

文章编号:1006-3080(2015)02-0144-07

基于贝叶斯理论与 Vine Copula 的化工过程异常事件数的预测

吕 成, 张子扬, 任 翔, 李绍军

(华东理工大学化工过程先进控制和优化技术教育部重点实验室, 上海 200237)

摘要:针对化工过程风险,提出了一种化工过程异常事件数的预测方法。化工生产过程中由于受到干扰,时常发生异常事件。异常事件如果得不到有效控制将引发生产事故,其发生次数越高表明发生生产事故的概率越大,因此,准确预测化工过程异常事件数有助于提高化工过程的风险管理水平。基于操作班组,采用贝叶斯理论与 Vine Copula 建立了动态预测模型,实现对化工过程一个轮班内异常事件数的预测。

关键词:风险管理; 异常事件; 操作班组; 贝叶斯理论; Vine Copula

中图分类号:X937

文献标志码:A

Forecasting Abnormal Event Numbers in Chemical Process with Bayesian Theory and Vine Copula

LÜ Cheng, ZHANG Zi-yang, REN Xiang, LI Shao-jun

(Key Laboratory of Advanced Control and Optimization for Chemical Processes, Ministry of Education, East China University of Science and Technology, Shanghai 200237, China)

Abstract: A prediction method is proposed to cope with abnormal event numbers, which often appear in chemical process due to external disturbance. If an abnormal event is not effectively controlled, it will probably result in an accident. The higher abnormal event numbers are, the greater the probability of production accidents is. Therefore, the precise prediction on the abnormal event numbers can effectively improve the risk management level on the chemical process. Usually, four operating teams work in a workshop and the number of abnormal event varies from team to team. Based on operating teams, the dynamic prediction model is constructed by using the Bayesian theory and Vine Copula such that the abnormal event numbers in a operating team can be predicted.

Key words: risk management; abnormal event; operating team; Bayesian theory; Vine Copula

近年来,随着人们对化工产品需求的提高,化工生产过程正朝着大型化、综合化和复杂化的方向发展,化工企业面临着日趋严重的化工过程风险。化工事故时有发生,造成了巨大的经济损失、生态环境污染以及人员伤亡。

为降低化工生产过程风险,化工生产研究者和

实际工作者提出了许多化工过程风险分析方法^[1-5]。这些风险分析方法可分为定量分析和定性分析方法,其中比较常用的是定量分析方法。定量分析方法通常使用故障树分析事件可能的起因,使用事件树分析异常事件可能造成的后果,采用概率、频率等指标量化风险。2006年,Seider等^[6]提出了一种动

收稿日期:2014-05-04

基金项目:国家自然科学基金(21176072)

作者简介:吕 成(1989-),男,安徽蚌埠人,硕士生,研究方向为化工过程风险评估。

通信联系人:李绍军,E-mail:lishaojun@ecust.edu.cn

态风险评估方法,该风险评估方法将事件树与化工安全系统相结合基于历史数据动态评估安全系统失败的概率以及事故发生概率。Huang 等^[7]将事件树与模糊理论相结合提出了一种评估大型系统风险的简便方法。这些方法都是在现有信息的基础上对当前系统的状态进行风险评估,没有预测系统未来的风险状态,并且这些方法在进行风险评估时,没有将操作班组的因素考虑在内。

化工事故发生之前通常有一定的先兆,其中最明显的现象是化工生产过程中异常事件在一个轮班时间内的次数增加。在流程工业中,当生产过程受到干扰或发生故障时,控制系统未能将过程变量控制在正常范围内,此时生产过程出现异常事件。图 1 为变量波动区域以及报警临界线图。当变量偏离正常区域,越过 High/Low 报警边界线,可视为一次异常事件发生。如果异常事件未演变成生产事故,即变量在偏离正常范围一段时间后,通过自动或手动控制恢复到正常区域,这类异常事件称为未遂事件。如果安全控制系统控制失败,异常事件将会造成生产停车,甚至造成生产事故。

