

基于过程分析技术和设计空间的金银花醇沉加醇过程终点检测

徐冰¹, 罗赣¹, 林兆洲¹, 艾路¹, 史新元^{1,2}, 乔延江^{1,2}

(1. 北京中医药大学中药信息工程研究中心, 2. 教育部中药制药与新药开发关键技术工程研究中心, 北京 100029)

摘要 建立了一种新的基于过程分析技术(PAT)和质量源于设计(QbD)设计空间的中药制药过程终点分析与控制方法. 以近红外(NIR)光谱技术为PAT工具,采集正常操作条件下制药过程的多批次NIR光谱;采用主成分分析结合移动块相对标准偏差(PCAMBRSD)法,确定每一批次过程理想终点样本(DEPs),由多批DEPs的光谱信息构成过程终点设计空间;在过程终点设计空间确定的范围内,建立多变量统计过程控制(MSPC)模型,利用多变量Hotelling T^2 和SPE控制图对过程终点进行判断.应用上述方法,进行了金银花醇沉加醇过程终点检测研究,结果表明该方法灵敏、准确,适用于中药制药过程终点检测.

关键词 过程分析技术;设计空间;近红外光谱法;醇沉过程;终点检测

中图分类号 O654.9

文献标志码 A

美国食品药品监督管理局(FDA)颁布的《过程分析技术(PAT)工业指南》^[1]鼓励制药工业采用新的过程分析工具,通过合理的过程设计、分析与控制,增强对工艺过程的理解,降低过程不确定性和风险,并保证持续生产出满足质量要求的药品.作为PAT工具之一的近红外光谱(NIR)技术因具备全息、快速及无损等特点,在制药过程分析中应用广泛^[2-5].通过NIR在线检测,采集大量反映过程变化的信息,结合过程设计、多变量统计分析和建模等方法,可实现制药过程的实时控制和决策.

近年来,NIR光谱技术在中药制药过程中的应用逐渐增多^[6-8].其中,基于NIR光谱技术的过程终点检测与判断是中药制药过程分析的重要内容之一.中药制药过程终点检测方法可分为定量分析和定性分析两类.定量终点分析方法需使用离线样品建立在线定量校正模型^[9],其优点是可以实时反映过程中有效成分或指标性成分含量的变化,缺点是费时费力,且模型维护成本较高;定性分析方法(如移动块标准偏差法)^[10,11]用于评价连续光谱之间的波动情况,其优势是简单方便,无需定量模型即可分析过程终点,但缺点是存在检测延时,且准确度较低^[12].本文以金银花醇沉加醇过程为研究对象,将质量源于设计(QbD)中“设计空间(Design space)”的理念和方法引入中药制药过程分析,从设计层次保证过程终点检测的质量,同时集成主成分分析、移动块相对标准偏差和多变量统计过程控制等多种化学计量学手段,建立了一种新的基于QbD设计空间的灵敏、准确的过程终点检测方法.

1 原理和方法

1.1 基于设计空间的过程终点分析策略

人用药品注册技术国际协调会(ICH)在其指南文件ICH Q8(R2)中将设计空间定义为:可以确保产品质量的输入变量(如:物料属性)和工艺参数之间的多维结合与相互作用^[13].对化学药品和生物制品的生产过程建立设计空间时,通常采用实验设计和多变量统计分析的方法确定过程输入、关键工艺参数和产品关键质量属性之间的关系模型,基于过程模型开发设计空间^[14,15].中药因具有来源多样性和成分复杂性等特点,其产品多采用固定工艺参数模式生产.但由于过程输入的原料药材的波动以

收稿日期: 2013-06-03.

基金项目: 国家“重大新药创制”科技重大专项(批准号: 2010ZX09502-002)资助.

