSILOW C B. Inferential control of processes: part 1. steady state analysis and design[J]. American Institute of Chemical Journal, 1978, 24(3): 485-508.
- [ 2 ] KANO M, NAKAGAWA Y. Data-based process monitoring, process control, and quality improvement: recent developments and applications in steel industry[J]. Computers & Chemical Engineering, 2008, 32(1-2): 12-24.
- [ 3 ] BISHOP C M. Pattern recognition and machine learning[M]. New York: Springer, 2007.
- [ 4 ] LI X L, SU H Y, CHU J. Multiple model soft sensor based on affinity propagation, gaussian process and bayesian committee machine[J]. Chinese Journal of Chemical Engineering, 2009, 17(1): 95-99.
- [ 5 ] FENG R, ZHANG Y Z, SONG C L, et al. A multiple model approach to modeling based on fuzzy support vector machines [J]. Journal of Shanghai Jiaotong University, 2003, 8(2): 137-141.
- [ 6 ] CHEN T, REN J H. Bagging for gaussian process regression[J]. Neurocomputing, 2009, 72(7-9): 1605-1610.
- [ 7 ] BREIMAN L. Bagging predictors [J]. Machine Learning, 1996, 24(2): 123-140.
- [ 8 ] JACOBS R A, JORDAN M I, NOWLAN S J, et al. Adaptive mixtures of local experts [J]. Neural Computation, 1991, 3(1): 79-87.
- [ 9 ] GILKS W R, RICHARDSON S, SPIEGELHALTER D J. Markov chain monte carlo in practice[M]. New York: Chapman & Hall, 1996.
- [10] NEAL R M. Bayesian learning for neural networks[M]. New York: Springer-Verlag, 1996.
- [11] NEAL R M. Regression and classification using gaussian process priors[J]. Bayesian Statistics, 1998, 6: 475-501.
- [12] RIGGS J B, KARIM M N. Chemical and bio-process control 2nd edition[M]. USA: Ferret Publishing, 2001.
- [13] 郭小萍,王福利. 基于滑动窗 PCA 的 PVC 聚合过程在线监测[J]. 聚氯乙烯, 2005(10): 39-41.  
GUO XIAO-PING, WANG FU-LI. On-line monitoring the polymerization process of PVC based on the moving windows PCA modeling[J]. Polyvinyl Chloride, 2005 (10): 39-41 .
- [14] BEZDEK J C. Pattern recognition with fuzzy objective function algorithms[M]. Norwell, MA, USA: Kluwer Academic Publishers, 1981.
- [15] MAULIK U, BANDYOPADHYAY S. Performance evaluation of some clustering algorithms and validity indices[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(12): 1650-1654.
- [16] GEISSER S, EDDY W F. A predictive approach to model selection[J]. Journal of the American Statistical Association, 1979, 74(365): 153-160.
- [17] GELFAND A E, DEY D K. Bayesian model choice: asymptotics and exact calculations[J]. Journal of the Royal Statistical Society Series B: Methodological, 1994, 56(3): 501-514.
- [18] RASMUSSEN C E, WILLIAMS C K I. Gaussian process for machine learning [M]. Cambridge: MIT, 2006.

- [13] ZHANG Y Q, SUN S J. Software reliability modeling based on unascertained theory[J]. Journal of Software, 2006, 17(8):1681-1687.
- [14] WILCOX R C, AYYUB B M. Uncertainty modeling of data and uncertainty propagation for risk studies[C]// 4th International Symposium on Uncertainty Modelling and Analysis. [S. l.]:IEEE, 2003, 184-191.
- [15] WANG Z. Reduction algorithms based on discernibility matrix: the ordered attributes method[J]. Journal of Computer Science and Technology, 2001, 16(6): 489-504.
- [16] 支天云, 苗夺谦. 二进制可辨矩阵的变换及高效的属性约简算法的构造[J]. 计算机科学, 2002, 29(2): 140-142.
- ZHI TIAN-YUN, MIAO DUO-QIAN. The binary discernibility matrix's transformation and high efficiency attributes reduction algorithm's conformation [J]. Computer Science, 2002, 29(2): 140-142.
- [17] 刘开第, 曹庆奎, 庞彦军. 基于未确知集合的故障诊断方法[J]. 自动化学, 2004, 30(5): 747-756.
- LIU KAI-DI, CAO QING-KUI, PANG YAN-JUN. A method of fault diagnosis based on unascertained set [J]. Acta Automatica Sinica, 2004, 30(5): 747-756.
- (编辑 侯 湘)



(上接第 146 页)

- [19] SUNDARAJAN S, KEERTHI S S. Predictive approaches for choosing hyperparameters in gaussian processes[J]. Neural Computation, 2001, 13(5): 1103-1118.
- BUGS 0.5 Examples Volume 1 (version i) [EB/OL]. (1999-06-08) [2010-03-11]. <http://www.stat.ufl.edu/system/man/BUGS/eg05vol1/>.
- [20] SPIEGELHALTER D J, THOMAS A, BEST N, et al. (编辑 侯 湘)