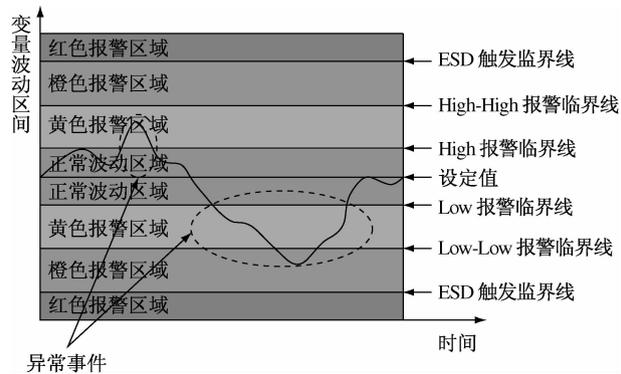


图 1 变量报警区域以及报警临界线

Fig. 1 Variable belt zones and alarm thresholds

对于复杂的化工生产过程,需要同时监控几百个变量,因此,DCS 和 ESD 系统被广泛运用于化工过程监控。当生产过程发生异常事件时,异常事件首先被记录在报警数据库中。一段时间内异常事件发生的次数反映了生产过程的风险水平。当异常事件数高于平均水平时,表明生产过程风险升高。

一段时间内异常事件的发生次数是一个变化的量,其大小不仅与自动控制系统相关,也与操作人员密切相关。化工生产过程需要 24 h 连续不断地监控,一般采用轮班制。本文以每个轮班时间段内出现的异常事件数作为研究对象,对其进行预测。对于人员技能娴熟、经验丰富的班组,其值班期间的异常事件数会低于平均水平,并且其操作结果会对接

下来的 3 个班组的异常事件数产生积极的影响。相反,对于新人较多的班组,在值班期间异常事件数会高于平均水平,并且其操作结果会对接下来的 3 个班组的异常事件数产生消极影响。本文在考虑操作班组的基础上,提出了一种化工过程异常事件数预测模型。该预测模型将贝叶斯理论与 Vine Copula 相结合,利用 DCS 报警数据库中记录的历史数据预测出下次轮班期间异常事件的发生次数。当预测的异常事件数高于正常水平时,值班班组对风险高的安全系统要加强管理,防止事故的发生。这种动态风险评估方法基于普遍存在的报警数据库中的历史数据,具有很强的可操作性,从整体上加强了生产过程安全管理能力,提高了生产过程的安全性。

1 贝叶斯理论

贝叶斯分析方法是一种统计推断方法,与经典的统计分析方法相比具有将先验信息与样本似然函数相结合的特点,使之更加灵活、直观并具有较明显的统计分析优势。贝叶斯预测模型是运用贝叶斯统计进行的一种预测,该预测方法是一种以动态模型为研究对象的时间序列预测方法。贝叶斯模型不仅利用了前期的数据信息,还加入了决策者的经验和判断等信息,并将客观因素和主观因素结合起来,对异常情况的发生具有较多的灵活性。目前,贝叶斯预测方法已经被广泛运用于各个科研领域^[8-10]。

假设化工过程某个班组在值班期间异常事件发生次数的历史观测数据为 $\text{data} = \{x(1), x(2), \dots, x(t-1), x(t)\}$, 异常事件数 x 的先验分布为 $f(x)$, 依据贝叶斯公式,其后验分布为

$$\underbrace{f(x | \text{data})}_{\text{后验分布}} = \frac{g(\text{data} | x) f(x)}{\int g(\text{data} | x) f(x) dx} \propto \frac{g(\text{data} | x) \underbrace{f(x)}_{\text{先验分布}}}{\underbrace{g(\text{data} | x)}_{\text{似然函数}}} \quad (1)$$

从式(1)可以看出,异常事件次数的后验分布 $f(x | \text{data})$ 正比于先验分布 $f(x)$ 与似然函数 $g(\text{data} | x)$ 的乘积。当贝叶斯统计分析被用于预测时,该后验分布即为预测分布。如果先验分布 $f(x)$ 和似然函数 $g(\text{data} | x)$ 可以使得先验分布和后验分布有相同的形式,那么就称先验分布与似然函数是共轭的,该先验分布称为共轭先验分布。采用共轭先验分布避免了因观测数据更新增加的计算量,增加了贝叶斯统计分析的实用性。正态分布的共轭先验分布依旧是正态分布。若 x 服从正态分布 $N(\theta, \sigma^2)$, 其中方差 σ^2 已知,期望 θ 为服从正态分布 $N(\mu, \tau^2)$ 随机

变量。根据贝叶斯理论,变量后验分布为

$$\bar{y} | \text{data} \sim N(\bar{\theta}, \sigma^2) \quad (2)$$

$$\bar{\theta} \sim N\left(\frac{\tau^2}{\sigma^2/n + \tau^2} \bar{x} + \frac{\sigma^2/n}{\sigma^2/n + \tau^2} u_0, \left(\frac{1}{\tau^2} + \frac{n}{\sigma^2}\right)^{-1}\right) \quad (3)$$