联系人简介: 乔延江,男,博士,教授,博士生导师,主要从事中药信息学研究. E-mail: yjqiao@263.net

史新元,女,博士,副教授,主要从事中药过程分析研究. E-mail: shixinyuan01@163.com

及生产过程中不可避免的扰动,通常会导致不同正常操作条件(NOCs)下的中药制药过程之间存在一定差异.当批次间差异不影响最终产品质量时,积累一定数量的NOCs批次就相当于在一定范围内获得了过程输入(存在批间变异)、过程扰动(主要指控制参数的不确定性和环境变化的影响)和终产品质量之间的相互作用关系,而这种关系与设计空间是一致的.据此,本文建立了基于NIR过程分析技术和设计空间的中药制药过程终点检测方法,如图1所示,该方法主要包括3个步骤:(1)采集一个制药过程中多个NOCs批次的NIR光谱数据,构成三维光谱矩阵 $X_{H \times K \times M}$ (H 代表批次, K 代表时间, M 代表变量);(2)基于光谱信息,采用主成分分析结合移动块相对标准偏差(PCA-MBRSD)法,确立各个批次中的理想终点(Desired end-points,DEPs),由各批DEPs样本光谱所包含的信息构成过程终点设计空间;(3)基于过程终点设计空间建立多变量统计过程控制(MSPC)模型,进行过程终点检测.

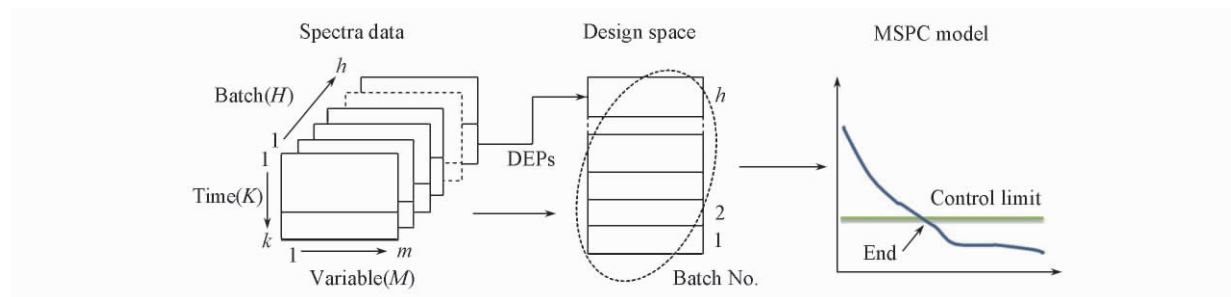


Fig. 1 A strategy for end-point detection based on design space and process analytical technology

1.2 PCA-MBRSD 法

将过程光谱矩阵 $X_{H \times K \times M}$ 按照变量方向重组成矩阵 $X_{HK \times M}$,对光谱矩阵 $X_{HK \times M}$ 进行主成分分析(PCA),选取前 A 个可以解释光谱数据中大部分变化的主成分.分别在每一批次光谱数据的主成分空间中,由公式(1)计算每个时间点的光谱和第1个时间点光谱之间的欧式距离:

$$d_i = \text{sqrt} \left[\sum_{j=1}^A (t_{ij} - t_{1j})^2 \right] \quad (1)$$

式中, t_{ij} 为第 i ($1 \leq i \leq k$)个时间点对应的第 j 个主成分的得分(Score)值.

对于第 h 批过程光谱数据,可以连续获得 k 个距离值($d_1 \sim d_k$).基于欧氏距离值,采用移动块相对标准偏差法评价该批中第1至第 k 时间点的光谱的波动情况.MBRSD值可由公式(2)~(4)计算:

$$\text{MBRSD} = \text{SD} / \bar{d} \quad (2)$$

$$\text{SD} = \text{sqrt} \left[\frac{1}{w-1} \sum_{i=1}^w (d_i - \bar{d})^2 \right] \quad (3)$$

$$\bar{d} = \frac{1}{w} \sum_{i=1}^w d_i \quad (4)$$

式中, w 为窗口宽度.对于第 h 批数据,可以获得($k-w+1$)个连续MBRSD值.合理设置MBRSD阈值,即可判断过程终点.选择各批中低于MBRSD阈值的光谱数据(即DEPs),并重新排列为矩阵 $X_{n \times m}$ (n 为样本个数, m 为变量数),则矩阵 $X_{n \times m}$ 包含的信息代表正常操作模式下过程终点的设计空间.