其中: \bar{x} 为历史数据均值; n 为历史数据长度。

2 异常事件数预测模型

化工生产车间通常采用四班三倒轮流值班的排班制度。假设化工生产车间有 A、B、C、D 4 个操作班组,通常其值班顺序为 A→B→C→D→A→B→C→D。由于班组操作能力不同,在每个班组值班期间生产过程发生的异常事件次数 x_A, x_B, x_C, x_D 不同。本文采用双层分布模型描述化工过程班组值班期间的异常事件数,如图 2 所示。

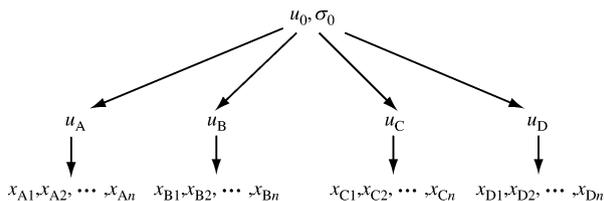


图 2 异常事件数贝叶斯双层分布模型

Fig. 2 Bayesian double distribution model of abnormal event numbers

该模型表示每个班组值班时异常事件数期望服从已知方差的正态分布,并且 4 个班组异常事件数的期望服从同一正态分布。

$$x_A \sim N(u_A, \sigma_A^2), \quad x_B \sim N(u_B, \sigma_B^2),$$

$$x_C \sim N(u_C, \sigma_C^2), \quad x_D \sim N(u_D, \sigma_D^2),$$

$$u_A, u_B, u_C, u_D \sim N(u_0, \sigma_0^2) \quad (4)$$

当存在各个班组的历史异常事件数 $\text{data} = (X_A, X_B, X_C, X_D)$ 时,依据式(3)和式(4),可得各个班组的异常事件数预测分布:

$$\tilde{x}_m | X_m \sim N(\tilde{u}_m, \sigma_m^2), m \text{ 表示 } A, B, C, D \text{ 班组}$$

$$\tilde{\mu}_m \sim N\left(\frac{\sigma_0^2}{\sigma_m^2/n_m + \sigma_0^2} \bar{x}_m + \frac{\sigma_m^2/n}{\sigma_m^2/n_m + \sigma_0^2} \mu_0, \left(\frac{1}{\sigma_0^2} + \frac{n}{\sigma_m^2}\right)^{-1}\right) \quad (5)$$

其中 n_m 是班组 m 的历史数据的个数。根据贝叶斯理论,预测值即为预测分布的期望:

$$\hat{x}_{m(n+1)} = \frac{\sigma_0^2}{\sigma_m^2/n_m + \sigma_0^2} \bar{x}_m + \frac{\sigma_m^2/n}{\sigma_m^2/n_m + \sigma_0^2} \mu_0 \quad (6)$$

然而,在实际值班过程中,由于班组操作先后顺序不同,先值班的班组对后值班的班组会产生影响。

例如,前面的班组操作能力较强,在其值班期间对监控系统进行有效调节,使得监控系统性能稳定,过程操作变量都在正常的生产范围,不会给后续的班组带来操作上的不利影响。相反若前面班组操作能力较弱,则会对后续班组带来较大的影响,也会造成后续班组的异常事件数高于正常水平。由于相同的两个班组轮班时间间隔不同,轮班间隔越长,影响程度越小,导致这种影响呈现非对称性。例如,A 班组的操作直接影响 B 班组,而 B 班组与 A 班组轮班时间间隔两次轮班,导致 A、B 两个班组之间存在非对称依赖关系,即 $\rho_{AB} \neq \rho_{BA}$ 。这种班组之间的非对称依赖关系体现在各个班组值班时异常事件数的非对称依赖关系。因此,如果将各个班组值班时异常事件数的非对称依赖关系融合到贝叶斯预测模型中,可以提高异常事件数的预测精度。Copula 函数是建立依赖模型的有效方法。当知道各个班组异常事件数的相关系数时,可先采用 Copula 求出异常事件数的联合概率分布函数,再采用贝叶斯理论求得后验联合概率密度函数,根据联合概率密度函数实现异常事件数预测。