1.3 多变量统计过程控制

将矩阵 $X_{n \times m}$ 进行均值标准化处理后,进行PCA分析,并建立PCA分析模型:

$$X = T_A P_A^T + E \quad (5)$$

式中, T_A 为由前 A 个主成分确立的得分矩阵; P_A 为负荷矩阵; E 为残差矩阵,表示矩阵 X 在 p_{A+1} 至 p_m 方向上的变化.经过PCA分析, m 维原始数据空间被2个正交且互补的子空间代替,即 A 维主成分空间和 $m-A$ 维残差子空间.主成分子空间反映正常数据的变化,而残差子空间则反映噪音的变化.分别采用Hotelling T^2 (或 D 统计量)和平方预测误差(SPE或 Q 统计量)统计量,在2个子空间中建立多变量统计过程控制模型,可实现多变量生产过程的统计监控. T^2 统计量的计算公式如下:

$$T^2 = \sum_{i=1}^A \frac{t_i^2}{\lambda_i} = x^T P_A A^{-1} P_A^T x \quad (6)$$

式中, $t_{(i)}$ 为得分向量 t 中的第 i 个元素, x 为样本, Λ 为由前 A 个主成分所对应的特征值组成的对角阵. SPE 统计量计算公式如下:

$$Q = \sum_{j=1}^m [x_{(j)} - \hat{x}_{(j)}]^2 = \|(I - P_A P_A^T) x\|^2 \quad (7)$$

式中, $x_{(j)}$ 为该样本第 j 个变量的测量值, $\hat{x}_{(j)}$ 为该样本第 j 个变量的 PCA 模型预测值, I 为单位矩阵.

T^2 统计量和 SPE 统计量从不同角度分析过程状态. T^2 统计量表征样本和主成分分子空间原点的距离, 主要通过主成分模型内部的主成分向量的波动来反映过程变化的情况; 而 SPE 统计量通过分析测量数据的残差, 即主成分模型外部数据变化进行过程监测, 主要衡量变量之间相关性被改变的程度. 基于正常过程数据建立 T^2 统计量和 SPE 统计量的控制限, 可对过程状态进行监控. T^2 统计量的控制限可以利用 F 分布按公式(8) 进行计算^[16]:

$$T_{\text{lim}}^2 = \frac{A(n^2 - 1)}{n(n - A)} F_{(A, n-A, \alpha)} \quad (8)$$

式中, α 为检验水平, $F_{(A, n-A, \alpha)}$ 为在自由度 A 和 $n-A$ 下的 F 分布临界值.

假设 SPE 统计量服从二项分布, 检验水平 α 下的 SPE 控制限的计算方法为^[17]:

$$\text{SPE}_\alpha = \theta_1 \left[\frac{c_\alpha h_0 \sqrt{2\theta_2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \right]^{1/h_0} \quad (9)$$

$$h_0 = 1 - \frac{2\theta_1 \theta_3}{3\theta_2^2}, \quad \theta_i = \sum_{j=A+1}^m \lambda_j^i$$

式中, c_α 为正态分布在检验水平为 α 下的临界值, λ 为特征值, m 为全部主成分个数.

2 实验部分

2.1 金银花醇沉过程

取金银花药材适量, 分别加 15 和 10 倍量水煎煮 2 次, 每次 0.5 h, 合并提取液, 过滤, 将滤液浓缩至指定密度, 取 400 mL 浓缩液按照表 1 中的工艺参数进行醇沉加醇操作.

Table 1 Operating conditions for different batches

Batch No.	Density of water extract(25 °C)	Concentration of alcohol precipitation(%)	Rotation speed of the pump/(r · min ⁻¹)
1—11	1.10	75	84
12—13	1.10	75	84
14	1.10	75	100(Abnormal)
15	1.10	77(Abnormal)	84
16	1.10	73(Abnormal)	84
17	1.15(Abnormal)	75	84

采用 YZ1515X 型蠕动泵和内径为 3 mm 的硅胶软管输送乙醇. 在加醇过程中, 从第 0.5 min 起每隔 0.5 min 在固定位置取样, 经 9000 r/min 转速离心后的真溶液用 Thermo Nicolet FTIR 仪进行分析, 光谱范围 4000 ~ 10000 cm⁻¹, 分辨率 4 cm⁻¹, 每个样品以空气为参比扫描 16 次后取平均光谱^[18], 每批连续取 60 个样品. 第 1 ~ 11 批用于建立金银花醇沉过程终点控制模型, 第 12 ~ 17 批用于模型的验证和应用.