根据 Sklar 定理^[11],当变量的相关系数已知时,变量的累积联合概率密度函数 $F(x_1, \dots, x_n)$ 可以通过 Copula 求得。

$$F(x_1, \dots, x_n) = C(u_1, u_2, u_3, \dots, u_n)$$

$$u_j = F_j(x_j) = \int_0^{x_j} f_j(x_j) dx_j, j = 1, 2, \dots, n \quad (7)$$

其中: C 为 Copula 函数; F_1, \dots, F_n 为各个随机变量的累积边缘分布函数。该联合分布函数只与变量的相关性有关,与变量的分布类型无关。相应的联合分布函数 $f(x_1, \dots, x_n)$ 为

$$f(x_1, \dots, x_n) = \sum_{i=1}^n f_i(x_i) c(u_1, u_2, u_3, \dots, u_n) \quad (8)$$

$$c = \frac{\partial^n C(u_1, \dots, u_n)}{\partial u_1, \dots, \partial u_n}$$

其中:随机变量 x_i 的 $f_i(x_i)$ 为边缘分布函数; c 为 Copula 密度函数。本文采用 Vine Copula^[12] 求取联合概率密度。Vine Copula 是以树型结构为基础的一种新型多元 Copula,它将多元 Copula 分解成多个 Pair-Copula^[13],比传统多元 Copula^[14] 有更强的灵活性,而 D-Vine Copula^[15] 是 Vine Copula 中最灵活的一种 Copula。图 3 为四元 D-Vine Copula 结构图,该 Copula 由 6 个二元 Copula 构成。

采用 D-Vine Copula 求取四元联合分布函数,其表达式为

$$f(x_1, x_2, x_3, x_4) = f_1(x_1) f_2(x_2) f_3(x_3) f_4(x_4) \times c_{12} \times c_{23} \times c_{34} \times c_{13|2} \times c_{24|2} \times c_{14|23} \quad (9)$$

其中 $c_{12}, c_{23}, c_{34}, c_{13|2}, c_{24|2}, c_{14|23}$ 为各个 Pair-Copula 密度函数。

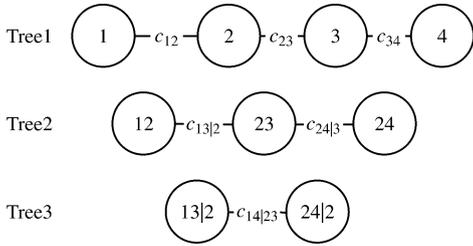


图 3 四元 D-Vine Copula 结构

Fig. 3 Construction of D-Vine Copula with four variables

本文选择二元高斯 Copula 作为 D-Vine Copula 的 Pair-Copula, 二元高斯 Copula 密度函数的表达式如下:

$$c(u_1, u_2) = \frac{1}{\sqrt{1-\rho^2}} \exp\left[-\frac{\rho^2(u_1^2 + u_2^2) - 2\rho u_1 u_2}{2(1-\rho^2)}\right] \quad (10)$$

二元高斯 Copula 中待定参数 ρ 为边缘分别为 u_1, u_2 的两个变量的线性相关系数, 该参数表征两个变量的相关性。依据式(9), 4 个班组的异常事件数联合概率密度函数为

$$f(x_A, x_B, x_C, x_D) = f_A(x_A) f_B(x_B) f_C(x_C) f_D(x_D) \times c_p | \rho_m \quad (11)$$

其中 c_p 为 Pair-Copula 密度函数的乘积, 由于班组之间的依赖关系非对称, 因此, c_p 的结构基于预测对象。

$$\underbrace{f(x_A, x_B, x_C, x_D | \text{data})}_{\text{后验联合分布}} = \frac{\prod_{j=1}^n g_A(x_{Aj}) g_B(x_{Bj}) g_C(x_{Cj}) g_D(x_{Dj}) \times f(x_A, x_B, x_C, x_D)}{\underbrace{\prod_{j=1}^n g_A(x_{Aj}) g_B(x_{Bj}) g_C(x_{Cj}) g_D(x_{Dj}) \times f(x_A, x_B, x_C, x_D)}_{\text{似然函数}} \underbrace{f(x_A, x_B, x_C, x_D)}_{\text{先验联合分布}}} \infty \quad (13)$$