2.2 红外光谱波段筛选

前文^[19, 20]对金银花醇沉过程的近红外光谱定量分析进行了系统研究, 采用移动窗口偏最小二乘法(mwPLS) 确定绿原酸的特征吸收波段为 8500 ~ 7300 cm⁻¹, 采用间隔偏最小二乘法(iPLS) 确定木樨草苷的特征吸收波段为 6201 ~ 6001 cm⁻¹, 以及

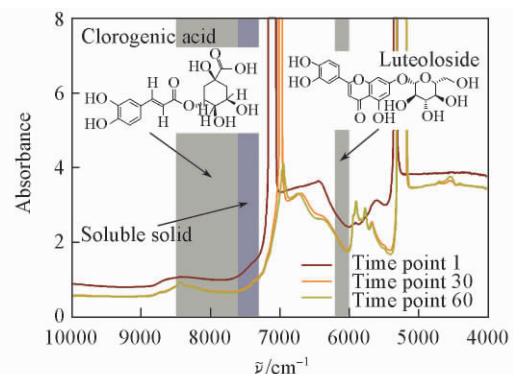


Fig. 2 Raw NIR spectra with the characteristic wave-band during the alcohol precipitation process

总固体的特征吸收波段 $7605 \sim 7308 \text{ cm}^{-1}$, 且均为原始光谱的定量效果最佳. 综合 NIR 定量分析结果 (图 2), 确定 $8500 \sim 7300 \text{ cm}^{-1}$ 和 $6201 \sim 6001 \text{ cm}^{-1}$ 两个波段下的 365 个波长变量, 用于开展加醇过程终点分析与检测研究.

2.3 统计软件

本文涉及的算法均由 Matlab 7.0 (MathWorks Inc., USA) 平台实现, PCA 算法使用 PLS Toolbox 2.1 (Eigenvektor Research Inc., USA), 其它程序自行编制.

3 结果与讨论

3.1 设计空间的构建

由表 1 中第 1~11 批数据构成光谱矩阵 $X_{660 \times 365}$ (即 11×60 个光谱样本, 365 个变量). 对 $X_{660 \times 365}$ 进行 PCA 分析发现, 前 2 个主成分可以解释原始数据 99.98% 的信息. 因此选择前 2 个主成分, 将窗口大小 w 设置为 6, 对各批次中的样本计算 MBRSD 值. 以批次 1, 5 和 9 为例, 3 个批次的 MBRSD 曲线如图 3 所示. 随着加醇过程的进行, MBRSD 值呈降低趋势. 参考经典高效液相色谱 (HPLC) 法的验证标准, 将 RSD 的阈值设置为 1%^[21]. 结果表明, PCA-MBRSD 法虽然简便, 但终点判断效果并不稳定, 例如批次 1 中第 36 时间点 MBRSD 值小于 1%, 但随后又超过阈值. 因此, 本研究中将连续 5 点的 MBRSD 值低于 1% 时确定为终点状态, 此 MBRSD 区间对应的样本为理想终点样本.

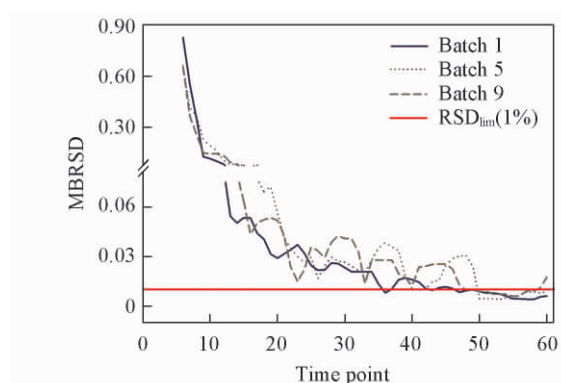


Fig. 3 Selection of end-points from the alcohol adding process under normal operating conditions by PCA-MBRSD

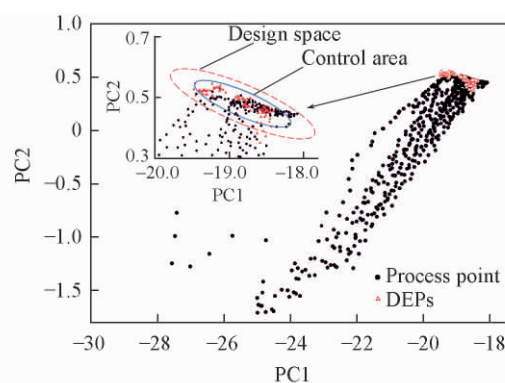


Fig. 4 Establishment of the design space for the end-point detection

从前 11 个批次中共选出 119 个 DEPs, 其在原始光谱确立的主成分空间中的位置如图 4 中蓝色三角形所示. 此 119 个样本包含的信息构成金银花醇沉过程终点检测的设计空间, 如图 4 中蓝色 99% 置信椭圆所示. 基于过程终点设计空间可建立相应的终点分析与控制策略, 并用于过程终点判断.

3.2 MSPC 模型的建立和应用

由 119 个 DEPs 组成终点光谱矩阵 $X_{119 \times 365}$, 对其进行了均值标准化处理. 以 15 折交叉验证和 PRESS 统计量优选主成分, 并确定主成分数为 4. 基于 PCA 分析结果, 将检验水平 α 设置为 0.95 (对应图 4 中的绿色置信椭圆), 按式 (8) 计算 Hotelling T^2 控制限为 10.1425, 按式 (9) 计算 SPE 控制限为 0.0389. 然后分别建立多变量 Hotelling T^2 控制图和 SPE 控制图. 在对新的批次进行监控前, 首先将新批次的过程光谱数据以矩阵 $X_{119 \times 365}$ 的均值和标准差进行标准化预处理, 再按式 (6) 和 (7) 计算相应 T^2 统计量和 SPE 统计量. T^2 统计量与 SPE 统计量具有互补作用: T^2 统计量主要在变量相互关系结构未发生变化的前提下监测系统工作点的变化; SPE 统计量则主要监测输入变量的数据结构是否变化, PCA 模型是否仍适用. 醇沉批次 13 为正常操作批次, 其加醇过程终点检测如图 5 黑色曲线所示. 在第 38 个时间点该批次的 T^2 统计量和 SPE 统计量均低于控制限, 因此第 38 个时间点为批次 12 的加醇终点. 在达到过程终点后, 第 49 个时间点的 T^2 值超过控制限, 而 SPE 值未超过控制限, 初步判断该点可能为偶然测量误差所致.

在醇沉批次 14 中, 加快了蠕动泵的转速, 即加快输送乙醇的速度. 结果表明, 通过 T^2 控制图和

SPE 控制图可以灵敏、准确地检测过程终点的改变. 如图 5 红色曲线所示, 第 34 个时间点为加醇过程终点, 并在达到终点后各点的 T^2 统计量和 SPE 统计量均处于受控状态. 醇沉批次 15 为模拟加醇量过多的故障批次, 即在正常加醇操作后继续加醇, 使最终醇沉浓度为 77%. 结果表明, 根据 T^2 控制图和 SPE 控制图可检测到加醇终点(第 38 个时间点), 如图 5 蓝色曲线所示. 而随着乙醇的继续加入, T^2 统计量从第 44 个时间点开始超过控制限, 同时 SPE 统计量保持正常, 表明过程出现异常, 但 MSPC 模型仍然适用. 醇沉批次 16 为模拟加醇量不足的故障批次, 即未达到指定加醇量, 结果如图 5 粉色曲线所示. 无论是 T^2 控制图还是 SPE 控制图均无法检测到过程终点, 提示过程存在异常. 醇沉批次 17 为模拟过程输入的异常, 即药液浓度密度高于正常批次, 结果如图 5 青色曲线所示. 虽然 T^2 控制图可以检测到过程终点(第 39 个时间点), 但各时间点的 SPE 统计量均高于控制限, 表明正常光谱变量所代表的关系结构已发生改变, MSPC 监控模型不再适用, 并且过程出现异常.