其中 g 为似然函数。实际情况中, 在 A 班组第 $n+1$ 次轮班之前, 其他 3 个班组异常事件数 $x_{Bn+1}, x_{Cn+1}, x_{Dn+1}$ 已知, 值班 A 班组异常事件数预测值可表示为

$$\tilde{x}_{An+1} = E(x_{An+1} | \text{data}, x_{Bn+1}, x_{Cn+1}, x_{Dn+1}) = \int_0^{+\infty} x_{A(n+1)} f(x_{Bn+1}, x_{Cn+1}, x_{Dn+1}, x_{An+1} | \text{data}) dx_{An+1} \quad (14)$$

$$c_p | \rho_m = \begin{cases} c_{BC} \times c_{CD} \times c_{DA} \times c_{BD} | C \times c_{CA|D} \times c_{BA|CD}; & \text{预测 A 班组异常事件频率, } \rho_m = \rho_A \\ c_{CD} \times c_{DA} \times c_{AB} \times c_{CA|D} \times c_{DB|A} \times c_{CB|DA}; & \text{预测 B 班组异常事件频率, } \rho_m = \rho_B \\ c_{DA} \times c_{AB} \times c_{BC} \times c_{DB|A} \times c_{AC|B} \times c_{DC|AB}; & \text{预测 C 班组异常事件频率, } \rho_m = \rho_C \\ c_{AB} \times c_{BC} \times c_{CD} \times c_{AC|B} \times c_{BD|C} \times c_{AD|BC}; & \text{预测 D 班组异常事件频率, } \rho_m = \rho_D \end{cases}$$

其中 ρ_m 是 Pair-copula 参数:

$$\begin{bmatrix} \rho_A \\ \rho_B \\ \rho_C \\ \rho_D \end{bmatrix} = \begin{bmatrix} \rho_{BC} & \rho_{CD} & \rho_{DA} & \rho_{BD|C} & \rho_{CA|D} & \rho_{BA|CD} \\ \rho_{CD} & \rho_{DA} & \rho_{AB} & \rho_{CA|D} & \rho_{DB|A} & \rho_{CB|DA} \\ \rho_{DA} & \rho_{AB} & \rho_{BC} & \rho_{DB|A} & \rho_{AC|B} & \rho_{DC|AB} \\ \rho_{AB} & \rho_{BC} & \rho_{CD} & \rho_{AC|B} & \rho_{BD|C} & \rho_{AD|BC} \end{bmatrix}$$

本文将该参数设为相对应的班组异常事件数的线性相关系数, 该系数可通过历史异常事件数求得:

$$\rho_{mn} = \frac{\sum_{i=1}^N (x_{mi} - \bar{x}_m)(x_{ni} - \bar{x}_n)}{\sqrt{\sum_{i=1}^N (x_{mi} - \bar{x}_m)^2 \sum_{i=1}^N (x_{ni} - \bar{x}_n)^2}} \quad (12)$$

其中 m, n 为 A, B, C, D 且 $m \neq n, N$ 为历史观测数据总数。

若预测对象为 A 班组异常事件数 \tilde{x}_{An+1} , 当存在历史数据为

$$\text{data} = \begin{bmatrix} X_B & X_C & X_D & X_A \\ x_{B1} & x_{C1} & x_{D1} & x_{A1} \\ x_{B2} & x_{C2} & x_{D2} & x_{A2} \\ \vdots & \vdots & \vdots & \vdots \\ x_{Bn} & x_{Cn} & x_{Dn} & x_{An} \end{bmatrix}$$

根据贝叶斯理论可求得后验联合概率分布函数为

由于联合概率密度函数表达式非常复杂, 采用通常的积分方法无法求得结果, 因此采用蒙特卡罗积分法^[15]计算以上积分, 计算结果如下:

$$\tilde{x}_{A(n+1)} = \frac{\sum_{k=1}^N x_{A(n+1)}^k f(x_{Bn+1}, x_{Cn+1}, x_{Dn+1}, x_{An+1}^k | \text{data})}{\sum_{k=1}^N f(x_{Bn+1}, x_{Cn+1}, x_{Dn+1}, x_{An+1}^k | \text{data})} \quad (15)$$