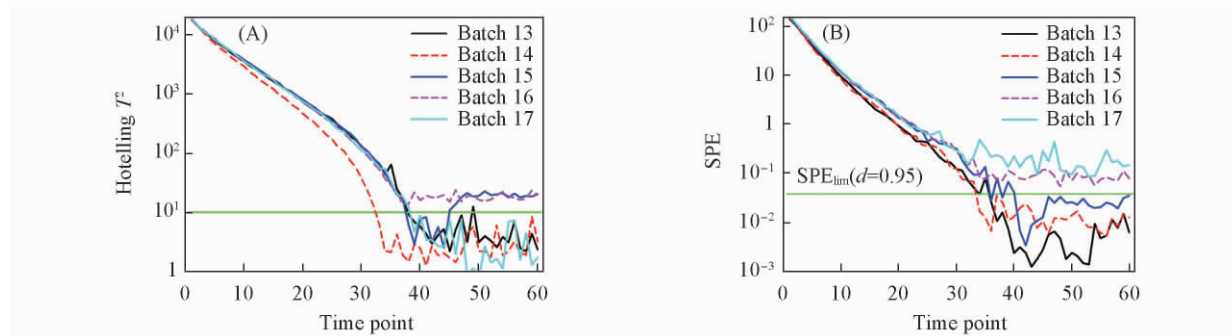


Fig. 5 End-point detection of the alcohol adding process for Batches 13 to 17

(A) Hotelling T^2 control chart; (B) SPE control chart.

3.3 HPLC 验证

采用 HPLC 法^[19]对批次 12, 13 和 14 中各时间点样品的绿原酸含量进行定量分析, 以验证过程终点检测的可靠性, 结果如图 6 所示. 在第 12 批中, MSPC 模型确定的过程终点(第 38 个时间点)对应的浓度为 1.61 mg/mL, 终点后的 23 个时间点绿原酸浓度的 RSD 值为 2.64%; 在第 13 批中, MSPC 模型确定的过程终点(第 38 个时间点)对应的浓度为 1.50 mg/mL, 终点后的 23 个时间点绿原酸浓度的 RSD 值为 2.37%; 在第 14 批中, MSPC 模型确定的过程终点(第 34 个时间点)对应的浓度为 1.56 mg/mL, 之后的 27 个时间点绿原酸浓度的 RSD 值为 1.30%, 表明使用本文方法进行金银花醇沉加醇过程终点检测时, 不仅终点判断准确, 而且没有时间延时.

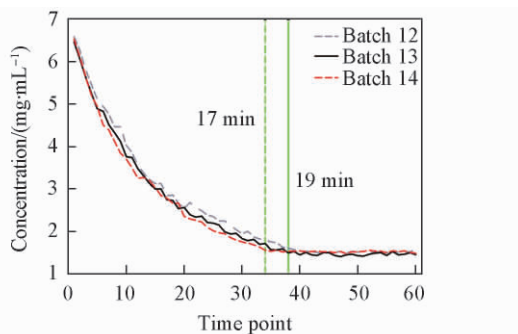


Fig. 6 HPLC validation for the end-point detection of the alcohol adding process

综上所述, 本文将 PAT 和 QbD 技术相结合用于检测中药制药过程终点. 以金银花醇沉过程为研究对象, 采用 NIR 光谱对多批次醇沉加醇过程进行快速分析, 用 PCA-MBRSD 法从正常加醇过程 NIR 数据中获得理想终点样本, 由理想终点样本构成加醇过程终点设计空间. 在设计空间确定的范围内, 采用 MSPC 方法建立了过程终点检测的多变量 Hotelling T^2 和 SPE 控制图, 对正常批次和异常操作条件下批次过程进行终点监控. 结果表明, 本文方法可以灵敏、准确地判断加醇过程终点, 且无延时, 可节省约 1/3 的加醇操作时间. 然而, 本文的研究对象为醇沉过程的真溶液部分, 实际醇沉过程中还存在大量的颗粒或絮凝物, 可以联合在线颗粒分析工具, 如聚焦光束反射测量(FBRM)和颗粒录影显微镜(PVM)等开展进一步研究.

参 考 文 献

- [1] US Food and Drug Administration, *Guidance for Industry: PAT — A Framework for Innovative Pharmaceutical Development*,