其中: k 为采样次数; $x_{A_{n+1}}^k$ 是预测分布函数 $N(\bar{u}_A, \sigma_A^2)$ 的第 k 次采样值; \bar{u}_A 为分布 $N\left(\frac{\sigma_0^2}{\sigma_A^2/n + \sigma_0^2} \bar{x}_A + \frac{\sigma_A^2/n}{\sigma_A^2/n + \sigma_0^2} \mu_0, \left(\frac{1}{\sigma_0^2} + \frac{n}{\sigma_A^2}\right)^{-1}\right)$, 当总采样次数逼近无穷时, 采用蒙特卡罗积分法计算得到的积分结果趋向真值, 本文取 $N=10\ 000$ 。其他的班组报警数期望值也按照相同的方法进行预测。

3 应用实例

以某苯酚制取苯乙烯化工过程为例说明本文方法的应用。表 1 为该过程 2012 年 12 月各个班组轮班期间异常事件数观测数据, 即每个班组在各自值班期间(8 h)生产过程发生的异常事件数, 共 93 个观测值, 涉及流量、液位、压力等 229 个监测变量。2012-12-1 早班为 A 班轮班, 轮班顺序为 A、B、C、D。将前 33 个观测值作为历史数据, 分别采用基于贝叶斯理论预测方法以及灰色预测方法和神经网络预测方法预测异常事件数, 并对结果进行比较。

3.1 化工过程异常事件预测

使用本文提出的预测方法预测异常事件数, 首先要确定该模型的参数。该模型参数分为两个部分, 其中一部分为 4 个班组异常事件数分布函数的

初始参数, 这些参数根据经验以及对历史数据的分析设定为 $(\sigma_A, \sigma_B, \sigma_C, \sigma_D, \sigma_0, \mu_0) = (27, 25, 17, 16, 18, 186)$; 另一部分为 Copula 函数参数, D-Vine Copula 最优参数可采用极大似然方法以及马尔可夫链蒙特卡罗(Markov Chain Monte Carlo, MCMC)方法^[16]基于历史数据估计求得, 随着历史数据增加亦可实现动态估计。本文采用历史异常事件数线性相关系数作为 Pair-Copula 参数。根据式(12), 以表 1 中前 40 个数据作为初始历史数据, 求得 Pair-Copula 参数如下:

$$\begin{bmatrix} \rho_A \\ \rho_B \\ \rho_C \\ \rho_D \end{bmatrix} = \begin{bmatrix} 0.78 & 0.67 & 0.55 & 0.53 & 0.41 & 0.51 \\ 0.67 & 0.55 & 0.74 & 0.41 & 0.30 & 0.20 \\ 0.55 & 0.75 & 0.75 & 0.30 & 0.72 & 0.41 \\ 0.77 & 0.78 & 0.67 & 0.79 & 0.53 & 0.60 \end{bmatrix} \quad (16)$$

灰色预测模型是根据灰色系统理论提出来的的一种预测方法, 由于其建模过程容易, 使用样本少, 预测公式简单, 易于求解, 因而该模型得到快速发展, 并被广泛应用于经济、管理和工程领域。本文采用改进的灰色预测方法 MGM(1, 1)^[17] 作为对比预测模型, 该模型已被有效地用于海浪能时间序列预测中。本文以最近 10 个轮班异常事件数作为 MGM(1, 1) 预测模型输入数据, 预测步数设为 1, 即仅预

表 1 各个班组异常事件数观测值

Table 1 Abnormal event number observations of teams

日期	异常事件数			日期	异常事件数		
	早班	中班	晚班		早班	中班	晚班
2012-12-1	120	155	133	2012-12-17	171	222	199
2012-12-2	141	147	178	2012-12-18	165	180	234
2012-12-3	149	121	151	2012-12-19	208	171	191
2012-12-4	191	170	136	2012-12-20	251	213	180
2012-12-5	172	207	193	2012-12-21	217	258	239
2012-12-6	147	187	256	2012-12-22	193	192	230
2012-12-7	221	138	169	2012-12-23	211	177	188
2012-12-8	196	212	151	2012-12-24	197	191	162
2012-12-9	172	205	228	2012-12-25	175	180	177
2012-12-10	167	185	237	2012-12-26	151	163	208
2012-12-11	197	177	201	2012-12-27	201	169	172
2012-12-12	247	237	175	2012-12-28	236	243	181
2012-12-13	178	180	191	2012-12-29	158	181	194
2012-12-14	158	180	201	2012-12-30	157	178	191
2012-12-15	210	143	167	2012-12-31	197	172	191
2012-12-16	195	183	160				

测下个轮班异常事件数。

神经网络是 20 世纪 40 年代兴起的一种预测决策技术,因其具有极强的非线性动态处理能力、强大的自适应、自学习功能而被广泛应用于不同领域。在众多神经网络中,BP 神经网络是最具代表性和应用最为广泛的一种网络模型^[18-19]。本文采用 BP 神经网络建立化工过程异常事件数预测模型作为基于贝叶斯理论的预测模型的对比模型。采用距离预测值最近的 4 个观测值作为神经网络的输入,BP 神经网络采用 4-5-1 结构(即输入层节点、隐含层节点和输出层节点分别为 4、5 和 1)。表 1 中前 40 个数据首先用于训练神经网络。

3.2 预测结果分析

图 4 形象地展示了各个模型的预测结果,横坐

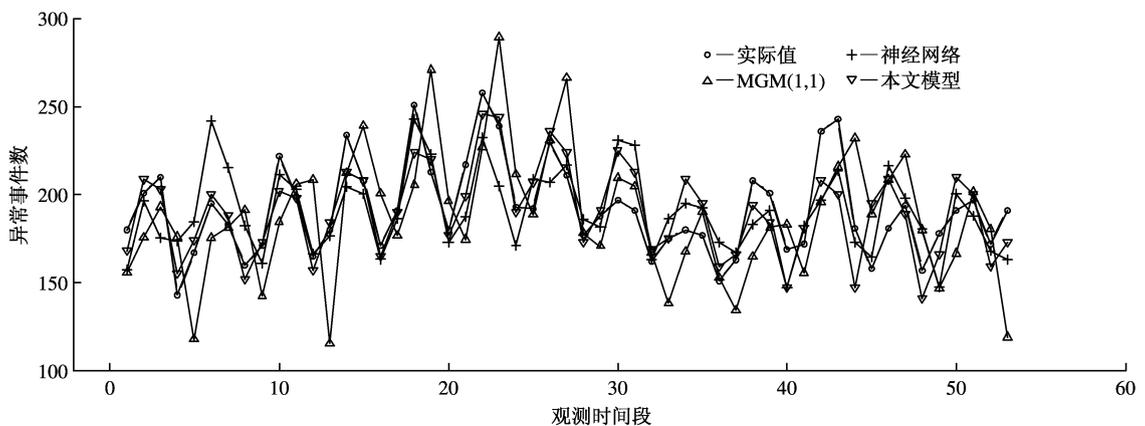


图 4 化工过程异常事件数预测结果

Fig. 4 Prediction result of abnormal event numbers in chemical process

表 2 预测结果评价

Table 2 Evaluation results

方法	MAE	RMSE	TIC	MAPE/ %	ME	N
MGM(1,1)	24.386 4	30.012 38	0.988 991 4	8.708 04	83.838 95	12
神经网络	17.548 65	39.463 28	0.991 358 6	10.000 28	112.584 1	9
本文模型	12.96	16.515 45	0.995 426	6.892 3	43	2

从表中可以看出,相对于 MGM(1,1)灰色预测模型和神经网络预测模型,基于贝叶斯理论的预测模型的预测结果各项指标均为最优值,与图 4 所示的预测结果相符。

准确预测化工过程异常事件数是实现动态风险管理的一种方法,在此基础之上,可以进一步对化工生产安全系统进行研究。由于班组能力不同导致安全管理子系统在不同班组操作时控制失效的概率不同,如果能够评估出在不同班组操作时各个安全管理子系统控制失效的概率,采用异常事件预测值乘

标为异常事件数统计时间段,本文基于操作班组预测异常事件数,因此采用每个轮班时间(8 h)作为一个统计时间段。纵坐标为每个时间段内异常事件发生次数,即异常事件数。从图 4 可以看出采用基于贝叶斯理论的预测模型预测值与实际值的拟合效果明显优于其他两个模型。为进一步定量检验基于贝叶斯理论的风险预测方法的有效性,本文采用平均误差(MAE)、均方根误差(RMSE)、平均绝对误差(MAPE)、希尔不等系数(TIC:值越大表明预测性能越好)、最大误差(ME)、失效次数(N:相对偏差大于 20%)6 项指标对预测结果进行评价。表 2 中列出了 3 种预测模型化工过程异常事件数的预测评价结果。

以子系统控制失败的概率可以实现停车、事故等次数预测。

4 结束语

化工事故造成的灾难性后果不断诠释着化工生产过程的高风险性,促使工厂积极完善生产风险管理系统。虽然化工事故数据稀少,但发展成为未遂事件的异常事件却大量存在于报警数据库中。本文提出的基于贝叶斯理论的化工过程异常事件数预测