Manufacturing and Quality Assurance, 2004

- [2] Rosas J. , Blanco M. , González J. , Alcalá M. , *Talanta* , 2012 , 97 , 163—170
- [3] Karande A. , Heng P. , Liew C. , *Int. J. Pharm.* , 2010 , 396(1/2) , 63—74
- [4] Qiao W. Q. , *Chem. Res. Chinese Universities* , 2012 , 28(2) , 239—241
- [5] Sulub Y. , LoBrutto R. , Vivilecchia R. , Wabuyele B. , *Anal. Chim. Acta* , 2008 , 611(2) , 143—150
- [6] Li W. , Xing L. , Fang L. , Wang J. , Qu H. , *J. Pharm. Biomed. Anal.* , 2010 , 53(3) , 350—358
- [7] Zhu X. R. , Li N. , Shi X. Y. , Qiao Y. J. , Zhang Z. Y. , *Chem. J. Chinese Universities* , 2008 , 29(5) , 906—911(朱向荣, 李娜, 史新元, 乔延江, 张卓勇. 高等学校化学学报, 2008, 29(5) , 906—911)
- [8] Rosa S. , Barata P. , Martins J. , Menezes J. , *J. Pharm. Biomed. Anal.* , 2008 , 47(2) , 320—327
- [9] Jin Y. , Ding H. , Liu X. , Wan X. , Luan L. , Wu Y. , *Spectrochim. Acta A Mol. Biomol. Spectr.* , 2013 , 109 , 68—78
- [10] Wu Z. , Tao O. , Dai X. , Du M. , Shi X. , Qiao Y. , *Vib. Spectrosc.* , 2012 , 63 , 371—379
- [11] Wu Y. , Jin Y. , Li Y. , Sun D. , Liu X. , Chen Y. , *Vib. Spectrosc.* , 2012 , 58 , 109—118
- [12] Momose W. , Imai K. , Yokota S. , Yonemochi E. , Terada K. , *Powder Technol.* , 2011 , 210(2) , 122—131
- [13] International Conference on Harmonisation , *ICH Q8(R2) : Pharmaceutical Development* , 2009
- [14] Wu H. , White M. , Khan M. , *Int. J. Pharm.* , 2011 , 405(1/2) , 63—78
- [15] Jiang C. , Flansburg L. , Ghose S. , Jorjorian P. , Shukla A. , *Biotechnol. Bioeng.* , 2010 , 107(6) , 985—997
- [16] Hu K. , Yuan J. , *J. Process Contr.* , 2008 , 18 , 797—807
- [17] Huang H. , Qu H. , *Anal. Chim. Acta* , 2011 , 707 , 47—56
- [18] Wu Z. , Xu B. , Du M. , Sui C. , Shi X. , Qiao Y. , *J. Pharm. Biomed. Anal.* , 2012 , 62 , 1—6
- [19] Xu B. , Wu Z. , Lin Z. , Sui C. , Shi X. , Qiao Y. , *Anal. Chim. Acta* , 2012 , 720 , 22—28
- [20] Xu B. , Lin Z. , Shi X. , Qiao Y. , Du M. , Wu Z. , *International Conference on Biomedical Engineering and Biotechnology* , IEEE Computer Society , Macau , 2012 , 1566—1571
- [21] Puchert T. , Holzhauser C. , Menezes J. , Lochmann D. , Reich G. , *Eur. J. Pharm. Biopharm.* , 2011 , 78(1) , 173—182

End-point Detection of the Alcohol Adding Process in Alcohol Precipitation of *Lonicerae Japonicae* Based on Design Space and Process Analytical Technology

XU Bing¹ , LUO Gan¹ , LIN Zhao-Zhou¹ , AI Lu¹ , SHI Xin-Yuan^{1 2*} , QIAO Yan-Jiang^{1 2*}

(1. Research Center of TCM Information Engineering , Beijing University of Chinese Medicine ,

2. Engineering Research Center of Key Technologies for Chinese Medicine Production and New Drug Development , Ministry of Education , Beijing 100029 , China)

Abstract Under the guidance of the FDA's process analytical technology(PAT) initiative and the ICH's quality by design(QbD) principles , a new approach was applied for the end-point detection of the pharmaceutical process of Chinese medicine preparations. Generally , three steps were included in the proposed approach: (1) perform the NIR analysis for the batch process under normal operating conditions(NOCs) , and collect the NIR spectra of different batches; (2) determine the desired end-points(DEPs) for every NOCs batch using the PCA based moving window relative standard deviation(PCA-MBRSD) method. Then , the QbD design space was defined , depending on these DEPs; (3) within the area specified by the design space , the multivariate statistical process control(MSPC) strategy was introduced , and the Hotelling T^2 and SPE control chart were established to monitor and detect the end-point of the manufacturing process. The above methods were tested on the end-point detection of the alcohol adding process of alcohol precipitation of *Lonicerae Japonicae* water extract. The results demonstrated that the sensitivity and accuracy of the proposed approach were satisfactory.

Keywords Process analytical technology(PAT) ; Design space; Near infrared spectroscopy(NIRS) ; Alcohol precipitation process; End-point detection (Ed. : N, K)