方法能够根据历史异常事件数较准确地提前预测出化工过程异常事件数。当预测值大于正常水平时,相应的操作人员要提高警惕,积极监控生产过程,甚至主动监测安全控制系统,这有助于降低生产风险,提高产品质量。该预测方法是一种动态预测方法,具有很强的可操作性,是实现动态风险管理的有效手段。

参考文献:

- [1] 贾伟,朱建新,高增梁,等. 区域定量风险评价方法及其在化工园区中的运用[J]. 中国安全科学学报, 2009, **19**(5): 140-146.
- [2] Reniers G L L, Sørensen K, Dullaert W. A multi-attribute systemic risk index for comparing and prioritizing chemical industrial areas[J]. Reliability Engineering & System Safety, 2012, **98**(1): 35-42.
- [3] Yu Qian, Jiang Juncheng, Yu Hanhua. Research on the emergency response system of major dangerous chemical accident on highway based on the GIS[J]. Procedia Engineering, 2012, **45**: 716-721.
- [4] He Guizhen, Zhang Lei, Lu Yonglong, *et al.* Managing major chemical accidents in China: Towards effective risk information[J]. Journal of Hazardous Materials, 2011, **187**(1): 171-181.
- [5] Cox P, Niewöhner J, Pidgeon N, *et al.* The use of mental models in chemical risk protection: Developing a generic workplace methodology[J]. Risk Analysis, 2003, **23**(2): 311-324.
- [6] Meel A, Seider W D. Plant-specific dynamic failure assessment using Bayesian theory [J]. Chemical Engineering Science, 2006, **61**(21): 7036-7056.
- [7] Huang David, Chen Toly, Wang M J J. A fuzzy set approach for event tree analysis[J]. Fuzzy Sets and Systems, 2001, **118**(1): 153-165.
- [8] Raftery A E, Gneiting T, Balabdaoui F, *et al.* Using Bayesian model averaging to calibrate forecast ensembles [J]. Monthly Weather Review, 2005, **133**(5): 1155-1174.
- [9] 林震, 杨浩. 基于车速的交通事故贝叶斯预测[J]. 中国安全科学学报, 2003, **13**(2): 34-36.
- [10] Duan Qingyun, Ajami N K, Gao Xiaogang, *et al.* Multi-model ensemble hydrologic prediction using Bayesian model averaging[J]. Advances in Water Resources, 2007, **30**(5): 1371-1386.
- [11] Sklar A. Random variables, distribution functions, and copulas: A personal look backward and forward[C]// Distributions with Fixed Marginals and Related Topics. Hayward: Institute of Mathematical Statistics, 1996: 1-14.
- [12] Bedford T, Cooke R M. Vines——A new graphical model for dependent random variables [J]. The Annals of Statistics, 2002, **30**(4): 1031-1068.
- [13] Aas K, Czado C, Frigessi A, *et al.* Pair-copula constructions of multiple dependence[J]. Insurance: Mathematics and Economics, 2009, **44**(2): 182-198.
- [14] 韦艳华, 张世英. Copula 理论及其在金融分析上的应用[M]. 北京:清华大学出版社, 2008.
- [15] Brechmann E C, Schepsmeier U. Modeling dependence with C-and D-vine copulas: The R-package CDVine[J]. Journal of Statistical Software, 2013, **52**(3): 1-27.
- [16] Yu Cindy L, Li Haitao, Wells M T. MCMC estimation of Lévy jump models using stock and option prices[J]. Mathematical Finance, 2011, **21**(3): 383-422.
- [17] Truong D Q, Ahn K K. Wave prediction based on a modified grey model MGM (1, 1) for real-time control of wave energy converters in irregular waves [J]. Renewable Energy, 2012, **43**: 242-255.
- [18] Guo Zhenhai, Wu Jie, Lu Haiyan, *et al.* A case study on a hybrid wind speed forecasting method using BP neural network[J]. Knowledge-Based Systems, 2011, **24**(7): 1048-1056.
- [19] 吴微, 陈维强. 用 BP 神经网络预测股票市场涨跌[J]. 大连理工大学学报, 2001, **41**(1): 9-15.

欢迎订阅

《华东理工大学学报(自然科学版)》

地址:上海市梅陇路 130 号 436 信箱 邮编:200237

邮发代号:4-382